

DNA Coding Method를 이용한 퍼지 규칙의 획득

Acquisition of Fuzzy Rules Using DNA Coding Method

윤 성용*, 박 종규*, 오 성권*, 안 태천*

Sungyong Yun, Jonggyu Park, Sungkwun Oh, and Taechon Ahn

원광대학교 공과대학 전기전자공학부

*School of Electrical & Electronics Engineering,
College of Engineering, Wonkwang University.

ABSTRACT

본 논문에서는 생물학적인 DNA와 인위적인 DNA의 개발 메커니즘에 근거를 둔 새로운 DNA coding method를 소개한다. 이 방법은 DNA 유전자의 Redundancy 성질과 Overlapping 성질을 분석하여 구성한 알고리즘으로, DNA 자체의 특성 때문에 염색체의 길이를 자유자재로 변화시킬 수 있는 장점을 가지고 있다. 즉 삽입과 삭제의 과정을 쉽게 처리할 수 있다.

따라서 본 연구에서는 이 DNA coding method를 사용하여 시스템 제어에 필요한 지식을 표현하는 적당한 퍼지규칙을, 전건부와 후건부의 구조 및 매개변수 동조를 통하여, 획득하고, 이 규칙을 새로운 유전자 알고리즘을 이용하여 최적화시킬 수 있는 새로운 퍼지 규칙 획득 알고리즘을 개발한다. 이 방법의 유용성을 입증하기 위하여 물탱크 시스템에 적용한 결과 효과적으로 퍼지규칙을 획득한다.

I. 서론

일반적으로 퍼지 컨트롤러와 같은 전문가 시스템을 구축할 때 가장 어려운 과정중의 하나가 전문가가 갖고 있는 지식을 문자화하는 것이다. 더구나 제어의 경우 지식을 언어로 표현하는 것이 대체로 정확하지 못할 뿐 아니라 일관성이 결여될 수도 있다. 따라서 적응제어와 같이 제어 규칙을 스스로 생성한다면 또 그것으로 제어가 가능하다면 매우 유용하게 쓰일 수 있을 것이다.

퍼지 제어를 설계함에 있어서 가장 어려운 부분은 퍼지 규칙을 구성함에 있다 해도 지나치지 않다. 그만큼 퍼지 규칙은 퍼지 제어에 있어서 핵심이 되는 부분이기도 하고 까다로운 부분이기도 하다. 만약에 이러한 부분을 자동으로 생성할 수 있다면 이는 어떠한 이론이 가질 수 있는 가장 큰 장점 중

의 하나라고 볼 수 있다.

본 연구에서는 생물학적인 DNA와 인위적인 DNA의 개발 메커니즘에 근거를 둔 DNA coding method를 소개한다. 이 방법은 지식을 융통성 있게 표현할 수 있다는 장점이 있어 퍼지 규칙에 있어서 언어적인 표현이 가능하다.

본 논문에서는 이 DNA coding method를 사용하여 퍼지 제어에 필요한 규칙을 자동 생성하고 그 실용성을 검토하였다.

II. 생물학적 DNA

염기, 당 및 인산이 결합된 것을 뉴클레오타이드라 하고 다수의 뉴클레오타이드가 결

합되어 생성된 고분자 화합물이 곧 핵산이다. 핵산은 라이보 핵산(RNA)과 디옥시라이보 핵산(DNA)의 두 가지가 있는데 이 중 DNA는 염색체에만 존재한다. DNA는 2중 나선의 구조를 가지고 있으며 A, T, G, C의 염기간의 수소 결합으로 결합되어 있는데 염기의 A는 T와, G는 C와 결합한다. 이들 중 세 염기의 연속된 배열을 codon(유전 정보의 최소단위)이라고 부르는데 DNA의 차이는 바로 A, T, G, C의 4종류의 염기 배열의 차이이다. 총 64종류의 codon은 20종류의 아미노산이 된다. 모든 생물의 유전자가 DNA라는 점에서 유전자의 특징은 4종의 염기 배열의 방법이라고 할 수 있다. 따라서 DNA 분자 내의 염기의 배열을 유전정보, 또는 유전 암호라고 한다. 밝혀진 바로는 ATG는 읽기 시작의 암호이고, TGA는 읽기 끝을 나타내는 암호이다. DNA는 세포의 핵 내에 있고 단백질이 합성되는 곳은 세포질 내의 리보솜 위에서도이다. 따라서 핵 내의 DNA의 염기 배열 등을 복사해서 그것을 세포질의 리보솜으로 전달하는 물질이 존재하는데 이것이 메신저(전령) RNA이며 mRNA라고 쓴다. 먼저 핵 내의 DNA의 염기 배열에 의해서 mRNA가 합성된다. 이 때 DNA의 2중 나선의 일부가 풀려 한 가닥의 사슬 형태가 되고 그 한가닥 사슬의 염기 배열에 대응하는 mRNA가 합성된다. 이 경우 DNA의 A에는 U가 G에는 C가, T에는 A가, C에는 G가 대응한다. 이것을 전사라고 한다.

III. 인위적인 DNA Coding Method

그림 1과 2는 제안된 DNA가 진화하는 순서를 보여준다. GA는 문제에 대한 명확한 고안과 이들이 여분의 부분을 가지지 않게 하기 위한 방법으로 보통 사용되었다. 이 전통적인 coding법은 완료된 mRNA로 간주할 수 있다.

여기에서는 DNA로부터 퍼지 규칙 집합이 진화하는 방법을 제시한다. 염색체는 A, G, C, T의 네 개의 단위로 구성되어 있다. 번역된 codon들은 각각의 아미노산에 대응하고 인위적인 아미노산들은 각자의 의미를 가지며 유전자의 의미는 이 아미노산의 결합에 의해 규정된다. 아미노산은 입력 변수나 소속함수의 형태 또는 기타 다른 것들로 변환될 수 있다. 아미노산의 순서는 퍼지 규칙을 만든다. 여기에서 DNA 염색체는 수위 제어

를 위한 퍼지 규칙의 집합을 만든다.

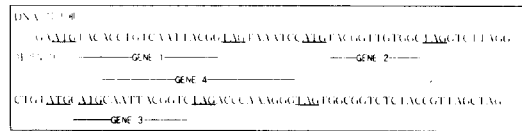


그림 1 염색체의 변환과 유전자의 중복

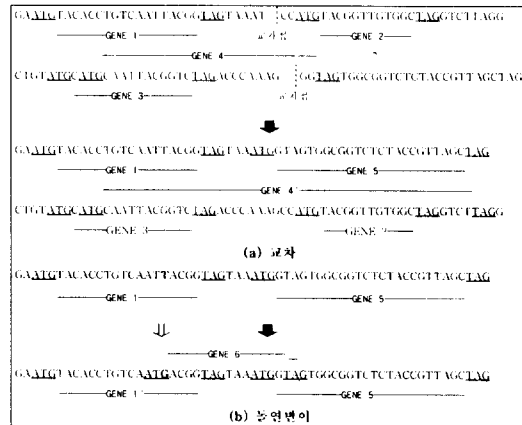


그림 2 교차와 돌연변이

그림 1은 DNA 염색체의 예와 변환 메커니즘을 보여준다. 이 그림에서 유전자는 시작 codon ATG로부터 시작하고 종료 codon TAG에서 끝나며 유전자 안의 codon들은 아미노산으로 변환된다: Ile, Tyr, Thr.... 각각의 아미노산들은 문제 해결을 위한 자신의 역할을 갖는다.

DNA의 진화 메커니즘에 의하여 시작점은 다른 codon으로부터 이동할 수 있고 어떤 유전자들은 다른 유전자와 중복되어 변환될 수 있다. 각각의 중복 유전자들은 중요한 역할을 수행한다. 그림 1은 이러한 중복된 표현을 보여준다. 이 그림에서 Gene 4는 Gene 3과 마찬가지로 DNA 염색체로부터 읽을 수 있다.

그림 2는 교차와 돌연변이의 예를 보여준다. 교차점으로부터 오른 편은 교체되고 새로운 Gene 4', Gene 5가 생성된다. 이 방법에서 교차점에 따라서 유전자는 완전히 바뀔 수 있다. 여기에서 교차점을 강제적으로 정하지는 않는다. 그림 2(b)는 돌연변이의 예를 보여준다. 그림에서 가리켜진 하나의 염기는 T에서 G로 바뀌었다. 결과에 따르면 Gene1은 Gene'1로 바뀌었다. 또한 이 교체에 의해 시작 codon ATG는 새로이 생성되고 새로운 Gene 6이 생성되었다.

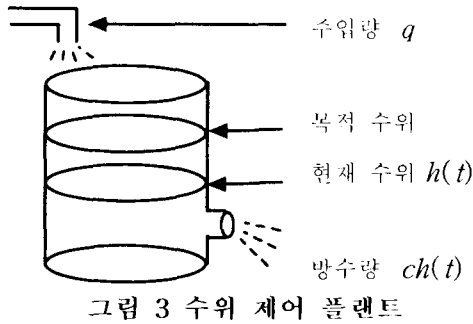
DNA coding method는 다음의 특징을 가진다.

- (a) 지식의 융통성 있는 표현이 가능하다.
- (b) coding에 여분과 중복이 있다.

- (c) 염색체의 길이가 가변적이다.
- (d) 교차점이 강제적으로 주어지지 않는다.

IV. 수위제어를 위한 DNA Coding Method

1. 모델링



- q = 유량
- t = 시간
- h = 수위
- T = 누수 계수 (= 시정수)
- H = 경과 시간 (= $t - t_0$)
- $h(t_0)$ = 1 단위 시간 전의 수위

그림 3은 제어할 플랜트를 나타내고 있고 이것에 대한 방정식은 다음과 같이 나타낼 수 있다.

$$h(t) = q + \frac{h(t_0) - q}{e^{\frac{H}{T}}}$$

2. 퍼지규칙 생성

수위 제어를 위한 퍼지규칙에는 두 개의 입력 변수와 하나의 출력 변수가 필요하다.

이를 위해서 먼저 Random하게 A, T, G, C의 염기배열을 1 개체당 300개씩 100 개체를 생성한다. 다음으로 생성된 개체로부터 규칙을 번역한다. 본 논문에서는 그림 2와 같이 시작 codon을 ATG 대신 A로 지정하고 종료 codon TGA이 없이, 그 후에 나오는 codon을 첫 번째 입력 변수, 다음 codon을 두 번째 입력 변수, 세 번째 codon을 출력 변수로 놓고 번역을 하였다. 이 때 입력 변수와 출력 변수는 다음과 같다.

- e = 현재 수위 - 목적 수위
 - Δe = 현재 수위 - 1 단위 시간 전의 수위
 - Δq = 수입량의 증감량
- 여기에서 e 와 Δe 는 두 개의 전건부 항이

되고 Δq 는 하나의 후건부 항이 된다.

표 1은 염색체의 배열을 조건에 맞는 언어적인 항으로 만들기 위한 전건부 번역표이고 후건부는 문자열에서 3문자를 4진수로 해석하고 10진수로 변환한 후 이하의 계산을 행하면 -6 ~ 6 사이의 실수 값이 얻어진다. (여기에서 A = 0, G = 1, C = 2, T = 3) 후건부 : 0.19047619×10 진수 - 6.0

표. 1 번역표

codon의 종류			멤버십 함수의 위치(位置)	
AA*	AG*	AC*	→	NB
AT*	GA*	GG*	→	NS
GC*	GT*	CA*	→	ZO
CG*	CC*	CT*	→	PS
TA*	TG*	TC*	→	PB
TT*	다음 문자부터 3개를 읽는다.			
*는 Don't Care를 나타냄				

그림 4는 이를 토대로 번역을 시행한 결과이다.

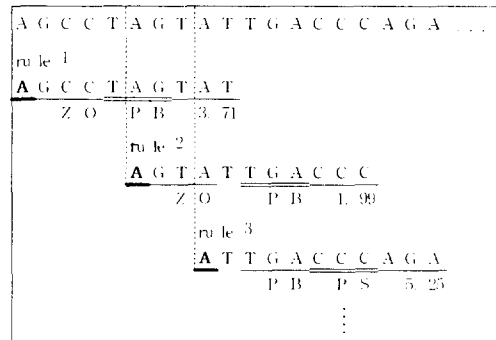


그림 4 시작 codon에 의한 문자열의 번역

여기에서 만들어진 각 개체를 다음의 평가 함수에 의해 평가한다.

$$H_i = \sum_{t=0}^n (|\text{목적수위} - h(t)|)$$

단 $i = 1, 2, \dots, 100$ (개체수)이다.

그 세대로 전부 개체의 평가치를 구한 후 그 중에서 평가치의 최대치(max)와 최소치(min)을 구하고 다음 식에 의해 각 개체의 그 세대에 대한 적합도를 구한다.

$$fit_i = \frac{\max - H_i}{\max - \min} \times 100$$

위 식에 의해 가장 뛰어난 개체에는 100이 가장 나쁜 개체에는 0이 주어지게 되고 가장 평가가 높은 개체가 돌연변이 등에 의해 소멸하는 것을 막기 위한 조치로 상위 5위까지 무조건 다음 세대에 전한다.

이제 정해진 순위에 따라 적응도가 낮은 개체를 그 세대에서 도태시킨다. 본 연구에서는 30개체가 도태된다.

다음으로 도태에 의해 전체 개체수가 감소하였기 때문에 원래 개체수가 될 때까지 남은 개체군 중에서 우수한 것을 Random하게 복사한다. 곧 우수한 개체를 증식시킨다.

또한 국소해에 빠질 위험성을 벗어나기 위해 가변의 교차점을 이용한 교차와 Random하게 얻어진 위치에서의 돌연변이를 시행한다.

V. 모의 실험

표 2 생성된 규칙 Δq

$e \backslash \Delta e$	NB	NS	ZO	PS	PB
NB	-4.19		4.66		
NS	-4.47	2.87			
ZO	-5.61				
PS			-0.38	-4.66	
PB	-4.28				

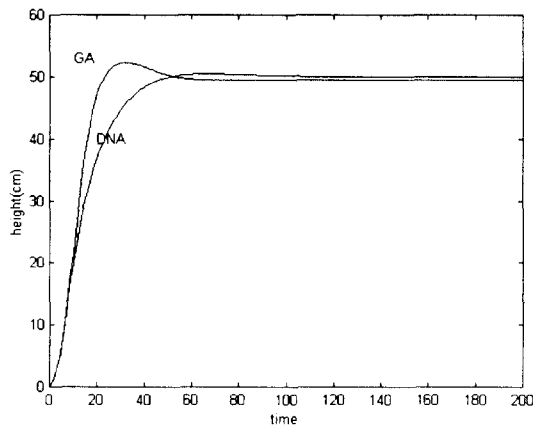


그림 5 수위 변화 그래프

본 논문에서는 플랜트의 시정수를 20으로, 목표 수위를 50cm로 주고 이상의 연산으로 얻어진 퍼지 규칙을 이용하여 교차율 40%, 돌연변이율 40%로 시뮬레이션을 행한 결과 위와 같은 수위 변화 그래프가 얻어졌다.

VI. 결론

본 연구에서는 GA에 DNA coding method

를 결합시켜 퍼지규칙을 획득하는 방법을 제안하였다. 여분(Redundancy) 특성과 유전자 중복의 허용성 때문에 지식의 언어적 표현에 적합한 DNA 알고리즘과 진화의 특성을 가지는 GA 알고리즘을 결합하여 새로운 알고리즘을 만들고, 이를 시뮬레이션 하였다. 통상적인 GA보다 어느 정도 개선된 결과를 얻을 수 있었고 이에 비추어 수위 제어에 필요한 퍼지 규칙의 획득에 있어 유효성을 가지고 있음이 확인되었다.

이 개선된 원인은 분석하면 GA에 비하여 제안한 알고리즘의 장점은 다음과 같다.

- 1) Random한 교차·돌연변이 연산에 대하여 길이에 대한 특별한 고안이 필요하지 않다.
- 2) 언어적인 표현이 다양하다
- 3) 중복 번역이 가능하다.

이후의 과제로는 시작 codon을 지정함에 있어서의 비율 결정, 염기의 종류 및 codon의 기본개수 증가의 연구와 MIMO 시스템에서의 적절한 적용 방법의 연구가 이루어져야 할 것이다.

VII. 참고 문헌

- [1] Tomohiro Yoshikawa, Takeshi Furuhashi and Yoshiaki Uchidawa. "The Effects of Combination of DNA Coding Method with Pseudo-Bacterial GA". Proceedings of the 1997 IEEE International InterMag '97 Magnetics Conference, 4/01/1997
- [2] Tomohiro Yoshikawa, Takeshi Furuhashi, Yoshiaki Uchikawa. "A Combination of DNA Coding Method with Pseudo-Bacterial GA for Acquisition of Fuzzy Control Rules". The 1st Online Workshop on Soft Computing (WCS1) Aug.19 (Mon) - Aug. 30 (Fri), 1996
- [3] T. Furuhashi, Y. Miyata, Y. Uchikawa. "Pseudo-Bacterial Genetic Algorithm and Finding of Fuzzy Rules". the Second Online Workshop on Evolutionary Computation (WEC2), pp.65-68 (1996)
- [4] Elie Sanchez. "DNA Biosoft Computing". 9th Annual Meeting of Biomedical Fuzzy Systems Association (Kurashiki, November 15-16, 1996)
- [5] Jee-Hyong Lee, Hyung-Kwang Lee. "A Study on Automatic Generation of Fuzzy Controller by Genetic Algorithm". Proceedings of KFIS Fall Conference, 1995 Vol. 5, No. 2
- [6] R. C. Murphy, R. Deaton, D. R. Franceschetti, S. E. Stevens Jr, and M. Garzon. "A New Algorithm for DNA Based Computation". Proceedings of the 1997 IEEE International InterMag., 1997 Magnetics Conference, 4/01/1997