ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 177-182

Mixture states in Potts neural networks

D. Bollé and J. Huyghebaert

Instituut voor Theoretische Fysica, K.U. Leuven

B-3001 Leuven, Belgium

Abstract

The presence and stability of mixture states in Q-state Potts neural networks is studied. Firstly a model is considered where a finite number $\tilde{p}$ of biased patterns is stored, using two different learning rules. A comparison of the retrieval properties is made. Secondly a model is studied where a finite number $\tilde{p}$ of biased and an infinite number $p - \tilde{p}$ of unbiased patterns are stored, using the Hebb learning rule. The storage capacity and the temperature-capacity phase diagram are calculated to study the retrieval behaviour.

1 Introduction

Neural networks with multi-state neurons [1]-[5] have attracted recent interest in order to study the storage and retrieval properties of grey-toned patterns. In [5] a systematic study of Q-state Potts neural networks with biased patterns and a generalized Hebb learning rule has been performed. In the present contribution the influence of the learning rule on the presence and stability of the symmetric and asymmetric mixture states at finite temperature, T, is considered.

2 Model

Consider the Hamiltonian

$$H = -\frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \sum_{k,l=1}^{Q} J_{ij}^{kl} u_{\sigma_i,k} u_{\sigma_j,l} , \tag{1}$$

with $u_{\sigma_i,k}$ the Potts operator, which takes the value $Q-1$ if $\sigma_i = k$ or -1 if $\sigma_i \neq k$, and $\sigma_i \in \{1, 2, ..., Q\}$ the Potts spins. J_{ij}^{kl} represents the synaptic coupling between neuron i in state k and neuron j in state l. Notice that the neurons are interconnected with all the others, $J_{ij}^{kl} \neq 0$ and that the couplings are symmetric: $J_{ij}^{kl} = J_{ji}^{lk}$.

ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 177-182

3 A finite number of biased patterns

All the stored patterns $\{\xi_i^\mu\}, \mu = 1, 2, ..., \tilde{p}$ ($\tilde{p}$=finite) are biased, i.e. the $\{\xi_i^\mu\}$ are chosen as independent random variables which can takes the values $k = 1, ..., Q$ with probability

$$P(\xi_i^\mu = k) = \frac{1 + B_k}{Q}, \qquad k = 1, ..., Q \tag{2}$$

where the B_k are the bias parameters.
Two different learning rules (see [6] for the Hopfield model) are studied

- Hebb learning rule

$$J_{ij}^{kl} = \frac{1}{Q^2 N} \sum_{\mu=1}^{\tilde{p}} u_{\xi_i^\mu, k} u_{\xi_j^\mu, l}, \tag{3}$$

- learning rule with mixture term

$$J_{ij}^{kl} = \frac{1}{Q^2 N} \sum_{\mu=1}^{\tilde{p}} u_{\xi_i^\mu, k} (u_{\xi_j^\mu, l} - \frac{1}{\tilde{p}} \sum_{\nu=1}^{\tilde{p}} u_{\xi_j^\nu, l}). \tag{4}$$

For both interactions one can calulate the free energy [7] which depends on the order parameters

$$m_\mu = \frac{1}{N} \sum_{i=1}^{N} \langle\langle < u_{\xi_i^\mu, \sigma_i} > \rangle\rangle. \tag{5}$$

They represent the overlap with the μ-th pattern. Here $< . >$ stands for the thermal average and $\langle\langle . \rangle\rangle$ for the average over the $\tilde{p}$ biased patterns.
These order parameters satisfy fixed-point equations, which can be solved numerically. These fixed-point equations have [7]

- asymmetric solutions $\mathbf{m} = (m_1, m_{\tilde{p}-1}, ..., m_{\tilde{p}-1})$ with retrieval properties $(m_1 >> m_{\tilde{p}-1})$
- symmetric solutions $\mathbf{m} = (m_{\tilde{p}}, ..., m_{\tilde{p}})$ which only exist for the Hebb learning rule and which represent confusion of the neural network
- a solution $\mathbf{m} = 0$ representing the disordered state

The example of a $Q = 3$ network with the bias types $\mathbf{B}_1 = a(2, -1, -1)$ and $\mathbf{B}_2 = a(1, 0, -1)$, $a \in [0, 1]$, is considered. For $\tilde{p} = 3$ biased patterns the $T - a$ diagram is presented in Fig. 1.

ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 177-182

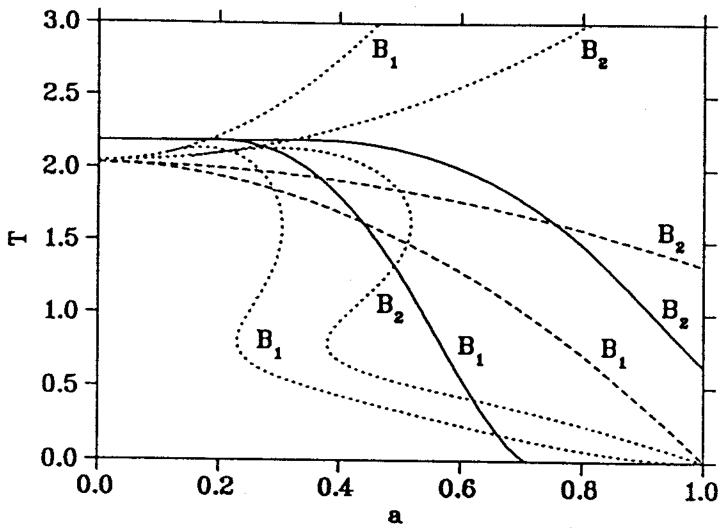

Fig.1. The $T-a$ diagram for the $Q=3$ $\mathbf{B}_1$ and $\mathbf{B}_2$ network with $\tilde{p}=3$ biased patterns.

The solid (respectively the dashed) lines represent the temperature below which stable asymmetric retrieval solutions appear in the model with the Hebb learning rule (respectively the model with the Hebb learning rule with mixture term). For the Hebb model with bias $\mathbf{B}_1$ retrieval is only possible for $a \leq 1/\sqrt{2}$.
In the region between the two dotted lines there exist stable symmetric solutions of the Hebb model. Below the lowest dotted line these symmetric solutions become unstable. At $a=1$ the symmetric solutions of the $\mathbf{B}_1$ ($\mathbf{B}_2$) model exists up to $T=6.56$ ($T=3.53$). Remark that for the learning rule with mixture term there are no symmetric solutions.
The $\mathbf{m}=0$ solution exists for all temperatures, and it become stable above $T=Q-1+(\tilde{p}-1)\sum_{k=1}^{Q}B_k^2/Q$.
In view of these results, one concludes that the Hebb learning rule with mixture term leads to asymmetric retrieval over the whole range af a-values. Further, it has no symmetric solutions. Consequently its retrieval properties are better than the model with the usual Hebb rule.

4 A combination of biased and unbiased patterns

In this model we consider p stored patterns $\{\xi_i^\mu\}, \mu=1,...,p$ of which

- $\tilde{p}$ (finite) are biased: $P(\xi_i^\mu=k)=\frac{1+B_k}{Q}$, $\qquad \mu=1,...,\tilde{p}$
- $p-\tilde{p}$ (infinite) are unbiased: $P(\xi_i^\mu=k)=\frac{1}{Q}$, $\qquad \mu=p-\tilde{p},...,p$

The synaptic couplings are given by the Hebb learning rule

$$J_{ij}^{kl}=\frac{1}{Q^2N}\sum_{\mu=1}^{p}u_{\xi_i^\mu,k}u_{\xi_j^\mu,l}. \tag{6}$$

ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 177-182

Using the mean field approximation and the replica technique, one finds that the free energy [1],[7] depends on

- the macroscopic overlap with a condensed pattern

$$m_\nu^\lambda = \frac{1}{N}\sum_{i=1}^{N}\langle\langle < u_{\xi_i^\nu,\sigma_i^\lambda} > \rangle\rangle, \tag{7}$$

- the Edwards-Anderson order parameter

$$q_{\lambda\rho} = \frac{1}{N}\sum_{i=1}^{N}\langle\langle\sum_{k=1}^{Q}\frac{1}{Q} < u_{k,\sigma_i^\lambda} >< u_{k,\sigma_i^\rho} > \rangle\rangle, \tag{8}$$

- the total mean-square random overlap with the non-condensed patterns

$$r_{\lambda\rho} = \frac{1}{\alpha}\sum_{\mu=s+1}^{p}\langle\langle m_\mu^\lambda m_\mu^\rho \rangle\rangle, \tag{9}$$

where λ and ρ are replica indices. Here the storage capacity, α, is given by $\alpha = p/N$. Within the replica symmetric approximation ($m^\lambda = m, q_{\lambda\rho} = q, r_{\lambda\rho} = r$) the order parameters (6)-(8) satisfy fixed-point equations [7] having

- asymmetric solutions $\mathbf{m} = (m_1, m_{\tilde{p}-1}, ..., m_{\tilde{p}-1}, 0, ..., 0)$ and $q \neq 0$ $r \neq 0$ with retrieval properties ($m_1 >> m_{\tilde{p}-1}$)
- symmetric solutions $\mathbf{m} = (m_{\tilde{p}}, ..., m_{\tilde{p}}, 0, ..., 0)$ and $q \neq 0$ $r \neq 0$ representing confusion in the network
- spin-glass solutions $\mathbf{m} = 0$ and $q \neq 0$ $r \neq 0$
- a paramagnetic solution $\mathbf{m} = 0$ and $q = r = 0$

The storage capacity α, at $T = 0$, of the $Q = 3$ $\mathbf{B}_1$ and $\mathbf{B}_2$ model with $\tilde{p} = 3$ correlated and $p - \tilde{p}$ uncorrelated patterns is given in Fig. 2, as a function of a for the asymmetric retrieval (full line) and symmetric solutions (dotted line). For the model with bias $\mathbf{B}_1$ retrieval is only possible for $a \leq 1/\sqrt{2}$. Comparing both the $\mathbf{B}_1$ and the $\mathbf{B}_2$ case one sees that the storage capacity of the asymmetric states is always the largest for the $\mathbf{B}_2$ model. Furthermore the storage capacity of the symmetric states grows slower for the $\mathbf{B}_2$ than for the $\mathbf{B}_1$ model. At $a = 1$ the storage capacity of the symmetric states of the $\mathbf{B}_1$ ($\mathbf{B}_2$) model is $\alpha = 6.40$ ($\alpha = 0.80$).

ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 177-182

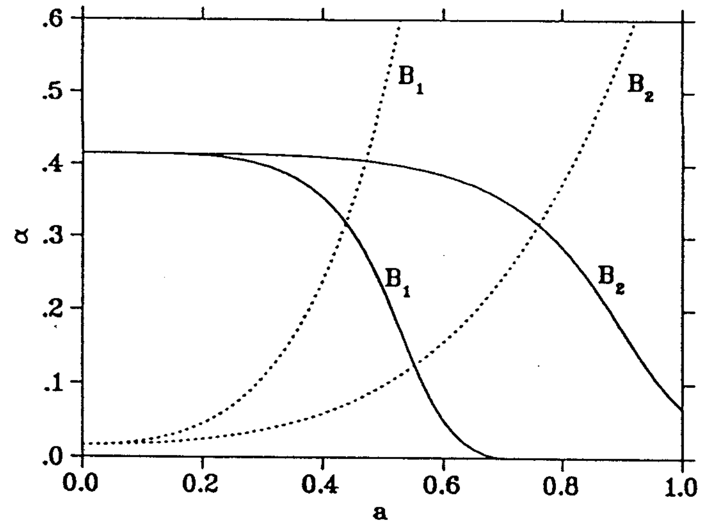

Fig.2. The $\alpha - a$ diagram for the $Q = 3$ $\mathbf{B}_1$ and $\mathbf{B}_2$ network with $\tilde{p} = 3$ biased patterns and $p - \tilde{p}$ unbiased patterns.

In Fig. 3. and 4. the $T - \alpha$ phase diagrams at $a = 0.4$ are shown.

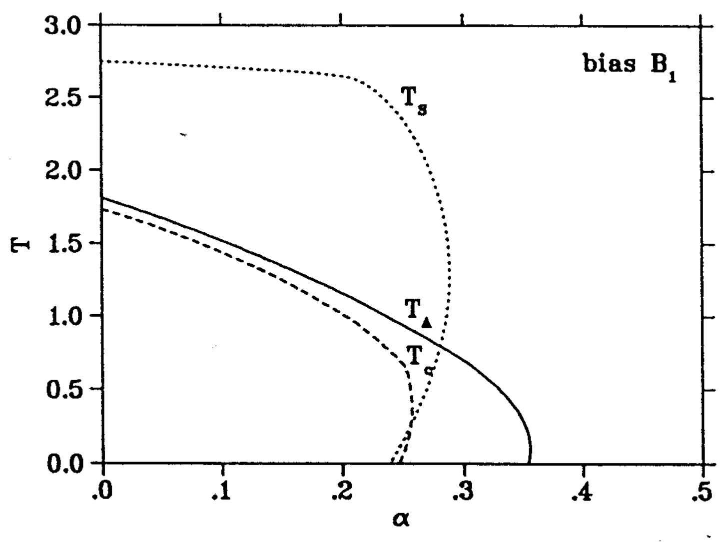

Fig.3. The $T - \alpha$ phase diagram for the $Q = 3$ $\mathbf{B}_1$ network with $\tilde{p} = 3$ biased patterns and $p - \tilde{p}$ unbiased patterns at $a = 0.4$.

ESANN'1993 proceedings - European Symposium on Artificial Neural Networks
Brussels (Belgium), 7-8-9 April 1993, D-Facto public., ISBN 2-9600049-0-6, pp. 177-182

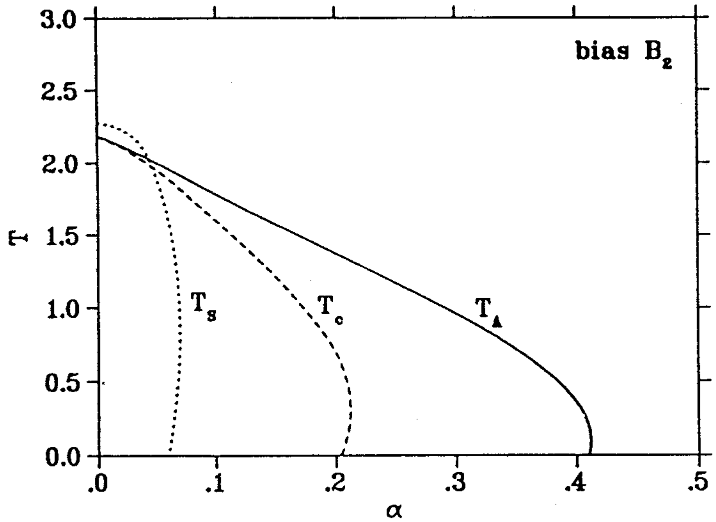

Fig.4. The $T-\alpha$ phase diagram for the $Q=3$ $\mathbf{B}_2$ network with $\tilde{p}=3$ biased patterns and $p-\tilde{p}$ unbiased patterns at $a=0.4$.

Below T_A (full line) asymmetric retrieval states show up as local minima of the free energy. Below T_S (dotted line) symmetric states appear as local minima of the free energy. Below T_c (dashed line) the asymmetric retrieval states become global minima of the free energy. The transition at T_c is a first order transition. Comparing Fig. 3. and 4., one concludes that the retrieval region for the $\mathbf{B}_2$ model is larger than for the $\mathbf{B}_1$ model. Furthermore, the $\mathbf{B}_2$ model has the smallest $T-\alpha$ region for existence of symmetric states.
Consequently the asymmetric retrieval properties are better for the set of bias parameters $\mathbf{B}_2$.

This work has been supported in part by the Research Fund of the K.U. Leuven (grant OT/91/13).

References

[1] Kanter I 1988 *Phys. Rev. A* **37** 2739

[2] Shim G M, Kim D, Choi Y M 1992 *Phys. Rev. A* **45** 1238

[3] Vogt H, Zippelius A 1992 *J. Phys. A* **25** 2209

[4] Bollé D, Dupont P, Huyghebaert J 1992 *Phys. Rev. A* **45** 4194

[5] Bollé D, Cools R, Dupont P, Huyghebaert J *J. Phys. A*, to appear

[6] Fontanari J F, Theumann W K 1990 *J. Phys. France* **51** 375

[7] Bollé D, Huyghebaert J "Storage of mixture states in Potts glass neural networks" preprint KUL-TF-92