Critical Temperatures for Intermittent Search in Self-Organizing Neural Networks

Peter Tiňo[1]

School Of Computer Science, University Of Birmingham
Birmingham B15 2TT, UK

Abstract. Kwok and Smith [1] recently proposed a new kind of optimization dynamics using self-organizing neural networks (SONN) driven by softmax weight renormalization. Such dynamics is capable of powerful intermittent search for high-quality solutions in difficult assignment optimization problems. However, the search is sensitive to temperature setting in the softmax renormalization step of the SONN algorithm. It has been hypothesized that the optimal temperature setting corresponds to symmetry breaking bifurcation of equilibria of the renormalization step, when viewed as an autonomous dynamical system called iterative softmax (ISM). We rigorously analyze equilibria of ISM by determining their number, position and stability types. Moreover, we offer *analytical* approximations to the critical symmetry breaking bifurcation temperatures that are in good agreement with those found by numerical investigations. So far the critical temperatures have been determined only via numerical simulations. On a set of N-queens problems for a wide range of problem sizes N, the analytically determined critical temperatures predict the optimal working temperatures for SONN intermittent search very well.

1 Introduction

Since the pioneering work of Hopfield [2], adaptation of neural computation techniques to solving difficult combinatorial optimization problems has proved useful in numerous application domains [3]. In particular, a self-organizing neural network (SONN) was proposed as a general methodology for solving 0-1 assignment problems in [4]. The methodology has been successfully applied in a wide variety of applications, from assembly line sequencing to frequency assignment in mobile communications (see e.g. [5]).

Searching for 0-1 solutions in general assignment optimization problems can be made more effective when performed in a continuous domain, with values in the interval $(0, 1)$ representing partial (soft) assignments. Typically the softmax function is employed to ensure that elements within a set of positive parameters sum up to one. When endowed with a physics-based Boltzmann distribution interpretation, the softmax function contains a free parameter (temperature, or inverse temperature). As the system cools down, the assignments become more and more crisp.

Recently, interesting observations have been made regarding the appropriate values of the temperature parameter when solving assignment problems with SONN endowed with softmax renormalization of the weight parameters [1]. There is a critical temperature T_* at which SONN is capable of powerful intermittent search through a multitude of high quality solutions represented as meta-stable states of the SONN dynamics. It has been suggested that the critical temperature may be closely related to the symmetry breaking bifurcation of equilibria in the autonomous softmax dynamics [6]. Kwok and Smith numerically studied global dynamical properties of SONN in the intermittent search mode and argued that such models display characteristics of systems at the edge-of chaos [7].

In this contribution we attempt to shed more light on the phenomenon of critical temperatures and intermittent search in SONN. In particular, since the critical temperature is closely related to bifurcations of equilibria in autonomous iterative softmax systems (ISM), we rigorously analyze the number, position and stability types of fixed points of ISM. Moreover, we offer analytical approximations to the critical temperature, as a function of ISM dimensionality. So far, critical temperatures have been determined only via numerical simulations. Due to space limitations, we cannot prove fully all the statements presented in this study. Detailed proofs can be found in [8].

2 SONN with softmax weight renormalization

Consider a finite set of elements $j \in \mathcal{J} = \{1, 2, ..., N\}$ that need to be assigned to elements $i \in \mathcal{I} = \{1, 2, ..., M\}$, so that a global cost (potential) $Q(\mathcal{A})$ of an assignment $\mathcal{A} : \mathcal{J} \rightarrow \mathcal{I}$ is minimized. Partial cost of assigning $j \in \mathcal{J}$ to $i \in \mathcal{I}$ is denoted by $V(i, j)$. The "strength" of assigning j to i is represented by the "assignment weight" $w_{i,j} \in (0, 1)$. The SONN consists of the following steps:

1. Initialize assignment weights $w_{i,j}$, $j \in \mathcal{J}$, $i \in \mathcal{I}$, to random values around 0.5.
2. Choose at random an input item $j_c \in \mathcal{J}$ and calculate partial costs $V(i, j_c)$, $i \in \mathcal{I}$, of all possible assignments of j_c.
3. The "winner" element $i(j_c) \in \mathcal{I}$ is the one that minimizes $V(i, j_c)$, i.e. $i(j_c) = \mathrm{argmin}_{i \in \mathcal{I}} V(i, j_c)$.
   The "neighborhood" $\mathcal{B}_L(i(j_c))$ of size L of the element $i(j_c)$ consist of L elements $i \neq i(j_c)$ that yield the smallest partial costs $V(i, j_c)$.
4. Assignment weights of nodes $i \in \mathcal{B}_L(i(j_c))$ get strengthened, those outside $\mathcal{B}_L(i(j_c))$ are left unchanged:

   $$w_{i,j_c} \leftarrow w_{i,j_c} + \eta(i)(1 - w_{i,j_c}), \quad i \in \mathcal{B}_L(i(j_c)),$$

   where

   $$\eta(i) = \exp\left\{ -\frac{|V(i(j_c), j_c) - V(i, j_c)|}{|V(k(j_c), j_c) - V(i, j_c)|} \right\},$$

   and $k(j_c) = \mathrm{argmax}_{i \in \mathcal{I}} V(i, j_c)$.

5. Weights $\mathbf{w}_i = (w_{i,1}, w_{i,2}, ..., w_{i,N})^T$ for each element $i \in \mathcal{I}$ are normalized using softmax
$$w_{i,j} \leftarrow \frac{\exp(\frac{w_{i,j}}{T})}{\sum_{k=1}^{N} \exp(\frac{w_{i,k}}{T})}.$$
6. Repeat from step 2 until all elements $j \in \mathcal{J}$ have been selected (one epoch).

Even though the soft assignments $w_{i,j}$ evolve in continuous space, when needed, a 0-1 assignment solution can be produced by imposing $\mathcal{A}(j) = i$ if and only if $j = \text{argmax}_{k \in \mathcal{J}} w_{i,k}$.

A frequently studied (NP-hard) assignment optimization problem in case of SOP is the N-queen problem: place N queens onto an $N \times N$ chessboard without attacking each other. In this case $\mathcal{J} = \{1, 2, ..., N\}$ and $\mathcal{I} = \{1, 2, ..., N\}$ index the columns and rows, respectively, of the chessboard. Partial cost $V(i, j)$ evaluates the diagonal and column contributions[1] to the global cost Q of placing a queen on column j of row i (see [1] for more details).

Kwok and Smith [1] argue that step 5 of the SONN algorithm is crucial for intermittent search by SONN for globally optimal assignment solutions. In particular, they note that temperatures at which symmetry breaking bifurcation of equilibria of the renormalization procedure in step 5 occurs correspond to temperatures at which optimal (both in terms of quality and quantity of found solutions) intermittent search takes place.

3 Iterative softmax

Denote the $(N-1)$-dimensional simplex in $\mathbb{R}^N$ by S_{N-1}, i.e.

$$S_{N-1} = \{\mathbf{w} = (w_1, w_2, ..., w_N)^T \in \mathbb{R}^N \mid w_i \geq 0, i = 1, 2, ..., N, \text{ and } \sum_{i=1}^{N} w_i = 1\}.$$

Given a parameter $T > 0$ (the "temperature"), the softmax maps S_{N-1} into its interior

$$\mathbf{w} \mapsto \mathbf{F}(\mathbf{w}; T) = (F_1(\mathbf{w}; T), F_2(\mathbf{w}; T), ..., F_N(\mathbf{w}; T))^T, \tag{1}$$

where

$$F_i(\mathbf{w}; T) = \frac{\exp(\frac{w_i}{T})}{Z(\mathbf{w}; T)}, \quad \text{and} \quad Z(\mathbf{w}; T) = \sum_{k=1}^{N} \exp(\frac{w_k}{T}). \tag{2}$$

The softmax map $\mathbf{F}$ induces on S_{N-1} a discrete time dynamics

$$\mathbf{w}(t+1) = \mathbf{F}(\mathbf{w}(t); T), \tag{3}$$

sometimes referred to as *Iterative Softmax* (ISM). Unless stated otherwise, we study systems for $N \geq 2$.

[1] in the sense of directions on the chessboard

4 Fixed points of Iterative Softmax

Recall that $\mathbf{w}$ is a fixed point (equilibrium) of the ISM dynamics driven by $\mathbf{F}$, if $\mathbf{w} = \mathbf{F}(\mathbf{w})$. It is easy to see that the maximum entropy point $\overline{\mathbf{w}} = (N^{-1}, N^{-1}, ..., N^{-1})^T \in S_{N-1}$ is a fixed point of ISM (3) for temperature setting T. In addition, there is a strong structure in the fixed points of ISM - coordinates of any fixed point of ISM can take on only two distinct values.

Theorem 1. *Except for the maximum entropy fixed point* $\overline{\mathbf{w}} = (N^{-1}, ..., N^{-1})^T$, *for all the other fixed points* $\mathbf{w} = (w_1, w_2, ..., w_N)^T$ *of ISM (3) it holds:* $w_i \in \{\gamma_1(\mathbf{w}; T), \gamma_2(\mathbf{w}; T)\}$, $i = 1, 2, ..., N$, *where* $\gamma_1(\mathbf{w}; T) > N^{-1}$ *and* $\gamma_2(\mathbf{w}; T) < N^{-1}$.

We will often write the larger of the two fixed-point coordinates as $\gamma_1(\mathbf{w}; T) = \alpha N^{-1}$, $\alpha \in (1, N)$.

Theorem 2. *Fix* $\alpha \in (1, N)$ *and write* $\gamma_1 = \alpha N^{-1}$. *Let* ℓ_{min} *be the smallest natural number greater than* $(\alpha - 1)/\gamma_1$. *Then, for* $\ell \in \{\ell_{min}, \ell_{min} + 1, ..., N - 1\}$, *at temperature*

$$T_e(\gamma_1; N, \ell) = (\alpha - 1) \left[-\ell \cdot \ln \left(1 - \frac{\alpha - 1}{\ell \gamma_1} \right) \right]^{-1}, \tag{4}$$

there exist $\binom{N}{\ell}$ *distinct fixed points of ISM (3), with* $(N - \ell)$ *coordinates having value* γ_1 *and* ℓ *coordinates equal to*

$$\gamma_2 = \frac{1 - \gamma_1(N - \ell)}{\ell}. \tag{5}$$

5 Fixed point stability in Iterative Softmax

Theorem 3. *Consider a fixed point* $\mathbf{w} \in S_{N-1}$ *of ISM (3) with one of its coordinates equal to* $N^{-1} \leq \gamma_1 < 1$. *Define*

$$T_s(\gamma_1) = \begin{cases} T_{s,2}(\gamma_1) = \gamma_1, & \text{if } \gamma_1 \in [N^{-1}, 1/2) \\ T_{s,1}(\gamma_1) = 2\gamma_1(1 - \gamma_1), & \text{if } \gamma_1 \in [1/2, 1). \end{cases} \tag{6}$$

Then, if $T > T_s(\gamma_1)$, *the fixed point* $\mathbf{w}$ *is stable.*

Theorem 4. *Consider a fixed point* $\mathbf{w} \in S_{N-1}$ *of ISM (3) with one of its coordinates equal to* $N^{-1} \leq \gamma_1 < 1$. *Let* $N - \ell$ *be the number of coordinates of value* γ_1. *Then if*

$$T \leq T_u(\gamma_1; N, \ell) = \gamma_1 \ (2 - N\gamma_1) \frac{N - \ell}{N\ell} + \frac{1}{N} - \frac{1}{N\ell}, \tag{7}$$

$\mathbf{w}$ *is not stable.*

An illustrative summary of the previous results for equilibria $\mathbf{w} \neq \overline{\mathbf{w}}$ is provided in figure 1. The ISM has $N = 17$ units. Coordinates of such fixed points can only take on two possible values, the larger of which we denote by γ_1. The number of coordinates with value γ_1 is denoted by N_1. Temperatures $T_s(\gamma_1)$ (6) above which equilibria with larger coordinate equal to γ_1 are guaranteed to be stable are shown with solid bold line. For $N_1 = N - \ell \in \{1, 2, 3, 4\}$, we show the temperatures $T_u(\gamma_1; N, \ell)$ (7) bellow which equilibria with larger coordinate equal to γ_1 are guaranteed to be unstable with solid normal lines. Temperatures $T_e(\gamma_1; N, \ell)$ (4) at which equilibria with the given number N_1 of coordinates of value γ_1 exist (dashed lines) are also marked according to stability type of the corresponding fixed points. The stability types were determined by eignanalysis of Jacobians of the ISM map $\mathbf{F}$ at the fixed points. Stable and unstable equilibria existing at temperature $T_e(\gamma_1; N, \ell)$ are shown as stars and circles, respectively. All the unstable equilibria are of saddle type. Horizontal dashed line shows a numerically determined temperature by Kwok and Smith [1] at which attractive equilibria of 17-dimensional ISM lose stability and the maximum entropy point $\overline{\mathbf{w}}$ remains the only stable stable fixed point. Position where $T_s(\gamma_1)$ crosses $T_e(\gamma_1; N, \ell)$ for $N_1 = 1$ is marked by bold circle. Note that no equilibrium with more than one coordinate greater than N^{-1} is stable.

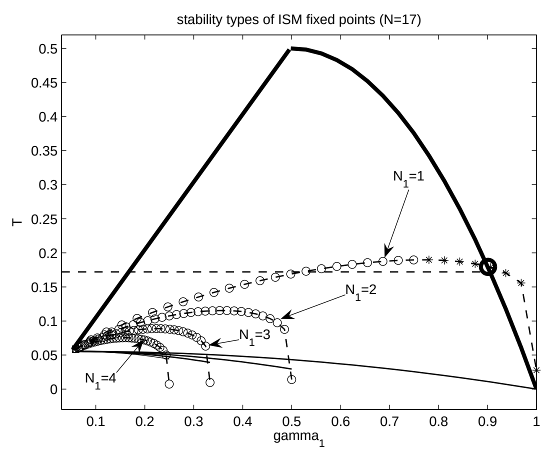

Fig. 1. Stability types for equilibria $\mathbf{w} \neq \overline{\mathbf{w}}$ of ISM with $N = 17$ units as a function of the larger coordinate γ_1 and temperature T. For a more detailed explanation, see the text.

6 Critical temperatures for intermittent search in SONN

It has been hypothesized that ISM provides an underlying driving force behind intermittent search in SONN with softmax weight renormalization [6, 1]. Kwok and Smith [1] argue that the critical temperature at which the intermittent search takes place corresponds to the "bifurcation point" of the *autonomous* ISM dynamics when the existing equilibria lose stability and only the maximum entropy point $\overline{\mathbf{w}}$ survives as the sole stable equilibrium. The authors numerically determined such bifurcation points for several ISM dimensionalities N. It was reported that bifurcation temperatures decreased with increasing N. Based on our analysis, the bifurcation points correspond to the case when equilibria near vertexes of the simplex S_{N-1} (equilibria with only one coordinate of large value γ_1) lose stability. We will call such fixed points *one-hot equilibria*. Based on the bound $T_s(\gamma_1)$ (6) and temperatures $T_e(\gamma_1; N, N-1)$ (4) at which equilibria of ISM exist, such a bifurcation point can be approximated by a bold circle in figure 1.

We present an analytical approximation to the critical temperature $T_*(N)$ at which the bifurcation occurs by expanding $T_e(\gamma_1; N, N-1)$ (4) around γ_1^0 as a second-order polynomial $T^{(2)}_{N-1}(\gamma_1)$. Based on figure 1 (and similar figures for other ISM dimensionalities N), a good choice for γ_1^0 is e.g. $\gamma_1^0 = 0.9$. Approximation to the critical temperature is then obtained by solving

$$T^{(2)}_{N-1}(\gamma_1) = T_s(\gamma_1)$$

for $\gamma_1 = \alpha/N$, and then plugging the result $\gamma_1^{(2)}$ back to bound (6), i.e. calculating $T_s(\gamma_1^{(2)})$.

We have

$$T_e(\gamma_1; N, N-1) = \frac{N\gamma_1 - 1}{(N-1)\ln\left(\frac{(N-1)\gamma_1}{1-\gamma_1}\right)}, \tag{8}$$

$$T'_e(\gamma_1; N, N-1) = \frac{N}{(N-1)\ln\left(\frac{(N-1)\gamma_1}{1-\gamma_1}\right)} - \frac{N\gamma_1 - 1}{(N-1)\gamma_1(1-\gamma_1)\ln^2\left(\frac{(N-1)\gamma_1}{1-\gamma_1}\right)} \tag{9}$$

and

$$T''_e(\gamma_1; N, N-1) = \frac{N\gamma_1^2 - 2\gamma_1 + 1 - 2\ln^{-1}\left(\frac{(N-1)\gamma_1}{1-\gamma_1}\right)}{(N-1)\gamma_1^2(1-\gamma_1)^2\ln^2\left(\frac{(N-1)\gamma_1}{1-\gamma_1}\right)}. \tag{10}$$

By solving

$$A\gamma_1^2 + B\gamma_1 + C = 0, \tag{11}$$

where

$$A = \frac{1}{2}T''_e(\gamma_1^0; N, N-1) + 2,$$

$$B = T'_e(\gamma_1^0; N, N-1) - \gamma_1^0\ T''_e(\gamma_1^0; N, N-1) - 2,$$

$$C = T_e(\gamma_1^0; N, N-1) - \gamma_1^0\ T'_e(\gamma_1^0; N, N-1) + \frac{1}{2}(\gamma_1^0)^2\ T''_e(\gamma_1^0; N, N-1)$$

and retaining a solution $\gamma_1^{(2)}$ that is compatible with the requirement that $\mathbf{w} \in S_{N-1}$, we obtain an analytical approximation $T_*^{(2)}(N)$ to the critical temperature $T_*(N)$:

$$T_*^{(2)}(N) = 2\gamma_1^{(2)}(1 - \gamma_1^{(2)}). \tag{12}$$

A cheaper approximation to $T_*^{(1)}(N)$ to $T_*(N)$ can be obtained by first-order expansions of both $T_s(\gamma_1)$ and $T_e(\gamma_1; N, N-1)$.

To illustrate the approximations $T_*^{(1)}(N)$ and $T_*^{(2)}(N)$ of the critical temperature $T_B = T_*(N)$, numerically found bifurcation temperatures for ISM systems with dimensions between 8 and 30 are shown in figure 2 as circles. The approximations based on quadratic and linear expansions, $T_*^{(2)}(N)$ and $T_*^{(1)}(N)$, are plotted with bold solid and dashed lines, respectively. Also shown are the temperatures $1/N$ above which the maximum entropy equilibrium $\overline{\mathbf{w}}$ is stable. At bifurcation temperature, $\overline{\mathbf{w}}$ is already stable and equilibria at vertexes of simplex S_{N-1} lose stability. The analytical solutions $T_*^{(2)}(N)$ and $T_*^{(1)}(N)$ appear to approximate the bifurcation temperatures well.

We also numerically determined optimal temperature settings for intermittent search by SONN in the N-queens problems. Following [1], the SONN parameter β was set to $\beta = 0.8$[2]. The optimal neighborhood size increased with the problem size N from $L = 2$ (for $N = 8$), through $L = 3$ ($N = 10, 13$), $L = 4$ ($N = 15, 17, 20$), to $L = 5$ ($N = 25, 30$)[3]. Based on our extensive experimentation, the best performing temperatures for intermittent search in the N-queens problems are shown as stars. Clearly, as suggested by Kwok and Smith, there is a marked correspondence between the bifurcation temperatures of ISM equilibria and the best performing temperatures in intermittent search by SONN. The *analytically* obtained approximations critical temperatures predict the optimal working temperatures for SONN intermittent search very well. So far the critical temperatures have been determined only via extensive numerical simulations.

7 Conclusions

Self-organizing neural networks driven by softmax weight renormalization are capable of powerful intermittent search for high-quality solutions in difficult assignment optimization problems [1]. We have rigorously analyzed equilibria of the underlying renormalization process and derive *analytical* approximations to their critical bifurcation temperatures necessary for optimal SONN performance. The approximations are in good agreement with the temperature values found by numerical investigations.

---

[2] Our experiments confirmed finding by Kwok and Smith that the choice of β is in general not sensitive to N.

[3] the increase of optimal neighborhood size L with increasing problem dimension N is in accordance with findings in [1]

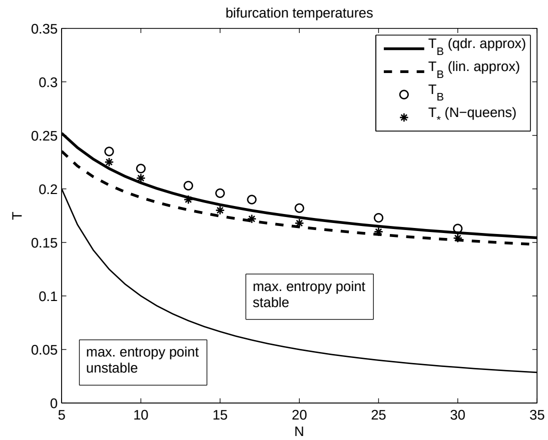

Fig. 2. Analytical approximations of the bifurcation temperature $T_B = T_*(N)$ for increasing ISM dimensionalities N. The approximations based on quadratic and linear expansions, $T_*^{(2)}(N)$ and $T_*^{(1)}(N)$, are plotted with bold solid and dashed lines, respectively. Numerically found bifurcation temperatures are shown as circles. The best performing temperatures for intermittent search in the N-queens problems are shown as stars.

References

1. Kwok, T., Smith, K.: Optimization via intermittency with a self-organizing neural network. Neural Computation **17** (2005) 2454–2481
2. Hopfield, J.: Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Science USA **79** (1982) 2554–2558
3. Smith, K.: Neural networks for combinatorial optimization: a review of more than a decade of research. INFORMS Journal on Computing **11** (1999) 15–34
4. Smith, K.: Solving the generalized quadratic assignment problem using a self-organizing process. In: Proceedings of the IEEE Int. Conf. on Neural Networks. Volume 4. (1995) 1876–1879
5. Smith, K., Palaniswami, M., Krishnamoorthy, M.: Neural techniques for combinatorial optimization with applications. IEEE Transactions on Neural Networks **9** (1998) 1301–1318
6. Kwok, T., Smith, K.: Performance-enhancing bifurcations in a self-organising neural network. In: Computational Methods in Neural Modeling: Proceedings of the 7th International Work-Conference on Artificial and Natural Neural Networks (IWANN'2003). Volume Lecture Notes in Computer Science, vol. 2686., Singapore, Springer-Verlag, Berlin (2003) 390–397
7. Langton, C.: Computation at the edge of chaos: Phase transitions and emergent computation. Physica D **42** (1990) 12–37
8. Tiňo, P.: Equilibria of iterative softmax and critical temperatures for intermittent search in self-organizing neural networks. Technical Report CSRP-06-07, University of Birmingham, School of Computer Science, http://www.cs.bham.ac.uk/~pxt/PAPERS/tino.tr06.pdf (2006)