

# Random 탐색법과 조합된 Tabu 탐색법을 이용한 신경회로망의 학습

°신 광 계°, 양 보 석°, 최 원 호°°

(Learning of Neural Network Using Tabu Search Method with Random Moves)

(Kwang-Jae Shin, Bo-Suk Yang, Won-Ho Choi)

## 1. 서론

계층형 신경회로망은 학습능력이나 비선형사상능력을 가지고 있고, 그 특징을 이용하여 패턴인식이나 동정 및 제어 등의 적용이 시도되어 성과를 올리고 있다<sup>(1)</sup>. 현재, 그 학습법으로 널리 이용되고 있는 것이 역전파학습법으로 최강강하법이나 공액경사법 등의 최적화 방법이 적용되고 있지만, 학습에 많은 시간이 걸리는 점, 국소적 최적해(local optimization)에 빠지는 수렴이 이루어져 오차가 충분히 작게 되지 않는 점이 문제점으로 지적되고 있다<sup>(2)</sup>.

신경회로망의 학습알고리즘은 최적설계문제에서 사용되고 있는 최적화 이론을 적용할 수 있고, 대국적 최적해를 구하는 알고리즘을 적용하여 신경회로망의 학습정도향상을 도모하고자 하는 많은 연구가 진행되고 있다. 최적화 이론에서 대국적 최적해를 구하는 알고리즘으로서 random 탐색법, simulated annealing법, GA법, 그리고 tabu 탐색법 등이 있다. 또, 시그모이드 함수의 구배를 최적화하는 알고리즘이나 의사역행렬을 이용한 고속학습법이 제안되었다<sup>(3)</sup>. 그러나, 이러한 개선방법에도 불구하고 여전히 문제점은 남아있다.

본 논문에서는 Hu에 의해 고안된 random 탐색법과 조합된 tabu 탐색법(random tabu 탐색법)을 결합계수를 구하는 학습알고리즘으로 직접 사용하여 국소적 최적해에 수렴하는 것을 방지하고, 수렴정도를 개선하는 새로운 방법을 제안한다. 이 방법을 배타적 논리합 문제에 적용하여 역전파법 및 tabu 탐색법을 이용한 오차역전파법<sup>(6)</sup>과 비교한다. 그리고, 각 파라미터가 오차함수의 수렴에 미치는 영향을 조사한다.

## 2. 신경회로망의 개요

신경회로망(neural network)이란, 인간의 신경세포(neuron)의 모델인 네트워크화 하고, 고도병렬분산처리를 실현한 정보처리 시스템의 총칭이다. 신경회로망은 그 구조에 의해 계층형과 상호결합형으로 대별된다. 그 중 계층형 신경회로망에 관해서는 진단, 문자인식, 로봇트 관철(manipulator)의 제어, 패턴인식 등의 응용에도 많은 연구가 활발히 진행되고 있다.

그 배경으로는 (1) 입력출력관계를 학습시키는 것 만으로 비선형성이 강한 사상을 네트워크 내의 결합계수로써 구축할

수 있다고 하는 간편성, (2) 학습패턴을 순조롭게 내삽하고 미학습의 입력에 대해서도 타당한 값을 출력할 수 있는 일반화(generalization)능력, (3) 학습이 종료하면 간단한 연산을 하는 것 만으로 출력이 얻어지는 응용시의 신속성, (4) 입력변수, 즉 입력층 유니트의 증가에 용이하게 임의 연속함수를 임의의 정밀도로서 근사화할 수 있다고 하는 높은 사상(mapping)능력 등의 여러가지 뛰어난 특징을 가지고 있기 때문이다<sup>(2)</sup>.

신경회로망의 학습이란, 정보처리 시스템의 목적에 맞도록 유니트간의 결합하중을 조절하는 것이다. 따라서 학습은 신경회로망에 있어서 가장 중요한 것이며, 최적 분류 패턴을 하기 위하여 결합하중을 수정하는 것이다. 이러한 결합하중의 수정에 의한 학습을 수행하기 위해서 평가기준이 필요하며, 외부에서 교사신호로 입력신호에 대한 이상적인 출력을 부여하여 비교평가하는 교사있는 학습방식과 평가기준을 내장해서 외부에서 하나하나 교사신호를 부여하지 않는 교사없는 학습방식으로 구분할 수 있다.

계층형 신경회로망은 유니트가 연결되어 Fig. 1과 같이 구성된다. 이 네트워크에서 입력유니트  $i$ 의 출력  $I_i$ 와 중간층 유니트  $j$ 의 출력  $H_j$  그리고 출력층 유니트  $k$ 의 출력  $O_k$ 의 관계는 다음과 같다.

$$\begin{aligned} H_j &= f\left(\sum_i W_{ji} \cdot I_i + W_{j0}\right) \\ O_k &= f\left(\sum_j V_{kj} \cdot H_j + V_{k0}\right) \end{aligned} \quad (1)$$

여기서

$W_{ji}$  : 입력층 유니트  $i$ 에서 중간층 유니트  $j$ 로의 결합계수

$V_{kj}$  : 중간층 유니트  $j$ 에서 출력층 유니트  $k$ 로의 결합계수

$W_{j0}$  : 중간층 유니트  $j$ 의 오프셋,

$V_{k0}$  : 출력층 유니트  $k$ 의 오프셋

$f(x)$  : 시그모이드 함수  $f(x)=1/(1+\exp(-2x/u_0))$

( $u_0$ 는 시그모이드 함수 온도)

교사있는 학습방식에서 널리 쓰이는 평가함수로는 출력 ( $O$ )과 교사신호( $T$ )와의 차의 2승, 즉 제곱 오차 함수( $E$ )가 이용된다.

$$E = \sum_p E_p = \sum_p \sum_k (T_{kp} - O_{kp})^2 / 2 \quad (2)$$

\* 부산수산대학교 대학원 기계공학과

\*\* 부산수산대학교 공과대학 기계공학과

\*\*\* 효성중공업(주) 기술연구소

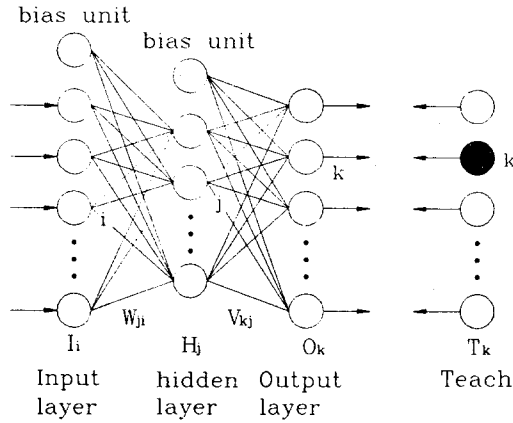


Fig. 1 계층형 신경회로망의 구조

### 3. Random Tabu 탐색법

Hu<sup>(4)</sup>는 Glover가 고안한 tabu 탐색법을 개량하고, 그 방법을 일반적인 제약조건이 있는 최적화 문제에 확장하였다. 어떤 목적함수  $f(x)$ 를 제약조건하에서 최소화 하는 문제의 경우, tabu 탐색법에서는 step수, count수의 2개의 새로운 정수를 정의한다. Step수는 탐색을 하는 근방영역의 개수이고, count수는 하나의 근방영역을 탐색하는 회수의 상한치를 나타낸다. 제약조건을 만족하는 해의 제 1근사해(초기치)를  $x_0$ 로 하고,  $x_0$  주위에 근방영역  $N(x_0, h_i)$ 을 다음과 같이 설정한다(Fig. 2).

$$N(x_0, h_1), N(x_0, h_2), \dots, N(x_0, h_r)$$

여기서  $h_i (i=1, 2, \dots, r)$ 는 step폭,  $r$ 은 step 수,  $P$ 는 step비이다.

$$H = \{h_1, h_2, \dots, h_r\}$$

$$\begin{aligned} h_1 &= b - a \\ h_2 &= h_1 P \\ h_3 &= h_2 P \\ &\vdots \\ h_r &= h_{r-1} P \end{aligned}$$

Tabu 탐색법의 idea는 각각의 근방영역 중에서  $x$ 를 임의(random)로 발생시키고, 그것에 의한  $f(x)$ 가  $f(x_0)$ 보다 적으면, 그 점을 그 영역내의 최소점으로서 기억하고 전 근방영역에서 선택된 최소점들 중에서 가장 작은  $x$ 를 제 2차 근사해  $x_1$ 으로 하고,  $x_1$ 주위에 다시 근방영역을 설정하여 탐색을 반복하는 것이다. 기본적인 방법은 Hooke-Jeeves들이 제안한 직접탐색법의 패턴탐색법과 격자탐색을 조합한 방법에 가깝지만<sup>(4)</sup>, 개량 step폭 대신에 근방영역을 설정한 것과 그 크기가 다른 것을 복수개 설정하는 점에서 다르다.

이 방법의 특징으로서는 다음과 같은 것을 들 수 있다.

- (1) 탐색영역을 크기가 다른 여러 영역으로 분할하여 해를 탐색하기 때문에, 맹목적인 탐색을 피할 수가 있고, 탐색에 요하는 회수를 줄일 수가 있다.
- (2) Random 탐색을 사용하기 때문에, 국소적 최적해에 빠지는 것을 막을 수가 있고, 대국적 최적해를 구하는 것이 가능하다.

- (3) 다른 최적화 방법과 조합하여 사용하므로서 보다 정도 좋고, 고속으로 최적해를 구하는 것이 가능하다.

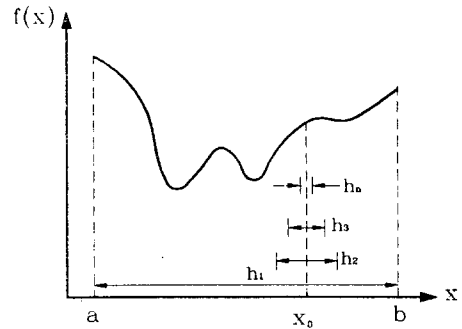


Fig. 2 1변수 문제에서의 근방영역의 설정

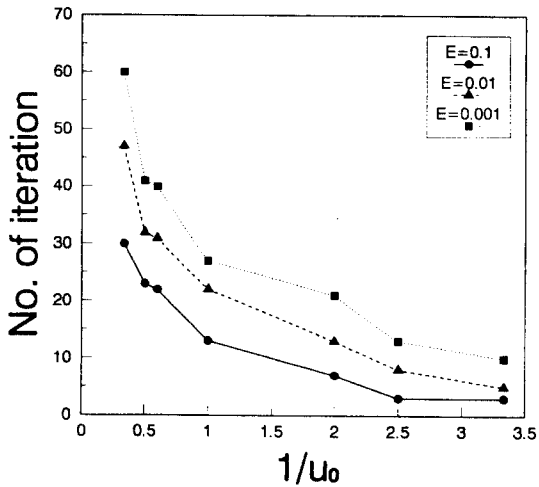
### 4. Random Tabu 탐색법을 이용한 신경회로망의 학습법

신경회로망의 학습은 주어진 입력패턴에 대해서 교사패턴에 가까운 출력패턴을 발생하도록 결합계수의 수정을 하고 이를 통해 결과적으로 평가기준을 만족하는 결합계수를 구하는 것이다. 따라서 평가기준이 되는 목적함수를 식 (2)의 출력과 교사신호의 제곱오차합수로 하고 설계변수를 입력층과 중간층( $W_{ji}$ ), 중간층과 출력층( $V_{kj}$ )의 결합계수로 할 때 제곱오차를 최소화 하는 결합계수를 구하는 문제와 같다.

$$\begin{aligned} E &= \sum_p E_p = \sum_p \sum_k (T_{kp} - O_{kp})^2 / 2 \\ &= f(W, V) \rightarrow \text{Minimize} \end{aligned} \quad (3)$$

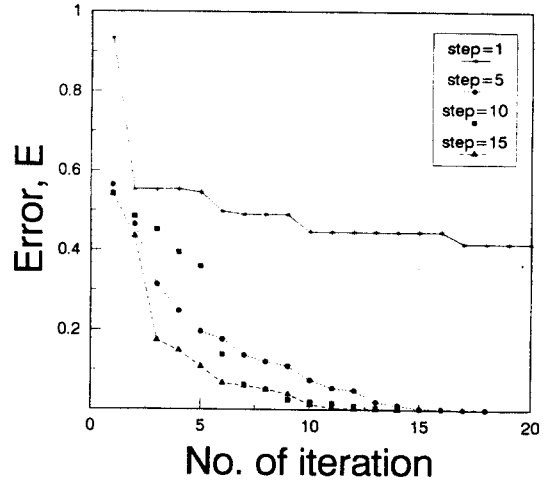
네트워크를 학습시키기 위해 step수와 count수 등의 파라미터를 정한 후 결합계수를 초기화하고 근방영역을 설정한다. 그리고, 결합계수의 행렬 각요소를 각각의 근방영역 중에서 random하게 발생시키고 이 결합계수에 의한 제곱오차합을 기존의 결합계수에 의한 제곱오차합과 비교하여 적으면 이 결합계수를 그 근방영역에서의 오차최소점으로 기억하고 전 근방영역에서 선택된 최소점들 중에서 가장 작은 오차를 가지는 결합계수를 근사최적결합계수로 하고, 이 결합계수의 각요소들 주위에 근방영역을 설정하여 탐색을 반복하는 것이다. 이 과정에서 크기가 다른 근방영역에서의 탐색이 국소적 최적해에의 수렴을 방지하고 random 탐색에서의 맹목적인 탐색을 피하도록 하고 있다. 본 연구에서 제안한 random tabu 탐색법을 이용한 신경회로망의 구체적인 학습은 다음과 순서로 이루어 진다.

- (1) 네트워크에 입력패턴과 교사패턴을 적입한다.
- (2) 파라미터의 값을 정한다. 즉 중간층 유닛수, 시그모이드 함수의 온도( $u_0$ ), step비, step수, count수, 오차한계치, 최대반복수 그리고 초기 난수발생 경계치를 정한다.
- (3) 결합계수를 초기 난수발생 경계치 내에서 난수로 초기화한다.
- (4) Step H를 정의한다.



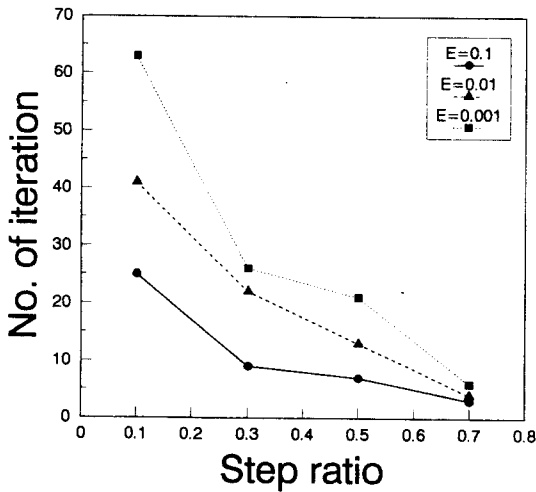
int. wgt.=1.0, hidden unit=10  
step ratio=0.5, step=5, count=10

Fig. 6 시그모이드함수의 기울기(1/u0)에 따른 수렴특성



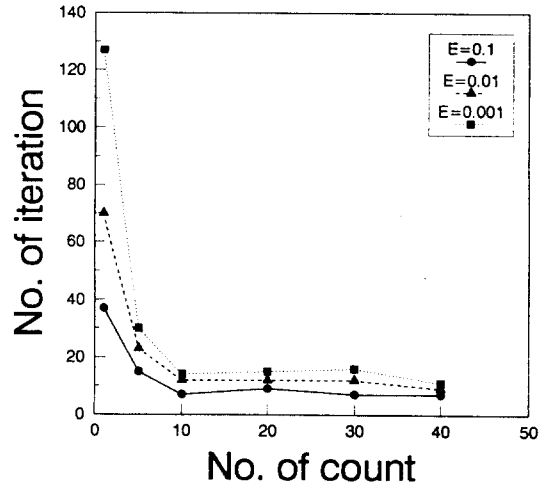
u0=0.5, int. wgt.=1.0, hidden unit=3  
count=10, step ratio=0.5

Fig. 8 Step수에 따른 수렴특성



int. wgt.=1.0, hidden unit=10  
u0=0.5, step=5, count=10

Fig. 7 Step비에 따른 수렴특성



int. wgt.=1.0, hidden-unit=3  
u0=0.5, step ratio=0.5, step=10

Fig. 9 Count수에 따른 수렴특성

Tabu 탐색법에 있어서 중요하게 고려되는 파라미터에는 step수와 count수가 있다. 이들은 각각 변경했을 때의 수렴특성을 비교한 결과를 Fig. 8과 Fig. 9에 표시한다. 이들로 부터 다음을 알 수 있다.

- (1) Step수가 증가하면, 보다 정도가 좋고 상세한 해의 탐색이 가능하게 된다. 그러나, step수가 너무 많이 증가하여 근방영역의 폭이 아주 작아진다면 이 근방영역 중에서의 탐색에 의한 평가함수의 개선이 크게 의미가 없어지는 경우가 있을 수 있다.
- (2) Count수가 증가하면, 오차가 최소, 즉 대국적 최적해에 도달할 확률이 증가하게 된다.

이 두가지에서 주의하지 않으면 안되는 것은 이 2개의 정수를 무한히 크게 하면, 불필요한 탐색시간이나 프로그램시 불필요한 메모리 영역을 사용해 버리기 때문에 이들 정수의 결정에는 시행착오를 요한다.

## 5. 결론

Random 탐색법과 조합된 tabu 탐색법(random tabu 탐색법)을 이용하여 제곱오차를 최소화하는 결합계수를 구하여 신경회로망을 구축하는 새로운 학습알고리즘을 제안하였다.

- (5) Step  $h_k$ 에 대한 결합계수의 근방영역을 정의하고 random 이동이 발생되어 진다.
- (6) Random 이동된 결합계수에 의한 오차와 기존의 결합계수에 의한 오차값을 비교한다. 오차가 작아지지 않았다면 그 근방영역에서 다시 임의 이동하여 오차값을 비교한다. 단, 하나의 근방영역내에서의 random 이동 횟수는 정해진 count수 만큼까지만 반복가능하다. 오차가 작아졌다면 기존의 결합계수를 random 이동된 결합계수로 교체하고 새로운 step  $h_k$ 를 정한 후 (5)의 단계로 간다. 단, step수 만큼 반복하고 오차값이 오차한계를 만족하면 종료하고 만족하지 않으면 (5)의 단계로 돌아간다.

Fig. 3은 이 학습알고리즘의 흐름도(flow chart)를 나타내고 있다.

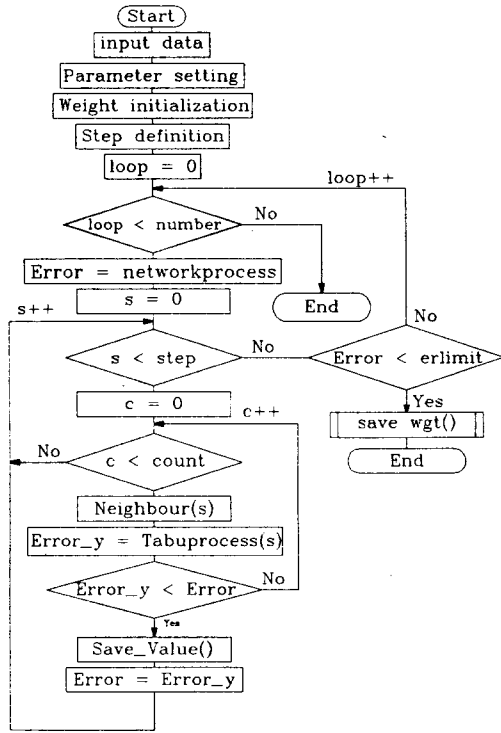


Fig. 3 Random tabu 탐색법을 이용한 학습알고리즘의 흐름도

## 5. 신경회로망의 학습에의 적용 예

본 연구에서는 적용 예로서 배타적 논리합(XOR)을 이용하였다. Fig. 4는 XOR 학습용 학습용 신경회로망이다. Table 1은 학습결과를 기존의 오차역전파법과 비교한 것이다. 좌단부에는 입출력관계가 있는 논리연산자이고, 우단부는 random tabu 탐색법에 의한 학습을 30회 학습한 후의 출력과 오차역전파 알고리즘에 의해 10000회 학습한 후의 출력이다. Random tabu 탐색법에 의한 학습이 학습반복회수가 오차 역전파법에 비해 훨씬 적은데도 오히려 학습결과가 더 좋은 결과를 보이고 있다.

Fig. 5는 기존의 오차역전파법, tabu 탐색법을 이용한 오차역전파법<sup>(6)</sup> 그리고 random tabu 탐색법을 이용한 학습법에 의한 제곱오차의 수렴특성을 보이고 있다. Random tabu 탐색법을 이용한 학습법이 기존의 방법보다 월등히 수렴속도가 빠르고 발산하지 않는다.

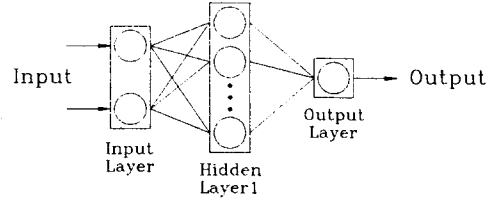


Fig. 4 XOR 학습용 계층형 신경회로망의 구조

Table 1 Learning results(XOR)

Input	Teach	Random Tabu Search			Back Propagation			
		Unit=2	Unit=6	Unit=10	Unit=2	Unit=6	Unit=10	
0	0	0.0	0.00328	0.00332	0.00262	0.00892	0.00385	0.00399
1	0	1.0	0.99592	0.99692	0.99700	0.98974	0.99320	0.99349
0	1	1.0	0.99733	0.99631	0.99653	0.99159	0.99330	0.99405
1	1	0.0	0.00192	0.00120	0.00325	0.00801	0.00814	0.00759

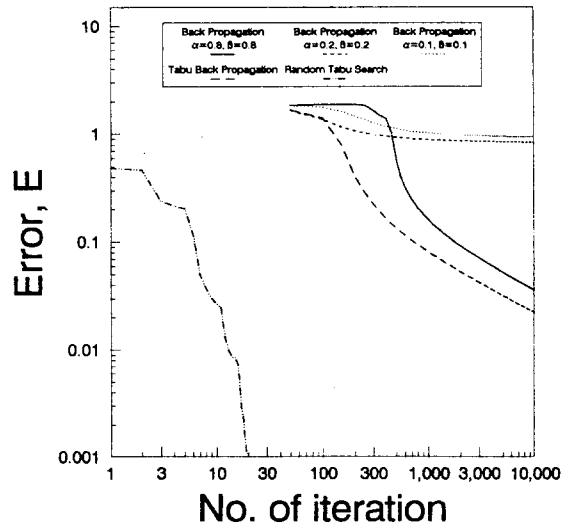


Fig. 5 학습에 의한 제곱오차의 수렴비교

Fig. 6은 시그모이드 함수의 기울기에 따른 오차함수의 수렴정도를 나타내고 있다. 시그모이드 함수의 기울기는 시그모이드 함수의 온도( $u_0$ )의 역수에 상당한다. 시그모이드 함수의 기울기 클수록 수렴정도가 좋으나 상대적으로 일반화 능력이 떨어진다.

Fig. 7은 근방영역의 폭을 단계별로 줄이는 비율, 즉 step비에 따른 학습도를 나타낸 것이다. step비가 클수록 수렴속도가 빠르다.

그리고 본 학습법이 기존의 방법에 비하여 국소적 최적해에 수렴하는 것을 방지하고, 수렴정도를 획기적으로 개선하는 것을 배타적 논리합 문제에 적용하여 확인하였다. 그 결과를 정리하면 다음과 같다.

- (1) 국소최적해에 빠지는 것을 방지할 수 있다.
- (2) 수렴속도가 기존의 학습법에 비해 획기적으로 빠르다.
- (3) Step비가 클 수록 수렴속도가 빠르다.
- (4) Step수와 count수가 증가할 수록 대체로 수렴속도가 증가하고, 대국적 최적해를 찾을 가능성이 커진다.
- (5) 시그모이드 함수의 기울기가 커질 수 수렴속도가 증가한다. 그러나 일반화 능력은 상대적으로 감소한다.

## 참고문헌

- (1) 中野聲 외 4인 著, 박민용, 최항식 譯 “뉴로 컴퓨터”, 大英社, 1991.
- (2) Casimir K., John G., and Garrett P. : “Neural Computing”, Neural Ware, Inc, 1989.
- (3) Shin Morishita., 横浜國立大學工學部, Tabu探索法によるニューラルネットワークの新しい學習法, 1993.
- (4) Hu, N., International Journal Numerical Methods in Engineering, Vol. 35, pp. 1055-1070, 1992.
- (5) Jasbir S. Arora., INTRODUCTION TO OPTIMUM DESIGN, McGrawHill, 1990.
- (6) 김유신, 양보석, 최성필, Tabu탐색법을 이용한 신경회로망의 학습개선에 관한 연구, 대한기계학회 추계학술대회, 1994.