

## "Connecting Mind (Brain) and Machine: Descartes' & Turing's Legacies and going beyond them"

(- / text in Korean -)

“마음과 기계, 뇌의 연결 역사: 튜링기계 이론, 체화된 마음”

by Jung-Mo Lee

(Sungkyunkwan University; <https://www.facebook.com/metapsy>)

**\*Copyright©2012, Jung-Mo Lee:** / 12 June, 2012

\* Contents: in English, pp. 2-4. / in Korean(한글), pp. 4-6.

Main texts : pp. 7-77. ; References : pp. 78-83.

\* The size of the pdf file : 1,227 K

\* Web Link: <http://blog.naver.com/metapsy/40161158790>

(A Korean Version of the PDF file is attached at this link.)

---

### A tribute to Professor Margaret A. Boden,

who inspired me through her writings to delve into the history of ideas on mechanism. specially in relation to the launching of Cognitive Science in the 1950s,

and to other young Korean students who are interested in Cognitive Science, or aspire to become cognitive scientists, and are thirsty for the relevant information on Cognitive Science written in Korean .

- - - - -

## [Contents in English]

1. **Mind-Machine Connections : Enterprises before the 17C**
  - 1.1. Early attempts to connect Mind and Machine: Greeks and ...
  - 1.2. Ramon Lull's Logic Machine
  - 1.3. Other attempts before the 17C, and homunculus
2. **Mind-Machine Connections in the 17C**
  - 2.a. Machine in HOBBS
  - 2.b. Machine in DESCARTES
  - 2.c. La METTRIE's Machine
  - 2.d. Calculators in the 17C
  - 2.e. PASCAL's calculator
  - 2.f. LEIBNIZ Machine
3. **The 19C & 20C's Formalization attempts before Turing Machine**
  - 3.a. BABBAGE's Analytic Machine  
-[Table 1]. Structure of Babbage's Analytic Machine -
  - 3.b. BOOLE and Formalization in Logic
4. **Phenomena, Formal Systems, Machine, & Mathematical Thinking**
  - 4.a. Phenomena & Formal Descriptions
  - 4.b. Machine & Formal Descriptions
  - 4.c. Mathematical Thinking & Formal Descriptions
  - 4.d. Mind and Computer: Computation and Formal System
5. **Turing Machine Theory**
  - 5.a. Turing and his Life
  - 5.b. Turing Machine Theory: Introduction
  - 5.c. Turing Machine Theory in details  
- [Fig. 2]. Turing Machine -  
- [Table 2]. An example of Operations in Turing Machine -
  - 5.d. Intelligent Machine and Turing Test  
- [Fig. 3]. Turing Test Situation
  - 5.e. Implications of Turing Machine Theory
6. **Scholars pioneered the Machine-Mind connections, after Turing**
  - 6.1. N. Wiener.
  - 6.2. C. Shannon.
  - 6.3. W. McCulloch
  - 6.4. Early Computer Hardware

6.5. John von Neuman

## **7. Conferences that spearheaded the new Paradigm Shift, by Connecting Mind and Machine**

- 7.1. The first Hixon Symposium in 1948
- 7.2. Dartmouth Conference in 1956
- 7.3. MIT Information Symposium in 1956
- 7.4. NPL Symposium in 1958
- 7.5. UNESCO Conference in Paris in 1959

## **8. The Birth of New Psychology in the '1950s and '1960s**

- 8.1. Early Physio-psychological Research
- 8.2. The Legacy of Functional Psychology
- 8.3. Gestalt Psychology
- 8.4. Psychometrics and other Research
- 8.5. Verbal Learning and Memory Research in the '50s
- 8.6. The Early Cognitive Scholars in Europe
- 8.7. The Early Cognitive Psychologists before U. Neisser
- 8.8. Human Engineering Research in Psychology
- 8.9. D. E. Broadbent's Contributions

## **9. The Paradigm Shift: The Birth of Cognitivism and Cognitive Science**

- 9.1. Linguistic Legacy
- 9.2. Impact of AI and Computer Science
- 9.3. Formal Launching of Cognitive Psychology by U. Neisser
- 9.4. Ideas of H. Simon and A. Newell: Computer Metaphor  
– [Table 2]. Intellectual Precursors (Disciplines) of Cognitivism –

## **10. Spreading and Reformulation of Cognitivism in the '20C: Connectionism and Brain Science**

- 10.1. Classical Cognitivism
- 10.2. Connectionism and Neural Nets
- 10.3. Cognitive Neuroscience
- 10.4. In Search for an Alternative Paradigm

## **11. Reformulating Cognitive Science : For Embodied Cognition Paradigm**

- 11.1. Backgrounds of the Rise of the Embodied Cognition Paradigm
- 11.2. Précis of 'Embodied Cognition' Approach
- 11.3. Radical Embodied Cognition Approach
- 11.4. Embodied Cognition and the Interactions of Human with Artifacts
- 11.5. Interactions I: with Artifacts in General (A-1)
- 11.6. Interactions II: with Agentive Artifacts (A-2)

11.7. Problems of Embodied Cognition Approach

11.8. Embodied Mind and Cognition: Interim Conclusion

## 12. Summary, Synthesis and Prospects

- [Fig. 4]. Embodied Cognition Paradigm and its Relation with Other Disciplines
  - [Table 3]. A Summary of the Intellectual Enterprises of Connecting Mind (Brain) and Machine: Retrospect and Prospect
- 

### [목차] (contents in Korean)

#### 1. 17 세기 이전의 기계와 마음의 연결시도

- 1.1. 초기의 기계와 마음의 연결시도: 자동기계로서의 몸과 마음
- 1.2. *Ramon Lull*의 논리기계
- 1.3. 17세기 이전의 다른 노력들: homonculus 개념

#### 2. 17세기의 [마음-기계] 연결 생각들

- 2.a. HOBBS의 기계론
- 2.b. DESCARTES의 기계론
- 2.c. La METTRIE의 기계론
- 2.d. 17세기의 계산기 개념
- 2.e. PASCAL의 계산기
- 2.f. LEIBNIZ의 기계론

#### 3. 튜링기계 이전 19세기와 20세기의 형식화 시도들

- 3.a. BABBAGE: 형식 기계론  
[표 1]. 배비지(Babbage)의 분석기계의 구성요소
- 3.b. BOOLE: 논리적 사고의 형식화

#### 4. 현상, 형식 체계, 기계, 수학적 사고

- 4.a. 현상과 형식적 기술(記述)
- 4.b. 기계와 형식적 기술(記述)
- 4.c. 수학적 사고와 형식적 기술(記述)
- 4.d. 마음과 컴퓨터: 계산과 형식체계

#### 5. 튜링기계 이론

- 5.a. 튜링의 삶
- 5.b. 튜링기계 이론: 서론

- 5.c. 튜링기계 이론: 세부
  - [그림2]. 튜링기계
  - [표2]. 튜링기계의 조작 수행의 한 예 :  $(3+2) = 5$
- 5.d. 지능적 기계와 튜링 테스트
  - [그림3]. 튜링테스트(검사) 상황
- 5.e. 튜링기계 이론의 의의;

## 6. 튜링 이후 기계-뇌-마음 연결의 개념적 변화를 주도한 학자들

- 6.1. Wiener.
- 6.2. Shannon .
- 6.3. W. McCulloch
- 6.4. 컴퓨터 하드웨어의 발전
- 6.5. 수학자 John von Neuman.

## 7. 마음 기계 연결 변혁의 기폭제가 된 학술 모임들

- 7.1. 1948년의 첫 HIXON 심포지엄 : 대뇌 메커니즘과 행동
- 7.2. 1956년 여름의 Dartmouth대학 모임
- 7.3. 1956년 가을의 MIT 정보이론 심포지엄
- 7.4. 1958년 가을의 영국의 NPL 심포지엄
- 7.5. 1959년의 파리 유네스코 정보처리 학술대회

## 8. 1960년대 이전의 심리학 내의 움직임: 새 심리학의 태동

- 8.1. 신경생리심리학 연구
- 8.2. 기능주의 심리학
- 8.3. 형태심리학
- 8.4. 심리측정, 기타
- 8.5. 언어학습과 기억연구
- 8.6. 유럽의 인지심리학 선구자
- 8.7. 미국의 인지심리학 선구자.
- 8.8. 인간공학 연구
- 8.9. 브로드벤트의 공헌

## 9. 인지주의, 인지과학의 탄생과 확산

- 9.1. 언어학의 영향
- 9.2. 인공지능과 컴퓨터 과학의 영향
- 9.3. 인지심리학의 공식적 출발
- 9.4. 사이먼과 뉴웰의 생각: 컴퓨터와 마음 간의 유추
  - [표2]. '인지과학'이라는 새 패러다임의 등장에 영향 준 사조들

## 10. 20세기의 인지과학의 확산과 재구성 시도: 신경망, 신경과학

- 10.1. 고전적 인지주의(Classical Cognitivism) 접근

- 10.2. 연결주의(Connectionism)
- 10.3. 뇌과학과 인지신경과학적 접근
- 10.4. 대안적 틀의 모색

## 11. 인지과학 재구성: 체화된 인지 접근

- 11.1. 체화된 마음 관점의 떠오름의 배경
- 11.2. 체화된 마음 접근의 요점
- 11.3. 급진적 체화된 인지 접근
- 11.4. 체화된 마음과 기계(인공물)와 인간
- 11.5. 인간과-인공물일반(A-1)의 상호작용
- 11.6. 인간과-행위주체자로서의 인공물(A-2)의 상호작용
- 11.7. 체화적 접근의 문제점
- 11.8. 체화된 인지: 맺는 말

## 12. 전체 종합

- [그림4]. ‘체화된 인지’ 접근을 포함하는 [인지과학]과 다른 학문과의 관계
- [표3]. 조망: 년대별 ‘기계-몸(뇌)-마음(지능)’ 연결 특성

## [Main Texts in Korean]

과학은 역사적으로 그 연구 대상에 대해 보다 효율적이고 엄밀하고 단순한 설명을 부여하려고 노력해왔다. 이러한 노력의 하나로 자연현상에서 대한 가장 과학적이고 엄밀하고 단순한 설명으로 대두된 것이 기계론적(mechanistic) 설명이었다. 기계론은 그 연원을 고대 희랍에서 찾을 수도 있고, 유럽에서 초보적 기계와 지능적 기계를 만들려는 여러 시도들이 있었지만, 희랍시대 이후의 서구 기계론은 처음에는 물리학에서 이론적으로 체계화되고 이것이 화학, 생물학, 생리학을 거쳐(Dijksterhuis, 1961) 심리학에 도입되었다(Cohen, 1980). 17세기 유럽에서는 공학의 발달과 갈릴레오와 뉴턴의 물리학이 대두됨에 따라, 우주물 기계론적으로 보려는 입장이 과학자들에게 널리 퍼졌다. 이들의 기계론적 입장은 기본적으로 물리적 자연현상이 기계적 결정론(Mechanistic Determinism)에 따라 일어나며 법칙적이고 예언 가능하며 관찰, 실험 및 측정이 가능하다는 생각으로 형성되었다.

뒤를 이어 이러한 입장을 인간에게도 적용하여 설명하려는 시도가 나타났고, 홉스(T. Hobbes), 데카르트와 라메트리(La Mettrie) 등의 기계론적 견해가 인간의 신체와 마음의 과정에 적용되었다(이정모, 1983, 2001, 2009). 17 및 18세기의 대륙의 기계론과 그 이후의 기계론은 크게 둘로 나누어 볼 수 있다. 하나는 데카르트와 말브랑쉬(N. de Malebranche)<sup>1)</sup> 등과 같이 심신 이원론을 전제하고 신체는 기계이나 마음(정신)은 기계가 아니라는 2원론적 입장이며, 다른 하나는 가상디(P. Gassendi)<sup>2)</sup>; 또는 라메트리 등의 입장처럼 신체 뿐 만 아니라 마음(심리적 과정)도 기계적이라는 일원론적 입장이다.

데카르트가 인간의 신체는 동물과 같은 하나의 기계라고 천명한 이후에, 인간의 신체를 기계로 보는 데에는 거의 이의가 없어 왔다. 하지만 인간 마음 또는 심적 과정을 기계적이라고 볼 수 있느냐에 대해서는 데카르트의 2원론적 기계론의 제시와 라메트리의 일원론적 주장 (저서 [L' Home Machine(기계인간)]) 이래 계속적으로 논쟁이 되어져 왔다. 20세기 중엽에 이르러서는 마음의 본질을 기계의 전형인 컴퓨터에 유추하여 설명하려는 고전적 인지주의 접근과 이에 대립되는 견해가(예: 체화된 인지(Embodied cognition) 입장) 심리철학과 인지과학에서 첨예화된 논제로 논의되고 있다 (Dennett, 1986, 이정모, 2010ㄴ, 2011ㄴ).

일반적으로 마음에 대한 일원론적 입장을 취하는 관점인 기계론은, 마음이 물질인 두뇌의 산물이며, 두뇌란 수많은 뉴런들이 모여 이루어지며, 이들 뉴런들은, 아비브(Arbib, 1964) 등이 기술했듯이, 기계적 법칙에 의해 상호작용하며 그 상호작용의 결

1) [http://en.wikipedia.org/wiki/Nicolas\\_Malebranche](http://en.wikipedia.org/wiki/Nicolas_Malebranche)

2) [http://en.wikipedia.org/wiki/Pierre\\_Gassendi](http://en.wikipedia.org/wiki/Pierre_Gassendi)

과로 나타나는 것이 마음의 내용이라는 것이다. 즉 마음이란 두뇌의 작용을 반영하기에, 두뇌의 작용이 기계적인 한, 마음의 작용도 기계적이며, 따라서 마음은 하나의 기계로 간주할 수 있다는 것이다. 또한 기계의 정의를 ‘유한히 기술할 수 있는 현상’(Jackson, 1985)과 같은 의미로 규정한다면, ‘기계적’이란 유한히 기술할 수 있다는 의미, 또는 알고리즘에 의해 효율적으로 규정할 수 있다는 의미가 된다(Cutland, 1980). 따라서 기계론에서 ‘마음이 기계적이다’라는 논지는 마음의 작용을 효율적으로 알고리즘에 의하여 유한히 기술할 수 있다는 의미가 된다.

이러한 입장의 기계론은, 20세기 이전에는 다분히 사변적이며 엄밀한 개념 규정이 결여되어 있었고 체계화되지 못했다. 이러한 결함이 1930년대를 기점으로 하여 튜링(A. Turing)을 비롯한 학자들에 의해 가다듬어져서 새로운 기계론이 제기되었다. 이 글에서는 마음과 기계론을 연결하려는 이러한 입장이, 역사적으로 어떻게 발전되어 왔는가, 그 핵심 이론인 튜링 기계론은 무엇이며 어떠한 영향을 미쳤는가를 살펴보기로 한다.

## 1. 17세기 이전의 기계와 마음의 연결 시도

20세기의 과학의 가장 큰 특징 중의 하나는 인간의 마음과 컴퓨터<sup>3)</sup>를 개념적으로 연결한 것이라 할 수 있다(이를 과학철학에서는 ‘인지혁명(Cognitive Revolution)’이라고 한다). 이러한 과학적 업적은 단숨에 이루어진 것이 아니다. 역사적으로 오랜 동안 여러 학자들에 의하여 마음, 신체, 기계, 계산의 개념 사이의 연결 생각이 시도되고, 세련화되어 온 결과의 산물이다.

그렇다면 어떻게 인간과 기계를 연결하여 ‘기계인간’을 만들 수 있으며, 인간과 기계가 본질적으로 같은 지적 능력을 지닌 것을 보일 수 있을 것인가? 어떻게 기계에 사람의 심적 능력을 부여할 수 있을까? 이러한 물음에 대한 대답은 예로부터 인간이 만들어 온 자동기계(automata), 계산기계(computing machines)의 여러 유형을 살펴봄으로써 찾아질 수 있다.

### 1.1. 초기의 기계와 마음의 연결 시도: 자동기계로서의 몸과 마음

인간이 ‘기계인간’을 만들려고 수없이 노력해 왔다는 것은 희랍신화(기술의 신 Hephaestus 관련 이야기), 이슬람 문화권의 역사, 유럽의 문화사 상에서 기발한 생각을 지녔던 사람들에게서 찾아 볼 수 있다(McCorduck, 1979).

아마도 인간을 기계로 처음 개념화하기 시작한 것은, 서양 중심으로<sup>4)</sup> 보았을 때에,

---

3) 컴퓨터는 인간이 마음을 작동시켜 만든 인공물(artefact)인 기계의 하나이다.



희랍 신화의 이야기들일 수 있다. 희랍 신중에서 대장장이 일과 불을 맡은 신인 헤파에스투스(Hephaestus)는 판도라 상자도 만들었지만 인간을 도울 수 있는 현대적 의미의 자동기계 인간을 만들었다고 할 수 있다. 대장장이 일을 도와주는 자동기계인간, 보초를 서는 자동기계인간을 만들었다고 신화에서 이야기되고 있다. 또한 희랍신화에 의하면 결혼 혐오자인 피그말리온(Pygmalion)은 상아로 여성을 만들기도 하였고, 대달루스(Daedalus)는 얼굴 표정을 변화시키는 동상을 만들기도 하였다고 한다.

희랍 이후 알렉산드리아가 번영하던 시대에는 자동기계 인형을 극장에서 사용하기도 하였다고 한다. 중세의 교황 실베스터 2세는 그가 교황이 되기 전에 말하는 머리를 지닌 동상을 만들기도 하였다고 한다. 또한 중세의 서구의 기독교 암흑시대에 과학적 명맥을 유지하던 이슬람에서는 ‘자이리야’라고 하는 ‘28개의 아랍 알파벳을 이슬람 철학의 기본 개념 범주에 연결하게 하는’ 생각하는 기계도 만들었다고 한다. 이러한 기계를 본 13세기의 스페인의 기독교도인 라몬 룰(Ramon Lull)은 아랍의 자이리야 기계에 대응되는 기독교용 생각하는 자동기계를 만들어야 한다고 생각하였다.

## 1.2. Ramon Lull의 논리기계

13세기의 스페인의 카탈론 왕국의 왕의 고문으로 있던 Ramon Lull은 아랍권을 여행하는 기회가 있었는데, 이때에 아랍의 이슬람 철학의 모든 문제에 대하여 답을 주는 자이리야(zairja; <http://en.wikipedia.org/wiki/Zairja>)라고 하는 기계를 보고나서, 이에 대응되는 기독교식의 기계를 만들어야겠다고 생각하였다. 그래서 그가 창안해 낸 것이 서양의 ‘논리기계’의 시초라고 볼 수 있는 'Ars Magna' 라는 자동기계였다.<sup>5)</sup> 이와는 달리, ‘Antikythera mechanism’이 인류 최초의 논리기계, 아날로그 기계라고 보는 입장도 있다.<sup>6)</sup>

'Ars Magna'라는 기계는 몇 개의 원판으로 되어 있는데 이 원판들은 각기 한 중심점을 공유하는 동심원들로 이루어졌다. 각 원들은 기호 또는 단어들인 몇 개씩 적혀 있었고, 이 단어들은 기독교적 세계관의 내용들이었다. 따라서 이 기계는 기독교적 세계관 내에서 모든 주제에 대하여 진리에 도달할 수 있게 하는 사실을 발견하는 ‘생각하는 기계’의 설계였다.<sup>7)</sup>

예를 들어서 이 기계를 단순화하여 그린 아래 그림에서 보이는 것처럼, 원반의 밖쪽 원에는 주어로 사용될 수 있는 단어들인 제시되고, 중간 원에는 동사로 사용될 수 있는 단어들인, 안쪽 원에는 목적어로 사용될 수 있는 단어들인 열거되어 있는 것과

4) 동양에서는 중국에서 기계의 발명과 발달이 동양의 다른 나라에 비하여 앞서 있었다고 할 수 있다. 중국의 기계의 발전에 대하여는, Bautista Paz, E., Ceccarelli, M., Echávarri Otero, J., Muñoz Sanz, J. L. (Ed) (2010). [A Brief Illustrated History of Machines and Mechanisms]. Springer. (Series: History of Mechanism and Machine Science, Vol. 10)의 2장에 중국의 발명과 기계에 대한 역사가 기술되어 있다 (2장. Chinese Inventions and Machines). 중국의 발명 역사에 관한 개괄적 내용은 다음에 있다: [http://en.wikipedia.org/wiki/List\\_of\\_Chinese\\_inventions](http://en.wikipedia.org/wiki/List_of_Chinese_inventions)

5) 이 Ars Magna와 관련하여, 이 기계의 자세한 그림들을 보려면 다음 사이트의 도해 그림들을 참조할 수 있다: <http://lullianarts.net/downloads.htm>.

같은 형태이다. 위의 원반의 예에서 ["I love mice," "You hate cats," and "They eat frogs."] 같은 내용을 나타낼 수 있다. 이러한 동심원의 한 원이 16개 정도 있어서 다른 단어나 기호를 나타내고, 한 원반에 14개 정도의 동심원들이 있고, 그리고 이러한 원반들이 여러 개가 있다면, 이들이 조합되면서 나타낼 수 있는 문장, 개념의 수는 상당한 수에 이른다.

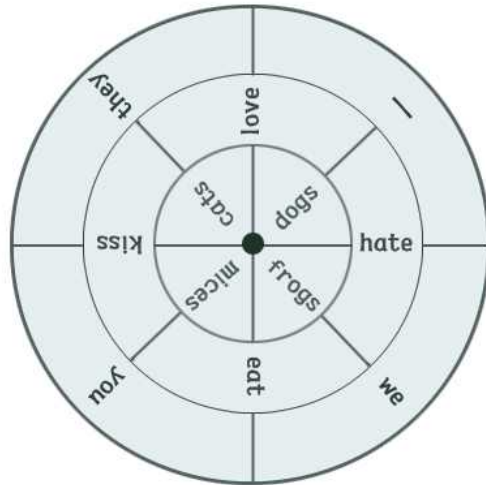
Lull은 성경에 있는 모든 기독교적 철학과 관점의 기본 내용들에 관한 지식이 ‘유한한’ 개수라고 생각하였고, 이 유한한 지식을 [예를 들어 "God's mercy is infinite," "God's mercy is mysterious," "God's mercy is just,"] 이러한 논리와 기호(단어나 어구를 포함한)가 연결되는 조합 방식을 통하여 모두 나타낼 수 있다고 보았다.



Antikythera mechanism는 천체의 위치를 계산하는 계산기계로 기원전 1세기 경에 만들어졌다고 보며, 최초의 과학적 계산기, 또는 인류 최초의 아날로그 컴퓨터라고 볼 수 있으며 인류문화사에서 다빈치의 그림 모나리자보다 더 중요한 위치를 차지한다고도 평가된다. 이 기계는 최고 72 개의 톱니를 지녔다고 추측된다. 이 기계는 3개의 주축 원반 있는데 앞에 있는 원반은 두 개의 둥근 고리가 있다. 365일을 나타내는 2종의 이집트 율력에 대응하는 원반들이다. 1년 중 일자와 달과 해의 위치를 나타내준다고 볼 수 있다. 또 당시에 알려져 있던 5개의 행성의 위치도 표시하는 메커니즘이 있던 것 같다. 뒤쪽 위쪽에는 별점에 상응하는 원반이 있는데 한번 회전이 47개로 나누어지며, 19년 주기의 235개 달을 나타낸다. 뒤쪽 아래쪽 원반은 223개로 나누어지고 이는 다시 54개로 나누어 질 수 있다. 이는 바빌로니아의 사로스 율력과 일식의 주기를 나타내주고 올림픽 게임 년도도 계산하게 하며 또 다른 종류의 당시의 율력을 나타낼 수 있게 하기도 한다. 웹의 [http://en.wikipedia.org/wiki/Antikythera\\_mechanism](http://en.wikipedia.org/wiki/Antikythera_mechanism) 에 관련 그림과 이 기계를 현대 시점에서 도해적으로 재구성한 사진이 있다. (자료원: [http://en.wikipedia.org/wiki/Antikythera\\_mechanism](http://en.wikipedia.org/wiki/Antikythera_mechanism))

7) <http://www.maxmon.com/1274ad.htm> 과 [http://en.wikipedia.org/wiki/Ramon\\_Llull#Ars\\_generalis\\_ultima..28Ars\\_Magna.29](http://en.wikipedia.org/wiki/Ramon_Llull#Ars_generalis_ultima..28Ars_Magna.29) 참조

Ars Magna (Ars generalis ultima) 원반



[그림 1]. 13세기의 Ramon Lull의 생각기계

이러한 Lull의 기독교적 사고의 논리기계에 대한 생각은 현대의 정보이론의 선구였다고 볼 수 있으며, 이후에 유럽의 많은 사람들에게 전파되었다. 걸리버 여행기의 작가 스유프트(Jonathan Swift)에게도 영향을 주어서 '걸리버 여행기'의 내용 중에 이런 기계와 관련된 이야기가 나온다.

그러나 현대 정보이론, 계산이론과의 연결에서 중요한 것은 이러한 Lull의 생각이 17세기에 라이프니츠(Gottfried von Leibniz)에게 영향을 주었다는 사실이다(라이프니츠 기계에 대하여는 후에 2.f.절에서 다시 설명한다). 라이프니츠는 기계적 계산기와 미적분체계를 발전시켰는데, 이는 Lull의 생각을 형식논리(formal logic)로 확장한 것이라고 할 수 있다.

### 1.3. 17세기 이전의 다른 노력들: homunculus 개념

이후에 16세기 초에 Paracelsus는 '작은 인간 (homunculus)'이라는 가상적 인간을 이야기하였고, 1580년 경에는 폴란드의 유대인들 사이에서 인공인간 골렘(Golem)에 대한 이야기가 나타났다. 프라하의 램비 뢰브(Judah Loew ben Bezalel)가 흙에 물, 불 등을 가하여 만들어서 유대인들을 외부 공격에서 지키고 여러 가지 잔심부름을 하게 하는 그런 인공인간 이야기가 만들어진 것이다.<sup>8)</sup>

8) 골렘에 관한 영문 자료 및 그림 소개 사이트: <http://en.wikipedia.org/wiki/Golem>

**[작은 인간 ["homunculus"] 개념의 선구]**. 16세기의 연금술학자 Paracelsus는 자신이 작은 인간을 만들었다고 하였다. 그는 이것을 homunculus라고 불렀다. 이 작은 인간은 12인치의 키에 자동인형이 하는 모든 일을 한다고 하였다. 이 작은 인간을 만드는 처방은 동물들의 뼈, 정충, 피부조각들, 털을 조합하는 것이며 말뚝으로 둘러싸인 땅에 40일간 놓아두면 배아가 자연히 생겨난다고 했다. 이것은 인간의 모습을 지녔지만, 투명하고 신체가 없다. 40일 이내에 인간의 피를 먹이면 살게 되고 인간의 지능을 지닌다는 것이다.

이 작은 인간 개념은 이후에 여러 사람들에 의하여 사용되었는데, 정충속의 자동인간의 개념으로도 사용되었다(<http://en.wikipedia.org/wiki/Homunculus>의 그림들을 참조). 현대에 이르러 심리철학적 논쟁에서는 인간 내부에 있어서 인간의 행동과 생각을 좌우 하는 여러 ‘작은 인간’의 개념으로 철학적 논쟁을 위한 개념으로 사용되어 왔다. 현대의 신경과학에서는 대뇌피질의 각 부분이 담당하는 기능이 감각과 운동, 특히 입과 손 부분에 상대적으로 많은 대뇌피질 영역을 할애하고 있다는 의미를 전달하기 위하여 하나의 비유적 개념으로 감각-운동 ‘작은 인간’이라는 용어를 사용하기도 한다.

## 2. 17세기의 [마음-기계] 연결 생각들

역사적으로 마음과 기계를 연결하는 구체적 산물을 만들기 보다는 개념적으로. 이론적으로 체계적으로 관련지으려 한 학자들 중에서 현재의 인지과학과 논의와 가장 크게 관련을 지을 수 있는 사람들을 언급한다면 홉스와 데카르트와 라메트리이다. 이들의 기계론을 살펴보자.

### 2.a. HOBBS의 기계론

17세기 중엽의 토마스 홉스(T. Hobbes; [http://en.wikipedia.org/wiki/Thomas\\_Hobbes](http://en.wikipedia.org/wiki/Thomas_Hobbes); 1588-1679)는 아리스토텔레스의 틀을 벗어나 갈릴레오와 케플러의 물리학과 하비(W. Harvey; [http://en.wikipedia.org/wiki/William\\_Harvey](http://en.wikipedia.org/wiki/William_Harvey))의 생리학에서 강조된 운동<sup>9)</sup>의 개념을 심리학 영역으로 확장하였다. 그는 감각과 사고가 감각기관과 두뇌 내의 운동으로 이해될 수 있다고 보았다. 뿐만 아니라 다른 종류의 심적 과정도 신체의 운동에 근거한다고 보았다.

---

9) 1628년에 출간된 "De Motu Cordis"(영문 이름 ‘심장과 혈액의 움직임’) 책에서 Harvey가 제시한 혈액순환의 운동 개념

그는 영혼이나 마음이 몸과는 별도로 이해될 수 있는 것이 아니라고 보았다. 인간은 근본적으로 기계라고 보았다. 심적 과정이란 완전히 물질적, 신체적 기초에 의존해 있고, 따라서 인간의 행동은 물리학과 마찬가지로 양적 과학의 기초 위에서 이해될 수 있다고 보았다. 인간은 물리적 법칙, 즉 인과법칙에 의하여 작동하며 양적 과학의 기초 위에서 이해될 수 있다고 그는 생각했다. 감각과 사고, 정서가 감각기관과 두뇌 내의 운동으로 이해될 수 있으며 뿐만 아니라 다른 종류의 심적 과정도 신체의 운동에 근거한다고 보았다. 인간은 쾌(快) 추구하고 고통 회피의 기계적 원리에 의하여 작동하는 유기체라고 보았다. 그는 철학을 움직이는 물체들의 연구로 환원하려 하였었다. 그렇기는 하지만 흄스는 어디까지나 철학자였지 경험적 연구자는 아니었다.

그의 유물론적 관점에는 ‘사고는 계산이다’라는 생각이 내재해 있었다. 그는 추론적 사고가 일종의 덧셈과 뺄셈과 같은 계산이라고 보았다. 그는 말하기를 ‘나는 추론적 사고라는 것은 계산이라고 생각한다. 그리고 계산이란 여러 가지를 동시에 더하여 합을 얻거나 뺄셈을 하여 나머지를 계산하는 하는 것과 같다. .... 삼단추론이란 두 명제를 합하는 것과 같다(Hobbes 1655, 1.2).’ 인간 사고를 계산이라고 보는 흄스의 관점은 후세에 라이프니츠를 거쳐 오늘날 인지과학적 관점에 흘러들었다. 특히 인공지능의 관점에 이러한 생각은 이어지고 있다.

마음과 기계에 대한 Hobbes의 기본 가정을 현대적 표현으로 바꾸어 본다면, 인간의 이성적 사고는 본질적으로 일정한 상징(기호)들을 명료히 규정할 수 있는 어떤 규칙에 따라 체계적으로 조작한다는 것이었다. 이것의 의의는 인간의 사고, 이성도 하나의 계산과정으로 볼 수 있다는 것인데, 유사한 관점의 현대 인지주의 등장에 300여년을 앞선 이러한 Hobbes의 생각은 20세기에 이르러 후대 학자들에 의하여 체계화되고 정교화된다.

## 2.b. DESCARTES의 기계론

이러한 흄스의 영향을 받은 데카르트는 그 나름대로의 기계론을 제시하였다. 그는 인간과 동물을 차별화 하였다. 동물은 마음이 없는 하나의 순수한 자동기계로 보았으나, 인간은 신체와 마음이 있으며, 신체는 동물과 같이 자동기계이지만 마음은 기계가 아니라고 생각하였다(동물기계 (bete machine)이론).

그는 예를 들어 원숭이와 같은 외모와 내부 기관을 지닌 기계를 만들었을 경우에 그 기계와 실제의 동물을 구별할 수 있는 방법이 없다고 하였다. 그렇지만, 인간의 경우는 아무리 어떤 기계를 만들어서 그 기계에 우리의 신체와 같은 외형과 기관을 갖도록 갖추어 준다고 하더라도 그 기계가 인간과 같을 수는 없다. 그 기계가 정말로 사람인지 아니면 단순히 인간의 외형을 지닌 기계인지를 알 수 있는 것은 두 개의 검사를 해보면 알 수 있다고 하였다(Gunderson, 1985). 그 검사의 하나는 ‘언어’ 검사이고, 다른 하나는 ‘지식-행위’ 검사이다. 아무리 기계를 인간처럼 만들어 놓더라도, 그 기계는 결코 우리와 같이 말이나, 생각이 담긴 상징부호(기호)를 사용할 수 없을 것이

며, 또한 아무리 인간과 유사하게 그럴싸하게 행동한다고 하더라도 단순히 신체 기관의 성향에서 행동이 나오는 것이지, 지식에서 행동이 나오지는 않을 것이다. 따라서 이 두 측면을 검사해보면, 그 대상이 인간인지 기계인지 알 수 있다고 하였다.

데카르트는 후에<sup>10)</sup> 전자, 즉 ‘언어’ 테스트만이 인간과 기계를 차별화 할 수 있는 것이라고 생각을 수정했다. 동물이나 기계가 인간을 따라 올 수 없는 이유는 그들이 언어를 인간처럼 여러 양상으로 배열할 수 없을 것이라는 것이다. 이 논지는 현대 인지과학에서 촘스키(N. Chomsky) 등의 인간 언어의 창의성, 생성성의 개념과 연관되어 다시 제시되고 있다.

## 2.c. La METTRIE의 기계론

데카르트 이후의 기계론을 이어 받은 J. La Mettrie<sup>11)</sup>는 데카르트의 기계론을 근본적으로 수정하였다. 데카르트가 동물은 순수한 기계라고 한 것은 옳지만, 인간은 기계가 아니라고 한 것은 틀렸다고 하며 그는 ‘인간기계(l’homme machine) 이론’의 관점을 다음과 같이 제시하였다<sup>12)</sup>

“데카르트가 ‘동물은 순수한 기계다’라고 논증하였다. 그러한 그의 생각은 옳다. 그러나 인간이 동물과 본질적으로 다르지 않다는 것을 보일 수 있다. 그리고 인간도 기계라는 것을 보일 수 있다. 동물과 인간 사이에 차이가 없다는 것을 보이는 방법의 하나는 인간에게 특유한 기술과 능력이라고 혹자들이 생각하는 것들을 동물도 (예를 들어 원숭이를 언어를 사용하도록 훈련시킨다거나), 기계도 (말하는 기계 사람) 공유할 수 있다는 것을 보여주는 것이다.

그렇다고 해서 동물과 인간이 동일한 정도의 지능을 갖고 있다는 것을 보여주는 것은 아닐 것이며, 그들과 인간의 지(知)에 본질적 차이는 없다는 것을 보여주는 것일 뿐이다. 인간의 메커니즘이 그저 더 크고, 더 미묘한 정도의 조직화와 복잡성을

10) 데카르트의 마음과 영혼에 대한 논의가 주로 담긴 책, [성찰]은, 초판 1641년, 재판 1642년, 3판(이전 책이 라틴어 판이고) 불어 판이 1647년에 출판되었다. 초판의 제목은 ‘제일철학에 관한 성찰, 여기서 신의 현존 및 인간 영혼의 불멸성이 증명됨’이고, 재판의 제목은 ‘제일철학에 관한 성찰, 여기서 신의 현존 및 인간 영혼과 신체의 상이성이 증명됨’, 불어판의 제목은 ‘제일철학에 관한 르네 데카르트의 성찰, 여기서 신의 현존 및 인간 영혼과 신체의 실제적 상이성이 증명됨’ 이었다. - 자료원: 데카르트(지음), 이원복(옮김)(2004) [성찰]에서.

11) La Mettrie 책 사이트:

[http://books.google.co.kr/books?id=LOVAAAAcAAJ&printsec=frontcover&hl=ko&source=gbs\\_ge\\_summary\\_r&cad=0#v=onepage&q&f=false](http://books.google.co.kr/books?id=LOVAAAAcAAJ&printsec=frontcover&hl=ko&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false) ; [불문]

[http://books.google.co.kr/books?id=hIRz71uLMdMC&printsec=frontcover&hl=ko&source=gbs\\_ge\\_summary\\_r&cad=0#v=onepage&q&f=false](http://books.google.co.kr/books?id=hIRz71uLMdMC&printsec=frontcover&hl=ko&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false) [영문1]; <http://cscs.umich.edu/~crshalizi/LaMettrie/Machine/> [영문2]

12) La Mettrie 에 관한 자료는 다음 웹 자료를 참조: [http://en.wikipedia.org/wiki/Julien\\_Offray\\_de\\_La\\_Mettrie](http://en.wikipedia.org/wiki/Julien_Offray_de_La_Mettrie) ; 한글자료 - <http://www.aistudy.co.kr/pioneer/LaMettrie.J.htm>0000

지니고 있다는 것뿐일 것이다. “우리는 결국 모두 기계인 것이다. 그러나 사고할 수 있는 기계라는 것이다.”

단순한 기계일 뿐이라고 데카르트가 명증했던 동물이 언어적 기술을 - 데카르트가 인간과 동물을 구별짓는 기준이라고 생각했던 언어를 - 보일 수 있게 함으로써 동물이 어떤 본질적인 측면에서 인간인 우리와 다르지 않다는 것과, 우리도 기계라는 것을 알게 된다. 따라서 동물기계는 인간기계로 변환되는 것이다. .... 데카르트는 동물이 기계임을 보였지만, 사고와 감정이 없는 순수한 기계임을 보인 것은 아니다. 동물은 우리 - 앎과 지능적인 창조물인 - 인간과 같고, 우리는 동물들과 같다. 모두 사고와 느낌을 갖게 되는 복잡한 기계인 것이다(La Mettrie (1774). *Oeuvres philosophiques*. Vol I. ; quoted in Gunderson (1985; pp. 23-24)).>

그렇다면 어떻게 <기계인간>을 만들 수 있으며, 기계에 사람의 심적 능력을 부여할 수 있을까? 이러한 물음에 대한 대답은 예로부터 인간이 만들어 온 자동기계, 계산기계의 발달 역사를 17세기를 중심으로 살펴봄으로써 찾아질 수 있다.

## 2.d. 17세기의 계산기 개념

인간이 <기계인간>을 만들려고 수없이 노력해 왔다는 것은 앞서 기술한 바와 같이 희랍신화, 이슬람권의 역사, 유럽의 기발한 생각을 지녔던 사람들에게서 찾아 볼 수 있다(McCorduck, 1979). 기계인간은 인류의 역사가 가진 오랜 꿈이었으며, 그 역사 속에서 시작품들이 만들어져 왔었다. 이러한 노력들이 단순한 시행착오식 수작업이 아니라, 이론적 기반을 지니기 시작한 것은 17세기부터였다. 신학자, 천문학자, 언어학자이며 동시에 수학자이었던 Wilhelm Schickard가 ([http://en.wikipedia.org/wiki/Wilhelm\\_Schickard](http://en.wikipedia.org/wiki/Wilhelm_Schickard)) 1623년에 “계산하는 시계”라는 세계 최초의 계산기를 만들었다. 10진법에 기초한 계산기였다. 그러나 이 기계는 당시와 후대에 영향을 못 주었고, 이후 파스칼, 라이프니츠 등이 새로운 계산기를 고안하여 냈다.<sup>13)</sup>

17세기 중엽에 이르러 유럽에서는 이후 20세기의 인지과학과 인지심리학에 영향을 줄 세 개의 중요한 발전이 있었다.

첫째는 체계적 규칙에 따라 수리적 계산을 하는 능력이 발전한 것이었다. 이러한 수리적 계산 체계의 발전은 과학적 법칙의 도출을 가능하게 하였고, 이것은 물리적 세계에 대한 이해도를 높였다. 둘째는 계산의 원리가 규칙적이고 엄밀히 규정될 수 있는 형태로 진행되기 때문에 이는 기계에 의해 수행될 수 있을 것이라는 생각의 발전이었다. 셋째는 이러한 체계적이며 명세화할 수 있는 계산과정이 인간의 이성적 사고의 본질과 닮았으리라는 생각(원래는 Hobbes의 생각이었음)의 떠오름이었다.

13) 1920년대까지의 계산기의 발전 역사를 개관한 책 자료는 다음을 참조 : [http://www.rechenmaschinen-illustrated.com/Martins\\_book/Ernst%20Martin%20-%20Rechen%20Maschinen%](http://www.rechenmaschinen-illustrated.com/Martins_book/Ernst%20Martin%20-%20Rechen%20Maschinen%20)

이러한 생각들은 이후 몇 세기를 거쳐 확장되고, 정교화되고 발전되었다. 계산기의 지속적인 발전은 마침내 20세기의 컴퓨터의 발명과 발전으로 이어졌고, 산술의 발전이 대수의 발전으로, 그리고 이것이 수학체계의 발전으로 이어졌다.

## 2.e. PASCAL의 계산기

1639년에서부터 6년 여 동안 초기 계산기를 만든 파스칼(B. Pascal)은 오늘날의 주유소 휘발유 카운터나, 수도 검침기와 같은 종류의 계산기를 1645년에 완성해 내었다. 주로 10진법에 기초하고 덧셈과 뺄셈만 할 수 있던 이 기계를 파스칼 기계(Pascalina 또는 the Arithmetique)라고 부른다.<sup>14)</sup> 파스칼은 데카르트와는 달리 인간의 사고도 비물질적 요소로만 되어 있는 것이 아니라 기계적 요소가 있는 것이 아닐까, 즉 데카르트의 생각은 틀렸고, 인간도 기계와 같을 수 있지 않은가 하고 생각하였다. 그러나 파스칼은 이러한 문제에 대하여 계속 생각한 결과, 결국은 부정적인 결론을 내렸다. 왜냐하면 인간의 마음은 이성과 감정, 의지로 구성되어 있는데, 이성만 기계적 요소를 지닐 수 있으나 감성과 의지는 이성에 의존하지 않고 작용하여 얹어 이루어지기에 인간의 마음이 기계적일 수는 없다고 결론지었다. 이는 오늘날의 일반인들이 컴퓨터와 마음의 관계에 대하여 기본적으로 지니고 있는 상식적 생각과 같다고 하겠다.

여기에서 파스칼 계산기와 관련하여서 우리는 의문이 들 수 있다. 그까짓 덧셈 정도나 하는 그 당시의 계산 기계를 근거로 해서 인간의 이성, 사고가 기계적인가를 판단하는 파스칼의 생각의 타당성 문제이다. 이에 대해서는 인간 지능 진화사의 상황적 설명을 제시할 수 있다. 17세기의 유럽의 일반인들은 현대의 초등학교 학생들이 하는 셈 정도도 잘 못하였다. 현대에는 초등학교 학생들이 백, 천, 만, 억 단위의 숫자를 사용해서 자유자재로 덧셈 뺄셈을 하고 있지만, 17세기 당시의 사람들은 아주 쉬운 덧셈 뺄셈도 하지 못하였다. 덧셈 뺄셈을 하는 것은 아주 지능이 높은 학자들이나 하는 것으로 되어 있었다. 그것도 숫자 열자리 범위 이내에서 그러했다. 그러한 이유는 이 당시에 유럽에서 아라비아 숫자가 사용된 것이 아니라, 로마자를 사용해서 수를 표현하였기 때문에 그에 따른 어려움 때문이었다.

이후 인도의 수 체계가 아랍을 통해 유럽에 수입되어 십진법과 영을 지닌 아라비아 수 체계를 사용하면서 일반인들의 산술계산 능력이 발달하였다. 이러한 바탕에서 갈릴레오는 과학은 수학의 언어로 설명해야 한다고 했고, 홉스는 인간 이성은 본질적으로 수리적 계산의 형태라는 주장을 제기하였다. 이에 따라 수리적 계산법, 계산체계가 발전되고, 계산은 규칙적이며, 그 단계를 세밀히 규정 가능하다는 생각과, 계산 원리

14) <http://en.wikipedia.org/wiki/Pascaline>



는 기계로 수행 가능할 것이라는 생각이 자리잡게 되었다. 체계적인 계산, 규정 가능한 계산이 곧 인간 이성적 사고의 특성일 수 있다는 생각의 바탕이 싹튼 것이다.

[그림 3]. 파스칼 계산기



## 2.f. LEIBNIZ의 기계론

“기계를 사용한다면 누구에게라도 맡길 수 있는 계산을 하기 위하여 계산하는 일에 노예처럼 수많은 시간을 소모하여야 한다는 것은 뛰어난 사람이 할 일이 못된다.”고 말하였던 라이프니츠(G. W. von Leibniz)는 언젠가는 계산기가 산술뿐만 아니라, 논리에서도 인간의 이성적 사고를 능가하리라고 생각하였다.

그는 다음과 같은 ‘보편언어’ 개념을 제시하였다. 숫자는 모든 종족, 국가, 문화에서 공용으로 쓰인다. 그렇다면 숫자와 마찬가지로 모든 국가, 민족에 공통적으로 쓸 수 있는 보편적 언어를 만들 수 없을까? 이러한 언어 특성은 본질적으로 논리적 위계 특성을 지닐 것이고, 그 언어의 개념들은 동식물 분류에서처럼 [내포]와 [배제] 관계로 표현할 수 있을 것이며, 이러한 내포, 배제의 개념적 조작이란 수학에서의 덧셈 뺄셈 조작과 같을 것이고, 그렇다면 모든 개념에 그 개념과 다른 개념의 포함 정도를 나타내는 숫자를 부여하여 그 개념을 나타낼(represent) 수 있지 않을까? 이러한 새 언어를 사용하여 모든 나라의 사람들이 의사소통을 하고 또 각종 문제해결을 할 수 있지 않을까? 국가 간의 논리적, 윤리적 문제 등이 생기면, 서로 논란할 필요없이, ‘우리 앞에서 계산합시다.’ 하는 식으로 그저 이 체계에서 계산만 하면 되지 않을까? 이러한 생각은 인간이 사고한다는 것은 계산하는 것 이상의 것이 아니다 라는 흙스의 생각을 발전시킨 것이었으며 현대 인지주의의 기본 생각과 상당히 근접해 있는 생각이었다.

이러한 생각을 하던 그는 1671년에 파스칼 계산기보다 한 단계 진보한 계산기를 고안하였다(제작은 1673년). [독어로 Staffelwalze, 영어로는 Step Reckoner]라고 불리는 이 계산기는 덧셈과 뺄셈만 하는 파스칼 계산기와는 달리 곱셈과 나눗셈도 할 수 있는 계산기였다.<sup>15)</sup> 이 기계는 반복된 덧셈과 자리 올리기(shifting)를 통하여 곱셈도 하는 계산기였다.

15) 라이프니츠 계산기 그림: <http://www.diycalculator.com/imgs/hist-step-reckoner.jpg>

라이프니츠의 또 다른 중요한 기여는 이진법 체계의 제시였는데 그는 계산기가 이진법을 도입하여야(오늘날의 디지털 컴퓨터와 같이) 한다고 주장하였으나, 당시 기술계의 기술 수준이 이진법 체계의 기계를 지원할 수 있는 수준이 되지 않아서 그의 기계는 10진법에 기초한 체계에 머물렀다. 라이프니츠는 2진법을 강력히 옹호하였다. 2진법은 두 개의 수자만 있기에 스위치의 개폐(On & Off)에 의하여 쉽게 나타낼 수 있다는 생각이었다. 그러나 이진법을 계산기에 구현하며, 규칙 기반의 기호논리 체계에 연결하는 것은 20세기에 이르러서 컴퓨터 발달과 기호체계이론의 발달이 이루어진 후에야 가능하였다. 계산기가 전자식으로 되자, 이 2진법과 전자회로의 'On-Off'를 연결하는 아이디어는 아주 잘 맞았다. 명제의 진위 표현에 회로의 'On'과 'Off'를 연결시켜 생각하면 되었다. 즉 계산기의 전기회로가 명제논리의 흐름을 나타낼 수 있게 되는 것이다. 이러한 생각은 아주 중요한 생각이기는 했지만 앞서 이야기한 대로 라이프니츠는 자신의 계산기에 이 2진법을 도입하지 않고 10진법을 사용하였었다. 300년 앞서 간 라이프니츠의 생각을 뒷 받쳐 줄 테크놀로지 분야의 생각이 모자랐던 것이다.

또한 계산기 기계 제작을 넘어선 라이프니츠의 생각의 핵심은 논리적 진술을 수학으로 환원하는 데 있었으나, 이후의 19세기의 논리학계의 연구는 다른 방향으로 진행되어 전통적 수학과 논리학을 기호논리학의 하위 영역으로 접근하는 방향으로 흘러갔다.

### 3. 튜링기계 이전 19세기와 20세기의 형식화 시도들

#### 3.a. BABBAGE: 형식 기계론

이러한 라이프니츠의 생각은 19세기 중반의 배비지(Charles C. Babbage)의 계산기계 이론에 의해 일보 진척이 되었다.

---

[표 1]. 배비지(Babbage)의 분석기계의 구성요소

---

1. 제 1 요소: - 자료 (data)
  - 지시 (instructions)
2. 제 2 요소: 계산 방앗간 (mill)
3. 제 3 요소: 제어기재 (Control Mechanism);
  - 입력을 받아 계산 방앗간에 전달하고 계산 방앗간이 계산하도록 통제
4. 제 4 요소: 기억 (memory) 요소

- 원래의 자료와 계산 결과를 저장

## 5. 제 5 요소: 출력기구 (output device)

---

배비지는<sup>16)</sup> 모든 수학적 계산표들을 기계에 의해 할 수 있으리라고 생각하고, 처음에는 [차이기계 (Difference Engine)]를, 그 다음에는 [분석기계 (Analytical Engine)]를 제안하였다. 후자의 이론은 100년 앞선 생각이었다. 이 기계는 이론적으로 어떤 유형의 계산도 가능한 계산기였으며, 데이터, 계산절차로서의 지시(instructions), 계산 순서에 따라 진행되는 실제 계산의 세 요소를 갖춘 이론적 기계였고, 하나의 보편 기계이론이었으며, 오늘날의 프로그램 가능한 디지털 컴퓨터의 이론적 원형이었다고 할 수 있다.

그러나 당시에는 실제로 이를 기계로 구현하는 공학기술이 없었다. 배비지의 생각은 당시의 하드웨어 공학기술이나 이론 수준보다 앞선 것이었기에 실현되지 못하였다. 아마도 그 당시에는 기관차 정도 크기의 계산기와 기관차 정도의 동력이 필요하였을지도 모른다. 그의 분석기계의 생각은 100년이나 앞선 생각이었다. 상응하는 공학기술이 당시에 없었기에 그의 창의적인 생각은 아이디어로만 존재할 수밖에 없었다 (물론 그가 만든 작은 모형은 있었지만).

예외적으로 그를 알았던 당시의 러브레이스(Ada Lovelace) 백작부인<sup>17)</sup>은 배비지의 기계의 이론을 이해하고 이 기계가 오늘날의 디지털 컴퓨터와 같이 여러 유형의 문제 해결에 적용될 수 있을 가능성을 인정하였으며(Babbage 자신은 이런 의의를 충분히 파악하지 못하고 수리적 처리 중심으로 생각하였던 것 같다.) 이론적 발전에 도움을

---

16) Babbage의 생애, 차이기계, 분석기계 그림 자료: <http://www.charlesbabbage.net/>

17) Augusta Ada (King) Lovelace 백작부인(1815 - 1852)은 마음-기계 연결의 역사에서 두드러진 인물이다. 그녀는 세계 최초의 컴퓨터 프로그래머라고 평가되고 있다. 영국 시인 바이론의 딸로서 세 아이의 엄마이기도 했고 그는 어릴 시절 병약하였고 36세에 암으로 사망하기는 했지만, 아마튜어 수학자이던 그의 어머니의 계획에 따라 여러 가정교사(튜터)에게서 음악과 수학을 배웠다. 17세에 한 파티에서 배비지의 [분석 엔진]에 대한 입문적 설명을 듣고는 그 아이디어에 매료되었다. 그는 1843년의 한 이태리의 공학도인 루이지 메나부리아(Luigi Menabrea)의 관련 논문을 불어로 번역하면서 거기에 자신의 주석 노트를 달았다. 바로 이 노트에서 러브레이스 백작부인은 수학문제를 푸는 단계적 과정 (오페레이션; step-wise program)을 세계 최초로 기술하였다. 이것이 후세대 학자들에 의해 세계 최초의 컴퓨터 프로그램 또는 컴퓨터 알고리즘이라고 평가되는 내용이다. 그와 친하였던 찰스 배비지는 자신의 [분석엔진] 이론 틀을 제시하며 그저 계산이나 수를 다루는 엔진으로만 생각하였지만, 러브레이스 백작부인은 이 이론적 기계가 그러한 수 계산을 넘어서 여러 가지의 큰 가능성과 의미를 지니고 있음을 깨달았다. 현대의 컴퓨터의 가능성을 백년 이상 앞서서 예견한 것이다. 현대적으로 이야기하자면 러브레이스 백작부인이 제시한 것은 20세기의 초기의 컴퓨터 시기에는 펀치카드에 구멍을 내어 입력되는 지시 프로그램 수준이라고 볼 수 있지만, 그가 현대적 컴퓨터 프로그래밍 아이디어를 20세기의 인지과학이나 인공지능이 출발하기 백 여 년 전에 이미 제시하였다는 것은 부정하기 힘들다. 이 러브레이스 백작에 대한 상세한 이야기를 보려면 다음 사이트의 그에 대한 년보, 일대기, 이야기 자료를 참조하기 바란다.: <http://www.ideafinder.com/history/inventors/lovelace.htm> . 위키피디아의 자료는 다음에 있다. : [http://en.wikipedia.org/wiki/Ada\\_Lovelace](http://en.wikipedia.org/wiki/Ada_Lovelace)

주었으며, 이 이론적 기계를 위하여 세계 최초의 프로그램(기호 조작)을 작성하였다.

### 3.b. BOOLE: 논리적 사고의 형식화

이어서 부울(G. Boole)은 모든 수학적 작업은 체계적 기호(상징) 조작의 한 형태이며, 수학은 수와 양의 학문이 아니라, 기호를 사용하는 한 방법이라고 규정하였다. 라이프니츠의 생각에 따라, 논리학과 계산 수학을 조합하여 기호논리학을 새롭게 출발시킨 것이다. 이러한 입장은 20세기 초의 철학자 화이트헤드(N. Whitehead)와 러셀(B. Russell)에 의하여 더 발전된다. 이러한 일련의 작업들은 ① 모든 전통적 수학/산출 계산과 ② 기호(상징)화 가능한 모든 논리와 추리의 문제를 계산해 낼 수 있는 보편적 기계(Universal Machine)가 가능하다는 생각의 바탕이 되는 작업이었다고 볼 수 있다.

다음 단계는 이러한 기계와 마음에 대한 생각의 가능성을 보다 구체적인 개념적 기계이론으로 정교화 하는 작업이었다. 이 작업은 20세기에 들어와서 1930년대에 튜링(A. Turing)을 중심으로 한 튜링기계 이론이 제시되어서야 결정적으로 이루어졌다.

물론 1930년대에 튜링기계이론이 제기되기 이전에도 마음과 기계를 연결하려는 여러 시도들이 20세기 초에 있었다. 스페인의 케베도(L. Torres Quevedo)는 20세기 초기에 스페인의 왕립 학회 회원으로서 서양장기를 두는 기계를 발명했으며, 생각하는 기계에 대한 아이디어를 갖고 있었다. 그는 상황의 변화에 대해서 적절하게 잘 대응하는 기계는 감각기관들과 팔다리와 같은 것들을 가져서 감각기관을 통해 들어온 정보들에 따라 일을 수행할 수 있어야 한다고 생각하였다. 그러나 당시의 기계 제작 수준과 제어 수준이 조잡했기에 이토록 지능적인 기계의 탄생은 후에 전자공학에서 집적 회로가 개발된 이후에나 가능해졌다.

또한 컴퓨터 미국, 영국 등이 가장 최초로 프로그램으로 제어되는 범용 디지털 컴퓨터를 만들었다고 생각하기 쉽겠으나 독일에서 이미 2차 대전이 끝나기 전에 이러한 기계가 기계공인 쭈제(Konrad Zuse)에 의해 제작되었다. 쭈제는 라이프니츠처럼 귀찮은 계산을 대신해 줄 기계를 고안하기 위해서 착수했으며 생각하는 기계에 대한 가능성에 대해선 분명한 관점을 갖고 있었다. 그는 1941년에 세상 처음으로 프로그램에 의해 제어되는 T3라는 튜링컴퓨터를 만들었다. 또 'Plankalkül'이라고 이름 붙인 프로그래밍 언어도 개발했다. 이외에도 마음과 기계를 연결하려 한 다른 여러 시도들도 있었다.

## 4. 현상, 형식 체계. 기계, 수학적 사고

#### 4.a. 현상과 형식적 기술(記述)

튜링 기계론을 설명하기에 앞서서 설명되어야 할 것은, 현상에 대한 수학적 기술이론(theory of mathematical description)이다. 이 이론에 의하면 과학적 방법은, 본질적으로, 자연현상을 수학적으로 기술하는 기술 유형의 하나를 선택하는 방법이라고 볼 수 있다.

자연현상은 무한할 수도 있고 유한할 수도 있다. 현상이 유한하면서 수학적으로 유한히 기술될 수도 있고 무한하면서 유한히 기술될 수도 있다. 자연현상을 수학적으로 기술하려는 입장의 핵심은 무한한 또는 유한한 현상을 유한히 기술할 수 있다고 가정하는 것이다.

수학에서는 현상에 대한 기술의 개념이 무한 기술(記述)의 필요성에서 유한 기술로, 유한 기술에서 함수들 또는 규칙 집합(rule-sets)의 기술의 개념으로 옮겨간다. 그런데 미적분학에서 보듯이, 모든 비연속적 현상은 특정 기호(스트링)에 자연수를 연결시켜주는 함수에 의해, 그리고 어떤 자연수도 기호들(스트링들)에 의해 나타내질 수 있다.

현상을 기술하기 위하여는 현상의 발생의 관찰이 선행되어야 한다. 현상이나 과정이란 현실에서 발생하는 사건들이며, 발생이란 '일련의 상황들'로 간주될 수 있으며 현상이란 '상황들이 발생할 수 있는 모든 가능한 양식의 집합'으로 규정될 수 있을 것이다. 이는 한 현상은 다른 현상으로 구성될 수 있음을 의미한다.

따라서 현상에 대한 수학적 기술의 첫 단계는 그 현상의 모든 가능한 상황의 집합  $X$ 와 시간 척도  $T$ 를 규정하는 것이다. 상황 집합을 구성하는 요소 상황을  $\theta_i$ , 시간 척도  $T$ 내의 한 요소 시간을  $T_i$ , 현상의 발생을  $\theta$ 라고 한다면 발생 $\theta$ 는  $T_i$ 와  $\theta_i$ 를 상응시켜주는 함수라고 할 수 있으며, 현상이란 곧  $\theta_i$ 의 집합, 즉  $\{\theta_1, \theta_2, \theta_3... \theta_i\}$ 이라 할 수 있다.

고로 어떤 현상의 기술이란 발생  $\theta$ 의 집합을 기술하는 것이다. 그렇게 되면 현상에 대한 완벽한 기술이란 발생  $\theta$ 의 가능한 무한 집합을 기술하는 것이며, 유한 기술을 그 근분으로 하는 수학적 기술이란 그 현상의 발생집합을 유한하게 기술하는 것이다. 이는 모든 발생  $\theta_i$ 들을 열거하거나 나열하는 것이라고 하기보다는, 이  $\theta_i$ 들을 임의의 정확수준 정도로 계산할 수 있게 하는 유한 규칙 또는 함수를 위한 '기술(descriptions)로서 받아들임을 의미한다. 이때의 계산이란 정확히 효율적으로 규정할 수 있다는 의미가 된다.

고교 시절의 수학에서 미적분과 한계 개념에서 제시되는 것에서 나오듯이, 확률이론에 의하면, 연속적 현상을 비연속적 현상을 사용하여 임의의 정도까지 추정하고 기술할 수 있다. 이는 연속함수 현상을 이산(비연속 discrete)함수인 계단 함수로 기술하는 것이며 후자는 다시 기호의 계열이나 기호의 집합 단위인 스트링(기호집합, strings)으로 기술할 수 있음을 의미한다.

다시 말하여, 연속적 현상은 비연속적 현상으로 기술할 수 있으며, 이 비연속적 현상의 발생은 유한히 기술될 수 있고, 이는 특정 기호 집합단위 또는 스트링의 집합을 유한히 기술하기만 하면 된다고 하겠다. 기술하고자 하는 현상의 발생 집합이 유한하다면 과학적 연구자는 그저 그에 해당하는 모든 스트링을 열거하면 된다. 그 집합이 무한하다면 그 모든 경우에 대해 제  $i$ 번째 스트링을 계산해 낼 수 있는 규칙(유한히 효율적으로 기술할 수 있다는 규칙), 또는 함수를 찾아내면 된다.

그렇다면 현상에 대한 기술(記述)의 개념이 무한 기술의 필요성에서 유한 기술로, 유한 기술에서 함수들 또는 규칙 집합(Rule-Sets)의 기술의 개념으로 옮겨간다. 그런데 모든 비연속적 현상은 특정 스트링에 자연수를 연결시켜주는 함수에 의해 나타낼(표상될) 수 있다. 그리고 어떤 자연수라도 스트링에 의해 표상될 수 있다. 따라서 결론적으로 말하자면, 현상의 기술의 문제란, 함수 또는 규칙집합에 대한 유한 기술의 문제, 다시 말하여 한 스트링 집합과 다른 스트링 집합을 대응(사상, mapping)시켜주는 문제이다.

결론적으로 말하자면, 과학에서 자연현상을 기술하는 문제란, 수학적으로 표현하자면, 함수 또는 규칙집합에 대한 기술의 문제, 다시 말하여 한 스트링 집합과 다른 스트링 집합을 대응(사상, mapping)시켜주는 문제이다.

그런데 한 스트링 집합과 다른 스트링 집합을 대응시키는 함수를 다루는 수학적 이론이 자동기계(automata)이론이며, 이 자동기계이론 중에 가장 대표적인 이론이 계산가능성(computability) 이론이며(이에 대한 더 자세한 설명은 이정모(2001)의 154-157 참조) 그의 대표적 형태가 튜링기계 이론이다.

#### 4.b. 기계와 형식적 기술(記述)

기계의 개념을 정의하자면, 명백한 일련의 규칙에 의해 일련의 조작을 수행하는 기구이다. 기계체계에 내장된 조작의 유형과 기본가정이 일정한 수라면 우리는 이들 모두를 적절한 기호로 표상하여 종이에 적어 볼 수 있다. 최초의 기본 가정들은 공리라든가 계와 같은 기본 공식으로 표상될 수 있으며, 하나의 조작은 그 조작이 일어나기

전 상태와 후의 상태를 나타내는 공식과, 어떠한 규칙이 적용되었는가를 명시함으로써 표상될 수 있다. 기계의 조작들이 아무리 많고 복잡하더라도 충분한 시간만 주어 진다면 이러한 조작계열의 아날로그를 기록할 수 있다. 그리고 이러한 아날로그는 형식적 증명이 된다.

즉 기계의 조작 하나 하나가 규칙의 적용에 의해 표상된다. 또 일정상황에서 기계가 어떤 조작을 수행할 것인가 여부를 결정하는 조건들은, 이 표상에서 일정 공식에 어떤 규칙이 적용 될 수 있는가 여부를 결정해 주는 조건, 즉 적용성의 형식적 조건(formal conditions of applicity)이 된다.

이러한 규칙을 추론의 규칙으로 간주함으로써, 하나하나의 공식이 이전의 공식 또는 공식들에 어떤 형식적 추론 규칙을 적용하여 도출되는 증명 계열을 획득 할 수 있다. 따라서 한 기계가 산출해 낼 수 있는 조작 계열, 즉 결론들이란 그 기계와 대응하는 형식체계 내에서 증명될 수 있는 정리와 상응한다. 즉 기계의 조작 결과인 출력은 한 형식체계에서 도출된 정리에 해당한다는 것이다.

기계에 있어서의 계산 가능성의 개념을 명확히 해주고 형식 체계를 활용하여 그것이 심리 현상의 기술에까지 적용된 대표적 이론이 바로 [튜링기계이론]이다. 튜링기계 이론을 설명하기에 앞서서 수학적 사고 과정을 정형화하여 기술하려고 노력한 시도들을 살펴보자.

#### 4.c. 수학적 사고와 형식적 기술(記述)

사람들은 각종의 논리적 추론을 하며 창의적 사고를 한다. 또한 수학자들은 각종 수학적 문제를 제기하고 이에 대하여 연역적으로 추론하여 문제를 해결한다. 그들은 어떤 한 명제의 참과 거짓을 증명을 통하여 밝힌다. 이러한 추론적 사고, 수학적 추론은 어떻게 이루어지는 것일까? 이러한 사고 과정을 이어가는 절차들을 명확히 설명할 수는 없을까? 즉 어떤 완전한 형식적(formal) 틀을 갖춘 일관성 있는 공리적 체계에 의해 이 과정들을 단계, 단계 정확하고 엄밀하게 기술할 수는 없을까?

이러한 물음은 예전부터 수학자, 논리학자, 철학자이 그리고 최근에는 인지심리학자와 인공지능학자들이 관심을 가져온 문제이다. 각종 수학적 문제의 제기와 증명의 추론과정을 완전히 형식화된 어떤 공리와 규칙의 체계에 의해 설명할 수 있다면 그것이 지니는 의의는 큰 것이다.

이러한 물음에 대하여 수학자들은 기계론적 입장을 취한다. 그들에 의하면 수학적 추론이란 기계적이다. 수학적 개념이나 명제를 형식화할 수 있고, 주어진 비공식적 증

명을 어떤 형식체계 내에서 점검할 수 있는 형태로 형식화할 수 있으며, 수학적 증명 방법을 형식체계 내의 잘 규정된 절차로 환원할 수 있다는 것이다.

그런데 이러한 형식체계는 기계로 간주할 수 있으며 (쇠로 만든 구조물만이 기계의 전부가 아니다. 입력과 출력 사이의 과정과 관계를 엄밀히 규정할 수 있는 모든 시스템이 기계인 것이다). 이 체계 내에서의 정리와 산출과정은 기계적 조작(연산; operations)으로 간주할 수 있다. 그러므로 모든 (수학적) 추론은 기계화, 형식화할 수 있다.

동시에 이를 확대하여 해석하면 수학적 추론의 상위 체계인 인간의 마음도 기계로 간주할 수 있으며, 마음의 작용도 기계화, 형식화할 수 있다고 주장할 수 있다. 이러한 기계론을 체계적으로 강력히 전개한 것이 튜링이었으며 이에 반대되는 입장으로 해석된 것이 괴델(K. Gödel) 정리의 확대 해석이다.

#### 4.d. 마음과 컴퓨터: 계산과 형식체계

다른 책에서도(이정모, 2009, 2010) 설명하면서도 언급하였지만 인지과학의 주요 두 핵심 개념은, 마음의 내용을 지칭하는 개념인 표상(representation)과, 마음의 과정을 지칭하는 개념인 계산(computation)이다. 그리고 이 두 개념의 밑바탕에 놓여있는 새 패러다임인 정보처리 패러다임의 생각은 표상의 양식 및 계산 과정의 표현 또는 규정을 일정한 형식을 부여하여 접근할 수 있다는 것이었다.

조금 단순하게 생각을 하여 보면 마음과 컴퓨터(일종의 기계)가 다 같이 ‘언어’적 특성을 지니고 있다고 할 수 있다. 내용 요소가(마음의 내용, 컴퓨터의 데이터) 있고 그 요소에 무언가의 작용이(심적 과정, 컴퓨터의 연산과정) 일정한 규칙에 의존하여 일어난다고 볼 수 있다. 정보처리적 패러다임(information processing paradigm)의 인지과학은 이 마음과 컴퓨터의 요소들과 작용과정들에 대하여, 객관성을 지닌, 어떤 같은 방식으로 표현하며, 기술할 수 있을 것이라고 가정하고 접근하였다. 과학적 객관성과 엄밀성, 단순성을 담보하는 어떤 방식으로 기술할 것인가를 생각하는 과정에서 도출된 것이 ‘형식성’이라 할 수 있다. 마음과 컴퓨터의 내용들, 그리고 그에 작동하는 과정들을 정보처리적 틀에서 기술하는 것의 첩경은 그 대상들을 형식적(formal)으로 기술하는 것이라고 본 것이다.

그러면 어떠한 틀의 형식적 기술을 할 것인가, 마음과 컴퓨터가 기본적으로 어떤 내용적, 작용적 특성에서, 형식적 기술 측면에서 관련된다고 볼 수 있을 것



인가 하는 물음과 관련하여 고전적(classical) 인지과학의 생각의 핵심을 이루고 있는 것이 ‘튜링기계’ 이론이다. 튜링기계이론은 영국의 수학자 튜링이 단독으로 이루어내었다고 하기보다는 수학 내에서, 그리고 논리학의 도움을 받아, 어떤 대상을 규정하고 기술한다는 것이 무엇인가, 무엇을 기술할 수 있는가 하는 문제에 대한 오랜 생각들이, 특히 수학자들(A. Church, D. Hilbert, K. Gödel 등)의 생각이 모아지고, 이를 발전시켜 새로운 형태로 전개한 것이라고 할 수 있다.

## 5. 튜링기계 이론

### 5.a. 튜링의 삶

튜링기계 이론을 제시한 수학자 앨런 튜링은 오늘날의 컴퓨터가 주판과 같은 단순히 숫자 계산 기계에 머물지 않고 - 21세기 현 시점에서 생각하자면 인터넷과 SNS를 가능하게 하고 모든 영역에서 디지털 컴퓨터가 각종 유형의 정보처리를 담당하는 것이 당연한 이야기이지만, 60 내지 80 여 년 전 20세기 초엽만 하여도 디지털 컴퓨터의 존재는 물론, 그 광범한 활용의 인간 세상이 아니었다 - 인간의 마음, 특히 사고를 흉내 내며 정보처리를 하는 기계로 바뀌는 개념적 전환, 발상의 전환의 기초를 놓은 사람이다.

그는 수학자이며 논리학자이었고, 수학 문제에 대한, 전통과 어긋나는 특이한 해결 방법을 갖고 있었다. 튜링은 2차 세계대전 중에는 울트라(Ultra) 암호 프로젝트에 참가하여 5bit 체계인 기계를 통해서 독일의 암호 장치인 Enigma의 암호 체계를 풀어냈다. 대전 후에 영국의 국립 물리 연구소에 ACE(Automatic Computing Engine)를 개발키 위해서 참여하였다. 세계 최초의 저장된 프로그램 컴퓨터인 “Madam”(Manchester Automatic Digital Machine)의 개발에 참여하기도 하였다.

그에게 암호 해독은 세상의 물리 법칙에 대한 하나의 비유에 해당되는 것이었고, 학습에서 기계의 학습과 인간의 학습은 유사하다고 보았으며, 기계에게 처벌과 강화에 의한 조건 형성, 동기의 부여를 할 수 있다고까지 생각하였다. 그는 인간을 흉내 낸 기계보다는 기계와 비슷한 인간이라는 생각을 가졌었다.

그는 비록 현대의 범용적 디지털 컴퓨터의 기초 이론을 제시하였고 디지털 세대를 여는 이론적, 개념적 기초를 놓은 그였지만, 동성애자였던 그는 결국은 영국의 반 동성애적 문화와 제도를 이겨내지 못하고 젊은 나이에 자살하였다.<sup>18)</sup>

### 5.b. 튜링기계 이론: 서론

---

18) 튜링에 대한 한글 자료는 <http://windshoes.new21.org/person-turing.htm> 에 자세히 소개되어 있다.

튜링은 1937년에 “On Computable Number with an Application to the Entscheidungs-problem(결정문제에 적용된 계산가능 수)”<sup>19)</sup>이라는 논문을 통해 어떤 수학 문제들은 고정되고 제한된 명료한 처리 과정들에 의해선 풀릴 수 있다는 것을 증명했다. 여기서 “제한되고 명료한 처리”란 그가 제안한 보편적이고 추상적인 기계가 할 수 있는 처리를 논한다.

그가 제안한 이론적 기계가 Turing기계인데, 이 기계는 뒤에서도 다시 설명되겠지만 무한히 긴 종이테이프를 갖고 있고 그 테이프는 다시 사각형 칸들로 구분되며, 각 칸들은 숫자나 빈 공백으로 채워져 있어서 한 번에 한 칸씩만 스캔할 수 있는 이 기계가 그 종이테이프를 앞으로 뒤로 왔다갔다 읽어 가거나 혹은 지우고 써가면서 작업을 하도록 개념화 되어 있다.

Turing은 이처럼 아주 단순한 기계가 2진 표기 체계로 되어 있는 그 어떠한 프로그램도 실행시킬 수 있음을 보였다. 어떤 종류의 과제이건 간에 그것의 해결을 수행하기 위해 필요한 단계 과정들을 명확히 표현할 수 있다면, 그러한 과제나 문제는 모두 그가 제안한 보편기계 튜링기계에 2진법 형식으로 나타내어 그 문제를 해결할 수 있다는 것이다.

이러한 튜링기계 이론은 수학과 계산이론, 컴퓨터과학, 인공지능 연구에 지대한 의의를 지니는 개념이었다. 이 개념은 배비지의 ‘분석엔진’에 유사한 것이었고 라이프니츠의 2진법적 계산기 개념을 발전시킨 것이기는 하지만, 이후의 모든 유형의 디지털 컴퓨터를 만들 수 있는 개념적 모델을 제공한 그러한 이론이며 생각이었다.

### 5.c. 튜링기계 이론: 세부

튜링기계 이론의 세부 내용을 살펴보면 다음과 같이 설명할 수 있을 것이다.<sup>20)</sup>

수학은 형식체계로 이루어졌다고 볼 수 있다. 일반적으로 형식체계란 일련의 기호 요소들(알파벳)의 집합과 이들을 구성하고 변환시키는 명백하고 유한한 알고리즘 규칙 또는 절차로 구성된다. 알고리즘 또는 효율적 절차(effective procedures)란 어떤 수리적 조작을 수행하는 기계적 규칙 또는 자동적 방법 또는 프로그램을 의미한다. 효율적 절차란 그 절차가 단계적으로 수행되면 일정한 유한 수의 단계를 거친 후에 출력이 나온다는 의미를 가진다.

19) <http://web2.comlab.ox.ac.uk/oucl/research/areas/ieg/e-library/sources/tp2-ie.pdf>

20) 이 내용들은 실체는 추상수학적(대수) 이론이라고 할 수 있다. 관련 한글자료는 국내 인공지능 연구자들의 사이트 [http://www.aistudy.co.kr/linguistics/turing\\_linz.htm](http://www.aistudy.co.kr/linguistics/turing_linz.htm)에 제시되어 있다.

형식체계란 유한히 기술될 수 있는 체계를 의미한다. 유한히 기술될 수 있다 라고 하는 것은 한 형식체계 내에서, 알고리즘적 유한 절차들을 기계적으로 적용하여 주어진 상징 스트링이 규칙에 잘 맞는가(well-formed), 또는 공리(axioms)인가를 결정할 수 있고, 또 어떤 진술이 규칙에 잘 맞는 유한 진술 집합에서 도출될 수 있는가를 결정할 수 있다는 것이다.

수학에서 거론되는 개념으로서의 기계는 자동차와 같은 쇠로 만든 하드웨어가 아니라, 어떤 명백한 일련의 규칙에 의해 일련의 조작을 수행하는 기구이다. 기계체계에 내장된 조작의 유형과 기본가정이 일정한 수라면 우리는 이들 모두를 적절한 기호로 표상하여 종이에 적어 볼 수 있다. 최초의 기본 가정들은 공리라든가 계와 같은 기본 공식으로 표상될 수 있으며, 하나의 현상은 조작이 일어나기 전 상태와 후의 상태를 나타내는 공식과, 어떠한 규칙이 적용되었는가를 명시함으로써 나타내어질 수 있다. 기계의 조작들이 아무리 많고 복잡하더라도 충분한 시간만 주어진다면 이러한 조작 계열의 아날로그를 기록할 수 있다. 그리고 이러한 아날로그는 형식적 증명이 된다. 즉 기계의 조작 하나 하나가 규칙의 적용에 의해 표상된다.

일반적으로 수학에서의 기계란 유한 자동기계 (finite automaton)를 지칭한다 (Arbib, 1964; McNaughton, 1982). 유한 자동기계란 어떤 유한 수의 입력을 받아들일 수 있고 유한 수의 내적 상태를 지니고 있으며, 어떤 유한 수의 출력을 내어놓을 수 있는 체계이다.

유한 자동기계를  $A$ , 유한 수의 입력을  $I$ , 유한 수의 내적상태를  $q$ , 유한 수의 출력을  $O$ ,  $q$ 와  $I$ 가 상호작용하여 결정하는  $A$ 의 다음상태 ( $q \times I \rightarrow q'$ )의 함수를  $\lambda$ , 다음 출력의 함수 ( $q \times I \rightarrow O$ )를  $\delta$ 라 한다면

$$\text{자동기계 } A = ( I, O, q, \lambda, \delta ).$$

의 형식으로 표현할 수 있다. 이러한 자동기계는 불연속적 시간 척도 상에서 작용하는데, 만일 지금 시간  $t$ 에서 상태  $q$ 에 있었고 입력  $a$ 를 받는다면 시간  $t+1$ 에서는 상태  $\lambda(q, a)$ 로 바뀌고  $\delta(q, a)$ 를 출력으로 내놓게 된다. 이러한 기계를 이산상태 (discrete state) 기계라고 할 수 있다.

이러한 형식체계  $T$ 의 알고리즘적 규칙들만 주어진다면 체계  $T$ 의 정리들을 기계적으로 하나씩 산출하여 열거할 수 있는 기계  $T_m$ 을 구성할 수 있다. 마찬가지로, 무한한 기억을 지닐 수 있는 컴퓨터가 있다면, 그 기계의 출력이 형식체계  $T_m$ 의 정리들과 동일한, 어떤 특정 형식체계  $T$ 를 발견할 수 있을 것이다. 그렇다면 어떤 형식체계도 기계로 간주할 수 있으며, 역으로 어떤 기계도 형식체계로 간주 할 수 있다는 정리를 세울 수 있다.

이러한 틀에 의하면, 어떤 튜링기계  $T_m$ 의 기계표에서  $\langle I_i, q_i, O_i, q_j, I_i \rangle$ 라는 내용, 즉

(IF 입력이  $I_i$ 이고, 현재상태가  $q_i$ 이면,  
THEN 출력은  $O_i$ 로 출력하고,  
현재 상태는  $q_j$ 로 바꾸고,  
다음에 테이프를  $I_i$ 로 옮기라.)

는 내용 전체를 0와 1를 조합하여 표현할 수 있을 것이다 (예: 01011). 이를  $D_m$ 이라고 하자. 보편튜링기계  $U_m$ 의 행동은, 한 튜링기계  $T_m$ 의 실제 현재상태  $q_i$ 와 입력 내용  $I_i$ 를 확인한 후에  $D_m$ 에 따라 수행될 수 있다. 즉

(If  $I_i, q_i, D_m$ :  
THEN DO  $\langle O_i, q_j, I_i \rangle$ .)

의 절차를 수행할 수 있을 것이다.

이러한 수행이란 인간의 사고가 튜링기계  $T_m$ 의 행동을 추적하여서 나타나는 것과 본질적으로 같은 절차에 의해 이루어진다고 하겠다. 그리고 보편 튜링기계란  $D_m$ 과 같은 상징내용에 의해 다른 튜링기계를 흉내 낼 수 있는 기계라고 할 수 있다. 이는 보편 튜링기계가 보편목적적(범용적) 디지털 컴퓨터와 대등하다는 의미가 된다.

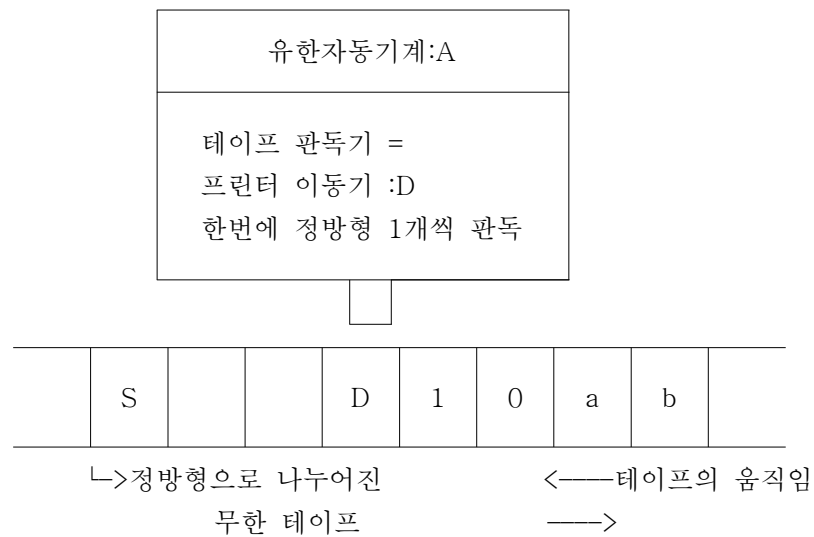
보편튜링기계에 주어진  $D_m$ 이란 스트링은 디지털 컴퓨터에 넣는 프로그램으로 볼 수 있고 이 프로그램이란 튜링기계  $T_m$ 의 한 스트링과 다른 스트링을 대응시키는 함수로 볼 수 있다. 그렇다면 이러한 프로그램을 지닌 컴퓨터란 유한히 기술할 수 있는 상징 조작과정을 구현화하는 보편 튜링기계라고 할 수 있게 된다.

튜링(Turing, 1936)은 이러한 이산상태기계의 논리를 근거로 그의 튜링기계의 이론을 제시하였다. 튜링기계란 현대적 용어로 디지털 컴퓨터가 할 수 있는 계산을 모두 할 수 있는 유한자동기계이다. 다시 이야기 하여 튜링기계에는 입력기구인 매체 테이프가 있고, 무한 용량의 기억창고가 있으며, 출력을 매체 테이프 위에 하고, 테이프 위에 이미 쓰여져 있는 기호들을 훑어서 읽고, 출력되어야 할 또 다른 기호를 테이프 위로 인쇄하며, 또 테이프를 좌우로 움직이게 하는 기구 (D)가 있다.

관독되는 사각형(정방향)의 내용과 각 순간의 기계의 상태가 그 순간의 기계의 전체 형태 (configuration)를 결정한다. 전체 형태란 한 순간의 기계의 상태와 입력된 정보

와 이를 관독하고 있는 입력기의 부분으로 구성되며, 유한 수의 내적 전체형태들이 있을 수 있다. 현재의 전체형태가 다음에 어떤 조작 (operations)을 할 것인지를 결정한다. 조작은 관독되는 정방형의 내용을 바꾸거나 그 정방형을 좌우로 움직이거나 현재상태를 다른 상태로 바꾸거나 정지하거나 한다. 정지한 때의 테이프의 내용을 출력이라 한다(그림 2 참조).

[그림 2]. 튜링기계



이러한 조작을 수행하는 튜링기계는 유한 수의 어떤 상태들 중의 한 상태에 처할 수 있다(표 2. 참조). 입력기구인 테이프는 선형테이프로서 좌우 양방향으로 무한하며 나뉘어져 있고, 테이프의 한 사각형은 공란이나 한 기호를 지닐 수 있다. 테이프는 이 자동기계가 매 순간 t에 사각형 하나만을 관독하여 하나의 일정한 상태에 있을 수 있도록 움직여진다.

[표2]. 튜링기계의 조작 수행의 한 예 :  $(3+2) = 5$

A: 조작과제 : 덧셈  $(3 + 2) \rightarrow (5)$

테이프 내용의 변화 : 0/1/1/1/0/1/1/0 --> 0/0/1/1/1/1/1/0

판독기의 위치 : (이후 테이프가 좌우로 움직임)

**B: 기계표 :**

『IF(현 상태, INPUT), THEN(OUTPUT, 변화된 새 상태, 테이프 움직임)』

현상태	INPUT :	0	1
S1		(0,S,좌)	(0,S2,좌)
S2		(1,S3,좌)	(1,S2,좌)
S3		(0,S0,정지)	(0,S3,좌)

**C: 기계표를 사용하여 (3+2) (5)의 조작을 수행한 단계 별 내용**

단계	현상태	입력(판독)	출력	새상태	테이프 움직인 방향	출력 후 테이프의 내용
1	S1	0	0	S1	좌	0
2	S1	1	0	S2	좌	0 0
3	S2	1	1	S2	좌	0 0 1
4	S2	1	1	S2	좌	0 0 1 1
5	S2	0	1	S3	좌	0 0 1 1 1
6	S3	1	1	S3	좌	0 0 1 1 1 1
7	S3	1	1	S3	좌	0 0 1 1 1 1 1
8	S3	0	0	S0	정지	0 0 1 1 1 1 1 0

튜링기계는 알파벳의 유한 집합에서 추출된 상징(기호)들이 테이프에 주어지면 이를 하나씩 판독하고 현재의 상태를 점검하고, 기억내의 기계표(machine table)에서 이 두 조건이 규정하는 지시, 알고리즘, 추론 규칙을 조회하여 출력  $\delta(q, a)$ 를 내어놓고, 다음 상태  $\lambda$ 로 옮겨가는 유한 자동기계이다.

이 기계는 임의의 복잡한 계산을 또는 기계적 조작을 몇 개의 단순한 기계적 조작들

의 조합 또는 반복에 의해 수행 할 수 있음을 보여준다. 이는 알고리즘의 복잡성이 질적으로 증가하는 것을 기억의 크기와 알고리즘을 수행하는 시간의 양적 증가로 대치한다. 튜링기계가 기계적인 까닭은 이 기계의 조직들이 본질적으로 순환함수라는 의미에서이다. 즉 효율적 알고리즘 절차가 있으며 순환적으로 셀 수 있는 (recursively enumerable) 함수이기 때문이다.

순환적으로 셀 수 있다는 것은, 어떤 기호 스트링들에 대해 각각이 특정 집합에 소속하는지를 가려낼 결정 절차가 있으며, 이 스트링 집합과 다른 스트링 집합을 대응시켜주는 알고리즘이 있으며, 계산함수에 의해 유한히 기술(記述) 가능함을 의미한다. 튜링기계가 순환적으로 셀 수 있는 함수를 다룬다는 것은, 어떠한 함수이건 한 스트링 집합과 다른 스트링 집합을 대응시킬 수 있는 계산함수 모두를 튜링기계가 다룰 수 있다는 의미이다(McNaughton, 1982, Cutland, 1980; Jackson, 1985). 또한 유한히 기술 가능하며 계산 가능한 어떠한 함수도 부분순환함수로 표현할 수 있다는 정리에 따르면, 유한히 기술 가능한 어떠한 계산가능 함수도 튜링기계가 계산할 수 있다는 결론에 도달하게 된다.

튜링기계란 실상은 하나의 절차 또는 그 집합이라고 볼 수 있다. 한 튜링기계  $T_m$ 은 그 기계표와 테이프 내용만 주어진다면 [표 2]의 C에서 기술한 바와 같이 그 절차를 손으로도 모사할 수 있다. 이는 그 모사 절차가 알고리즘적이며 그 알고리즘적 절차가 다른 기계에 의해 수행될 수 있음을 의미한다. 이러한 알고리즘적 수행은 한 튜링기계  $T_m$ 의 전체 형태(configuration)를 새로운 상징들(예를 들어 0와 1의 조합)로 표상함에 의해서 이루어질 수 있다(이와 같이 스트링들에 다른 기호(상징) 또는 수를 부여하는 것을 피델화라고 한다).

튜링은 이에서 한 걸음 더 나아가, [보편 튜링기계(Universal Turing Machine) 정리]를 제시하고 이를 증명하였다. 보편 튜링기계 정리란 어떠한 튜링기계  $T_m$ 에 대해서도 이를 모사(simulate)할 수 있는 보편 튜링기계  $U_m$ 이 존재한다는 것이다.

이러한 보편 튜링기계가 존재한다는 것은 상당히 큰 의의를 지닌다. 덧셈하는 기계, 장기 두는 기계, 문자 복사기계, 문제 해결 기계 등을 그 하는 일에 따라 서로 다른 기계를 따로 만들어야 할 필요가 없어지기 때문이다. 서로 다른 여러 일들을 하는 기계들을 모사할 수 있는 하나의 보편튜링기계만 있으면 되기 때문이다. 각 튜링기계들의 행동이 유한히 기술될 수 있으며 보편 튜링기계가 각 튜링기계를 합한 기억 및 처리능력만 보유하고 있으면 되는 것이다. 또한 그러한 보편튜링기계로서 인간의 마음을 접근할 가능성이 시사될 수 있다.

#### 5.d. 지능적 기계와 튜링 테스트

튜링은 보편튜링기계의 구성이론을 제시한 후, 더 나아가 인간의 마음을 기계로서 보는 기계론을 강력히 전개했다.

**지능적 기계.** 인간의 지능, 그리고 인간의 사고라는 것은 가능한 대상이나 대안이나 답들을 나타내는 상징(기호)에 대한 탐색이라 생각하였던 튜링은 1950년에 쓴 논문, ‘계산기계와 지능’에서<sup>21)</sup> ‘지능적 기계’의 관점과 ‘튜링테스트(튜링검사)’라는 생각을 제시하였다.

그는 먼저 다음과 같은 생각을 하였다. 우리가 타인의 마음을 이해한다는 것은 타인의 외현적 행동을 보고 아는 것이다. 그런데 인간의 행동을 그대로 모사할 수 있는 어떤 기계가 있을 수 있다. 더욱이 인간의 행동이 기호(상징)로써 표출된다고 했을 때, 그 상징 표출 조작과정을 충분히 흉내 내어 모사할 수 있는 기계가 있을 수 있다. 즉 인간의 행동을 그대로 모사 할 수 있는 유한 자동기계가 있을 수 있다.

그런데 기계란 앞서 논의한 바에 의하여 형식체계이다. 그렇다면 인간도 상징을 받아들여 이에 대해 추론 규칙을 적용하여 출력을 내어놓는 형식체계로 간주할 수 있으며 형식 체계에 적용되는 논리와 법칙을 사용하여 인간의 마음을 설명할 수 있지 않는가?

그렇다면 수학적 직관을 포함한 인간의 모든 마음의 과정은 기계화, 형식화할 수 있으며 곧 인간은 기계 이상의 무엇을 할 수 있는 체계가 아니라고 할 수 있다. 즉 ‘인간은 기계이다’라고 결론지을 수 있다. 튜링자신은 ‘인간은 기계이다’라고 강력하게 주장하지는 않았고, 단지 ‘인간의 마음의 작용들을 모두 모사할 수 있는 기계를 구성할 수 있다’는 약한 표현을 사용하였지만, 그는 인간의 마음의 작동을 지능적 기계의 작동 과정과 유사하다고 보았던 것이다.

**뇌, 마음, 기계.** 또한 그는 어떻게 기계가 인간의 지능을 나타내게 할 수 있는지에 대해서 검토하면서 인간의 뇌를 그 지침으로 삼았다. 인간을 닮은 생각하는 기계를 만들기 위해서 인간의 뇌의 작용을 게임 등에서 모사하는 기계를 만드는 것이 도움이 될 것이라고 생각하였다. 서양장기인 ‘chess, checker, tic-tac-toe, bridge, poker’ 등의 게임을 할 수 있는 그러한 뇌를 만드는 것과 이를 기계와 같은 형식체계로 표현함에 대하여 그는 계속 생각하였다.

그는 뇌의 기능을 규명하는데 있어서 해부학적인 접근보다는 수학적 접근이 더

---

21) Turing, A. (1950). Computing machinery and intelligence. Mind, LIX, 2236, 33-60.



유용할 것이라고 생각하였고, 뇌와 기계는 정보 처리적인 관점에서는 서로 유사하다고 보았다. 따라서 기계라는 하나의 구조를 아는 것이 뇌라는 다른 하나의 구조를 아는 것에 도움을 줄 것이라고 생각하였고, 인간과 기계의 작동하는 차이를 질적(유형) 차이이기보다는 양적(정도)의 차이인 것이라고 생각하였다.

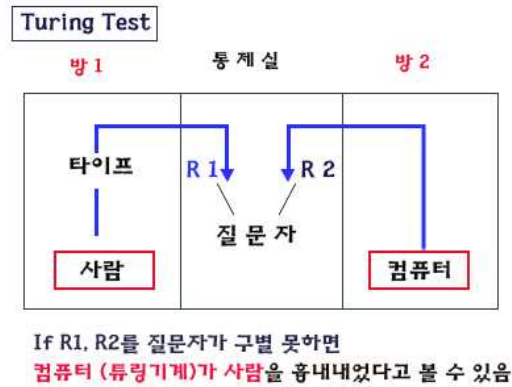
물론 그 당시의 컴퓨터가 기능이 아주 제한적이라는 점에서 튜링의 생각이 더 발전하지는 못하였지만 그는 다른 사람들이 컴퓨터와 인간 신경계가 물리적으로 다른 속성이기에 컴퓨터와 인간이 서로 다를 것이라고 주장하는 반론에 별로 귀를 기울이지 않았고, 그런 물리적 수준보다는 다른 수준에서 (오늘날의 용어로 ‘정보처리 수준’에서; 그러나 당시에는 정보처리 수준이란 개념이 제안되어 있지 않았다) 컴퓨터와 뇌가 공통적인 것이 많으리라고 생각하였다.

**튜링 테스트.** 과연 기계는 생각할 수 있는가 하는 물음에서 출발한 그의 ‘튜링 테스트’ 개념은 인간과 기계가 과연 같이 ‘생각을 하는가’ 여부를 확인하는 방법에 대한 이론의 전개이었다. 그는 먼저 ‘생각한다’는 것을 정의하기가 어렵기 때문에 이 물음을 비교적 덜 애매한 “모사 게임에서 (사람의 마음처럼) 잘 할 수 있는 그러한 상상 가능한 디지털 컴퓨터가 있는가(“Are there imaginable digital computers which would do well in the imitation game”)?“라는 물음으로 바꾸어 놓고 그의 튜링테스트 개념을 제시하였다.

그가 제시한 튜링테스트(Turing Test)란, 한 튜링기계와, 한 인간 피험자, 그리고 한 질문자의 세 개체 각각 서로 다른 방에 있으며, 이 세 개체 사이의 의사소통이 텔레타이프를 통해서만 이루어진다고 할 때, 질문자가 자신의 질문에 대한 답이 튜링기계에서 나오는 것인지, 인간 피험자에게서 오는지를 구분할 수 없다면 그 튜링기계는 튜링테스트를 통과한 것이며 그 기계는 인간과 같이 사고할 수 있다고 결론짓는 것이라고 하는 생각이다. 만약 질문자가 자기가 얘기하는 대상이 사람인지 기계인지를 구분해 낼 수 없다면, 그렇다면 그 기계를 사람과 같게 생각할 수밖에 없을 것이고 그것이 자연스러운 것이라고 튜링은 생각하였다.

이후에 이 튜링테스트 개념은 인공지능의 성공가능성을 논할 때나 철학에서 마음과 기계를 연결하여 심신론을 논할 때에, 그리고 디지털 테크놀로지 세상의 미래를 논할 때에 하나의 필수적 핵심 개념으로 등장하며 후대의 사람들에게 많은 상상과 논쟁의 단초가 되었다.

### [그림 3]. 튜링테스트(검사) 상황



### 5.e. 튜링기계 이론의 의의

튜링기계론과 튜링테스트(검사)의 개념은 인지주의 및 인공지능의 형성과, 인지과학에서의 정보처리 패러다임의 발전, 심리철학에서의 기계론의 대두 등에 직접 간접으로 커다란 영향을 주었다. 튜링기계 이론은 배비지와 러브레이스가 생각했던 바인, 보편적 계산 기계가 이론적으로 가능하다는 것을 형식적으로 증명한 것이었다.

그렇기는 하지만 수학자 앨런 튜링의 ‘튜링기계’는 실용적 의의가 부족하였다. 왜냐하면 튜링기계를 튜링의 이야기 그대로 구현하여서는 아주 단순한 문제 해결에도 굉장히 많은 단계의 계산 절차를 거쳐야 했던 때문이다. 튜링의 이론적 튜링기계처럼 단순하면서도 상징(기호) 정보처리 능력이 효율적인 구체적, 실용적 계산기의 개발이 필요하였다.

비슷한 시기에 마음, 뇌, 계산, 인공뇌, 지능적 기계 등에 대하여 여러 개념과 이론이 제시되었던 것과, 구체적인 디지털 컴퓨터의 개발 및 발전은 이러한 필요성을 채워줄 수 있는 개념적 전환의 바탕이 되었다.<sup>22)</sup>

## 6. 튜링 이후 기계-뇌-마음 연결의 개념적 변화를 주도한 학자들

### 6.1. Wiener.

미국의 수학자이며 공학자였던(이런 영역적 명칭이 그에게는 오히려 부적합하다.) 와이너(Norbert Wiener, 1894-1964; [http://en.wikipedia.org/wiki/Norbert\\_Wiener](http://en.wikipedia.org/wiki/Norbert_Wiener))는 기계와 인간(뇌신경 과정, 마음, 사회)의 연결을 핵심 주제로 하는 사이버네틱스

22) 컴퓨터과학의 발전역사: (<http://www.cs.uwaterloo.ca/~shallit/Courses/134/history.html>)

(Cybernetics)(<http://en.wikipedia.org/wiki/Cybernetics>) 영역의 창시자이었고, 또한 피드백(되먹임, feedback) 및 제어 개념을 형식화(formalization) 하는 등의 창의적 아이디어를 제시함을 통해 공학일반, 시스템 제어, 컴퓨터과학, 생물학 철학, 사회조직학, 인지과학 등의 분야에 많은 영향을 준 학자이었다.<sup>23)</sup>

그는 여러 논문 발표 후에 1948년 3월에 사이버네틱스(Cybernetics)라는 책 - ([사이버네틱스: 또는 동물과 기계에서의 제어와 커뮤니케이션](Cybernetics: Or the Control and Communication in the Animal and the Machine)) - 을 출판하여 전자공학과 동물 및 인간(인간의 신경계)을 연결하는 사이버네틱스라는 새 분야를 열었다. 이 영역의 생각들은 컴퓨터와 인간, 인간과 기계의 연결에 대한 사람들의 생각을 온통 뒤바꾸어 놓았으며 기계와 인간의 연결은 아주 새로운 개념 틀로 접근 가능하게 되었다.

수학자였지만 메시지 전달 현상에 학문적 관심을 지녔던 그는 메시지를 통한 기계의 제어와 인간 사회에서의 언어의 역할을 유사한 것으로 생각하였다. 인간이건, 동물이건, 기계이건 간에 공학에서의 제어이론은 본질적으로 메시지 이론의 한 영역이라는 관점을 지녔다. 그는 그가 창안한 사이버네틱스라는 영역의 목표가, 제어와 커뮤니케이션의 문제를 다루며, 여러 영역에서 그것들이 드러나는 것을 분류하는 적절한 아이디어와 테크놀로지를 찾아내게 하는 언어와 기술을 개발하는 것이라고 생각하였다.

앞서 언급한 17세기의 라이프니츠의 기계론적 생각의 영향을 받았던 그는 뉴턴 이래로 전통적 과학 패러다임의 중심개념이었던 ‘에너지’ 개념 대신 ‘정보(information)’ 개념이 과학계와 기술계의 핵심 개념으로 대두되게 한 한 배경을 제시한 장본인이다. 그는 1948년에 쓴 책 [사이버네틱스]에서 “정보는 정보일 뿐, 물질이나 에너지가 아니다. 어떠한 유물론적 관점도 이를 인정하지 않으면 오늘날에 생존할 수 없다 (*"Information is information, not matter or energy. No materialism which does not admit this can survive at the present day."*)” 라고 하여 정보를 강조하고 정

---

23) Wiener라는 이름은 영어로는 와이너라고 읽지만 독일 발음으로 비이너임. 그는 14세에 학부를 졸업하고 17세에 수리논리학 논문으로 하버드대 수학과 박사학위를 취득하고 25세에 MIT 수학과 강사가 되었다. 2차 대전 중 대공 유도탄 기술 관련 연구를 하면서 전자와 정보, 제어, 피드백에 관심을 갖게 되었고, 전후의 MIT대학의 인지과학 팀의 형성과 그 발전의 핵심인물이 되었다. 그의 학문적 영향을 받은 학자로는 그레고리 베이트슨, 마가렛미드, 맥컬러 등이 있다. 그는 그가 ‘boundary regions of science’라고 이름붙인 ‘학문간 융합’의 영역에서 연구하며 아이디어를 제시하였던 20세기 초의 최고의 융합적 연구자였다.

나는 개인적으로 Pamela McCorduck이 그녀의 책, [Machines who think] (1972)에서 Wiener를 이야기하면서 “Wiener loved to work in what he called the boundary regions of science, the no-man's land between the various established fields. Specialization, as necessary as it was, seemed to him only one way of doing science.” 라고 쓴 표현에 사로잡혔었다. 아마도 이것이 인지과학을 하면서 내가 따르고자 한 길이 아니었든가 생각된다.

보 위주의 패러다임으로의 과학 체제의 변혁을 시작하였다.

아마도 20세기 초반에 커뮤니케이션학, 수학, 전자공학, 공학일반, 컴퓨터과학, 철학, 심리학, 인지과학, 인공지능, 로봇틱스, 등의 여러 분야에서 사람들에게 창의적 아이디어들을 촉발시키고 인간과 기계의 연결에 대하여 가장 융합적 접근을 하였고, 인문학, 수학, 공학을 연결하여 새로운 개념들을 제시한 천재가 아니었던가 생각된다. 그는 정보 중심의 인류문화와 인지주의가 20세기의 인간 사회에서 등장하여야 하는 바탕 배경을 놓았다고 할 수 있다. 그의 커뮤니케이션 및 '정보이론'은 다음에 언급할 새넨(Shannon)의 정보이론의 영향을 받지 않고 독자적으로 형성된 것이며 그의 관점을 후에 언급할 맥컬러(McCulloch) 등에게 영향을 주었다.

## 6.2. Shannon .

미국 수학자이며 양화된 '정보이론'의 아버지라고 지칭되는 새넨(Claude Elwood Shannon. 1916-2001; [http://en.wikipedia.org/wiki/Claude\\_Shannon](http://en.wikipedia.org/wiki/Claude_Shannon))은 21세의 MIT 대학원생이던 1937년에 '전기회로 릴레이와 스위치의 상징적 분석'이라는 석사논문(학술지 발표는 1938임)에서 전기회로를 보편적 계산에 쓸 수 있다는 이론을 제기하였다.

그는 이진법 기호로 일반적 산수 문제를 표현할 수 있을 뿐만 아니라, 부울(Boole) 대수와 기호논리의 보다 보편적 문제들을 전자 회로로 표현할 수 있다는 생각을 제시하였다. 그는 이진법 부호가 진위 관계나 조건적 양자택일(이접적, either/or) 관계를 나타낼 수 있음을 보였으며, 모든 이진법 부호를 개/폐만 가능한 (전화의) 전기회로의 연쇄로 기계적으로 표현할 수 있음을 지적하였다. 그렇다면 논리학, 철학, 수학에서 거론되는 모든 논리적 명제를 2진법적 전기회로로 표현 가능한 것이 되며, 전자공학과 논리학, 철학, 수학을 연결하는, 따라서 부울 2진법 대수에 기초하고 있는 현대 디지털 컴퓨터 전자회로 설계와 디지털 문화를 가능하게 하는, 전혀 새로운 가능성이 열리게 되는 것이다.

프린스턴대학의 고등과학기술연구원에서 폰노이만, 아인슈타인, 튜링 등과 교분을 쌓은 (튜링이 보여준 '보편튜링기계'에 관한 1936년 논문에서 제시된 아이디어들을 흡수하여 새넨은 그의 커뮤니케이션(정보) 이론을 보완하였다고 한다.) 새넨은 1948년 후반에, '커뮤니케이션의 수학적 이론(A Mathematical Theory of Communication)' (<http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf>)이라는 논문을 발표하여 - 노버트 와이너의 이론적 생각에 영향받은 것이 아니라, 독자적으로 제시하였음 - 정보의 개념이 기존의 주먹구구식 추상적 개념이 아니라 엄밀히 측정할 수 있는 메시지의 양의 개념으로 (물리학의 엔트로피 개념을 사용하여) 과학계에서 널리 퍼지고 자리 잡는 데에 결정적 역할을 하였다. 그는 커뮤니케이션 시스템에서의 '불확

실성'과 'redundancy', '엔트로피'의 개념을 사용하여 '정보'를 양적으로 측정할 수 있게 하는 수학적 모델을 제시하여 커뮤니케이션 이론(현재는 '정보이론'이라고 불린다.)의 영역이 과학 영역으로 자리 잡게 한 것이다.

후에 그의 동료들과 쓴 논문에서는<sup>24)</sup> 배경 잡음으로부터 신호를 분리하는 처리와 제어의 개념을 논하여 '정보처리' 개념이 포함된 정보이론의 시대를 열었다. 그는 또한 언어 분석을 확률적 기초에서 수행하는 1951년의 그의 논문, '활자화된 영어에서의 예측과 엔트로피'는 언어처리, 컴퓨터언어학에 정보이론을 연결하게 하여 후에 언어(처리) 연구와 컴퓨터과학의 연결을 가능하게 하기도 하였다. 그가 제시한 논문들에서의 일련의 아이디어들에 의하여 인류 문화가(적어도 텔레커뮤니케이션, 마이크로 프로세서 테크놀로지가) 아날로그 시대에서 디지털 시대로 넘어가는 개념적 변혁이 촉발된 것이다.<sup>25)</sup>

### 6.3. W. McCulloch

다음으로 미국의 신경생리학자 맥컬러와 핏츠(Warren S. McCulloch, 1898-1969; [http://en.wikipedia.org/wiki/Warren\\_Sturgis\\_McCulloch](http://en.wikipedia.org/wiki/Warren_Sturgis_McCulloch); & Walter Harry Pitts, 1923-1969; [http://en.wikipedia.org/wiki/Walter\\_Pitts](http://en.wikipedia.org/wiki/Walter_Pitts))<sup>26)</sup>가 이러한 개념을 한 단계

24) Shannon, Claude E. & Warren Weaver, (1949): A Mathematical Model of Communication. Urbana, IL: University of Illinois Press. 이 정리는 후에 Shannon-Hartley 정리로 재구성됨.

25) 와이너와 새넌의 '정보'개념은 다소 차이가 있다. 스베이비 (Karl-Erik Sveiby, 1994; <http://www.sveiby.com/articles/Information.html#Cybernetics>)에 의하면, 와이너는 정보양, 엔트로피, 피드백, 배경잡음 등의 개념을 인간 뇌의 신경기능의 본질적 특성으로 도입하며, 정보양을 한 시스템 내의 조직화 정도로 규정하였다(자연히 엔트로피는 그 시스템의 비조직화 정도로 규정되고, 정보의 양은 부정(negative) 엔트로피가 됨). 정보의 가장 단순한 형태는 동등 확률을 지닌 선택대안들 사이의 선택 확률로 표현된다. 확률로 표현되는 정보에 의하여 시스템은 기준과 주어진 입력 사이의 차이를 계산하고, 이 차이를 줄이기 위한 부정(negative) 피드백이 필요하게 되고 제어가 생긴다. 한편 새넌의 정보의 개념은 조금 달랐다. 커뮤니케이션 시스템 내에서의 신호전달과 커뮤니케이션 채널의 용량 한계에 관심이 있던 그는 정보의 양을 확률의 합의 부정(negative) 지수함수(the negative of the logarithm of a sum of probabilities)로 규정하였다. 이는 정보라는 것은 메시지를 선택하는 자유의 척도이며 이 자유가 클수록 불확실성이 큰 것이며 정보가 많은 것이라고 본 것이다. 새넌의 정보 개념은 정적(positive) 엔트로피와 같은 것이었다. 그에겐 정보란 한 메시지에 관한 것이 아니라 그 메시지가 전하여지는 상황 전체에 관한 것이었다. 상황 하에서 메시지 선택이 이루어지며 그러한 연관에서 정보와 정보의 양이 측정될 수 있고 거론되는 것이다. 반면 와이너의 정보 개념은 부정(negative) 엔트로피(체계의 구조화 정도)의 개념이었다.

26) Walter Pitts에 관한 자료([http://www.csulb.edu/~cwallis/artificialn/walter\\_pitts.html](http://www.csulb.edu/~cwallis/artificialn/walter_pitts.html))에 의하면 그는 15세에 서커스 단원이 되기 위하여 가출하지만 그 대신 쉬카고 대학에 가서 버트랜드 러셀의 강의를 듣고 저명한 논리학자 카르납의 원고에 수정되어야 할 부분을 지적하고 카르납이 그렇게 강권했음에도 불구하고 박사학위를 받지 않았고, 쉬카고에 있는 동안 McCulloch 교수와 그 유명한 신경세포의 논리적 대수에 관한 논문을 썼고 또 후에는 시각 인지과학의 대표적 학자인 Jerome Lettvin과 함께 시각 인지과학의 고전적 대표 논문인 개구리의 뇌의 시각정보처리에 관한 논문을 썼다. 그가 가깝게 지냈던 사람들은 러셀, 카르납, 맥컬러, 그리고 사이버네틱스를 창안한 노버트 비이너, 그리고 인지과학자 레트빈 등이었다. 그는 이들이 그를 MIT대의 교수로 시키려는 여러 노력에도 불구하고 거절하고 학위가 없는 채로 인지과학-인공지능-뇌-수학을 연결하는 아주 중요한 논문들을 내는 한 연구자로

더 발전시켰다. McCulloch는 학부에서 철학을 먼저 전공하고, 다음에 심리학을 전공한 후, 의학으로 가서 신경생리학자 겸 사이버네틱스 학자로서 활동하였다. 그는 지식이란 무엇이며, 그 지식을 알 수 있는 인간이란 어떤 존재인지에 대해서 궁금해 했다. 그는 철학과 심리학, 의학, 신경학 등 여러 측면을 연결하여 생각을 했다.

로젠블루스 등(Rosenblueth, Wiener, & Bielow)이 [행동, 의도, 목적(Behavior, Purpose, and Teleology)]라는 책을 써 낸 것과 같은 해인 1943년에 맥컬러는 핏츠와 함께 ‘신경활동에 내재하는 아이디어의 논리적 계산(A Logical Calculus of the Ideas Immanent in Nervous Activity)’이라는 논문을 발표했는데 이 두 글 모두 ‘정보 처리 모형’을 인간의 마음에 대한 유용한 가설로서 제시했다.

맥컬러와 핏츠 논문의 의의는 사람들에게 인간의 ‘두뇌’를 ‘생각하는 기계’로 보는 새로운 시각을 소개했다는 것과, 인간의 뇌와 지식 간의 관계가 밝혀져야 할 것이라는 점을 지적했다는 데 있다. 그들은 인간 두뇌까지도 정보를 2진법으로 처리하는 기계로 개념화 할 수 있음을 논하였다. 뇌와 신경계가 수많은 신경세포들의 망으로 이루어졌으며, 신경세포들 하나하나가 기본적으로 활성화되거나, 안 되거나 하며, 신경 흥분을 인접 신경세포에 전달하거나 안 하는 것이 신경계의 기본 활동이며, 이는 이진법적 스위치와 같은 양상이다 라는 것이다.

이는 전기(전자) 컴퓨터의 기계적 과정이 뇌와 신경계에 대한 좋은 모델이 될 수 있음을 시사하는 것이었다. 그 당시의 일반적 연구들은 신경 세포의 작용 형태가 실무적(전무율, all-or-none)인 것으로 간주하였으나, 그 이후 연구에 의하면 신경 세포들은 대개 비선형적인(nonlinear) 반응 형태라는 사실이 발견됐다. 맥컬러와 핏츠의 신경세포 모델(MCP Neurons; McCulloch-Pitts Neurons) (<http://www.mind.ilstu.edu/curriculum/modOverview.php?modGUI=212>)은 오늘날에 보아서는 단순성의 문제점도 있었으나 그 당시에는 분석도구가 되는 수학 기법이 충분히 발전되어 있지 않았기 때문에 그의 단순한 신경망 이론은 당시에는 아주 타당한 것으로 간주되었었다.

맥컬러는 철학의 문제들도 신경 체계의 해부학이나 생리학적인 관점에서 해결되어야 한다고 생각했다. 또 신경망이 학습이나 기억, 예측, 의도 따위를 어떻게 구현하는지에 대해서도 나름대로의 안목을 갖고 있었다. 그는 인간의 마음을 과학적으로 규명될 수 있다는 신념을 표명함으로써 그의 주변의 많은 사람들이 그러한 연구에 자신을 갖고 도전할 수 있게끔 했다. 이러한 영향을 받은 사람 중의 한 사람이 후에 인공지능의 대가가 된 민스키(M. Minsky)와 같은 학자였다.

---

삶을 마쳤다. 인지과학 분야에는 상당히 창의적인 사람들이 많이 있었지만, 그는 후에 시각정보처리 이론을 제시한 데이빗 마(D. Marr)처럼 젊은 나이에 세상을 뜬 아까운 천재적 젊은 인지과학도였다고 생각된다.

#### 6.4. 컴퓨터 하드웨어의 발전

마음과 두뇌의 과정을, 기계적 계산과정으로, 그리고 2진법적 논리적 계산관계로 개념화 할 수 있다는 이론적 틀이 이와 같이 급격히 발전되는 것과 병행하여, 이러한 개념화를 실제의 컴퓨터로 구현할 수 있는 바탕을 제공할 수 있는 컴퓨터 하드웨어(계산력)의 발전도 계속 진행되었다. 1943년에 하버드 Harvard대학과 IBM의 과학자들은 전자기 릴레이를 사용하여 [Harvard Mark I]이라는 범용적 초대형 컴퓨터를 개발하였다. 모든 수리적 계산을 할 수 있는 컴퓨터였다.

그 직후에 펜실바니아대학 연구자들이 전자튜브를 사용한 [ENIAC]이라는 컴퓨터를 만들었다. 이 팀의 연구자들이 수학자 노이만(J. von Neumann)과 토론하고 협동하는 과정에서, '계산'의 본질 개념을 영원히 바꾸어 놓는, 과학사적으로 중대한 일이 발생하였다.

이 당시의 컴퓨터는 계산 속도는 놀라울 만큼 빨라졌는데, 그 계산을 수행하라는 계산 지시 방법이 뒤쳐져 있었다. 매번 다른 계산을 할 때마다 새로운 지시들을 주어야 하고, 이 지시가 입력되어 처리될 때까지 컴퓨터는 놀고 있어야 했다. 물론 이 당시에는 데이터나 지시가 모두 구멍뚫린(핀치, punched) 카드를 통해 이루어졌다.

#### 6.5. 수학자 John von Neuman.

이러한 비효율성의 문제를 해결하는 틀을 향가리 태생의 미국 수학자 폰 노이만(J. von Neumann; 1903-1957; [http://en.wikipedia.org/wiki/John\\_von\\_Neumann](http://en.wikipedia.org/wiki/John_von_Neumann))이 제시하였다.

그는 대부분의 복잡한 계산들이 동일한 몇 개의 하위 계산처리 과정들(현대 의미로 서브루틴)이 다른 순서와 형태로 조합되어서 이루어진다는 것을 발견하고, 이러한 하위계산처리 과정들을 미리 일정한 순서를 지닌 이진법 부호들의 연쇄로, 즉 '지시'로 컴퓨터의 기억에, 자료와 함께, 저장하였다가 필요한 시점에 사용하게 하는 '**저장된 프로그램(stored program)**'의 개념을 생각해 내었다. 2진법 부호의 컴퓨터 소프트웨어 열개(architecture) 이론을 제시한 것이다. 이것은 대단한 발견이요 공헌이었다. '저장된 프로그램'의 개념은 마음과 기계의 유추에서 심적 과정 절차의 개념에 바로 적용될 수 있는 개념이었다.

노이만의 공헌에 의하여, 컴퓨터 계산 작업 지시와 관련된 고차적 언어인 프로그래밍 어셈블리 언어가 가능해져서 'BASIC' 등의 어셈블리 언어들이 탄생했고, 이후의 컴퓨터의 하드웨어적, 개념적 발전이 급속도로 이루어지게 되었다. 하드웨어적 측면에서는 이후에, 트랜지스터의 발명, 실리콘 칩의 발명이 이어져서 컴퓨터 처리 속도의 개선이 급격히 개선되었고, 개념적 측면에서는 [프로그래밍]의 이론과 도구 개발이

전반적으로 급속도로 이루어졌다.

당시에는 미국과 영국에 기계-뇌-마음을 연결하는 주제인 사이버네틱스를 발표 및 토론의 주제로 삼는 학자들의 모임이 여럿 있었다(예: Ratio Club, Teleological Society, Macy Conference). 이러한 학술적 모임들 통해 사이버네틱스와 새로운 과학 패러다임의 가능성에 관심있는 학자들이 서로에게 영향을 주고받았다. 이 모임들이 잘 되었던 것은 그 당시에 과학적 패러다임이 이미 ‘에너지’에서 ‘정보’로(from ENERGY to INFORMATION) 바뀌고 있는 역동적 시대이었음을 나타낸다. 그러나 새로운 과학적 패러다임의 결정적 형태가 드러난 것은 아마도 다음의 7절에서 제시하는 몇 개의 학술모임에 의해서이라고 하는 것이 옳을 것 같다.

## 7. 마음 기계 연결 변혁의 기폭제가 된 학술 모임들

### 7.1. 1948년의 첫 HIXON 심포지엄 : 대뇌 메커니즘과 행동

이러한 기폭제가 된 학술모임의 첫째 모임은 미국 캘리포니아의 칼텍 공과대학에서 1948년 9월 20-25일에 개최된 “대뇌 메커니즘과 행동” 1회 심포지엄이었다(<http://calteches.library.caltech.edu/923/1/Pump.pdf> 하단자료). 이 사이트 자료와 이후에 이 심포지엄의 내용이 책으로 나온 후에 이 책을 평한 자료(<http://calteches.library.caltech.edu/1314/1/books.pdf>)에 의하면, 이 심포지엄에는 당시에 오르는 분야인 ‘심리생물학’ 분야와 관련된 학자들이, 심리학, 신경심리학, 심리병학(정신병학), 동물학, 수학, 물리학 등의 분야에서 참석하였었다. 주관은 캘리포니아 공대 생물학부장 비이들(George W. Beadle)과 그 대학 심리생물학과의 방문교수였던 제프리스(Lloyd A. Jeffress) 담당하였고 7개의 초청발표가 있었다. 그 7개 발표는 두정엽과 측두엽의 차이 문제에서부터 ‘마음이 왜 뇌에 있는가(‘Why the Mind is in the Head’) 하는 주제에 이르는 것이었다.<sup>27)</sup>

자동기계에 대한 수리-형식적 이론을 기계와 인간과 동물의 중추신경계 이론으로 적용할 수 있다는 유명한 수학자 폰노이만<sup>28)</sup>의 발표로 시작된 이 심포지엄은, 비록 종합적 통일된 결론을 도출해 내지는 않았지만 ‘기계(컴퓨터)-뇌-마음-지능의 연결’이 생물학, 심리학, 사이버네틱스(그 당시는 컴퓨터 과학이 아직 하나의 독립된 분야로 형성되지 못하였었다) 등의 여러 분야의 공동주제이며 접근하여 앞으로 학제적으로 연구하여야 할 주제라는 공통의(그러나 그러한 생각이 어떠한 새 통일된 패러다

27) 발표자는 Ward C. Halstead(쉬카고대), Heinrich Kluver(쉬카고대), Wolfgang Koehler(형태심리학 주창자), Lorente de No, Rockefeller(의학연구소), K. S. Lashley(하바드대), R. Lorente de No(록펠러의학연구소), John von Neumann(고등과학연구원) 등이었고 (<http://calteches.library.caltech.edu/923/1/Pump.pdf>) Warren S. McCulloch(일리노이대)를 비롯한 19명의 학계 선두 연구자들이 참석하였다.



임을 가져오는가에 대한 구체적 개념 형성이 없이 끝난 모임이었고 그러한) 자각을, 그리고 그 주제가 상당히 복잡한 주제이라는 인식을 참석자들에게 심어주는 데에 성공한 모임이었고 이어서 이어진 다음의 두 모임의 지적 배경을 제공한 모임이었다.

프린스턴대로 옮겨서 박사학위를 받기 이전에 캘리포니아 공과대학 대학원생이던 맥카시는 이 HIXON 심포지엄에 참석하여 수학자 튜링, 새년, 심리학자 라슬레이 등의 이야기를 들으면서 인공지능에 대한 첫 개념과 다음의 Dartmouth 모임의 기획 아이디어의 바탕 생각을 하였던 것 같다.

## 7.2. 1956년 여름의 Dartmouth대 모임

8년이 지난 1956년 여름의 Dartmouth 대학에서 아주 중요한 학술모임이 열렸다. 1955년 당시에 Dartmouth 대학에 있던 수학자 맥카시<sup>28)</sup>는 다음 해에 열릴 Dartmouth 학회를 조직하기 위한 연구 프로젝트 제안서<sup>29)</sup>를 썼다. 이 제안서에서 McCarthy는 ‘인공지능(Artificial Intelligence: AI)’이라는 용어를 최초로 사용하였다. 이 날이 인공지능(AI)의 공식적인 탄생일로 간주되고 있으며 이후 그의 계속된 노력의 결과로 이 용어가 20세기 말까지 마음과 기계를 연결하는 첨단 핵심 분야를 지칭하는 용어로 학계에서 공식적으로 사용되었다.<sup>30)</sup>

28) 맥카시(John McCarthy, 1927-2011; [http://en.wikipedia.org/wiki/John\\_McCarthy\\_\(computer\\_scientist\)](http://en.wikipedia.org/wiki/John_McCarthy_(computer_scientist)); <http://www.nytimes.com/2011/10/26/science/26mccarthy.html?pagewanted=all>)는 컴퓨터과학 출판을 도왔고, 인공지능 분야를 새로 창시하였고, MIT와 스탠포드대에서 인지과학의 형성기에 중요한 역할을 하였다. 수리논리학과 인공지능을 연결하려던 그는 인공지능 발전을 돕는 프로그래밍 언어 ALGOL, LISP 등도 개발하였다. 기계-마음 연결의 커다란 계기가 된 것이 Dartmouth 모임이었다면 그 모임을 기획하고 제안서를 썼던 그의 역할을 높이 평가하여야 하리라 본다.

29) 제안서: <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>: 제안서 내용 = : 제목; ‘A Proposal for the Dartmouth summer research project on artificial intelligence. / 연구제안자: J. McCarthy, M. L. Minsky, N. Rochester, C.E. Shannon; / 제안일자: 1955. 08. 31./ 모임제안 연구 요약: “We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire.”/ 이 연구모임의 목적: “to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. (이후 내용 생략): 모임 설명: [http://www.livinginternet.com/i/ii\\_ai.htm](http://www.livinginternet.com/i/ii_ai.htm)

30) 이 Dartmouth 학술 모임의 참석자는 John McCarthy, Marvin Minsky, Nathaniel Rochester, Claude Shannon, Ray Solomonoff, Oliver Selfridge, Trenchard More, Arthur Samuel, Herbert Simon<sup>1)</sup>, Allen Newell 등이었다. 그들은 대학에 혹은, 기업에 종사하는 사람들이었다. McCarthy는 Dartmouth대학에 종사하는 수학교수였고, Minsky는 Harvard에서 수학과 신경학을 전공하고, Rochester는 IBM에서 Shannon은 Bell연구소에서 일하고 있는 학자이었다. Dartmouth 학술모임은 프로젝트 모임이라고 하기 보다는 두 달에 걸친 여름학교라고 보는 것이 더 좋은 이 모임에서는 몇 사람의 초청강연자가 있기는 하였지만 실질적으로는 강연자와 학생의 구별이 없는 자유로운 지적 교환의 모임이었다.

과연 기계가 생각을 할 수 있을까? 17세기 데카르트 이래 과거의 과학 전통에서는 마음을 인간의 몸과 분리시켜서 실재하지 않는 영적인 것으로 생각했기 때문에 인간의 내적 세계, 즉 마음은 과학의 대상에서는 제외되어왔다.

그러나 1956년 여름의 Dartmouth대학 학술모임에서 모인 여러 명의 과학자들의 생각은 그것과 확실히 달라 보였다. 그들은 여러 가지 다른 배경들을 각자가 가지고 있었다. 수학자, 심리학자, 전기공학자, 심지어 학교뿐만 아니라 기업에 종사하는 사람들까지 함께 모였다. 이렇게 다른 배경을 가지고 있었지만, 그들은 한 가지 공통된 생각을 지니고 있었다.

그것은 다름이 아니라, 기계가 인간의 지능을 가질 수 있다는 것이고, 그것을 위해서는 인간의 마음을 과학적이고, 형식적으로 설명할 수 있어야 하고, 디지털 컴퓨터는 그것을 위한 가장 좋은 기계라는 생각이었다.

그들의 이러한 주장이 있기까지는 그 배경이 되는 이론적 토대가 학문적으로 이루어져 있었다. Turing의 자동기계이론으로 대표되는 수리논리학, ‘McCulloch & Pitts’의 신경망 모델, von Neumann의 프로그램 내장식 컴퓨터의 설계, Wiener의 사이버네틱스이론, Shannon의 정보이론, von Neumann의 자기증식 automata 이론, 캐나다 생리심리학자 D. E. Hebb(후에 8절에서 다시 언급됨)의 학습규칙(learning rules) 등의 새로운 생각들은 인지과학의 탄생에 기여한 핵심적인 아이디어이며, 기계와 마음의 연결에 대하여 위의 주장을 할 수 있게 하는 근거들이다. 이 Dartmouth 모임 후에 Minsky<sup>31)</sup>는 Herbert Simon, Allen Newell과 모임을 가지고, 인간처럼 지능적으로 사고할 수 있는 컴퓨터 프로그램의 개발 가능성을 검토하기 시작했다.

‘지능’에 관한 언급이 명시적으로 거의 되지 않았음에도 불구하고 이 모임에 참석한 사람들은 상징(기호) 조작 과정이 무엇인지를 이해하게 되었고, 기계와 인간의 사고 과정을 연결하여 새 틀에서 생각할 수 있다는 것을 이해하게 되었고, 계산심리학이라는 영역<sup>32)</sup>이 새롭게 형성될 수 있는 가능성을 인식하게도 되었다. 그러나 참석하였던 심리학 분야의 연구자들과 대학원생들에게 더 영향을 준 것은 다음의 MIT 정보이론 모임이었다.

---

31) MIT의 인공지능연구실 창립자인 인지과학자 민스키(Marvin Minsky, 1927-; 위키백과 소개자료 [http://en.wikipedia.org/wiki/Marvin\\_Minsky](http://en.wikipedia.org/wiki/Marvin_Minsky); 홈, <http://web.media.mit.edu/~minsky/>)가 1950년대에 ‘뇌는 고기 덩어리로 된 기계다(Brain happens to be a meat machine.)’라고 주장함으로써 당대의 사람들을 경악케 했다. 아마도 사람들은 인간을 기계와는 다른 그 어떤 존재로 인식하기에 그렇게 당황했을 것이다. 인지주의의 정보처리 모형은 인간 뇌가 담당하는 모든 지적 활동을 정보 처리로서 본다. 따라서 인간의 뇌도 하나의 기계처럼 인식될 수 있을 것이다. 민스키 교수는 후에 (1969년) 페이퍼트 박사(Seymour Papert)와 [퍼셉트론]이라는 책을 저술하여 인공신경망 이론과 연결주의 이론의 발전에 큰 한 획을 그었다. (비록 이 책이 신경망, 인공지능 역사에서 가장 심각한 논란거리가 되었고 소위 ‘인공지능의 겨울’이 찾아오게 하였지만). 1974년에 제시된 그의 프레임(A framework for representing knowledge) 개념은 인공지능과 컴퓨터 프로그래밍에 새로운 패러다임을 가져다 주었다.

32) 계산심리학(computational psychology)에 대한 최근의 정의 참고 사이트:  
<http://www.encyclopedia.com/doc/1O11-computationalpsychology.html>

### 7.3. 1956년 가을의 MIT 정보이론 심포지엄

이어서 같은 해 9월 11일부터 사흘간 MIT에서 “정보이론 심포지엄(Symposium on Information Theory)”이 열렸다. 이 학술모임에서 발표된 주요 논문은 다음과 같다.

1. Newell & Simon : "Logic Theory Machine" (; 계산 기계에서 사용되는 수학적 정리의 증명 제시)
2. Chomsky: "Three models of Language" (; Shannon의 정보이론적 접근이 'natural language'에는 적용될 수 없다는 점에서 변형문법,transformational grammar, 에 기초한 언어 산출 모델1 제시)
3. George Miller: "Magic Number 7 + or - 2" (; 인간의 기억이 정보처리적 과정을 보임과 처리용량의 중요성 및 한계 제시)

이 학술모임이 인공지능은 물론, 인지주의의 출발에 큰 획을 긋고, 인지과학이 출발하게 되는 촉발적 모임이었다라고 할 수 있다. 이 학술 모임은 인지과학이 출발하는데 아주 중요한 역할을 하였다. 여러 아이디어들이 토론되고 수렴되고 결국은 정보처리 패러다임의 인지주의와 인지과학이 탄생하는 기폭제의 역할을 한 것이다.

뉴웰과 사이먼은 1954년 여름에 인간의 문제해결에 대한 이론을 컴퓨터에 모사할 수 없을가에 관해 이야기를 나누었고, 1954년 패턴인식에 관한 연구를 RAND회사에서 보고하였다. 이들은 단순한 패턴 인식(재인) 장치가 아니라 정보의 변환을 수행하고, 몇 가지 논리를 가진 인공지능 프로그램을 만들었다. 1956년 여름에 그들은 ‘Thinking Machine’ 프로그램을 개발하였고, Dartmouth 학술 모임에서는 그들은 인간의 지능이 요구되는 과제에 적용되는 정보처리모델이 작용하는 예인, ‘논리이론가 (Logic Theorist)’라는 인공지능 프로그램을 발표하여, '인공지능'이 하나의 과학이라는 주장에 형식적 정당화를 처음으로 제공하였다.

이 모임의 참석자, 발표자 중에서도 인지주의와 인지과학의 입장에서 보면 사이먼<sup>33)</sup>과 뉴웰(Newell)이 끼친 영향을 높이 평가할 수 있다. 그들은 기계와 마음의 연결의 형식적 산물인 ‘Thinking Machine’을 창안하였던 것이다. 그들은 1950년대에 이미 컴퓨터는 수뿐만 아니라 정보를 처리하는 장치로 볼 수 있으며, 컴퓨터는 비수학적

---

33)사이먼(Herbert A. Simon)은 인공두뇌학에 관심을 지녔던 아버지의 영향을 받아서 행정학에 관심을 가졌다. 1948년에 출발한 카네기 공대(Carnegie Tech)의 GSIA(산업경영 대학원)에서 행정학, 경제학 관련 연구를 하며 회사, 지방자치, 정부정책, 체계이론(System Theory)를 발달시켰다. 뉴웰(Allen Newell)을 만난 이후 1952년 뉴웰과의 이야기에서 도출된 아이디어에 흥미를 느끼고, 행정조직의 의사결정의 매우 적절한 대응책으로서 문제해결을 생각하기 시작하였다. 그들은(A. Newell, J. C. Shaw, H. A. Simon 세 사람) 미국 기업 RAND 회사에서 50년 초기에 팀을 형성하였는데, 이 당시의 RAND의 프로젝트는 대공방어의 모의실험으로 인간-기계의 상호작용을 연구하는 것, 인간 worker의 실질적인 수행(performance) 증가를 보려고 하는 데에 있었다.

(nonnumerical) 상징을(기호를) 조작할 수 있는 능력이 있다고 보았다.

마음도 어떤 전제를 가지고 그것을 바탕으로 결론을 처리하기에, 즉 마음은 입력을 받고 자료를 어떤 처리과정을 거쳐서 처리하고, 그 결과를 출력하는 것으로 볼 수 있다고 생각하였다. 실생활의 각종의 지적 관여를 요구하는 문제를 다루는데 있어 두뇌와 컴퓨터가 유사하게 정보처리 한다고 생각하였던 것이다. 그들에 의해서 마음과 기계의 연결이 형식적 체계로 구체화된 것이다.

#### 7.4. 1958년 가을의 영국의 NPL 심포지엄

이어서 영국에서는 국립물리학실험실 중심으로 ‘사고과정의 기계화("Mechanization of the thinking process)"라는 심포지엄이 기획되고 이 심포지엄이 1958년 11월 24-27일에 영국 테딩턴(Teddington)에서 개최되었다.

이 학술모임에서는 사고과정 기계화 관련 일반적 원리, 자동적 프로그래밍, 언어 번역, 말 인식, 기계학습, 생물학에의 의의, 산업에의 의의 등의 주제가 발표되었고, 초빙강연자들이 발표한 내용의 일부(발표자)를 들면 다음과 같다: 상식으로 프로그래밍 하기(맥카시), 지능의 오페레이션 측면 (맥케이), 연속 변환기 퍼셉트론 (로젠블랫), 기계의 자아 의식: 내적 아날로그의 기능과 통합-휴리스틱 프로그래밍 (민스키), 가설 만들기에 관하여 (유틀리), 팬더모니엄 패턴인식 모델(셀프리지), 휴리스틱 프로그래밍 등. 또 도서관 문헌 인출, 말 인식, 조건확률, 동물행동을 흉내내는 블랙박스, 자동 패턴인식, 자동 컴퓨팅엔진, 개념의 성장과 그들의 물리적 아날로그 등도 논의되었다.

참석한 사람들 중에는 저명한 미국의 수학-사이버네틱스 학자 맥컬러, 영국 시각 심리학자 그래고리, 영국 사회인지심리학자 바틀레이(Bartlett), 영국 심리(정신)병의 학자-사이버네틱스학자 로스 애쉬비(Ashby) 등이 있었다. 이 모임은 미국의 Dartmouth 모임이나 MIT 정보이론 모임에 이어서, 급변하는 기계(컴퓨터)-마음 연결에의 소용돌이 사조의 떠오름과 그에 따른 지적흥분을 재확인하는 모임이었다고 할 수 있다.

#### 7.5. 1959년의 파리 유네스코 정보처리 학술대회

정보이론의 대두, 디지털 컴퓨터의 발전, 컴퓨터 과학이라는 새로운 분야의 떠오르려는 움직임, 인공지능 영역의 떠오름 등의 새 변화를 감지한 미국의 관련 여러 학회들로 구성된 미국컴퓨터 연합위원회(Joint Computer Committee, USA)는 이러한 새 변화에의 흐름을 수렴할 필요성을 인식하여 UNESCO에 ‘국제적 정보처리 학술대회’ 개최를 제안하게 되었다. UNESCO는 그 기관의 자연과학 국제자문위원회의 심의 결과 이 제안을 승인을 하고 2년여의 기획 후에 1959년 6월 15일에서 20일의 옛세에 걸친 학술모임을 개최하였다.<sup>34)</sup>

34) 소개: <http://comjnl.oxfordjournals.org/content/1/3/105.full.pdf+html>

이 모임은 6개의 학술 주제를 초점으로 발표와 토론이 이루어졌다. 그 주제들은 다음과 같았다: 디지털 컴퓨팅과 수리적 분석, 디지털컴퓨터의 논리적 디자인, 디지털 컴퓨터를 위한 보통(Common) 기호언어, 언어의 자동번역, 패턴인식 및 역학(기계학), 정보의 수집, 저장, 인출. 50여개의 초청 강연 발표와, '미래의 컴퓨터 테크닉' 등의 12개의 소 심포지엄으로 구성된 이 학술모임에 유럽을 비롯한 전 세계의 정보, 컴퓨터 전문 연구자들이 모여서 발표를 하였다.

이 모임에서 다섯째 날인 6월19일에 패턴인식 및 기계학습 초청강연 세션에서 뉴웰 등(Newell, Shaw, Simon)은 범용문제해결(General Problem Solver, GPS) 프로그램을 발표하였다. 이 프로그램은 아마도 '문제에 대한 지식(서술적 지식)'과 '그 문제를 해결하는 전략에 대한 지식(절차적 지식)'을 나누어 접근한 인류 최초의 컴퓨터프로그램이었다고 할수 있다. 이 GPS 프로그램은 오늘날에는 널리 알려진 '하노이탑'과 같은 단순한 사고(추리) 문제를 '수단-목표' 같은 휴리스틱을 적용하여 하위 목표 상태를 검색하여 문제를 해결하는 논리 프로그램이었다. 인간의 마음 작동에 의한 문제해결적 사고(추리) 과정이, 컴퓨터라는 기계에서 문제내의 가능한 상태들을 검색하는 과정들로서 컴퓨터 언어(IPL프로그래밍 언어)로 형식화되어 표현될 수 있음을 보인 것이다.

이러한 기계(컴퓨터)와 마음(문제해결적 추리)을 연결하는 뉴웰 등의 시도, 그리고 이 학술모임에서 거론되고 토론된 많은 내용들은 전세계적으로 정보, 컴퓨터, 형식적 언어, 인지 등에 대한 학문적 인식을 높이고 인지주의, 인공지능, 컴퓨터과학 등이 새로운 학문적 연구 영역으로서 자리잡는 데에 기여하였다고 생각된다.

## 8. 1960년대 이전의 심리학 내의 움직임: 새 심리학의 태동

기계, 뇌, 마음(주로 논리적 사고)을 개념적으로 그리고 이론적으로 연결하려는 이러한 움직임과 디지털 컴퓨터와 프로그래밍 기법의 발전이 공학자들을 중심으로 진행되고 있는 사이에, 그와는 독립적으로, 기계-마음 연결의 다른 한 측면인 '마음' 영역에서 경험과학적 연구를 진행하여 오던 심리학에서는 그 나름대로 많은 변천이 있었다. 이러한 심리학계의 움직임을 살펴보면 다음과 같다.

### 8.1. 신경생리심리학 연구

첫째는 19세기 초의 신경생리심리학적 연구의 영향이다. 신경생리심리학적 연구의 주 입장인 생리심리학적 기계론은 라메트리(La Mettrie) 등의 관점을 이어 받은 기계

---

제안서: <http://unesdoc.unesco.org/images/0012/001267/126713EB.pdf>

론으로, ‘생리적 메카니즘’이 ‘심리적 메커니즘과 구조’에 선행됨을 심리학자들로 하여금 믿게 했다. 이러한 생리적, 기계론적 메커니즘(기제)이 나타나는 심적구조(외현적 행동이 아니라)를 이론적으로 찾는 바탕을 제공했다. 1940년대의 신경심리학자 헤브(D. Hebb)의 세포군(cell assembly) 개념이나 집단활동(mass action) 개념은 조직화된 구조를 지니는 내적 표상의 개념을 함축하고 있었던 것이다.

캐나다 신경심리학자 헤브(Donald O. Hebb; 1904-1985; [http://en.wikipedia.org/wiki/Donald\\_O.\\_Hebb](http://en.wikipedia.org/wiki/Donald_O._Hebb))의 신경세포들에 대한 이론, 특히 학습 이론은 신경과학에 상당한 영향을 주었다. 그는 행동주의 심리학 이론이 신경계에서 신경신호가 일방향적으로만 전달된다는 잘못된 가설에 기초하고 있었다는데 있음을 지적하였고, 1949년에 출판된 그의 저서, [행동의 조직화]에서 세포군집이론(cell assembly theory)을 제시하여 신경계의 시냅스에서 신경흥분이 지속되거나 반복되면 그것이 시냅스의 가소성을 변화시키고, 따라서 함께 발화하는 세포들은 함께 연결되게 된다는 요지의 학습이론을 제시하였다. 이 개념과 이론은 1980년대와 그 이후의 신경망 이론에서 기본 개념과 이론으로 사용되어 왔다. 그의 이러한 개념은 자극-반응의 단순한 연결만을 강조하는 행동주의 심리학 이론에서 심리학이 벗어나서 정보처리 이론을 추구할 수 있게 하는 하나의 배경을 제시한 것이다.

## 8.2 기능주의 심리학

다른 영향은 기능주의(functionalism in psychology)가 행동주의로 변형되기 이전에 제기했던 바의 심적 과정에 대한 생각들이다. 초기 기능주의의 관점을 지닌 윌리엄 제임스(W. James) 등의 심리학자들은 마음의 적응적 특성을 강조하며, 마음이 조직화되어 있고 여러 다른 기능 또는 구조로 구성되어있다고 보고, 심리학의 과제란 의식의 전형적인 기능적 작용을 발견하는 것이라고 보았다. 이들의 생각에는 정보처리적 인지 심리학의 ‘구조’, ‘과정’ 등의 개념들이 함축되어 있었던 것이다.

## 8.3. 형태심리학

이러한 변화의 배경이 되었던 다른 한 흐름은 형태주의(Gestalt) 학파 (이정모, 1983)의 영향이다. 형태주의 심리학자들이 심리적 환경과 마음의 구조적 특징을 강조한 것은 인지심리학적 개념들의 선구적 배경이 되었다. 형태주의를 행동주의와 조합한 톨만(E. C. Tolman)의 목적적 행동 인지도(cognitive maps)의 개념도 인지심리학 개념의 선구적 개념이다. 또한 레빈(K. Lewin)이 시작한 장(場, field)이론 중심의 사회심리학도 인지적 구조와 과정 개념의 발전을 촉진시켰다.

## 8.4. 심리측정, 기타

심리측정학(psychometrics) 분야에서 지능, 적성, 성격 특성들을 분석하고 검사를 개발한 것도 인지적 기능, 구조, 과정 등에 대한 개념을 발전시키는데 도움을 주었다고 할 수 있다. 자극을 지각함에 있어서 개인의 판단과 결정 경향을 강조한 신호탐지이론(signal detection theory: SDT)을 중심으로 한 정신물리학(psychophysics) 이론과 개념들도 인지주의적, 정보처리적 심리학을 형성시키는 데에 촉진적 역할을 했다고 할 수 있다.

한편, 프로이트(S. Freud)의 이론도 비록 그 과학적 타당성은 현재의 심리학자들에 의해 논란되고 있으나, 심적 구조와 역동적 과정 그리고 기제의 개념을 이론적으로 부각시키고 가다듬었다는 점에서, 인지 심리학의 개념들이 나타날 수 있는 간접적 배경을 제시해 주었다고 할 수도 있다.

### 8.5. 언어학습과 기억연구

행동주의 심리학 틀에 반발하여 인지심리학의 출발을 자극한 전통적 심리학 내의 언어학습 및 기억 연구에서의 선구적 개념들과 입장들이 다음과 같이 있었다.

마음의 구조나 내용을 인정하지 않는 행동주의 접근의 부적절성에 대한 자각은 심리학 안에서 언어 학습과 기억에 관한 연구를 중심으로도 강하게 일어났었다. 행동주의의 자극-반응 연합 이론에 의하면 우리가 학습하는 것은 접속성에 의한 자극과 반응의 단순한 기계적 연합이다. 따라서 학습된 양식(특히 학습된 순서)대로 자극-반응 연합을 기억해 내야 할 것이다. 그런데 1950년대의 언어학습과 기억에 관한 심리학적 연구 결과는 행동주의 이론의 예언과는 어긋나는 실험 결과들을 보고하기 시작했다.

50년대 초에 활발히 이루어졌던 언어학습(verbal learning)에서의 덩이짓기(clustering) 연구에 의하면, 심리학 실험에서 피험자들에게 여러 단어들을 무선적으로 제시하여도, 피험자들은 제시된 순서대로 기억해 내는 것이 아니라 단어들 사이의 일정한 관계, 범주적 관계나 연상관계에 의해 몇 개의 덩이로 다시 덩이지어 회상해 낸다는 현상이 드러났다. 즉 주어진 자극들의 물리적 시공간의 접속성에 의해 기계적으로 자극과 반응을 수동적으로 연합 학습하였다가 그대로 기억해내는 것이 아니라, 인간은 자극 간의 어떤 의미적 관계성을 파악하여 그 관계성에 의거하여 새로운 순서로 재구성하여, 조직화하여 기억해 낸다는 것이다. 이러한 결과는, 자극과 반응을 재구성하고 조직화하는 어떤 고등 정신 작용을 상징하지 않을 수 없게 했으며, 이러한 작용이 상당히 체계적으로 조직화된 인지 구조를 근거로 자극의 수용과 산출에 계속 관련되고 있음을 인정하지 않을 수 없게 하였다.

이러한 입장은 수동적이고 기계적인 연결상자로서의 과정만을 자극-반응 사이에 인정하던 행동주의 접근방법에 본질적으로 어긋나는 것이다. 지나치게 강조된 객관적 방법론과 극단적으로 편협화되고 단순화된 개념체계의 불균형으로 인해 행동주의가 심적 과정(앞을 중심으로 한)을 설명하는 데에 부적합함을 시사하는 결과이었다. 행동

주의에 대한 이러한 비판적 논의가 언어 학습과 기억 과정의 연구를 중심으로 1950년대에 제기되었지만, 실상 행동주의에 대한 비판과 대안적 접근은 그 이전부터 심리학 내에서, 그리고 후술할 인접 학문들의 영향을 받아 1950년대에서 1960년대에 걸쳐서 면면히 이어져 왔었다.

## 8.6. 유럽의 인지심리학 선구자

20세기 초엽의 미국 중심의 행동주의의 영향을 덜 받은 유럽에서는 인지주의적 및 비행동주의적 개념과 경험적 심리학적 연구가 비교적 뚜렷한 형태를 유지하였다.

그 대표적 연구의 하나가 1930년대의 영국의 바틀레이(F. C. Bartlett)를 중심으로 한 기억 및 사고 연구이다. 바틀레이는 기억이 수동적 재생 과정이 아니라 능동적 구성 및 재구성 과정이며 의미에의 노력(effort after meaning)임을 기억 실험을 통해 보였고, 사고도 여러 가지 신념과 태도 등에 의해 구성되는 것임을 보였다. 이에 따라 능동적 구성으로서의 사회적 마음 개념과 능동적 구성을 결정하는 사전 지식 또는 스키마(schema)라는 심적 구조의 중요성이 부각되었다.

한편 스위스에서 피아제(J. Piaget)는 지적 발달의 조작적(operational) 측면들을 강조하면서 인지 과정, 인지 구조, 생득적 지식 등의 측면들을 부각시켰으며, 사고와 지능에 관한 경험적, 이론적 연구의 발전을 자극시켰다.

한편 독일의 형태심리학 전통의 둥커(Karl Duncker), 셀츠(Otto Seltz)와 드그룻(Adriaann de Groot) 등은<sup>35)</sup> 서양장기 두기 등의 문제해결 과정 연구를 제시하여 인지적 연구의 가능성을 보였고, 소련의 비고스키(L. S. Vygotsky), 루리아(A. Luria) 등은 언어, 사고, 기억 과정들을 분석하고 인지적, 사회적 측면이 강조된 이론을 제기하여서 또 다른 배경을 제시하였다.

## 8.7. 미국의 인지심리학 선구자

정작 정보처리적 인지 심리학이 탄생한 미국에서의 인지 심리학의 직접적 선구는 1950년대 이전에는 드물었다. 1939년에 무어(T. V. Moore)라는 심리학자가 [인지 심리학]이라는 책을 써서 반응시간 방법을 적용한 인지 심리학적 연구를 주장하였으나, 별 영향을 주지 못하였다. 1949년과 1950년대의 초에 포스트맨(L. Postman), 밀러(G. A. Miller), 가너(W. R. Garner) 등이 낱개 논문이나 책에서 '인지'나 '정보'의 개념을 제기하였으나, 이것도 직접적 영향을 초래하지 못하였다. 특히, 밀러나 애트니브(F. Attneave) 등이 정보이론을 심리학에 도입하였음에도 불구하고 그들은 인지적,

35) 독일의 형태주의 심리학 전통의 연구와 미국 기능주의 등 학문 전통이 사이먼과 뉴웰의 문제해결 인지시뮬레이션 연구에 준 영향의 흐름도를 Herbert Simon (1999). 'Karl Duncker and Cognitive Science'의 7쪽 그림1에서 참고할 수 있다:

(<http://octopus.library.cmu.edu/cgi-bin/tiff2pdf/simon/box00071/fld05480/bdl0001/doc0001/simon.pdf>)



심리적 개념을 심리학에서 사용하는 것을 비과학적 접근으로 보았거나, 정보의 분석 접근을 심리학적 접근으로는 그리 중요하지 않다고 보았다.

이러한 비교적 활발하지 못한 인지적 접근 시도에 비교적 큰 영향을 준 것은 브루너 등(Jerome Seymour Bruner, Jacqueline J. Goodnow, & George A. Austin)의 책, [사고의 연구(A Study of Thinking)](1956)와, 밀러 등(G. A. Miller, E. Galanter, & K. H. Pribram, 1960)의 책, [행동의 계획과 구조(Plans and the Structure of Behavior)](1960)라는 두 책의 발간 및 그 영향과, 브루너와 밀러가 주축이 되어 처음으로 하버드 대학에 ‘인지연구 센터(Harvard Center for Cognitive Studies)’를 1960년에 설립한 것의 영향이었다.

브루너는 [사고의 연구] 책을 통해 행동주의자들이 ‘자극-반응’ 연쇄에 의해 사고과정(특히 개념형성과정)을 설명하려 한 것이 부적절함을 보이고, 인지자 개인의 인지적 가설형성 과정과 인지적 전략의 중요성을 제시하여 행동주의에 대한 대안적 접근의 가능성을 보였다. 브루너 교수의 당시의 관심 주제이었던 개념 형성 및 발달과 가설 검증의 합리적 사고 행동 주제 중심으로 전개된 이 책은 기존의 행동주의 심리학 대신에 새 심리학의 틀을 추구하여야 할 배경 근거를 제시하였다.

다음으로, 1960년의 밀러 등의 책은 행동주의 심리학의 ‘자극-반응’ 연결 강조 대신에 ‘계획’을 행동의 기본 단위로 대체해야 함을 주장하였다. 행동주의 심리학자들이 주장하듯이 ‘자극-반응’의 연결에 의해서 행동이 이루어진다고 하기보다는, 자기 조절하는 체계가 정보의 피드백(되먹임, feedback)에 의해 행위를 인도해 나아가는 것이라고 보았다.

즉 인간이 문제해결하는 것은 계획이라고 하는 다양한 전략을 사용하는 것이다. 계획이란 기초적 특정 행위단위와 하위목표 및 상위목표들이 위계적으로 조직화된 체계가, 지향하는 목표 상태와 현재 상태를 비교하고, 다음에 이 둘 사이의 차이를 감소시키는 행위단위 연쇄를 작동시켜, 계속되는 정보적 피드백 내용을 참고하여 행동을 외현적으로 내놓는 것이라고 보았다.

이러한 관점은 행동주의적 틀을 완전히 버지는 못했으나 심적 구조, 정보처리 프로그램, 서브루틴, 정보의 흐름, 통제 등의 개념들의 선구를 제시해 준 것이며, 뉴웰 Newell 과 사이먼(1961)이 정보처리체계를 구체적으로 개념화하고 정보처리적 접근을 정식으로 진수시키는 데에 결정적 바탕을 제공해 주었다.

이 두 책과 함께 마음-기계의 연결에의 ‘인지혁명’이라는 새 패러다임의 형성을 촉진한 다른 중요한 문헌은 다음의 다섯을 들 수 있다. : 밀러(George A. Miller)의 1956년 [Psychological Review]지 논문 “마술적 숫자  $7 + - 2$ (The Magical Number Seven, Plus or Minus Two)”, 영국 심리학자 브로드벤트(Donald

Broadbent)의 1958년 책, [지각과 커뮤니케이션(Perception and Communication)]. 미국 언어학자 촘스키(Noam Chomsky)의 1959년 논문 “스키너의 책 ‘언어행동’에 대한 서평(Review of Verbal Behavior, by B.F. Skinner)”, 뉴웰 등(Newell, Shaw, and Simon)의 “인간의 문제해결 이론의 요소들(Elements of a Theory of Human Problem Solving)”, 나이서(Ulric Neisser)의 1967년 책 [인지심리학(Cognitive Psychology)].

이 문헌들이 마음-기계 연결과 인지주의 형성에 미치는 영향들은 여기의 다음 절들에서 설명된다.

한편, 1960년에 설립된 [하바드대 인지연구센터]는 아마도 인문과학, 사회과학, 자연과학, 공학 등 사이에 전문 연구자들 사이에 생각을 나누는 학문간 진정한 융합 연구 센터의 역할을 하게 하는 그러한 연구센터였다. 이 연구센터를 통하여 마음과 뇌, 기계의 연결, 인지주의의 형성이 크게 조장되었다.<sup>36)</sup>

이러한 심리학 내의 반(反)행동주의적 움직임에 박차를 가하여 행동주의 심리학을 붕괴시키고, 대안으로서 정보처리적 접근 방법에 기초한 인지주의를 형성시켜 인지과학이 출발할 수 있는 배경을 심리학 밖에서 제공한 여러 인접 학문들의 영향, 사조들이 있었다. 이 인접학문들이 위에서 기술한 철학, 신경 모형, 커뮤니케이션 이론, 사이버네틱스, 정보 이론, 언어학, 컴퓨터 과학 등이었으며, 이에 첨가하여 상응하는 움직

36) 정보 또는 정보처리의 개념을 중심으로 1950-1960년대 일어난 마음-뇌-컴퓨터(기계) 연결의 과학적 혁명인 ‘인지혁명(Cognitive revolution)’과 그와 관련된 하바드대 인지연구센터의 활동과 역할 자료를 찾는다면 앞서 언급한 Howard Gardner(1985), Baars(1986), Boden(2006)의 책 내용 이외에도 다음 자료들을 참고할 수 있을 것이다:

- [1]. 위키피디아 자료 인지혁명: [http://en.wikipedia.org/wiki/Cognitive\\_revolution](http://en.wikipedia.org/wiki/Cognitive_revolution)
- [2]. 인지혁명에 대한 하바드대 심리학과 밀러 교수의 회고  
(‘The cognitive revolution: a historical perspective’ / George A. Miller / TRENDS in Cognitive Sciences Vol.7 No.3 March 2003, pp. 141-144/ (<http://www.cs.princeton.edu/~rit/geo/Miller.pdf>)
- [3]. 하바드대 역사학자 코헨콜 교수의 글: (‘Instituting the science of mind: intellectual economies and disciplinary exchange at Harvard’s Center for Cognitive Studies’/ by Jamie Cohen-Cole/  
- BJHS 40(4): 567-597, December 2007./ doi:10.1017/S0007087407000283/  
- 이 글에는 110여개에 달하는 각주가 있는데 그 상당수가 참고고문헌들이다.  
<http://www.polthought.cam.ac.uk/seminars/intros2008-2009/Cohen-Cole.pdf>
- [4]. ‘인지혁명’에 대한 하바드대 심리학과 교수 스티븐 핑커의 설명 동영상 자료:  
<http://news.harvard.edu/gazette/story/multimedia/the-cognitive-revolution/>  
- 그 동영상 강연에 사용한 ppt 자료 내용  
[http://pinker.wjh.harvard.edu/lectures/Rediscovery\\_Labels\\_revised.pdf](http://pinker.wjh.harvard.edu/lectures/Rediscovery_Labels_revised.pdf)
- [5]. 인지심리학 소개 사이트에서 설명하는 인지혁명  
[http://findarticles.com/p/articles/mi\\_g2699/is\\_0000/ai\\_2699000063/](http://findarticles.com/p/articles/mi_g2699/is_0000/ai_2699000063/)
- [6]. ‘Karl Duncker와 인지과학’  
- (사이먼이 형태심리학자 칼 덩커 등이 자신들의 문제해결 프로그램 도출에 영향 준 관계를 설명하는 글)  
<http://octopus.library.cmu.edu/cgi-bin/tiff2pdf/simon/box00071/fld05480/bdl0001/doc0001/simon.pdf>
- [7]. 구글 검색: ‘Harvard Center for Cognitive Studies’

임으로 심리학 내에서 나타난 것이 인간공학 연구, 브로드벤트의 정보처리 이론 등을 들 수 있다.

## 8.8. 인간공학 연구

정보처리 또는 정보유통의 개념은 2차 세계대전을 전후로 인간공학적 연구를 통하여 심리학에 도입되었다. 2차대전 중에 종군하면서 어떻게 하면 각종 병기나 기계를 가장 효율적으로 다룰 수 있는가, 효율적인 기기로 고안할 수 있는가 하는 문제를 연구 하던 심리학자들은 종래의 행동주의 심리학의 자극-반응 학습에 대한 연구가 심리현상의 설명에 아무런 도움을 줄 수 없음을 깨달았다.

그들은 인간을 단순한 자극-반응 학습자로서가 아니라, 기계와 상호작용함에 있어서 기계의 계측판과 기계의 통제 사이에 개재하면서 계측판 등에서 각종 정보를 입수하고 처리하여 결정을 내려 기계를 제어하게 하는 정보제어 체계 또는 정보전달 체계로 볼 필요성을 느끼게 되었다. 따라서 그들은 외부의 정보가 어떻게 인간에게 들어와 흘러가며 처리되는가에 관심을 지니게 되었다. 그 결과, 학습 과정에 대한 연구보다는 지각, 기억, 주의, 신호탐지, 결정 등의 문제와 이러한 문제들을 포괄하는 인지 과정에 대한 연구가 활발히 이루어지기 시작하였다.

## 8.9. 브로드벤트의 공헌

이러한 배경 위에서 체계 및 정보 유통 개념과 이론을 가다듬어 심리학에 체계적인 적용을 시도한 학자들은 브로드벤트와 그의 동료들이다. 브로드벤트(Broadbent, 1958)는 인간의 지각과 의사소통이 즉각적으로 이뤄지는 것이 아니라 단계적인 자극 정보처리 과정에 의해 이뤄진다고 보았다.

그에 따르면 자극 정보처리 과정은 다음과 같다. 자극이 감각기관의 통로를 통해 정보처리 체계 안으로 들어오면, 들어온 자극은 임시 저장고인 단기기억에 머물러 있게 되고, 이것이 선택적 여과기에 의해 통로별로 또는 자극 유형별로 선택되어 제한된 용량을 지닌 중앙처리기로 넘어가며 여기에서 보다 의미있는 정보로 추상화 처리된다. 그러한 처리 결과는 기억에 저장되든지 아니면 반응 생성을 통제하는 체계를 거쳐 어떤 반응으로 산출된다. 구체적으로 브로드벤트는 신호탐지 상황 등에서 선택적 주의와 관련하여 이러한 정보처리 과정이 어떻게 이론화될 수 있는가를 실험적으로 보였다.

이러한 브로드벤트 모형은, 자극과 반응의 단순한 연합에 의해 인지 과정을 설명하던 종래의 행동주의 이론의 부족함을 드러내 주었고, 인간 내부에서 어떻게 자극이 지각, 분석, 조직, 해석되어 처리되는가 하는 내적 처리 과정의 중요성에 대해 학자들의 주의를 환기시켰으며, 이러한 과정들을 설명하는데 사용할 수 있는 정보처리적 개념들을 체계화하여 제시함으로써 인지심리학이 어떠한 방향으로 형성될 수 있는가 하

는 본보기를 보여주었다.

이와 같이 심리학이 정보처리적 인지과정의 문제를 중심으로 인지주의와 마음-기계의 연결을 지향하는 패러다임적 전환 방향으로 움직여가는 단계에서, 심리학 밖에서 결정적인 박차를 가하며 사이버네틱스가 추구하던 생각들을 결집시키어 하나의 새로운 과학적 패러다임으로 마음-기계의 연결의 ‘인지주의(Cognitivism)’를 탄생시키게 한 직접적인 자극은 심리학의 인접학문인 언어학과 컴퓨터과학으로부터였다.<sup>37)</sup>

## 9. 인지주의, 인지과학의 탄생과 확산

### 9.1. 언어학의 영향

1950년대 후반에서 시작하여 촘스키(Chomsky)를 비롯한 언어학자들과 언어심리학자들은 이전에 심리학계의 전형적 틀이었던 행동주의 접근이 인간의 언어 행동을 설명함에 있어서 아주 부적절하며 새로운 접근이 필요함을 보여주었다.

그들은, 언어란 본질적으로 규칙적이며 구조를 지니고 있으며, 과거의 경험에 구속되지 않고 문법 규칙에 따라 새로운 언어 표현을 얼마든지 변형 생성해 낼 수 있고, 언어의 외적 표현 행동인 수행(performance)이 중요한 것이 아니라, 그것을 가능하게 하는 내적 언어능력(competence)이 중요하며, 이러한 언어 능력은 경험적으로 자극-반응 조건형성에 의해 습득되는 것이 아니라 생득적으로 주어지는 것임을 주장하였다.

그들은 이러한 주장들에 대한 증거로 아주 단순하고도 명백한 언어 표현 또는 이해의 예를 제시하였다(Chomsky, 1959). 이러한 논의는 언어를 중심으로 한 인간의 심적 과정이 단순한 자극-반응 연결의 집적 또는 그것의 단순한 재현이 아니라, 어떤 규칙에 의하여 변형되고 생성, 처리되는 특질을 지니고 있는 인지적 구조와 기능을 중심으로 이루어지는 과정임을 보여주었다.

촘스키는 심리학 내에서 언어에 대한 행동주의적 접근, 특히 스키너(B. F. Skinner)의 언어행동이론이 잘못된 것이며, 행동주의자들이 주장하듯이 외현적 자극-반응의 연결에 의해 언어를 설명할 수 있는 것이 아니라 심적구조, 문법적 규칙, 생득적 지식 등에 의해서 인간의 언어행동을 설명할 수 있음을 주장하였다. 촘스키가 행동주의 심리학적 접근의 문제점을 어떻게 공격하였는가에 대하여 이정모(1988c, 51-52쪽)는 상세한 설명을 제시한 바 있다.

---

37) 다음 사이트에서 제시된 바와 같이, [컴퓨터과학]의 초기형성단계는 [인지과학] 및 [인지주의]의 초기 형성단계와 튜링기계이론, 사이버네틱스, 인공지능 등의 영역에서 그 이론적 주제가 중첩되어 있다. 컴퓨터과학의 형성을 이야기 하면서 인지과학의 형성 배경을 언급 안한다면 그 분야의 바탕을 잘 모르고 있는 것이다. (A Very Brief History of Computer Science):

<http://www.cs.uwaterloo.ca/~shallit/Courses/134/history.html>

이러한 촘스키의 비판적 논변의 제기와 그에 대한 학자들의 수용의 결과로 행동주의 접근 방법이 언어 및 사고의 연구 방법 이론으로서 부적절함의 인식이 급격히 확산되었다. 이에 따라서, 이전에 행동주의자에 의해 추방되었던 심성적(mentalist) 개념들이 - 즉, 내적 구조, 심적 과정, 심적 능력, 관념, 의식, 사고, 인지 등 - 소생될 바탕이 이루어졌고, 지각, 기억, 언어, 사고 등에 관한 새로운 개념화와 방법론이 대두되었다.

이러한 새 움직임에 따라서 학자들은 규칙, 관계성, 구조와 같은 개념들을 재도입하여 심적 과정을 생각하게 되었다. 또한 정보처리, 정보유통의 개념들을 이에 융합하여, 언어이해, 언어생성, 사고, 문제해결 등의 과정들을 정보처리 과정들로 간주하고, 언어 자극의 지각에서부터 이들의 분석 및 의미 이해, 언어 처리 규칙들을 기억 속에 저장하는 과정 및 이러한 저장된 정보의 활용 과정 등에 대하여 정보처리적 모형을 세우고 이를 검증하기 시작하였다.

이러한 새로운 연구 접근의 시도와 진행 과정에서 언어, 사고, 및 인지과정들의 이해를 위해 보다 명확한 개념을 제시하고, 이론화하는 방법을 제시하며, 그러한 이론의 객관적 검증의 도구를 제공하며 커다란 영향을 준 것이 컴퓨터 과학이었다.

## 9.2. 인공지능과 컴퓨터 과학의 영향

앞서 언급한 바와 같이 1950년대 초의 디지털 컴퓨터의 발달과 더불어 튜링기계 이론이 확산됨으로써, 컴퓨터가 단순히 수리적 계산만 할 수 있는 것이 아니라 상징 표상도 가능하다는 것이 컴퓨터과학자들에 의해 인식되기 시작하였다.

일반적으로 인간이 어려운 수학 문제를 풀 때에 어떤 도약적이고 창조적인 지적 과정을 거쳐서 문제를 해결한다고 여겨져 왔으며, 이러한 과정은 확실히 밝혀 지적할 수 없고 설명하기 곤란한 과정이라고 생각되어 왔다. 그런데 수리논리학에서는 이러한 도약적 인지 과정을 명백하게 규정될 수 있는 (형식화(formalized) 할 수 있는) 단계 절차들로 밝히려고 노력해 왔다.

그런데 앞서 기술된 바처럼 1930년대 초에 수학자 튜링은 소수의 기본적인 규칙들에 근거한 어떤 추상적인 기본 논리체계와 몇 개의 실행 능력을 지닌 자동기계(automata)를 만든다면, 이러한 자동기계는 창의적인 인간 지능의 개입 없이도 모든 수학적 문제를 자동적으로 풀 수 있을 것이며 또 문제를 푸는 효율적인 단계 절차들이 명료히 규정될 수 있을 것이라고 보았다. 이러한 자동기계를 튜링은 튜링(보편)기계라고 불렀고, 이러한 보편 기계의 추상적 원칙이 디지털 컴퓨터와 같은 다른 기계 체계에도 적용될 수 있으며, 이러한 원칙이 적용되는 기계들은 서로 상대방 기계 체계의 행동 절차를 충분히 흉내(simulate)있으리라고 시사했다.

## 9.3. 인지심리학의 공식적 출발

이러한 변화의 움직임이 종합되며, 인접 분야 연구자들과의 협응적, 학제적 연구가 왕성하게 일어남에 따라, 또, 앞에서 기술한 바와 같이, 브루너 등의 [사고의 연구] 책과 밀러 등의 [행동의 계획과 구조]라는 두 책의 발간과, 하바드대학의 '인지연구센터' 설립 등, 그리고 기타 심리학 내의 변화에의 다른 움직임이 하나로 수렴되어 정보처리적 패러다임의 인지과학의 출발이 촉진되게 되었다.

1967년에 나이서 교수가 [인지심리학]이라는 책을 출간함에 따라 '인지주의'는 공식적으로 이름이 인정, 확산되고, '인지심리학'이 심리학 내의 독자적인 분야 및 연구 패러다임 틀로 자리 잡게 되었다. 이러한 정보처리패러다임의 인지심리학 및 인지과학이 출발하여 확고한 위치를 차지하는 데는 사이먼과 뉴웰의 기호(상징) 조작 정보처리체계 이론과 그들의 컴퓨터 시뮬레이션을 중심으로 한 형식적(formal) 연구 전략이 든든한 바탕이 되었다.

#### 9.4. 사이먼과 뉴웰의 생각: 컴퓨터와 마음 간의 유추.

보편기계인 튜링기계가 주는 의의는 상당히 큰 것이었으나 그러한 의의는 1950년대 후반까지는 올바르게 파악되지 못하고 있었다. 1956년 여름에 위에서 언급된 다트머스 대학 심포지엄과 가을의 MIT대학의 심포지엄에 모인 과학자들은, 수학자, 심리학자, 언어학자, 전기공학자 등으로 각자 여러 다른 배경을 가지고 있었지만, 그들의 생각은 점차 한 가지 공통된 점으로 모아지고 있었다. 그것은 다름이 아니라, 기계가 인간의 지능을 가질 수 있다는 것이고, 인간의 마음을 과학적이고, 형식적(formal)으로 설명할 수 있으며 또 그래야 하고, 튜링기계에론에 바탕을 두어서 이루어진 디지털 컴퓨터는 그것을 위한 가장 좋은 기계라는 생각이었다.

1958년 히슨(Hixon) 심포지엄에 참석하였던 사이먼과 뉴웰은 컴퓨터를 숫자이건 아니건 모든 종류의 기호를 조작할 수 있는 기계로 보았다. 대부분의 사람들이 컴퓨터를 연산장치로 보고 있던 그 당시로는 그들의 직관은 혁신적인 것이었다. 그들은 단순한 숫자의 조작기계로만 생각하던 컴퓨터를 숫자의 계산 수준을 넘어서 범용(汎用) 목적의 상징(기호)조작 체계(general purpose symbol manipulation system)로 볼 수 있으며, 튜링기계도 숫자 조작 체계의 수준을 넘어서 상징(기호)조작 체계로서 볼 수 있음을 주장하였다.

그들은 이어서 한 걸음 더 나아가 인간의 마음도 컴퓨터와 같이 기호(상징) 조작 체계로서 개념화할 수 있으며, 따라서 인간의 마음을 일반목적 기호(상징) 조작 체계인 튜링 기계의 현실적 구현으로써 간주할 수 있음을 제안하였다. 인간의 마음을 정보처리 체계로 본 뉴웰의 생각과 인간의 마음을 기호조작체계로 본 사이먼의 생각은 서로 사용하는 어휘가 달랐지만 그 의미는 똑 같았다.

이러한 사이먼 등의 생각은 혁명적인 것이었다. 인간의 사고 과정이 실체가 없는 추상적이고 불확정한 과정이 아니라, 튜링 기계라는 추상 체계를 지닌 자동기계 내에서 수리기호라는 상징을 조작하는 과정으로서, 명확하고 구체적인 절차와 과정에 의해 기술될 수 있고 구체적으로 조작될 수 있음을, 다시 말하여 수리기호 조작 과정이 물체나 물리적 과정과 같이 구체적으로 연구할 수 있는 대상임을 튜링이 보여주었다.

사이먼 등은 한 걸음 더 나아가 그러한 튜링 기계를 수리기호 처리만이 아닌 일반 목적의 기호 조작 체계로 개념화할 수 있고, 디지털 컴퓨터와 인간의 마음도 이러한 체계의 추상적 원칙을 구현하여 기호를 조작하는 체계 또는 기계로 볼 수 있다고 본 것이다. 그러므로 동일한 원칙을 구현하는 한(限), 컴퓨터라는 한 체계는 다른 한 상징(기호) 조작 체계인 인간의 마음의 작용을 충분히 모사(simulate)할 수 있다고 보았고, 따라서 인간의 마음의 내용과 과정을 컴퓨터 유추 개념과 컴퓨터 모의실험을 통해 명확한 구조와 절차로 기술하고 구체적으로 조작할 수 있으리라는 가능성을 제시한 것이다.

이러한 이론적 보는틀의 테두리가 일단 형성되자, 사이먼은 그의 동료들과 함께 이러한 정보처리적 보는틀의 개념을 가다듬고, 그러한 개념들과 컴퓨터 모의실험(computer simulation)을 통하여 인간의 심적 과정을 유추할 수 있는 세부적 방법론을 제시하였다(Newell & Simon, 1972). 그들은 컴퓨터와 인간의 마음의 공통 특징인 기호(상징) 조작 과정이란, 근본적으로 입력된 자극 내용에서 정보를 추출 또는 조작하여 출력하는 정보처리 과정이라고 간주한다. 이러한 정보처리 과정을 내포하고 있는 것이 정보처리 체계(information processing system: IPS)이며 그 기본 특성은 그들에 의하면 다음과 같다.

정보처리 체계란 기본적으로 사물을 기호(상징)로 표상(表象, representation)하는 능력과 기호를 조작하는 능력을 지니고 있다고 본다. 사물의 표상은 정보처리 체계의 기본 요소인 기호(상징)와 개개의 기호가 관계성(relation)에 의해 연결된 것들이 모여 이루어진 기호구조(symbol structure)로 표상된다.

여기에서 유의할 것은 사물에 대한 기호 표상 이외에 이러한 기호구조를 조작하는 처리 과정 자체도 기호의 형태로서 표상된다는 점이다. 이러한 처리 과정의 기호 표상이 곧 정보처리 프로그램이라 할 수 있다. 이러한 기호, 기호구조, 프로그램과 각종 자료들이 기억이라는 하위 구조 체계에 저장된다. 기호의 조작이란, 기호와 기호 구조를 입력으로 받아 이를 처리하여 다른 기호와 기호구조로 출력하는 과정을 일컫는데, 이것이 곧 정보처리 과정이다(Newell & Simon, 1972; Simon, 1981).

뉴웰과 사이먼은 어떤 정보처리 체계가 지니는 기초 정보처리 과정(Elementary Information Process: EIP)에는 다음과 같은 과정이 있다고 본다: 변별 과정(상이한 기호 또는 기호구조에 따라 다른 정보처리를 함), 비교, 검증 과정(기호의 예 : 토큰

token들이 동일한 개념 또는 기호의 원형(prototype) 또는 유형에 속하는가를 검증하는 과정), 새 기호의 창출 과정, 기호구조의 산출 과정, 기호구조의 복사 및 변형 과정, 외계 자극에 내적 기호를 부여하거나 또는 내적 기호구조에 외현 반응을 생성하는 과정, 기호의 참조 준거 지목 또는 연결하는 과정, 기호구조의 저장 과정 등. 뉴웰과 사이먼은 이러한 기초 정보처리 과정이 처리구조(processing structures) 요소 사이에서 작용함으로써 각종 수준의 정보처리가 이루어진다고 보았다.

다음에 그들은 인간의 심적 과정을 이러한 정보처리 체계의 구조와 과정들에 어떻게 대응시킬 수 있는가를 기술하였다. 인간이 대상을 지각하고 경험에 의하여 학습하며 기억하고, 언어를 이해하고 구사하며 사고한다는 것은, 인간의 감각기관, 기억, 중앙처리구조, 반응(운동)기관 등의 각 정보처리 구조들 사이와 각 구조의 내부에서 기호(상징)를 조작, 처리하는 정보처리 과정들이 일어난다는 것을 말한다. 그 결과로 자극에 대한 감각 정보가 부호화(바꾸어 풀어 넣기 작업, encoding)되어 기호로 표상된 다음 기억에 저장되고 저장된 기호 내용들이 후에 되꺼내어져(인출, retrieval) 내적인 기호구조의 변화, 또는 반응기관과 연결된 외적 행동의 변화로써 내어놓아진다(output).

그들은 이러한 심적 정보처리과정을 컴퓨터 모의실험을 통해 구현하여 연구할 수 있다고 보았다. 이러한 컴퓨터 모의실험(computer simulation) 방법의 도입으로 인한 영향은 매우 컸다. 이 방법은, 합리주의 전통에서 강조되었으며, 실험심리학 전통에서 배척되어 왔던 개념들(생득적 능력, 심적 구조, 도식(schema) 규칙, 통제 등과 같은 정신 기능적 개념들)을 구체적으로 조작, 기술할 수 있는 컴퓨터 체계의 상징조작의 개념들(통제 과정, heuristics, subroutines, production rules, encoding, retrieval 등의 개념들)로 재개념화하여 이성적 추론에 의해 명확히 규정할 수 있도록 했다. 또한 개인의 내성보고를 근거로 정보처리 세부 과정들을 프로그래밍하여 이를 모의실험을 통해 검증함으로써 합리주의 전통의 주요 방법이던 내성법을 부활시켰다. 동시에, 경험주의에서 강조하였던 경험적 관찰을 컴퓨터 모의실험 비교 방법으로 구현시킴으로써, 합리주의의 방법론과 경험주의 방법론을 창의적으로 결합한 새로운 방법론을 제시한 것이다. 이러한 컴퓨터 모의실험 접근 방법은 자연히 고등 정신 과정을 이전 보다는 폭넓게, 그리고 엄밀하게 기술하여 설명할 수 있게 하였고, 입력과 출력 사이에 개재하는 인간 마음의 작용에 대한 추론이 이전의 이론적 접근들에서보다 실제적인 마음에 더 근접한 모형으로 도출되게 하였다.

물론 컴퓨터 유추에 기초한 정보처리적 접근 방법의 기본 요소가 컴퓨터 모의실험만으로 이루어지는 것은 아니다. 모의실험 이외에도 프로토콜 분석(protocol analysis; 사임먼 등이 개발한 연구 방법; 실험 참여자의 내성 언어 보고 내용의 분석방법) 등 다른 방법론적 요소가 첨가되어 있다. 또한 컴퓨터 유추 모형에 의해 정보처리 과정을 설명하려고 하면서도 기본적 개념과 일반 절차만 도입하여 적용하고, 실제의 검증



은 전통적 실험 방법을 사용할 수도 있다. 특히 정보처리 과정의 개개 단계를 분리시켜 주는 힘을 가지고 있다고 간주되는 반응시간(reaction time) 측정 방법 등을 사용하여 실험 연구를 할 수도 있다.

이상에 기술된 컴퓨터와 은유와 기본 개념과 이론이 학계에서 처음부터 그대로 받아들여진 것은 아니었다. 최초에는 그저 단순한 하나의 유추 개념으로서 받아들여졌고, 1970년대에 이르러서야 비로소 컴퓨터 모형이 인지 과정에 대한 이론 형성 및 그 검증에의 도구가 될 수 있음에 대한 인식이 학계에 자리잡게 되었다.

사이먼과 뉴웰 등의 접근 방법의 기본 개념들과 방법론은 새로운 돌출구를 마련해 주었다. 여기에 정보 이론, 커뮤니케이션 공학, 언어학 등의 영향이 융합됨으로서 마음과 인지에 대한 기존의 개념 체계와 이론이 대거 수정되고 세련화되었으며, 정보처리적 접근이라는 새로운 보는틀에(인지주의: cognitivism) 의해 인지 과정과 마음 전반을 연구하는 인지과학(Cognitive Science)과 인지심리학이 인공지능 연구와 함께 과학계에서 새로운 그리고 활발한 연구 영역으로 자리 잡게 되었다. 물론 인지과학이 여러 학문들의 개념, 이론, 방법론들을 모두 다 수용한 바는 아니었다. 인지과학적 개념화와 연구 수행에 적절하다고 판단되는 부분만 수용되었고, 다른 부분들을 배제되었다. 마음-기계 연결의 새로운 틀인 인지주의의 구체적 학문으로 인지고학이 형성되는 과정에 영향을 준 개념, 이론 등을 축약하여 표로 정리하자면 다음 표와 같다.

**[표2]. ‘인지과학’이라는 새 패러다임의 등장에 영향준 사조들**

근원	새 과학 패러다임 형성에 영향 준 흐름들의 영향 내용 (개념, 이론, 방법 등)
구성주의	<ol style="list-style-type: none"> <li>1. 능동적 과정으로서의 마음과 심적 활동 개념 제공</li> <li>2. 기억, 주의, 의식 내용의 강조</li> <li>3. 심적과정을 분석, 통합하는 접근 제공</li> <li>4. 내성법 제공. 내성보고분석법 + 반응시간측정법 제공</li> </ol>
언어 학습 연구	<ol style="list-style-type: none"> <li>1. 자극-반응 학습보다는 심적 구조의 적용(조직화)가 주요 문제임을 제시</li> <li>2. 언어학습의 경험적 자료들 일부와 실험실 절차와 측정기술 등 제공</li> <li>3. 기억 분야의 강조와 연구활동이 활발한 실험심리학자들 제공</li> </ol>
철학	<ol style="list-style-type: none"> <li>1. 인식론, 심신론, 과학철학, 논리학의 바탕을 제공</li> <li>2. 마음과 컴퓨터의 은유 논의, 마음과 두뇌(물질)와의 관계 문제, 지향성(intentionality), 표상의 미의 파생 문제, 각종 심리 기능의 분화와 통일성, 단원성, 인지과학의 과학철학적 기초 등의 문제에 대한 논의와 개념들 제시</li> </ol>
인간 공학	<ol style="list-style-type: none"> <li>1. 인간을 의사결정자, 정보전달자로 유추하는 틀 제공</li> <li>2. 정보 개념의 중요성, 정보전달의 한계, 신호탐지이론 등 제공</li> <li>3. 공학과 물리과학의 개념들과 심리학의 개념들의 접근 마당 제공</li> <li>5. 고도화된 실험실 기법 제공</li> </ol>
커뮤니케이션 공학, 정보이론	<ol style="list-style-type: none"> <li>1. 커뮤니케이션 채널의 유추 틀 제공</li> <li>2. 정보 이론 및 채널 용량의 한계 개념 제공</li> <li>3. 직렬 대 병렬처리 등과 부호화, 해독, 인출 등의 처리과정 개념</li> <li>4. 불확정성 개념 등 제공</li> </ol>
컴퓨터 과학	<ol style="list-style-type: none"> <li>1. 인간이 상징조작 체계라는 틀 제공</li> <li>2. 디지털 컴퓨터와 컴퓨터 유추 (은유) 제공</li> <li>3. 저장된 프로그램 개념과 계산 개념 제공: 계산주의(computationalism) 틀 제공</li> <li>4. 컴퓨터 모의실험 (computer simulation) 방법을 과학의 한 주 연구방법으로 제공</li> <li>5. 복잡한 지적 행동에의 강조와 이를 형식적(formal)으로 분석하는 틀 제공</li> <li>6. 인공지능학 (후에 로보틱스) 영역의 제공</li> </ol>
언어학	<ol style="list-style-type: none"> <li>1. 심리학의 행동주의 틀을 거부할 이유와 논리를 제공</li> <li>2. 형식문법론을 중심으로 한 고차인지과정인 언어에 대한 형식적 접근 이론을 제공,</li> <li>2. 퍼포먼스 대신 내적규칙과 심적능력(competence)의 개념을 부상시킴</li> <li>4. 언어 영역의 중요성 강조</li> <li>5. 창조성과 규칙성에의 관심 되살림</li> <li>6. 능력과 수행의 구분</li> </ol>
신경과학	<ol style="list-style-type: none"> <li>1. 뇌와 마음(인지)의 관계 논의를 위한 경험적 바탕 제공 (1980년대 이전은 영향력 약했음)</li> <li>2. 인지신경과학 분야 형성 + 뇌영상기법 등 인지신경기법 및 자료, 이론 제공</li> <li>3. 뇌-마음 은유 제공; 1980년대의 신경망(연결주의) 접근의 기초 제공</li> <li>4. 여러 인지 이론의 신경적 타당성 근거 검증 및 제공 ; 의식 연구의 부활</li> </ol>
인지심리학	<ol style="list-style-type: none"> <li>1. 인지주의의 핵심적 경험과학으로서의 기초 역할</li> <li>2. 인지과정에 대한 경험적 자료, 개념, 이론 제공; 주제 제공</li> <li>3. 인공지능과의 밀접한 연결로 인공지능 연구의 이론적, 개념적 배경 제공</li> <li>4. 신경과학에 어떤 심적과정을 어떻게 개념/이론화하여 접근할 것인가의 개념, 이론, 자료 제공</li> <li>5. 철학, 신경과학, 인공지능 등 및 진화생물학, 응용공학 등 학문간 융합적 연결 고리 제공</li> </ol>

## 10. 20세기의 인지과학의 확산과 재구성 시도: 신경망, 신경과학

### 10.1. 고전적 인지주의(Classical Cognitivism) 접근

방법론적 객관성을 강조한 행동주의 심리학이 심리학에서 마음, 인지, 심적 능력 등의 개념을 배제하고 설명력이 극히 제한된 '자극-반응' 연결의 객관적 기술에 그치는 접근을 전개하는 데에 대한 반발에서 1950년대 후반에 탄생한 고전적 인지주의(Classical Cognitivism)는 심리학에 '마음', '인지' 개념을 되살려 놓았다. 사이먼(H. A. Simon) 등의 인지과학 창시자들에 의하여 1950년대 후반에 등장하여 1970년대부터 인지과학에서 그 입지를 확고히 한 고전적 인지주의는 인간 마음과 컴퓨터를 유사한 정보처리적 원리를 지닌 시스템으로 보고, 이 정보처리시스템의 정보처리 과정과 구조를 밝힘을 통하여 마음의 문제를 접근할 수 있다고 보았다.

이러한 틀은 고전적 인지주의의 이론적 기반을 제공하였던 철학의 기능주의의 이론적 배경에 기초하고 있었기에, 기능주의 철학에서 강조하는 다중구현원리(Multiple realizability)<sup>38)</sup>의 철학적 관점을 따라서 그 이론 틀을 구축하였었다. 즉, 추상적 정보처리 원리의 중요성을 강조한 나머지 그 정보처리 원리가 구현되는 물리적 실체는 중요하지 않다고 보아 정보처리체계가 구현되는 실체가 넘어진, 전자 칩이건 별로 중요하지 않다는 입장을 전개하였다. 그리고 마음의 작동 원리를 논리적 형태의 규칙의 도출과 같은 형식적 접근에 의하여 계열적인 과정으로 기술할 수 있다고 보았다. 모든 심적 현상은 상징(기호적) 표상(representation)을 형성, 저장, 활용하는 계산(computation)적 정보처리라는 기본 입장이 그 핵심을 이루었다.

### 10.2. 연결주의(Connectionism)

인지주의가 출발한지 20년 후에 마음이 실제로 구현되기 위하여 전제되어야 하는 신경적 기초를 무시하는 초기의 고전적 인지주의가 드러내는 한계점을 극복하기 위하여 1980년대 중반에 대두된 연결주의(connectionism)는 고전적 인지주의와는 다른

---

38) 다중구현원리에 대하여는 <http://plato.stanford.edu/entries/multiple-realizability/>의 설명이나 [http://en.wikipedia.org/wiki/Multiple\\_realizability](http://en.wikipedia.org/wiki/Multiple_realizability)의 설명에서 보는바와 같이 컴퓨터라는 기계와 마음 작용을 동류의 정보처리 과정을 지니는 시스템으로 유추할 수 있게하는 (1950년대와 1960년 대의 학계으 경향) 배경이 된 철학적 기능주의 이론 틀이었다. 이 원리에 대한 논의는 작고한 서울대 철학과 김영정 교수의 1996년의 책 [심리철학과 인지과학]과 이 논의의 대표적 세계적 철학자인 김재권 (Jaekwon Kim) 교수의 영문 문헌과 국내 서적들을(예: [수반과 마음]) 참고하기 바란다. 이 다중구현 원리는 초기에는 (뇌를 거의 배제하고) 컴퓨터와 마음의 작용을 유추하는 배경 근거이론으로 거론되었지만 ('강한 인공지능' 관점이지 이론들) 최근에는 이와 반대되는 입장으로도 이용되고 있다. 김재권 교수에 대한 소개는 다음에 있다: ([http://article.joinsmsn.com/news/article/article.asp?Total\\_Id=3061770](http://article.joinsmsn.com/news/article/article.asp?Total_Id=3061770)).

관점을 전개하였다.

연결주의는<sup>39)</sup> 마음의 작동이 그 신경적 기반 구조인 뇌의 특성에 의하여 결정된다고 보고 뇌의 기본 단위인 세포들 간의 연결강도의 조정 중심으로 마음의 작동 특성을 이론적으로 개념화하였다. 연결주의는 고전적 인지주의처럼 미리 내장된 알고리즘적 규칙이나 지식표상을 전제하지 않으며, 신경망적 분산적, 병렬적 확률적 계산에 의한 상징(기호)이하 수준에서의 정보처리 메커니즘을 강조하였다. 그런데 비록 1980년대에 기존의 고전적 인지주의를 대체할 것처럼 보였던 이 연결주의는 실제의 뇌 특성 중심으로 이론적 모델을 전개하였기 보다는 추상화된 뇌, 이상화된 이론적 뇌의 신경망적 특성을 중심으로 마음의 작동 메커니즘을 모델링하였다는 한계점을 지니고 있었다.

### 10.3. 뇌과학과 인지신경과학적 접근

연결주의적 접근의 확산 다음에 인지적 패러다임 틀의 중심 위치를 점유한 것은 인지신경과학적 접근이었다. 인지신경과학적 접근은 앞에서 언급한 여러 보는틀이 심리적 기능과 과정의 측면에 대한 접근들임과는 달리, 마음을 실체의 뇌의 신경생물적 구조와 과정에 기초하게 하는 존재론적 차원의 접근이기에 앞의 접근들과는 다른 차원의 접근이라고도 할 수 있고 별도의 접근으로 범주화하는 것 자체가 부적절할 수 있다.

하여간 이 인지신경과학적 접근은 1980년대의 연결주의가 지니고 있는 제약점, 즉 실제의 뇌보다는 이상화된 이론적 뇌를 상정하여 놓고 이를 모델링한다는 측면과, 연결주의가 마음의 고차 수준이 아닌 상징이하(subsymbolic) 수준에서의 처리를 이야기하고는 있지만, 기본적으로 형식적 계산주의의 입장을 벗어나지 못하고 있다는 측면의 결함을 보완해주는 접근이다. 또한 연결주의 이전의 기존 모든 접근의 2원론적 존재론 가정의 미흡함을 마음의 작동 원리를 생물적, 신경적 바탕에 기초하도록 하는 일원론적 환원주의적 접근에 의하여 보완하려는 접근으로써 인지과학의 대세적 접근으로써의 위치를 현재 차지하고 있다.

이전에는 전통적 심리학 및 인지과학 영역의 연구에서는 심적 과정이나 표상체계에 대한 개념, 이론, 가설적 예언 등은 실험실 내에서의 인지과정 실험에 의해 그 타당성을 검증 받고 세련화되었다. 그러나 인지신경과학적 접근이 도입되고 확산되면서 하위 심적구조와 단계적 과정을 제시한 심리이론들 특히, 주의, 지각, 기억, 언어, 사고 등에 관한 인지이론들, 개념들은 그 이론적 구성개념의 타당성과 예언의 타당성이 신경생리학적, 신경생물학적 기반(neural correlates)에 의해 검증되고, 재구성되고 있다. 마음의 과정과 구조에 대한 어떤 아이디어가 있으면 이제는 최종 확인과 검증을 인지신경과학적 실험을 통해 확인하는 절차가 추가되거나 그것으로 보완되고 있는 경향이

---

39) 연결주의 개요 설명: <http://plato.stanford.edu/entries/connectionism/>  
<http://en.wikipedia.org/wiki/Connectionism>

다.

#### 10.4. 대안적 틀의 모색

그런데 고전적 인지주의나 연결주의나 뇌과학의 인지신경과학적 접근이, 마음의 다양한 측면을 충분히 설명하지 못하는 제한적

접근이라는 반론이 1980년대 중반 이래로 꾸준히 제기되어 왔다. 정보처리적 소프트웨어 시스템으로서의 인지체계를 강조하며 마음을 계산주의와 표상주의, 기호(상징)체계의 틀에서 접근한 고전적 인지주의나, 추상화된 이론적 뇌의 세포 수준의 하위 단위들의 활성화와 연결 원리의 모델링을 중심으로 상징이하 체계를 강조하며 접근한 연결주의(신경망적) 접근이나, 실제의 구체적인 뇌의 신경적 구조와 과정적 활동의 측면을 강조하며 환원주의적 입장을 취한 인지신경과학적 접근들이, 인지 현상의 많은 특성을 밝혀주기는 하였지만 실상은 마음의 본질적 측면을 충분히 다루지 못하고 있으며(Hollnagel, 2007)<sup>40)</sup> 다른 차원의 접근이 필요하다는 논의가 계속 제기되었다. 이는 주로 몸의 문제를 중심으로 탈데카르트적 관점을 제시하는 새로운 접근이었다.

1950년대에서 1980년대까지의 ‘고전적 인지주의’는 데카르트의 심신이원론에서 탈피하지 못하고 생물적 측면의 중요성을 격하시켜 뇌의 탐구를 소홀히 하였다. 한편 1980년대 초에 등장한 연결주의나, 그 이후에 등장하여 현대 인지과학의 성과를 이끌어가고 있는 인지신경과학은 물질로의 환원주의적, 일원론적인 입장에서 뇌를 강조하기는 하였지만, 근본적으로는 현상을 경험하는 주체와 그 대상인 객체를 이분법적으로 보는 데카르트의 존재론을 벗어나지 못하고 있었다고 비판받을 수 있다. 즉 마음의 작용을 곧 뇌의 신경적 활동 상태로 환원하는 단순한 일원론적 시도 이외에는 마음과 몸의 관계의 본질적 측면에 대한 어떤 시사를 주지 못하였다고 볼 수 있다.

현재 진행되고 있는 인지과학적, 신경과학적 연구들은 ‘환경’과 독립된 별개의 실체로서 작용하는 뇌와, 뇌의 신경적 상태로 개념화된 마음의 개념적 틀의 타당성, 충분성에 대하여 그 개념적 기초를 엄밀히 체계적 분석을 하지 않은 채 진행되어 온 것이라고 비판 받을 수 있다. 고전적 인지주의가 뇌의 중요성을 경시한 채, 추상적인 표상 체계로 마음을 개념화하였던 데카르트 식 접근이었음에 대하여, 인지신경과학적 접근은 뇌라는 물리적 구체성만 되찾아 준 것일 뿐, 마음의 바탕이 되는 몸 기반 전체를 되찾아 주거나 환경과의 역동적 상호작용적 연결 본질의 이론적 의의를 살린 것이 아니었다 라는 비판이다.

---

40) "Although the widespread acceptance of the analogy between the brain and the digital computer overcame the constraining influence of behaviorism, the computer analogy itself soon become as constraining as behaviorism. ... The study of cognition became the study of 'cognition in the mind'.(351쪽)... 그러나 그것이 아니라 "... cognition emerged as a result of the interaction between the brain, the body and the world. The need to understand cognition as more than a mental process and to study it under more natural conditions led to a development that in the 1980s and 1990s culminated in the formulation of the principles for 'cognition in the wild'...(352쪽)"

이러한 비판적 관점을 가진 여러 분야에서의 생각들이 연결되고 수렴되어 지난 몇 년 사이에 인지과학에서 하나의 영향력 있는 대안적 틀로 떠오른 것이, 마음에 몸의 바탕을 연결하여 주는 틀인 ‘연장(확장)된 마음(Extended Mind)’, ‘체화된 마음(Embodied mind)’, ‘환경에 들어있는 마음(Embedded Mind)’ 또는 ‘체화된 인지(Embodied Cognition)’<sup>41)</sup>라고 불리는 접근이다.

이 틀은 심리학과 인지과학의 이론적 틀을 현재 변화시키고 있고, 앞으로도 계속 그 영향이 증대하리라 본다. 과거 1950년대의 인지주의의 출발과 떠오름 시점보다도 더 드라마틱한 전기를 현재 심리학과 인지주의가 아마도 맞고 있다는 뎀 등의 (Bem & Keijzer, 1996) 논의가 타당할 수 있다. 이러한 움직임은 과거에 철학과 심리학, 인지과학을 지배하여 온 전통적인 데카르트적 존재론과 인식론에 바탕을 둔 ‘마음(mind)’의 개념으로부터 탈피하여, 구체적인 ‘몸’이라는 실체를 통한 ‘환경’과의 상호작용 속에서 출현하는 인간의 적응 ‘행위’로서의 ‘마음’의 관점으로 전환하는 탈데카르트적 관점의 움직임의 추세라고 볼 수 있다.

## 11. 인지과학 재구성: 체화된 인지 접근

제3의 인지과학 패러다임이라고<sup>42)</sup> 일컬어지는 이 틀은 인간의 마음이나 인지가, 개인 내의 뇌 속에 추상적 언어적 명제 형태로 표상된 내용과, 그를 기초로 한 고전적 인지주의의 정보처리적 계산적 과정이라고 하기보다는, 구체적인 몸을 가지고(embodied) 환경에 구현, 내재되어(embedded) 사회환경에 적응하는 유기체(organism)가 환경(environments)과의 순간 순간적 상호작용(interaction) 행위 역동(dynamics) 상에서 비로소 존재하게 되는, 즉 유기체의 몸과, 문화, 역사, 사회의 맥락에 의해 구성되고 결정되는 그러한 역동적 활동으로서의 마음임을 강조하는 접근이다.

이러한 접근은 아직은 통일적 틀을 이루지 못하고 다소 산만하게 여러 이름들로 전개되고 있지만, 고전적 인지주의에서 배제되었던 ‘몸’을 마음의 바탕으로 되찾게 하며(embodied mind), 체화된 마음과 분리될 수 없는 ‘환경’을 인지과학과 심리학에 되살려 놓게 하며, 공간적 연장(extension)이 없는 ‘심리적(정신적) 실체’라는 마음이 아니라 ‘몸을 통해’환경에 연장된, 확장된 마음(extended mind)으로 마음 개념을 재개념화할 가능성을, 아니 그래야 하는 필연성(Bickhard, 2008)을 제시하고 있는 것이다.

41) 철학에서 주로 ‘연장(확장)된 마음(Extended Mind)’으로, 인지과학 등에서 ‘Embodied Mind’, ‘Embodied Cognition’, ‘Embedded Mind’, ‘Embedded Cognition’ 등으로 언급되고, 철학에서 ‘embodied’와 ‘embedded’를 조합하여 ‘Amalgamated Cognition’이라고 지칭되기도 하는 이 접근은, 이러한 명칭 하나하나가 엄밀히 규정하자면 서로 다른 의미를 지닐 수 있지만, 여기에서는 편의상 맥락에 따라 ‘체화적 인지’, ‘체화된 인지’, 또는 ‘체화적 마음’으로 혼용하여 표현하겠다.

42) 또는 제6(또는 7)의 심리학적 패러다임이라고도 지칭되는

이 움직임은 종래의 일반인들이나 과학자들이 갖고 있던 마음과 몸에 대한 데카르트 식의 이원론적 생각을 벗어나려는 것이다. 즉 심신이원론이나, ‘마음은 곧 뇌의 신경과정이다’ 라는 환원주의적 일원론<sup>43)</sup>을 벗어나려는 새로운 보는틀이다. ‘체화된 인지’의 보는틀은 고전적 인지주의의 정보처리 접근이 지니는 제한점을 벗어나려 한다. 즉 환경과는 독립적으로 한 개인 뇌 내부에서 일어나는 정보의 인지적 표상이나 처리가 아니라, 몸으로 환경 속에 구체화되며, 몸의 활동을 통하여 환경과 연결되어 ‘상호작용’하며 살아가는 인간의 ‘행위’로서 마음을 설명하고자 하며(Hollnagel, 2007)<sup>44)</sup>, 그리고 환경 내의 다른 인간의 마음이나 각종 인공물에 분산표상된 마음, 그리고 문화적, 사회적, 역사적으로 상황 지워지며 행위로 구성되는 마음으로서 보려는 것이다. 구체적 몸으로 환경에 체화된다는 것(embodied 또는 embedded)은 표상, 인지, 마음을 거론함에 있어서 보조적 개념화가 아니라 필수적 개념화라고 볼 수 있다. 이러한 대안적 접근은 앞으로 인지과학적, 심리학적 탐구에, 그리고 자연히 인접 학문에 상당한 변화를 가져오는 시사를 지닌다고 볼 수 있다.

### 11.1. 체화된 마음 관점의 떠오름의 배경

체화된 인지 접근의 요점은 본질적으로 데카르트의 이원론적 존재론과 그에 기반한 인식론을 벗어나자는 탈 데카르트적 움직임의 일환이라고 볼 수 있으며, 이러한 시도는 이미 일찍이 17세기의 B. Spinoza에 의하여 이루어졌었다(다마지오, 2007). 몸에 대한 강조는 이후에 유럽의 현상학적 철학자들(하이데거(1998), 메를로 폰티(2002) 등)에 의하여 이어져 왔다고 할 수 있다. 추상화된 마음의 측면이 강조되는 데카르트의 이원론적 존재론을 넘어서서 몸과 마음을 둘로 나눌 수 없다는 입장에서 전개된 프랑스의 메를로 폰티 등의 관점에서는 인간의 의식적 경험의 뿌리가 몸에 있음과, 몸과 마음과 환경이 하나의 단위를 이루는 입장이 제시된다. 몸이 환경(세상)과 일체가 되어 적응하는 과정에서 몸의 행위 하나하나가 마음을 구성한다고 보는 것이다.

(급진적) 체화적 마음 접근의 연원을 미국 내에서는 19세기 말의 철학자 듀이(J. Dewey)나 철학자이면서 심리학자이었던 제임스(W. James)에서 찾을 수도 있다(Chemero, 2009). 하지만, 심리학 내에서 이러한 접근의 현대적 심리학이론 제기는 깁슨(J. J. Gibson, 1979) 등의 생태학적 입장 중심으로 전개되었다고 할 수 있다. 20세기 중엽부터 감각과 지각의 심리현상을 중심으로 깁슨 등의 생태심리학적 연구는 이미 이러한 관점을 설득력 있게 펼쳤다. 이외에도 심리학을 넘어서 인지과학 내에서 다음과 같은 주변학문 사조들의 이론적 영향들이 수렴되어 “제3의 인지 혁명”이라고도 지칭되는 이러한 움직임을 이루어 내었다고 할 수 있다.

43) Chemero(2009)는 이를 ‘무차비한 환원주의(ruthless reductionism)’이라고 까지 지칭하기도 한다.

44) "Yet cognition can also be seen as an aspect of what people do, rather than as something that goes on in their individual or collective minds(352 쪽)... If we accept that cognition is an aspect of what people do, the focus should be on problems that are representative of human performance. (353쪽)

인지과학 내에서 재현된 하이데거적 존재론-인식론 논의(Dreyfus(1991) 등의 논의), 언어학의 써얼(J. Searle) 등의 언어행위 논의, 상황의미론 논의, 인지언어학자 레이코프(G. Lakoff) 등의 은유와 마음, 체험적 실재론 논의(Lakoff, & Johnson, 1999), 마투라나(Maturana와 Varela, 1983)를 중심으로 한 인지생물학(biology of cognition) 및 체화된 마음(embodied mind) 논의(Varela, Thompson, & Rosch, 1991; Thompson, 2007), 인공지능학과 로봇 연구에서의 브룩스(R. Brooks, 1991) 등의 반사적 로봇(reactive robotics) 논의 및 최근의 인지로봇 연구, 심리학에서의 비표상체계 논의, 인지인류학에서의 지식의 사회 문화적 제약 이론 및 인공물과 외적 분산 표상 개념에 관한 논의(예: Hutchins, 1995; Roy D'Andrade, 1995)), 심리학 등에서의 분산적 인지, 상황적 인지(situated cognition) 논의, 행위로서의 마음(mind as acts)에 대한 논의, 담화적 마음(discursive mind) 논의, 동역학체계 논의, 심리학과 로보틱스 등에서의 인지의 감각운동 기반에 대한 진화론적 논의, 상호작용론(interactivism) 논의 등이 이러한 제 2의 패러다임적 변혁의 배경이 되었다고 할 수 있다(이정모, 2001, 2007, 2009 ㄱ).

철학에서 체화된 인지의 입장을 전개하는 것으로는 먼저 연장(확장)된 마음(Extended Mind)과 관련한 클라크 등(Clark과 Chalmers, 1998; Clark, 1997, 2008)의 논의 등이 있으며, 인간과 컴퓨터의 상호작용을 체화된 상호작용의 관점에서 본 두리쉬(Dourish, 2001), 마음은 뇌 안에 있거나 개인 안에 있는 것이 아니라 뇌를 넘어서, 개인을 넘어서 있다는 윌슨(Wilson, 2002), 마음은 뇌 자체도, 기계속의 도깨비도 아니다 라는 주제로 강하게 ‘마음=뇌’ 관점을 비판한 락웰(T. Rockwell, 2005), 뇌 속의 마음이 아니라 몸과 괴리되지 않으며 세상과 괴리되지 않은 마음으로, 인지로 재개념화하여야 한다는 윌러(M. Wheeler, 2005), 몸 이미지가 아닌 몸 스키마의 개념을 사용하여 ‘몸이 마음을 어떻게 조형하는가’ 하는 주제를 다룬 갤러거(S. Gallagher, 2007), 지각도 사고도 감각-운동적 신체적 행위에 바탕하고 있다는 철학자 A. Noe(알바 노에, 2009) 등의 주장 등이 있고, 정대현(2001), 이영의(2008), 윤보석(2009; 4장 1절), 이정모(2010) 등의 국내 학자들의 마음 논의에서도 신체성 내지는 체화된 인지(마음) 개념의 논의가 전개되었다.

20세기 초에 행동주의심리학이 마음을 심리학에서 배제하였고, 1960년대 이후의 고전적 인지주의가 ‘마음’을 심리학에 되찾아주었지만 뇌의 역할을 무시하였고, 1980년대 이후의 인지신경과학이 ‘마음’을 다시 ‘뇌’ 속으로 넣어주었지만 뇌를 제외한 ‘몸’과 ‘환경’의 역할을 무시하고 데카르트 식 “마음 = 뇌신경과정 상태”의 환원주의적 일원론의 관점을 전개하였다면, 이제 21세기에서 제 삼의 대안적 관점을 통하여 그 뇌를 몸으로, 그리고 다시 그 몸을 환경으로 통합시키는 작업을 하여야 한다고 볼 수 있다.



‘마음’은 ‘뇌’ 속에서 일어나는 신경적 상태나 과정이라고 하기보다는 신경적 기능구조인 뇌와, 뇌 이외의 몸, 그리고 환경의 삼자가 괴리되지 않은 통일체적 총합체(nexus) 상에서 이루어지는 역동적 행위 중심으로 재개념화되어야 한다는 것이다. 몸을 배제한, 체화되지 않은 상호작용의 개념으로는 인간과 인간간의 상호작용을 포함한 인간과 환경 간의 상호작용을 설명할 수 없다(Seifert, 2008).

이외에도 최근에 다음과 같은 학자들의 논의에 의해 이 접근이 틀을 갖추어 전개되고 있다: Bickhard(2008), Calvo & Gomila(2008), Chemero(2009), Menary(2010), Rowlands(2010), Wallace, Ross, Davies, & Anderson(2007).

### 11.2. 체화된 마음 접근의 요점

그러면 인지과학의 제3의 대안이라는 ‘체화된 마음’ 접근의 핵심적 주장은 무엇인가? 전통적 인지주의(cognitivism)에 대한 대안 관점으로(post-cognitivism) 제시된 체화된 마음(또는 인지) 접근의 중심 주장을 다음과 같이 다시 정리할 수 있을 것이다.

체화된 마음의 보는틀은 “미시적, 신경적 또는 생물적 단위 수준에서 모든 것을 설명하려는 ‘연결주의’와 같은 낮은 설명 수준의 접근이 지니는 한계, 그리고 그보다 한 수준 위에서 명제 중심으로 논리적 체계에 의해 설명하려는 ‘고전적 인지주의’의 정보 처리 접근이 지니는 제한점을 벗어나려 한다. 즉 환경과는 독립적으로 한 개인 마음 내부에서 일어나는 정보의 인지적 표상이나 처리가 아니라, 환경 속에서 살며, 이와 상호작용하며 살아가는 인간의 행위로서 인지를 설명하고자 하며, 환경이 인간의 인지의 특성, 한계를 규정, 제약하고 인간의 인지구조가 환경을 규정하고 변화시키는 그러한 상호작용의 관계 속에서의 인지를 연구하고자 한다.<sup>45)</sup> 세상 속에서 적응하며 활동하는 존재이며 세상의 일부로서의 한 개인이 사회적 상호작용 속에서, 그리고 물리적 환경의 자연물과 인공물과의 상호작용에서 이루어지는 담화에 의해 구성되고 의미를 지니는, 그리고 구체적인 신체에 구현된(embodied) 실체로서의 인간 마음, 그리고 환경의 다른 인간의 마음이나 각종 인공물에 분산표상된 마음, 문화적, 사회적, 역사적으로 상황 지워지며 행위로 구성되는 마음으로서, 인지로서 보려는 것이다(이정모, 2001).”<sup>46)</sup> 이러한 ‘체화된 인지’ 접근의 요점은 밴다이크 등(van Dijk, Kerkhofs, van Rooij와 Hasellager, 2008)과 빅카드(Bickhard, 2008) 등의 글에서도 ‘잘 표현되어 있다.<sup>47)</sup>

45) "... behaviour in real world conditions clashes with the assumptions of 'cognitions in the mind'. ... it is necessary not only to focus on 'cognition in the world' but to extend the meaning of cognition far beyond the notions of human information processing... . Cognition therefore cannot be isolated in the mind of a thoughtful individual, but involves people and artefacts distributed in space of time, and ... it is more important to understand what cognition does, than to understand what it is. (Hollnagel, 2007, p. 354)"

46) 이정모(2001)의 643-644쪽; 일부 내용 삭제 및 보완함.

47) "..... we question the very assumption that cognizing is something that the brain does... but we

더 나아가서 칼보 등(Calvo와 Gomila, 2008, 12-13 쪽)에 의하면 체화적 마음 접근의 요체는 체화됨(embodiment)보다는 상호작용주의(interactivism)와 동역학주의(dynamicism)이며, 이 관점은 다음과 같은 핵심 주장을 전개한다.

환경과의 심적 역동적 상호작용은 몸에 의존하며, 따라서 감각운동적 측면이 인지의, 마음의 핵심이 되며, 고차 심적 기능도 이러한 기초의 제약과 허용 틀에서 이해되어야 하며, 지각은 능동적이며, 행위는 지각에 의해 인도되며, 신경계, 몸, 환경 요인이 실시간 상에서 상호작용하는 것을 이해함을 통하여 과학적 설명이 주어지고, 전반적 계획이나 통제가 없이 분산된 단위들의 지엽적 상호작용에 의하여 자가조직적(autopoeitic)으로, 창발적으로 출현하는 것이 심적 현상이며, 마음은 환경에 확장된(extended), 상황지어진(situated) 것으로 분석, 이해되어야 하며, 자연적, 생태적 상황에서 맥락이 고려되어서 이해되어야 하며, 전통적 논리적 형식적 접근보다는 역동적 시간 경과와 상호작용성을 다루기에 적절한 접근을 통하여 탐구되어야 하며, 신경생물학적 가능성(plausibility)이 반드시 고려되며, 현상 과정이 어떻게 (주관적으로) 체험되는가 하는 가에 대한 현상학적 접근도 설명적 구성요소로 반드시 들어가야 한다.

즉, [1] 뇌를 포함하는 몸과, [2] 환경(각종 물리적, 사회적, 심리적)과, [3] 그리고 이들이 연결되는 상호작용적 활동(interactivity)의 세 측면이 서로 괴리되지 않고, 표상이 없이 하나의 역동적 전체로서 개념화되는 그러한 접근을 하여야 하는 것이다

### 11.3. 급진적 체화된 인지 접근

체화된 인지를 주장하는 입장은 그 접근의 강력함의 정도에 따라서 일반적으로 다음의 두 개의 접근으로 나누어 볼 수 있다. 하나는 ‘약한(온건한) 체화된 인지’ 입장이고 다른 하나는 강한(급진적) ‘체화된 인지’ 입장이다(Chemero, 2009).

‘약한(온건한) 체화된 인지’ 입장은 고전적 인지주의나 인지신경과학적 접근의 부족함을 보완하여 표상을 형성하고 저장하고 이를 활용하는 계산(computation) 과정으로 인지(마음)의 본질을 보는 입장이다. 즉 기존의 고전적 인지주의 관점을 기본 틀로 인

---

believe there is more to cognition than the processes inside the brain alone. Besides the brain, the organism's body and the world form part of the physical substrate that underlies cognition and behavior. ... cognition and behavior emerge from the bodily interaction of an organism with its environment. ... the physical structure of body and world, and the internal milieu of the organism's body, all provide important constraints that govern behavioral interactions. From this perspective, cognitive states are best explained by a physical system of interacting components, where the brain is only one such component. In other words, the brain is best viewed not as a commander or director of behavior, but rather as only one of the players among equally important others ... higher cognitive functions cannot directly mapped onto brain structures (van Dijk, Kerhofs, van Rooij, & Haselager, 2008, p. 298; 밑줄은 필자가 추가함)."

"If representation, thus cognition, is derived from (inter)action, then some sort of embodiment is required in order for such action and interaction to be possible: actions, thus embodiment, are not mere auxiliaries to representation, but, instead, are essential to it (Bickhard, 2008; p. 35)

정하고 단지 그 표상이 어디에서 오는가 하는 표상의 원천(source)을, 추상적인 인지주의에서 강조하는 형식(formal) 명제적 상징(기호)이 아니라, 감각운동(sensory-motor)적으로 기반함에, 즉 몸의 활동 정보에 근거함을 강조하려는 입장이다. 칼보 등(Calvo와 Gomila, 2008; Clark, 1997, 2008; Clark와 Chalmers, 1998, Thompson, 2007)이 이러한 입장을 취하는 것으로 분류될 수 있을 것이다.

‘강한(급진적) 체화된 인지(Radical Embodied Cognition)’ 입장은 그것이 아니라 기존의 고전적 인지주의적 접근의 전반적 문제점 특히 기존의 고전적 인지주의가 특정 시점의 정적인(static) 표상, 그리고 추상화된 표상, 통사적 언어적 명제 표상, 그리고 이러한 의미의 고전주의적 ‘계산(computation)’을 강조하는 문제점을 지적한다. 기존의 고전적 인지주의 틀로서는 실제로 일어나는 인간 활동의 역동적인 측면에 기초한 인간 마음(인지)의 본질을 살릴 수 없고, 이론적으로 접근할 수 없으며, 그보다는 한 시점에서의 표상이 아니라, 이어지는 연속된 시점 상에서 역동적으로 변화하는 궤적을 그리는 역동적 활동으로서의 인지로 보아야 하며, 뇌, 마음, 환경이 불가분적으로 결합되어 하나의 통합적 단위의 행위적 거시체계를 이루는 것으로 보아야 한다는 것이다. 즉 마음(mind)을, 인지를 역동적 동역학체계(dynamic systems)로 보아야 한다는 것이다.

고전적 인지주의에서처럼 뇌가 환경을 표상하기만 하는 것이 맞다면, 환경과 뇌를 분리하여 생각할 수 있고, 외적 행위는 뇌 속에 ‘표상된 환경’(environment-as-represented)이 이끌어낸다고 볼 수 있고, 그러기에 환경을 인지 체계의 한 요소로 보는 것은 잘못이며 오직 내적 표상만 있을 뿐이라고 할 수 있다. 그러나 인지의 주체인 생명체와 환경이 비선형적으로 결합되어(coupled)되어 있는 동역학체계를 이룬다면, 인지와 환경은 별개의 괴리된 요소로 분리될 수 없다는 것이다. 동물은 정보가 풍부한 환경에서 능동적으로 지각하고 행동하는 ‘행위주체자’이며, 동물과 환경은 분해불가능한 단위체로 체계를 이루고 있으며, 동물은 환경의 객관적 대상 그 자체를 지각, 인지하는 것이라고 하기보다는 환경에서 제공되는 ‘행위 가능성(affordances)을 지각, 인지하면 그에 대하여 행위를 발하는 것이라는 입장이다.

Chemero(2009) 등에 의하면 ‘체화된 인지’ 접근 중에서 ‘표상’과 ‘계산’ 관점을 유지하는 온건한 접근은 잘못되었다. 체화된 인지 접근은 본질적으로 동역학체계 이론을 포함하는 틀에서, 심성표상을 전제하지 않고 인지를 설명하여야 한다. 인지란 본질적으로, 함께 하나의 역동적 체계를 이루고 있는 ‘환경’ 속에서의 그 환경이 제공하는 ‘행위 가능성(affordances)’에 의해 촉발되는 활동(actions in the environment)을 통해서 이해될 수 있는 것이지, ‘내적 표상’이나, Chemero가 ‘심적 체조’라고 부르는<sup>48)</sup> 고전적 인지주의 식의 ‘계산’을 통해 이해될 수 있는 것이 아니라는 것이다. ‘급진적 체화된 인지’ 입장은 인지과학의 목적이 비선형적으로 연합되어 있는 ‘동물-환경 시스템

48) ‘급진적 체화된 인지과학은 심적 표상을 전제하지 않는 설명적 도구를 사용하여 지각, 인지, 행위에 대한 과학적 탐구를 하는 학문이다. 심적 체조(gymnastics)가 없는 인지과학이다. (Chemero, 2009; 29 쪽)’

템'을 연구하는 것이라고 본다. 따라서 행위체인 동물의 부분인 뇌만 따로 떼어서 설명하거나, 또는 그에 의해 조작되는 표상만으로 설명하여서는 과학적 설명의 충분성이 없다는 것이다.

#### 11.4. 체화된 마음과 기계(인공물)와 인간

마음을 단순히 뇌 속에서 일어나는 신경적 과정의 결과로써, 그리고 알고리즘적 또는 확률적인 정적 계산적 정보처리로써 개념화하지 않고, 몸과 괴리되지 않은 마음이 몸을 통하여 환경에 공간적 확장, 연장의 특성을 지닌 역동적인 활동(행위)에 존재하는 것으로 개념화 한다면, 위에서 논하였듯이 체화된 자체보다는 '상호작용성', 그리고 그 '역동성'이 더 핵심적인 개념이어야 함은 당연한 것 같다.

체화된(embodied)이라는 것은 인간과 환경의 상호작용 행위가 출현할 수 있는 상황 조건, 가능성 조성 기반에 지나지 않으며, 심리학이나 인지과학이 관심 가져야 할 것은 그러한 상황지위점을 통하여 마음-몸-환경의 합일체가 무엇을 이루어 내는가 하는 '역동적 활동'이어야 하는 것이다. 그것이 바로 마음-몸-환경의, 상호 괴리될 수도 없는 실체들이 함께 이루어내는 상호작용, 아니 '함께 존재하는 방식'일 것이다. 환경과 인간의 상호작용이 심리학, 인지과학의 핵심 개념으로 떠오르게 되는 논리적 기반이 여기서 제공된다.

몸에 구현된 마음이 환경과 상호작용함의 특성을 규명하고자 함에 있어서, 인간과 (다분히 수동적 물질인) 자연물과의 상호작용은 김슨(J. J. Gibson) 류의 생태심리학이 다룰 수 있지만, [인간]과 [의도를 지닌 행위 주체로서의 다른 인간] 또는 [행위 주체로서의 다른 동물과 인공물(로봇, 인공시스템)]과의 상호작용은 김슨 류의 기존의 생태심리학이 충분한 설명적 틀을 제공하여 주지 못한다.

[인간-〈일반 물질적 대상〉]의 상호작용은 생태심리학적 접근이, [인간-인간]의 상호작용은 인지심리학, 인지신경심리학, 사회심리학과 인지인류학 등의 인지과학이, [인간-〈동물〉]의 상호작용은 동물심리학, 인지심리학, 사회심리학, 인지인류학, 동물행동학, 인지신경과학이 다룬다면, [인간-〈행위주체자의 역할을 하는 인공생물시스템 또는 인공인지시스템으로서의 로봇이나 다른 인공물(Artifacts 2)〉]과의 상호작용의 본질에 대한 개념화의 이론적 정리, 재구성이 [인간-〈수동적 대상인(도구나, 핸드폰과 같은 기기 등) 일반 인공물 일반(Artifacts 1)〉]과의 상호작용과는 다른 관점에서 필요하게 된다(이정모, 2009ㄱ).

#### 11.5. 인간과-인공물일반(A-1)의 상호작용

'행위주체자(agents)로서의 인공물'에 대한 논의에 앞서서 행위주체자가 아닌 각종

인공물 일반(artifacts in general: 언어, 문화, 제도와 같은 소프트인공물, 그리고 칼 같은 하드인공물, 컴퓨터나 스마트폰과 같은 혼합인공물 포함)과 마음의 상호작용 관계를 진화 역사의 측면에서 되생각하여 본다면 다음과 같이 논할 수 있을 것이다.

인류 진화의 역사를 본다면 인간이라는 종의 진화는 인간의 신체적 진화, 마음의 진화의 역사라고 하기 보다는 인간이 만들어낸 인공물과 인간의 마음, 그리고 몸이 공진화해 온 역사라고 볼 수 있다. 인간이 인공물을 만들고 활용한다는 일방향적 활동에 의하여 인간의 진화가 이루어졌기보다는, 인공물이 인간의 신체적, 심리적 활동을 확장시키고 또 제약하기도 하는 쌍방향적 상호작용 과정으로 진행되어 왔다고 볼 수 있다. 인간의 마음속의 어떤 내적 표상 구조, 특히 외부 세계와 자신의 문제 상황간의 관계에 대한 가설적 구성개념들이 외현화되어 물리적 환경에 구현되어 인공물(도구)이 되었을 것이다. 그리고 이러한 외현화 및 구현 과정 속에서 인간의 뇌와 마음은 끊임없이 외부 환경의 구조와 역동적 변화와 상호작용하며 그 환경의 인공물과 함께 공진화하여 왔을 것이다(이정모, 2007, Seifert, 2008). 따라서 인간 마음의 본질을 탐구함에 있어서 몸을 통한 환경의 인공물 세계와의 상호작용(interactivity)의 측면을 제쳐놓고 생각하기 힘들다.

그런데 과거의 1990년대 전반까지의 ‘인간-인공물’ 상호작용의 연구는(J. Gibson 류의 생태심리학적 접근과 D. Norman의 전통적 인지공학도 포함하여) 전통적인 데카르트적 인식론에 기초한 이론들에서 벗어나지 못했었다. 마음은 환경과는 독립적으로 존재하고, 독자적인 표상을 지닌다는 데카르트적 입장에 바탕을 두고서, 내적으로 뇌 속에 표상화된 개별 지식의 전달과 이를 표상으로 수용하는 것으로, 인지적 활동뿐만 아니라 인간-인공물 상호작용을 개념화했던 과거의 전통적 관점은 역동적인 인간-인공물 상호작용, 특히 연속적 시간 궤적 상에서의 역동적 심리적 현상을 설명하기에는 부적절하며, 몸을 통하여 구현되는 활동(activities)으로서의 마음의 작용 특성을 반영하기 어렵다.

인간의 마음이 ‘뇌 속에 갇힌 인지’가 아니라 능동적으로 환경과 상호작용하는 활동과정상에서 역동적 시간 궤적 상에서 나타나는 것인데, 이러한 상호작용적 활동성을 무시하고 정적인 상징(기호)표상의 저장과 이에 대한 고전적 계산으로서의 ‘마음’으로 개념화함으로써, 인지활동의 상황의존성, 맥락의존성, 사회문화요인에 의한 결정성 등이 무시되었고, 실제 장면에서 여러 가지 문제를 유발하는 인공물을 디자인하게 하였다. 즉 인간과 환경 인공물간의 변증법적 통일성(dialectic unity in activity) 측면을 파악하지도, 살리지도 못하였다(이정모, 2007).

이제 앞서 제시된 마음의 새로운 개념, 즉 뇌와 몸과 환경이 하나로 엮어진 통합체에서의 능동적 상호작용 활동으로 재구성된 마음 개념의 체화적 인지 틀을 도입한다면, 인공물이, 그

리고 이들이 구성하는 현실공간이나 가상공간이, ‘확장된 마음’으로서, 그리고 마음의 특성을 형성, 조성하는 기능 단위 또는 공간, 대상 및 사건으로서 작용하며, 마음과 인공물이 하나의 통합적 단위를 형성한다고 볼 수 있다. 그렇게 마음과 인공물의 관계를 재구성한다면, 인간의 마음의 작동 특성 본질의 심리학적, 인지과학적 탐구는 물론, 인간의 각종의 적응, 부적응의 이해와 이러한 변화의 각종 응용심리학적, 응용인지과학적 적용 실제(practice)에서 새로운 좋은 틀을 이루어 낼 수 있을 것이다.

### 11.6. 인간과-행위주체자로서의 인공물(A-2)의 상호작용

환경을 이루는 것이 일반 인공물뿐만이 아니라, 행위를 스스로 낼 수 있고 인간과 쌍방향적으로 상호작용할 수 있는 미래의 로봇이나 다른 인지시스템 같은 행위주체자(대행자; agents)들이라고 한다면, 인간을 포함한 이 행위 주체자들 사이의 쌍방향적, 사회문화적 상호작용의 측면의 중요성을 고려하지 않을 수 없다. 체화된 마음이 전개하는 상호작용 상황은 Gibson 류의 생태지각심리학적 접근이 주로 다루어 온 물리적 세상 상황을 넘어서는 것이다. 체화된 마음의 접근이 데카르트의 존재론(49)을 넘어서고, Gibson의 생태심리학적 이론틀을 넘어서서 정립되어야 하는 필요성이 여기에서 제기된다.

더구나 R. Kurzweil(2005)등이 주장하듯이 인간과 인공물의 신체적 그리고 심리적(인지적) 경계가 무너지는 특이점(the Singularity)이 (비록 약소한 형태이더라도) 30년 내에 도달하게 될 수도 있다는 가능성을 고려한다면 데카르트 류의 존재론에서 개념화한 환경의 물질적 대상, 도구와의 상호작용을 넘어서, 인간의 동반자(partners)(Seifert, 2008)로서 인간과 더불어 (또는 인공물끼리) 사회적, 문화적 관계를 창출해 낼 행위주체자로서의 인공물과의 새로운 유형의 심적, 행위적 상호작용이 전개될 것을 예측하고 이에 대한 개념화가 이루어져야 할 것이다. 신체적으로 그리고 심리적 과정 측면에서 인공물과 인간의 경계가 무너지는 가능성 등을 생각한다면 마음에 대한 개념화와 탐구에서 인간 마음과 공진화해갈 새로운 형태의 인공물, 새로운 형태의 상호작용에 대한 심리학적, 인지과학적 개념적 재구성 작업이 필요하게 된다.

### 11.7. 체화적 접근의 문제점

고전적 인지주의에 대한 제 3의 대안으로도 지칭되는 체화적 접근도 그 자체가 지니고 있는 문제점도 있다. 우선 개념적으로 정리되어야 할 것은 ‘체화된 마음(EM)’ 접

---

49) 데카르트 틀의 부족함은 리콰르가 다음과 같이 잘 지적하고 있다(윤성우, 2005).[데카르트의 코기토는 ... “사유하는 어떤 것”이지, 타자와 담화를 교환하고 행위를 주도하는 “누구”도 아니며, 그 행위의 상대자가 될 수 있는 누구도 아니며, 한 인간의 진술한 이야기 속에 등장하는 인물이 될 수도 있는 그 누구도 아니며.. 너무나 강한 1인칭 중심의 코기토이기에, ‘나’뿐만 아니라 ‘너’, ‘그/그녀(들)’ 심지어는 각자의 주체성을 지칭할 수 없을 뿐 아니라..... (223쪽)]

근에서의 행위 개념과 고전적 행동주의의 행위 개념의 차별화의 문제이다. 체화된 마음 접근이 행위를 강조하기에 이 접근과 행동주의심리학이 기본적으로 같지 않는다고 의문을 제기할 수 있지만 Keijzer(2005)에 의하면 두 접근들은 차이가 있다. 행동주의를 보완하여 이론적 행동주의로 재구성하더라도, 행동주의의 행동 개념과 체화된 마음 접근의 행동 개념과 차이가 있다. 행동주의는 일방향적이며 단순 행동 개념 상에서 입장을 전개하며 인지를 배제하지만, 체화된 마음 접근은 복잡한 구조를 지닌 행위 개념을 이론적으로 전제하며, 그 행위가 쌍방향적이며, 행동적 맥락을 인정하면서도 인지의 개념을 재구성한 형태로 살린다. 두 접근의 행위 개념은 같은 개념이 아닌 것이다.

체화적 접근에 대한 다른 중요한 문제점도 있다. 아담즈 등(Adams와 Aizawa, 2008, 2010)은 그들의 책 [인지의 경계]라는 책과 논문들에서 ‘체화적 인지’ 또는 ‘체화적 마음’ 접근에 대한 강력한 이의를 제기하였다. 그들이 체화적 접근을 비판한 핵심은 체화적 인지 접근이 지니는 등가성(parity) 상정의 오류와 결합(coupled) 논지의 오류를 지적한 것이었다. 한편 국내에서 장대익(2010)은 발생계이론에 대한 비판적 논의의 전개와 체화된 인지 이론의 비판적 논의의 전개 과정의 유사성을 제시하며, 이 문제를 다시 거론한다. 체화된 마음 접근은<sup>50)</sup> 환경이 인지 과정의 일부분처럼 기능을 한다고 하여, 환경을 뇌와 마찬가지로 인지체계의 부분이라는 등가성 원리를 제시하는 데, 장대익 등은 오랜 진화적, 발생적 역사를 지닌 뇌와 그렇지 않은 외적 자극을 등가적으로 간주하는 것이 오류라고 지적한다. 또한 체화적 인지 접근은 뇌, 몸, 환경이 하나의 불가분적 결합된 체계(coupled system)를 이루기에 환경을 인지체계의 부분으로 간주하는 것은 타당하다고 보는데, 이에 대하여 장대익은 인지가 뇌, 몸, 환경간의 복잡한 인과적 상호작용의 산물임을 기꺼이 받아들인다 하더라도, 뇌를 인지의 특권적 담지자 또는 표지(mark)로 받아들이는 데는 문제가 없으며, 환경을 인지체계의 부분으로 간주하는 것은 틀림없는 오해이다 라는 등의 논지를 전개하였다. 아담즈 등이나 장대익의 지적처럼 환경은 인지의 내재적(intrinsic) 내용을 제시하는 것이 아니라 단지 인지의 파생적(derived) 내용을 제시할 뿐이라고 비판받을 수 있다.

이러한 비판의 논지에 대하여 Chemero(2009) 등은 그러한 지적은 동물과 환경을 하나의 선형적 체계 틀 내에서 연결할 때면 타당할지 몰라도, 비선형적 체계, 동역학적 체계에서는 문제가 되지 않는다고 반박한다. 또한 Clark(2010)는 체화된 인지 접근의 핵심은 미세 수준에서 접근하려는 신경과학적 시도를 시스템 수준, 즉 거시 수준으로 상향시켜 접근하여야 한다는 과학적 입장을 제시하려고 한 것이며, 환경이 곧 뇌와 등가적으로 인지적 표지를 지닌다는 언급을 한 적이 없으며, 환경이 뇌, 몸과 함께 커다란 인지체계를 이룬다는 것을 주장했을 뿐이라고 대응한다. 그에 의하면 인지적 과정이 과연 무엇인가의 규정과, 파생되지(derived) 않은 내재적(표상)내용의 규정

---

50) 여기서 언급되는 체화적 접근은 ‘급진적 체화된 인지 접근’을 지칭하는 것이 아님.

등은 아담즈 등이 상정하듯이 그리 단순한 문제가 아니라고 한다.

우리는 여기에서, 무엇이 실제의 '세상'인가, 무엇이 '인지적'인가에 대한 상이한 존재론적, 인식론적 입장들에 바탕을 둔, 반-체화적입장 및 체화적 입장을 보게 된다. 체화적 인지 접근에 대한 찬반론은 과학적 설명의 논의에 앞서, 존재론적 가정에 대한 논의가 혼재되어 있어서 단순한 찬반의 과학적 논지의 전개로는 해결되기 어렵다고 하겠다.

### 11.8. 체화된 인지: 맺는 말

체화된 마음 입장에 의하면 마음과 뇌가 동일한 것이 아니며 마음은 [뇌를 포함하는 몸]과 [환경]의 집합체상에서 일어나는 어떤 활동으로 개념화해야 한다는 것이다. 환경에 공간을 차지하는(extended) 몸이 있는, 그에 바탕을 둔 마음 개념을 재구성하여야 하는 것이다. 마음은 한 사람 뇌 속에만 갇혀있는 그 무엇, 한 개인의 그 무엇이 아니라, 환경과 통합되며 여러 다른 사람의 마음, 그리고 다른 인공물들에 의하여 지원을 받거나 상호작용하면서 그들과 함께 진화되며 문화사회적으로 구성되며 공유되는 것에 의해 특징지어지는 활동으로서의 마음인 것이다.

이러한 접근이 가져오는 인지과학적, 심리학적 시사는 상당히 크다. 그 하나는 이 접근이 기존의 심리학, 인지과학의 암묵적 가정으로 전제되어온 데카르트 식 존재론과 인식론적 가정들에서 벗어나게 하고, 생체가 환경과 상호작용하는 체험적 행위의 측면까지도 다루며, 심리학과 인지과학이 현상학적 체험을 포함한 더 넓은 영역과 관점의 과학이 될 수 있게 한다.

물론 이와 관련하여 부수적으로 문제되는 것은 방법론적 세련화의 문제이다. 몸을 통하여 환경과 상호작용하는 역동적이고 다면적인 인간(몸)-환경의 상호작용 측면을 어떻게 경험과학적 탐구의 면면으로 방법론적으로 객관화하는가가 큰 과제로 남을 수 있다. 또한 주관적 현상학적 측면을 전통적 경험과학적 연구 틀에 객관화 가능한 형태로 어떻게 도입할 것인가 등의 문제들이 남는다.

체화된 인지 접근의 강한(급진적) 입장, 약한(온건한) 두 부류 중에서 어떤 접근이 우세할 것인가는 앞으로 이론적 측면에서 더 많은 논의의 대상이 되겠지만 필자의 직관으로 예상되기는 앞으로의 심리학, 인지과학이 '급진적 체화된 인지'의 접근보다는 '약한(온건한) 체화된 인지' 접근을 취하리라 생각된다. 그렇기는 하지만 '급진적 체화된 인지' 접근의 버팀목인 '동역학적 체계'로서의 인지 또는 마음에 대한 개념적 틀은 실제 연구를 수행하는 과학자들에 의해 더욱 도입되고 확산되리라 본다. 그리고 고전적 인지주의는 변형된 형태로 잔존하여 체화된 인지 접근이 설명하지 못하는 인지/마음의 영역을 설명하는 틀로 남아 있을 것 같다. 이러한 측면을 윤성우(2005) 교수가



리콤피르를 논의하며 전개한 글을 인용하여 표현하자면 다음과 같다(괄호 안은 필자의 첨언).

“달리 표현하자면 나의 사유함이, 나의 존재함보다 우월한 사태라는 것이 데카르트의 입장이라면, 나의 존재함이 - 리콤피르에게서는 보다 구체적으로 몸에 근거한 의지하기, 상징읽기, 말 주고받기, 이야기하기, 윤리적으로 행위 하기 등이다 (즉 체화된 마음 접근의 입장) - 나의 사유함보다 우선하고 그것을 넘어선다는 것이 리콤피르의 입장이다. 그렇다고 해서 극단의 반(反)표상적이거나 반(反)재현적인 포스트모던적 테제에 동의하는 것은 아니다. 사유와 존재 사이의 불일치 또는 사유로부터의 존재의 탈주를 주목하고자 하는 것이지, 대상화하고 표상하는 사유의 무용성을 주장하는 것은 아니다(225쪽).”

'체화된 인지' 접근의 떠오름은 심리학, 인지과학의 응용 영역에서 기존의 각종 인간-인공물 사이의 상호작용에 대한 관점을 이론적 접근이나 연구들을 크게 바꾸어 놓으리라고 본다. 또한 많은 사회-문화적 인간 활동 및 제도에 대한 설명과 이해가 이 새 틀에 의하여 달라지리라 본다.

## 12. 전체 종합

이 글에서 마음과 기계의 연결을 위한 시도들이 역사적으로 면면히 이어져 왔음을 보았다. 마음과 기계의 전반적 연결이라고 하기 보다는 기계와 인지적 과정인 사고(지능)의 연결이 수학자들을 비롯한 여러 학자들의 관심을 끌었던 주제였다. 앞에서 이미 언급되+었지만, 마음과 기계를 연결하여 20세기에 인류에게 디지털 문화를 가능하게 한 새로운 패러다임의 형성에 결정적으로 공헌한 여러 생각의 흐름들이 있었다.

첫째는 철학과 수학에서의 형식이론과 계산이론의 공헌이다. 특히 앞에서 기술한 바와 같이 수학자 튜링은 <튜링 자동기계>이론을 제시하여, 이전에는 신비하고 과학적 연구가 불가능하다고 생각해 온 수학문제 풀이의 사고과정들을 엄밀히 형식화할 수 있는 가능성을 제시하였다.

또한 디지털 컴퓨터의 발달과, 폰노이만을 중심으로 한 저장된 프로그램의 개념의 발달, 컴퓨터 프로그램을 부울 식의 형식체계로 표현가능하다는 새년의 생각과 그의 정보이론, 그리고 두뇌는 하나의 논리계로 간주할 수 있으며 신경세포간의 작용을 명제논리체계로 표현할 수 있다는 맥컬르력과 피츠의 생각, 와이너 등의 사이버네틱스 이론, 디지털 컴퓨터는 단순한 숫자조작기계라기보다 범용적 목적의 상징조작 체계(general purpose symbol manipulation system)인 튜링계로 간주할 수 있다는 사이먼 등의 생각, 그리고 행동주의에 대한 날카로운 비판의 제기과 인지적 능력과

형식적 통사론을 강조한 촘스키의 언어학이론, 심리학 내에서의 인지심리학의 새로운 출범과 인지심리학적 연구의 실험결과들의 집적, 두뇌 손상자들에 대한 신경학적 연구결과들의 집적, 새로운 과학철학과 심리철학의 떠오름, 인류학, 사회학에서의 민속방법론, 및 새로운 인지인류학, 인지사회학의 떠오름 등이 모두 하나의 시대정신(정보, 정보처리, 계산, 인지적 마음, 신경적 활동 중심 등의 Zeitgeist) 배경을 형성하면서 수렴되어서 인지적 패러다임의 형성이 가능하였다.

그 당시인, 1950년대 중반에 이러한 생각들과 연구결과들이 하나의 통일적 틀을 형성하지 못한 채 소용돌이로 있다가 1956년 MIT에서 개최된 정보이론 심포지엄을 기폭제로 하여 하나의 새로운 과학적 패러다임으로 형성되었다.

앞서 언급한 바와 같이, 이 MIT 심포지엄에 참여한 심리학자, 언어학자, 전기공학자, 신경과학자, 정보이론학자 등이 공통적으로 생각했고, 깨달았던 개념은 기계와 마음은 별개라 아니라 하나의 공통적인 개념틀에 의하여 연결될 수 있는 것이다. 오랜 역사를 통하여 여러 학자들이 가다듬고 발전시켜온 '마음-기계'의 연결 시도가 구체화 가능하다는 깨달음이었다. 일반인들이 신비하다고 생각하여온 심적 작용이란 일종의 정보 변환 및 조작이며(계산; computation), 컴퓨터 은유의 개념적 틀에 바탕하여 마음의 본질과 그 작동 특성을 접근할 수 있으리라는 생각이었다. <마음이란 본질적으로 컴퓨터와 같이 정보처리의 문제이며, 이는 여러 상이한 학문들이 공통적으로, 협동적으로 접근할 수 있고, 그렇게 해야 한다>고 그들의 생각이 모아졌던 것이다. 이러한 생각이 바로 인지주의 패러다임적 변혁을 가져온 핵심적 개념, 생각이었던 것이다.

이러한 바탕에서 기계와 인간의 마음에 대한 관점의 변혁, 발상의 전환을 하나의 과학적 패러다임으로 구체화하며 그 순수이론적, 응용적 의의와 가능성을 탐구하는 새로운 융합적 종합과학으로서 수렴적으로 샘솟듯 생겨난 것이 '인지과학(cognitive science)'이요 그러한 생각의 바탕틀이 고전적 인지주의(Classical Cognitivism)인 것이다. 마음과 기계의 관계성에 대한 선대들의 생각을 발상의 전환을 통하여 하나의 중요한 자연과학적 연구 주제로 구체화하고, 그를 통하여 기계-마음의 연결을 탐구하는 마음의 과학을 자연과학의 반열에 올려놓을 수 있다는 과학적 패러다임의 변혁이 1950년대 후반과 1960년대 초에 이루어진 것이고, 그에 의해 인류에게는 디지털 사회가 시작되었다.

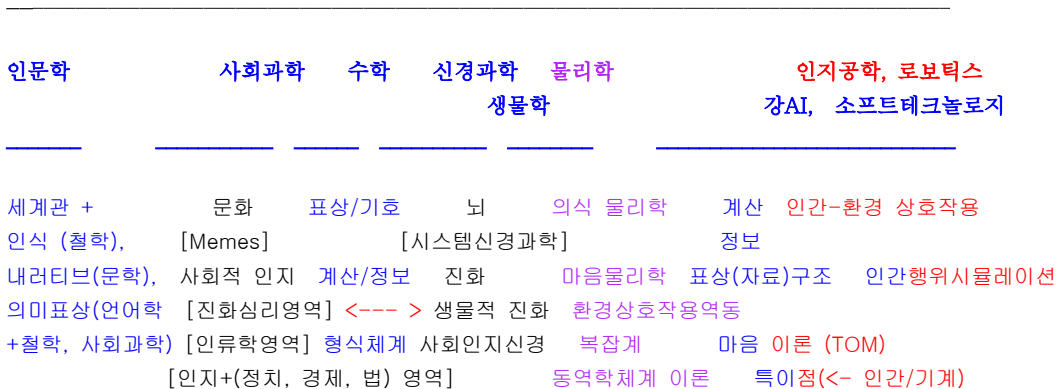
그러나 이러한 새로운 패러다임인 '고전적 인지주의'는 마음과 기계의 연결을 형식화된 언어로 엄밀히 기술하고 디지털 컴퓨터의 활용을 강조한 이론 틀을 추구 하다 보니, 그 나름대로의 문제점들이 있었다. 초기의 기능주의 철학의 다중구현가능성 개념을 강조하며 뇌를 소홀히 취급하였던 것이다. 이것이 1980년대 초중반에 신경망적 접근을 강조한 연결주의에 접근에 의하여 1차적으로 수정보완되었다. 1980년대 후반

에 시작되어 지금까지 계속되고 있는 뇌과학적, 신경과학적 접근의 연구들은 마음을 뇌신경활동으로 환원시켜 접근하는 대안적 설명 가능성을 주장하고 있다. 그러나 이 모두들은 데카르트의 존재론 전통을 벗어나지 못하고 있었기에 그에 따른 문제점들이 잔존하였다.

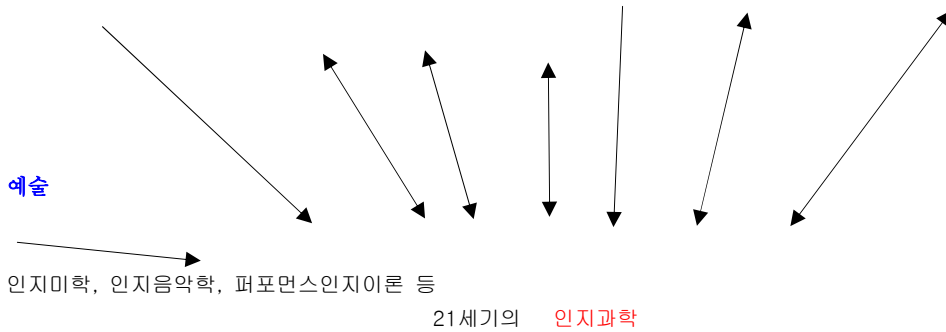
그러한 상황에서 탈데카르트적 관점과 가능성을 제시한 것이 ‘체화된 인지(마음)’이라는 보완적 새 틀이었다. 탈데카르트적인 관점의 심리학, 인지과학을 추구하는 체화적 인지 접근은 심리학, 인지과학의 주제와 영역을 확장시키며 일면 마음을 기계에 유추하면서도 ‘마음-환경(기계)-몸(뇌)’의 연결에의 새로운 접근의 가능성을 제시한다. 그 결과로, 환경 내의 인간, 그리고 행위주체(agents)로서 존재할 로봇 등의 인공물과 몸을 통하여 상호작용하는 행위 현상 일반이 심리학과 인지과학의 주요 분석대상이 될 수 있고, 미래의 심리학/ 인지과학은 생체로서의 인간 및 동물 자체뿐만 아니라 현재, 그리고 미래에 존재하게 되는 온갖 유형의 인공물, 특히 행위주체자(대행자)로서 작동할 여러 유형의 인공물, 인간의 몸이나 인지와 경계가 없는 그러한 미래 인공물과 인간이 상호작용 현상을 탐구하는 학문으로 그 탐구 영역이 확장되리라 본다.

이에 따라 기존의 많은 사회과학, 공학이 다루어 온 역동적 상황들, 연구영역들이 미래에는 심리학, 인지과학의 영역으로 포섭, 확장될 수 있는 것이다. 예를 들어 로봇 연구에서 로봇의 인지적, 정서적 반응, 로봇-로봇 상호작용, 로봇-인간 상호작용, 인간-로봇매개-인간 상호작용 등의 영역이 당연히 심리학, 인지과학의 영역이 되게 되는 것이다. 심리학, 인지과학의 내연과 외연이 확장되는 것이다. 체화된 인지의 틀을 도입하여, 확장된 인지과학과 인문학, 사회과학, 자연과학, 테크놀로지의 수렴 및 융합적 상호작용 관계를 그림으로 나타내자면 [그림 1]과 같다.

[그림 4]. ‘체화된 인지’ 접근을 포함하는 [인지과학]과 다른 학문과의 관계



인지(뇌)공학/ BCI <---> 인지기능향상(CE) 미래테크  
 -> Human V. 2.0



\* 인지과학은 AI, Internet과, IT, Robotics의 이론-개념적 모-학문이며 (enabled the Emergence of the Digital Age, AI & IT, and Robotics in Human History), 기존의 여러 학문들은 그 본질상 ‘체화적 인지(Embodied Cognition)’ 또는 ‘확장된(공간을 지닌) 마음 (Extended Mind)’ 틀을 내포한 새 인지과학에 연결된다. 인지과학은 그 본래적 특성이 학제적, 수렴적이며 (Multi-Disciplinarity), 자연히 인문학, 사회과학, 자연과학, 테크놀로지 모두를 수렴하는 융합과학이며 (A Convergence of Humanities, Social Sciences, Natural Sciences, & future Technologies), 인지과학의 형성, 확산, 재구성 등은 필연적으로 세계관, 인간관의 변화를 시사한다(implies a Paradigm Shift from the Old Conception of Sciences & Technologies, and World View. 인지과학은 인간-인공물(기계)의 구분을 넘어선 [제 2의 계몽시대]를 함의한다(going beyond the Singularity, and implies the 2nd **Enlightenment** in human history).

또한 커즈와일(R. Kurzweil)의 주장대로 가까운 미래에 특이점의 시점이 도래하여 인간과 인공물, 인간의 마음과 인공물의 지능 간의 경계가 어느 정도라도 허물어진다면, 마음, 지능의 개념뿐만 아니라, 인간 존재(being)의 개념이 밀뿌리부터 재구성되어야 하는 새로운 개념의 휴머니티 시점이 닥아 오는 것이다.

이러한 변혁은 17세기의 ‘신 중심에서 -> 인간 중심으로’의 ‘제1의 계몽시대’의 생각들의 변혁에 못하지 않은 생각들의 변혁을 시사하는 것이다. 이러한 변혁의 도래를 ‘제2의 계몽(깨달음) 시대’의 도래라고도 할 수 있을 것이다(이정모, 2008, 2009ㄴ). 20세기 중반의 마음(사고)-기계(컴퓨터) 연결에 의하여 인류 디지털 문화가 탄생되고 인간 인지적 능력을 컴퓨터의 형식 언어로 표현한 인류의 시도를 제2의 르네상스라고 한다면. 마음-몸-환경(인공물)을 하나의 괴리되지 않은 통일체로 개념화하며 그 역동적 상호작용과 새 휴머니티 개념에 초점 맞추는 것은 ‘제3의 르네상스’라고 할 수 있을 것이다. 이러한 개념을 도해로 나타내자면 [표2]와 같다.

[표3]. 조망: 년대별 ‘기계-몸(뇌)-마음(지능)’의 연결 특성

1930-1960년대; 사이버네틱스: 시스템[기계-뇌(신경계)-마음-사회] 간의 연결  
 1960-1970년대; 고전적 인지주의: 수학(형식체계)-컴퓨터(기계)-마음(인지)의 연결  
 1980-1990년대; 연결주의: (추상적 이론적) 뇌 - 컴퓨터 형식체계의 연결  
 1990-2010년대; 신경과학 : (실제적) 뇌의 신경적 과정으로 환원 /(마음=뇌; 기계= 부수적)  
 2000-2020년대: 체화된 인지 : 몸(뇌)-마음-환경(인공물: 기계)의 연결  
 2020-미래; ? : 몸(뇌)-[환경 => 사회(인간적, 인공물적) ] + [마음]의 연결

- ∴ new Humanity 틀
- 앞부분 <= 자연과학적, 공학적, 형식화 접근
- 뒷부분 <= 형식화, 인문사회과학적 접근
- 전 체 <= 인문, 이과의 구별 없이 융합, 수렴적 접근 통한 통합

20세기까지는 생물로서의 ‘인간의 마음’과 ‘인간이 만든 인공물’의 지능을 이분법적으로 구분하여 생각했는데 미래에 가서는 그것을 쉽게 나누기 어려운 시점이 오게 되고 두 개를 포괄하는 관점이 필요하게 된다. 그러한 관점은 말하자면 통합적 유물론이라 볼 수 있다.

이제 몸과 마음을, 또는 마음과 기계를 이분법적으로 생각해왔던 관점을 넘어서, 그리고 마음을 뇌로 환원시키려는 일원론적 유물론도 넘어서, 또 인간과 인공물(기계)을 이분법적으로 생각하였던 관점도 넘어서, 뇌, 몸, 환경(기계를 포함한 인공물도 환경의 일부임)이 괴리되지 않은 하나의 합체적 단위로서 작동하여 내는 그러한 틀로 마음의 개념을 재구성하여야 한다는 것이다.

그러한 변화가 강하게 부각될 미래의 시점에서 인지과학이나 심리학이 전통적 학문체제의 틀을 유지할 수 있을까, 아니면 기존의 심리학/인지과학보다 더 넓은 틀로써 사회과학, 공학, 인문학, 생명과학 등을 연결하는 종합적 ‘심리-인지과학 (psychological-cognitive science)’으로 변형될 것인가 궁금하여진다. 인지과학이나 심리학에서 전통적 객관적 실험 중심의 연구방법론 집착이 과연 타당한가, 다원적 방법론과 다원적 메타포를 도입하여야 하지 않는가, 과연 고전적 인지주의 또는 그의 변형이 21세기의 심리학, 인지과학의 나아가소는 인간학의 중심틀이어야 하는가 하는 물음이 자연히 던져지게 된다(Wallace & Ross, 2007).

다가오는 가까운 미래, 21세기 중반, 후반에 새로운 틀에서 심리학, 인지과학을 재구성해 낼 미래의 과학자들의 역할을 기대하여 본다.

## [참고 문헌]

- 김영정 (1996). **심리철학과 인지과학**. 서울: 철학과 현실사.
- 김재권 (1994). **수반과 심리철학**. 서울: 철학과 현실사.
- 다마지오, A. R. (저), 임지원(역). (2007). **스피노자의 뇌**. 사이언스 북스 (원제 Looking For Spinoza : Joy, Sorrow And The Felling Brain (2003))
- 마르틴 하이데거 (지음), 이기상 (옮김) (1998, 1927). 『존재와 시간』. 서울: 까치 (까치 글방 138)
- 모리스 메를로-퐁티 (지음), 류의근 (옮김) (2002, 1945). 『지각의 현상학』. 서울: 문학과 지성사.
- 안토니오 다마지오 (지음), 임지원 (번역) (2007). 『스피노자의 뇌: 기쁨, 슬픔, 느낌의 뇌과학』. 사이언스북스
- 알바 노에 (저), 김미선 (역) (2009). 『뇌과학의 함정』 (원저 제목: Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness.) 서울: 갈리온.
- 윤보석 (2009). 『컴퓨터와 마음: 물리 세계에서 마음의 위상』. 서울: 아카넷. (대우 학술총서 594)
- 윤성우 (2005). 『해석의 갈등』: 인간 실존과 의미의 낙원. 서울: 살림
- 이영의 (2008). 「체화된 마음과 마음의 병」, 철학탐구, 23집, 5-37.
- 이정모 (1983). Gestalt 개념의 형성사 (I): Descartes 에서 Hamilton 까지. **한국심리학 회지**, 4, 97-118.
- 이정모 (1998ㄱ). 언어심리학의 형성사 (I): Chomsky 이전 언어학과의 상호작용. 이정모, 이재호 (편), **인지심리학의 제 문제 (II) : 언어와 인지**. 서울: 학지사. (15-47 쪽)
- 이정모 (1998ㄴ). 「언어심리학의 형성사 (II): Chomsky 언어학과의 상호작용」. 이정모, 이재호 (편), **인지심리학의 제 문제 (II) : 언어와 인지**. 서울: 학지사. (49-75 쪽)
- 이정모 (2001). 『인지심리학: 형성사, 개념적 기초, 조망』. 서울: 아카넷 (대우학술총서 511).
- 이정모 (2001). **인지심리학: 형성사, 개념적 기초, 조망**. 서울: 아카넷. (대우학술총서 511)
- 이정모 (2002). 인지과학의 과거, 현재, 미래: 한국적 조망. **인지과학**, 13, 69-79.
- 이정모 (2003ㄱ). 언어와 사고. 조명한, 이광오, 김정오, 이양, 이광오, 신현정, 김영진, 고성룡, 김소영, 정혜선, 도경수, 이정모, 이현진 (지음). **언어심리학**. 서울: 학지사. (12장; 375-409쪽).
- 이정모 (2007). 「심리학의 개념적 기초의 재구성(2): 인지과학적 접근에서 본 ‘마음’ 개념의 재구성과 심리학 외연의 확장」, 한국심리학회지: 일반, 26, 2, 1-38.

- 이정모 (2008). 「마음의 체화적(embodied) 접근: 심리학 패러다임의 제6의 변혁」, 한국실험및인지심리학회 2008년 겨울 제43차 학술대회 논문집, 143-152.
- 이정모 (2009ㄱ). 『인지과학: 학문간 융합의 원리와 응용』. 서울: 성균관대학교출판부.
- 이정모 (2009ㄴ). 「인지로 모인다: 인지과학의 전개와 미래 융합학문」, 김광웅 외 지음. 『우리는 미래에 무엇을 공부할 것인가』. 서울: 생각의 나무. (103-159 쪽)
- 이정모(2010). 「‘체화된 인지(Embodied Cognition)’ 접근과 학문간 융합: 인지과학 새 패러다임과 철학의 연결이 주는 시사」, 철학사상, 38, 27-66.
- 이정모, 이건효, 이재호 (2004). 「사이버 인지심리학의 개념적 재구성: 인공물과 인지의 공진화」, 한국심리학회지: 실험, 16, 4, 365-391.
- 장대익 (2005). **이보다보 관점에서 본 유전자, 선택, 그리고 마음**. 서울대학교 박사학위 논문.
- 장대익 (2010). 「뇌를 넘어서? 체화된 마음 이론에 대한 비판」, 홍성욱, 장대익 (역음). (2010). 『뇌 속의 인간, 인간 속의 뇌』. 서울: 바다출판사. (339-367 쪽)
- 정대현 (2001). 『심성내용의 신체성』. 서울: 아카넷. (대우학술총서 528)
- Adams, F. R., & Aizawa, K. (2008). *The Bounds of Cognition*. Chichester, West Sussex: Blackwell Pub..
- Adams, F. R., & Aizawa, K. (2010). “Defending the bounds of cognition.” In R. Menary (Ed.). *The extended mind*. Cambridge, MA: MIT Press.
- Arbib. M. A. (1964). *Brains, machines, and mathematics*. New York: McGraw-Hill.
- Atneave F. (1959). *Applications of information theory to psychology: A summary of basic concepts, methods, and results*. New York: Holt, Rinehart, and Winston.
- Baars, B. J. (1986). *The cognitive revolution in psychology*. New York: Guilford Press.
- Bartlett, F. C. (1932). *Remembering*. Cambridge: Cambridge University Press.
- Bem, S., & Keijzer, F. (1996). “Recent changes in the concept of cognition.” *Theory & Psychology, 6*, 449-469.
- Bickhard, M. H. (2008). “Is embodiment necessary?” In P. Calvo & A. Gomila (Eds.). *Handbook of cognitive science: An embodied approach*. Amsterdam: Elsevier. (pp. 29-40).
- Boden, M. A. (2006). *Mind as machine: A history of cognitive science*. Oxford: Clarendon Press.
- Broadbent, D. E. (1958). *Perception and communication*. Oxford: Pergamon.
- Brooks, R. (1991). “Intelligence without representation.” *Artificial Intelligence, 47*, 139-159.

- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: John Wiley & Sons.
- Calvo, P. & Gomila, T. (2008). *Handbook of cognitive science: An embodied approach*. Amsterdam: Elsevier.
- Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chalmers, D. (2004). How can we construct a science of consciousness? In M. Gazzaniga, (ed) *The cognitive neurosciences III*. MIT Press
- Chemero, A. (2009). *Radical embodied cognitive science*. MIT Press.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton
- Chomsky, N. (1959). A review of B. F. Skinner's verbal behavior. *Language*, 35, 26-58.
- Churchland, P. M. (1986). *Neurophysiology: Toward the unified science of the mind/ brain*. Cambridge, MA: MIT Press.
- Clark, A. & Chalmers, D. (1998). The extended mind. *Analysis*, 58, 10-23.
- Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.
- Clark, A. (2001). Reasons, Robots and The extended mind. *Mind and Language*, 16, 121-145.
- Clark, A. (2003). *Natural - born cyborg: Minds, technologies, and the future of human intelligence*. London: Oxford University Press.
- Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford: Oxford University Press.
- Clark, A. (2010). "Coupling, constitution, and the cognitive kind: A reply to Adams and Aizawa." In R. Menary (Ed.). *The extended mind*. Cambridge, MA: MIT Press.
- Clark, A. (in press). Memento's revenge: Objections and replies to the extended mind" to appear in R. Menary (ed.) *Papers On The Extended Mind*.
- Cutland, N. (1980). *Computability: An introduction to recursive function theory*. Cambridge: Cambridge University Press.
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam. [번역서: 데카르트의 오류, 역자: 김린, 중앙문화사, 1999.]
- de Groot, A. D. & Gobet, F. (1996). *Perception and memory in chess: Studies in the heuristics of the professional eye*. Assen: Van Gorcum.
- de Groot, A. D. (1965). *Thought and choice in chess*. The Hague: Mouton.
- Dewey, J. (1896). "The reflex arc concept in psychology." *Psychological Review*, 3,



357-370.

- Dourish, P. (2001). *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge, MA: MIT Press.
- Dreyfus, H. (1991). *Being-in-the world: A commentary on heidegger's being and time, division I*. Cambridge, MA: MIT Press.
- Dreyfus, H. L. (2006). Why Heideggerian AI failed and how fixing it would require making it more Heideggerian (with the help of Merleau-Ponty). Special Address, 2006 *American Philosophical Association Conference*.
- Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge: Cambridge University Press.
- Evans, J. St. B. T. (1989). *Bias in human reasoning*. Hillsdale, NJ: Erlbaum.
- Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning, *Memory & Cognition*, 11, 293-306.
- Evans, J. St. B. T., Over, D. E., & Manktelow, K. I. (1993). Reasoning, decision making and rationality. *Cognition*, 49, 165-187.
- evolution*. New York: Basic Books.
- Fodor, J. A. (1983). *The modularity of mind: An essay of faculty psychology*. Cambridge, MA: MIT Press.
- Gallagher, S. (2005). *How the body shapes the mind*. Oxford: Oxford University press.
- George A. Miller (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences*. 7, 141-144.
- George A. Miller (2003). The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences*. 7, 3, 141-144.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.
- Gunderson, K. (1985). *Mentality and machines* (2nd ed.), London: Croom Helm.
- Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.
- Hollnagel, E. (2007). "The elusiveness of cognition." In B. Wallace, A. Ross, J. Davies, & T. Anderson, T. (Eds.) (2007). *The mind, the body and the world: Psychology after Cognitivism?* Exeter: Imprint Academic. (pp. 346-355).
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Jackson, Jr. P. C. (1985). *Introduction to Artificial Intelligence*. New York: Dover.
- Keijzer, F. (2005). "Theoretical behaviorism meets embodied cognition: Two

- theoretical analyses of behavior.” *Philosophical Psychology*. 18, 1, 123-143.
- Kurzweil, R. (2005). *The Singularity Is Near*. New York: Viking. [한국어 번역본: 레이 커즈와일 (지음), 김명남, 장시형 (옮김). 특이점이 온다: 기술이 인간을 초월하는 순간. 서울: 김영사. 2007.]
- Lakoff, G. & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic Books.
- Maturana, H. & Varela, F. (1987). *The Tree of Knowledge: The Biological Roots of Human Understanding*. Boston: Shambhala.
- Menary, R. (ed.) (2010). *The extended mind*. Cambridge, MA: MIT Press.
- Miller G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt.
- Minsky, M., & Papert, S. (1969). *Perceptrons*. Cambridge, MA: MIT Press.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Rockwell, T. (2005). *Neither Brain nor Ghost: A nondualist alternative to the mind-brain identity theory*. Cambridge, MA: MIT Press.
- Rowlands, M. (2010). *The New Science of the Mind: From Extended Mind to Embodied Phenomenology*. Cambridge, MA: MIT Press.
- Rumelhart, D. E., McClelland, J. J., & PDP Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.
- Seifert, U. (2008). “The co-evolution of human and machines: A paradox of interactivity.” In U. Seifert, J. H. Kim, & A. Moore (Eds.) *Paradoxes of interactivity: Perspectives for media theory, Human-Computer Interaction, and Artistic Investigation*. transcript Verlag, Bielefeld. (pp. 8-23).
- Selfridge, O. (1956). Pandemonium : A paradigm for learning. In Blake, D. V. & Uttley, A. M., (Eds.), *Proceedings of the Symposium on Mechanization of Thought Processes*, 511-529.
- Skinner, B. F. (1957). *Verbal Behavior*. Acton, MI: Copley Publishing Group.
- Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Cambridge: MA: Harvard University Press.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.
- Wallace, B., & Ross, A (2007). “Conclusion: The future of an illusion.” In B. Wallace, A, Ross, J. Davies, & T. Anderson, T. (Eds.) (2007). *The mind, the body and the world: Psychology after Cognitivism?* Exeter: Imprint Academic. (pp. 356-361).

- Wallace, B., Ross, A., Davies, J., & Anderson, T. (Eds.) (2007). *The mind, the body and the world: Psychology after Cognitivism?* Exeter: Imprint Academic.
- Wheeler, M. (2005). *Reconstructing the Cognitive World: the Next Step*. Cambridge, MA: MIT Press
- Wilson, M. (2002). "Six views of embodied cognition." *Psychonomic Bulletin & Review*, 9, 625–636.