

## 인공에이전트를 이용한 교섭게임에 관한 연구

장석철<sup>1\*</sup> · 석상문<sup>2</sup> · 윤정일<sup>3</sup> · 윤정원<sup>4</sup> · 안병하<sup>5</sup>

<sup>1</sup>동부정보기술 컨설팅사업부 / <sup>2</sup>특허청 정보심사팀 / <sup>3</sup>한국전자통신연구원 전파방송연구단  
<sup>4</sup>경상대학교 기계항공공학부 & 항공기부품연구소 / <sup>5</sup>광주과학기술원 기전공학과 시스템통합 연구실

### Analysis on the Bargaining Game Using Artificial Agents

Seok-cheol Chang<sup>1</sup> · Sang-moon Soak<sup>2</sup> · Joung-il Yun<sup>3</sup> · Jung-won Yoon<sup>4</sup> · Byung-ha Ahn<sup>5</sup>

<sup>1</sup>Dongbu Information Technology Co., Ltd., Seoul 135-879

<sup>2</sup>Korean Intellectual Property Office, Gov. Complex Daejeon Bldg.4, Daejeon 302-701

<sup>3</sup>Electronics and Telecommunications Research Institute (ETRI), Radio & Broadcasting Research Division,  
Broadcasting System Research Group, Daejeon 305-700

<sup>4</sup>School of Mechanical and Aerospace Engineering and ReCAPT, Gyeongsang National University

<sup>5</sup>Mechatronics Department, GIST, Gwangju 500-712

Over the past few years, a considerable number of studies have been conducted on modeling the bargaining game using artificial agents on *within-model* interaction. However, very few attempts have been made at study on *between-model* interaction. This paper investigates the interaction and co-evolutionary process among heterogeneous artificial agents in the bargaining game.

We present two kinds of the artificial agents participating in the bargaining game. They play some bargaining games with their strategies based on genetic algorithm (GA) and reinforcement learning (RL). We compare agents' performance between two agents under various conditions which are the changes of the parameters of artificial agents and the maximal number of round in the bargaining game. Finally, we discuss which agents show better performance and why the results are produced.

**Keywords:** Bargaining Game, Artificial Agent

#### 1. 서론

Ståhl(1971)과 Rubinstein(1982)이 이론적 기반을 설립한 이후로 교섭게임(bargaining game)은 전자상거래(Omato *et al.*, 2002), 노사협상(Berninghaus *et al.*, 2002), 분쟁해결(Nakayama, 2000)등에서 양자간 협상 연구에 중요한 분석도구로 게임이론, 경제학, 심리학, 컴퓨터공학 등의 분야에서 많은 연구가 진행되어져 왔다.

교섭게임이란 두 명의 경기가가 일정량의 재화를 나누는 게임으로 수많은 내쉬균형(Nash Equilibriums)이 존재하지만 이론

상으로 마지막 단계의 제안자(proponent)가 상대방(respondent)에게 소량의 몫( $\epsilon > 0$ )을 제안하고, 응답자는 이를 수락하게 되는데 이를 부분게임완전균형(sub-game perfect equilibrium)이라 한다. 그러나 실제 실험결과는 이론적 예측과는 많은 차이를 보이고 있다. 즉, 제안자가 소량의  $\epsilon$ 을 제안하는 경우, 응답자는 불평등한 거래로 인식하고 제안을 거부하는 경향을 보인다. 그래서 대부분의 실험에서 양쪽 경기가가 재화를 거의 동등하게 나누어 갖게 된다. Page *et al.*의 조사에 의하면 제안자의 60~80%는 재화의 0.4~0.5 부분을 상대방에게 제안을 하였고, 오직 3%만이 0.2이하 부분의 재화를 상대방에게 제안하는 경향이 있었

이 논문은 2단계 BK21사업 및 2006년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임(KRF-2006-003-D00022).

\* 연락저자 : 장석철 박사, 135-879 서울 강남구 삼성동 154-17 동부정보기술 컨설팅사업부, Tel : 02-3449-2224, Fax : 02-3449-2392

E-mail : scchang@dongbu.com

2005년 3월 접수; 2006년 1월 수정본 접수; 2006년 3월 게재 확정.

다. 또한 응답자의 경우 50% 이상이 전체 몫의 1/3 이하의 제안은 거부하는 것으로 드러났다(Page et al., 2000; Roth and Erev, 1995; Kagel et al., 1996; Burnell et al., 1999). 이러한 이론적인 예측과 실험결과의 차이는 공정성(fairness)과 모든 경기자들의 완전합리성(common knowledge of rationality) 등에 기인하는 것으로 밝혀졌다(Stanley 1998; Thaler 1988; Suleiman 1996).

위와 같은 연구는 경제 및 심리학 분야에서 사람이 직접 실험에 참여하여 결과를 산출하였으나 최근 들어 컴퓨터공학 분야에서는 인공에이전트(artificial agent)를 이용하여 여러 가지 게임이론을 분석하려는 시도가 이루어지고 있다. 특히 교섭게임의 경우 진화알고리즘(evolutionary algorithm)이나 강화학습이론(reinforcement learning : RL)등을 이용한 연구가 활발히 진행되었다. Matwin et al.(1991)은 유전자 알고리즘(Genetic Algorithm : GA)을 이용한 교섭게임에서의 협상지원시스템(Negotiation Support System : NSS)을 개발하였고, Page et al. (2000)은 진화전략(evolution strategy : ES)을 교섭게임에 적용하여 적응시스템을 구축하였고, Bragt et al.(2002)은 다수의 상대 참가자들을 구별하고, 그에 따른 전략을 세우기 위하여 GA로 진화하는 유한 오토마타(finite automata)를 이용하였다. Fang Zhong et al.(2002)은 RL로 전략을 진화시키는 에이전트와 고정규칙을 가진 에이전트사이의 게임을 통해서 진화과정을 분석하였고, Cooper et al.(2003)은 RL로 진화하는 에이전트에서 제안자와 응답자사이의 상대적인 학습속도를 측정하고 그것이 결과에 미치는 효과를 예측하였다. Grosskopf(2003)는 RL과 지향성학습(directional learning : DL)으로 전략을 진화시키는 에이전트를 이용하여, 한 명의 제안자와 한 명의 응답자사이의 게임과 한 명의 제안자와 세 명의 응답자사이의 게임을 비교하고, 각 에이전트의 상호공진화 과정을 연구하였다. Takadama et al.(2003)은 ES 모델, RL 모델, 학습 분류 시스템(learning classifier system : LCS) 모델과 같이 3가지 교섭게임 모델을 제시하고 그것의 유용성을 검토하였다.

이와 같이 인공에이전트를 이용한 교섭게임은 최근 수년간 상당히 많은 연구가 진행되어져 왔으나 몇 가지 한계점들도 드러내고 있다. 과거의 연구방향은 경제 및 심리학분야에서 사람이 참여하여 실시했던 실험을 인공에이전트를 이용하여 재현하고 그 결과를 비교 분석하여 모델의 타당성을 증명하는 수준이었다. 그러나 이러한 실험들은 동일한 전략을 가진 인공에이전트간의 상호작용(within-model interactions)을 분석한 것으로, 다양한 성향과 전략을 가진 참여자들 사이에 상호작용(between-model interactions)이 발생하는 실제 사회에서의 교섭 게임을 분석하기에는 다소 무리가 따른다. 또한 사용되는 인공에이전트의 주요 모수(principal parameters)나 행동양식(patterns of action)등에 대한 사전연구는 교섭게임의 정확한 분석을 위한 필수 조건이라 할 수 있다.

본 논문에서는 유전자알고리즘 기반의 에이전트(GA-based agent)와 강화학습 기반의 에이전트(RL-based agent)를 도입하고 이질적인 두 진화 에이전트간의 교섭 게임을 실시하였다. 이를 통하여 실제 사회에서 발생하는 단체교섭게임(노사협상

등)에서 상호 성향과 전략이 다른 두 집단 간의 교섭과정 및 상호대응과정을 인공에이전트를 이용하여 분석하기 위한 하나의 방안을 제시하고자 하였다. 이와 같은 목적을 위하여 본 논문은 다음과 같은 사항을 조사하였다.

- 각 에이전트의 주요한 모수는 무엇이고 행동특성은 어떤가
- 두 에이전트간의 상호작용(interaction) 및 공진화(co-evolution) 과정은 어떻게 이루어지나?
- 단일 교섭게임(one-shot bargaining game)과 다단계교섭게임(sequential bargaining game)의 차이는 무엇인가?

## 2. 교섭게임 및 인공에이전트의 설계

본 장에서는 위에서 간단히 언급한 교섭게임이 어떻게 구현되었었으며, 이에 참여하는 인공에이전트들은 주요한 모수가 무엇이며 무슨 특징들을 가지고 있는지에 관하여 토론한다.

### 2.1 교섭게임 설계

앞에서 언급하였듯이 교섭게임은 두 명의 경기자가 일정량의 재화를 나누는 게임으로<Figure 1>은 전체적인 교섭게임 과정을 나타낸 것이다. 그림에서 실험모수 설정(setting parameters) 단계는 실험결과에 영향을 미칠 수 있는 GA-based agent와 RL-based agent의 모수 값들을 설정하는 단계이다. 이때 우리는 각 에이전트의 모수특성 들을 알아보기 위해 한 쪽 에이전트의 모수를 고정시켜놓고 다른 에이전트의 모수들을 변화시켜가며 이러한 변화량이 결과에 미치는 영향을 조사한다(Chang, 2004). 또한 단일 교섭게임( $mr = 1$ )과 다단계교섭게임( $mr = 2,3,4$ )에서의 각 에이전트의 성능차이를 비교분석하기

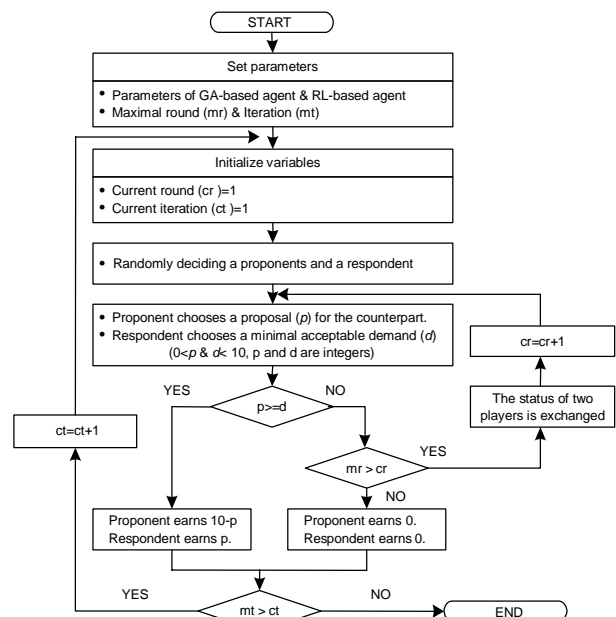


Figure 1. The procedure of the bargaining game.

위하여 최대단계(maximal round: mr)의 값을  $mr = 1, 2, 3, 4$ 로 변화시킨다. 이후 변수 값을 초기화하고 임의로(randomly) 제안자와 응답자를 결정하였다. 제안자의 경우 일정량의 재화(본 실험에서는 10으로 설정)중 상대방에게 양보할 수 있는 재화의 양  $p$ 를 제시하고 응답자는 자신이 받아야 할 최소의 재화의 양  $d$ 를 결정한다. 그리고 둘의 값을 비교하여  $p-d \geq 0$ 가 만족하면 거래가 성립되고, 그렇지 않으면 응답자와 제안자의 위치가 바뀌게 된다. 마지막 단계까지 거래가 성립되지 않으면 두 거래참여자는 아무런 재화도 받지 못하게 된다. 모든 실험은 최대 반복횟수(maximal iteration : mt)를 100,000번으로 하고 각 모수마다 10번씩 실험을 실시한 후 이를 평균 내어 그 결과를 산출하였다. 이러한 결과는 분산분석(ANOVA)을 통하여 각 모수의 유의성(significance)을 조사하였다.

### 2.2 GA-based agent 설계

유전자 알고리즘(Genetic Algorithm : GA)은 적자생존의 법칙(the law of the survival of the fittest)을 이용한 탐색 알고리즘으로 좀더 나은 해(solution)를 찾아가기 위하여 적합도 값(fitness value)에 따라 선택(selection), 교차(crossover), 돌연변이(mutation) 등의 연산자를 이용한다. 본 실험에서 적합도 값은 GA-based agent가 교섭게임을 통하여 얻은 수입(payload)으로 측정하였다. <Figure 2>는 다양한 전략들로 이루어진 GA-based agent의 모집단과 여기서 추출된 전략의 모형을 나타내고 있다. 각각의 전략은 2개의 행벡터(row vector)로 이루어졌는데 첫 단계에서 GA-based agent가 제안자로 시작할 경우 첫 번째 행의 전략을 사용하고, 응답자로 시작할 경우 2번째 행의 전략을 사용하게 된다. 예를 들어 <Figure 2>의 경우에 GA-based agent가 제안자로 첫 단계를 시작할 경우 각 단계별 전략은  $p_1 = 8, d_2 = 7, p_3 = 3, d_4 = 5$ 이 된다.

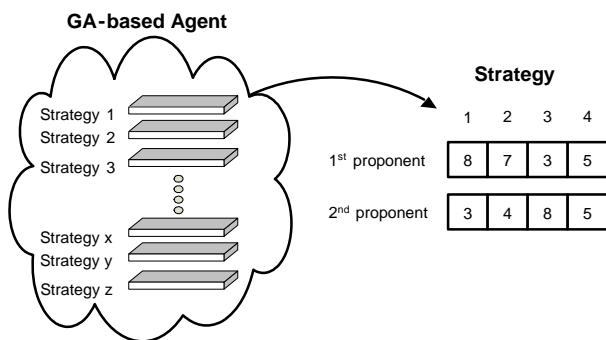


Figure 2. Schematic of GA-based agent.

일반적으로 GA-based agent는 교섭게임에서 모집단의 크기(population size :  $n$ ), 교차확률(probability of crossover :  $\alpha$ ), 돌연변이확률(probability of mutation :  $\beta$ )과 같이 3가지의 주요한 모수를 가지고 있다. 만약  $\alpha = 0.8$ 이면 모집단에서 다음 세대(offspring) 전략의 80%는 선택된 부모(parents)의 교차로 만

들어지고, 나머지 20%는 정예주의(elitism)로 부모 전략이 그대로 전승된다. 다음은 본 실험에서 조사한 모수들과 그 수준(level)을 나타내고 있다.

- Population size :  $A_1 (n = 10), A_2 (n = 30), A_3 (n = 50)$
- Crossover :  $B_1 (\alpha = 0.6), B_2 (\alpha = 0.7), B_3 (\alpha = 0.8), B_4 (\alpha = 0.9), B_5 (\alpha = 1.0)$
- Mutation :  $C_1 (\beta = 0), C_2 (\beta = 0.05), B_3 (\beta = 0.1)$

### 2.3 RL-based agent 설계

RL-based agent는 2개의 행렬, 즉  $p$ 값을 나타내는 제안행렬(offer matrix)과  $d$ 값을 나타내는 임계행렬(threshold matrix)로 구성된다. <Figure 3>은 RL-based agent 모델로 첫 단계(first round)에서 제안자로 시작할 경우 제안행렬을 사용하고, 응답자로 시작할 경우 임계행렬을 사용하여 전략을 세우게 되는데 그 방법은 다음과 같다.

Offer Matrix				
Round Value	1	2	3	4
1	4.20	7.50	0.61	0
2	33.00	2.64	8.27	0
3	2.89	0.24	4.27	0
4	8.83	6.49	0.65	0
5	1.83	6.27	0.92	0
6	0.60	0.25	0.49	0
7	0.35	0.06	0	4.62
8	0.08	0.18	0.03	0
9	0.14	0.02	0	0.09

Threshold Matrix				
Round Value	1	2	3	4
1	3.93	12.89	4.34	5.53
2	0.70	1.64	1.16	0.06
3	3.12	0.21	0	4.67
4	7.51	2.11	5.24	0
5	0.57	3.70	5.56	0
6	12.42	9.72	0.26	0
7	16.69	4.93	0.12	0
8	1.03	1.12	0.15	0
9	0.49	2.49	0.02	0

Figure 3. Schematic of RL-based agent : The values of the propensity are experimental values at  $\mu = 0.010, \kappa = 0.005$ .

RL-based agent가 반복횟수(iteration)  $t$ 일 때  $k$ 번째 단계(round)에서  $i$ 번째 순수전략(pure strategy)을 가지려는 성향(propensity)을  $Q_k^i(t)$ 라 하고 이것으로부터 받은 수입(payload)을  $W_k^i(t)$ , 학습률(learning rate)을  $\delta$ 라 할 때 에이전트의 성향은

$$Q_k^i(t+1) = Q_k^i(t) + \delta W_k^i(t)$$

와 같이 갱신한다.

이러한 성향값을 기반으로 우리는 RL-based agent의 전략을 선택하게 되는데, 반복횟수(iteration)  $t+1$ 일 때  $k$ 번째 단계(round)에서  $i$ 번째 순수전략(pure strategy)을 선택할 확률  $P_k^i(t)$ 은

$$P_k^i(t) = \frac{Q_k^i(t)}{\sum_{i=1}^n Q_k^i(t)}$$

와 같이 결정된다.

전략선택의 메커니즘은  $\epsilon$ -greedy 방법을 이용하였다. 처음 게임을 시작할 때는 높은 확률로 임의의 전략을 선택하다가 반

복횡수가 증가할수록 강화학습법에 의한 전략을 선택하게 된다. 이는 초반에 광범위한 범위의 해공간(solution space)을 탐색하다가 시간이 지나면서 점차 해가 수렴해 가는 특성을 지닌다

RL-based agent의 모수는 여러 가지가 있지만 본 실험에서는 절삭값(cutoff value :  $\mu$ )과 건망성 값(forgetting value :  $\kappa$ )으로 한정한다. 여기서 절삭값은 잡음신호(noise)와 같은 사건들이 결과에 영향을 미치는 것을 방지하기 위한 모수로 만약  $P_k^i(t) < \mu$ 이면  $P_k^i(t) = Q_k^i(t) = 0$ 으로 설정한다. 건망성 값은  $Q_k^i(t)$ 가 한없이 증가하는 것을 방지하기 위하여 교섭 게임을 반복할 때마다 성향값  $Q_k^i(t)$ 에  $1 - \kappa$ 를 곱하여 그 값을 감소시킨다. 이러한 모수들은 Roth and Erev(1995)가 단일교섭 게임에 처음으로 적용했던 것이다. 그러나 이 연구는 통계적인 분석이 없이 그러한 모수들이 결과에 미치는 영향을 그래프 상에 표현해주는 것으로 그쳤다. 우리는 분산분석을 이용하여 통계적으로 모수가 결과에 미치는 영향을 조사할 것이다. 또한 이러한 모수들을 단일교섭게임 뿐만 아니라 다단계 교섭게임에도 적용하여 그 유의성을 파악할 것이다 다음은 본 실험에서 조사한 모수들과 그 수준(level)을 나타내고 있다.

- Cutoff value :  $D_1(\mu = 0), D_2(\mu = 0.005), D_3(\mu = 0.010)$
- Forgetting value :  $E_1(\kappa = 0), E_2(\kappa = 0.001), E_3(\kappa = 0.005)$

### 3. 결과분석

#### 3.1 GA-based agent와 RL-based agent의 성능비교

본 절에서는 GA-based agent와 RL-based agent가 교섭 게임을 할 때 상호작용 및 공진화 과정을 토론했다. <Figure 4>는 모의 실험 결과 각 에이전트가 받은 수입(payoff)을 누적 평균한 것으로 모든 데이터는 최대단계(maximal round)를  $ms = 4$ 로 고정시키고, 동일조건에서 10번씩 실험하여 이를 평균한 값이다. 여기서 X축은 10,000회까지의 반복횡수(iteration)를, Y축은 1~9사이의 누적평균수입(cumulative average payoff)을 나타낸 것이다. 또한 붉은 점은 반복횡수가 100,000회일 때의 수입을 나타낸 것이다. 그림에서 윗선(upper curve)은 10번의 실험 중 최대치를, 중간선(middle curve)은 평균치를, 그리고 아랫선(lower curve)은 최소치를 나타낸다. 그리고 좌측 두열의 그림은 RL-based agent의 모수를  $\mu = 0, \kappa = 0$ 으로 고정시킨 후 GA-based agent의 모수를 변화시켰을 때의 결과 값이고, 우측 두열의 그림은 GA-based agent의 모수를  $n = 30, \alpha = 0.8, \beta = 0.05$ 로 고정시킨 후 RL-based agent의 모수를 변화시켰을 때의 결과 값이다.

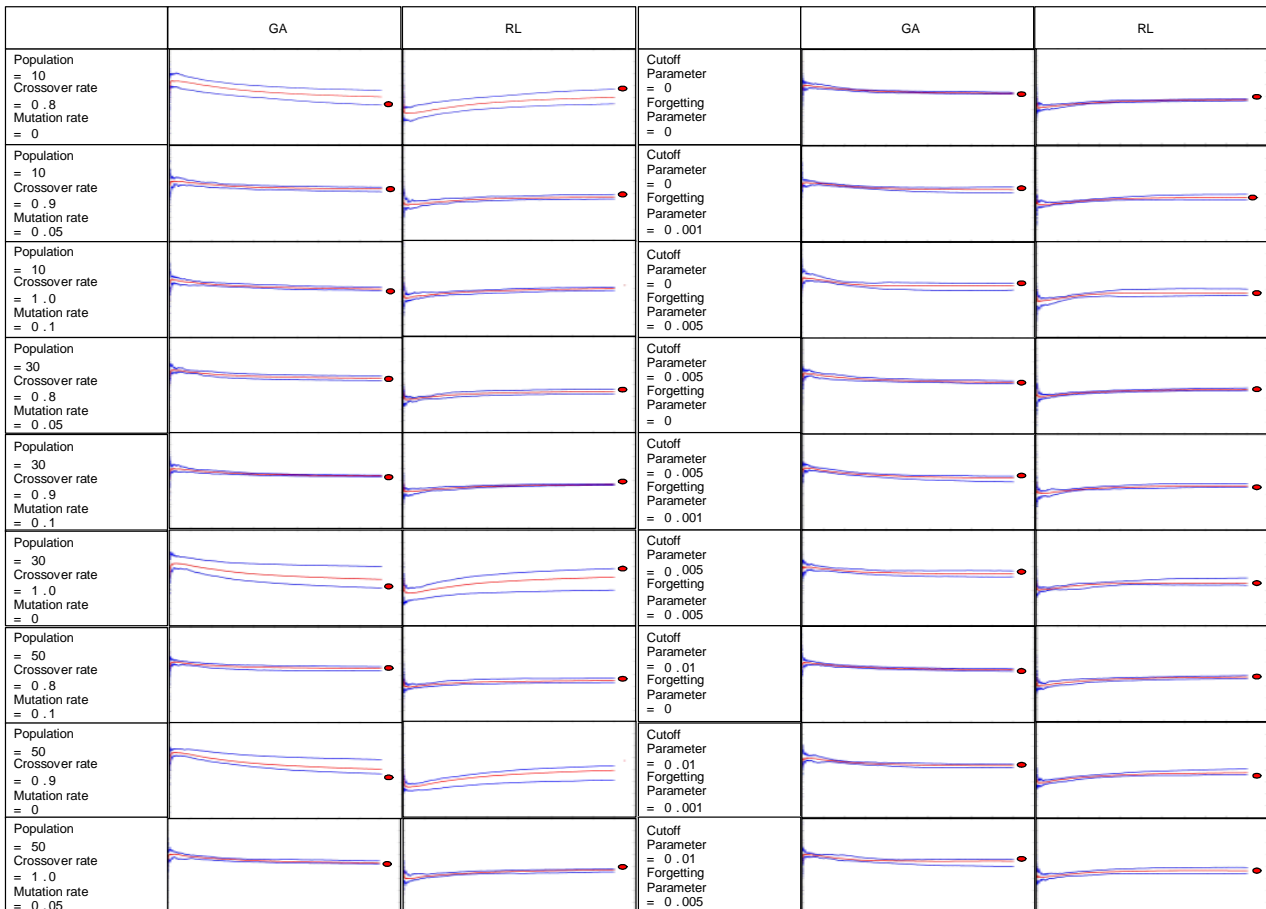


Figure 4. cumulative average payoffs of agents with various parameters.

위의 실험결과 에이전트의 성능측면에서는 게임의 초반부에 *GA-based agent*가 *RL-based agent*보다 전체적으로 우수한 성능을 보이다가 반복횟수가 1000회 이상이 되면 그 차이는 점점 줄어드는 것을 관찰할 수 있다. 이는 *GA-based agent*의 학습속도가 매우 빨라 좋은 해(good solution)에 빨리 수렴하는 경향을 보이며, *RL-based agent*는 수렴속도는 느리나 어느 정도 시간이 지나면 좋은 해에 수렴한다는 것을 알 수 있다.

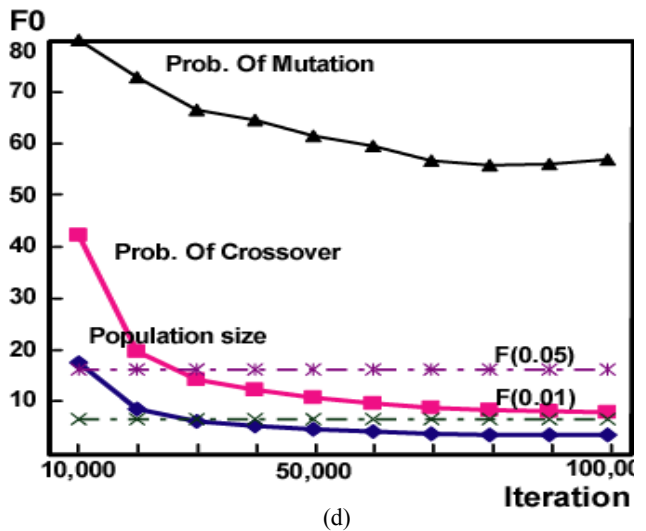
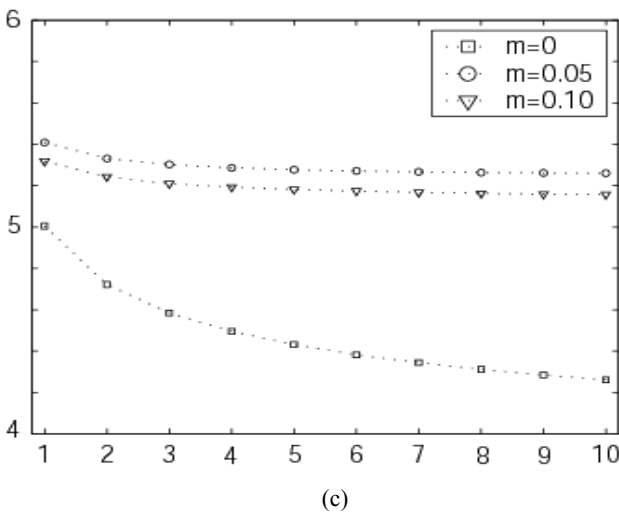
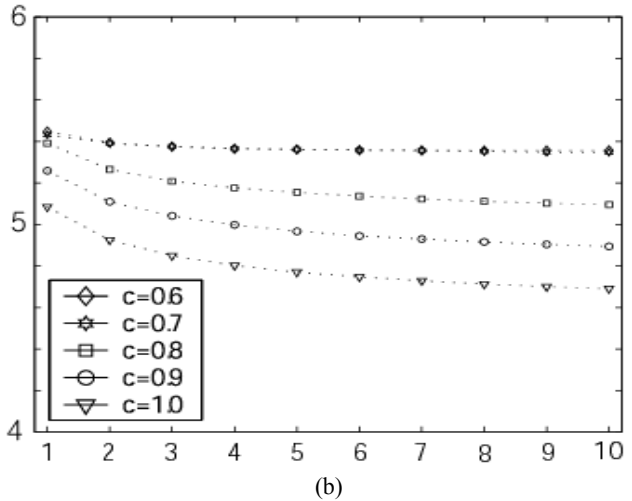
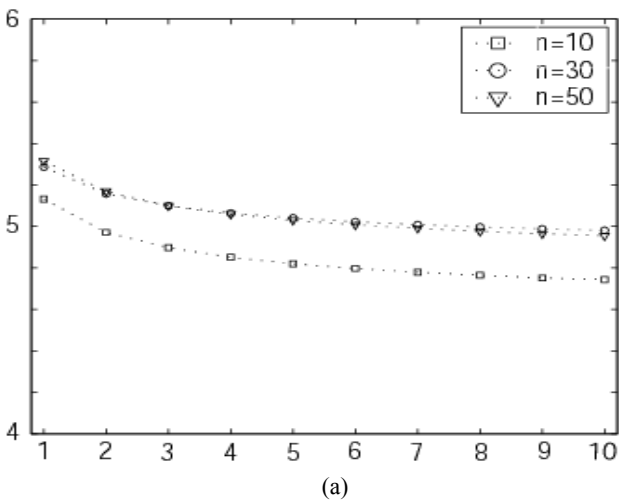
모수에 따른 행동특성으로 수입편차(deviation of average payoffs), 즉 수입의 최대치와 최소치의 차이를 관찰해보면 대부분의 경우에서 편차가 매우 적다. 그러나 *GA-based agent*는  $\beta = 0$ 일 때 큰 편차를 보이고 있는데, 이는 초기 모집단이 임의로 형성된 이후 해를 변화시키는 돌연변이 연산의 부재로 초기 모집단의 전략에 따라 성능이 많은 영향을 받기 때문이다. *RL-based agent*의 경우에는 큰 차이는 없지만  $\kappa = 0.005$ 인

경우에 편차가 약간 큰 것을 알 수 있다. 이는 건망성값이 클수록 순수 전략 간의 선택확률의 차이가 줄어들기 때문에 높은 성향값을 가진 소수의 전략에 집중하기보다는 다양한 전략을 택하기 때문인 것으로 예측된다.

### 3.2 GA-based agent의 모수분석

본 절에서는 *RL-based agent*의 모수를  $\mu = 0, \kappa = 0$ 으로 고정시키고 *GA-based agent*의 모수를 변화시킬 때, 실험결과에 어떠한 영향을 주는가를 통계적 기법을 통하여 분석하였다. <Figure 5>는 *GA-based agent*의 3개 모수값들의 변화에 따른 결과비교 및 모수들이 결과에 미치는 영향력을 나타내고 있다.

<Figure 5-(a)>는 모집단 크기에 따른 결과의 변화량을 나타낸 것으로  $n = 30$ 과  $n = 50$ 은 거의 비슷한 수입을 획득했지만  $n = 10$ 은 낮은 수입을 나타내고 있다. 이는 최대단계



Note : In (a), (b) and (c), x-axis means the number of iterations (unit : ten-thousand iterations) and y-axis the amount of payoff which the average payoff of ten-times experiments.

Figure 7. The results of *GA-based agent*. (a) Population size (b) Probability of crossover (c) Probability of Mutation (d) Analysis of variance.

( $m_s = 4$ )와 관계있는 것으로,  $n = 30$  이상은  $m_s = 4$ 에서의 해공간(solution space)을 충분히 표현할 수 있으나,  $n = 10$ 은 그렇지 못하다. 그러므로 만약 최대단계가 늘어나면 모집단의 크기도 늘어나야 할 것이다 <Figure 5-(b)>는 교차확률(the probability of crossover)의 변화에 따른 결과를 보여주는 것으로 이는 유전자 알고리즘의 정예주의와 관련이 있는데 이것의 도입은 해의 개선에 많은 도움이 되는 것을 알 수가 있다. 즉  $\alpha = 0.7$ 까지는 정예주의의 적용비율이 높을수록 수입이 증가하다가 이후는 해의 개선이 나타나지 않는다 <Figure 5-(c)>는 돌연변이확률(the probability of mutation)을 나타내는 것으로  $\beta = 0.05$ 에서 가장 좋은 성능을 나타낸다.  $\beta = 0$ 에서는 해의 탐색구간이 한정되어 있어서 가장 낮은 성능을 나타내고 있으며  $\beta = 0.10$ 에서는 최적해 근처에서 발산(divergence)하는 경향을 나타낸다. <Figure 5-(d)>는 각 모수들의  $p$ -value를 나타낸 것으로 대부분의 모수들이 결과에 미치는 영향력은 매우 유의한 것으로 나타났다. 즉  $F_0 > F(0.05)$ 와  $F(0.01)$ 을 나타내고 있다. 이러한 영향력은 반복횟수가 증가할수록 점차 감소하는 경향을 보이며, 특히 돌연변이 확률은 결과에 가장 큰 영향력을 미치는 것으로 드러났다.

### 3.3 RL-based agent의 모수분석

여기서는 GA-based agent의 모수를  $n = 30, \alpha = 0.8, \beta = 0.05$ 으로 고정시키고 RL-based agent의 모수를 조사하였다. <Figure 6-(a)>은 절삭값(cutoff value)의 변화가 결과에 아무런 영향을 주지 못하는 것으로 나타났다. 우리는 좀더 정확한 해석을 위하여 기존의 실험( $\mu = 0.005$  and  $\mu = 0.010$ ) 이외에 값을 변화시켜  $\mu = 0.05$  and  $\mu = 0.10$ 도 실험해 보았지만 역시 결과에는 아무런 변화를 주지 못했다. 이것은 에이전트가  $P_k^i(t)$ 의 값에 비례하여 그들의 전략을 선택하기 때문이다. 즉  $P_k^i(t)$ 은 초기에는 모두가 낮은 값에서 시작을 하지만 좋은 해의  $P_k^i(t)$  값은 매우 빠르게 증가한다. 그러나  $P_k^i(t)$ 이 절삭값 근처에 있는 전략들은 선택될 확률이 거의 0에 가까워서 결과에 별다른 영향을 주지 못하는 것으로 파악된다. <Figure 6-(b)>은 건망성값이  $\kappa = 0$ 일때 최대의 성능을 발휘하고  $\kappa$ 의 값이 증가할수록 낮은 성능을 보인다. 이는  $\kappa$ 값이 낮을수록 광범위한 해구간을 탐색하기 보다는 좋은 해로 빨리 수렴하기 때문이다. 만약, 최대단계가 커지면 해구간도 기하급수적으로 늘어나므로 빠른 수렴은 오히려 안 좋은 결과를 산출할 수 있다.

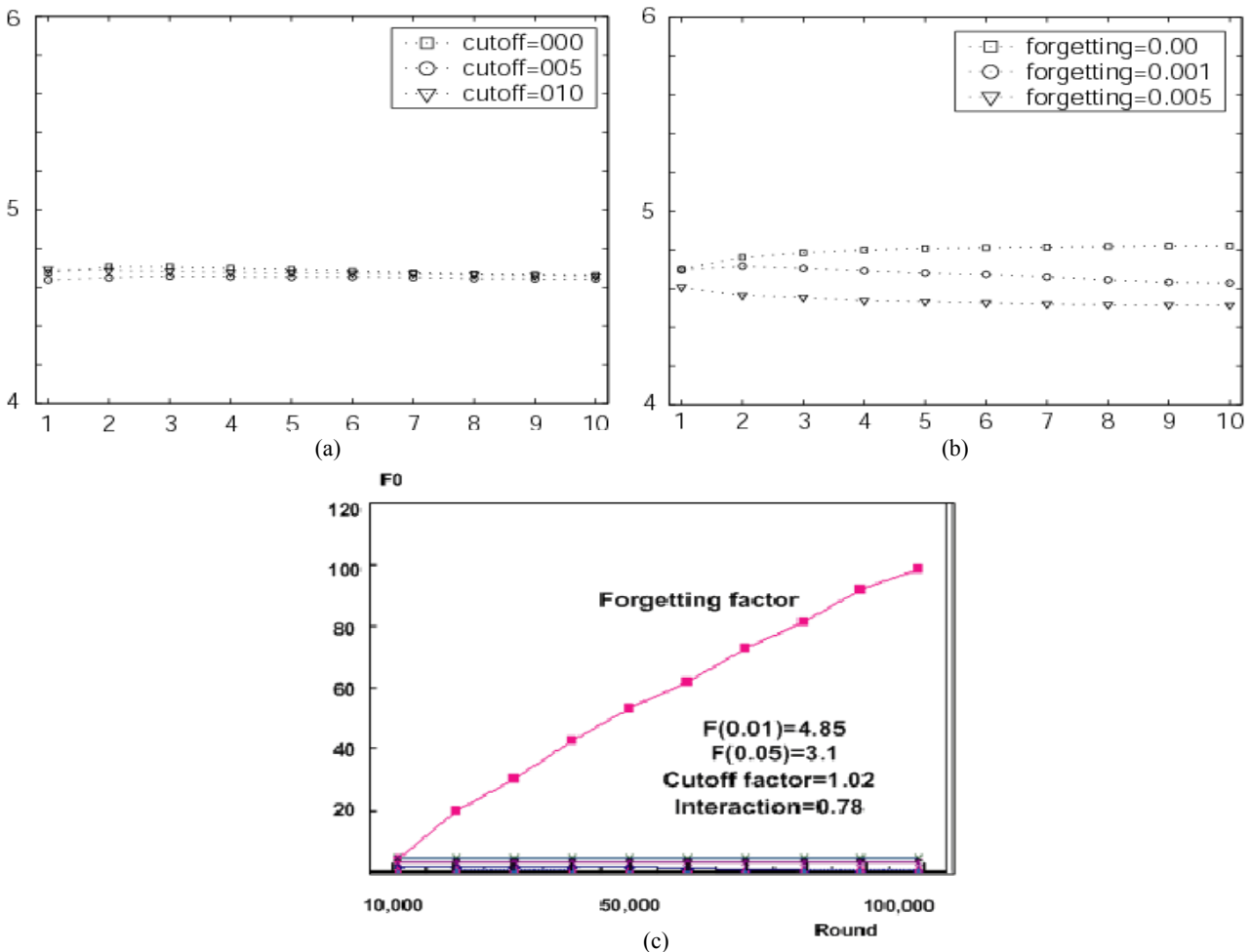


Figure 12. The results of RL-based agent. (a) Cutoff parameter (b) Forgetting parameter (c) The analysis of variance.

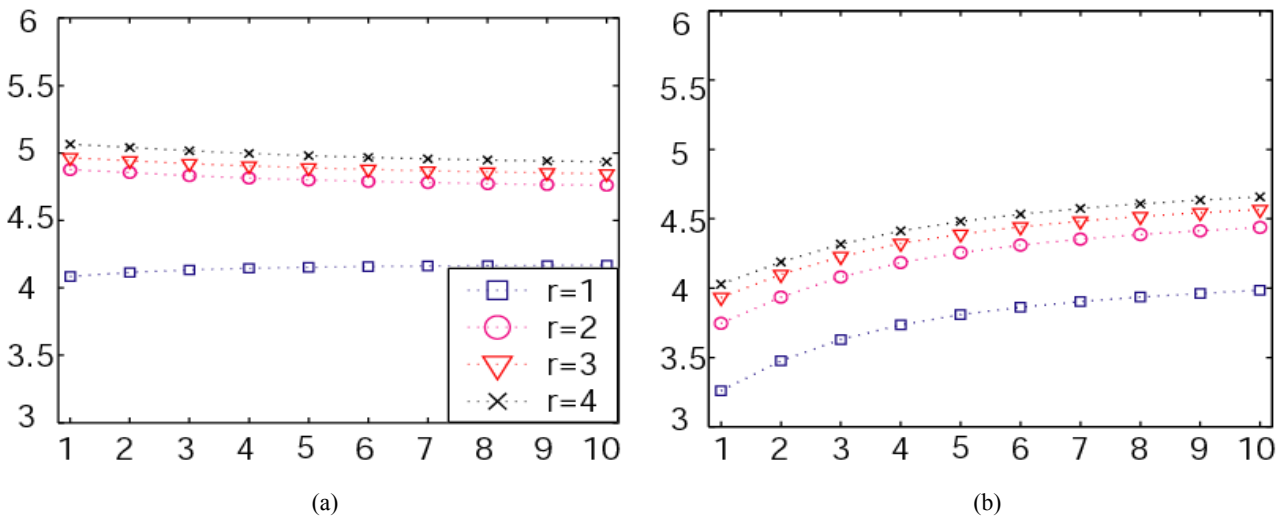


Figure 7. Payoff in the variation of maximal round : (a) GA-based agent (b) RL-based agent.

### 3.4 단일교섭게임과 다단계 교섭게임의 차이분석

<Figure 7>은 최대라운드가  $mr = 1$ 에서  $mr = 4$ 까지 변할 때 수입의 변화를 나타내고 있는 것으로 이를 통하여 우리는 단일교섭게임(one-shot bargaining game :  $mr = 1$ )과 다단계교섭게임(sequential bargaining game :  $mr = 2, 3, 4$ )의 차이를 확인한다. 여기서  $mr = 1$  and 2 사이의 수입의 차이는 다른 경우( $mr = 2$  and 3 또는  $mr = 3$  and 4)보다 훨씬 크다는 것을 알 수 있다. 이는 전략을 스스로 진화하는 에이전트는 마지막 단계에 이르면 거래를 성사시켜서 소량의 재화라도 얻으려고 노력한다. 그러므로 단일교섭게임은 항상 낮은 수입만을 얻게 된다. 그러나 다단계교섭게임은 마지막 단계에 도달하기 전까지는 많은 재화를 상대방으로부터 취하려고 노력하므로 그 차이가 발생하는 것으로 예측된다.

## 4. 결론

최근 들어 게임이론을 분석하고 시장경제에 이를 적용하기 위하여 인공지능을 이용한 분석기법이 많이 연구되고 있다. 그러나 인공지능을 게임에 적용함에 있어서 인공지능의 성능에 영향을 줄 수 있는 모수와 그 행동양식에 대한 연구나 또는 단일교섭게임과 다단계 교섭게임의 성능상의 차이를 조사한 연구는 거의 찾아볼 수 없다. 이에 우리는 교섭게임 분석에 있어서 널리 사용되는 진화에이전트 중에서 GA-based agent와 RL-based agent에 대해서 조사하였다. 그 주요한 결과는 다음과 같다.

- 게임초반에는 GA-based agent의 성능이 RL-based agent 보다 우수하지만 반복횟수가 증가할수록 그 차이는 감소한다.
- GA-based agent의 모수에서 모집단의 크기 교차연산/ 돌연변이연산 모두가 성능에 유의한 영향력을 발휘하였고

특히 돌연변이 연산이 가장 큰 영향력을 발휘하였다.

- GA-based agent는  $n = 30, \alpha = 0.7, \beta = 0.05$ 에서 최적의 성능을 나타낸다.
- GA-based agent의 반복횟수가 증가할수록 결과에 미치는 모수의 영향력은 감소한다.
- RL-based agent의 모수 중 절삭값은 결과에 아무런 영향력을 발휘하지 못하지만 건망성값은 반복횟수가 늘어날수록 큰 영향력을 발휘하고,  $\kappa = 0$ 일 때 가장 좋은 성능을 보인다.
- 단일교섭게임과 다단계교섭게임은 에이전트의 수입측면에서 차이를 보이고 있다.

이제까지 우리는 인공지능을 이용한 교섭게임에서의 모수분석 및 에이전트간 성능비교를 실시하였다. 앞으로의 연구방향은 본 논문에서 실시하였던 두 에이전트사이의 상호반응 연구를 확장하여 실제 사회에서와 같이 다양한 성향과 속성을 가진 여러 에이전트간의 상호반응을 연구할 계획이다. 본 논문은 GA-based agent 및 RL-based agent의 모수특성을 제공함으로써 해당 인공지능을 이용한 게임이론 연구시 활용 가능할 것이다. 또한 성향과 전략이 다른 이질적 인공지능 에이전트간의 게임 모델링에 참고자료가 될 수 있을 것이다.

## 참고문헌

Ståhl, I. (1971), *Bargaining Theory*, Stockholm, Stockholm School of Economics.  
 Rubinstein, A. (1982), Perfect equilibria in a bargaining model, *Econometrica*, **50**, 97-109.  
 Omoto, T., Kobayashi, K., and Onishi, M. (2002), Bargaining model of construction dispute resolution, *IEEE International Conference on Systems, Man and Cybernetics*, **7**, 7-12.  
 Berninghaus, S., Güth, W., Lechler, R., and Ramser, H.-J. (2002),

- Decentralized versus collective bargaining - An experimental study, *International journal of game theory*, **7**(3), 437-448.
- Nakayama, M. (2000), E-commerce and firm bargaining power shift in grocery marketing channels : A case of wholesalers' structured document exchanges, *Journal of information technology(JIT)*, **15**(3), 195-210.
- Page, K. M., Nowak, M. A., and Sigmund, K. (2000), The spatial ultimatum game, *Proceedings, Biological sciences*, **267**(1458), 2177-2182.
- Roth, A. and Erev, I. (1995), Learning in Extensive Form Games : Experimental Data and Simple Dynamic Models in the Intermediate Term, *Games and Economics Behavior*, **8**, 164-212.
- Kagel, J. H., Kim, C., and Moser, D. (1996), Fairness in Ultimatum Games with Asymmetric Information and Asymmetric Payoffs, *Games and Economic Behavior*, **13**, 100-110.
- Burnell, S. J., Evans, L., and Yao, S. (1999), The Ultimatum Game : Optimal Strategies without Fairness, *Games and economic behavior*, **26**(2), 221-252.
- Stanley, T. D. and Tran, U. (1998), Economics Students Need Not Be Greedy : Fairness and the Ultimatum Game, *The Journal of socio-economics*, **27**(6), 657-664.
- Thaler, R. H. (1988), Anomalies : The ultimatum game, *Journal of Economic Perspectives*, **2**, 195-206.
- Suleiman, R. (1996), Expectations and fairness in a modified Ultimatum game, *Journal of economic psychology*, **17**(5), 531-554.
- Matwin, S., Szapiro, T., and Haigh, K. (1991), Genetic algorithms approach to a negotiation support system, *IEEE Trans. Systems, Man and Cybernetics*, **21**, Issue 1, 102-114.
- Van Bragt, D. D. B. and La Poutrè, J. A. (2002), Co-evolving automata negotiate with a variety of opponents, *Proceedings of the 2002 Congress on Evolutionary Computation*, **2**, 1426-1431.
- Zhong, Fang, Kimbrough, Steven O., and Wu, D. J. (2002), Cooperative agent systems: artificial agents play the ultimatum game, *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, 2169-2177.
- David J. Cooper, Nick Feltovich, Alvin E. Roth, and Rami Zwick (2003), Relative versus Absolute Speed of Adjustment in Strategic Environments ; Responder Behavior in Ultimatum Games, *Experimental economics, a journal of the Economic Science Association*, **6**(2), 181-207.
- Grosskopf, Brit (2003), Reinforcement and Directional Learning in the Ultimatum Game with Responder Competition, *Experimental economics, a journal of the Economic Science Association*, **6**(2), 141-158.
- Takadama, K., Suematsu, Y. L., Sugimoto, N., Nawa, N. E., and Shimohara, K. (2003), Towards Verification and Validation in Multiagent-Based Systems and Simulations : Analyzing Different Learning Bargaining Agents, *The 4th Workshop on Multi-Agent Based Simulation*, 18-32.
- Chang, Soek-Cheol, Yun, Joung-II, Lee, Ju-Sang, Lee, Sang-Uk, Mahalik, N. P., and Ahn, Byung-Ha (2005), Analysis on the Parameters of the Evolving Artificial Agents in Sequential Bargaining Game, *The special issue on Software Agent and its Applications, IEICE*, **E88-D**(9), Sept.