

지능로봇의 동기 기반 행동선택을 위한 베이지안 행동유발성 모델

Motivation-Based Action Selection Mechanism with Bayesian Affordance Models for Intelligence Robot

*손광희, **이상형, ***서일홍

*Gwang Hee Son, **Sang Hyoung Lee, ***Il Hong Huh

Abstract - A skill is defined as the special ability to do something well, especially as acquired by learning and practice. To learn a skill, a Bayesian network model for representing the skill is first learned. We will regard the Bayesian network for a skill as an affordance. We propose a soft Behavior Motivation (BM)switch as a method for ordering affordances to accomplish a task. Then, a skill is constructed as a combination of an affordance and a soft BM switch. To demonstrate the validity of our proposed method, some experiments were performed with GENIBO (Pet robot) performing a task using skills of Search-a-target-object, Approach-a-target-object, Push-up-in front of -a-target-object..

Key Words : Probabilistic Affordance Models, soft Behavior Motivation switch, Bayesian network, Action Selection Mechanism

1. 서 론

지능 로봇은 임무를 보다 수월하게 완수하기 위한 능력인 기술(skill)을 학습할 수 있어야 하며, 이를 합리적으로 선택하여 임무에 적용해야 한다. 이렇게 학습된 기술은 재사용 가능해야 하며 수정이 용이해야 한다. 우리는 기술을 학습하기 위해 베이지안 네트워크로 표현된 확률모델을 하나의 기술로 나타내고 이를 학습하는 방법을 제안한다. 여기서 하나의 기술에 대한 베이지안 네트워크는 행동유발성(affordance)으로 나타낼 수 있는데, 이는 로봇의 모터와 감지능력을 통한 지능로봇과 임무를 수행하는 환경사이의 연결 관계로써 표현될 수 있다. 베이지안 네트워크를 생성하기 위해선 인지와 관련된 변수, 기초 행동 그리고 행동 결과로 구성된 임무를 완수하는데 필요한 변수들이 선택되어야 한다. 또한 지능로봇은 임무를 완수하기 위해 상황 적응적 행동의 선택이 가능해야 할 것이다. 하지만 대부분의 임무들은 시간에 따른 행동선택이 요구되며, 임의로 배열된 행동 유발성으로는 임무를 완수하기 어렵다. 이를 해결하기 위해 로봇은 확률 또는 결정론적 모델을 학습해야하는데, 우리는 소프트 행동-동기 스위치(soft Behavior-Motivation switch)를 이용하여 새로운 모델의 학습 없이 행동 유발성에 순위를 부여한다. 우리는 행동 유발성과 소프트 행동-동기 스위치의 조합을 기술로 정의하였으며, 이를 행동 선택 메커니즘에 적용하여 때 다양한 임무에 대하여 시간에 따른 행동선택이 가능했으며 이를 실

험을 통하여 증명하였다.

2. 베이지안 행동 유발성 모델의 학습

행동 유발성은 모터와 감지능력을 통한 로봇과 환경과의 관계로 정의된다. 행동 유발성을 생성하기 위해 우리는 인지, 기초 행동 그리고 행동의 결과들의 연관관계를 베이지안 네트워크를 이용하여 표현했다.

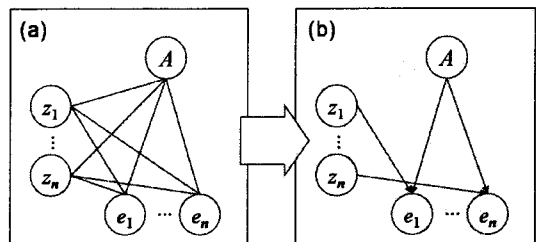


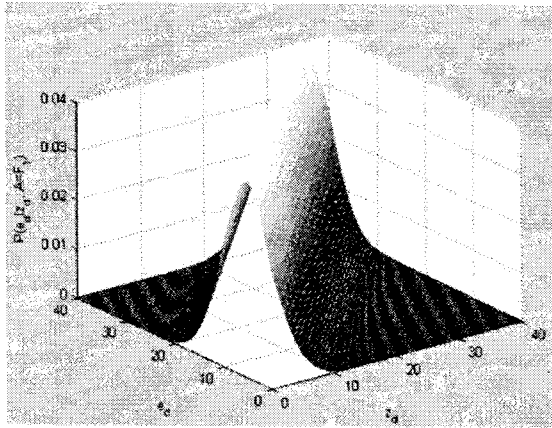
그림 1. 베이지안 네트워크의 예제 (a) 완전 결합 확률(full joint distribution)을 위한 베이지안 네트워크 (b) 행동 유발성을 위한 베이지안 네트워크

제안된 베이지안 네트워크의 구조는 두 가지 가정을 기반으로 한다. 첫째, 다수의 훈련 데이터를 수집하여 네트워크의 구조를 이미 알고 있다. 둘째, 인지, 기초행동, 그리고 행동 결과에 대한 변수들을 조건부 독립이다. 이는 행동 결과와 대한 변수 e_i 는 오직 인지와 관련된 변수 z_i 와 기초 행동에 대한 변수 A 에 의해서만 독립이다. [그림 1 (a)]에 완전 결합 분포는

$$P(z_1, \dots, z_n, A, e_1, \dots, e_n | \delta, \pi), \quad (1)$$

저자 소개

- * 손 광 희 : 한양대학교 전자컴퓨터통신공학 석사과정
- ** 이 상 형 : 한양대학교 전자컴퓨터통신공학 박사과정
- ***서 일 홍 : 한양대학교 전자컴퓨터통신공학 교수



[그림 2. 선형 가우시안 분포의 예]

와 같이 정의 된다. z_i 은 인지, A 는 기초 행동, e_i 은 행동결과에 대한 변수들을 나타내며, δ 는 훈련 데이터의 집합, 그리고 π 는 선행적 지식에 대한 변수들을 나타낸다. 완전 결합 분포는 행동 유발성을 나타내는 [그림 2 (b)]와 같은 베이지안 네트워크로 간략화 될 수 있으며,

$$P(Z, A, E|\delta, \pi) = P(Z|\delta, \pi) \cdot P(A|Z, \delta, \pi) \cdot P(E|A, Z, \delta, \pi) \quad (2)$$

$$= P(Z|\delta, \pi) \cdot P(A|\delta, \pi) \cdot P(E|A, Z, \delta, \pi),$$

$Z = \{z_1, \dots, z_n\}$ 와 $E = \{e_1, \dots, e_n\}$ 로 정의된다. 완전 결합 분포를 [그림 2 (b)]를 이용하여 계산하기 위해선 각각의 확률 모델이 학습되어야 한다. 우리의 베이지안 네트워크의 확률 분포는 선형 분포로 나타내기 충분하며, 이를 위해 선형 가우시안 분포[그림 2]를 사용하였다.

3. 소프트 행동-동기 스위치(Soft Behavior-Motivation Switch) 를 이용한 기술의 생성과 베이지안 행동 유발성 모델을 이용한 행동선택

앞에서 우리는 행동 유발성 모델의 생성에 관해 살펴보았다. 로봇이 임무를 완수하기 위해 시간에 따른 상황 적응적 행동을 임의로 배열된 행동 유발성 하에서 선택하기 어렵기 때문에 로봇은 순차적인 행동을 수행하기 위한 모델을 학습해야 한다. 이를 위해 우리는 소프트 행동-동기 스위치[그림 3]를 사용한다. 또한 행동 유발성을 순차적으로 정렬하기 위해 소프트 행동-동기 스위치와 행동 유발성을 하나의 쌍으로 구성[그림 3]하였다.

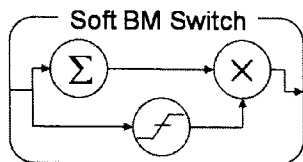


그림 3. 소프트 행동-동기 스위치 (soft Behavior Motivation switch)

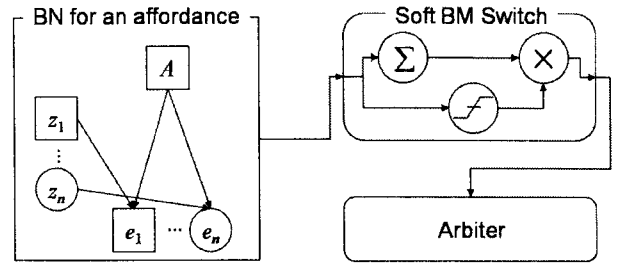


그림 4. 행동 유발성과 행동-동기 스위치의 쌍

로봇은 임무의 완수를 위하여 기술을 학습 및 수행이 합리적인 필요가 있다. 이를 위해 우리는 소프트 행동-동기 네트워크를 생성하는 방법을 제안한다. 로봇은 소프트 행동-동기 네트워크[그림 5]를 기반으로 현재 인지한 정보를 사용하여 각 스킬의 기초 행동에 대한 평가를 2-튜플(2-tuple)을 계산하며 수행하며 2-튜플은 아래와 같이 정의된다.

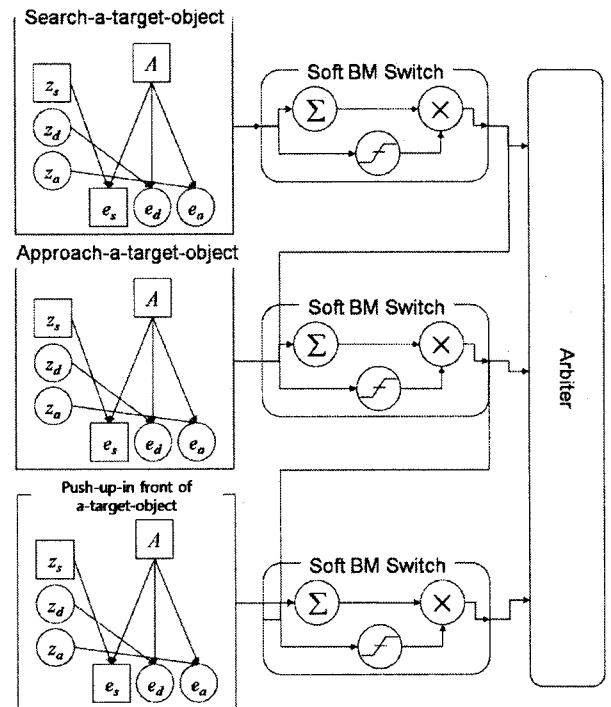


그림 5. 단일 임무에 대한 소프트 행동-동기 네트워크의 예제

$$T_i = [v_i, a_i^*],$$

$$\begin{cases} v_i = \max_{P_{ij}(a_j)} \{(\sum_{k=1}^{i-1} v_k + P_{ij}(a_j)) \cdot P_{ij}(a_j)\} \\ a_i^* = \arg \max_{a_j} \{(\sum_{k=1}^{i-1} v_k + P_{ij}(a_j)) \cdot P_{ij}(a_j)\}, \end{cases} \quad (3)$$

v_i 는 i^{th} 기술에 대한 i^{th} 기초 행동을 사용한 확률 분포의 최대값을 나타내며, a_i^* 는 i^{th} 기술에서 값을 최대로 하기 위한 i^{th} 의 행동을 나타낸다. 각 행동 유발성을 이용한 각 행동의 확률은 아래와 같이 구해질 수 있다.

$$P(A = a_i | z_1, \dots, z_n, \delta, \pi) \quad (4)$$

2-튜플을 기반으로 임무를 실행하기 위한 행동이 조정기 (arbiter)에 의해 2-튜플에 대한 v_i 를 정규화 하며 이 값을 이용하여 행동을 선택한다.

조정기는 선택된 $v_{i,N}$ 에 대응하는 2-튜플 T_i 를 선택할 수 있으며,

$$A = \Phi(T_i), \quad (5)$$

함수 Φ 는 선택된 2-튜플로부터 행동을 선택한다. 로봇은 행동 유발성과 소프트 행동-동기 스위치의 쌍으로 구성된 기술을 이용하여 임무에 적합한 행동을 선택한다.

또한 로봇은, 현재 상황에 따라 지식의 재구성 과 증가를 필요로 한다.

4. 실험

우리가 제안한 방법을 4족 로봇인 제니보(GENIBO)를 적용했으며, 탐색, 접근, 팔굽혀펴기를 기술을 이용하여 목표 물체 앞에서 팔굽혀 펴기를 하는 임무를 실험적으로 증명하였다.



그림 6. 행동 유발성을 학습하기 위한 학습 데이터의 수집

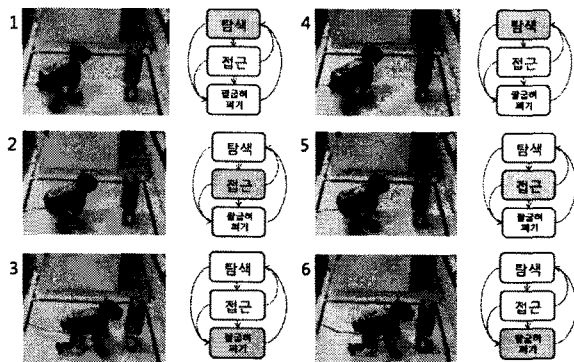


그림 7. 상태들간의 이동이 자유로운 유한상태기계(fully connected FSM)

학습 데이터를 수집[그림.6]하여 행동 유발성의 확률 모델이 계산되며, 이 확률 모델은 소프트 행동-동기 스위치의 쌍으로 결합되어 스킬을 구성한다. 탐색 기술에 대한 행동 유발성은 목표 물체를 찾기 전후 상황과 로봇의 좌우 방향전환으로 구성된 기본행동 간의 관계로 나타낼 수 있으며, 접근하는 기술은 목표 물체에 접근 하기 전후 상황, 전진하는 기초 행동 간의 연결관계, 그리고 팔굽혀 펴기 기술은 목표 물체

앞에서 팔굽혀 펴기를 수행하기 전후 상황, 팔굽혀 펴기로 연결관계로 행동 유발성을 나타낼 수 있다.

5. 결 론

우리는 베이저안 행동 유발성을 이용한 동기 기반 행동선택 메커니즘을 제안하였으며 이를 이용하여 불확실하고 변화하는 동적인 실제 환경에서 임무를 수행했을 때[그림 7] 로봇이 완전히 결합된 유한상태기계와 동일한 역할을 수행할 수 있었다.

6. Acknowledgement

본 연구는 지식경제부 및 정보통신연구진흥원의 IT 핵심 기술개발 사업의 일환으로 수행하였음. [2008-F-038-01, 상황 적응 로봇 인지 기술 개발].

참 고 문 헌

- [1] S. Russell and P. Norvig, Artificial Intelligence: A Modern Approach, Prentice Hall, 2003, ch. 14.
- [2] O. Leibel, P. Bessiere, J. Diard, and E. Mazer, "Bayesian Robot Programming," Autonomous Robots, pp.49-79, 2004.
- [3] A. Dearden and Y. Demiris, "Learning Forward Models for Robots," in Proc. of the International Joint Conference on Artificial Intelligence, Edinburgh, pp.1440-1445, 2005
- [4] E. Rome, J. Hertzberg, G. Dorffner, and P. Doherty, "Towards Affordance-based Robot Control," in Proc. of Dagstuhl Seminar 06231, June 5-9, 2006.