Malaysian Journal of Computer Science, Vol. 15 No. 2, December 2002, pp. 28-33

USING AN EASTERN PHILOSOPHY FOR PROVIDING A THEORETICAL BASIS FOR SOME HEURISTICS USED IN ARTIFICIAL NEURAL NETWORKS

Asoka S. Karunananda
Department of Mathematics & Computer Science
The Open University of Sri Lanka
Nawala
Nugegoda
Sri Lanka
Tel: 94 01 853777, Ext. 225
Fax: 94 01 436858
email: asoka@maths.ou.ac.lk

ABSTRACT

Artificial Neural Networks technology has been recognised as a promising approach to solve a variety of complex problems, which could not be solved otherwise. Despite showing promising results, this technology has often been criticised for the lack of a theoretical basis. This paper presents a theoretical basis for some heuristics used in designing and training of Artificial Neural Networks. Our research exploits Theravada Buddhist theory of mind, which is a well-known Eastern philosophy and also falls under the context of alternative systems of knowledge. The theoretical basis is derived using the concept of thought process in Buddhism. In this sense, firstly, our theory supports the heuristics that the number of neurons in the input layer should be the same as the number of components in an input. The second heuristic, the use of small values as initial weights has been supported. A theoretically based mechanism for initialisations of weights has also been introduced. Further work on the project introduces the novel idea of recursive training of Artificial Neural Networks. Hence, it is concluded that the integration of Western model of Artificial Neural Networks with the Eastern model of thought process in Buddhism will be immensely be beneficial for the development in the field.

Keywords: ***Artificial Neural Networks, Buddhist theory of mind, Thought process, Bhavanga***

1.0 INTRODUCTION

Artificial Neural Networks has been a major application area of Artificial Intelligence (AI). However, Artificial Neural Networks was born several years prior to the birth of Artificial Intelligence. Over the last few decades, Artificial Neural Networks has shown its potential to solve a variety of problems that could not be solved by classical techniques. In particular, applications of Artificial Neural Networks for business forecasting, engineering design, medicine and various industries have been very successful. Although Artificial Neural Networks has generally been considered as a branch of AI, some researchers argue that this is not the case. This is mainly due to the fact that AI does not particularly consider numeric manipulation while numeric manipulation is at the core of Artificial Neural Networks. However, the primary aim of Artificial Neural Networks is to make models of the human brain, which is the theme of AI too. So, Artificial Neural Networks falls easily under AI rather than under any other discipline.

In spite of the fact that Artificial Neural Networks is widely applicable, there is a frequent complaint about this technology. That is, Artificial Neural Networks lacks a theoretical basis in contrast to other disciplines. For example, design of Artificial Neural Networks for solving problems is not guided by an underlying theory. At present, it happens just as a result of heuristic based *ad-hoc* process. This paper reports on research work carried out to provide a theoretical basis for some heuristics used in design and training of Artificial Neural Networks. The research work exploits Buddhist theory of the human mind. As an Eastern philosophy, Buddhist philosophy has won much recognition in the context of alternative systems of knowledge. Therefore, the theoretical basis presented here can be seen as an attempt to merge Western and Eastern thoughts towards a development of a better basis for some parts of research in Artificial Neural Networks. This effort will clearly ensure the advancement of the field of Artificial Neural Networks.

In order to present the research work, the rest of the paper is organised as follows. Section 2 provides an overview of Artificial Neural Network technology. Section 3 describes heuristic-based practices of designing and training of

Neural Networks. Section 4 presents the theory for supporting two major heuristics for designing and training of Neural Networks. This also gives a mechanism for initialising the weights for training of a Neural Network. Section 5 provides a summary and a brief introduction on further work.

2.0 OVERVIEW: NEURAL NETWORKS TECHNOLOGY

Artificial Neural Networks (simply the Neural Networks) is seen as an attempt to model the functionality of the human brain. As Zurada describes, Artificial Neural Networks can be defined as a physical cellular system, which can acquire, store and utilise experiential knowledge [10]. The structure of Artificial Neural Networks is defined as a simple analogy to the human brain. In this sense, an Artificial Neural Network consists of a network of neurones. The neurones are considered as processing units while the connections store processed information in the form of weights connecting neurones. Artificial Neural Networks are designed as several layers of neurones. Training of an Artificial Neural Network begins by assigning initial arbitrary weights for the connections. During the process of training of the neural network, the weights will be adjusted to represent the entire set of the input data. A trained neural network will be represented as a form of a weight matrix. This matrix or matrices will then be used for solving problems. Thus the basic principle behind neural network is that the network can be trained to represent a known set of information and then the trained network can be used for solving problems.

In view of the above, the first phase of applying Artificial Neural Network can be seen as the *training* of the network such that weights should be adjusted to accommodate the whole input set. The training process is dependent on a variety of factors including design of the network, representation of input data and how you change the weights in each cycle. Major portion of the study of neural networks falls under investigation of techniques for weight changing during the process of training. These techniques are known as learning algorithms. Depending on the network structure and the mode of training, there are several training algorithms including Hebb's, Perceptron, Delta learning rule, Backpropagation, Counter propagation, Statistical training, Hopfiled, etc.

Among those, Hebb's learning rule is considered as the first formal training algorithm proposed; it was introduced in 1946. When the structure of a neural network is complicated, the training also becomes difficult. For example, training of multi-layer neural networks is difficult and need more sophisticated training algorithms than Hebbian training. At present, Backpropagation training algorithm is seen as the most successful training algorithm for multi-layer neural networks. It has also noticed that multi-layer neural networks with minimum of three layers can be used to model any real-world problem on the average.

3.0 SOME HEURISTICS USED IN NEURAL NETWORKS

Design and training of neural networks are very much based on some heuristic decisions. Stated in another way, there is no underlying theory to the design of a neural network with the appropriate number of layers, number of neurons in the input layer, selecting the initial weight matrix, selecting a training algorithm, etc. The appropriate decisions with regard to above factors are very much influential in training efficiency and the reliability of the solutions provided by the network.

There are various heuristics used at various levels in designing and training of neural networks. In this sense, two major heuristics used in designing and training of neural networks is listed below.

- **Heuristics 1:** The number of neurons in the input layer can be determined as the number of components required to represent an input.

- **Heuristics 2:** The weight values of initial weight matrix should be as small as possible.

The validity of above heuristics has already been established by means of a variety of experiments in the field of Artificial Neural Networks. Nevertheless, such heuristics are not supported by any theoretical foundation. So, in our research, we explain that the above heuristics based decisions can be supported with a theoretical basis, which exploits the Buddhist theory of mind.

4.0 BUDDHIST THEORY OF MIND

This section provides a brief description of Buddhist philosophy, with an intention of using it to define a theoretical basis for heuristics in Artificial Neural Networks. Buddhism is popular among modern researchers due to its sound theories pertaining to phenomena like behaviour of human mind, concept of time, concept of space, etc. So, Buddhism plays a major role in research in alternative systems of knowledge. Here, the Buddhist theory of mind is exploited not from a religious viewpoint but from a philosophical viewpoint. The reader is reminded not to consider the difference between brain and mind at this level. In the current context we may consider these terms as synonyms. However, the reader is still free to think of the mind as an abstract process. It should be noted that there is no harm to treat the mind as an abstract concept such as an atom. In that context, it is significant that people research into atoms based on various observations but none of them has seen an atom or an electron. Similar line of thinking is valid for the mind too.

According to Buddhism, the thinking process itself is defined as the mind. The mind operates as a sequence of conditional flow of thoughts or "citta". Thus the entire mind can be t -space'. A thought consists of mental factors or emotions, which characterise the nature of the particular thought. A group of thoughts are bundled into a phenomenon called *thought process*. The output of one thought is fed to the pool of thoughts and that contributes to generate the next thought. Thought processes are formed in response to the inputs accepted by the flow of thoughts. It should be noted even without an input, thought process can continue in terms of a particular thought which is called "*bhavanga*". This thought operates during a dreamless sleep. It keeps the continuation of the flow of thoughts or the mind. All thought processes begin and end up with bhavanga. Fig. 1 illustrates the concept of thought process as a collection of thoughts.

Fig. 1: Thoughts and thought process

At this point one major aspect is highlighted according to which the mind behaves with regard to a given input. According to Buddhism, mind has a probabilistic behaviour to a great extent. This fact is exploited for this project as a key result from Buddhism. Events, over which mind acts probabilistically are, in fact, what is known as emotions. For example, anger can be such an emotion pertaining to the input. Happiness is also an emotion. The mind behaves probabilistically in terms of emotions, with regard to an input.

As mentioned earlier, a thought process is generated in response to an input. Once an input has entered to the thought-flow, there will be a sequence of thoughts forming several thought processes before accepting the next input. In other words, due to each input the thought-flow reaches respective limiting states. At these states one can introduce a new input. Limiting states can be continued without any change until we introduce an input. Fig. 2 illustrates thought processes and continuation of limiting states, which are denoted by black spots.

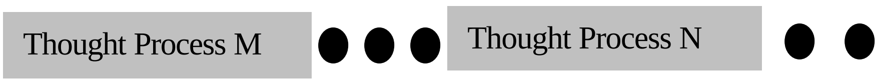

Fig. 2: Thought Process and Limiting states

4.1 A Formal Model for Theory of the Mind

This section formally presents the theory of the mind discussed in the previous section. With the postulate of probabilistic behaviour of the mind the next important aspect is calculating those values of probabilities. Here, Buddhism's interpretation that mind deals with a conditional flow of thoughts is used. Stated in another way, if you are angry at the time of receiving i^{th} input there is a higher probability of being angry when you receive the j^{th} input, where $j = i+1$. Emotional effect of each thought process is stored and gives a cumulative result later when we decide upon an input.

Thus, a theory of the mind with the following axioms is postulated, which are exploited from the Buddhist theory of mind.

Axiom 1: Mind behaves probabilistically.

Axiom 2: Probability of events over the emotions can be represented as a matrix $P = (p_{ij})_{nxn}$

where $\sum_{j=1}^{n} p_{ij} = 1$ and

$p_{ij} < p_{ii}$ for all i = 1,2,…..n and
j = 1,2,……n, n is the number of emotional factors.

Using the above axioms, three main results have been deductively obtained. These results exhibit some characteristics, which explains the behaviour of the human mind. Next the derivation of results using the above axioms is discussed.

Result 1: By axiom 2, it follows that P is a square matrix. Note that both *i* and *j* vary from 1 to *n*

Result 2: Axiom 2 also implies that there is *m* such that $P^m = P^{m+1}$. This result shows that the probabilistic behaviour of human minds has the characteristics of Markov chains in statistics.

Result 3 Result 2 implies that the thought-flow converges into a limiting state with a limiting matrix P^m before accepting the next input. This result agrees with Buddhism, as it says that there will be several thought processes with regard to a given input before accepting the next input.

Result 4: Results 1 and 2 follow that mind can be modelled as a flow of thoughts with Markov chains M_1, M_2, ..., M_{n-1}, $M_{n.}$ Hence the conditional cumulative effect with regard to emotions at a given moment can be obtained from Buddhism. However, this result is not required to explain a theoretical basis for neural network. The result can be used for knowledge modelling. Any interested reader may refer to [1] for further application of Result 4.

The above results are presented in its most general form with *n* emotional factors. In many instances, the model for a specific case when n=3 is evaluated. In this particular case only three emotional factors are considered, namely, happiness, unhappiness and neutrality. Computer modelling and testing of the above theory for n=3 can be found in [1] and [2]. Further, the above theory was applied for Knowledge Modelling for Knowledge-Based Systems. They are presented in [5] and [6].

4.2 Using Theory from Buddhism

This section explains how the above results can be used to provide a theoretical basis for two heuristics used in Artificial Neural Network designing and training. Section 3 has mentioned the two very common heuristics. With the above results, below are the explanations how these two heuristics can be supported by the theory developed by exploiting Buddhism.

4.2.1 Explanation for Heuristics 1

The direct implication of heuristics 1 is that it is preferable to use square matrices as weight matrix of the input layer. Stated in another way, many people have practically noticed that it is convenient to represent the number of components of an input equal to the number of neurones in the input layer. So, this heuristic is supported by Result 1. It is a direct implication of axioms 1 and 2. Therefore, selecting the number of components of an input similar to the number of neurones in the input layer has a theoretical justification from Buddhism. The choice has been attributed to the actual behaviour of human mind according to Buddhism.

4.2.2 Explanation for Heuristics 2

In training the Neural Networks, people generally select very small values for initialising the weight matrices. Without a theoretical support, experiments show that this choice makes efficient convergence of the training process. However, Result 3 in the theory shows that matrix P converges to a limiting matrix with the *smallest* possible values in the matrix. Therefore, using small values in the initial weight matrix also has theoretical justification from Buddhism. This result is attributed to other advantages, which is described in the next section.

4.3 Novel Way for Weight Initialisation

It should be important to mention that integrating Artificial Neural Network concepts with Buddhist theory of mind, provides not only a theoretical justification for some heuristics but also a novel way to find the initial weight matrix. Without doubt, a technique for deciding on the initial weight matrix will be very useful as this is generally done in an *ad-hoc* manner. The approach is extremely simple and described as follows in a two-step process.

Step 1: Define an arbitrary matrix P such that $P = (p_{ij})_{nxn}$

where $\sum_{j=1}^{n} p_{ij} = 1$ and

$p_{ij} < p_{ii}$ for all i = 1,2,…..n and j = 1,2,……n, with n is the number of neurons in the input layer. Note also that the number of components in an input is also considered as *n*. This follows from the heuristics, for which now has been provided a theoretical basis.

Step 2: Compute the limiting matrix L such that $L = P^m = P^{m+1}$. The matrix L, can easily be computed using simple multiplication of matrices. In modelling of human mind, both *m* and *L* are important. However, in this case only L is important and that can be treated as the initial weight matrix W.

Note that L is the matrix with smallest vales, which can be generated using P. At present Artificial Neural Networks technology does not provide any mechanism to generate a smallest set of values using a given set of weights. So, our theoretically based initialisation of weights has an added value.

5.0 SUMMARY AND FURTHER WORK

This paper has presented a theoretical basis for some heuristics used in Neural Networks designing and training. It is provided in response to the usual criticism that Artificial Neural Networks technology has no theoretical basis for some aspects of designing and training. The Buddhist theory of mind is exploited for this purpose. Note that the above theory has originally been developed as an attempt to model human behaviour. There have also been varieties of application of the exploited theory. This paper has shown how the above theory can be used to address some issues in Artificial Neural Networks. Here, it has also presented a theoretical approach to initialisation of weights for training of Artificial Neural Networks. Therefore, this paper has presented two important contributions, namely:

- provision of a theoretical basis for some heuristics used in neural networks
- a method for initialising of weights for training of neural networks.

Future work of the project aims at an area that is neglected in Neural Network technology at the moment. That is, that the current technique in Neural Network has neglected the fact that human brains are liable to get fatigued when the same form of input is applied continuously. Stated in another way, if an input is applied repeatedly, the mind attempts to accept the input at the beginning but loses the interest as time goes on. It should be noted that many concepts in Neural Networks and Buddhist theory of mind have some parallel ideas. For example, according to Buddhism, input cannot enter the thought flow immediately without adjusting the mind for it. This has some similarity with changing weights. Further, the idea of terminating a training session has a meaning with regard to the concept of limiting state of a thought-process.

In this sense, future work will be on the study of how Neural Networks behave beyond the training. In the modern Neural Networks technology, once the network is trained the only option available is the use of it for solving problems. However, Buddhist theory of mind allows us to think of recursive training using the same input. This line of research will be of great interest due to the fact that, although Artificial Neural Networks can give an answer to a problem, there is no justification to say, it is the best answer.

REFERENCES

[1] A. S. Karunananda, "Computer Modelling of the Theravadic Buddhist Theory of Thought Processes - A Model of Human Behaviour". *MPhil Thesis*, The Open University of Sri Lanka, 1992.

[2] A. S. Karunananda, "Computer Modelling of Thought Processes". *International Journal of Computer Application in Technology*, Vol. 6, No 2/3, 1993.

[3] A. S Karunananda, "Investigating the Use of Metaphors for Knowledge Acquisition". *PhD Thesis*, University of Keele, UK, 1995.

[4] A. S. Karunananda, "Buddhist Thought Process: A New Approach to Knowledge Modelling". *Proceedings of 53rd Annual Session of Sri Lanka Association for the Advancement of Science*, December, Colombo, Sri Lanka, 1997, p. 303.

[5] A. S. Karunananda, "Response-Driven Knowledge Acquisition for Knowledge Modelling". *KAW98,* April, 1988, Banff, Alberta, Canada, pp. 18-23.

[6] A. S. Karunananda, "A Theoretical Basis for Weight Initialisation in Neural Networks". *Proceedings of 5th Annual Session of Sri Lanka Association for Advancement of Science*, December, 1998, Sri Lanka, pp. 243, 14-19.

[7] P. Picton, *Introduction to Neural Networks*. MacMillan, 1994.

[8] D. E. Rumelhart and J. L. McClelland, *Parallel Distributed Processing*. The MIT Press, Vol. 1, 1986.

[9] D. Wasserman, *Neural Computing Theory and Practice*. Van Nostrand Reinhold, New York, 1989.

[10] J. M. Zurada, *Introduction to Artificial Neural Systems*. West Publishing Company, New York, 1992.

BIOGRAPHY

Asoka S. Karunananda obtained his Ph.D. degree from University of Keele, UK. Currently, he is working as a Senior Lecturer in Computer Science in the Open University of Sri Lanka. His research areas include Artificial Intelligence, Expert Systems, Genetic Algorithms, Fuzzy Logics, Theory of Computation, Artificial Neural Networks and E-systems. He is very much keen on the use of Eastern system of knowledge for the development of areas in Artificial Intelligence. He is a founder member of Sri Lanka Association for Artificial Intelligence (SLAAI) and also works as the Vice-President of SLAAI. He is a member of Computer Society of Sri Lanka and Sri Lanka Association for Advancement of Science. He has published and presented more than 60 papers locally and internationally.