

유전자 알고리즘에 대한 수렴특성의 개선

Improvement of Convergence Properties for Genetic Algorithms

이홍규*

Hong-Kyu Lee*

요 약

유전자 알고리즘은 효과적으로 최적의 해를 구하는 기법이나 진화연산자의 선정에 따라 조기에 국부 최적해에 고착되어 전역 최적해로의 탐색을 어렵게 하는 문제점을 가지고 있다. 본 논문에서는 국부 최적해로 수렴하게 되는 원인을 분석하고, 국부 최적해에서 벗어나 전역 최적해로의 천이가 가능하도록 하는 방법을 제안하였다. 본 논문에서 사용한 방법은 평균 해밍거리에 따라 진화연산자를 가변시키는 방법으로서 국부 최적해에 고착되지 않도록 유전자에 다양성을 부여하여 지속적으로 모집단의 진화 특성을 유지하는 방법이다. 제안된 방법은 시뮬레이션을 통하여 효용성을 입증하였다.

Abstract

Genetic algorithms are efficient techniques for searching optimum solutions but have the premature convergence problem getting stuck in the local optimum according to the evolutionary operator. In this paper we analyzed the reason for converging to the local optimum and proposed the method which able transit to the global optimum from the local optimum. In these methods we used the variable evolutionary operator with the average hamming distance, to maintain the genetic diversity of the population for getting out of the local optimum. The theoretical results are proved by the simulation experiments.

Key words : Genetic Algorithms, Premature Convergence Properties, Average Hamming Distance

I. 서 론

유전자 알고리즘(GA(genetic algorithm))은 최적해를 탐색하는 매우 유용한 기법으로서 점차 응용영역이 확장되고 있다. 유전자 알고리즘에서 모집단은 다수의 염색체들로 구성되어 있고, 각 염색체의 비트배열들은 통상 초기화 과정에서 임의로 생성된다. 유전자 알고리즘은 높은 적합도 함수 값을 가진 염색체가 생존 확률이 높은 특성을 가지고 있기 때문에 세대가 경과함에 따라 모집단에 있는 모든 염색체들이 높은 적합도

값을 가지게 되고 비트배열 또한 유사하게 된다[1].

최근에 연구되고 있는 분야중의 하나는 유전자 알고리즘의 수렴특성에 대한 체계적인 분석과 조기에 수렴하여 국부 최적해에 고착되는 것을 방지하는 방법이다. Goldberg는 스키마이론에서 building block growth 개념을 제안하였고, uniform crossover가 고적합도 schemata를 파괴할 수도 있다는 점을 지적하였다[2]. Rudolph는 CGA(canonical genetic algorithm)의 기본 특성은 전역 최적해에 수렴하는 것을 보장하는 것이 아니고 최선의 해를 구하는 것이며, 돌연변이

* 한국기술교육대학교(School of Information Technology, Korea University of Technology and Education)

- 제1저자 (First Author) : 이홍규
- 투고일자 : 2008년 9월 22일
- 심사(수정)일자 : 2008년 9월 23일 (수정일자 : 2008년 10월 16일)
- 게재일자 : 2008년 10월 30일

연산자만이 전역 최적해로 수렴할 가능성을 부여하는 요소라는 것을 주장하였다[3]. 엘리트선택(elitist selection) 방법을 사용하는 CGA인 경우에 전역 최적해로 수렴할 수 있다는 것을 증명하였다[4]. Leung은 기본적인 선택과 진화연산자의 특성들을 분석하여 abstract evolutionary algorithm에 대한 몇 개의 수렴정리와 수렴속도를 예측하는 방법을 제안하였다[5]. Chakraborty는 돌연변이 확률이 높고 모집단(population)의 규모가 작으면 엘리트방법을 사용하는 유전자 알고리즘(EGA(genetic algorithms with elitism))이 국부 최적해에 빠질 수 있으나, 작은 돌연변이 확률의 경우에는 EGA는 안정된 값으로 점차로 감소되어 국부 최적해에 빠지지 않는다는 것을 발견하였다[6]. Altenberg 등은 비트배열들과 전역 최적해에 대한 비트배열사이의 해밍거리들을 적합도(fitness) 함수들과 비교하는 fitness distance correlation 접근방법을 제안하였다[7, 8]. Goldberg 등은 세대와 세대 사이의 평균 적합도 값의 변화 정도를 이용하여 수렴에 소요되는 시간을 예측하였고, 수렴을 위해서는 모집단의 크기가 어느 정도 규모 이상이 되어야 한다는 것을 제안하였다[9].

유전자 알고리즘(GA)에서 개체군의 다양성(diversity)과 수렴(convergence) 특성을 유지하기 위해서는 교차(crossover)와 돌연변이(mutation) 연산 과정이 필요하다. 또한 높은 적합도(fitness)를 갖는 개체는 교차나 돌연변이에 의하여 개체가 변형되지 않도록 보호되는 것이 요구된다. 유전자 알고리즘의 궁극적인 목적은 전역 최적해(global optimum solution)를 구하는 것인데, 조기수렴으로 인한 국부 최적해(local optimum solution)에 고착되는 것을 방지하는 것이 필요하다[10]. 적응 GA에서는 표준적인 GA에 비하여 복잡한 정도의 변화가 매우 심한 multimodal 함수를 최적화하는데 좋은 성능을 갖는다. 대부분의 함수들에 대하여 적응 GA는 표준적인 GA에 비하여 전역 최적해에 수렴하는 속도가 빠르고, 국부 최적해에서 빠져나오지 못하는 경우도 줄어들게 된다[11, 12]. 국부 최적해로 수렴하지 않고 전역 최적해로 수렴하도록 하는 방법으로 모집단의 크기를 증대하는 방법이 사용되기도 하였다[10, 13]. 이는 모집단의 크기를 크게 함으로서 돌연변이 연산자에 의하여 국부 최적해

가 아닌 다른 위치가 선택될 수 있는 확률이 높아지게 하는 것으로 해석된다. Affenzeller는 조기수렴 문제를 해결하기 위하여 모집단으로부터 reproduction에 의한 자손세대 중에서 부모세대보다 적합도가 높은 개체들을 이용하여 먼저 모집단을 구성하고 나머지 부분은 부모세대보다 적합도가 낮은 개체들로부터 선택하여 모집단을 구성하는 적응선택모델을 제안하였다[13].

진화 로봇에 유전자 알고리즘을 이용하여 로봇 스스로 최적의 경로를 탐색하고[14-16], 항공관제시스템 등에 응용한 예로서는 활주로와 공항 도착 예정시간, 항공기의 크기 등으로부터 대기시간을 최소화 하는 항공기의 착륙순서와 활주로를 배정하는 방법으로 유전자 알고리즘이 사용되었다[17]. 또한 로봇의 최적 이동경로를 결정하는데 응용한 예로서 장애물에 충돌 가능성을 최소화 하고, 이동 소요시간을 최소화 하는 경로를 결정하는 방법으로 유전자 알고리즘이 사용되었다[18].

본 논문에서는 국부 최적해로 수렴하지 않고 전역 최적해로 수렴하도록 하는 방법으로 평균 해밍거리에 기반을 두고 가변의 교차 연산자 및 가변의 돌연변이 연산자를 사용하는 방법을 제안하였다.

II. GA의 수렴특성 분석

2-1. 해밍거리와 스키마

비트배열 $[a]$ 와 배열 $[b]$ 사이의 해밍거리 $h(t)$ 는 다음과 같이 정의된다[1].

$$h(a(t), b(t)) = \sum_{i=1}^L |a_i(t) - b_i(t)| \quad (1)$$

여기서 i 는 비트 좌표이고, L 은 염색체 비트배열의 길이이다. t 세대에서 모집단의 평균 해밍거리 $Ah(t)$ 는 다음과 같이 정의된다.

$$Ah(t) = (\sum_{i=1}^L |a_i(t) - b_i(t)|) / (N(N-1)/2) \quad (2)$$

여기서 N 은 개체(염색체)의 수, t 는 세대를 나타낸다. 돌연변이가 없는 경우에는, 수렴된 모집단의

평균 해밍거리는 영(0)이다. 그러나 돌연변이가 있는 경우에는 돌연변이 확률의 크기에 따라 수렴된 모집단의 평균 해밍거리는 증가하게 된다. 수렴속도가 느릴수록 진화특성이 잘 나타나기 때문에 통상 높은 교차 확률이 선택된다.

스키마 이론에 따라 building block이 성장하기 위한 조건은 다음과 같다[1, 3].

$$\frac{m(H,t+1)}{m(H,t)} \geq \frac{f(H,t)}{f(t)} [1 - P_c \frac{\delta(H)}{L-1} - o(H)P_m] \quad (3)$$

여기서 $m(H,t+1)$ 은 $t+1$ 세대에서의 스키마 인스턴스 H 의 수, $m(H,t)$ 는 t 세대에서의 스키마 인스턴스 H 의 수를 나타낸다. $f(H,t)$ 는 t 세대에서 측정된 스키마의 평균 적합도이고, $f(t)$ 는 t 세대에서 모집단내에 있는 모든 개체의 평균 적합도이다. $\{\delta(H), o(H), P_c, P_m\}$ 들은 각각 스키마의 길이, 스키마의 차수, 교차 확률 및 돌연변이 확률을 나타낸다. 알고리즘이 수렴하기 위해서는 선택압(selection pressure)을 나타내는 (4)식과 (5)식을 만족해야 한다.

$$\frac{f(H,t)}{f(t)} \geq 1 \quad (4)$$

$$1 - P_c \frac{\delta(H)}{L-1} - o(H)P_m > 0 \quad (5)$$

모집단이 수렴하는 경우에는 스키마 차수 $o(H)$ 가 최대값 L 까지 증가하게 된다. 그러므로 성장속도(growth rate)가 1보다 커지게 하기 위해서는 P_m 을 아주 작은 값으로 선택할 필요가 있다. 수렴특성만을 고려하면 돌연변이 확률을 작은 값으로 정해야 하나 유전자의 다양성에 문제가 발생되며, 유전자의 다양성을 유지하기 위해서는 적절한 크기의 돌연변이 확률이 적용되어야 한다.

돌연변이가 없으면 GA의 평균 해밍거리는 세대가 진행됨에 따라 비트배열 길이의 1/2에서 0으로 감소된다. 선택된 부모개체에 대하여 교차와 돌연변이 연산을 통하여 진화된 자손사이에는 큰 해밍거리 값을 가지나, 교차에 의한 진화인 경우에는 단지 선택압에만 영향을 미치게 되고, 해당 스키마가 선택될 확률이 높아져 해당 블록이 모집단내에 많아지게 되

는 효과를 가져 온다. 따라서 수렴특성을 관찰하는 데는 모집단 전체에 대한 평균 해밍거리를 수렴특성의 지표로 삼는 것이 효과적이다.

2-2. 국부 최적해와 전역 최적해의 조건

Rudolph는 엘리트 선택방법을 사용하는 영이 아닌 작은 값을 갖는 목적함수에 대하여 전역 최적해에 수렴하는 자체 적응방법(self-adaptation method)을 제안하였다[4].

모집단의 각 세대가 최적해로 수렴했다면 그 모집단의 평균 해밍거리는 인접한 세대사이에 아주 작은 크기만이 변화될 것이다. 일반적으로 교차 연산자만을 사용하는 경우에 모집단은 빠르게 수렴하고, 국부 최적해에 빠지기 쉬워진다. 따라서 돌연변이는 조기 수렴을 방지하고 전역 최적해로 수렴할 수 있도록 유전자의 다양성을 유지시켜주는 매우 중요한 파라미터이다. 그러나 높은 돌연변이 확률은 좋은 스키마를 나쁜 것으로 파괴하기 쉽기 때문에 수렴하지 않고 발산될 가능성이 높아진다. 그러나 돌연변이는 유용한 스키마를 유용하지 않은 것으로 파괴하기 쉽기 때문에 엘리트 선택방법 등의 사용으로 모집단에서 가장 유용한 개체를 유지하기도 한다. 실제로 이 방법은 매우 효과적이거나 모든 형질(allele)들이 나타나도록 할 수도 없고, 보장할 수도 없다.

III. 파라미터 선정

본 논문에서 사용한 GA 알고리즘을 단계별로 정리하면 다음과 같다.

- ① 모집단의 염색체의 수와 비트수를 정하고, random하게 초기화한다.
- ② 목적함수(objective function)로부터 적합도함수(fitness function)를 정하고, 모집단의 염색체들의 적합도를 평가한다.
- ③ 선택기법을 이용하여 가용한 모집단으로부터 부모(parents)를 선정한다.
- ④ 전체 모집단에서 적합도가 높은 상위 일정비율의 개체(elitist)들을 reproduction과정을 거치시

않고 직접 다음 세대로 천이시키는 방법을 사용한다.

- ⑤ 선택된 부모들(mating pool)로부터 평균해밍거리를 기반으로 가변의 교차와 돌연변이 연산과정을 거쳐 자손인 새로운 개체를 생성한다.
- ⑥ 현재세대의 개체들을 새로운 세대의 개체들로 교체한다.
- ⑦ 절차 “②”에서부터 반복한다.

GA 알고리즘 수렴특성의 결과로서 최종 세대 (generation)에서는 모집단에 있는 모든 염색체들이 유사하게 된다. 이 유사성은 거리기반(distance-based) 방법이나 순서기반(sequence-based) 방법으로 측정된다.

유전자 알고리즘이 근본적으로 가지고 있는 두 종류의 수렴문제 중 조기수렴의 경우에는 각 염색체의 비트배열들이 동일한 형태를 가지게 되기 때문에 이들이 국부 최적해에 수렴하게 되면 돌연변이에 의한 모집단의 다양성이 보장되지 않는 한 전역 최적해로 수렴하지 않을 뿐만 아니라 진화 특성도 잃게 된다. 따라서 모집단이 가지고 있는 유전자의 다양성 (diversity)이 수렴특성에 따라 가변적으로 변화되어 항상 어느 정도의 다양성을 유지시키는 것이 중요하다.

모집단의 관점에서 보면 교차 연산자에 의한 진화는 모집단내에 동일한 비트배열의 블록이 증가하는 결과를 가져오게 되고, 따라서 평균 해밍거리와 스키마 이론을 이용하면 유전자 알고리즘의 수렴특성들을 관찰할 수 있는 가능한 접근방법이 될 수 있다.

모집단에 속한 개체 각각에 대하여 국부 최적해에서 전역 최적해로 수렴하는 과정을 검토해 보기로 한다. 최적해의 비트배열에서 적어도 하나 이상의 비트가 변경(flip)될 확률 P_F 는 식 (6)과 같다.

$$P_F = 1 - (1 - P_m)^L \quad (6)$$

모든 N 개의 개체 중에서 k 개의 최적해에서 비트 배열이 변경될 확률은

$$P_F^k (1 - P_F)^{N-k} \quad (7)$$

이다. 또한 모집단 내에서 모든 비트배열이 변경되지 않을 확률은

$$(1 - P_m)^{N \cdot L} = (1 - P_F)^N \quad (8)$$

이 되고, 돌연변이에 의하여 개체 n_i 가 최적해 n_o 로 변환될 확률 P_{n_o} 는

$$P_{n_o} = P_m^{h(n_i, n_o)} (1 - P_m)^{L-h(n_i, n_o)} \quad (9)$$

가 되고, $P_{n_o} > 0$ 이 항상 성립해야 한다.

국부 최적해에서의 스키마 길이 $\delta(H)$ 는 0으로 수렴되고 스키마 차수 $o(H)$ 는 개체의 비트배열 길이 L 로 수렴된다. 따라서

$$1 - P_c \frac{\delta(H)}{L-1} \cong 1 \quad (10)$$

$$1 - o(H)P_m \cong 1 - LP_m \quad (11)$$

(10)식과 (11)식을 (5)식에 적용하면

$$P_m < \frac{1}{L} \quad (12)$$

을 만족하도록 돌연변이 확률이 선정되어야 한다. 또한, 국부 최적해에서 전역 최적해로 천이가 이루어지기 위해서는 적어도 하나 이상의 개체가 돌연변이에 의하여 다른 해로의 천이가 이루어져 국부 최적해로부터 빠져 나와야 한다. 즉, 모집단을 구성하고 있는 개체가 10개이고, 각 개체가 100개의 비트배열로 구성되어 있다면 돌연변이 확률이 $1/(10 \times 100)$ 이상이 되면 적어도 하나의 개체는 국부 최적해로부터 새로운 해로 천이가 이루어진다. 즉,

$$P_m > \frac{1}{NL} \quad (13)$$

(12)식과 (13)식으로부터 국부 최적해로부터 천이가 일어나고, 계속하여 전역 최적해로 수렴하려면 돌연변이 확률 P_m 은 다음의 관계를 만족해야 한다.

$$\frac{1}{NL} < P_m < \frac{1}{L} \quad (14)$$

결과적으로 모집단의 수가 작을수록 사용 가능한 돌연변이 확률의 범위가 좁아지고, 모집단의 수가 클수록 사용 가능한 돌연변이 확률의 범위가 넓어진다.

이상의 결과로부터 돌연변이 확률을 가변적으로 적용하는 것이 효과적이라는 결과를 도출할 수 있다. 즉, 어느 해로 수렴될 때까지는 낮은 돌연변이 확률을 사용하여 수렴속도를 좋게 하고, 수렴이 된 이후에는 돌연변이 확률을 높게 하여 국부 최적해에서 빠져 나올 수 있도록 하는 것이다. 수렴이 이루어진 이후에는 평균 해밍거리가 0에 가까워지므로 최소화 하나 이상의 개체가 돌연변이에 의하여 적합도가 더 높은 개체가 선택되어야 전역 최적해로의 천이를 보장할 수 있게 된다.

평균 해밍거리는 모집단의 각 개체가 부모세대로부터 자손세대로 진화하면서 비트배열 중 평균적으로 역전된 비트의 개수와 같다. 이 경우에 변화되지 않은 비트배열은 부모세대와 자손세대의 비트배열 중 평균적으로 동일한 비트의 수가 되므로 $L - Ah(t)$ 는 스키마의 차수 $o(H)$ 와 같게 되고, 변화된 비트배열의 수는 평균 해밍거리 $Ah(t)$ 를 나타내며 스키마 거리 $\delta(H)$ 와 같게 된다.

만약 모집단이 돌연변이에 의하여 항상 일정한 평균 해밍거리를 가지도록 할 수 있다면, 평균 해밍거리는 (5)식에서 스키마 거리 $\delta(H)$ 와 유사하게 작용을 하게 되고, 비트배열의 길이 L 과 평균 해밍거리 $Ah(t)$ 의 차이는 스키마 차수 $o(H)$ 와 유사한 작용을 하게 된다. 즉, t 세대에서 교차와 돌연변이에 의한 영향을 스키마 거리와 스키마 차수 대신에 평균 해밍거리에 의한 표현으로 다음의 관계를 유도할 수 있다.

$$1 - P_c \frac{\delta(H,t)}{L-1} \cong 1 - P_c \frac{Ah(t)}{L-1} \quad (15)$$

$$1 - o(H,t)P_m \cong 1 - (L - Ah(t))P_m \quad (16)$$

유전자 알고리즘이 수렴하려면 (15)식과 (16)식은 항상 0보다 커야 한다. 여기서 교차확률 P_c 와 돌연변이 확률 P_m 은 평균 해밍거리 $Ah(t)$ 에 영향을 받게 되는 것을 알 수 있다. 즉,

$$P_c < \frac{L-1}{Ah(t)} \quad (17)$$

$$P_m < \frac{1}{L - Ah(t)} \quad (18)$$

여기서 P_c 는 0과 1사이의 값을 사용하기 때문에 (17)식은 수렴하는 경우 항상 성립한다. P_m 은 (14)식과 (18)식을 동시에 만족해야 하므로 P_m 은 (18)식에 의하여 제한을 받고, (14)식과 (18)식으로부터 선정된 P_m 은 (6)-(9)식에서 동일하게 적용되게 된다.

3-1. 모집단의 크기와 선택 연산자

모집단내에 있는 염색체의 수와 염색체의 길이(encoding length)는 수렴속도와 수렴범위(bound)에 영향을 미친다. 모집단의 크기가 충분히 크면 유전자 알고리즘은 전역 최적해로 수렴할 확률이 높아지고 알고리즘의 성능 또한 좋아진다. 이는 돌연변이 확률이 작더라도 모집단이 커지면 모집단의 다양성이 유지되기 때문이다.

엘리트 선택방법은 자손의 일부를 부모세대에서 적합도가 가장 높은 일부로 대체하여 적합도가 높은 개체가 교차나 돌연변이 없이 다음 세대로 직접 천이하는 GA방법으로서 적합도가 높은 개체가 파괴되지 않기 때문에 세대사이에 평균 적합도를 지속적으로 향상시킬 수 있게 된다. 따라서 엘리트 염색체는 진화 연산자에 의한 파괴가 일어나지 않기 때문에 전역 최적해로 수렴할 확률이 높아진다.

3-2. 교차 연산자

교차 연산자만을 사용하는 경우에는 알고리즘은 빠르게 수렴하며 국부 최적해에 수렴할 확률이 높아진다. 이 경우 평균 해밍거리가 0으로 수렴되고 더 이상의 진화는 이루어지지 않는다. 평균 해밍거리에 기반을 두고 가변의 교차확률을 다음 식과 같이 사용하는 방법을 제안하였다.

$$P_c(t+1) = P_c \cdot f_c(Ah(t)) \quad (19)$$

$f_c(Ah(t))$ 는 t 세대에서의 평균 해밍거리를 기반

으로 교차확률을 가변시키는 함수이다.

3-3. 돌연변이 연산자

본 논문에서는 탐색의 진행에 따라 파라미터의 변화를 감지하여 이에 적합한 돌연변이 확률로 변화시키는 가변 돌연변이 방법을 사용하였다. 돌연변이 연산자는 탐색영역을 넓게 하여 국부 최대값 또는 최소값에 수렴되는 확률을 감소시키는 배경(background) 연산자로서의 역할이 지배적이다. 국부 최대값 또는 최소값에 조기 수렴하는 것을 방지하기 위해서는 GA의 초기단계에는 교차 연산자가 탐색에 주로 작용을 하고, 돌연변이의 영향은 최소로 하여야 한다. 세대가 진행됨에 따라 교차 연산자보다 돌연변이 확률의 영향이 크게 한다. 본 논문에서는 이와 같은 개념을 구현하기 위하여 평균 해밍거리에 기반을 두고 가변의 돌연변이 확률을 다음 식과 같이 사용하는 방법을 제안하였다.

$$P_m(t+1) = P_m \cdot f_m(Ah(t)) \quad (20)$$

$f_m(Ah(t))$ 는 t 세대에서의 평균 해밍거리를 기반으로 돌연변이 확률을 가변시키는 함수이다. 결과적으로 유전자의 다양성이 증가되도록 하여 초기에 국부 최적해에 수렴이 되었다 하더라도 국부 최적해에서 천이하여 전역 최적해로 수렴될 수 있도록 하였다.

IV. 모의실험

모의실험에서는 앞에서 제안한 가변의 엘리트 개체의 천이비율, 가변의 교차 확률, 가변의 돌연변이 확률 등 유전자 알고리즘 파라미터의 유용성을 증명하였다. 탐색의 대상으로는 함수의 최대값을 찾는 문제로서 다음과 같이 침두값이 다수 존재하고 복잡도가 높은 목적함수를 사용하였다.

$$f_{object}(x) = x + 10\sin(5x) + 7\cos(4x), \quad (21)$$

$$0 \leq x \leq 6.$$

일반적으로 많이 사용되는 유전자 알고리즘에 적용해 보기 위하여 선택방법으로는 stochastic universal 선택방법을 사용하였으며, 진화 연산자로는 one point 교차 연산자와 bit inversion 형태의 돌연변이 연산자가 사용되었다. 염색체들의 수 N , 모집단내의 염색체의 비트길이 L 는 수렴특성을 확인하기 위하여 임의로 선정하였으며, (13)식, (17)식 및 (18)식의 수렴범위를 검증하기 위하여 교차 확률 P_c 와 돌연변이 확률 P_m 등은 평균 해밍거리에 따라 조정이 가능하도록 하였다.

예를 들면, P_c 는 평균 해밍거리에 따라 0.6과 0.9 사이에서 변화되도록 설정하였고, P_m 은 $N=20, L=100$ 이면 $0.0005 < P_m < 0.01$, $N=10, L=100$ 이면 $0.001 < P_m < 0.01$, $N=10, L=50$ 이면 $0.002 < P_m < 0.02$, $N=10, L=25$ 이면 $0.004 < P_m < 0.04$ 에서 전역 최적해로의 수렴특성이 나타나는 것을 확인할 수 있다.

그림 1은 적절한 기준 P_m 의 선정으로 전역 최적해로 수렴되는 경우이며 모집단의 모든 개체가 최대값의 위치에 수렴되어 있는 것을 보여 주고 있으나, 그림 2는 P_m 의 기준 값을 앞에서의 수렴범위보다 높게 설정한 경우로서 모든 개체가 수렴하지 않고 서로 다른 위치에 존재하고 적합도 값이 세대에 따라 변화가 심한 것을 확인할 수 있다. 그림 3은 P_m 의 기준 값을 수렴범위보다 낮게 설정한 경우로서 진화특성은 유지하고 있으나 전역 최적해로의 수렴이 보장되지 않는다는 것을 확인할 수 있다.

또한 고정 돌연변이 확률을 사용하는 경우와 평균 해밍거리에 의한 가변 돌연변이 확률을 사용하는 경우를 비교하여 실험을 수행해본 결과 가변 돌연변이 확률을 이용하는 방법이 전역 최적해로의 수렴에 이상적임이 확인되었다. 교차확률 P_c 와 돌연변이 확률 P_m 이 큰 값을 가지면 (5)식의 값이 0 근처의 아주 작은 값을 가지게 되어, (4)식의 선택압이 아주 큰 값을 가지더라도 알고리즘이 전역 최적해로 수렴할 확률은 매우 작아진다.

실험결과를 종합해 보면 교차 연산자는 수렴특성에 미치는 영향이 크지 않고, 배경(background) 연산자로서의 돌연변이 확률은 수렴특성에 중요한 역할을 하고 있다는 것이 확인되었다. 적절한 확률의 선

택이 중요하고 국부 최적해가 아닌 전역 최적해로 수렴을 보장하기 위해서는 유전자가 일정한 다양성 (diversity)을 유지하는 것이 매우 중요하다는 결과를 얻을 수 있었고, 이는 평균 해밍거리로서 확인이 가능하였다.

V. 결 론

유전자 알고리즘에서 사용하고 있는 파라미터들은 교차 연산자, 돌연변이 연산자, 엘리트 개체 천이 비율, 모집단의 크기 등이 있다. 특히 돌연변이 연산자는 조기 수렴을 방지하고 전역 최적해로 수렴할 수 있도록 유전자의 다양성을 제공하는 매우 중요한 파라미터로서 적절한 돌연변이 확률을 정하는 방법을 제안하였다.

평균 해밍거리를 기반으로 가변의 유전자 알고리즘 파라미터를 사용했을 때 국부 최적해가 아닌 전역 최적해로 수렴할 확률이 높아지고, 그 중에서도 본 논문에서 제안한 가변의 돌연변이 확률에 의하여 유전자의 다양성을 유지하는 것이 매우 중요하다는 결과를 얻을 수 있었다. 즉, 모집단이 적당한 다양성이 유지될 때 전역 최적해에 도달할 확률이 현저하게 높아진다는 것을 확인할 수 있었다.

본 논문에서 분석하고 실험에 의하여 확인된 돌연변이 확률을 정하는 방법은 유전자 알고리즘을 응용한 시스템에 유용하게 적용할 수 있다. 앞으로는 제안된 유전자 알고리즘을 이용하여 이동로봇의 경로 제어, 항공관제 시스템의 운영 최적화 등에 응용하는 방법을 연구할 계획이다.

VI. 참 고 문 헌

[1] Hong K.. Lee, and Gordon K. Lee, "Convergence Properties of Genetic Algorithms", *Proc. of the ISCA*, pp.72-176, Mar. 2004.
 [2] D.E. Goldberg, and K. Sastry, "A Practical Schema Theorem for Genetic Algorithm Design and Tuning", *GECCO*, San Fransisco, 2001.
 [3] G. Rudolph, "Convergence Analysis of Canonical Genetic Algorithms", *IEEE Trans. on Neural Networks*, special issue on Evolutionary Computation, vol. 5, no. 1, pages 96-101, Jan. 1994.
 [4] G. Rudolph, "Self Adaptive Mutation May Lead to Premature Convergence", *IEEE Trans. Evolutionary Computation*, vol. 5, no. 4, pp. 410-414, Aug, 2001.

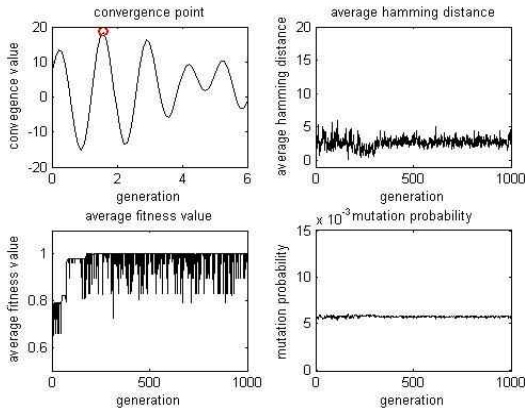


그림 1. 전역 최적해에 수렴하는 경우
 Fig 1. Convergence to global optimum.

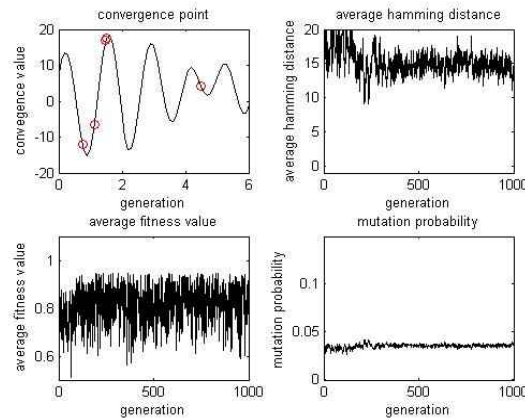


그림 2. 발산의 경우
 Fig 2. Divergence.

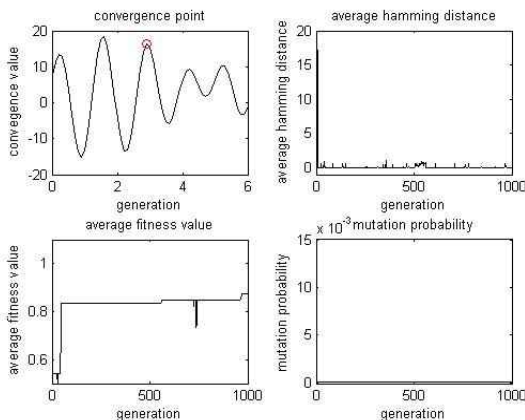


그림 3. 국부 최적해에 수렴하는 경우
 Fig 3. Convergence to local optimum.

- [5] K.S. Leung, Q.H.Duan, Z.B. Xu, and C.K. Wong, "A New Model of Simulated Evolutionary Computation-Convergence Analysis and Specifications", *IEEE Trans. Evolutionary Computation*, vol. 5, no.1, pp. 3-16, Feb. 2001.
- [6] B. Chakraborty, and P. Chaudhuri, "On the Use of Genetic Algorithm with Elitism in Robust and Nonparametric Multivariate Analysis", *Austrian Journal of Statistics*, vol. 32, no. 1&2, pp 13-27, 2003.
- [7] L. Altenberg, "Fitness Distance Correlation Analysis: An Instructive Counter Example", *ICGA97*, Morgan Kaufmann, San Fransisco, 1997.
- [8] C.W. Ahn, and R.S. Ramakrishna, "Elitism-Based Compact Genetic Algorithms", *IEEE Trans. Evolutionary Computation*, vol. 7, no. 4, pp.367-385 Aug. 2003.
- [9] M. Pelikan, D.E. Goldberg, and E. Cantu-Paz, "Bayesian Optimization Algorithm, Population Sizing, and Time to Convergence", *Genetic and Evolutionary Computation Conference*, Las Vegas, NV, July 8-12, 2000.
- [10] Jin M. Won, and Jin S. Lee, "Premature Convergence Avoidance of Self-adaptive Evolution Strategy", *The 5th International Conference on Simulated Evolution. And Learning*, Busan, Korea, Oct. 2004.
- [11] David Metzgar, and Christopher Wills, "Evidence for the Adaptive Evolution Minireview of Mutation Rates", *Cell*, vol. 101, pp. 581-584, June 9, 2000.
- [12] M. Srinivas, and L. M. Patnaik, "Adaptive Probabilities of Crossover and Mutation in Genetic Algorithms", *IEEE Trans., System, Man and Cybernetics*, vol. 24, no. 4, April 1994.
- [13] M. Affenzeller, and S. Wagner, "SASEGASA: An Evolutionary Algorithm for Retarding Premature Convergence by Self-Adaptive Selection Pressure Steering", *Computational Methods in Neural Modelling*, Springer LNCS, vol. 2686, pp. 438-445, 2003.
- [14] Damion Gastelum, and Hong K. Lee, and Gordon K. Lee, et al, "The Development of a Testbed for Evolutionary Learning Algorithms For Mobile Robotic Colonies", *2004 International Symposium on Collaborative Technologies and Systems*, pp. 212-217, Jan. 2004.
- [15] S. Nolfi, and D. Floreano, "Evolutionary Robotics", MIT Press, 2000.
- [16] R. Pfeifer, and C. Scheier, "Understanding Intelligence", MIT Press, 1999.
- [17] V. H. L. Cheng, L. S. Crawford, and P. K. Menon, "Air Traffic Control Using Genetic Search Techniques", *1999 IEEE International Conference on Control Applications*, Hawaii, HA, August 22-27.
- [18] Yanrong Hu and Simon X. Yang, "A Knowledge Based Genetic Algorithm for Path Planning of a Mobile Robot", *Proceedings of the 2004 IEEE International Conference on Robotics & Automation*, New Orleans, LA April 2004.

이 홍 규 (李弘珪)



1977년 : 서울대학교 전자공학과 공학사

1979년 : 서울대학교 전자공학과 공학석사

1989년 : 서울대학교 전자공학과 공학박사

1979년 3월 ~ 1992년 2월 : 국방과학연구소 선임연구원

1994년 3월 ~ 1995년 2월 : 독일 뮌헨공대 방문교수

2003년 1월 ~ 2004년 1월 : 샌디에고 주립대 방문교수

1992년 3월 ~ 현재 : 한국기술교육대학교 교수

관심분야 : 진화로봇, 유전자 알고리즘, 로봇틱스, 전자전