

# 실시간 적응 학습 제어를 위한 진화연산(I)

장성욱\* · 이진걸\*\*

## Evolutionary Computation for the Real-Time Adaptive Learning Control(I)

Sung-Ouk Chang, Jin-Kul Lee

**Key Words:** Evolutionary Algorithms(진화알고리즘), Real-time adaptive learning(실시간 적응학습 제어), mutation(돌연변이), estimation(평가), selection(선택).

### Abstract

This paper discusses the composition of the theory of reinforcement learning, which is applied in real-time learning, and evolutionary strategy, which proves its the superiority in the finding of the optimal solution at the off-line learning method. The individuals are reduced in order to learn the evolutionary strategy in real-time, and new method that guarantee the convergence of evolutionary mutations are proposed. It possible to control the control object varied as time changes. As the state value of the control object is generated, applied evolutionary strategy each sampling time because the learning process of an estimation, selection, mutation in real-time

These algorithms can be applied, the people who do not have knowledge about the technical tuning of dynamic systems could design the controller or problems in which the characteristics of the system dynamics are slightly varied as time changes. In the future, studies are needed on the proof of the theory through experiments and the characteristic considerations of the robustness against the outside disturbances.

### 1. 서 론

실제 산업현장에서 널리 사용되고 있는 PID와 같은 고전제어기의 경우, 시스템에 대한 정확한 동적모델이 필요하며, 시스템의 성능을 우수하게 하기 위한 모델에 대한 이해를 바탕으로 한 제어기의 튜닝(turning)등이 수반된다[1][2][3]. 인공지능적 기법을 이용한 제어시스템의 설계는 이러한 불편함을 해소하고, 실제 동적모델의 불확실성 뿐만아니라 제어기의 인위적인 튜닝을 학습을 통하여 자동으로 할수 있지만, 학습시간 및 전역범위에서의 데이터의 탐색이 필요하므로 실시간으로 학습하는 것이 매우 힘들다[4][5]. 한편, 인

공지능 기법을 이용한 학습방법은 지도학습(supervised learning), 비지도학습(unsupervised learning), 강화학습(reinforcement learning)으로 나누어지며, 지도학습과 비지도 학습의 경우 각각의 알고리즘 적용시 실시간으로 제어기를 구성하는 것이 매우 어려우며, 실시간 학습 제어기로 강화학습(reinforcement learning)에 대한 많은 연구가 행해지고 있으나, 강화학습의 경우 성공률의 지속성이 현저히 낮고, 연속된 성공률의 보장이 어렵다는 단점을 가지고 있다[6][7][8]. 진화알고리즘(Evolutionary Algorithm)은 전역적 범위에서 최적의 해를 탐색하는 데 널리 사용되고 있으나, 수 많은 개체속에서 개체간의 자연적인 경쟁을 통해 최적의 해를 구하는 알고리즘 특성상 많은 계산량을 요구하기 때문에 오프라인(off-line)상태에서 주로 사용되었으며 실시간으로 사용되지 못하는 단점을 지니고 있다[9][10][11][12].

\* 부산대학교 지능기계공학 박사과정

\*\* 부산대학교 기계공학부 교수

본 연구에서는 오프라인 학습법에서 최적의 해를 찾는 우수성이 입증된 진화전략 알고리즘과 실시간에서의 강화학습적 개념을 복합적으로 구성하였으며, 진화전략알고리즘의 실시간 학습을 위한 개체수를 조정함으로써 실시간 시에 필요한 계산량을 줄이고, 돌연변이 진화시의 수렴성을 높이기 위한 새로운 기법을 제안한다. 따라서 제어대상의 상태 값이 발생하는 매 순간마다 진화 알고리즘을 적용하여 평가 와 선택 그리고 재생산(mutation)과정을 실시간으로 학습하기 때문에 제어대상의 변화에 능동적으로 대처 할 수 있다.

## 2. 이론적 배경

### 2.1 진화전략 알고리즘을 이용한 실시간 신경망 학습

자연세계의 진화과정을 기본모델로 하는 진화 전략 알고리즘은 부모의 교배에 의한 새로운 개체의 생성과정과 생성된 자식과 부모개체간의 경쟁을 통한 우성인자의 적자생존인 선택과정으로 크게 나누어진다[13]. 자연세계의 진화과정을 살펴보면 교배에 의한 진화보다는 돌연변이에 대한 영향이 더 큰 것으로 연구보고 되고 있다[14]. 진화전략 알고리즘은 기본적인 알고리즘 특성상 부모개체와 자식개체간의 경쟁으로 인하여 자연스런 탐색범위조정과 평가, 선택이 이루어지는 구조이므로, 경쟁할 수 있는 부모와 자식간의 개체수가 미리 형성되어야 하고, 개체수를 줄이면 국소탐색에 그치거나 많은 학습시간을 요구하게 된다. 이러한 많은 경쟁할 수 있는 개체의 형성 및 탐색 방법 때문에 진화전략 알고리즘 적용 시 오프라인(off-line)학습으로 최적의 해를 탐색하는 연구가 주로 행하여 졌다[15][16]. 한편, 진화전략 알고리즘을 실시간으로 학습적용하기 위해서는 피드백 시스템(feedback system)에서 샘플링타임 안에 모든 학습이 완료되어야 하므로 개체수가 부모, 자식 각각 1개로 한정되어야 하며 이러한 개체수의 감소에 따른 학습시간의 증가, 탐색범위의 조정문제, 수렴성에 대한 고찰 등의 문제가 해결되어야 한다[17][18][19].

본 연구에서는 실시간으로 학습이 가능하고 부모·자식간의 개체수를 각각1개로 한정하였을 때, 개체수의 감소에 따른 탐색범위 조정의 어려움을 해결하기 위하여 시스템의 제어시 발생하는

오차(error)량을 이용하여 다음과 같은 수식을 제안한다.

$$\sigma(t+1) = \sigma(t) \cdot \exp(N(0, e(t))) \quad (1)$$

$$x(t+1) = x(t) + N(0, \sigma(t+1)) \quad (2)$$

진화전략 알고리즘은 해의 탐색범위를 가우스 분포를 따르는 정규분포를 이용하여 랜덤하게 탐색하는 특성을 지니고 있으며, 개체간의 경쟁에 의해 자연스런 탐색범위 조정을 그 특징으로 한다. 그러나 실시간 학습으로 개체 수를 줄일 경우, 많은 개체간의 자연스런 경쟁을 통한 탐색범위의 자동조정이 불가능하다[20][21].

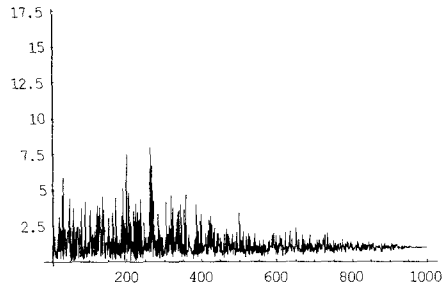


Fig. 1 The diagram of sigma according to decrease error

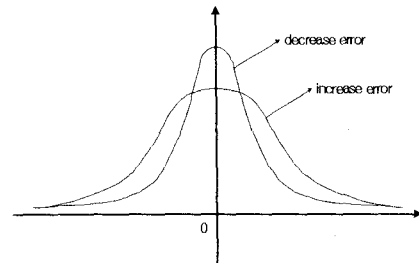


Fig. 2 The Gaussian distribution for the increase or decrease error

수식(1)에서 보면 평균이 0, 분산이  $(\sqrt{e})^2$ 인 정규 분포의 값의 영향을 받아  $\sigma$ 가  $\sigma(t+1)$ 로 변화하는 것을 알 수 있다. Fig.1 은 수식(1)의 오차량의 감소에 따른  $\sigma(t+1)$ 의 변화량을 도식적으로 나타내기 위하여 수학적 모델식의 해석에 널리 사용되고 있는 매스매티카(mathematica) 프로그램을 사용하여 나타낸 것이다[22].

Fig.1에서 나타낸 바와같이 오차량이 점점 감소함에 따라 편차값이 일정한 값으로 수렴됨을 알 수 있다. 즉 수식(1)과 Fig.1, Fig.2에서 나타낸 바와 같이 시스템의 오차값에 대하여 오차량의 증가에 따라 탐색범위조정이 자동으로 이루어지며,

오차값이 증가하면  $\alpha(t+1)$ 값이 증가함으로써 전역적인 탐색특성이 증가하며, 오차값이 감소함에 따라  $\alpha(t+1)$ 값이 감소함으로써 국소적인 탐색이 가능하다.

결국 출력  $x(t)$ 는  $\alpha(t+1)$ 값에 따라 새로운  $x(t+1)$ 값을 발생시키므로,  $\alpha(t+1)$ 값이 0을 기준으로 대칭인 정규분포인 곡선을 나타내므로  $x(t)$ 의 값에 양의 값이나 음의 값이 더해질 확률은 확률적으로 항상 50%이다. 즉 샘플링 타임마다  $x(t+1)$ 값이  $x(t)$ 값보다 증가하거나 감소할 확률이 정확히 50%이기 때문에 실제시스템은 샘플링 시간마다 제어입력을 수식적으로 자동적으로 가감하게 된다.

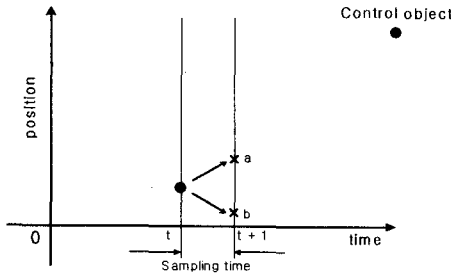


Fig. 3 The position Control of real-time learning over each sampling time

발생하는 시스템의 출력은 Fig.3에서 나타낸 바와 같이 a지점 혹은 b지점으로 출력을 나타내게 되고, a값의 경우만을 선택하고 재생산함으로써 실시간으로 진화 학습하게 된다.

### 3. 전기 유압 서보 시스템

#### 3.1 유압서보장치

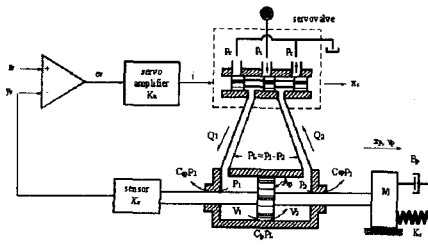


Fig. 4 Schematic diagram of electro-hydraulic servo system

유압 서보장치는 위치제어의 경우 Fig.4에서 나타낸 바와같이 유압공급장치, 서보밸브, 부하,

위치센서, 밸브 구동용 앰프 등으로 구성된다

#### 3.2 서보밸브의 수학적 모델

서보밸브는 사용되는 주파수역에 대하여 시스템의 주파수 응답특성을 고려하여 2차 지연 요소로 모델링 하였다[23].

$$\frac{d^2x_s(t)}{dt^2} + 2\xi_v \omega_v \frac{dx_s(t)}{dt} + \omega_v^2 x_s(t) = \omega_v^2 \cdot K_t \cdot i(t) \quad (3)$$

여기서,  $K_t$ 는 토크 모터 게인[cm/mA]이고,  $\xi_v$ 는 서보밸브의 감쇠비,  $\omega_v$ 는 서보밸브의 비감쇠 고유진동수,  $x_s$ 는 서보밸브의 스톱변위,  $i$ 는 서보앰프의 입력전류이다.

#### 3.3 서보 밸브의 유량 특성

Fig.3에 표시한 화살표 방향이 실린더 및 서보 밸브 스톱 변위의 양의 방향이라고 하면 식(3)의 서보 밸브 스톱의 변위  $x_s(t)$  [cm]로 인하여 스톱을 통과하는 기름의 공급측 유량  $Q_1(t)$  [cm<sup>3</sup>/s] 및 귀환측 유량  $Q_2(t)$  [cm<sup>3</sup>/s]는 스톱 오리피스 유량 방정식으로부터 각각 식(4), 식(5)와 같다[24].

$$Q_1(t) = C_d w x_s(t) \sqrt{\frac{2}{\rho} (P_s - P_1(t))} \quad (4)$$

$$Q_2(t) = C_d w x_s(t) \sqrt{\frac{2}{\rho} P_2} \quad (5)$$

여기서  $C_d$ 는 밸브 스톱의 오리피스 유량 계수,  $w$ 는 밸브 포트의 단면적 구배[cm],  $\rho$ 는 작동유의 밀도[kg·s<sup>2</sup>/cm<sup>4</sup>],  $P_s$ 는 공급 압력[kg/cm<sup>2</sup>] 그리고  $P_1$  및  $P_2$ 는 각각 유압 실린더 내의 공급측과 귀환측의 압력[kg/cm<sup>2</sup>]을 나타낸다.

#### 3.4 양로드 유압 실린더의 부하 유량식

서로 대칭인 양로드 유압 실린더의 좌우 양실에 연속 방정식을 각각 적용하면 실린더의 공급측 유량  $Q_1(t)$  과 귀환측 유량  $Q_2(t)$ 는 각각 식(6)과 식(7)과 같이 나타낼 수 있다.

$$Q_1(t) = C_{ip}(P_1(t) - P_2(t)) + C_{ep}P_1(t) + \frac{dV_1(t)}{dt} + \frac{V_1(t)}{\beta_e} \frac{dP_1(t)}{dt} \quad (6)$$

$$Q_2(t) = C_{ip}(P_1(t) - P_2(t)) - C_{ep}P_2(t) - \frac{dV_2(t)}{dt} - \frac{V_2(t)}{\beta_e} \frac{dP_2(t)}{dt} \quad (7)$$

여기서  $C_{ip}$  및  $C_{ep}$ 는 각각 내부, 외부 누설 계수[ $\text{cm}^3/\text{kg}/\text{s}$ ]이고,  $\beta_e$ 는 작동유의 체적 탄성 계수[ $\text{kg}/\text{cm}^2$ ]이다. 양로드 유압 실린더에서의 부하 유량 방정식은 식(8)과 같이 정리된다.

$$Q_L(t) = A_p \frac{dx_p(t)}{dt} + C_t P_L(t) + \frac{V_t}{4\beta_e} \frac{dP_L(t)}{dt} \quad (8)$$

여기서,  $V_t$ 는 밸브, 유압관로 및 실린더 전체의 체적[ $\text{cm}^3$ ],  $C_t$ 는 피스톤 전체 누유 계수로

$$C_t = (C_{ip} + \frac{C_{ep}}{2}) [\text{cm}^3/\text{kg}/\text{s}] \text{이다.}$$

### 3.5 양로드 유압 실린더의 운동 방정식

한편 양로드 유압실린더의 운동 방정식은 식(9)과 같다.

$$A_p P_L(t) = M \frac{d^2 x_p(t)}{dt^2} + B_p \frac{dx_p(t)}{dt} + K_s x_p(t) + F_L \quad (9)$$

여기서  $A_p$ 는 실린더의 수압 면적[ $\text{cm}^2$ ]이고,  $M$ 은 관성 질량 부하 [kg·s<sup>2</sup>/cm],  $B_p$ 는 점성 부하[ $\text{kgf} \cdot \text{s}/\text{cm}$ ],  $K_s$ 는 탄성 부하 [kg/cm], 그리고  $F_L$ 은 외란[ $\text{kg}$ ]의 항으로서 부하에 의한 외란과 쿨롱 마찰력 및 고체 마찰력에 의한 외란 등으로 구성된다.

## 4. 시뮬레이션 결과 및 고찰

진화전략 알고리즘을 이용하여 실시간으로 신경망의 가중치를 업데이트 하는 과정을 Fig.5에 나타내었다. 그림에서 나타낸 바와 같이 신경망에서 학습한 제어신호가 유압서보장치를 제어하도록 구성하였다. 유압서보장치에서 나온 출력 상태값(state)은 샘플링 시간마다 신경망 제어기로 입력되며, 신경망의 각각의 가중치는 출력 상태값을 근거로 실시간으로 진화전략알고리즘을 적용하여 매 순간 업데이트(update)됨으로써 실시간

으로 학습제어가 가능하다. 실시간 학습을 위하여 매 샘플링 시간마다 시스템의 출력값을 근거로 학습의 성공, 실패를 판단하며, 성공시에 1의 값을 실패시에 -1값을 리턴(return)하도록 함으로써 매순간 학습의 성공과 실패를 체크하도록 프로그래밍하였다[25][26].

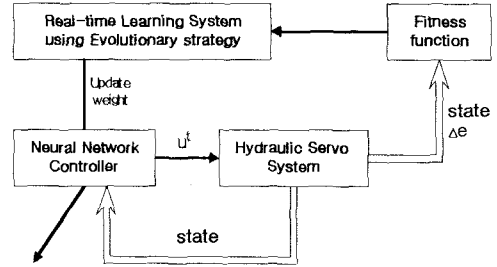


Fig. 5 Block diagram of the real-time learning system using evolutionary algorithm.

즉, 실시간 학습을 위해서 매 샘플링 시간마다 시스템의 출력 상태값을 근거로 학습의 성공과 실패를 평가함수(fitness function)에서 판단하여 진화알고리즘에서 선택 재생산의 과정을 거친후 신경망의 새로운 가중치를 업데이트하여 제어입력을 발생시킨다. 컴퓨터 시뮬레이션은 샘플링 시간은 1[ms]이고, 오차량과 오차량의 변화율을 신경망의 입력으로 하였으며, 시스템의 출력값을 근거로 오차량의 변화율을 평가의 기준으로 적용하였다.

자연세계에서 진화시에 매 세대마다 진화하기 위해서는 새로운 개체의 생성이 반드시 필요하다. 따라서 실시간 학습을 위해서 진화전략 알고리즘을 적용시에 매 샘플링 타임마다 새로운 개체를 생성하는 것이 매우 중요하며, 이러한 개체의 생성과 평가에 따라 시스템의 학습능력이 결정되어 진다. 그러므로 시뮬레이션은 매 샘플링 타임마다 성공·실패를 체크하고 성공·실패에 따른 개체의 평가 선택에 의하여, 적자 생존한 우수개체를 매 순간마다 돌연변이를 발생시킨다. 이러한 돌연변이 연산은 정규 가오스분포(Normalized Gaussian Distribution)를 따라 랜덤(random)하게 이루어지지만,  $\sigma(t+1)$ 가 0을 중심으로 대칭이므로 양의 돌연변이와 음의 돌연변이 발생률이 확률적으로 같기 때문에, 돌연변이에 의한 학습은 2스텝마다 최소한 진척을 나타낼 것으로 예상할 수 있다.

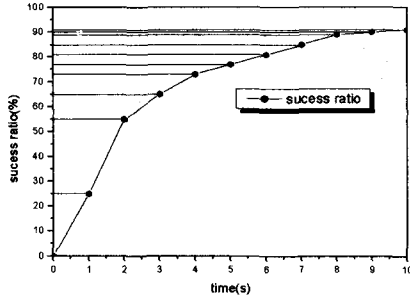


Fig. 6 The simulation result of the success ratio over learning

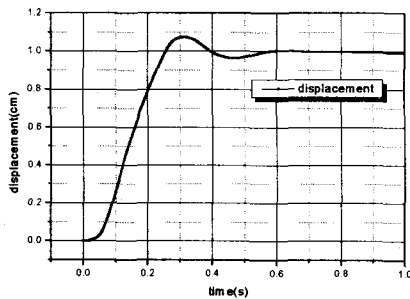


Fig. 7 The simulation result of real-time learning system using evolutionary algorithm.

Fig. 6 는 100번의 시뮬레이션 수행 시 학습시간에 따른 성공률을 나타낸 것이다. 샘플링 시간이 1[ms]이므로 시간1초는 1000번의 학습 및 돌연변이 발생, 평가가 이루어 졌음을 알 수 있다. 그림에서 나타낸 바와 같이 학습시간 2초 이내에 성공률이 55%로 매우 뛰어나며, 10초 이내의 성공률이 91%정도이므로 확률적으로 제시한 오차량에 의한 탐색범위 조정에 의한 실시간 학습의 타당성을 확인할 수 있었다. 한편 Fig. 7은 실시간 학습에 의한 유압시스템의 위치제어를 행하였을 때의 결과를 나타낸 것이다. 그림에서 나타낸 바와 같이 시스템의 출력이 상승시간 0.2[sec], 정착시간 0.3[sec]로 매우 뛰어난을 알 수 있다.

Fig. 6, Fig. 7 결과로 유추해 볼 때 실제시스템에서 실시간으로 학습제어기를 적용 가능함을 확인할 수 있었으며, 시스템에 대한 정확한 지식이 없더라도 이러한 학습기법을 적용함으로써 실제시스템의 제어가 가능함을 알 수 있었다. 또한

제어기의 학습이 시스템의 매 순간 순간의 출력을 근거로 이루어지므로 시간에 따라 시스템의 특성이 변화하는 곳에서도 이러한 기법을 적용할 수 있음을 확인할 수 있었다.

#### 4. 결론

본 연구에서는 진화전략 알고리즘을 신경망에 적용하여 실시간으로 시스템의 특성에 맞는 학습을 수행하여 제어기를 설계하였으며, 그 타당성을 컴퓨터 시뮬레이션을 통하여 확인하였다. 신경망을 이용한 학습은 제어와 플랜트(plant)을 함께 고려하여 학습하는 것이 힘들며, 이러한 것을 보완하기 위해 최근에 실시간으로 제어기를 학습하는 강화학습에 대한 연구가 활발하다.

본 연구에서 제시한 실시간 학습은 널리 쓰이고 있는 신경망의 기본구조를 그대로 사용하였으며, 신경망의 가중치를 업데이트 하기 위해 진화전략 알고리즘을 적용하였다. 한편 진화전략 알고리즘의 실시간 학습을 위하여 개체수를 줄이고 매 샘플링 시간마다 재생산, 평가, 선택의 과정을 행하였다. 또한 진화전략 알고리즘의 특성인 많은 개체간의 자연스런 경쟁에 의한 자동 탐색 범위조정을 시스템에서 발생하는 오차량을 이용함으로써 실시간성 보장을 위한 경쟁 개체 수를 줄였을 때의 문제점을 오차에 의한 자동 탐색으로 범위를 조정함으로써 해결하였다. 따라서 제시한 제어알고리즘의 적용으로 인하여 시스템에 대한 전문적인 튜닝지식이 없는 사람이 제어기를 설계하고자 할 때나, 시간에 따라 플랜트의 특성이 조금씩 변화하는 시스템의 제어기의 설계에 널리 사용 가능함을 확인할 수 있었다. 앞으로 실험을 통한 이론의 검증과 외부부하에 대한 제어기의 강인성에 대한 특성고찰에 대한 연구가 지속적으로 필요하다.

#### 참고문헌

- [1] S.O.Chang, J.K.Lee, "Development of a Hydraulic Servo System Real-Time Simulator Using a One-Board Microprocessor and personal computer," Proceedings of the 4th Asia-Pacific conference on Control and Measurement, pp128-132, 9-12 July, 2000.
- [2] 장성욱, 이진걸, "원보드 마이크로 프로세서 제어기 및 PC를 이용한 유압서보시스템의

실시간 시뮬레이터 개발," 한국정밀공학회지, 제17권, 제8호, pp94-99, 2000.

- [3] S.O.Chang, J.K.Lee, "A Study on the Application of the Real-Time Simulator," Proceedings of the 15th Korea Automatic Control Conference, October 19-21, 2000
- [4] A. K. Jain, et al, "Artificial Neural Networks: A Tutorial," IEEE Computer, vol. 26, no. 3, pp. 31-44, 1996.
- [5] R.P.Lippmann, "An Introduction to Computing with Neural Nets," IEEE ASSP Magazine, Vol.4, No.2, pp4-22, Apr. 1987
- [6] Simon Haykin, Neural Networks : A comprehensive foundation, second Ed. 1999, Prentice-Hall Inc.
- [7] Charles W. Anderson, "Learning to Control an Inverted Pendulum Using Neural Networks", IEEE Control Systems Magazine, Vol. 9, No. 2, pp. 31-37, April 1989.
- [8] A. G. Barto, R. S. Sutton and C. W. Anderson, "Neuronlike Adaptive Elements That Can Solve Difficult Learning Control Problem," IEEE Transactions on Systems, Vol.13, No.5, pp. 835-846, September/October 1983.
- [9] D.B. Fogel, "The Advantages of Evolutionary Computation," Proc. of BCEC97: BioComputing and Emergent Computation, World Scientific, Singapore, pp.1-11, 1997.
- [10] D.B. Fogel and L.J. Fogel, "Preliminary Experiments on Discriminating between Chaotic Signals and Noise Using Evolutionary Programming," Proceedings of the First Annual Conference, MIT Press, pp.512-520, 1996.
- [11] K. Chellapilla and D. Fogel, "Two New Mutation Operators for Enhanced Search and Optimization in Evolutionary Programming," in Applications of Soft Computing, Proc. SPIE Vol. 3165, pp. 260-269, 1997.
- [12] Back, T and Schwefel, H.-P, "Evolutionary Computation: An Overview", Proceedings of IEEE International Conference on Evolutionary Computation, pp20-29, 1996.
- [13] Fogel D B, "Evolutionary Computation. Toward a New Philosophy of Machine Intelligence," Pisoataway, NJ : IEEE Press. 1995
- [14] E Vonk, L -C Jain, R -P Johnson, "Automatic Generation of Neural Network Architecture Using Evolutionary Computation," World Scientific Publishing Co., 1997
- [15] D.B. Fogel and J.W. Atmar, "Comparing Genetic Operators with Gaussian Mutations in Simulated Evolutionary Processes Using Linear Systems," Biological Cybernetics, Vol. 63:2, pp. 111-114, 1990.
- [16] Back, T," Evolutionary Algorithms in Theory and Practice," Oxford, NY, 1996.
- [17] Schwefel, H.P." Evolution and Optimum Seeking", John Wiley, NJ, 1995.
- [18] Z. Michalewicz and D. B. Fogel, How to Solve It: Modern Heuristic, Springer-Verlag, Berlin, pp. 161-184/335-341, 2000.
- [19] E. Sanchez and M. Tomassini, eds., Toward Evolvable Hardware: The Evolutionary Engineering Approach, Springer-Verlag, Berlin, pp. 19-47/221-249, 1996.
- [20]. T. Back, U. Hammel and H. P. Schwefel, "Evolutionary Computation: Commnets on the History and Current State," IEEE Transaction on Evolutionary Computation, Vol. 1, No. 1, pp. 3-17, 1997.
- [21] D. B. Fogel, "An Introduction to Simulated Evolutionary Optimization," IEEE Transaction on Neural Network, Vol. 5, No. 1, pp. 3-14, 1994.
- [22] Martha L. Abell et all, "The mathematica handbook," Boston: Academic Press. 1992.
- [23] Moog, "Technical Bulletin," Moog Controls Division, East Autora N.Y, 1981.
- [24] Herbert. E. Merrit, " Hydraulic Control Systems", John & Wiley & sons Inc, 1967.
- [25] S. Geva and J. Sitte, "A Cart Experiment Benchmark for Trainable Controllers," IEEE Control Systems, Vol. 13, No. 5, pp. 40-51, 1995.
- [26]. R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, The MIT Press, Cambrigde, 1998.