

베이지안 네트워크를 이용한 행동기반 에이전트의 목적추론

김정중, 조성배
연세대학교 컴퓨터과학과

uribyul@candy.yonsei.ac.kr sbcho@cs.yonsei.ac.kr

Goal Inference of Behavior-Based Agent Using Bayesian Network

Kyung-Joong Kim and Sung-Bae Cho
Dept. of Computer Science, Yonsei University

요약

베이지안 네트워크는 변수들간의 원인-결과 관계를 확률적으로 모델링하기 위한 도구로서 소프트웨어 사용자의 목적을 추론하기 위해 널리 이용된다. 행동기반 로봇 설계는 반응적(reactive) 행동 모델을 효과적으로 결합하여 복잡한 행동을 생성하기 위한 접근 방법이다. 행동의 결합은 로봇의 목표, 외부환경, 행동들 사이의 관계를 종합적으로 고려하여 동적으로 이루어진다. 그러나 현재의 결합 모델은 사전에 설계자에 의해 구조가 결정되는 고정적인 형태이기 때문에 환경의 변화에 맞게 목표를 변화시키지 못한다. 본 연구에서는 베이지안 네트워크를 이용하여 현재 상황에 가장 적합한 로봇의 목표를 설정하여 유연한 행동 선택을 유도한다. Khepera 이동로봇 시뮬레이터를 이용하여 실험을 수행해 본 결과 베이지안 네트워크를 적용한 모델이 상황에 적합하게 목적을 선택하여 문제를 해결하는 것을 알 수 있었다.

1. 서론

유연한 에이전트의 행동을 생성하기 위해 인공지능 연구자들은 다양한 방법을 사용한다. 전통적인 계획기법은 계산시간이 많이 걸리고 환경의 변화에 취약하다는 단점을 가지고 있다. 이러한 문제를 보완하기 위해 입력신호를 반응행동(reactive behavior)과 직접 연결시키는 구조가 등장했으며, Rodney Brooks는 로봇을 위해 반응행동(reactive behavior)을 계층적으로 결합하는 포섭구조를 제안하여 행동기반 로봇제어 연구를 시작하였다.

행동기반 에이전트 제어는 기존의 전통적인 계획기법이 가지고 있던 문제를 해결했지만 행동선택, 목적부파, 최적경로탐색 등에서 해결해야 하는 많은 어려움을 가지고 있다. 최근, 기존의 계획기법과 행동기반 시스템을 결합하여 서로의 단점을 보완하려는 시도가 이루어지고 있으며 하이브리드 구조라고 한다. 즉, 기존의 계획기법처럼 앞으로 진행될 행동선택에 대해 대략의 계획을 가지면서도 행동기반 구조를 기초로 하는 방식이다. 이러한 기본적인 행동생성을 다양한 구조와 함께 deliberation, motivation, emotion 등의 기능을 결합하려는 연구도 함께 진행중이다[1].

행동기반 에이전트의 중요한 구성요소인 행동 선택 알고리즘에 관해서도 다양한 연구가 진행되어 왔다. 그중 기존의 계획기법과 반응구조의 중간에 위치하는 행동 네트워크 방식이 있다. 기존의 행동선택 방식과 틀리게 이 방법은 외부환경과 함께 로봇의 내부 목표도 함께 고려하여 행동기반 로봇이 특정한 목표를 달성할 수 있도록 되어있다.

그러나, 행동 네트워크의 구조가 사전에 고정되기 때문에 변화하는 환경에 쉽게 적응하지 못하는 단점이 있다. 즉, 변화하는 상황에 적합한 목표를 설정해 주기 위한 deliberation 기능을 가지고 있지 못하다. 이러한 부분을 보완하기 위해 목적추론에 널리 사용되는 베이지안 네트워크를 적용하여 에이전트의 유연한 행동 생성을 시도하며 이동로봇 에이전트를 이용하여 모델의 타당성을 검증한다(그림 1).

2. 행동 네트워크

행동 네트워크는 행동들 사이의 관계, 목표, 외부환경을 구성요소로 하여 현재 상황에 가장 적합한 행동을 선택하는 모델이다. 행동 네트워크의 각 노드는 기본행동을 나타내며 예지는 기본행동들 사이의 관계를 나타낸다. 노드와 노드 사이의 관계는

predecessor, successor, conflictor 등의 세 가지로 구성되며 노드의 삭제조건, 추가조건, 선행조건 등의 구성요소 관계로부터 설정된다. 노드는 상태와 관련되어 있는 선행조건, 추가조건, 삭제조건과 실행코드, 활성도로 이루어진다.

행동 네트워크의 역할은 고차원의 행동을 기본행동들의 결합으로 구현할 수 있는 방법을 제공하는 것이다. 행동 네트워크는 로봇의 외부환경과 현재 목표에 가장 적합한 하나의 행동을 선택한다. 반복적인 행동선택을 통해 로봇은 정해진 목표를 달성할 수 있다. 행동 네트워크의 구조는 로봇이 실제 작업을 수행하기 전에 설계자에 의해 결정되며 실행되는 동안에 변하지 않는다.

그림 2는 행동 네트워크의 예를 보여준다. B_1, B_2, B_3, B_4, B_5 는 기본행동을 나타내고, S_1, S_2, S_3, S_4 는 환경센서를 나타내며, G_1, G_2, G_3, G_4 는 목표를 나타낸다. 행동들 사이의 연결관계를 나타내는 실선과 점선은 각각 predecessor 연결과 successor 연결을 나타낸다. 각 행동은 활성도 a_1, a_2, a_3, a_4, a_5 를 갖는다. 상태집합 $Z=(S_1, S_2, S_3, S_4, G_1, G_2, G_3, G_4)$ 의 부분집합을 이용하여 각 행동의 선행조건, 삭제조건, 추가조건을 결정한다. 선행조건은 행동이 실행되기 위해 반드시 참이어야 하는 조건이고, 추가조건은 행동이 실행되었을 때 참이 되기 쉬운 조건이며, 삭제조건은 행동이 실행되었을 때 거짓이 되기 쉬운 조건이다.

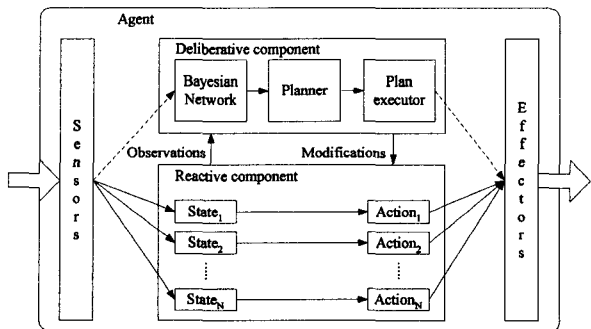


그림 1. 베이지안 네트워크를 이용한 행동기반 구조

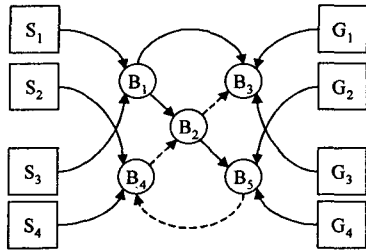


그림 2. 행동 네트워크의 예

```

while (true) {
  /* signal input from sensors and goals */
  for (i=0; i<behavior_number; i++) {
    for (j=0; j<sensor_number; j++) {
      if (sensor[j] is precondition of behavior[i] &&
          sensor[j]==true) alpha[i]+=sensor[j]*weight_a;
    }
    for (j=0; j<goal_number; j++){
      if(goal[j] has link to behavior[i] && goal[j]!=0)
        alpha[i]+=goal[j]*weight_b;}}
  /* signal propagation among behaviors */
  for(i=0;i<behavior_number;i++){
    for(j=0;j<behavior_number;j++){
      if(behavior[i] is predecessor of behavior[j])
        alpha[j]+=alpha[i]*weight_c;
      if(behavior[i] is successor of behavior[j])
        alpha[i]+=alpha[j]*weight_d;
      if(behavior[i] is conflictor of behavior[j])
        alpha[j]-=alpha[i]*weight_e;}}
  normalization(alpha);
  /* behavior selection */
  index = 0;
  for(i=0;i<behavior_number;i++){
    if(alpha[i]>threshold && preconditions of behavior[i]
are true)
      candidates[index++]=behavior[i];}
  if(index==0){threshold*=0.9;}
  else{execute_highest_alpha_behavior(candidates); break;}}
    
```

그림 3. 행동 선택 알고리즘

환경센서가 행동의 선행조건이라면 환경센서와 행동 사이의 연결이 이루어진다. 행동의 실행이 목표를 달성하는데 도움이 된다면 행동과 목표 사이의 연결이 이루어진다. 행동들 사이의 연결은 세 종류로 나뉘며 다음과 같이 정의한다(그림3).

- B_1 is predecessor of B_2 : B_1 의 선행조건 Z_1 이 거짓이고, Z_1 이 B_2 의 추가조건이다.
- B_1 is successor of B_2 : B_1 의 추가조건 Z_1 이 거짓이고, B_1 이 실행가능하며, Z_1 이 B_2 의 선행조건이다.
- B_1 is conflictor of B_2 : B_1 의 선행조건 Z_1 이 참이고, Z_1 이 B_2 의 삭제조건이다.

그림 3은 행동선택 과정을 의사코드로 표현한 것이다. 우선, 환경센서와 목적으로부터 연결이 존재하는 행동의 활성화 값을 올려준다. 두 번째 단계는 행동들 사이의 연결을 통해 활성도를 교환하는 것이다. 활성화도 교환은 행동들 사이에 존재하는 predecessor, successor, conflictor 연결을 통해 이루어진다. 실행조건을 만족하는 행동들 중에서 가장 높은 활성화도를 가지는 행동을 선택한다. Predecessor 연결은 $A \rightarrow B$ 일 경우 A의 활성화도를 B의 활성화도에 더한다. Successor 연결은 $A \rightarrow B$ 의 경우 B

의 활성화도를 A의 활성화도에 더해준다. Conflictor 연결의 경우 $A \rightarrow B$ 인 경우 B의 활성화도를 A의 활성화도 만큼 빼준다.

weight_a, weight_b, weight_c, weight_d, weight_e는 활성화도를 높여주거나 낮추어줄 때 연결의 특성에 맞게 곱해주는 가중치이다[2].

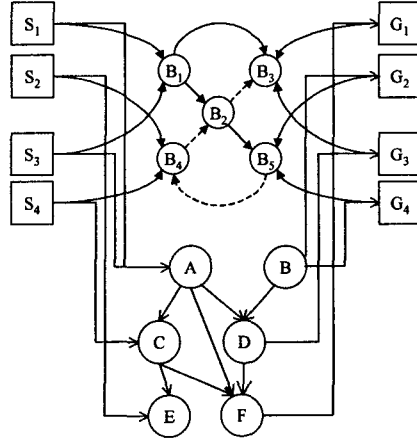


그림 4. 제한하는 모델

3. Deliberative 행동 네트워크 모델

행동 네트워크의 목적은 로봇의 행동선택 순서에 영향을 주어 전통적인 계획기법처럼 작동하도록 한다. 그러나 한번 설정되어진 목적은 로봇이 실행하는 동안 변함없기 때문에 상황의 변화에 적절하게 대처하기 어렵다.

3.1 베이저안 네트워크

베이저안 네트워크는 변수사이의 확률적 의존성을 정의하는 방향성 그래프(DAG: Directed Acyclic Graph) 이다[3]. 그래프의 노드는 변수를 나타내며 각 노드는 조건부 확률 테이블을 가지고 있다(그림 4). 그림 4에서 조건 D는 조건 A와 B의 값이 주어졌을 때, 네트워크상의 모든 다른 노드에 대해 조건적으로 독립이다.

$$\begin{aligned}
 (\forall x_i, y_j, z_k) P(X=x_i | Y=y_j, Z=z_k) &= P(X=x_i | Z=z_k) \\
 P(X_1 \dots X_n | Y_1 \dots Y_m, Z_1 \dots Z_n) &= P(X_1 \dots X_n | Z_1 \dots Z_n) \\
 P(A_1, A_2 | V) &= P(A_1 | A_2, V) P(A_2 | V) = P(A_1 | V) P(A_2 | V) \\
 P(Y_1, \dots, Y_n) &= \prod_{i=1}^n P(y_i | Parents(Y_i))
 \end{aligned}$$

3.2 제한하는 모델

그림 4에서 A, B, C, D, E, F는 deliberation을 수행하기 위해 필요한 변수들이다. 즉, 로봇의 방향이 목표물과 일치하고 있는가, 로봇의 현재위치가 목표물과 가까운가, 로봇이 진행하고 있는 방향으로 장애물이 있는가 등이 된다. 베이저안 네트워크는 가장 높은 확률값을 가지고 있는 목적을 결정하고 연결이 성립되도록 한다.

4. 실험 및 결과

제한하는 에이전트 행동제어 모델의 타당성을 평가하기 위해 이동에이전트를 사용한다. Khepera 이동로봇 시뮬레이터를 사용하여 제한하는 모델의 타당성을 평가한다. 그림 6에서 로봇

은 출발점에서 장애물을 피해 A위치로 이동한 후 B위치로 가는 것을 목표로 한다. A위치에는 빛이 놓여져 있으며 B위치에는 작은 코르크가 놓여 있다. 로봇이 사용할 수 있는 행동은 “빛 따라가기,” “장애물 피하기,” “직진하기,” “오른쪽으로 가기,” “아래쪽으로 가기,” 등이다. “빛 따라가기”와 “장애물 피하기”는 셀룰라 오토마타 기반 신경망을 이용하여 개발하였고, 다른 행동은 직접 프로그래밍하였다.

그림 5은 사용한 행동 네트워크 모델을 보여준다. 행동 네트워크만을 사용한 실험에서 로봇은 A위치에 도달했지만, A위치의 빛에 걸려서 움직이지 못하는 상태에 이르고 말았다. 이러한 문제를 해결하기 위해 그림 6의 베이지안 네트워크를 행동 네트워크와 결합하였다. 반면 제한한 행동 네트워크와 베이지안 네트워크를 결합한 경우 그림 7에서 보듯이 A와 B 모두 지나가는 것을 확인할 수 있었다.

행동 네트워크와 베이지안 네트워크를 결합한 경우, 로봇이 정지했는지 여부를 확률로 계산하여 목적을 결정하는데 반영하였다. 그림 8은 행동 선택 순서를 분석해 본 것이다. 베이지안 네트워크를 사용한 경우가 정보의 양이 더욱 많아 적은 행동만으로 목적을 달성한다.

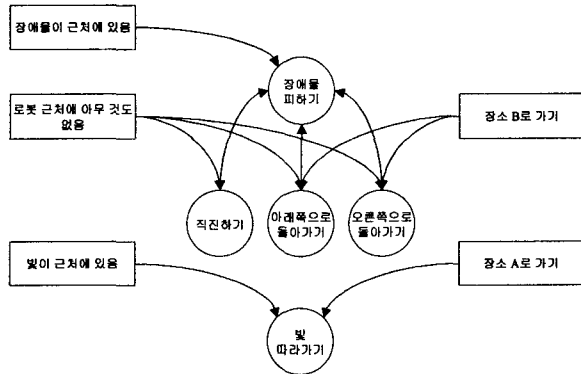


그림 5. 행동 네트워크

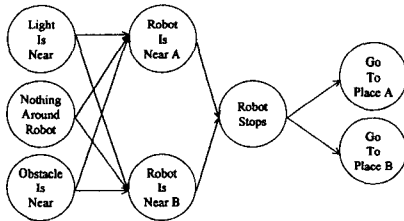


그림 6. 베이지안 네트워크

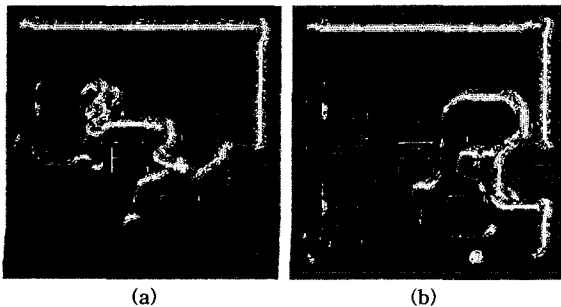


그림 7. 로봇의 이동경로 (a) 행동 네트워크만을 사용한 경우 (b) 베이지안 네트워크와 행동 네트워크를 함께 사용한 경우

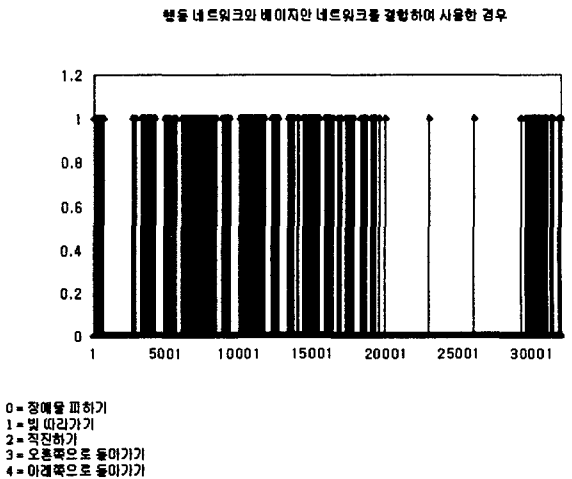
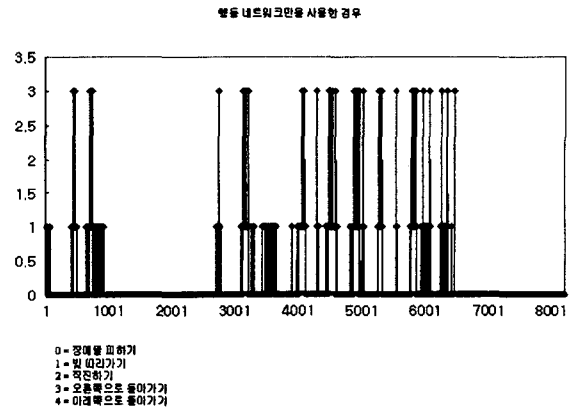


그림 8. 행동 선택 순서

5. 결론 및 향후연구

베이지안 네트워크는 현재 알려진 상황정보를 바탕으로 알려지지 않은 변수의 값을 예측하는데 유용하게 쓰인다. 행동 네트워크는 복잡한 행동을 생성하는데 유용하게 사용된다. 하지만 실제 환경에서 부딪히게 되는 상황을 반영하기에는 어려움이 있다. 이러한 문제를 해결하기 위해 베이지안 네트워크와 행동 네트워크의 결합 모델을 제안하였으며 실험을 통해 결합 모델이 행동 네트워크 모델이 해결하지 못하는 문제를 풀 수 있음을 보였다.

감사의 글

이 연구는 과학기술부가 지원한 뇌과학 연구 프로그램에 의해 지원되었음.

참고문헌

- [1] R. C. Arkin and D. C. Mackenzie, "Planning to behave: A hybrid deliberative/reactive robot control architecture for mobile manipulation," *Proc. 1994 Int. Symp. on Robotics and Manufacturing*, pp. 5-12, August 1994.
- [2] T. Tyrrell, "An evaluation of Maes's bottom-up mechanism for behavior selection," *Adaptive Behavior*, vol. 2, pp. 307-348, 1994.
- [3] D. Heckerman. "A tutorial on learning with Bayesian networks," *Technical Report MSR-TR-95-06*, Microsoft Research, March 1995.