Sequence Processing Neural Network with $Q$-States Monotonic Transfer Function

Katsuki Katayama$^1$ and Tsuyoshi Horiguchi$^2$

$^1$Department of Computer Science and Systems Engineering, Muroran Institute of Technology, Hokkaido 050-8585, Japan
$^2$Department of Computer and Mathematical Sciences, Graduate School of Information Sciences, Tohoku University, Sendai 980-8579, Japan

Storage capacity as for retrieval of sequences of binary patterns is investigated for a fully connected neural network with $Q$ ($>2$)-states. By using a generating-function method of path-integral representation, we find that the network with $Q$-states monotonic transfer function retrieves more sequences of the stored patterns than that with a binary monotonic transfer function at zero temperature, if the control parameter is chosen optimally. We compare the results obtained by the analytic method with those by numerical simulations.

§1. Introduction

We investigate storage capacity of a fully connected neural network with a $Q$ ($>2$)-states monotonic transfer function. We focus on a problem for retrieval of sequences of binary patterns with synchronous dynamics. As for an analytic method, it is difficult to use conventional methods of the equilibrium statistical mechanics for the investigation of the present network, because of asymmetry of interactions. We then use a generating-function method of path-integral representation, and obtain equations for order parameters in the stationary state. We also perform numerical simulations in order to compare the obtained analytic results.

The present paper is organized as follows. In §2, we give an analysis by means of the generating function of path-integral representation. In §3, we present the obtained results. Concluding remarks are given in §4.

§2. Generating function method of path-integral representation

We define a generating function of path-integral representation $Z[\psi]$ as follows:

$$Z[\psi] = \left[ \sum_{\sigma(0)} \cdots \sum_{\sigma(t)} P[\sigma(0), \cdots, \sigma(t)] \exp \left\{ -i \sum_{s=0}^{t-1} \psi(s) \cdot \sigma(s) \right\} \right]_{\xi}.$$  \hspace{1cm} (2.1)

$P[\sigma(0), \cdots, \sigma(t)]$ represents a probability defined for a configuration $(\sigma(0), \cdots, \sigma(t))$, and $[Y]_{\xi}$ denotes a pattern average of quantity $Y$. From the generating function $Z[\psi]$, we can derive equations of dynamical order parameters, namely a sequence overlap $m(s)$, an activity $a(s)$, a response function $G(s, s')$ and a correlation function $C(s, s')$. When the updating rule is synchronous, the probability $P[\sigma(0), \cdots, \sigma(t)]$
Sequence Processing Neural Network

267

is expressed as follows:

\[ P[\sigma(0), \ldots, \sigma(t)] = P[\sigma(0)] \prod_{s=0}^{t-1} \prod_{i=1}^{N} W_i[\sigma_i(s+1)|\sigma_i(s)]. \quad (2.2) \]

The transition probability \( W_i[\sigma_i(s+1)|\sigma_i(s)] \) is given for the \( Q \)-states monotonic transfer function\(^1\) as follows:

\[
W_i[\sigma_i(s+1)|\sigma_i(s)] = \frac{\exp \left[ \beta \left\{ \sum_{j=1}^{N} J_{ij} \sigma_j(s) \right\} \sigma_i(s+1) - \beta \theta \sigma_i(s+1)^2 \right]}{\sum_{\sigma_{i(s+1)}} \exp \left[ \beta \left\{ \sum_{j=1}^{N} J_{ij} \sigma_j(s) \right\} \sigma_i(s+1) - \beta \theta \sigma_i(s+1)^2 \right]}, \quad (2.3)
\]

where \( \theta \) denotes a threshold parameter and \( \beta = 1/T; T \) is a network noise level. The Hebb rule for the interaction is defined for the sequence processing neural network as follows:

\[
J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_{i}^{\mu+1} \xi_{j}^{\mu}, \quad \xi_{i}^{p+1} = \xi_{i}^{1}, \quad (2.4)
\]

where each of \( \xi_{i}^{\mu} \) takes \( \pm 1 \) randomly. We notice that the interaction (2.4) is asymmetry.

After a long calculation, we obtain the overlap \( m \), the activity \( a \), the response function \( G \), the correlation function \( C \) and a parameter \( \rho \) in the stationary state, which is independent of \( s \) and \( s' \), as follows:

\[
m = \int Dz \left[ \frac{\sum_{\sigma} \sigma \exp \left\{ \beta (m + \sqrt{\alpha \rho z}) \sigma - \beta \theta \sigma^2 \right\}}{\sum_{\sigma} \exp \left\{ \beta (m + \sqrt{\alpha \rho z}) \sigma - \beta \theta \sigma^2 \right\}} \right], \quad (2.5)
\]

\[
a = \int Dz \left[ \frac{\sum_{\sigma} \sigma^2 \exp \left\{ \beta (m + \sqrt{\alpha \rho z}) \sigma - \beta \theta \sigma^2 \right\}}{\sum_{\sigma} \exp \left\{ \beta (m + \sqrt{\alpha \rho z}) \sigma - \beta \theta \sigma^2 \right\}} \right], \quad (2.6)
\]

\[
G = \beta \left[ a - \int Dz \left( \frac{\sum_{\sigma} \sigma \exp \left\{ \beta (m + \sqrt{\alpha \rho z}) \sigma - \beta \theta \sigma^2 \right\}}{\sum_{\sigma} \exp \left\{ \beta (m + \sqrt{\alpha \rho z}) \sigma - \beta \theta \sigma^2 \right\}} \right)^2 \right], \quad (2.7)
\]

\[
\rho = \frac{a}{1 - G^2}, \quad (2.8)
\]

where \( \alpha \) is a load parameter, and \( Dz \) is a Gaussian measure, and we get \( C = 0 \).

§3. Results

We show phase diagrams on \( \theta-\alpha \) plane at \( T = 0 \) for \( Q = 3 \) and \( Q = 4 \) in Figs. 1(a) and (b), respectively; a line denotes the analytic result, and a circle denotes results by numerical simulations with 4000 neurons. For \( Q = 3 \), we have a retrieval phase, a paramagnetic phase 1 and a paramagnetic phase 2. The paramagnetic phase 1 is characterized by the activity \( a \neq 0 \), and the paramagnetic phase 2 is characterized by the activity \( a = 0 \). We obtain the maximum storage capacity 0.2752 at \( \theta = 0.10 \).
When $\theta = 0$, the storage capacity becomes 0.2690, and this is the value of the storage capacity obtained for the binary monotonic transfer function. For $Q = 4$, we have a retrieval phase 1, a retrieval phase 2 and a paramagnetic phase. The retrieval phase 1 is characterized by the overlap $m \sim 1$, and the retrieval phase 2 is characterized by the overlap $m \sim 1/3$. We obtain the maximum storage capacity 0.2761 at $\theta = 0.10$. When $\theta = 0$, the storage capacity becomes 0.2690, and this is again the value of the storage capacity for the binary monotonic transfer function. We find that the analytic results are in excellent agreement with those obtained by numerical simulations for $Q = 3$ and $Q = 4$. We show a phase diagram on $\theta$-$\alpha$ plane at $T = 0$ for $Q = \infty$ in Fig. 2; a line denotes the analytic result, and a circle denotes results by numerical simulations with 4000 neurons. For $Q = \infty$, the maximum storage capacity is 0.2690 at $\theta = 0$, and this is again the value of the storage capacity for the binary monotonic transfer function.

The maximum storage capacity becomes 0.2690 for $Q = 2$, 0.2752 for $Q = 3$, 0.2761 for $Q = 4$, 0.2760 for $Q = 5$, 0.2759 for $Q = 6$ and 0.2690 for $Q = \infty$. Then, we find that the system with $Q = 4$ has the largest storage capacity among all the value of $Q$.

§4. Concluding remarks

We have investigated the storage capacity of the fully connected neural network with a $Q$ ($> 2$)-states monotonic transfer function. We have focused on a problem for the retrieval of the sequences of the binary patterns with synchronous dynamics. As for the analytic method, it is difficult to use conventional methods of the equilib-
rium statistical mechanics for the investigation of the present network, because the interactions are asymmetric. Then, we have used the generating-function method of path-integral representation, and obtain the equations for the order parameters in the stationary state. We have found that the network with the \( Q > 2 \) -states monotonic transfer function retrieves more sequences of the binary patterns than that with the binary monotonic transfer function at zero temperature, when the control parameter for the \( Q > 2 \) -states monotonic transfer function is selected optimally. We have showed that the analytic results are in excellent agreement with those obtained by numerical simulations.

References