단순 유전 알고리즘에서 단속평형의 지속시간에 대한 연구

(The Duration of Punctuated Equilibria in Simple Genetic Algorithms)

오 상 엽 †

(Sangyeop Oh)

요 약 유전 알고리즘에서 개체군은 지역 최적치에 빠질 수 있지만 긴 지속시간이 지난 후에는 여기에서 빠져나올 수 있으며 이러한 현상을 단속평형 (punctuated equilibrium)이라고 한다. 자연계나 컴퓨터 생태계 (computational ecosystems) 에서 관찰되는 단속평형은 확산 방정식 (diffusion equation)으로 잘 설명된다. 본 연구에서는 단순 유전 알고리즘을 이론적으로 분석하여 개체군의 움직임이 확산 방정식으로 표현될 수 있다는 것을 보인다. 또한 적합도 (fitness) 함수를 단위화 (unitation) 함수로 국한하면 이 분석을 더 구체화하여 유전 알고리즘의 주요 변수들이 이 방정식에 나타나도록 할 수 있다. 이 경우 확산 방정식에 대한 이론적 결과를 이용하면 지역 최적치에서 빠져 나오기까지의 지속시간이 개체군의 크기, 1/(돌연변이 확률), 그리고 지역 최적치의 깊이에 대해 지수적으로 증가한다는 것을 알 수 있다. 이러한 이론적 결과는 이중안정 지형 (bistable landscapes)에서의 시뮬레이션 결과와 일치한다.

키워드 : 유전 알고리즘, 단속평형

Abstract For genetic algorithms, the population may get stuck in a local optimum. The population can escape from this after a long duration. This phenomenon is called punctuated equilibrium. The punctuated equilibria observed in nature and computational ecosystems are known to be well described by diffusion equations. In this paper, simple genetic algorithms are theoretically analyzed to show that they can also be described by a diffusion equation. When fitness is the function of unitation, this analysis can be further refined to make the parameters of genetic algorithms appear in this equation. Using theoretical results on the diffusion equation, the duration of equilibrium is shown to be exponential of such parameters as population size, 1/(mutation probability), and potential barrier. This is corroborated by simulation results for bistable potential landscapes with one local optimum and one global optimum.

Key words : genetic algorithms, punctuated equilibria

1. 서 론

유전 알고리즘(Genetic Algorithms: GAs)은 자연선택 과정을 모델로 하여 만들어진 최적화 방법이다[1]. 홀랜드에 의해 원래의 아이디어가 제안된 이래[2], 여러 가지 종류의 GA들이 다양한 분야에서 성공적으로 적용돼 왔다. 이 중에서 단순 유전 알고리즘(Simple Genetic Algorithms: SGAs)은 룰렛휠 선택(roulette wheel selection), 단순 돌연변이(simple mutation), 단순 교차(simple crossover) 등 유전 알고리즘의 핵심적인 특성을 가지고 있는 가장 단순한 알고리즘이다.

다른 많은 휴리스틱 최적화 방법과 같이 단순 유전 알고리즘은 지역 최적치에 빠지는 문제를 가지고 있다. 다행히 유전 알고리즘에는 지역 최적치에서 빠져나올 수 있게 해 주는 두 가지 특성이 있는데, 그 중 첫 번째, 미시적 효과(microscopic effect)는 개체(individual)가 돌연변이(mutation)나 교차(crossover)를 통해서 탈출하는 경우에 관찰된다. 이것은 개체군의 분산에 관련된 특성인데, 이와 달리 두 번째 특성인 확률적 요동(probabilistic fluctuation)은 개체군의 평균치의 요동을 가리킨다.

유전 알고리즘의 적합도 함수(fitness function)는 개체군의 크기가 무한한 경우의 움직임에 따라 두 개의

† 정 회 원 : 금오공과대학교 컴퓨터공학부 교수
syoh@kumoh.ac.kr
논문접수 : 2005년 3월 31일
심사완료 : 2005년 9월 8일

클래스로 분류할 수 있는데[3,4], 그 중에서도 한 개의 지역 최적치와 한 개의 전역 최적치를 갖는 이중안정(bistable) 적합도 함수로 범위를 좁혀서 생각해 보자. 개체군 상태가 초기에 어떤 값을 갖더라도 결국에는 전역 최적치로 수렴하는 경우의 적합도 함수를 클래스 I이라고 한다. 반대로, 초기 상태에서 개체군의 일정 비율이 이미 전역 최적치의 영역에 있지 않으면 지역 최적치에서 영원히 벗어나지 못하는 경우의 적합도 함수를 클래스 II라고 한다. 그런데 이러한 지역 최적치에서의 평형(equilibrium)은 개체군의 크기가 유한한 경우에 깨질 수 있다. 각 GA 연산자들은 확률적 변수들을 포함하기 때문에, 개체군의 크기가 유한하면 GA 연산의 결과가 매번 달라질 수 있다. 이것이 앞에서 언급한 확률적 요동을 유도하고 결과적으로 단속평형을 일으킨다.

단속평형은 준안정상태(meta-stable state)에 있는 시스템이 평형상태에 오랫동안 머물다가 이웃에 있는 더 안정된 상태로 갑작스럽게 전이해 가는 현상이다. 전역 최적치는 안정상태에 대응되고 그에 이웃한 지역 최적치는 준안정상태에 대응된다. 그런데 여기에서 최적치는 적합도 함수의 값으로서 선택과정만 고려한 것에 비해 안정상태는 돌연변이 등의 다른 영향도 모두 고려한 것이므로 최적치와 안정상태가 정확하게 일치하는 것은 아니고, 안정상태는 포텐셜(potential) 함수의 값으로서 표현된다. 단속평형은 컴퓨터 생태계(computational ecosystems: CEs)[5], 신다윈주의 진화 모델[6], GA[7] 등에서 관찰되고 분석되어 왔다. 이 시스템들은 적합도 공간이나 포텐셜 지형(potential landscape) 상에서 개체들의 집단인 개체군(population)이 돌아다닌다는 공통점을 가지고 있다.

만약 개체군을 두 종류의 개체로 나눈다면 시스템 상태는 특정 종류의 개체의 비율 z로 표현된다. 이 경우 CE의 움직임(dynamics)은 다음과 같은 확산 방정식에 따른다.

$$dz(t) = -dtF'(z) + \varepsilon dB(t) \quad (1)$$

여기에서 F는 포텐셜 지형(potential landscape), $B(t)$는 표준 브라운 프로세스(standard Brownian process), ε은 작은 상수다[5]. 이 확산 방정식에 대한 수학적 결과를 이용하면[8], 단속평형 현상이 존재한다는 것과 여러 주요 변수들과 준안정상태의 지속시간 T의 관계에 대해서 설명할 수 있다. 이 지속시간은 개체군의 크기나 준안정상태와 안정상태 사이에 있는 포텐셜 장벽의 높이에 지수적(exponential)으로 증가하는 관계를 갖는다.

신다윈주의 모델에서는, 개체의 특성 x를 고려할 때 z $\boldsymbol{z}$ 를 $\bar{x}$로 대체하고 $(-F)$를 $\bar{x}$의 적합도로 대체하면 개체군 평균치 $\bar{x}$의 움직임이 식 (1)에 따르게 된다. 따라서 신다윈주의 모델은 자연계의 진화과정에서 관찰되는 단속평형과 지수적으로 긴 지속시간을 설명할 수 있다[6].

GA에서는, Vose가 전역 최적치가 아닌 지역 최적치는 안정적이지 않다는 것을 보임으로써 유전적 탐색 과정에서 관찰되는 단속평형을 설명하려고 했다[7]. 하지만 준안정상태의 지속시간과 같은 양적인 성질에 대해서는 고려하지 않았다. 한편 Royal Road 적합도 함수에 대해서 단속평형과 유사한 현상이 관찰되고 분석되기도 했다[9]. 그러나 이것은 Royal Road 적합도 함수가 계단 함수와 같은 모양을 가지고 있어서 적합도 지형이 평평하기 때문에 나타나는 현상으로 적합도 장벽으로 둘러싸여 있는 지역 최적치에서 나타나는 현상과는 다르다. 따라서 여기에서는 평평한 영역을 통과하는 시간이 1/(개체군의 크기)에 비례한다.

한편, 클래스 I의 적합도 함수를 갖는 GA의 움직임에 대한 연구가 있었는데[10,11] 이 경우, 일단 엘리트 개체가 전역 최적치에 생성되면 개체군의 이동이 신속하게 일어난다. 전역 최적치에 첫 번째 엘리트가 생성되기까지 걸리는 세대수를 T_1이라고 하면 클래스 I 적합도 함수에 대해서 $T_1 \simeq T$가 된다. 개체군 크기가 무한하면 엘리트는 전역 최적치에 즉시 생성되고 따라서 T_1은 1이다. 하지만 개체군 크기가 유한하면 T_1은 훨씬 커질 수 있다. 이러한 T_1에 대한 '유한 개체군 효과'는 [11]에서 연구됐으며 이러한 분석은 [12-14]에서 개발된 GA 동역학에 기반을 둔 것이다. 이들의 결과에 따르면 개체군 크기가 커짐에 따라 개체군 분산이 증가하고 따라서 T_1은 감소한다. 이들은 또한, 개체군 내에서의 유사도로서 유전형의 '상관관계(correlations)'를 계산했다. 개체군 크기가 크면 이 상관관계가 작아지고, 교차는 개체군 분산을 크게 만든다[10]. 그러나 이러한 효과들은 단속평형에 대한 자연계의 연구결과들에 반대된다. 이것은 이들이 적합도 함수를 클래스 I으로 한정하여 연구했기 때문인 것으로 보인다.

클래스 II의 적합도 함수를 갖는 GA의 경우에도 개체군 크기가 유한하면 GA 연산의 결과가 매번 달라질 수 있기 때문에 확률적인 요동이 발생하고 평형을 중단시킬 (punctuate) 수 있다. 이 경우 개체군 크기가 작아지면 요동의 크기가 커지고 따라서 T는 작아진다. 이 요동은 주로 개체군 분산으로부터 발생하는 게 아니라 개체군 평균의 앙상블 분산으로부터 발생한다. 따라서 유한 개체군 효과는 클래스 I의 경우와 반대로 나타나게 될 것이다.

일반적으로 GA에서는 두 종류 이상의 유전형(genotype)이 있으므로, 신다윈주의 모델에서와 같이 두 종류의 개체들을 다루기 위해 본 연구에서는 표현형

(phenotype)의 개체군 평균에 촛점을 맞춘다. 먼저 중심극한정리(central limit theorem)[15]를 이용해 SGA의 움직임이 식 (1)과 유사한 식으로 표현된다는 것을 보인다. 적합도를 클래스 II의 단위화 함수로 국한하면 이 식을 더 구체적으로 유도할 수 있다. 그 다음에 식 (1)에 대한 수학적 결과를 적용하면 GA의 단속평형에 대한 특성들을 얻을 수 있다. 단위화 함수는 적합도 함수의 일부분에 불과하지만 분석과 이해가 용이하기 때문에 여러 GA 연구에 이용돼 왔다[4,16]. 이렇게 얻은 이론적 결과를 이중안정 포텐셜 지형에서 얻은 시뮬레이션 결과와 비교한다.

정리하면 다음과 같다. 유전 알고리즘은 자연계의 진화과정을 모델로 하는 것인데도 불구하고, 유전 알고리즘에서 단속평형과 관련된 기존의 연구들은 자연계에서의 연구와 반대되는 결과를 보여주고 있다. 본 연구에서는, 적합도 함수를 적절히 분류하고 특정 종류의 적합도 함수에 대해 새로운 분석 방법을 적용하면 이러한 모순이 생기지 않는다는 것을 보이고자 한다. 또한 시뮬레이션을 통해 이러한 이론적 분석이 잘 적용된다는 것을 확인하고자 한다. 본 논문의 구성은 다음과 같다. 먼저 2절에서는, 분석의 배경으로서 식 (1)을 더 자세히 살펴본다. 그 다음 3절에서 SGA의 움직임을 분석한다. 4절에서는 이중안정 적합도 지형을 소개하고 시뮬레이션 결과를 보여준다. 5절에서는 교차가 결과에 미치는 영향에 대해 논한다. 마지막으로 결론과 향후 연구에 대해서는 6절에서 다룬다.

2. 관련연구

2.1 확산 과정(diffusion processes)

식 (1)을 만족시키는 1-차원 확산 과정 $z(t)$를 고려할 때 이중안정 포텐셜 지형 F가 다음의 조건들을 만족시킨다고 하자.

1. F는 $-\infty<z<\infty$에서 정의된 미분가능한 함수다.
2. F가 $(-\infty,z_1]\cap[z_2,z_3]$에서는 순감소(strictly decreasing)하고 $[z_1,z_2]\cup[z_3,\infty)$에서는 순증가(strictly increasing)하는 $z_1<z_2<z_3$가 있다.
3. $F(z_1)>F(z_3)$

그러면 z_1이 준안정상태이고 z_3가 안정상태, 그리고 그 사이 z_2에 장벽이 있게 된다. 여기에서 장벽의 높이를 $D=F(z_2)-F(z_2)$라고 한다. 결국, 두 개의 골짜기가 있는 능선 위에서 공이 굴러다니는 양상이 된다.

이러한 이중안정 시스템의 움직임에 대한 수학적 연구에 따르면[8], 시스템의 초기상태가 z_1 근처인 경우에 단속평형이 관찰되고, 준안정상태의 지속시간을 T라고 하면 $\varepsilon\to 0$에 따라서 다음이 성립한다.

$$T \propto \exp\left(\frac{2D}{\varepsilon^2}\right) \tag{2}$$

또한, 이러한 전이는 $F(z_0)$, $F(z_4)>F(z_2)$를 만족시키는 $z_0<z_1$와 $z_4>z_3$가 존재하는 경우에 $t\to\infty$에 따라서 시스템이 안정상태에 계속 머물게 되기 때문에 일방향적(unidirectional)이라고 한다.

본 연구의 목표는 SGA의 움직임이 식 (1)에 의해 표현될 때 ε이 SGA 변수들과 어떤 관계를 갖는지 보이는 것이다. 그러면 식 (2)를 통해 지속시간 T와 SGA 변수들의 관계를 얻을 수 있다.

2.2 컴퓨터 생태계(Computational Ecosystems)

CE의 시스템 상태와 상태 전이 규칙의 정의는 GA와 유사하다[17]. CE에서는 한 개체군이 N 개의 에이전트로 구성되고 각 에이전트는 높은 보수를 얻기 위해 R 개의 전략 중에서 한 개를 선택한다. 보수 함수(payment function) f는 선택된 전략과 개체군의 상태에 따라 보수를 결정한다. 개체군의 상태는 벡터 $\mathbf{n}=(n_1,n_2,...,n_R)$ 또는 $\mathbf{r}=(r_1,r_2,...r_R)$로 표현되는데 여기에서 n_i는 전략 i를 사용하는 에이전트의 수이고 $\mathbf{r}=\mathbf{n}/N$이다. 각 에이전트는 전략을 바꿀 기회를 단위 시간에 α번 가지는데 어느 전략으로 바꿀지는 전략 i가 최선의 전략으로 간주될 확률 $\rho_i(\mathbf{r})$에 의해 결정된다.

시스템 상태는 개체군 상태로 표현된다. $P(\mathbf{n},t)$를 시각 t에 개체군 상태가 n일 확률이라고 하면 특정 t에서의 함수 P는 시각 t에서 가능한 모든 시스템 상태를 포함하는 특정한 앙상블 (ensemble)을 표현한다. 오직 한 번의 전략 변화가 있을 수 있을 정도로 짧은 시간 간격 Δt를 고려하면, $\mathbf{n}$에 아무런 변화가 없거나 한 에이전트가 전략 j에서 i로 바꿨을 가능성이 있다. 그러면 확률의 변화는 다음과 같다.

$$\frac{P(\mathbf{n},t+\Delta t)-P(\mathbf{n},t)}{\alpha\Delta t} = -P(\mathbf{n},t)\sum_{i\neq j} n_j\rho_i + \sum_{i\neq j} P(\mathbf{n}^{[j,i]},t)(n_j+1)\rho_i^{[j,i]} \tag{3}$$

여기에서 $\mathbf{n}^{[j,i]}$는 $n_j^{[j,i]}=n_j+1$, $n_i^{[j,i]}=n_i-1$를 만족시키고 다른 원소들은 $\mathbf{n}$과 동일한 벡터를 가리킨다. 또한, ρ_i와 $\rho_i^{[j,i]}$는 각각 n과 $\mathbf{n}^{[j,i]}$에서 얻어지는 값이다.

비율 r_i의 앙상블 평균 $\langle r_i\rangle$는 다음과 같이 계산된다: $\langle r_i\rangle=\sum_{\mathbf{n}} r_i(\mathbf{n},t)P(\mathbf{n},t)$. 식 (3)에 $\Delta t\to 0$를 적용한 다음 이 계산식을 이용하면 $\langle r_i\rangle$의 미분계수(derivative)는 다음과 같이 얻어진다.

$$\frac{d\langle r_i\rangle}{\alpha dt} = \langle\rho_i\rangle - \langle r_i\rangle \tag{4}$$

결국, 식 (4)는 비율 r_i가 평균적으로 ρ_i에 접근해 간

다는 것을 의미한다.

2.3 단순 유전 알고리즘

SGA는 N 개체로 구성되는 개체군을 다룬다. 각 개체는 L 비트의 이진끈(binary string)이고 각 비트는 0이나 1의 값 중 하나를 가진다. 따라서 $R=2^L$ 개의 유전형이 가능하다. 유전형 i에 대한 표현형을 x_i라고 하고 적합도 f는 표현형 영역에서의 함수라고 하자. 개체군 상태는 $\mathbf{n}=(n_1,n_2,...,n_R)$ 또는 $\mathbf{r}=(r_1,r_2,...r_R)$로 표현되는데 여기에서 n_i는 유전형 i를 갖는 개체들의 수이고 $\mathbf{r}=\mathbf{n}/N$이다.

다음 세대의 개체군은 룰렛휠 선택, 단순 돌연변이, 단순 교차의 SGA 연산자들을 현 세대의 개체군에 적용해 얻는다[1]. 현 개체군에서 룰렛휠 선택으로 두 개의 개체를 얻은 다음 단순 돌연변이는 두 개체의 각 비트를 p_m의 확률로 토글(toggle)시킨다. 그 다음에 단순 교차가 이 두 개체에 p_c의 확률로 적용된다. 단순 교차에서는 각 개체를 일정 지점에서 두 개의 부분끈(substring)으로 나누고 두 번째 부분끈을 서로 교환해서 다시 붙인다. 이 자르는 지점은 가능한 L 지점들 중에서 균등 확률에 따라(uniform-randomly) 선택된다. 교차 이후에는, 만들어진 두 개의 자식 개체 중에서 한 개를 무작위로 선택하여 다음 세대 개체군에 집어넣는다. 이 과정을 N 번 반복하면 세대차(generation gap) 1의 새로운 세대를 얻게 된다[1].

한편, 본 연구의 분석과정에서는 개체가 한 개의 특성만 가지고 있는 단순한 경우, 즉, 표현형이 한 개의 스칼라 값을 나타내는 경우를 다룬다. 표현형을 부호화하는 방법에는 여러 가지가 있는데, 유전형 i에서 1 값을 갖는 비트의 수를 l_i라고 할 때 $x_i=l_i/L$이면 적합도 함수를 단위화 함수의 일종이라고 한다[16].

2.4 유전 알고리즘에서의 단속평형

GA에서 단속평형에 대한 분석은 Vose에 의해 처음 이루어졌다[7]. Vose는 행렬방정식을 이용해 GA에서 전역 최적치가 아닌 지역 최적치는 안정적이지 않다는 것을 보임으로써 GA가 결국 전역 최적치에 도달할 수 있다는 것을 보였다. 그런데 이 과정에서 단속평형, 즉, 준안정상태로부터의 탈출이 존재한다는 것을 이용했지만 이 준안정상태의 지속시간과 같은 양적인 성질에 대해서는 고려하지 않았다. 한편 Royal Road 적합도 함수에 대해서 단속평형과 유사한 현상이 관찰되고 분석되기도 했다[9]. 그러나 이것은 Royal Road 적합도 함수가 계단 함수와 같은 모양을 가지고 있어서 적합도 지형이 평평한 부분들이 있기 때문에 나타나는 현상으로 적합도 장벽으로 둘러싸여 있는 지역 최적치에서 나타나는 현상과는 다르다. 이들은 개체군이 평평한 영역을 통과하는 시간이 1/(개체군의 크기)에 비례한다고 주장하고 있으며, 지역 최적치에 빠진 경우는 다루고 있지 않다.

한편, 클래스 I의 적합도 함수를 갖는 GA의 움직임에 대한 연구가 있었는데[10,11] 이 경우, 일단 엘리트 개체가 전역 최적치에 생성되면 이것이 다른 개체들에 영향을 미쳐서 개체군 전체의 이동이 신속하게 일어난다. 전역 최적치에 첫 번째 엘리트가 생성되기까지 걸리는 세대수를 T_1이라고 하면 클래스 I 적합도 함수에 대해서 $T_1 \simeq T$가 된다. 개체군 크기가 무한하면 엘리트는 전역 최적치에 즉시 생성되고 따라서 T_1은 1이다. 하지만 개체군 크기가 유한하면 T_1은 훨씬 커질 수 있다. 이렇게 T_1에 개체군의 크기가 영향을 미치는 것을 '유한 개체군 효과'라고 하는데[11], 이러한 분석은 [12-14]에서 개발된 GA 동역학에 기반을 둔 것이다. 이들의 결과에 따르면 개체군 크기가 커짐에 따라 개체군 분산, 즉, 개체군을 이루는 개체들의 분산이 증가하고 따라서 T_1은 감소한다. 이들은 또한, 개체군 내에서의 유사도로서 유전형의 '상관관계(correlations)'를 계산했다. 개체군 크기가 커지면 이 상관관계가 작아지고, 교차는 개체군 분산을 크게 만든다[10]. 따라서 큰 개체군 크기나 교차가 지역 최적치로부터의 탈출을 촉진하는 것이다. 그러나 이러한 효과들은 단속평형에 대한 자연계의 관찰결과들에 반대된다. 이것은 자연계에서 단속평형이란 대부분 클래스 II의 적합도 함수에서 관찰되기 때문인 것 같다. 하지만, 아직까지 클래스 II의 적합도 함수에서 생기는 단속평형에 대한 연구는 이루어지지 않았다.

3. 분 석

본 연구에서 확률변수 Z의 평균과 분산은 각각 $E(Z)$와 $V(Z)$로 표현한다. 하지만 개체 특성의 개체군에서의 평균과 개체군 특성의 앙상블에서의 평균을 구분하기 위한 경우에는 $\overline{Z}$와 $\langle Z\rangle$의 표현을 각각 사용한다. 개체 특성 x의 개체군 평균 $\overline{x}$는 $\overline{x}(\mathbf{r})=\sum_i x_i r_i$로 표현되고 개체군 상태로부터 개체군 평균을 $\overline{X}=\overline{x}(\mathbf{r})$와 같이 얻을 수 있다. 여기에서 r_i는 하나의 확률변수다. 한 개체군에서 개체들은 서로 독립적으로 생성됐으므로 개체군들의 앙상블은 이 X의 앙상블을 개체군 단위로 묶어 놓은 것에 대응되고 정수 k에 대해서 $E(X^k)\simeq\langle\overline{X^k}\rangle$이다. 그런데 $\langle\overline{x^k}(\mathbf{r})\rangle=\overline{x^k}(\langle\mathbf{r}\rangle)$이므로

$$E(X)=\langle\overline{X}\rangle=\overline{x}(\langle\mathbf{r}\rangle) \tag{5}$$

와

$$V(X)=E(X^2)-E^2(X)=s_X^2(\langle\mathbf{r}\rangle) \tag{6}$$

를 얻을 수 있다. 여기에서 $s_X^2(\mathbf{r}) = \overline{x^2}(\mathbf{r}) - \overline{x}^2(\mathbf{r})$는 개체군 상태가 $\mathbf{r}$일 때 개체 특성 X의 개체군 분산이다.

한편, $(\tau+1)$ 번째 세대에서의 X가 서로 독립적인 확률변수 $X_s(\tau)$, $X_m(\tau)$, 그리고 $X_c(\tau)$로 구성된다고 하자. 여기에서 $X(\tau+1) = X_s(\tau) + X_m(\tau) + X_c(\tau)$인데 X_s는 선택(selection)의 결과이고 X_m과 X_c는 각각 돌연변이와 교차에 의한 변화량이라고 하자. 이렇게 하면 SGA의 각 연산자가 어떤 영향을 미치는지 분석하여 전체 결과를 얻을 수 있다.

3.1 확산방정식

시각 t에 개체군 상태가 r일 확률을 $P(\mathbf{r},t)$라고 하면 특정 t에서의 P는 시각 t에 가능한 모든 개체군 상태를 포함하는 개체군들의 앙상블을 표현한다. 확률변수 Z가 개체군의 한 특성을 나타낼 때 그 앙상블 평균은 $\langle Z \rangle = \sum_{\mathbf{r}} Z_{\mathbf{r}} P(\mathbf{r},t)$로 표현된다. 이제 α를 단위 시간당 평균 세대수라고 하자. 그러면 SGA는 시간 Δt 동안 개체군 상태를 평균적으로 $\alpha\Delta t$번 바꾼다. 여기에서 Δt는 $\alpha\Delta t$가 0이나 1이 될 정도로 충분히 작다고 가정한다.

SGA 연산자들은 확률적 사건들을 포함하고 있으므로 이 표현형은 확률변수 X라고 표현할 수 있다. 단순하게 말하자면 개체군 앙상블은 이 X의 앙상블을 개체군 단위로 묶어놓은 것에 대응된다. 세대차가 1이므로 각 자식 개체의 생성은 다른 것들과 독립적이고 따라서 자식 개체들로 구성되는 새로운 세대의 개체군은 X에 대한 크기 N의 표본(sample)이라고 할 수 있다. 중심극한정리를 사용하면 개체군 평균 $\overline{X}$가 근사적으로 평균 $E(X)$와 분산 $V(X)/N$를 갖는 가우스 분포(Gaussian distribution)를 갖는다[15]. 결국, SGA가 시각 t_0에 시작한다면 $\overline{X}$는 다음과 같이 표현된다.

$$\overline{X}(t) = \langle \overline{X}(t) \rangle + \alpha(t-t_0)G(t) \tag{7}$$

여기에서 $G(t)$는 평균이 0이고 분산이 $V(\overline{X}(t)) = V(X)/N$인 가우스 분포를 따르는 확률변수다. 다시 말하면, t_0의 한 개체군으로부터 t의 여러 개체군이 만들어질 수 있는데 이 가능한 개체군들이 앙상블을 구성하는 것이다. $\Delta\overline{X}$는 다음과 같이 얻는다.

$$\overline{X}(t+\Delta t) - \overline{X}(t) = \langle \overline{X}(t+\Delta t) \rangle - \langle \overline{X}(t) \rangle + \alpha\Delta t\hat{G}(t) \tag{8}$$

여기에서 $\Delta t\hat{G}(t) = (t+\Delta t-t_0)G(t+\Delta t) - (t-t_0)G(t)$이다. 가우스 확률변수들을 선형으로 더한 것이므로 $\hat{G}(t)$는 평균 0, 분산 $\hat{V}(\overline{X}(t))$인 가우스 확률변수가 되는데 이 분산은 $V(\overline{X}(t))$와 $V(\overline{X}(t+\Delta t))$ 사이의 값을 갖는다.

그런데, 각각 분산 1이고 서로 독립적인 가우스 확률변수들을 누적해서 더하면 표준 브라운 프로세스(standard Brownian process) $B(t)$를 얻는다고 알려져 있다. $\hat{G}$의 확률적 성질은 SGA 연산자들에 들어있는 확률적 사건들에서 비롯된다. 이러한 확률적 사건들은 연속된 세대들 사이에서 서로 독립적이고 따라서 이 세대들의 $\hat{G}$들은 서로 독립적이다. $\hat{G}/\sqrt{\hat{V}(\overline{X})}$는 평균 0, 분산 1인 가우스 분포를 따르므로 Δt 동안 이들을 누적해서 얻는 B의 변화량은 $\Delta B = (\alpha\Delta t)\hat{G}/\sqrt{\hat{V}(\overline{X})}$가 된다.

또한, $\langle \Delta\overline{X}(t) \rangle = \langle \overline{X}(t+\Delta t) \rangle - \langle \overline{X}(t) \rangle$의 앙상블 평균 부분은 $a(\overline{X}(t_0)) = \langle \Delta\overline{X}(t_0) \rangle / \Delta t$와 같이 정의되는 함수 $a(\overline{X})$를 찾은 다음, $\overline{X}(t_0)$를 $\overline{X}(t)$로 치환하여 구할 수 있다. 이 방법은 다음의 분석 조건이 성립할 때 유효하다고 알려져 있다[18].

1. 적합도 지형이 $\overline{X}$ 상에서 충분히 매끄러워야(smooth) 한다.
2. 연속한 두 세대의 $\overline{X}$ 값의 변이가 충분히 작아야 한다.

이러한 조건이 만족되면 식 (8)은 $\Delta t \to 0$에 따라 다음과 같은 확률론적 미분 방정식이 된다[18].

$$\frac{d\overline{X}(t)}{dt} = a(\overline{X}(t)) + \sqrt{\frac{V(X)}{N}}\frac{dB(t)}{dt} \tag{9}$$

여기에서 $\Delta t \to 0$에 따라 $\hat{V}(\overline{X}) = V(\overline{X}) = V(X)/N$이 되는 것을 이용했다. 포텐셜 함수를 F라고 하고 $F'(\overline{X}) = -a(\overline{X})$의 식에 의해 구하면 최종적으로 다음과 같이 식 (1)과 같은 모양을 얻게 된다.

$$\frac{d\overline{X}}{dt} = -F'(\overline{X}) + \sqrt{\frac{V(X)}{N}}\frac{dB}{dt} \tag{10}$$

한편 각 개체가 개체군에서 특정 위치를 점유하고 있다고 생각하면, 자식 개체가 SGA 연산자들에 의해 생성되는 과정을 현 세대의 한 개체를 교체하면서 개체군에 삽입되는 과정으로 간주할 수 있다. 자식개체의 유전형과 교체되는 개체의 유전형을 각각 i와 j라고 하면 이러한 교체는 CE에서 한 에이전트의 전략이 j에서 i로 바뀌는 것에 대응된다. 또한, 적합도 함수는, 일반적으로 개체군 상태에 독립적이긴 하지만, CE의 보수 함수에 대응된다. ρ_i를 유전형 i를 갖는 자식 개체가 생성될 확률이라고 해석하면 세대차가 $1/N$인 SGA는 CE의 특별한 경우가 된다. 세대차가 1인 경우에도 각 자식개체의

생성이 서로 독립적이므로 $\overline{X}$가 한 세대 동안에 상당히 변하지 않는다면 식 (4)를 그대로 사용할 수 있다. 여기에서 α는 CE에서와 같이 한 개체의 평균 교체수다. 따라서 식 (4)로부터 $\sum_{i=1}^{R} x_i \langle r_i(t)\rangle = \langle \overline{X} \rangle = E(X)$를 이용해서 다음을 구할 수 있다.

$$\frac{d\langle \overline{X}(t)\rangle}{\alpha dt} = \sum_{i=1}^{R} x_i(\langle \rho_i(\mathbf{r}(t))\rangle - \langle r_i(t)\rangle)$$

$$= E(X_s + X_m + X_c) - E(X) \quad (11)$$

그런데 여기에서 돌연변이나 교차에 대한 부분은 표현형의 부호화(coding) 방식이나 적합도 함수의 정의에 따라 크게 달라질 수 있다. 따라서 먼저 룰렛휠 선택만을 고려하면, $\rho_i(\mathbf{r}(t)) = f(x_i)r_i/\overline{f}$이고, 여기에서 $\overline{f} = \sum_{i=1}^{R} f(x_i)r_i$이다. 특정 개체가 유전형 i를 가지는 것은 베르누이 분포 (Bernoulli distribution)의 확률적 사건이고, 세대차가 1이므로 각 개체의 선택 과정은 다른 것들과 독립적이다. 그러면 n_i는 이진 분포를 가지고, 충분히 큰 N에 대해 비율 $r_i=n_i/N$는 분산이 $1/N$에 비례하는 가우스 분포로 근사적으로 표현될 수 있다. 이것은 개체군의 크기 N이 충분히 큰 경우에 r_i는 $<r_i>$ 주위의 좁은 영역에 들어가고 따라서 평균장 근사(mean field approximation)를 적용할 수 있다는 것을 의미한다. 평균장 근사는 각 개체가 각각의 다른 개체들과 전역 변수를 통해 서로 영향을 미치고 개체군의 크기가 큰 경우에 잘 적용된다[19]. 그러면 식 (11)의 선택 과정 부분은 다음과 같게 된다.

$$E(X_s) - E(X) = \sum_{i=1}^{R} x_i \left(\frac{f(x_i)\langle r_i\rangle}{\overline{f}(\langle \mathbf{r}\rangle)} - \langle r_i\rangle \right) = \frac{\overline{(x-\overline{x})f}}{\overline{f}} \quad (12)$$

여기에서 개체군 평균은 $<\mathbf{r}>$에 대해서 구한다. 한편 위의 식 (9)에 대한 조건들이 만족되고 $V(X)$가 충분히 작으면 f는 다음과 같이 선형적으로 확장될 수 있다.

$$f(x) \simeq f(\overline{x}) + (x-\overline{x})f'(\overline{x}) \quad (13)$$

결국, 식 (12)에서 f를 식 (13)으로 치환하면 다음을 얻게 된다.

$$\sum_{i=1}^{R} x_i \left(\frac{f(x_i)\langle r_i\rangle}{\overline{f}(\langle \mathbf{r}\rangle)} - \langle r_i\rangle \right) \simeq s_X^2(\langle \mathbf{r}\rangle) \frac{f'(\overline{x}(\langle \mathbf{r}\rangle))}{f(\overline{x}(\langle \mathbf{r}\rangle))} \quad (14)$$

3.2 단위화 함수

2.1절의 확산 작용에 대한 이론을 식 (10)에 적용하기 위해서는 SGA의 실행 중 $V(X)$가 상수여야 한다. 그러나 상수가 아닌 경우에도, $V(X)$가 일정한 영역 안에 놓여있다면 지속시간 T는 지수적인 일정 영역 안에 놓이게 된다. 이 절에서는 $V(X)$에 촛점을 맞추면서 SGA 연산자들의 영향을 각각 분리해서 살펴보도록 한다. SGA에 대한 수학적 분석은 상당히 복잡하지만 적합도 함수를 단위화 함수로 국한함으로써 아래와 같은 결과를 얻을 수 있다.

한 개체의 표현형은 $x_i = l_i/L$로 정의된다. 먼저 단순 돌연변이를 생각해 보면, 한 비트에 돌연변이를 적용했을 때 한 비트의 변화량 (예를 들어 1⇒0에 대해서는 -1)은 평균 $p_m(1-2l(t)/L)$, 분산 $p_m - p_m^2(1-2l(t)/L)^2$인 확률 분포에 따른다. 한 개체의 돌연변이는 한 비트 돌연변이의 크기 L인 표본이므로 중심극한정리를 적용하면 $p_m \ll 1$이고 $L \gg 1$인 경우에 X_m은 근사적으로 평균이

$$E(X_m) = p_m(1-2l(t)/L) = p_m(1-2E(X_s)) \quad (15)$$

이고 분산이

$$V(X_m) = [p_m - p_m^2(1-2l(t)/L)^2]/L \simeq p_m/L \quad (16)$$

인 가우스 분포를 갖는다.

다음에 교차에 대해 생각해 보자. 단순 교차는 두 부모 개체에 있는 1의 수를 바꾸지 않는다. 따라서 $\overline{X}$ 그리고 더 나아가 식 (5)을 이용하면, $E(X)$는 교차로 인해 변하지 않는다는 것을 알 수 있다. $V(X)$에 대한 영향도 상당히 복잡한 과정을 거치면 계산할 수 있지만 본 연구에서는 이를 이용하지 않는다.

$<\overline{X}>$에 대한 교차의 영향은 0이므로 $<\overline{X}>$에 대한 돌연변이의 영향을 고려하면 식 (11)은 다음과 같게 된다.

$$\frac{d\langle \overline{X}(t)\rangle}{\alpha dt} = (E(X_s) - E(X)) + E(X_m)$$

$$\simeq s_X^2(\langle \mathbf{r}\rangle) \frac{f'(\overline{x}(\langle \mathbf{r}\rangle))}{f(\overline{x}(\langle \mathbf{r}\rangle))} + p_m(1-2\overline{x}(\langle \mathbf{r}\rangle)) \quad (17)$$

여기에서 $p_m \ll 1$의 가정하에 $p_m^2(E(X_s) - E(X))$ 항은 제거되고 $E(X) = \overline{x}(\langle \mathbf{r}\rangle)$로 치환됐다. 이 식의 우변에서 첫 번째 항은 선택에 의해 적합도가 높은 방향으로 개체군이 이동하려는 경향을 나타낸다. 선택에 의해 변하는 양은 분산이 클수록, 적합도 함수의 기울기가 클수록 크다. 두 번째 항은 개체군이 돌연변이에 의해 $\overline{x}=1/2$의 방향으로 이동하려는 경향을 나타낸다.

$t = t_0$라고 하면 $\overline{x}(\langle \mathbf{r}\rangle) = \langle \overline{X} \rangle = \overline{X}$가 되고 다음과 같은 결과를 얻는다.

$$a(\overline{X}(t_0)) = \frac{d\langle \overline{X}(t)\rangle}{dt} = \alpha \left[V(X) \frac{f'(\overline{X}(t_0))}{f(\overline{X}(t_0))} + p_m(1-2\overline{X}(t_0)) \right] \quad (18)$$

여기에서 $s_X^2(\langle \mathbf{r} \rangle)$은 식 (6)을 이용해 $V(X(t_0))$로 치환됐다. 이제 $F'(\overline{X}) = -a(\overline{X})$라고 정의한 다음 $V(X)$가 $\overline{X}$에 독립적이라고 가정하면 F의 모양이 근사적으로 어떻게 될 것인지에 대한 아이디어를 다음과 같이 얻을 수 있다. 하지만 이 가정이 일반적으로 성립하지는 않는다.

$$F(\overline{X}) \sim -\alpha[V(X)\log f(\overline{X}) + p_m(\overline{X} - \overline{X}^2)] \quad (19)$$

한편, 준안정상태 즉, 평형상태에서의 앙상블 분산을 V_e라고 하면 개체군이 평형상태 근처에서 요동치고 있으므로 $V(X) \sim V_e$라고 할 수 있다. 식 (1)과 식 (10)을 적용해 얻은 $\varepsilon = \sqrt{V(X)/N}$을 식 (2)에 치환하면 다음을 얻는다.

$$T \propto \exp\left(\frac{2DN}{V_e}\right) \quad (20)$$

평형상태에서는 $d\langle \overline{X} \rangle = 0$이므로 식 (17)로부터 다음을 구할 수 있다.

$$V_e \simeq p_m(2E_e - 1)\frac{f(E_e)}{f'(E_e)} \quad (21)$$

이 식의 우변에서 p_m에 곱해지는 나머지 부분은 p_m에 독립적이지는 않지만 f가 완만하게 변한다는 가정하에 p_m을 이용한 선형 확장으로 근사될 수 있으며 이때 차수(order) 0인 항이 나오게 된다. 또한 SGA의 각 연산자는 앞에서 보았듯이 N에 무관하므로, 결국 V_e의 특성은 다음과 같다.

1. V_e는 N에 독립적이다.
2. $V_e \propto p_m$.

결국, 지속시간은 다음과 같이 결정된다.

$$T \propto \exp\left(\frac{2DN}{K_m p_m}\right) \quad (22)$$

여기에서 K_m은 δ의 함수이고, δ는 적합도 장벽의 깊이를 나타내는데 4절에서 구체적으로 정의한다. D도 δ의 함수다.

4. 시뮬레이션

4.1 단위화 함수

4.1.1 이중안정 지형

단속평형은 F의 지형이 식 (2)의 조건들을 만족시키는 경우에 관찰된다. 또한 f가 식 (9)의 조건들을 만족시키면 앞 절의 이론적 분석을 적용시킬 수 있다. 이러한 조건들을 고려하여 적합도 함수는 다음과 같이 정의된다.

$$f(x) = \begin{cases} 1 - 3\delta x & \text{if} \quad 0 \le x < 1/3 \\ 1 - 2\delta + 3\delta x & \text{else if} \quad 1/3 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

여기에서 δ는 f 지형에서 장벽의 깊이다. 결국 f는 경사도의 절대값이 동일한 두 개의 선형 경사로 구성된다.

시뮬레이션에서 SGA 변수들의 영역에 대해 다음과 같은 지침을 사용한다. 첫째, f가 연속함수로 근사되고 중심극한정리가 식 (15)와 (16)을 구하기 위해 적용될 수 있도록 L이 충분히 커야 한다. 둘째, 돌연변이 확률 p_m의 첫 번째 차수까지의 근사가 분석에 사용되므로 $p_m \ll 1$을 만족시켜야 한다. 셋째, 3.1절에서 중심극한정리를 적용하기 위해 개체군 크기 N이 충분히 커야 한다. 그림 1은 위의 적합도 함수에 대한 포텐셜 지형을 식 (19)에 의해 구한 것이다.

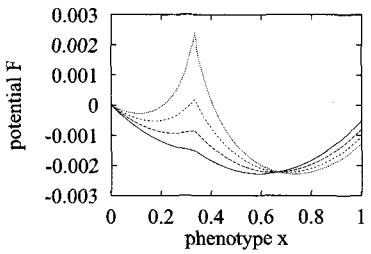

그림 1 다양한 적합도 장벽 깊이에 대해 식 (19)로부터 예상되는, 표현형(phenotype)에 따른 포텐셜 함수들(potential functions). y-축의 단위는 α다. 아래에서부터 위로, 각각 δ=0.3, 0.5, 0.7, 0.9이고 다른 변수들은 $s_X^2 = 0.002$, $p_m = 0.01$이다.

δ가 작을 때에는 포텐셜 지형에 오직 한 개의 최소치가 존재한다는 것을 고려해야 한다. 앞에서의 분석이 적용되기 위해서는 포텐셜 지형이 이중안정적이여야 한다. 즉, 두 개의 최소치가 존재해야 한다. 따라서 시뮬레이션에서는 (0.4,1]의 영역에서 δ 값을 택한다.

4.1.2 단속평형

그림 2는 각각 p_c=0이거나 p_c=1인 SGA의 실행과정에서 관찰되는 전형적인 단속평형을 보여준다. 개체군은 $\overline{x}$=0의 상태에서 출발해서 급속하게 준안정상태로 수렴하고 그 주위에서 요동친다. 긴 지속시간이 지나고 $\overline{x}$는 갑자기 포텐셜 장벽을 넘어서 안정상태 주위에서 요동친다.

한편, 개체군 분산에 대한 유한 개체군 효과와 교차 효과는 그림 3에서 찾을 수 있다. 개체군이 작은 영역에서는 N=32와 N=64의 그래프에서 보이듯이 개체군 크

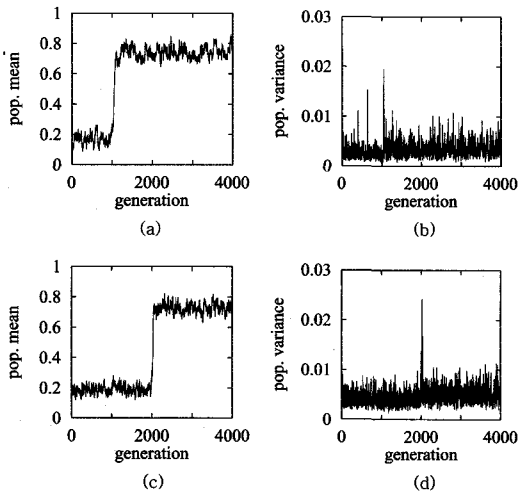

그림 2 세대(generation)에 따라 변하는 표현형의 개체군 평균 (population mean) $\bar{x}$의 그래프에 나타나는 단속평형. 그래프 (a)와 (b)는 L=30, N=40, p_c=0, p_m=0.013, δ=0.7인 경우, 각각 $\bar{x}$와 s_X^2를 나타낸다. 그래프 (c)와 (d)는 L=30, N=40, p_c=1.0, p_m=0.02, δ=0.7인 경우, 각각 $\bar{x}$와 s_X^2를 나타낸다.

기가 증가하면 개체군 분산이 약간 증가한다. 그러나 N=64와 N=320의 그래프를 비교하면 개체군이 큰 영역에서 개체군 크기 증가에 따른 개체군 분산의 증가가 무시할 만큼 작다는 것을 알 수 있다. 교차가 사용되면 개체군 분산이 약간 증가하지만 N=320에 대한 그래프에서 보이듯이 개체군이 큰 경우에는 그 영향은 무시할 만큼 작은 양이다. 더 중요한 변화는 개체군 크기가 증가함에 따라 그래프에 나타나는 요동의 크기가 작아진다는 것이다. 이것은 평형상태에서 $\langle s_X^2 \rangle$의 감소를 의미하고 중심극한정리에 의한 $\langle \bar{X} \rangle$의 감소와 관련돼 있다[15].

4.1.3 준안정상태의 지속시간

몇 가지 중요 변수에 대해, $\bar{x}$=0의 개체군 상태에서 시작해서 전이가 일어나기까지 준안정상태의 지속시간 T를 세대단위로 기록한다. 고려하는 변수들은 개체군 크기 N, 돌연변이 확률 p_m, 적합도 장벽의 깊이 δ의 세 가지다. 그림 4(a), (b) 그리고 (d)는 T가 N과 δ에 대해 급속히 증가하는 함수라는 것을 보여준다. 그리고 그림 4(c)는 T가 p_m에 대해 급속히 감소하는 함수라는 것을 보여준다. 그림 4는 또한 교차가 이 지속시간을 더 길게 만든다는 것을 보여준다.

그림 4(a)와 (b)는 p_c=1에 대한 로그-스케일 그래프

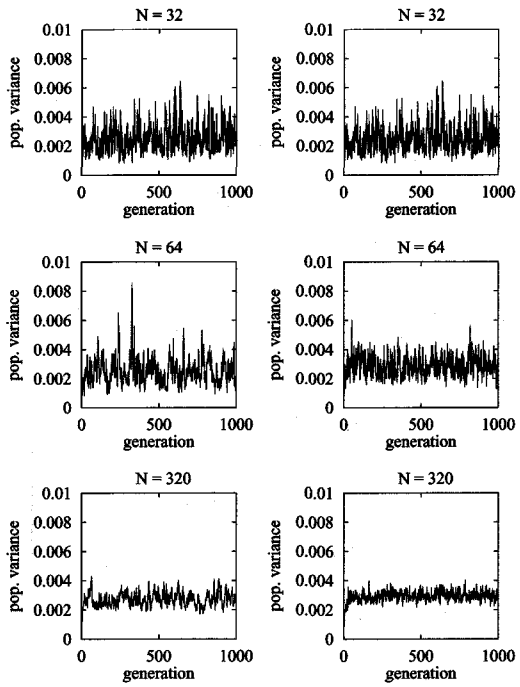

그림 3 다양한 개체군 크기(N)에 대해 평형을 이룬 전형적인 SGA의 개체군 분산(population variance) s_X^2. 기본 변수값들은 L=30, p_m=0.01, δ=0.7. 단, (a)와 (b)에 대해 각각 p_c=0이고 p_c=0.

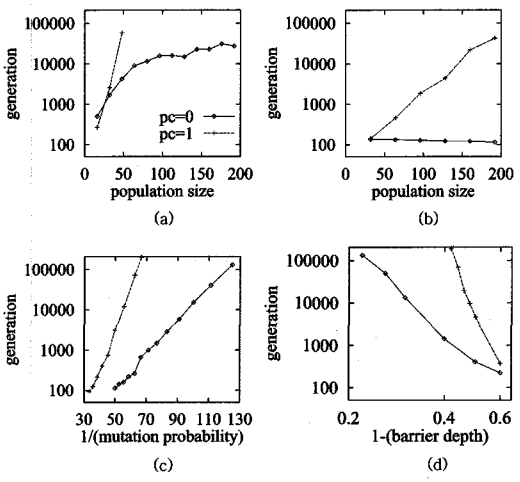

그림 4 변수들에 대한 준안정상태의 지속시간 T. 지속시간은 세대(generation) 단위이고 변수들은 개체군 크기(population size), 돌연변이 확률(mutation probability), 장벽의 깊이(barrier depth)이며 각 점은 100회 실행한 결과의 평균치다. 기본 변수값들은 L=30, N=40, p_m=0.01, δ=0.7이다. 단, (a)와 (b)에서는 각각 p_m=0.012, p_m=0.022이다.

가 선형이므로 T가 개체군 크기 N에 대해 지수적으로 증가한다는 이론적 결과를 뒷받침한다. 하지만 p_c=0인 경우 $N\simeq 60$ 이상의 영역에서 그래프가 아래로 휘어지는 것을 보면 지수적 성질이 약화되는 걸 알 수 있다. 본 연구의 이론적 분석은 개체군이 거시적으로는 해답공간(solution space)에서 굴러다니는 강체구(rigid ball)와 같이 간주될 수 있다는 가정 아래 이루어진다. 이러한 가정은 개체군의 크기가 준안정상태의 인력 웅덩이의 크기에 비교할 만큼 큰 경우에 깨지게 되는데 이 경우 $1024=2^{L/3}$ 정도가 된다. N이 큰 경우, 돌연변이에 의해 더 많은 개체들이 안정상태의 영역에 나타나게 되고 이것이 전이를 촉진시킨다. 이것이 서론에서 언급한 SGA 연산자들의 미시적 효과다. 하지만 p_c=1인 경우에 안정 영역에 나타난 엘리트 개체들은 5.1 절에서 논의되듯이 교차에 의해 즉시 제거되고, 따라서 이 미시적 효과는 상당히 약화된다. 따라서 N이 상당히 크더라도 p_c=1인 경우에는 지수적 성질이 명확히 나타나게 된다. 돌연변이 확률 p_m에 대해서는, 그림 3(c)의 로그-스케일 그래프가 선형이므로 식 (22)에 일치한다. 그러나 p_m이 큰 경우에는 이러한 선형성이 약화된다.

그림 4(d)에 대해서는, 장벽의 깊이 δ가 f 지형의 값이기 때문에 식 (22)가 직접 사용될 수 없다. 포텐셜 장벽은 $D=F(1/3)-F(x_e)$다. δ가 변해도 $V(X)$가 거의 변하지 않는다고 가정하면 T에 대한 δ의 기여는 주로 $F(1/3)$ 항으로부터 나온다. 이 가정은 시뮬레이션을 통해 실험적으로 확인할 수 있다. 따라서 지속시간을 식 (19)로부터 구해보면 δ가 대략 0.5 이상일 때 $T\sim\exp(-K_\delta\log(1-\delta))=(1-\delta)^{-K_\delta}$가 된다. 여기에서 K_δ는 하나의 상수다. 이것은 아주 대략적인 근사지만 그림 4(d)의 로그-로그 그래프가, $(1-\delta)$가 대략 0.5 보다 작은 영역에서, 선형이라는 것과 일치한다. δ가 작은 경우에는 D도 작아져서 안정상태와 준안정상태 사이에 실질적인 장벽이 없어지고 포텐셜 지형이 더 이상 이중 안정성을 갖지 않게 된다.

4.1.4 2-차원 단위화 함수

앞에서 이론적으로 분석한 결과는 적합도 함수가 다차원 상에서 정의되는 경우에도 적용될 수 있는 가능성이 있다[8,20]. 이것을 보이기 위해 2-차원 단위화 함수에 대해 시뮬레이션 결과를 얻었다. 여기에서 적합도 함수는 다음과 같이 정의된다.

$$f(x,y)=\begin{cases}1-\frac{3\delta}{\sqrt{2}}\sqrt{x^2+y^2} & \text{if} \quad 0\le\sqrt{x^2+y^2}<\sqrt{2}/3\\ 1-2\delta+\frac{3\delta}{\sqrt{2}}\sqrt{x^2+y^2} & \text{else if} \quad \sqrt{2}/3\le\sqrt{x^2+y^2}\le\sqrt{2}\\ 0 & \text{otherwise}\end{cases} \tag{24}$$

시뮬레이션 결과는 그림 5에 나오는데 이 그래프들은 1-차원 단위화 함수의 경우와 거의 같은 모양을 갖는다.

4.2 다른 적합도 함수들

4.2.1 비부호 이진 코드(unsigned binary code)

이 절에서는 식 (23)과 동일한 적합도 함수를 갖지만 x가 단위화 코드가 아니라 비부호 이진 코드에 의해 부호화된 경우의 시뮬레이션 결과를 보여준다. 즉, 한 개체는 비부호 이진수를 표현한다.

그림 6에서 시뮬레이션 결과가 나오는데 준안정상태의 지속시간 T가 개체군 크기나 돌연변이 확률에 따라 지수적으로 증가하지 않는다는 걸 알 수 있다. 다만 적합도 장벽의 깊이와 T의 관계는 단위화 함수에 대한 결과와 유사하다.

4.2.2 그레이 코드(Gray code)

이 절에서는 개체 유전자의 표현형이 유전형으로 부호화될 때 그레이 코드에 따르는 경우를 다룬다. 그런데 그레이 코드의 특성상 식 (23)의 적합도 지형을 유전형 정의구역에서 그려보면 준안정상태가 적합도 장벽으로 둘러 쌓여 있지 않다는 걸 알 수 있다. 따라서 다음과 같이 새로운 적합도 함수를 정의한다.

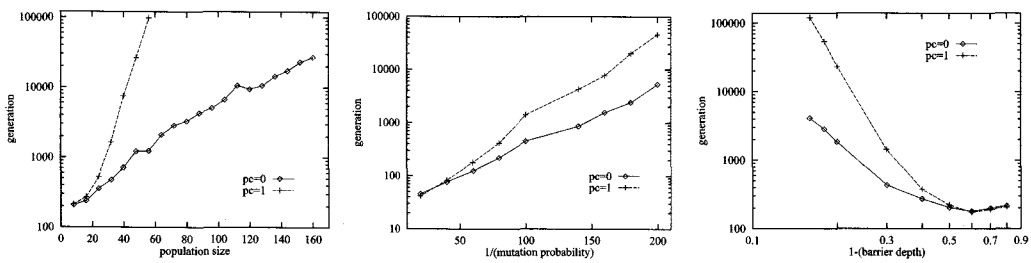

그림 5 변수들에 대한 준안정상태의 지속시간 T. 지속시간은 세대(generation) 단위이고 변수들은 개체군 크기(population size), 돌연변이 확률(mutation probability), 장벽의 깊이(barrier depth)이며 각 점은 100회 실행한 결과의 평균치다. 기본 변수값들은 L=30, N=40, p_m=0.01, δ=0.7이다.

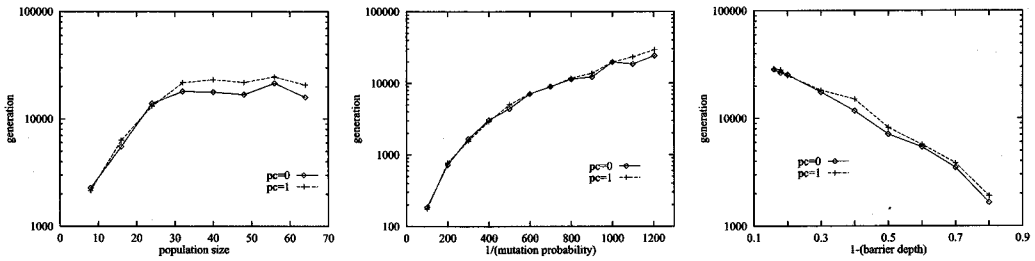

그림 6 변수들에 대한 준안정상태의 지속시간 T. 지속시간은 세대(generation) 단위이고 변수들은 개체군 크기(population size), 돌연변이 확률(mutation probability), 장벽의 깊이(barrier depth)이며 각 점은 100회 실행한 결과의 평균치다. 기본 변수값들은 L=30, N=30, p_m=0.001, δ=0.7이다.

$$f(x)=\begin{cases} -\frac{3\delta}{x_{go}-x_{lo}}x+\frac{3x_{lo}\delta}{x_{go}-x_{lo}}+1 & \text{if} \quad x_{lo}\le x<\frac{x_{go}+2x_{lo}}{3} \\ \frac{3\delta}{x_{go}-x_{lo}}x+\frac{(2x_{go}+x_{lo})\delta}{x_{go}-x_{lo}}+1 & \text{else if} \quad \frac{x_{go}+2x_{lo}}{3}\le x\le x_{go} \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

여기에서 x_{lo}와 x_{go}는 각각 지역 최적치와 전역 최적치를 갖는 표현형의 값을 나타낸다. 이렇게 정의하면 x_{lo}와 1 비트 만큼 다른 모든 유전형들이 x_{lo}보다 낮은 적합도를 갖게 된다. 또한 식 (23)에서와 비슷하게 적합도의 지역 최적치와 전역 최적치는 각각 1과 1+δ가 되고 장벽의 적합도는 $f((x_{go}+2x_{lo})/3)=1-\delta$가 된다. x_{lo}와 x_{go}의 유전형의 비트들이 서로 다르게 되도록 해서 유전형 영역에서도 두 지점이 멀리 떨어지도록 한다. 시뮬레이션에서는 L=6인 경우에 x_{lo}=5=000111$_{(2)}$과 x_{go}=47=111000$_{(2)}$를 사용한다. L값이 이렇게 작은 것은 L이 큰 경우에 이중안정성이 유지되도록 정의하는 것이 쉽지 않기 때문이다.

그림 7에서 시뮬레이션 결과가 나오는데 준안정상태의 지속시간 T가 개체군 크기나 돌연변이 확률에 따라 지수적으로 증가하지 않는다는 걸 알 수 있다. 적합도 장벽의 깊이와 T의 관계도 단위화 함수에 대한 결과와 다르다.

5. 토 의

단위화 적합도 함수를 사용하는 이중안정 문제에서 개체들은 어느 인력 웅덩이에 들어있는가에 따라 두 가지 종류로 나눌 수 있다. 준안정상태와 안정상태에 있는 개체들의 종류를 각각 A와 B라고 하자. r_A와 f_A를 각각 A 종류의 비율과 평균 적합도라고 하고 r_B와 f_B도 마찬가지로 정의한다. 개체군이 오직 A 종류만 가지고 있으면 B 종류는 선택만으로는 생성되지 않는다. B 종류의 개체가 돌연변이나 교차에 의해 생성될 확률이 아주 작아서 대부분의 세대에서 $r_B=0$라고 하자. 일단 r_B가 τ번째 세대에서 0보다 커지면 그 다음 세대의 r_B는 다음과 같이 구해진다.

$$r_B(\tau+1)\simeq\frac{f_B}{\bar{f}}r_B(\tau)-\frac{f_B}{\bar{f}}r_B(\tau)\frac{f_A}{\bar{f}}r_A(\tau) \tag{26}$$

여기에서 $\bar{f}=f_Ar_A+f_Br_B$이고 $r_A^2(\tau)$ 항과 돌연변이의 기여는 무시할 만큼 작으므로 생략됐다. 이 식은 A 개체와 B 개체가 교차하면 그 자식들은 장벽 근처에 생성

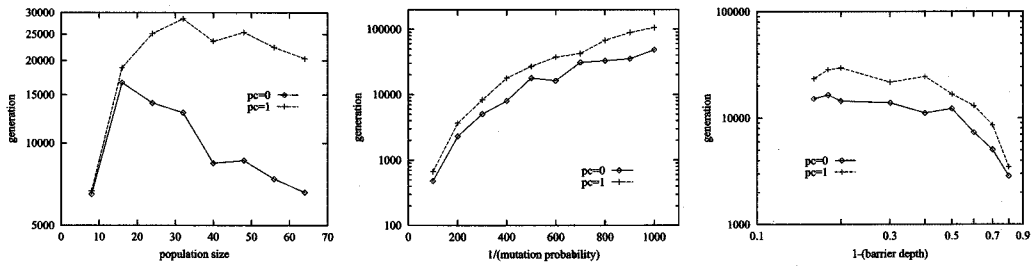

그림 7 변수들에 대한 준안정상태의 지속시간 T. 지속시간은 세대(generation) 단위이고 변수들은 개체군 크기(population size), 돌연변이 확률(mutation probability), 장벽의 깊이(barrier depth)이며 각 점은 100회 실행한 결과의 평균치다. 기본 변수값들은 L=6, N=30, p_m=0.002, δ=0.7이다.

되고 그 다음에 이어지는 선택에 의해 제거된다는 것을 의미한다. 즉, 적합도가 단위화 함수인 경우에 교차는 B 종류의 개체가 생성되는 것을 진작시키기도 하지만 또한 그것을 방해하기도 한다. 식 (26)에 따르면 r_B가 감소할 조건은 다음과 같다.

$$f_B\sqrt{r_B} < \overline{f} \tag{27}$$

이것은 교차가 다수파의 경쟁자를 제거할 충분조건이라고 할 수 있다. 앞에서 $r_B \ll 1$이라고 가정했으므로 f_B가 f_A보다 상당히 크더라도 이 조건은 성립한다.

표 1은 시뮬레이션에서 각 SGA 연산들의 기여도를 보여준다. 정의구역 [0,1/3)로 돌아오는 것은 주로 선택에 의한 것인데 정의구역 [1/3,7/15]로 돌아오는 것은 주로 교차에 의한 것이다. 전체적으로 보자면 교차가 B 종류의 출현을 저지하는 데에 중요한 역할을 하고 있다. 이것은 그림 4에서 $p_c = 1$에 대한 그래프가 더 긴 지속시간을 보여주는 것을 설명해 준다. 결국, 단속평형의 평형상태는 지역 최적치와 전역 최적치가 교차해서 생긴 자식 개체들이 적합도 장벽 위에 생기고 이들이 선택과정에서 제거되는 과정이 계속 반복되는 것이다.

표 1 포텐셜 지형에서 SGA 연산들에 의해 지역 최적치에서 탈출하거나 다시 지역 최적치로 돌아오는 횟수. 매 SGA 연산 이후, 개체군에서 최대의 표현형을 가진 개체를 추적한다. 이 개체는 각 연산자에 의해 장벽을 나타내는 특정 표현형의 기준치를 넘어서기도 하고 다시 돌아오기도 한다. 이 기준치는 (a)에 대해서는 1/3이고 (b)에 대해서는 7/15이다. 개체군의 전이가 일어날 때까지 각 SGA 연산에 의해 이 최대 개체가 몇 번이나 탈출하고 몇 번이나 돌아오는지 그 횟수를 기록하고 100회의 SGA 실행에 대해 이 횟수의 평균치를 구한다. 변수들의 값은 L=30, N=40, p_c=1, p_m=0.02, δ=0.7이다.

(a)

탈출 연산자	탈출 횟수	돌아오게 만드는 연산자 및 횟수		
		선택	교차	돌연변이
선택	0	0	0	0
돌연변이	333.9	217.8	108.2	7.4
교차	190.5	110.8	59.7	19.5
합계	524.4	328.6	167.9	26.9

(b)

탈출 연산자	탈출 횟수	돌아오게 만드는 연산자 및 횟수		
		선택	교차	돌연변이
선택	0	0	0	0
돌연변이	12.5	4.8	6.8	0.5
교차	8.8	2.9	3.7	1.6
합계	21.3	7.7	10.5	2.1

6. 결론 및 향후 연구

자연계의 진화에서 관찰되는 단속평형은 강체구가 이중안정 포텐셜 지형에서 아래쪽으로 당기는 힘과 작은 요동에 따라 움직이는 확률론적 모델로 잘 설명된다. 이 구의 중심은 개체의 특성 x의 개체군 평균 $\overline{x}$를 나타내는데, x는 확률변수 X에 대응된다. 여기에서 작은 요동은 상수 ε의 크기를 갖는다. 구가 준안정상태에 빠져있을 때 시스템은 일종의 평형 상태를 갖지만 이 작은 요동이 구를 준안정상태의 인력 웅덩이로부터 탈출할 수 있도록 해 준다. 이러한 확률론적 과정에 대한 수학적 연구에 따르면 평형의 지속시간 T는 $\exp(1/\varepsilon^2)$에 비례한다[8].

이러한 단속평형의 이론을 SGA에 적용하기 위해서는 몇 가지 조건이 만족되어야 한다. 첫 번째, 이산적인(discrete) SGA로부터 연속적인 모델을 만들기 위해 식 (9)에 대한 조건들이 필요하다. 두 번째, 미시적 효과를 배제하여 개체군을 강체구로 근사시키기 위해 단위화 적합도 함수가 사용된다.

이러한 조건들 아래, 본 연구에서는 이론적 분석을 통해 다음과 같은 결과를 얻었다. 첫째, 개체군 크기 N이 중심극한정리를 적용할 수 있을 정도로 큰 경우에 $\varepsilon^2 \simeq V(X)/N$이다. 둘째, 준안정상태에서의 $V(X)$ 값인 V_e는 N에 독립적이고 돌연변이 확률 p_m에 비례한다. 마지막으로, $T \propto \exp(N/p_m)$이라는 결과를 얻었다.

시뮬레이션에서는 지속시간 T를 탈출이 일어나기까지의 세대수로 정의하고 이것이 N과 p_m에 대해 어떻게 변하는지 기록했다. 시뮬레이션은 이론적 결과와 상응하는 결과를 생성했다. T와 적합도 장벽의 깊이 사이의 관계도 대략적인 이론적 분석에 따른 예상과 일관성이 있음을 보여주었다.

만약 미시적 효과만 고려한다면 지속시간 T는 개체군 크기가 커짐에 따라 줄어들 것이다. 단속평형에서는 그 영향력이 상대적으로 작으므로 본 연구에서는 미시적 효과를 무시했다. 하지만, 더 정확한 분석을 위해서는 향후 미시적 효과가 고려되어야 할 것이다. 큰 개체군을 서로 격리된 작은 개체군들로 나누는 것이 미시적 효과와 거시적 효과를 모두 얻는 방법이 될 수 있을 것이다. 이것은 병렬(parallel) GA들이 좋은 성능을 보이는 이유에 대한 이론적인 설명이 될 수 있다. 한편, 식 (22)는 더 큰 돌연변이 확률과 더 작은 적합도 장벽의 깊이가 SGA의 전역 최적화를 촉진한다는 것을 설명한다. 이것은 시뮬레이티드 어닐링(simulated annealing)의 기본적인 아이디어이기도 하다[21].

본 연구의 결과를 실용적인 측면에서 보자면, 지역 최적치로부터 빨리 탈출할 수 있도록 하려면 알고리즘의 주

요 변수들을 어떻게 설정해야 할 지에 대한 지침으로 이용할 수 있다. 또한 본 연구의 결과는 여러 가지로 확장될 수 있을 것이다. 첫째, 포텐셜 지형에 약간의 제한을 가하면 지속시간에 대한 결과를 F가 두 개 이상의 최적치를 가지고 있는 경우나 x가 다차원의 값인 경우에도 적용할 수 있다[8,20]. 또한, GA는 CE와 직접적인 관계를 가지므로 CE에 대해서 얻은 결과들을 GA에 적용할 수 있다. 여기에는 협동(cooperation), 경쟁(competition), 혼돈(chaos), 그리고 전역 정보의 지연(delayed global information) 등의 주제가 포함된다[19,22].

정리하면 다음과 같다. 유전 알고리즘은 자연계의 진화과정을 모델로 하는 것인데도 불구하고, 유전 알고리즘에서 단속평형의 지속시간과 관련된 기존의 연구들은 자연계에서의 연구와 반대되는 결과를 보여줘 왔다. 본 연구에서는 적합도 함수의 종류에 따라 다른 결과를 얻을 수 있다는 것을 지적하고 확산방정식을 적용하여 유전 알고리즘에서도 자연계에서와 같은 결과를 얻을 수 있다는 것을 수학적으로 보였다. 하지만 이러한 분석이 가능하도록 하기 위해 적합도 함수를 특정 종류로 국한할 수 밖에 없었다. 한편 추가로, 이러한 특정 종류의 적합도 함수에 대한 시뮬레이션 결과가 이론적 분석과 일치한다는 것을 보였다.

참 고 문 헌

[1] Goldberg, D. E., Genetic algorithms in search, optimization, and machine learning, Addison-Wesley, New York, 1989.

[2] Holland, J. H., Adaptation in natural and artificial systems, The MIT Press, Cambridge, MA, 2nd Ed., 1992.

[3] Goldberg, D. E., "Simple genetic algorithms and the minimal, deceptive problem," in Genetic Algorithms and Simulated Annealing, Morgan Kaufmann, Los Altos, CA, 1987.

[4] Shapiro, J. and Prügel-Bennett, A., "Genetic algorithm dynamics in a two-well potential," in FOGA4 edited by Belew, R. K. and Vose, M. D., Morgan Kaufmann, San Francisco, 1997.

[5] Ceccatto, H. A. and Huberman, B. A., "Persistence of nonoptimal strategies," Proceedings of the National Academy of Sciences of the United States of America, pp. 3443 - 3446, 1989.

[6] Newman, C. M., Cohen, J. E. and Kipnis, C., "Neo-darwinian evolution implies punctuated equilibria," Nature, pp. 400 - 401, May 1985.

[7] Vose, M. D. and Liepins, G. E., "Punctuated equilibria in genetic search," Complex Systems, Vol.5, No.1, pp. 31 - 44, 1991.

[8] Kipnis, C. and Newman, C. M. "The metastable behavior of infrequently observed, weakly random, one-dimensional diffusion processes," SIAM Journal on Applied Mathematics, Vol.45, No.6, pp. 972 - 982, 1985.

[9] Nimwegen, E. van, Crutchfield, J. P. and Mitchell, M., "Statistical dynamics of the royal road genetic algorithm," Santa Fe Institute Working Paper 97-04-035, 1997.

[10] Jansen, T. and Wegener, I., "On the analysis of evolutionary algorithms - a proof that crossover really can help," in Proceedings of the 7th Annual European Symposium on Algorithms (ESA'99) edited by Nesetril, J., Springer, Berlin, 1999.

[11] Rogers, A. and Prügel-Bennett, A., "The dynamics of a genetic algorithm on a model hard optimization problem," Complex Systems, Vol.11, No.6, pp. 437-464, 2000.

[12] Prügel-Bennett, A. and Shapiro, J. L., "An analysis of genetic algorithms using statistical mechanics," Physical Review Letters, Vol.72, No.9, pp. 1305 - 1309, 1994.

[13] Prügel-Bennett, A. and. Shapiro, J. L., "The dynamics of a genetic algorithm for simple random ising systems," Physica D, Vol.104, No.1, pp. 75 - 114, 1997.

[14] Rattray, M., "The dynamics of a genetic algorithm under stabilizing selection," Complex Systems, Vol.6, No.3, pp. 213 - 234, 1995.

[15] Scheaffer, R. L. and McClave, J. T., Probability and statistics for engineers, PWS-KENT, Boston, 1990.

[16] Srinivas, M. and Patnaik, L. M., "On modeling genetic algorithms for functions of unitation," IEEE transactions on Systems, Man, and Cybernetics, Vol.26, No.6, pp. 809 - 821, 1996.

[17] Huberman, B. A. and Hogg, T., "The behavior of computational ecologies.," in The Ecology of Computation edited by Huberman, B. A., North-Holland, Amsterdam, 1988.

[18] Kampen, N. V., Stochastic processes in physics and chemistry, Elsevier Science, Amsterdam, 1981.

[19] Kephart, J. O., Hogg, T. and Huberman, B. A., "Dynamics of computational ecosystems," Physical Review A, Vol.40, No.1, pp. 404 - 421, 1989.

[20] Galves, A., Olivieri, E. and Vares, M. E., "Metastability for a class of dynamical systems subject to small random perturbations," The Annals of Probability, Vol.15, No.4, pp. 1288 - 1305, 1987.

[21] Aarts, E. and Corst, J., Simulated Annealing and Boltzmann Machines, John Wiley & Sons, New York, 1989.

[22] Hogg, T. and Huberman, B. A., "Controlling chaos in distributed systems," IEEE transactions on Systems, Man, and Cybernetics, Vol.21, No.6, pp. 1325 - 1332, 1991.

오 상 엽

정보과학회논문지 : 소프트웨어 및 응용
제 32 권 제 9 호 참조