

퍼지 추론 기반의 유전알고리즘 선택 연산자

Fuzzy Reasoning based Selection Operator for Genetic Algorithm

서기성, 현수환

Kisung Seo, Soohwan Hyun

서경대학교 전자공학과

요 약

본 논문은 퍼지추론을 통해 개체의 유사성과 적합도의 종합적 평가를 이용한 유전알고리즘의 선택연산자를 제안한다. 일반적으로 많이 쓰이는 적합도에 의한 선택 방법에 비해서 유사성에 대한 요소를 추가함으로써 조기에 수렴하는 현상의 감소와 성능향상을 얻을 수 있다. 또한 기존의 세대형(generational)에서 점진형(steady-state)으로 진화 수행방식의 변형을 통해 보조적인 향상을 제공할 수 있다. 제안된 방법을 f3deceptive 와 f5trap 등의 기만적 문제에 대해서 실험하였으며, 다른 연산자를 이용한 결과에 비하여 우수한 성능을 얻을 수 있었다.

키워드 : 유전알고리즘, 퍼지추론기반 선택연산자, 유사도, 적합도, 기만적 문제

Abstract

This paper introduces a selection operator which utilized similarity and fitness of individuals based on fuzzy inference. Adding similarity feature to fitness, proposed selector obtained the decrease of premature convergence and better performances than other selectors. Moreover, an adoption of steady-state evolution provided enhancement of performances additionally. Experiments of proposed method for deceptive problems were tested and showed better performances than conventional methods

Key Words : Genetic Algorithm, Fuzzy Reasoning based Selection Operator, Similarity, Fitness, Deceptive Problem

1. 서 론

유전알고리즘[1]의 수행에 있어서 조기 수렴 문제는 매우 중요한 문제로 다루어지고 있다. 조기 수렴(premature convergence)은 전역 최적점에 도달하지 못하고, 대신 지역 최적 값으로 수렴되는 현상을 의미한다.

조기 수렴문제의 해결책으로 제시된 여러 기법 중의 하나는 복수개 집단(multi-pop) 모델을 사용하여 개체의 다양성을 유지하도록 하는 것이다. 그러나 단순한 복수개 집단(multi-pop) 모델은 조기 수렴 개선에 기대보다 큰 효과가 없으며, 이를 개선하기 위하여, 복수개 집단을 기반으로 하여 발전된 모델인 HFC (Hierarchical Fair Competition)[2,3], ALPS (Age-Layered Population Structure)[4] 등이 조기 수렴을 방지하는 데 매우 높은 효율성을 가지고 있음이 보고되었다.

이들 모델의 공통점은 개체들을 계층적으로 나누어 각 계층별로 공정한 선택을 하는 것이다. HFC 모델은 적합도를 기반으로 했고, ALPS 모델에서는 각각의 개체들의 발전되어온 역사(나이)를 기반으로 했다.

하지만 이들 모델은 섬 모델을 기반으로 한 다층군집화에 기초하고 있기 때문에, 많은 수의 개체를 확보해야 좀 더 효

율적인 연산이 가능하다는 단점이 있다. 즉, 임베디드 장비나 로봇 등의 소규모 메모리의 이용만이 가능한 환경에서 온라인으로 유전알고리즘을 적용하기에는 적합하지 않다.

따라서 단일집단에 대해서 기존의 조기수렴 문제의 해결을 위한 계층적 구분의 모델을 기본 원리를 이용하면서, 더 적은 개체수로서 효율적 연산이 가능한 퍼지추론 기반의 선택 연산자를 제안한다. 단일 집단을 가상적으로 임의의 n 개의 개체군을 나누고, 개체의 적합도와 유사도에 기반한 퍼지추론을 통해, 효율적인 계층화를 구성하고자 한다.

유사도는 해밍 거리를 이용해서 해당 군집, 혹은 전체 개체에 대한 각 개체의 평균적 유사도를 측정하여 사용하였으며, 적합도와 유사도 모두 스케일링을 통해 0부터 1까지의 범위의 수치로 변환하였다.

부가적으로 제한된 자원의 사용과 연산량 측면에서 세대형(generational) 진화방식보다 유리한 점진형(steady-state) 진화방식을 적용하였다.

2. 계층적 고립 모델의 문제점

HFC나 ALPS 모두 집단의 계층화 모델로 조기수렴 문제를 크게 개선하였으나, 아직도 다음과 같은 문제점들이 남아 있다.

첫째, 적합도나 나이등의 하나의 평가 기준으로 계층화를

접수일자 : 2007년 10월 25일

완료일자 : 2007년 12월 10일

하므로, 같은 계층 내에서 우수 개체들에 의해서 지배되는 지역수렴이 발생 될 수 있다. 즉, 적합도는 낮지만 가능성 있는 형질을 가진 개체들이 상위 군집으로 이동되기 전에 선택에서 탈락될 확률이 매우 높기 때문에 다양성을 잃을 수가 있다.

둘째, 특히, 세대형(generational) 진화방식의 경우에 매 세대마다 너무 많은 개체들이 유전연산에 참여함으로 이러한 현상을 가속화 시킬 수 있다.

셋째, 복수개 집단(multi-pop) 모델에 기반을 두고 있기 때문에, 많은 수의 개체를 확보하지 못할 경우에 조기수렴을 막는 효율이 매우 낮아지게 된다.

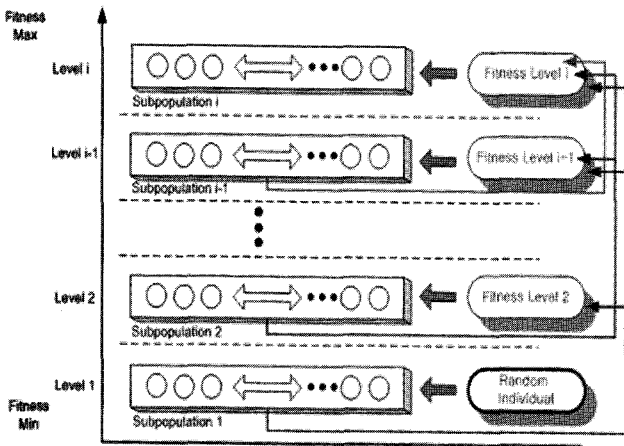


그림 1. 적합도에 의한 HFC 모델의 계층화 방식

따라서, 본 연구에서는 새로운 퍼지 추론 기반의 조합 선택 연산자를 제안한다. 기존의 HFC(그림 1, [2])의 복수 계층화 모델 대신에 하나의 군집에서 퍼지를 통한 여러 가지 정보의 조합으로 간접적 계층화를 이루어내고, 추가적으로 점진형(steady-state) 진화방식을 적용해서 기존의 대규모의 군집을 필요로 하던 연산을 적은 수의 연산으로도 가능하게 한다.

3. 퍼지 추론 기반의 조합 선택 연산자

제안하는 퍼지 추론 기반의 조합 선택연산자의 모델은 아래의 그림 2 와 같은 흐름으로 이루어진다. N 개의 개체를 통한 임의의 선택을 한 후에 유사성과 적합도에 대해 각각 재조정 작업이 이루어진 후에 퍼지추론 모델을 통해 퍼지수를 결과로 나타내고, 비퍼지화를 통해 실제 이용 가능한 개체의 소속도를 평가한다. 그리고 이를 기반으로 적합도 뿐 아니라 유사도에 대한 비교를 조합시킨 형태로 선택연산이 이루어진다.

유사도의 재조정 작업은 해밍거리를 이용해서 이루어진다. 임의로 선택된 개체의 그룹 내에서 모든 개체에 대해 해밍거리를 구한 후, 그 평균값을 문제 전체 비트수에 대한 비율로 환산하는 방법을 사용했다.

선택은 토너먼트(크기 = 9) 방식을 사용하였고, 토너먼트에 속한 개체들은 퍼지추론 모델을 통해 계산된 퍼지값을 기준으로 경쟁한다. 그림 3은 적합도와 유사도를 기반으로 한 퍼지 추론을 통해 부모가 될 개체를 선택함을 보여준다.

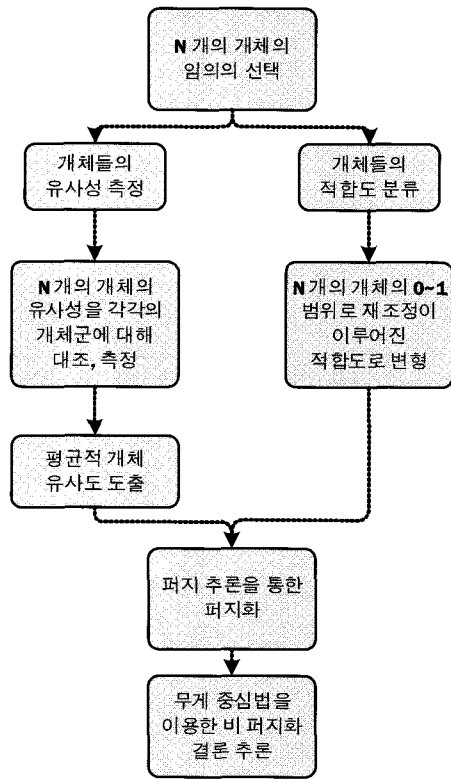


그림 2. 퍼지를 이용한 조합방식

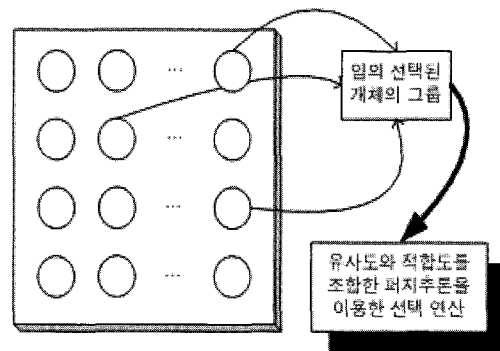


그림 3. 퍼지추론 기반의 선택연산자

4. 퍼지 모델

퍼지 모델은 전반부 변수 2개와 후반부 변수 1개로 구성했다. 전반부 변수 X_1, X_2 는 각각 적합도와 유사도를 나타내고, 후반부 변수는 G_i 은 소속계층을 의미한다. 전반부 변수의 멤버십 함수는 각각 Big, Middle, Small 로 구성했으며, 다음과 같은 가우시안 분포를 따른다.

$$A = \int_x \frac{e^{-a(x-b)^2}}{x}$$

후반부 변수는 다섯 개의 상수 값으로 구성된다. 각 변수

들의 범위는 다음과 같다.

$$0 \leq X1, X2 \leq 1$$

$$Y \in \{G1 = 1, G2 = 2, G3 = 3, G4 = 4, G5 = 5\}$$

추론 방식은 간략 추론방식을 이용했으며, 각 규칙은 다음과 같이 구성하였다.[5]

- R1 : IF X1 is Big and X2 is Big,
Then Y is G5
- R2 : IF X1 is Big and X2 is Middle,
Then Y is G3
- R3 : IF X1 is Big and X2 is Small,
Then Y is G1
- R4 : IF X1 is Middle and X2 is Big,
Then Y is G4
- R5 : IF X1 is Middle and X2 is Middle,
Then Y is G3
- R6 : IF X1 is Middle and X2 is Small,
Then Y is G2
- R7 : IF X1 is Small and X2 is Big,
Then Y is G3
- R8 : IF X1 is Small and X2 is Middle,
Then Y is G2
- R9 : IF X1 is Small and X2 is Small,
Then Y is G1

이는 각각의 적합도와 유사도에 따라 계층화 시킨 방식으로, G1 은 개선될 확률이 가장 적은 개체가 속한 계층을 의미하고, G5 는 개선될 확률이 가장 높은 개체가 속한 계층을 의미한다.

본 논문에서 테스트에 이용한 가우시안 분포는 반복 실험을 통해 가장 좋은 결과를 보인 것을 이용했다.

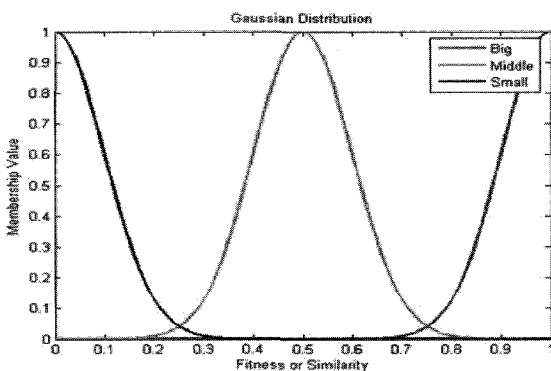


그림 4. 실험에서 이용된 가우시안 분포함수

4. 시뮬레이션 결과 및 고찰

4.1 테스트 함수

실험에서 이용된 문제는 f3deceptive 와 f5trap 함수이다 [6,7]. 다음의 수식(1), (2)와 같이 정의된 함수로서 각각의 3

비트, 5비트가 하나의 부분문제로 정의된다. f3deceptive 는 연결되어 있는 3개의 비트로 구성된 스트링에 포함되어있는 1의 개수로 적합도를 평가하고, f5trap 은 연결되어 있는 5개의 비트에 속해있는 1의 개수로 적합도를 평가 한다. 연속된 단위문제를 묶어서 문제의 크기가 커질수록 해결이 매우 힘들어지는 문제로 알려져 있다.

$$f3deceptive(u) = \begin{cases} 0.9 & u=0 \\ 0.8 & u=1 \\ 0 & u=2 \\ 1 & otherwise \end{cases} \quad (1)$$

$$f5trap(u) = \begin{cases} 4-u & \text{if } u < 5 \\ 5 & otherwise \end{cases} \quad (2)$$

여기서, u값은 한 단위문제당 포함되어 있는 1의 개수이다. 즉, f3deceptive 는 3비트를 한 단위로 볼 때, 1의 개수가 1,2 개로 증가될 때 점점 낮아지기 때문에 모든 비트가 서서히 0으로 수렴해 나가는 현상을 볼 수 있다. 그리고 전역 최적점은 반대방향에 존재하기 때문에 알고리즘이 지역 최적점을 회피해 전역 최적점의 탐색 능력을 실험해보기 좋은 함수로 널리 이용되고 있다.[6,7] GA 코드는 Gallops[8]를 수정하여 사용하였다.

4.2 실험결과

f3deceptive 문제의 실험에 대한 GA 파라미터는 pop-size=30, pcross=0.7, pmutation=0.1를 이용했고, 사용된 연산자는 교배연산자는 2점 교배 연산자, 변이연산자는 다중비트 변이 연산자를 사용했다. 각각의 실험방법은 FC, T+DC, T, R 로 구분하였는데, FC는 제안된 퍼지선택 연산자이고, T는 토너먼트 선택연산자이다. R은 룰렛휠 선택 연산자를 나타내고, DC는 DeJong Crowding[9] 기법을 의미한다. 여기서 제안된 연산자 FC는 점진형(steady-state) 진화방식을 사용하였으며, 나머지 연산자는 일반적인 세대형(generational) 진화방식을 사용하였다.

표 1. 테스트 결과(연산량, f3deceptive)

	R	T	T+DC	FC
20	548,632	44,254	24,819	15,017
40	1,476,106	103,766	67,562	52,170
80	3,842,609	320,227	222,893	185,723
160	5,842,967	783,645	534,458	471,572
200	8,986,015	1,243,961	975,250	732,389

각 실험은 10회씩 반복 실험 되었으며, 최고 적합도를 찾는 데 걸린 연산량의 평균값을 기록하였다. 즉, 연산량(number of evaluations) 이 작을수록 성능이 우수함을 의미한다. 표 1 에 각 선택연산자에 대한 연산량의 수치가 나와 있고, 왼쪽의 숫자(20, 40, 60, ...)는 블록(단위 묶음)의 크기로서 f3deceptive 에서는 3개의 비트가 하나의 블록으로 구성된다. 즉, 20 블록은 유전자 스트링의 크기가 60 비트임을 나타낸다.

제안된 퍼지기반 선택연산(FC)이 다른 3가지 방법보다 성능이 우수함을 알 수 있었고, DC 가 두 번째로 성능이 우수한 것은 개체의 유사도를 이용하므로 인해 얻어진다고 사료된다. 롤렛휠(R)이나 토너먼트(T)에 비하여는 월등한 성능향상을 보여주고 있음을 확인할 수 있다.

그림 5의 그래프에 연산량의 증가 추이가 나와 있다. 제안된 방법이 문제의 크기에 따라 증가되는 연산량의 기울기가 다른 알고리즘에 비해서 낮은 것을 확인할 수 있다. 롤렛휠의 경우 성능의 차이가 너무 커서 그래프에는 표시하지 않았다.

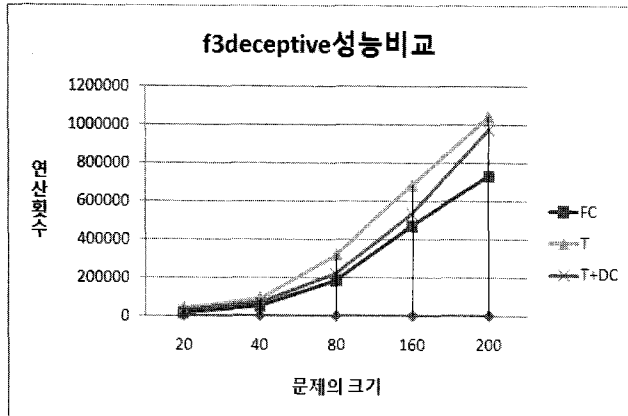


그림 5. f3deceptive 함수를 통한 성능비교

f5trap 문제에 대한 실험도 f3deceptive 와 마찬가지로 popsize=30, pcross=0.7, pmutation=0.1 의 환경에서 수행되었다. 블록 크기별 연산량의 결과가 표 2 에 나와 있다.

표 2. 테스트 결과(연산량, f5trap)

	R	T	T+DC	FC
20	6,846,248	533,344	371,889	282,461
40	X	1,678,021	963,200	678,476
80	X	3,662,225	2,285,588	1,245,239
160	X	8,080,191	6,087,087	3,554,323

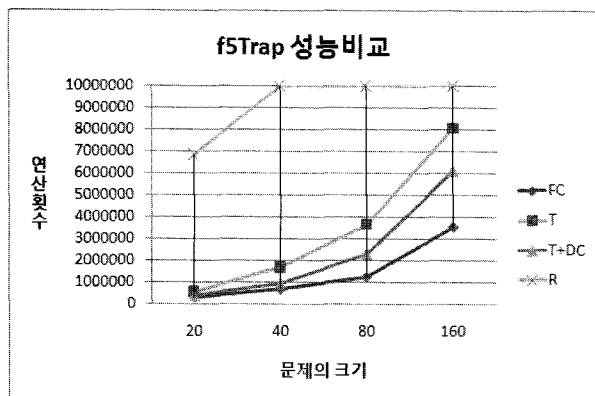


그림 6. f5trap 함수를 통한 성능비교

표 2의 결과에 표시된 X 는 천만번 이상의 연산을 거치고도 전역 최적점을 찾아내지 못한 경우이다. f5trap 문제의 경우에는 5개나 되는 비트를 하나의 블록으로서 보기 때문에, f3deceptive 문제보다 탐색의 난이도가 더 높다.

표 2와 그림 6에서 보는 바와 같이, 제안된 방법이 다른 알고리즘과 성능차이가 f3deceptive에 비해서 더 커진 것을 알 수 있다. 더 어려운 문제에 대해 더 좋은 성능을 발휘하고 있는 것을 확인할 수 있었다.

4.3 연산횟수에 따른 적합도와 유사도의 변화

그림 7 아래의 흐리게 표시된 부분은 두 부모의 유사도(개체 집단에서 나머지 개체에 대한) 및 부모간의 차이를 표현한다. 연산이 증가함에 따라 유사도의 차이가 줄어들지만 연산 종료시까지 일정 거리를 유지함을 알 수 있다.

그림 8 은 200 블록의 f3deceptive 문제에 대한 결과로서 80 블록 문제에 비해서 적합도 및 유사도의 차가 감소하였음을 볼 수 있다.

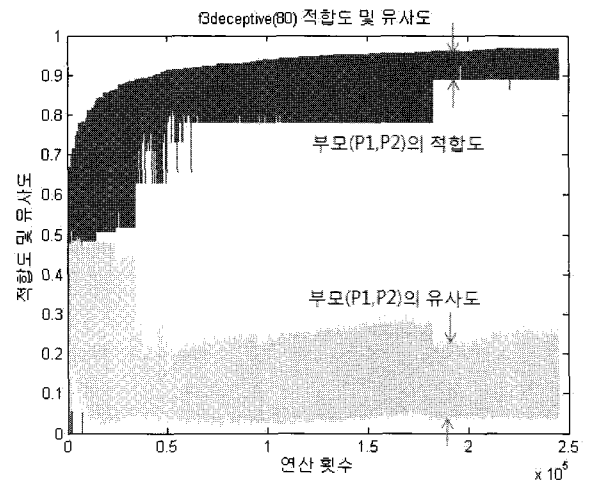


그림 7. f3deceptive(80)의 적합도 및 유사도

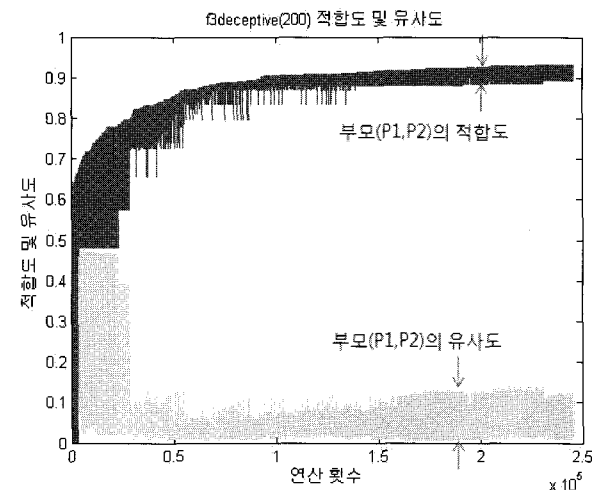


그림 8. f3deceptive(200)의 적합도 및 유사도

그림 9 과 10 의 그래프는 f5trap 문제에 대한 적합도 및

유사도 그래프로서 f3deceptive 문제 결과와 유사한 양상을 보여준다.

f3deceptive 결과에 비해서 적합도 곡선의 증가추이가 느림을 볼 수 있는데 이는 f5trap의 문제의 난이도가 높음을 반증하는 것으로 해석할 수 있다.

유사도도 f3deceptive 결과 그래프에 비해서 낮게 표시되며, 유사도 차이의 폭도 더 감소했음을 알 수 있다.

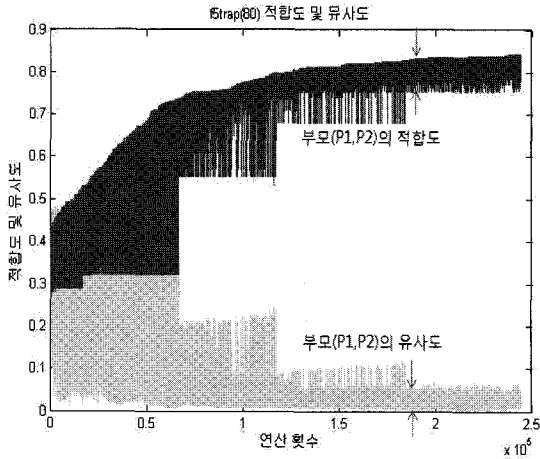


그림 9. f5trap(80)의 적합도 및 유사도

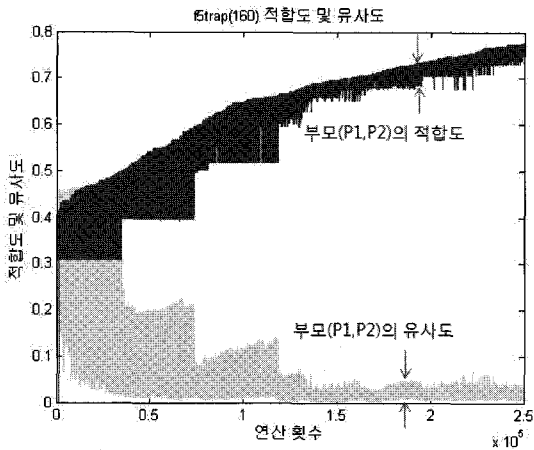


그림 10. f5trap(160)의 적합도 및 유사도

5. 결 론

본 연구에서는 퍼지추론을 기반으로 하여 적합도와 개체 간의 유사성을 통합한 유전 알고리즘의 선택 연산자를 구현하였고, 이를 기존의 대표적인 선택 방식들과 비교하였다. 대규모의 군집 및 개체수를 사용한 방식과는 달리, 일반적으로 이용되는 개체보다도 훨씬 적은 30개의 개체만으로 기만적인 어려운 문제를 효율적으로 탐색할 수 있음을 보였다. 따라서 온라인이나 작은 계산 용량만을 제공하는 시스템에서 효율적인 진화연산을 할 수 있는 가능성을 제공한다.

보다 많은 실험을 통해서 적합도 및 유사도가 조기수렴에 미치는 영향에 대한 분석적 고찰이 필요하고, 이를 반영한

개선된 선택 연산자의 구현이 요구된다. 또한 적합도와 유사도 외에 추가적으로 개체들의 나이, 개체의 개선속도등의 다른 특성 요소를 포함하는 방법도 고려해 볼 수 있다. 아울러 진보된 퍼지 추론법의 구현등이 필요하다고 본다.

참 고 문 헌

- [1] J. H. Holland, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA, 1989.
- [2] J. Hu, E. Goodman, K. Seo, Z. Fan, R. Rosenberg, "The Hierarchical Fair Competition(HFC) Framework for Continuing Evolutionary Algorithms", *Evolutionary Computation*, The MIT Press, Vol. 13, Issue 2, pp 241-277, 2005.
- [3] J. Hu, E. Goodman, and K. Seo, "Continuous hierarchical fair competition model for sustainable innovation in genetic programming." In Riolo, R. and Worzel, W., editors, *Genetic Programming Theory and Practice*, chapter 6. Kluwer Publishers, Boston, MA, 2003.
- [4] G. S. Hornby. "ALPS: The Age-Layered Population Structure for Reducing the Problem of Premature Convergence", In Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation, GECCO'06, July, Seattle, Washington, USA. pp. 815 - 822, 2006
- [5] 오성권, *프로그래밍에 의한 컴퓨터 지능*, 내하출판사, 2002.
- [6] M. Pelikan, D. E. Goldberg, and E. Cant' u-Paz, "BOA: The Bayesian Optimization Algorithm", In Proceedings of the Genetic and Evolutionary Computation Conference, GECCO'99, July, Orlando, Florida USA, pp. 525 - 32, 1999.
- [7] J. Hu and E. Goodman, "Robust and Efficient Genetic Algorithms with Hierarchical Niching and a Sustainable Evolutionary Computation Model", In Proceedings of the Genetic and Evolutionary Computation Conference, GECCO'04, Lecture Notes in Computer Science, Springer, June, Seattle WA, Part I, pp. 1220-1232, 2004.
- [8] E. D. Goodman, *Gallops User's Manual*. Michigan State University, July 1996.
- [9] K. A. De Jong, *An analysis of the behavior of a class of genetic adaptive systems*, (Doctoral dissertation, University of Michigan). Dissertation Abstracts International, 36(10), 5140B (UMI 76-9381), 1995.

저 자 소 개



서기성(Kisung Seo)

1986년 : 연세대학교 전기공학과 공학사

1988년 : 연세대학교 전기공학과 공학석사

1993년 : 연세대학교 전기공학과 공학박사

1993~1998년 : 서경대학교 산업공학과
전임강사, 조교수

1999~2002년 : Michigan State
University, GARAGe (Genetic Algorithms

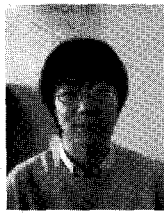
Research and Applications Group), Research Associate

2002년~2003년 : Michigan State University, Electrical &

Computer Engineering, Visiting Assistant Professor

2003~2004년 : 서경대학교 전자공학과 조교수

2004~현재 : 서경대학교 전자공학과 부교수



현수환(Soohwan Hyun)

2003~현재 : 서경대학교 전자공학과 학사
과정