DEPARTMENT OF COMPUTER SCIENCE

ALBERT EINSTEIN SCHOOL OF PHYSICAL SCIENCES

ASSAM UNIVERSITY SILCHAR
(A Central University constituted under Act XIII of 1989)
Silchar - 788 011, Assam, India

DECLARATION

I **Rakesh Kumar** bearing Registration No**. Ph.D/730/09** dated **11/02/2009**, hereby declare that the subject matter of the thesis entitled "**Study of evolving neural network for word sense disambiguation"**is the record of work done by me and that the contents of this thesis did not form the basis for award of any degree to me or to anybody else to the best of my knowledge. The thesis has not been submitted in any other University / Institute.

Place :Silchar

Date : Rakesh Kumar

ACKNOWLEDGEMENTS

I express my sincere gratitude to my research supervisor Prof. Bipul Shyam Purkayastha. Professor and Head, Department of computer science , Dean School of physical science, Assam university for his expert guidance, invaluable suggestions, encouragement and keen interest in the conception planning and execution of the study. It is a great honor and privilege to complete the present study under his supervision.

I extend my sincere thanks to the entire faculty and staff of the Computer Science Department for their cooperation and encouragement during my research.

Express my deep sense of gratitude to my wife Nisha Baiswar, who has always encouraged me and supported me through every step of my life.

My heartfelt acknowledgement and thanks extended to my relatives, teachers, friends and well wishers of personal and professional arena who have assisted, guided and cooperated directly and indirectly and for constantly encouraging me to go head.

And last but not the least, I proposed my sincere thanks to the author(s) who have contributed a number of research papers, articles and other forms of literature, especially those I Have consulted in my study.

Above all, I express my deep sense of gratitude to " God" the guiding source behind all human efforts , who guide and direct life and who always have been a great source of strength, wisdom and enlightening force behind all efforts in making this thesis a success.

I Finally dedicate the thesis to my beloved parents.

(Rakesh Kumar)

List of Figures

Chapter 3

Fig 3.1 Structure of a single population evolutionary algorithm.

Fig 3.2 Structure of an extended multi population evolutionary algorithm.

Fig 3.3 Presents the corresponding selection pool.

Fig 3.4 Global population model.

Fig 3.5 Local model.

Fig 3.6 Unrestricted migration topology.

Fig 3.7 Scheme for migration of individuals between subpopulation.

Fig 3.8 Ring migration topology.

Fig 3.9 Neighborhood migration topology.

Chapter 7

Fig 7.1 Example of coding

Fig 7.2 Input layer presentation of boy.

Fig7.3 Input layer presentation of runs.

Fig 7.4 Output layer presentation of boy.

Fig 7.5 Sample of genes 1-30.

Fig7.6 Sample of genes 30-60.

Chapter 8

Fig8.1 Fitness value for SYSTEM-A.

Fig 8.2 Fitness value for SYSTEM-C.

Fig8.3 Fitness values for SYSTEM-C under higher mutation values.

Fig 8.4 Fitness values for SYSTEM-B under higher mutation values

Certificate……I

Declaration……II

Abstract……..III

Acknowledgement……..IV

List of figures……..V

Published Papers……VI

Chapter 1: Introduction ……..……2

1.1 Problem area……...4

1.2 Problem statement……..4

1.1 Objective of research……..….4

1.2 Significance of research…….… 5

1.3 Organization of Thesis……..……..6

Chapter 2: Neural Network

2.1 Introduction…….…….. 8

2.2 The basic of neural network……9

2.3 How do Neural network differ from conventional techniques……..…….12

2.4 What application should neural network be used for?.....13

2.5 What are their limitations?.....13

2.6 What are their advantages over conventional techniques?.....14

Chapter 3: Genetic Algorithm

3.1 Introduction……..…….16

3.2 Selection……18

3.3 Recombination……..19

3.4 Mutation……20

3.5 Reinsertion……20

3.6 Population models—parallel implementation of evolutionary algorithm…….21

3.7 Application of multiple/different strategies and competition between subpopulation……..21

3.8 Population model—parallel implementation…….22

3.9 Global model—worker/ farmer…….23

3.10 Local model - Diffusion model……..24

3.11 Regional model—Migration……25

Chapter 4: Word sense disambiguation

4.1 Approaches of word sense disambiguation……..30
4.2 Knowledge based method…….30
4.2.1 Lesk Algorithm……..30
4.2.2 Measure of semantic similarity computed over semantic work……30
4.2.3 Heuristic Method……...…31
4.2.4 Most frequent sense……31
4.2.5 One sense per discourse……..…31
4.2.6 One sense per collocation……..31

4.3 Automatically or semi-automatically acquired……..…31
4.3.1 Unsupervised corpus base method……..31
4.3.2 Supervised corpus based method……32
4.3.3 Semi supervised method…….……32
4.4 Application of word sense disambiguation……..32
4.4.1 Machine translation…….32
4.4.2 Information extraction…….33
4.4.3 Information retrieval……33
a)Cross language information retrieval……33
b)Question answering…….....33
4.5 Word Senses…….35
4.6 Context…….35
4.7 Evaluation……36
4.8 Language Philosophy……37
4.9 Enumeration Vs generation…….38
4.10 The origin of features……39
4.11 Recording word senses……..40
4.12 Frequency distribution……..41
4.13 Acceptability in context……42
4.14 Existing sense inventories…….44

4.14.1 Wordnet…….....44
4.14.2 Finnish Synonym dictionary……46
4.14.3 Cross lingual terminology……47

Chapter 5 : Dimension of context

5.1Introdution……49
5.2 Context size……..49
5.3 Context modality…….50
5.4 Depth of preprocessing……51

Chapter 6: Disambiguation

6.1Inroduction…….55
6.2 Manually developed resources……...55
6.3 Early Ideas…….55
6.4 Artificial Intelligence ideas……56
6.5 Knowledge Based ideas…….57
6.6 Machine learning methods…….57
6.6.1 Supervised……..58
6.6.2 Semi- supervised……58
6.6.3 Unsupervised……..59
6.6.4 Combining classifiers…….59
6.6.5 Degree of supervision…….59

Chapter 7: Methodology and system design

7.1 Introduction……..……62
7.2 Genetic algorithm for neural network optimization…….63
7.3A natural language processing task for neural network……65
7.4 Genetic definition of neural network topology……67
7.5 Schema disruption computation……..70
8.Empirical results……73
9 Conclusion and future work……79
10 References…… 81