

지능형 에이전트의 움직이는 장애물 충돌 회피를 위한 베이지안 추론 주도형 행동 네트워크 구조

(Bayesian Inference driven Behavior-Network Architecture for Intelligent Agent to Avoid Collision with Moving Obstacles)

민 현 정 [†] 조 성 배 ^{**}

(Hyeun-Jeong Min) (Sung-Bae Cho)

요약 본 논문에서는 변화하는 환경에서 에이전트의 인지 정보로부터 움직이는 물체의 운동모델을 미리 알 수 없는 경우에도 적용할 수 있는 적응적인 행동을 생성하는 방법을 제안한다. 전통적인 에이전트의 지능제어 방법은 환경에 대해 알고 있는 정보를 이용한다는 제약 때문에 강건하지만 다양하고 복잡한 환경에 적용할 수 없었다. 환경에 대한 정보가 없는 상황에서 에이전트가 자율적으로 행동하기 위해서는 행동 기반의 방법이 적합하며, 실제와 같은 변화하는 환경에서 에이전트의 적응적 행동을 위해서는 상황을 미리 추론하고 대처하는 능력이 필요하다. 움직이는 장애물 피하기는 변화하는 환경에서의 적응적 행동생성의 가능성을 보여줄 수 있는 문제이기 때문에 다양한 방법으로 연구되고 있다. 본 논문에서는 고정된 장애물뿐만 아니라 움직이는 장애물을 인지하고 피하는 적응적인 행동을 생성하기 위한 2단계의 제어 구조를 제안한다. 1단계는 상황을 인지하고 자율적으로 행동을 생성하는 행동 네트워크 구조이고 2단계는 변화하는 상황을 추론하고 제어정보를 1단계로 전달하는 베이지안 네트워크 구조이다. 시뮬레이터를 이용한 실험을 통해 제안한 방법으로 고정된 장애물과 움직이는 장애물을 피하고 목적지를 찾아가는 것을 확인할 수 있었다.

키워드 : 행동 네트워크, 베이지안 네트워크, 움직이는 장애물 피하기, 행동기반 AI

Abstract This paper presents a technique for an agent to adaptively behave to unforeseen and dynamic circumstances. Since the traditional methods utilized the information about an environment to control intelligent agents, they were robust but could not behave adaptively in a complex and dynamic world. A behavior-based method is suitable for generating adaptive behaviors within environments, but it is necessary to devise a hybrid control architecture that incorporates the capabilities of inference, learning and planning for high-level abstract behaviors. This paper proposes a 2-level control architecture for generating adaptive behaviors to perceive and avoid dynamic moving obstacles as well as static obstacles. The first level is behavior-network for generating reflexive and autonomous behaviors, and the second level is to infer dynamic situation of agents. Through simulation, it has been confirmed that the agent reaches a goal point while avoiding static and moving obstacles with the proposed method.

Key words : behavior network, Bayesian network, avoiding moving obstacles, behavior-based AI

1. 서론

전통적인 인공지능은 상대적으로 잡음이나 불확실성이 배제된 환경에서 지도 등의 환경정보를 갖고 작동하는 시스템을 고려해 왔다. 미리 정의된 환경에서 최적화된 순서를 계획하고 수행하는 전통적인 계획기반의 시스템과는 달리 행동에 기반한 시스템은 환경 조건과 그

[†] 학생회원 : 연세대학교 컴퓨터학과
solusea@candy.yonsei.ac.kr
^{**} 종신회원 : 연세대학교 컴퓨터산업공학부 교수
sbcho@csai.yonsei.ac.kr
논문접수 : 2004년 1월 29일
심사완료 : 2004년 6월 11일

에 따른 행동의 수행으로 복잡하고 불확실한 환경에서 빠르게 반응할 수 있다[1]. 행동 네트워크는 이러한 행동 기반 시스템의 환경 자극에 즉각적으로 반응하여 빠르게 대응할 수 있다는 장점과 목적을 부여하여 행동 기반의 시스템보다 고 차원의 추상적인 행동을 생성할 수 있다는 장점이 있다[2]. 또한, 전통적인 시스템은 예측 능력과 정확성 면에서 좋은 효과를 얻을 수 있는 반면에 행동 기반의 시스템은 빠르게 반응할 수 있고 환경에 독립적이며 연산이 비교적 간단하다는 장점이 있다[3].

전통적으로 고정된 장애물 피하기의 문제는 최적 경로 탐색이나 지도 제작의 방법으로 해결되어 왔다. 환경에 대한 정보가 있는 경우에 장애물 피하기 문제는 계획 방법으로 최적의 경로를 찾을 수 있다. 그리고 Brooks가 제안한 행동 기반의 시스템을 이용하면 환경에 대한 지도를 가지고 있지 않은 경우도 로봇의 센서 정보와 행동선택 방법을 이용해서 알려지지 않은 환경에서 고정된 장애물 피하기 문제를 해결할 수 있다[4]. 그러나 이러한 방법만으로는 환경이 변하거나 움직이는 장애물과 같은 다양한 상황에는 장애물이 움직이는 방향을 예측할 수 없기 때문에 대처하기 어렵다. 움직이는 장애물 피하기 문제는 미리 정의되어 있는 환경정보를 이용할 수 없고 실시간으로 센서 정보에 따라 상황을 인지해야 하기 때문에 어려움이 따른다.

변화하는 환경은 실시간으로 얻을 수 있는 센서 정보에만 의존해서는 정확한 상황을 추론할 수 없기 때문에, 환경정보를 가지고 있지 않은 에이전트가 자율적으로 행동할 수 있도록 보장해야 한다. 따라서 본 논문에서는 에이전트의 자율적 행동생성을 위한 행동 네트워크 구조와 변화하는 상황에 대처하고 행동 네트워크를 제어하는 베이지안 네트워크의 하이브리드 구조를 제안한다. 에이전트의 센서 정보와 이전 행동, 그리고 움직이는 장애물의 이동방향 등의 정보가 상황 추론을 위한 베이지안 네트워크의 원인 노드가 된다.

이 논문은 다른 하이브리드 방법과 비교하여 세가지 장점을 가지고 있다. 1) 행동 네트워크로 목적을 가진 에이전트가 목적을 달성하기까지 자율적으로 행동을 생성할 수 있다. 2) 베이지안 추론을 이용하여 움직이는 상황을 추론하고 대처할 수 있다. 3) 환경에 독립적으로 적용할 수 있다. 장애물 회피의 문제는 학습이나 계획을 이용한 다른 연구에서도 정교한 결과를 유도해 낼 수 있지만 전처리 시간이 필요하고 환경이 바뀔 때마다 또 다시 학습하거나 계획해야 한다. 이런 점에서 본 논문에서 제안하는 방법은 자율적인 로봇의 행동 생성뿐만 아니라 환경에 독립적으로 추론하고 실시간으로 행동을 생성할 수 있다.

실험을 통해 즉각적인 행동 생성에 적합한 행동 네트워크만으로는 다양한 방향으로 움직이는 장애물에서 피하기 행동을 생성할 수 없음과 제안하는 방법을 비교 분석하고 규칙기반으로 문제를 풀었을 경우의 성공률을 비교함으로써 제안하는 방법의 타당성을 보인다. 논문의 구성은 다음과 같다. 2장에서는 다양한 환경에서의 적응적인 행동 생성에 관한 연구를 살펴보고 3장에서는 제안하는 방법인 행동 네트워크와 베이지안 네트워크, 그리고 그 결합방법에 대하여 소개한다. 4장에서는 제안한 방법의 실험과 그 결과를 분석한 후 마지막으로 5장에서는 결론 및 향후 연구에 대하여 언급한다.

2. 관련 연구

에이전트의 지능적인 행동 연구는 전통적인 지식 기반의 시스템과는 달리 환경에서 자율적, 적응적 행동을 생성하는 능력을 많이 고려하기 때문에, 다양한 환경에서 사람과 같은 행동을 생성하는 에이전트 구조가 제안되었다. 이러한 행동 기반 시스템의 다양한 응용 연구로는 Burgard와 Thrun의 행동 기반 시스템에 계획기능을 부여한 박물관 가이드 로봇이나 Bennett의 자율적인 수중 탐사선 등이 있다[5,6]. 이 행동 기반 시스템은 명확한 지식을 표현하는 전통적인 인공지능 방법과는 달리 에이전트와 환경과의 직접적인 상호작용을 통해 행동을 생성한다. 이러한 센서와 행동의 연결로 즉각적 행동을 생성할 수 있는 행동 기반 시스템에 목적을 부여하여 전체적인 목적에 따라 자율적으로 행동을 선택할 수 있도록 Maes가 제안한 방법이 행동 네트워크이다.

행동 네트워크와 관련된 연구로 Nicolescu와 Mataric은 행동 기반 구조를 바탕으로 순차적인 문제 또는 좀더 복잡한 문제를 해결하기 위해서 행동 기반 구조를 확장시킨 독자적인 행동 네트워크 시스템을 제안하였다[7]. Khoo와 Zubek은 컴퓨터 게임에서의 캐릭터 행동을 생성하기 위해 고정된 행동 네트워크를 이용하였고, 이를 주어진 규칙을 가지고 행동을 생성하는 기존 시스템과 비교하였다[8]. Weigel 등은 축구 로봇의 행동 생성을 위하여 행동 네트워크에 기반한 행동 선택 방법을 적용하였고 다중 로봇의 환경으로 확장하여 연구하였다[9]. Matsuura는 두발로 걷는 로봇의 걷기 동작을 위해서 센서 자극에 반응하여 움직임을 생성하는 조형 행동 네트워크(Formative Behavior Network)라는 제어 시스템을 제안하였다[10].

전통적으로 장애물 피하기는 널리 연구되어 온 문제인데 미리 주어진 정보를 이용한 고정된 장애물 피하기 문제를 주로 다루었다. 그러나 실제 환경에서의 자율적이고 적응적인 행동 생성이 중요한 문제로 대두되면서 변화하는 환경에서의 움직이는 장애물 피하기의 행동

표 1 행동 기반 시스템의 응용 연구

대상	내용	연구자	연도	특징
행동기반 시스템	박물관 가이드 로봇	Burgard와 Thrun[5]	2001	행동기반 시스템에 계획기능 부여로 강건하면서 반사적인 행동 생성
행동 네트워크	수중 탐사선	Bennett[6]	2000	알려지지 않은 환경에서 자율적으로 탐사
	행동 네트워크	Nicolescu와 Mataric[7]	2001	행동기반 시스템에서 순차적인 문제 또는 복잡한 문제를 해결하기 위한 방법 제안
	게임 캐릭터 행동	Khoo와 Zubek[8]	2002	미리 설계된 행동 네트워크 시스템 적용
움직이는 장애물 피하기	GA+추론	Inoue[15]	2000	베이지안 네트워크와 유전자 알고리즘을 이용한 행동생성
	조형 행동네트워크	Matsuura[10]	2000	이족 로봇의 걷기 행동 생성
	퍼지제어 시스템	Mucientes와 Barro[11]	2001	퍼지와 퍼지 추론을 이용한 적응적 행동생성, 많은 퍼지룰을 생성해야 하는 제약
	계획	Lane과 Kaelbling[17]	2001	환경을 state space로 분류하여 계획을 계층적으로 분해 해석하는 방법을 제안
	행동기반 + 학습	Smart[13]	2002	각각의 상황을 학습시켜 움직이는 장애물 피하기의 행동 생성, 학습된 행동에 국한
	축구 로봇	Weigel 등[9]	2002	축구로봇의 행동생성을 위한 행동 네트워크 적용

생성에 많은 연구가 집중되었다. Mucientes와 Barro는 퍼지 추론의 방법을 이용하여 움직이는 장애물 피하기의 퍼지 제어 시스템을 제안하였다[11]. Fujimori와 Tani는 움직이는 장애물을 피하고 목적지까지의 네비게이션을 위하여 다중 로봇이 협력하는 충돌 방지 기법을 이용하였다[12].

움직이는 장애물 피하기 문제를 해결하기 위해 표 1과 같이 행동기반 시스템에 학습[13,14], 추론[11,15], 계획[13,17] 등을 보장하는 연구가 다양하게 진행되고 있다. 본 논문에서는 센서를 이용하여 상황을 추론하는 방법으로 베이지안 네트워크를 이용하였다. 학습이나 계획을 이용한 문제 해결 방법은 복잡한 상황의 정교한 회피에 유용하지만 환경에 종속적이고 다른 환경에 적용하기 위한 복잡성을 감수해야 하는 제약이 따른다. 퍼지 규칙과 진화를 이용한 움직이는 장애물 피하기에 대한 Hasimoto의 연구도 있지만 이 또한 상황에 따라 퍼지 규칙을 정의해야 하고 환경에 많은 변수가 존재하면 할수록 더 많은 퍼지 규칙을 생성해야 한다. 움직이는 장애물 피하기 문제는 불확실성이 내포되어 있는데, 규칙 기반의 방법은 전문가가 모든 정보를 고려하여 작성해야 하고 정보의 변경시 모든 규칙을 고려해야 하기 때문에 불확실성이 존재하는 상황에 대처하기 어렵다. 한편, 베이지안 네트워크는 로보틱스 분야에서의 지도 제작을 위한 추론 방법으로 연구되기도 하였고[18] 대화형 에이전트의 감성기반 추론에서도 연구되었다[19].

3. 베이지안 네트워크 주도형 행동 네트워크

행동 네트워크는 센서 자극과 선행 조건에 따라 행동 state에서 가장 우선순위가 높은 행동을 선택하기 때문에 반사적 행동을 생성할 수 있고, 선행조건, 목적, 그리고 행동의 모델링에 따라 환경에 독립적으로 적용할 수

있으며 다시 모델링하는 불편을 겪지 않아도 된다. 그러나 예측하지 못하는 다양한 상황의 변화에 적응적인 행동을 생성하기는 어렵다. 베이지안 네트워크를 이용한 변화하는 상황의 추론은 기본적인 센서 정보, 장애물의 방향 인지 정보, 이전의 행동정보 등을 이용하므로 환경에 독립적으로 적용할 수 있는 장점이 있다. 따라서 자율적 행동 생성을 위한 1단계의 행동 제어와 적응적인 행동생성을 도와주는 2단계의 상황추론 구조를 제안하며 그림 1은 시스템의 전체적인 구조이다.

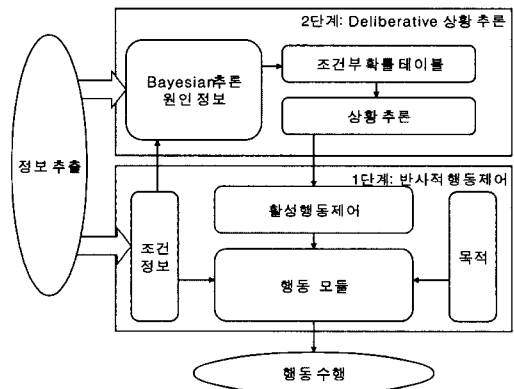


그림 1 행동 네트워크에 베이지안 추론기능을 부여한 구조

3.1 정의

행동 네트워크에서 베이지안 추론 과정을 수학적으로 표현하기 위해 다음과 같이 변수를 정의하여, 제어구조에서의 활성 노드를 선택하기 위한 함수, 각 노드의 활성도 값을 계산하는 함수 및 베이지안 추론의 함수에 사용한다.

- α : 활성화도
- θ : 임계값으로 행동 노드가 활성화되면 값이 초기화되고 행동이 선택되지 않을 때마다 10%씩 감소
- ϕ : 조건이 참이 되도록 하기 위해 센서정보로부터 유도된 활성화 에너지의 양
- γ : 목적을 만족하기 위해 목적정보로부터 유도된 활성화 에너지의 양
- δ : 보호된 목적(protected goal)을 만족시키기 위해 제거된 활성화 에너지의 양
- t : 현재 상태를 나타내는 시간
- B : 행동 네트워크에서 정의된 행동 노드들의 집합
- B_{PS} : 행동 노드 b 의 선행조건들의 집합
- B_{AS} : 행동 노드 b 의 추가조건들의 집합
- B_{DS} : 행동 노드 b 의 삭제조건들의 집합

제어구조에서 현재 상태 t 에서의 활성화 노드 b_i 를 선택하기 위한 함수는 다음과 같이 정의된다. 즉, b_i 의 활성화도 α_b 가 임계값 θ 보다 크고 수행가능하며 다른 노드보다 큰 활성화도를 가질 때 b_i 가 활성화된다.

$$b_i(t) = \begin{cases} 1 & \begin{cases} \alpha_{b_i}(t) \geq \theta \\ \text{if } executable(b_i, t) = 1 \\ \alpha_{b_i}(t) \geq \alpha_{b_j}(t), \forall j: \exists (1) \text{ and } (2) \end{cases} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$executable(b_i, t) = \begin{cases} 1 & \text{if } b_i \text{ is executable at time } t \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

그리고 이 행동 네트워크에서 활성화 노드를 선택하기 위한 행동 노드 b_i 에서의 활성화도 α_b 는 다음과 같다.

$$\alpha_b(t) = \begin{cases} 0 & \text{if } t = 0 \\ D(b_i, t) + \beta(b_i, t) & \text{otherwise} \end{cases}$$

이로부터 행동 네트워크의 활성화도 $D(b_i, t)$ 는 다음과 같이 정의된다.

$$D(b_i, t) = S(b_i, t) + G(b_i, t) - P(b_i, t) + \sum_{b_j, b_k} (SP_{EW}(b_j, b_k, t) + SP_{FW}(b_j, b_k, t) - SP_{FC}(b_j, b_k, t)) \quad (3)$$

$$SP_{EW}(b_j, b_k, t) = \begin{cases} 1 & \text{if } b_j \in B_{PS}, b_k \in B_{AS} \\ 0 & \text{otherwise} \end{cases}$$

$$SP_{FW}(b_j, b_k, t) = \begin{cases} 1 & \text{if } b_j \in B_{AS}, b_k \in B_{PS} \\ 0 & \text{otherwise} \end{cases}$$

$$SP_{FC}(b_j, b_k, t) = \begin{cases} 1 & \text{if } b_j \in B_{DS}, b_k \in B_{PS} \\ 0 & \text{otherwise} \end{cases}$$

여기에서 $\beta(b_i, t)$ 는 베이지안 네트워크에서의 추론결과로 행동 네트워크의 행동 노드 활성화도에 영향을 주는

가중치이다.

$$\beta(b_i, t) = \begin{cases} \xi & \text{if } b_i \in r_{k,i}, i \in I, \Psi(k, i) = \sigma \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$r_{k,i} = \{s \mid s \in \text{effect_node}(k), s: \text{ith state}\}$$

$$I = \{x \mid x \in S_j(t), 1 \leq x \leq \#(\text{Effect Nodes}), S_j(t): \text{causes at time } t\}$$

$$\Psi(k, i) = P(b_m \mid c_{i_1} \dots c_{i_m}): \text{conditional probability, } m: \#(\text{cause nodes})$$

$$\sigma = \text{Max}\{\Psi(k, i) \mid \Psi: \text{conditional probability, } 1 \leq i \leq \#(\text{states in } r_i), k: \text{effect node}\}$$

식 (3)에서 $S(b_i, t)$ 는 시간 t 에 센서로부터 활성화된 입력 정보를, $G(b_i, t)$ 는 시간 t 에 목적으로부터 활성화된 입력을, 그리고 $P(b_i, t)$ 는 시간 t 에 보호된 목적으로부터 활성화의 제거를 나타낸다. 그리고 $SP_{EW}(b_j, b_k, t)$ 는 노드 b_j 에서 b_k 로의 후향 확산(backward spreading)을 나타내는데 이는 b_j 의 선행 조건이면서 b_k 의 추가 조건인 경우이고, $SP_{FW}(b_j, b_k, t)$ 는 b_j 에서 b_k 로의 전향 확산(forward spreading)을 나타내는데 이는 b_j 의 추가 조건이면서 b_k 의 선행 조건인 경우이다. 마지막으로 $SP_{FC}(b_j, b_k, t)$ 는 b_j 에서 b_k 로의 제거를 나타내며 이는 b_j 의 삭제 조건이면서 b_k 의 선행 조건인 경우이다. 식 (4)에서 $r_{k,i}$ 는 베이지안 네트워크에서의 k 번째 결과 노드에서 i 번째의 상태를 의미하고 I 는 t 에서 조건 정보들을 만족하는 결과 노드들의 인덱스로 구성된 집합이다. $\psi(k, i)$ 는 k 번째 결과 노드의 i 번째 상태의 조건부 확률이고, σ 는 k 번째 결과 노드에서 가장 확률이 높은 상태를 나타낸다.

3.2 행동 네트워크

시스템에서 사용된 행동 네트워크는 그림 2와 같이 구성되며 기본적인 행동은 ‘빛 따라가기’, ‘직진하기’, ‘왼쪽으로 돌기’, ‘오른쪽으로 돌기’이다. Maes가 제안한 행동 네트워크는 기본 행동을 구성하는 행동 노드가 각각 선행조건, 추가조건, 삭제조건, 활성화도, 수행 코드를 가지고 있고, 노드들 간에는 내부 연결과 외부 연결을 가지고 있다[2]. 내부 연결에는 선행조건 연결(predecessor link), 후계자 연결(successor link), 충돌자 연결(conflictor link)이 있고 외부 연결에는 센서와 목적으로의 연결이 있다.

행동 네트워크를 구성하기 위해서는 먼저 주어진 문제에 적합한 기본 행동을 정의하고, 센서로부터 얻을 수 있는 정보와 달성하고자 하는 목적을 정의한다. 그림 2의 예에서 센서로 구성된 정보는 ‘빛 근처’, ‘아무것도 없음’, ‘장애물 근처’, ‘그림자 영역’이고, 목적 정보는 ‘목적지 도달’과 ‘장애물 피하기’이다. 그리고 행동 노드는 행동 노드가 활성화되기 위한 선행 조건, 추가 조건, 삭제 조건 및 수행 코드와 활성화도 정보로 구성된다. 수행 코드는 각 행동 노드에 있는 행동을 직접 프로그래밍하

거나 다른 연산을 이용하여 구현할 수 있고, 활성화는 센서 정보와 행동 노드와의 연결 관계에 따라 매번 갱신되는 가중치이다. 표 2는 행동 네트워크를 구성하는 각 행동 노드의 정보이다.

각 행동 노드를 정의한 후에 행동 네트워크의 센서 정보와 목적 정보를 연결하는 연결 정보를 정의해야 한다. 그림 2의 예에서 내부 연결은 선행자 연결, 후행자 연결, 충돌자 연결로 구성되어 있고, 외부 연결은 정의된 센서와 목적 정보를 연결한다. 여기에서 '빛 따라가기' 행동의 선행자 연결은 '빛 근처'의 센서 정보와 '직진하기' 행동이고 후행자 연결은 '목적지 도달'이 된다. '직진하기' 행동의 선행자 연결은 '아무것도 없음'의 센서 정보와 '오른쪽으로 돌기'와 '왼쪽으로 돌기'의 내부 행동으로 구성되며, 후행자 연결은 '목적지 도달'과 '장애물 피하기'의 목적 정보와 '빛 따라가기'의 행동으로 구성되어 있다. '오른쪽으로 돌기' 행동 노드는 '그림자 영역'과 '장애물 근처'의 센서 정보와 '직진하기'의 행동을 선행자 연결로 구성하고 '장애물 피하기'의 목적 정보와 '직진하기'의 행동으로 후행자 연결을 구성한다.

표 2는 내부 연결과 외부 연결이 활성화 되는 조건을 나타낸다. 이 표에서 A와 B는 행동 네트워크를 구성하는 행동 노드이고 p 는 현재 state의 조건이다. 내부연결이 A에서 B로 활성화 될 때 이 활성화가 선행자 연결이면 조건 p 가 거짓이고 A의 선행조건이며 B의 추가조건이어야 한다. 외부연결인 센서정보로의 활성화가 되려면 조건 p 가 참이면서 A의 선행조건이어야 하고 목적정

표 2 활성화 조건

내부 연결 (A → B로의 활성화)	선행자 연결	$(p=false) \wedge (p \in A \text{의 선행 조건}) \wedge (p \in B \text{의 추가 조건})$
	후계자 연결	$(p=false) \wedge (p \in A \text{의 추가 조건}) \wedge (A \text{가 수행 가능}) \wedge (p \in B \text{의 선행 조건})$
	충돌자 연결	$(p=true) \wedge (p \in A \text{의 선행 조건}) \wedge (p \in B \text{의 삭제 조건})$
외부 연결 ($p \rightarrow A$ 로의 활성화)	센서 정보	$(p=true) \wedge (p \in A \text{의 선행 조건})$
	목적 정보	$(p \geq 0) \wedge (p \in A \text{의 추가 조건})$

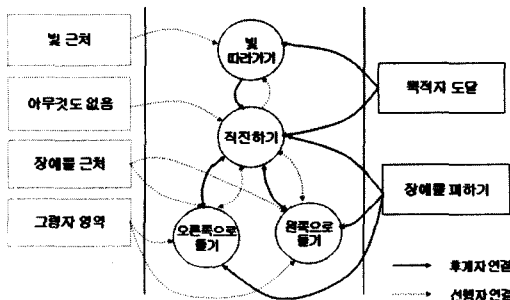


그림 2 행동 네트워크 구조

보로 활성화되려면 조건 p 가 0보다 크거나 같고 또한 A의 추가조건이어야 한다.

실험에 사용된 에이전트인 YAKS라는 시뮬레이션 로봇은 8개의 빛 감지 센서, 8개의 적외선 센서, 색 구분 센서를 가지고 있다. 움직이는 장애물 피하기 실험을 위하여 8개의 적외선 센서를 0부터 7까지의 거리감지 정보 및 에이전트의 이전 행동에 따른 거리변화와 장애물의 위치 변화를 기본 정보로 이용하였다. 1단계의 반사적 행동을 구성하기 위해서는 현재 상황에 적합한 행동을 선택할 수 있도록 거리감지 센서와 빛 감지 센서의 정보를 이용하고, 2단계의 상황 추론을 위해서 이전 행동에 따른 정보를 함께 분석하여, 추론하기 위한 기본 정보로 이용하였다. 그리고 행동의 수행을 위해 에이전트는 두개의 모터를 가지고 있고 이 모터의 속도 제어로 직진하기, 왼쪽으로 돌기, 오른쪽으로 돌기의 행동을 수행할 수 있도록 구성하였다.

3.3 베이지안 추론

베이지안 네트워크는 모든 사건들을 DAG(Directed Acyclic Graph)로 표시하는데 그래프의 각 노드는 부모 노드들과 독립적으로 모델링 된다[20,21]. 즉, 확률 값으로 표시된 현상들의 조건들로 원인과 결과의 관계를 네트워크로 구성하여 원인들로부터 결과를 추론하도록 구성된다. 그림 3은 움직이는 장애물 피하기를 위해 설계된 베이지안 네트워크의 구조이다. 센서 정보로부터 원인으로 구성되는 노드가 '거리감지 0' ~ '거리감지 7'이고, 이전 정보를 이용해서 상황 예측에 필요한 원인으로 구성한 노드가 '장애물과의 거리변화', '장애물 위치변화', 그리고 '이전 행동'과 각 방향에서의 장애물의 변화에 따른 '왼쪽 장애물 변화', '오른쪽 장애물 변화', '정면 장애물 변화', '후면 장애물 변화'이다. 초기값으로 모든 입력정보에 대한 확률은 각 노드에 있는 조건의 개수 n 에 대하여 $\frac{1}{n}$ 로 정의했다. 따라서 '거리감지 0'의 초기 확

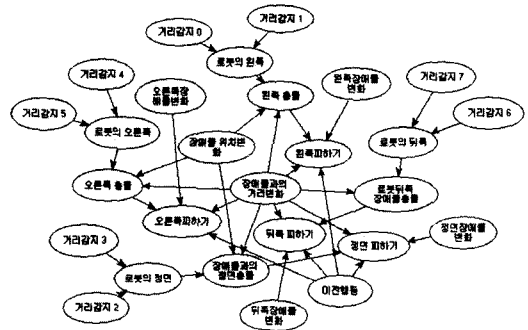


그림 3 움직이는 장애물 피하기를 위한 베이지안 네트워크

물은 $P(\neq ar0) = 0.5$, $P(Far0) = 0.5$ 로 정의된다. 센서 정보에 대한 예측정보도 중요하지만 다양한 환경을 예측하기 위해서는 ‘장애물 위치변화’, ‘이전 행동’, 그리고 각 방향에서의 이전 행동에 따른 장애물의 위치 변화인 ‘왼쪽 장애물 변화’, ‘오른쪽 장애물 변화’, ‘앞쪽 장애물 변화’, ‘뒤쪽 장애물 변화’의 예측 정보가 중요한 인자로 작용한다.

예를 들어서 현재 상태에서 왼쪽에 장애물이 감지되면 ‘거리감지 0’ 또는 ‘거리감지 1’의 노드에서 $P(Near0)$ 와 $P(Near1)$ 의 확률이 높아진다. 그리고 이전 정보에 따라 ‘장애물과의 거리변화’ 노드에서 $P(Approach)$ 와 $P(GoAway)$ 의 확률과 ‘왼쪽 장애물 변화’ 노드의 $P(Front2Left)$, $P(Left)$, $P(Rear2Left)$ 의 확률이 결정된다. 센서 정보로 왼쪽에 장애물이 감지되면서 ‘장애물과의 거리변화’노드에서 $P(Approach) > P(GoAway)$ 이면 ‘왼쪽 장애물 변화’와 ‘이전 행동’ 노드의 확률 값에 따라 ‘왼쪽 피하기’ 노드의 $P(NoTurn)$, $P(LeftTurn)$, $P(RightTurn)$ 중에서 가장 확률이 높은 조건이 선택된다. ‘왼쪽 장애물 변화’노드는 $P(Front2Left)$, $P(Left)$, $P(Rear2Left)$ 의 세 가지 조건으로 구성되어 있는데, 각각 이전에 장애물 감지가 에이전트의 정면에 있다가 이전 행동으로 왼쪽으로 변경되었을 경우, 이전에도 왼쪽에 있었고 현재도 왼쪽에 있는 경우, 그리고 마지막으로 이전에 에이전트의 뒤쪽에 있었는데 현재 왼쪽에서 감지되는 경우를 나타낸다. 세 가지 경우에 따라 마지막 결과 노드인 ‘왼쪽 피하기’의 조건부 확률이 결정된다.

그림 4는 베이지안 네트워크에서 “왼쪽피하기” 노드의 원인이 되는 노드들로 조건부 확률 테이블을 작성한 예를 보여준다. 본 논문에서 필요한 모든 노드로 구성하지 않았지만 이 예로 베이지안 네트워크를 구성하는데 필요한 조건부 확률 테이블에 대해 알 수 있다. “거리감지 n” 노드는 “far”와 “near”의 매개변수를 가지고 있고

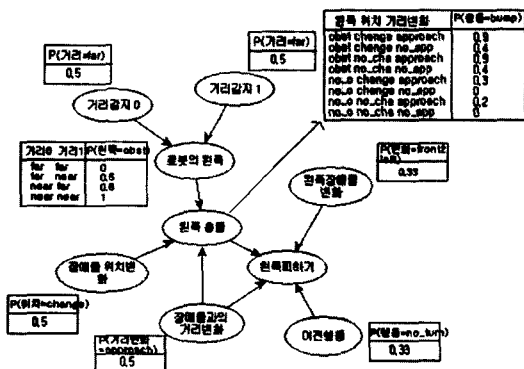


그림 4 “왼쪽피하기”를 위한 베이지안 네트워크에서의 조건부 확률 테이블 작성 예

각각 0.5로 초기화된다. 이 정보는 실험을 하면서 실제 센서 정보로 변경되며, “로봇의 왼쪽” 노드는 실험 로봇의 왼쪽에 장애물이 있을 확률을 계산한다. 조건부 확률 테이블은 그림에서와 같이 두개의 센서가 모두 “far”인 경우, 하나가 “far”이고 다른 하나는 “near”인 경우, 그리고 둘 다 “near”인 경우가 있다. 이 테이블도 각각 왼쪽에 장애물이 있을 확률이 0, 0.6, 1로 작성되어 있고 이 확률은 다른 노드의 정보에 따라 변경된다.

3.4 행동 네트워크에서의 상황 추론기능

베이지안 네트워크를 이용한 상황의 추론을 환경에 독립적으로 적용하기 위해 가능한 센서 정보와 이전 상황과의 관계로 구성된 정보를 행동 네트워크의 행동 선택에 가중치로 제공한다. 본 논문에서 베이지안 추론에 영향을 주는 정보로는 8개의 거리감지 센서, 장애물과의 거리변화, 에이전트에서 장애물의 위치 변화, 에이전트의 이전 행동, 그리고 목적지 방향이 있다. 8개의 거리감지 센서를 이용해서 에이전트의 위치에서 어느 방향에 장애물이 있는지를 추론할 수 있다. 그리고 이전 상태에서 감지된 장애물의 위치와 현재 상태에서 감지된 장애물의 위치 변화에 따라, 또한 에이전트가 선택했던 이전 상태에서의 행동에 따라 다른 조건이 주어진다.

그림 5는 행동 네트워크 구조에서 베이지안 추론을 이용하여 행동 네트워크의 활성도를 갱신함으로써 상황을 예측하는 과정을 나타낸다. 환경과 동기로부터 자극을 계산하고 행동 네트워크를 구성하는 연결자로부터 행동 사이의 자극이 확산되어 활성도 값을 계산한다. 이 행동 네트워크의 과정에 베이지안 추론의 조건부 확률 값을 이용해서 얻은 추론 정보로 행동 네트워크의 활성도 값을 갱신한다. 다시 행동 네트워크는 이 활성도와 선행조건으로 수행 가능한 노드를 체크하고 행동을 수행하도록 한다. 만약 조건을 만족하지 않거나 활성도가 임계 값보다 작으면 임계 값을 조정하고 다시 행동 네트워크의 자극을 계산하는 과정을 반복한다.

앞서 서술한 바와 같이 행동 네트워크는 활성도와 선행 조건을 만족하는 행동을 선택하여 해당하는 수행 코

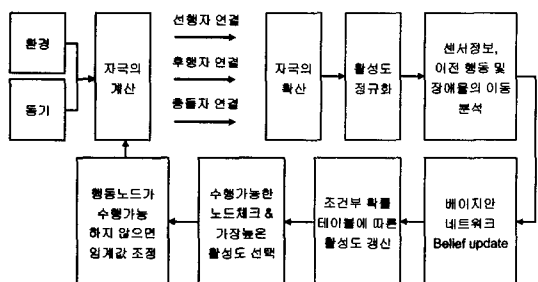


그림 5 행동 네트워크에서의 상태 추론 과정

드를 실행한다. 이 과정에서 베이지안 추론으로 조건부 확률 테이블을 구성해서 현재의 상황을 예측하는 절차를 삽입하고 행동 네트워크의 활성도를 갱신함으로써 에이전트의 행동 선택에 상황 예측의 결과를 반영하도록 한다. 베이지안 네트워크에서 마지막으로 확률 값이 결정되는 노드는 ‘오른쪽 피하기’, ‘왼쪽 피하기’, ‘앞쪽 피하기’, ‘뒤쪽 피하기’이다. 각각의 노드는 ‘No_Turn’, ‘Left_Turn’, ‘Right_Turn’의 확률 변수를 가지고 있고 다양한 정보로 구성된 부모 노드들의 확률에 따라 이 4개 노드의 조건부 확률 테이블이 갱신된다.

베이지안 네트워크의 확률 테이블로부터 행동 네트워크의 활성도를 갱신하는 방법은 먼저 거리 감지 센서로부터 현재 상태에서 감지된 장애물의 위치에 따라서 베이지안 네트워크의 노드를 결정한다. 예를 들어 현재 상태에서 장애물이 감지된 방향이 왼쪽이라면 베이지안 네트워크의 ‘왼쪽 피하기’ 노드에서 갱신된 확률 테이블인 조건 $P(\text{NoTurn})$, $P(\text{LeftTurn})$, $P(\text{RightTurn})$ 중에서 가장 확률이 높은 조건을 선택한다. 그리고 행동 네트워크의 ‘왼쪽으로 돌기’, ‘오른쪽으로 돌기’, ‘직진하기’ 행동 노드에서 $P(\text{NoTurn})$ 인 경우는 ‘직진하기’ 행동에 있는 활성도를 갱신하고 $P(\text{LeftTurn})$ 이 선택된 경우는 ‘왼쪽으로 돌기’ 행동에 있는 활성도를 갱신한다.

4. 실험 및 결과

본 실험에서는 YAKS라는 3D 로봇 시뮬레이터를 사용하고, 실험 환경은 그림 6과 같이 구성한다. 그림에서 4면은 각각 울타리로 막혀 있으며 2개의 모서리에 빛 영역의 목적지가 있다. 시뮬레이션 환경에는 제안된 방법으로 행동을 생성하는 로봇(0번)과 움직이는 장애물 역할을 하는 두 대의 로봇(1번과 2번)이 있다. 이 두 로봇은 울타리만 감지하고 피하며 로봇과 마주쳤을 경우에 움직이는 장애물 역할만 하기 때문에 피하기 행동을 하지 않도록 프로그램 되었다. 그림에서 흰색 원으로 작게 표시된 것이 고정된 장애물인데, 실험에 따라 고정된 장애물과 움직이는 장애물의 개수는 변경될 수 있다. 로봇은 임의의 위치에서 시작되며 행동 네트워크에서의

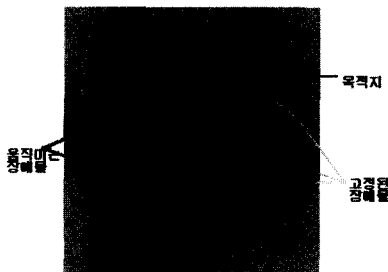


그림 6 시뮬레이터의 실험환경

제안된 제어구조에서의 비교 분석을 위해 움직이는 장애물의 진행 방향을 고정시켜서 실험하고, 제안된 제어구조에서의 분석을 위해서는 다양한 방향으로 움직이도록 실험환경을 구성하였다.

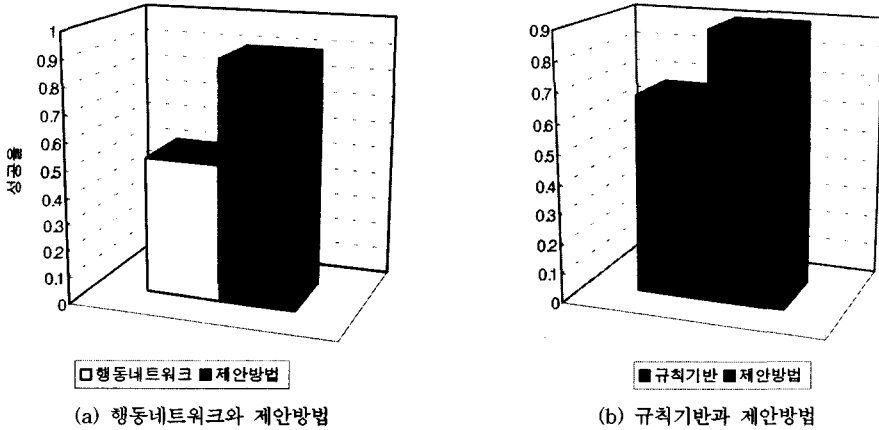
실험의 목적은 로봇이 임의 위치에서부터 고정된 장애물과 움직이는 장애물을 피하면서 대각선 방향의 빛으로 표시된 목적지로 찾아가는 것이다. 행동 네트워크를 구성하는 기본 행동으로는 직진하기, 왼쪽 피하기, 오른쪽 피하기, 빛 따라가기가 있으며, 로봇의 센서로는 장애물과의 거리 감지를 위한 거리감지 센서와 목적지의 빛을 찾기 위한 빛 감지 센서가 있고, 행동을 위한 모터가 있다. 실제 환경에서의 예측할 수 없는 센서의 잡음을 고려하여 시뮬레이션에서 센서정보 값을 읽을 때 난수발생으로 5%의 센서잡음을 넣었다.

4.1 제안한 방법의 비교분석

행동 네트워크와 제안한 제어구조를 이용한 움직이는 장애물 피하기 실험을 한 결과, 행동 네트워크 로봇과 장애물이 같은 방향에서 진행 중일 경우에는 피하기 행동을 할 수 있었으나 서로 다른 방향에서 진행 중일 경우에는 피하기 행동을 할 수 없음을 볼 수 있었다. 또한 ‘오른쪽으로 돌기’와 ‘왼쪽으로 돌기’의 행동 선택에 따라 실패와 성공의 경우가 나뉘었다. 같은 위치에서 로봇을 수행시키고 움직이는 장애물도 같은 위치에서 시작시켜 두 방법을 비교해 보았다. 행동 네트워크에서의 장애물 피하기는 약 70번의 행동 선택에서 움직이는 장애물과 충돌해서 더 이상 진행할 수 없었으나 같은 환경에서 제안한 제어구조를 이용한 로봇은 두개의 움직이는 장애물을 피하고 목적지를 찾아가았다.

행동 네트워크와 제안한 제어구조에서의 성공률과 규칙기반과 제안방법의 성공율을 분석하기 위하여 로봇을 임의의 위치와 각도로 시작하여 장애물과 만났을 때 피하기 행동을 60회 반복 실험하였다. 규칙기반 시스템은 베이지안 네트워크를 이용하여 추론하는 부분을 규칙을 적용하도록 실험하였고 사용된 규칙은 거리감지 센서 정보 4개, 이전 행동 6개, 접근/멀어짐 2개로 48개로 정의하였다. 행동 네트워크와 제안한 방법에서 총 60번의 시도 중 각각 52%와 90%의 성공률을 그리고 규칙기반 방법과 제안한 방법에서 각각 67%와 90%의 성공률을 확인할 수 있었다. 그림 7(a)는 행동 네트워크와 제안한 방법에서, (b)는 규칙기반 방법과 제안한 방법에서 각각 랜덤하게 움직이는 장애물 피하기의 성공률을 나타낸 그래프이다. 그림에서 y좌표가 성공할 확률을 나타내며 0부터 1까지의 값으로 표시된다.

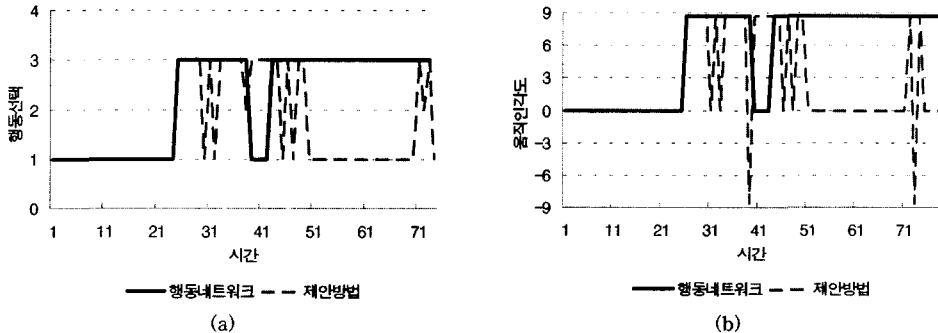
그림 8은 행동 네트워크와 제안한 제어구조에서의 행동선택을 비교한 그림이다. (a)는 두 방법이 선택한 행동의 시퀀스 비교이고 (b)는 두 방법 각각의 행동에 따



(a) 행동네트워크와 제안방법

(b) 규칙기반과 제안방법

그림 7 움직이는 장애물 피하기 성공률 비교



(a)

(b)

그림 8 움직이는 장애물 피하기의 행동 네트워크와 제안한 방법에서의 행동 선택 비교(a)와 각도 분석(b)

른 각도를 분석한 것이다. 그림에서 실선은 행동 네트워크에서의 행동 선택과 각도 분석이고 점선은 제안한 방법에서의 행동 선택과 각도 분석이며, (a)의 행동선택에서 1은 “직진하기”, 2는 “왼쪽으로 돌기”, 3은 “오른쪽으로 돌기”이다. 그림에서와 같이 제안한 방법에서는 상황에 따라 “왼쪽으로 돌기”와 “오른쪽으로 돌기” 행동을 선택했고 행동 네트워크에서는 장애물에 대하여 “오른쪽으로 돌기” 행동을 선택했음을 알 수 있다.

4.2 각도에 따른 움직이는 장애물 피하기

실험 환경에서와 같이 2개의 움직이는 장애물 환경에서 제안한 방법을 이용한 로봇과 장애물과의 이동 각도에 따른 로봇의 행동 선택을 분석해 보았다. 제안한 제어구조에서는 그림 9와 같이 2번의 움직이는 장애물과 3번의 벽 피하기, 그리고 2번의 고정된 장애물 피하기를 하면서 목적지에 도달했음을 알 수 있다. 그림에서 좌측에 있는 것이 로봇과 움직이는 장애물의 trajectory를 그래프로 표시한 것이고 우측에 있는 것이 좌측 로봇의 목적지 도달까지의 행동선택을 나타낸 그래프이다. trajectory에서 동그라미로 표시된 부분이 장애물을 만나서 피하기 행동을 선택한 부분이다. 이 부분에 대해서 우측

의 목적지 도달까지의 행동 선택에서 같은 색의 동그라미로 표시해 놓았다. 장애물을 만났을 경우에 상황에 따라 왼쪽으로 돌기와 오른쪽으로 돌기의 행동을 적절히 선택하면서 움직이는 장애물을 피할 수 있었다. 표 3은 상황 변화에 따른 행동 선택에 대하여 베이지안 네트워크에서 추론한 결과의 예를 보여준다. 거리감지 센서에서 “far”와 “near”의 기준은 실험에 따른 임계값으로 정의하였다. 이 예는 추론에 결정적인 영향을 주는 원인 노드들의 몇가지 정보에 따른 확률 테이블의 변화를 나타낸 것이다. 실제로 베이지안 네트워크의 구성은 규칙기반 방법에서처럼 모든 조건들을 고려하지 않고 모델링할 수 있고 변화하는 조건에 유연하게 대처할 수 있음을 규칙기반 방법과의 비교 실험으로 알 수 있다.

마지막으로 그림 10은 제안한 방법에서 로봇과 다른 방향에서 진행해 오는 움직이는 장애물과 마주칠 경우의 피하기 행동을 보여준다. 이 실험은 로봇의 왼쪽에서 움직이는 장애물과 마주친 경우와 오른쪽에서 움직이는 장애물과 마주친 경우, 그리고 정면에서 오는 경우의 세가지를 분석하였다. 그림 10에서 (a)는 로봇의 왼쪽에서 장애물 2와 같은 방향으로 진행하다가 마주친 경우에

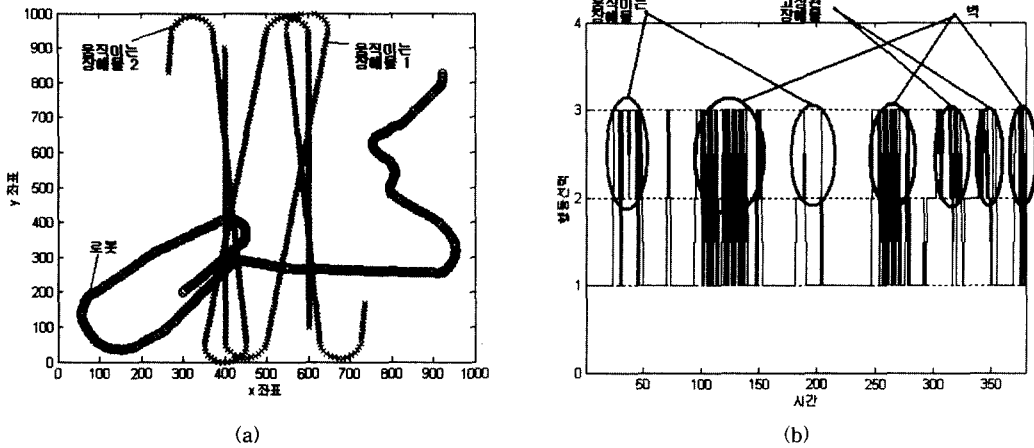


그림 9 제안한 방법에서의 움직이는 장애물 피하기의 trajectory(a)와 행동 선택(b)

표 3 상황 변화에 따른 베이지안 네트워크의 추론결과

원인노드								결과노드						
센서0	센서1	센서2	센서3	센서4	센서5	센서6	센서7	거리 변화	위치 변화	이전 행동	장애물 변화	직진	왼쪽턴	오른쪽턴
far	far	far	far	far	far	far	far	없음	없음	직진	없음	0.8	0.1	0.1
near	near	far	far	far	far	far	far	접근	변화	직진	원→원	0.16	0.12	0.72
near	far	far	far	far	far	far	far	접근	변화	직진	앞→원	0.39	0.06	0.55
far	far	near	near	far	far	far	far	접근	변화	직진	앞→앞	0.08	0.46	0.46
far	far	near	near	far	far	far	far	접근	변화	직진	원→앞	0.11	0.79	0.10
far	far	near	near	far	far	far	far	접근	변화	왼쪽	원→앞	0.11	0.01	0.88
far	far	far	far	near	near	far	far	접근	변화	왼쪽	오른쪽	0.10	0.84	0.06

장애물 피하기를 한 경우이고 (c)는 로봇의 오른쪽에서 장애물 2와 같은 방향으로 진행하다가 마주친 경우에 장애물 피하기를 한 경우이다. (b)는 (a)에서의 시간에 따른 행동선택 시퀀스를 나타낸 그래프이다. x좌표는 시간이고 y좌표는 시간별로 선택된 행동이다. 마지막으로 (d)는 (c)에서의 시간에 따른 행동 선택을 나타낸 그래프이다. 그리고 (e)와 (f)는 로봇의 정면에서 움직이는 장애물과 마주친 경우의 trajectory 그래프와 그에 따른 행동 선택을 각각 나타낸다. (b), (d), (f)에서 x좌표는 시간이고 y좌표는 각 단계별로 선택된 행동이며, y좌표에서 행동은 1번이 '직진하기', 2번이 '왼쪽으로 돌기', 3번이 '오른쪽으로 돌기'를 나타낸다.

5. 결론 및 향후연구

로봇 시뮬레이터를 이용한 실험을 통해서 행동 네트워크와 베이지안 네트워크를 결합한 방법으로 움직이는 장애물 및 고정된 장애물 피하기를 하며 목적지를 찾아갈 수 있음을 확인했다. 행동 네트워크의 기본 행동과 센서와 목적을 연결함으로써 로봇의 자율적인 행동을 생성할 수 있었으며, 로봇의 센서 정보를 이용하여 상황

을 추론하고 그 확률 값을 행동 네트워크의 행동 선택에 가중치로 제공함으로써 변화하는 상황에 적응하는 행동을 생성할 수 있었다.

향후 움직이는 장애물 피하기의 적응적인 행동 생성을 위하여 움직이는 장애물의 속도에 다양성을 부가하고 로봇에도 다양한 속도를 부여하는 것으로 확장할 필요가 있다. 또한 서비스 로봇의 지능적인 행동 생성으로의 확장을 위하여 본 논문에서와 같이 다양한 환경에 적용하면서 순차적으로 주어진 문제를 수행할 수 있는 계획과 추론의 하이브리드 시스템에 대한 연구가 필요할 것이다.

참고 문헌

[1] M. J. Mataric, "Interaction and intelligent behavior," MIT AI Lab Tech Report, 1994.
 [2] T. Tyrrell, "An evaluation of Maes's bottom-up mechanism for behavior selection," *Adaptive Behavior*, vol. 2, no. 4, pp. 307-348, 1994.
 [3] R. A. Brooks, "A robust layered control system for a mobile robot," *IEEE Transactions on Robotics and Automation*, RA-2-1, pp. 14-23, 1986.

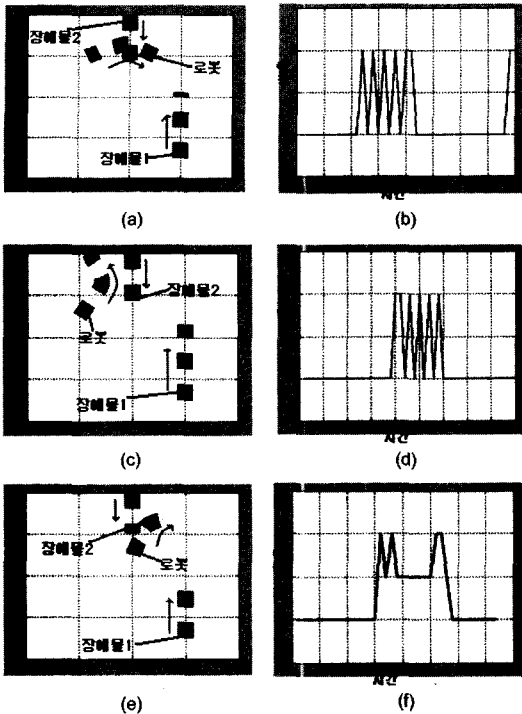


그림 10 제한한 방법에서의 실험 로봇과 다른 방향으로 움직이는 장애물 피하기. (a,c,e)는 각각 로봇의 왼쪽, 오른쪽, 정면 장애물 피하기의 trajectory 이고 (b,d,f)는 각각에 대한 행동 선택

- [4] R. C. Arkin, *Behavior-Based Robotics*, MIT Press, 1998.
- [5] M. Beetz, T. Arbuckle, T. Belker, A. B. Cremers, D. Schulz, M. Bennewitz, W. Burgard, D. Hahnel, D. Fox, and H. Grosskreutz, "Integrated, plan-based control of autonomous robot in human environments," *IEEE Intelligent Systems*, vol. 16, no. 5, pp. 56-65, 2001.
- [6] A. A. Bennett, "A behavior-based approach to adaptive feature detection and following with autonomous underwater vehicles," *IEEE Oceanic Engineering*, vol. 25, no. 2, pp. 213-226, 2000.
- [7] M. N. Nicolescu and M. J. Mataric, "Extending behavior-based systems capabilities using an abstract behavior representation," *AAAI Fall Symposium on Parallel Cognition*, pp. 27-34, 2000.
- [8] A. Khoo and R. Zubek, "Applying inexpensive AI techniques to computer games," *IEEE Intelligent Systems*, vol. 17, no. 4, pp. 48-53, 2002.
- [9] T. Weigel, J.-S. Gutmann, M. Dietl, A. Kleiner, and B. Nebel, "CS Freiburg: Coordinating robots for successful soccer playing," *IEEE Trans. on Robotics and Automation*, vol. 19, no. 5, pp. 685-699, 2002.
- [10] M. Matsuura and M. Wada, "Formative behavior

network for a biped robot: A control system in consideration of motor development," *Robot and Human Interactive Communication*, pp. 101-106, 2000.

- [11] M. Mucientes, R. Iglesias, C. V. Regueiro, A. Bugarin, P. Carinena, and S. Barro, "Fuzzy temporal rules for mobile robot guidance in dynamic environments," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 31, no. 3, pp. 391-398, 2001.
- [12] A. Fujimori and S. Tani, "A navigation of mobile robots with collision avoidance for moving obstacles," *Int. Conference on Industrial Technology*, pp. 1-6, 2002.
- [13] W. D. Smart, "Making reinforcement learning work on real robots," Ph.D. Thesis, Department of Computer Science, Brown University, 2002.
- [14] M. N. Nicolescu and M. J. Mataric, "A hierarchical architecture for behavior-based robots," *Autonomous Agents and Multi-Agent Systems*, pp. 227-233, 2002.
- [15] T. Inamura, M. Inaba, and H. Inoue, "User adaptation of human-robot interaction model based on Bayesian network and introspection of interaction experience," *Proc. of the 2000 IEEE/RSJ Int. Conference on Intelligent Robots and Systems*, pp. 2139-2144, 2000.
- [16] S. Hashimoto, F. Kojima, and N. Kubota, "Perceptual system for a mobile robot under a dynamic environment," *Proc. of 2003 IEEE Int. Symposium on Computational Intelligence in Robotics and Automation*, pp. 747-752, 2003.
- [17] T. Lane and L. P. Kaelbling, "Toward hierarchical decomposition for planning in uncertain environments," *Proc. of the 2001 IJCAI Workshop on Planning under Uncertainty and Incomplete Information*, pp. 1-7, 2001.
- [18] K. Basye, T. Dean, J. Kirman, and M. Lejter, "A decision-theoretic approach to planning, perception, and control," *IEEE Expert*, vol. 7, no. 4, pp. 58-65, 1992.
- [19] T. Akiba and H. Tanaka, "A Bayesian approach for user modeling in dialogue systems," Technical Report of Tokyo Institute of Technology, 1994.
- [20] R. E. Neapolitan, *Learning Bayesian Network*, Prentice Hall Series in Artificial Intelligence, 2003.
- [21] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kauffman, 1988.

민 현 정

정보과학회논문지: 소프트웨어 및 응용
제 31 권 제 3 호 참조

조 성 배

정보과학회논문지: 소프트웨어 및 응용
제 31 권 제 1 호 참조