

유전 알고리즘과 퍼지 추론 시스템의 합성

Fusion of Genetic Algorithms and Fuzzy Inference System

黃 熹 秀* · 吳 聖 權* · 禹 廣 芳**
(Hee-Soo Hwang · Sung-Kwun Oh · Kwang-Bang Woo)

Abstract - An approach to fuse the fuzzy inference system which is able to deal with imprecise and uncertain information and genetic algorithms which display the excellent robustness in complex optimization problems is presented in this paper. In order to combine genetic algorithms and fuzzy inference engine effectively the new reasoning method is suggested. The efficient identification method of fuzzy rules is proposed through the adjustment of search areas of genetic algorithms. The feasibility of the proposed approach is evaluated through simulation.

Key Words : Fuzzy Inference System(퍼지추론엔진), Genetic Algorithms(유전알고리즘)

1. 서 론

비선형적이고 다변수인 시스템을 대상으로한 퍼지 제어기와 퍼지 모델링 기법의 유용성은 이미 알려진 사실이며 이들은 퍼지 추론 시스템에 기초하고 있다. 퍼지 추론 시스템은 크게 퍼지 규칙과 추론 엔진으로 구성되어 있다. 퍼지 제어기나 퍼지 모델의 성능은 퍼지 규칙의 구성 방법에 의존하며 보다 좋은 성능을 위해서는 퍼지 규칙의 동정(identification)이 필연적이다. Sugeno[2]는 다

차원 추론 방식에 기초한 동정 기법을 제안했다. 이 방법은 퍼지 규칙의 수를 최소로 하면서도 상당히 정확성을 보임이 입증되었으나 입출력 변수의 증가에 따라 계산량의 급증이라는 어려움을 야기한다.

최근에는 신경 회로망의 학습 기능을 이용하여 퍼지 규칙을 연고자 하는 방향으로 연구가 진행되고 있다. Hopfield 신경 회로망은 그 특성상 제어 규칙의 후건부 파라미터 동정이라는 제한적인 문제에만 적용이 되므로[3] 다층 신경 회로망을 이용한 퍼지규칙 동정에 대한 연구가 보다 활발히 진행되고 있다[4, 5]. 그러나 이 방법 역시 퍼지 규칙을 구하기 위해 신경 회로망의 학습율과 시그

*正 會 員 : 延世大 大學院 電氣工學科 博士課程 修了
**正 會 員 : 延世大 工大 電氣工學科 教授 · 工博
接 受 日 字 : 1992年 2月 19日
1 次 修 正 : 1992年 7月 8日

모이드 함수의 기울기 조정이 필요하며 국부 수렴 역시 문제가 된다.

본 논문에서는 퍼지 규칙을 동정하기 위하여 유전 알고리즘을 사용한다. 유전 알고리즘은 자연 도태와 유전적 성질에 기초한 탐색 알고리즘으로 기존의 최적화 및 탐색 절차와는 다음과 같은 점에서 차이가 있다[1]. 첫째, 유전 알고리즘은 최적화 문제의 파라미터 집합을 알파벳으로 구성된 유한 길이의 문자열(string)로 코딩한다. 그 결과로 파라미터 영역의 연속성이나 미분치의 존재와 같은 제약을 받지 않는다. 둘째, 많은 최적화 방법에서 사용되는 점에서 점으로의 이동 방법은 다수 극점(multimodal)을 갖는 탐색 영역에서는 국부적인 극점을 찾을 위험이 있다. 반면에 유전 알고리즘은 점들의 집단(population of strings)으로부터 동작하여 다수의 최고점을 병렬로 올라가기 때문에 잘못된 최고점을 찾을 확률이 줄어든다. 셋째, 많은 탐색 기법이 효과적인 수행을 위해 보조적 정보를 필요로 하는데 비해 유전 알고리즘은 이러한 보조적인 정보를 필요로 않으며 단지 개별적인 문자열과 관련된 목적 함수값만을 필요로 한다. 넷째, 유전 알고리즘은 탐색 영역에 대한 보다 효율적인 탐색 방향을 유도하기 위해 확률적 천이 규칙을 사용한다. 이상의 4가지 차이점으로 인해 유전 알고리즘은 보다 강인하게 된다.

본 논문에서는 위에 언급된 것처럼 복잡한 문제에서 뛰어난 강인성을 발휘하는 유전 알고리즘과 부정확하고 불확실한 정보를 다룰 수 있는 퍼지 추론 시스템의 결합이 기술된다. 퍼지 추론 시스템은 크게 추론 엔진과 퍼지 규칙으로 구성되는데, 유전 알고리즘과의 효과적 결합을 위해 후건부가 비선형적인 새로운 추론 방식이 2절에서 제안되고 유전 알고리즘이 3절에서 기술되며 4절에서는 유전 알고리즘의 탐색 영역 조절을 통한 퍼지 규칙의 효율적 동정 수법이 제안된다. 5절에서는 Box와 Jenkins가 사용한 가스로의 시계열 데이터[6] 및 Sugeno가 사용한 수치적 데이터[11]를 사용하여 본 논문에서 제안된 방법의 타당성 및 동정의 정확도 등이 컴퓨터 시뮬레이션을 통해 비교 검토된다.

2. 퍼지 추론 시스템

본 논문에서는 퍼지 추론 시스템과 유전 알고리즘의 합성을 용이하게 하기 위하여 전건부는 unimodal형의 퍼지 변수와 후건부는 시그모이드 함수 형태의 비선형식으로 구성된 퍼지 임프리카

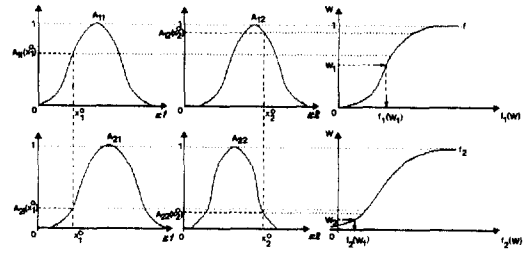


그림 1 유전 알고리즘을 위한 퍼지 추론 시스템
Fig. 1 Fuzzy inference system for genetic algorithms

이션(implication)을 제안한다. 두개의 입력 변수 x_1, x_2 와 하나의 출력 변수 y 에 의해 2개의 퍼지 임프리케이션으로 구성된 다음과 같은 퍼지 규칙을 고려하자.

$$R^1 : \text{If } x_1 \text{ is } A_{11} \text{ and } x_2 \text{ is } A_{12} \text{ then } y=f_1(W_1).$$

$$R^2 : \text{If } x_1 \text{ is } A_{21} \text{ and } x_2 \text{ is } A_{22} \text{ then } y=f_2(W_2).$$

입력 (x_1^0, x_2^0) 에 대한 퍼지 추론치 y^* 의 계산 과정은 그림 1과 같다. 그림 1에서 A_{ij} 는 전건부 퍼지 변수로 식(1)과 같이 기술되고 W_i 는 i 번째 퍼지 임프리케이션(implication)의 전건부 적합도로 식(2)에 의해 계산된다. 후건부의 f_i 는 전건부 적합도 W_i 에 의해 출력이 결정되는 비선형식으로 식(3)과 같이 정의된다. 그러나 $W_i=1$ 인 경우에 대해서는 식(3)로부터 $f_i(W_i)$ 를 구할 수 없게 되는데, $W_i=1$ 인 경우가 크리슘(crisp)한 경우로(퍼지 변수가 unimodal형인 경우 오직 하나의 정점만이 존재) x_1, x_2 의 논리영역 모든 조합 수를 고려한다면 그 경우가 일어날 가능성은 극히 희박하지만 이 경우가 발생하면 $W_i=1.0-\epsilon(\epsilon=10^{-6})$ 로 처리한다. 이와같이 퍼지 임프리케이션의 후건부에 비선형식을 도입함으로써 전건부 적합도 W_i 에 대한 유일한 $f_i(W_i)$ 를 얻을 수 있으므로 기존의 방식에서와 같은 적분 연산을 필요로 하지 않으면서도 퍼지 임프리케이션의 후건부가 선형적인 경우[2]보다 더 많은 비선형성을 내포할 수 있다.

$$A_{ij} = \frac{1.0}{1.0 + a_{ij} \cdot (x_j - b_{ij})^2} \quad (1)$$

$$W_i = A_{i1}(x_1^0) \cdot A_{i2}(x_2^0) \quad (2)$$

$$f_i(W_i) = e_i - \frac{1}{d_i} \cdot \ln \left[\frac{1.0 - W_i}{c_i \cdot W_i} \right] \quad (3)$$

윗 식에서 i 는 i 번째 퍼지 임프리케이션을 나타내고 j 는 j 번째 입력 변수를 표시한다. 입력(x_1^0, x_2^0)

에 대해 두개의 퍼지 임프리케이션에 대한 전건부 퍼지 변수 값 $A_{11}(x_1^0)$, $A_{12}(x_2^0)$, $A_{21}(x_1^0)$, $A_{22}(x_2^0)$ 이 식(1)에 의해 결정되면 퍼지 임프리케이션의 전건부 적합도 W_1 , W_2 가 식(2)에 의해 계산되고, 이 값에 의해 식(3)으로부터 $f_1(W_1)$, $f_2(W_2)$ 가 얻어지면 최종적인 퍼지 추론 값 y^* 는 식(4)와 같이 계산된다.

$$y^* = \frac{W_1 \cdot f_1(W_1) + W_2 \cdot f_2(W_2)}{W_1 + W_2} \quad (4)$$

3. 유전 알고리즘

인공 유전 시스템의 문자열(string)은 생물학적 시스템에서의 염색체(chromosomes)에 해당하는 것으로 자연 시스템에서 하나 또는 그 이상의 염색체가 어떤 유기체의 구성과 동작을 위해 전체 유전적 규범을 형성하도록 결합되며 결합된 전체 유전적 패키지가 환경과의 상호 작용을 통해 유기체를 형성하듯이 인공 유전 시스템에서는 문자열들이 결합되어 구조(structure)를 이루며 그 구조는 목적 함수에 맞도록 적용되어 파라미터 집합으로 디코딩된다. 인공 유전 시스템의 핵심인 유전 알고리즘은 코딩, 집단으로부터의 탐색, 보조 정보의 불필요, 랜덤화된 연산자들의 직접 사용이라는 특성으로 인해 복잡한 최적화 문제나 탐색에 있어서 뛰어난 강인성을 발휘하고 있다.

초기 집단으로부터 시간이 경과함에 따라 개선된 집단을 발생시키는 유전 알고리즘은 유전 연산자에 의해 구성이 되는데 많은 실용적인 문제에 있어서 좋은 결과를 낳는 단순한 유전 알고리즘은 재생(reproduction), 교배(crossover), 돌연변이(mutation) 연산자로 구성되어 있다.

3.1 재생 연산자

재생은 적합도 함수값에 의해서 개별적인 문자열이 복제되는 과정이다. 이 연산자는 문자열 특성 가운데 다윈의 적자 생존 즉 자연 도태의 인공판으로 볼 수 있으며 이 연산자를 통해 보다 높은 적합도를 갖는 문자열이 이후의 세대에서 더 많은 수의 자손을 갖게된다. 일단 문자열이 재생을 위해 선택되어지면 문자열의 정확한 복제가 이루어지고 복제된 문자열은 다음의 유전 연산자 연산을 위해 임시적인 집단인 교배 풀(mating pool)안으로 삽입된다.

3.2 교배 연산자

재생에 이어 교배가 진행된다. 교배 풀에서 새

로이 재생된 문자열이 임의로 짝을 이루면 각 쌍의 문자열은 임의로 선택된 정수 위치 k ([1, L-1] 사이, L은 문자열 길이)에 대해 $k+1$ 과 L사이의 모든 문자열이 교환된 새로운 문자열로 형성된다. 예를들어 다음의 문자열 A_1 과 A_2 를 고려하자.

$$A_1 = 0 \ 1 \ 1 \ 0 : 1 \ 1$$

$$A_2 = 1 \ 1 \ 0 \ 0 : 1 \ 0$$

1과 4(L=5) 사이의 임의의 수인 $k=4$ 를 선택했다고 하면 교배 연산 결과 새로운 문자열 A_1' 과 A_2' 가 발생한다.

$$A_1' = 0 \ 1 \ 1 \ 0 \ 0$$

$$A_2' = 1 \ 1 \ 0 \ 0 \ 1$$

3.3 돌연변이 연산자

돌연변이율은 자연 집단에서도 대단히 적기 때문에 돌연변이는 유전 알고리즘에서 부차적인 메카니즘으로 간주된다. 비록 재생과 교배가 효과적으로 탐색할 수 있지만 그들만으로는 잠정적으로 유용한 유전 물질(특별한 위치에서의 0이나 1)을 잃어버리게 될 수도 있기 때문에 돌연변이가 필요하다. 2진 코드상에서 돌연변이는 어떤 문자열의 위치에서 1을 0으로, 0을 1로 바꾸는 것을 의미한다. 본질적으로 돌연변이는 문자열 공간을 통한 랜덤 작업으로, 재생 및 교배와 함께 사용될 때 중요한 개념(notion)의 조기 손실을 막아준다.

유전 알고리즘은 파라미터 집합을 나타내는 코딩된 문자열들의 집단에서 동작하며 재생, 교배와 돌연변이는 새로운 문자열 집단을 만들어 내기 위해 연속적으로 문자열 집단에 적용된다. 이들 연산자는 랜덤수 발생, 문자열 복제와 부분 문자열 교환이라는 단순한 형태이지만 그 단순성에도 불구하고 뛰어난 탐색 기능을 발휘하는데 이는 스키마 이론[1]에 의하여 설명된다. 그 이론에 의하면 재생 연산자는 관측된 최적 유사 패턴에 존재하는 특별한 스키마(어떤 문자열 위치에서 유사성을 갖는 문자열의 부분 집합을 기술하는 유사 형판)가 보다 많은 표본을 다음 세대에 제공하도록 하고 교배 연산자는 매우 긴 길이를 갖는 스키마가 다른 문자열을 갖는 스키마와 결합되도록 하여 짧은 길이를 갖는 스키마가 관측된 최적의 것에 많은 표본을 제공하도록 한다. 즉 교배와 돌연변이에 의하여 길고 고차인 스키마의 분열에도 불구하고 유전 알고리즘은 비교적 적은 양의 문자열을 처리하는 동안 본질적으로 많은 양의 스키마를 처리하게 된다. 이러한 내재적 병렬성이 유전 알고리즘을 다양한 최적화 문제에 효과적으로 적용되

```

Define Metric:
max generation      maxgen,      population size      popsize.
probabilities       Pc, Pm,      length of individuals  lchrom
Initialization:
Random Generation of inial population P(0) of popsize:
Evaluate fitness of individuals in P(0):
Produce New Generation:
For(gen=1 to maxgen) OR until convergence do /* some end criterion */
  Reproduce individuals with reproduction probability, Pi in P(gen-1);
  Allocate reproduced individuals into mating pool:
  For(i=1 to popsize) do
    Randomly select 2 individuals as parents from mating pool:
    Crossover the parents with crossover probability, Pc:
    Randomly choose an individual from the offspring;
    Mutate the individual with mutation probability, Pm.
  endfor
  Evaluate the fitness of individuals in P(gen).
  Retain the fittest(gen).
endfor
    
```

그림 2 유전 알고리즘의 순차
Fig. 2 Sequence of genetic algorithms

도록 한다. 3개의 유전 연산자를 사용한 알고리즘 순차가 그림 2에 보여진다.

그림 2에서 유전 알고리즘은 초기 파라미터 집단(population)을 임의로 발생하고 적합도에 비례하는 수의 문자열을 재생 연산자에 의해 교배 풀(mating pool)안으로 복제한다. 교배 풀안의 재생된 문자열은 교배 연산자와 돌연변이 연산자에 의해 새로운 문자열을 발생하고 이 과정이 진행됨에 따라 집단은 보다 적합한 문자열로 구성된다.

4. 유전 알고리즘과 퍼지 추론 시스템의 합성

본 절에서는 유전 알고리즘과 퍼지 추론 시스템의 합성 기법이 제시되는데, 이는 2절에서 제안된 퍼지 추론 방식을 기초로 퍼지 규칙을 동정하기 위해서다. 퍼지 규칙의 동정은 규칙을 구성하고 있는 각 퍼지 임프리케이션의 전건부 퍼지 변수 및 후건부의 비선형식을 정의하는 파라미터를 동정하는 일이 된다. 2절의 그림 1에서 제시된 2개의 퍼지 임프리케이션을 고려한다면 4개의 전건부 및 2개의 후건부 비선형식을 동정하는 것으로 전건부 퍼지 변수는 식(1)에 의해 후건부 비선형식은 식(3)에 의해 정의되었으므로 결국 전건부 변수의 파라미터인 a_{ij}, b_i 와 후건부 파라미터인 c_i, d_i, e_i 를 동정하는 것이다. 본 절에서는 유전 알고리즘을 사용하여 이들 파라미터값을 효과적으로 찾아내기 위해서 적합도 함수 형태로 목적 함수의 변환이 이루어지며, 유전 알고리즘의 바람직하지 못한 조기 수렴을 방지하기 위한 적합도 함수값의 환산이 사용되고, 멀티 파라미터 탐색을 위해 사용된 코딩 방식과 유전 알고리즘의 탐색 영역 조정을 통한 최적한 값의 보다 빠른 탐색 기법이 제시된다.

4.1 목적 함수를 적합도(fitness)함수 형태로 변환

퍼지 규칙의 동정은 궁극적으로 식(5)의 목적 함수값 E 를 최소화시키는 퍼지 변수와 비선형식의 파라미터를 찾아내는 것이다. 반면에 유전 알고리즘에서는 적합도 함수를 최대화시키도록 진화가 이루어지므로 최소화 문제를 최대화 문제로 변경하여야 한다. 이들의 이중성(duality)이 이미 널리 알려져 있듯이 최소화 문제를 유전 알고리즘에 맞도록 최대화 문제로 쉽게 변환할 수 있다. 그 변환 방법에는 여러 가지가 있지만 본 논문에서는 식(6)에서와 같이 목적 함수의 역수를 적합도 함수로 결정한다.

$$E = \frac{1.0}{N} \cdot \sum_{i=1}^N \frac{|y_i^0 - y_i^*|}{y_i^0} \cdot 100(\%) \quad (5)$$

$$f = 1.0 / E \quad (6)$$

여기서 N 은 데이터 수, y_i^0 는 데이터 출력값이고 y_i^* 는 퍼지 추론값이다.

4.2 적합도 함수값의 선형 환산

소수 집단으로 구성되는 유전 알고리즘에서 만족할만한 결과를 얻기 위해서는 집단내의 경쟁 수준을 조정하는 것이 필요하며 이는 적합도 환산을 통해 복제되는 문자열 수를 조정함으로써 이루어진다. 적합도 환산 절차에는 여러가지가 있지만 본 논문에서는 식(7)과 같은 선형 환산[1]을 사용하여 식(7)은 식(8)을 만족시켜야 한다.

$$f' = af + b \quad (7)$$

여기서 f 는 본래 적합도, f' 는 f 의 환산된 적합도이고 계수 a, b 는 평균 환산 적합도 f'_{avg} 와 본래의 평균 적합도 f_{avg} 가 같도록 결정된다. 이는 각 집단내의 평균적인 구성원이 다음 세대에 단지 하나의 자손만을 낳을 수 있음을 의미한다.

$$f'_{max} = C_{mut} * f_{avg} \quad (8)$$

여기서 C_{mut} 는 집단내의 최적 구성원에 대한 복제수로 본 논문에서는 2.0이다. 이러한 적합도 환산은 유전 알고리즘의 초기 수행에서 몇개의 특별한 개체들의 적합도를 낮추고 낮은 적합도를 갖는 개체들은 적합도를 증가시킴으로써 바람직하지 않은 조기수렴을 방지할 수 있으나 유전 알고리즘이 상당히 진화되어 집단의 적합도 평균치의 최대치가 상대적으로 비슷해지면 낮은 값을 갖는 적합도가 환산된 후에 음이 되는 경우가 발생하는데 이를 방지하기 위하여 본래의 적합도 최소치 f_{min} 을

환산된 적합도 $f'_{min}=0$ 에 대응하도록 선형 환산한다.

4.3 멀티 파라미타 코딩

멀티 파라미타로 구성된 최적화 문제를 푸는데 성공적으로 사용된 한 가지 코딩 방식은 다수의 단일 파라미타를 직렬로 연결하는 코딩 방식이다. 이를 위해 디코드된 양의 정수 $[0, 2^L]$ 를 규정된 구간 $[P_{min}, P_{max}]$ 로 선형적으로 사상시킨다. 그림 3에서 단일 파라미타 P 는 이진수로 된 문자열(문자열 길이: $L=4$)로 코딩되는데 코드 0000은 파라미타 P 가 갖을 수 있는 최소값 P_{min} 으로 1111은 최대값 P_{max} 로 디코드되며 0000과 1111사이의 코드는 P_{min} 과 P_{max} 사이의 값으로 선형적으로 디코드된다. 2절에서 제시된 2개의 퍼지 임프리케이션을 구성하는 4개의 전건부 퍼지 변수를 정의하는 8개의 전건부 파라미타와 2개의 후건부 비선형식을 정의하는 6개의 파라미타를 동정하기 위해서는 단일 파라미타 코드를 필요한 수(14개) 만큼 직렬로 연결하고 각각의 파라미타는 자신의 코드 길이(L_i)와 최대값인 $max_{a_i}, max_{b_i}, max_{c_i}, max_{d_i}, max_{e_i}$ 와 최소 값인 $min_{a_i}, min_{b_i}, min_{c_i}, min_{d_i}, min_{e_i}$ 을 갖는다. 따라서 구하고자 하는 파라미타 값은 문자열중 그 파라미타에 해당되는 코드를 지정된 파라미타 최대값과 최소값의 사이 값으로 선형적으로 디코드된다.

4.4 유전 알고리즘의 탐색 영역 조정

유전 알고리즘은 여러 방향에서 탐색을 통해 언덕을 올라감으로 국부 수렴의 문제가 발생하지 않

단일 파라미타 P(코드 길이 L=4)	
0 0 0 0	-> P_{min}
1 1 1 1	-> P_{max}

멀티 파라미타				
a_{11}	b_{11}	d_2	e_2
L_1	L_2	L_{13}	L_{14}
$min_{a_{11}}$	$min_{b_{11}}$	min_{d_2}	min_{e_2}
$max_{a_{11}}$	$max_{b_{11}}$	max_{d_2}	max_{e_3}
000...001	011...010	110...100	111...111

그림 3 단일 파라미타와 멀티 파라미타의 코딩
Fig. 3 Coding of single parameter and multiparameters

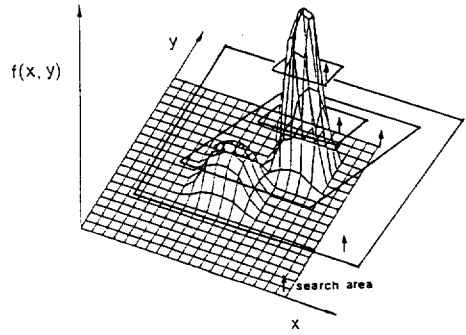


그림 4 유전 알고리즘의 탐색 영역 조정
Fig. 4 Adjustment of search area of genetic algorithms

```

Define Metric:
max generation      maxgen,      population size     popsize,
probabilities       Pc, Pm,      no. of parameters  nparm,
length of chromosome(individuals) lchrom,
length of strings in parameters  lparam;

Initialization:
Initialize min and max values of parameters: /* Decide search area */
Random Generation of initial population P(0) of popsize:
Evaluate fitness of individuals in P(0):
For(gen=1 to maxgen) /* Produce New Generation */
  Reproduce individuals with reproduction probability, Pr in P(gen-1):
  Allocate reproduced individuals into mating pool:
  For(i=1 to popsize) do
    Randomly select 2 individuals as parents from mating pool:
    Crossover the parents with crossover probability, Pc:
    Randomly choose a site for mutation:
    Mutate the individual with mutation probability, Pm:
  endfor
  while(some adjusting criterion) /* Adjustment of Search Area */
    Decode individuals:
    if(no. of decoded individuals)
      Diminish max values and increase min values of parameters:
      Reduce the adjusting criterion:
    do:
      Evaluate the fitness of individuals in P(gen):
      Scale the fitness of individuals in P(gen):
      Retain the fittest(gen):
  endfor
  
```

그림 5 탐색 영역 조정에 의한 유전 알고리즘의 순차
Fig. 5 Sequency of genetic algorithms by adjustmet of search area

지만 최적한 값을 보증하지 못하며 적합도 함수를 극대화 시키기 위한 파라미타 수의 증가는 유전 알고리즘의 탐색 영역을 증가시키고 결과적으로는 언덕의 정점값 또는 정점 근방의 값을 찾는데 많은 시간이 걸리게 만든다. 이러한 문제점을 해결하는 한 수단으로 탐색영역 축소를 통한 유전 알고리즘의 수렴성을 개선하는 방법이 제시되는데 이는 그림 4에서 점차 줄어드는 사각형의 면적이 유전 알고리즘의 탐색 영역에 해당하며 마지막 단계에서는 최고치 언덕의 정점 근처에서 언덕을 오르도록 함으로써 보다 빠른 시간내에 정점의 근방

에 오르도록 하는 방법이다.

그림 5에 탐색 영역의 조정을 통한 유전 알고리즘의 순차가 보여지는데 여기서 유전 알고리즘의 기본 순차는 그림 2에 보여진 것과 동일하며, 다만 유전 알고리즘이 진화되어 집단내의 적합도 함수 최대값이 α 보다 큰 개체(문자열)가 m 개 발생하면 그 m 개의 문자열로부터 파라미터 값들을 디코드하고 이 값에 기초하여 파라미터의 최소값을 증가시키고 최대값은 감소시킴으로써 유전 알고리즘의 탐색 영역을 점차 줄여나가는 방법이 추가된다. 탐색 영역이 조정되어 줄어들수록 α 는 커지고 m 은 작아진다.

5. 시뮬레이션

본 절에서는 이상에서 제시된 기법의 타당성 및 유용성을 비교 평가하기 위하여 시스템 입출력 데이터로부터 시스템을 기술할 수 있는 퍼지 규칙을 동정한다.

5.1 가스로 데이타인 경우

본 절에서는 Box와 Jenkins[6]의 가스로 시계열 입출력 데이터를 사용하여 가스로의 연소된 이산화탄소의 농도를 퍼지 모델링 한다. 퍼지 모델링을 위한 입력 변수는 $u(t-4)$ 와 $y(t-1)$ 이고 출력 변수는 $y(t)$ 이다. 여기서 $u(t)$ 는 가스 흐름율이고 $y(t)$ 는 연소된 이산화탄소의 농도이다. 본 논문에서는 총 299쌍의 데이터[6] 가운데 퍼지 규칙을 동정하기 위하여 144개를 사용하고 동정된 퍼지 규칙을 평가하기 위하여 또다른 144의 데이터를 사용한다. 동정용 데이터를 사용하여 동정할

그림 6에 보여진 퍼지 규칙은 6개의 퍼지 임프리케이션으로 구성되며 입력 변수 $u(t-4)$ 는 3개의 퍼지 변수 N (Negative), Z (Zero)와 P (Positive)로 입력변수 $y(t-1)$ 은 2개의 퍼지 변수 S (Small)과 B (Big)으로 출력 변수 $y(t)$ 는 6개의 비선형 방정식인 f_1, \dots, f_6 으로 구성된다.

위 그림에서 식(1)에 의해 정의된 전건부 퍼지 변수 N, Z, P, S, B 는 동정해야할 파라미터가 각 퍼지 변수에 대해 a, b 이므로 10개이고 식(3)에 의해 정의된 후건부 비선형식은 동정해야할 파라미터가 각 비선형식에 대해 c, d, e 이므로 18개이다. 유전 알고리즘에 의해 파라미터를 동정하기 위하여 그림 7에 보여진 것처럼 총 28개의 파라미

R1	if $u(t-4)$ is N & $y(t-1)$ is S	then $y(t) = f_1(W_1)$
R2	if $u(t-4)$ is N & $y(t-1)$ is B	then $y(t) = f_2(W_2)$
R3	if $u(t-4)$ is Z & $y(t-1)$ is S	then $y(t) = f_3(W_3)$
R4	if $u(t-4)$ is Z & $y(t-1)$ is B	then $y(t) = f_4(W_4)$
R5	if $u(t-4)$ is P & $y(t-1)$ is S	then $y(t) = f_5(W_5)$
R6	if $u(t-4)$ is P & $y(t-1)$ is B	then $y(t) = f_6(W_6)$

그림 6 가스로의 이산화탄소 농도를 기술하는 퍼지 규칙

Fig. 6 Fuzzy rules to describe CO₂ concentration of gas furnace

28개의 파라미터(npara=28)											
N		Z		P		f _s			f ₆		
a	b	a	b	a	d	c	d	e	c	d	e
1000010011	0001110101					1100010101	0001111000				
코드길이 (lpara=10)											
문자열 길이(lchro=280)											

그림 7 멀티 파라미터 동정을 위한 코딩

Fig. 7 Coding for multiparameters indetification

표 1 동정된 파라미터 값

Table 1 Values of identified parameters

전건부 파라미터									
N		Z		P		S		B	
a	b	a	b	a	b	a	b	a	b
5.622	-0.81	25.35	-0.39	4.924	1.090	84.06	0.546	62.11	1.356
후건부 파라미터									
	f ₁	f ₂	f ₃	f ₄	f ₅	f ₆			
c	40.6700	50.1827	91.1817	60.9488	67.3203	61.1003			
d	26.2371	-44.267	-63.267	-39.015	-63.149	-68.124			
e	47.7133	68.0462	47.1573	62.6879	43.4825	58.3204			

다가 질렬로 연결되며 이들은 총 280개의 이진수로 구성된 문자열로부터 디코드화 된다. 유전 알고리즘이 진화됨에 따라 각각의 파라미터는 지정된 위치에 있는 10개의 코드로부터 그 파라미터의 최소값 및 최대값의 범위내로 선형적으로 디코드 된다. 본 논문에서 사용된 유전 알고리즘을 위한 파라미터 값들은 진화 세대 수(maxgen) 300, 집단의 크기(popsiz) 100, 교배 확률(P_c) 0.6, 돌연변이 확률(P_m) 0.0333이다.

그림 8에 유전 알고리즘의 진화 과정에 있어서 탐색 영역이 조정되지 않은 경우와 조정된 경우에 대하여 집단내의 최대 적합도를 갖는 문자열의 디코드된 파라미터 값으로부터 식(5)에 의해 계산된 E 가 보이며 그림에서 세대 106($\alpha=1.18, m=12$), 117($\alpha=1.33, m=7$), 159($\alpha=1.75, m=5$)와 163($\alpha=1.88, m=3$)에서 탐색 영역이 조정된 경우가 더 낮은 E 값(더 높은 적합도 함수값)으로 수렴함을 볼 수 있다.

표1에 정규화된 입력값에 대한 동정된 파라미터 값이 보여지고 동정용 데이터를 사용해서 식(9)에 의해 계산된 값이 동일한 데이터를 사용하여 다른 기법에 의해 동정된 퍼지 모델과의 비교를 위해 표2에 보여진다. 이 표에 의하면 본 논문에서 제시된 기법에 의해 동정된 퍼지 모델이 가장 적은 퍼지 규칙을 사용하여 가장 정확하게 동정되었음을 알 수 있다. 평가용 데이터를 사용하여 식(9)에 의해 계산된 평균 제곱 오차는 0.32로 표2에 보여진 다른 어떤 방법의 동정 오차 보다도 적음을 알 수 있다.

$$E = 1/N * \sum_{k=1}^N [y(k) - \hat{y}(k)]^2 \quad (9)$$

여기서 N 은 사용된 데이터 수, $y(k)$ 는 실제 데

표 2 가스로의 입출력 데이터를 사용한 퍼지 모델링 기법의 비교
Table 2 Comparison of modeling techniques using gas furnace data

모델링 기법	평균 제곱 오차	사용된 규칙 수
Tong's Model[7]	0.469	19
Pedrycz's Model[8]	0.776	20
Xu's Model[9]	0.328	25
Sugeno's Model[10]	0.355	6
Our Model	0.134	6

이타의 출력값이고 $\hat{y}(k)$ 는 동정된 모델로부터 추론된 값이다.

5.2 수치 데이터의 경우

본 절에서는 식(10)에 의해 기술된 시스템으로부터 추출된 40개의 데이터[11]를 사용한다. 이 데이터는 Sugeno가 제안한 퍼지 추론 및 동정 기법에 적용된 데이터로 본 논문에서 사용된 기법에 이 데이터를 적용함으로써 제안된 기법의 타당성 및 정확도를 비교 검토할 수 있다. 40개의 데이터 가운데 1~20까지의 데이터는 퍼지 규칙을 동정하기 위한 데이터이고 21~40까지는 동정된 퍼지 규칙을 평가하기 위한 데이터이다.

$$y = (1.0 + x_1^{0.5} + x_2^{-1} + x_3^{-1.5})^2 \quad (10)$$

동정용인 1~20까지의 데이터를 사용하여 동정할 그림 9에 보여진 퍼지 규칙은 8개의 퍼지 임프리케이션으로 구성되며 각 입력 변수는 2개의 퍼지 변수를 갖고 출력 변수는 8개의 비선형식을 갖는다.

전건부 퍼지 변수 S_1, S_2, S_3 및 B_1, B_2, B_3 에서 S 는 Small을 B 는 Big을 의미하며 숫자는 각기 입력 변수에 대응한다. 식(1)에 의해 정의된 이 퍼지 변수는 12개의 a, b 파라미터로 식(3)에 의해 정의된 비선형식 f_1, f_2, \dots, f_8 은 24개의 c, d, e 파라미터로 구성된다. 따라서 총 36개의 파라미터가 5-1절에 기술된 방식과 동일하게 유전 알고리즘에 의해 구해진다. 표 3에 정규화된 입력값에

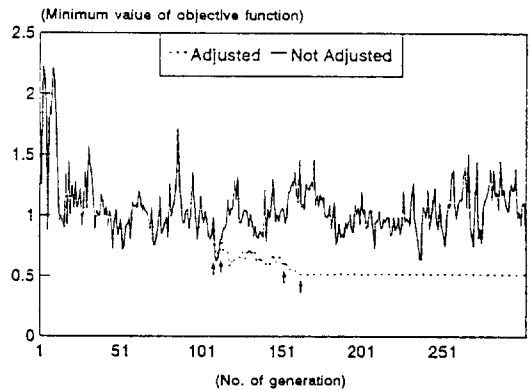


그림 8 진화 과정에서 집단내의 최적 문자열로부터 계산된 목적 함수값
Fig. 8 Objective function values calculated from the best string in evolving populations

대한 동정된 파라미터 값이 보여지고 표 4에는 동정용 데이터를 사용해서 식(5)에 의해 계산된 E_1 과 평가용 데이터를 사용해서 계산된 E_2 가 보여진다. 표 4에서 동정용 데이터를 사용하여 본 논문에서 제시된 기법에 의해 계산된 목적 함수값 E_1 은 Sugeno[11]가 제안한 두가지 모델의 E_1 값의 중간 정도인 반면 평가용 데이터를 사용하여 계산된 E_2 는 Sugeno의 것 보다 작다. 시스템을 퍼지 규칙에 의하여 모델링하거나 퍼지 제어를 수행할 경우 시스템이나 제어 데이터중의 일부분을 사용하여 퍼지 규칙을 동정한다. 즉 한정된 데이터를 사용하여 동정된 퍼지 규칙은 동정에 사용되지 않은 데이터에 대해서도 만족해야 한다. 따라서 본 논문에 제시된 기법에 의해 계산된 E_2 가 훨씬 작다는 것은 동정된 규칙이 동정에 사용되지 않은 데이터에 대해서도 잘 맞음을 의미하며 Sugeno가 제안한 방법이 전건부 및 후건부의 구조 및 파라미터 동정이라는 반복적인 절차에 의해서 이루어진 반면 본 논문에서는 가장 기본적인 구조(입력 변수가 2개의 퍼지 변수를 갖음)를 가정한 후 파라미터만을 동정한 것임을 고려한다면 제안된 기

R1	If x_1 is S_1 & x_2 is S_2 & x_3 is S_3 then $y = f_1(W_1)$
R2	If x_1 is S_1 & x_2 is S_2 & x_3 is B_3 then $y = f_2(W_2)$
R3	If x_1 is S_1 & x_2 is B_2 & x_3 is S_3 then $y = f_3(W_3)$
R4	If x_1 is S_1 & x_2 is B_2 & x_3 is B_3 then $y = f_4(W_4)$
R5	If x_1 is B_1 & x_2 is S_2 & x_3 is S_3 then $y = f_5(W_5)$
R6	If x_1 is B_1 & x_2 is S_2 & x_3 is B_3 then $y = f_6(W_6)$
R7	If x_1 is B_1 & x_2 is B_2 & x_3 is S_3 then $y = f_7(W_7)$
R8	If x_1 is B_1 & x_2 is B_2 & x_3 is B_3 then $y = f_8(W_8)$

그림 9 시스템을 기술하는 퍼지 규칙
Fig. 9 Fuzzy rules to describe the system

표 3 동정된 파라미터 값
Table 3 Values of identified parameters

전건부 파라미터											
S 1		B 1		S 2		B 2		S 3		B 3	
a	b	a	b	a	b	a	b	a	b	a	b
59.13	0.503	82.87	0.951	60.78	-0.11	23.73	0.893	87.39	0.144	8.215	0.573
후건부 파라미터											
	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8			
c	37.531	17.829	54.106	4.464	3.4234	29.711	51.284	2.669			
d	-5.91	-6.532	-14.24	-28.59	-1.959	-43.0	-0.708	-5.162			
e	20.217	10.497	11.105	3.8178	35.51	22.878	30.547	12.663			

표 4 목적 함수 값
Table 4 Objective function values

	E_1	E_2
Our Model	1.26	1.36
Sugeno's Model 1	1.5	2.1
Sugeno's Model 2	1.1	3.6

법이 간단하면서도 우수하다고 볼 수 있다.

6. 결 론

본 논문에서는 불확실하고 부정확한 정보를 처리할 수 있는 퍼지 추론 시스템과 복잡한 최적화 문제에서 강인성을 발휘하는 유전 알고리즘을 결합하는 방법이 제시되었다. 이를 위해 유전 알고리즘과의 효과적 결합을 위한 새로운 추론 방식이 사용되고 유전 알고리즘의 탐색 영역 조절을 통해 퍼지 규칙을 효율적으로 동정하는 방법이 제시된다. 가스로의 시계열 입출력 데이터 및 수치적 데이터를 사용하여 제안된 방법의 타당성 및 정확성이 비교 검토되었고, 그 결과 제안된 기법이 간단하면서도 정확함이 입증되었다. 새로운 추론 방식에서 퍼지 임프리케이션의 후건부가 시그모이드 함수의 역함수를 사용하고 있기 때문에 전건부 적합도가 1인 경우의 처리에 문제가 있으므로 이에 대한 해결 방안과 온라인 파라미터 동정에 있어서 유전 알고리즘의 수렴성을 증가시키기 위한 탐색 영역의 동적 조절에 대한 연구가 이루어져야 할

것 같다.

본 연구의 일부는 한국과학재단 지정 제어
계측신기술 연구센터(서울대학교)의 연구비
지원에 의해 수행되었습니다.

참 고 문 헌

- [1] D.E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning," 1989.
- [2] T. Takagi, and M. Sugeno, "Fuzzy Identification of Systems and Its Applications to Modeling and Control," IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-15, No. 1, pp. 116~132, 1985.
- [3] H. Nomura, et al., "Self-Tuning Method of Fuzzy Reasoning by Hopfield Neural Network," 5th Fuzzy System Symposium, pp. 177~181, KOBE, June 2~3, 1989.
- [4] M.M. Gupta and M.B. Gorzaczany, "Fuzzy Neuro-Computational Technique and Its Application to Modeling and Control," IFSA' 91 Brussels, Artificial Intelligence, pp. 46~49, 1991.
- [5] W. Pedrycz, "Neurocomputation in Relational Systems," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 13, No. 3, pp. 289~296, 1991.
- [6] G.E.P. Box and G.M. Jenkins, "Time Series Analysis, Forecasting and Control," Holden Day, San Francisco, 1970.
- [7] R.M. Tong, "The evaluation of fuzzy models derived from experimental data," Fuzzy Sets and Systems, Vol. 4, pp. 1~12, 1980.
- [8] W. Pedrycz, "An identification algorithm in fuzzy relational systems," Fuzzy Sets and Systems, Vol. 13, pp. 153~167, 1984.
- [9] Chen-Wei Xu and Yong-Zailu, "Fuzzy Model Identification and Self-Learning for Dynamic Systems," IEEE Trans. on Syst., Man, And Cybern., Vol. SMC-17, No. 4, pp. 683~689, 1987.
- [10] M. Sugeno and T. Yasukawa, "Linguistic Modeling Based on Numerical Data," IFSA' 91 Brussels, Computer, Management & System Science, pp. 264~267, 1991.
- [11]菅野道夫, "フアジイ制御," pp. 148~149, 日刊工業社, 1988.



황희수(黃熹秀)

1963년 6월 5일생. 1986년 연세대 공대 전기공학과 졸업. 1988년 동 대학원 전기공학과 졸업(석사). 현재 연세대 대학원 전기공학과 박사과정 수료.



오성권(吳聖權)

1957년 2월 13일생. 1981년 연세대 공대 전기공학과 졸업. 1983년 동 대학원 졸업. 1983~89년 금성산전(주) 연구소 선임연구원. 현재 연세대 대학원 전기공학과 박사과정 수료.



우광방(禹廣芳)

1934년 1월 25일생. 1957년 연세대 이공대 전기공학과 졸업. 1964년 미국 오레곤 대학원 졸업(공박). 1966년 미국 워싱턴대 조교수. 1971년 미국립 암연구소 책임연구원. 1979년 미국 존슨 홉킨스대 의대 연구원. 1980년 재미 한국과학기술자협회장. 현재 연세대 공대 전기공학과 교수. 당학회 평의원.