

인공 유기체의 학습 행동이 게임 캐릭터의 전략에 미치는 영향

* 박사준, 김성환, 김기태
중앙대학교 컴퓨터학과

{phdjoon,xfile119,ktkim}@ailab.cse.cau.ac.kr

Influence of a Game Character's Strategies On Artificial organism's learning behavior

* Sa-Joon Park, Sung-Hwan Kim, Ki-Tae Kim
Dept. of Computer Science & Engineering, Chung-Ang University

요 약

컴퓨터 게임에서의 인공지능은 규칙 기반 추론을 기반으로 한 추론 엔진을 사용하고 있다. 이 규칙 기반 추론 엔진은 비교적 간단하고 구현하기 쉽지만 규칙이 몇 가지 되지 않는다는 것과 규칙 변화가 없는 단점으로 게임 플레이어가 그 규칙들을 쉽게 알아버린다는 문제가 있다. 게임 제작자들은 이런 단점을 극복하고자 게임 플레이어끼리 경쟁을 벌이기 위해서 배틀 넷 등 네트워크 쪽으로 그 단점을 보완하려고 하고 있다. 하지만 오히려 네트워크로의 발전은 더욱 더 인간에 가까운 게임 캐릭터 인공지능을 요구하게 되었으며 규칙 기반 추론 방법으로는 이러한 요구를 충족할 수 없기 때문에 새로운 방법이 필요하게 된 것이다. 이 논문에서는 그 새로운 방법에 대한 대책으로 신경망 알고리즘과 유전자 알고리즘을 사용한 인공생명 방법론으로 그 해결책을 모색하려 한다.

1. 서론

최근의 게임은 단순한 순발력에 의거한 조작이나 단순히 게임을 오래하는 것을 즐기기 보다는 복잡한 환경하에서 다양한 전략과 전술을 구사하여야 하는 방향으로 개발되어지고 있다. 기존의 캐릭터의 움직임은 단순한 규칙기반 추론에 의하여 만들어 진 것으로써, 복잡한 게임 환경에서는 규칙기반 추론에 의하여 모든 상황에 적절히 대처하여 캐릭터의 행동을 규칙화한다는 것은 매우 힘든 일이다. 또한 규칙으로 정해진 캐릭터의 움직임은 사용자가 쉽게 간파할 수 있어서, 그 게임을 조금만 해보아도 그 캐릭터가 어떻게 움직이는 지 알 수 있으므로, 게임의 흥미를 반감시키는 요인이 된다.

본 논문에서는 복잡한 환경하에서 살아있는 듯한 캐릭터를 구현하기 위하여, 인공생명 시뮬레이터를 제안한다.

인공생명은 생명체의 특성인 적응성, 창발성등을 연구하여 이를 응용하고자 하는 학문으로, 게임에 이를 이용하게 되면 복잡한 환경이나 사용자의 조작에 적응하거나 전혀 예상치 못한 창발적인 행동으로 인하여 게임의 흥미와 재미를 높일 수 있을 것이다.

본 논문에서는 게임이론의 고전적인 테마인 추격과 회피에 대하여 인공생명 시뮬레이션을 통하여 각 캐릭터들의 움직임을 진화하는 시뮬레이션을 수행하였다.

2. 기반 연구

2.1. 인공생명

인공생명이란 생명체가 나타내는 현상을 컴퓨터, 로봇과 같은 인공 매체 상에 재현함으로써 생명의 일반적인 특성에 대해 연구하는 학문이다. 생물학이 생물 현상에 대해 분석적 방법으로 접근하였다면, 인공생명은 종합적 방법으로 접근한다. 인공생명은 생명 현상에 대한 이론적 이해를 돕고, 나아가 생명의 원리를 여러 응용 분야에 적용한다.

인공 생명은 현재 생물학의 경계를 이루는 알고 있는 생명(life-as-we-know-it)의 한계를 넘어, 가능한 형태로서의 생명(life-as-it-could-be)에 대한 접근을 가능하게 한다[1].

2.2. 인공생명 시뮬레이션

가능한 형태로서의 인공 생명을 연구하는 데는 그 방법론으로 인공생명 시뮬레이션을 사용한다. 인공생명 시뮬레이션을 위하여 가상의 실험 환경을 구축하고 그 안에서 인공 유기체들의 움직임을 분석하여 생물학 이론에 대한 실험 데이터를 제공하거나, 생명 원리를 탐구한다 [2][3]. 시뮬레이션 환경 안에서 유기체의 일생은 출생, 환경과의 상호작용, 교배와 번식, 죽음 등 일련의 사건들이나 과정으로 표현된다.

인공생명 시뮬레이션은 다음 세 가지 주요 부분으로 구성된다.

- ① 진화를 이루는 유전 알고리즘
- ② 프로그램으로 표현되는 인공 유기체
- ③ 유기체가 살아가는 환경

2.2. 강화 학습

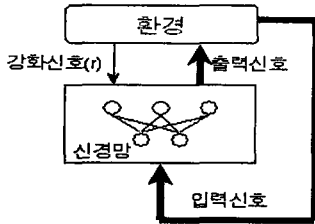
일반적으로 신경망의 학습 방법은 감독자 학습 방법(supervised learning)을 사용한다. [입력 패턴, 목적 패턴]형태의 완전한 학습 패턴을 얻는 것은 매우 어려우며, 때로는 불가능하다. 이런 목적 패턴 없이 자율적으로 학습할 수 있는 모델이 강화 학습(reinforcement learning)이다. 강화 학습의 목적은 강화 신호를 극대화 하도록 입력과 이에 대한 출력을 대응시키는 방법을 학습하는 것이다.

강화 학습의 가장 큰 특징은 여러 번의 시도와 실패를 반복하는 시행착오 탐색(trial-and-error search) 이라는 것과 강화 신호는 행동이 일어난 후에 주어진다는 것이다.

3. 역 전파를 이용한 강화 학습

각 캐릭터에게 학습을 시키기 위해서 오차를 수정하는 방법으로 오류

역 전과 방법을 사용하였다. 목적 패턴과의 오차값을 이용하여 가중치를 조절하지만 여기서는 입력 패턴에 대한 목적 패턴을 정해놓지 않고 자신이 취했던 행동으로 인해 발생한 환경의 변화를 인지하고 행동의 유용성을 평가하여 보상 모델에 다른 강화 신호를 결정되며, 강화 신호로써 가중치 값이 결정되며 이는 환경에 대한 학습을 통해 유기체의 행동이 최적에 방향과 효과를 가지게 한다. 여기 실험에서는 약자와 강자 유기체를 두고서 유리하다고 판단되는 행동은(+)학습을 하여 강화했으며, 불리하다고 판단되는 행동은(-)학습을 하여 억제했다. 이처럼 강화신호로써 학습을 하게 된 것은 오차를 수정 하는 법이 기존처럼 정적이 환경이 아니고 동적이기 때문에 보상모델을 두어 강화 학습을 하게 하였다.



[그림 1] 강화신호

보상 모델은 두 가지로 만들었다. 첫번째 보상모델은 강자 캐릭터와 약자 캐릭터 모두 같은 보상 모델로 학습을 한다. 따라서 어떤 캐릭터든 환경에 대한 공격 성공 시 자기편, 상대편의 구분 없이 (+1) 강화 신호를 받으며, 환경으로부터 공격을 받을 때에는 (-1)의 강화 신호를 받는다. 반면 보상 모델2는 환경을 인식하여 자기편과 상대편으로 분리한다. 따라서, 환경에 대한 공격 성공 시 상대편인 경우는 (+) 강화 신호를, 자기편인 경우에는(-)강화 신호를 받는다. 보상 모델 1과 보상 모델2의 큰 차이점은 약자캐릭터와 강자캐릭터의 보상 값을 달리하여 환경을 인식 시키고 캐릭터 특성을 살려 학습을 하여 캐릭터에 대한 행동 및 성향을 확실하게 제어할 수 있다.

[표 1] 강화신호

종류		보상1	보상2(강자/약자)	
기본값		0	0	
정지		0	0	
이동	성공	+1	+1	
	실패		-1	
공격	성공	+1	자기편	-3
			상대편	+3
	실패	-1	-2	
피 공격		-1	-3	

4. 시뮬레이션

4.1 신경망

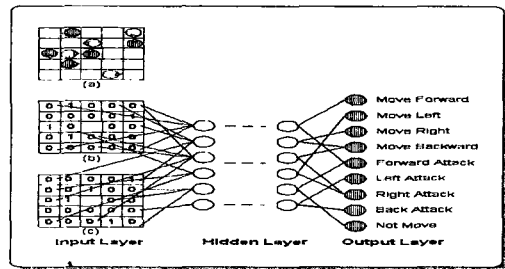
본 시뮬레이션에서는 유전 정보를 이용하여 신경망을 구성하며 이를

캐릭터의 행동함수로 이용하였다. 신경망의 입력으로는 현재 캐릭터가 위치한 곳을 중심으로 5x5의 격자 환경을 인식한다. 자신의 위치를 제외한 24곳에 대하여 상대편과 자신편의 유무에 따라 각각 1,0의 값으로 입력한다. 이런 형식으로 입력 값이 정해지면 은닉층을 통하여 출력결과가 나온다. 출력 결과는 총 9가지로 구분이 된다. 다음 표는 신경망의 각 입력/은닉/출력 단자에 대한 설명이다.

[표 2]신경망 구성 요소

Input Unit	24개의 자신 편에 대한 존재 유무
	24개의 상대편에 대한 존재 유무
Hidden Unit	71개로 내부적으로 개체의 기억장소로 사용
Output Unit	4개의 움직임 방향에 대한 유닛
	1개의 움직이지 않음
	4개의 각 방향에 대한 공격 유닛

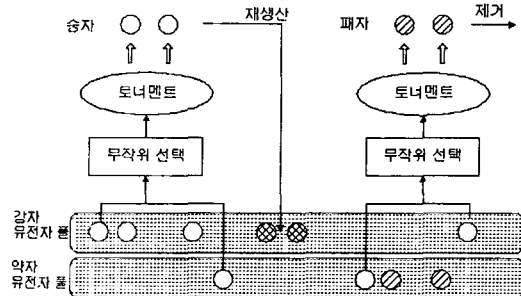
입력 유닛이 48개, 은닉유닛이 71개, 출력 유닛이 9개로 전체 128개의 유닛으로 신경망이 구성되어 진다.



[그림 2] 캐릭터의 상황인식

4.2 경쟁 알고리즘

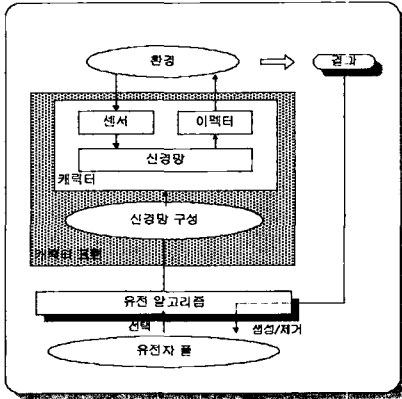
경쟁은 승자 결정 때 두 번, 패자 결정 때 두 번, 총 네 번의 토너먼트를 통해 진행된다. 각 토너먼트마다. 우전자 풀로부터 경쟁에 참여하는 캐릭터 집단을 임의로 선택한 후 토너먼트를 거쳐 패자를 제거하고 승자의 자손으로 대체한다.



[그림 3]경쟁 알고리즘

4.3 시뮬레이션 구조

시물레이션의 전체 구조는 크게 두 부분으로 나뉜다. 첫째는 학습을 하는 부분이다. 강자 유기체와 약자 유기체가 격자 안에서 서로 경쟁을 하는 과정에서 보상 모델을 바탕으로 강화 학습을 통해 강화 신호를 이용하여 신경망에서 앞으로 유기체가 어떻게 행동을 해야 되는지 행동을 설정하도록 하는 과정이다. 둘째는 학습을 통해 우열에 있는 개체를 선별하여, 세대를 재생성하고 열성에 있는 개체를 대체하는 과정이다. 본 시물레이션의 구조는 다음 그림과 같다.



[그림 4] 시물레이션 구조

5. 실험 결과

시물레이션은 A, B, C 세가지 모델로 A-보상 모델을 적용하지 않고, B-보상 모델1을 적용하였으며 C-보상 모델2를 적용하여 실험한 결과 약자 캐릭터와 강자 캐릭터는 공통적으로 처음부터 일정 세대에 이르기까지는 학습과 인공지능적인 진화를 통해서 불필요한 움직임을 줄여서 생존율을 높이려는 모습을 보였다.

약자 캐릭터는 실험 모델 A, B, C에서 공통적으로 강자 캐릭터를 찾아서 공격하는 공격 전략보다는 정지 하여 공격하는 수비 전략을 택해서 생존율을 높이고자 했다. 이는 약자 캐릭터 각각의 생명력이 강자 캐릭터와 비교해서 절반 밖에 안되고 공격력도 떨어지기 때문에 약자 캐릭터를 약자 캐릭터 각각 떨어져서 움직이기보다는 모여서 방어하는 전략이 강자 캐릭터를 상대하는데 좋다는 것을 학습과 인공지능적인 진화를 통해서 알게 된 것이다.

강자 캐릭터는 약자 캐릭터와는 반대로 그 수가 적고 각각의 생명력과 공격력이 강하기 때문에 계속 이동하면서 약자 캐릭터의 빈틈을 노려서 공격하는 모습을 보였다. 개체 수가 약자 캐릭터의 절반임에도 불구하고 이동 횟수 비율이 크고 공격 횟수도 약자 캐릭터 집단과 거의 비슷하다는 것이 약자 캐릭터 집단보다는 강자 캐릭터 집단이 보다 많이 움직이고 공격했음을 의미한다.

실험 모델 A, B, C에서 어떤 의도로 학습을 시킴에 따라서 약자 캐릭터 집단이나 강자 캐릭터 집단의 성격이 많이 달라지는 면을 보였다. 학습을 시키지 않은 실험 모델 A에 비해서 실험 모델 B의 약자 캐릭터 집단과 강자 캐릭터 집단 모두 수비적인 입장을 취했으며 실험 모델 C에서는 약자 캐릭터 집단과 강자 캐릭터 집단가 공격적인 면을 강화시켰다. 또한 실험 모델 A에 비해서 실험 모델 B, C는 불필요한 행동을 줄이는 속도가 빠르게 진행되어서 시간적인 면에서 학습이 효과적임을 나타냈다.

집단간의 우열 비교한 결과를 볼 때 약자 캐릭터가 강자

캐릭터에 비해서 약한 모습을 보여주고 있다. 또한 학습을 한 모델쪽이 학습을 하지 않을 쪽에 비해 강함을 알 수 있다. 그러나 반드시 학습을 한다고 해서 그 결과가 좋은 것은 아니고 적의 전략 구사 방향과 맞지 않은 전략을 구사할 경우 학습된 경우에도 승률이 저조함을 알 수 있다. 이것을 볼 때 각 캐릭터 집단간의 밸런스를 맞추는 데는 캐릭터 각각의 물리적인 체력이나 공격력 혹은 숫자도 중요하지만 그 캐릭터들이 구사하는 전략의 방향도 중요함을 알 수 있다.

실험 결과로 학습을 시킨 모델은 학습을 시키지 않은 모델보다 보다 빠르고 목적성 있게 변화하였다. 또한 세대를 거듭하면서 상대방의 행동 전략에 따라서 수비와 공격을 번갈아 하는 상호 대응 전략을 보였다.

6. 결론 및 향후 과제

본 논문에서는 복잡 다양한 게임환경에서 실제로 살아있는 듯한 캐릭터의 행동을 구현하기 위하여, 인공지능 시물레이터를 이용하는 방법론을 제안하였다.

시물레이션의 결과로 각각의 캐릭터 집단상 상대방의 행동 전략에 따른 다양한 상호 대응 전략을 구사하게 되었으며, 복잡한 게임환경일수록 규칙기반의 추론으로는 어려운 캐릭터의 다양한 행동을 구현할 수 있었다. 즉, 인공지능 시물레이터를 사용하여 캐릭터의 행동을 진화시킴으로써 적응적이고 창의적인 행동을 구현할 수 있다는 것을 증명하였다.

향후 연구 과제로는 인공지능 시물레이션으로 구현된 캐릭터의 행동을 보다 게임적인 요소를 가미하기 위하여 부분적인 규칙의 도입과 게임 기획자가 원하는 행동을 구현하는 방법을 고려하여야 한다.

또한 본 시물레이션에서는 단순한 두 종류의 캐릭터만을 사용하였으나, 보다 다양한 캐릭터를 등장시켜 복잡한 상호작용 환경이 가능하도록 하는 연구가 필요하다.

참고문헌

[1] Langton, C., "Studying artificial life with cellular automata", Physica D, vol 22, pp. 120-149, 1986.
 [2] Miller, G.F. & Cliff, D.(1995) Co-Evolution of Pursuit and Evasion I : Biological and Game-Theoretic Foundations. Cognitive Science Research Paper 311 School of Cognitive and Computing Sciences, University of Sussex.
 [3] Miller, G.F (1995) Artificial Life as Theoretical Biology: How to do real science with computer simulation. Cognitive Science Research Paper 378 School of Cognitive and Computing Sciences, University of Sussex.
 [4] Collins, R.J., Jefferson, D.R., (1991) AntFarm: Towards Simulated Evolution. In Farmer, J.D., Langton, C., Rasmussen, S., and Taylor, C. (Ed.) Artificial Life II, Addison-Wesley.
 [5] 배환국, "인공 유기체 집단간의 경쟁을 통한 상호진화에 관한 연구", 중앙대학교 석사학위 논문