

에이전트의 최적 행동 생성을 위한 행동선택 네트워크의 계획 기능

민현정, 김경중, 조성배
연세대학교 컴퓨터과학과

{solusea, uribyl}@candy.yonsei.ac.kr, sbcho@cs.yonsei.ac.kr

Planning Capability of Action Selection Network for Generating Optimal Behaviors of Agent

Hyeun-Jeong Min, Kyung-Joong Kim and Sung-Bae Cho

Dept. of Computer Science, Yonsei University

요 약

최근 빠른 시간에 행동을 표현할 수 있는 장점을 가진 반응형 시스템과 최적화된 시퀀스를 생성할 수 있는 계획에 기반한 시스템을 통합하기 위한 하이브리드 시스템의 연구가 활발히 진행되고 있다. 행동 네트워크 구조는 센서와 목적에 대한 외부연결과 행동들 사이의 내부연결을 통해 수동적으로 설계되지만, 자동적으로 행동을 생성할 수 있고 복잡한 문제에 적용할 수 있는 장점이 있다. 본 논문에서는 이동 에이전트의 행동을 생성하기 위한 최적화된 방법을 찾는 문제에 대해 이 행동 네트워크에 계획 기능을 추가함으로써 행동 시퀀스를 최적화하는 방법을 제안한다. 행동 네트워크는 입력된 정보와 목적 정보를 가지고 다음에 수행할 행동을 선택하여 각 상황에 가장 높은 우선순위를 가지는 행동만을 선택한다. 이 행동 네트워크에서 선택된 모든 행동들을 몇 단계 앞서 수행시켜 가장 좋은 결과를 가져올 행동으로 다음의 행동을 선택하는 방법을 통하여 복잡하고 불확실한 환경에서 주어진 목표를 달성하기 위한 전체적인 최적 행동 시퀀스를 생성할 수 있다. Khepera 이동 로봇을 이용한 실험을 통해 제안한 행동 네트워크에 계획을 이용한 방법이 행동 네트워크 구조에서보다 더 적은 행동 시퀀스로 목적을 달성함을 알 수 있었다.

1. 서 론

기계에 인간 수준의 지능을 심어주고자 연구해 온 전통적인 인공지능은 상대적으로 잡음이나 불확실성이 배제된 비교적 간단한 환경에서 작동하는 시스템을 고려해 왔다. 미리 정의된 환경에서 최적화된 시퀀스를 계획하고 수행하는 전통적인 계획 기반의 시스템과는 달리 행동에 기반한 시스템은 복잡하고 불확실한 환경에서 빠르게 반응하고 인지할 수 있다[1,2]. 행동 기반의 접근방법은 기능적으로 조건과 행동을 연결지어 빠른 시간에 행동할 수 있는 반응형 시스템의 간단한 연결에만 국한하지 않고 계산할 수 있도록 구체화 한 방법이다. 행동에 기반한 시스템은 빠른 시간에 행동할 수 있도록 다음에 선택할 행동을 결정하기 위해 계산할 수 있도록 했다[3]. 기존의 행동기반 시스템은 환경의 자극에 즉각적으로 반응하여 불확실한 상황에도 효과적으로 대응할 수 있었으나, 복잡한 문제에 적용하기 부적합하고 초기화시에 수동으로 설계되는 제약점이 있었다. 이러한 제약점을 극복하기 위해 행동들을 연결하여 수동과 자동으로 행동을 생성할 수 있는 행동 네트워크가 연구되어 왔다.

동물의 습성을 연구하는 생태학과 기계에 지능을 부여하려는 인공지능 분야가 협력하여, 관찰된 동물들의 행동에 기초한 행동 또는 로봇에서 적절한 행동을 만들어 낼 수 있는 행동 선택 방법(action selection mechanism)을 제안했다[4]. 이 행동 선택 구조에 기반한 행동 네트워크는 센서를 통해 얻어진 정보와 내부 목표를 가지고 하위 레벨의 행동들을 연결하여 환경에서 가장 적절한 행동을 선택하여 실행하는 방법으로 상위 수준의 행동을 효과적으로 생성하기에 유용한 방법이다. 행동 네트워크는 외부 환경의 센서로부터 정보를 얻거나 목적이나 동기로부터 정보를 얻을 수 있다. 외부적인 환경이나 목적 또는 동기로부터의 정보에 부가적으로 내부적인 행동들 사이의 정보를 이

용해서 가장 큰 활성화 값을 갖는 행동을 선택한다.

불확실한 환경에서 반응형 시스템과 계획의 연결 방법인 하이브리드 구조로 유연한 행동을 만들 수 있다[5]. 이 연구에서는 네트워크에 임의의 시작점과 목적점이 주어졌을 때 행동 네트워크 구조보다 더 빠른 시간 내에 혹은 더 짧은 행동 시퀀스를 찾기 위해 행동 네트워크에 계획을 이용하는 방법을 시도했다. 행동 네트워크 구조에 계획을 이용한 접근 방법은 목적지까지의 최적화된 행동 시퀀스를 찾을 수 있도록 한다. 결합하는 방법은 먼저 행동 네트워크에서 n 개의 행동을 선택한다고 하면, 행동 네트워크의 각 단계에서 선택되거나 선택되지 않은 모든 후보 행동들을 고려해서 몇 단계 미리 실행시킴으로 가장 좋은 행동을 선택한다. 그리고 이 행동으로 최적화된 시퀀스를 만든다. 실험을 통해 행동 네트워크에 계획기능을 부가한 방법으로 최적화된 행동 시퀀스를 생성함을 보여준다.

2. 행동 네트워크

행동 네트워크는 그림 1과 같이 구성된다. 그림 1에서와 같이 행동 네트워크는 각각의 기본적인 행동을 가지고 있다. 조건들은 각 행동을 위한 선행조건이며, 목적들은 행동 네트워크를 이용해서 도달하고자 하는 결과이다.

행동 네트워크를 만들기 위해 먼저 각각의 기본 행동들을 정의하고, 센서로부터 얻을 수 있는 정보들로 선행조건을 만든다. 그리고 도달하고자 하는 목적들을 만든다. 외부 연결은 선행조건과 행동을 연결하는 링크와 행동과 목적을 연결하는 링크의 두 가지가 있고, 내부연결은 기본 행동들을 연결하는 링크이다. 이렇게 선행조건 연결(predecessor link)과 후계자 연결(successor link)을 만든다. 각 행동들은 선행조건을 가지고 있

고, 모든 선행조건을 만족해야 그 행동을 선택할 수 있다. 선행 조건은 행동을 수행하기 위해 참이어야 하는 논리적인 조건들이다[6]. 행동 네트워크의 각 행동은 고유의 활성화 값을 가지고 있고, 이 활성화 값이 가장 큰 행동으로 다음 행동을 선택한다.

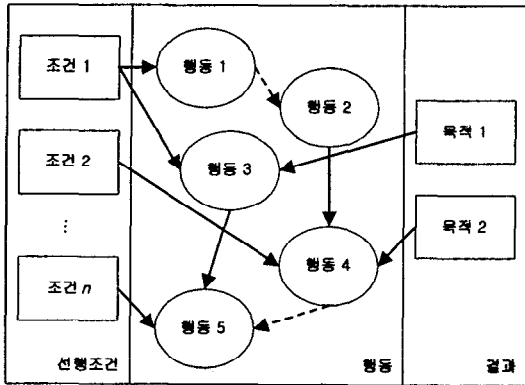


그림 1. 행동 네트워크. 점선은 후계자 연결이고 실선은 선행조건 연결이다.

3. 행동 네트워크와 계획의 결합

행동 네트워크는 센서를 이용하여 목적을 달성하기 위한 행동을 선택하고, 계획은 이 행동 네트워크에서 행동 탐색 트리를 만들고 최적화된 행동 시퀀스를 생성한다. 행동 네트워크와 계획의 결합은 먼저 행동 네트워크를 만들고 그 행동 네트워크를 유지하면서 선택할 수 있는 행동들 중에서 가장 최적화된 행동을 선택하는 것이다. 행동 네트워크의 활성화 값이 큰 행동을 선택해서 적절한 단계까지 행동을 실행시켜보고 마지막 단계에서 가장 최적화된 적절한 행동 시퀀스를 생성한다. 단계를 설정하기 위해서 수행시간을 고려해야 한다. 최대 n 개의 행동이 선택 가능한 행동 네트워크에서 k 단계를 수행시킨다고 가정할 때 수행시간은 최대 $O(nk)$ 가 된다. 이 수행시간을 고려하여 k 를 정의하도록 한다. 그리고 k 단계까지 모든 후보 행동들을 수행해서 가장 목적에 가까운 행동을 선택해서 그 부모 노드로 행동 시퀀스를 생성한다. 이러한 방법을 목적을 달성할 때까지 계속적으로 반복하여 전체적으로 최적화된 행동 시퀀스를 생성한다.

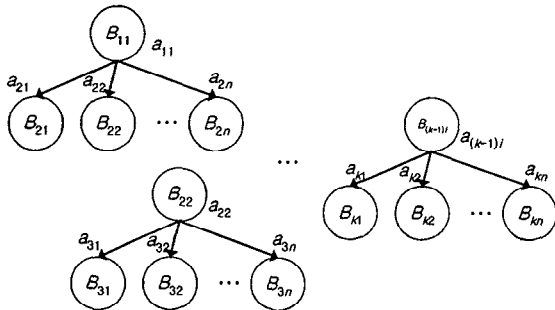


그림 2. 행동탐색 트리의 생성 과정. 최대 n 개의 행동들을 선택하고 k 단계까지 수행한다. 마지막 단계에서 최대 n^k 행동 시퀀스를 얻을 수 있다.

그림 2는 k 단계까지 선택된 행동들을 저장하는 행동탐색 트리의 생성 과정을 보여준다. 그림에서 B_{ij} 는 i 단계의 j 행동, a_{ij} 는 B_{ij} 의 활성화 값을 나타낸다. k 단계에서 목적에 가장 가까운 행동을 선택하고 이 트리를 역으로 탐색해서 최적화된 행동 시퀀스를 찾을 수 있다. 다음은 k 단계에서 최적 행동을 선택하는 함수이다.

$$B_{optimize} = \underset{c}{argmax} (f_{goal}(B_c) + a_{max})$$

여기에서 $a_{max} = \max(a_{ij} \mid i=k \text{ and } 0 < j \leq n)$ 이고 $f_{goal}(B_c)$ 는 행동이 목적에 가장 접근하고 참이 되는 조건을 찾는다. 그림 2에서 행동 시퀀스는 $B_{11} \rightarrow B_{22} \rightarrow \dots \rightarrow B_{(k-1)n}$ 로 결정되었음을 보여준다. 행동 네트워크의 k 단계에서 선택된 최대 n^k 개의 행동들 중에서 가장 목적에 가까운 행동을 선택하고 그 부모 노드를 연결해서 최적화된 행동 시퀀스를 생성한다. 이 임의로 설정한 k 단계까지의 행동 탐색트리를 목적을 달성할 때까지 계속적으로 반복해서 전체적으로 최적화된 행동 시퀀스를 찾을 수 있다.

그림 3은 목적을 달성할 때까지의 최적화된 행동 시퀀스를 생성하기 위한 순서도를 보여준다. 이것은 행동 네트워크에서 적절한 단계의 행동을 선택해서 행동 탐색 트리를 만들고 이 방법을 반복하여 전체적으로 최적화된 시퀀스를 찾음을 나타낸다. 이 순서도는 이동 로봇이 목적을 달성하면 종료된다.

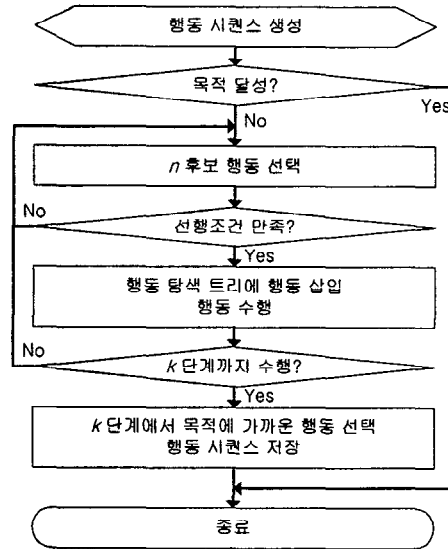


그림 3. 행동 시퀀스를 만드는 순서도.

행동의 선택은 행동 네트워크에서 행동의 활성화 값을 이용하는 Maes의 행동 선택 메커니즘을 따른다. 따라서, 선택된 행동들은 그 행동의 선행조건을 만족해야 하며 또한 활성화 값이 커야 한다. 이 과정으로 선택된 행동들은 행동 탐색 트리에 삽입하고, 각 선택된 모든 행동을 수행한다. 이 행동 탐색 트리의 마지막 단계에서 목적에 가까운 행동을 선택해서 트리를 탐색해서 최적화된 시퀀스를 만든다. 이것을 전체적인 목적을 달성할 때까지 반복한다. 목적을 달성하면 전체적인 최적 시퀀스를 생성한다.

4. 실험 및 결론

실험을 위해 4개의 기본 행동모듈을 선택하는 Khepera 이동 로봇 시뮬레이터를 사용한다. 이 로봇은 배터리를 충전해야만

살아가는 상황으로 설정했고, 로봇은 반드시 배터리 충전영역으로 가서 배터리를 충전해야만 오래 살아남을 수 있다. 행동 네트워크에서의 행동은 CAM-Brain에서 진화된 것과 프로그램된 것을 이용한다[7]. 로봇이 배터리 충전영역으로 이동하기 위한 최적의 행동 시퀀스를 생성하기 위해 계획기능을 결합한다. 행동 네트워크만을 이용하는 경우와 계획기능을 결합한 방법을 비교하여 분석한다. Khepera 로봇은 지역에 대한 정보를 활용하지 않고 오직 센서 정보만을 활용하여 행동 시퀀스를 최적화한다.

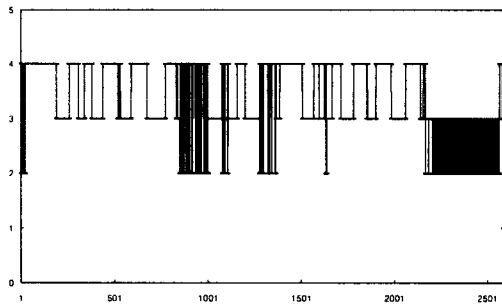


그림 4. 행동 네트워크를 이용한 시뮬레이션에서의 행동 시퀀스. x축은 시간이고 y축은 각 시간에 선택된 행동을 나타낸다.

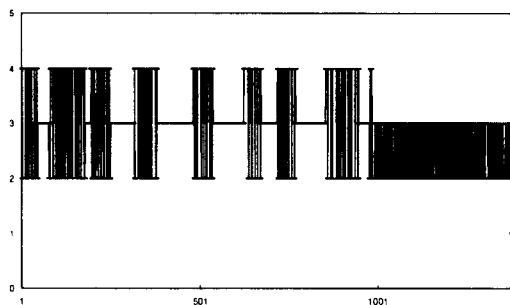


그림 5. 행동 네트워크와 계획을 결합한 방법을 이용한 시뮬레이션에서의 행동 시퀀스. x축은 시간이고 y축은 각 시간에 선택된 행동을 나타낸다.

행동 네트워크를 이용한 방법과 행동 네트워크와 계획의 결합 방법을 비교하기 위해 같은 환경에서 실험하여 행동 선택 횟수를 비교하였다. 그림 4는 행동 네트워크에서 로봇이 배터리를 충전하기까지의 행동 선택 과정을 보여준다. 이 환경에서 로봇은 배터리를 충전하기까지 2586번의 행동 선택이 발생했다. 그림에서 행동 1은 "배터리 충전하기," 2는 "빛 따라가기," 3은 "장애물 피하기," 4는 "직진하기"로 정의한다. 그림 5는 행동 네트워크와 계획의 결합을 이용한 방법에서 로봇이 배터리를 충전하기까지의 행동 선택 과정을 보여준다. 이 환경에서는 배터리를 충전하기까지 1376번의 행동 선택이 발생했다. 계획은 8단계($k=8$)로 설정했다.

그림 6은 행동 네트워크에서와 계획을 결합한 구조에서의 Khepera 이동로봇의 이동경로를 각각 보여주고 있다. 동일한 실험 환경에서 임의의 영역을 선택했다. 그림 6에서 (a)는 약 670번의 행동 선택이 발생했고, (b)는 약 200번의 행동 선택이

발생했다. 실험결과 계획을 결합한 행동 네트워크 구조는 행동 네트워크를 이용한 방법에 비해 더 적은 행동선택으로 목적을 달성하는 것을 알 수 있다.

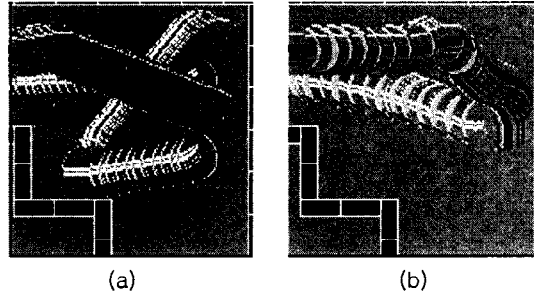


그림 6. 행동 네트워크에서의 로봇의 이동 경로(a)와 행동 네트워크와 계획의 결합에서의 로봇의 이동 경로(b). 동일한 실험환경에서 임의의 영역을 보여준다.

실험을 통해서 행동 네트워크에서의 행동 시퀀스와 계획기능을 부여한 행동 네트워크에서의 행동 시퀀스를 비교해 보았다. 실험결과 계획기능을 부여하여 더 빠른 시간에 행동 시퀀스를 생성함을 알 수 있었다. 행동 네트워크에 계획기능을 이용하기 위해 미리 가능한 모든 행동들을 수행시켜보고 마지막 단계에서 가장 목적에 가까운 행동으로 최적 행동 시퀀스를 생성한다. 향후 동적인 다른 환경에서 행동생성 네트워크의 계획기능으로 최적행동 시퀀스를 만들어 낼 수 있고 공식화된 계획 방법을 찾고자 한다.

감사의 글

이 논문은 한국학술진흥재단의 연구과제(2002-005-H20002)에 의해 지원되었음.

참고문헌

- [1] R. A. Brooks, "Intelligence without representation," *Artificial Intelligence Journal*, vol. 47, pp. 139-159, 1991.
- [2] M. J. Mataric, "Interaction and intelligent behavior," *Ph.D. Thesis*, Univ. of MIT, 1994.
- [3] M. J. Mataric, "Behavior-based control: Main properties and implications," *IEEE International Conference on Robotics and Automation*, pp. 46-54, 1992.
- [4] P. Maes, "Bottom-up mechanism for behavior selection in an artificial creature," In *Proceedings of the First International Conference on Simulation of Adaptive Behavior*, pp. 238-246, 1994.
- [5] R. C. Arkin and D. C. Mackenzie, "Planning to behave: A hybrid deliberative/reactive robot control architecture for mobile manipulation," In *Proceedings of International Symposium on Robotics and Manufacturing*, pp. 5-12, 1994.
- [6] K.-J. Kim and S.-B. Cho, "Dynamic selection of neural network modules based on cellular automata for complex behaviors," *Journal of the Korean Institute of Electrical Engineers*, vol. 51, pp. 160-168, 2001.
- [7] S.-B. Cho and G.-B. Song, "Evolving CAM-Brain to control a mobile robot," *Applied Mathematics and Computation*, vol. 111, pp. 147-162, 2000.