

논문 2009-46SC-4-2

# 베이저안 행동유발성 모델을 이용한 행동동기 기반 행동 선택 메커니즘

(Behavioral motivation-based Action Selection Mechanism  
with Bayesian Affordance Models)

이 상 형\*, 서 일 홍\*\*

(Sang Hyoung Lee and Il Hong Suh)

## 요 약

로봇이 지능적이고 합리적으로 임무를 수행하기 위해서는 다양한 솜씨(skill)가 필요하다. 우리는 솜씨를 생성하기 위해 우선 행동유발성(affordance)을 학습한다. 행동유발성은 행동을 유발하게 하는 물체 또는 환경의 성질로써 솜씨를 생성하는데 유용하게 사용될 수 있다. 로봇이 수행하는 대부분의 임무는 순차적이고 목표 지향적인 행동을 필요로 한다. 그러나 행동유발성만을 이용하여 이러한 임무를 수행하는 것은 쉽지 않다. 이를 위해 우리는 행동유발성과 목표 지향적 요소를 반영하기 위한 소프트 행동동기 스위치(soft behavioral motivation switch)를 이용하여 솜씨를 생성한다. 솜씨는 현재 인지된 정보와 목표 지향적 요소를 결합하여 행동동기를 생성한다. 여기서 행동동기는 목표 지향적인 행동을 활성화시키기 위한 내부 상태를 말한다. 또한, 로봇은 임무 수행을 위해 순차적인 행동 선택을 필요로 한다. 우리는 목표 지향적이고 순차적인 행동 선택이 가능하도록 솜씨를 이용하여 솜씨 네트워크(skill network)를 생성한다. 로봇은 솜씨 네트워크를 이용하여 목표 지향적이고 순차적인 행동을 선택할 수 있다. 본 논문에서는 베이저안 네트워크를 이용한 행동유발성 모델링 및 학습 방법, 행동유발성과 소프트 행동동기 스위치를 이용한 솜씨 및 솜씨 네트워크 생성 방법, 마지막으로 솜씨 네트워크를 이용한 목표 지향적 행동 선택 방법을 제안한다. 우리의 방법을 증명하기 위해 제니보(애완 로봇)를 이용한 교시 기반 학습 방법을 통해 “물체 찾기”, “물체에 접근하기”, “물체의 냄새 맡기”, 그리고 “물체를 발로 차기” 행동유발성들을 학습하였다. 또한, 이들을 이용하여 솜씨 및 솜씨 네트워크를 생성하여 제니보에 적용하고 실험하였다.

## Abstract

A robot must be able to generate various skills to achieve given tasks intelligently and reasonably. The robot must first learn affordances to generate the skills. An affordance is defined as qualities of objects or environments that induce actions. Affordances can be usefully used to generate skills. Most tasks require sequential and goal-oriented behaviors. However, it is usually difficult to accomplish such tasks with affordances alone. To accomplish such tasks, a skill is constructed with an affordance and a soft behavioral motivation switch for reflecting goal-oriented elements. A skill calculates a behavioral motivation as a combination of both presently perceived information and goal-oriented elements. Here, a behavioral motivation is the internal condition that activates a goal-oriented behavior. In addition, a robot must be able to execute sequential behaviors. We construct skill networks by using generated skills that make action selection feasible to accomplish a task. A robot can select sequential and a goal-oriented behaviors using the skill network. For this, we will first propose a method for modeling and learning Bayesian networks that are used to generate affordances. To select sequential and goal-oriented behaviors, we construct skills using affordances and soft behavioral motivation switches. We also propose a method to generate the skill networks using the skills to execute given tasks. Finally, we will propose action-selection-mechanism to select sequential and goal-oriented behaviors using the skill network. To demonstrate the validity of our proposed methods, “Searching-for-a-target-object”, “Approaching-a-target-object”, “Sniffing-a-target-object”, and “Kicking-a-target-object” affordances have been learned with GENIBO(pet robot) based on the human teaching method. Some experiments have also been performed with GENIBO using the skills and the skill networks.

**Keywords :** Action-Selection-Mechanism, Bayesian network, Affordance, Skill, Behavioral motivation

\* 학생회원, \*\* 정회원, 한양대학교 컴퓨터공학부

(Department of Computer Science & Engineering, Hanyang University)

※ 본 논문은 지식경제부 기술혁신사업으로 지원된 연구임. 논문의 모든 교신은 ihsuh@hanyang.ac.kr로 하여야 함.

접수일자 : 2009년6월26일, 수정완료일: 2009년7월20일

## I. 서론

사람은 지능적이고 합리적으로 임무를 수행하기 위해 다양한 솜씨(skill)를 필요로 한다. 솜씨는 학습과 연습을 통해 주어진 임무를 잘 수행할 수 있도록 하는 특별한 능력을 말한다<sup>[1]</sup>. 로봇도 지능적이고 합리적으로 임무를 수행하기 위해서는 다양한 솜씨가 필요하다. 또한, 솜씨는 쉽게 재사용되거나 수정될 수 있어야 한다. 우리는 이러한 솜씨를 생성하기 위해 우선 행동유발성(affordance)을 학습한다. 행동유발성은 에이전트에게 행동을 유발하게 하는 물체 또는 환경의 성질을 말한다<sup>[2]</sup>. 행동유발성은 에이전트가 지니고 있는 인지와 행동 능력을 통해 에이전트 자신과 환경 사이의 관계를 표현한 정보를 의미한다<sup>[3]</sup>. 이러한 행동유발성은 로봇이 솜씨를 생성하는데 유용하게 사용될 수 있고 재사용하는 것이 용이하다<sup>[4]</sup>. 그러나 행동유발성은 임무의 목적과 합쳐졌을 때 의미 있고 효율적으로 사용될 수 있다<sup>[5]</sup>. 로봇이 수행하는 대부분의 임무는 순차적이고 목표 지향적인 행동을 필요로 한다. 그러나 행동유발성만을 이용하여 목표 지향적인 행동이 필요한 임무를 수행하는 것은 쉽지 않다. 임무 수행을 위해서는 행동유발성에 임무의 목적이 반영되도록 모델링 되어야 한다. 이를 위해 우리는 솜씨를 모델링한다. 솜씨는 현재 인지된 정보와 목표 지향적 요소들을 이용하여 행동동기를 생성한다. 여기서, 행동동기는 목표 지향적인 행동을 활성화시키기 위한 내부 상태로 정의된다<sup>[6]</sup>. 더욱이 로봇이 수행하는 대부분의 임무는 순차적인 행동을 필요로 한다. 우리는 이를 위해 솜씨를 이용한 솜씨 네트워크(skill network)를 생성한다. 로봇은 이를 기반으로 목표 지향적이고 순차적인 행동을 선택할 수 있다. 솜씨를 생성하기 위해 우선 베이지안 네트워크를 기반으로 행동유발성을 모델링하고 학습한다. 베이지안 네트워크를 이용하여 행동유발성을 모델링하고 학습하는 이유는 다음과 같다. 베이지안 네트워크는 각 노드에 확실적인 정보량을 할당한 방향성이 존재하는 그래프를 말한다<sup>[7]</sup>. 베이지안 네트워크에서 토폴로지 구조와 조건적 확률 분포를 생성하면 모든 변수에 대한 완전 결합 확률 분포를 구하는 것이 용이하다<sup>[8]</sup>. 완전 결합 확률 분포를 알게 되면 몇몇의 데이터가 누락된 상황에서도 확률 분포를 구할 수 있다. 더욱이 이것은 인과 관계를 배우는데 사용될 수 있고, 불확실하거나 부정확한 정보 하에서도 문제를 다룰 수 있다<sup>[9]</sup>. 행동유발성을 표현하기 위

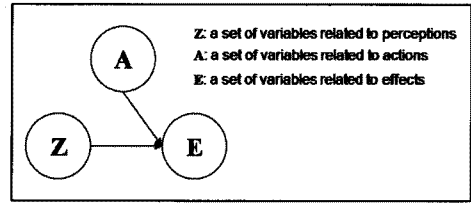


그림 1. 행동유발성을 위한 베이지안 네트워크의 토폴로지 그래프

Fig. 1. A topology graph of a Bayesian network for representing an affordance.

한 베이지안 네트워크의 토폴로지 구조는 그림 1과 같이 인지(perception), 행동(action), 그리고 행동 결과(action effect) 변수들의 연관관계로 구성된다. 여기서 행동 결과는 현재 인지된 정보와 행동에만 영향을 받는다고 가정한다. 우리는 행동유발성과 목표 지향적 요소를 반영하기 위한 소프트 행동동기 스위치(soft behavioral motivation switch)를 이용하여 솜씨를 생성한다. 솜씨는 현재 인지된 정보와 목표 지향적 요소를 이용하여 행동동기를 계산한다. 그리고 이러한 솜씨들을 이용하여 순차적 행동이 가능하도록 솜씨 네트워크를 생성한다. 이렇게 생성된 솜씨 네트워크는 변화하는 환경, 부정확한 인지, 확실적인 행동 결과 하에서도 순차적이고 목표 지향적인 행동이 필요한 다양한 임무에 사용될 수 있다. 또한, 생성된 솜씨들의 다양한 조합(추가/삭제/변경)을 통해 다양한 솜씨 네트워크를 생성할 수 있다. 더욱이 하나의 솜씨는 다른 솜씨에 전혀 영향을 받지 않기 때문에 수정 및 추가가 용이하다.

본 논문은 다음과 같이 구성된다. II장에서 관련 연구들을 소개한다. III장에서는 베이지안 네트워크 기반의 행동유발성 표현 및 학습 방법, 행동유발성과 소프트 행동동기 스위치를 이용한 행동동기 생성 방법, 행동동기를 이용한 솜씨 생성 방법, 그리고 솜씨를 이용한 행동선택 방법을 설명한다. IV장에서는 제니보<sup>[10]</sup>를 이용한 교시 기반 학습 방법을 통해 “물체 찾기”, “물체에 다가가기”, “물체의 냄새 맡기”, 그리고 “물체를 발로 차기” 행동유발성들을 생성하는 과정을 설명한다. 또한, 이를 이용하여 “물체에 다가가서 냄새 맡기”, “물체에 다가가서 발로 차기” 솜씨를 생성하고 제니보에 적용한 실험을 설명한다. 마지막으로 V장에서는 결론과 향후 계획을 설명한다.

## II. 관련 연구

지능 로봇은 불확실한 환경, 부정확한 인지, 그리고 확률적인 행동 결과 하에서 합리적인 행동 선택을 하기 위한 솜씨를 필요로 한다. 더욱이 로봇은 학습된 솜씨를 쉽고 빠르게 재사용 또는 수정할 수 있어야 한다.

Lee 등은 임무를 프로그래밍하는 방법을 통해 완전 결합된 유한 상태 머신과 같이 행동을 수행할 수 있는 행동 선택 메커니즘을 제안하였다<sup>[11]</sup>. 더욱이 이 방법은 강화학습을 통해 임무 수행을 위한 행동동기들을 쉽게 추가/삭제할 수 있도록 하였다. 그러나 이들은 인지와 행동의 결과에 대한 불확실성을 고려하지 않았다. 그리고 이 방법은 인지와 행동 사이의 연관관계를 미리 프로그래밍 하였다. Sevin 등은 가상 휴먼이 가상 환경에서 자율적으로 행동하기 위한 동기 기반 행동 선택 모델을 제안하였다<sup>[12]</sup>. 이 방법은 계층적으로 구성된 시스템의 동기를 계산하는 과정에서 각 계층이 평행하게 계산되도록 하였다. 로봇은 동기를 기반으로 행동을 선택하도록 하였다. Beaudry 등은 행동 선택을 위해 각 모듈의 개별적인 동기에 따라 행동 생성 모듈을 선택하는 구조적인 방법론을 제안하였다<sup>[13]</sup>. 그러나 이 방법들 모두 인지와 행동의 연관관계를 사람이 생성하였고, 인지와 행동의 결과에 대한 불확실성을 고려하지 않았다.

Lazkano 등은 이동 로봇에서 임무 수행을 위한 학습 방법으로 베이지안 네트워크를 사용하는 방법을 제안하였다<sup>[14]</sup>. 이 연구에서 베이지안 네트워크가 불확실성과 다양성을 다루는데 가치 있는 학습 방법이라는 것을 증명하였다. Dearden 등은 로봇의 모터와 주변 환경에 대한 행동 결과를 예측하는 모델을 제안하였다<sup>[15]</sup>. 이 연구는 베이지안 네트워크를 이용해 행동을 학습하고 표현하였다. 학습된 모델은 사람의 행동을 모방하기 위해 사용되었다. 위의 연구들은 오직 행동 명령과 관측된 로봇의 상태 사이의 인과 관계를 학습하였다. 이 방법들은 임무 지향적 행동을 필요로 하는 복잡한 임무를 수행하는 것에 대해 다루지 않았다. 다시 말해, 이 연구들은 임무 지향적 행동이 필요한 임무 수행을 위한 모델을 고려하지 않았다. Lebeltel 등은 베이지안 추론과 베이지안 로봇 프로그래밍 기반의 로봇 프로그래밍 방법을 제안하였다<sup>[16]</sup>. 베이지안 로봇 프로그래밍은 불확실성이 존재하는 환경에서 로봇 프로그래밍을 하기 적합한 단순하고 일반적인 프레임워크이다. 이들은 단순한 확률 분포의 곱으로 완전 연결 확률 분포를 단순화

하였다. 그러나 이 방법은 임무 지향적 행동이 필요한 임무 수행이 어렵고, 이를 위해서는 새로운 확률 모델을 학습해야 한다. 임무 지향적인 행동을 수행하기 위한 확률 모델을 학습하는 것은 쉬운 일이 아니다.

Montesano 등은 베이지안 네트워크 기반의 행동유발성 모델을 제안하였고, 불연속적인 변수들을 사용하여 특정한 상황에 대해 표현하였다<sup>[17]</sup>. 그러나 이들은 임무 지향적 행동이 필요한 복잡한 임무 수행을 위한 행동유발성들 사이의 전위들에 대해서는 고려하지 않았다. Raubal 등은 물리적, 사회 제도적, 그리고 정신적 행동 유발성과 로봇 제어 시스템을 위한 계층적 행동유발성 기반 표현 방법을 이용한 로봇 구조를 제안하였다<sup>[18]</sup>. 그러나 이들은 이 개념을 증명하기 위한 로봇 적용 실험을 하지 않았고, 인지와 행동의 결과에 대한 불확실성을 고려하지 않았다. Stoychev는 로봇이 도구를 사용하기 위한 행동유발성 학습에 대한 새로운 방법을 제안하였다<sup>[19]</sup>. 여기서 행동유발성은 인지, 로봇의 행동, 행동의 결과에 대한 조합으로 표현되었다. 그러나 이들은 복잡한 임무 수행을 위한 순차적 행동 선택과 인지 및 행동 결과의 불확실성을 고려하지 않았다.

## III. 베이지안 행동유발성 모델을 이용한 동기 기반 행동 선택 메커니즘

임무 수행을 위한 순차적이고 목표 지향적인 행동을 선택하기 위해 베이지안 네트워크 기반 행동유발성 표현 및 학습 방법, 행동유발성과 소프트 행동동기 스위치를 이용한 솜씨 생성 방법, 솜씨들을 이용한 솜씨 네트워크 생성 방법, 그리고 솜씨 네트워크를 이용한 행동 선택 방법을 제안한다. 이 방법은 최소한 네 가지의 장점을 가진다. 첫째, 이 방법은 솜씨들 사이의 전이가 완전 결합된 유한 상태 머신처럼 일어날 수 있다. 때문에 상황의 변화에 따라 적합한 행동을 선택하는 것이 가능하다. 둘째, 이것은 각 솜씨를 구성하는데 필요한 행동유발성들을 독립적인 베이지안 네트워크로 표현하기 때문에 단순한 확률 분포로 표현할 수 있고 수정하는 것이 용이하다. 셋째, 솜씨들이 독립적으로 생성되기 때문에 추가/삭제 또는 점진적 학습 시 다른 솜씨에 영향을 주지 않는다. 마지막으로 학습된 행동유발성과 솜씨들은 쉽게 재사용될 수 있다.

### 1. 베이지안 행동유발성 모델 학습

앞에서 언급한 바와 같이 행동유발성은 에이전트가 게 행동을 유발하게 하는 물체 또는 환경의 성질을 말한다. 행동유발성은 에이전트가 지니고 있는 인지와 행동 능력을 통해 에이전트 자신과 환경 사이의 관계에 대한 정보이다. 행동유발성을 생성하기 위해 우리는 인지, 행동, 행동 결과에 대한 변수들을 이용한 베이지안 네트워크를 학습한다. 베이지안 네트워크를 사용하는 이유는 다음과 같다. 완전 연결 확률 분포가 주어지면 추론을 사용하여 언제든지 임의의 질문에 대한 답을 계산할 수 있다<sup>[6]</sup>. 그러나 완전 연결 확률 분포를 계산하는 것은 쉽지 않다. 우리는 단순화된 확률 분포의 곱을 이용하여 완전 연결 확률 분포를 계산하기 위해 베이지안 네트워크를 사용한다. 베이지안 네트워크의 토폴로지 구조를 생성하기 위해 두 가지 가정을 하였다. 첫째, 인지, 행동, 그리고 행동 결과에 관한 모든 변수들은 서로 독립적이다. 둘째, 행동 결과에 대한 변수  $e_i$ 는 인지 에 대한 변수  $z_i$ 와 행동에 대한 변수  $A$ 에만 의존한다. 완전 결합 확률 분포의 식은 다음과 같다.

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

여기서  $z_i$ 는 인지와 관련된 변수,  $A$ 는 행동과 관련된 변수,  $e_i$ 는 행동 결과와 관련된 변수,  $\delta$ 는 학습 데이터, 그리고  $\pi$ 는 사전 지식을 나타낸다. 행동유발성의 완전 결합 확률 분포는 가정에 의해 그림 2와 같이 구성된다. 그리고 베이지안 네트워크에 의해 완전 결합 확률 분포 식은 다음과 같이 단순화 할 수 있다.

$$\begin{aligned} P(Z, A, E | \delta, \pi) &= P(Z | \delta, \pi) \cdot P(A | Z, \delta, \pi) \\ &\quad \cdot P(E | A, Z, \delta, \pi) \\ &= P(Z | \delta, \pi) \cdot P(A | \delta, \pi) \cdot P(E | A, Z, \delta, \pi), \end{aligned} \tag{2}$$

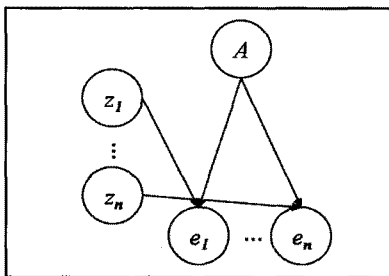


그림 2. 행동유발성을 위한 베이지안 네트워크의 토폴로지 구조  
Fig. 2. A topology structure of a Bayesian network for an affordance.

여기서  $Z = \{z_1, \dots, z_n\}$ 는 인지와 관련된 변수들의 집합이고,  $E = \{e_1, \dots, e_n\}$ 는 행동 결과와 관련된 변수들의 집합을 나타낸다. 식 (2)를 이용하여 완전 결합 확률 분포를 계산하기 위해서는 센서 모델에 대한 확률 분포  $P(Z | \delta, \pi)$ , 행동 모델에 대한 확률 분포  $P(A | \delta, \pi)$ , 그리고 인지와 행동이 주어진 상황에서 행동 결과에 대한 확률 분포  $P(E | A, Z, \delta, \pi)$ 를 학습해야 한다. 우선 각 행동유발성을 학습시키기 위한 학습 데이터를 수집한다. 우리의 행동유발성 학습을 위한 베이지안 네트워크는 그림 3과 같이 하이브리드(hybrid) 베이지안 네트워크로 정의한다. 행동과 관련된 변수  $A$ 는 불연속적인 변수로 정의한다. 그리고 인지와 관련된 변수들의 집합인  $Z$ 와 행동 결과와 관련된 변수들의 집합인  $E$ 는 불연속적 변수와 연속적 변수들로 구성된다. 행동유발성의 베이지안 네트워크의 확률 분포를 학습하는 과정에서 불연속적 변수들로 구성된 확률 분포에 대해서는 조건 확률 테이블(conditional probability table)을 학습하고, 불연속적 변수와 연속적인 변수들로 구성된 확률 분포에 대해서는 선형 확률 분포를 사용하여 학습시킨다. 선형 확률 분포 중 가장 자주 사용되는 것은 선형 가우시안 확률 분포이다<sup>[7]</sup>. 선형 가우시안 확률 분포는 부모 노드의 값에 의해 평균  $\mu$ 과 표준 편차  $\sigma$ 가 선형으로 변화하는 분포를 말한다. 선형 가우시안 확률 분포는 다음과 같이 정의된다.

$$\begin{aligned} P(e_i | z_i, A = a_i, \delta, \pi) &= N(a \cdot z_i + b, \sigma^2)(e_i) \\ &= \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{e_i - (a \cdot z_i + b)}{\sigma} \right)^2}, \end{aligned} \tag{3}$$

여기서  $e_i$ 는 행동 결과와 관련된 변수,  $z_i$ 는 인지와 관련된 변수,  $A$ 는 행동과 관련된 변수이다. 식(3)은 행동

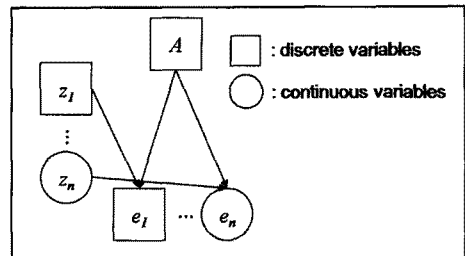


그림 3. 불연속적 변수와 연속적 변수로 구성된 하이브리드 베이지안 네트워크  
Fig. 3. A hybrid Bayesian network composed of both discrete and continuous variables.

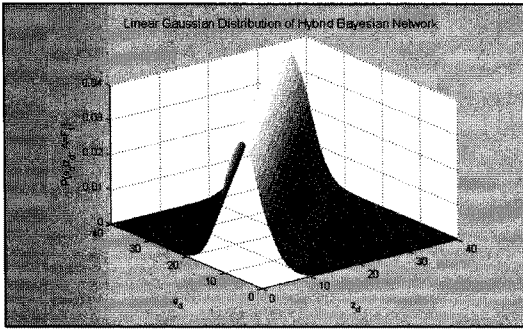


그림 4. 행동 관련 변수  $A = F_1$ (전진 10cm)인 경우 선형 가우시안 확률 분포의 예;  $a = 1, b = 10, \sigma = 10$ , 그리고  $c = 0$   
 Fig. 4. An example of linear Gaussian distribution given an action value as  $A = F_1$ (Forward 10cm);  $a = 1, b = 10, \sigma = 10$ , and  $c = 0$ .

과 관련된 변수  $A$ 에 대한 열거법에 의해 계산된다. 다시 말해, 행동과 관련된 변수  $A$ 의 모든 값  $a_i$ 에 대한 확률값을 계산할 수 있다. 여기서 식(3)의 각 파라미터들은 다음과 같이 정의된다.

$$a = 1, b = \frac{\sum_{i=1}^N (e_i - z_i)}{N}, \sigma = \frac{\sum_{i=1}^N ((e_i - z_i) - \mu)^2}{N} + c \quad (4)$$

여기서  $c$ 값은 확률 분포의 표준 편차를 조절하는데 사용된다. 그림 4는 식(3)과 식(4)를 이용하여 물체와 로봇 사이의 거리에 관련된 변수  $z_d$ , 행동 결과에 관련된 변수  $e_d$ , 행동과 관련된 변수  $A$ 의 값이  $F_1$ (전진 10cm)인 경우에 대한 선형 가우시안 확률 분포를 보여준다. 이렇게 학습된 행동유발성은 유사한 조건에서 재사용하는 것이 용이하다. 우리는 식(2), 식(3), 그리고 식(4)를 이용하여 베이지안 네트워크를 이용한 행동유발성의 완전 결합된 확률 분포를 학습한다.

## 2. 행동유발성 모델을 이용한 숨씨 생성

앞에서 언급한 바와 같이 행동유발성은 로봇이 임무 수행을 위한 숨씨를 학습하는데 유용하게 사용될 수 있다. 그러나 행동유발성은 임무의 목적과 합쳐졌을 때 의미 있고 효율적으로 사용될 수 있다. 로봇은 임무 수행을 위해 순차적이고 목표 지향적인 행동을 선택할 수 있어야 한다. 임무를 수행하기 위해서는 행동유발성에 임무의 목적이 반영되도록 모델링 되어야 한다. 행동동

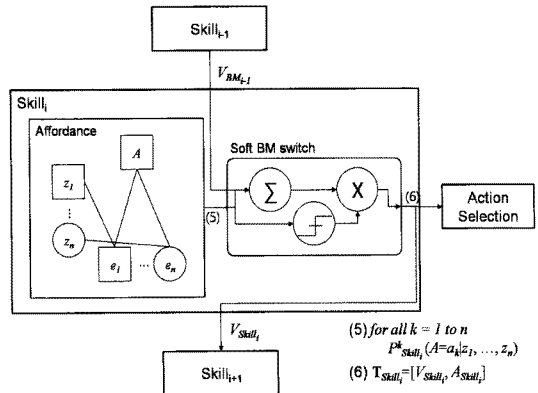


그림 5. 행동유발성과 소프트 행동동기 스위치로 구성된 행동동기 구조  
 Fig. 5. A structure of a behavioral motivation composed of an affordance and a soft BM switch.

기는 목표 지향적인 행동을 활성화시키는 내부 상태를 의미한다. 우리는 이러한 행동동기를 계산하기 위해 행동유발성과 목표 지향적 요소를 반영하기 위한 소프트 행동동기 스위치를 이용하여 숨씨를 생성한다. 숨씨는 그림 5와 같이 행동유발성과 소프트 행동동기 스위치에 의해 구성된다. 소프트 행동동기 스위치는 Lee 등이 제안한 동기 기반 행동 선택 구조<sup>[11]</sup>에서 영감을 얻어 정의하였다. 그러나 기존 방법과는 다르게 현재 인지된 센서 정보에 의해 상위 행동동기가 논리 스위치처럼 ON/OFF로 하위 행동동기에 영향을 주는 것이 아니라 현재 인지된 센서 정보에 의해 전달되는 양을 조절하여 하위 행동동기에 전달한다. 이는 행동동기가 행동 선택에 더 합리적으로 사용되도록 할 수 있도록 한다. 숨씨의 행동동기를 계산하기 위해 현재 들어온 센서 값을 기반으로 다음과 같이 행동유발성 값을 계산한다.

$$P_{Skill_i}^k(A = a_k | z_1, \dots, z_n) = \alpha \cdot \int_{e_1, \dots, e_n} P(z_1, \dots, z_n, A = a_k, e_1, \dots, e_n) \quad (5) \\ = \alpha \cdot P(A = a_k) \cdot P(z_1) \cdot \dots \cdot P(z_n) \\ \cdot \int_{e_1} P(e_1 | A = a_k, z_1) \cdot \dots \cdot \int_{e_n} P(e_n | A = a_k, z_n)$$

여기서  $z_i$ 는 인지와 관련된 변수들,  $A$ 는 행동과 관련된 변수,  $e_i$ 는 행동결과와 관련된 변수를 의미한다. 그리고 인덱스  $k$ 는 행동 변수의  $k$ 번째 행동 값의 인덱스를 의미하고,  $Skill_i$ 는  $i$ 번째 숨씨를 의미한다. 각 숨씨는 식(5)를 이용하여 각 숨씨의 행동유발성 값을 병렬

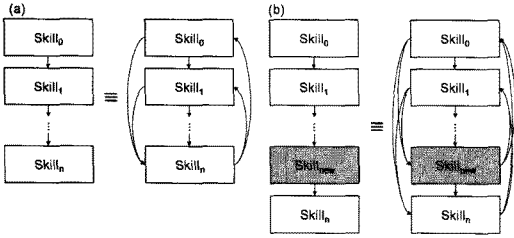


그림 6. 완전 결합된 유한 상태 머신처럼 동작되는 솜씨의 예. (a) 완전 결합된 유한 상태 머신과 같은 행동동기의 전이, (b) 새로운 행동동기와 기존 행동동기 사이의 관계 생성

Fig. 6. An example of a skill operating like a fully connected finite state machine. (a) transition among behavioral motivations like a fully connected finite state machine, (b) generation of relationships among exist behavioral motivation and an additional behavioral motivation.

적으로 구한 후에 목표 지향적 행동 선택을 위해 행동동기 값을 계산하고 각 솜씨에서 가장 적합한 행동을 추천한다. 이를 위해 우리는 2-튜플을 생성한다. 2-튜플은 다음과 같이 정의된다.

$$\mathbf{T}_i = [V_{Skill_i}, A_{Skill_i}], \quad (6)$$

$$\begin{cases} V_{Skill_i} = \max_{a_k} ((V_{Skill_{i-1}} + P_{Skill_i}^k(A = a_k | z_1, \dots, z_n)) \\ \quad \cdot P_{Skill_i}^k(A = a_k | z_1, \dots, z_n)) STEP(P_{Skill_i}^k), \\ A_{Skill_i} = \arg \max_{a_k} \\ \quad ((V_{Skill_{i-1}} + P_{Skill_i}^k(A = a_k | z_1, \dots, z_n)) \\ \quad \cdot P_{Skill_i}^k(A = a_k | z_1, \dots, z_n)) STEP(P_{Skill_i}^k), \end{cases}$$

여기서  $V_{Skill_i}$ 는  $i$ 번째 행동동기의 행동동기 값을 의미하고,  $A_{Skill_i}$ 는  $i$ 번째 행동동기에서 가장 적합한 행동 값을 의미한다. 그리고  $V_{Skill_{i-1}}$ 는 이전 행동동기 값,  $P_{Skill_i}^k(A = a_k | z_1, \dots, z_n)$ 는 식(5)에 의해 계산된 행동유발성 값,  $STEP(P_{Skill_i}^k)$ 는 행동유발성 값이 일정 값 이상일 때 값을 통과시키는 필터 역할을 수행한다. 이렇게 병렬적으로 계산된 각 솜씨의 행동동기는 행동을 선택하는데 사용된다.

### 3. 솜씨 네트워크 생성

앞에서 우리는 행동유발성과 소프트 행동동기 스위치를 이용하여 솜씨를 생성하는 방법을 설명하였다. 로봇이 수행하는 대부분의 임무들은 순차적인 행동이 필요하다. 때문에 솜씨 하나만 가지고 임무를 수행하는

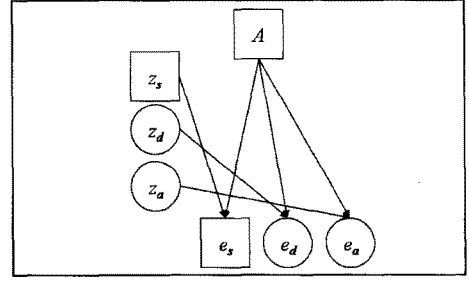


그림 7. 우리의 행동유발성들을 생성하기 위한 하이브리드 베이지안 네트워크

Fig. 7. A hybrid Bayesian network for generating our affordances.

것은 쉽지 않다. 때문에 임무를 수행하기 위해서는 솜씨들을 조합하여 순차적인 행동이 일어날 수 있도록 모델링 되어야 한다. 우리는 순차적 행동이 가능하도록 그림 5와 같이 솜씨들을 이용하여 솜씨 네트워크를 생성한다. 각 솜씨들은 현재 인지된 센서 값과 상위 행동동기 값을 기반으로 식(6)과 같이 2-튜플을 계산한다. 로봇은 이렇게 계산된 각 솜씨의 2-튜플값을 이용하여 다음과 같은 식을 이용하여 하나의 튜플을 선택한다.

$$\mathbf{T}^* = \text{Draw} \left( \frac{V_{\mathbf{T}_i}}{\sum_{i=1}^N V_{\mathbf{T}_i}} \right), \quad (7)$$

여기서 우리는 2-튜플을  $\text{Draw}(\cdot)$  함수를 이용하여 뽑는 방식을 사용하기 위해 전체  $N$ 개의 행동동기 값으로 행동동기 값들을 정규화(normalization)하여 계산하였다. 로봇은 목표 지향적 행동 선택을 위해 다음 식과 같이 선택된 2-튜플에 속한 행동을 수행한다.

$$A^* = \Phi(\mathbf{T}^*) \quad (8)$$

여기서  $\Phi(\cdot)$ 는 선택된 2-튜플에 속한 행동을 추출하는 단순한 함수이다. 솜씨 네트워크는 로봇이 임무수행을 위한 순차적이고 목표 지향적인 행동을 선택할 수 있도록 도와준다. 여기서 우리는 솜씨 네트워크를 구성하기 위한 솜씨들의 배치 순서는 사람이 결정해 주었다고 가정하였다. 생성된 솜씨 네트워크는 그림 6-(a)과 같이 완전 결합된 유한 상태 머신처럼 솜씨들 사이의 전이가 발생될 수 있다. 로봇은 솜씨 네트워크를 구성하고 있는 각 솜씨들을 병렬적으로 계산하고, 이를 기반으로 순차적이고 목표 지향적인 행동을 선택하기 때문에 완전 결합된 유한 상태 머신처럼 전이가 발생할

표 1. 행동유발성 학습을 위한 학습 데이터  
Table 1. Training data for learning our affordances.

종류	학습 데이터 수	사용 행동
물체 찾기	70(개)	몸통-좌/우(10도, 20도, ..., 40도, 50도)
물체 다가가기	70(개)	전진(10cm, 20cm, ..., 50cm)
물체 냄새 맡기	30(개)	머리-좌/우/상/하 3도, 6도, 9도)
발로 차기	30(개)	앞 다리-좌/우 차기

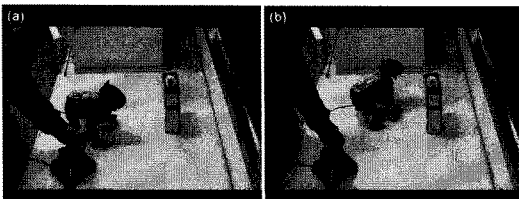


그림 8. 조이스틱을 사용하여 행동유발성 학습하기 위한 학습 데이터를 수집하는 예. (a) 물체 찾기, (b) 물체에 다가가기

Fig. 8. Examples of collecting training data for learning affordances using a joystick. (a) Searching-a-target-object, (b) Approaching-a-target-object.

수 있다. 또한, 이러한 특징 때문에 그림 6-(b)와 같이 숨씨를 추가하거나 삭제할 때 해당 숨씨와 관련된 숨씨들과의 연결만을 통해 숨씨 네트워크를 재구성하는 것이 가능하다. 이 외에도 각 숨씨들은 서로 독립적으로 생성되기 때문에 점진적 학습 및 수정 시에 다른 숨씨에 전혀 영향을 주지 않는다. 또한, 학습 데이터를 이용한 행동유발성의 수정만으로 숨씨를 수정할 수 있다.

#### IV. 실험

우리가 제안한 방법을 증명하기 위해 우리는 제니보를 이용하여 “물체 찾기”, “물체에 다가가기”, “물체 냄새 맡기”, 그리고 “물체를 발로 차기” 행동유발성을 학습하였다. 그리고 학습된 행동유발성을 이용하여 “물체에 다가가서 냄새 맡기”, “물체에 다가가서 발로 차기” 두 개의 숨씨를 생성하였다. 우리는 우선 행동유발성을 생성하기 위해 그림 7과 같이 베이지안 네트워크를 생성하였다. 우리가 생성한 네 가지의 행동유발성은 그림 7과 같이 동일한 베이지안 네트워크의 토폴로지 구조를

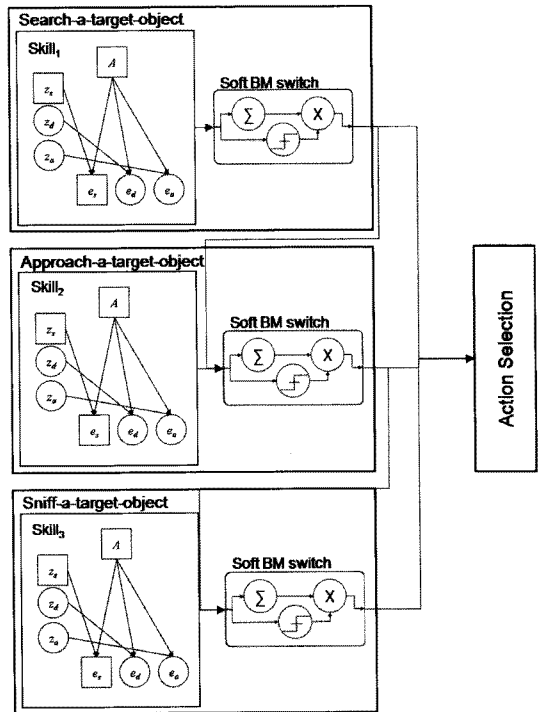


그림 9. “물체 찾기”, “물체에 다가가기”, “물체 냄새 맡기” 행동동기들을 이용한 “물체에 다가가서 냄새 맡기” 숨씨의 예

Fig. 9. An example of a skill using behavioral motivations of “Search-a-target-object”, “Approach-a-target-object”, and “Sniff-a-target-object”.

사용하였다. 여기서  $z_s$ 는 로봇의 시야에 목표하는 물체가 들어왔는지 여부를 나타내는 변수,  $z_d$ 는 로봇과 물체 사이의 상대적인 거리를 나타내는 변수,  $z_a$ 는 로봇과 물체 사이의 상대적인 각도를 나타내는 변수를 말한다. 각 행동유발성의 베이지안 네트워크의 파라미터를 학습하기 위해 우리는 표 1과 같은 데이터들을 수집하였다. 그림 8과 같이 조이스틱을 이용하여 제니보의 학습 데이터를 수집하였고, 이를 이용하여 베이지안 네트워크를 학습하였다.

학습된 행동유발성들과 소프트 행동동기 스위치들을 이용하여 “물체 찾기”, “물체에 다가가기”, “물체 냄새 맡기”, “물체를 발로 차기” 숨씨들을 생성하였다. 이렇게 생성된 숨씨들을 이용하여 임무 수행을 위해 순차적인 행동이 필요한 “물체에 다가가서 냄새 맡기”와 “물체에 다가가서 발로 차기” 숨씨 네트워크를 그림 9와 같이 생성하였다. 그림 9는 “물체에 다가가서 냄새 맡

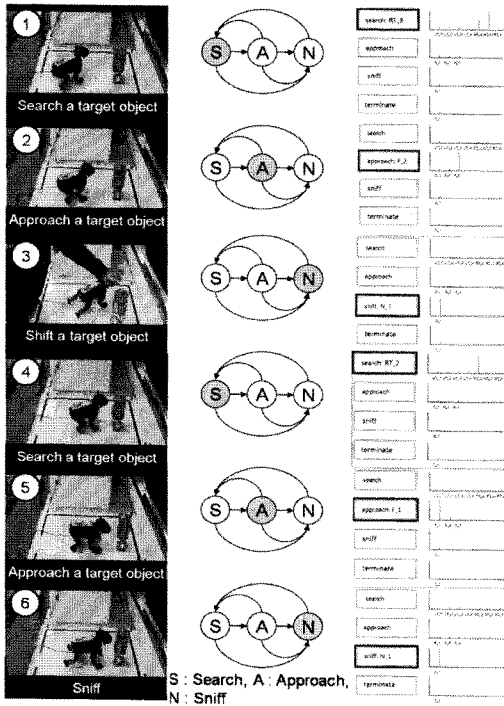


그림 10. “물체에 다가가서 냄새 맡기” 숨씨가 완전 결합된 유한 상태 머신처럼 동작되는 예  
 Fig. 10. An example of a skill operating like fully connected finite state machine.

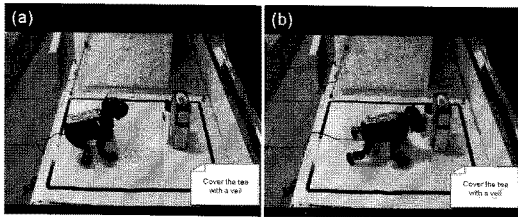


그림 11. 목표 물체가 장애물로 가려진 상황에서 임무 수행 예. (a) 장애물이 목표 물체를 가린 상황, (b) 냄새 맡는 행동을 수행하는 예  
 Fig. 11. An example of executing a task in the situation of partial occlusion of the target object. (a) The situation of partial occlusion of the target object, (b) An example of executing “Sniff-a-target-object”.

기” 숨씨 네트워크에 대한 예를 보여준다. “물체에 다가가서 발로 차기” 숨씨 네트워크는 “물체에 다가가서 냄새 맡기” 숨씨 네트워크에서 “물체 냄새 맡기” 숨씨를 “물체를 발로 차기” 숨씨로 교체하는 것만으로 생성될 수 있다. 그림 10은 제니보가 “물체에 다가가서 냄새 맡기” 숨씨 네트워크를 이용하여 임무를 수행하는 과정을

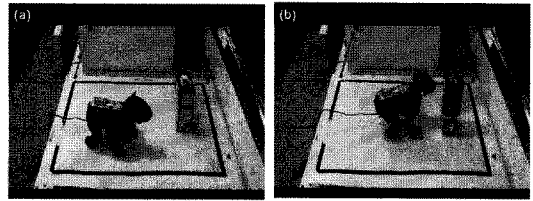


그림 12. 행동유발성이 일반적으로 사용되는 예. (a) “Strawberry tea” 상자, (b) “Butter waffle” 상자  
 Fig. 12. Examples using affordance generally. (a) “Strawberry tea” box, (b) “Butter waffle” box.

An algorithm for BM-based Soft ASM with Bayesian Affordance Models

1. Construct a Bayesian network for learning an affordance as shown in Fig. 2.
2. Collect some training data of each affordance as shown in Fig. 9.
3. Learn probabilistic models for calculating full joint distribution using training data.
4. Generate a skill with an affordance and a soft BM switch as shown in Fig. 6.
5. Establish a soft BM network for executing a task with soft BM switches of each skill as shown in Fig. 7.
6. Compute the 2-tuples of each skill with current perceived information and affordances using (5) and (6).
7. Draw a skill based on the 2-tuples of each skill using (7).
8. Extract a behavior for achieving a task based on a selected skill using (8).

그림 13. 베이지안 행동유발성 모델을 이용한 행동 동기의 행동 선택 방법의 알고리즘  
 Fig. 13. An algorithm of the behavioral motivation-based action selection mechanism with Bayesian affordance models.

보여준다. 제니보는 그림 10에서처럼 로봇 또는 물체의 위치를 변경하는 경우에도 95% 이상의 성공률로 임무를 완수하였다. 또한 그림 11에서처럼 목표 물체가 장애물에 의해 가려진 경우에도 임무를 완수하는 것을 확인할 수 있었다. 그림 11에서는 목표 물체를 장애물을 이용하여 부분적으로 가려놓은 상태에서 임무를 수행하는 과정을 보여준다. 그림 12는 하나의 물체에 대해 학습된 행동유발성을 이용하여 학습하지 않은 물체에도 사용될 수 있음을 보여주는 예이다. 로봇은 “Strawberry tea” 상자에 대해 행동유발성 모델을 학습하였으나 “Butter waffle” 상자에 대해서도 동일한 행동을 유발하는 것을 확인할 수 있었다. 그림 13은 우리가 제안한 방법의 전체 흐름도를 보여준다.

V. 결론

우리는 로봇의 임무 수행을 위해 베이지안 네트워크 기반의 행동유발성 모델링 및 학습 방법, 임무 지향적



행동 선택을 위해 행동유발성을 사용한 숨씨 생성 방법, 순차적 행동을 수행하기 위한 숨씨를 사용한 숨씨 네트워크 생성 방법, 숨씨 네트워크를 이용한 행동 선택 방법들을 제안하였다. 이 방법은 현재 들어오는 센서 값에 따라 각 숨씨 모듈들이 병렬적으로 행동동기 값을 계산하고 이를 기반으로 행동을 선택하기 때문에 변화하는 환경, 불확실한 인지, 확률적 행동 선택 하에서도 강인하게 목표 지향적 행동 선택이 가능하였다. 앞으로 우리는 서비스 로봇에 필요한 임무들을 수행하기 위해 다양한 숨씨들을 생성하여 실험할 예정이다. 또한, 연속적인 행동들에 대해 자동으로 데이터를 분리해서 중간목표를 생성하고, 이를 배치하는 방법을 연구할 예정이다.

### 참 고 문 헌

- [1] P. Perez-Paredes and M. Sanchez-Tornel, *Handbook of Research on E-Learning Methodologies for Language Acquisition*, Information Science Reference, pp. 3, 2008.
- [2] J. J. Gibson, *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston, 1979.
- [3] E. Rome, J. Hertzberg, G. Dorffiner, and P. Doherty, "Towards Affordance-based Robot Control," in *Proc. of Dagstuhl Seminar 062031*, June 5-9, 2006.
- [4] T. Kiyoharu, K. Kenta, and I. Katsumi, "Approach to Artificial Skill from Affordance Theory-Memory and Embodiment," in *Proc. of Journal of the Robotics Society of Japan*, vol. 22, no. 7, pp. 892-900, 2004.
- [5] W. Lidwell, K. Holden, and J. Butler, *Universal Principles of Design*, Rockport Publishers, pp. 20, 2003.
- [6] Dweck and S. Carol, "Motivational processes affecting learning," in *Proc. of American Psychological Association*, vol. 41, pp. 1040-1048, 1986.
- [7] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, ch. 14, 2003.
- [8] D. Heckerman, "A tutorial on learning with Bayesian network," *Technical report, MS Research Advanced Technology Division*, 1995.
- [9] <http://www.dasarobot.com>
- [10] S. Lee, I. H. Suh, and W. Y. Kwon, "A Motivation-based Action-Selection-Mechanism involving Reinforcement Learning," in *Proc. of International Journal of Control, Automation, and Systems*, vol. 6, no. 6, pp. 904-914, December, 2008.
- [11] E. de Sevin and D. Thalmann, "A Motivation Model of Action Selection for Virtual Humans," in *Proc. of Computer Graphics International 2005 (CGI'05)*, Stony Brook, NY, USA, June 22-24, 2005.
- [12] E. Beaudry, Y. Brosseau, C. Cote, C. Raievsky, D. Letourneau, F. Kabanza, and F. Michaud, "Reactive Planning in a Motivated Behavioral Architecture," in *Proc. of American Association for Artificial Intelligence*, pp.1242-1247, 2005.
- [13] E. Lazkano, B. Sierra, A. Astigarrage, and J. M. Ozteta, "On the use of Bayesian networks to develop behaviours for mobile robots," *Robotics and Autonomous Systems*, vol. 55, pp. 253-265, 2007.
- [14] A. Dearden and Y. Demiris, "Learning Forward Models for Robots," in *Proc. of the International Joint Conference on Artificial Intelligence*, pp. 1440-1445, Edinburgh, , 2005.
- [15] O. Lebeltel, P. Bessiere, J. Diard, and E. Mazer, "Bayesian Robot Programming," *Autonomous Robots*, pp. 49-79, 2004.
- [16] L. Montesano, M. Lopes, A. Bernardino, and J. Santos-Victor, "Modeling Affordances using Bayesian networks," in *Proc. of IEEE/RSJ Int. Conference in Intelligent Robots and Systems*, San Diego, USA, October, 2007.
- [17] M. Raubal and R. Moratz, "A Functional Model for Affordance-Based Agents," In: *E. Rome et al. (Eds.): Affordance-Based Robot Control, LNAI 4760*, pp. 91-105, 2008.
- [18] A. Stoychev, "Behavior-Grounded Representation of Tool Affordances," in *Proc. of the 2005 IEEE International Conference on Robotics and Automation*, Barcelona, Spain, April, 2005.

## — 저 자 소 개 —



이 상 형(학생회원)  
 2005년 한양대학교 전자및컴퓨터  
 공학과(공학사) 졸업.  
 2007년 한양대학교 전자컴퓨터  
 통신공학과 석사 졸업.  
 2007년~현재 한양대학교  
 전자컴퓨터통신공학과  
 박사 재학 중.

<주관심분야 : 인공지능, 로봇공학, 지능응용시스  
 템>



서 일 홍(정회원)  
 1982년 한국과학기술원 전기및  
 전자공학과(공학박사)  
 졸업.  
 1985년 대우중공업 연구소.  
 1995년~현재 한양대학교 컴퓨터  
 공학과 교수.

<주관심분야 : 지능응용시스템, 인공지능, 로봇공  
 학>