

Hence

$$\lim_{k \rightarrow \infty} \sum_{j=0}^k \alpha_j \prod_{i=j+1}^k (1 - \alpha_i) = 1 - \prod_{i=0}^{\infty} (1 - \alpha_i).$$

Therefore, by (7), $\lim_{k \rightarrow \infty} W_{k+1}$ exists. \square

III. CONCLUSIONS

In this paper we briefly discussed the celebrated Widrow-Hoff algorithm and concluded that this algorithm is robust. However, the weight vectors do not necessarily converge in the presence of measurement noise. We proposed a modified version of this algorithm in which the reduction factors are allowed to vary with time and showed that the algorithm is robust and the weight vectors converge in the presence of bounded noise. We used in our analysis only deterministic-type arguments. We also obtained an ultimate bound on the error in terms of a convex combination of the initial error and the bound on the noise.

REFERENCES

- [1] R. G. Bartle, *The Elements of Real Analysis*, 2nd ed. New York: Wiley, 1976.
- [2] S. Haykin, *Introduction to Adaptive Filters*. New York: Macmillan, 1984.
- [3] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Magazine*, vol. 4, no. 2, pp. 4-22, Apr. 1987.
- [4] D. E. Rumelhart, J. L. McClelland, and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1, *Foundations* (A Bradford book). Cambridge, MA: MIT Press, 1986.
- [5] B. Widrow and M. A. Lehr, "30 years of adaptive neural networks: Perceptron, Madaline, and backpropagation," *Proc. IEEE*, vol. 78, pp. 1415-1442, Sept. 1990.
- [6] B. Widrow and R. Winter, "Neural nets for adaptive filtering and adaptive pattern recognition," *IEEE Computer*, vol. 21, pp. 25-39, Mar. 1988.

Technical Comments

Comments on "Parallel Algorithms for Finding a Near-Maximum Independent Set of a Circle Graph"

Evan W. Steeg

Readers of the above article¹ by Takefuji *et al.* in the September 1990 issue may be interested to learn of earlier work on parallel distributed algorithms for RNA secondary structure prediction. In my M.Sc. thesis [1], [2], completed at the University of Toronto in January 1989 and presented at the Fourth International Symposium on Biological and Artificial Intelligence Systems in Trento, Italy, in September 1988, I describe a model and experiments very similar to those of Takefuji *et al.*

Manuscript received September 4, 1990.

The author is with the Department of Computer Science, University of Toronto, Toronto, Ontario, Canada M5S 1A4.

IEEE Log Number 9042023.

¹Y. Takefuji, L.-L. Chen, K.-C. Lee, and J. Huffman, *IEEE Trans. Neural Networks*, vol. 1, pp. 263-267, Sept. 1990.

In my work I define a mapping of the RNA secondary structure prediction problem onto mean field theory neural networks [3] and specify a way for the networks to improve their structure prediction performance over time by learning the sequence-to-structure mapping as represented in a training set. Experiments in training the networks on small tRNA subsequences led to networks that performed comparably to those described by Takefuji *et al.*: very accurate structure prediction for the small RNA molecules tested, with convergence within several hundred iterations.

Some important similarities and differences between the two approaches are:

- The mappings of the problem onto neural networks are similar. Both approaches are based, more or less directly, on the RNA secondary structure matrix, which is a two-dimensional binary vector representation of base-pairing in RNA molecules. In both approaches the classical "Tinoco" rules [4] for determining the thermodynamically optimal structure are assumed, i.e., favor the formation of "stem structures" (sequences of adjacent edges in the circular graph) and remove "knot structures" (edges that intersect with other edges). Takefuji *et al.* place more emphasis on the circular graph as the basis for their representation. Both their representation and mine require $O(n)$ units, where n is the number of possible base pairs. (However, a molecular biologist would more likely think of this as $O(m^2)$, where m is the length of the RNA sequence).
- The method outlined by Takefuji *et al.* is based on McCulloch-Pitts neurons, and there is no provision made for training the networks. My method is based on the use of MFT networks and the selection of four adaptable parameters, instead of hundreds or thousands of individual connection weights, that make up the small optimization space for the training process.
- Both models can be implemented in parallel hardware. The MFT network may also be implemented in analog hardware, as described by [5].

In the search for fast and accurate RNA structure prediction methods, three things are paramount. First is the need for expert knowledge, theoretical and empirical, to be built into the prediction algorithms [6]-[8]. There are subtle relationships between free energies of particular base pairings (Watson-Crick and others) and other substructures, and a purely graph-theory algorithm for maximizing stacked base pairs and removing knots will not work on longer RNA molecules. Second is the need for parallelism. Some of the classical approaches to sequence comparison may be recast easily into parallel algorithms [7], leading, for example, to $O(km)$ algorithms on $O(k)$ -processor machines. Neural networks approaches, also, of course, offer hope of fine-grained parallelism implemented in efficient hardware. Third, machine learning techniques should be exploited, to refine and extend the representations of expert knowledge on the basis of trial and error over large amounts of real data. Takefuji *et al.* outline an elegant way to map the problem onto neural architectures. If such mappings can be augmented with empirical knowledge (e.g., free energy values of base pairs and substructures) and the ability to learn, as demonstrated in my work, then extremely fast and accurate structure prediction may be within reach.

REFERENCES

- [1] E. W. Steeg, "Neural network algorithms for RNA secondary structure prediction," Master's thesis, Computer Science Dept., University of Toronto, Toronto, Canada, 1989.
- [2] E. W. Steeg, "Neural network algorithms for RNA secondary structure prediction," Tech. Rep. CRG-TR-90-4, Computer Science Dept., University of Toronto, Toronto, Canada, 1990.

- [3] C. Peterson and J. R. Anderson, "A mean field theory learning algorithm for neural networks," MCC Tech. Rep. EI-259-87, Microelectronics and Computer Technology Corporation, Austin, TX, Aug. 1987.
- [4] I. Tinoco, O. C. Uhlenbeck, and M. D. Levine, "Estimation of secondary structure in ribonucleic acids," *Nature*, pp. 230-362, 1971.
- [5] J. J. Hopfield and D. W. Tank, "Neural computation of decisions in optimization problems," *Biol. Cybern.* vol. 52, pp. 141-152, 1985.
- [6] M. Zuker and P. Stiegler, "Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information," *Nucleic Acids Research*, vol. 9, pp. 133-148, 1981.
- [7] D. Sankoff, J. B. Kruskal, S. Mainville, and R. J. Cedergren, "Fast algorithms to determine RNA secondary structures containing multiple loops," in *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*, D. Sankoff and J. B. Kruskal, Eds. Reading, MA: Addison-Wesley, 1983.
- [8] H. M. Martinez, "An RNA folding rule," *Nucleic Acids Research*, vol. 12, pp. 323-334, 1984.

Authors' Reply

Y. Takefuji and K.-C. Lee

Steeg mapped the RNA secondary prediction problem onto a Hopfield neural network [1]. The mapping procedure is similar to ours. The major difference is in the Lyapunov energy function, where it plays a role for solving the problem. We have also presented another new mapping procedure for the RNA prediction problem, where $n(n-1)/2$ processing elements are used for a sequence of n bases [2]. In [1], the energy function has four parameters, α , β , γ and μ , which can be adjusted during learning based on the mean field theory. Steeg made an assumption that there are a few or fewer than ten parameters which are obtainable through learning from examples. He claimed that the optimization space with these four adaptable parameters is smaller than that of individual connection weights. Remember that these four parameters actually represent the individual connection weights.

Consider two simple Lyapunov energy functions:

$$E = \frac{1}{2} \sum_x \sum_i \sum_{j \neq i} A(x, i, j) V_{xi} V_{xj} \quad (1)$$

$$E = \frac{A}{2} \sum_x \sum_i \sum_{j \neq i} V_{xi} V_{xj}. \quad (2)$$

For learning, if the network has the maximum connectivity, the Lyapunov function will have many variables $A(x, i, j)$, where the number of variables will be determined by the ranges of x , i , and j . In (1), the weight strength is given by $T_{xi, yj} = -A(x, i, j) \delta_{xy}(1 - \delta_{ij})$, where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. To reduce the optimization space, the Lyapunov function is given by (2), where it has only one variable, A . The weight strength is given by $T_{xi, yj} = -A \delta_{xy}(1 - \delta_{ij})$, where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise. Our question with regard to Steeg's learning scheme is, Are four adaptable parameters enough to satisfy general learning in predicting the RNA secondary structure?

Manuscript received October 1, 1990.

The authors are with the Department of Electrical Engineering and Applied Physics, Case Western Reserve University, Cleveland, OH 44106. IEEE Log Number 9042024.

The convergence speed for learning might be faster if the learning space were reduced. However, it is not yet known how many variable parameters are required to guarantee the satisfactory learning. In (1) of our paper, two coefficients, A and B , are also trainable if the same learning scheme [1] is used. Although the ability to learn is quite attractive, the determination of the Lyapunov function is more important. We believe that the thermodynamic model is the most reliable and consistent for predicting the RNA secondary structure. In other words, the learning capability is not necessary as long as important thermodynamic properties are embedded or exploited in the model. Our model, based on a part of Tinoco's model, is still lacking in several considerations, for example, the different strengths of G-C and A-U bonds, for example 2.4 kcal between G and C and 1.2 kcal between A and U. In another thermodynamic model of ours [2], the different strengths of G-C and A-U bonds and a hairpin loop constraint are considered. Instability forces, including interior looping, bulges, and other hairpin loops, must be also considered in the Lyapunov function of future neural network models.

We can conclude that the necessity of the learning capability for the RNA secondary structure prediction is questionable. We believe that the task is to build a robust parallel algorithm considering more thermodynamic properties in the model.

REFERENCES

- [1] E. W. Steeg, "Neural network algorithms for RNA secondary structure prediction," Master's thesis, Computer Science Dept., University of Toronto, Toronto, Canada, 1989.
- [2] Y. Takefuji, C. W. Lin, and K. C. Lee, "A parallel algorithm for estimating the secondary structure in ribonucleic acids," *Biol. Cybern.* vol. 63, no. 5, pp. 337-340, 1990.

Corrections to "Sufficient Condition for Convergence of a Relaxation Algorithm in Actual Single-Layer Neural Networks"

J. M. Zurada and W. Shen

Because of a printer's error, the wrong captions were published for the two figures in the above paper.¹ The captions that should have appeared are as follows:

Fig. 1. Two-bit A/D converter used for example convergence evaluation.

Fig. 2. Relaxation algorithm for network from Fig. 1 ($x = 1.3$): (a) stable solution for $\lambda = 3.5$; (b) unstable solution for $\lambda = 5$.

Manuscript received January 7, 1991.

The authors are with the Department of Electrical Engineering, University of Louisville, Louisville, KY 40292.

IEEE Log Number 9143146.

¹J. M. Zurada and W. Shen, *IEEE Trans. Neural Networks*, vol. 1, pp. 300-303, Dec. 1990.