

교배방법의 개선을 통한 변형 실수형 유전알고리즘 개발

論 文
49D - 12 - 4

Development of a Modified Real-valued Genetic Algorithm with an Improved Crossover

李 德 揆* · 李 聖 煥** · 禹 天 熙*** · 金 鶴 培[§]
(Deog-Kyoo Lee · Sung-Hwan Lee · Chun-Hee Woo · Hag-Bae Kim)

Abstract - In this paper, a modified real-valued genetic algorithm is developed by using the meiosis for human's chromosome. Unlike common crossover methods adapted in the conventional genetic algorithms, our suggested modified real-valued genetic algorithm makes gametes by conducting the meiosis for individuals composed of chromosomes, and then generates a new individual through crossovers among those. Ultimately, when applying it for the gas data of Box-Jenkin, model and parameter identifications can be concurrently done to construct the optimal model of a neural network in terms of minimizing with the structure and the error.

Key Words : 유전알고리즘, 지능형 신경망, 구조동정, 감수분열 유전알고리즘

1. 서 론

유전알고리즘(Genetic Algorithm, GA)은 모든 진화 기반의 탐색 알고리즘 중에서도 가장 널리 알려진 기법 중 하나로 홀란드(Holland)[1]에 의해 제안되었다. 유전알고리즘은 여타의 진화알고리즘에 비해 상대적으로 자연계의 진화원리에 근접하며, 비록 특정 문제에 대해 어떤 알고리즘이 적합한지에 대해서는 여러 관점에 의해 종종 논란의 여지가 있을 수 있지만, 다양한 종류의 문제에 대해 유연한 우수한 해를 제공한다. 골드버그(Goldberg)[2]는 선택과 돌연변이가 결합되면 등반(hill-climbing) 메커니즘 역할을 수행하고 이는 국부적인 환경에서 주로 이루어지므로 이를 극복할 수단으로 선택과 재결합(교배) 연산을 통해 이를 수 있다고 하였다. 그러나 유전알고리즘에서는 실수형 파라미터를 유전자로 이진 코딩해야 하는 변환문제와 염색체에서 파라미터 간의 상관성에 따라 유전자의 위치를 적절히 할당해야 하는 또 다른 문제가 적절한 알고리즘의 설계시에 고려되어야 한다. 물론 실수형 유전알고리즘을 통해 변환문제는 해결 가능하다 해도 이진교배보다 작은 폭의 탐색영역에 제한되는 단점이 있으며, 또한 파라미터의 상관성을 고려하여 좋은 형질의 유전자를 보존할 수 있는 교배방법을 결정해야 하는 문제는 여전히 남게 된다. 교배와 돌연변이의 확률에 대한

적절한 선택이 최적해를 탐색하는 알고리즘의 성능을 좌우한다. 이에 대해 뮐렌바인(Muhlenbein)은 최적한 돌연변이 확률값에 대해 개체의 길이의 역수값을 제시하였다[3]. 그렇지만 결국 최적화 문제에서 교배나 돌연변이 확률값은 문제의 성격에 따라 시행착오에 의해 결정될 수밖에 없다.

이러한 문제점을 해결하기 위해 실수형 유전자들로 구성된 개체(individuals)에 대해 인간 내에 존재하는 감수분열(meiosis)을 적용하여 개체들을 생식체(gamete)들로 만들고, 이 생식체들의 랜덤한 선택과 교배를 통해, 세대가 진화함에 따라 다양한 개체로 전역탐색을 수행하고 우수개체에 대한 보존 및 변이를 통해 국부탐색도 동시에 수행하는 감수분열 유전알고리즘(Meiosis Genetic Algorithm, MGA)을 제안하였는데[4], 본 논문에서는 이와함께 감수분열 알고리즘의 결과로 생성되는 경계값의 발생을 억제하기 위해 각 세대의 최우수 개체의 유전자열 방향을 고려하여 min/max 연산을 적용함으로써 무의미할 수 있는 연산량을 줄이는 방법을 추가한 알고리즘에 대해 기술한다.

지능형 신경망은 생물의 신경계와 같은 복잡한 시스템을 비교적 단순한 소자의 집단으로 모델화하면서 여러가지 정보처리의 가능성을 시뮬레이션에 의해 구성할 수 있다는 내재적인 장점이 있다. 특히 신경망의 특성인 은닉층을 이용한 블랙박스 방식의 접근으로 선형계는 물론 비선형계에 대한 모델링도 가능하다. 이때, 신경망을 구성함에 있어 최적의 신경망을 구성하고자 할 때, 파라미터를 동정하여 오차를 줄이는 것뿐만 아니라 신경망 구조 자체를 최적화시키는 것도 포함하여 고려하여야 한다. 신경망 구조의 최적화는 결국 입력에 대한 출력의 연산속도 및 연산량과 깊은 연관이 있으므로 파라미터의 최적화와 함께 중요한 요소이다. 경험에 의한 신경망 구조 선택으로는 새로운 신경망을 구성해야 할 경우에 한계에 부딪히게 되며 결국 구조동정을 수행해야

* 正 會 員 : 延世大 工大 電氣電子工學科 博士課程

** 正 會 員 : 現代重工業 機電研究所 前任研究員

*** 正 會 員 : 明知專門大 電氣科 教授

[§] 正 會 員 : 延世大 工大 電氣電子工學科 副教授

接受日字 : 2000年 11月 6日

最終完了 : 2000年 12月 7日

만 하는 상황에 직면하게 된다. 따라서 이러한 경우에 신경망의 구조동정과 함께 파라미터 동정을 동시에 수행하여 최적화된 신경망 모델을 구성하는 것이 중요하다. 따라서 본 논문에서는 제안한 MGA에 대하여 먼저 Dejong 함수와 비선형 함수에 대해 알고리즘을 평가하고 Box-Jenkin의 가스로 데이터에 대해 자율적으로 구조동정과 학습오차를 최소화 하는 지능형 신경망 모델을 시뮬레이션을 통해 구현한다.

2. 감수분열 유전알고리즘

기존의 유전알고리즘에 있어서, 실수형과 이진형 모두, 교배시에는 개체간에 어떤 특정 교배선택위치에서 이루어지게 되며, 돌연변이는 개체를 이루는 유전자들 중 일부에 확률적으로 발생하게 한다 (그림1-(a)).

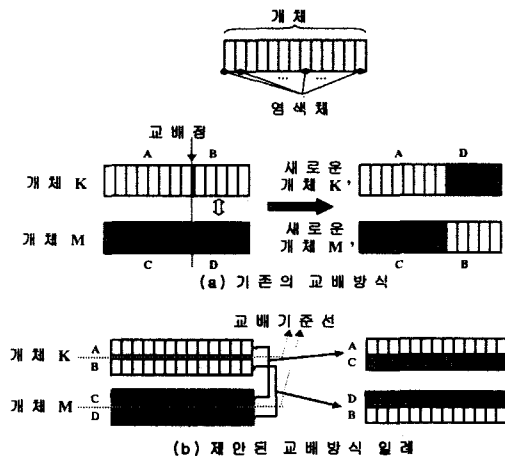


그림 1. 교배방식의 비교
Fig. 1. Comparison between two crossover types

이러한 경우, 물론 세대수 증가에 따라 최적해로의 수렴이 될 수도 있으나, 현재의 한정된 개체집단 내에서 우수한 개체가 많은 수의 자손을 낳게되므로 다른 개체들이 줄어들게 되고 이로 인해 공간탐색의 다양성 상실로 국부적 해로 수렴할 가능성도 있으며, 또한 교배나 돌연변이를 수행함으로써 부모의 형질을 잃어버리는 경우가 발생하기도 한다. 유전알고리즘에서 중요하게 다루어야 할 사항이 이와 같은 선택의 강도와 개체의 다양성의 적절한 조화문제이다. 이러한 문제는 기존의 유전알고리즘에 있어서 개체간의 일부 유전자 교환이라는 한정된 교배방식에 이유가 있다. 또한 실수계에서 이진스트링으로의 코딩으로 인한, 즉 재조합공간 (recombination space)과 평가공간(evaluation space)[5]과 같은 두 개의 서로 다른 공간사이의 변환에서 일어나는 오차와 연산시간의 증가라는 문제를 야기 시킬 수도 있다. 이와 같은 교배방식의 문제점을 개선하기 위해 생명체의 자손보존 전략인 자율적 교배상대 선택과 감수분열에 의해 생성되는 생식체의 교배를 통한 자손생성의 원리(그림1-(b))를 알고리즘에 적용하여 개체집단내의 각 개체의 특성을 비율적으로 보존하고 다양성을 갖는 자손들을 그 개체집단의 세대 진화에 따라 자연적으로 유지할 수 있게 하였다.

2.1 생식체의 생성 유형

교배방법에 대한 연구가 많이 수행되어오고 있는 가운데 유전알고리즘에서의 산술적 교배[6][7], 또는 진화전략에서의 중간교배[8]라는 산술적 개체의 교배방법을 변형한 새로운 교배방식을 제안하고자 한다. 교배를 위한 기존의 방식과는 다르게 생식체(반개체)의 개념을 도입한다. 즉, 개체를 가로절로 나누어 생식체를 생성시키고 이들을 랜덤 선택하여 교배를 수행함으로써 다음세대의 자식개체를 생성하는 방법이다. 이때 생식체의 생성유형, 즉, 개체의 감수분열 유형은 그림 2 에서와 같이 세 가지로 나누어 나타내었다. 여기서 XX, XY 타입은 성염색체의 외형적 특징을 이용하여 명명한 것으로, 인간과 같이 성염색체를 따로 갖고 있는 것이 아니라 외형적으로 감수분열되는 비율에 따라 생성되는 생식체를 기준으로 명명한 것이다. 또한, MM 타입 감수분열은 돌연변이(mutation) 유형을 나타내며 인간에 있어서 염색체의 이상현상을 모델로 제안한 것이다.

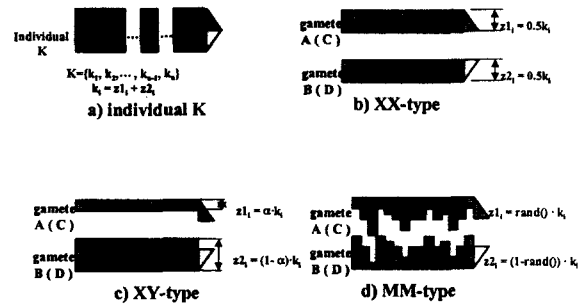


그림 2 생식체의 생성유형
Fig. 2 Types of Gametes

그림 2에서 rand()는 랜덤값을 발생시키는 랜덤함수를 의미하며, 개체에 대한 생식체 생성유형은 아래와 같이 정의한 감수분열의 기본 형식에 의해 나타낼 수 있다.

먼저 개체 K는 유전자 k_i ($1 \leq i \leq n$)의 1차원 배열로 구성되어 있으며, 파라미터 α 는 개체에 대한 감수비율을, θ_i 는 개체를 구성하는 유전자 각각에 대한 랜덤 감수비율을 나타낸다. 이때 파라미터 θ 가 반영된 개체 K를 K'라 나타내었다.

$$\text{개체 } K = \{k_1, k_2, \dots, k_{n-1}, k_n\} \quad (2.1)$$

$$\text{개체 } K' = \{k'_1, k'_2, \dots, k'_{n-1}, k'_n\} \quad (2.2)$$

$$\text{파라미터 } \theta = \{\theta_1, \theta_2, \dots, \theta_{n-1}, \theta_n\} \quad (2.3)$$

하나의 개체가 분리되어 생성되는 생식체는 두 개가 생성되는데 다음과 같은 식으로 나타낼 수 있으며 이때 파라미터의 비율에 따라 유형이 결정되게 된다.

$$\text{생식체 } A = \{z_{11}, z_{12}, \dots, z_{1n-1}, z_{1n}\} \quad (2.4)$$

$$\text{생식체 } B = \{z_{21}, z_{22}, \dots, z_{2n-1}, z_{2n}\} \quad (2.5)$$

$$k'_i = z_{1i} + z_{2i} \quad (2.6)$$

$$z_{1i} = \alpha k_i(1 + \theta_i) \quad (2.7)$$

$$z_{2i} = k_i - \alpha k_i(1 + \theta_i) \quad (2.8)$$

2.1.1 XX type

하나의 개체를 1/2로 감수분열하여 두 개의 동일한 생식체를 생성하는 경우로 이 유형들끼리 교배가 이루어지면 두 개의 온전한 개체의 결합시 수행되는 중간 산술 결합의 효과를 나타내게 된다. 이러한 0.5의 감수비율을 갖는 경우 전역적인 최적해가 존재하는 우수한 개체들이 있는 엘리트 집단 내에서 수행되게 되면 완만하게 전역 최적해에 도달할 수 있게 된다.

제안된 알고리즘에서도 엘리트집단에 대해 XX 유형으로 개체를 나누어 각 유전자의 파라미터 θ_i 에 대해 미세한 랜덤변화값 만을 부가, 엘리트풀에서의 랜덤 선택 교배를 수행하여 최적해에 접근할 수 있도록 하였다. 이때 최적해에 가까워짐에 따라 유전자들에 가해지는 변화의 폭을 수렴하기에 적합하도록 해야하는데 본 논문에서는 수렴에 따른 변화의 폭을 결정하기 위한 파라미터로 i_{factor} 을 정의하고 새로운 최적해의 발생유무를 기준으로 하여 변화량을 가하도록 하였다. 이때 $Eval$ 은 개체의 적합도를 평가한 회수이다. 이렇게 적용한 이유는 세대수의 증가에 따른 변화폭의 일정한 감소율로 인하여 변화량의 크기가 작아서 새로운 최우수개체의 수렴속도가 늦어지는 것을 방지하기 위함이다. $rand(\alpha)$ 는 유전자 각각에 대한 감수비율의 랜덤함수로 0과 0.5사이의 랜덤값을 발생한다.

$$\alpha = 0.5 \quad (2.9)$$

$$i_{factor} = \begin{cases} 0 & ; \text{새로운 해가 발생한 경우} \\ i_{factor} + + & ; \text{새로운 해가 발생하지 않은 경우} \end{cases} \quad (2.10)$$

$$\theta_i = rand(\alpha) / (Eval - i_{factor}) \quad , 1 \leq i \leq n \quad (2.11)$$

2.1.2 XY type

동적풀의 생식체들의 일부를 생성시키는 XY 타입 개체 분리의 경우, XX 타입에서의 경우와는 달리, 유전자 각각에 대한 감수비율은 없으며, 개체에 대한 감수비율 α 만을 적용하여 생성된다. 따라서 생식체는 개체내의 유전자들에 적용된 비율이 일률적이며 이러한 특성을 지닌 생식체들간의 교배를 통하여 현재의 국소값이 있는 부근을 크게 이탈하는 수단으로서 이용될 수 있으며 최적해로 접근시 수렴의 속도를 증가시키는 효과가 있다. 즉, 작은값이나 큰값의 개체군을 쉽게 형성함으로써 개체내의 형질특성을 유지하면서, 공간상의 다른 범위를 탐색을 가능하게 하는 효과를 나타낸다. 아래의 식에서 개체내의 유전자 각각에 대한 변화가 없으므로 θ_i 는 0을 갖게 되며, α 는 0에서 1사이의 랜덤값을 갖게 된다.

$$\alpha = rand(1) \quad (2.12)$$

$$\theta_i = 0 \quad , 1 \leq i \leq n \quad (2.13)$$

2.1.3 MM type

돌연변이 생식체 유형인 MM-타입은 동적풀의 또 다른

생식체 일부로서, 개체내의 모든 유전자들에 대해 랜덤하게 감수비율을 적용함과 동시에 개체에 대해서도 랜덤한 감수비율을 적용하여 개체를 감수분열하여 생성된 생식체이다. 랜덤하게 개체를 생성해서 적용하는 것과 마찬가지로의 효과를 나타내게 된다. 진화전략에서의 돌연변이와 같은 효과를 얻을 수 있으며, 이들의 랜덤 선택 교배시 생성되는 개체는 새로운 값으로, 공간적 이동폭이 크며 이에 따라 전역적 탐색의 수단으로서 효과를 나타낸다. 기존의 유전알고리즘에서의 개체집단의 진화에 따라 집단전체가 최적해를 갖는 개체들로 채워져 나가게 됨으로써 국부적 해에서의 탈출이 어려운 점을 이것으로 보완할 수 있다. 개체 감수비율인 α 는 0에서 1사이의 랜덤값, 유전자에 대한 감수비율 θ_i 는 구간의 상한값과 하한값 사이의 랜덤값을 갖게된다. 이때, ω_i 는 상한값과 하한값 사이의 구간폭을 의미한다.

$$\alpha = rand(1) \quad (2.14)$$

$$\omega_i = upper\ bound - lower\ bound \quad (2.15)$$

$$\theta_i = rand() \leq 0.5 * (\omega_i) \quad , 1 \leq i \leq n \quad (2.16)$$

2.1.4 경계값 발생량의 억제

생식체들간의 랜덤 선택에 의한 교배결과, 그림 3과 같이 XY, MM타입의 영향으로 경계값을 초과하는 개체가 생성될 수 있는데 이들에 대한 억제가 필요하다. 즉, 최적해가 경계값에 없는 경우에는 무의미한 연산으로 수행시간을 증가시킬 뿐이다.

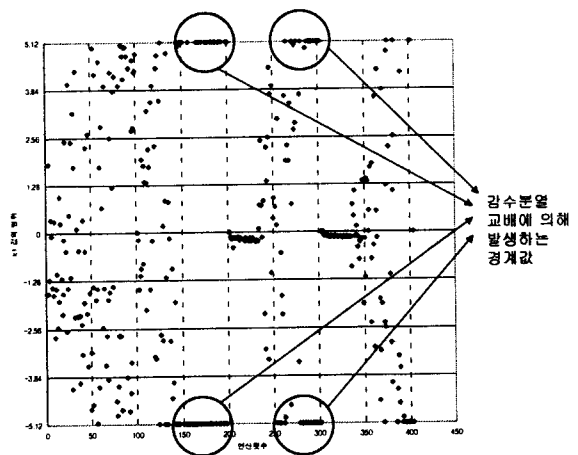


그림 3 생식체의 교배에 의한 경계값 발생
Fig. 3 Boundary values from crossover of gametes

경계값이 어느 범위를 초과한 경우 이에 대한 억제방법으로 그림 4와 같이 상태비트를 이용한 min/max 연산을 적용하였다. 이는 경계값을 갖는 개체들에 대해 랜덤한 개체로 대체하는 효과를 나타낸다. 상태비트는 세대가 진화해감에 따라 발생하는 각 세대에서의 최우수개체의 유전자값의 변화 방향을 나타내며 이 방향의 유전자값을 갖는 개체들을 발생시키는 것이다.

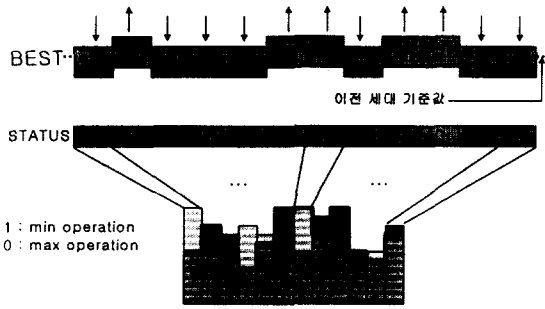


그림 4 상태비트에 의한 유전자값의 증감방향
Fig. 4 Direction of genes by status bits

2.2 감수분열 유전알고리즘의 구성

감수분열을 이용한 변형된 유전알고리즘은 그림 5와 같이 나타낼 수 있다.

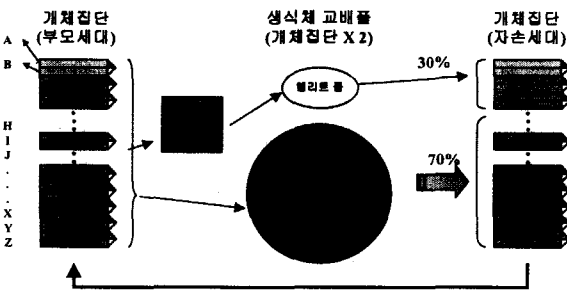


그림 5. 감수분열 유전알고리즘의 구성
Fig. 5. Structure of a meiosis genetic algorithm

2.2.1 개체집단

개체집단은 세대가 진화됨에 따라 정해진 개체집단의 크기를 갖는 부모세대(parent generation)와 자손세대(child generation)가 번갈아 존재하게 된다. 현세대의 부모세대내의 개체들은 생식체로 분리되어 교배풀에 전달되며 교배풀에서 랜덤하게 선택되어 자손세대를 구성하게 된다. 초기 개체집단은 랜덤한 값들로 구성되며 이 개체집단의 개체들에 대한 적합도를 평가하여 초기 최우수 개체를 선정하며 자손세대의 엘리트 집단을 구성하는 엘리트풀의 종자(seed)가 되어 최우수개체 버퍼에 저장된다. 적합도 계산은 개체집단내의 개체들에 대해 수행되며 이는 곧 연산시간과 비례하게 되므로 작은 개체집단의 크기를 가지는 것이 유리하다. 새로운 개체집단인 자손세대를 구성하기 위해 부모세대내의 개체들은 XX-타입을 기준으로 한 최우수개체 변형그룹과 XY-타입 감수분열 유형을 적용할 개체그룹, MM-타입 감수분열을 적용할 그룹의 세 가지로 나누어지게 된다. 개체집단과는 별개이지만 최우수개체를 저장하는 버퍼를 두어 그 세대에서 최우수 개체값을 보존하게 하였다.

2.2.2 생식체 교배풀

생식체 교배풀은 엘리트풀과 동적풀 두 가지로 구성되어 있다. 엘리트풀은 부모세대의 최우수개체에 대해 XX-타입 감수분열 유형을 적용하여 생식체로 분리하고 그 감수분열된 생식체내의 유전자 각각에 미세한 가중치를 부여하여 생성된 생식체들이 있는 저장장소이다. 이 엘리트풀 내의 생식체들 간에 교배를 수행하게 함으로써 현재상태의 적합도 부근에서 더 좋은 적합도값을 탐색할 수 있도록 하는 미세수렴탐색 방법이다. 엘리트 풀에서 교배를 통해 생성된 개체들은 다음 세대인 자손세대 개체집단의 30%를 채우도록 하였다.

동적풀은 부모세대의 개체들을 세 가지 유형으로 분리한 생식체들로 이루어져 있으며, 교배 상대의 선택은 생대개의 자율적인 선택에 의거 랜덤한 선택에 의하여 교배를 수행하도록 하였다. 자손세대 개체집단의 70%는 동적풀에 존재하는 모든 생식체들간의 랜덤교배에 의해 생성된 개체들로 채워지게 된다.

2.2.3 자손세대를 구성하는 개체유형

생식체들간의 교배결과 얻을 수 있는 새로운 개체들은 다음 표 1과 같으며 최종적으로 구성된 상태비트에 의한 min/max 연산을 적용한 감수분열 유전알고리즘의 흐름도는 그림 6과 같다.

표 1. 생식체간의 결합으로 생성되는 개체유형
Table 1. Each individual type by composing gametes

Type	X	Y	M
X	XX	XY	XM
Y	XY	YY	YM
M	XM	YM	MM

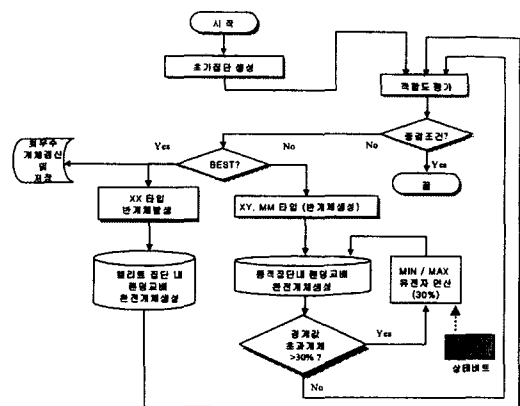


그림 6. MIN/MAX 감수분열 유전알고리즘
Fig. 6. MIN/MAX Meiosis Genetic Algorithm

3. 시뮬레이션

감수분열에 의해 구성된 변형 유전알고리즘의 성능평가를 위해 De Jong이 제시한 함수 최적화를 위한 시험용 함수와 비선형 함수들[4][9]에 대해 시뮬레이션을 수행하였고 얻은 결과는 표2와 같다. 이 알고리즘을 이용하여 Box-Jenkin의 가스로 데이터에 적용하여 신경망을 생성하였다.

3.1 감수분열 유전알고리즘에 대한 성능평가

본 시뮬레이션에서는 개체집단의 크기를 모두 10으로 고정시키면서 개체집단에서 엘리트집단은 30%를 할당하였다. 물론 이와 같은 가정은 결과의 심각한 변화없이 일반화시킬 수 있다. 부모세대로부터 동적 교배풀로 전달되는 생식체 유형은 XY 유형과 MM 유형을 각각 50%씩 생성시켰다. 엘리트풀에 대해서 감수비율을 0.5, 동적풀을 구성하기 위해 XY유형과 돌연변이 유형에 대해서 감수비율을 0에서 1사이의 랜덤값을 갖게 하였으며 함수에 따라 조정해야 할 파라미터를 별도로 두지 않았다.

특히, 변수에 대한 제한값이 있을 때 감수분열 유전알고리즘의 경우, 교배에 의해 중간값으로 수렴하는 기본적인 성질과 함께 감수분열의 효과로서, XY나 MM 간의 교배에 의해 구간을 이탈하게 되는 경우가 생기는데 이러한 경우에는 구간경계값을 갖도록 하였다. 구간경계값을 갖는 개체가 발생하더라도 교배풀에 들어갈 때 생식체 상태에 변화가 가지므로 다음세대에서 경계값으로부터 벗어날 수 있다. 따라서 교배풀은 항상 동적인 생식체들로 구성되게 된다.

다음 그림 7은 Dejong 함수 2에 대해 최적해에 수렴하는 과정을 나타낸 것으로 세대가 진화함에 따라 최적해로 수렴하는 적합도 값과 각 세대에서의 최우수개체가 발생한 위치를 나타낸다. 여기서 발생개체위치란, 개체집단의 크기가 10일 때 엘리트그룹은 1~3(30%), 나머지 동적그룹은 4~10(70%)의 위치를 의미하며, 동적그룹에서의 최우수개체 발생빈도가 높아질수록 최적해로의 수렴폭이 커짐을 알 수 있다. 개체발생위치가 0인 경우는 다음세대에서 현세대 적합도값 보다 좋은 값을 갖는 개체가 발견되지 않았음을 의미한다. 결과적으로 동적개체 그룹에서의 최우수개체의 발생으로 보통 최적해로의 빠른 접근을, 엘리트그룹에서의 최우수개체의 발생으로 완만하게 최적해로 수렴하는 특징을 나타내고 있다.

표 2. 전역최소값을 찾는데 요구되는 평균함수평가횟수
Table 2 Averaged number of function evaluations(nfe) required for finding the global minimum.

$f_i(x)$	ANM	ASA	DE1	DE2 (F=1)	MGA
F1	95	397	490	392	295
F2	106	11275	746	615	4154
F3	90258	354	915	1300	76
F4	X	4812	2378	2873	25
F5	X	1379	735	828	2235
F6	X	3581	834	1125	475
F7	X	X	22167	12804	554
F8	2116	11864	1559	1076	12

ANM, ASA, DE1 및 DE2 의 결과는 [9]에서 인용하였으며, ANM은 심플렉스(simplex), ASA 는 어닐링(annealing)에 근거한 알고리즘으로 최적값을 탐색하는 기법이며, DE1과 DE2는 차분진화알고리즘(Differential Evolution)으로 개체에 돌연변이 벡터를 적용하여 탐색하는 기법이다. [9]에서 제시한 함수들에 대해 MGA를 적용하여 얻은 결과를 이들 기법의 결과들과 함께 표 2에 나타내었다. 표2에서 F3, F4, F7, F8의 함수에 대해서 빠른 수렴결과를 나타내었는데 이는 감수분열 유전알고리즘의 교배로 인한 경계값을 갖는 개체에 의한 효과이다.

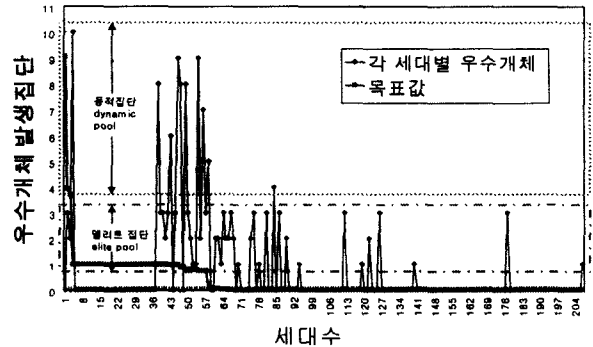
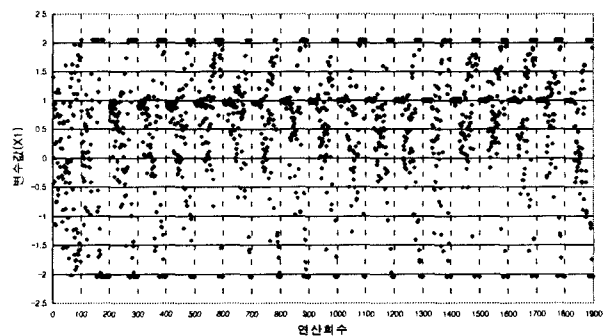
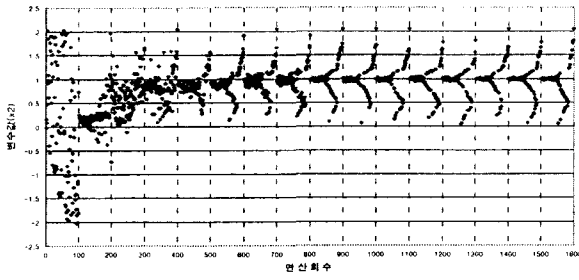


그림 7. 세대진화에 따른 집단별 우수개체 발생상태
Fig. 7. Position of finding the individual according to generation

그림 8은 De Jong 함수 F2 에 대해 MGA 와 min/max를 적용한 MGA하였을 때 세대진화에 따른 변수값의 분포를 나타낸다. 변수와 함수평가 값의 분포를 확인하기 위하여 개체집단의 크기는 100으로 하였다. MGA에서 변수들의 경계값이 많이 생성되는 반면에 min/max 연산을 적용한 MGA에서는 억제되었음을 알 수 있다. 그림 9는 세대진화에 따른 함수평가값에 대한 분포를 MGA와 min/max 방법을 적용한 MGA에 대해 비교한 것이며, 그림 10은 이때 MGA와 min/max MGA에서 변수에 따른 최적값으로의 수렴상황을 나타낸다.

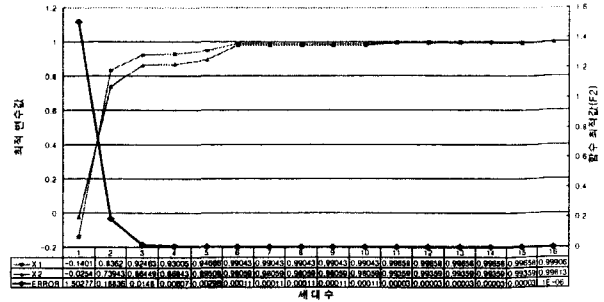


(a) MGA



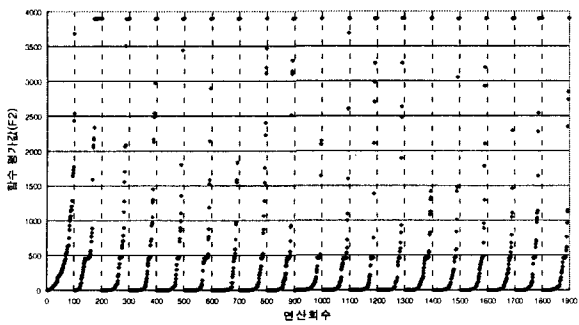
(b) MIN/MAX_MGA

그림 8 세대진화에 따른 최적변수 X2의 분포(F2)
Fig. 8 Distribution of optimal variable X2 according to generation(F2)

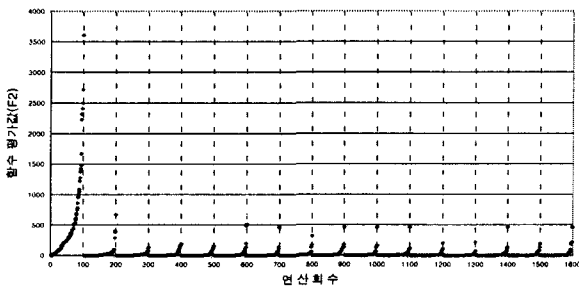


(b) MIN/MAX_MGA

그림 10. 함수 F2 에 대한 함수 최적값 및 변수값의 진화
Fig. 10. Evolution of optimal value and variable for function F2



(a) MGA



(b) MIN/MAX_MGA

그림 9 세대진화에 따른 함수 최적값 분포(F2)
Fig. 9 Distribution of optimal values for F2 according to generation

3.2 Box-Jenkin의 가스로 데이터에 대한 신경망 생성

가스로 프로세스는 입력 $u(k)$ 로 가스로 흐름률(gas flow rate)와 출력 $y(k)$ 로 CO_2 의 농도를 갖는 다이나믹 프로세스이다. 이 예는 296개의 입/출력 데이터 쌍으로 구성되며 $X = \{x_5, x_6, \dots, x_{296} | x_l = (u(l-4), y(l-1), y(l))\}$ 의 형태를 가지는 입/출력 데이터 쌍으로 정의된다.

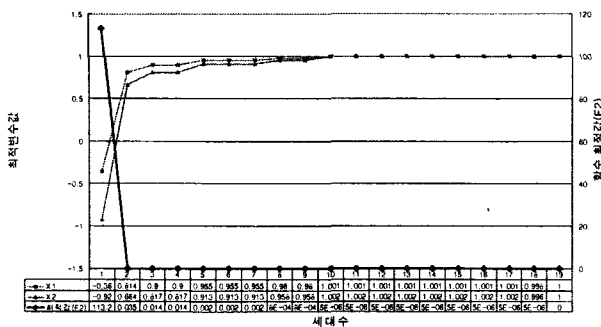
이에 대한 모델링은 신경망 모델로 구성하였으며 뉴럴모델의 파라미터 및 구조동정을 수행하였다. 적용한 신경망 모델은 3층의 순방향 신경망 구조로 최대값과 최소값에 의한 정규화된 데이터를 학습할 수 있게 구성하였다. 이때의 은닉층 뉴런과 출력층 뉴런은 각각 (3.1-2)와 같은 시그모이드 함수 특성을 가지며, 각 뉴런의 출력은 (3.3-4)와 같이 나타낼 수 있다. 여기서 H_j 및 O_k 는 각각 은닉층 및 출력층 뉴런의 시그모이드 함수의 기울기이며 W_{IHij} 는 i 번째 입력층 뉴런과 j 번째 은닉층 뉴런간의 연결하중치, W_{IOik} 는 i 번째 입력층 뉴런과 k 번째 출력층 뉴런간의 연결하중치, W_{HOjk} 는 j 번째 은닉층 뉴런과 k 번째 출력층 뉴런사이의 연결하중치를 나타내며 B_{Hj} 와 B_{Ok} 는 각각 은닉층 뉴런과 출력층 뉴런의 바이어스 값이다.

$$S_{Hj}(x) = \frac{1}{1 + \exp(-H_j x)} \quad (3.1)$$

$$S_{Ok}(x) = \frac{1}{1 + \exp(-O_k x)} \quad (3.2)$$

$$N_j = S_{Hj} \left(\sum_{i=1}^l X_i W_{IHij} + B_{Hj} \right) \quad (3.3)$$

$$Y_k = S_{Ok} \left(\sum_{j=1}^m N_j W_{HOjk} + B_{Ok} \right) + \sum_{i=1}^l X_i W_{IOik} + B_{Ok} \quad (3.4)$$



(a) MGA

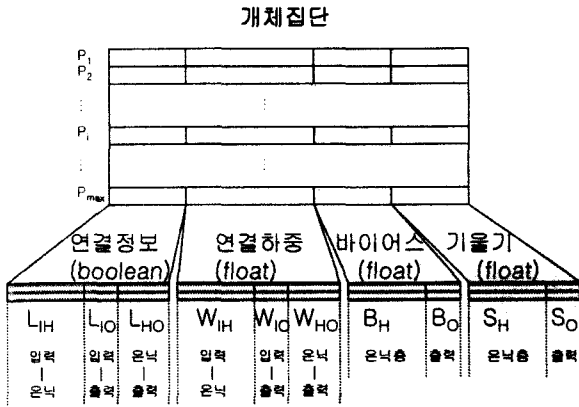


그림 11. 개체집단 및 개체내의 유전자 정보
 Fig. 11. Population and information of chromosomes in an individual

그림 11은 개체집단과 하나의 개체 내 유전자 구성을 나타내는 것으로 연결정보는 이진값으로서 링크의 제거나 추가 상태를 나타내는 정보이다. Box-Jenkin의 데이터에 대하여 시뮬레이션을 수행할 때 은닉층이 하나인 3층 구조를 바탕으로 최적의 구조동정과 파라미터 동정을 동시에 수행하였다. 신경회로망에서 초기 구조는 입력노드 2개와 은닉층노드 12개 출력노드 1개로 구성하였다. 개체집단의 크기는 20, 진화세대수는 20000회를 수행하였으며, 그때의 최적 구조 및 에러값을 나타내었다. 각 경우에서 생성된 신경망 구조에서 은닉층에서 발생한 노드중에는 G1과 G2 같이 입력과는 무관하게 출력노드에 이득(gain)으로 더해지는 노드가 존재하는데 이는 출력노드에 가해지는 산술적 이득으로 생각할 수 있다.

- 최종학습오차 : 0.18751
- 은닉층 노드수 : 3개
- 입력층-은닉층(IH)간 링크수 : 5개
- 입력층-출력층(IO)간 링크수 : 2개
- 은닉층-출력층(HO)간 링크수 : 3개

그림 12는 에러를 최적화 대상으로 했을 때 학습오차를 나타내며 그림 13은 20000세대후 최종적으로 최적적으로 구성된 신경회로망을 나타낸다.

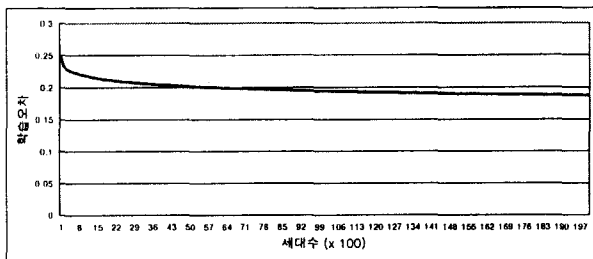


그림 12. 에러를 최적화 대상으로 했을 때 오차 변화
 Fig. 12. Error change in case of the optimal target achieving the minimum error

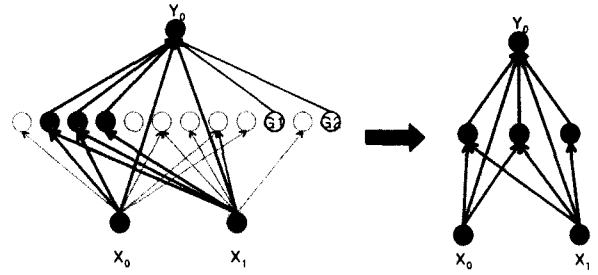


그림 13. 에러를 최적화 대상으로 했을 때 자율적으로 구성된 최적의 신경망 구조
 Fig. 13. Automatically-generated structure of a neural network in case of the optimal target achieving the minimum error

그림 14는 최소학습오차를 갖도록 구성된 신경회로망에 대해서 Box-Jenkin 데이터의 실제값과 모델출력값을 비교한 그래프이다.

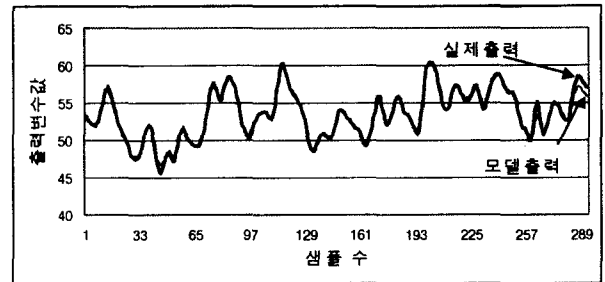


그림 14. 최소오차를 가질 때 Box-Jenkin 데이터의 실제값과 모델값
 Fig. 14. A reference output and a model output for the optimal target of minimum error

4. 결 론

제안된 감수분열 유전알고리즘은 기존의 유전알고리즘과 다음과 같은 점에서 차별화된다. 첫째로, 어떠한 교배점이 정해지지 않는다는 사실에서 기존의 교배방식과는 다르다는 점이다. 따라서, 개체집단의 모든 개체에 대해서 교배가 이루어지게 된다. 둘째로, 모든 개체가 교배를 수행하기 위해 생식체를 구성하고자 할 때 집단에 따라서 돌연변이성 생식체를 나타낼 수 있다는 점이다. 기존의 알고리즘에서는 세대가 진화해 나감에 따라 우수개체집단 보호 및 유지의 차원에서 돌연변이의 확률을 작게 두었으나 본 논문에서 제안된 알고리즘에서는 개체들의 모든 유전자에 대해서 돌연변이를 수행하고 있다. 모든 개체의 교배와 모든 유전자에 대한 돌연변이의 수행을 통해서 전역적인 탐색이 원활하게 수행됨을 보였다. 특히, 비선형 함수에 대한 최적값의 탐색에 있어서는 작은 수의 개체집단으로도 경계값에 해가 존재하는 경우 빠른 수렴속도를 나타내었는데, 이는 제안된 알고리즘에서 동적풀의 개체들의 교배에 의해 경계값을 용이하게 얻을 수 있기 때문이다. Box-Jenkin 데이터에 대한 지능형 신경망 모델 구성에서는 최적화하고자하는 대상에 따라 구

조 및 학습오차의 차이를 나타내었는데 이는 학습오차와 구조간의 설계비중에 따라 적절한 선택이 이루어질 수 있음을 나타내고 있다. 따라서 모델구성에 대한 최적화의 관점에 따라 자율적으로 효과적인 모델을 구성할 수 있음을 보였다.

참 고 문 헌

[1] J. H. Holland, 'Adaptation in Natural and Artificial Systems', The University of Michigan Press, Ann Arbor, 1975.

[2] D. E. Glodberg, "The Design of Innovation: Lessons from Genetic Algorithms", Lessons for the Real World, IlliGAL Report No. 98004, February, 1998.

[3] Thomas Baeck, "Optimal Mutation Rates in Genetic Search", *ibid*, pp. 2-8, 1993.

[4] 이덕규 외 4인, "염색체의 감수분열을 응용한 변형 유전알고리즘에 대한 연구", 한국 정보처리학회 논문지 제 7권 제 6호, pp. 1844-1851., 2000. 6.

[5] P. J. Angeline, G. M. Saunders and J. B. Pollack, "An evolutionary algorithm that constructs recurrent neural networks", *IEEE Trans. Neural Networks*, vol. 5, no. 1, pp. 54-65, Jan., 1994.

[6] Z. Michalewicz and C. Janikow, "Handling Constraints in Genetic Algorithms", *Proceedings of the Fourth Int., Conf. on Genetic Algorithms Morgan Kaufmann Publishers, Los Altos, CA*, pp. 151-157, 1991.

[7] Z. Michalewicz and C. Janikow, "GENOCOP : A Genetic Algorithm for Numerical Optimization Problems with Linear Constraints", *Communications of the ACM*, 1992.

[8] H. P. Schwefel, Private communication, 1991.

[9] R. Storn and K. Price, "Differential Evolution - A simple and efficient adaptive scheme for global optimization over continuous spaces", *Technical Report TR-95-012, ICSI, March 1995.*

감사의 글

이 논문은 2000년도 두뇌한국21 사업에 의하여 지원되었음.

저 자 소 개



이 덕 규 (李 德 揆)

1965년 1월 3일생. 1987년 연세대 전기공학과 졸업. 1989년 동 대학원 졸업(석사). 1991년-1996년 쌍용자동차 중앙연구소 연구원. 1996년~현재 동 대학원 전기전자공학과 박사과정

Tel : 02-2123-2778, Fax : 02-362-2780

E-mail : 2theq@yonsei.ac.kr

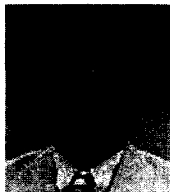


이 성 환 (李 聖 煥)

1966년 8월 10일생. 1990년 연세대 전기공학과 졸업. 1992년 동 대학원 전기공학과 졸업(석사). 1998년 동 대학원 전기공학과 졸업(공학박). 1995년~현재 현대중공업(주) 기전연구소 선임연구원

Tel : 031-289-5125, Fax : 031-289-5115

E-mail : hwaney@hhi.co.kr



우 천 희 (禹 天 熙)

1961년 6월 5일생. 1985년 연세대 전기공학과 졸업. 1993년 동 대학원 졸업(석사). 1985년-1995년 현대중공업 중앙연구소 선임연구원. 2000년 동 대학원 졸업(공학박). 1995년~현재 명지전문대학 전기과 교수

Tel : 02-300-1266, Fax : 02-362-2780

E-mail : chwoo@mail.mjc.ac.kr



김 학 배 (金 鶴 培)

1965년 10월 10일생. 1988년 서울대 전자공학과 졸업. 1990년 미국 미시간대 대학원 전기공학과 (EECS) 졸업(석사). 1994년 8월 동 대학원 졸업(공학박). 1994년 9월-1996년 9월 미국 National Research Council

(NRC) Research Associate at NASA Langley Research Center. 1996년 9월~현재 연세대학교 기계전자공학부 부교수

Tel : 02-2123-2778, Fax : 02-362-2780

E-mail : hbkim@yonsei.ac.kr