

진화의 정량화를 이용한 해 진화 분석

이승익, 조성배

연세대학교 컴퓨터과학과

e-mail:cypher@candy.yonsei.ac.kr, sbcho@csai.yonsei.ac.kr

An Analysis on the Evolution of Solutions Using Quantification of Evolution

Seung-Ik Lee, Sung-Bae Cho

Dept. of Computer Science, Yonsei University

요약

진화알고리즘이 주어진 문제에 대하여 만족스런 해를 도출할 수 있음이 많은 연구결과를 통하여 알려졌지만, 그러한 해가 진화과정에서 형성되는 과정에 대한 분석은 미비한 실정이다. 진화현상에는 적용적 진화 외에도 다양한 현상들이 동시에 존재하기 때문에, 비록 진화알고리즘을 적용하여 해를 도출하였다 하더라도 그것이 적용적 진화의 결과임을 단정하기는 어렵다. 이 논문에서는 적용적 진화의 정량화를 통하여 최종해가 다른 유전적 현상에 의하여 도출되어진 것이 아니라 주어진 문제에 잘 적용한 적용성에서 비롯된 결과임을 보인다. 이를 위하여 하드웨어 에이전트의 제어기를 유전자 알고리즘을 이용하여 진화적으로 구축하고, 적용성 측정을 통하여 최종해의 형성과정을 분석한다. 실험결과 최종해는 적용적 진화의 결과임을 알 수 있었다.

1. 서론

진화연산, 인공생명, 그리고 진화생물학 분야의 많은 연구들에서 알 수 있듯이 진화에는 적용적 진화 외에도 우연이나, 유전적 표류 같은 많은 현상이 존재한다[1-3]. 이러한 다양한 진화적 현상으로 인하여, 진화알고리즘으로 획득된 해가 주어진 문제에 적용적인 능력을 보임으로써 획득된 해인지 아니면 다른 진화적 현상의 산물인지를 단정짓기가 어렵다.

따라서 진화알고리즘을 통하여 좋은 해를 지속적으로 획득하기 위해서는 좋은 해를 생성하기 위한 능력, 즉 진화성이 진화과정에서 유지되어야 하며 이러한 조건하에서만 좋은 해의 지속적인 획득이 가능하다. 높은 진화성을 유지하고 있는 집단은 그렇지 못한 집단에 비하여 좋은 해를 생성해낼 가능성이 높지만 그러한 가능성이 최종적으로 형성된 해가 적용적인 진화의 결과임을 보장하지는 않는다.

그러므로 최종적인 해가 적용적인 진화의 결과임을 보이기 위해서는 진화성의 측정만으로는 부족하며 진화과정에서 최종적인 해가 어떻게 형성되었는지를 보이는 것이 필요하다. 이러한 과정은 그동안 블랙박스처럼 여겨져왔던 진화과정에서의 진화알고리즘의 역할을 규명한다는 의미

를 지니고 이후에 보다 진화성이 높은 진화를 설계하기 위한 필요조건으로 여겨진다.

그러나 이러한 부분에 대한 연구는 최종적인 해의 응용에 관심이 많았던 기존의 연구들에서 간과되어 온 것이 사실이다. 이 논문에서는 적용성에 기반한 진화성 측정방법과 스키마 분석을 통하여 진화과정에서의 적용적 개체를 인식하고 이러한 적용적 개체가 최종해의 형성에 기여함을 보임으로써 최종해가 적용적인 진화의 산물임을 보인다.

2. 관련연구

Bedau와 Packard는 진화성을 측정하는 방법을 제안하였다[4]. 이들이 제안한 진화활동성 통계방법과 그와 연관된 다이어그램은 적용적 진화의 역할을 잘 설명하고 있으며, 이러한 특성으로 진화 시스템의 진화성을 측정하는데 이용될 수 있다. 진화활동성 통계방법은 여러 가지 문제에 적용될 수 있는데, 적용적 진화를 시각화하거나[4-6], 오랜 시간에 걸친 진화의 경향을 알아내는 문제[7, 8], 그리고 진화를 분류하는 문제에 적용되었다[9]. 또한 이러한 진화활동성 기법을 이용하여 퍼지 제어기의 진화성을 측정한 연구결과도 있다[10].

3. 진화성의 측정

적용성은 다른 유전현상과 비교하여 새로운 기능을 수행하는 구조의 계속적 생성과 그 구조가 지속적으로 진화 과정에서 존재한다는 특성을 보인다. 이러한 특성을 측정하기 위하여 Bedau는 진화활동성이라는 측정기법을 제안하였다[4].

진화활동성 측정을 위해서는 진화의 각 구성요소에 카운터를 부착한다. 여기서 구성요소는 문제에 따라 달라질 수가 있는데, 보통 개체나 유전자가 구성요소로서 정의될 수 있다. 시간 t 에서 i 번째 구성요소의 카운터 $a_i(t)$ 는 식 (1)에 의하여 변경된다.

$$a_i(t) = \sum_k \Delta_i(k) \quad (1)$$

여기서 $\Delta_i(k)$ 는 증가함수로서 문제에 따라서 다양하게 정의될 수 있다. 시간 t 에서 활동성 a 를 갖는 구성요소의 개수를 나타내는 구성요소 활동분포 $C(t, a)$ 는 식 (2)와 같이 정의된다.

$$C(t, a) = \sum_i \delta(a - a_i(t)) \quad (2)$$

여기서 $\delta(a - a_i(t))$ 는 $a = a_i(t)$ 이면 1이고 그렇지 않으면 0인 함수이다. 이에 기반하여 여러 가지 통계적 단위가 정의될 수 있다. 주어진 시간에서 구성요소들의 지속적인 적응성을 평가하기 위한 총 진화활동성, 평균 진화활동성, 진화시스템에 유입되는 새로운 적응적 개체의 비율을 표시하는 신진화활동성 등이 정의된다. 이에 대한 자세한 사항은 참고문헌[4-7, 9, 10]을 참조하기 바란다.

4. 진화시스템

진화 대상 시스템의 진화성 측정 및 해 형성 과정의 분석을 위하여 이 논문에서는 소형 이동로봇을 위한 퍼지 제어를 유전자 알고리즘[11]을 적용하여 진화적으로 구축한다. 퍼지 제어를 진화적으로 구성하기 위하여 8개의 입력변수와 2개의 출력변수, 그리고 최대 20개의 규칙을 가질 수 있도록 유전자를 그림 1과 같이 인코딩하였다.

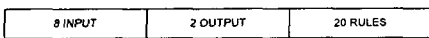


그림 1. 퍼지 파라미터의 코딩

각 입력변수에는 (VF, F, M, C, VC)의 다섯 개의 퍼지 집합이 정의되었고 출력변수에는 (BF, B, S, F, FF)가 정의되었다. 각 퍼지 집합은 삼각형 형태의 소속함수이며 다섯 개의 퍼지집합 중에서 두 개의 중점은 그림 2에서처럼 고정되어 있으므로 유전자 코드에는 3개만 코딩하면 된다.

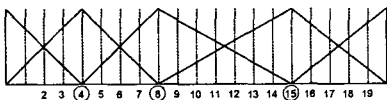


그림 2. 소속함수의 코딩

퍼지 제어기의 규칙은 그림 3과 같이 구성하였다.

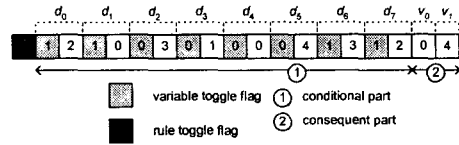


그림 3. 규칙의 인코딩

여기서 rule toggle flag은 해당규칙이 퍼지 제어기에 참여할지 여부를 표시하고 variable toggle flag은 해당변수가 규칙의 조건부에 참여하는지 여부를 표시한다. 따라서 그림 3은 다음과 같이 해석된다.

IF ($d_0=M$) and ($d_1=VF$) and ($d_6=C$) and ($d_7=M$)
THEN ($v_0=BF$) and ($v_1=FF$)

유전자 알고리즘을 적용하기 위하여 s 번의 동작시간을 가지고 n 개의 주요지점을 지난 퍼지 제어기의 적합도 합수 Θ 를 다음과 같이 정의하였다.

$$\Theta = \left(\frac{n}{N} \right) \sum_{i=0}^n V_i (1 - \sqrt{D_i}) (0.5 + 0.5(1 - I_i)) \quad (3)$$

여기서 N 은 주요지점의 총 개수이고 V_i 는 단계 t 에서의 로봇 모터의 평균 회전속도, D_i 는 두 모터의 속도차를 정규화한 값, I_i 는 가장 큰 값을 센서값을 정규화한 값이다.

5. 실험결과

5.1 환경설정

퍼지 제어기의 진화를 위하여 유전자 알고리즘을 사용하였고[11] 초기에 50개의 개체를 생성한 후 임의로 초기화하였다. 최대 1000세대까지 실험하였으며, 최우수개체 보존전략과 함께 50%의 개체들이 다음세대에 남도록 하는 overlap 기법을 적용하였다. 교차율은 0.5, 돌연변이율은 0.01로 하였다.

진화활동성 측정을 위한 증가함수는 식 (4)와 같이 정의하였다.

$$\Delta_i(t) = \begin{cases} \int_t^0 n_i(t) & \text{유전자 } i \text{가 시간 } t \text{에 존재할 때} \\ 0 & \text{그외} \end{cases}$$

여기서 $n_i(t)$ 는 유전자 i 에 해당하는 규칙의 수를 나타낸다.

5.2 시뮬레이션 결과

그림 4는 세대에 따른 적합도의 변화를 보여준다. 평균 적합도는 세대가 지남에 따라 서서히 증가하는 모습을 보이며, 우수 개체의 적합도는 초기에 서서히 증가하다가 550 세대 근처에서 급격하게 증가하는 모습을 보인다. 이 중에서 최대 적합도를 가진 개체는 746 세대에서 발생하였으며 이 개체의 이동모습이 그림 5에 나타나있다.

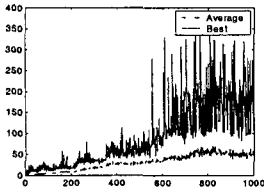


그림 4. 적합도 변화

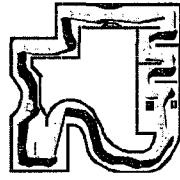


그림 5. 최개체의 케적

그림 6은 시간에 따른 각 유전자형의 진화활동성을 표시한다. 진화활동성의 정의에 의하여, 각 진화활동성을 표시하는 선은 항상 증가하는 형태로 나타나며, 어느 순간에 유전자형이 소멸하면 선이 끊기는 형태로 나타난다. 그림 6에서는 각 유전자형의 발생, 소멸, 집단에서의 분포정도, 그리고 적응정도를 알 수 있다

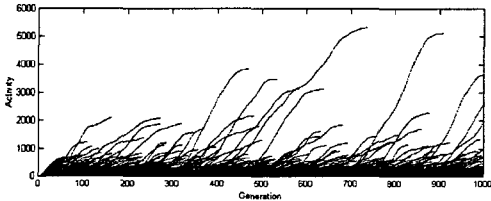


그림 6. 진화활동성 분포 C(t,a)

6. 최종해의 형성과정

그림 7은 일정 수준 이상의 높은 진화활동성 값을 갖는 주요규칙들을(salient rules) 보여준다. 표 1은 주요규칙들과 그에 상응하는 스키마를 보여주고, 표 2는 최우수개체의 규칙들과 해당하는 스키마를 보여준다.

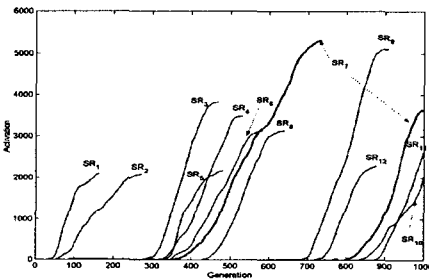


그림 7. 주요규칙의 진화활동성

각 스키마는 규칙에서 토글 플래그가 '1'인 부분만 남기고 나머지는 마스크되었다. 표 1과 2에서 알 수 있듯이 SH_7 은 BH_2 와 정확히 일치하고 SH_2 는 BH_4 와 일치한다. 따라서 최우수 개체의 규칙 BR_6, BR_{10} 은 주요규칙으로부터 형성되었음을 알 수 있다.

최우수 개체의 규칙 BR_2 는 주요규칙의 스키마와 비교를 통하여 유추될 수 있다. BR_2 의 스키마 BH_1 (표 2)을 주

요 규칙의 스키마 SH_1, SH_3 과 비교해 보면 그들 사이에 상당한 유사점이 있음을 알 수 있다.

표 1. 주요 규칙의 스키마

Name	Rule	Gene Code	Schema
SH_1	SR_1	1030202010010031113	#####10##1113
SH_2	SR_2	1011010040004010344	##1010#####44
	SR_3	1031010040002010344	
	SR_7	1001010040002010344	
	SR_{12}	1011010040002010344	
SH_3	SR_4	1030200011100031113	#####11###1113
SH_4	SR_5	1030201041401010041	#####14#####41
	SR_8	1000201011401010041	
SH_5	SR_6	1021410030404030401	###1410#####01
SH_6	SR_9	1100201011401010341	#10#####14#####41
	SR_{11}	1100001011401010341	
SH_7	SR_{10}	1000404040202031102	#####1102

표 2. 최우수 개체 규칙의 스키마

Name	Rule	Gene Code	Schema
BH_1	BR_2	1010303111203001133	#####1112###1133
BH_2	BR_6	1000200030100031102	#####1102
BH_3	BR_7	1021411000404110401	###1411#####11##01
BH_4	BR_{10}	1001010040002020044	##1010#####44
BH_5	BR_{11}	1100201041401011041	#10#####14###1041
BH_6	BR_{15}	1121300101102021022	#1213##1011###1022
BH_7	BR_{18}	1041302040003120021	###13#####12##21

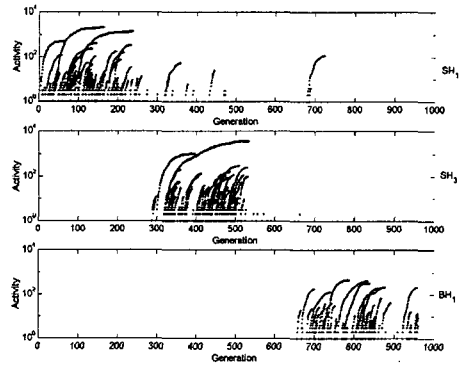


그림 8. SH_1, SH_3, BH_1 의 스키마 비교

그림 8은 이들 스키마 사이의 연관관계를 보다 명확히 보여주고 있다. 이 그림은 각 스키마에 속하는 규칙들의 진화활동성을 표시하고 있다. 그림에서 볼 수 있듯이 SH_1 의 소멸 혹은 감소가 SH_3 의 출현을 유도한 것을 볼 수 있으며, 다시 SH_3 의 소멸 혹은 감소는 BH_1 의 출현 혹은 증가로 이어짐을 볼 수 있다. 따라서 이를 통하여 우리는 최우수 개체의 규칙 BR_2 가 이전의 우수한 스키마로부터 형성되었음을 알 수 있다.

마찬가지로, 최우수 개체의 규칙 BR_7 의 경우에도 이전의 우수한 스키마 SH_5 로부터 형성되었으며(그림 9 참조), 규칙 BR_{11} 은 이전의 우수한 스키마 SH_4, SH_6 으로부터 형성되었음을 알 수 있다(그림 10 참조).

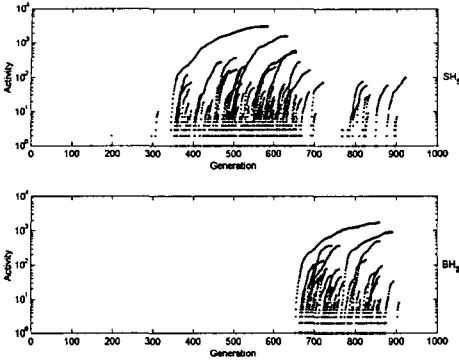


그림 9. SH_5, BH_3 의 스키마 비교

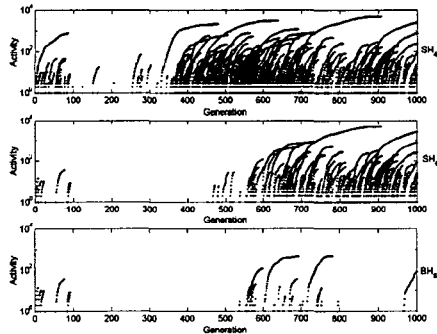


그림 10. SH_4, SH_6, BH_5 의 스키마 비교

7. 결론

이 논문에서는 진화의 적응성을 측정하는 한 방법인 진화활동성과 스키마 분석을 통하여 진화과정에서 적응적인 개체를 인지하고 이를 이용하여 최우수 개체의 형성 과정에 대한 분석을 수행하였다.

이로부터 우리는 최우수 개체의 구성 요소들이 우연히 형성된 것이 아니라 진화과정을 통하여 적응성을 검증 받은 요소들이 최우수 개체의 구성요소로서 참여하고 있음을 보였다.

따라서 진화알고리즘의 결과 형성된 최우수 개체는 다른 진화적 현상보다 적응이라는 진화적 현상을 통하여 형성되었다고 결론지을 수 있으며, 진화 활동성과 스키마 분석을 이용한 분석방법이 이러한 점을 보이는데 효과적으로 적용될 수 있음을 알 수 있다.

참고문헌

- [1] S.-B. Cho and T.S. Ray, "An evolutionary approach to program transformation and synthesis," *International Journal of Software Engineering and Knowledge Engineering*, vol. 5, pp. 179-192, 1995.
- [2] E. Mayr, *Towards a New Philosophy of Biology*, Harvard University Press, 1988.
- [3] M.J. West-Eberhard, "Adaptation: current uses," *Keywords in Evolutionary Biology*, Harvard University Press, pp. 13-18, 1992.
- [4] M.A. Bedau and N.H. Packard, "Measurement of evolutionary activity, teleology, and life," *Artificial Life 2*, Addison-Wesley, Redwood City, CA, pp. 431-461, 1992.
- [5] M.A. Bedau and C.T. Brown, "Visualizing evolutionary activity of genotypes," *Artificial Life 5*, pp. 17-35, 1997.
- [6] M.A. Bedau, S. Joshi and B. Lillie, "Visualizing waves of evolutionary activity of alleles," *Proceedings of the 1999 GECCO Workshop on Evolutionary Computation Visualization*, 1998
- [7] M.A. Bedau, E. Snyder, C.T. Brown and N. Packard, "A comparison of evolutionary activity in artificial evolving systems and in the biosphere," *Proceedings of the Fourth European Conference on Artificial Life, ECAL97*, MIT Press/Bradford Books, 1997.
- [8] C.C. Maley, "Four steps toward open-ended evolution," *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 1336-1343, 1999.
- [9] M.A. Bedau, E. Snyder and N. Packard, "A classification of long-term evolutionary dynamics," *Artificial Life 6*, pp. 228-237, 1998.
- [10] 이승익, 조성배, "진화활동성을 이용한 퍼지 제어기의 진화분석," *한국퍼지 및 지능시스템학회 2001년도 춘계 학술대회 학술발표논문집*, vol. 11, number 1, pp. 113-116, 2001.
- [11] D.E. Goldberg, *Genetic Algorithms in Search, Optimization & Machine Learning*, Addison Wesley, 1989.