

학습기법을 이용한 인공 생명이 진화에 미치는 영향

이말례

전북대학교 전자정보공학부

목 차

- I. 서 론
- II. 인공 유기체의 학습행동
- III. 인공생명의 진화
- IV. 인공생명 시뮬레이션
- V. 결 론

I. 서 론

인공생명을 연구하는데 있어서 가장 많이 사용되는 방법은 인공 진화이다. 인공진화는 인위적으로 주어진 환경 내에서 인공 유기체들을 재생산, 교배, 돌연변이와 같은 연산을 통해 세대 단위로 진화된다. 인공진화를 통한 학습은 유전자 수준에서 일어나는 본능 행동의 학습으로 자연계 생명체와 같이 다음 세대로 유전된다. 본능 행동은 인공진화에서 선택(selection) 연산을 통해 수정된다. 본능 행동은 개체가 형성될 때 결정되며, 한 세대 내에서는 변경되지 않는다. 본능 행동은 주변 환경의 변경 사항에 대해 즉각적인 대처가 어려우며, 이전 경험을 다음 행동 결정시에 반영하지 못한다. 따라서, 환경변화에 보다 효과적으로 대처하거나 이전의 경험을 활용하기 위해서는 인공진화에 학습 행동의 개념을 도입해야 한다. 학습 행동을 도입함으로써 인공 유기체는 행동 결정시에 자신이 쌓아온 경험을 활용할 수 있으며 자신의 적합도 값을 능동적으로 변화시킬 수 있다.

인공진화에 학습기법을 도입한 연구 논문으로 Nolfi는 개체 수준의 학습행동을 도입하여 인공생명 시뮬레이션을 통해 실험 결과를 보였다. 인공생명 시뮬레이션 수행은 2차원 격자 공간에서 수

행되었고 먹이를 수집하는 인공 유기체는 은닉층을 가지는 다층 신경망을 통해 표현되었다. 신경망의 입력으로는 가장 가깝게 위치한 먹이의 방향과 거리가 주어지며, 신경망 출력층에서는 인공 유기체의 행동이 출력된다. 개체 수준의 학습행동은 출력 패턴과 목적 패턴 사이의 오차값을 이용한 오차 역전파 알고리즘을 사용하였다[2]. 하지만 Nolfi의 연구는 인공 유기체의 행동에 대한 정확한 목적 패턴을 계산할 수 있는 단순한 환경에서만 수행되었다.

문헌 [1]은 인공 유기체 집단간의 경쟁을 통한 상호진화에 관해 연구 하였다. 여기서 인공 유기체는 유전자 정보를 이용하여 신경망을 구성하였으며, 집단을 이루어 간단한 규칙 하에 서로 경쟁하였다. 인공유기체 집단간의 상호진화를 통해 여러 가지 다양한 경쟁 전략이 나타났다[1]. 하지만 이 연구에서는 학습 기능을 추가 하지 않는 상태이다.

본 연구에서는 기존의 인공진화 시뮬레이션에 인공 유기체 수준의 학습 기능을 추가하여 목적 패턴을 정확히 결정할 수 없는 복잡하고 동적인 환경 하에서 학습 행동의 기능을 갖는 인공 유기체의 진화상의 특징을 살펴본다. 이를 위해 복잡한 환경과 다양한 행동을 갖는 유기체 집단간의

경쟁 실험에[1] 학습 행동을 도입하고 학습 행동의 영향을 분석 평가할 것이다.

때에 사용된다. 개선된 Hebb의 규칙은 다음과 같은 단계를 거쳐 학습한다.

II. 인공 유기체의 학습행동

본 장에서는 인공생명 시뮬레이션에 인공 유기체 수준의 학습 행동을 추가하기 위해 개선된 Hebb규칙을 제안하고, 동적으로 결정되는 복잡한 환경에서의 학습을 위해 강화 신호를 생성하는 보상 모델을 제안했다.

2.1 개선된 Hebb의 규칙

Hebb의 규칙은 서로 연결된 ij 신경세포가 동시에 활성화하는 (a)의 경우에만 연결 가중치를 조정한다. 이를 수식으로 표현하면 다음과 같다.

$$W(\neq w)_{i,j} = W(ol d)_{i,j} + \alpha a_i a_j$$

(α 학습률, a_i, a_j : 신경세포 i, j 의 활성화값)

Hebb의 규칙은 실제 신경 세포가 반응하는 방법을 가장 잘 표현하는 학습 규칙이다. Hebb의 규칙을 이용한 신경망은 목적 패턴이 없이 연결된 두개 신경세포의 활성화값과 그것들간의 연결 가중치에 의해서 학습 할 수 있다. 학습의 감독자가 존재하지 않는 무감독자 학습(unsupervised learning)에 적용될 수 있으며, 실제 생물체의 신경 세포 반응과 유사하므로 인공생명 시뮬레이션에 적합한 학습 모델이다. 반면에 연결 가중치에 대한 (-) 조정이 힘들고, 학습률로써 정해진 크기의 값만을 조정할 수 있다. Hebb 규칙의 이러한 제한 사항으로 인해 신경망의 학습속도가 저하될 수 있으며, 시간이 제한되는 경우에는 충분한 학습의 효과를 보장할 수 없다. 이러한 단점을 극복하기 위해 본 논문에서 제안한 것이 개선된 Hebb의 규칙이다. 개선된 Hebb의 규칙에서는 오차 역전파 알고리즘과 같이 목적 패턴과의 초기 오차값을 입력받으며, 이를 은닉층과 입력층을 향해 전파시킨다. 전파된 오차값은 이후 신경망의 가중치를 조절할

(단계 1) : 출력 노드와 초기 오차값($\delta 0$)을 입력 받는다.

(단계 2) : 연결 링크를 따라 오차값을 전파한다.

$$\text{출력 노드인 경우 : } \delta i = \delta 0$$

은닉 노드인 경우 :

$$\delta i = \sum_k \delta k$$

(단계 3) : 연결된 두 신경 세포가 모두 활성화 일 때만 조절값의 계산

$$\epsilon_{i,j} = \alpha \cdot a_i \cdot a_j \cdot \delta_j$$

(a_i, a_j : 신경세포 i, j 의 활성화값)

(단계 4) : 연결 가중치의 조절

$$w(new)_{ij} = w(ol d)_{ij} + \epsilon_{ij}$$

2.2 보상 모델

본 시뮬레이션에서는 신경망의 입력으로 현재 유기체가 위치한 곳을 중심으로 한 5×5 크기 격자 환경의 상태가 주어진다. 하나의 격자는 {아무도 없음, 자기편 있음, 상대편 있음}과 같은 3가지의 상태가 가능하므로 신경망이 받을 수 있는 서로 다른 입력 패턴의 수는 대략 324개 정도로 매우 많다. 또한, 신경망에 입력되는 패턴은 고정된 정적 패턴이 아닌 수시로 변하는 동적 패턴이므로 [입력패턴, 목적 패턴] 형태의 학습 패턴을 정의하는 것은 비현실적이다. 따라서, 학습패턴을 이용한 감독자 학습을 수행할 수 없으며 본 시뮬레이션에서는 대안으로 보상함수(reward function)를 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 보상 모델에 따른 강화 신호를 결정한다. 강화 신호는 인공 유기체 신경망 학습의 초기 오차값으로 입력되며, 은닉층과 입력층으로 전파된다.

목적 패턴의 역할을 대체하는 보상 함수는 인

본 시뮬레이션에서는 유전 정보를 이용하여 신경망을 구성하고, 이를 인공 유기체의 행동 함수로 이용하였다. 인공 유기체의 신경망은 입력 노드 48개, 은닉 노드 71개, 출력 노드 9개로 전체 128개의 노드로 구성된다. 신경망의 입력으로는 현재 유기체가 위치한 곳을 중심으로 55 크기의 격자 환경이 주어진다. 인공 유기체의 신경망은 자기 위치를 제외한 24곳에 대하여 상대편과 자기편의 유무를 1(존재함) 또는 0(존재하지 않음)의 형태로 입력 받는다[그림 7]. 이와 같이 비교적 넓은 범위의 주변 환경을 인식하는 것은 다른 응용 영역에 적용하는데 보다 일반적인 방법이 된다.

표 1. 신경망의 구성 요소

Input node	24개의 주변 격자에 대한 자신편의 존재 여부 24개의 주변 격자에 대한 상대편의 존재 여부
Hidden node	71개의 내부적으로 개체의 기억장소로 사용
Output node	4개 방향에 대한 이동 행동 1개의 움직이지 않음 4개 방향에 대한 공격 행동

인공 유기체의 신경망은 9개의 출력 노드를 가지며 각 출력 노드는 표 2에 나열된 인공 유기체가 취할 수 있는 기본 행동(primitive behavior)을 의미한다.

표 2. 인공 유기체의 행동

Move	Move Forward	Attack	Attack Forward
	Move Backward		Attack Backward
	Move Left		Attack Left
	Move Right		Attack Right
	Stop		

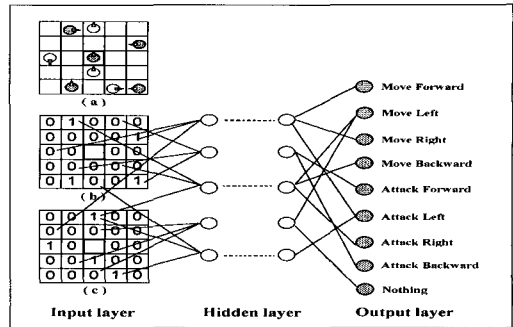


그림 2. 인공 유기체의 신경망 구성

경쟁은 승자 결정, 패자 결정 두 번의 토너먼트를 통해 진행된다. 각 토너먼트마다 유전자 풀(Gene Pool)로부터 경쟁에 참여하는 유기체 집단을 임의로 선택한 후 토너먼트를 거쳐 패자를 제거하고 승자의 자손으로 대체한다.

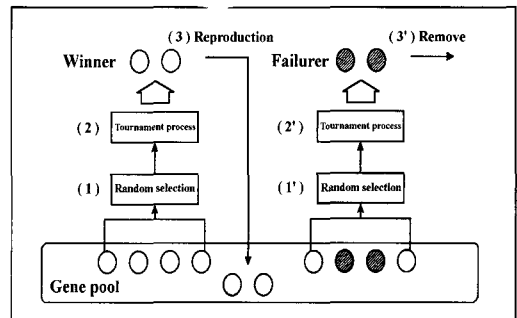


그림 3. 시뮬레이션의 진행

본 시뮬레이션은 진화를 위한 방법으로써 정상 상태 유전 알고리즘(Steady-State Genetic Algorithm : SSGA) 기법을 적용한다. 정상 상태 유전 알고리즘은 전체 개체군에 대한 적합도를 유지하지 않으며, 적합도로는 실제적인 경쟁 결과를 이용한다. 동적 경쟁 환경 하에서의 적합도는 일정 수치 형태로 표현될 수 없으며, 오직 실제 경쟁의 결과만을 판단 기준으로 한다.

IV. 인공생명 시뮬레이션

본 논문의 인공생명 시뮬레이션에서 한 집단에

포함된 개체의 수는 1000개, 돌연변이율은 3%로 고정하였다. 시뮬레이션은 표 3과 같이 학습 기능의 유무, 학습률, 보상 모델을 달리하여 5가지의 실험 모델에 대해 수행할것이다.

표 3. 시뮬레이션 모델

실험 모델	학습 기능	보상 모델	학습률
A	없음	-	-
B	있음	1	2
C	있음	1	4
D	있음	2	2
E	있음	2	4

A 실험 모델은 이전 실험[1]에서 사용하였던 시뮬레이션을 다시 수행하여 얻은 결과로서, 본문에서 제안하는 시뮬레이션 방법의 비교 대상이 된다.

보상 모델은 인공 유기체의 신경망 학습을 위한 강화 신호를 계산하는 방법으로서, 보상 모델 1은 이동 행동과 공격 행동의 강화를 목표로 하였으며, 보상 모델 2는 이동 실패 행동과 공격 실패 행동에 대한 억제와 이를 통한 보다 정확한 행동의 강화를 목표로 하였다. 또한, 학습률은 신경망 학습시에 보상 모델을 통해 주어진 강화신호의 학습 강도를 조절한다. 학습률 4는 학습률 2에 비해 2배의 강도로 강화신호를 학습한다. 이를 통해 인공 유기체의 신경망은 강화신호의 종류와 크기의 영향을 보다 크게 받는다. 학습률은 같은 조건의 실험 환경에 [1, 2, 4, 7.5]의 학습률을 미리 적용하여 보고, 그 중 효과가 좋은 학습률 [2, 4]를 선택하여 사용하였다.

시뮬레이션 수행 중에 100세대 단위로 행동 자료를 누적, 파일로 기록한다. 행동 자료는 이동 성공, 이동 실패, 총공격, 상대편 공격, 자기편 공격, 공격 실패의 6가지 항목에 대한 행동 횟수이며, 이는 100세대 동안의 진화에 대한 분석 자료를 제공한다. 또한, 이 자료는 유기체 집단 전체의 성격에 대한 추정을 가능케 한다. 각 실험 모델에서 얻은 행동 자료를 비교, 분석함으로써 해당 모델의 유기체 집단이 진화한 방향과 속도를

비교한다. 이를 통해 개체 수준의 학습 기능이 전체 집단의 진화에 미친 영향을 평가한다.

100세대 단위로 기록된 행동 자료의 비교, 분석만으로는 인공 유기체의 학습 행동이 실제로 경쟁 전략을 향상시켰는지 확인할 수 없다. 이에 각 실험 모델에서 얻은 특정 세대의 유기체 집단들 사이의 실제 경쟁을 통해 각 방법의 비교 우위를 판정한다.

V. 결 론

본 연구에서는 집단간의 상호진화 시뮬레이션에 인공 유기체 수준의 학습 행동을 도입하고, 그 영향을 평가하고 현재 진행 중에 있다. 인공 유기체의 학습 행동을 도입하기 위해 헤브의 규칙을 개선하고, 강화 학습을 적용하고 있는 상태이다. 개선된 헤브의 규칙은 기존 헤브 규칙에 비해 특정 크기의 오차를 학습할 수 있을거라 예상되며, 이는 신경망의 은닉층과 입력층으로 전파할 수 있다. 또한, 복잡하고 동적으로 결정되는 환경에서의 강화 학습 적용을 위한 보상 모델을 설계하고 있다.

인공 유기체의 학습 행동을 도입하였을 때 각 실험 모델 집단의 변화 속도는 빨라질거라 예상하고, 전체적인 진화의 성능이 향상될것이다. 또한, 학습 행동을 추가하였을 때에 일정한 시간에 보다 우수한 경쟁 전략을 탐색할 수 있을것이다. 진화의 전체적인 방향은 학습의 영향보다는 환경의 영향을 받을것이며, 인공 유기체 수준의 학습 행동을 통해 진화의 형태를 조절할 수 있을거라 확신한다.

참고문헌

- [1] Baldwin, J.M., "A New Factor in Evolution", Adaptive Individuals in Evolving Populations: Models and Algorithms, Belew, R.K., Mitchell, M. (Ed.), Addison-Wesley, 1996.
- [2] Collins, R.J., Jefferson, D.R., "AntFarm: Towards

- Simulated Evolution”, In Farmet, J.D., Langton, C., Rasmussen, S., and Taylor, C. (Ed.) Artificial Life II, Addison- Wesley, 1991.
- [3] Collins, R.J., “Studies in Artificial Evolution”, Phd Thesis Philosophy in Computer Science, University of California, Los Angeles, 1992.
- [4] Filho, J.R., Alippi, C., Treleaven, P., “Genetic Algorithm Programming Environment”, IEEE COMPUTER Journal.
- [5] Goldberg, D.E., “Genetic Algorithms in Search, Optimization and Machine Learning”, Addison - Wesley. ISBN 201 - 15767 - 5, 1989.
- [6] Gruau, F., Whitley, D., “Adding learning to the cellular development of neural networks: Evolution and the Baldwin effect”, Evolutionary Computation I, vol. 3, pp. 213 - 233, 1993.
- [7] Holland, J.H., “Adaptation in Natural and Artificial Systems”, University of Michigan Press, 1975, Reprinted by MIT Press, 1992.
- [8] Kaelbling, L.P., Littman, M.L., Moore, A.W., “Reinforcement Learning: A Survey”, Journal of Artificial Intelligence Research 4, AI Access Foundation and Morgan Kaufmann Publishers, 1996.
- [9] Koza, J.R., “Genetic Programming: On the programming of computers by means of natural selection”, MIT Press, ISBN 0 - 262 - 11170 - 5, 1992.
- [10] Langton, C., “Studying artificial life with cellular automata”, Physica D, vol 22, pp. 120 - 149, 1986.
- [11] Langton, C., “Artificial Life II”, C. Langton (Ed.), Addison - Wesley, pp. 1 - 47, 1989.
- [12] Lin, C.T., Lee, G., “Neural Fuzzy Systems: A Neuro - Fuzzy Synergism to Intelligent System”, Prentice Hall, ISBN 0 - 13 - 235169 - 2, 1996.
- [13] Miller, G.F., Cliff, D., “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, 1995.
- [14] Nolfi, S., Elman, J.L., Parisi, D., “Learning and evolution in neural networks”. CRL Technical Report 9019, University of California, San Diego, 1990.
- [15] Odum, E.P., “Ecology and Our Endangered Life - Support Systems”, chapter 7. co - evolution.
- [16] Parisi, D., Nolfi, S., Ceconi, F., “Learning, Behavior, and Evolution”, Toward a practice of autonomous systems, IT Press, 1991.
- [17] Reynolds, C., “Competition, Coevolution and the Game of Tag”, Artificial Life IV, R. Brooks and Pattie Maes(Ed.), MIT Press, Cambridge, Massachusetts, 1994.
- [18] Sims, K., “Evolving 3D Morphology and Behavior by Competition”, Artificial Life IV, R. Brooks and Pattie Maes(Ed.), MIT Press, Cambridge, Massachusetts, 1994.
- [19] Whitley, D., Gordon, S., Mathias, K., “Lamarckian, The Baldwin Effect and Function Optimization”..

저자소개

이말레



중앙대학교 컴퓨터공학과(공학박사)

USC Post-Doc. Fellow

현재 : 전북대학교 전자정보공학부 교수

※ 관심분야 : 인공지능, 로봇, 멀티미디어, 게임 등