Activity-conserving dynamics for optimisation networks

Postma, E.O.; van den Herik, H.J.; Hudson, P.T.W.

Published in:
Journal of Artificial Neural Networks

Publication date:
1995

Citation for published version (APA):
Postma, E. O., van den Herik, H. J., & Hudson, P. T. W. (1995). Activity-conserving dynamics for optimisation networks. *Journal of Artificial Neural Networks*, *2*(4), 437-442.

Activity-Conserving Dynamics for Optimisation Networks

Eric O. Postma
H. Jaap van den Herik
Patrick T.W. Hudson
Department of Computer Science
University of Limburg, Maastricht, The Netherlands

A novel type of dynamics conserving activity in neural networks is presented. We distinguish stochastic and deterministic activity-conserving dynamics. As an example, deterministic (mean-field) activity-conserving dynamics is applied to the N-queens problem. The new dynamics are more succesful in finding valid solutions than the standard dynamics.

1 INTRODUCTION

Conservation of some quantity is a common constraint in optimisation problems. For instance, in valid solutions of the Travelling Sales-Person (TSP) problem, the number of city-stop assignments equals the number of cities in the tour. In optimisation neural networks, assignments may be represented by active neurons, so that conservation constraints translate into conservation of activation. In their optimisation network for solving the TSP problem, Hopfield and Tank [2] incorporated the following conservation term E_c in the energy function:

$$E_c(\mathbf{a}) = \frac{\alpha}{2}\left(\sum_c \sum_s a_{cs} - A\right)^2, \tag{1}$$

where $\mathbf{a}$ is the activation vector with elements $a_{cs} \in [0,1]$ ($c, s \in \{1, 2, ..., A\}$). An active neuron, $a_{cs} = 1$, represents the assignment of city c to stop s. The positive parameter α weights the conservation constraint relative to other constraints.

[0]The authors thank Jos Uiterwijk for his help and discussions.
Correspondence and requests for reprints should be sent to Eric Postma, Department of Computer Science, University of Limburg, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Email: postma@cs.rulimburg.nl.

Minimisation of the conservation term leads to a network state in which the total activity ($\sum_c \sum_s a_{cs}$) is equal to A.

The conservation term (1) is meant to represent *soft enforcement* of the conservation constraint (cf. [10]). Network states violating the constraint are discouraged, but nevertheless possible. An alternative approach is to employ *strong enforcement*, i.e., only network states that obey the constraint are admitted. Peterson and Söderberg [7] defined dynamics for A-state (Potts) neurons which always satisfy the condition $\sum_i^A a_i = 1$. They defined the TSP conservation term as

$$E_c(\mathbf{a}) = \frac{\alpha}{2} \sum_c^A \left(\sum_s^A a_{cs} - 1 \right)^2 , \tag{2}$$

and used a Potts neuron satisfying $\sum_s a_{cs} = 1$ for each city c. Consequently, the conservation term (1) is equal to zero and can be discarded from the energy function. Optimisation networks with Potts neurons have been shown to perform very well on TSP problems as well as on other problems involving a conservation constraint. Unfortunately, Peterson and Söderberg's method [7] can only be applied when the conservation constraint can be expressed as (2). We present an alternative strong-enforcement method that does not suffer from such a limitation. Central to our method is the use of activity-conserving (AC) dynamics based on Kawasaki's updating scheme from statistical mechanics [4]. Under AC dynamics only transitions that leave the total activity in the network invariant are allowed. While the initial state of the network fixes the activity, our method allows for strong enforcement of the conservation constraint for any possible value of the total activity.

In this note we present stochastic and deterministic AC dynamics. In Section 2, we present the stochastic AC dynamics. Section 3 introduces the deterministic (mean-field) AC dynamics. In Section 4, its effectiveness is illustrated by its application to the N-queens problem. The results obtained with deterministic AC dynamics are succesfully compared with the results of standard dynamics as reported in the literature. Finally, Section 5 concludes on dynamics and complexity.

2 STOCHASTIC ACTIVITY-CONSERVING DYNAMICS

In stochastic Glauber updating [1], a neuron's state is *flipped* (from 0 to 1 or vice versa) with a probability proportional to the associated decrease in energy. In AC dynamics the states of two neurons are *exchanged* (from 0 1 to 1 0 or vice versa) also with a probability proportional to the associated decrease in energy. More specifically, given two neurons, i and j, one active and the other inactive, their states are exchanged with probability (cf. [4])

$$P(n_i \leftrightarrow n_j) = \frac{1}{2} \left[\tanh \left(\frac{(2n_i - 1)\Delta E(n_i \leftrightarrow n_j)}{T} \right) + 1 \right] . \tag{3}$$

In this equation, $n_i \in \{0, 1\}$ is the binary state variable of neuron i; and $\Delta E(n_i \leftrightarrow n_j)$ is the change in energy associated with the exchange of the states of neurons i and j. T is the temperature parameter ($T > 0$).

The stochastic dynamics may be combined with simulated annealing [5] in order to solve optimisation problems. In applying this combination to a molecular-configuration problem, we obtained results competitive with traditional algorithmic approaches [8]. Analogous to the mean-field formulation of Glauber neural-network dynamics [3], below we present a mean-field formulation of Kawasaki dynamics in neural networks.

3 DETERMINISTIC ACTIVITY-CONSERVING DYNAMICS

Mean-field Kawasaki dynamics were derived by Penrose [6] in the context of modelling the dynamics of Ising spins. We reformulate his results in neural-network terms and then apply the mean-field equations so obtained to an optimisation task. In the mean-field approximation to stochastic dynamics, continuous activation variables are defined as the average values of stochastic variables. In our case, the continuous variable a_i is defined as $a_i = \langle n_i \rangle$, where the brackets represent the average over time. For a standard neural network (e.g., [3]) with activation variables $a_i \in [0, 1]$ and weights $w_{ij} = w_{ji}$ and $w_{ii} = 0$, for all $i, j \in K$, mean-field AC dynamics are defined as (cf. [6]):

$$\frac{d}{dt} a_i = \frac{1}{2} \sum_j \left[(-2a_i a_j + a_i + a_j) \tanh \left(\frac{1}{T} \sum_{k \neq i,j} (w_{ik} - w_{jk}) a_k \right) \right]. \qquad (4)$$

For (4), Penrose formulated a Lyapunov function for the special case that all non-zero weights have the same value.

In contrast to the Glauber mean-field dynamics, it is not immediately clear how to realise deterministic AC dynamics, which involves the mutual exchange of activation, within a neural network. Any implementation allowing such exchanges and obeying the conservation constraint can be described by (4). However, the implementation of deterministic AC dynamics in a specific neural network requires a definition of how individual neural elements communicate. Fortunately, it is not necessary to detail the communication in order to evaluate the performance of a system with AC dynamics as compared, for instance, to one using standard dynamics. Below, we evaluate our deterministic AC dynamics using the N-queens problem.

4 APPLICATION TO THE N-QUEENS PROBLEM

The N-queens problem entails the placement of N queens on an $N \times N$ chess-board in a way that no two queens are placed on the same row, column, or diagonal.

Shagrir [9] formulated the N-queens energy function

$$
\begin{aligned}
E(\mathbf{a}) \quad &= \quad \sum_{xy} \left(\sum_{i \neq x} a_{xy} a_{iy} + \sum_{j \neq y} a_{xy} a_{xj} + \sum_{m \neq 0} a_{xy} a_{x+m,y+m} + \sum_{m \neq 0} a_{xy} a_{x+m,y-m} \right) \\
&\quad + \alpha \left(\sum_{xy} a_{xy} - N \right)^2 , \qquad (5)
\end{aligned}
$$

where a_{xy} is the activation of the neuron at position (x, y) on the $N \times N$ chessboard $(x, y, i, j \in \{1, 2, ..., N\})$. An active neuron $(a_{xy} = 1)$ represents the placement of a queen on position (x, y). The first right-hand term penalizes placement of two (or more) queens on the same row, column, left-upward diagonal and right-upward diagonals, respectively. The last term conserves the total activity (i.e., the number of queens) to N. With proper initial values, this term can be discarded when employing AC dynamics.

We tested deterministic AC dynamics on the N-queens problems for $N = 4$ to $N = 9$. At $t = 0$, we set the activations to $a_{xy} = 1/N$. In order to break the symmetry of the initial state, we superimposed small random noise within the range $[-0.01, +0.01]$ on the initial activations, while keeping the total activity fixed at N. During a single simulation run, the mean-field equations are integrated for 50 iterations with the temperature fixed at $T = 0.001$. (To allow for comparison with the results of standard dynamics, no mean-field annealing was employed.) Each run was repeated 100 times with different random initial values. In table 1, the results are listed together with results obtained by standard deterministic dynamics in a Hopfield and Tank network (taken from [9]). (The total number of solutions is also listed. For $N = 6$, it can be seen that the performance drops because there are relatively few valid solutions.) It is observed that AC dynamics are more succesful in finding valid solutions than the standard dynamics.

Table 1: Proportions valid solutions for AC dynamics and standard dynamics.

N	# solutions	*AC dynamics*	*standard dynamics*
4	2	1.00	0.18 (600 runs)
5	10	0.98	0.20 (600 runs)
6	4	0.10	0.03 (300 runs)
7	40	0.54	-
8	92	0.60	0.03 (300 runs)
9	352	0.45	-

5 CONCLUSION

AC dynamics offers a viable alternative to standard neural dynamics. For optimisation problems involving a conservation constraint the proposed dynamics reduce problem complexity considerably, through strong enforcement of the constraint. In those cases where the conservation quantity cannot be reduced to one, AC dynamics may serve as an alternative to Potts dynamics. Where they do, there should be a systematic comparison of the two dynamics to assess their relative advantages.

References

[1] R.J. Glauber. Time-dependent statistics of the Ising model. *Journal of Physics*, 4: 294–307, 1963.

[2] J.J. Hopfield and D.W. Tank. "Neural" computation of decisions in optimization problems. *Biological Cybernetics*, 52: 141–152, 1985.

[3] J.J. Hopfield. Neurons with graded response have collective computational properties like those of a two-state neuron. *Proceedings of the National Academy of Sciences, U.S.A.*, 81: 3088–3092, 1984.

[4] K. Kawasaki. Kinetics of Ising models. In C. Domb and M.S. Green, editors, *Phase transitions and critical phenomena, Volume 2*, pages 443–501. Academic Press, London, 1972.

[5] S. Kirkpatrick, C.D. Gelatt Jr., and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220: 671–680, 1983.

[6] O. Penrose. A mean-field equation of motion for the dynamic Ising model. *Journal of Statistical Physics*, 63: 975–986, 1991.

[7] C. Peterson and B. Söderberg. A new method for mapping optimization problems onto neural networks. *International Journal of Neural Systems*, 1: 3–22, 1989.

[8] E.O. Postma, H.J. Van den Herik, and P.T.W. Hudson. Activity-conserving dynamics for neural networks. In S. Gielen and B. Kappen, editors, *Proceedings of the International Conference on Artificial Neural Networks, ICANN'93* pages 539–544. Springer-Verlag, London, 1993.

[9] O. Shagrir. A neural net with self-inhibiting units for the N-queens problem. *International Journal of Neural Systems*, 3: 249–252, 1992.

[10] P.D. Simic. Statistical mechanics as the underlying theory of 'elastic' and 'neural' optimisation. *Network*, 1: 89–103, 1990.