

# GA-Hard 문제를 풀기 위한 공진화 모델

## Co-Evolutionary Model for Solving the GA-Hard Problem

박창현 · 이동욱 · 심귀보  
 중앙대학교 전자전기공학부

Chang-Hyun Park, Dong-Wook Lee, and Kwee-Bo Sim  
 School of Electrical and Electronic Engineering, Chung-Ang University  
 E-mail: kbsim@cau.ac.kr

### 요 약

공진화 알고리즘은 두 개 이상의 개체군이 상호작용하며 진화하는 알고리즘이다. 기존의 진화 알고리즘이 하나의 개체군으로 구성된 정적인 적합도 지형에서 해를 찾는 방식임에 반해 공진화 알고리즘은 두 개 이상의 개체군이 동적인 적합도 지형을 제공하여 더 강건하고 빠른 수렴성을 보인다. 본 논문에서는 GA가 풀기 어려운 GA-hard problem을 풀기 위하여 저자가 제안한 3가지 공진화 모델을 설명한다. 첫 번째 모델은 찾고자하는 해와 환경을 각각 경쟁하는 개체군으로 구성해 진화하는 방법으로 사용자의 환경 설정에 의해 지역적 해를 찾는 것을 방지하는 경쟁적 공진화 알고리즘이다. 두 번째 모델은 찾고자하는 해와 이를 보조하는 스키마를 각각 개체군으로 구성해 진화하는 스키마 공진화 알고리즘이다. 세 번째 알고리즘은 해를 구성하는 부분을 두 개의 개체군으로 나누고 두 개체군이 서로 게임을 통해 진화하도록 하는 게임이론에 기반한 공진화 알고리즘이다.

**Keywords :** GA-Hard Problem, 경쟁적 공진화 모델, 협조적 공진화 모델, 기생 공진화 모델

### 1. 서 론

유전자 알고리즘(GA)은 자연계의 진화과정을 컴퓨터 상에서 시뮬레이션 함으로써 복잡한 실세계의 문제를 해결하고자 하는 계산모델이다[1][2]. GA는 구조가 간단하고 방법이 일반적이어서 응용범위가 매우 넓고, 특히 적응적 탐색과 학습 및 최적화를 통한 공학적인 문제의 해결에 많이 이용되고 있다. 그런데, GA가 모든 최적화 문제에 좋은 결과를 가져오는 것은 아니다. 예를 들면, 최적해를 포함하는 스키마의 적합도가, 경쟁하는, 최적해를 포함하지 않는 스키마의 적합도보다 낮아서 결국 최적해를 포함하지 않는 스키마 쪽으로 수렴하기 때문에 최적해를 찾지 못하고 국소해에 머무는 경우가 발생한다. 이와 같은 경우의 문제를 GA-hard problem이라고 한다[2]. 다시 말하면 GA는 코딩방법 및 적합도 설정에 의해 결정되는 적합도 지형(fitness landscape)에 의해 성능의 제약을 받는다. 따라서 적합도 지형이 GA가 문제를 풀기 어렵게 결정되면 GA는 국소해에 조기 수렴하거나 해를 찾는데 시간이 오래 걸린다. 이와 같은 문제점을 해결하기 위하여 좀더 자연계의 진화 현상과 가깝게 모델링 한 것이 공진화 알고리즘이다 [3]-[12]. 공진화란 두 개 이상의 개체군이 서로 영향을 주고받으며 진화하는 것을 말한다. 여러 개체군간의 상호 작용은 동적인 적합도 환경을 제공하여 해가 조기 수렴하는 것을 막고 진화 속도를 향상 시킬 수 있다.

공진화 알고리즘은 개체군 간에 영향을 주고받는 방식에 따라 경쟁적 공진화 모델[3]-[6], 협조적 공진화 모델[4][7][8], 기생 공진화 모델[9]-[12]로 나누어 볼 수 있다. Hillis[3]는 sorting network과 data의 경쟁적 공진화 모델을 제안하였고 이 모델을 기반으로한 신경망과 학습패턴의 경쟁적 공진화 모델도 개발되었다[4][5]. Sim의 게임이론에 기반한 공진화알고리즘[6]도 경쟁적 공진화 알고리즘의 하나이다. Paredis는 SYMBIOT[4]라는 협조하는 두 개체군의 공진화 모델을 제안하였고 Potter와 DeJong[7]은  $n$ 개의 협조적으로 영향을 주고받는 공진화 모델을 개발하였다. Pena-Reyes와 Sipper[8]는 Potter의 공진화 모델을 퍼지 모델링에 적용하였다. 기생 개체군으로 virus 개체군[9]과 schema 개체군[10]-[12]을 도입한 기생 공진화 모델도 개발되었다. 한편 생물학자 Valen은 진화적 혁신과 적응의 주요한 원동력으로서 공진화적 군비확장경쟁(arms race)에서 유도되는 "Red-Queen effect"를 제안하였다[13].

GA가 풀기 어려운 GA-hard problem을 풀기 위한 방법으로써 본 저자는 경쟁적 공진화 모델, 스키마 공진화 모델 그리고 게임이론에 기반한 공진화 모델을 제안하였다. 경쟁적 공진화 모델은 해와 환경을 각각 개체군으로 구성하여 강건한 해를 찾기 위한 모델이고, 스키마 공진화 모델은 기생 개체군인 스키마 개체군을 이용하여 해 개체군의 진화적 탐색을 돕는 알고리즘이다. 게임이론에 기반한 공진화 모델은 해의 부분을 두 개의 개체군에 나누고 각 개체군이 게임을 통해 진화하는 알고리즘이다. 2, 3, 4장에서 각각 알고리

즘에 대하여 자세히 설명한다.

## 2. 경쟁적 공진화 알고리즘

### 2.1 공진화 알고리즘

본 알고리즘은 풀고자 하는 해와 환경을 함께 공진화 시켜 변화하는 환경에 대해서도 강건한 해를 찾기 위한 알고리즘이다. 주어진 문제를 본 알고리즘에 적용하기 위하여 목적하는 해를 주 개체군으로 구성하고 환경이나 외부 요인을 부 개체군으로 구성한다. 그림 1은 신경망 제어기 외를 주 개체군으로 학습패턴을 부 개체군으로 구성한 경쟁적 공진화를 나타낸 그림이다.

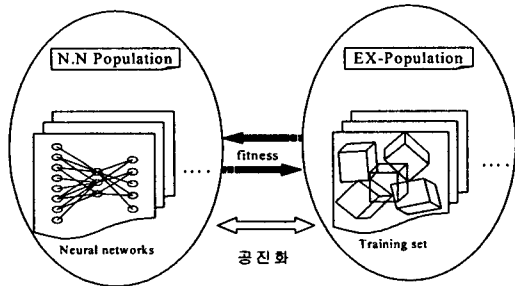


그림 1. 신경망과 학습패턴의 공진화

두 개체군이 진화하는 경쟁적 공진화 알고리즘은 다음과 같다.

#### [경쟁적 공진화 알고리즘]

##### (1) 개체군 구성

주목적이 되는 대상을 주개체군(primary population)으로 구성하고 환경과 같은 외부 요인을 부개체군(secondary population)으로 구성한다.

##### (2) 적합도 평가

주 개체군의 개체와 부 개체군의 개체의 적합도는 다음의 식 (1)과 (2)에 의해 구한다.

$$f(P_i) = \frac{\sum_{j \in A} f(P_i, S_j)}{L} \quad (1)$$

$$f(S_j) = \frac{\sum_{i \in B} f(S_j, P_i)}{K_j} \quad (2)$$

이때, 사용되는 변수는 다음과 같다.

- $N$ : 주개체군의 개체수
- $P_i$ : 주개체군의  $i$ 번째 개체(individual), 단,  $i=0 \dots N$
- $M$ : 부개체군의 개체수
- $S_j$ : 부개체군의  $j$ 번째 개체, 단,  $j=0 \dots M$
- $f(P_i)$ :  $P_i$ 의 적합도
- $f(S_j)$ :  $S_j$ 의 적합도
- $f(P_i, S_j)$ : 주개체  $i$ 와 부개체  $j$ 의 상호관계에 의한 적합도  
( $= 1 - f(S_j, P_i)$ )
- $A$ : 주개체  $i$ 가 선택한 부개체들의 인덱스 집합
- $L$ : 주개체의 평가회수
- $B$ : 부개체  $j$ 를 선택한 주개체들의 인덱스 집합
- $K_j$ : 부개체  $j$ 가 선택된 회수

즉, 주개체는 부개체에서 선택한  $L$ 개의 개체에 대하여 상호 적합도를 구한 후 평균값을 취하여 자신의 적합도로 가지며 부개체도 선택되었던  $K_j$ 번의 상호 적합도를 평균하여 자신의 적합도로 갖는다.

##### (3) 유전 연산

각각의 개체군에 대하여 일반적인 진화 알고리즘과 같은

선택, 교차, 돌연변이의 과정을 거쳐 다음 세대의 개체군을 구성한다.

##### (4) 종료조건

지정된 최대 세대가 지나거나 만족한 결과가 나올 때까지 (2)와 (3)을 반복한다.

### 2.2 공진화의 유효성 측정

경쟁적 공진화에서 진화적 적응과 개선이 일어나는 주요한 원인은 종간의 군비경쟁(arms race) 때문이다[12]. 즉 한 개체군은 다른 개체군의 적합도를 감소시키는 방향으로 진화를 하고 적합도의 영향을 받은 개체군은 떨어진 적합도를 유지하고자 진화를 하는 것이다. 이와 같이 적합도가 상대적으로 정의가 되어있기 때문에 공진화하는 개체의 동적인 적합도 지형(dynamic fitness landscape)을 갖는다. 따라서 적합도의 변화를 이용해 진화의 진행을 알아보기 힘들다. 따라서 공진화가 유효하게 진행되고 있는가를 알아보기 위하여 '상대 개체 조상들과의 경쟁'이라는 방법을 제안하였다[5].

현재  $g$  세대의 진화의 진행과정을 알아보기 위하여 부개체군의  $g$ 세대 이전의 각 세대들의 최고 개체들과 주개체군의  $g$ 세대 최고 개체와 순차적으로 경쟁을 시켜서 얻어진 적합도의 변화 모양을 살펴봄으로서 진화의 진행 과정을 알아볼 수 있다. 그림 2는 '상대 개체 조상들과의 경쟁'을 통한 공진화의 진행과정을 알아보는 방법을 나타낸 것이다.

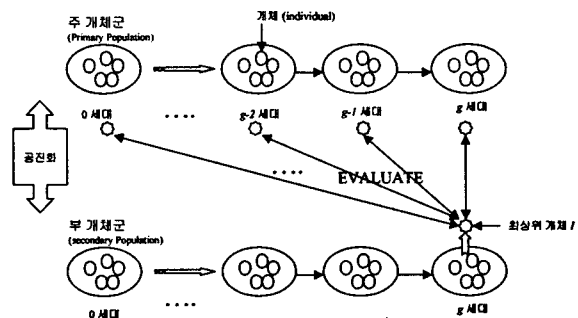


그림 2. 상대 개체 조상들과의 경쟁 개념

## 3. 스키마 공진화 알고리즘

### 3.1 스키마 공진화 알고리즘

다른 공진화 알고리즘과 같이 스키마 공진화 알고리즘(SCEA)은 두 개의 협조적으로 작용하는 개체군을 갖는다. 본 알고리즘은 기생 공진화 알고리즘의 일종으로서 해의 후보 개체군을 호스트 개체군(host population) 스키마로 구성된 개체군을 기생 개체군(parasite population)이라고 한다. 호스트 개체군은 일반적인 유전 알고리즘과 같은 방식으로 수행되고, 기생 개체군은 스키마의 집합으로서, '빌딩 블록'이라 불리는 유용한 스키마를 찾는데 사용된다. 그림 3은 SCEA 프로세스를 나타낸다.

GA는 한 세대동안 네 가지 중요한 단계를 갖는데, 그것은 적합도 평가, 선택, 교차, 돌연변이이다. 그러나 SCEA는 선택과정 전에 기생 프로세스를 추가로 포함한다. 호스트-개체군의 모든 개체들이 평가되어진 후에, 그 중의 몇 개가 기생-개체군의 스키마를 위해 임의로 선택된 다음, 기생 프로세스가 수행된다. 이것을 통해 새롭게 생성된 개체를 평가하고, 원래의 개체와 기생 프로세스에 의해 변한 개체 사이의 적합도 개선 정도를 측정한다. 가장 큰 개선

값을 갖는 기생 개체를 각각의 스키마에 대응하는 개체와 대치한다. 개선된 양을 이용하여, 기생 개체군에서 각각의 스키마의 적합도를 부여할 수 있다. 그러므로 기생 개체군의 스키마의 적합도는 스키마의 유용성을 가리킨다. 적합도 평가 후의 기생 개체군에는 SGA와 동일한 과정을 적용한다. 기생-개체군의 개체의 적합도는 기생 프로세스에 의해 계산된다. 따라서 기생 개체군은 기생, 선택, 교차, 돌연변이의 네 단계를 거친다.

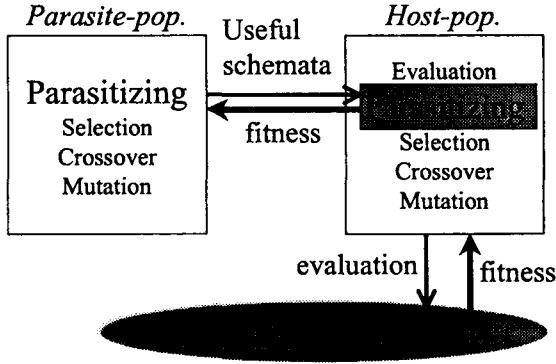


그림 3. SCEA 프로세스

3.2 기생 프로세스

기생 개체군은 유용한 스키마를 탐색하고 기생 프로세스에 의해 유전 정보를 호스트 개체군으로 전달한다. 그림 4는 기생 프로세스를 보여준다.

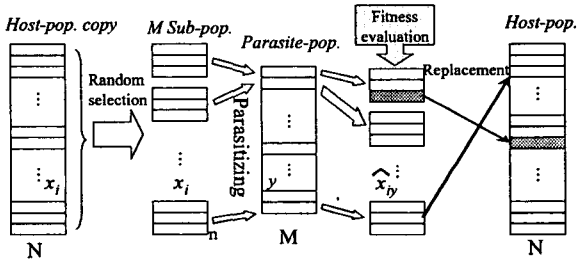


그림 4. 기생 프로세스. N은 호스트-개체군 크기, M은 기생-개체군 크기, n은 M개의 sub-개체군 크기.

기생 개체군의 스키마의 적합도는 호스트 개체군에서 추출된 n 개의 스트링에 의해 결정된다. 기생 프로세스는 스트링의 유전자가 스키마의 특정한 유전자에 의해 대체되어지는 것을 의미한다. 스키마 y의 p번째 위치가 고정 위치이면 그 값을, 그렇지 않고 \* (don't care)이면 자신의 값을 갖는다 (식(3)).

$$\hat{x}_i^p = \begin{cases} y^p, & \text{if } p\text{th character of } y \text{ is fixed} \\ x_i^p, & \text{otherwise} \end{cases} \quad (3)$$

여기에서 p는 유전자좌(0 ≤ p ≤ l-1), x\_i^p는 스트링 x\_i의 p 번째 자리 값, l은 스트링의 비트 개수를 나타낸다. 그림 5는 parasitizing의 한 예를 보여준다.

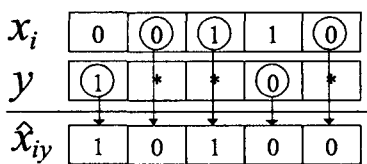


그림 5. Parasitizing의 예: x\_i는 스키마 y에 의해  $\hat{x}_{iy}$ 로 parasitizing 됨.

SCEA의 주요 프로세스는 기생 개체군에서 찾은 유용한 스키마를 적합도에 따라 호스트 개체군에 전달하는 것이고, 기생 개체군의 진화 방향은 호스트-개체군에 의해 결정된다. 기생 개체군에서 스트링 y의 적합도 F\_y는 다음 단계에 의해 결정된다. 다음은 스키마 공진화 알고리즘이다

[스키마 공진화 알고리즘]

- (1) 호스트 개체군에서 n 개의 스트링이 임의로 선택되어 복사된다.
- (2) 그림 5와 같이 parasitizing이 수행된다.
- (3) 기생-개체군에서 스트링 y의 적합도를 결정하기 위해, 적합도 함수를 다음 식과 같이 적합도의 개선 값으로 설정한다.

$$\hat{f}_{iy} = f(\hat{x}_{iy}) - f(x_i), \quad (i=1, \dots, n) \quad (4)$$

이 식에서 x\_1, ..., x\_n는 추출된 스트링,  $\hat{x}_{1y}, \dots, \hat{x}_{ny}$ 는 parasitizing된 스트링, f(x\_i)는 스트링 x\_i의 적합도, f( $\hat{x}_{iy}$ )는  $\hat{x}_{iy}$ 의 적합도를 나타낸다.

- (4) 각각 M개의 스키마에 대해, 가장 큰 개선 값을 갖는 최적 개체가 원래의 호스트-개체군의 대응되는 개체에 대체된다. 대체 결과, 호스트-개체군의 유용한 스키마는 증가하고 유용하지 않은 스키마는 감소한다. 이때, 원래의 호스트-개체군에 존재하지 않는 새로운 스키마가 추가될 수 있다.
- (5) 기생 개체군이 유용한 스키마를 찾는 역할을 하기 때문에 스키마 y의 적합도 F\_y는 (3)과 같이 적합도 개선 값의 합으로 정의된다.

$$F_y \equiv \sum_{i=1}^n \max[0, \hat{f}_{iy}] \quad (5)$$

식 (5)는 기생 개체군에서 스키마의 적합도가 호스트 개체군에서 대체된 스트링에 의존하고 있는 것을 의미한다. 본 알고리즘에서 기생 개체군은 호스트 개체군에 유용한 스키마를 전달하여 전역적 해의 수렴성을 보장하고 수렴 속도를 개선한다.

4. 게임이론에 기반한 공진화 알고리즘

4.1 게임이론에 기반한 공진화 알고리즘

진화적 게임의 형태를 가지는 공진화는 적정 세대가 지나면 안정된 균형 상태에 도달하고 이 상태는 '지배 특성 (dominance property)' 관점에서 게임의 최적해에 해당하므로 수학적 관점에서 이러한 공진화는 게임 이론적인 특성과 다이나믹스를 가진다. 이러한 특성들로부터 공진화 알고리즘을 게임 행렬을 통하여 구성하고 진화적 게임의 형태를 가지는 문제들에 대한 해법으로 제시될 수 있음을 알 수 있다. 본 논문에서는 '다목적 함수 최적화 문제 (Multi-objective Optimization Problems: MOPs)'의 풀이를 통한 알고리즘을 제안한다. 공학자들이 접하는 실제 최적화 문제의 대부분은 서로 결합하는 두개 이상의 목적 함수를 동시에 최적화해야 하는 문제를 포함하게 되는데 하나의 목적 함수를 가진 '단목적 함수 최적화 문제 (Single-objective Optimization Problems: SOPs)'의 경우에는 명확한 최적해가 존재하지만 다목적 함수의 최적화 문제의 경우는 그렇지 않다. 일반적으로 MOPs의 경우에는 하나의 최적해가 존재하는 것이 아니라 '파레토 최적해 집합 (Pareto optimal solution set)'이라고 알려진 해들의 집합이 존재하는데 이는 탐색 공간에서 모든 목적 함수를 고려할

때 이러한 해들의 집합보다 우위를 점하는 해가 존재하지 않는다는 점에서 최적해임을 의미한다.

4.2 MOPs를 풀기위한 GCEA의 설계

일반적인 ‘다목적 함수 최적화 문제(Multi-objective Optimization Problems: MOPs)’는  $n$ 개의 ‘결정 변수(decision variable)’,  $k$ 개의 ‘목적함수(objective function)’ 그리고  $m$ 개의 ‘제약조건(constraint)’을 포함한다.  $x, y$  두 개의 결정 변수를 가지고 두 개의 목적 함수  $f_1(x, y), f_2(x, y)$ 를 최적화 하는 풀기위한 게임이론에 기반한 공진화 알고리즘(GCEA)은 다음과 같다.

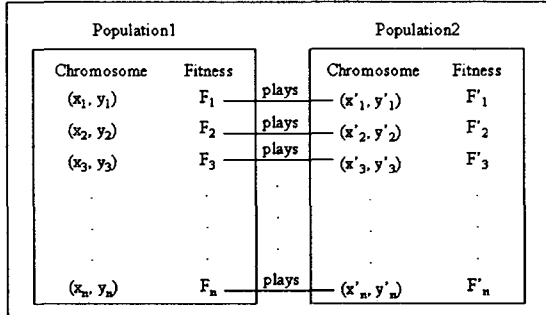


그림 6. GCEA의 개체군들 구조도

그림 6에서 적합도  $F_n$ 은 게임 행렬(G)로부터 계산되어 진 게임의 Pay-off,  $G_n$ 으로부터 결정되는데 (6)식과 같이  $G_n$ 은 두개의 목적함수 값 이용해 계산한다.

$$G_1((x_n, y_n), (x'_n, y'_n)) = f_1(x_n, y_n) - f_2(x'_n, y'_n) \quad (6)$$

$$G_2((x_n, y_n), (x'_n, y'_n)) = f_2(x'_n, y'_n) - f_1(x_n, y_n)$$

식(6)으로부터 두 게임 참가자의 보상값이 결정되면 두 개체의 적합도는 (7)에 의해 계산된다.

$$F_n = 100 \times G_1((x_n, y_n), (x'_n, y'_n)) / \alpha \quad (7)$$

$$F'_n = 100 \times G_2((x_n, y_n), (x'_n, y'_n)) / \alpha$$

식(7)에서  $\alpha$ 는 적합도  $F_n$ 과  $F'_n$ 을 normalize하기위한 상수이므로  $\max |G_k((x_n, y_n), (x'_n, y'_n))|$ 이 되어야 한다. 이 식들을 이용하여 GCEA는 다음과 같이 구성된다.

[게임이론에 기반한 공진화 알고리즘]

- (1) 그림 1과 같이 유전적으로 독립된 두개의 개체군 Population 1과 Population 2를 생성한다.
- (2) 첫 번째 개체군에서 선택된 첫 번째 게임 참가자는 두 번째 개체군에서 선택되어진 두 번째 게임 참가자와 게임을 수행하고 (3)식으로부터 보상값을 지불 받는다.
- (3) 두 번째 개체군의 게임 참가자는 표 2와 (3)식으로부터 보상값을 지불 받는다.
- (4) 두 참가자의 적합도가 (4)식에 의해 결정된다.
- (5) 두 개체군의 모든 참가자에 대하여 (2)와 (3)의 과정을 수행한다.
- (6) 이렇게 평가된 적합도에 기반하여 각 개체군의 자손 세대가 유전자 알고리즘을 통하여 독립적으로 재생산된다.
- (7) 두 개체군의 진화가 static state에 도달할 때까지 (2)에서 (6)까지의 과정을 반복한다.

위 알고리즘을 통하여 static state에 도달한 후 얻어진 모든 개체들은 MOPs의 파레토 평면에 존재하는 해 집합

이 된다.

5. 결 론

본 논문에서는 GA로 풀기 어려운 문제를 위한 공진화 알고리즘을 제안하였다. 공진화 알고리즘은 두 개 이상의 개체군이 상호 작용하여 진화함으로서 동적인 적합도 지형을 제공하고 이를 통해 지역적 최적해에 조기수렴하는 것을 막으며 수렴속도를 향상시킨다. 주 목적인 대상과 환경을 각각 개체군으로 구성하여 경쟁시키는 경쟁적 공진화 알고리즘과 스키마 개체들로 구성된 기생 개체군을 도입한 스키마 공진화 알고리즘 MOP문제를 위한 게임이론에 기반한 알고리즘을 제안하고 설명하였다.

감사의 글 : 본 연구는 과학기술부의 뇌신경정보학연구사업의 연구비 지원으로 수행되었습니다. 연구비 지원에 감사드립니다.

참 고 문 헌

- [1] J.H. Holland, *Adaptation in Natural and Artificial Systems*, Univ. of Michigan Press, 1975.
- [2] D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.
- [3] W.D. Hillis, "Co-evolving parasites improve simulated evolution as an optimization procedure," *Artificial Life II*, Addison-Wesley, pp. 313-324, 1991.
- [4] J. Paredis, "Coevolutionary Computation," *Artificial Life*, vol. 2, no. 4, pp. 355-375, 1995.
- [5] 정치선, 이동욱, 전효병, 심귀보, "경쟁적 공진화법에 의한 신경망의 구조와 학습패턴의 진화," *대한전자공학회 논문지*, 제36권 S편, 제1호, pp. 29-37, 1999. 1.
- [6] K.B. Sim, D.W. Lee, and J.Y. Kim, "Game theory based coevolutionary algorithm: A new computational coevolutionary approach," *Int. J. of Control, Automation, and Systems*, vol. 2, no. 4, pp. 463-474, 2004.
- [7] M.A. Potter and K.A. Dejong, "Cooperative coevolution: An architecture for evolving co-adapted subcomponents," *Evolutionary Computation*, vol. 8, no. 1, pp. 1-29, 2000.
- [8] C.A. Pena-Reyes and M. Sipper, "Fuzzy CoCo: a cooperative-coevolutionary approach to fuzzy modeling," *IEEE Trans. on Fuzzy Systems*, vol. 9, no. 5, pp. 727-737, 2001.
- [9] T. Fukuda, N. Kubota, and K. Shimojima, "Virus-evolutionary genetic algorithm and its application to traveling salesman problem," *Evolutionary Computation: Theory and Applications*, World Scientific, ch. 7, pp.235-255, 2000.
- [10] H. Handa, T. Horiuchi, O. Katai, and M. Bada, "A novel hybrid framework of coevolutionary GA and machine learning," *Int. J. Computational Intelligence and Applications*, vol. 2, no. 1, pp. 33-52, 2002.
- [11] K.B. Sim and D.W. Lee, "Schema co-evolutionary algorithm (SCEA)," *IEICE Trans on Information and Systems*, vol. E87-D, no. 2, pp. 416-425, 2004.
- [12] K.B. Sim, K.S. Byun, and D.W. Lee, "Design of fuzzy controller using schema coevolutionary algorithm," *IEEE Trans. on Fuzzy Systems*, vol. 12, no. 4, 2004.
- [13] D. Cliff, G.F. Miller, "Tracking the red queen: Measurements of adaptive progress in co-evolution," *COGS Technical Report CSRP 363*, Univ. of Sussex, 1995.