

# 지능로봇: 지능형 이동로봇 연구

## Intelligent Robot: Trends of Intelligent Mobile Robot

강 진 식

제주대학교 공과대학 통신컴퓨터 공학부

Jin Shig Kang

Faculty of Telecommunication & Computer Eng. Cheju Nat'l Univ.

### 요 약

최근 로봇에 지능을 부여하기 위한 다양한 연구가 진행되어 지고 있으며, 생물학-곤충들의 행동과 지능에 기초를 둔 로봇 지능-행동 연구가 중요한 연구 줄거리를 형성하게 되었다. 이 글은 로봇의 제어 구조-속고형, 반사행동형-에 대한 소개와 각각의 장·단점에 대한 고찰, 생물학-곤충들의 행동과 지능-에 기초를 둔 로봇 지능-행동 연구 분야의 최근의 연구 성과들을 고찰하고 소개한다. 그리고 지능을 부여하기 위한 학습 방법으로서의 강화학습(reinforcement learning)의 연구주제를 소개하고 향후의 로봇 연구 방향에 대하여 고찰한다.

### 1. 서론

수 천년동안 사람은 인간과 비슷한 형태의 피조물을 만들어 왔으며 그 피조물로 하여 인간이 하는 것과 유사한 작업을 하도록 하여 왔다. 로봇이라는 용어가 처음 사용된 것은 극작가 Karel Capek 의 희곡 "R.U.R(or Rossum's universal robots)에서부터 이다. 지능을 갖는 로봇이 처음 일반에 소개된 것은 1977년 영화 스타워즈부터 이다. 영화 스타워즈에는 휴머노이드 형태의 C3PO로봇과 R2D2로봇이 출연하여 인간을 돕는 역할을 담당하였다. 그러나 1990년대 이전에는 로봇이 걷고 사고하고 말을 하는 로봇(우리는 이러한 로봇을 android라 부른다)이 아니라 공장에서 자동차를 조립하고 용접하고 부품을 옮기는 역할을 담당 하였었다. 최근에 와서 로봇은 공장을 탈출하여 좀더 우리 주변에 가까이 다가왔다. android의 기능에 대한 연구는 현재에도 진행되고 있으며, 걷고 간단한 대화를 하는 로봇은 이미 등장하였다. 로봇은 인간이 감당하기 어려운 일들(예를 들면 원자력 발전소, 심해 탐사, 화성 탐사 등)을 해 내고 있다. 현재의 로봇 연구는 공장자동화 분야에서 ① 오락용, 장난감로봇 ② 개인 서비스로봇 ③ 의학용 수술로봇 ④ 환자보조로봇 ⑤ 가이드 경비로봇 ⑥ 위험한 환경에서의 탐사로봇 등과 같은 분야로 연구범위가 확대되어 왔다.

최초의 지능형 로봇은 W. Grey Walter의

Tortoises이다. W. Grey Walter는 두개의 센서(광센서, 범퍼)와, 두개의 모터, 두개의 뉴런-증폭기-로 구성된 로봇 Tortoises를 발표하였다. 이 로봇은 광센서를 이용하여 빛이 들어오는 방향을 scan하고 빛이 들어오는 방향으로 달려가도록 되어있다. (<http://www.newscientist.com/hottopics/ai/walter.jsp> 참조)

이 글은 지능형 로봇에 대한 연구 성과들을 소개하고 향후의 연구 방향에 대해 고찰 한다. 이글의 구성은 2장에서 로봇의 제어구조인 행동기반 제어시스템과 속고형 시스템, 그리고 복합형 구조를 소개하며, 3장에서는 생물학에 기초한 로봇에 대하여 소개한다. 그리고 4장에서는 로봇의 학습과 진화형 로봇에 대하여 서술하고 5장에서 지능형 로봇에 대한 연구 추세에 대하여 고찰

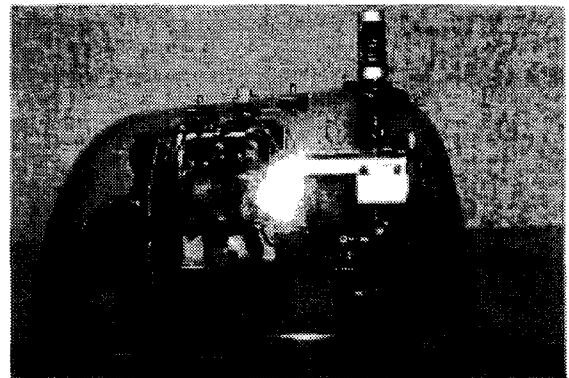
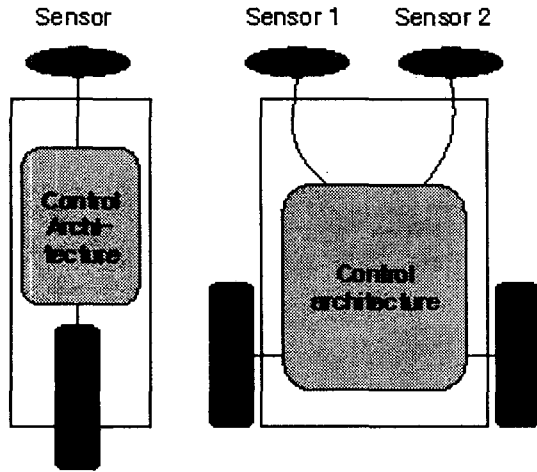


그림 1. Grev Walter의 Tortoises



(a) type 1 (b) type 2, 3, 4, ...  
그림 2. Braitenberg vehicle

하며 6장에서 이 글의 결론을 맺는다.

## 2. 로봇 제어구조

로봇의 제어구조는 숙고(deliberative)형과 행동기반(behaviour based) 구조 대표되며 두 구조의 장점을 조합한 조합형(hybrid)구조가 있다. 이 장에서는 지능형 이동 로봇의 시초가 된다고 볼 수 있는 Braitenberg vehicle에 대한 소개, 숙고형 제어구조, 행동기반 제어구조에 대하여 서술하고 장·단점을 고찰한다.

### 2.1 Braitenberg vehicle

Valentin Braitenberg[1, 2],(1984)은 새로운 인공 지능(New AI)의 개념인 심리합성(synthetic psychology)을 여러 형태의 이동로봇을 이용하여 제시하였다. 로봇은 센서들과 로봇의 구동을 위한 모터, 센서와 모터를 연결하는 제어구조(control architecture)로 구성되는 단순한 구조이며(그림 2 참조) 제어구조에 따라 이동 로봇의 행동이 다양한 형태로 진화할 수 있음을 보였다. 그림 2 (a)의 type 1은 센서에서 감지되는 신호에 따라 로봇은 전진과 후진을 할 수 있도록 되어 있으며 (b)는 Type 2, 3, 4에서 채용된 구조로 제어구조의 연결 방법에 따라 로봇은 다양한 행동을 하게 된다. 이러한 행동은 연결 방법에 의하여 결정되며 이 연결은 신경망이 담당한다. 각 뉴런은 서로 연결되어 상호작용을 하며 센서로부터 인식된 환경 신호를 이용하여 모터 출력을 결정한다. 이 방법의 중요한 점은 모터 출력을 결정하거나 환경 인식을 위한 어떠한 프로그램도 작성하지 않아도 되며 단지 뉴런에 의하여 학습된 만큼 센서신호를 증폭시키는 것에 불과하다는

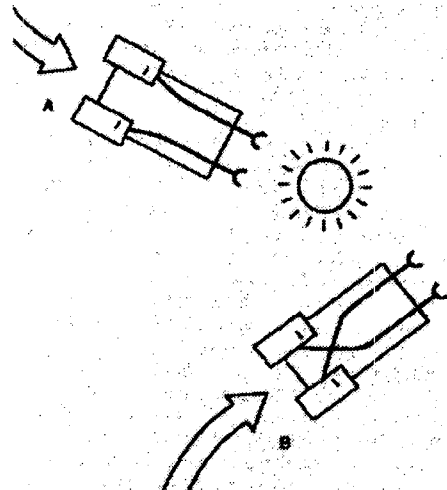


그림 3. Braitenberg vehicle의 행동  
A: 빛을 따라감(공격),  
B: 빛을 회피함(두려움)

것이다. 로봇의 행동은 내부 연결 방법과 외부환경의 상호 작용에 의하여 결정된다. Braitenberg vehicle은 메모리가 없는 단순 반사행동제어방식이며 이 로봇에 메모리(기억능력)가 추가되면 로봇의 행동 아주 복잡하게 된다. V. Braitenberg은 인공신경망의 Hebbian 학습 규칙을 제시하여 로봇이 다윈 진화론에 따른 진화를 할 수 있음을 보였다. 예를 들면 그림 3의 A는 빛-자극-에 대하여 공격적이고 B는 빛에 대하여 두려워하는 행동을 보인다. Valentin Braitenberg은 아주 복잡하고 동기 부여된 어떠한 행동도 단순한 정보 전달 체계로 되어 있으며 단순한 구조로 구성된 animatic creature도 예상된 행동을 하도록 할 수 있음을 보였다.

Braitenberg vehicle의 확장으로 MAVRIC[3, 4] (mobile, autonomous vehicle for reach in intelligent control)이 있다. (그림 4, <http://faculty.washington.edu/gmobus/AdaptiveAgents/mavric.html>)

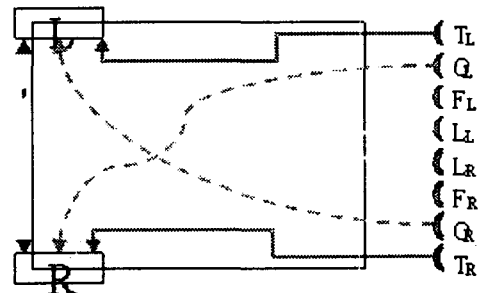


그림 4. 다중센서를 채용한 Braitenberg vehicle: T-touch sensor, L- light sensor and O-Ordor sensors

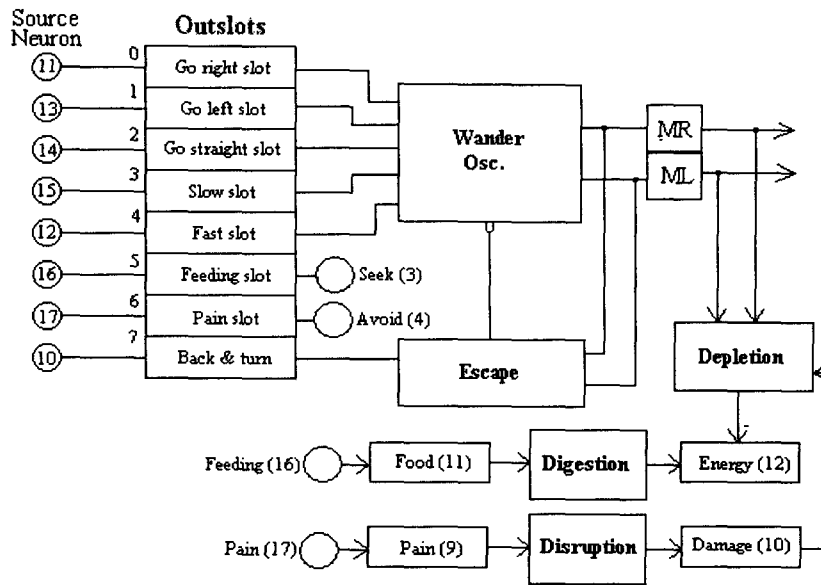


그림 5. MAVRIC 두뇌구조

MAVRIC은 다중센서 를 채용한 Braitenberg vehicle로 센서-인지 부시스템, 학습-행동선택에 관련된 부 시스템, 행동선택 부시스템과 모터제어 부시스템으로 구성되며 각 부 시스템은 동적 뉴런[5](dynamic neuron - adaptrode based)으로 구성되었다. 이 로봇은 다음절에 설명될 Brooks의 포섭(subsumption) 구조로 행동을 선택한다. 즉, 두 온도센서가 모두 감지되면 로봇은 전진하고 그렇지 않으면 빛 센서가 감지되어 지시하는 대로 행동한다. 실제 구현된 로봇은 액티브미디어사의 pioneer I을 바다로 구현하였다. (자세한 내용은 <http://faculty.washington.edu/gmobus/AdaptiveAgents/MAVRIC-EBA.html> 를 참조)

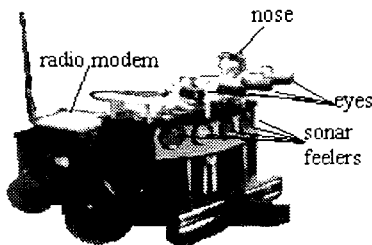
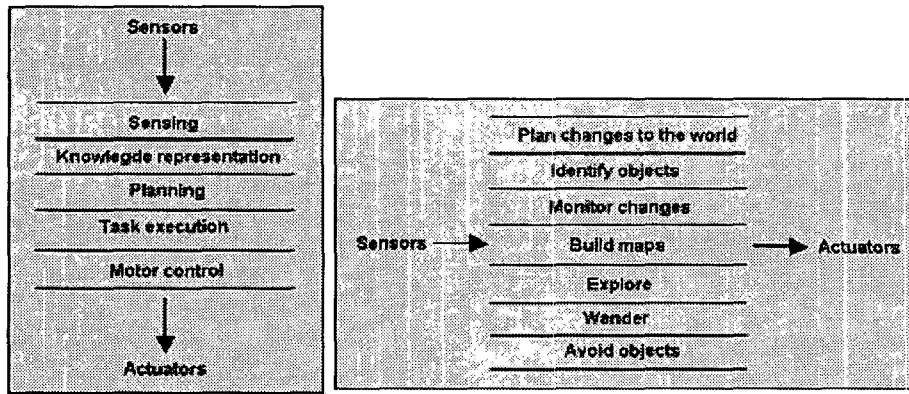


그림 6. MAVRIC - an embodied EBA agent. The robot is based on an ActivMedia Pioneer I platform. Light sensors

## 2.2 제어구조

고전적인 로봇 제어구조로서 **숙고형 (deliberative)** 구조는 분리된 일련의 모듈들의 정보처리 과정을 통하여 지능적인 행동을 하도록 한다. 센서에서 감지된 환경신호는 모든 모듈을 통과하면서 행동을 계산하는 구조로 신호의 처리 과정은 그림 7 (a)에서 보여 지는 바와 같다. 즉, 숙고형 제어구조는 모든 행동이 **환경감지 (sensing)-계획(planning)-행동(action)**의 주기로 수행되어 로봇의 반응속도가 매우 느리고 정적이며, 구조화된 환경(structured environment) 하에서만 정상적인 성능을 보여준다. 그러나 로봇이 실제로 작업하는 공간은 대부분 구조화되지 않으며 동적으로 변화하는 환경이다. 만일 로봇이 현재에 처해있는 상황에 대하여 계획(planning)이 완료되어 행동이 수행되고 있다 하더라도 갑작스런 환경 변화가 발생하면 새로운 계획을 수립하여야 한다. 외부의 도움 없이 모든 행동을 결정하고 수행 하여야 하는 자율이동로봇이라면 계획연산에 대한 하드웨어적인 부담이 매우 커지게 된다. 과거에는 이러한 문제를 해결하기 위한 방법으로 더 빠른 프로세서, 보다 효율적인 연산 알고리즘, 강인한(robust) 하드웨어를 사용하는 방법들이 연구되었지만 성능-가격측면에서 효율적이지 못한 것으로 알려졌다.

다른 한편으로는 **행동기반(behaviour based)** 제어구조가 R. Brooks[7]에 의하여 제시되어 최근 많은 연구가 수행 되고 있다.[7~12] 행동기반

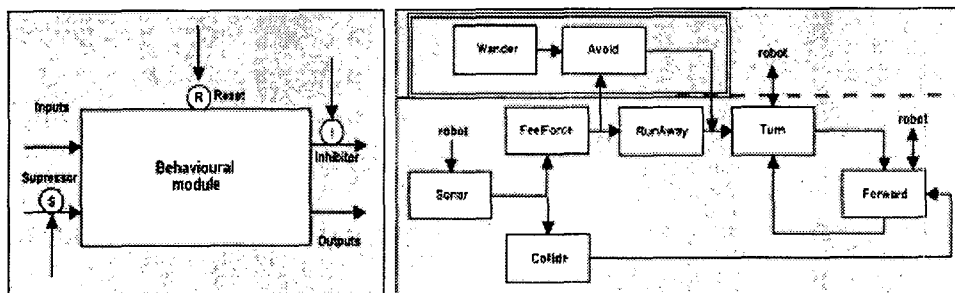


(a) 속고형 제어구조      (b) 행동기반 제어구조  
 그림 7. 속고형 제어구조 및 행동기반 제어구조

구조는 행동분해(behaviour decomposition)를 토대로 하며 목표(task)를 성취하기 위해 로봇은 다른 모든 행동(behaviour)을 알아야 한다. 이 행동들은 모두 로봇에 구현 되어져야 한다. 주어진 상황 아래서 각 행동은 그 행동을 수행하는데 필요한 센서 데이터에 접근할 수 있어야 하며 중심이 되는 계획과 제어원칙 없이 서로 병렬로 구현되어야 한다. 그리고 어떠한 돌발 상황에 대해서도 에이전트의 행동은 구현되어 있는 각 행동의 상호 작용에 의하여 생성될 수 있어야 한다. 이러한 행동기반 로봇 제어에서의 주요 이슈는 중재문제(arbitration problem)로 대표되는 각각의 행동 모듈의 출력이 서로 상반된 행동을 요구할 경우 어떤 행동을 로봇이 실제 수행 하게 할 것인가를 결정하여야 하는 경우이다. R. Brooks[7, 8]는 목표성취를 위한 행동으로 각각의 행동을 수평으로 구분하고(그림 7. b), 권한레벨(levels of competence)을 부여한 포섭(subsumption)구조를 제시하였다. 권한레벨은 가장 기초가 되는 행동을 가장 아래에 두고 각각의 상위구조들은 하위 구조와 시험을 거쳐 개별적으로 추가 된다. 또한 각각의 계층은 하위계층의 정보를 이용할 수 있으며 하위계층의 행동을 억제함으로써 하위계층의 행동에 영향을 줄 수 있다.

각각의 행동 모듈은 확장된 FSM(augmented finite state machine)으로 구현된다. FSM은 현재의 상태  $s(t)$ 와 입력  $i(t)$ 로부터 다음 상태  $s(t+1)$ 가 결정되며 되며 제한된 수의 상태만 갖는다. AFSM은 FSM에 시간지연을 고려한 timing mechanism과 내부 데이터 구조를 유지할 수 있도록 하는 기능이 부가된 것으로 그림 8(a)와 같은 구조이다. 그림에서 I는 상위 행동에 의한 억제 신호이며 S는 suppressor, R은 리셋 신호이며 이들 신호는 다른 행동을 구현한 AFSM의 출력 신호들이다. 그림 8 (b)는 장애물 회피 행동에 wander 모듈 추가하는 경우를 설명하고 있다. 그림에서 겹줄 상자는 추가되는 wander 행동이다. 장애물 회피 행동은 다른 물체가 로봇을 향해 접근하면 돌아서 이동하여야 하며 충돌이 일어나기 직전이라면 정지하여야 한다. wander 행동의 avoid 블록은 장애물 회피 행동의 FeelFree 블록으로부터 센서신호를 전달 받으며 RunAway 블록의 출력을 억제하여 자신의 출력으로 대체함으로써 로봇의 진행방향을 결정하게 된다.

[http://www.ai.mit.edu/people/brooks/publications\\_shtml](http://www.ai.mit.edu/people/brooks/publications_shtml)



(a) 행동모듈      (b) 장애물 회피 행동에 wander 모듈 추가  
 그림 8. 행동모듈 및 행동모듈 추가

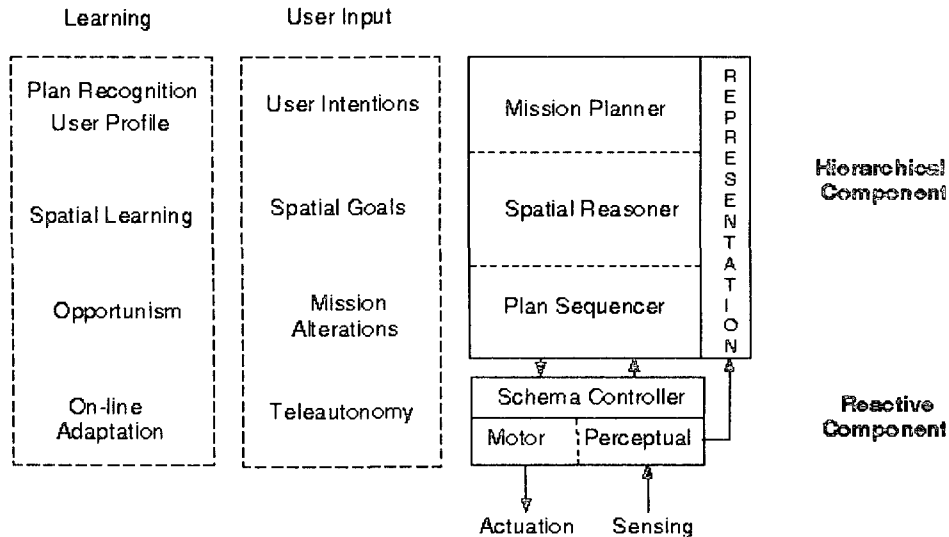


그림 10. AuRA - A Hybrid Deliberative/Reactive Architecture

### 2.3 Hybrid Architecture

행동기반 제어구조인 포섭 구조는 저가이면서 강인하지 않은 low-level의 하드웨어를 사용하여 다양한 시스템에 적용되어 성공사례가 발표되었다. 그러나 행동계층(hierarchy of behaviour)이 고정되어 하드웨어 구현이 엔지니어에 의존적이 될 수밖에 없다. 행동기반 제어구조-반사행동 제어-의 로봇은 장애물을 회피하고 벽을 따라가며, 좁은 통로(문: doorway)를 통과할 수 있지만 복잡한 행동들에 대해서는 제한된 성능만을 보인다. R. Brooks의 포섭(subsumption) 구조는 이러한 문제를 해결하기 위한 방안으로 제시되었지만 완전한 해결 방법이 되지는 못하였다. R. Arkin[10]은 반사행동 네비게이션은 제한된 네비게이션 성능을 보이며 다양하고 일반적인 장치-로봇-에는 적응성(adaptability)이 요구됨을 지적하였다. 하이브리드(hybrid: reactive/planning) 제어구조의 반사행동 계층은 계획(planning) 또는 sequencer의 통제를 받도록 하는 구조로 최근 많이 연구되어지고 있다. 기본 구조는 그림 9와 같다.[11]

Georgia Institute of Tec.(G-Tec)의 Aura [15]는 반사제어 시스템, 네비게이션 및 미션 계획기의 계층 제어구조를 갖는다.(그림 10 및 홈페이지 참조). Aura는 학습(learning)을 통하여 반사행동과 계획을 조합하는 하이브리드 구조로 되어 있다. SRI(stanford research inst.)의 SAPHIRA[16, 17]는 로봇 인지-행동(perception and action)의 복합구조로 퍼지제어방식의 반사행동 계획 시스템, 행동 sequencer, 숙고형 계획기를 포함하며 초음파센서를 이용한 지도제작과 네비게이션 기능이 있다.(그림 11참조)

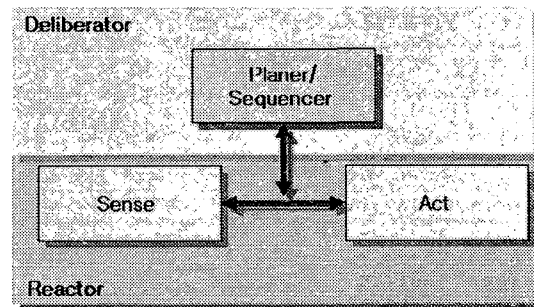


그림 9 하이브리드 제어구조

Bonasso 와 Kortenkamp의 3T[18, 19]는 3-layer로 구성되어있다. 3-layer는 반사제어시스템, 목표 sequencer, 숙고형 모듈로 구조는 그림 12와 같다. 3T 구조의 주 아이디어는 계획과 sequencing, 그리고 제어를 구분하는 것으로 상

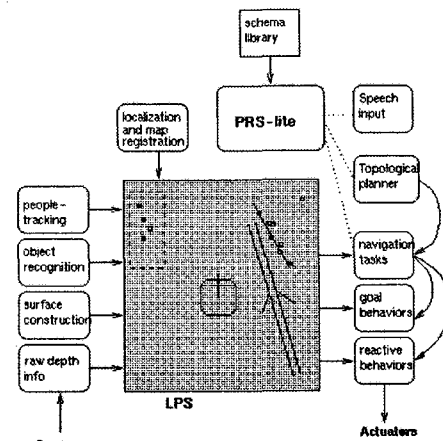


그림 11 Architecture: SAPHIRA

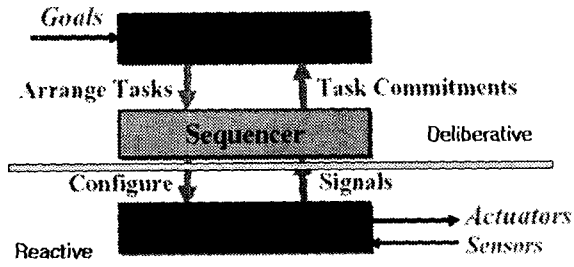


그림 12. Bonasso 와 Kortenkamp의 3T구조

위계층은 하위계층의 제어흐름을 제공하고 하위 계층은 상태변화와 테스크 성취의 성공 실패의 동기신호를 상위계층에 제공한다. 이는 각 계층이 자신의 역할만 담당 하도록 튜닝 되는 이질적인(heterogeneous) 구조이다.

### III. 생물학에 기초한 로봇

많은 학자들은 자연 현상에서 영감을 얻어 어려운 공학문제를 해결 하여왔으며 특히 최근의 로봇 연구에서는 그러한 현상이 두드러지게 나타나고 있다. 이 장에서는 생물학(biology)-곤충학(ento-mology)에 토대를 둔 로봇지능연구에 대한 소개를 한다.

#### 3.1 Phonotaxis in crickets and robots

phonotaxis는 음원을 찾아가는 능력으로 암 귀뚜라미는 수 귀뚜라미가 내는 소리를 이용하여 그 위치를 찾아간다. 귀뚜라미의 귀는 앞다리에 위치하며 tracheal 관으로 연결되어 있어 음압차를 이용한 방향 판단이 가능하도록 되어있다. (그림 13). Reeve와 Webb[21~23]은 ANN을 이용한 Robot을 제시하였다. 뉴런을 고정된 membrane capacitance와 membrane conductance를 갖는 RC 회로로 모델링하고, membrane 양단

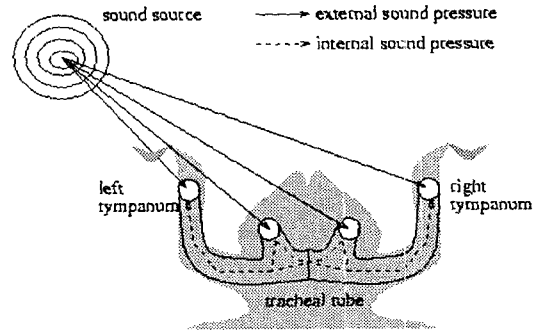


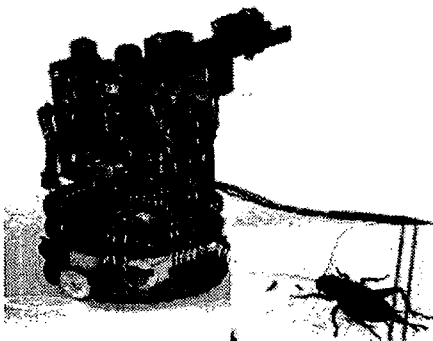
그림 13. 귀뚜라미 귀의 구조

의 전압은 외부 입력이 주어지지 않으면 지수적으로 감소하도록 한다. 만일 membrane 전위가 threshold보다 높으면 뉴런이 점화되며 synapses에 점화 신호를 출력한다. synapses는 전위과 같은 전위를 갖도록 하는 variable conductance로 모델링되며 synaptic neurotransmitter에 의하여 개방되는 ion channel의 역-전위 역할을 한다. 따라서 synapses는 역-전위와 시변 시간지연으로 특성화되며, ion channel의 컨덕턴스 변화에 의하여 weighting을 결정할 수 있게 된다. 그림 14에 Reeve와 Webb의 로봇과 로봇구조를 보였다.

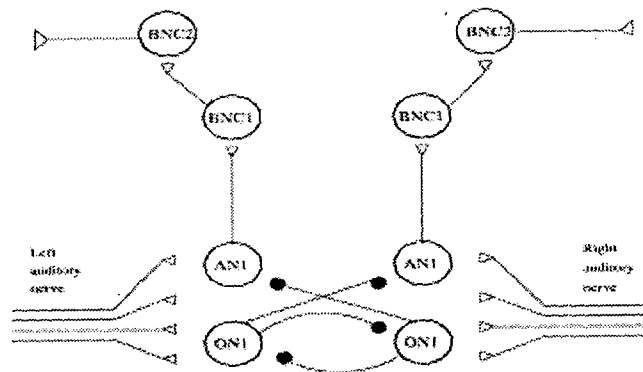
#### 3.2 네비게이션을 위한 곤충의 전략

Cataglyphis는 큰개미의 일종으로 집(nest)를 찾아올 때 pheromone을 사용하지 않는다. Cataglyphis의 homming 네비게이션에[23~25] 대한 연구결과로 path-integration, visual piloting과 구조적 탐색(systematic search)에 의하여 집을 찾는다. 하늘의 극화패턴(polarization pattern, 그림 15참조)으로 부터 얻어지는 방향정보를 이용한path-integration뿐만 아니라 집 근처의 landmark들을 이용한다.

그림 16은 구현된 로봇과 로봇에 이용된 센서



(a) 귀뚜라미 로봇



(b) 로봇 제어구조

그림 14 Reeve와 Webb의 로봇과 로봇구조

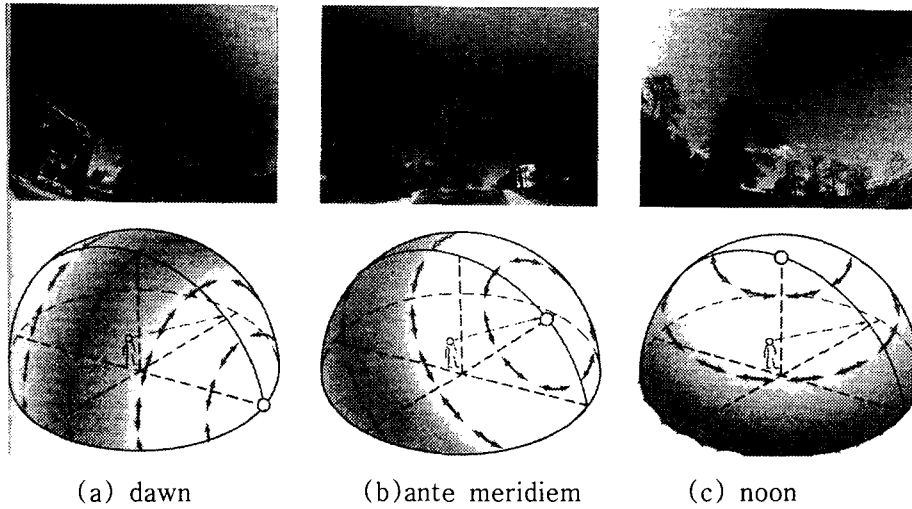


그림 15. polarized pattern of the sky

이다. 이 로봇에는 3가지의 시각센서-panoramic visual system, polarized - light sensors, ambient-light sensors-가 이용되었다. 이 로봇에 이용된 시각 시스템은 CCD 카메라와 con형 거울로 구성되며 로봇을 중심으로 하는 360° 전 방향의 화상이 얻어지며 로봇은 이 화상을 이용한 네비게이션을 수행한다.(그림 17)

### 3.3 ROBONOSE

대부분의 개미들은 자신이 지나가는 경로에 화학 물질을 분비하여 지나간 자취를 표시하고

이를 추적하여 집까지 돌아온다. 이러한 특성은 electro-chemical 센서를 이용한 로봇 네비게이션에 응용될 수 있다. A. Lilienthal과 T. Duckett Braitenberg는 vehicle형태의 로봇(그림 18참조)으로 냄새를 추적하고[26, 27] 건물 내 환경에서 Gas 농도 그리드-맵을 구하는 방법을 제시하였다. electro-chemical 센서를 이용한 Braitenberg vehicle에서 센서-모터 연결은 바퀴의 속도와 감지된 신호세기 사이의 선형 전달함수 형태로 구하여 지며 이동 로봇의 localization에 적용될 수 있음을 보였다.

### 3.4 Biomimetic Robot[28~31]

Stanford 대학의 M. Cutkosky와 T. Kenny, Berkeley의 R. Full과 H. KazerooniU.C., Harvard 의 R. Howe, 그리고 Johns Hopkins 대학의 R. Shadmehr이 팀을 구성하여 생태 모방로봇 (biomimetic robot)에 대한 연구를 1999년부터 수행하고 있다. 생태 모방로봇(biomimetic robot)은

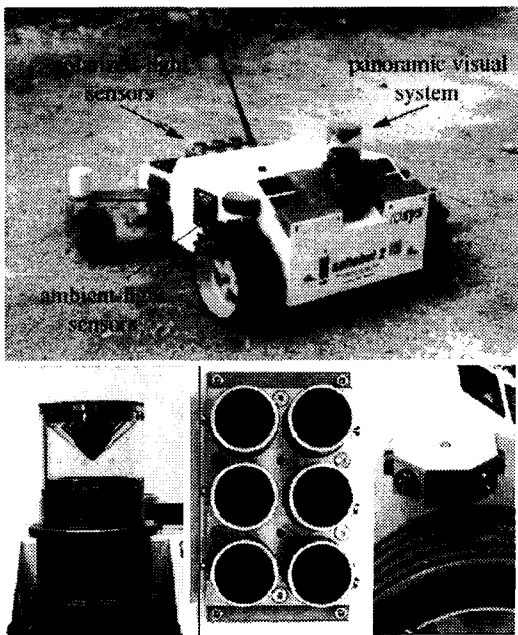


그림 16. The mobile robot Sahabot 2 and Robot sensors. From left to right: the panoramic visual system, polarized light sensors, ambient-light sensors.

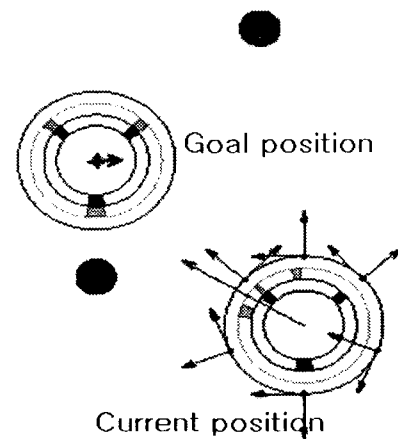


그림 17 네비게이션

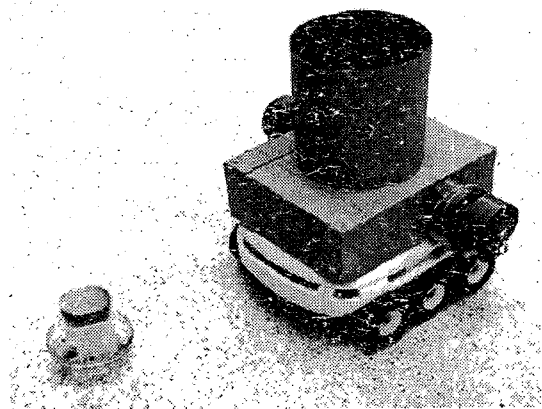


그림 18. " Orebro Mark III mobile nose.

곤충의 기계적 특성을 연구하는 low level control, 곤충을 모방한 구조 및 내부장치를 연구하는 Shape Deposition Manufacturing, 그리고 곤충의 locomotion, 지능을 로봇에 구현하는 High-Level Control로 나눌 수 있다. Low level control은 ① What passive properties are found in nature? ② What properties in mechanical design? ③ How should properties be varied for changing tasks, conditions?를 주제로 생물학적 물질, 요소, locomotion에서의 역할, SDM (shape decomposition manufacturing) 물질 및 요소로부터 material behaviour의 모델링과 SDM 구조제작을 위한 설계규칙을 연구하고 있다. 현재까지의 연구 결과로 ① 근육은 광범위의 포텐셜 함수를 가지며 ②근육의 기능은 신경 활성화로 예측할 수 없으며 같은 모터 뉴런에 의하여 활성화되는 근육이라도 유사한 기능을 하지 않을 수 있다. High-Level Control은 ① How is Compliance used in Locomotion? ② What Compliance Strategies in Human-level Tasks?을 주제로 연구하여 곤충들을 자연주파수 10 Hz로 달리며 곤충의 다리는 spring and damper system과 유사한 동작을 하며 다리의 임피던스

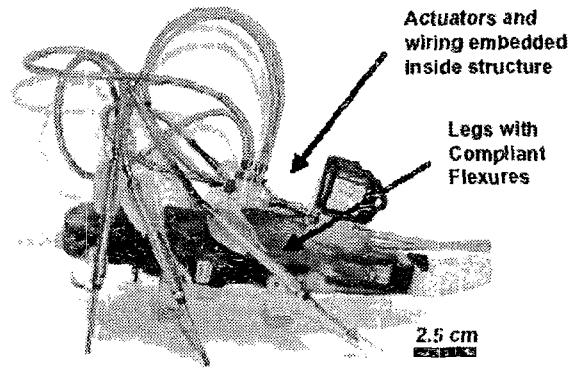


그림 20. Sprawlita.

는 10 Hz까지 주파수에 따라 증가함을 보였다. ([http:// www- cdr.stanford.edu/biomimetics/](http://www-cdr.stanford.edu/biomimetics/) 참조)

#### IV. 학습-진화 및 다중로봇

로봇은 매우 동적이며 구조화 되지 않은 환경에서 작업한다. 따라서 환경 변화에 따른 적응성 (adaptability)을 요구되며 이 적응성은 로봇이 학습기능을 갖거나 스스로 진화할 수 있도록 설계함으로써 성취될 수 있다. 이 장에서는 로봇의 학습기능과 진화, 다중로봇에 대하여 설명한다.

##### 4.1 로봇의 학습

로봇에 있어서의 적응성은 로봇의 행동이 환경변화에 의하여 조정되어야 하는 행동적응성 (behaviour adaptability), 시간이 경과함에 따라 이전에 실패했던 행동을 변화시키는 진화적 적응성 (evolutionary adaptability), 그리고 로봇의 지각시스템이 환경변화에 적응해야 하는 센서 적응성 (sensor adaptation)으로 구분할 수 있다. 로봇은 학습을 통하여 수행성능 (performance)을 개선할 수 있으며, 접근방법은 ① 강화학습 (reinforcement learning), ② 신경망

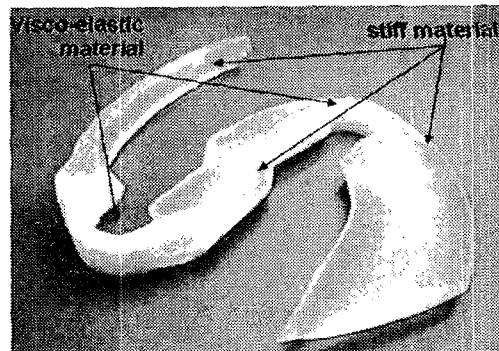
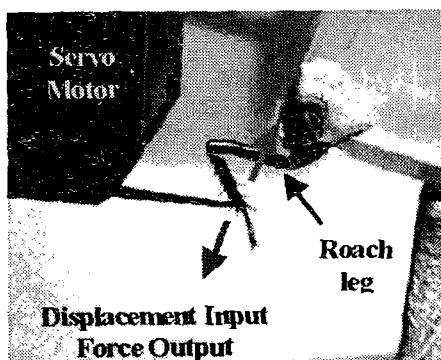
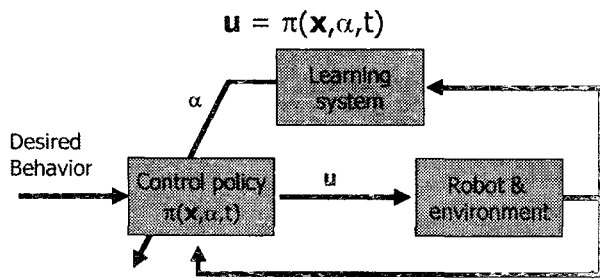
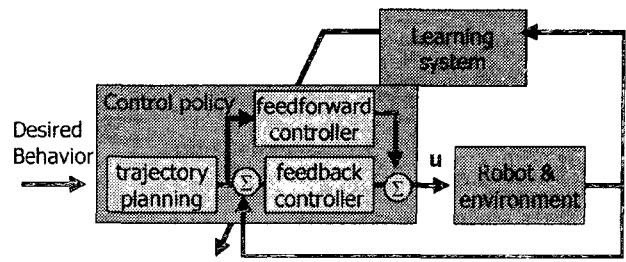


그림 19 "Sprawlita"의 dynamic locomotion과 flexible leg.





(a) Task learning



(b) indirect learning

그림 21. Task learning and indirect learning

학습, ③ 진화학습(evolutionary learning), ④ Instance Based learning, ⑤ Case base reasoning으로 구분된다.(학습유형: 표 참조).

|                                                                                           |                                                                                           |
|-------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|
| Numeric<br>Ex: logical assertions, production rules, semantic rules                       | Symbolic<br>Ex: neural network, statistical methods                                       |
| Inductive<br>Learning from experiment or examples                                         | Deductive<br>Produce more efficient concept from initial one provided by robot            |
| Continuous (online)<br>Learning while robot is interacting with environment               | Batch (offline)<br>Processing large batch of data prior to making behavior change         |
| Supervised<br>Explicit notion of correctness i.t.o. optimal behavior established a priori | Unsupervised<br>No notion of particular target goal state the system is trying to achieve |

그림 21(a)는 task learning의 개념을 설명한다. 연속상태  $s$ 로부터 입력  $u$ 로의 mapping 정책(policy)  $\pi$ 를 찾는 과정을 보여주고 있으며 (b)는 제어정책(control policy)를 찾는 간접학습의 개념을 설명하고 있다. 간접학습 과정은 task를 planning과 execution stage로 분리하고 planning으로부터 원하는 기구학 경로(kinematic trajectory)를 생성하며 계획(plan)을 모터 명령으로 변환하고 역 기구학 모델을 이용하여 학습을 수행한다.

#### 4.2 강화학습

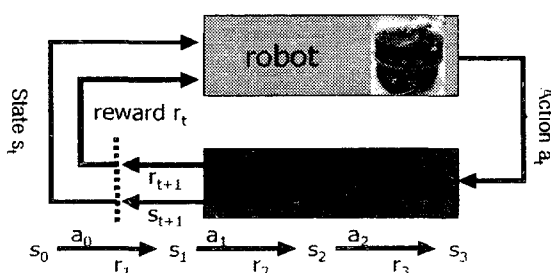


그림 22 강화학습

강화학습(RL: reinforcement learning)은 그림 22와 같은 순서로 학습이 이루어진다. 현재상태  $s_t$ 에서 행동  $a_t$ 를 할 경우 상·벌은  $r_{t+1}$ 이며 식 (1)로 표현되는 할인된 누적 상·벌(discounted accumulated reward)을 최대화 하는 제어정책  $a = \pi(s)$ 을 학습에 의하여 구한다.

$$R = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots + \dots$$

$$= \sum_{i=0}^{\infty} r_{t+i} \gamma^i \quad (1)$$

식 (1)에서  $\gamma: 0 < \gamma < 1$ 인 할인 인자(discounting factor)로 예측 공간(estimation horizon)을 무한 공간(infinite horizon)이 아닌 유한공간(finite horizon)으로 만들어 준다. 여기서  $\pi_t(s): S^t \rightarrow A$ 를 출발상태  $s_t$ 로부터 학습된 정책이라 하면 이 정책은 향후의 reward 기대치  $E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots]$ 를 최대화 한다. 상태 값 함수(state value function)  $V^\pi(s)$ 와 행동 값 함수(action value function)  $Q^\pi(s, a)$ 를 다음과 같이 정의하자.

$$V^\pi(s) = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots$$

$$= \sum_{i=0}^{\infty} r_{t+i} \gamma^i \quad (2)$$

$$Q^\pi(s_t, a_t) = r(s_t, a_t) + \gamma \max_a Q^\pi(s_{t+1}, a^*) \quad (3)$$

상태 값 함수(state value function)  $V^\pi(s)$ 는 상태  $s$ 를 출발하여 정책  $\pi$ 를 따를 때 누적 상·벌의 예측치이며, 행동 값 함수(action value function)  $Q^\pi(s, a)$ 는 상태  $s_t$ 에서 행동  $a_t$ 를 하고 이후에는 정책  $\pi$ 를 따를 때 누적 상·벌의 예측치이다. 이때 최적 정책은

$$\pi(s) = \operatorname{argmax}_a Q^\pi(s_{t+1}, a^*) \quad (4)$$

이다.

강화학습에 있어서의 딜레마는 완전한

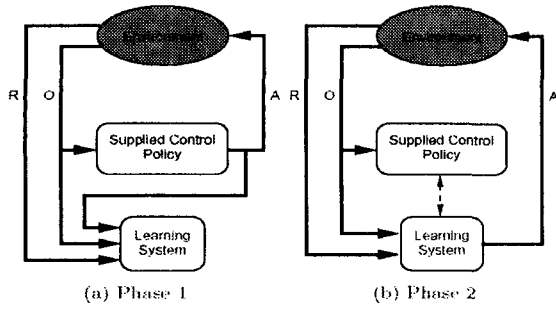


그림 23 OPL

unsupervised 학습이기 때문에 발생된다. 간단한 가이드라인(정보)을 추가하는 것(supervised)만으로도 쉽게 해결할 수 있는 단순한 문제도 RL을 이용하면 학습하는데 많은 시간이 걸릴 수 있다. 자연계에는 선행지식-정보-가 없는 학습은 존재하지 않으며 자연계에서 일어나는 모든 학습은 경험을 통하여 개선되어진다(학습 또는 진화). Smart와 Kaelbling[41]은 OPL(Off-polish Passive Learning)알고리즘을 제시하였다. OPL은 시스템 동력학식에 대한 정보와 센서 없이 단순한 경로(trajectory)만으로 학습을 수행한다. 학습은 2단계고 나누어지며 그림 23(a)의 첫 번째 학습 단계에서 로봇은 실제 하드웨어 코딩된 코드나 조이스틱을 이용한 제어행동인 제어정책(supplied control policy)에 의하여 제어되며, 이 동안에 RL 시스템은 상태, reward, action을 수동적으로 관측만하고 실제로 로봇을 제어하지는 않는다. 제어정책은 RL시스템에 0아닌 reward를 제공하여야 하며, 이는 trajectory를 학습하는 것이 아니라 경험을 생성해 내는 것이다. VFA(value function approximation.)이 종료되면 그림 23(b)의 두 번째 단계가 시작되며, 이 단계에서 제어정책은 표준 RL에서처럼 로봇을 제어

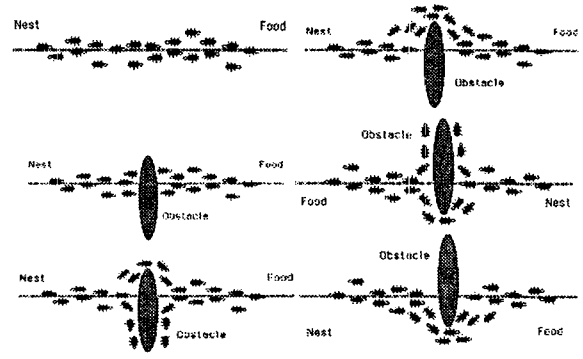


그림 24 개미의 집단행동

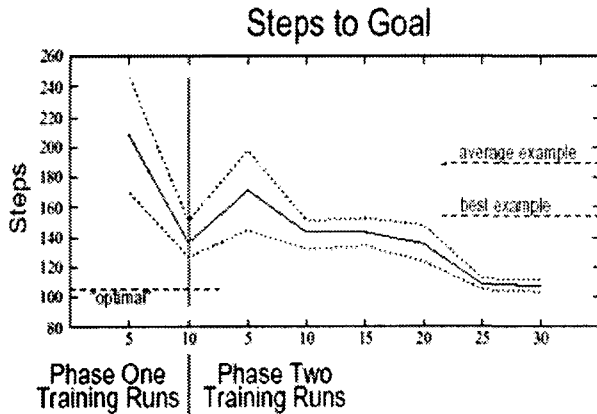
한다. 만일 제어정책이 software라면 제어결정(control decisions)들을 보완하기 위하여 동작이 유지된다. 학습을 이와 같이 두 단계로 나누면 로봇을 제어하기 위한 학습된 정책을 찾기 이전이라도 VFA로부터 bootstrap 정보를 알아낼 수 있으며, 두 번째 단계로 넘어가면 로봇은 reward가 주어지는 상태들을 알 수 있게 되며 학습이 이루어진다. 실험결과로 Supervision에 의한 수행 성능 저하는 빠르게 회복되며 보다 나은 정책을 찾음을 보였다.(그림 24) 이 과정에서 정보(knowledge)가 주어지지 않으면 목표 찾기(goal finding)는 학습 진행과정보다 더 많은 시간이 걸릴 수 있다.

### 4.3. 다중로봇

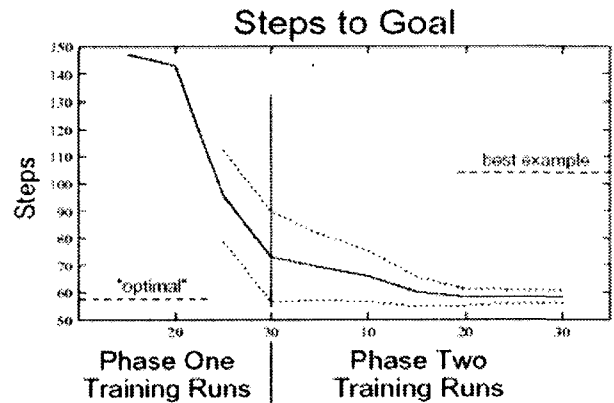
다중로봇에 대한 연구는 그룹학습 문제, 지식의 공유, 통신, 중재 문제 등 다양한 주제를 포함하고 있다.

#### (1) Swam Intelligence[43]

자연계에는 무리지어 행동하는 곤충들이 있으며 리더가 없이도 서로 협동하고 환경에 잘 적응

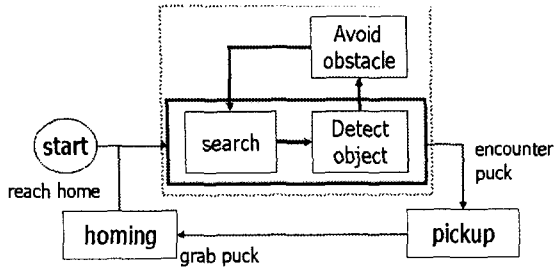


(a) Wall Following



(b) Obstacle Avoidance

그림 25. OPL 실험결과



(a) 로봇 제어흐름



(b) 실험

그림 26 Khepera 로봇을 이용한 swam intelligence

한다. 벌, 말벌, 개미 집단의 행동은 환경 변화에 대하여 유연(flexible)·강인(robust)하고 분포(decentralized)되어있으면서 Self-organized인 특징을 보인다. 사회적인 곤충은 무리를 이끄는 리더가 없어도 복잡하고 지능적인 self-organization 특징을 보인다. 이 특징은 지능시스템 설계에 잘 적용될 수 있으며, self-organizing에 의한 사회적 곤충 모델로부터 분포 문제를 해결하는 장치의 설계에 도움이 된다. 이것을 **swam intelligence**라 한다. 그림 25은 개미집단의 행동을 나타내었다. 개미집단이 갑자기 장애물이 나타났을 때 최적경로를 찾아가는 과정을 보여주고 있다. swam intelligence를 로봇에 적용할 때의 문제점은 문제 정의가 어려우며 이에 따라 프로그램이 어렵다는 점이다. swan

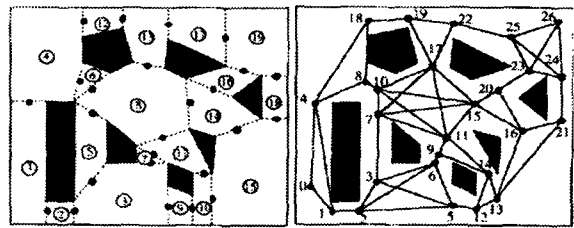
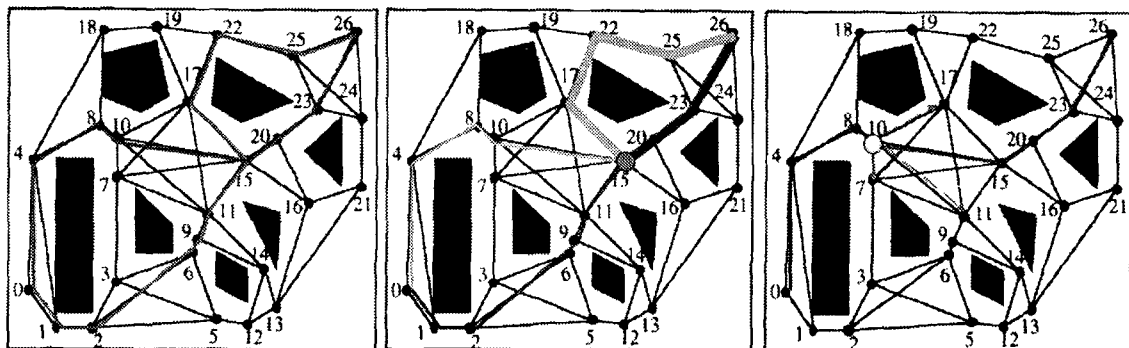


그림 27 Initial setup

intelligence를 갖는 시스템을 구현은 집단행동(collective behavior)에 대한 catalog를 만들고 사회적 곤충의 목표접근 방식을 모델링 함으로서 가능해 질 수 있다. FSA(finite state automation)인 반사행동 제어형태의 다중 로봇을 이용한 foraging 시나리오(물건 모으기)를 고려하면 각각의 로봇의 상태-행동은 그림 26(a)로 묘사될 수 있다. 그림 26(b)는 Khepera 로봇을 이용한 실험 과정을 보였다.

(2) Learning and planning of Multi-robot

여러 로봇이 같은 공간에서 task를 수행한다면 서로 협력, 경쟁, 공생인 관계에 있게 된다. 다중로봇에서의 학습은 개개의 로봇에서 일어나며 다른 로봇과 협력, 경쟁, 공생관계가 유지되어야 한다. M. Mataric[44~46]은 PE: progress estimator를 도입한 강화학습으로 다중로봇에 대한 학습 방법을 제시하였다. PE는 부분내부회로로 구성되며 현재의 목표에 대한 양(+), 또는 부(-)의 reward를 제공한다. behavior가 진행되는 동안에는 다른 behavior로 전환하지 않음으로서 탐색을 보장하여 하나의 behavior를 선택할 수 있게 한다. Task학습은 모든 behavior는 알려져 있을 때 behavior 간의 switching 함수를 학습하게 된다. M. Mataric의 다중로봇 학습 방법은



**Chromosome**  
0-1-2-6-9-11-15-17-22-25-26  
0-4-8-10-15-20-23-26

**Crossover**  
Parent1: 0-1-2-6-9-11-15-17-22-25-26  
Parent2: 0-4-8-10-15-20-23-26  
Child1: 0-1-2-6-9-11-15-17-22-25-26  
Child2: 0-4-8-10-15-17-22-25-26

**Mutation:**  
Randomly selects one gene (node) to mutate. Its following nodes are picked randomly from sequentially connected nodes

그림 28 T. Shibata 와 T. Fukuda의 알고리즘

단일 로봇에 대한 학습 방법으로 로봇 상호관계(interaction)를 고려하지 않았으며 숨겨져 있는 환경(hidden environment)과 로봇 상호작용에 의한 숨겨진 상태를 고려하지 않았다.

T. Shibata 와 T. Fukuda[47]는 대기시간을 줄이고 이동경로가 짧으며 각 로봇이 서로 충돌하지 않으면서 최적의 경로를 찾는 방법을 GA(genetic Algorithm)을 통하여 제시하였다. 그림 27는 초기 setup이며 그림 28은 GA에 의한 경로계획 결과이다. T. Shibata 와 T. Fukuda의 알고리즘은 각 노드를 문자형 유전자로 정의 하여 진화하도록 하였으며 evaluation 적합도 함수로

$$\text{Fitness} = \text{Tour Length} + \text{Waiting Time}$$

를 사용하여 다른 로봇과 충돌을 방지하기 위한 진화 전략으로 사용하였다. T. Shibata 와 T. Fukuda의 연구 결과는 로봇 상호간에 통신을 하지 않아도 GA를 이용하면 로봇 상호간에 충돌이 일어나지 않으면서 최적경로를 구할 수 있음을 보였다. 그러나 환경이 알려져 있어야 하며 복잡한 task를 수행 하는 데에는 어려움이 따를 것으로 여겨진다.

이동하는 장애물이 존재하는 경우 동적환경에서의 경로계획은 NP-hard 문제로 단순한 반사행동 경로계획(reactive motion planning)은 수렴성을 보장하지 못한다. Azarm과 Schmidt[48], Carpin과Pagello[49]는 충돌문제를 제거하기 위하여 협상(negotiation) 또는 랜덤 시간지연을 포함하도록 하는 분포(distribute) 해를 구하였지만 최적성을 고려하지 않았으며, LaValle과 Hutchinson[60~52]은 집중연산(centralized computing)에 의한 최적 해를 구하였으나 실시간 재-계획이 불가능 하다.

Y. Guo와 L. Parker[53]는 실외공간에 적용가능하며 실시간으로 re-planning되는 분포최적(distributed optimal) 경로계획법을 제시하였다. Y. Guo와 L. Parker의 경로 계획에서 각 로봇은 할당된 목표를 가지고 있으며 출발위치와 목표위치를 알고 있다. 선-정의(pre-defined)된 지도가 주어지며 다른 로봇과 통신이 가능하다. 로봇 motion-계획문제는 출발위치에서 목표위치까지의 충돌이 없는 최적 경로를 찾는 문제로 평가할수는 다음과 같다.

$$K = \gamma_1 \max(T_1, T_2, \dots, T_N) + \gamma_2 \sum_{i=1}^N I_i$$

여기서,  $T_1, T_2, \dots, T_N$ 은 각 로봇이 목표에 도달하는데 걸리는 위한 시간이며  $I_i$ 는  $i$ 번째 로봇의 idle 시간,  $\gamma_1, \gamma_2$ 는 양의 가중치(weighting)이

다. Y. Guo와 L. Parker의 알고리즘은 로봇 motion-planning을 경로계획(path-planning)과 속도계획으로 나누어 수행하며 다른 로봇이 접근하면 통신모듈을 통한 방송(broadcasting)을 통하여 경고한다. 이 알고리즘 역시 map을 알고 있어야 하며 동적환경의 불확실성을 고려하지 않았다.

#### 4.4 진화로봇[54~55]

진화로봇은 인공진화(artificial evolution)과정을 포함하는 설계에 의하여 로봇이나 로봇 제어 시스템을 개발하려는 시도이다. 서로 다른 염색체를 갖는 초기 개체군에 로봇 제어코드가 확률적으로 생성되도록 하여 환경에 투입하면 로봇은 자유롭게 운동하며 수행성능을 평가하게 된다. 이때 적합도가 높은 로봇은 수행성능 평가가 우수한 로봇의 유전인자를 복제함으로서 생성되며 개체의 수행성능이 만족할 정도가 되도록 여러 세대를 거쳐 진화를 반복한다. 이때 로봇 행동은 로봇 제어시스템이 비교적 고정되고 유전적으로 명기된 메커니즘을 갖는 타고난(innate)행동과 plastic-제어시스템으로부터 생성되는 학습된 행동(learned behavior)으로 구분할 수 있다. 앞에서 서술한 행동기반 시스템의 행동들은 고정되어 있지만 진화로봇에서의 행동은 진화에 의하여 변화하게 된다.

로봇에서의 모터동작은 부분적으로 환경으로부터 받는 센서신호패턴으로 결정되며 정확한 환경 인지를 위하여 센서 감지주기를 빠르게 하면 일반적으로 다른 성능이 저하되게 된다. 환경과의 상호작용과 다른 개체 로봇과의 상호작용을 원활하게 하기 위해서는 공-진화(co-evolution)가 요구된다. 공진화는 개체, 혹은 개체군 간의 협동, 경쟁, 공생 등의 공동목표를 위한 개체 간 공진화(inter-individual co evolution), 개체들과 환경과의 상호작용, 환경변화에 적응하기 위한 공진화로 나눌 수 있다.

Cliff et al.[56~59]과 Harvey et al.[59~60]은 유전형질(genotype)은 제어시스템과 센서-모폴로지(morphology)에 의하여 결정되며 진화된 로봇은 과 전략적 위치에 위치한 두개의 시각 장치의 상관된 행동을 비교하는 것과 같은 단순한 전략에 의존함을 보였다. Ferenc Havasi et al[61]은 2002 EvoNet Summer Schools 강연 노트에서 ① What happens if we add more knowledge in the evolution process? ② Are we reducing capability of evolution of exploring? ③ Are we reducing capability of evolution of generalizing? 에 대한 방법으로 두 형태의 구조를 제시하였다.

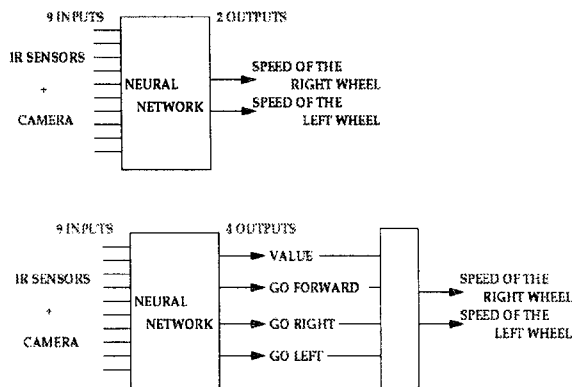


그림 29 Two different architecture

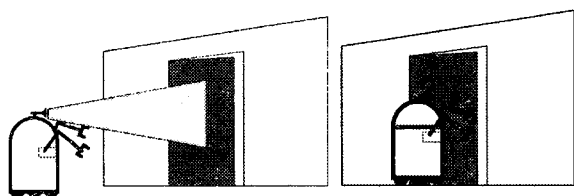


그림 30 Simple Example of State Estimation

(그림 29: 센서-신경망-모테제어신호(그림 위), 센서-신경망-행동명령-변환기-모터제어신호). 실험은 장애물 회피에 대한 진화와 이동하는 물체 따라가기에 대한 진화로 적합도 함수는

$$fitness = high\ speed \times strait\ motion \times low\ sensor\ activation$$

이며 카메라가 켜지면 속도를 줄이고 직진하도록 하였다. 실험 결과로 두 제어기 모두 이동하는 물체를 따라가도록 하는 실험에 대하여 성공적으로 진화하며 두 번째 제어기(그림 아래)가 빨리 진화하며 좋은 수행성을 보였다.

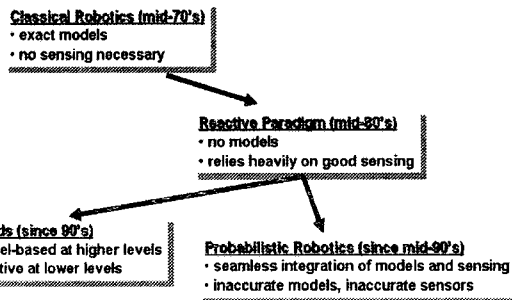


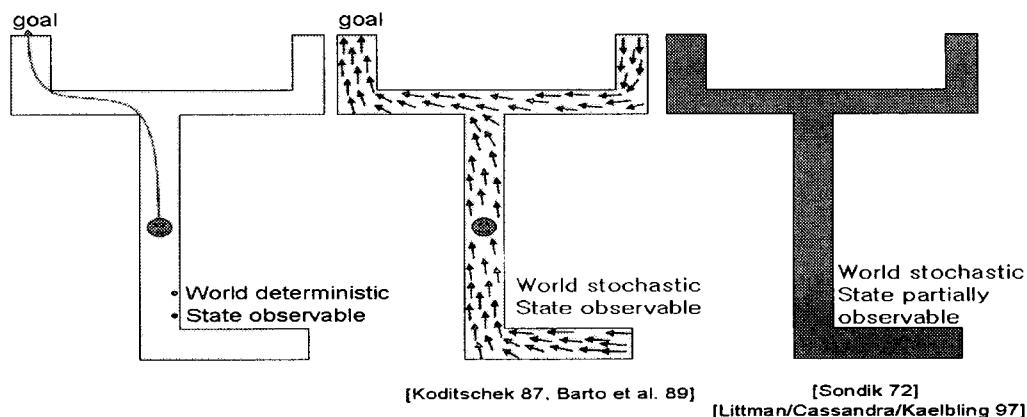
그림 32 로봇연구변화 추의

## V. 연구동향

### 5.1 Probabilistic Robot[62]

로봇은 산업현장과 이야기 속, 영화 스크린을 떠나 우리 주변-오락, 게임, 장난감, 서비스로봇, 수술용 로봇, 가이드 로봇-에 와 있다. 로봇이 작업(task)을 수행하여야하는 공간이 매우 다양해졌으며, 이 공간들은 예측할 수 없고 확률적이다. 로봇의 행동 또한 확률적이 될 수밖에 없으며, 환경과 로봇 모델 또한 부정확하고 완전하지 못하다. 로봇에 부착된 센서도 역시 환경과 노이즈의 영향으로 제한된 성능만을 보인다. 이러한 제약들은 1990년대에 들어서면서 확률 로봇 공학 (PR: probabilistic robotics)이라는 새로운 주제를 가져왔다. PR의 주 아이디어는 환경인식과 로봇

행동은 확률이론에 의하여 표현 가능하며 상태추정(state estimation)에 의한 환경인식, utility 최적화에 의한 행동선택 으로 불확실성이 존재하는 환경에서의 로봇공학 문제를 해결 할 수 있다. 로봇공학에서 확률 이론을 이용하면 부정확한 모델, 센서 노이즈에 의한 불완전한 센서신호에 대하여 강인한 성능을 보이지만 확률을 계산하여야



(a) Classical Optimization

(b) MDP

(c) POMDP

그림 31 Classical Optimization, MDP and POMDP

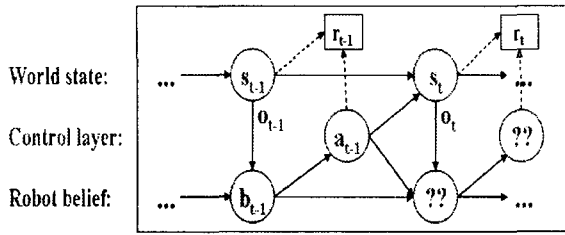


그림 33 POMDP 구조

하는 계산 부하문제와 확률가정에 있어서의 오류, 근사 된 방법이라는 단점을 포함하고 있다.

상태추정에 대한 간단한 예로 그림 30를 고려하자. 로봇의 센서로부터 문까지 거리  $z$ 는 측정으로 얻을 수 있다. 이때  $z$ 에 대하여 문이 open 일 확률  $P(\text{open} | z)$ 는 Bayes 룰로부터 쉽게 구할 수 있다. 즉,  $P(\text{open} | z)$ 는 diagnostic 확률 변수라서 쉽게 구할 수 없지만  $P(z | \text{open})$ 은 causal인 확률로 쉽게 구하여진다. Bayes 룰은 쉽게 확장이 가능하며 로봇이 다른 변수들  $z_1, z_2, \dots, z_n$ 를 측정 한 경우, 확률은 다음과 같이 구할 수 있다.

$$P(x | z_1, z_2, \dots, z_n) = \frac{P(z_n | x, z_1, z_2, \dots, z_{n-1})P(x | z_1, z_2, \dots, z_{n-1})}{P(z_n | z_1, z_2, \dots, z_{n-1})}$$

로봇 action은 확실성(certainty)에 의한 것이 아니며 일련의 action이 진행됨에 따라 불확실성(uncertainty)은 커지게 된다. 따라서 로봇의 다음 행동은 불확정적이며 Task를 수행하기 위한 최적의 정책을 구할 필요가 있다. 예를 들어 그림 31에서 로봇의 다음 행동을 고려하여보자. 그림에서 (a)는 비-확률적이며 상태를 예측할 수 있는 경우이므로 고전적인 최적화에 의해 최적-경로를 찾을 수 있다. (b)는 확률 world이지만 상태를 예측할 수 있으므로 목표를 행하는 경로를 찾을 수 있다. 그러나 (c)의 경우는 확률적이며 상태는 부분적으로 예측 가능한 경우로 목표를 찾는다는 보장을 할 수 없다. (b)의 경우에는 MDP(markov decision process), (c)는 POMDP(partially observable MDP)라 하며 이들을 이용하는 연구 결과들이 발표되었다.[64, 65] POMDP의 기본 구조는 그림 33과 같으며 task는 로봇이 action  $a_t$ 를 행하고 난 이후의 상태 결정과 다음 action을 결정하는 정책-계산이다. 정책-계산과정은 매우 복잡하며 이를 개선하기 위한 방법으로 S. Thrun et al.은 계층구조를 갖는(hierarchical)

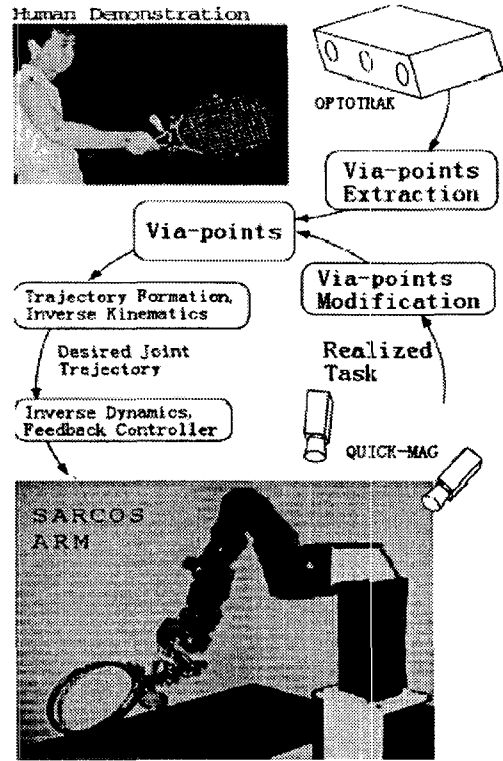


그림 34 테니스 서브동작

POMDP 제시하였다. (<http://robots.stanford.edu/index.html>).

### 5.2 Imitation Learning

모방학습(imitation learning)은 보는 것에 의한 학습(learning by watching), 시범학습(learning from demonstration)이라 할 수 있으며 행동조교(Mentor, teacher)가 시범을 보임으로서 학습 속도를 높이고 강화학습을 개선할 수 있다. 모방학습은 인간형 로봇(humanoid robot)과 복잡한 동작에 대한 1회 시범학습에 있어서 아주 우수한 학습효과를 보인다. Schaal[66]과 Schaal et al.[67]은 그림 35의 모방학습 스키마를 제시하였다. 그림의 오른쪽은 기본지각인자와 시각정보가 공간정보와 객체인식정보로 변환되는 과정을 보여주며, 왼쪽은 모터인자로 각 모터의 명령은 각 모터에 대응되는 primitive로부터 입력된다. 학습은 MP: movement primitive와 모터명령 생성기를 조정한다. 이 구조에서의 모방 학습은 시범동작 인식(perception)-공간 변환에 의한 물체 인식 및 공간정보 추출-mapping MP-primitive 조정 과정을 통하여 학습한다. Schaal et al.[68]에서는 supervised 모터학습 알고리즘을 제시하였다.(그림 34). Miyamoto와 Kawato[69]는 스플라인 함수를 이용한 모방학습 알고리즘을 이용한

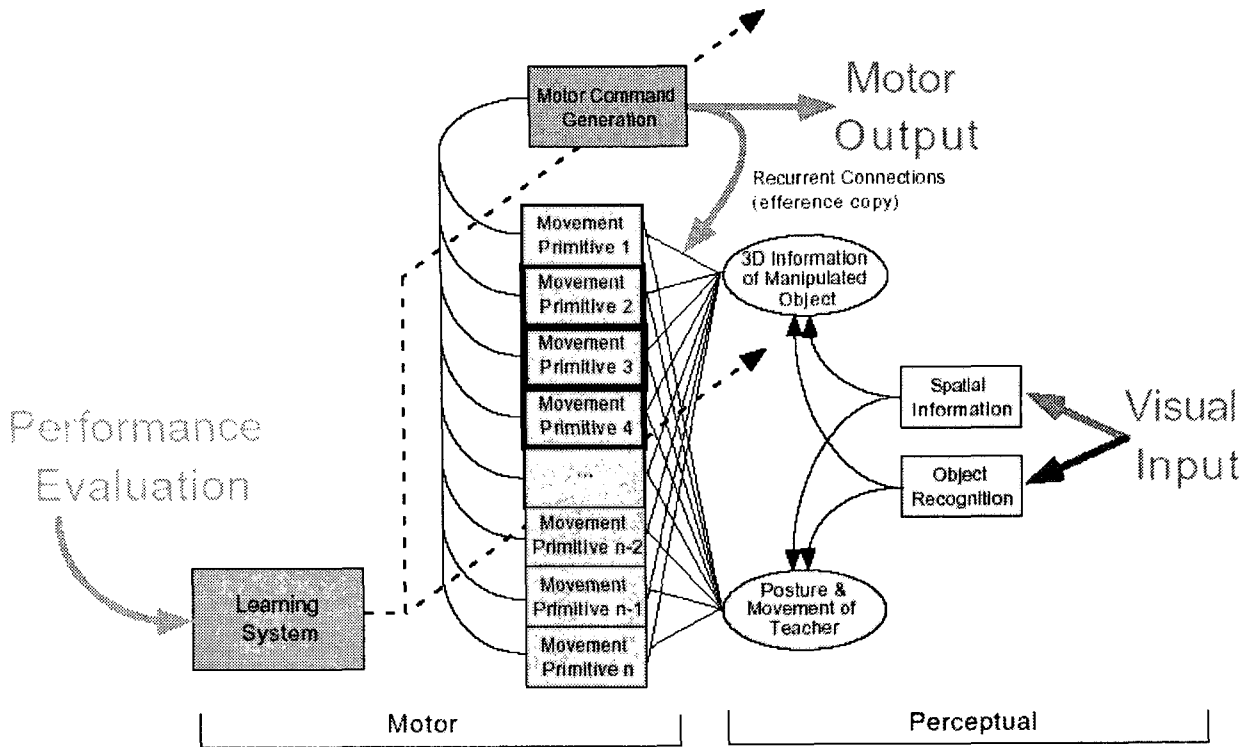


그림 35 모방학습 스키마

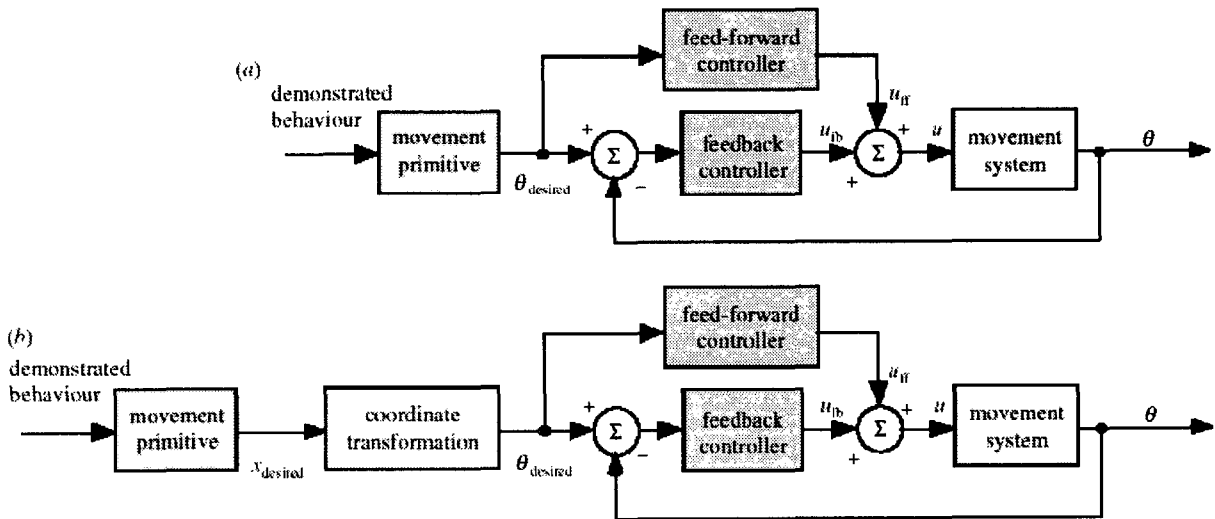


그림 36 supervised 모터 학습 알고리즘

경유 점(via-point)에 대한 학습으로부터 로봇 매니플레이터의 테니스 서브동작과 업스윙동작 학습을 보였다. Schaal et al.[70]은 휴머노이드 로봇의 테니스 동작 학습과 연결 동작학습 결과를 보였다.

### 5.3 Social Robot

21C 로봇은 어디에서도 만날 수 있게 될 것이다. 로봇의 역할이 확대되어 청소와 같은 가사 일을 도우고 오락, 게임 등의 일을 할 수게 되었다. 사람은 로봇을 이용하여 정밀한 수술을 하고

로봇과 경쟁하여 게임을 즐기며 관광지나 도서관 등의 공공장소에서 로봇의 안내를 받게 될 것이다. 사람은 사회적 상호작용(social interaction)에 전문가이지만 로봇은 상호작용에 익숙하지 못하다. 로봇은 사람과 상호작용이 가능한, 통신(사람과 이야기 할 수 있는)이 가능한 형태로 설계되어야 한다. 백과사전[72]에 따르면 사회적 로봇은

A social robot is an autonomous robot that interacts and communicates with humans by following the social rules attached to its role.

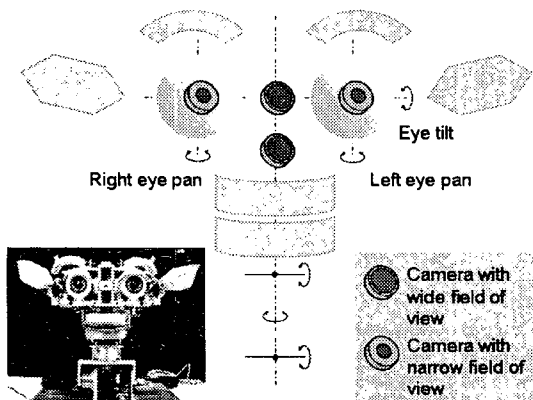


그림 37 Kismet

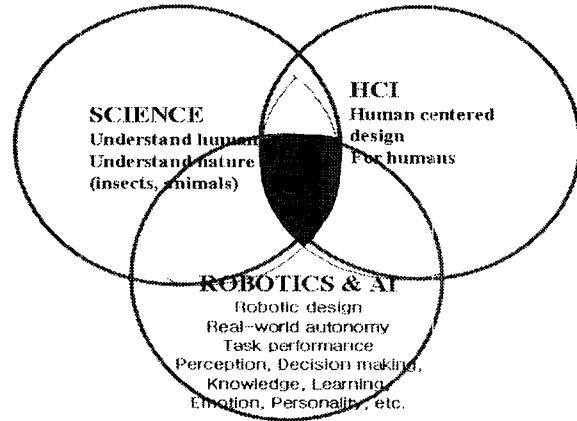


그림 38 Future Robotics

으로 정의할 수 있다. 로봇은 기구나 장치가 아닌, 스스로 지능을 가지고 있으면서 지식-지능을 높일 수 있어야 하며 스스로 동기부여하며 (self-motivated), 자발적으로 행동하는(pro-active) 생명체와 같은 것이라야 한다. 사회적 생명체인 로봇 설계에서 고려하여야 하는 사항들은 용도를 고려한 형태, 외관, 인간성 등이 고려되어야 한다.

Kismet[73,74]은 유아 수준의 사회적 상호작용을 목적으로 설계된 로봇이다. 구조는 그림 37과 같다. 간단한 감성표현과 움직이는 물체를 응시할 수 있는 정도의 지능이 구현되어 있다.

## VI. 결론

21C는 자연과학과 로봇공학, 인공지능, HCI (human computer inter face)등이 복합된 로봇 연구가 진행 될 것이다. 사람을 보다 더 이해하고 동물, 곤충들의 행동을 이해함으로써 보다 인간과 친숙한 로봇을 구현 할 수 있을 것이다. 인간과 로봇과의 상호작용과, 로봇과 로봇의 상호작용, 더 지능적이고 더 사교적인 로봇을 구현하기 위해서는 자연에 대한 이해, 사람과 로봇의 상호작용, 의사소통에 대한 구현방법, 많은 로봇이 공존하기 위한 진화능력 등이 부여되어야 한다.

## Reference

1. Braitenberg, V. (1984). *Vehicles: Experiments in Synthetic Psychology*, The MIT Press, Cambridge MA.
2. Hogg, D. W., Martin, F. & Resnick, M. (1991). *Braitenberg Creatures*, MIT Media Laboratory .
3. Mobus, G.E., "Toward a theory of learning and representing causal inferences in neural networks", in Levine, D.S. and Aparicio, M (Eds.),

- Neural Networks for Knowledge Representation and Inference, Lawrence Erlbaum Associates, 1994.
4. Mobus, G.E. and Paul S. Fisher, "Foraging Search at the Edge of Chaos", Presented at Metroplex Institute of Neurodynamics Conference on Oscillations in Neural Networks, May, 1994.
5. Mobus, G.E. and Paul S. Fisher, "MAVRIC's Brain", Presented at IEA/AIE-94. 1994.
6. Pfeifer, R. & Scheier, C. (1999). *Understanding Intelligence*, The MIT Press, Cambridge MA.
7. R. Brooks. "A Robust Layered Control System for a Mobile Robot." IEEE Journal of Robotics and Automation. RA 2, pp 14 23, April 1986.
8. Brooks, R.A., "How to build complete creatures rather than isolated cognitive simulators," in K. VanLehn (ed.), *Architectures for Intelligence*, pp. 225-239, Lawrence Erlbaum Associates, Hillsdale, NJ, 1991.
9. R. C. Arkin, *Behavior Based Robotics*, Cambridge, MA: MIT Press, 1998.
10. Arkin, Ronald C. Towards the Unification of Navigational Planning and Reactive Control. Working Notes of the AAAI Spring Symposium on Robot Navigation. Stanford University, March 28-30, 1989.
11. R. C. Arkin and D. MacKenzie, "Planning to Behave: A Hybrid Deliberative/Reactive Control Architecture for Mobile Manipulation", 1994 international symposium on Robotics and Manufacturing, Maui, August 1994.
12. E. Sorouchyari, "Mobile Robot Navigation : A neural network approach,", Proceedings Art Coll.Neuro., Eco. Poly., Lausanne, 1989, p. 159-175.
13. M. J. Mataric. Integration of representation into goal driven behavior based robots. IEEE Transactions on Robotics and Automation,



- 8(3):304 - 312, June 1992.
14. R. C. Atkin, G. W. King, D. L. Westbrook, B. Heeringa, A. Hannon, and P. Cohen. Spt: Hierarchical agent control: a framework for defining agent behavior. In Proc., Intl. Conf. on Autonomous Agents, pages 425 - 432, May 2001.
  15. Kian Hsiang Low, Wee Kheng Leow, and Marcelo H. Ang Jr., "A Hybrid Mobile Robot Architecture with Integrated Planning and Control", Proc. 1st AAMAS'02 , pages 219-226, July 15-19, 2002, Bologna, Italy.
  16. Bonasso, R. Peter and David Kortenkamp, Using a Layered Control Architecture to Alleviate Planning with Incomplete Information. 1996.
  17. <http://www.cc.gatech.edu/ai/robot-lab/online-publications/jetai-final.pdf>
  18. K. Konolige, K. L. Myers, E. H. Ruspini, and A. Saffiotti " The Saphira Architecture: A Design for Autonomy", Journal of Experimental and Theoretical Artificial Intelligence 9(1), 1997, pp. 215-235 .
  19. K. Konolige and K. L. Myers , "The Saphira Architecture for autonomous mobile robots", In AI-based Mobile Robots: Case studies of successful robot systems (D. Kortenkamp, R. P. Bonasso, and R. Murphy, eds.), MIT Press, 1996.
  20. R. Peter Bonasso, David Kortenkamp, and Robin Murphy, "Mobile Robots: A Proving Ground for Artificial Intelligence" in Mobile Robots and Artificial Intelligence,, ed. David Kortenkamp, R. Peter Bonasso, and Robin Murphy, AAAI Press, 1998.
  21. R. P. Bonasso, R. J. Firby, E. Gat, D. K. D. Miller, and M. Slack. Experiences with an architecture for intelligent, reactive systems. Journal of Experimental and Theoretical Artificial Intelligence, 9(2 - 3):237 - 256, 1997.
  22. Webb, B., & Scutt, T. (2000) A simple latency dependent spiking neuron model of cricket phonotaxis. Biological Cybernetics 82(3):247-269
  23. Webb, B. & Harrison, R. (2000) Integrating sensorimotor systems in a robot model of cricket behaviour, Sensor Fusion And Decentralized Control In Robotic Systems III November 5-8, Boston. (Editors: McKEE GT, Schenker PS) Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE), 4196, 113-124. 2000.
  24. Reeve, R. & Webb, B. New neural networks for robot phonotaxis. (submitted to special WGW edition of) Philosophical Transactions of the Royal Society, (2002).
  25. Dimitrios Lambrinos, Ralf Möller, Thomas Labhart, Rolf Pfeifer& Rüdiger Wehner " A mobile robot employing insect strategies for navigation" Robotics and Autonomous Systems, special issue on Biomimetic Robots, Vol. 30, 2000, 39-64
  26. Dimitrios Lambrinos, Ralf Möller, Rolf Pfeifer& Rüdiger Wehner, Landmark Navigation without Snapshots: the Average Landmark Vector Model, Proc. 26th Göttingen Neurobiology Conference (N. Elsner; R. Wehner, eds.), no. 30a, Thieme Verlag, 1998.
  27. Ralf Möller "Insect Visual Homing Strategies in a Robot with Analog Processing", Biological Cybernetics, special issue: Navigation in Biological and Artificial Systems, Vol. 83, No. 3, 2000, 231-243
  28. Achim Lilienthal and Tom Duckett, Creating Gas Concentration Gridmaps with a Mobile Robot Proc. IROS-2003, IEEE/RSJ Int. Conference on Intelligent Robots and Systems, Las Vegas, October 27-31, 2003.
  29. Achim Lilienthal and Tom Duckett, Experimental Analysis of Smelling Braitenberg Vehicles Proc. ICAR-2003, IEEE International Conference on Advanced Robotics, Coimbra, Portugal, June 30-July 3, 2003.
  30. Cham, J. G., Bailey, S. A., Cutkosky, M. R., "Robust Dynamic Locomotion Through Feedforward-Preflex Interaction" (PDF Format) ASME IMECE Proceedings, Orlando, Florida, November 5-10, 2000.
  31. Bailey, S. A., Cham, J. G., Cutkosky, M. R., Full, R. J. "Comparing the Locomotion Dynamics of a Cockroach and a Shape Deposition Manufactured Biomimetic Hexapod" International Symposium on Experimental Robotics (ISER2000), Honolulu, HI, December 10-13, 2000.
  32. Cham, J. G., Bailey, S. A., Cutkosky, M. R., "Robust Dynamic Locomotion Through Feedforward-Preflex Interaction" (PDF Format) ASME IMECE Proceedings, Orlando, Florida, November 5-10, 2000.
  33. Xu, X., Cheng, W., Dudek, D., Hatanaka, M., Cutkosky, M. R. and Robert J. Full, "Material Modeling for Shape Deposition Manufacturing of Biomimetic Components." ASME Proceedings, DETC/DFM 2000, Baltimore, Maryland, September 10-14, 2000.
  34. Leslie Pack Kaelbling, Michael L. Littman, and Andrew W. Moore, "Reinforcement Learning: A Survey", Journal of Artificial Intelligence Research 4 (1996) 237-285 1996.
  35. M. N. Nicolescu and M. J. Matarić. Learning and interacting in human robot domains. In C. C. White and K. Dautenhahn, editors. IEEE

- Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, Special Issue on Socially Intelligent Agents The Human in the Loop, volume 31, pages 419-430. IEEE, September 2001.
36. Ronald C. Arkin, Yoichiro Endo, Brian Lee, Douglas C. MacKenzie, and Eric Martinson; A Reactive Robot Architecture with Planning on Demand; Ananth Ranganathan and Sven Koenig; 2003.
  37. J. Brian Lee, Maxim Likhachev, and Ronald C. Arkin Selection of Behavioral Parameters: Integration of Discontinuous Switching via Case-Based Reasoning with Continuous Adaptation via Learning Momentum; ; 2001.
  38. Eric Martinson, Alexander Stoytchev, and Ronald C. Arkin; Robot Behavioral Selection Using Q-learning; 2001
  39. Richard S. Sutton and Andrew G. Reinforcement Learning: An Introduction Barto MIT Press , Cambridge, MA, 1998.
  40. L. Kaelbling, M. Littman, and A. Moore. Reinforcement learning: A survey. Journal of Artificial Intelligence Research, 4:237-285, 1996.
  41. W. D. Smart and L. P. Kaelbling. Effective reinforcement learning for mobile robots. In Proceedings of the International Conference on Robotics and Automation (ICRA- 2002), volume 4, pages 3404-3410, 2002.
  42. R. Sutton and A. Barto. Reinforcement learning: an introduction. Adaptive computation and machine learning. MIT Press, Cambridge, Mass., 1998.
  43. E. Bonabeau, M. Dorigo, and G. Theraulaz, "Swarm Intelligence: From Natural to Artificial Systems", Santa Fe Studies in the Sciences of Complexity, Oxford University Press, 1999.
  44. M. J. Mataric. Distributed approaches to behavior control. In SPIE - Sensor Fusion V, vol. 1828, 373-382, 1992.
  45. M. Mataric. Interaction and Intelligent Behavior. MIT AI Lab Technical Report, AI-TR-1495, August 1994.
  46. M. Mataric. Reward functions for accelerated learning. In Proceedings of the Int. Machine Learning Conf., 181-189, 1994.
  47. T. Fukuda and T. Shibata, "Hierarchical Intelligent Control For Robotic Motion By Using Fuzzy, Artificial Intelligence, and Neural Network," IEEE Technology Update Series, Fuzzy Logic Technology and Applications, Technical Activities, pp.129-134 (1994).
  48. Azarm, K. and Schmidt, G.: A Decentralized Approach for the Conflict-Free Motion of Multiple Mobile Robots . - In: Proceedings of the IEEE/RSJ International Conference on Robots and Systems, Osaka, 1996, Vol. 3, page 1667-1675.
  49. S. Carpin, E. Pagello. A Distributed Algorithm for Multi-robot Motion Planning , Proceedings of the Fourth European Workshop on Advanced Mobile Robots (EUROBOT01), 207-214, Lund (Sweden), September 2001.
  50. R. Sharma, S. M. LaValle, and S. A. Hutchinson. Optimizing robot motion strategies for assembly with stochastic models of the assembly process. IEEE Trans. on Robotics and Automation, 12(2):160--174, April 1996.
  51. S. M. LaValle and S. A. Hutchinson. Optimal motion planning for multiple robots having independent goals. IEEE Trans. on Robotics and Automation, 14(6):912--925, December 1998.
  52. S. M. LaValle and S. A. Hutchinson. An objective-based framework for motion planning under sensing and control uncertainties. International Journal of Robotics Research, 17(1):19--42, January 1998.
  53. Y. Guo and L. E. Parker, "A Distributed and Optimal Motion Planning Approach for Multiple Mobile Robots", Proceedings of IEEE International Conference on Robotics and Automation, 2002.
  54. Nolfi, S. and Floreano, D. Evolutionary robotics, MIT Press, 2000.
  55. Mattias Wahde, Evolutionary Robotics, The Use of Artificial Evolution in Robotics A tutorial presented at IROS 2004, 2004.
  56. Cliff, D., Harvey, I., and Husbands, P. Explorations in Evolutionary Robotics, Adaptive Behavior 2:1, pp. 73-110, 1993.
  57. Cliff, D., Husbands, P., and Harvey, I. Evolving visually guided robots In: Proc. of the 2nd Int. Conf. on Simulation of Adaptive Behavior (SAB92), MIT Press/Bradford books pp 374-383, 1993.
  58. Cliff, D. and Miller, G.F. Co-evolution of pursuit and evasion II: Simulation methods and results, in Proc. of the 4th Int. Conf. on the Simulation of Adaptive Behavior (SAB96), MIT Press, pp. 506-515, 1996.
  59. Harvey, I., Husbands, P., and Cliff, D. Issues in Evolutionary Robotics, In: From Animals to Animals II: Proc. of the 2nd Int. Conf. on Simulation of Adaptive Behavior (SAB92), pp. 364-373, 1992.
  60. Harvey, I., Husbands, P., Cliff, D., Thompson, A., and Jakobi, N. Evolutionary robotics: the Sussex approach, Robotics and Autonomous Systems, 20,

- No 2-4, pp. 205-224, 1997.
61. Vincenzo Giordano, Ferenc Havasi, Michael Schwarz, Gregory Valigiani, Stefan Wiegand, Evolutionary robotics, 5th EvoNet Summer Schools Szeged, Hungary ,2002.
  62. S. Thrun. Probabilistic algorithms in robotics. *AI Magazine*, 21(4):93-109, 2000.
  63. S. Thrun, M. Beetz, M. Bennewitz, W. Burgard, A.B. Cremers, F. Dellaert, D. Fox, D. Hähnel, C. Rosenberg, N. Roy, J. Schulte, and D. Schulz. Probabilistic algorithms and the interactive museum tour-guide robot minerva. *International Journal of Robotics Research*, 19(11):972-999, 2000.
  64. J. Pineau, G. Gordon, and S. Thrun. Point-based value iteration: An anytime algorithm for POMDPs. *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI) 2003*
  65. N. Roy, G. Gordon, and S. Thrun. Finding Approximate POMDP Solutions Through Belief Compression. Submitted to *JAIR*, 2004.
  66. S. Schaal. Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences*, 3(6):233-242, 1999.
  67. S. Schaal. The sl simulation and real-time control software package. Technical Report Computer Science Tech Report, <http://wwwslab.usc.edu/publications/schaal-TRSL>, University of Southern California, 2001.
  68. Aude Billard, Yann Epars, Gordon Cheng and Stefan Schaal Discovering Imitation Strategies through Categorization of Multi-Dimensional Data *Proceedings of the 2003 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems Las Vegas, Nevada · October 2003.*
  69. Miyamoto, H., & Kawato, M. (1998). A tennis serve and upswing learning robot based on dynamic optimization theory. *Neural Networks*, 11(7-8), 1331-1344.
  70. Miyamoto, H., Schaal, S., Gandolfo, F., Gomi, H., Koike, Y., Osu, R., Nakano, E., Wada, Y., & Kawato, M. (1996). A Kendama learning robot based on dynamic optimization theory. *Neural Networks*, 9(8), 1281-1302.
  71. Morimoto, J., & Doya, K. (1998). Reinforcement learning of dynamic motor sequence: learning to stand up. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 3, 1721-1726.
  72. <http://encyclopedia.thefreedictionary.com/Social%20robot>
  73. Breazeal(Ferrell), C. and Scassellati, B. (2000). "Infant-like Social Interactions Between a Robot and a Human Caretaker". To appear in *Special issue of Adaptive Behavior on Simulation Models of Social Agents*, guest editor Kerstin Dautenhahn.
  74. Breazeal(Ferrell), C. and Velasquez, J. (1998), "Toward Teaching a Robot 'Infant' using Emotive Communication Acts". In *Proceedings of 1998 Simulation of Adaptive Behavior, workshop on Socially Situated Intelligence*, Zurich Switzerland. 25-40
- 로봇 제어구조 관련 홈페이지**  
[www3.humnet.unipi.it/cybernetics/docs/](http://www3.humnet.unipi.it/cybernetics/docs/)  
<http://users.aber.ac.uk/jnw/CS364/9.php>  
<http://www.ifi.unizh.ch/groups/ailab/people/lambri/mitbook/braitenberg/braitenberg.html>  
<http://faculty.washington.edu/gmobus/AdaptiveAgents/MAVRIC-EBA.html>  
<http://www.ai.mit.edu/people/brooks/index.shtml>  
<http://www.cc.gatech.edu/ai/robot-lab/>
- Biorobotics 관련 홈페이지**  
<http://homepages.inf.ed.ac.uk/bwebb/cricket/main.html>  
<http://www.daimi.au.dk/~hhl/cricket.html>  
<http://www.ifi.unizh.ch/groups/ailab/projects/sahabot/>  
[http://www-ra.informatik.uni-tuebingen.de/forschung/robonose/welcome\\_e.html](http://www-ra.informatik.uni-tuebingen.de/forschung/robonose/welcome_e.html)  
[http://aass.oru.se/Research/Learning/proj\\_nbs.html](http://aass.oru.se/Research/Learning/proj_nbs.html)  
<http://www-cdr.stanford.edu/biomimetics/>
- RL:**  
<http://www.cs.ualberta.ca/~sutton/book/the-book.html>  
<http://evonet.lri.fr/summerschool2002/>