

반복적 죄수의 딜레마 게임에서 전략적 연합의 의사결정 방법에 따른 협동의 진화

양승룡⁰ 노현길 조성배
연세대학교 컴퓨터과학과
(saddo⁰, knowhunger, sbcho)⁰@candy.yonsei.ac.kr

The Evolution of Cooperation according to Decision Making Methods of Strategic Coalition in Iterated Prisoner's Dilemma Game

Sung-Lyong Yang⁰, Hyun-Gul Noh, Sung-Bae Cho
Dept. of Computer Science, Yonsei University

요 약

반복적 죄수의 딜레마 게임은 복잡한 사회현상들을 모델링 하기 위하여 주로 사용되는 방법이다. 본 논문에서는 집단 을 협동으로 진화시키는데 있어서 전략적 연합을 이용하는 방법과 전략적 연합에 속해 있는 전략들이 연합의 의사결정 방법 때 어떠한 결함방법을 선택할 경우 더 빠르고 안정적으로 진화하는가에 대한 실험 결과를 제시한다. 의사결정 방법으로는 신경망의 결합에 주로 사용되는 투표방법, Borda 함수, Condorcet 함수, 평균방법 그리고 최고이득 함수 방법 등을 사용하였다. 실험결과 최고이득 함수 방법이 가장 좋은 결과를 보였으며 결합 방법에 따라 다소 상이한 결과가 나타났다.

1. 서 론

오늘날의 사회는 개인과 개인, 개인과 그룹 그리고 그룹과 그룹간의 상호작용을 바탕으로 관계를 형성한다. 그러나 이들의 행동양상을 예측하고 이에 맞춰 적절한 의사결정을 하기란 쉬운 일이 아니다. 반복적 죄수의 딜레마 게임은 이러한 복잡한 사회 현상들을 모델링하고 보다 나은 선택을 함으로써 협동으로 진화하기 위한 연구방법이다. 기존의 반복적 죄수의 딜레마 게임은 주로 전략들의 표현방법에 따른 협동으로의 진화 정도를 연구하거나 또는 죄수들의 수적 증가에 따른 협동으로의 진화 정도를 연구하였다. 본 논문에서는 전략들이 진화를 하는데 있어서 전략적 연합을 결성하고 이 연합에서 어떠한 의사결정 방법을 따를 때 가장 협동으로 잘 진화하는지에 대한 실험결과를 제시한다.

2. 협동으로의 진화를 위한 진화적 접근 방법

진화연산(Evolutionary Computation)은 자연계에서 생물체의 진화 방식을 컴퓨터 시스템에 적용한 것으로 복잡한 사회 현상들을 모델링 하고 문제를 해결하는데 많이 이용되는 방법이다[4]. 진화적인 방법은 동적인 문제해결이나 보다 정확한 해를 구하고자 할 경우 또는 낯선 환경에 적용될 경우 문제해결을 용이하게 하는 장점이 있다.

2.1 반복적 죄수의 딜레마 게임(IPD)

전통적인 반복적 죄수의 딜레마 게임은 두 명의 죄수들이 각각 협동(C) 또는 배반(D)을 선택하여 이득을 돌려받는 게임이다[1]. 이득표에 따르면 이 게임에서 죄수들이 가장 선호하는 선택은 직관적으로 배반임을 알 수 있다(표 1).

표 1. 반복적 죄수의 딜레마 게임 이득표.
단, $T > R > P > S$, $2R > (T + S)$

Choice	Cooperate	Defect
Cooperate	R / R	T / S
Defect	S / T	P / P

그러나 게임이 무한히 반복될 경우, 죄수들은 서로에게 이익이 되는 협동을 선택하는 것이 더 유리하다는 사실을 알게 된다. 따라서 전략들은 어느 정도 협동으로 진화하게 된다.

2.2 전략의 진화를 위한 유전자 알고리즘

유전자 알고리즘(Genetic Algorithm)은 자연 생태계의 유전자 선택과 결합을 기초로 하는 탐색 알고리즘이다[3]. 또한 유전자 알고리즘은 방대한 탐색공간에서 해를 찾을 수 있는 강력한 기법이다[2]. 그리고 이 알고리즘은 전략을 표현하기 위하여 2진 bit 표현을 사용한다. 즉, 1,024개의 점으로 구성된 공간을 표현하기 위해서는 10개의 bit가 필요하다($2^{10}=1,024$).

유전자 알고리즘은 새로운 전략의 생성을 위하여 다음과 같은 단계를 거치게 된다[2].

평가(Evaluation): 이득함수에 따라 전략의 적합도를 평가하는 단계

선택(Selection): 다음세대로 진화하게 될 우수한 전략을 선택하는 단계

재생산(Reproduction): 교차와 돌연변이에 의하여 새로운 전략이 생성되는 단계

3. 전략적 연합의 의사결정을 위한 결합방법

집단에 속한 전략은 반복적 죄수의 딜레마 게임을 진행함에 따라 전략적 연합을 결성하게 된다. 물론 전략적 연합은 연합을 결성했을 때 서로에게 이익이 될 때만 결성되게 된다. 이것은 일반적인 사회와 동일한 환경설정에서 비롯된다. 전략들은 게임을 함으로써 연합을 형성하게 된다. 본 연구의 실험에서 연합형성 조건은 다음과 같다.

- $s_{mn} > 0.8$ (s_{mn} : 게임하는 전략들 간의 유사도)
- $P_c > 1.0 \frac{(T+S)}{2}$ (P_c : 연합했을 때의 이득)
- $P_{mn} < 1.0 \frac{(T+S)}{2}$ (P_{mn} : 전략 상호 간의 이득)

전략적 연합의 결과를 결합하기 위하여 추상레벨에서는 투표방법, 순위레벨에서는 Borda 함수, Condorcet 함수를 사용하였다. 그리고 즉정치 레벨에서는 평균방법을 사용하였으며, 기타 방법으로 최고이득(Highest Payoff) 함수방법을 사용하였다.

다중 신경망이 M개의 클래스 ($A = \{1, \dots, M\}$)로 이루어진 문제에서 입력된 미지의 패턴 x 가 어느 클래스에 속하는 지를 결정하는 것으로 볼 때, 신경망 결합은 K개의 신경망이 출력하는 결과 e_k 를 이용하여 결합된 결과 $F(e_k(x))$ 를 결정하는 것이다.

3.1 추상레벨(투표방법)

투표방법은 전략적 연합에 속해 있는 에이전트들이 다수결로 많이 선택한 결과를 연합의 결정으로 따르는 방법이다. 이를 수식을 표현하면 다음과 같다.

$$F(e_k(x)) = j \text{ if } S(x \in C_j) = \max(S(x \in C_i))$$

$$S(x \in C_i) = \sum_{k=1}^K G_k(x \in C_i), \quad i \in A$$

$$G_k(x \in C_i) = \begin{cases} 1 & \text{if } e_k(x) = i \text{ and } i \in A \\ 0 & \text{otherwise} \end{cases}$$

3.2 순위레벨

3.2.1 Borda 함수

클래스 i 에 대하여 신경망 k 가 출력한 순위 r_{ki} 에 따라 r_{ki} 를 구하고 이 값들을 모든 신경망에 대해 합산한 것을 Borda 점수로 정하여, 이 점수가 가장 큰 클래스를 결과로 결정하는 방식이다. Borda 함수는 다음과 같다.

$$F(e_k(x)) = \max_{i \in A} A(B_i(e_k(x)))$$

$$B_i(e_k(x)) = \sum_{k=1}^K (M - r'_k(x))$$

3.2.2 Condorect 함수

각 클래스에 대하여 자신을 제외한 다른 모든 클래스와 순위를 비교하여 높은 순위의 신경망의 개수를 합하여 최소 값을 구해서 그 값을 Condorect 값으로 정하고, 이 값을 비교하여 높은 값을 가지는 클래스가 상위의 순위를 가지는 방식이다. Condorect 함수는 다음과 같다.

$$F(e_k(x)) = \max_{i \in A} (Con_i(e_k(x)))$$

$$Con_i(e_k(x)) = \min_{j \in M-i} \#(K : r'_k > r'_k)$$

3.3 측정치 레벨

3.3.1 평균방법

비교적 간단한 방법으로 각 클래스 i 마다 주어진 신경망의 출력값 m_i 를 모두 합하여 평균을 구해서 값이 가장 큰 클래스를 다중 신경망의 결과로 사용하는 방법이다. 수식으로 표현하면 다음과 같다.

$$F(e_k(x)) = \max_{j \in A} S_j \quad \text{단, } S_i = \left(\sum_{k=1}^K m'_k \right) / K$$

3.3.2 최고이득 방법

이 방법은 전략적 연합에 속해 있는 전략들 중 가장 이득이 높은 전략이 선택한 결정을 따르는 방법이다. 따라서 게임의 결과가 가장 좋은 전략의 결정을 따르므로 한편으로는 독단적인 결정이 될 수도 있다.

$$F(e_k(x)) = \max_{i \in A} (P_i(e_k(x)))$$

4. 실험결과

4.1 실험환경

본 논문의 실험에서는 전략들이 연합을 형성하고 연합 내에서의 각 전략들이 선택한 결정을 결합하는 방법을 비교함으로써 협동으로의 진화에 어떠한 영향을 미치는가를 실험하였다. 본 논문에서 사용된 전략의 표현방법은 X.Yao의 방식을 따랐으며 전략들이 집단 내에 있는 다른 전략들에 의해 동적으로 평가받는 공진화(Co-evolution) 방법을 사용하였다. 세부적인 실험환경은 다음과 같다(표2).

표2. 실험을 위한 환경설정

Factors	Value
Number of Player	2
Population Size	100
Coalition Size	Population / 3
Crossover Rate	0.6
Mutation Rate	0.01
Initial Population	Random
Iterations	400
Rounds in each game	100

본 논문의 실험은 전략의 집단이 협동으로 진화하는데 있어 전략적 연합을 이용하고자 하는 것이다. 그림1은 전략적 연합을 사용하지 않는 일반적인 반복적 죄수의 딜레마 게임을 실험한 결과이다.

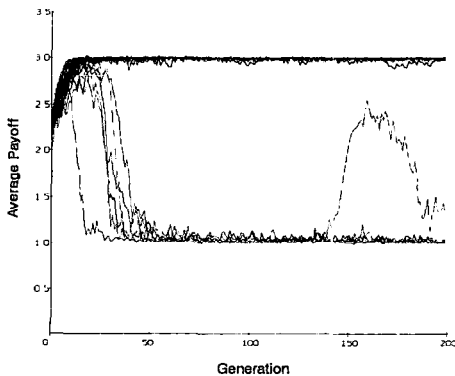


그림1. 전통적 죄수의 딜레마 게임 실험 결과

이 실험에서 30번의 실험을 수행한 결과 6번은 협동으로 진화하지 못하였으며 그 이외에는 주로 10~20세대에서 협동으로 진화하였다.

4.2 추상 레벨-가중 투표 방법

연합에 속한 에이전트들을 이득 순으로 나열하고 순위에 따라 가중치를 부여하여 가중치가 높은 결과를 따르는 방법으로 30번의 실험 중 5번이 협동으로의 진화에 실패하였다. 따라서 전략적 연합을 사용하지 않는 결과와 유사하였으며 전체적인 게임의 진행상태도 유사하였다(그림2).

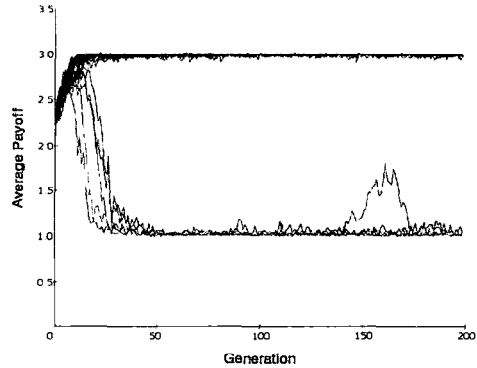


그림2. 가중 투표 방법에 의한 실험 결과

4.3 순위 레벨

4.3.1 Borda 함수 방법

연합에 속한 에이전트들을 이득 순으로 나열하고 순위를 정하여 그 순위의 역수를 Borda 값으로 정한 후, 그 값이 큰 결과를 따르는 방법이며 실험결과 20%(6번)가 협동으로의 진화에 실패하였다. 또한 150세대 부위에 상호배반에서 상호협동으로 진화하는 전략이 있었다(그림3).

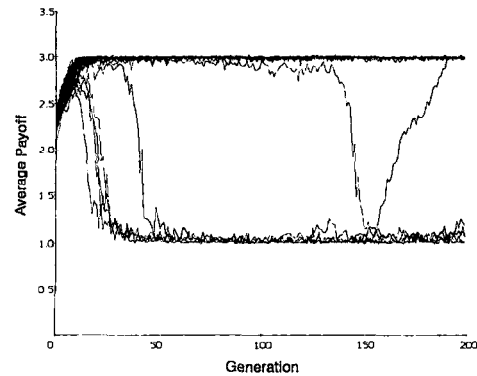


그림3. Borda 함수 방법에 의한 실험 결과

4.3.2 Condorect 함수 방법

이 방법은 자신을 제외한 나머지 연합에 속한 에이전트들을 각각 비교하여 자신보다 높은 이득을 가진 에이전트의 수를 Condorect 값으로 가지고, 그 값이 적은 결과를 따르는 방법으로 실험결과 약 13%인 4번이 협동으로 진화하지 못하였다(그림4). 따라서 전략적 연합을 사용하지 않는 게임보다 협동으로 진화하는 비율이 높았다.

일반적으로 집단은 초기에 상호배반으로 진화하게 되면 다시 상호협동으로 진화하기는 어렵다. 이는 전체적인 집단내의 분위기가 배반으로 결정되기 때문인데 이럴 경우 극심한 돌연변이에 의한 유전자의 변형이 발생하기 전에는 협동으로 진화하기 어렵다.

4.4 측정치 레벨(평균 방법)

이 방법은 비교적 간단한 방법으로 각 에이전트들이 얻은 이득을 평균하여 높은 평균을 가지는 결과를 연합의 결정으로 채택한다. 이 실험에서 실험결과 Condorect 함수를 이용한 결합방법과 유사한 결과가 도출되었으며, 초기에 4번의 실험이 상호배반으로 진화하여 안정되는 결과를 보여주고 있다(그림5).

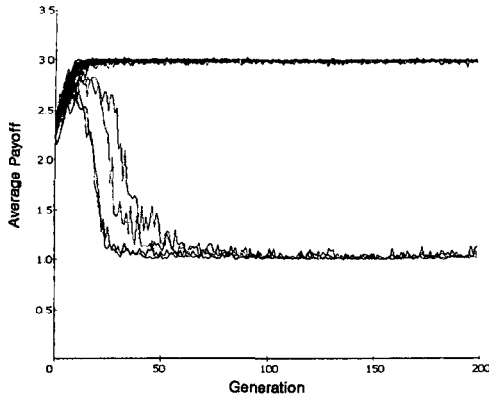


그림 4. Condorect 함수 방법에 의한 실험 결과

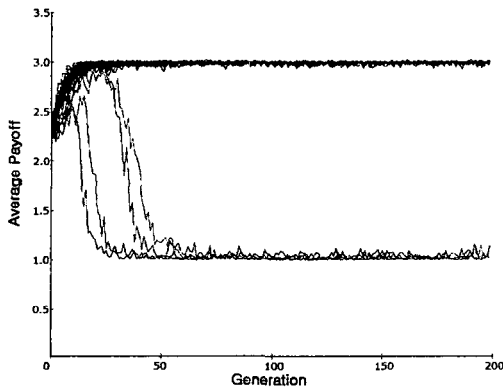


그림 5. 평균 방법에 의한 실험 결과

4.5 최고 이득(Highest Payoff)

가장 간단한 방법으로 연합에 속한 에이전트들 중 가장 이득이 높은 에이전트의 결과를 따르는 방법이다. 그럼에도 전략들은 빠르게 협동으로 진화하였으며 30번의 실험결과 중 2번만이 협동으로 진화하지 못하였다(그림 6).

이 결합방법은 상당히 독단적인 의사결정 방법일 수도 있다. 그러나 가장 이득이 높은 전략이 일반적으로 가장 게임을 잘하는 전략이라고 볼 때 직관적으로는 단순하면서도 우수한 결합방법임을 알 수 있다.

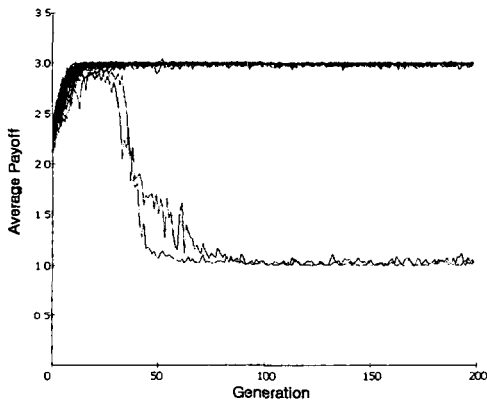


그림 6. Highest Payoff에 의한 실험 결과

본 논문에서는 연합에 속한 전략들의 의사결정 방법으로 일반적으로 널리 알려진 신경망의 결합방법을 이용하였다. 표 3은 각 결합방법 별로 100회를 실험하여 도출된 결과를 바탕으로 협동으로의 진화 정도를 살펴본 것이다.

표 3. 결합방법에 따른 협동으로의 진화율

결합 방법	협동으로의 진화율(%)
연합 없음	80
가중 투표 방법	89
Borda 함수 방법	84
Condorect 함수 방법	87
평균 방법	86
최고이득 방법	94

5. 결론 및 향후 연구

실험결과에 따르면 전략의 진화를 위하여 연합을 이용했을 경우, 그렇지 않을 경우에 비해 협동으로 진화하는 비율이 높음을 알 수 있으며 10 세대를 전후하여 협동으로 진화함을 알 수 있다. 또한 전략적 연합에서의 결합방법에 따라 다소 상이한 결과를 보이나, 최고이득 함수방법을 이용한 결합방법에서 가장 좋은 결과를 보이고 있다.

전략적 연합은 반복적 죄수의 딜레마 게임에서 새롭게 시도되는 협동으로의 진화를 위한 한 방법이다. 이러한 전략적 연합은 일반적인 사회에서도 두드러지게 나타나는 현상이며, 이를 이용할 경우 어느 정도 협동으로의 진화를 개선할 수 있을 것이다.

결론적으로 전략의 집단을 협동으로 진화시키기 위하여 전략들 간의 연합은 충분한 고려요소가 될 수 있으며, 이를 보다 구체적으로 증명하기 위해서는 보다 다양한 결합방법과 이득함수의 결정 및 죄수의 수적 증가에 대한 실험이 추가적으로 이루어져야 할 것이다.

참고 문헌

- [1] R. Axelrod, *The Evolution of Cooperation*. New York: Basic Books, 1984.
- [2] P. J. Darwen and X. Yao, On evolving robust strategies for iterated prisoner's dilemma in *Progress in Evolutionary Computation, Lecture Notes in Artificial Intelligence*, Vol. 956(X. Yao, ed.), (Heidelberg, Germany), pp. 276-292, Springer-Verlag, 1995.
- [3] K. A. Dejong, Genetic algorithm based learning. In Y. Kodratoff and R. Michalski, editors, *Machine Learning*, pp. 611-638. Morgan Kaufmann, 1990.
- [4] X. Yao, *Evolutionary Computation: Theory and Applications*, World Scientific, 1999.
- [5] P. J. Darwen and X. Yao, An experimental study of n-Person iterated prisoner's dilemma games, *Informatica*, Vol. 18, pp. 435-450, 1994.
- [6] Y. C. Seo, S. B. Cho and X. Yao, The impact of payoff function and local interaction on the n-player iterated prisoner's dilemma, *Knowledge and Information Systems: An International Journal*, 2(4), pp. 461-478, 2000.
- [7] D. Ashlock, N. Leahy and B. Wagner, A representational sensitivity study of game theoretic simulations, *Congress on Evolutionary Computation*, 2000.
- [8] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, 1989.
- [9] D. Ashlock, M. D. Smucker and E. A. Stanley, Preferential partner selection in an evolutionary study of prisoner's dilemma, *Biosystems* 37, pp. 99-125, 1996.