

# 학습기법을 이용한 멀티에이전트 시스템 개발에 관한 연구

이향란\*, 박성수\*\*, 이말레\*\*\*

\*군산대학교 정보통신학과

\*\*전북대학교 정보통신대학원

\*\*\*여수대학교 멀티미디어학부

e-mail:ilmarekim@naver.com

## A Study on for Multi-Agent System Development using Learning Method

Lee hang-Ran\*, Park seng-Su\*\*, Mal-rey Lee\*\*\*

\*Dept. of Information Communication, Kunsan National University

\*\*Dept. of Computer Engineering, Jeon-Buk National University

\*\*\*Dept. of Multimedia, Yosu National University

### 요 약

멀티에이전트 시스템은 분산적이고 개방적인 인터넷 환경에 잘 부합된다. 멀티 에이전트 시스템에서는 각 에이전트들이 자신의 목적을 위해 행동하기 때문에 에이전트간 충돌이 발생하는 경우에 조정을 통해 협력할 수 있어야 한다. 그러나 기존의 멀티 에이전트 시스템에서의 에이전트 간 협력 방법에 관한 연구 방법들은 동적 환경에서 서로 다른 목적을 갖는 에이전트간의 협동 문제를 올바르게 해결할 수 없다는 문제가 있었다. 본 논문에서는 신경망과 강화학습을 이용하여 목적 패턴을 정확히 결정할 수 없는 복잡하고 동적인 환경에

멀티 에이전트의 시스템 모델을 제안한다.

### 1. 서론

멀티 에이전트 시스템은 그 연구의 시작을 분산 인공지능에 두고 있으며, 에이전트들 사이의 공동 작업을 통해서 각 에이전트의 능력 이상을 요구하는 문제의 해결 방법을 찾으려는 연구 분야이다[2, 3]. 멀티 에이전트 시스템에 관한 연구는 에이전트들이 해결해야 할 공통의 목표를 설정하고, 각 에이전트들 간의 협력과 조정 과정을 거쳐 주어진 문제를 해결하려는 방식을 사용한다. 멀티 에이전트 시스템에서는 각 에이전트들이 병렬적으로 수행되기 때문에 본질적으로 병렬성을 요구하는 많은 문제 영역에 쉽게 적용될 수 있으며 이런 병렬성은 특정 부분의 오 동작으로 인한 시스템의 전체적인 성능

저하를 방지 할 수 있을 뿐만 아니라, 문제가 여러 부분으로 나뉘기 용이할 경우에는 전체 문제 해결 시간을 줄일 수 있다. 또한 전체 문제를 여러 개의 부분 문제로 나누어 각 부분문제를 해결하기 때문에 에이전트의 독립적으로 설계하므로, 중앙 집중적인 시스템과 달리 간단한 프로그래밍 모델을 통한 문제의 해결이 가능하며, 새로운 기능을 가지는 에이전트를 추가함으로써 쉽게 시스템을 확장할 수 있다.

멀티 에이전트 시스템에서 가장 중심적인 연구 과제는 에이전트간의 조정, 협력에 대한 것이다. 에이전트들은 자신에게 할당된 문제를 풀어 나가는 과정에서 부분 관찰(local view), 다중목표(multiple goal), 분산된 정보 등의 제약 때문에 혼란을 겪어

나, 에이전트간의 충돌을 일으킬 수 있다. 또한, 에이전트들이 해결하고자 하는 문제에 대한 제약을 만나거나, 개 개의 에이전트가 가지고 있는 각기 다른 능력과 특별한 지식들을 서로 공유할 필요가 있을 경우 에이전트간의 협력이나 조정이 필요하게 되며, 다른 에이전트의 행동에 따라 자신의 행동이 결정될 경우나, 전체 시스템의 효율을 높이기 위해서도 조정이나 협력이 필요하게 된다. 그러나 이를 위하여 제안된 여러 방법들은 각 에이전트들의 역할이 한 번 주어지면 고정되어 동적으로 변하는 개방 환경에의 적용에 적합하지 않다는 문제점을 가지고 있었다.

본 연구에서는 이런 문제점을 극복할 수 있는 에이전트 사이의 역할 조정 모델을 제안한다. 이 모델에서는 에이전트들 사이에 역할 충돌이 일어나는 경우에 강화학습을 이용하여 자신의 역할을 수정한다. 본 논문의 구성은 다음과 같다. 먼저 2 장에서는 기존의 에이전트간 협력 방법에 관한 연구 방법들과 그 제약점 들을 살펴보고, 3 장에서는 동적인 환경에 적용 가능한 역할 조정 모델을 제안한다. 4 장에서는 제안한 모델을 인공생명 경쟁 문제에 적용한 실험 결과를 살펴보고 5 장에서 결론을 내린다.

## 2. 기존에이전트간 협력 방법

기존의 멀티 에이전트 시스템에 대한 연구에서 제안하고 있는 에이전트간의 조정 기술에는, 에이전트의 유기적인 구조를 설계하는 방법, 에이전트간의 계약을 이용하는 방법, 멀티 에이전트 계획을 이용하는 방법, 협상을 이용하는 방법 등이 있다. 우선 에이전트 유기적인 구조를 구성하는 방법은 가장 간단한 조정방법[5]으로, 에이전트들은 master/slave 혹은 client/server 의 구조를 가지도록 설계한 다음, 에이전트간의 계층적인 관계를 통해 협력과 조정을 수행하게 된다. 이 방법에서 에이전트들은 주로 흑판 구조[6]를 이용해서 서로간의 통신을 하게 된다. 그러나 이 경우에는 에이전트들의 구조로 인한 추가적인 제어가 필요하며, 서로간의 통신을 위해서 흑판을 사용하기 때문에 병목 현상을 발생시키는 등, 멀티 에이전트 시스템의 이점을 저하시킬 수 있으며 에이전트들이 단순한 구조를 가져야 한다는 제약을 가지게 된다. 제약[7]을 통한 에이전트들간의 조정은 주로 contract net protocol(CNP)을 사용하며, 이는 분산된 환경에서 에이전트에 대한 자원 할당이나 문제의 분배에 주로 사용된다. 우선 관리자로 설정된 에이전트는 자신에게 할당된 문제를 작은 부분 문제들로 나누어 이를 수행할 다른 에이전트를 찾게 된다. 그럼 관리자가 요구하는 문제를 수행할 수 있는 에이전트는 계약자가 되어서 자신이 선택한 부분 문제를 해결하게 된다. 이런 과정이 재귀적으로 이루어지며 계약자가 되었던 에이전트가 다시 관리자가 되어 문제를 새로이 다른 에이전트에게 할당하게 된다. 에이전트간의 계약에 의한 조정에서는 에이전트간

의 계층적인 관계를 자동으로 만들어 내며, 수행할 문제에 대한 자원의 동적 할당이 가능하고 자연스러운 부하조절이 가능하다는 장점을 지니고 있다. 그러나 에이전트간의 조정이 수동적이며, 조정을 위한 통신이 이루어지기 어려울 경우 전체적인 시스템 자체에 대한 문제를 불러일으킬 수 있다는 단점을 가지고 있다.

또 멀티 에이전트 계획을 이용하는 방법에는 중앙 집중적인 계획 방법과 분산적인 계획 방법이 있다[8]. 중앙 집중적인 계획 방법은 하나의 조정 에이전트가 전체 시스템을 관찰하며, 에이전트간의 충돌이 발생했을 경우, 각 에이전트의 계획을 조정하도록 지시를 내려 전체 시스템의 계획을 수정하게 된다. 분산적인 계획에서는 각 에이전트가 다른 에이전트의 계획에 대한 모델을 가지고 있어 서로 통신하며 작업 수행 중 발생하는 충돌이나 여러 가지 제약사항을 해결하기 위해 각자 가지고 있는 계획을 스스로 수정하게 된다. 이 방법의 경우에는 에이전트가 상당히 많은 양의 정보를 서로 공유해야 하며, 에이전트간의 통신을 위한 추가적인 시간이 필요하게 되어 전체 시스템의 복잡도가 증가하게 되는 문제점을 가지고 있다.

협상을 이용한 에이전트간의 조정은 분산 인공지능에서 조정과 협력을 위한 중요한 방법 중의 하나로 사용되었으며, 기본적으로 에이전트간의 통신을 통해서 상호간에 동의 할 수 있는 결과를 이끌어 내는 것을 의미한다. 이 방법에는 게임이론을 기반으로 하는 협상[9], 계획에 기반을 둔 협상[10], 휴리스틱에 의한 협상[11] 등의 방법이 있다. 게임이론을 이용하는 경우 수익 행렬을 사용하여, 에이전트 행동에 따른 이익에 따라 에이전트의 행동을 조정하는 방법을 사용한다. 이 경우에는 에이전트들이 공유해야 하는 서로에 대한 완전한 지식을 가정하기 때문에 현실 세계의 문제에 적용하기 적당하지 않을 뿐만 아니라, 셋 이상의 에이전트의 경우에는 게임이론을 손쉽게 확장할 수 없다는 단점을 가지고 있다. 계획에 기반한 협상의 경우에는 계획에 대한 사전지식에 따라서 에이전트들이 자동적으로 협상을 수행하게 된다. 우선 에이전트들이 자동적으로 협상을 수행하게 된다. 우선 에이전트는 각자의 행동을 계획하고, 그 계획을 분리된 조정 에이전트가 에이전트의 상태나, 메시지 형태, 대화 방법을 이용해서 협상을 수행하게 된다.

그 밖에도 에이전트간 상호 메시지 전달 없이 사전 협약을 통한 협력 작업을 수행하는 방법[12]이나, 개미나 꿀벌 등 곤충들의 행동 유형을 분석하여 에이전트의 행동 결정에 적용하는 방법들이 있었으나, 모두 각 에이전트들에게 한번 할당된 목표는 고정되어 변하지 않았기 때문에 에이전트간의 역할에 충돌이 발생하는 경우 이를 해결할 수 없다는 문제를 가지고 있었다.

## 3. 조정모델

본 연구에서 제안한 조정 모델은 신경망의 입

력으로 현재 에이전트가 위치한 곳을 중심으로 한 5\*5 크기 격자 환경의 상태가 주어진다. 하나의 격자는 {아무도 없음, 자기편 있음, 상대편 있음}과 같은 3 가지의 상태가 가능하므로 신경망이 받을 수 있는 서로 다른 입력 패턴의 수는 대략  $3^{24}$  개 정도로 매우 많다. 또한, 신경망에 입력되는 패턴은 고정된 정적 패턴이 아닌 수시로 변하는 동적 패턴이므로 [입력패턴, 목적패턴] 형태의 학습 패턴을 정하는 것은 비 현실적이다. 따라서, 학습패턴을 이용한 감독자 학습을 수행할 수 없으며 본 조정모델에서는 대안으로 조정함수를 이용하여 강화 신호를 계산하는 강화 학습 모델을 사용한다. 입력 패턴에 대한 목적 패턴을 정해 놓지 않고 이전에 자신이 취했던 행동으로 인해 발생한 환경의 변화를 인지하고 행동의 유용성을 평가하여 조정 모델에 따른 강화 신호를 결정한다. 강화 신호는 인공 유기체 신경망 학습의 초기 오차값으로 입력되며, 은닉층과 입력층으로 전파된다.

목적 패턴의 역할을 대체하는 조정 함수는 멀티 에이전트 학습 행동의 방향과 효과를 좌우한다. 조정함수는 멀티 에이전트의 창발성 원리에 따라 시스템의 하부 구조인 멀티 에이전트의 시점에서 설계한다. 저 수준에서 조정 함수를 적용하더라도 창발성의 원리를 통해 복잡한 집단 학습의 총체적 양상이 나타난다. 본 조정 모델에서는 두가지의 상이한 조정 모델을 적용하였다. 각 조정 모델은 행동의 강화와 억제 두가지 측면에서 설계되었다. 멀티 에이전트의 경쟁 환경에서 유리하다고 판단되는 행동은 (+) 학습을 하여 강화 했으며, 불리하다고 판단되는 행동은 (-) 학습을 하여 억제했다. 멀티 에이전트의 행동 판단은 학습 행동 적용의 초기 단계로서 전문가의 경험에 의해 판단된다. 또한, 행동의 강화와 억제 효과가 있어 그 강도를 달리하기 위해 여러 가지 이산적인 값을 이용하였다. 멀티 에이전트가 환경으로부터 두번의 공격을 받게 되는 행동을 하였다면, 한번의 공격을 받게 되는 행동 보다 두배 크기의 억제 학습을 수행하였다.

멀티 에이전트는 자신으로부터 시작한 “한 스텝”이 모두 끝났을 때 이전 스텝의 행동에 대한 강화 신호를 학습한다. 이를 통해 다른 유기체와 상호 작용하여 받게 되는 강화 신호를 모든 유기체에 대해 집단 내에서의 순서와 관계없이 평등하게 학습한다.

조정모델 1 에서의 멀티 에이전트는 자신을 제외한 모든 것을 환경으로 인식한다. 따라서, 개체는 환경에 대한 공격 성공 시 자기편, 상대편의 구분 없이 (+2) 강화 신호를 받으며, 환경으로부터 공격을 받을 때는 (-)의 강화 신호를 받는다. 조정모델 1 은 멀티 에이전트 유기체의 움직임을 활성화하고 멀티 에이전트의 공격성을 유도함으로써 보다 많은 경쟁 전략들이 출현할 수 있도록 하는

것이 목적이다. 또한 이 모델은 개체의 시점에서 자신에게 주어지는 손익을 따져 강화 신호를 계산한다. 조정모델 1 의 기본 값은 (+1)로서 멀티 에이전트가 환경과의 상호 작용을 하지 않는 경우에는 행동을 강화 시킨다.

이는 멀티 에이전트의 단순한 이동이나, 움직이지 않는 경우까지 행동 강화 학습을 수행한다. 반면에 환경으로부터 공격을 받는 경우에는 피공격 회수에 따라 행동 억제 학습 강도를 달리한다. 행동 강화 학습과 행동 억제 학습 사이의 균형을 위해 공격 성공시에는 (+2)(기본값+ 행동 판단값)의 다소 높은 강화 신호를 제공한다. 또한, 공격 실패 행동은 기본값에 (-1)을 더하여 아무런 학습도 하지 않는다.

이동과 공격은 동시에 성립할 수 없지만 이동과 피공격, 공격과 피공격은 동시에 성립될 수 있다. 예를 들어, 멀티 에이전트가 앞으로 한번 이동한 후 상대편으로부터 두번의 공격을 받으면 최종적으로(-1)=(1+0-2)의 강화 신호를 받는다. 조정모델 1 의 알고리즘은 그림 1 과 같다.

(단계 1): 강화 신호를 기본값으로 초기화  
 (단계 2): if(공격 성공) then 강화신호 = 강화신호 +1  
 (단계 3): if(공격 실패) then 강화신호 = 강화신호 -1  
 (단계 4): if(피공격) then 강화신호 = 강화신호 - (피공격 회수)

[그림 1] 조정모델 1 알고리즘

#### 4. 멀티에이전트 시뮬레이션

본 논문의 멀티에이전트 시뮬레이션에서 한 집단에 포함된 개체의 수는 1000 개, 돌연변이율은 3%로 고정하였다. 시뮬레이션은 [표 2] 와 같이 학습 기능의 유무, 학습률, 조정 모델을 달리하여 5 가지의 실험 모델에 대해 수행하였다.

[표 2] 시뮬레이션 모델

실험 모델	학습 기능	조정 모델	학습률
A	없음	—	—
B	있음	1	2
C	있음	1	4
D	있음	2	2
E	있음	2	4

#### 5. 결론

본 논문에서는 집단간의 상호진화 시뮬레이션에 멀티에이전트 수준의 학습 행동을 도입하고, 그 영향을

평가하였다. 에이전트의 학습 행동을 도입하기 위해 신경망과, 강화 학습을 적용하여 복잡하고 동적으로 결정되는 환경에서의 강화 학습 적용을 위한 조정 모델을 설계하였다.

에이전트의 학습 행동을 도입하였을 때 각 실험 모델 집단의 변화 속도는 빨라졌으며, 전체적인 진화의 성능이 향상되었다. 또한, 학습 행동을 추가하였을 때에 일정한 시간에 보다 우수한 경쟁 전략을 탐색할 수 있었다. 진화의 전체적인 방향은 학습의 영향보다는 환경의 영향을 받았으며, 에이전트 수준의 학습 행동을 통해 진화의 형태를 조절할 수 있었다. 에이전트의 학습 행동의 적용은 특정 목적을 가지는 에이전트 집단의 진화를 가능하게 하며, 이는 학습기법을 여러 분야의 특성에 맞게 적용할 수 있는 메커니즘을 제공한다. 예를 들면, 빨라진 집단의 변화 속도를 통해 환경에 대한 시스템의 적응성을 높일 수 있다. 또는 진화의 형태 조절을 통해 무인 탱크 시뮬레이션에서 공격 성향이 두드러진 경쟁 전략을 초기에 탐색하던지, 방어 성향이 강한 무인 탱크 행위자와 같은 특정 목적을 가지는 에이전트를 진화시킬 수 있다. 또한, 여러 개의 로봇이 공동 작업을 해야 하는 경우 그 행동 전략을 탐색할 수 있다.

본 논문의 시뮬레이션 결과는 강화 신호를 생성시키는 조정 모델에 따라 많은 영향을 받았다. 보다 적응적 시스템을 만들기 위해서는 본 논문에서 학습 행동 적용의 초기 단계로서 휴리스틱에 의거해 설계하였던 조정 모델을 인공 진화 시뮬레이션을 통해 스스로 환경에 적합한 값을 찾아야 하며, 앞으로는 이에 관한 좀더 심도 있는 연구가 필요하다. 또한, 본 논문은 조직체가 환경과의 상호 작용에 의해 그 행동을 변경시키는 에이전트의 개별적 학습에 대해 연구하였다. 앞으로는 대규모의 에이전트들이 전체적인 제어 없이 지역적으로 상호 작용하여 에이전트가 형성되는 에이전트의 발상(morphogenesis)이나 진화의 과정에 의해 가능한 행동의 수가 늘어나는 종족의 발생(phylogenesis)에 관한 연구도 진행되어야 할 것이다.

## [참고 문헌]

- [1] 정보윤 외 2 인, "강화학습을 이용한 멀티에이전트 시스템의 자동 협력 조정 모델", 한국 인지과학회 논문지, 10 권 1 호, pp. 1-11, 1999.3.
- [2] Katia P. Sycara, "Multiagent System", AI MAGAZINE, Summer, 1998.
- [3] H. S. Nwana, L. Lee. and N. R. Jennings, "Co-ordination in Multi-agent systems", Software Agents and Soft Computing, Towards Enhancing Machine Intelligence, Concepts and Applications, Springer, 1997.
- [4] Peter Stone and Manuela Veloso, " Multi-agent Systems: A Survey from a Machine Learning Perspective", IEEE Trans. on Knowledge and Data Engineering, June 1996.
- [5] Durfee E. H., Lesser V. R., and Corkill D. D., "Coherent Cooperation among Communication Problem Solvers", IEEE Trans. Computers, Vol. 11, pp. 1275-1291, 1987.
- [6] Hayes-Roth B, "A Blackboard Architecture for Control", Artificial Intelligence, No. 25, pp 251-321, 1985.
- [7] Smith R. G., " The Contract Net Protocol: High-Level Communication and control in a Distributed Problem Solver", IEEE Trans. on Computers. 29. No .12, 1980.
- [8] Georgeff M., "A Theory of Action for Multi-Agent Planning", Proc. 1984 National Conf. Artificial Intelligence, pp. 121-125, August, 1984.
- [9] David C. Parkes and Lyle H. Ungar, "Learning and Adaption in Multiagent Systems ", AAAI workshop on Multiagent Learning Providence, June 30, 1997.
- [10] Kreifelt T. and von Martial F., " A negotiation framework for autonomous agents", in Demazeau Y. and Muller J. P.(Eds): Decentralized AI2, Elsevier Science, 1991.
- [11] Werkman K. J., "Knowledge-based model of negotiation using shareable perspectives", Proc. Of the 10<sup>th</sup> International Workshop on DAI, Texas, 1990.
- [12] Sandip Sen and Edmund H. Durfee, "The role of commitment in cooperative negotiation ", International Journal on Intelligent & Cooperative Information System, Vol. 3, No. 1, pp. 67-81, 1994.