

BSB 신경망을 위한 최적 설계 방안

An Optimal Design Procedure for Brain-state-in-a-box Neural Network

임 영 희*, 박 대 희*, 박 주 영**

Younghee Im*, Daihee Park*, Jooyoung Park**

※이 논문은 1996년도 한국학술진흥재단의 공모과제 연구비에 의하여 연구되었음.

요 약

본 논문에서는 BSB 신경망을 위한 최적의 설계 방안을 제시하고자 한다. 제안된 방법은 크게 해공간의 매개변수화와 진화 프로그램을 이용한 매개변수의 최적화과정으로 나뉜다. 또한 DOA 근사해석에 기초한 성능지수는 대규모 BSB 신경망으로의 적용을 가능하게 한다.

ABSTRACT

This paper presents an optimal design procedure to realize an BSB neural networks by means of the parametrization of solution space and optimization of parameters using evaluation program. In particular, the performance index based on DOA analysis may make an associative memory implementation reach on the level of practical success.

I. 서 론

뉴론(neuron)의 상호 결합에 의한 폐환 시스템(feed-back system)이 Hebb 규칙에 의해 학습되면 새로운 개념의 기억장치로 활용될 수 있음[1]이 Hopfield에 의해 보여진 이후, 연상 메모리(associative memory)는 신경망의 주요 응용 분야의 하나로 자리잡아 왔다. 그러나 Hopfield 신경망[1]은 집합적 계산 능력(collective computation capability)과 대규모 병렬 처리(massively parallel processing)등과 같은 많은 장점에도 불

구하고, 뉴론 사이의 연결 강도(weight)가 대칭(symmetric)이고, 각각의 원형 패턴(prototype pattern)들이 전체 시스템의 고정점(fixed point)으로 저장되는 것을 보장하지 못하는 등 현실성 있는 연상 메모리로 사용되기에는 그 역량이 아직 부족한 실정이다. 따라서 효과적인 연상 메모리의 구현을 위해 여러 가지 신경망 모델 및 연결강도 결정 방법들이 제시되어 왔다[2-10]. 다음은 이러한 기존의 연구들 중 본 논문과 직접적인 관련이 있는 연구들이다.

Yen 등에 의해 제안된 eigenstructure 방법[2, 3, 4]은 각각의 원형 패턴들을 점근적 안정인 평형점(equilibrium point)으로 기억할 뿐 아니라, 전체 시스템의 전역적 안정성(global stability)까지 보장해 준다. 또

*고려대학교 전산학과

**고려대학교 제어계측공학과

한 신경망의 연결 강도를 찾는 과정에서 두 개의 매개변수 τ_1, τ_2 를 사용하므로 이들에 대한 적절한 선택으로부터 DOA(domain of attraction) 등의 성능 지수도 어느 정도 향상시킬 수 있다. 실제로 eigenstructure 방법은 성능면에서 기존의 outer product 방법[1], projection learning rule[5] 등 보다 우수함이 입증되었다. 그러나 이 방법은 원형 패턴의 성공적인 되부름(recall)과 전역적 안정성을 보장하기 위한 절차의 확립 과정에서 단 두 개의 매개변수(τ_1, τ_2)만을 허용했기 때문에, 이들을 보장하는 전체의 해공간에 대한 광범위한 탐색이 이루어질 수 없고, 연결 강도 행렬이 대칭이어야 한다는 문제점을 안고 있다.

Zak 등은 BSB(brain-state-in-a-box) 신경망에 대한 동특성 해석[6]과 BSB 신경망을 이용한 연상 메모리의 설계 방법론[7]을 제시하였다. 이들에 의해 제시된 설계 방법론은 Yen 등의 eigenstructure 방법과는 달리 연결 강도 행렬이 대칭이어야 한다는 제약 조건이 불필요하고, 각 원형 패턴에 대한 점근적 안정성(asymptotical stability)이 보장되며, 의사 원형 패턴(spurious state)의 생성이 최대한 억제된다. 그러나 이들은 연상 메모리의 설계 과정에서 연상 메모리의 기본 요건 중 하나인 각 평형점에 대한 DOA에 대해서는 고려하지 않았다는 단점을 갖는다.

보다 효과적인 연상 메모리의 구현을 추구하는 이러한 방향의 연구들이 궁극적으로 실용성 있는 대안을 제시할 수 있기 위해서는, 다음과 같은 연상 메모리의 기본 요건들을 충족시킬 수 있어야 한다[8, 9]:

- ① 각각의 원형 패턴들이 점근적 안정인 평형점(asymptotically stable equilibrium point)으로 기억될 수 있어야 한다.
- ② 비대칭 상호 연결 구조(nonsymmetric interconnection structure)를 가져야 한다.
- ③ 저장된 각 평형점에 대한 DOA가 조절 가능하고 충분히 커야 한다.
- ④ 원형 패턴이 아닌 고정점, 즉 의사 원형 패턴의 수가 적어야 한다.
- ⑤ 그밖에 전역적 안정성, 학습 및 망각 능력(learning & forgetting capability), 고저장율(high storage efficiency) 등을 갖추어야 한다.

신경망을 이용한 연상 메모리의 구현을 다루는 대부분의 연구들은 위의 여러 가지 조건중 일부 항목에

만 초점을 맞추어 적당한 해(solution) 중 하나를 찾아내는 방식을 취하고 있다. 그러나 기존의 접근 방식들과는 달리, 설계 과정에서 일부 조건을 만족하는 해집단을 매개변수화(parametrization)할 수 있고, 이를 최적화할 수 있다면 연상 메모리의 성능을 향상시킬 수 있다.

본 논문에서는 연상 메모리의 설계 문제에 있어 중요한 위치를 차지하고 있는 BSB 신경망을 이용하여 위의 주요 기본 요건들(①~④)을 모두 만족하는 연상 메모리를 구현하고자 한다. 즉 연상 메모리의 설계 문제를 일종의 다목적 최적화 문제(multi-objective optimization problem)로 이해하고, 우선 위의 기본 요건 ①과 ②를 보장하는 해의 집합을 매개변수화하고, 나머지 요건들은 진화 프로그램(evolution program)을 이용한 매개변수의 최적화 과정에서 충족시킴으로써 보다 완성된 연상 메모리를 구현하고자 한다. 본 논문에서 제안된 설계 방법은 크게 i)해공간의 매개변수화와 ii)진화 프로그램을 이용한 우수해 탐색의 두 과정으로 구분할 수 있다. 매개변수화 단계에서는 원형 패턴들이 점근적 안정인 평형점으로 저장되고, 연결 강도 행렬이 비대칭 구조를 갖도록 연결 강도 행렬과 바이어스 벡터(bias vector)들로 이루어진 해공간을 다루기 쉬운 형태로 매개변수화한다. 이어 탐색 단계에서는 최근 매개변수와 목적 함수사이에 미분 가능성 등이 확보되기 어려운 문제의 최적화에 널리 사용되고 있는 진화 프로그램을 이용하여 각 평형점의 DOA는 크고, 의사 원형 패턴의 수는 최소가 되도록 매개변수를 최적화한다. 이때 퍼지 시스템의 동특성 해석에도 유용하게 사용된 바 있는 unravelling algorithm[11]을 응용한 상태 전이 과정 해석을 통해 각 평형점의 DOA와 의사 원형 패턴의 수를 정확히 구할 수 있다. 그러나 이러한 DOA 예측 방법은 뉴런의 수가 많은 대규모의 신경망의 경우, 상태 공간 내의 정점의 수가 지나치게 많아지므로 현실적인 적용이 어렵다는 문제점을 안고 있다. 따라서 본 논문에서는 각 원형 패턴이 갖는 DOA를 구하기 위해 연결 강도 행렬의 원소들이 DOA의 크기에 미치는 영향을 직접 파악하여 DOA 측면에서 유리한 방향으로 연결 강도를 조정함으로써 대규모 신경망에서의 연상 메모리 구현 가능성도 함께 제시하고자 한다.

본 논문의 구성은 다음과 같다:1장 서론에 이어 2

장에서는 본 논문에서 연상 메모리의 구현을 위해 사용되는 BSB 신경망에 대해 기술하고, 3장에서는 연결 강도 행렬이 비대칭 구조를 가지며, 각각의 원형 패턴이 점근적 안정인 평형점으로 저장됨을 보장하는 해공간의 매개변수화에 대해 설명한다. 4장에서는 각 평형점의 DOA 예측을 위한 두 가지 방법론, 상태전이 과정 해석과 DOA 근사해석에 대해 설명하고, 이어 5장에서는 진화 프로그램을 이용한 매개변수의 최적화 과정을 다룬다. 6장에서는 제안된 방법의 유용성을 보이기 위한 실험 방법 및 결과를 분석하고, 마지막으로 7장에서는 결론 및 향후 연구과제에 대해 논한다.

II. BSB 신경망

1977년 Anderson에 의해 제안된 BSB 신경망은 연상 메모리의 설계 문제에 있어 중요한 위치를 차지하는 모델 중에 하나이며, n 개의 뉴론으로 구성된 BSB 신경망의 동특성은 다음의 상태 방정식으로 표현될 수 있다[10]:

$$x(k+1) = g(x(k) + \alpha Wx(k)), \quad x(0) = x_0 \quad (1)$$

여기서 $x(k)$ 는 k 번째 시간 스텝에서의 상태 벡터이고, W 는 연결 강도 행렬이며, α 는 $x(k)$ 의 하중합(weighted sum) $Wx(k)$ 가 다음 시간 스텝에서의 $x(k+1)$ 값에 영향을 주는 정도를 제어하는 양수이다. 또한 $g: R^n \rightarrow R^n$ 는 출력의 i 번째 원소 $g_i(\cdot)$ 가 다음과 같이 정의되는 선형 포화 함수(linear saturating function)이다:

$$g_i = \begin{cases} 1 & \text{if } y_i \geq 1 \\ y_i & \text{if } -1 < y_i < 1 \\ -1 & \text{if } y_i \leq -1 \end{cases} \quad (2)$$

위의 선형 포화 함수가 갖는 특성에 따라, BSB 신경망의 상태 궤적(state trajectories)은 집합 $H_n = [-1, 1]^n$ 위를 움직이게 된다. (1)식으로 표현되는 원래의 BSB 신경망은, 다른 신경망 모델과 마찬가지로 뉴론 출력값의 하중합에 바이어스(bias)를 더해 줄 경우, 더욱 다양한 함수를 나타낼 수 있게 된다. 이러한 효과를 고려하여 바이어스 벡터 b 를 도입하면 식 (1)은 다음

과 같이 변환된다:

$$\begin{aligned} x(k+1) &= g(x(k) + \alpha(Wx(k) + b)) \\ &= g((I + \alpha W)x(k) + \alpha b) \end{aligned} \quad (3)$$

식 (1)의 BSB 신경망과 구별하기 위하여, 식 (3)의 신경망 모델은 일반화된 BSB(generalized BSB, GBSB)라 불리기도 한다. 본 논문에서는 바이어스 벡터 개념이 추가된 (3)식의 모델을 BSB 신경망이라는 명칭하에 일관되게 사용한다.

III. 해공간의 매개변수화

다음과 같이 정의되는 매핑(mapping) $L: R^n \rightarrow R^n$ 과 $T: R^n \rightarrow R^n$ 을 생각하자.

$$\begin{aligned} L(x) &= (I + \alpha W)x + \alpha b \\ T(x) &= g((I + \alpha W)x + \alpha b) = g(L(x)) \end{aligned}$$

벡터 공간 R^n 에서의 유클리디안 노름(Euclidean norm)을 $\|\cdot\|$ 으로 표기하고, 중심 $y \in R^n$ 과 반경 $\varepsilon > 0$ 을 갖는 구를

$$B_\varepsilon(y) = \{x \in R^n : \|x - y\| < \varepsilon\}$$

와 같이 표현하면, BSB 신경망의 평형점과 안정성 및 점근적 안정성 등의 개념은 다음과 같이 정의될 수 있다.

정의 1. $[-1, 1]^n$ 위의 점 x 가 $T(x) = x$ 를 만족하면, “ x 는 BSB 신경망의 평형점이다”라고 한다.

정의 2. BSB 신경망의 평형점 x^* 가 임의의 $\varepsilon > 0$ 에 대하여

$$x(0) \in B_\delta(x^*) \Rightarrow \forall k \geq 0, x(k) \in B_\varepsilon(x^*)$$

를 만족하는 $\delta > 0$ 가 존재하면, “평형점 x^* 는 안정이다”라고 한다.

정의 3. BSB 신경망의 평형점 x^* 가 안정이고, $x(0) \in B_\delta(x^*)$ 이면 $k \rightarrow \infty$ 일 때 $x(k) \rightarrow x^*$ 인 양수 δ 가 존재하면, “평형점 x^* 는 점근적 안정이다”라고 한다.

일반적으로 (3)식으로 표현되는 BSB 신경망에 있어서 각 평형점이 점근적 안정인지를 판별하는 것은

간단한 문제가 아니지만, $[-1, 1]^n$ 위의 정점에 자리 잡은 평형점 x^* 에 대해서는 손쉬운 판별 방법이 존재한다. 즉 장점 $x^* \in \{-1, 1\}^n$ 가

$$(L(x^*))_i x_i^* > 1, \quad i = 1, \dots, n \quad (4)$$

를 만족하면, x^* 는 점근적 안정인 평형점으로 기억된다[7]. 따라서 $\{-1, 1\}^n$ 위의 한 점으로 주어지게 되는 각 원형 패턴이 조건 (4)를 만족하도록 연결 강도 행렬 W 와 바이어스 벡터 b 를 선택하면, 주어진 각각의 원형 패턴이 신경망의 점근적 안정인 평형점으로 자리잡게 됨을 보장할 수 있다. 이와 같이 원형 패턴들이 점근적 안정인 평형점으로 저장될 수 있도록 해주는 W 와 b 의 집합을 일정한 매개변수를 이용하여 표현해 주는 작업이 본 논문의 일차 목적인 해공간의 매개변수화이다.

조건 (4)로부터 상태 전이 과정을 결정해 주는 연결 강도 행렬 W 와 바이어스 벡터 b 가 다음의 조건 (5)를 만족하면 주어진 원형 패턴 $s_m, m=1, \dots, p$ 는 모두 점근적 안정인 평형점이 됨을 보장할 수 있게 된다.

$$(L(s^m))_i s_i^m > 1, \quad i = 1, \dots, n \quad m = 1, \dots, p \quad (5)$$

Zak 등[7]에 의해 수식화된, 각각의 원형 패턴들이 점근적 안정인 평형점으로 저장되는 해공간의 매개변수화 과정은 다음과 같다. 이때 뉴런의 개수는 n , 기억하고자 하는 원형패턴의 수는 p 라 하자. 행렬 B 와 V 가 각각

$$B = [b, \dots, b] \in R^{n \times p}, \quad V = [s^1, \dots, s^p] \in \{-1, 1\}^{n \times p}$$

와 같이 정의되고 V 의 pseudo-inverse가 V^+ 로 표기될 때, W 와 b 가 다음의 식 (6)을 만족하면 원형패턴 $s_m, m=1, \dots, p$ 는 점근적 안정인 평형점이 된다[7].

$$W = (DV - B)V^+ + \Lambda(I - VV^+) \quad (6)$$

여기서 행렬 D 와 Λ , 벡터 b 는

$$\forall i, d_{ii} > \sum_{j \neq i} |d_{ij}| \quad (7)$$

$$\forall i, \lambda_{ii} < -\sum_{j \neq i} |\lambda_{ij}| - |b_i| \quad (8)$$

$$\forall i, d_{ii} < \sum_{j \neq i} |d_{ij}| + |b_i| \quad (9)$$

를 만족한다. 따라서 W 와 b 의 해공간은 d_{ij} 와 λ_{ij} , 그리고 b_i 에 의해 매개변수화된다.

그러나 위의 매개변수화 과정을 BSB 신경망을 이용한 연상 메모리의 최적화 과정에 곧바로 이용할 경우 탐색 공간이 고차원이 되는 문제점이 발생한다. 이러한 문제점을 해결하기 위해 본 논문에서는 전체 해공간을 다루기 쉬운 일부 해공간으로 축소하고자 한다. 이는 행렬 D 와 Λ 를 각각

$$D = \tau_1 I, \quad (10)$$

$$\Lambda = -\tau_2 I \quad (11)$$

로 변환함으로써 가능해진다. 따라서 식 (6)과 조건 (7), (8), (9)는 다음과 같이 변형된다.

$$W = (\tau_1 V - B)V^+ - \tau_2(I - VV^+) \quad (12)$$

$$0 < \tau_1 < |b_i| < \tau_2, \quad i = 1, \dots, n \quad (13)$$

즉, 식 (12)와 조건 (13)에 의해 n 개의 뉴런으로 이루어진 BSB 신경망이 p 개의 원형 패턴들을 각각 점근적 안정인 평형점으로 저장하도록 연결 강도 행렬 W 와 바이어스 벡터 b 의 해공간을 다루기 쉬운 형태로 매개변수화하였다.

IV. 상태 전이 과정 해석과 DOA의 근사해석

BSB 신경망은 식 (12), (13)에 의해 비대칭 상호 연결 구조를 가지며, 모든 원형 패턴들이 점근적 안정인 평형점으로 저장됨이 보장된다. 그러나 DOA의 크기 및 의사 원형 패턴의 수 등의 성능은 매개변수가 조건 (13)으로 표현되는 공간내의 어떤 위치를 차지하느냐에 따라 달라진다. 따라서 각 매개변수들은 진화 프로그램을 이용한 최적화 과정에서 충분히 큰 DOA와 의사 원형 패턴의 수가 최소가 되도록 조정된다. 그러나 이를 위해서는 각 평형점의 DOA의 크기와 의사 원형 패턴의 수가 고려된 성능 지수(performance index)를 정의할 필요가 있으며, 이는 상태 전이 과정

해석과 DOA의 근사해석에 의해 가능하다.

4.1 상태 전이 과정 해석

연상 메모리의 구현을 위한 신경망의 뉴런의 수가 많지 않을 경우, $\{-1, 1\}^n$ 위의 각 정점들을 초기로 하여 $x(k)$ 의 전이 과정을 관찰함으로써 각 평형점의 DOA와 의사 원형 패턴의 수, 리미트 사이클(limit cycle)의 수 등 연상 메모리의 성능 평가를 위한 기본 자료를 제공받을 수 있다. 이러한 전이 과정의 해석을 위해 퍼지 시스템의 동특성 해석[13]에도 유용하게 사용된 바 있는 unravelling algorithm[11]을 일부 수정하여 사용할 수 있다. 즉, $\{-1, 1\}^n$ 위의 각 정점에서 출발한 상태 궤적(state trajectory)이 최종적으로 어떠한 평형점 또는 리미트 사이클에 도달하느냐에 따라 장점들을 그룹(group)으로 나누고, 각 정점 $z \in \{-1, 1\}^n$ 에는 그룹수(group number) $G(z)$ 를 할당한다. 또한 최종적으로 평형점 z^* 에 도달하는 그룹에 속하는 초기 정점 z 들에 대해서는 해밍 거리(hamming distance) $HD(z, z^*)$ 와 초기 정점들이 평형점이나 리미트 사이클에 도달하기까지의 상태 전이 횟수를 나타내는 스텝수(step number) $S(z)$, 그리고 최종적으로 도달한 주기 궤적(periodic motion)을 나타내는 주기수(periodicity number) $P(z)$ 등을 할당한다. 따라서 모든 초기 정점들에 대한 상태 전이 과정이 끝나면, $G(z)$ 에 의해 의사 원형 패턴의 수를, 그리고 $HD(z, z^*)$ 에 의해 각 평형점의 DOA를 유추해 낼 수 있다.

이렇게 구해진 평형점의 DOA와 의사 원형 패턴의 수는 연상 메모리의 성능 지수의 정의에 활용된다. 즉 진화 프로그램을 이용한 최적화 과정에서, 다음과 같은 성능지수를 사용하면 매개변수들로 이루어진 염색체들은 각 평형점의 DOA는 최대가 되고, 의사 원형 패턴의 수는 최소가 되는 방향으로 진화하게 된다.

$$J1 = c_1 \times \sum_{m=1}^b w_i N_i^m - c_2 \times F \quad (14)$$

여기서 N_i^m , F , w_i 는 다음과 같이 정의되는 양수이다:

N_i^m : 원형 패턴 s^m 과 해밍 거리 i 만큼의 거리를 유지하는 정점들 중 그 상태 궤적이 원형 패턴 s^m 으로 수렴하는 정점들의 수.

F : 의사 원형 패턴의 수.

w_i : N_i^m 에 곱해지는 무게 상수(weighting factor).

(예: $w_1 = 2, w_2 = 0.4, w_3 = 0.1, w_4, \dots, w_n = 0.05$)

또한 c_1, c_2 는 사용자의 주관 및 적용되는 모델에 따라 변화하는 무게 상수이다.

4.2 DOA의 근사해석

앞서 살펴보았듯이 신경망을 이루는 뉴런의 수가 많지 않을 경우, $\{-1, 1\}^n$ 위의 정점들 초기치로 한 $x(k)$ 의 전이 과정을 관찰함으로써 DOA의 크기, 의사 원형 패턴의 수와 같은 신경망의 성능 평가를 위한 유용한 정보를 제공받을 수 있다. 그러나 대규모의 신경망을 다룰 경우, 상태 공간내의 정점의 수가 지나치게 많아지므로 unravelling algorithm[11]을 이용한 상태 전이 과정 해석은 적용이 불가능해진다. 따라서 본 논문에서는 이러한 문제점을 해결하기 위하여 DOA를 정확하게 예측하는 대신, 연결 강도 행렬의 원소들이 DOA 크기에 미치는 영향을 직접 파악하여 DOA 측면에서 유리한 방향으로 연결 강도들을 변화시킨다. 즉 해공간의 매개변수화 과정에 이용되었던 조건 (5)에 대한 관찰로부터, DOA는 조건의 좌변의 크기가 우변보다 얼마나 큰가에 따라 영향을 받게 될 수 있다. 마찬가지로 조건 (5)에 대한 관찰로부터 연결 강도 행렬의 원소들 중 일부가 다른 항들에 비해 매우 큰 절대치를 갖는 경우에는, 큰 절대치를 갖는 항에 곱해지게 될 원형 패턴의 원소들 중 한 비트만 바뀌어도 조건 (5)가 더 이상 성립하지 않을 수도 있다. 그러므로 연결 강도 행렬의 표준 편차는 DOA에 영향을 주는 중요한 요소가 될 수 있다. 이러한 정성적인 관찰은 궁극적으로 매개변수가 DOA의 크기에 미치게 되는 영향을 수식화할 수 있게 해준다. 따라서 연상 메모리 구현을 위해 사용된 신경망이 어느 정도의 우수성을 갖는지를 정량적으로 표현해 주는 성능 지수는 다음과 같이 정의할 수 있다.

$$J2 = c_1 \times \min\{(L(s^m); s^m)\} - c_2 \times (W \text{의 표준편차}) \quad (15)$$

이때 c_1, c_2 는 사용자의 주관 및 적용 모델에 따라 변화되는 가중치이다.

이후 본 논문에서는 정의된 두 개의 성능 지수 $J1, J2$ 를 각각 사용하여 연상 메모리를 구현해 보고, 두 방법론의 장·단점을 비교·분석한다.

V. 진화 프로그램을 이용한 매개변수의 최적화

진화 프로그램의 기본 원리는 자연 생태계의 적자 생존의 원리를 모방한 유전자 알고리즘과 동일하다. 즉, 문제의 해를 표현하는 염색체(chromosome)로 구성된 개체집단(population)을 점진적으로 개선시켜 나감으로써 최적의 해를 얻게 된다. 그러나 고전적인 유전자 알고리즘에서는 문제의 해를 표현하는 염색체가 고정된 길이의 이진 문자열(fixed-length binary string)로 구성되므로 주어진 문제가 복잡할 경우 주어진 문제를 이진 문자열로 표현 가능하도록 변형해야 하는 어려움이 있다. 반면 진화 프로그램의 경우에는 주어진 문제의 변환 과정 없이 문제에 맞게 염색체를 표현하고, 그에 따라 적절히 변형된 유전 연산자를 사용함으로써 보다 복잡한 문제에 적용이 가능하다[14].

5.1 염색체의 표현

BSB 신경망을 위한 진화 프로그램의 염색체는 실수 형태이며 다음과 같이 매개변수 τ_1, τ_2, β_i 로 이루어진다.

τ_1	τ_2	β_1	...	β_n
----------	----------	-----------	-----	-----------

이때 바이어스 벡터 b_i 는 조건 (13)에 의하여 각 염색체의 τ_1 과 τ_2 값에 따라 그 범위가 달라지므로, b_i 대신 각 염색체의 유전인자에 $[-1, 1]$ 구간의 값을 가지며, 0이 아닌 β_i 를 사용한다. 따라서 바이어스 벡터 b_i 는 β_i 로부터 조건 (13)을 만족하도록 다음과 같이 결정된다.

$$b_i = \begin{cases} (\tau_2 - \tau_1)\beta_i - \tau_1, & \text{if } \beta_i < 0 \\ (\tau_2 - \tau_1)\beta_i + \tau_1, & \text{if } \beta_i > 0 \end{cases} \quad (16)$$

본 논문에서 고전적인 유전자 알고리즘에서 일반적으로 사용되는 단순 교배(simple crossover)를 사용하고, 돌연변이(mutation) 연산의 경우에는 비균등 돌연변이(non-uniform mutation)를 수행함으로써 진화 프로그램에 지역적 미세 조정 능력(local fine-tuning capability)을 제공한다[14].

5.2 유전 연산자

5.2.1 단순 교배(Simple Crossover)

교배는 두 염색체간의 임의의 교차점(cross site)에서 교차가 발생하는 것을 말하며, 교차점은 난수 발생에 의해 구한다. 예를 들어 두 염색체 $s'_v = \langle v_1, \dots, v_m \rangle$ 와 $s'_w = \langle w_1, \dots, w_m \rangle$ 가 교배를 위해 선택되고, k 번째 유전인자부터 교차가 수행된다면 두 염색체는

$$s'_v{}^{t+1} = \langle v_1, \dots, w_k, \dots, w_m \rangle \text{와 } s'_w{}^{t+1} = \langle w_1, \dots, v_k, \dots, v_m \rangle$$

로 진화된다.

5.2.2 비균등 돌연변이(Non-uniform Mutation)

비균등 돌연변이 연산은 다음과 같이 정의된다. 만약 염색체 $s'_v = \langle v_1, \dots, v_m \rangle$ 의 k 번째 유전인자 $v_k \in [l_{(k)}^v, u_{(k)}^v]$ 가 돌연변이 연산을 위해 선택되면, $s'_v{}^{t+1} = \langle v_1, \dots, v'_k, \dots, v_m \rangle$ 이 되고, v'_k 은 다음과 같다.

$$v'_k = \begin{cases} v_k + \Delta(t, u_{(k)}^v - v_k), & \text{if random digit is 0} \\ v_k - \Delta(t, v_k - l_{(k)}^v), & \text{if random digit is 1} \end{cases} \quad (17)$$

식 (17)의 함수 $\Delta(t, y)$ 는 $[0, y]$ 사이의 값을 가지며, 시간 스텝 t 가 증가함에 따라 0에 가까운 값을 갖는다.

$$\Delta(t, y) = y \cdot \left(1 - r^{(1 - \frac{t}{T})^r}\right) \quad (18)$$

이때 r 은 $[0, 1]$ 사이의 난수이고, T 는 최대 진화 횟수, b 는 비균등정도(non-uniformity)를 나타내는 시스템 매개변수이다. 따라서 비균등 돌연변이 연산은 시간 스텝 t 가 작을 때, 즉 진화 초기에는 전체 해공간에 대해 균등하게 탐색을 수행하고, 진화가 어느 정도 진행된 후에는 지역적 탐색을 수행함으로써 진화 프로그램에 지역적 미세 조정 메커니즘을 제공한다.

VI. 실험 방법 및 결과 분석

6.1 실험 방법

본 논문에서 제안된 방법의 유용성을 보이기 위해 각각 10개의 뉴론을 갖는 BSB 신경망에 6개의 원형 패턴을 저장하는 연상 메모리[7]를 구현하였다. 이때 식 (14)과 (15)의 성능 지수 $J1, J2$ 를 각각 적용하여 그

결과를 비교·분석해 본다. 기억하고자 하는 원형 패턴은 다음과 같다:

$$\begin{aligned}
 s^1 &= [-1 \ 1 \ 1 \ 1 \ 1 \ 1 \ -1 \ -1 \ -1 \ -1]^T \\
 s^2 &= [1 \ 1 \ -1 \ -1 \ -1 \ 1 \ -1 \ -1 \ 1 \ -1]^T \\
 s^3 &= [-1 \ 1 \ 1 \ 1 \ -1 \ -1 \ 1 \ -1 \ 1 \ -1]^T \\
 s^4 &= [1 \ -1 \ -1 \ 1 \ 1 \ -1 \ 1 \ 1 \ 1 \ -1]^T \\
 s^5 &= [1 \ -1 \ -1 \ 1 \ 1 \ -1 \ 1 \ 1 \ 1 \ -1]^T \\
 s^6 &= [1 \ 1 \ -1 \ 1 \ -1 \ 1 \ 1 \ 1 \ -1 \ -1]^T
 \end{aligned}$$

편의상 위의 이진 벡터를 그에 대응하는 십진수를 이용하여 표기하기로 한다. 즉 벡터의 엔트리(entry) 중 -1을 0으로 바꾼 후, 십진수로 바꾸면 위의 원형 패턴들은 각각 267, 456, 496, 622, 786, 860이 됨을 알 수 있다.

DOA 근사 해석에 사용된 성능 지수 J_2 는 다음과 같다.

$$J_2 = 10 \times \min\{(L(s^m); s_i^m) - 9.09 \times (W \text{의 표준편차})\}$$

위의 성능 지수들이 적용된 진화 프로그램의 매개변수는 다음과 같다.

집단의 크기	50
최대 진화 횟수	1000
교배 확률	0.5
돌연변이 확률	$\tau_1, \tau_2: 0.1$
	b_i

또한 초기 집단의 각 염색체는 난수 발생에 의해 $\tau_1 \in [3, 12], \tau_2 \in [7.5, 30], \beta_i \in [-1, 1], \beta_i \neq 0$ 이 되도록 초기화하였다.

6.2 실험 결과 분석

다음은 성능 지수 J_2 를 이용한 BSB 신경망이 충분한 크기의 DOA를 갖도록 최적화된 매개변수이다.

$$\tau_1 = 7.047$$

$$\tau_2 = 9.166$$

$$b = [-7.06 \ 7.05 \ -7.09 \ 7.12 \ 7.05 \ 7.05 \ 7.06 \ -7.05 \ 7.05 \ -7.05]^T$$

다음의 표 1은 Zak 등[7]에 의해 제시된 연결 강도 행렬 W 와 바이어스 벡터 b 를 BSB 신경망에 적용한 결과 발생한 각 평형점에 대한 DOA이다. 표 2는 상태의 전이 과정을 모두 관찰하여 성능 지수 J_1 을 구한 후 유전자 알고리즘을 적용한 결과이고, 표 3은 DOA 근사해석에 기반을 둔 성능 지수 J_2 를 이용한 결과값이다.

표 1. Zak[7]등에 의해 설계된 BSB 신경망의 DOA

HD	0	1	2	3	4	5	6	7	8	9	10
원형 패턴	267	1	4	0	0	0	0	0	0	0	0
	456	1	10	16	15	12	7	0	0	0	0
	496	1	4	2	1	0	0	0	0	0	0
	622	1	6	4	1	0	0	0	0	0	0
	786	1	7	6	3	0	0	0	0	0	0
	860	1	7	5	13	13	8	3	1	0	0

표 2. 성능 지수 J_1 를 이용한 BSB 신경망의 DOA

HD	0	1	2	3	4	5	6	7	8	9	10
원형 패턴	267	1	10	37	44	31	15	1	1	0	0
	456	1	10	35	43	26	3	1	0	0	0
	496	1	10	32	22	8	3	0	1	0	0
	622	1	10	41	56	37	11	0	0	0	0
	786	1	10	35	29	11	0	0	0	0	0
	860	1	10	32	26	5	1	0	0	0	0

표 3. 성능 지수 J_2 를 이용한 BSB 신경망의 DOA

HD	0	1	2	3	4	5	6	7	8	9	10
원형 패턴	267	1	10	15	10	3	0	0	0	0	0
	456	1	10	15	5	1	0	0	0	0	0
	496	1	10	29	38	27	11	0	0	0	0
	622	1	10	28	28	11	2	0	0	0	0
	786	1	10	16	10	2	0	0	0	0	0
	860	1	10	16	2	3	0	0	0	0	0

실험 결과 본 논문에서 제안된 방법의 경우 Zak 등 [7]의 결과보다 DOA가 현저히 커졌음을 알 수 있다. 즉, 논문에서 제안된 방법이 전체 해공간을 다루기 쉬운 일부 해공간으로 축소하였음에도 불구하고, Zak 등의 방법보다 우수한 성능을 나타낸다. 이는 진화 프로그램을 사용함으로써, 단일 해가 아닌 해집단을 이용한 우수해 탐색과 Zak 등의 방법에서는 고려되지 않았던 DOA와 의사 원형 패턴의 적절한 고려 때문이다. 그리고 모든 상태의 전이과정을 빠짐없이 관찰하는 대신, DOA 근사해석을 이용하여 최적화를 수행한 경우(표 3)의 성능이 기존의 결과(표 1)보다 우수함은 매우 고무적이다. 왜냐하면 대규모의 신경망에서는 상태 전이 과정의 관찰이 불가능하므로 DOA 근사해석에 의존하는 표 3의 방법론이 유일한 대안이 되기 때문이다. 한편, 의사 원형 패턴의 경우에는 세 가지 방법 모두 같은 결과(2개)를 제공한다.

다음의 표 4는 Yen 등[3, 4]에 의해 제안된 eigenstructure 방법과 본 논문에서 제안된 성능 지수 J 를 이용한 방법의 실험 결과를 1024개 초기 정점들에 대하여 이 정점들이 최종적으로 어떠한 위치에 도달하는가를 관찰하여 구분한 것이다.

표 4. Eigenstructure 방법과 본 논문에서 제안된 방법과의 비교

	A 그룹	B 그룹	C 그룹
• 성능 지수 J_1 을 이용한 BSB 신경망	608	45	371
• 성능 지수 J_2 를 이용한 BSB 신경망	338	92	594
• 망각 알고리즘을 사용하지 않은 eigenstructure 방법	240	102	682
• 망각 알고리즘을 사용한 eigenstructure 방법	338	114	572

A 그룹은 1024개의 초기 정점들 중 해밍 거리가 가장 가까운 원형 패턴으로 수렴하는 정점들의 수이고, B 그룹은 원형 패턴으로 수렴하되 최종적으로 도달한 정점이 초기치에 가장 가까운 원형 패턴이 아닌 정점들의 수이며, C 그룹은 의사 원형 패턴 또는 리미트 사이클에 도달하는 정점들의 수이다. 따라서 A 그룹의 크기가 클수록, C 그룹의 수가 작을수록 우수

한 연상 메모리라 할 수 있다.

표 4를 통해 본 논문에서 제안된 방법이 eigenstructure 방법에 비해 우수한 연상 메모리를 구현할 수 있음을 알 수 있다. 이는 연상 메모리의 설계 과정에 해공간의 매개변수화와 진화 프로그램을 이용한 매개변수의 최적화 과정을 포함할 경우, 상태 전이 과정 해석을 이용한 방법과 DOA 근사해석을 이용한 방법 모두 망각 알고리즘이 포함된 eigenstructure 방법보다도 우수한 성능을 얻을 수 있음을 보여준다.

Ⅶ. 결론 및 향후 연구과제

본 논문에서 연상 메모리의 설계 문제를 일종의 다목적 최적화 문제로 이해하고, BSB 신경망에 진화 프로그램을 적용하여 보다 완성된 연상 메모리를 구현하였다. 본 논문에서 제안된 방법은 크게 해공간의 매개변수화와 진화 프로그램을 이용한 우수해 탐색의 두 과정으로 나눌 수 있다. 우선 비대칭 상호 연결 구조를 가지며, 각각의 원형 패턴이 점근적 안정인 평형점으로 저장되도록 하는 해공간을 다루기 쉬운 형태로 매개변수화하고, 이후 이러한 매개 변수들을 DOA의 크기와 의사 원형 패턴의 수 등을 고려한 성능 지수를 이용한 진화 프로그램에 의해 DOA는 충분히 크게, 의사 원형 패턴의 수는 최소가 되도록 최적화하였다. 그리고 정확한 DOA의 계산 대신, 연결 강도 행렬 W 의 원소들이 DOA 크기에 미치는 영향을 파악, 진화 프로그램을 이용하여 DOA 측면에서 유리한 방향으로 연결 강도를 조정함으로써 대규모의 BSB 신경망을 위한 설계 최적화가 가능하도록 하였다. 또한 실험을 통하여 본 논문에서 제안된 방법이 Zak 등에 의해 제안된 방법과 eigenstructure 방법에 비해 우수함을 보였다.

이후 연구에서는 DOA 근사해석과정에서 직접적으로 고려할 수 없었던 의사 원형 패턴의 수를 망각 알고리즘(forgetting algorithm[3])을 통해 최소화하는 방법을 생각해 볼 수 있다.

참 고 문 헌

1. J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abili-

ties," in Proc. Nat. Acad. Sciences, vol. 79, pp. 2554-2558, 1982.

2. A. N. Michel, J. Si, and G. Yen, "Analysis and synthesis of a class of discrete-time neural networks described on a closed hypercube," IEEE Trans. Neural Networks, vol. 2, pp. 32-46, 1991.
3. G. G. Yen and A. N. Michel, "A learning and forgetting algorithm in associative memory: the eigen-structure method," IEEE Trans. Circuit System, vol. 39, pp. 212-225, 1992.
4. G. G. Yen, "Unlearning algorithm in associative memory," Proc. of World Congress on Neural Networks, pp. 1466-1470, 1995.
5. L. Personnaz, I. Guyon, and G. Dreyfus, "Information storage and retrieval in spin-glass like neural networks," J. Physique Letters, vol. 46, pp. L359-L365, 1985.
6. S. Hui and S. H. Zak, "Dynamical analysis of the brain-state-in-a-box(BSB) neural models," IEEE Trans. Neural Networks, vol. 3, pp. 86-94, 1992.
7. W. E. Lillo, D. C. Miller, S. Hui and S. H. Zak, "Synthesis of brain-state-in-a-box(BSB) based associative memories," IEEE Trans. Neural Networks, vol. 5, pp. 730-737, 1994.
8. A. N. Michel and J. A. Farrell, "Associative memories via artificial neural networks," IEEE Contr. Syst. Mag., vol. 10, pp. 6-17, 1990.
9. G. G. Yen and A. N. Michel, "A learning and forgetting algorithm in associative memories: Results involving pseudo-inverses," IEEE Trans. Circuits System, vol. 38, pp. 1193-1205, 1991.
10. J. A. Anderson, J. W. Silverstein, S. A. Ritz, and R. S. Jones, "Distinctive features, categorical perception, and probability learning: Some applications of a neural model," in Neurocomputing: Foundations of Research, J. A. Anderson and E. Rosenfeld, Eds. Cambridge, MA: MIT Press, 1988.
11. C. S. Hsu and R. S. Guttalu, "An unravelling algorithm for global analysis of dynamical systems: An application of cell-to-cell mappings," ASME Journal of Applied Mechanics, vol. 47, pp.

940-948, 1980.

12. 김대수, **신경망 이론과 응용 I, II**, 하이테크 정보, 1993.
13. Y. Y. Chen and T. C. Tsao, "A description of the dynamical behavior of fuzzy systems," IEEE Trnas. Syst., Man and Cybernetics, vol. 19, pp. 745-755, 1989.
14. Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, 2nd ed. Springer-Verlag, New York, 1992.
15. 윤성식, 박주영, 박대희, "해공간의 매개변수화와 유전자 알고리즘을 이용한 BSB 신경망의 설계," 대한전자공학회 논문지(B), No. 2, pp. 178-186, 1996.

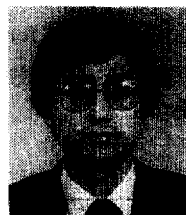


임 영 희(Younghee Im) 정회원
 1994년: 고려대 전산학과(학사)
 1996년: 고려대 전산학과(석사)
 1996년~현재: 고려대학교 전산학과 시간강사
 1997년 9월: 고려대 전산학과 박사과정 입학
 ※주관심분야: 퍼지시스템, 신경망, 유전자 알고리즘 등임



박 대 희(Daihee Park) 정회원
 1982년: 고려대 수학과(학사)
 1984년: 고려대 수학과(석사)
 1989년: 플로리다 주립대 전산학과(석사)
 1992년: 플로리다 주립대 전산학과(박사)
 1993년~현재: 고려대 전산학과 조교수

※주관심분야: 퍼지 시스템, 신경망, 유전자 알고리즘, 데이터베이스 등임



박 주 영(Jooyoung Park) 정회원
 1983년: 서울대 전기공학과(학사)
 1992년: 텍사스대 전기 및 컴퓨터공학과(박사)
 1993년~현재: 고려대 제어계측공학과 조교수
 ※주관심분야: 신경망, 제어 및 시스템 이론 등임