

논문 2007-44CI-2-2

패리티 판별을 위한 유전자 알고리즘을 사용한 신경회로망의 학습법

(Learning method of a Neural Network using Genetic Algorithm for 3
Bit Parity Discrimination)

최재승*, 김정화**

(Jae-Seung Choi and Chung-Hwa Kim)

요약

신경회로망의 학습에 널리 사용되고 있는 오차역전파 알고리즘은 최급하강법을 기초로 하고 있기 때문에 초기값에 따라서는 극소값에 떨어지거나, 신경회로망을 학습시킬 때 중간층 유닛수를 얼마로 설정하는 등의 문제점이 있다. 따라서 이러한 문제점을 해결하기 위하여, 본 논문에서는 3비트 패리티 판별을 위하여 신경회로망의 학습에 교차법, 돌연변이법에 새로운 기법을 도입한 개량형 유전적 알고리즘을 제안한다. 본 논문에서는 세대차이, 중간층 유닛수의 차이, 집단의 개체수의 차이에 대하여 실험을 실시하여, 본 방식이 학습 속도의 면에서 유효하다는 것을 나타낸다.

Abstract

Back propagation algorithm based on a gradient-decent method has been widely used to the training of a neural network. However, this algorithm have some problems such as dropping the minimum value in a local area according to an initial value and setting the number of units in a hidden layer when training the neural network. Accordingly, to solve the above-mentioned problems, this paper proposes a genetic algorithm using the training method of the neural network. Thus, the improved genetic algorithm using a new crossover and mutation method is proposed to discriminate 3 bit parity. Experiments confirm that the proposed system is effective for training speed after demonstrating for generation gap, the number of units in the hidden layer, and the number of individuals.

Keywords : Genetic algorithm, neural network, back propagation algorithm, parity discrimination,
the number of units in a hidden layer.

I. 서론

신경회로망(Neural Network: NN)은 비선형 근사능력 및 학습능력을 가지고 있으며 다양한 분야로 이에 대한 응용이 광범위하게 연구되고 있다^[1~2]. 또한, 신경회로망은 새로운 학습법의 능력판별을 위하여 XOR 회

로 및 패리티 판별 회로와 같은 기준회로로서 자주 사용되고 있다^[3~4]. 계층형 NN의 학습에 광범위하게 사용되고 있는 오차역전파 (Back Propagation : BP) 알고리즘^[5]은 네트워크가 초기 난수(random number)에 의존하여 수렴율이 저하되는 등의 문제 및 학습을 종료시키는 타이밍이 경험적으로 실시되어지는 문제들이 있다. 또한 BP 학습법은 최급하강법을 기초로 하고 있기 때문에 초기값에 따라서는 극소값에 떨어지는 경우가 있으며, NN을 학습시킬 때 중간층 유닛수를 얼마로 결정하는 등이 항상 문제가 된다. 중간층이 적으면 학습이 충분히 되지 않으며, 또한 많으면 시간이 걸린다. 따라서 본 논문은 계층형 NN의 학습에 유전자 알고리즘

* 정회원, 신라대학교 전자공학과
(Department of Electronics Engineering, Silla University)

** 정회원, 조선대학교 전자공학과
(Division of Electronics Engineering, Chosun University)

접수일자: 2006년6월29일, 수정완료일: 2007년1월1일

(Genetic Algorithm: GA)을 사용하여 돌연변이의 일 형태로서 유닛수의 증감의 기능을 수용함으로써 이 문제가 해결 가능한 것을 제안한다.

GA는 Holland에 의해서 고안된 계산이론이며^[6], 공학적 시스템에 대해서 상당히 강력한 최적화 능력을 가지고 있으며 최근 다양한 분야에서 이용되고 있다^[7-8]. GA는 생물의 진화과정을 추상화한 알고리즘이며, 문제를 인공적인 유전자열로 표현하여, 선택, 교차, 돌연변이라는 생존경쟁에 기초한 탐색원리에 의하여 조합한 최적화 문제의 해를 효율적으로 탐색하는 것이 가능하다. 따라서 본 논문에서는 3비트 패리티 판별을 위하여 NN의 학습에 교차법, 돌연변이법에 새로운 기법을 도입한 개량형 GA를 제안하여 단순 GA와 비교 실험한다. 본 논문은 NN의 학습에 GA와 BP 학습법을 병용한 방식이 학습 속도의 면에서 유효하다는 것을 나타낸다.

II. 신경회로망

본 논문에서는 3층 구조의 퍼셉트론(perceptron)형^[9], [10]의 NN을 사용하여 BP 알고리즘으로 학습시켰다. 퍼셉트론형 네트워크는 일반적으로 몇 개의 층으로 구성되는 계층형이며, 각층은 적당한 수의 유닛(unit)으로 구성되며, 각층 간의 결합은 입력층부터 출력층으로 향하는 결합이다. 본 논문에서는 m층으로 구성되는 네트워크를 고려하여, k층의 제i 유닛에의 입력의 총합을 $i^{(k)}$, 출력력을 $o^{(k)}$ 로 하며, k-1층의 제j 유닛으로부터 k층의 제i 유닛에의 결합의 가중치를 $w_{ij}^{(k)}$ 로 한다. 또한 입력층 이외에는 문턱값(threshold)을 가지며, 유닛의 입출력 특성은 비선형(sigmoid) 함수로 하며 f로 나타낸다. 학습계수 및 가속도계수는 각각 0.1, 0.9로 하며, 가중치의 초기치는 ± 0.5 의 범위 내에 들어오도록 난수를 사용하였다. 따라서 이러한 변수 간의 관계를 식 (1)과 식 (2)와 같이 나타낸다.

$$o^{(k)}_i = f(i^{(k)}_i) \quad (1)$$

$$i^{(k)}_i = \sum_j w_{ij}^{(k)} o^{(k-1)}_j + \theta^{(k)}_i \quad (2)$$

여기서 $\theta^{(k)}$ 는 k층의 제i 유닛의 문턱값이다.

1장에서 기술한 BP 학습법의 문제점을 해결하기 위하여 본 논문에서는 GA에 의한 학습 방법을 제안한다.

III. 유전자 알고리즘

본 논문에서는 NN의 학습 문제를 문제 고유의 성질에 적용시킨 유전자 조작을 도입한 개량형 GA를 사용한 방법을 제안하여, 세대차이, 중간층 유닛수의 차이, 집단 개체수의 차이에 대하여 단순 GA^[6]에 의한 방법과 비교 실험을 실시한다. 학습에 사용한 2종류의 GA의 동작은 각각 다음과 같다. (1) 단순 GA에서는 선택은 적응도 비례 전략 방법을, 교차는 일정 교차법을, 돌연변이는 적응변이를 각각 채용한다. 또한, 1회의 세대 교체에서는 적응도에 따라서 2개의 집단으로 분류하여 나쁜 쪽의 개체 집단을 사멸시킨다. (2) 개량형 GA의 변경점 중의 하나는, 단순 GA의 일정 교차에서는 비트마다 교환하지만 개량형 GA의 교차는 결합 가중치를 나타내는 비트열마다 교환하는 방식을 채택한다. 돌연변이에 대해서는 단순 GA에서는 그 위치를 특별히 지정하지 않지만 개량형 GA에서는 결합가중치를 나타내는 비트열의 하위 비트마다 돌연변이 하는 확률이 높게 되도록 변이율을 설정한다.

1. 패리티 판별을 위한 학습 조건

본 논문에서는 3비트 패리티 판별문제를 위하여 퍼셉트론형의 NN에 GA를 사용하여 학습을 실시하였다. 패리티 문제는 입력 패턴에 포함되는 비트 1의 수만이 문제가 되며 일반적으로 길이 n의 부호어 X는 식 (3)과 같이 나타낸다.

$$X = x_1 x_2 x_3 \cdots x_n \quad (x_i = 0 \text{ or } 1) \quad (3)$$

여기에서 X에 부호 c를 부가하여 길이 n+1의 부호어 X'를 식 (4)와 같이 만든다.

$$X' = x_1 x_2 x_3 \cdots x_n c \quad (c = 0 \text{ or } 1) \quad (4)$$

여기에서 $x_1 \oplus x_2 \oplus x_3 \oplus \cdots \oplus x_n \oplus c = 0$ 이 되도록 c를 결정한다. 이것이 패리티 판별 문제이며, x_i 를 NN의 입력 패턴으로 하며 c를 출력 패턴으로 한다.

본 논문에서는 GA 학습법을 사용하여 NN에 짝수 패리티를 학습시킨다. NN의 입력 데이터는 0 ~ 7을 2진수로 표현한 000~111의 8개의 비트패턴을 사용하였으며, 학습 데이터는 각 입력 데이터에 대응하는 짝수 패리티를 사용하였다.

2. 유전자형의 설정

일반적으로 유전자형의 기호열로서는 벡터가 사용되어지며 각 요소는 비트열로 표현된다. 본 논문에서는 NN과 유전자와의 관계를 명확하게 하기 위하여 유전자열로서는 NN의 종합계수 행렬을 사용하여 이 요소로서 비트열을 사용한다. 본 논문에서 고려하고 있는 NN은 m 층으로 구성되는 네트워크이므로, 종합계수 행렬은 II장에서 기술한바와 같이 모두 $m-1$ 개 존재한다. 본 논문에서는 그림 1에 나타낸 것과 같이 NN의 가중치 w_{ij} 와 문턱치 θ_i 을 요소로 하는 $m-1$ 개의 행렬 w 를 하나의 유전자로 한다. 네트워크의 가중치의 각 비트들을 입력층으로부터 출력층으로 향하게 직렬로 연결하여 1차원 배열을 만든 것을 하나의 개체로 한다. 그림 1의 행렬 중의 각 결합 가중치는, 교차와 돌연변이를 실시할 경우에, 각 요소를 소수부분에 5비트를 사용한 총 10비트의 부호 있는 절대치로서 2진수 고정소수점 표현으로 나타내며, 부의 수는 2의 보수표현을 사용한다.

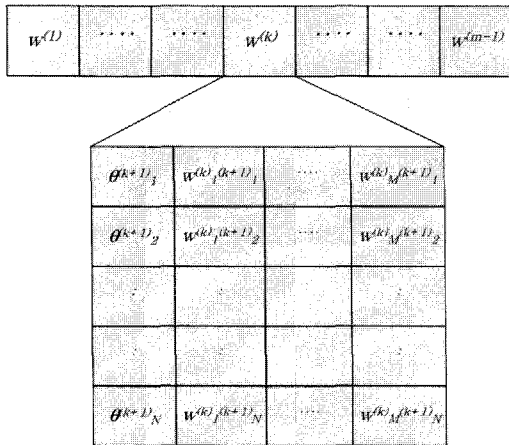


그림 1. m 층으로 구성되는 NN의 유전자형
Fig. 1. Genotype of an m -layer NN.

IV. GA를 사용한 NN의 학습법

1. 단순 GA에 의한 학습

가. 적응도의 설정

일반적으로 적응도가 결정되면 이 값을 그대로 선택시의 확률에 반영하는 것 보다 어떤 함수를 도입하는 스케일링을 사용하여 적응도의 차이를 확대 혹은 축소하는 방법이 수렴의 효율성을 높게 한다. 본 실험에서는 원래의 적응도를 f , 새로운 적응도를 f' 라고 하면, $f' = f^2$ 로 나타내는 2승 스케일링을 채택한다. 2승 스케일링은 적응도가 크면 클수록 적응도가 확대되는 확

률이 커지므로 본 논문에서는 적응도가 큰 것이 확실히 선택되도록 이 방법을 제안한다.

나. 선택 교차

선택교차는 환경에 적응한 개체를 적응도에 따라서 확률적으로 선택하는 조작이며, 본 실험에서는 적응도 비례 전략을 사용한다. 적응도 비례 전략은 룰렛 모델 혹은 몬테카를로 모델로 불리는데, 각 개체의 적응도에 비례한 확률로 자손을 남길 가능성이 있는 모델이다. 어떤 개체 i 가 각각의 선택정도에 따라 선택될 확률 $pselect_i$ 는 식 (5)와 같이 나타낸다.

$$pselect_i = \frac{f_i}{\sum_{j=1}^n f_j} \tag{5}$$

다. 교차

본 단순 GA에서는 자식의 염색체를 어느 쪽의 부모의 유전자를 계승할 것인가를 비트마다 결정하는 일정 교차법을 채택하며, 특별히 교차위치를 지정하지 않는다. 이 교차법은 세대교체의 초기단계에 있어서 부모의 염색체의 배열은 모두 랜덤하기 때문에 광역탐색이 되지만, 세대교체가 진행되어 집단이 수렴하게 되면 부모의 염색체배열 등에는 뛰어난 공통의 부분이 증가하기 때문에 이 부분을 목표로 한 극소탐색이 된다.

라. 돌연변이

본 단순 GA에서는 교차결과로부터 나온 2개의 개체의 근사도를 해밍거리로 측정하여, 거리가 가까울수록 높은 변이율로 하는 적응변이를 도입한다. 또한 이 돌연변이의 구체적인 조작은 먼저 이 돌연변이율에 의해서 개체마다 돌연변이의 실행여부를 결정하며, 돌연변이 시키는 개체의 염색체 중의 1비트만을 대립유전자(원래가 0이면 1, 원래가 1이면 0)로 교체한다. 그 위치는 특별히 지정하지 않고 난수에 의해서 결정된다. 본 실험에서는 식 (6)과 같은 돌연변이율 M_r 을 설정한다.

$$M_r = M_{\max} - (M_{\max} - M_{\min}) \times \frac{d_H}{l_c} \tag{6}$$

여기에서 M_{\max} 는 최대돌연변이율, M_{\min} 는 최소돌연변이율, d_H 는 해밍거리, l_c 는 염색체길이를 각각 나타낸다. 본 실험에서는 $M_{\min} = 0.05$, $M_{\max} = 0.5$ 로 하며,

또한 정규화 처리를 위해서 해밍거리를 염색체길이로 나눈다.

마. 세대모델

본 실험에서는 연속 세대모델을 사용하였으며, 현 세대 중의 몇 퍼센트가 교체될 것인가를 나타내는 세대차이라고 하는 매개변수를 도입한다. 단순 GA는 집단 중의 모든 개체가 한번에 다음세대로 교체되므로 어떤 지점에서는 비교적 우수한 해인 개체가 교차 및 돌연변이에 의해서 파괴되어 버리는 원하지 않는 현상이 일어난다. 따라서 모든 개체를 한번에 다음 세대로 교체하는 것이 아니라 집단 중에서 비교적 높은 평가를 얻은 개체에 대해서는 그대로 다음 세대에 남김으로서 이 문제를 해결할 수 있다. 본 논문에서 제안하는 세대차이를 다음 식과 같이 나타낸다.

$$\text{세대차이} = \frac{\text{교체하는 개체수}}{\text{집단의 개체수}}$$

2. 개량형 GA에 의한 학습

NN의 학습을 GA에 의해서 실시할 경우에 통상의 유전자 조작을 사용하면 여러 가지 문제가 발생한다. 예를 들면, 하나의 가중치 및 문턱치를 1차원으로 표현했을 때 이것들은 분산되어 버리기 때문에 교차에 의해서 파괴되어 버리는 경우가 많다. 본 논문에서는 이러한 문제를 해결하기 위하여 교차방법 및 돌연변이 방법에 대해서 수정작업을 한 개량형 GA를 제안하여 NN의 학습 문제에 적용시킨 유전자 조작을 실시한다. NN의 구조, 유전자형의 설정, 적응도의 설정, 개체의 선택 방법 등은 앞장에서 기술한 단순 GA의 유전자 조작을 실시하는 경우와 동일한 설정조건으로 한다.

가. 스키마타(schema)

염색체가 1차원으로 표현되어 있는 경우 이 중에는 의미있는 부분패턴이 발생하는데 이와 같은 부분패턴들이 다음 세대에 어느 정도 살아남을 수 있는가를 나타내는 것이 스키마타이다. 스키마타에서 유전자의 특징을 나타내는 값으로 정의길이와 차수를 사용된다. 정의 길이는 스키마타의 최초 고정부분과 최후 고정부분 사이의 거리를 나타내며, 차수는 스키마타 중에서 값이 결정된 부분의 수를 나타낸다. 예를 들면, 10****01의 스키마타는 정의길이가 5이고 차수는 4이다. 여기에서 '*'로 표시된 기호를 무관기호라 하고, '*'가 아닌 기호

를 특정기호라 한다. 따라서 이러한 스키마타를 사용함에 따라서 각 개별의 방대한 유전자로부터 유전자집단이 가지는 특성을 파악하여 표현가능하게 된다.

나. 교차

NN의 결합가중치의 학습문제에 GA를 적용할 경우, 1차원적으로 표현된 결합가중치는 스키마타라고 생각할 수 있다. 그런데 교차를 할 때에 교차위치가 스키마타의 도중일 경우에는 스키마타가 파괴되어 다음 세대의 개체에 부모의 형질이 유전이 어려워지므로 교차위치를 특별히 지정할 필요가 있다. 본 실험에서는 가중치 및 문턱치의 요소가 파괴되지 않도록 가중치 및 문턱치마다 구분하는 기법을 도입한 일정교차를 사용한다.

다. 돌연변이

가중치를 비트열로 표현한 경우에 상위부터 하위까지의 비트열 중의 각 비트에는 각각 가중치 차이가 있기 때문에, 가중치가 큰 비트를 돌연변이 시키면 가중치의 값이 크게 변한다. 따라서 모든 비트를 동일한 확률로 돌연변이 시키는 것은 적당하지 않으므로, 본 실험에서는 돌연변이 하는 확률을 상위비트는 적게 하고 하위비트는 크게 함으로서 이 문제를 해결한다. 그림 2는 본 논문에서 제안하는 개량형 GA의 돌연변이 방식을 나타낸다.

본 논문에서는 염색체 중의 하나의 가중치 혹은 문턱치를 표현하는 연속되는 2진수 열을 고려한다. 이 요소를 표현하는 2진수 열에 대해서 최상위로부터 계산하여 *i*번째의 비트가 돌연변이 하는 확률을 식 (7)과 같이 설정한다.

$$M_i = M_{\min} + \left(\left(\frac{M_{\min} - M_{\max}}{l_p - 1} \right) \times (i - 1) \right) \quad (7)$$

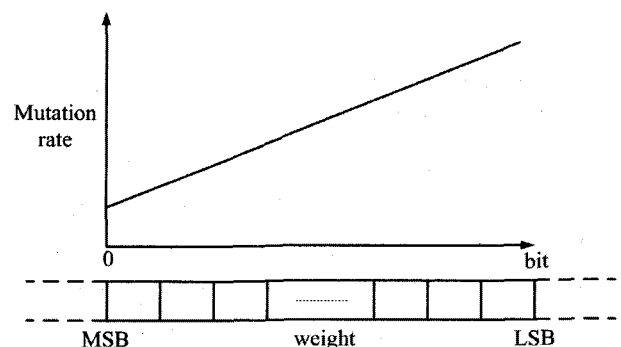


그림 2. 개량형 GA의 돌연변이
Fig. 2. Mutation of an improved GA.

여기에서 l_p 는 요소의 비트길이를 나타내며, $M_{min} = 0.05$, $M_{max} = 0.4$ 로 고정하여 실험을 실시한다. 그리고 M_{min} 과 M_{max} 의 평균돌연변이율 $M_{r.ave}$ 은 식 (8)과 같이 나타낸다.

$$M_{r.ave} = \frac{M_{min} + M_{max}}{2} \quad (8)$$

이 $M_{r.ave}$ 에 의하여 한 개체의 염색체 중의 각각의 요소에 대해서 돌연변이 여부를 결정하며, 어떤 요소가 돌연변이 될 것이 정해지면 각 비트마다의 돌연변이율에 의해서 돌연변이 되는 비트가 결정된다. 이 조작에 의해서 상위비트일 수록 돌연변이 하는 확률이 낮아지며 하위비트일수록 돌연변이 하는 확률은 높아진다.

V. 실험 결과

1. 패리티판별 학습

본 논문에서는 NN의 학습 문제에 대한 GA의 유효성을 검증하기 위하여 개량형 GA에 의한 알고리즘을 사용하여 단순 GA와 비교 실험을 실시하였으며, 본 실험의 3비트 패리티 판별을 위한 조건은 다음과 같다. 1) 학습에 사용한 NN은 입력층 3, 중간층 5, 출력층 1의 3층으로 구성된다. 2) GA에 대한 집단의 개체수는 50이다. 3) GA의 1주기를 학습횟수로 하여 세대수가 2000을 넘게 된 것에 대해서는 학습실패로 간주하며, 평균2층 오차가 10^{-4} 까지 되도록 학습을 실시하였다. 3비트 패리티 판별을 위하여 적응도의 평가 → 선택 → 교차 → 돌연변이를 한 주기로 하여 BP 학습법을 15회 실시하였다. 교차가 일어나는 확률 $Pr\{\text{교차}\} = (1\text{학습 주기 중의 교차의 횟수}/\text{집단의 개체수})$, 돌연변이가 일어나는 확률 $Pr\{\text{돌연변이}\} = (1\text{학습 주기 중에 돌연변이를 하는 유전자의 수}/\text{집단의 개체수})$ 로 정의하며, 본 실험에서는 각각 0.1 및 0.3으로 하였다.

2. 실험 1 : 세대 차이에 의한 실험

본 실험은 단순 GA 및 개량형 GA의 유전적 조작을 실시한 경우에 대하여 어느 정도 탐색능력이 향상되는가를 확인하기 위하여 세대 차이를 0.2, 0.4, 0.6, 0.8, 1.0으로 변경하여 3비트 패리티판별 실험을 실시하였으며, 실험결과로부터 최적한 세대 차이를 조사할 수 있었다.

표 1. 세대차이에 대한 효과

Table 1. The effect for the generation gap.

세대차이	성공/시행 횟수	평균 수렴세대수	평균계산 시간
0.2	18/20	259.2	25.0
0.4	19/20	135.0	13.7
0.6	19/20	82.3	9.1
0.8	19/20	57.4	5.1
1.0	20/20	125.0	13.1

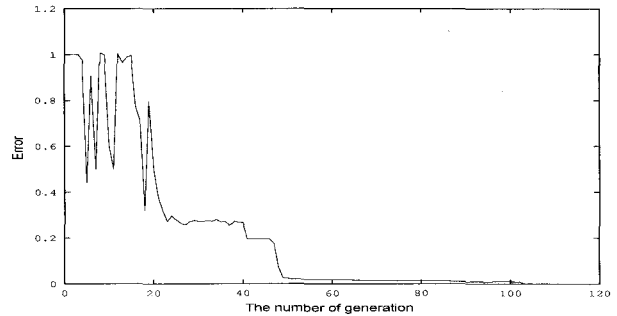


그림 3. 학습시의 오차곡선 (세대차이 0.8)

Fig. 3. An error curve for NN training (where the generation gap is 0.8).

가. 단순 GA에 의한 실험 결과

표 1은 각각의 세대 차이에 대해서 시행횟수 및 이 중 학습에 성공한 횟수, 학습이 성공한 경우의 평균 수렴세대수, 수렴까지의 평균 계산시간(CPU time)을 나타낸다. 표의 평균 계산시간에서 알 수 있듯이 세대차이가 0.2인 경우에 진화하는 모양이 완만하다. 이것은 1회의 세대교체로 생기는 새로운 개체수가 충분하지 않기 때문에 이 중에서 우수한 개체가 나타날 확률이 낮아지기 때문이라고 판단된다. 본 단순 GA에서는 특별히 교차의 위치를 지정하지 않았기 때문에 교차에 의해서 부모의 형질을 잃어버리는 경우가 생길 수 있다. 그림 3은 세대차이가 0.8인 경우의 NN의 학습시의 오차곡선을 나타내며, 가로축은 세대수, 세로축은 각 세대수에 대해서 가장 우수한 개체를 나타내는 네트워크의 평균자승오차이다.

나. 개량형 GA에 의한 실험 결과

단순 GA의 방법과 동일하게 세대 차이를 5종류로 변경한 경우의 결과를 표 2에 나타내며, 세대 차이가 0.8인 경우의 학습시의 오차곡선을 그림 4에 나타낸다. 세대 차이를 0.8로 설정하였을 때의 그림 3과 그림 4의 학습 오차곡선을 비교하면, 교차법의 차이에 의해서 다음 세대에 현 세대보다 나쁜 개체가 나타날 비율의 차이 및 수렴의 차이가 생길 가능성이 있다는 것을 나타내고 있다. 각 세대교체에 대해서는 모든 개체가 도태되기

표 2. 세대 차이에 대한 효과

Table 2. The effect for the generation gap.

세대 차이	성공/시행횟수	평균수렴 세대수	평균 계산시간
0.2	20/20	120.1	12.1
0.4	20/20	76.5	7.6
0.6	20/20	26.4	2.5
0.8	20/20	23.1	2.0
1.0	20/20	44.0	4.5

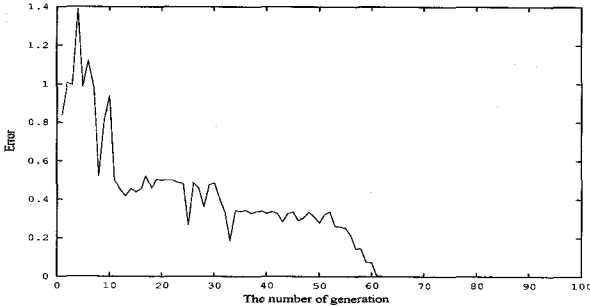


그림 4. 학습시의 오차곡선 (세대차이 0.8)

Fig. 4. An error curve for NN training (where the generation gap is 0.8).

때문에 어떤 개체가 몇 세대 동안 계속하여 생존하는 경우는 드물다. 따라서 각 개체는 현 세대의 형질을 교차에 의해서만 다음 세대에 남길 수 있다. 이상의 세대 차이에 의한 결과로부터, 돌연변이법과 교차법에 수정 작업을 한 경우의 개량형 GA의 유전자 조작법은 단순 GA의 유전자 조작을 실시할 경우보다도 강력한 탐색능력을 가지고 있다는 것을 알 수 있다. 세대차이가 0.8인 경우가 최적이라는 본 실험결과로부터 이하의 실험에서는 세대 차이를 0.8로 하여 실험을 실시한다.

3. 실험 II : 중간층 유닛수의 차이에 의한 실험

NN을 학습시킬 경우, 중간 유닛수를 얼마로 하는가에 따라서 NN의 학습이 크게 변할 수 있다. 중간 유닛수가 너무 적으면 학습이 충분히 이루어지지 않고 너무 많으면 시간이 너무 걸릴 수 있으며, 또한 입력의 오차에 대한 범용화가 잘 이루어지지 않는 경우가 있다. 따라서 이러한 문제에 대한 영향을 검토하기 위하여 GA 학습법을 사용하여 집단 개체수를 50, 세대차이를 0.8로 한 경우의 중간 유닛수에 대한 효과를 이하의 실험에 나타낸다.

가. 단순 GA에 의한 실험 결과

표 3은 단순 GA에 대한 중간층의 유닛수가 미치는 영향을 나타낸다. 표의 결과로부터 중간층의 유닛수가

표 3. 중간층 유닛수에 대한 효과

Table 3. The effect for the number of units in the hidden layer.

중간층의 유닛수	성공/시행 횟수	평균 수렴세대수	평균 계산시간
3	14/20	153.5	17.3
5	19/20	57.4	5.1
7	20/20	65.5	5.4
9	20/20	69.3	6.5

5인 경우, 학습이 가장 빨리 수렴된 것을 알 수 있다. 중간층의 유닛수가 증가하면 그만큼 네트워크에 대한 정보가 증가하여 한 개체의 염색체의 길이도 크게 된다. 따라서 이에 대응하여 집단의 개체수를 일정 부분 크게 하지 않으면 충분한 탐색을 할 수 없으며 극소 해에 빠질 가능성이 높게 된다.

나. 개량형 GA에 의한 실험 결과

표 4는 개량형 GA에 대한 중간층의 유닛수가 미치는 영향을 나타낸다. 표 3의 단순 GA의 조작을 실시한 경우의 결과와 비교하면, 전체 학습이 수렴할 때까지의 세대수는 단순 GA의 조작을 실시한 경우보다도 감소하고 있다. 본 실험에는 나타나지 않았지만 개체수가 20, 70, 100인 경우에도 개체수가 50인 경우와 비슷하게 최적의 중간층수는 5라는 결과를 구할 수 있었다. 따라서 중간층 유닛수가 3인 경우에 상당히 나쁜 결과를 나타내고 있는 것은 네트워크 자신이 충분한 문제해결 능력을 가지고 있지 않는 것이 그 원인이라고 할 수 있다.

표 4. 중간층 유닛수에 대한 효과

Table 4. The effect for the number of units in the hidden layer.

중간층의 유닛수	성공/시행 횟수	평균 종료세대수	평균 계산시간
3	16/20	129.8	10.2
5	20/20	23.1	2.0
7	20/20	23.7	3.8
9	20/20	28.0	4.4

4. 실험 III : 집단 개체수의 차이에 의한 실험

가. 단순 GA에 의한 실험 결과

일반적으로 개체수가 적으면 GA가 가지는 광역탐색 능력을 충분히 발휘하지 못하게 되어 극소 해에 빠질 가능성이 있으며, 너무 크면 랜덤 탐색이 되어서 효율이 나쁘게 된다. 따라서 NN의 학습문제를 취급하는 경우에는 중간층의 유닛수를 크게 하면 염색체의 길이도

표 5. 개체수에 대한 효과

Table 5. The effect for the number of individuals.

개체수	성공/시행횟수	평균 수렴세대수	평균 계산시간
20	18/20	60.2	5.9
50	19/20	57.4	5.1
70	20/20	76.0	6.7
100	20/20	116.3	7.3

길게 되므로 이에 맞추어서 개체수를 증가시킬 필요가 있다. 표 5는 세대 차이 0.8에 대하여 중간층의 유닛수가 5인 경우의 결과를 나타낸다. 표 5로부터, 개체수가 20인 경우에 학습이 수렴할 때까지의 평균 계산시간은 비교적 적지만 2,000세대가 될 때까지 학습이 성공한 것은 약 90%로 가장 좋지 않은 값이 되었다. 이것은 개체수가 비교적 적기 때문에 GA의 학습 알고리즘 상, 1세대분의 계산시간이 적게 끝나기 때문에 수렴은 빨리 끝나지만, 반대로 충분한 탐색공간을 가지지 않기 때문에 극소 해에 빠지는 경우가 많은 것을 나타내고 있다. 또한 개체수가 100인 경우에 2,000세대까지에 학습의 수렴율은 상당히 높아졌지만 반면에 탐색시간이 가장 길게 되었다. 이것은 개체수가 비교적 크기 때문에 1회의 탐색에 필요로 하는 계산시간이 길게 되어 전체적으로 계산시간이 길게 되었다고 판단된다.

나. 개량형 GA에 의한 실험 결과

표 6은 세대 차이를 0.8로 하였을 때의 개량형 GA에 대한 중간층 유닛수가 5인 경우의 결과를 나타낸다. 표 5의 단순 GA의 실험 결과와 비교하면, 개체수에 상관 없이 단순 GA보다도 개량형 GA가 충분히 좋은 결과를 남기고 있다. 그리고 중간층 유닛수가 5가 되면, 네트워크 내부에 학습에 필요한 정보를 충분히 가지므로 네트워크 자신이 충분한 문제해결 능력을 갖추므로 모든 경우의 개체수에 대해서 세대수가 2,000이 될 때까지 학습은 충분히 수렴한다.

이상의 결과로부터 본 실험에서는 이들의 개체수 중에서 50의 경우가 최적이라는 결과를 구할 수 있었으며, 또한 개량형 GA에 의한 속도개선의 효과가 현저하다는 것을 알 수 있었다.

표 6. 개체수에 대한 효과

Table 6. The effect for the number of individuals.

개체수	성공/시행횟수	평균 수렴 세대수	평균 계산 시간
20	20/20	26.0	2.5
50	20/20	23.1	2.0
70	20/20	38.3	4.3
100	20/20	40.1	4.5

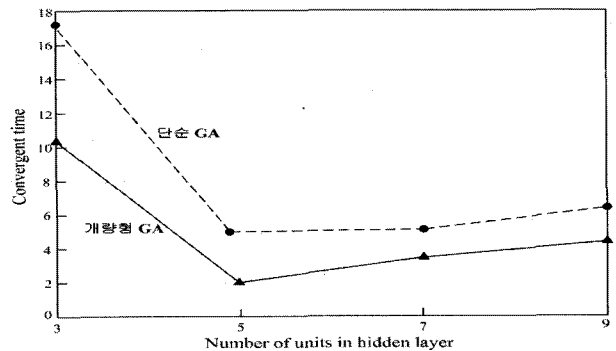


그림 5. 중간층 유닛수와 수렴시간의 관계

Fig. 5. Relation between the number of units in the hidden layer and convergence time.

5. 단순 GA와 개량형 GA의 비교 실험 결과

그림 5는 개체수 50일 때, 중간층의 유닛수 3, 5, 7, 9의 4종류를 사용하여 단순 GA 및 개량형 GA에 의한 중간층 유닛수와 수렴시간과의 관계를 나타낸다. 중간층의 유닛수가 5인 경우에 세대수가 2,000이 될 때까지 가장 빨리 학습이 수렴된 것을 알 수 있다. 반면에 중간층의 유닛수가 증가하면 그만큼 네트워크에 대한 정보가 증가하여 한 개체의 염색체의 길이도 크게 되므로, 이에 대응하여 집단의 개체수를 일정 부분 크게 하지 않으면 충분한 탐색을 할 수 없으며 극소 해에 빠질 가능성이 높게 된다.

이상의 결과로부터, 돌연변이법과 교차법에 수정작업을 한 경우의 개량형 GA의 유전자 조작법은 단순 GA의 유전자 조작을 실시할 경우와 비교하여 강력한 탐색 능력을 가지고 있으며, 또한 최종적으로 계산량의 감소라는 양호한 결과를 가져온다는 것을 알 수 있었다. 따라서 본 실험에서는 중간 유닛수 5의 경우가 최적이라는 결론을 구할 수 있었다.

VI. 결 론

본 논문에서는 NN의 학습방법에 주목하여 강력한 탐색에 의한 최적화 알고리즘으로 알려진 유전자 알고리즘을 사용하여 3비트 패리티 판별문제를 연구하였다.

본 논문에서는 종래부터 사용되어진 BP법에 기인하는 문제점 및 NN의 학습문제를 통상의 GA로 취급한 경우의 문제점을 나타내, 유전적 알고리즘의 필요성을 기술하였다. 이러한 문제를 해결하기 위하여 개량형 GA에서는 교차법과 돌연변이법에 다음과 같이 수정작업을 하여 실험을 실시하였다. (1) 통상의 GA가 교차의 위치를 지정하지 않는 것인데 반해, 적어도 가중치 및

문턱치를 나타내는 스키마타가 파괴되지 않도록 요소의 도중에 교차위치가 오지 않도록 교차위치에 제한을 설정하는 교차법을 실시하였다. (2) 통상의 GA가 교차위치와 동일하게 돌연변이의 위치를 지정하지 않는 것에 반해, 가중치가 큰 비트는 돌연변이 하는 확률을 적게 하고 가중치가 적은 비트는 확률을 크게 함으로서 돌연변이 하는 위치에 기울기를 주어 근방탐색의 수단으로서의 성격을 강하게 하였다.

결론적으로 본 개량형 GA를 사용한 네트워크의 학습은 전체적으로 통상의 GA에 의한 학습보다도 양호한 결과가 구해졌다. 본 실험의 GA의 개량은 위에서 기술한 것과 같이 NN 학습문제뿐만 아니라 그 외의 실수치의 조합 문제에 본 개량 GA를 적용한 경우에도 충분한 성능 향상을 기대할 수 있다고 본다.

참 고 문 헌

[1] W. G. Knecht, M. E. Schenkel, G. S. Moschytz, "Neural network filters for speech enhancement", *IEEE Trans. Speech and Audio Processing*, vol. 3, no. 6, pp. 433-438, 1995.
 [2] 우동현, 엄일규, 김유신, "공간 주파수 특성을 가지는 다중 신경 회로망을 이용한 영상 보간", *대한전자공학회 논문지*, 제41권 5호 SP편, pp. 135-141, 2004. 9.

[3] D. G. Stork and J. D. Allen, "How to solve the N-bit parity problem with two hidden units", *Neural Networks*, vol. 5, no. 6, pp. 923-926, 1992.
 [4] J. M. Minor, "Parity with two layer feedforward nets", *Neural Networks*, vol. 6, no. 5, pp. 705-707, 1993.
 [5] A. V. Ooyen and B. Nienhuis, "Improving the convergence of the back-propagation algorithm", *Neural Networks*, vol. 5, no. 3, pp. 465-471, 1992.
 [6] J. H. Holland, "Adaptation in Natural and Artificial Systems", University of Michigan Press, 1975.
 [7] J. J. Grefenstette, R. Gopal, B. J. Rosmaita and D. Van Gucht, "Genetic algorithms for the traveling salesman problem", *Proc. of an International Conference on Genetic Algorithms and Their Applications*, pp. 160-168, 1985.
 [8] 이상열, 윤기천, "유전자 알고리즘을 이용한 다중 레이더 펄스열 분리", *대한전자공학회 논문지*, 제 40권 6호 SP편, pp. 98-105, 2003. 11.
 [9] S. K. Pal, S. Mitra, "Multilayer perceptron, fuzzy sets, and classification", *IEEE Transaction on Neural Networks*, vol. 3, no. 5, pp. 683-697, 1992.
 [10] 이백영, 이택승, 황병원, "학습속도 개선과 학습데이터 축소를 통한 MLP 기반 화자증명 시스템의 등록속도 향상방법", *대한전자공학회 논문지 SP편* 제39권 제3호, pp. 87-97, 2002. 5.

저 자 소 개



최 재 승(정회원)
 1989년 조선대학교 전자공학과 학사졸업
 1995년 일본 오사카시립대학 정보통신공학과 석사졸업
 1999년 일본 오사카시립대학 정보통신공학과 박사졸업
 2000년~2001년 일본 마쯔시타 전기산업주식회사 AVC사 연구원
 2002년~2007년 경북대학교 디지털기술연구소 연구원, 프로젝트 리더
 2007년~현재 신라대학교 전자공학과 교수
 <주관심분야: 음성신호 및 영상신호처리, 유전알고리즘, 디지털통신 및 디지털 TV>



김 정 화(정회원)
 1979년 조선대학교 전자공학과 학사 졸업
 1981년 조선대학교 전자공학과 석사 졸업
 1991년 숭실대학교 전기공학과 박사 졸업
 1979년~현재 조선대학교 전자공학과 교수
 <주관심분야: 유비쿼터스, 신호처리 및 시스템, 영상처리, 영상 워터마킹>