

# (시스템구성적)인공지능과 마음

(System-Constructive) Artificial Intelligence and Mind



영동대학교 컴퓨터공학

박충식

(leciel@youngdong.ac.kr)

# 목차

---

- 들어가면서
- 인공지능과 구성주의
- 시스템 구성적 인공지능의 개념
- 시스템 구성적 인공지능의 정보처리
- 시스템 구성적 인공지능에서의 마음
- 시스템 구성적 인공지능과 학제적 마음연구
- 마치면서
- 참고문헌

# 들어가면서

---

- 인공지능
  - 기계를 인간과 같은 지능을 가진 존재로 만드는 것
- 지능적 기계의 꿈
  - 그리스 신화 - 헤파이스토스(불칸)
  - 아랍, 중국, 이집트 신화
  - 로봇 - R.U.R.(카를 차פק)
- 컴퓨터의 등장...그러나
  - 수치계산과는 다르다.
  - 좀 더 빠른 컴퓨터
  - 좀 더 많은 지식

# 인공지능과 구성주의

---

- 인공지능의 탄생
  - 1956년, 다트머스대학
  - 맥카시, 민스키, 뉴웰, 사이먼, 등
- 신경망 이론 vs. 기호처리 이론
- 인공지능 초기 프로그램
  - LT – 프린키피아 매스매티가 논리증명
  - Bacon – 케플러 법칙
  - Toy Problem
  - GPS

# Tip: 인공지능 연구분야

---

- Search and Problem Solving
- Knowledge Representation & Reasoning
- Learning
- Commonsense Reasoning
  
- Speech Recognition
- Vision
- NLP
- Game
- Planning
- Robotics

# 인공지능과 구성주의

---

- 어려운 프로그램 - 인간에겐 쉬운
  - 자연언어처리 - 언어번역, 대화
  - 화상인식
- 행위적 인지(비표상 방법, 층위적 구조, 포섭 구조, 행동기반) 방법
  - 신경망 아니고, 기호도 아닌
  - 로봇제어에 효과적
  - 화상, 자연언어 까지는 아닌
  - 생태학적 접근
  - 반데카르트적, 심신일원론적, 물리주의적
  - 구성주의적

# Tip: 행위적 인지

---

## □ 행동기반 로봇의 기본

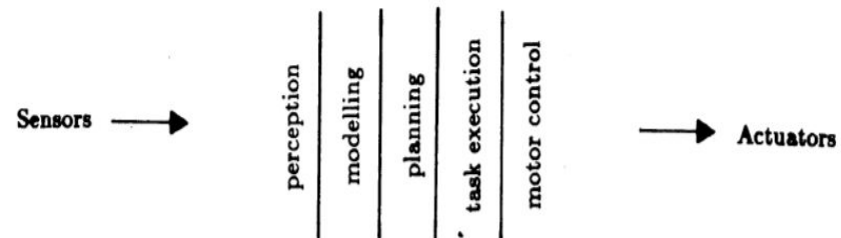
- Situatedness
- Embodiment
- Intelligence
- Emergence

## □ 포섭 구조

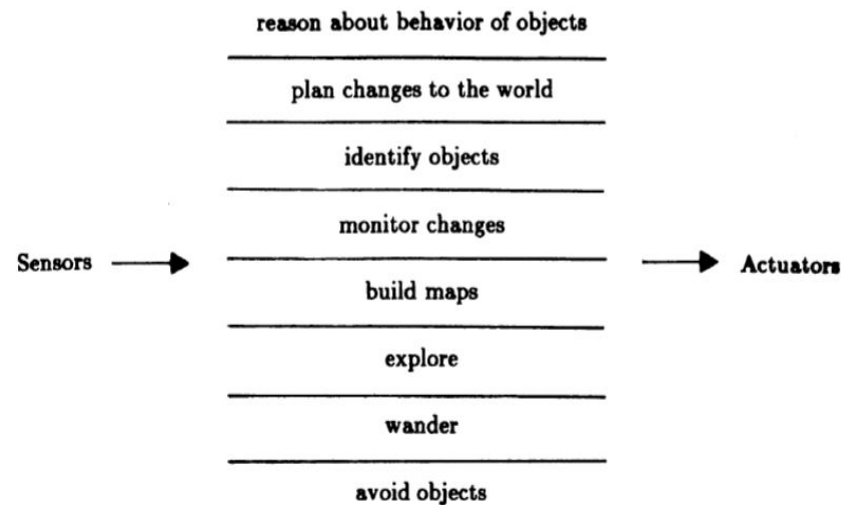
- 행동들의 계층적 구조
- 각 계층은 단위 행동(AFSM)들의 집합으로 구성
- 여러 층이 모여 과제 수행
- 각 층은 독립적인 목표(Goal)을 가짐
- 각 층의 행동들은 병렬적으로 수행됨
- 하위 층은 상위 층에 의해 제어됨(하위층 포섭)

# Tip: 행위적 인지

- Data fusion



- Layered Architecture





# 인공지능과 구성주의

---

- 구성주의
  - 베이트슨, 켈리, 피스터, 파스크, 마투라나, 바렐라
  - 삐아제, 웨스켈, 글라저스펠트, 슈미트, 루만, 레이코프
  - 소크라테스, 버클리, 칸트, 비코, 메를로퐁티
- 최근 구성주의
  - 생물학, 심리학, 컴퓨터과학, 시스템이론
  - 뇌과학, 복잡계이론
- 구성주의적 담론과의 연결
- 행위적 인지과학(Enactive Cognitive Science)
- 급진적 구성주의 - 상대주의인가?

# Tip: 구성주의 이론들

---

- 버트란피의 일반시스템이론(General System Theory)
- 삐아제의 스키마이론
- 글라저스펠트의 급진적 구성주의(Radical Constructivism)
- 슈미트의 경험구성적 문예학
- 피스터의 2차 사이버네틱스 이론
- 마투라나와 발레라의 오토포이에시스(Autopoiesis)
- 켈리의 PCP(Personal Construct Psychology)
- 파스크의 대화이론
- 루만의 커뮤니케이션 이론
- 라코프의 은유이론
- 요아스의 사회학 구성이론
- 웨스켈의 움벨트(Umwelt)이론
- 마투라나와 바렐라의 오토포이에시스, 신경현상학
- 투르친의 메타시스템전이 이론(Meta-System Transition Theory)
- 호프스태터의 유동개념(Fluid Concept)

# 인공지능과 구성주의

---

## □ 구성주의의 내용

- 인식의 과정
- 인지는 생물학적 현상 - 신경시스템의 변화로만 간여
- 관찰자의 구성 - 생물학적, 인지적, 문화적 조건하에서
- 인지체계는 조작적 폐쇄 또는 조작적 재귀지시성
- 진위문제 -> 정합성문제
- 생물학적 현상 -> 심신일원론
- 객관적 진리의 인식 불가능성
- 가치는 인간 삶의 유용성 - 실용적 노선 추구
- 행위와 인지에 대한 책임

# 시스템구성적 인공지능의 개념

---

- 시스템구성적 인공지능
- 구성주의적 관점의 지능
  - 불확실한 환경에 대한 개체의 적응능력의 일종
  - 자연선택과 자기조직화
  - 인간의 생존전략 - 지능
- 지능, 마음
  - 지능 - 추론, 계획, 문제해결, 추상적 생각, 이해, 언어 사용, 학습등 - 마음의 속성
  - 마음 - 생각, 인지, 기억, 감정, 의지 - 지각, 사유, 판단, 통제 - 육신의 상대개념
  - 만들수 있는 정도의 마음에 대한 이해

# 시스템구성적인공지능의 개념

---

## □ 시스템구성적 인공지능

- 지능 = 총체적 주재자
- 지능 = 마음
- 심신일원

## □ 마음

- 재료 - 뉴론, 세포, 분자, 물질, ..
- 중요기관 - 두뇌
- 분산처리 모듈
- 확장된 지능, 마음 개념

# 시스템구성적인공지능의 개념

---

- 시스템구성적 인공지능의 관점
  - 정서, 의지, 자아, 의식, 자유, 등
  - 진리, 윤리, 예술, 종교, 등
  - 구성주의 - 작업가설
  - 지능에 대한 이해
  - 삶에 대한 이해

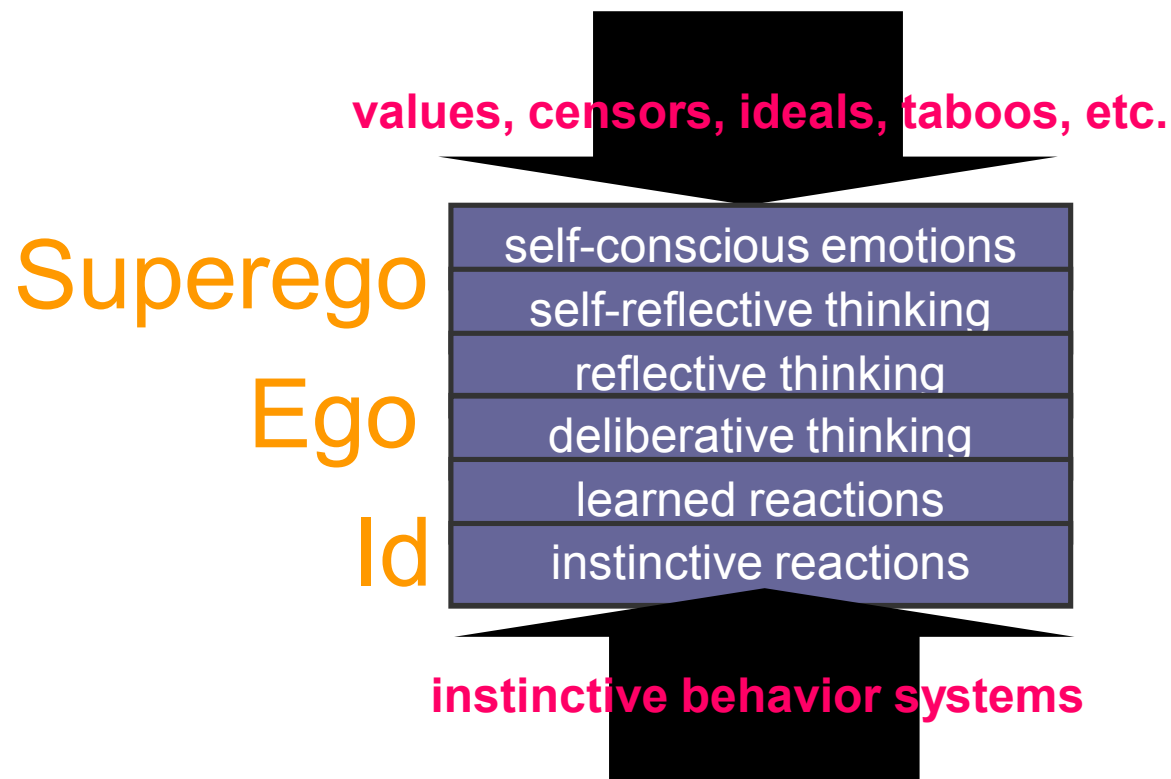
# 시스템구성적 인공지능의 정보처리

---

- 지능-마음 - 정보처리시스템
  - 계산주의적 마음연구
- 시스템과학적 접근
  - 환경과의 구분
  - 항상성 유지
  - 적응의 결과 - 인지
  - 상호작용 - 의미생성
  - 체화된 인지이론, 심볼 그라운드링 문제
- 은유적 사고
  - 몸에 기반한 인지구조
  - 은유 이론, 유동이론

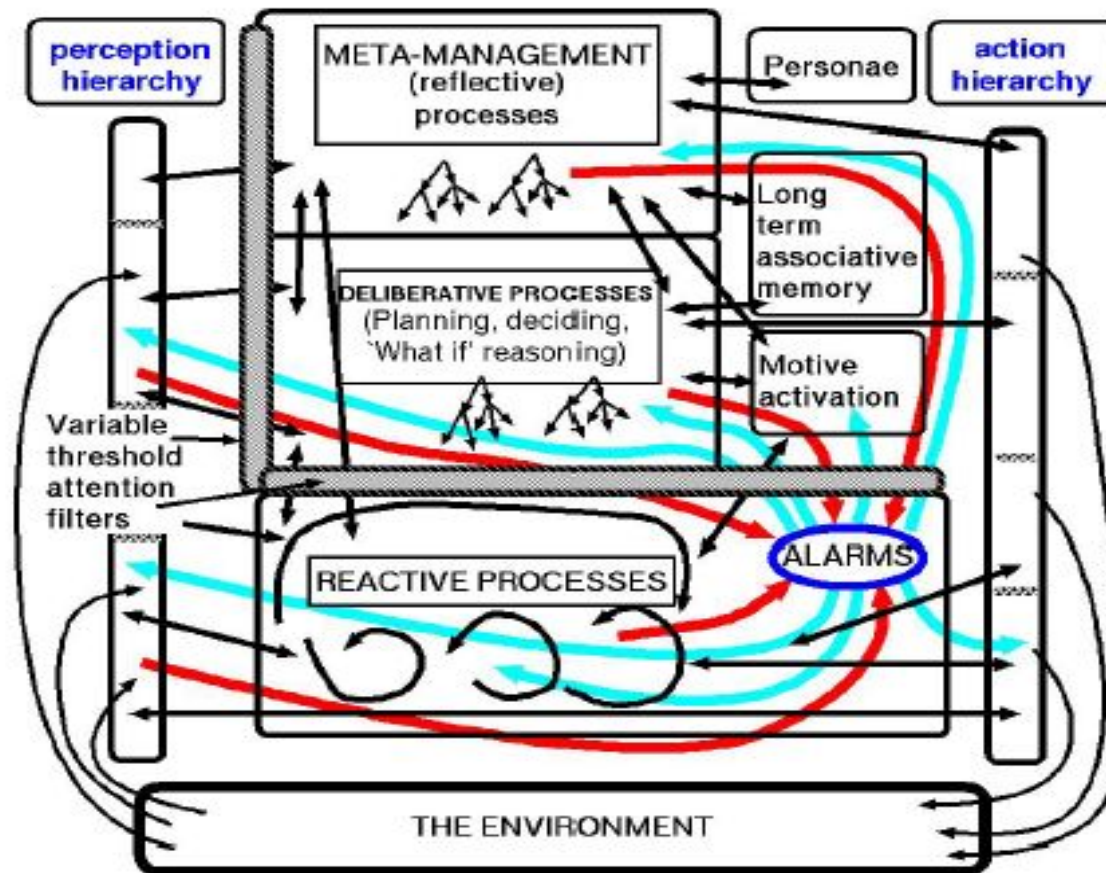
# Tip: Minsky – Emotion Machine

---





# Tip: Sloman - H-CogAff Architecture



# 시스템구성적 인공지능의 마음

---

- 실체는 없고, 모델링만 존재
- 감각-운동 능력의 추상화, 재추상화 - 언어
- 관찰자 자신의 추상화 - 의식, 자유의지
- 의식 - 통합적 의사결정을 위한 추상적 수준의 정보처리, 재추상화된 환경시뮬레이션
- 자의식 - 자기 시뮬레이션
- 자유의지 - 의사결정을 내리는 의식적 정보처리 과정에 자신을 포함
- 정서, 감정 - 자극에 대한 통합적 평가
- 욕구, 동기 - 연속적인 목표 생성 시스템

# 시스템구성적 인공지능의 학제적 마음연구

---

## □ 계산인문사회학

- 언어 - 사회학습적 기제
- 의식, 자아 - 사회 현상속의 역할 모델링
- 정서, 동기 - 사회적 맥락속에서 역동적 생성

## □ 시스템구성적인공지능과의 학제연구

- 구성적 성격
- 감성적 성격
- 컴퓨터 시뮬레이션

# 시스템구성적인공지능의 학제적 마음연구

---

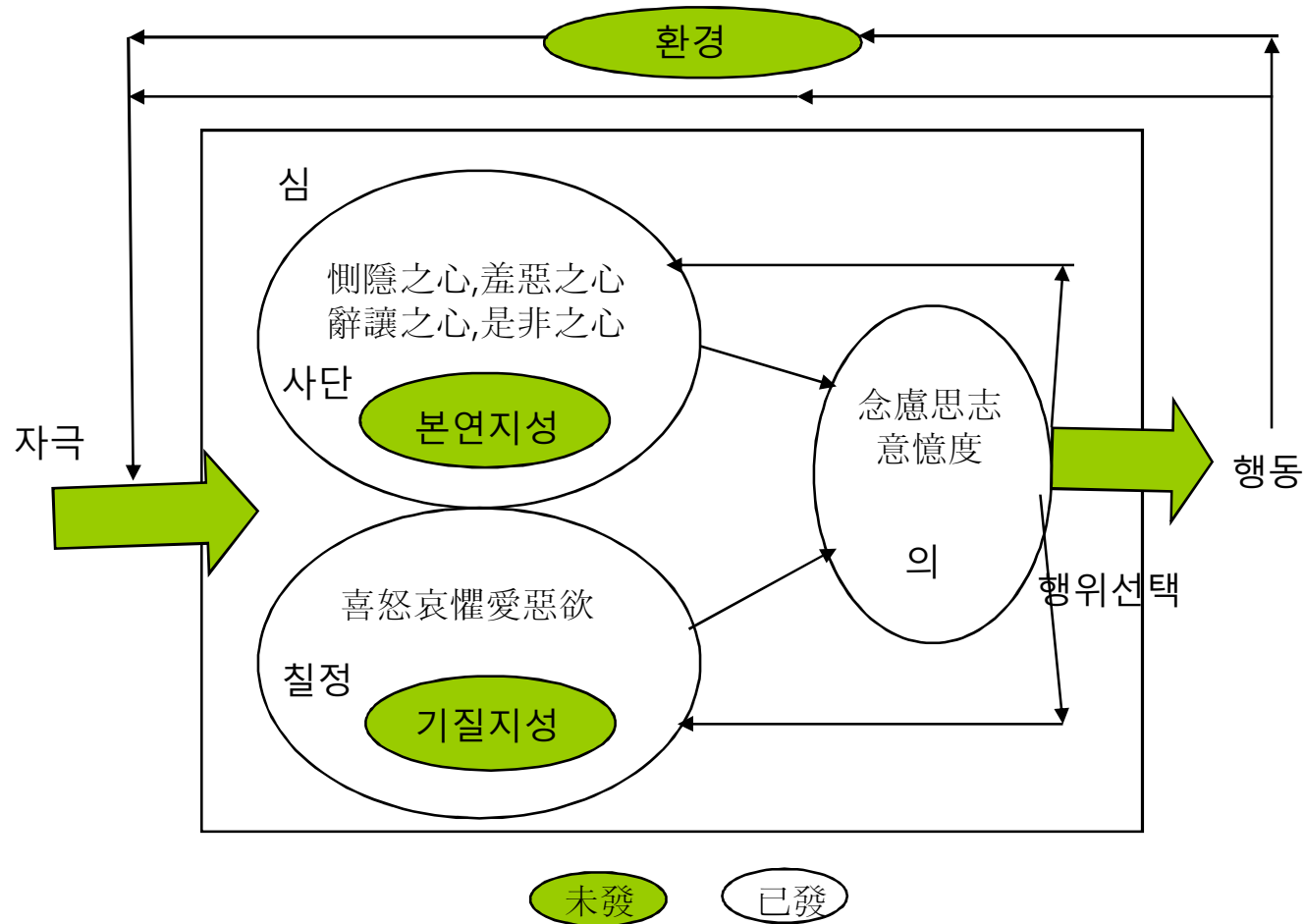
- 유교 성리학
  - 논어와 인공지능?
- 종교학
- 지향성, 퀴리아, 의식, 자의식, 자유의지
  - 1인칭 과학 - 신경현상학
- 행위경제학
- 교육학

## 시스템구성적인공지능의 학제적 마음연구

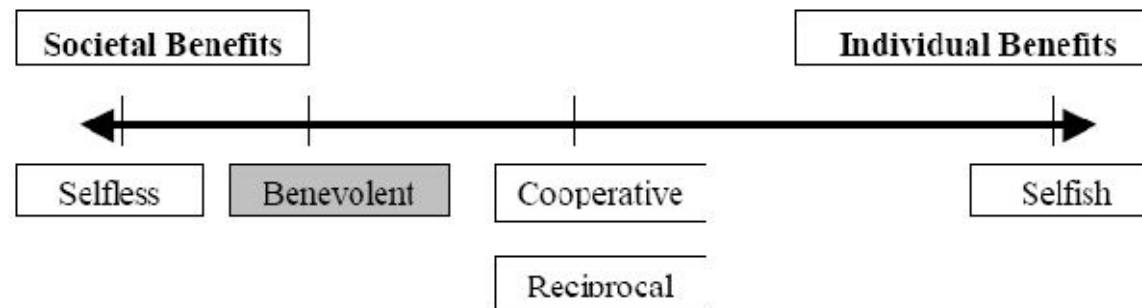
---

- 그레마스의 기호학
- 라캉의 욕망이론
- 스피노자와 유학
- 니클라스 루만-사회시스템이론
- 라투르의 사회연결망이론, 행위자연결망이론

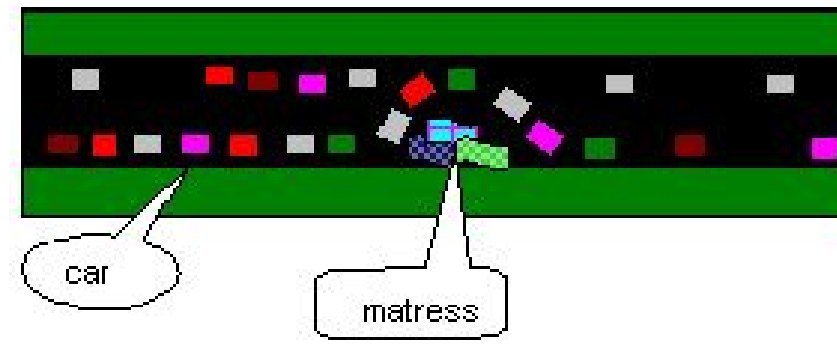
# Tip: 사단칠정 도덕심성모델



# Tip: Benevolent agents



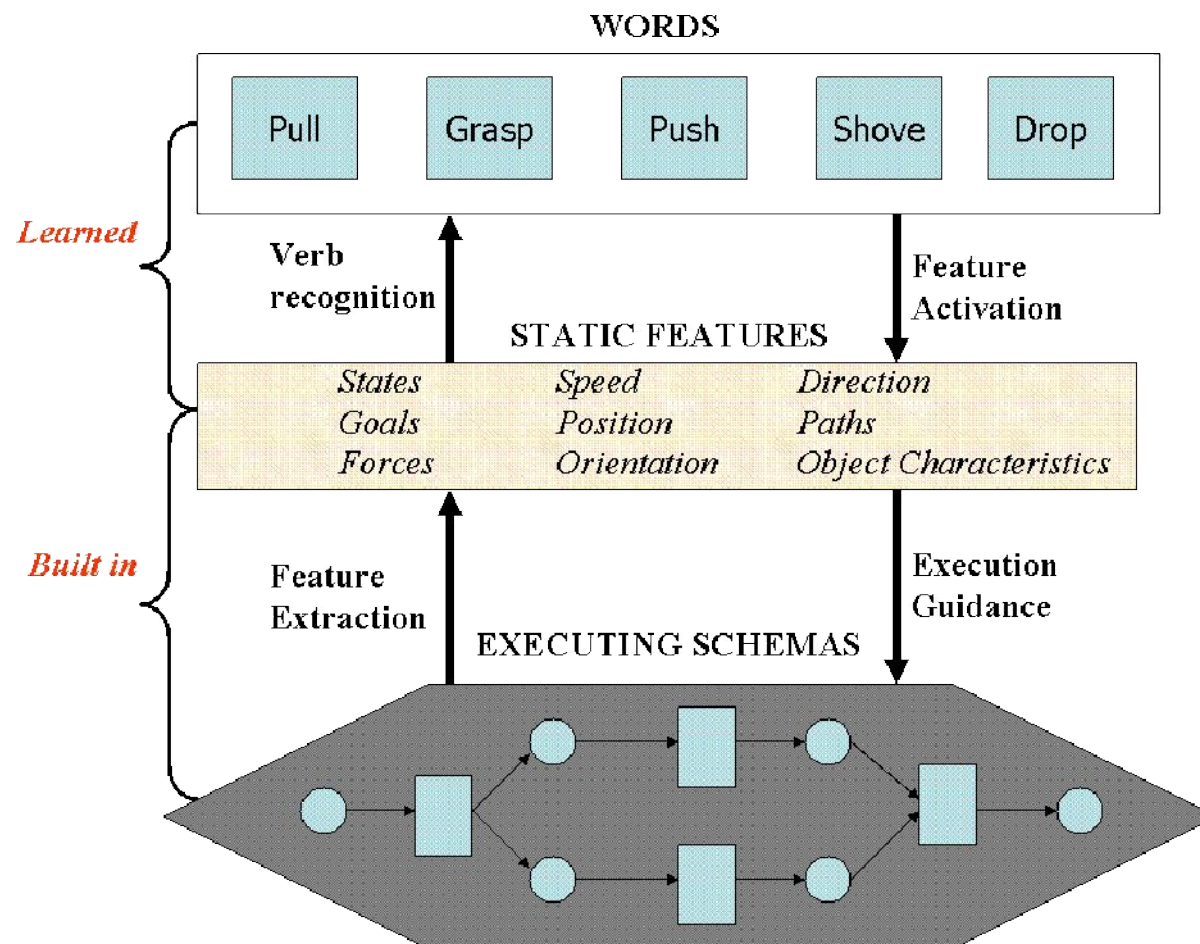
Relationship between benevolent, cooperative, and reciprocal agents



Mattress in the road

# Tip: Lakoff and Feldman - NTL

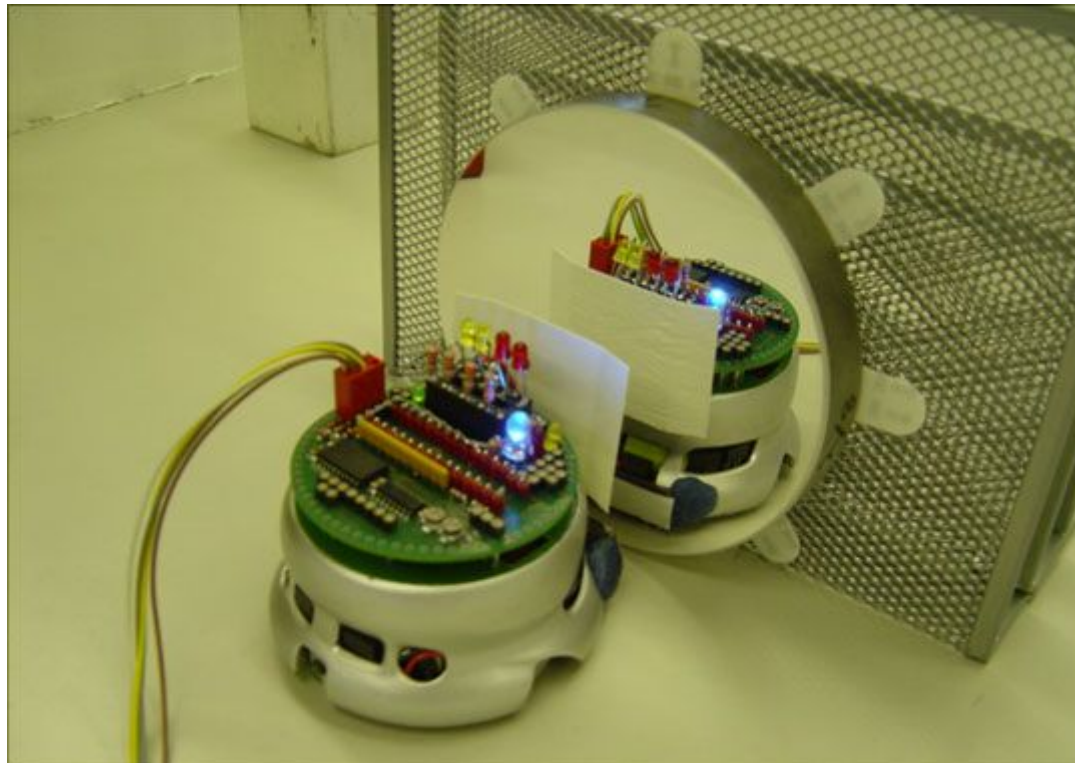
## Bailey's Model for Learning Action Verbs





## Tip: Junichi Takeno - Mirror Image Cognition

---



# Tip: Niclas Luhman

---

- 사회체계로서 커뮤니케이션을 자동생산개념으로 설명
- 인지와 커뮤니케이션이 작동상 폐쇄된, 자기생산적으로 작동하는 그래서 전적으로 독자적인 체계들로 구상된다.
- 커뮤니케이션은 하나의 사회적 단위의 구성원들 사이의 상호 조율된 행동 방식의 유발로 규정된다.
- 커뮤니케이션은 불가능하다. 인간은 커뮤니케이션할 수 없고 커뮤니케이션만 커뮤니케이션한다.
- 의식과 커뮤니케이션 사이의 구조적 접속이 이루어지려면 반드시 언어가 개입되어야 한다.
- 커뮤니케이션은 서로 다른 세가지 선택조정으로 이루어지는 체계-정보의 선택, 전달방법의 선택, 선택적 이해
- 기본적 현실로서 사회를 구성하는 과정은 커뮤니케이션과정이다.
- 심리체계에서 의미를 생산하는 것은 사유이며, 사회체계에서 의미를 생산하는 체계는 커뮤니케이션이다.
- 사회체계는 커뮤니케이션을 기초로 하여 조직되기 때문에 사회체계의 성공 여부는 커뮤니케이션의 선택과 합의의 가능성에 달려 있다.

# 마치면서

---

- 물리주의 시스템구성적 인공지능
- 마음의 자연화 - 마음의 공학화
- 유전자에 의한 맹목적 생존기계
- 인간의 가치와 개인의 존엄성
- 유전자의 긴팔 vs. 인간숙주의 반란
- 자기조직화와 창발성 - 층위별 내적원리와 목적
- 인간-마음과 유전자의 공존
- 인간-마음 - 생존본능과 자유의 창발적 시스템
- 인간-마음의 공학적 이해?

# 마치면서(아주 개인적인)

---

- 삶에 대한 이해/삶에 대한 이해에 대한 이해
- 물리(物理) vs. 문리(文理)
- 의미 만들기-가치 만들기(이성과 감성)
  - 마음공부
  - 체화된 인지 - 심볼 그라운드

# Tip: 물리(物理) vs. 문리(文理)

## □ 엠파스 국어사전

### □ 물리<sup>2</sup>(物理) [명사] [명사]

1. 모든 사물의 이치. 물리를 밝히다.

2. 사물에 대한 이해나 판단의 힘. 물리가 나다

- 물리가 트이다

- 도편수도 그제야 물리가 조금 터진 듯 알겠다는 듯이 빙긋 입술에 웃음을 띠워 씩그러진다. <<박종화, 임진왜란>>

3. <물리> = 물리학(物理學).

### □ 문리<sup>2</sup>(文理) [명사]

발음 [물-] [명사]

1. 글의 뜻을 깨달아 아는 힘.

2. 사물의 이치를 깨달아 아는 힘. 문리가 트이다

- 문리가 훤하다

- 여자 쪽으로는 문리가 트였고 추리력이 밝다 보니 노 순경이 불현듯 그 생각을 떠올린 모양이라고 그는 생각했다.<<김원일, 불의 제전>>

3. 문과(文科)와 이과(理科)를 아울러 이르는 말.

4. <언어> = 문맥<sup>1</sup>(文脈). 문리가 통하다

- 문리를 파악하다.

# 참고문헌(1)

---

- Anderson (2003), "Embodied Cognition: A Field Guide", *Artificial Intelligence* 149, 91-130.
- Barber, Blanchard, Buchinger, Cessac, and Streit(2006), Expectation-driven interaction: a model based on Luhmann's contingency approach.
- Brooks (1991), "Intelligence without Representation," *Artificial Intelligence* 47(1/3), 139-159.
- Bainbridge William Sims (1994), "Artificial Social Intelligence", *Annual Review of Sociology*, 1994, Vol. 20. 20 Issue 1, 407-436.
- Bainbridge, William Sims (2006), *God from the machine: Artificial Intelligence of Religious Cognition*, AltaMira Press.
- Choong-Shik, Park and Gwon-Jong, Yoo (2003), "A New Approach to Neo-Confucian Mind Model from Constructivism and Artificial Intelligence", *proceedings of 20th World Congress of Philosophy*, 12-12.
- Churchland (1989), *Neurophilosophy: Toward a Unified Science of the Mind-Brain*, MIT press.
- Damasio (2003), *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*, Harvest Books.
- Feldman and Narayanan (2003), "Embodied Meaning in Neural Theory of Language, Brain and Language", 2003. <http://www.icsi.berkeley.edu/NTL/publications.html>
- Foester (1995), "Ethics and Second-order Cybernetics", *Stanford Electronic Humanities Review* 4-2. <http://www.stanford.edu/group/SHR/4-2/text/foerster.html>
- Hofstadter (1995), *Fluid Concepts and Creative Analogies*, Basic Books.
- Klamka, Spaniol and Denev, "PALADIN: A Pattern Based Approach to Knowledge Discovery in Digital Social Networks",
- Mingers (1995), *Self-Producing Systems: Implication and Application of Autopoiesis*, Plenum Press.

# 참고문헌(2)

---

- Lakoff and Johnson (1999), *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*, Basic Books.
- Macy and Willer(2002), "From Factors to Actors: Computational Sociology and Agent-Based Modeling", *Annual Review of Sociology*, 2002, Vol. 28. 20 Issue 1, 143-166.
- McDermott (2001), *Mind and Mechanism*, MIT Press.
- McGee (2005), "Enactive Cognitive Science Part 1: Background and Research Themes", *Constructive Foundations* vol. 1, no. 1, 19-34
- McGee (2006), "Enactive Cognitive Science Part 2: Methods, Insights, and Potential", *Constructive Foundations* vol. 1, no. 2, 73-82.
- Minsky (2004), *Emotion Machine*, draft, 2004. <http://web.media.mit.edu/~minsky/>
- Minsky, Push Singh, and Aaron Sloman (2004), "The St. Thomas Common Sense Symposium: Designing Architectures for Human-Level Intelligence", *AI Magazine*, 25-2: Summer 2004, 113-124.
- Reilly (1996), "Believable Social and Emotional Agents," *Technical Report CMS-96-138*, School of Computer Science, Carnegie Mellon University, PA.
- Sloman and Chrisley (2003), "Virtual Machines and Consciousness", *Journal of Consciousness Studies* 10, 4-5.
- Thompson(2004), "Life and mind: From autopoiesis to neurophenomenology. A tribute to Francisco Varela", *Phenomenology and the Cognitive Sciences* 3: 381-398.
- Varela (1999), *Ethical Know-how*, Stanford Univ. Press.
- Varela and Shear(1999), "First-person Methodologies:, What, Why, How?", *Journal of Consciousness Studies*, 6, No. 2-3, 1999, pp. 1-14
- Ziemke (1999), Rethinking Grounding, in *Understanding Representation in the Cognitive Sciences - Does Representation Need Reality?*, Riegler, etc(eds.).

# 참고문헌(3)

- 고병권, 「스프노자의 코뮌주의: 자유를 향한 욕망의 아쌍블라주」, 『인문연구』, 277-299
- 골드먼(저), 석봉래(역) (1998), 『철학과 인지과학』, 서광사, 서울
- 글라저스펠트(저), 김판수 외 6명(역) (1999), 『급진적 구성주의』, 원미사, 서울.
- 김관도, 유청봉(저), 김수중, 박동현, 유원준(역) (2003), 『중국문화의 시스템론적 해석』, 천지, 서울
- 김기현(저) (1998), 『현대인식론』, 민음사, 서울.
- 김성도(저) (2003), 『구조에서 감성으로』, 고려대학교 출판부, 서울
- 김영정(저) (1996), 『심리철학과 인지과학』, 철학과 현실사, 서울
- 김재권(저), 하종호, 김선희(역) (1999), 『심리철학』, 철학과 현실사, 서울.
- 네그리(저), 이기웅(역) (2005), 『전복적 스피노자』, 그린비, 서울
- 다마지오(저), 김린(역) (1999), 『데카르트의 오류』, 중앙문화사, 서울.
- 데블린(저), 전대호(역) (2002), 『수학 유전자』, 까치, 서울.
- 라깡(저) (1998), 『욕망이론』, 문예출판사, 서울
- 루만(저), 이남복(역) (1996), 『생태학적 커뮤니케이션』, 유명사, 서울.
- 루이스 윌퍼트(저), 황소연(역) (2007) 『믿음의 엔진』, 에코의 서재, 서울
- 르두(저), 최준식(역) (2007), 『느끼는 뇌』, 학지사, 서울
- 마뚜라나, 바렐라(저), 최호영(역) (2007), 『삶의 나무』, 갈무리, 서울
- 마이어스(저), 박정수(역) (2005), 『누가 슬라보예 지젝을 미워하는가』, 도서출판 앨피, 서울
- 박충식 (2005), 「종교적 기계 만들기: 계산종교학」, 『인지과학학회 하계 워크샵』, 미발간원고
- 박충식(2004), 「구성적 인공지능」, 『인지과학』, 제 15 권 제 4 호, 61-66.
- 박충식, 유권종 (2000), 「새로운 도덕 심성모델: 퇴계학, 구성주의, 인공지능」, 『한국철학자대회 논문집』, 12-12.
- 박충식, 유권종 (2002a), 「性理學 심성모델 시뮬레이션을 이용한 유교 예 교육 효용성 분석」, 『동양철학회』, 21-21.



# 참고문헌(4)

---

- 박충식, 유권종, 장숙필 (2002), 「서구심리모델과 조선 성리학의 심성모델 비교: 구성주의의 관점에 의하여」, 『세계한국학대회 논문집』, 12-12.
- 발레라, 톰슨, 로쉬(저), 석봉래(역) (1997), 『인지과학의 철학적 이해』, 옥토, 서울.
- 발리바르(저), 진태원(역) (2005), 『스피노자와 정치』, 이제에북스, 서울
- 브룩스(저) (2005), 『로봇만들기』, 바다출판사, 서울
- 쉐퍼(저), 이남복(역) (2002), 『니클라스 루만의 사회사상』, 백의, 서울
- 슈미트(저), 박여성(역) (1995b), 『구성주의』, 까치, 서울.
- 요아스(저), 신진욱(역) (2003), 『행위의 창조성』, 한울아카데미, 서울.
- 유동운(저), 신제도주의 경제학(1999), 『신제도주의 경제학』, 선학사, 서울
- 이나스(저), 김미선(역), (2007), 『꿈꾸는 기계의 진화』, 북센스, 서울
- 이초식(저) (1993), 『인공지능의 철학』, 고려대학교 출판부, 서울
- 존슨(저), 노양진(역) (2000), 『마음속의 몸』, 철학과 현실, 서울.
- 카우프만(저), 국형태(역) (2002), 『혼돈의 가장자리』, 사이언스북스, 서울.
- 커즈와일(저), 김명남, 장시형(역) (2007), 『특이점이 온다』, 김영사, 파주.
- 페트리리, 폰지오(저), 김수철(역) (2001), 『토머스 시벅과 생명의 기호』, 이제이북스, 서울.
- 홀랜드(저), 김희봉(역) (2001), 『숨겨진 질서』, 사이언스북스,