OPTIMIZATION WITH NEURAL NETWORKS

by

Jayadeva

Department of Electrical Engineering

A THESIS SUBMITTED
IN FULFILMENT OF THE REQUIREMENTS OF
THE DEGREE OF

DOCTOR OF PHILOSOPHY

to the

INDIAN INSTITUTE OF TECHNOLOGY, DELHI

JUNE 1993

This thesis is dedicated to

my late grandmother, my parents, and my teachers

CERTIFICATE

This is to certify that the thesis entitled "Optimization with Neural Networks", which is being submitted by Mr. Jayadeva to the Department of Electrical Engineering, Indian Institute of Technology, Delhi, for the award of the degree of Doctor of Philosophy, is a record of bonafide research work he has carried out under our supervision and guidance, and, in our opinion, it has reached the standard fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted to any other university or institute for the award of a degree or a diploma.

Dr. Basabi Bhaumik

Prof. S.C. Dutta Roy

Department of Electrical Engineering

Indian Institute of Technology

Hauz Khas, New Delhi - 110016.

INDIA

Acknowledgements

It is said that the battle of Kurukshetra was won because of the Arjuna's *sarathi* or charioteer. The role of a Ph.D supervisor is perhaps similar in many ways; he is the one who guides you about where to attack, what course (s!) to take, which traps to avoid, and of course, about when it is time to call it a day. I have been very fortunate to have had, as my supervisors, not one, but two *gurus*, the like of whom are scarce today. I will forever be in debt to Prof. S.C. Dutta Roy and Dr. Basabi Bhaumik for their encouragement, guidance, interest, advice, and support, and for their understanding and help when I needed them; it is a debt I am privileged to have received, and I hope that I shall continue to have their blessings in the future.

I am grateful to Prof. Dutta Roy for having motivated, inspired, and helped me to stay on and pursue a doctoral programme here; and to Dr. Basabi Bhaumik and Prof. Dutta Roy for having given so much of their time and effort. I owe a debt of gratitude to Dr. Bhaumik for the numerous valuable discussions, which also made research enjoyable and worthwhile.

I would like to express my gratitude to Prof. Surendra Prasad for his generosity, help and advice; to Profs. P.C.P. Bhatt, Suresh Chandra and P.K. Bhattacharyya for their help and for interesting discussions, and to Dr. G.S. Visweswaran, Prof. D. Nagchoudhuri, and Prof. S.S. Jamuar for their help at various times. I am particularly thankful to Mr. Vikram Gadre for his help, especially when this manuscript was being prepared. Thanks are also due to Dr. Pankaj Sharan for his cups of tea and his witty conversations, which were very welcome at the end of a tiring day; also to Mr. M.V.B. Sastry, Dr. Akhila Sinha, Mr. Paul Francis and Mr. A. Varshney. I would also like to acknowledge many other colleagues and associates with whom I interacted during my Ph.D, including Dr. Saeid Noshabadi, Mr. K.C. Sharma, Mr. Kaushik Saha, Mr. C.M. Markan, and Mr. Rajiv Jain.

I am grateful to Dr. J.E. Beasley of the Imperial College, London, for providing the test examples used in Chapters 3 and 4; to Dr. D.J. Willshaw and Mr. M. Simmen for providing the data for the 30 and 100-point example used therein; and to so many others who responded to my reprint requests.

I owe a debt of gratitude to my parents and my sister, for their understanding, support, love and patience, without which this thesis could not have seen the light of day.

Finally, I ask the forgiveness of those whom I omitted to mention in the haste of submitting my thesis.

Jayadeva
(JAYADEVA)

ABSTRACT

The recent resurgence of interest in neural networks, is in a substantial measure, due to the recognition of the fact that self-organization of dynamical systems can be used to compute the solutions of many problems. Hopfield, in his papers in 1982 and 1984, showed that a network of simple interconnected elements (called the Hopfield net) could be used to solve an optimization problem. Since then, researchers have used neural nets for solving a large number of applications from diverse fields such as control engineering, signal processing and combinatorial optimization, by formulating various tasks as optimization problems. The network generally chosen for such purposes is the Hopfield net.

There are, however, difficulties with this computational paradigm. Frequently, the network converges to an invalid or spurious solution. Convergence may require several thousand iterations. The large number of interconnection weights required for many applications makes implementation in hardware difficult and simulation on computers expensive. The interconnection complexity often severely restricts the size of the problem that can be solved.

This thesis is concerned with the problem of optimization using neural networks. We show that most neural optimization approaches perform only one iteration of a Sequential Unconstrained Minimization Technique (SUMT); this may not be sufficient to obtain a valid solution. We suggest a remedial measure using SUMTs, in which the interconnection weights change with time. This is in contrast to most neural optimization methods, which use fixed weights. Through simulations, we show that choosing different initial

values for the weights leads to different solutions; the results indicate the existence of a space-time tradeoff in neural optimization, i.e. better solutions can be found at the expense of additional computation. We use this observation to develop a heuristic method for improving the quality of solutions.

Though the Hopfield net can be used for a very wide range of applications, other models may be more appropriate for solving a specific category of problems. We propose a class of neural networks, to be called Internet in the sequel, which is particularly suitable for interconnection problems. We specifically apply the Internet to the Euclidean Steiner Tree, the Rectilinear Steiner Tree, and the Steiner Circuit problems. To the best of our knowledge, the proposed neural networks are the first to be reported in the literature. The solutions found by these Internets for examples from the literature compare very favourably with the best known or optimal ones.

Most neural networks perform a steepest descent to find a minimum of the cost or energy function. As a consequence, the initial state of the network considerably influences the solution found. We show that the Internet has the ability to escape from local minima, which enables it to explore many different topologies in a network design problem, and reduces the significance of the network's starting configuration. The mechanism of escaping from a local minimum is based on gradually altering the curvature of the energy surface at such a point, so that it becomes a local maximum. Such a technique is very different from that used in probabilistic search methods like simulated annealing.

We use SUMTs to extend the Internet to solve the Travelling Salesman Problem (TSP). It is shown that the Elastic Net algorithm for the TSP can be obtained using SUMTs. A comparison between the TSP Internet and the

Elastic Net is made by simulating both networks over a set of randomly generated examples; the results show that in comparison to the Elastic Net, the TSP Internet converges to a valid solution in more instances and requires fewer iterations to do so. It also finds better solutions in many cases. In the sequel, we also briefly discuss how subjective contours in vision can be determined using a variation of the Internet.

While continuous-time networks have generally been used for neural optimization, binary networks can sometimes offer many advantages over them. For example, they may be easier to implement and less expensive to simulate on digital computers. One of the earliest binary nets to be proposed in the literature is the feedforward, single layer perceptron. It can represent simple geometric structures like lines and hyperplanes. We use such a network to find the Convex Hull of a set of points in a plane; it is then used in a scheme to solve the TSP.

We propose a binary network, to be called the Modified Hopfield Network (MHOP) in the sequel. We present a MHOP Analog-to-Digital Converter (ADC) and compare it with Hopfield's ADC, which is based on a continuous-time network. The MHOP ADC does not converge to spurious solutions, and requires an order of magnitude fewer weights than Hopfield's ADC. We next study some relations between a MHOP and its continuous-time equivalent, and illustrate, with the help of an example, that a MHOP can perform the same task as its continuous-time equivalent. However, the dynamics of a MHOP is easier to analyze because its state evolves in discrete steps, and it is less expensive to simulate on a digital computer.

Finally, we present an overview of the thesis and suggest some problems for further research.

TABLE OF CONTENTS

CHAPTER 3 Internet - A Class of Neural Networks and its Applications

CHAPTER 4 Further Applications of the Internet