

직접조작을 위한 NK-landscape 기반 IGA 인터페이스

민현정^o, 조성배
연세대학교 컴퓨터과학과
{solusea, sbcho}@cs.yonsei.ac.kr

An IGA Interface based on NK-landscape for Direct Manipulation

Hyeun-Jeong Min and Sung-Bae Cho
Dept. of Computer Science, Yonsei University

요 약

본 논문에서는 직접조작 방법을 이용하여 대화형 유전자 알고리즘(IGA)의 진화성을 향상시키고 세대를 반복하면서 다양한 해를 얻을 수 있음을 NK-landscape를 이용해 분석한다. 또한 NK 모델을 이용하여 직접조작의 진화 과정을 분석하기 위한 IGA 인터페이스를 생성한다. IGA를 이용한 진화는 특성상 적은 수의 개체와 세대수로 제한되기 때문에 지역해에 빠질 수 있는 문제점이 있다. 이러한 IGA의 제약사항을 극복하는 방법으로 직접조작을 이용할 수 있으며 이 방법은 사용자가 원하는 개체를 생성할 수 있도록 진화과정에서 유전자를 직접 변경한다. IGA에서 진화성의 향상과 다양한 해의 생성을 분석하기 위하여 진화연산자인 교차, 돌연변이와 직접조작 방법의 성능을 비교 분석한다. Kauffman이 제안한 NK-landscape로 진화과정에서 해공간이 얼마나 "ruggedness"한지와 다양한 진화연산자에서 어떤 성능을 갖는지 분석할 수 있다. 실험을 통해 3D 꽃 디자인 문제에서 NK-landscape 기반 IGA 인터페이스를 이용하여 IGA 해공간을 분석하고 교차, 돌연변이와 직접조작 방법의 성능을 비교함으로써 직접조작 방법으로 더 빠른 시간에 다양한 최적해를 찾을 수 있음을 알 수 있었다.

1. 서론

진화 알고리즘은 인공지능 분야에서 최적화 문제와 같은 다양한 응용범위를 가지고 있으며, 진화 알고리즘의 성능은 진화연산에 적용되는 변수인 돌연변이율, 교차율, 개체수, 세대수에 따라 좌우된다. 대화형 유전자 알고리즘(IGA)은 적합도 함수가 명시적으로 정의되지 않고 평가가 애매한 문제인 디자인이나 예술, 음악, 영상 생성, 애니메이션 등의 응용분야에 사용되는 최적화 방법이다[1]. 이 알고리즘은 적합도 평가에 대하여 사용자가 개체를 직접 평가하도록 변형한 것이며 따라서 사용자의 감성이 진화과정의 중요한 인자가 된다. IGA에서는 적합도 평가 때문에 사용자의 피로도가 증가할 수 있기 때문에 적은 수의 세대내에서 개체를 평가하게 된다. 이것은 많은 세대와 개체수를 가질수록 좋은 진화 성능을 얻을 수 있는 진화 알고리즘에 비해 국소탐색과 지역해에 빠질 수 있는 문제점을 남기게 된다.

직접조작 인터페이스는 사용자가 개체를 평가하는 과정에서 개체를 직접 조작하게 함으로써 사용자가 원하는 개체를 생성할 수 있도록 하는 방법이다[2]. 사용자의 감성이 평가에 중요한 역할을 하는 IGA 문제에서 직접조작 방법은 사용자가 만족할만한 개체를 생성하도록 한다. 이 방법은 HCI, 소프트웨어 공학, 심리학, 언어학, 예술 등의 다양한 분야에서 이용되고 있다. 직접조작 방법은 IGA의 평가과정에서 개체의 형질을 직접조작 함으로써 그 성능에 영향을 준다. 진화과정에 영향을 주는 다른 전략으로 다른 두 개체의 유전자를 바꾸는 교차연산과 한 개체의 유전자를 바꾸는 돌연변이 연산이 있다.

Fitness Landscape 분석방법은 개체를 표현하는 유전자의 해 공간이 갖는 적합도를 보여주는 도구이며, 개체가 진화를 거쳐 어떻게 다음 세대에 살아남는지를 보여준다. 이 개념

은 1932년 Wright가 처음 도입했는데, 개체를 이루고 있는 유전자의 유전인자들이 적합도의 평가에 얼마나 기여하느냐에 따라 진화의 다양성에 영향을 준다는 것을 발견하였다[3]. NK-landscape는 Kauffman에 의해 제안된 Fitness Landscape 모델이며 진화 알고리즘에서 이 모델은 다른 인코딩 방법에 의한 영향과 진화 연산의 성능을 분석하는 방법으로 유용하다[4]. NK-landscape 모델은 문제에 종속적으로 생성되며 NK 모델을 생성하여 해공간의 "ruggedness"를 알 수 있도록 해공간을 보여준다. NK 모델에서 landscape는 N 개의 유전인자를 가지고 있으며 각각의 유전인자는 개체의 적합도 평가에 기여하는 다른 K 개의 유전인자와의 관계를 적합도로 표현한 방법이다[5].

본 논문에서는 적은 수의 세대와 개체수를 갖는 IGA의 제약사항을 보완할 수 있도록 직접조작 방법을 이용하며 성능 향상에 중요한 역할을 하는 다른 진화연산과의 비교를 통해 직접조작 방법이 IGA에서 유용함을 분석한다. 이때 NK 모델을 통해 "ruggedness" 해공간의 이론적인 뒷받침을 제시한다. 실험에서는 먼저 꽃을 진화기법으로 자동 생성하는 IGA 문제[6]에 원하는 패턴을 가진 우성형질을 생성하기 위한 진화전략으로 직접조작 방법을 이용한다. 또한 진화연산인 교차와 돌연변이와의 성능을 비교 분석함으로써 직접조작 방법이 다양성을 가지면서 최적해를 찾는 유용한 방법임을 보인다. 마지막으로 NK-landscape를 이용하여 "ruggedness" 해공간을 분석하고 직접조작의 유용성을 보이기 위해 적합도 테이블로 IGA 인터페이스를 생성한다.

2. NK-landscape 기반 IGA 인터페이스

2.1. IGA 기반 꽃 생성

본 논문에서는 기존연구인 방향성 그래프와 IGA를

이용하여 사용자 평가가 필요한 꽃 모양 디자인 문제에 제안하는 방법을 적용하였다 [6]. 기존연구에서는 꽃 모양 디자인을 위해 실제 꽃 모양의 구조를 유전자형으로 표현하는 구조적 방향성 그래프와 자연스러운 모양을 생성하기 위해 IGA 를 사용하였다. 이 시스템에서 유전자형은 비트 문자열이 아닌 방향성 그래프로 5×5 행렬로 표현된다. 이 방향성 그래프의 유전자는 그림 1의 예에서 보는 것과 같이 정의된다. 그래프의 각 노드는 크기와 색 정보를 가지고 있다. 방향성 그래프에서 노드 변수와 에지의 개수를 교차와 돌연변이 연산으로 변경함으로써 꽃의 기본구조로부터 다양한 모양이 생성된다.

노드	줄기	가지	꽃술	꽃잎	잎
줄기	3	2	1	0	0
가지	0	2	1	0	1
꽃술	0	0	0	10	0
꽃잎	0	0	0	0	0
잎	0	0	0	0	0

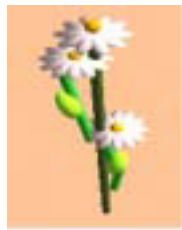


그림 1. 유전자형과 표현형 예

대화형 유전자 알고리즘을 이용하여 방향성 그래프의 노드를 유지하고 에지의 개수와 패러미터 값을 변경함으로써 꽃의 기본구조로부터 성장된 다양한 모양을 생성한다. 예를 들어서 {줄기, 가지, 꽃술, 꽃잎, 잎}으로 정의된 5 개의 노드에서 {줄기, 잎}과 {가지, 꽃술, 꽃잎}에 교차 연산이 발생한다면 두 개체에서 각각의 노드를 교차시키고 에지의 개수와 에지에 관련된 패러미터가 같이 교차한다. 이 교차 연산으로 첫 번째 부모의 {가지, 꽃술, 꽃잎}의 형질과 두 번째 부모의 {줄기, 잎}의 형질이 첫 번째 자식에게 나타나고, 첫 번째 부모의 {줄기, 잎}의 형질과 두 번째 부모의 {가지, 꽃술, 꽃잎}의 형질이 두 번째 자식에게 나타나게 된다.

2.2. 적합도 테이블과 진화전략

0~8	0~8	0~11	0~11	0~3	0~3	0~3	0~7	0~2	0~2	0~2	0~11
줄기	가지	잎	꽃잎	줄기	가지	잎	꽃잎	줄기	가지	잎	꽃잎
크기			색				에지				

꽃잎 크기 (0~11)		줄기 색 (0~3)		줄기 에지 (0~2)	
0	{0.45, 0.05, 0.15}	0	DarkGreen	0	1
1	{0.45, 0.05, 0.18}	1	DarkOliveGreen	1	2
2	{0.45, 0.05, 0.2}	2	PaleGreen	2	3
⋮	⋮	3	Green		

그림 2. NK 모델 생성을 위한 유전자

NK 모델을 생성하기 위하여 방향성 그래프로 정의된 유전자에 대하여 개체의 적합도에 영향을 줄 수 있는 유전자를 뽑아서 그림 2와 같이 표현하였다. 방향성 그래프의 유전자형을 갖는 꽃 모양 생성에 영향을 주는 유전자로 각 노드의 크기와 색, 그리고 연결 개수와 줄기, 가지, 잎의 형질로 구성하였다. 꽃을 구성하는 유전자는 그림에서와 같이 각각 정의된 개수의 유전자가 있는데 이것은 방향성 그래프의

유전자형 표현을 NK-landscape로 분석하기 위해 재 정의한 형태이다. 꽃 모양의 디자인을 생성하는 유전인자의 종류는 그림 2와 같이 정의된다. 그림은 꽃잎 크기와 줄기 색, 줄기의 에지수에 대한 유전인자의 종류를 예로 나타냈다. 정의된 유전인자에 따른 형질변화 각각에 대하여 직접 평가하는 방법으로 적합도를 평가하여 적합도 테이블을 생성하도록 한다. 적합도 테이블은 각 유전인자에 대하여 인자의 값과 적합도로 구성된다. 이 적합도 테이블은 돌연변이 연산과 직접조작의 진화연산에서 각 개체의 유전자에 대하여 평가하는 기준으로 사용된다.

S ₁	S ₂	S ₃	F ₁	F ₂	F ₃	F
0	0	0	2	6	5	4
0	0	1	4	8	6	6
0	0	2	5	8	7	6
0	1	0	6	6	3	5
0	1	1	2	1	4	2
0	1	2	3	1	2	2
0	2	0	6	8	6	6
0	2	1	7	6	8	7
0	2	2	8	9	7	8
⋮	⋮	⋮	⋮	⋮	⋮	⋮

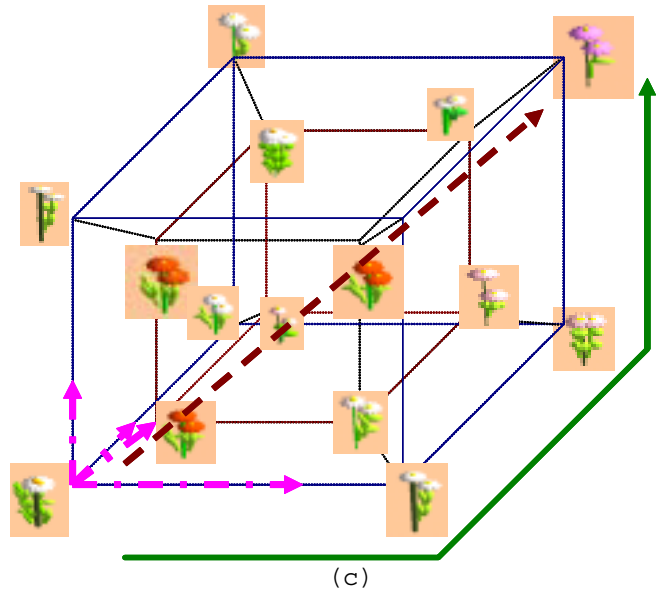
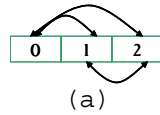


그림 3. 꽃 생성의 유전자로 NK 모델 적합도 테이블 생성 (a) 유전인자와의 관계, (b) 유전인자의 형질에 따른 적합도, (c) 형질변화에 따른 개체의 표현형

그림 3은 꽃 생성의 방향성 그래프로 표현된 유전자형을 그림 2와 같은 방법으로 유전인자를 표현하고 NK 모델 적합도 테이블을 생성한 예이다. 이 예는 N=3, K=2인 경우의 NK-landscape를 생성하는 과정을 보여준다. 여기에서 N과 K는 비트 문자열인 다른 유전자형과는 달리 여러 종류의 값을 가질 수 있다. 그림에서 S₁, S₂, S₃는 각각 0~2의 3개 값을 가지고 있다. (a)는 유전인자의 개수가 3이고 K가 2인 경우 각각의 유전인자에 영향을 주는 상관관계를 나타낸다. (b)는 유전인자의 적합도를 평가하고 이것으로 개체의 적합도를 구한 예를 보여준다. NK 모델은 유전자를 이루는 각각의 유전인자에 대해 n^{K+1}개의 적합도 테이블을 갖는다. 여기에서 n은 유전인자를 구성하는 형질의 개수이고 K는 유전인자의 개수이다. (c)에서 개체의 적합도 평가함수는 다음 식으로 구할 수 있다.

$$F = \frac{1}{N} \sum_{i=1, j=1}^{N, n^{k+1}} F_i^{(j)}$$

그림 4는 유전인자와 유전인자를 가지고 있는 개체의 적합도를 테이블로 나타낸 예이다. 여기에서 유전인자에 나타나는 숫자는 정의된 인자의 종류이다. 이 예는 0번 줄기크기, 0번 가지 크기, 9번 잎 크기의 유전인자에서 적합도 8, 1번 줄기 색, 1번 가지 색, 2번 잎 색의 유전인자에서 적합도 9, 그리고 3번 가지 예지, 3번 잎 예지의 유전인자에서 적합도가 9임을 보여준다. 이 적합도 테이블은 앞서 기술한 n^{k+1} 개의 항목을 가지고 있다.

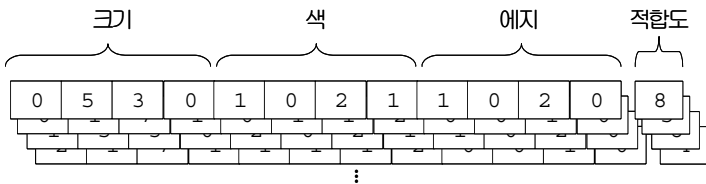


그림 4. 적합도 관계 테이블로 나타낸 nk 모델의 예

돌연변이 연산은 개체의 각 노드와 예지의 개수 및 예지를 이루고 있는 변수에 변이가 발생하도록 한다. 직접조작 방법은 IGA 과정에서 생성된 NK 모델의 적합도 테이블을 참조하여 구성된 형질에 따라 적합도를 평가하는 과정에서 평가자가 임의로 직접 형질을 바꾼다. 위의 적합도 테이블을 이용하여 돌연변이율의 변화와 직접조작과의 성능변화를 비교할 수 있다.

3. 실험 및 결론

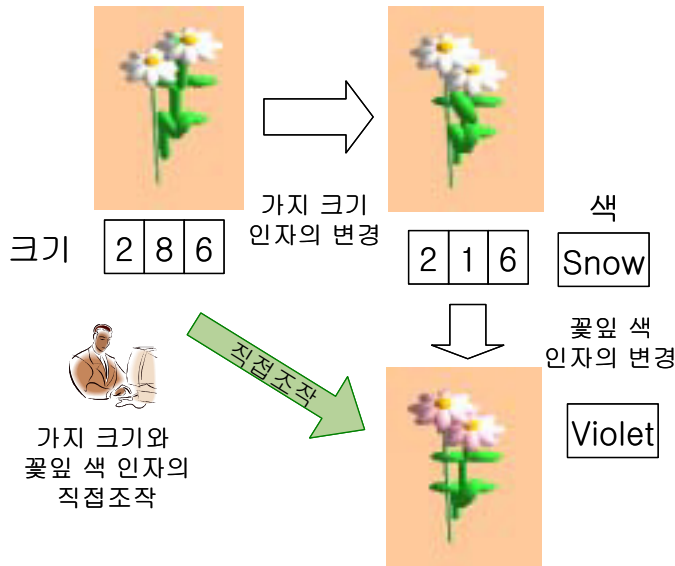


그림 5. IGA 인터페이스를 이용한 직접조작의 예

실험은 3D 꽃 모양 생성 문제 [6]에 제안한 방법으로 NK 모델을 생성하고 그 적합도 테이블로 돌연변이율의 변화와 직접조작의 적합도를 평가하고 결과를 비교한다. 이를 위하여 두 가지로 나누어서 각각 수행하였다. 첫 번째는 NK 모델을 생성하기 위한 실험이고 두 번째는 적합도 테이블을 참조하여 진화과정을 평가하는 실험이다. NK 모델을 생성하기 위한 적합도 평가는 각각의 형질변화에 따른 모든 개체에 대하여

직접 평가하였다. 두 번째 실험에서 첫 번째 실험에서 생성한 적합도 테이블을 참조하여 진화연산의 적합도를 평가하였다.

그림 5는 IGA 인터페이스를 이용하여 직접조작으로 유전인자를 바꾼 예를 보여준다. 줄기, 가지, 잎의 유전인자 각각의 종류가 2, 8, 6으로 정의된 개체에 가지크기 유전인자를 1로 바꾼 후의 개체 표현형과 "Snow"의 꽃잎 색 유전인자를 "Violet" 꽃잎 색으로 바꾼 후의 개체를 보여준다. 이 인터페이스를 이용해 직접조작 방법에서 적합도를 평가한 것과 돌연변이율을 바꾸어 적합도를 평가한 것을 비교하여 직접조작이 유용함을 알 수 있었다. IGA의 제약점인 적은 수의 개체와 세대수 때문에 지역해에 빠질 수 있는데 직접조작으로 지역해에 빠지는 것을 극복할 수 있고 다양한 해를 생성할 수 있다.

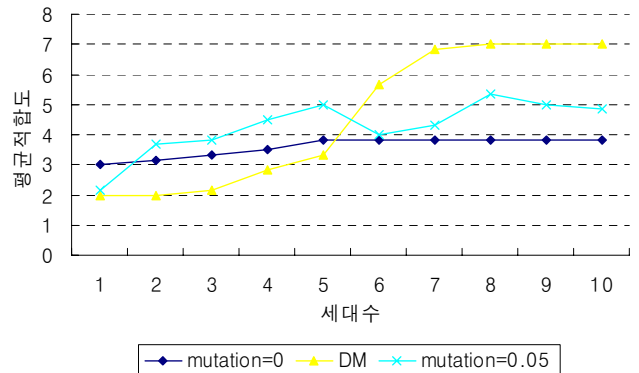


그림 6. 진화의 성능 분석

그림 6은 돌연변이율 0과 0.05 및 직접조작에서의 평균 적합도를 그래프로 나타낸 것이다. 돌연변이가 발생하지 않을 경우는 5세대 정도에서 최적의 해를 찾았고 마지막 세대에서 하나의 해에 수렴한다는 것을 알 수 있다. 돌연변이율이 0.05일 경우는 진화과정에서 다양하게 적합도 평가를 받으면서 수렴해 가는데 마지막에 돌연변이가 없을 때보다는 다양한 해를 찾을 수 있었다. 그러나 최적의 해라기 보다는 다양한 해를 생성함을 알 수 있다. NK 모델을 이용한 IGA 인터페이스를 통해 직접조작 방법은 빠른 속도로 수렴하면서 마지막 세대에서 다양한 해를 생성함을 알 수 있다. 또한 마지막 세대에 생성한 해가 최적의 적합도를 받았던 우성형질임을 알 수 있었다.

감사의 글

본 연구는 프론티어 연구사업의 지원에 의한 것임.

참고문헌

- [1] H.-S. Kim and S.-B. Cho, "Application of interactive genetic algorithm to fashion design," *Engineering Applications of Artificial Intelligence*, vol. 13, no. 6, pp. 635-644, 2000.
- [2] B. Shneiderman, "Direct manipulation: A step beyond programming languages," *IEEE Computer*, vol. 16, no. 8, pp. 57-69, 1983.
- [3] S. Hill and C. O'Riordan, "Genetic algorithms, their operators and the NK model," *Technical Report*, NUI, Ireland, 2001.
- [4] Y. Gao, "Threshold phenomena in NK landscapes," Master's Thesis, University of Alberta, Canada, 2001.
- [5] S. A. Kauffman, *The Origins of Order*, Oxford University Press, 1993.
- [6] H. J. Min and S. B. Cho, "Automatic creation of 3D artificial flowers with interactive evaluation on evolutionary engine," *Proc. of 4th International Symposium on Advanced Intelligent Systems*, pp. 702-705, 2003.