

지연시냅스를 가진 계열 연상 메모리의 상기 다이나믹스 해석

김 응 수[†]

요 약

신경 회로에는 반드시 어떤 의미의 피드백이 존재한다. 신경 회로의 기본적인 정보처리 능력을 연구하는 측면에서는 피드백이 없는 계층형 회로의 동작 특성 및 제 성질에 대한 논의의 중요성에 의문의 여지는 없으나, 여기서는 다이내믹한 시스템으로서의 신경회로망 모델의 본질에 보다 직접적으로 다가갈 목적으로 지연 시냅스를 가진 귀환형 신경 회로를 대상으로 그 다이나믹스를 이론적으로 해석한 결과를 기술한다. 본 논문에서 제안한 신경회로망 모델의 상기 다이나믹스를 설명하기 위하여 통계적 해석법을 도입하였으며, 이러한 이론 해석의 결과를 컴퓨터 시뮬레이션과 비교하여 그 타당성을 입증하였다.

Analysis of Recall Dynamics of Sequential Associative Memory with Delay Synapses

EungSoo Kim[†]

ABSTRACT

Every neural network has some kind of feedback. For the sake of analyzing fundamental aspects of information processing in neural nets, a net without feedbacks is an important theoretical model. But here we focus on a recurrent neural net with delay synapses as a realistic dynamical model of nervous systems.

Synaptic connections are determined by a version of the Hebb rule (correlation type rule). We use a statistical neurodynamic method to explain the retrieval dynamics of the network. The result of the analysis for the sequential associative memory with delay synapses is compared with computer simulation. We have succeeded in explaining the dynamics of this network by theoretical analyses.

1. 서 론

분산 표현된 정보의 기억이나 상기라고 하는 문제는 신경회로망 연구에 있어서 중요한 주제이다. 이러한 신경회로망 모델의 동작에는 정보의 축적 과정(학습)과 처리 과정(인식이나 판별 등)이 있으며, 그것들이 실질적으로 분리되어 있는 경우와 분리 불가

능한 경우가 있다. 정보 축적 과정의 알고리즘에 중점을 둔 연구는 BP알고리즘 [1]으로 대표되는 것과 같이, 알고리즘을 어떻게 변경하면 성능이 얼마나 변하는가 혹은 어느정도의 시간으로 학습이 완성되는가 하는 부분에 대한 많은 연구가 진행되어 왔다 [2] [3]. 또한 정보의 처리과정에 중점을 둔 연구로는 연상기억의 다이나믹스와 평형상태의 해석에 대한 연구가 있다 [4][5]. 즉, 기억을 어느정도 잘 저장할 수 있는가 혹은 기억을 재현하기 위해서는 상태 오차를

[†] 정. 회 원: 선문대학교 공과대학 전자공학과
논문접수: 1995년 12월 20일, 심사완료: 1996년 3월 28일

어느정도까지 허용하는가 하는 연구 등이 있다 [6].

지금까지 이러한 문제들에 대한 연구는 동적 과정이 본질적이 아닌 모델이나 또는 그것이 본질적이라 하더라도 결과적으로 달성되는 평형 상태만이 의미가 있는 모델을 대상으로 주로 연구되어 왔다.

연상 기억의 다이내믹스에 관한 연구가 성황을 이룰 때 자기 상기형 (Auto-correlation associative memory)의 모델이 주로 연구되었다 [4]. 자기 상기형이라는 것은 기억패턴을 신경회로 모델의 평형상태로 기억하는 것으로서 정적인 패턴을 기억하는 것이다. 기억패턴에서 벗어난 신호가 입력되더라도 그것에 가장 가까운 기억패턴을 출력할지 어떨지는 이 모델의 성능을 평가하는 척도의 하나이다. 이 경우 시냅스 결합으로서는 상관행렬 (Correlation matrix)과 함께 직교사영형 (Orthogonal projection type)이 쓰이며, 각 시냅스 결합에 관한 여러가지 성질이 밝혀졌으며 [5] 지금도 연구는 진행되고 있다.

그러나 이러한 모델들에 자주 나타나는 시냅스 값의 대칭성에 대한 제약이라든지 혹은 단순히 평형 상태만을 구하는 것은 정보처리의 측면에서 볼 때 부자연스러워 보인다. 왜냐하면 모든 정보처리는 결국 어떠한 형태의 기호의 흐름 (계열)이기 때문이다.

자기 연상과 대조되는 것이 계열 연상 (Sequential associative memory)이다. 계열연상 모델은 주로 다음의 두가지 이유에서 중요하다. 하나는 정보의 형태에 관한 것이다. 사람을 시작으로 하는 현실시스템에 관한 모든 정보는 반드시 어떤 의미의 계열이므로, 이러한 성질을 신경회로망 모델에서 다루는 것은 평형 상태에 수렴하는 자기상기형 모델에만 그치지 않고 다이내믹하게 변하는 계열연상의 의미를 정보처리의 입장에서 밝힘과 동시에 모델의 연구를 보다 발전시키기 위해서도 중요하다. 또 다른 하나는 수학적 해석의 문제이다. 위에서도 언급하였듯이, 연구의 초기에는 자기상기형의 모델이 주로 대상이 되어 왔으나 그것을 통해서 밝혀진 것은 자기상기형 모델에 대한 이론해석의 본질적인 어려움이였다. 그러나 계열연상형 모델에는 그런 류의 어려움이 없으며 엄밀한 해석이 가능하다는 것이 알려져 있다 [7]. 이론적으로 엄밀한 해석이 가능하다는 것은 기존모델의 해석에 그치지 않고 이론을 기초로 모델을 발전시킬 수 있을 뿐만 아니라 다른 해석곤란한 모델들에 대한 이해도

증진시킬 수 있다.

이러한 관점에서 비대칭 시냅스 값을 가진 회로망으로 실현되는 동적인 연상 (계열 연상)은 중요하게 여겨지며, 이러한 모델은 보다 뇌의 기능에 가까운 정보처리를 실현하고 이해하는데 도움이 될 것으로 본다. 또한 실제의 생물학적 입장에서 본 신호의 지연은 기억과 상기의 측면에 대단히 중요한 역할을 한다는 것이 알려져 있다 [8]. 실제의 신경계에는 시냅스에 지연이 존재하며, 그것은 어떠한 형태로든 신경회로의 동작에 영향을 미치고 있다. 뉴런에 가까운 수상돌기에 붙어있는 시냅스가 뉴런에 미치는 영향은 빠르며 반대로 먼 시냅스는 천천히라고 하는 효과를 나타낸다. 이러한 차이에 기능적인 의미가 있을까, 아니면 단순히 구조상의 타협의 산물인 것인가. 여기서는 그것까지 밝히기는 어렵지만 그러한 구조의 의미를 밝히는 연구와 관련지어 갈것이다.

따라서 여기서는 지연 시냅스를 가진 계열 연상형 신경회로망 모델의 다이내믹스를 정량적으로 해석한다. 수치실험을 정확히 설명하는 상태천이식을 나타낸 다음 이러한 수식으로 표현되는 동작을 시뮬레이션 결과와 비교함으로써 고찰하고자 하는 신경회로망 모델의 기억과 상기 다이내믹스에 관한 제 특성에 대하여 기술하고자 한다.

2. 계열 연상 모델의 배경

신경회로망 모델에서 계열을 포함한 패턴 다이내믹스에 관하여 연구한 것은 Caianiello (1961)까지 거슬러 간다 [9]. 그러나 그의 연구는 모델의 제안이 주였으며 신경회로망 모델의 다이내믹스에 대한 구체적인 수리 해석은 Amari(1972)와 Fukushima(1973)에서 시작되었다고 할 수 있다 [4][8]. 이들의 연구에서는 학습 과정도 포함하여 어느 정도의 다이내믹스가 해석되었다.

Hopfield(1982)가 자기 연상 기억 모델에 관한 논문을 발표한 이후 [10], 여러 연구자들에 의해 이 모델에 대한 통계역학적 해석 가능성이 밝혀지면서 신경회로망 모델의 다이내믹스에 관한 정량적인 이론이 본격적으로 나타나기 시작했다. 계층형 회로에서의 계열 연상 기억 모델을 이론적으로 해석한 것은 Meir와 Domany(1987)이다. 그들은 다이내믹스의 엄밀해를

해석한 다음 기억용량을 비롯하여 계층형회로망의 특성에 관한 여러가지 측면을 이론적으로 밝혔다. 그들의 이론에 따르면 자기상기형 모델의 기억용량(안정적으로 보존할 수 있는 기억패턴의 최대수를 뉴런의 수로 나눈 것)이 0.15 임에 대하여 계층형회로에서의 기억용량은 0.27인 것을 밝혔다 [11].

자기 상기형 모델에 대한 엄밀한 이론이 극히 특수한 조건을 제외하고는 원리적으로 불가능한 것임에 대하여 계열연상 모델(계층형모델 포함)은 시냅스 결합의 불완전성 등과 같은 여러가지 조건하에서도 엄밀한 해석이 가능하다.

뉴런은 다양한 형태로 과거 동작에 영향을 받으며 또한 시냅스에는 여러 가지의 지연 시간이 있다. 이것에 관한 연구로서는 상관관이 존재하는 계열 연상에 대하여 뉴런의 자기 결합(히스테리시스)의 역할에 관한 연구가 있다 [12]. 또한 지연 시냅스에 관해서는 지연이 1인 자기 상관 시냅스와 지연이 1에 비해서 큰 계열 상관 시냅스를 조합하여 하나의 패턴에서의 체제 시간을 길게 하면서 동기 없이 계열을 재생할 수 있는 모델이나 [13], 연속적으로 분포한 지연 시간을 가진 시냅스와 적응적인 역치를 이용하여 여러 가지 속도로 계열을 재생할 수 있는 모델 등도 해석되어 있다 [14].

여기서 다루는 모델은 그것들에 비하면 꽤 단순한 동기로 시작되었으며, 보다 알기 쉬운 구조를 하고 있다. 즉, 예를 들어 시각 t 에 있어서 회로망이 계열

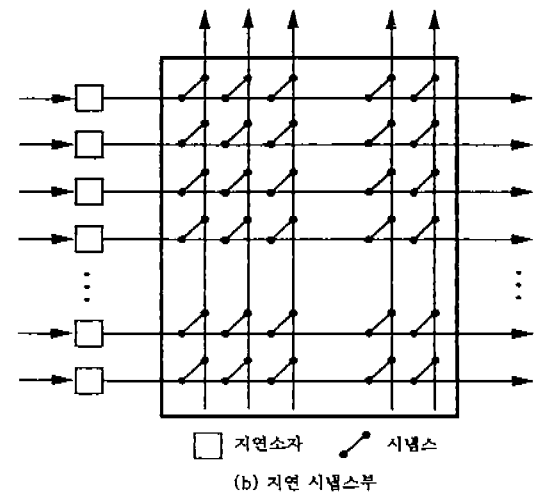
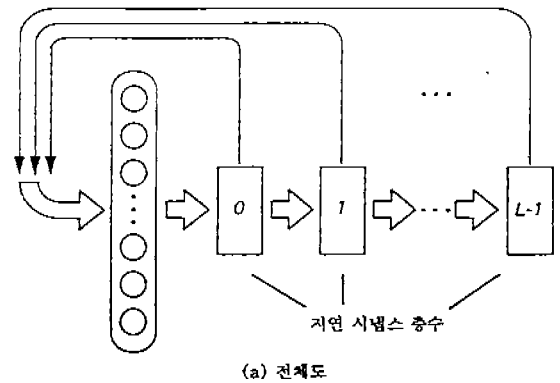
$$A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$$

의 계열 패턴 중 패턴 D 의 상태에 있었다고 할 때, 상태의 갱신에 있어서 지연 1의 정보 $D \rightarrow E$ 뿐만 아니라 지연 2의 정보 $C \rightarrow E$ 나 지연 3의 정보 $B \rightarrow E$ 등도 함께 이용한다면 보다 바른 계열을 재생할 수 있다는 기대이다. 랜덤패턴의 계열에 관하여 단순히 생각해 보면, 덧붙이는 신호가 L 개로 되면 S/N 비가 \sqrt{L} 배로 되므로 지연 시냅스를 다중으로 할수록 연상 성능이 향상될 것으로 여겨진다. 그러나 실제로는 그렇게 단순하지 않다는 사실이 수치 시뮬레이션에 의해 지적되어 이론적 해석이 시도되었지만 시뮬레이션에 나타나는 기묘한 움직임을 충분히 설명하지는 못하였다 [15].

3. 모델의 개요

지금부터 논의하는 신경회로망 모델의 개념도는 (그림 1)과 같다. 뉴런에서의 출력은 복수의 지연 시냅스 층을 거쳐서 귀환된다. 지연 시냅스에는 지연된 시간만큼 늦어진 뉴런 활동에 관한 상관(correlation)이 축적되어 있다(상세한 것은 4절). 뉴런의 출력은 2차이며, 시간도 정수값을 취하는 이산 다이내믹스를 가정한다. 시냅스는 모든 지연 시냅스 층을 거쳐서 뉴런층과 결합되어 있다. 뉴런의 수를 n , 지연 시냅스 층의 수를 L 로 하면 전체 Ln^2 개의 시냅스 결합이 있는 것으로 된다.

한편 이러한 신경회로망 모델에서는 기억패턴의



(그림 1) 지연 시냅스를 가진 신경회로망 모델
(Fig. 1) Neural network model with delayed synapses

정보를 분산적이면서도 다중적으로 축적 하기 때문에 회로의 성능은 기억패턴 상호의 상관에 깊이 의존한다. 따라서 여기서는 랜덤한 패턴 계열에 대한 성질을 통계적인 방법으로 논의하고자 한다. 이러한 방법은 모델의 평균적인 성질을 파악하는데 중요한 역할을 수행하기 때문이다.

4. 상기 다이내믹스의 해석

여기서의 목표는 회로망의 평균적인 유사도 (overlap)를 계산하는 것이다. 구체적인 수식화를 행하면 다음과 같다. 회로망의 다이내믹스는 각 뉴런에 대하여 다음의 식을 동기적으로 적용하면서 진행한다.

$$x_i(t+1) = f \left(\sum_{l=0}^{L-1} \sum_{j=1}^n w_{ij}^l x_j(t-l) \right) \quad (1)$$

- 여기서,
- n : 뉴런의 수
- L : 지연시냅스 층의 수
- t : 시각 (0, 1, 2, ...)
- l : 지연 시각 (0, 1, 2, ..., L-1)
- $x_i(t)$: 시각 t 에 있어서 뉴런 i 의 상태 (1 혹은 -1)
- w_{ij}^l : 뉴런 j 에서 뉴런 i 로의 시냅스 결합 중에 지연 시간이 l 인 것

$$f(u) = \begin{cases} -1, & u < 0 \\ +1, & u \geq 0 \end{cases}$$

즉, 이산적인 지연 시간을 가진 시냅스로 구성된 이산 다이내믹스의 특성을 가지는 신경회로망 모델을 생각한다. 기억시키고 싶은 패턴 계열을

$$S(0) \rightarrow S(1) \rightarrow S(2) \rightarrow \dots \rightarrow S(m-1) \rightarrow S(m) \quad (2)$$

로 한다. 단, $S(\mu)$ 는 성분이 +1의 n 차원 벡터이다. 일반적으로는 계열이 위와 같이 하나인 것이 아니라 복수개 있을 수도 있으며 또한 주기적인 계열이 있을 수도 있지만 계열이 복수개 있더라도 주기 계열의 주기가 $2L+1$ 이상이면 이후 기술하고자 하는 논의에는 기본적으로 영향이 없다. 따라서 해석을 단순화하기 위하여 지금부터는 식 (2)로 주어지는 계열을 대상

으로 논의를 한다. 단, 여기서는 계열의 분기에 관해서는 다루지 않는다.

식 (2)의 계열에 대해서 다음의 상관 행렬형 시냅스 결합을 이용한다.

$$w_{ij}^l = \begin{cases} c_l \sum_{\mu=l}^{m-1} s_i(\mu+1) s_j(\mu-l), & i \neq j \\ 0, & i = j \end{cases} \quad (3)$$

- 단, $s_i(\mu)$: 계열 패턴 $S(\mu)$ 의 제 i 번째 성분(1 혹은 -1 각각의 확률이 1/2인 랜덤패턴)
- c_l : l 지연 시냅스의 크기

논의를 필요 이상으로 복잡하게 만들지 않기 위하여 자기 결합은 0으로 한다. 상관행렬형의 시냅스 결합은 뉴런이 같은 상태 일때에는 두 뉴런간의 시냅스 결합을 증가시키고 다른 상태 일때에는 줄어든다는 자연적인 학습법칙에 기초하여 구성되는 것으로서 기본적으로는 Hebb 학습규칙과 같은 것이다. 다만 여기서는 여러가지 지연시간을 고려하므로 뉴런이 같은 상태인가 어떤가를 지연시간을 포함시킨 의미로 생각한다. 이것을 학습규칙의 형으로 쓰면 다음과 같이 된다.

$$w_{ij}^l(new) \leftarrow w_{ij}^l(old) + c_l x_i(\tau+1) x_j(\tau-l) \quad (4)$$

여기서 τ 는 학습 과정에 있어서의 이산 시각이다.

5. 연상 다이내믹스 이론

회로망의 초기 상태는 시각 $t=0$ 에서 계열 패턴의 처음 패턴 $S(0)$ 와만 상관이 있으며 그 외의 패턴과는 상관이 없는 것으로 한다. 이 상태를 초기 상태로서 회로에 설정한 다음, 그 후의 시각 $t \geq 1$ 에서의 회로의 상태와 계열 패턴과의 관계를 조사한다.

지연 시냅스 층수 L 이 2 이상인 경우 $X(1)$ 을 정하는 것은 $X(0)$ 뿐만 아니라 $X(-1), X(-2), \dots, X(-L+1)$ 도 필요할 것이나, 여기서는 시각 $t=0$ 의 경우에는 $X(t)$ ($t < 0$)는 효력을 가지지 않는 것으로 한다. 마찬가지로 일반의 시각 t 에서도 $X(t)$ ($t < 0$)는 효력을 가지지 않는 것으로 하며, 따라서 l 지연 시냅스는 시각 $t=l$ 에서 처음으로 효력을 가지기 시작하는 것으로 한다.

초기 조건으로서는 이외에도 연속하는 패턴 두개 $S(0), S(1)$ 혹은 패턴 세개 $S(0), S(1), S(2)$ 를 초기 설정하는 방법 등도 있을 수 있으나 여기서는 이러한 조건은 고려하지 않는다.

그러면 먼저 시각 t 에 있어서의 회로의 상태 $X(t) = (x_1(t), x_2(t), \dots, x_n(t))$ 와 그 시각에 상기되어야 하는 패턴 $S(t)$ 와의 거리를 나타내는 척도로서 유사도 a_t 를 다음과 같이 정한다.

$$a_t = \frac{1}{n} S(t) \bullet X(t) = \frac{1}{n} \sum_{i=1}^n s_i(t)x_i(t) \quad (5)$$

여기서 \bullet 는 내적(inner product)을 뜻한다. 유사도가 a_t 라면 패턴과 일치하고 있는 뉴런의 수는 $n/2(1+a_t)$ 이며 일치하고 있지 않은 수는 $n/2(1-a_t)$ 이다. 이 유사도 a_t 가 패턴의 수 m , 뉴런의 수 n , 지연 시냅스의 층수 L 이 주어진 조건에서 어떻게 나타나는가를 조사한다. 한편 지금부터는 n 이나 m 은 충분히 큰 반면 L 은 이들과 비교해서 충분히 작다고 가정하고서 통계적으로 해석해 간다. 이 경우, 다이내믹스는 L 과 비율 m/n 으로 규정된다. 여기서는 비율 m/n 을 r 로 표시하고 부하율 (loading ratio)로 부른다.

시각 $t, t-1, t-2, \dots, t-L+1$ 에 있어서 회로의 상태가 $X(t), X(t-1), X(t-2), \dots, X(t-L+1)$ 이며, 이러한 각각의 회로 상태는 그 시각에서 상기되어야 할 패턴 $S(t), S(t-1), S(t-2), \dots, S(t-L+1)$ 에 가깝다고 하자. 이때, a_{t+1} 즉, $X(t+1)$ 과 $S(t+1)$ 사이의 유사도는 (1)식과 (3)식을 이용하여 다음과 같이 나타낼 수 있다.

$$\begin{aligned} & \sum_{l=0}^{L-1} \sum_{j=1}^n w_{ij}^l x_j(t-l) \\ &= \sum_{l=0}^{L-1} \sum_{j=1, j \neq i}^n c_l \sum_{\mu=1}^{m-1} s_i(\mu+1) s_j(\mu-l) x_j(t-l) \\ &= \sum_{l=0}^{L-1} c_l \left\{ s_i(t+1) \sum_{j \neq i}^n s_j(t-l) x_j(t-l) \right. \\ & \quad \left. + \sum_{\mu \neq t}^{m-1} \sum_{j \neq i}^n s_i(\mu+1) s_j(\mu-l) x_j(t-l) \right\} \\ &= n s_i(t+1) \sum_{l=0}^{L-1} c_l (a_{t-l} + N_i(t, l)) \\ &= n s_i(t+1) \left(\sum_{l=0}^{L-1} c_l a_{t-l} + \sum_{l=0}^{L-1} c_l N_i(t, l) \right) \quad (6) \end{aligned}$$

단, a_{t-l} 에 있어서 $j \neq i$ 의 부분은 $1/n$ 의 영향이므로 무시한다. 또한

$$N_i(t, l) = \frac{1}{n} s_i(t+1) s_i(\mu+1) \sum_{\mu \neq t} \sum_{j \neq i} s_j(\mu-l) x_j(t-l) \quad (7)$$

식 (1)과 함수 $f(u)$ 의 정의에 따르면, 식 (6)으로 기술된 식들 가운데 마지막 식의 괄호 안이 0 이상 일 때에 $x_j(t+1) = s_j(t+1)$ 이 된다. 이러한 상황에서는 중심 극한 정리 (central limit theorem)가 성립하며, $N_i(t, l)$ 혹은 $\sum c_l N_i(t, l)$ 은 정규 분포에 따른다고 간주할 수 있다. 따라서 $\sum c_l N_i(t, l)$ 의 분산을 구하면 유사도 a_{t+1} 을 다음과 같이 구할 수가 있다.

여기서는 계산의 흐름만을 설명한다. 먼저 평균은 $s_i(t+1), s_i(\mu+1)$ 이 $x_j(t-l)$ 과 상관이 없다는 것으로부터,

$$\begin{aligned} E[N_i(t, l)] &= \frac{1}{n} E[s_i(t+1)] E[s_i(\mu+1)] \\ & \quad \sum_{\mu \neq t} \sum_{j \neq i} E[s_j(\mu-l) x_j(t-l)] \\ &= 0 \end{aligned}$$

가 된다. 분산에 대해서는, 특히 $L=1$ 의 경우에는 이미 알려져 있는 것과 같이,

$$\begin{aligned} V[N_i(t, 0)] &= E[N_i(t, 0)^2] \\ &= r + 4\gamma (a_{t-1}/\sigma_{t-1})^2 \quad (8) \end{aligned}$$

$$\sigma_{t-1} = \sqrt{V[N_i(t-1, 0)]}$$

이며, $L=1$ 이므로 l 은 0을 취한다:

$$\text{단, } \gamma(u) = \frac{1}{\sqrt{2\pi}} \exp(-u^2/2). \quad (9)$$

이것을 똑같은 방법으로 확장하면,

$$\begin{aligned} V[N_i(t, l)] &= E[N_i(t, l)^2] \quad (10) \\ &= r + 4\gamma \left(\sum_{k=0}^{L-1} c_k a_{t-l-k-1} / \sigma_{t-l-1} \right)^2, \end{aligned}$$

$$\sigma_{t-l-1} = \sqrt{V \left[\sum_{k=0}^{L-1} c_k N_i(t-l-1, k) \right]}$$

이 되는 것은 쉽게 알 수 있다. 따라서 가령 $N_i(t, l)$ 과 $N_i(t, l')$ ($l \neq l'$)의 공분산 (covariance)이 0, 다시 말해 다른 지연 시냅스 층으로부터의 입력들 사이에 상관성이 없다면 각 l 에 대하여 c_l 의 크기를 주어서 $V[N_i(t, l)]$ 의 최종적인 분산 σ_l^2 를 구할 수 있다.

그렇지만 이와 같이 구해진 상태천이식에서는 회로의 기억 능력은 L 에 비례해서 증대하며 수치 실험에서 확인되었던 것 같은 기억 능력의 포화성은 재현할 수 없었다. 실제, 다른 지연 시냅스층에서의 입력의 공분산 $cov[N_i(t, l), N_i(t, l')] = E[N_i(t, l)N_i(t, l')]$ (단, $l \neq l'$)을 계산하면 양의 값이 되는 것을 알 수 있다. 그 결과를 포함해서 초기의 과도적인 지연 시냅스 층이 가지는 효과도 고려하면 최종적으로 다음과 같은 상태천이식이 구해진다.

$$a_{l+1} = \phi\left(\frac{A_l}{\sigma_l}\right) \tag{11}$$

여기서,

$$\phi(u) = \int_{-u}^u \gamma(t) dt$$

$$A_l = \sum_{l'=0}^{L-1} c_l a_{l-l'}$$

$$\sigma_l^2 = \sum_{l'=0}^{L-1} \sum_{l''=0}^{L-1} c_l c_{l'} v_{l-l', l-l''}$$

$$v_{l-l', l-l''} = r \delta_{l, l'} + 2rc_{l-l'-1} \overline{\gamma_{l-l'-1}} + 2rc_{l-l''-1} \overline{\gamma_{l-l''-1}}$$

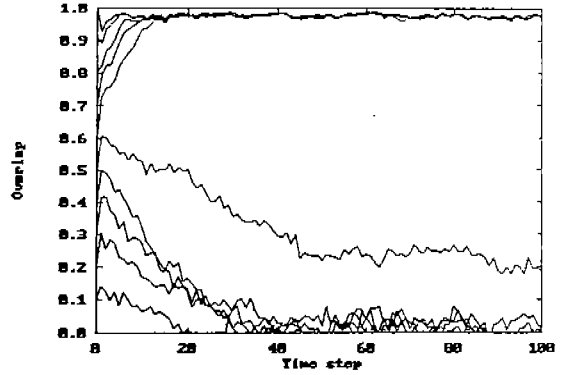
$$+ 4\overline{\gamma_{l-l'-1}} \overline{\gamma_{l-l''-1}} \sum_{k=0}^{L-1} \sum_{k'=0}^{L-1} c_k c_{k'} v_{l-l-k-1, l-l'-k'-1}$$

$$\overline{\gamma_k} = \gamma\left(\frac{A_k}{\sigma_k}\right) / \sigma_k$$

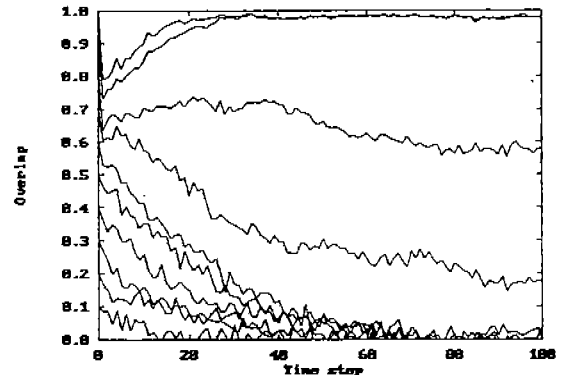
단, $\delta_{l, l'}$ 는 크로네커 델타(Kronecker delta)이고 초기 조건은 $a_0 = r$, $V_{0,0} = r$ 로 두며 이외에는 모두 0이다. 또한 모든 변수와 파라메터 a_k , $\overline{\gamma_k}$, v_k, k' , c_k 는 짐자중 적어도 하나가 음의 값이면 0으로 정의한다.

6. 상태천이식과 수치 실험의 비교

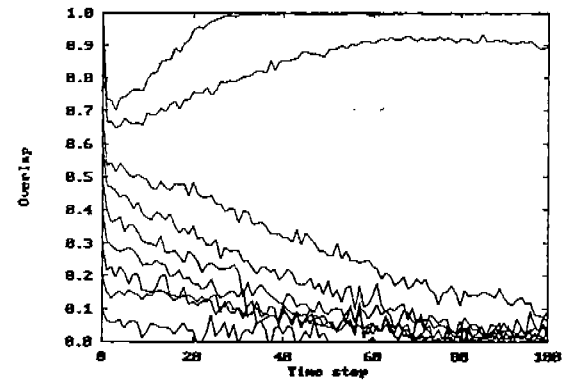
회로망의 다이내믹스는 초기 조건과 패턴상호간의 상관 및 지연 시냅스의 층수에 따라 달라진다. 이러한 상기 다이내믹스에 대하여 수치실험한 결과를 (그



(a) $l = 1, m = 175, r = 0.35$

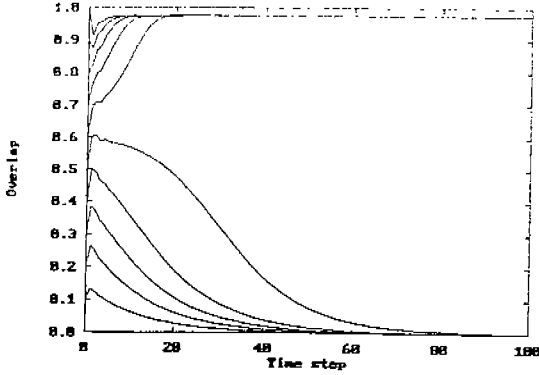


(b) $l = 3, m = 325, r = 0.65$

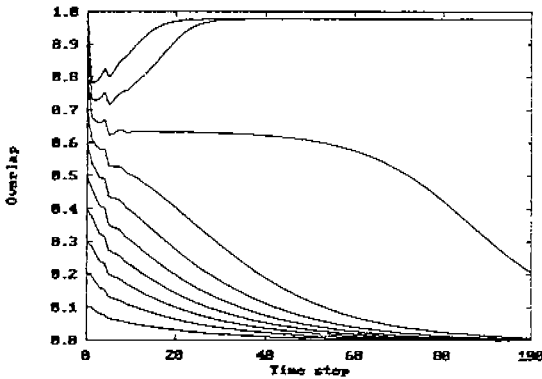


(c) $l = 6, m = 425, r = 0.85$

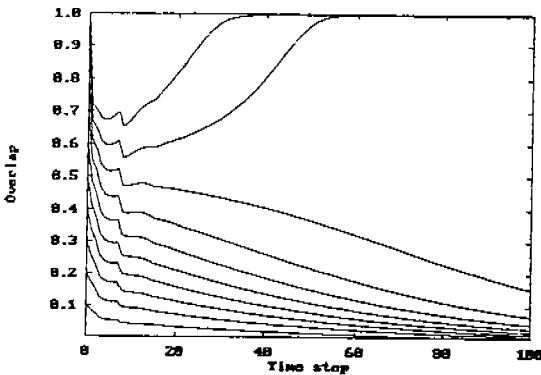
(그림 2) 수치실험에 의한 다이내믹스의 천이 (Fig. 2) Dynamic retrieval of the simulation



(a) $l = 1, r = 0.35$



(b) $l = 3, r = 0.65$



(c) $l = 6, r = 0.85$

(그림 3) 해석식에 의한 다이내믹스의 천이
(Fig. 3) Dynamic retrieval of the theory

림 2)에 나타낸다. (그림 2 (a))는 지연 시냅스 총 $l=1$ 인 회로망에서의 상기 다이내믹스 특성이며, 이때 $n=500, m=175$ 로서 부하율 $r=0.35$ 이다. 또한 각 실험결과는 10번의 평균을 취한 값이며, 가로축은 시간 t 를 세로축은 유사도 a_t 를 나타낸다. 따라서 그림에 나타난 각 선들은 여러가지 초기값으로 동작하는 계열연상의 다이내믹스를 나타내 보이고 있다. (그림 2 (b))는 $l=3, m=325$, (그림 2 (c))는 $l=6, m=425$ 일때의 상기 다이내믹스를 각각 나타내고 있다.

(그림 2)에서 초기 유사도 a_0 가 회로망의 역치값 $a_{th}(r, L)$ 보다 크면 상기 특성은 높은 유사도를 나타내는 값 1로 접근하며, $a_{th}(r, L)$ 보다 작으면 상기 특성은 0를 향해서 감소해 가는 것을 볼 수 있다. 이것은 회로망의 상태가 그 시점에 상기되어야 할 기억패턴과 전체적으로 다르다는 것을 나타낸다. 만약 초기 유사도 a_0 가 $a_{th}(r, L)$ 보다 크다면 계열 패턴중 상기되어야 할 기억패턴의 상기는 상대적으로 빠르게 수행될 것이다.

(그림 3)에 상태천이식으로 계산한 회로망의 다이내믹스를 나타내었다. 각각의 동작 조건은 (그림 2)에서의 수치 실험과 동일한 조건하에서 수행되었다. 이러한 상기 다이내믹스의 상태천이식에 의한 계열 패턴 상기 결과는 수치실험과 거의 일치하고 있음을 볼 수 있다.

여기서 설치 시뮬레이션의 상황은 열역학적 극한 (여기서는 $N \rightarrow \infty, t \rightarrow \infty$)과 거리가 있으므로, 수치 실험에서 뉴런의 수를 늘리거나 평균치를 확대하여 회로망의 다이내믹스를 구하면 더욱 좋은 결과를 얻을 수 있다.

7. 결 론

지연 시냅스를 가진 계열 연상형 신경회로망 모델의 다이내믹스를 이론적으로 조사, 지연 시냅스총과 연상 성능의 관계를 규명하였다. 그것에 의해 시냅스의 다중화가 반드시 그에 비례해 성능 향상을 가져다 주는 것은 아니라는 것을 알 수 있었다.

또한 회로망의 다이내믹스는 초기 조건, 패턴 상호간의 상관, 지연 시냅스 총수 등에 따라 달라지며, 초기 유사도의 값에 따라 정확한 연상이 수행되는 범위 및 각 지연 회로망의 역치값을 알 수 있다. 또한 신경

통계학적 방법을 도입한 상기 다이내믹스 해석을 통하여 회로망의 성능 평가를 수치해석적 방법으로 실행할 수 있음은 물론이거니와 이러한 이론해석의 결과를 통하여 신경회로망 모델의 보편적 성질을 이해할 수 있다.

앞으로는 이 모델의 기억용량, 베이슨(basin)의 크기, 초기 조건이 다를 경우 및 계열이 복수개 있는 경우 등에 대하여 그 특성을 조사할 것이다. 또한 각 지연 시냅스층의 크기 c_i 를 바꾸어 보는 것과 시냅스가 전결합이 아닌 경우의 영향에 대해서도 구체적으로 조사해 볼 필요가 있다.

참 고 문 헌

[1] E. Rumelhart, G. Hinton, "Learning internal representation by Error Propagation", in PDP, Vol. 1. MIT press, 1986.
 [2] T. S. Wilkinson, D. Mighell and J. Goodman, "Backpropagation and its Application to Hand-written Signature Verification", IEEE Conference on Neural Information Processing Systems-Nature and Synthetic, 1989.
 [3] T. Sejnowski, C. Rosenberg, "Parallel networks that learn to pronounce English text", Complex Systems, Vol. 1, 1987.
 [4] S. Amari, "Learning pattern and pattern sequence by self-organizing nets of threshold elements", IEEE Trans. Computer, C-21, pp. 1197-1206, 1972.
 [5] S. Amari, "Associative memory and its statistical neurodynamical analysis", Neural and Synergetic Computers, H. Haken ed., New York, Springer Series in Synergetics, Vol. 42, pp. 85-99, 1988.
 [6] D. Garduer, "Noise modulation of synaptic weights in a biological neural network", Neural Networks, Vol. 2, pp. 69-76, 1989.
 [7] E. Domany, W. Kinzel and R. Meir, "Layered neural networks", J. Phys. A:Math. Gen., Vol. 22, pp. 2018-2102, 1989.
 [8] K. Fukushima, "A model of associative memory in the brain", Kybernetik, Vol. 12, pp. 58-63, 1973.
 [9]. E. Caianiello, "Outline of a theory of thought

processes and thinking machines", J. Theoretical Biology, Vol. 2, pp. 204-235, 1961.
 [10] J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities", Proceedings of Nat'l. Acad. Sci., Vol. 79, pp. 2554-2588, 1982.
 [11] R. Meir, E. Domany, "Exact solution of a layered neural network model", Phys. Rev. Lett., Vol. 59, pp. 359-362, 1987.
 [12] H. Yanai, Y. Sawada, "Associative memory network composed of neurons with hysteretic property", Neural Networks, Vol. 3, pp. 223-228, 1990.
 [13] Sompolinsky, I. Kanter, "Temporal association in asymmetric neural networks", Phys. Rev. Lett., Vol. 57, pp. 2861-2864, 1986.
 [14] Kleinfeld, "Sequential state generator by model neural networks", Proc. Nat'l. Acad. Sci. USA, Vol. 81, pp. 9496-9473, 1986.
 [15] E. S. Kim, Y. Kondou, Y. Sawada, "Retrieval ability of sequential patterns in a multi-delayed feedback neural network", IEICE technical report, NC 91-162, pp. 229-233, 1992.
 [16] M. Bartholomeus, A. Coolen, "Sequences of smoothly correlated patterns in neural networks with random transmission delays", Biol. Cybern. Vol. 67, 1992.



김응수

1977년 부산대학교 전자공학과 졸업(학사)
 1979년 부산대학교 대학원 전자공학과 졸업(공학석사)
 1993년 Tohoku Univ.(일) 공학연구과 졸업(공학박사)
 1986년~1987년 Waseda Univ. 객원연구원

1984년~1993년 한국전자통신연구소 책임연구원
 1994년~현 재 선문대학교 전자공학과 교수
 관심분야:(Neural Network)

Statistical Physics of neural network,
 Chaos and fractal theory in neural network,
 Modeling biological function with neural network,
 Learning and generalization algorithm