진화적 반복 죄수 딜레마 게임의 수렴 특성 분석

(Analysis on the Convergence Properties of Evolutionary IPD Games)

지도교수 : 장병탁

이 논문을 공학학사 학위 논문으로 제출함.

2003 년 6 월 30 일

서울대학교 공과대학

컴 퓨 터 공 학 부

이 태 양

2003 년 8 월

1. 서론

1.1 연구 목적

N명이 참가하는 반복적 죄수 딜레마 게임(NIPD game)의 전략은, 임의의 알고리즘으로부터 유전자 알고리즘(GA, Genetic Algorithm)을 사용해 진화적으로(evolutionarily) 만들어 질 수 있다. NIPD 게임의 대표적인 연구자 Axelrod(, Robert M.)는 몇 가지 실험을 통해 주고 받기 전략(TFT, Tit-For-Tat)이 평균적으로 가장 우수한 전략임을 보였으며, Yao와 Darwen은 NIPD 게임에 GA를 적용할 때, 개체 군이 협동으로 수렴할 수 있음을 보여주는 실험을 했다.
죄수 딜레마가 '배신'을 유도한다는 점에서 '딜레마'로 불린다는 사실을 생각하면, 이기적인 개체들이 GA를 통해 협동으로 수렴할 수 있다는 사실은 시사하는 바가 크다. 그러나 GA는 수행 환경의 미세한 차이에도 결과의 변화가 크다는 약점이 있으며, 따라서 협동으로 수렴하는 과정을 자세히 연구해 볼 필요가 있다.

1.2 연구 내용

죄수 딜레마 게임(prisoner's dilemma game)은 게임 이론에서 다루는 가장 간단한 형태의 게임이지만, 사회학, 경제학, 철학 등 많은 분야에서 흔히 발견되는 형태의 게임이기도 하다. 본 논문에서는 GA를 사용해 NIPD 게임의 전략을 구하는 데, 있음직 하지 않은 '최적 전략'을 찾기 보다는 다음의 질문에 대답한다.

1) 참가자들이 자기의 이익에 최선을 다 할 때 그들의 평균 이익은 어떤 방식으로 변화하는가?
2) NIPD 게임에 GA를 사용할 때, history count와 population 및 game count는 어떤 역할을 하는가?
3) 개체 군이 협동으로 수렴할 때의 우위 전략들은 발견적(heuristic) 전략들과 어떤 차이를 보이는가?

오로지 자기의 이익에만 최선을 다하는 IPD 게임의 모델은 GA를 사용하여 비슷하게 모사(simulate)될 수 있는데, Yao와 Darwen의 실험[YaDa2000]에서 알 수 있는 것처럼 이기적 참가자들 사이에서도 상호 협력 관계가 구축될 수 있으며, 이 관계가 구축되는 확률은 게임 크기(N)에 반비례한다. history count는 개체 군이 협동으로 수렴하는 것에 영향을 주지 않으나, 진화 과정을 좀 더 안정적으로 만들어 준다. population과 game count는 진화적 NIPD 게임을 시뮬레이션하는 데 있어서 매우 중요한 파라미터이며, 충분한 수의 population과 game count로 실험해야만 결과를 제대로 해석할 수 있다. 마지막으로, 협동으로 수렴하는 개체 군의 우위 전략들은, 대표적인 발견적 전략인 트리거 전략과 주고 받기 전략과 비교할 때 상대와 나의 행동을 종합적으로 검토한다는 점에서 한층 복잡하다는 것을 실험을 통해 알 수 있다.

1.3 연구 범위

본 논문에서, NIPD 게임의 '전략'은 상대와 나의 과거 행동(협동/배반), 다시 말해 '히스토리'에만 의존하는 것으로 가정한다. 다시 말해, 각 개체에게 다른 종류의 정보는 제공되지 않는다. 상대적으로 기억 장소를 많이 사용하는 GA의 특성으로 인해, 히스토리의 수는 보통 2~4개 정도로 제한된다.

또한, 넓은 의미의 진화 연산(Evolutionary computation)에는 다음의 네 가지 종류가 있다.

유전 알고리즘 (GA, Genetic Algorithm)
유전 프로그래밍 (GP, Genetic Programming)
진화 프로그래밍 (EP, Evolutionary Programming)
진화 전략 (ES, Evolution Strategy)

본 논문에서는 일반적인 GA를 사용하여 실험을 수행한다.

1.4 논문의 구성

2장에서는 죄수 딜레마 게임의 특성과, 그것에 대한 진화적 접근 방법에 대해 설명한다. 3장에서는 2장에서 설명한 방법을 사용해 Yao와 Darwen의 실험을 재현하고, history count와 game count를 바꾸어서 다시 실험하며, 4장에서는 3장의 실험 결과들을 분석한다. 5장은 결론이다.

2. 관련 연구

2.1 죄수 딜레마 게임

1980년대 한국, 전두환 독재 정권 시절을 생각해 보자. 두 명의 학생이 국가 보안법 위반으로 안기부에 검거 되었다. 둘은 서로 다른 방에서 각자 자백을 강요 당한다. 만일 두 학생이 모두 혐의를 인정하지 않는다면(협동), 괘씸죄에 걸려 5년간 징역을 살게 된다. 한 명만이 인정(배신)한다면 그는 풀려나고, 다른 학생은 종신형을 언도 받는다. 둘 모두 혐의를 인정(배신)하면, 둘 다 20년 형을 받는다고 하자.
가장 좋은 방법은 둘 모두 혐의를 부인하는 것이다. 그러나 서로 상대방의 선택을 알 수 없으므로, 결국 둘 다 배신을 하게 된다. 이 경우 배신은 게임 이론에서 말하는 '절대 우위 전략', 즉 상대가 어떤 선택을 하더라도 자기에게 유리한 전략이 된다. 하지만 둘 모두 이 전략을 택한 결과 각자 20년형 살게 되었다. 이것을 '죄수 딜레마'라고 한다.
죄수 딜레마 게임은 참가자들이 사전에 모의할 수 없는(즉, 서로의 선택을 알지 못하는) 상태에서 이루어지는 non-zerosum 게임이며, 게임의 참가자들은 매 순간마다 '배반' 혹은 '협동' 중 하나를 택한다.

참가자가 얻는 이득을 다음과 같이 표로 나타낼 수 있다.

	상대의 협동	상대의 배반
나의 협동	R	S
나의 배반	T	P

표 1 – 2IPD 게임의 이득표

단, T > R > P > S 이고 2R > T + P 이다.

죄수 딜레마 게임은 N명의 참가자들에 대해서도 확장될 수 있다. 이 경우 각 참가자가 얻는 이익은 다음과 같다.

나를 제외한 협동자의 수	0	1	…	i	…	N – 1
나의 협동	C(0)	C(1)		C(i)		C(N-1)
나의 배반	D(0)	D(1)		D(i)		D(N-1)

표 2 – NIPD 게임의 이득표

단, x = 0, 1, 2, …, N-1인 모든 x에 대해,
D(x) > C(x) 이고 D(x+1) > D(x), C(x+1) > C(x) 이며 C(x) > (D(x) + C(x-1)) / 2 이다.

함수 C와 D는 여러 가지가 있을 수 있다. 예를 들어

C(x) = 2x
D(x) = 2x + 1

는 위의 제약 조건들을 만족한다. 또한 이 경우, N명의 참가자 중 k명이 협동했다면(따라서 N-k명이 배신), 참가자의 전체 이익 S(k)는

S(k) = 2(k – 1)k + (2k + 1)(N – k)
S(k+1) = 2k(k + 1) + (2k + 3)(N – k – 1)
S(k)의 계차 수열 B(k) = S(k+1) – S(k) = 2N – 3
따라서, S(n) = S(0) + (B(0) + B(1) + … + B(n-1)) = (2N – 3)n + N

그러므로, 참가자들의 평균 이익 P(n)은

$P(n) = S(n) / N = 1 + (2N - 3)\, n / N$

이 결과로부터, P(n)을 조사하기만 하면 협동하는 비율(n / N)을 구할 수 있는 장점이 있다.

2.2 진화적 접근

유전자 알고리즘(GA, Genetic Algorithm)이란 생물학의 진화 현상을 흉내낸 계산 모델이다. 이것은 풀고자 하는 문제에 대한 가능한 해들을 정해진 자료 구조로 표현한 다음, 이들을 무작위로 서서히 변화시키면서 좋은 해들을 남기고 나쁜 해들을 제거하는 방법이다. 따라서 자료 구조로 표현된 하나의 해는 진화 생물학에서의 개체(individual)에 대응하며, 나쁜 해들을 제거하는 것은 자연 선택에, 해를 변화시키는 것은 교배와 돌연변이에 각각 대응한다. 유전자 알고리즘의 대략적인 흐름은 다음과 같다.

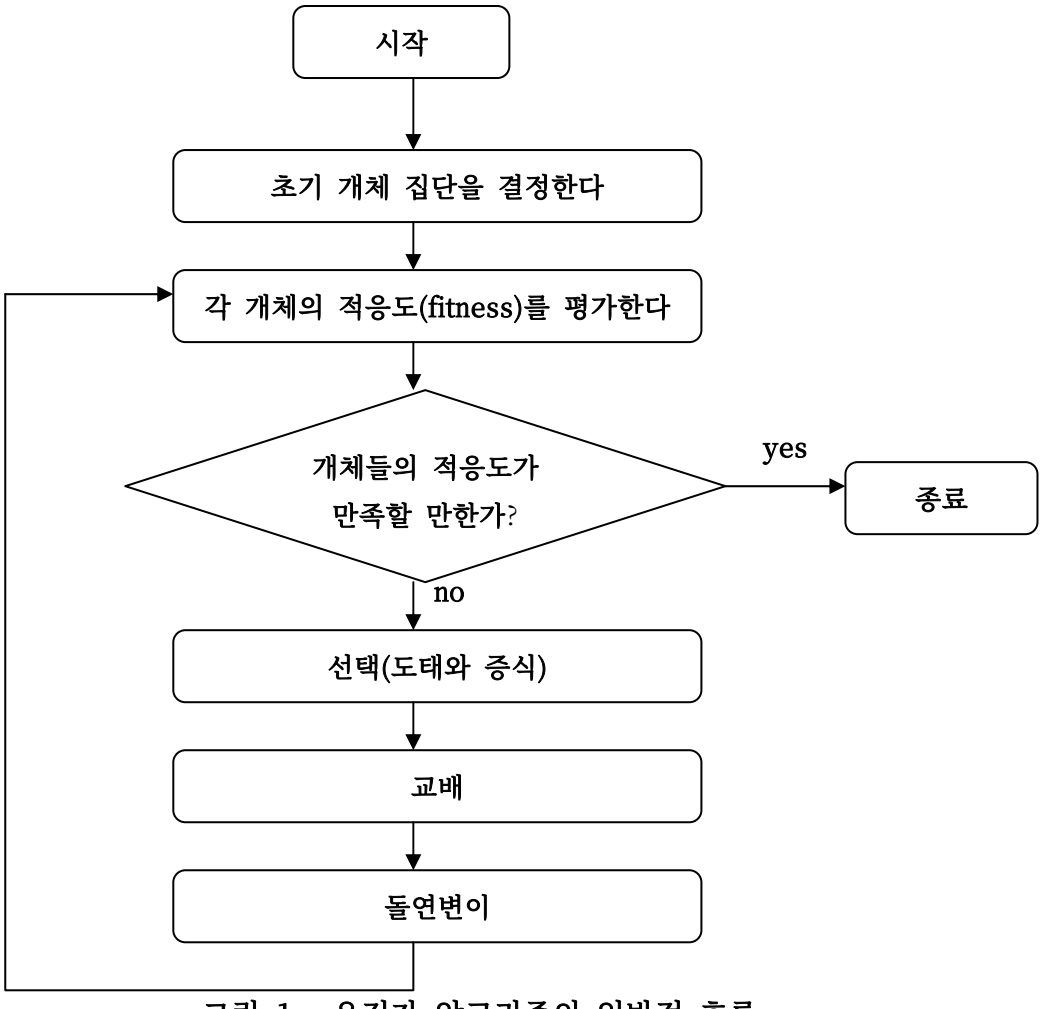


그림 1 – 유전자 알고리즘의 일반적 흐름

GA는 원하는 해를 구하는 결정적(deterministic) 방법이 없거나, 혹은 비 효율적일 경우 사용한다. GA는 본질적으로 무작위 스크리닝(random screening)이며, 복수의 개체들 사이에 선택이나 교배 등의 유전적 조작에 의해 상호 협력적으로 해를 탐색하므로 단순한 무작위 스크리닝보다 좋은 결과를 낳을 수 있다. NIPD 게임의 최적 해를 구하는 결정적 방법은 알려져 있지 않으며, 최적 해가 있는지 조차 의심스럽다는 점에서 GA는 NIPD와 매우 잘 어울리는 알고리즘이라 하겠다

.

유전자 알고리즘에서 중요한 매개 변수들(GA parameters)은 다음과 같다

개체 수(population)

: 개체 수가 많으면 그만큼 다양한 해(유전자)들이 존재하게 되지만, 계산 시간이 많이 걸린다.

적응도(fitness) 및 평가

: 개체의 적응도를 평가하는 방법은 유전자 알고리즘의 핵심이다.

선택(selection)

: 적응도가 높은 개체는 많은 자손을, 적응도가 낮은 개체는 상대적으로 적은 자손을 가지게 한다. 선택 방법에도 여러 가지가 있는데, 적응도 비례 방식, 엘리트 보존 방식, 기대치 방식, 순위 방식 등이 있으며, 이들의 혼합하여 사용할 수도 있다.

교배(crossover)

: 교배는 복수의 부모 개체들로부터 자식 개체를 구성하는 방법이다. 부모 개체들의 유전자를 섞음으로써 다양한 자식 개체를 만들어 낼 수 있다. 교배 방법에도 단순 교배, 부분 일치 교배, 순서 교배 등 여러 가지가 있다.

돌연변이(mutation)

: 돌연변이는 개체의 유전자에 변화를 줌으로써 유전자를 좀 더 다양하게 하는 방법을 말한다. 유전자의 일부분을 임의적으로(randomly) 바꿀 수도 있고, 순서를 바꾸거나 부분적으로 중복 혹은 삭제 하는 등 여러 가지 방법이 있다.

2.3 Axelrod의 표현 방법

GA를 사용하기 위해서는 찾고자 하는 해(solution)를 정해진 자료 구조(유전자)로 표현할 수 있어야 한다. 죄수 딜레마 게임에서 찾고자 하는 것은 '전략'이다. 각 참가자는 이전의 결과들을 기억할 수 있으며, 또한 각 참가자는 단 하나의 '전략'을 가지고 있다. '전략'은 참가자들의 기억(히스토리, history)만을 사용하여 다음 행동을 결정한다. N명의 참가자가 게임을 할 때, 하나의 결과(라운드, round)는 다음과 같이 N-bit로 표현할 수 있다.

0110110 ... 0011 k번째 비트는 k번째 참가자의 배신(1) 혹은 협동(0)을 나타낸다.

만일 각 참가자가 최근 h개의 결과를 기억한다면, 히스토리는 H = hN개의 bit로 표현된다. 하나의 '전략'은 가능한 H-bit의 모든 조합에 대해 다음 행동을 결정할 수 있어야 하므로, 2^H개의 bit로 표현된다. 그런데 게임을 처음 시작할 때에는 히스토리가 없으므로, 히스토리를 임의로 설정해 주어야 한다. 따라서 하나의 '전략'은 모두

$2^H + H$

개의 bit로 표현할 수 있다.

2.4 Yao와 Darwen의 표현 방법

Axelrod의 표현 방법은 참가자의 수에 따라 사용하는 메모리가 기하 급수적으로 커진다. 또한, 각 참가자에게는 누가 협동했는가 보다는 몇 명이 협동했는가가 중요하므로, 라운드를 다음과 같이 표현할 수 있다.

0 01101... 첫번째 비트는 자신의 배신(1) 혹은 협동(0), 나머지 비트들은 협동자들의 수

이렇게 하면, $N = 2^n$ 명의 참가자가 있는 경우 하나의 라운드를 n + 1 비트로 표현이 가능하다. h개의 라운드를 기억한다면, 히스토리는 H = h(n + 1) 비트가 되어 Axelrod의 방법보다 훨씬 적은 메모리를 사용하게 된다.

2.5 발견적(heuristic) 전략

진화적 접근에서는 고려하지 않지만, 죄수 딜레마 게임에서 가장 널리 알려진 전략에는 트리거 전략(trigger strategy)과 주고 받기 전략(TFT, Tit-For-Tat strategy)이 있다. 두 전략 모두 처음에는 협동으로 시작한다. 그러다가 상대가 배반을 하면 자기도 배반을 하기 시작하는데, 트리거 전략은 협동으로 돌아 오는 일이 없고, 주고 받기 전략은 상대가 협동을 하면 곧바로 협동으로 돌아 온다.
실제에 있어, 이 두가지 전략은 극단적인 전략에 속한다. 보통 사람의 경우, 상대의 배반을 몇 번 눈감아 줄 수 있으며, 자기가 배반을 시작했다면 트리거 전략처럼 영원히 협동으로 돌아오지 않는 것도, 주고 받기 전략처럼 곧바로 협동으로 돌아 오는 것도 아니다. 적절한 전략은 이 둘 사이의 어딘가에 위치하리라고 예상되며, 개체군이 서로 협동하는 방향으로 수렴하는 경우 이 '적절한 전략'이 대다수를 이룰 것이라고 기대한다. 이 문제는 4장에서 다시 논의된다.

참고로, 히스토리 수가 1인 2IPD 게임의 트리거 전략을 Axelrod의 방법으로 표현하면 다음과 같다.

히스토리	다음 행동
00	0
01	1
10	1
11	1

표 3 – 히스토리 1인 2IPD 게임의 트리거 전략

주고 받기 전략은 다음과 같다

히스토리	다음 행동
00	0
01	1
10	0
11	1

표 4 – 히스토리 1인 2IPD 게임의 주고 받기 전략

3. 실험

3.1 진화적 NIPD 게임의 흐름

NIPD 게임을 진화적으로 구성할 때, GA의 주요 개념은 다음과 같이 설명될 수 있다.

개체 수(population)
: 진화적 NIPD 게임에서, 하나의 개체에 해당하는 것은 하나의 '전략'이다. 개체 수가 많을수록 GA는 좀 더 다양한 전략을 탐색할 수 있게 된다.

적응도(fitness) 및 평가
: 개체 군에 속한 임의의 개체들과 일정 횟수 이상의 IPD 게임을 해서 높은 payoff를 가질수록 높은 적응도를 가진 것으로 평가한다. 또한, 개체 군의 평균 적응도가 일정 조건을 만족하면, 해를 찾은 것으로 간주하고 GA를 종료한다.

GA를 사용한 NIPD 게임의 흐름을 표로 나타내면 다음과 같다.

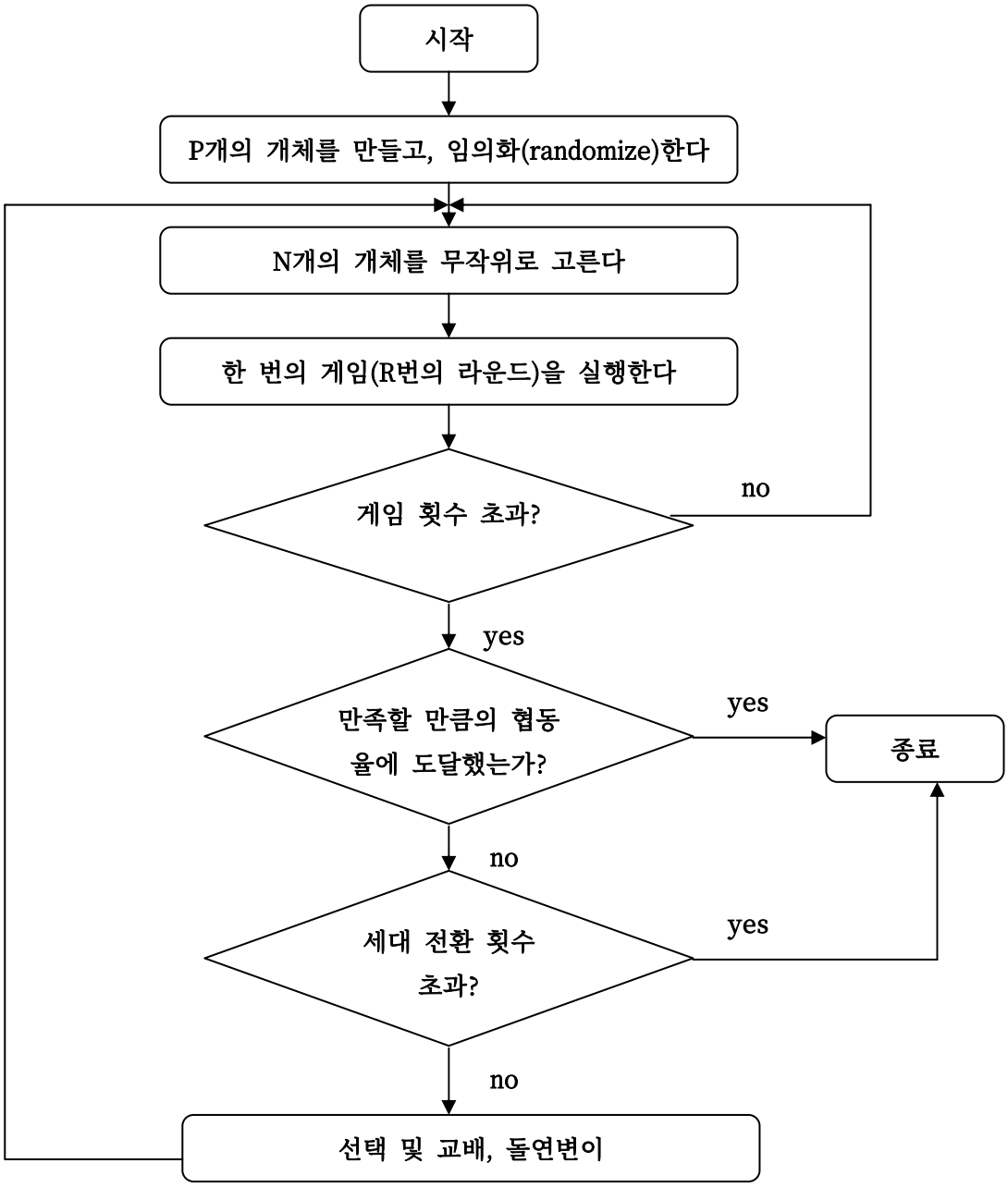
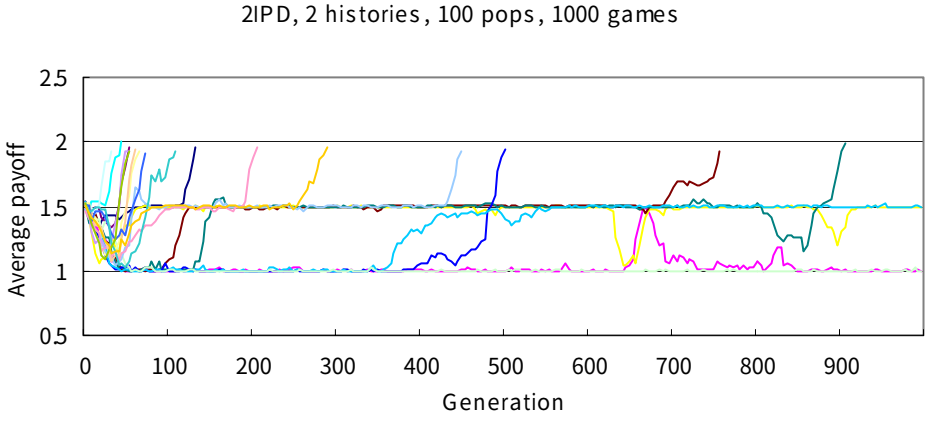


그림 2 – 진화적 NIPD 게임의 일반적 흐름

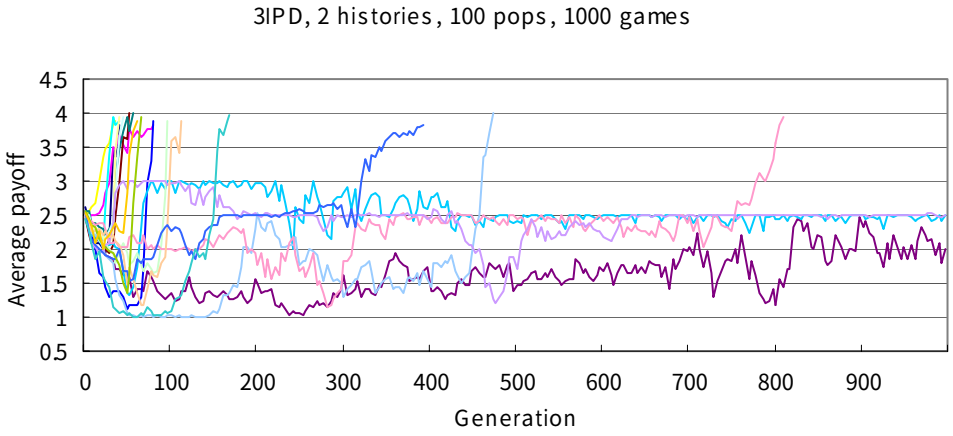
3.2 첫번째 실험 : history 2, population 100, game count 1000

본 논문에서는 Yao-Darwen의 방법을 사용하며, 매개 변수는 다음과 같다
개체 수 : 100
최대 세대 전환 횟수 : 1000
세대당 게임 횟수 : 1000
게임당 라운드 수 : 100
히스토리 수 : 2
돌연변이율 : 0.001
교배율 : 0.6
교배 방법 : 임의의 위치에서의 1-point crossover
선택 방법 : 순위 기반 선택(rank-based selection). 개체 당 최대 1.25, 최소 0.75개의 자손.
Payoff : C(x) = 2x, D(x) = 2x + 1

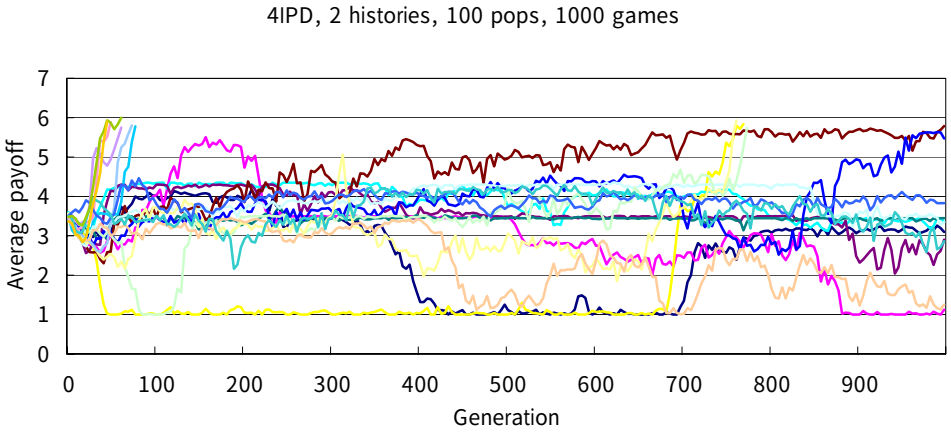
위의 매개 변수를 사용해 2, 3, 4, 5, 6, 8, 16 IPD 각각 20번씩 실행(run)한다. 각 run은 협동율이 5세대 연속 95%를 넘는 경우 즉시 중지되며, 그렇지 않은 경우 최대 세대 전환 횟수까지 실행된다.



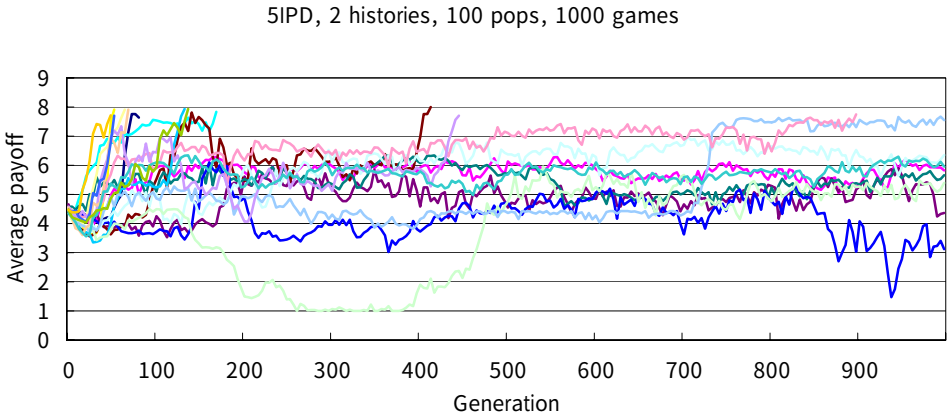
실험 결과 1 – 20회 중 16회가 협동으로 수렴, 평균 수렴 시간 191.6 세대



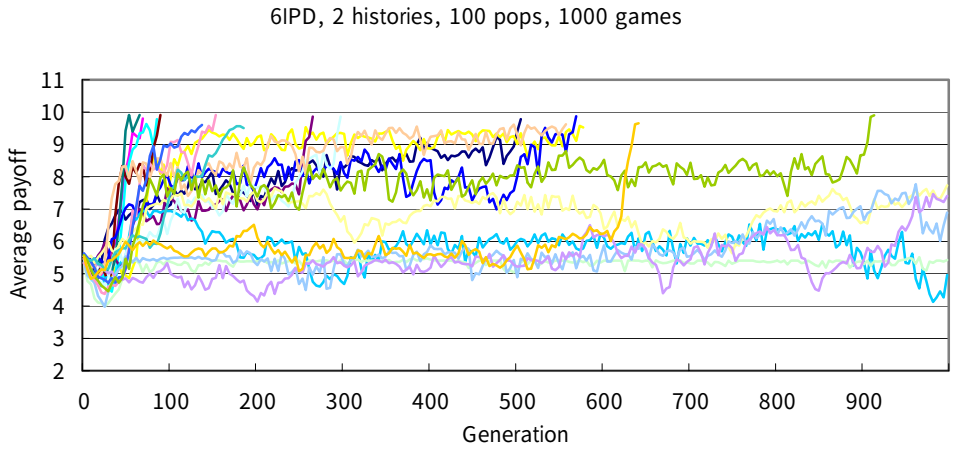
실험 결과 2 - 20회 중 17회가 협동으로 수렴, 평균 수렴 시간 135.4 세대



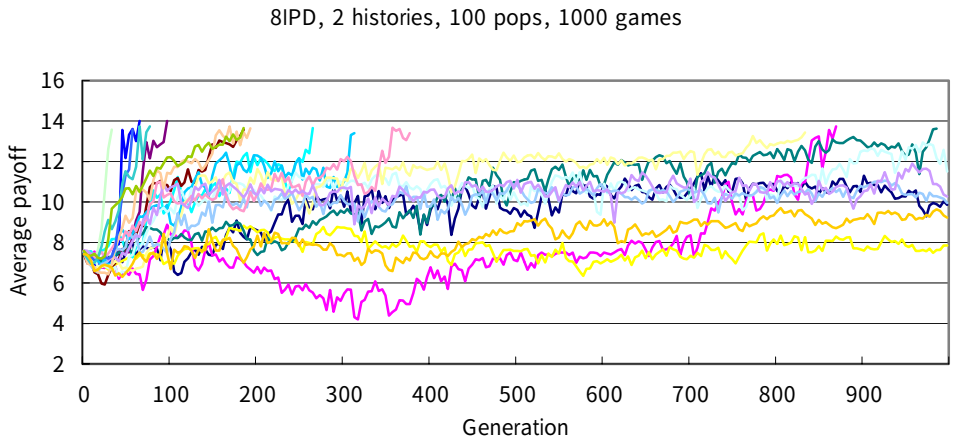
실험 결과 3 - 20회 중 9회가 협동으로 수렴, 평균 수렴 시간 134.8 세대



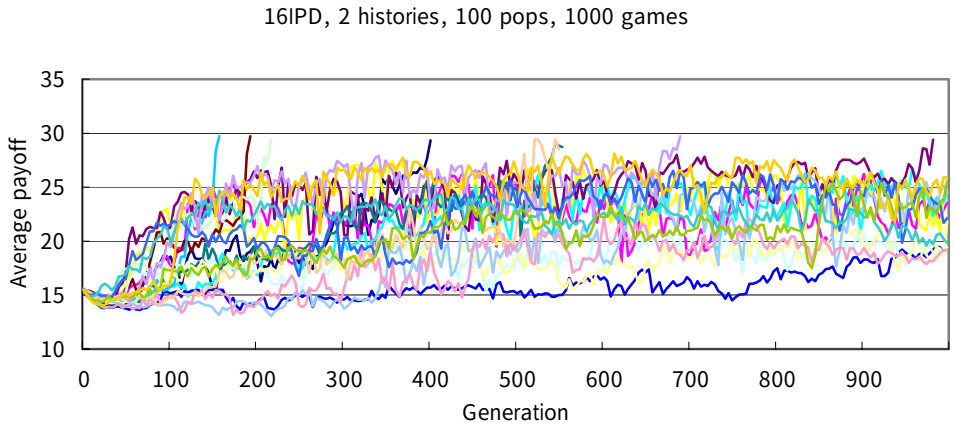
실험 결과 4 - 20회 중 12회가 협동으로 수렴, 평균 수렴 시간 130.0 세대



실험 결과 5 - 20회 중 15회가 협동으로 수렴, 평균 수렴 시간 257.6 세대



실험 결과 6 - 20회 중 14회가 협동으로 수렴, 평균 수렴 시간 228.8 세대



실험 결과 7 - 20회 중 8회가 협동으로 수렴, 평균 수렴 시간 188.2 세대

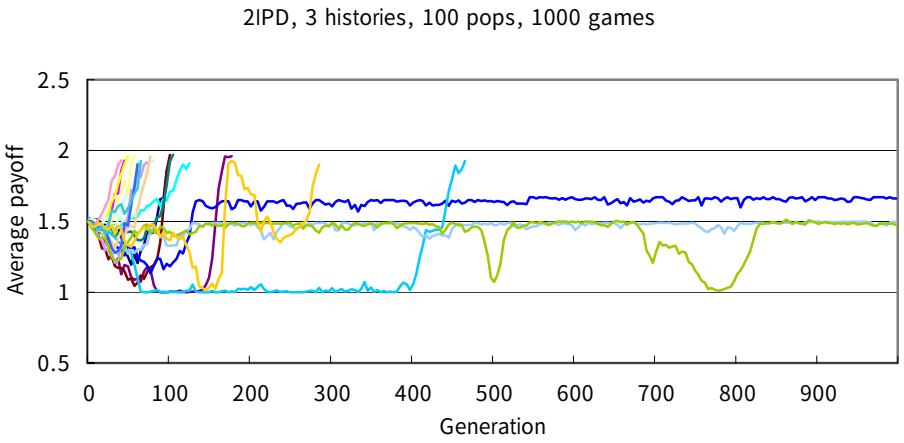
실험 결과를 정리하면 다음과 같다.

게임 크기	배신으로의 수렴 횟수 (total 20 runs)	평균 수렴 시간(세대) (수렴하지 않은 run 은 제외)
2IPD	4	191.6
3IPD	3	135.4
4IPD	11	134.8
5IPD	8	130.0
6IPD	5	257.6
8IPD	6	228.8
16IPD	12	188.2

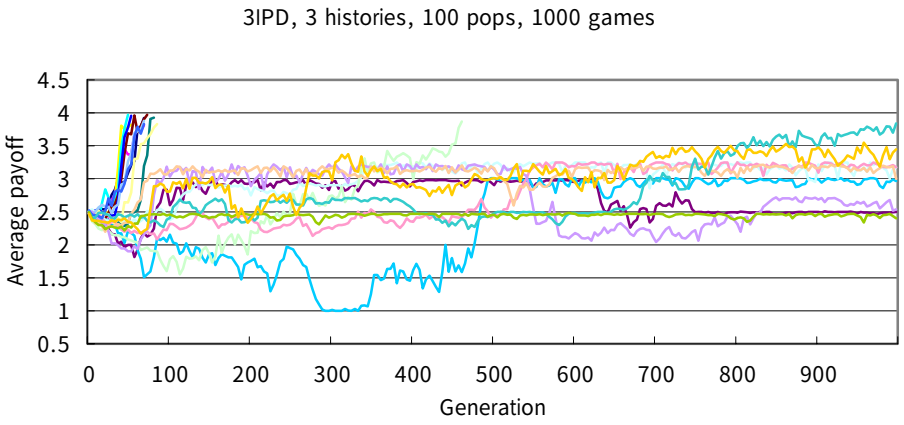
표 5 – history 2, population 100, game count 1000일 때의 실험 결과

3.3 두번째 실험 : history count 3, population 100, game count 1000

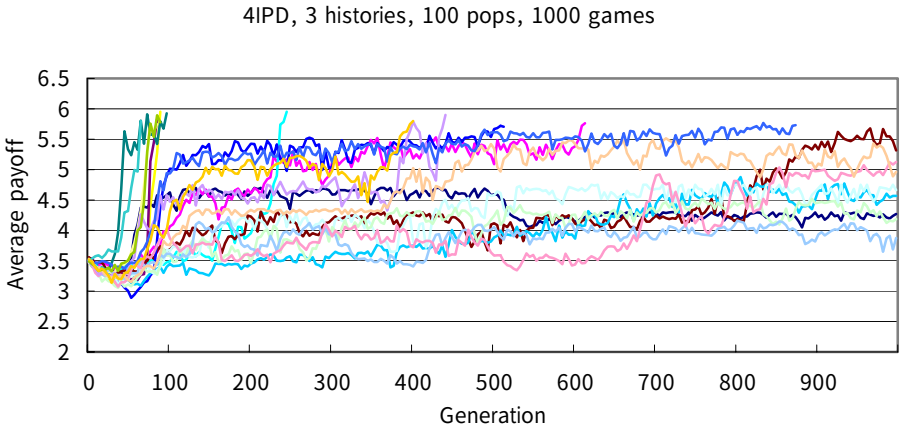
결과를 3개까지 기억하도록 한 후 실험을 반복했다.



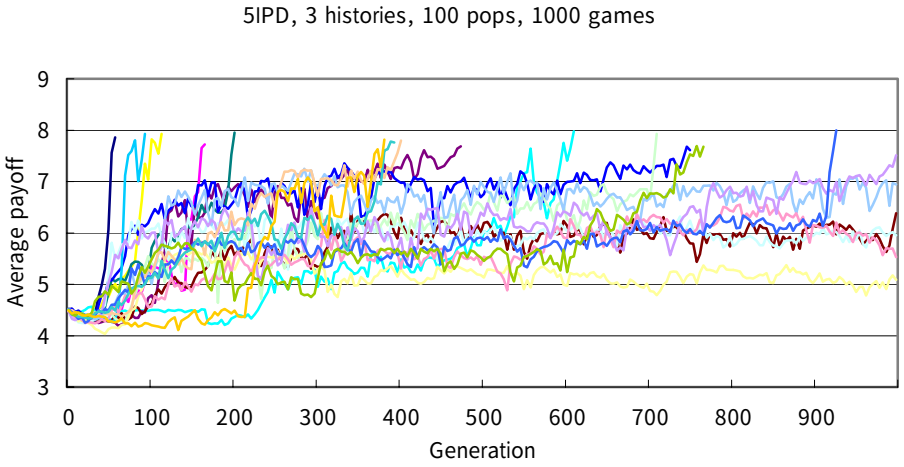
실험 결과 8 - 20회 중 17회가 협동으로 수렴, 평균 수렴 시간 101.2 세대



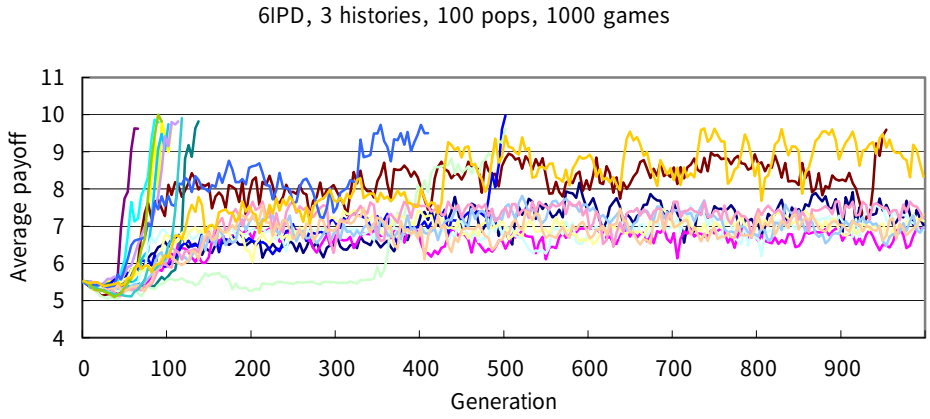
실험 결과 9 - 20회 중 11회가 협동으로 수렴, 평균 수렴 시간 58.2 세대



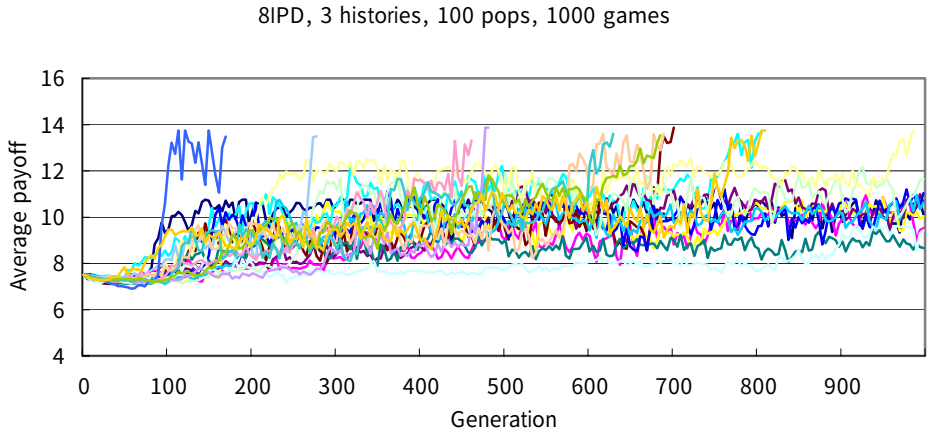
실험 결과 10 - 20회 중 12회가 협동으로 수렴, 평균 수렴 시간 181.6 세대



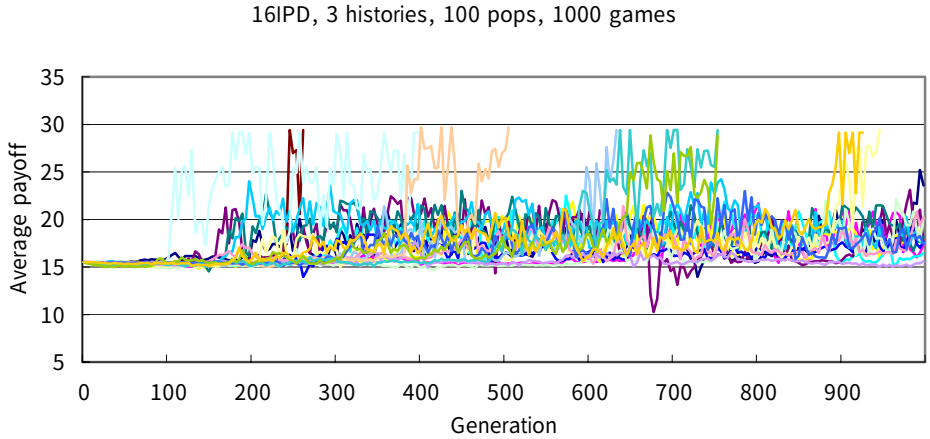
실험 결과 11 - 20회 중 14회가 협동으로 수렴, 평균 수렴 시간 303.8 세대



실험 결과 12 - 20회 중 12회가 협동으로 수렴, 평균 수렴 시간 161.0 세대



실험 결과 13 - 20회 중 11회가 협동으로 수렴, 평균 수렴 시간 336.2 세대



실험 결과 14 - 20회 중 8회가 협동으로 수렴, 평균 수렴 시간 259.8 세대

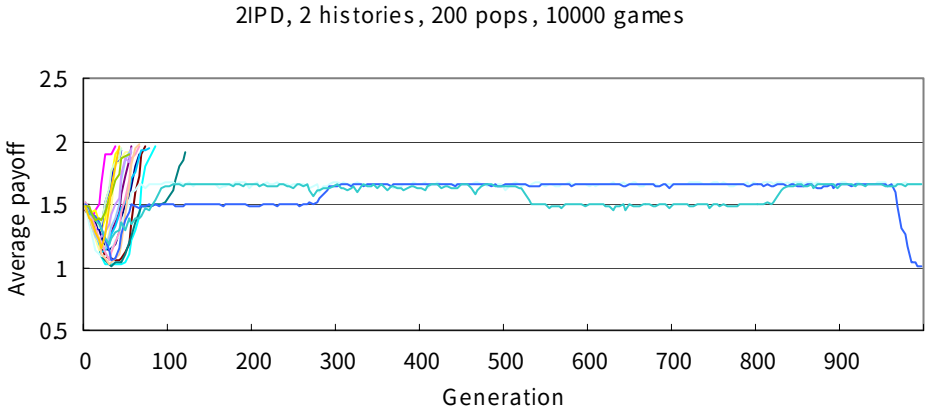
실험 결과를 정리하면 다음과 같다.

게임 크기	배신으로의 수렴 횟수 (total 20 runs)	평균 수렴 시간(세대) (수렴하지 않은 run 은 제외)
2IPD	3	101.2
3IPD	9	58.2
4IPD	8	181.6
5IPD	6	303.8
6IPD	8	161.0
8IPD	9	336.2
16IPD	12	259.8

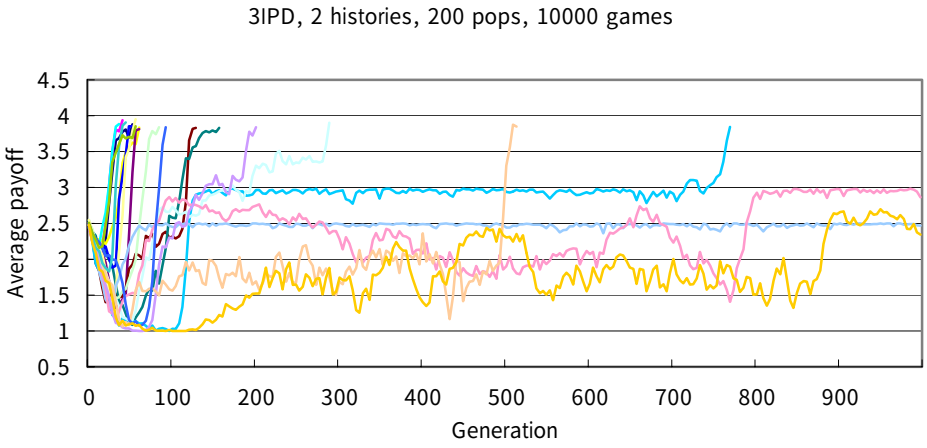
표 6 – history 3, population 100, game count 1000일 때의 실험 결과

3.4 세번째 실험 : history count 2, population 200, game count 10000

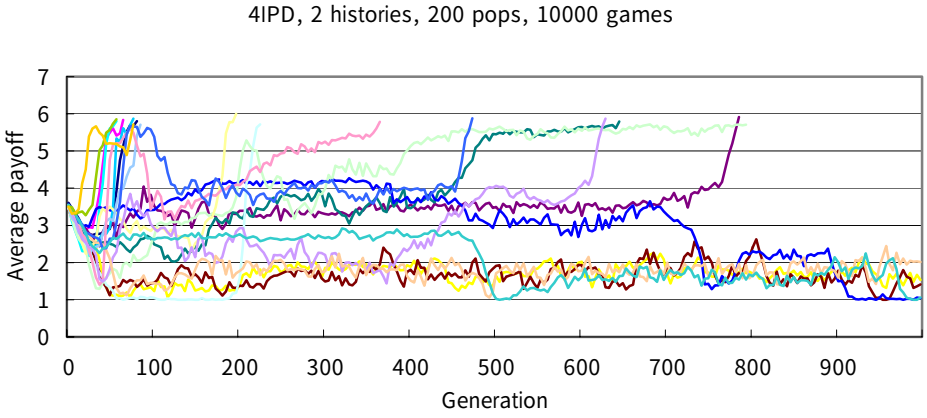
히스토리 수를 2개로 되돌려 놓고, 개체 수를 200으로, 세대 당 게임 수를 10000으로 증가시켜 다시 실험했다.



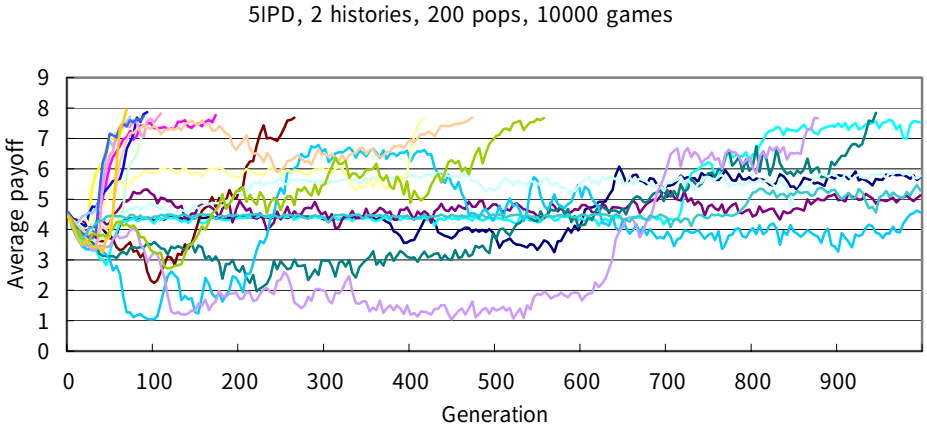
실험 결과 15 - 20회 중 17회가 협동으로 수렴, 평균 수렴 시간 54.4 세대



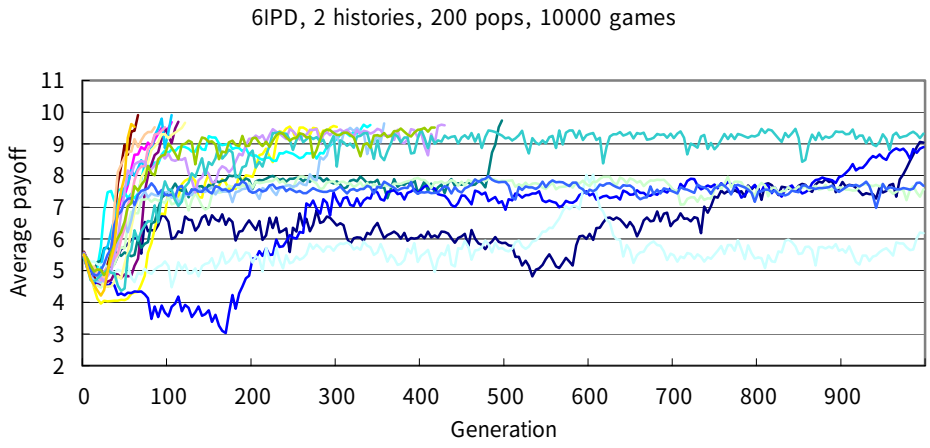
실험 결과 16 - 20회 중 17회가 협동으로 수렴, 평균 수렴 시간 137.6 세대



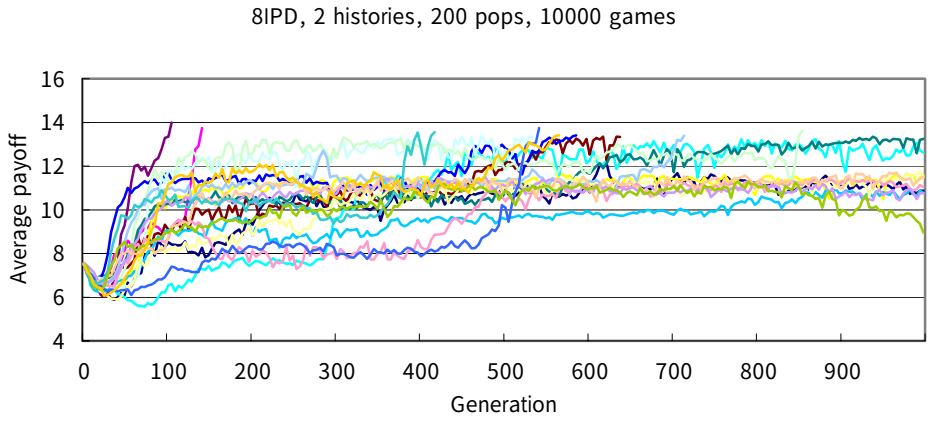
실험 결과 17 - 20회 중 15회가 협동으로 수렴, 평균 수렴 시간 232.8 세대



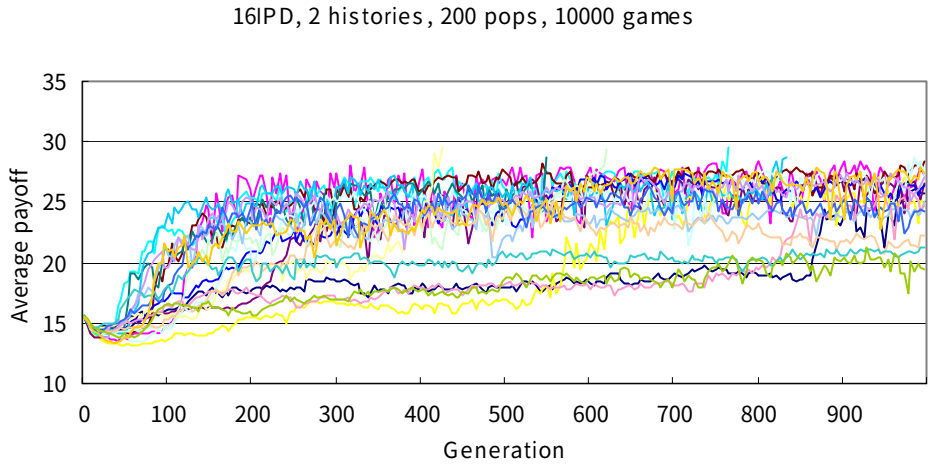
실험 결과 18 - 20회 중 14회가 협동으로 수렴, 평균 수렴 시간 218.0 세대



실험 결과 19 - 20회 중 14회가 협동으로 수렴, 평균 수렴 시간 156.8 세대



실험 결과 20 - 20회 중 10회가 협동으로 수렴, 평균 수렴 시간 256.0 세대



실험 결과 21 - 20회 중 5회가 협동으로 수렴, 평균 수렴 시간 160.4 세대

실험 결과를 정리하면 다음과 같다.

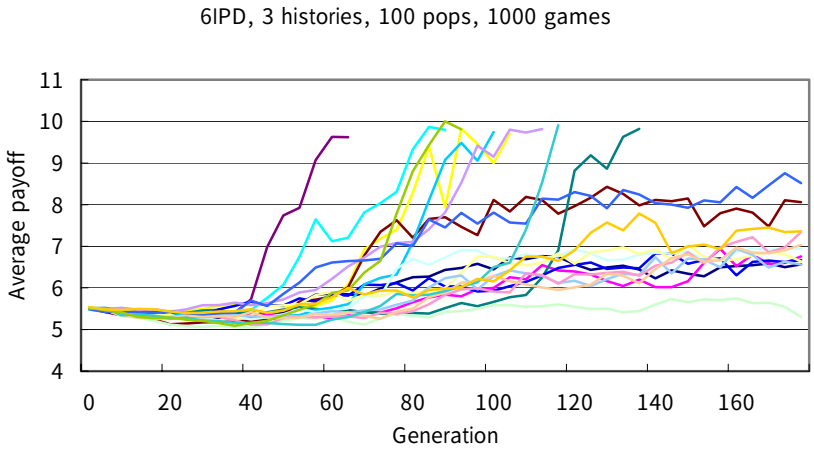
게임 크기	배신으로의 수렴 횟수 (total 20 runs)	평균 수렴 시간(세대) (수렴하지 않은 run 은 제외)
2IPD	3	54.4
3IPD	3	137.6
4IPD	5	232.8
5IPD	6	218.0
6IPD	6	156.8
8IPD	10	256.0
16IPD	15	160.4

표 7 – history 2, population 200, game count 10000일 때의 실험 결과

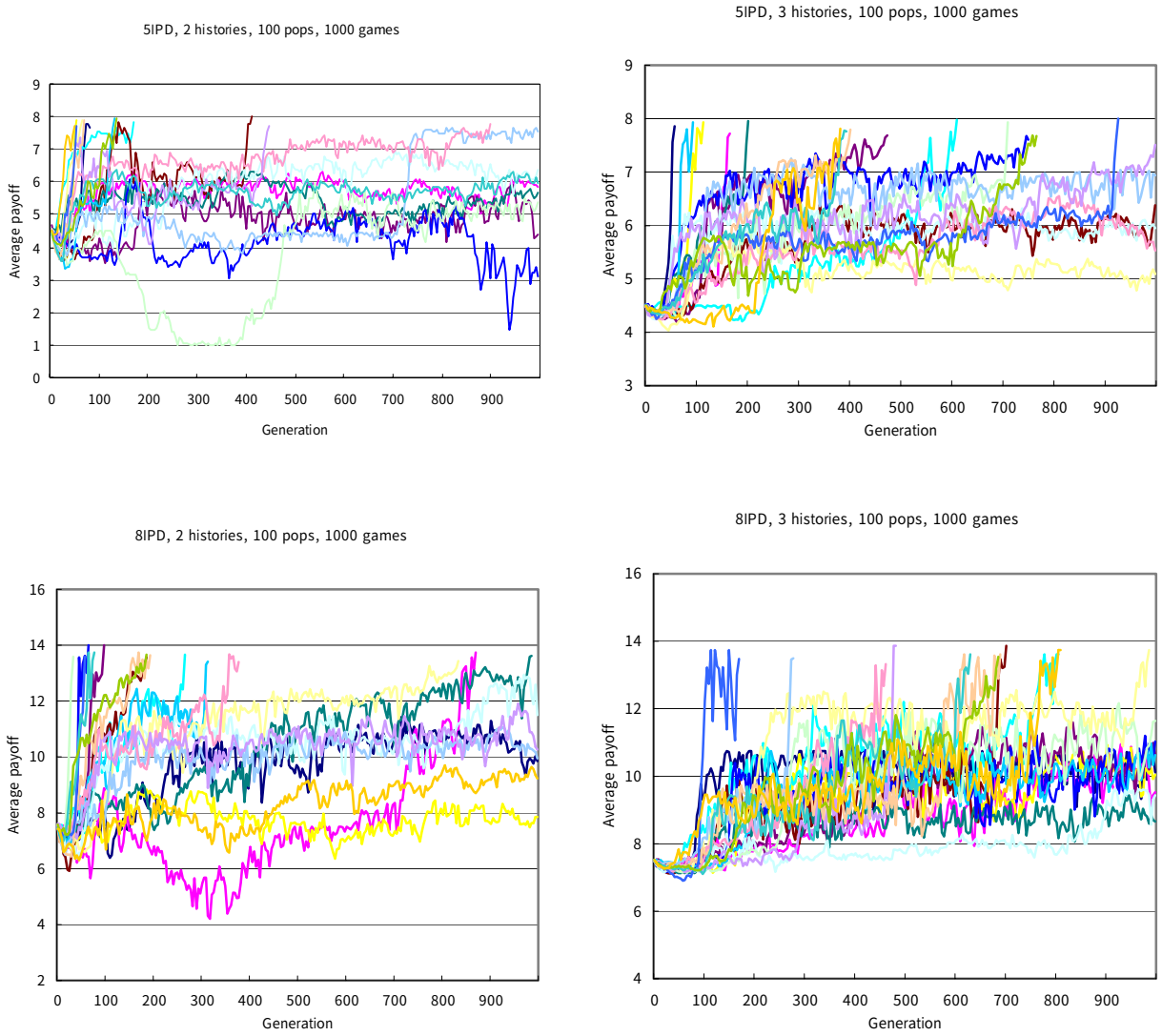
4. 결과 분석

4.1 history count와 population, game count의 영향

실험 결과, 히스토리 수는 협동으로 수렴하는 것에 별다른 영향을 끼치지 않았다. 다만, 히스토리 수가 3인 경우, 밑의 그림에서처럼 초반에 배반으로 기우는 경향이 다소 보이는 데, 이것도 미미한 수준이어서 전체 수렴 시간에 별다른 영향을 끼친다고 볼 수 없다.



한 가지, history count 3인 경우, history count 2인 경우보다 그래프의 상하 폭이 약간이지만 좁다. 이것은 history count가 늘어나면 개체 군의 진화가 좀 더 안정적인 것으로 해석할 수 있으나, 협동으로의 수렴을 강화하는 것은 아니다. 다음 그림은 5IPD(위)와 6IPD(아래)의 경우, history count 2(왼쪽)인 경우와 history count 3(오른쪽)인 경우를 비교한 것이다.



population 100, game count 1000인 경우(3.1, 4.1), 게임 크기(참가자의 수)가 작을수록 협동으로 수렴하는 경우가 많고, 평균 수렴 시간도 긴 것을 볼 수 있다. 그러나 이 경향은 그다지 뚜렷하게 나타나지는 않았으며, 예외가 많고 불규칙적이다.

population 200, game count 10000인 경우(4.2), 게임의 크기와 수렴 정도가 반비례(수학적인 반비례는 아니다)하는 경향이 더욱 뚜렷한데, (3.1, 4.1)과 비교해 2IPD는 좀 더 **협동으로** 수렴하고, 16IPD는 좀 더 **배반으로** 수렴했다.

게임 크기	배신으로의 수렴 횟수 (total 20 runs)
2IPD	3
3IPD	3
4IPD	5
5IPD	6
6IPD	6
8IPD	10
16IPD	15

표 8 - population 200, game count 10000일 때 게임 크기와 배신으로 수렴하는 정도의 관계

개체 군이 협동으로 진화하기 위해서는 배반을 잘하는 개체가 불리한 환경을 만들어야 하는데, 그러기 위해서는 게임 중 배반이 나왔을 경우 같이 배반(처벌)을 하고, 협력을 하면 같이 협력(보상)하는 분위기가 만들어져야 한다. 그런데 참가자가 많으면, 소수의 배반자를 처벌하기 위해 다수의 협력자들이 같이 손해를 보아야 하므로, 될 수 있는 한 처벌을 기피하게 될 것이다. 따라서 게임 크기와 협동으로의 수렴이 반비례 하는 경향은 이해될 수 있으며, 사실일 것으로 생각된다.

population이 100일 때 game count가 1000이면, 2IPD의 경우 한 세대에 개체 당 평균 20회의 게임을 하게 된다. 한 개체가 개체 군의 1/5과만 상호 작용하는 것은 충분하다고 볼 수 없으며, game count가 증가했을 때 반비례 관계가 뚜렷해지는 것으로 보아 game count가 매우 중요하다고 할 수 있다.

4.2 우위 전략의 분석

다음은 2IPD, 2 histories, 200 pops, 10000 games가 협동으로 수렴했을 경우, 상위 10개의 전략을 표시한 예이다. 괄호안의 것은 게임을 시작할 때의 히스토리(initial history)이다.
주의 : 계산을 조금이라도 빠르게 하기 위해, Yao와 Darwen의 유전자 표현 방식과는 약간 다르게 자기의 최근 행동을 각 히스토리의 맨 '마지막' 비트에 기록했다(Yao와 Darwen은 맨 처음 비트에 기록한다).

1111011111001100 (1010)
1110011110011100 (1010)
1110011111011101 (1010)

1110011111001100 (1010)
1110011111001111 (1010)
1111011111001100 (1010)
1110011111001100 (1010)
1110011111001111 (1010)
1110011111011101 (1010)
0110011111001101 (1010)

협동으로 수렴한 10번의 run들 각각의 상위 10개 전략, 총합 100개 전략을 분석한 결과, '전혀' 예외 없이 초기에는 협동으로 시작한다(initial history가 협동을 하게끔 setting 되었다). 또한, '거의 항상' 나타나는 비트들이 있는데, 다음과 같다.

1 1 ? ? ? 1 ? 1 1 ? 0 ? ? 1 ? ?

이것을 분석하면 다음과 같다.
(이하의 H(n, m)은 상대가 n번 배신하고 내가 m번 배신했다는 히스토리를 말한다. 예를 들어 H(1, 0)이면 상대가 최근 2번의 라운드 중 1번 배신하고, 나는 0번 배신했다는 뜻이다)

0번 비트 1은 H(2, 0)일 경우 반드시 배신한다는 뜻이다.
1번 비트 1은 H(2, 1)일 경우 반드시 배신한다는 뜻이다.
5번 비트 1은 H(2, 2)일 경우 반드시 배신한다는 뜻이다.
7번 비트 1은 H(1, 2)일 경우 반드시 배신한다는 뜻이다.
8번 비트 1은 H(1, 0)일 경우 반드시 배신한다는 뜻이다.
10번 비트 0은 H(0, 0)일 경우 반드시 협동한다는 뜻이다.
13번 비트 1은 H(1, 2)일 경우 반드시 배신한다는 뜻이다.

위의 결과를 일반화 하면 다음과 같다.

a. H(0, 0)이면 협동이므로, 상대와 내가 협동하고 있을 동안에는 무조건 협동한다.
b. H(1, 2), H(2, 2)이면 배신이므로, 내가 두 번 다 배신했을 경우에는 상대의 행동에 관계 없이 배신한다. 그러나 H(0, 2)의 경우 배신이 많지만 협동하는 경우도 있으므로, 내가 두 번 모두 배신을 해도 상대가 계속 협조적인 경우라면 협동할 수도 있다고 하겠다.[1]
c. H(2, 0), H(2, 1), H(2, 2)이면 배신, 즉 상대가 계속 배신하면 나 역시 무조건 배신이다.

[1] H(0, 2)의 경우에도 배신이 많은 것은, history count가 2로 제한적이기 때문인 듯하다. 내가 두 번 모두 배신한 것에는 이유가 있을 것이고(history count가 2이기 때문에 그 이유를 모두 기억하지는 못하지만), 상대가 겨우 2번 협동했다고 해서 자기도 협동으로 돌아서는 것은 현명하지 못한 것일 수 있다. 마찬가지 이유로, H(1, 2)는 거의 예외 없이 배신을 유지한다.

d. H(1, 0), H(1, 2)이면 배신이다. 즉, 내가 매우 협조적인데 상대가 한 번이라도 배신하거나, 혹은 내가 매우 비협조적인데 상대가 겨우 한 번 협조한 경우는 무조건 배신이다.
e. H(1, 1)은 정해지지 않았다. 즉, 상대가 배신했어도 나의 행동이 모호하다면, 협동의 가능성도 있다.

위의 일반화를 트리거 전략과 주고 받기 전략에 비교하면 이렇다. 진화적 전략은 트리거 전략처럼 반드시 배신을 유지하는 것은 아니다(H(0, 2)가 협동인 경우도 있다). 또한 주고 받기 전략처럼 상대가 협조한다고 자기도 바로 협조하는 것도 아니다. 진화적 전략은 비교적 배신에 민감하지만, 상대의 배신에 반드시 배신으로 되갚는 것은 아니며, 한 번 시작한 배신을 멈추는 것이 가능하지만 주고 받기 전략보다 그 조건이 엄격하다.

5. 결론

이제까지의 실험으로, history count가 증가함에 따라 개체의 진화가 다소 안정적이 되며, population 및 game count가 증가함에 따라 게임 크기와 협동으로의 수렴 경향의 반비례 관계가 좀 더 뚜렷해짐을 보았다. 그러나 population 및 game count가 충분한가를 평가하는 '정량적 기준'은 발견하지 못한 상태인데, 이것은 population 이나 game count 외 다른 많은 parameter들 역시 부족한 채로 실험했을 가능성이 있으며, 이런 경우 GA의 특성상 하나의 parameter가 다른 parameter에 끼치는 영향을 예측하기가 힘들기 때문이다.
진화적 전략이 협동으로 수렴하는 경우 우위 전략은 트리거 전략과 주고 받기 전략, 둘 사이에 위치하며, 상대의 행동과 나의 행동을 종합적으로 검토한다는 것을 발견했다. '사실상 모든' 우위 전략들은 협동으로 게임을 시작하며, '거의 모든' 우위 전략들은 상대의 배신을 배신으로 되갚는다. 그러나 트리거 전략과는 달리 배신에서 협동으로 옮겨오는 우위 전략도 존재하는데, 그 조건이 주고 받기 전략보다 훨씬 엄격하다. 그 외 한 가지 재미있는 점은, 개체 군이 우위 전략으로 수렴할 때 우위 전략들이 대체로 비슷하다는 점이다. 예를 들어 4.2절의 H(0, 2)나 H(1, 1)은 여러 번의 실험을 통해 '정해지지 않음'으로 분석되었지만, 한 번의 실험(하나의 개체 군)에서는 0이나 1로 거의 고정되는 것을 발견할 수 있다. 이것은 우위 전략들이 공통의 조상을 가질 확률이 높다는 점에서 당연한 것으로 생각되지만, 우위 전략을 정성적으로 분석할 때 함정에 빠질 가능성이 많다는 얘기이기도 하다.
GA가 각종 parameter에 매우 민감하므로, 가능한 각 parameter들의 변화와 그 조합을 모두 실험했다면 훨씬 믿을만한 결과를 얻을 수 있었을 것이나 시간상 그러지 못한 것이 아쉬웠다. 또한 협동이나 배신으로의 수렴을 좀 더 강화하는 여러 방법들 – 학습의 지역화, 상호 작용의 지역화, 선택(selection) 방법의 변화 등등 – 을 실험한다면, IPD 게임의 응용에도 많은 도움이 될 듯하다. 그리고 이 논문에서는 payoff function으로 C(x) = 2x, D(x) = 2x + 1만을 사용했는데, 다른 종류의 payoff를 사용한다면 어떤 결과가 나올 것인지도 궁금하다.
복잡한 조건 아래에서 IPD 게임의 해를 구하기 위해 진화적 접근은 일단 유용한 듯이 보인다. 그러나 조건의 변화에 해들이 매우 민감하므로, 어떤 조건이 어떤 변화를 일으키는 지에 대한 연구가 반드시 필요하다고 하겠다.

참고 문헌

X. Yao and P. Darwen, Genetic algorithms and evolutionary games, *Commerce, Complexity and Evolution*, W. Barnett, C. Chiarella, S. Keen, R. Marks, and H. Schnabl (eds.), Chapter 16, pp.313-333, Cambridge University Press, 2000.

서연규, *진화적 방식을 이용한 일반화된 전략 학습에 관한 연구*, 연세대학교 대학원 컴퓨터과학과 석사 학위 논문, 1999년 6월.

A. K. Dixit and B. Nalebuff, *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life*, W.W. Norton & Company, 1993.

P. Darwen and X. Yao, Co-evolution in iterated prisoner's dilemma with intermediate levels of cooperation: Application to missile defense, *International Journal of Computational Intelligence and Applications*, 2(1): 87-107, March 2002.

D. E. Goldberg, *Genetic algorithms in Search, Optimization and Machine Learning*, Addison-Wesely, 1989.