

감수분열 유전알고리즘을 이용한 퍼지 모델의 자동 설계

고택범*, 이덕규**

*경주대학교 컴퓨터전자공학부 **연세대학교 기계전자공학부

Design of fuzzy model using meiosis-genetic algorithm

Taek-Beom Koh*, DeogKyo Lee**

*School of CEE, Kyongju University **School of MEE, Yonsei University

Abstract - 본 연구에서는 실수형 염색체들로 구성된 개체에 대해 감수분열을 적용하여 개체를 만들고, 이 생식체들의 랜덤한 선택과 교배에 의해 세대가 진화함에 따라 탐색을 수행하는 감수분열 유전알고리즘을 이용하여 퍼지모델의 최적 구조와 파라미터를 탐색하고 Gradient Descent 알고리즘으로 파라미터를 정밀 조정하는 방안을 제안한다. 제안된 방안을 적용하여 Box-Jenkins의 가스로 데이터에 대한 퍼지모델을 구성하고 그 적용 가능성을 보인다.

1. 서론

홀란드(Holland)[1]에 의해 제안된 유전 알고리즘은 자원의 진화원리를 토대로 유전적 성질에 기초한 탐색 알고리즘으로 다양한 문제에 대해 우수한 해를 제공한다. 골드베르크(Goldberg)[2]는 선택과 돌연변이의 결합에 의해 등반(Hill-climbing) 메커니즘 역할을 수행하고, 선택과 교배(재결합) 연산을 통해 국부 수렴을 방지할 수 있다고 하였다. 하지만, 실수형 파라미터를 2진 문자열 코딩의 염색체로 변환해야 하므로 탐색 정밀도에 영향을 주고, 교배와 돌연변이에 의해 부모의 형질을 잃어버리는 경우가 발생하며, 파라미터의 상관관계에 따라 유전자의 위치를 설계 시에 고려해야 하는 문제가 발생한다. 이러한 문제점을 해결하기 위해 본 논문에서는 실수형 염색체들로 구성된 개체(individuals)에 대해 인간 내에 존재하는 감수분열(meiosis)을 적용하여 개체를 생식체(gamete)로 만들고 이 생식체들의 랜덤한 선택과 교배를 통해 세대가 진화함에 따라 다양한 개체로 전역탐색을 수행하고 우수개체에 대한 보전 및 변이를 통해 국부탐색도 동시에 수행하는 감수분열 유전알고리즘(meiosis genetic algorithm, MGA)을 제시한다. 제안된 감수분열 유전알고리즘을 이용하여 퍼지모델의 최적구조와 파라미터를 탐색하고 gradient descent 알고리즘으로 파라미터를 정밀 조정하여 모델 오차를 최소화하는 퍼지모델을 구성하고 Box-Jenkins[3]의 가스로 데이터에 대한 시뮬레이션을 통해 성능평가 및 적용 가능성을 검증한다.

2. 감수분열 유전알고리즘

2.1 생식체의 생성 유형

생식체의 생성유형을 그림 1과 같이 3 가지의 개체 감수분열 유형으로 나누어 나타낸다.

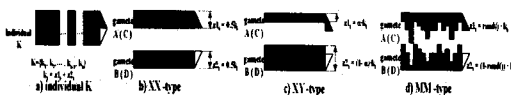


그림 1. 생식체의 생성 유형

여기서 XX, XY 유형은 외형적으로 감수분열되는 비율에 따라 생성되는 생식체를 기준으로 구분되고, MM 유형의 감수분열은 돌연변이(mutation)를 나타내며 인간에 있어서 염색체의 이상현상을 모델로 한 것이다. 개체 K는 염색체 k_i ($1 \leq i \leq n$)의 1차원 배열로 구성되어 있으며, 파라미터 α 는 개체에 대한 감수비율을, θ_i 는 개체를 구성하는 염색체 각각에 대한 감수비율을, rand()는 랜덤 값을 발생시키는 랜덤함수를 나타낸다. 개체 K가 파라미터 θ 만큼 변형된 개체를 K'라고 하면

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

$$\text{개체 } K' = \{k_1', k_2', \dots, k_n'\} \quad (2)$$

$$\text{파라미터 } \theta = \{\theta_1, \theta_2, \dots, \theta_n\} \quad (3)$$

하나의 개체가 분리되어 생성되는 2 개의 생식체가 생성되는데 다음과 같은 식으로 나타낼 수 있으며 파라미터의 비율에 따라 유형이 결정된다.

$$\text{생식체 } A = \{z1_1, z1_2, \dots, z1_n\} \quad (4)$$

$$\text{생식체 } B = \{z2_1, z2_2, \dots, z2_n\} \quad (5)$$

$$k_i' = z1_i + z2_i \quad (6)$$

2.1.1 XX 유형

하나의 개체를 1/2로 감수분열하여 2 개의 동일한 생식체를 생성하는 경우이며, 우수개체집단에 대해 XX 유형으로 개체를 나누어 각 염색체의 파라미터 θ_i 를 미세하게 랜덤 값으로 변화시키고 우수개체 풀에서 랜덤 선택 교배를 수행하여 최적 해에 접근할 수 있도록 한다. 최적 해에 접근함에 따라 염색체들에 가해지는 변화의 폭을 줄여 수렴하기에 적합하도록 하기 위해 새로운 최적 해의 발생유무를 기준으로 수렴에 따른 변화량을 결정하기 위한 파라미터로 i_{factor} 를 정의한다. 하나의 개체에서 생성되는 2 개의 생식체는 다음과 같이 나타낼 수 있다.

$$\alpha = 0.5 \quad (7)$$

$$z1_i = \alpha k_i(1 + \theta_i) \quad (8)$$

$$z2_i = (1 - \alpha)k_i(1 - \theta_i) \quad (9)$$

$$\theta_i = \text{rand}(0.5) / i_{factor} \quad (10)$$

$$i_{factor} = \begin{cases} 1 & ; \text{새로운 해가 발생} \\ i_{factor}++ & ; \text{새로운 해가 없음} \end{cases} \quad (11)$$

2.1.2 XY 유형

XY 유형 개체 분리는 염색체 각각에 대한 감수비율은 없고, 개체에 대한 감수비율만을 적용하여 생식체들이 생성된다. 개체내의 염색체들에 적용된 비율이 일률적인 특성을 지닌 생식체들간의 교배를 통하여 현재의 국소값이 있는 부근을 크게 이탈하는 수단으로 이용될 수 있으며, 최적해로 접근 시 수렴 속도를 증가시키는 효과가 있다.

$$\alpha = \text{rand}(1) \quad (12)$$

$$\theta_i = 0, 1 \leq i \leq n \quad (13)$$

$$z1_i = \alpha k_i \quad (14)$$

$$z2_i = (1 - \alpha)k_i \quad (15)$$

2.1.3 MM 유형

MM 유형은 개체내의 모든 염색체에 대해 랜덤하게 감수비율을 적용하면서 개체에 대해서도 랜덤한 감수비율을 적용하여 생식체를 생성한다. 이 생식체들의 랜덤 선택 교배에 의해 생성되는 개체는 공간적 이동 폭이 크므로 전역 탐색의 수단으로 효과를 나타낸다.

$$\alpha = \text{rand}(1) \quad (16)$$

$$w_i = \text{upper bound} - \text{lower bound} \quad (17)$$

$$-0.5w_i \leq \theta_i \leq 0.5w_i \quad (18)$$

$$z1_i = \alpha k_i(1 + \theta_i) \quad (19)$$

$$z2_i = (1 - \alpha)k_i(1 + \theta_i) \quad (20)$$

2.2 감수분열 유전알고리즘의 구성

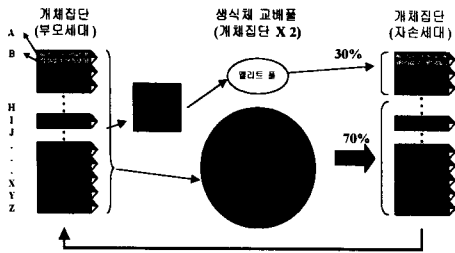


그림 2. 감수분열 유전알고리즘의 구성

2.2.1 개체 집단

개체집단은 세대가 진화함에 따라 정해진 개체집단의 크기를 갖는 부모세대(parent generation)와 자손세대(children generation)가 번갈아 존재한다. 현세대의 부모세대내의 개체들은 생식체로 분리되어 교배 풀에 전달되며 교배 풀에서 랜덤하게 선택되어 교배와 돌연변이에 의해 자손세대를 구성한다. 초기 개체집단은 랜덤한 값으로 구성되며 이 개체집단의 각 개체에 대한 적합도를 평가하여 최우수 개체를 선정하고 자손세대의 엘리트 집단을 구성하는 종자로 최우수 개체 버퍼에 저장된다. 자손세대를 구성하기 위해 부모세대내의 개체들은 XX 유형을 기준으로 한 최우수개체 변형그룹과 XY 유형과 MM 유형의 감수분열을 적용할 개체그룹으로 나뉘어 생식체로 분리되고 교배 풀에 전달되며 랜덤하게 선택되어 새로운 개체집단인 자손세대를 구성하게 된다. 개체집단과는 별도로 최우수개체를 저장하는 버퍼를 두어 현 세대의 최우수개체 값을 보존한다.

2.2.2 생식체 교배 풀

생식체 교배 풀은 엘리트 풀과 동적 풀로 구성된다. 엘리트 풀은 부모세대의 최우수개체에 대해 XX 유형을 적용하여 생식체로 분리하고 분리된 생식체내의 염색체 각각을 미세하게 랜덤하게 변화시켜 생성된 생식체들의 저장소이다. 엘리트 풀에서 교배를 통해 생성된 개체는 자손세대 개체집단의 30%를 채우도록 한다.

동적 풀은 XX, XY, MM 유형으로 분리된 생식체들로 이루어져 있으며 교배 상대의 선택은 생태계의 자율적인 선택에 따라 랜덤하게 교배할 수 있도록 한다. 동적 풀에 존재하는 최우수 개체를 포함한 부모세대로부터 생식체들간의 랜덤교배에 의해 생성된 개체들은 전역 탐색이 국부탐색에 비하여 더 큰 비중을 갖도록 자손세대 개체집단의 70%를 채운다.

2.2.3 교배 풀에서 생성되는 개체유형

생식체들간의 교배결과 얻을 수 있는 새로운 개체들은 표 1과 같다. 개체내의 염색체 값이 제한범위를 벗어나는 경우에는 경계 값을 갖도록 한다.

표 1. 생식체간의 결합으로 생성되는 개체유형

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

3. 퍼지모델의 구조와 파라미터

그림 3과 같이 본 논문에서는 식(21)로 표현되는 Takagi-Sugeno형 퍼지모델을 이용하여 비선형 시계열 다입력 단출력 시스템을 모델링 한다.

$$R^i: \text{If } x_1 \text{ is } A_{1i} \text{ and } x_2 \text{ is } A_{2i}, \dots, \text{ and } x_m \text{ is } A_{mi} \text{ then } y_i = a_{0i} + a_{1i}x_1 + \dots + a_{mi}x_m \quad (21)$$

$$y^* = \frac{\sum_{i=1}^m w_i y_i}{\sum_{i=1}^m w_i}, \quad w_i = \text{MIN}_{j=1}^m A_{ji}(x_j) \quad (22)$$

여기서 R^i 는 i 번째 퍼지규칙, x_j ($j=1, \dots, m$)는 입력 변수, $A_{1i}, A_{2i}, \dots, A_{mi}$ 는 퍼지공간을 분할하는 종형, 사다리형, 삼각형 등의 소속함수, y_i 는 퍼지규칙 R^i 의 출력이다.

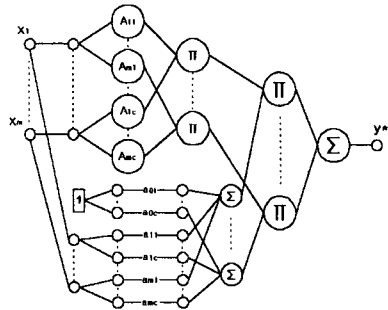


그림 3. Takagi-Sugeno형 퍼지모델

그림 4에 개체집단과 개체의 염색체 구성정보를 나타내며, 입력변수 정보는 이진 값으로 해당 입력변수의 유무 정보를 의미한다.

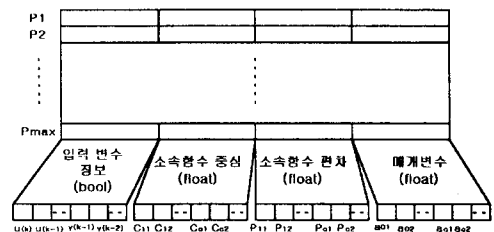


그림 4. 개체집단과 개체의 염색체 구성정보

4. 시뮬레이션 결과

Box-Jenkins의 가스로 데이터에 대하여 감수분열 유전알고리즘을 이용하여 모델 오차를 최소로 하는 Takagi-Sugeno형 퍼지모델의 구조와 파라미터를 동시에 탐색한다. 개체집단의 크기는 20, 진화세대수는

2000회를 수행하였으며, 그림 5에 모델 오차를 최소로 하는 경우 세대수의 증가에 따른 모델 오차의 변화를 보인다.

(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. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning", Addison-Wesley, 1989
 [3] G. E. P. Box and G. M. Jenkins, "Time Series Analysis, Forecasting and Control", San Francisco, CA: Holden Day, 1970
 [4] 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.
 [5] D. E. Goldberg, "The Design of Innovation: Lesson from Genetic Algorithms", Lessons for th Real World, IlliGAL Report No. 98004, February, 1998.

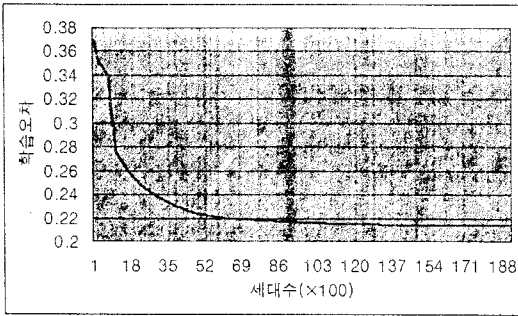


그림 5. 세대수의 증가에 따른 모델 오차의 변화

모델 오차를 더 줄이기 위해 감수분열 유전알고리즘에 의해 탐색된 퍼지모델의 파라미터를 gradient descent 알고리즘으로 정밀조종한다. 그림 6에 Box-Jenkins의 가스로 데이터에 대하여 감수분열 유전알고리즘과 gradient descent 알고리즘으로 최소학습오차를 갖도록 구성된 퍼지모델의 출력 값을 실제 값과 비교하여 나타낸다.

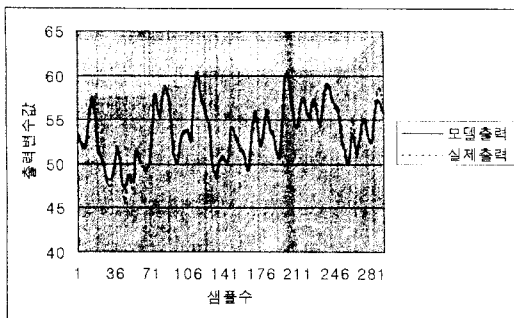


그림 6. Box-Jenkins 가스로 데이터의 실제 값과 동정된 퍼지모델의 출력 값

5. 결 론

기존의 유전알고리즘과 비교해 볼 때 감수분열 유전알고리즘의 특징은 ① 기존의 교배방식과 달리 교배점을 정하지 않으므로 파라미터간의 상관 관계에 따라 염색체의 위치를 적절하게 설정할 필요가 없고, ② 개체를 구성하는 전체 유전자에 교배와 돌연변이를 동시에 랜덤하게 적용할 수 있으므로 전역 탐색의 수행이 원활하다. 감수분해 유전알고리즘에 의해 모델 오차를 최소로 하는 퍼지모델의 구조와 파라미터의 탐색은 ① 엘리트 풀에서의 생식체의 작은 변화와 랜덤 교배에 의한 국부탐색과, ② 동적 풀을 구성하는 생식체의 돌연변이와 랜덤 교배에 의한 전역탐색으로 이루어져 효율적으로 수행된다. 그리고 감수분해 유전알고리즘으로 탐색된 퍼지모델의 파라미터를 초기 값으로 하여 gradient descent 알고리즘에 의해 최적 파라미터를 탐색하는 방안은 상호 보완적이고 국부탐색에 의해 모델 학습오차가 줄어들어 모델 정밀도가 향상된다. 본 논문에서는 제안된 기법을 Box-Jenkins 가스로 데이터의 퍼지모델에 적용하여 본 결과 대체적으로 만족스러운 학습 오차를 산출하였다.