An

Abstract

on

Phd. Thesis entitled

STUDY OF EVOLVING NEURAL NETWORK

FOR

WORD SENSE DISAMBIGUATION

By
RAKESH KUMAR
Ph.D. Registration No. : Ph.D/730 /09 dated 11/02/2009

Under The Supervision
Prof. Bipul Syam Purkayastha

DEPARTMENT OF COMPUTER SCIENCE
ALBERT EINSTEIN SCHOOL OF PHYSICAL SCIENCES
ASSAM UNIVERSITY
SILCHAR - 788 011, INDIA
2015

Abstract

The automatic disambiguation of word senses consists in assigning the most appropriate meaning to a polysemous word, i.e., a word with more than one meanings, within a given context.

This consists of two steps:

(i) considering the possible senses of the given word; and

(ii) (ii) assigning each occurrence of the word to its appropriate sense.

We propose a supervised approach to word sense disambiguation based on neural networks (NNs) combined with evolutionary algorithms (EAs). We dispose of large tagged datasets describing the contexts in which every sense of a polysemous word occurs, and use them to evolve an optimized NN that correctly disambiguates the sense of a word given its context. To this aim, we use an EA to automatically design NNs, with two kinds of distributed encoding schemes, based on the way words are written, to represent the context in which a word occurs.

One approach used by researchers trying to develop computer systems capable of understanding natural languages is that of training a neural network (NN) for the task. For this type of approach, one of the key questions becomes how to best configure NN parameters such as topology, learning rates, training data, and other. How to choose values for these parameters is still an open question, especially since the effect of these variables have on each other is not completely understood. Genetic algorithms (GA) are particularly well suited for finding optimal combinations of parameters, since they make no assumption about the problem being solved. Different NN configurations are coded as genomes, which have a fitness function based on how well they an solve a particular task. Genomes are paired and recombined in the hope that the offspring of good solutions will be even better.

Word sense disambiguation is the task of selecting the appropriate senses of a word in a given context. An excellent survey of the history of ideas used in word sense disambiguation is provided by Ide and Veronis (1998). Word sense disambiguation is an intermediate task which is necessary in order to accomplish some other natural language processing task, e.g.,

1) translation selection in machine translation,
2) eliminating irrelevant hits in information retrieval,
3) analyzing the distribution of predefined categories in thematic analysis,
4) part-of-speech tagging, prepositional phrase attachment and parsing space restriction in grammatical analysis,
5) phonetization of words in speech synthesis and homophone discrimination in speech recognition, and
6) spelling correction, case changes and lexical access in text processing.

Word sense disambiguation (WSD) involves the association of a given word in a text or discourse with a definition or meaning which is distinguishable from other meanings potentially attributable to that word. The task therefore necessarily involves two steps according to Ide and Veronis (1998)[1]. The *first* step is to determine all the different senses for every word relevant to the text or discourse under consideration, i.e., to choose a sense inventory, e.g., from the lists of senses in everyday dictionaries, from the synonyms in a thesaurus, or from the translations in a translation dictionary.

The *second* step involves a means to assign the appropriate sense to each occurrence of a word in context. All disambiguation work involves matching the context of an instance of the word to be disambiguated either with information from external knowledge sources or with contexts of previously disambiguated instances of the word. For both of these sources we need preprocessing or knowledge-extraction procedures[2] representing the information as context features.

For some disambiguation tasks, there are already well-known procedures such as morpho-syntactic disambiguation and therefore WSD has largely focused on distinguishing senses among homographs belonging to the same syntactic category.

However, it is useful to recognize that a *third* step is also involved: the computer needs to learn how to associate a word sense with a word in context using either machine learning or manual creation of rules or metrics. It is the third step which is the focus of this work and especially the machine learning aspect. Unless the associations between word senses and context features are given explicitly in the form of rules by a human being, the computer will need to use machine learning techniques to infer the associations from some training material. In order to avoid confusion, we will speak of manually created disambiguation techniques as a separate category and only divide the machine learning techniques into the subcategories of supervised, semi-supervised and unsupervised.

Problem area:

The automatic disambiguation of word senses has been an interest and concern since the earliest days of computer treatment of language in the 1950's. Sense disambiguation is an "intermediate task" (Wilks and Stevenson, 1996) which is not an end in itself, but rather is necessary at one level or another to accomplish most natural language processing tasks. It is obviously essential for language understanding applications such as message understanding, man-machine communication, etc.; it is at least helpful, and in some instances required, for applications whose aim is not language understanding.

Problem statement:

The idea of using simulated evolution to create neural networks that learn faster and generalize better is becoming increasingly widespread.

However, such evolutionary processes are usually extremely computationally intensive. We propose a supervised approach to word sense disambiguation based on neural networks combined with evolutionary algorithms. Large tagged datasets for every sense of a polysemous word are considered, and used to evolve an optimized neural network that correctly disambiguates the sense of the given word considering the context in which it occurs.

Objective of research:

1. Study the word sense disambiguation problem.
2. How the genetic algorithm and neural network can be used for word sense disambiguation.
3. To develop a hybrid method for the word sense disambiguation using neural network and genetic algorithm.

Significance of research:

Word sense disambiguation (WSD) is the task of determining the meaning of an ambiguous word in its context. It is an important problem in natural language processing (NLP) because effective WSD can improve systems for tasks such as information retrieval, machine translation, and summarization.

Neural network approaches to WSD have been suggested. These models consist of networks in which the nodes ("neurons") represent words or concepts, connected by "activatory" links: the words activate the concepts to which they are semantically related, and vice versa. In addition, "lateral" inhibitory links usually interconnect competing senses of a given word. Initially, the nodes corresponding to the words in the sentence to be analyzed are activated. These words activate their neighbours in the next cycle in turn, these neighbours activate their immediate neighbours, and so on. After a number of cycles, the network stabilizes in a state in which one sense for each input word is more activated than the others, using a parallel, analog, relaxation process.