

통합적 인지 모형의 가능성

이 영 의*

인지과학에서 최근 논의되고 있는 인지 이론들은 인지에 대한 적절한 모형을 제공하지 못하고 있다. 전통적인 인공지능 이론은 추리나 문제 해결과 같은 과제에는 적절한 것처럼 보이지만 문자와 음성 인식과 같은 패턴 인식 분야에서는 여전히 비효율적이다. 연결주의는 전통적인 인공지능 이론과는 정반대의 양상을 보이고 있다. 연결주의 체계는 패턴 인식에는 강하지만 추리에는 약하다. 한편 최근에 제시된 상황화된 행동 이론은 전통적인 인공지능과 연결주의에서 기본적으로 전제되고 있는 표상의 개념을 부정하고 실제 세계에서 직접 유래되는 지각에 바탕을 둔 모형을 제시하지만 인간의 인지를 효과적으로 설명하고 있지 못하다. 인지 모형들이 갖고 있는 이러한 한계점들을 강조하여 나는 이 글에서 인공지능, 연결주의, 상황화된 행동 이론을 각각 좌뇌 모형, 우뇌 모형, 로봇 모형이라고 부르고 그러한 한계 상황을 벗어날 수 있는 방법으로서 모형들간의 양립가능성을 이용한 통합적 인지 모형의 구축을 모색한다.

[주제어] 인지 모형, 인공지능, 연결주의, 상황화된 행동 이론, 표상, 물리적 뿌리내림, 통합적 인지 모형

1. 통합적 인지 모형

인간의 인지에 대한 적절한 모형은 무엇인가? 이러한 질문은 다시 인지를 구성하는 기본 요소들이 무엇이고 어떻게 그것들이 작용하는가와 같은 보다 구체적인 질문들로 세분화될 수 있다. 철학자들은 전통적으로 마음과 그와 관련된 심리 현상을 중요한 주제로 다루어 왔고 심리철학과 인식론에서 볼

* 고려대학교 철학과 강사
전자우편 : rheeye@korea.ac.kr

수 있듯이 다양한 이론들을 제시해 왔다. 철학자들은 두 가지 방향에서 인지에 접근해왔다. 첫 번째 방향은 마음의 실체성과 본성에 대한 존재론적 탐구이며, 심신 일원론, 심신 이원론, 그리고 기능주의와 같은 이론들은 존재론적 탐구의 결과물이다. 두 번째 방향은 마음의 인식적 체계와 작용에 관한 인식론적 탐구인데 추리와 발견의 논리에 관한 이론들은 이러한 탐구의 결과이다. 물론 철학자들이 제시한 이론에서 존재론적 탐구와 인식론적 탐구를 명확하게 구분하는 것은 어렵고, 경우에 따라서는 사실상 불가능하기도 하다. 또한 존재론적 탐구가 없이 진정한 인식론적 탐구는 불가능할 수도 있다. 그러나 우리는 인지 모형을 다루는데 있어서 논의의 편의상 그러한 두 가지 접근을 구분해 볼 수 있고, 특정한 존재론적 입장을 전제로 하거나 또는 그러한 전제가 없이 인식론적 논의를 할 수 있을 것이다.

이 글에서 논의되는 인지에 대한 모형들은 특정한 존재론적 입장을 전제로 제시된 것들이다. 그 모형들은 기능주의를 전제로 인지과학에서 개발된 이론들이다. 우리의 질문, 즉 적절한 인지 모형은 무엇인가라는 질문에 대해 인지과학 분야에서 주로 논의되는 입장들은 인공지능(artificial intelligence, 이하 AI), 연결주의(connectionism), 그리고 상황화된 행동 이론(situated action theory)이다. 이러한 이론들은 인지에 대한 서로 다른 모형을 제시하기 때문에, 그것들의 적절성을 판단하기 위해서는 비교와 평가 작업이 선행되어야 할 것이다. AI, 연결주의, 그리고 상황화된 행동 이론을 비교 평가하는 많은 문헌들이 있다. 예를 들어, 쿤이 제시한 과학혁명론을 도입하여 AI 패러다임이 몰락하고 연결주의 또는 상황화된 행동 이론이 새로운 패러다임이 될 것이라고 주장하는 입장이 있다. 모형들의 비교 평가의 문제에 있어서 두 가지 가능성이 있다. 즉, 세 가지 이론 중 특정한 이론이 이러저러한 이유로 다른 이론에 비하여 보다 적절한 모형을 제시하거나 아니면 그 중 어떤 것도 이러저러한 이유로 적절한 모형을 제공할 수 없다. 이론들의 적절성을 판단하는데 있어서 가장 필요한 작업은 위의 두 가지 가능성에서 나타나는 “이러저러한 이유”를 제시하는 것이다.

나는 이 글에서 적절한 인지 모형이 다루어야 할 몇 가지 항목들을 설정하

고 그러한 항목들이 어떻게 AI, 연결주의, 그리고 상황화된 행동 이론에서 설명되는가를 검토함으로써 인지 모형을 효과적으로 평가하는 한 가지 방식을 제시한다. 여기서 제시된 방식은 인지 모형을 비교 평가하는데 필요한 “이러저러한 이유”를 뒷받침하게 될 것이다. 신경 생리학자들의 연구 결과에 따르면, 인간의 뇌는 좌뇌와 우뇌로 구분되는데, 좌뇌는 언어와 논리와 같은 추상적 활동의 중추이고 우뇌는 패턴 인식과 같은 공간 인식의 활동의 중추이다. 물론 인지는 좌뇌와 우뇌의 활동이 통합적으로 나타난 결과이다. 그런데 AI, 연결주의, 상황화된 행동 이론은 인지 전체를 총괄적으로 설명하려고 제안되었지만 그 동안의 많은 논의를 통하여 어떠한 이론도 다른 이론에 비해서 이론적으로 우월하지 못하다는 점이 드러났다. AI 모형을 채택한 컴퓨터는 언어와 논리 활동에서 우수하고 연결주의를 채택한 컴퓨터는 패턴 인식에서 우수하다. 또한 상황화된 행동 이론을 채택한 로봇들은 실제 세계에서 유래한 자극에 반응하면서 주어진 과제를 수행한다. 나는 인지 모형들에서 나타나는 이러한 부분적인 성공을 강조하여 세 가지 이론들이 제공하는 인지 모형을 각각 좌뇌 모형, 우뇌 모형, 그리고 로봇 모형이라고 부르겠다. 적절한 인지 모형은 인지를 설명하기 위해서 심성의 내용, 심성의 내용이 변화되고 처리되는 과정, 그러한 변화를 구성하는 기본 요소, 기본 요소들이 연결되고 조합되는 규칙, 그리고 인지를 연구하는데 적합한 분석 수준을 제시해야 한다. 또한 이러한 요소들이 인지모형을 구성하는 방식이 설명되어야 한다. 나는 이러한 6가지 항목, 즉 심성 현상의 내용, 과정, 기본 요소, 규칙, 분석 수준, 그리고 인지 구성을 이용하여 AI, 연결주의, 상황화된 행동 이론의 이론적 양립 가능성을 검토할 것이다. 우리는 인지를 전반적으로 설명할 수 있는 적절한 인지 모형이 필요하며, 그러한 모형은 현재 부분적으로 성공을 거두고 있는 인지 모형들을 비교하고 통합하는 과정에서 나타날 것이다.

2. AI의 좌뇌 모형

“인공지능”이라는 용어는 일반적으로 넓은 의미로 사용되어 전통적인 기호론적 접근뿐만 아니라, 연결주의, 상황화된 행동 이론, 그리고 전문가 체계 등과 같은 인공 지능을 연구하는 모든 분야를 포함한다. 그러나 나는 이 글에서 인공지능에 대한 매우 제한된 용법을 채택하여 1956년에 개최된 다스머쓰 회의(Conference of Dartmouth)에서 추구된 인공지능에 대한 기호론적 접근을 토대로 하는 기호주의(symbolism) 만을 AI라고 보기로 한다.¹⁾ AI가 제공하는 좌뇌 모형은 그 동안 추리, 탐색, 문제 해결, 그리고 전문가 체계와 같은 분야들에서 많은 성과를 거두어 왔다. 인지 이론으로서 AI를 구성하는 몇 가지 가설들이 있는데, 그 중에서 AI의 성격과 특징을 가장 잘 드러내는 가설들은 사고언어 가설(language of thought hypothesis), 물리적 기호체계 가설(physical symbol system hypothesis), 그리고 인지에 대한 기호적 구성(symbolic architecture of cognition)이다.²⁾ 그러한 세 가지 가설들을 다음과 같이 규정하기로 한다.

- ① 사고언어 가설: 사고는 심성 언어에서 표상되며, 그러한 표상은 생산성, 체계성, 조합성을 갖는다(Fodor, 1975).
- ② 물리적 기호체계 가설: 물리적 기호체계는 일반적 지능 행위를 위한 필요 충분조건이다.
지능은 구문론적 규칙들에 의한 기호 조작의 과정이다(Newell and Simon, 1976).

1) 이러한 구분은 인공지능과 관련된 학술지들에 실리는 논문의 내용을 살펴보면 정당화된다. 예를 들어, 인공지능 분야의 대표적인 학술지인 「인공지능」(Artificial Intelligence)에는 대부분 기호론적 AI에 관한 논문들만이 게재되고 있으며 연결주의와 관련된 논의들은 찾아보기 힘들다.

2) AI에는 위에서 제시된 가설 이외에도 다른 종류의 근본 가설들이 있다. 예를 들어, 처치와 튜링 논제(Church-Turing thesis), 마음에 대한 표상 이론(representational theory of mind)과 기능주의(functionalism)가 그러한 경우에 속한다. 그러나 처치와 튜링 논제는 인지에 대한 기호적 구성과 관련하여 설명할 수 있고, 마음에 대한 표상 이론과 기능주의는 사고언어 가설과 관련하여 설명할 수 있으므로 생략한다.

- ③ 기호적 구성 가설: 컴퓨터의 구성 개념은 인지를 구성하는데 적용될 수 있다(Anderson, 1995; Newell, Rosenbloom, and Laird, 1989).

이상의 세 가지 가설을 통하여 AI가 제시하는 좌뇌 모형의 골격을 파악해보자. 사고언어 가설에 따르면, 사고는 인간의 뇌에서 물리적으로 구현된 기호적 체계로서의 심성 언어에서 이루어지고, 명제태도(propositional attitude)를 통하여 외부 세계와 관련된다. 예를 들어 “어떤 사람 S가 명제 P를 믿는다”는 문장에서 S는 P에 대한 명제 태도를 갖는다고 한다. 사고언어 가설을 지지하는 사람들은 명제 태도에 의하여 인간의 사고와 행동간의 관련성을 설명할 수 있다고 생각한다. 예를 들어, 철수가 집에 들어오자 말자 냉장고의 문을 여는 행동은 그가 시원한 물을 마시고 싶기 때문이라고 설명되었다고 가정하자. 여기서 “왜 철수가 집에 들어오자 말자 냉장고의 문을 열었는가?”라는 질문은 “철수가 시원한 물을 마시고 싶다”는 심리 상태에 있다”는 것과 “철수는 ‘냉장고에 시원한 음료수가 있다’고 믿는다”라는 명제 태도를 이용하여 대답된다. 우리가 갖고 있는 다양한 개별적인 명제태도는 사고언어에서 표상된다고 가정되는데 그러한 표상을 심성 표상이라고 한다. 사고언어 가설에 따르면, 심성 표상은 그 본질적 특성으로서 생산성(productivity), 체계성(systemacity), 조합성(compositionality)을 갖는다. 여기서 생산성은 사고언어에서 무한히 많은 사고들이 생산될 수 있다는 것을 의미하고, 체계성은 특정한 문장을 생산하는 능력은 본질적으로 다른 문장을 생산하고 이해하는 능력과 체계적으로 관련되어 있다는 것을 의미한다. 예를 들어, 어떤 사람이 “존은 메리를 사랑한다”를 표현할 수 있다면 그는 “메리는 존을 사랑한다”는 문장을 표현할 수 있다는 것이다. 체계성이 구문론적 특징이라면 조합성은 의미론적 특징이다. 조합성은 체계적으로 연관된 사고들은 의미론적으로도 연관되어 동일한 의미론적 요소들로부터 구성된다는 것을 의미한다. 예를 들어, “존이 메리를 사랑한다”를 생각하는 능력은 “메리는 존을 사랑한다”를 생각하는 능력과 연관되지만 “ $2 + 3 = 5$ ”를 생각하는 능력과는 무관하다.

AI가 제공하는 좌뇌 모형의 토대를 이루는 두 번째 가설은 물리적 기호체

계 가설이다. 사고언어 가설이 철학적인 배경에서 제시된 이론인 반면에 물리적 기호체계 가설은 공학적 배경에서 제시되었기 때문에 전자에 비하여 보다 더 경험적이다. 물리적 기호체계 가설에 따르면, 컴퓨터(정확히 말하면, 폰 노이만 방식의 컴퓨터)는 물리적 기호체계인데, 물리적 패턴으로서의 기호, 기호들의 구조로서의 표현, 표현에 대한 작동으로서의 과정으로 구성된다. 물리적 기호체계에 대한 의미론은 지시(designation)와 해석(interpretation)의 두 단계로 구분된다. 즉 표현이 체계에 주어져서 체계가 대상에 영향을 미치거나 그 대상에 의존하는 방식으로 작용할 경우 표현은 대상을 지시한다고 정의된다. 또한 표현이 과정을 지시하고 표현이 주어지면 체계가 과정을 처리할 때 체계는 표현을 해석한다고 정의된다. 이러한 개념적 틀을 토대로, 뉴웰과 싸이먼은 물리적 기호 체계는 일반적인 지능적 행동을 위한 필요 충분조건이라고 주장한다(Newell and Simon, 1976).

이제 사고언어 가설과 물리적 기호체계 가설을 이용하여 좌뇌 모형이 서론에서 제기된 사항들에 대해 어떠한 대답을 제시하는가를 검토해 보자. 첫째, 인지 상태는 표상 상태이거나(사고언어 가설) 기호처리 상태이다(물리적 기호체계 가설). 여기서 표상 상태와 기호처리 상태는 그 내용에 있어서 본질적으로 차이가 없다. 예를 들어, 사고언어 가설을 지지하는 사람들에게는 명제태도의 대상으로서의 표상 체계는 일종의 기호 체계에 해당하기 때문에, 인지 상태는 곧 표상 상태이다(Fodor and Pylyshyn, 1988: 12-13). 둘째, 표상들이 처리되는 심성 과정은 개별적인 표상들의 인과적 연속이거나(사고언어 가설) 기호들로 구성된 표현의 처리 과정이다(물리적 기호체계 가설). 그러므로 심성 과정은 표상 또는 표현들의 계산 과정이다. 셋째, 사고언어 가설과 물리적 기호체계 가설에 따르면 표상 또는 표현들은 기호들로 구성되기 때문에 심성 상태의 기본 요소는 기호이다. 여기서 기호들이 계산되는 방식에 대한 질문이 제기되는데, 이에 대한 AI의 표준적 대답은 계산은 명시적 규칙에 따른다는 것이다. 그러나 사고언어 가설과 물리적 기호체계 가설은 이 점에서 약간의 차이점을 갖는다. 사고언어 가설에 따르면 계산은 논리적 규칙에 따르는 과정이지만 물리적 기호체계 가설에 의하면 그것은 보다

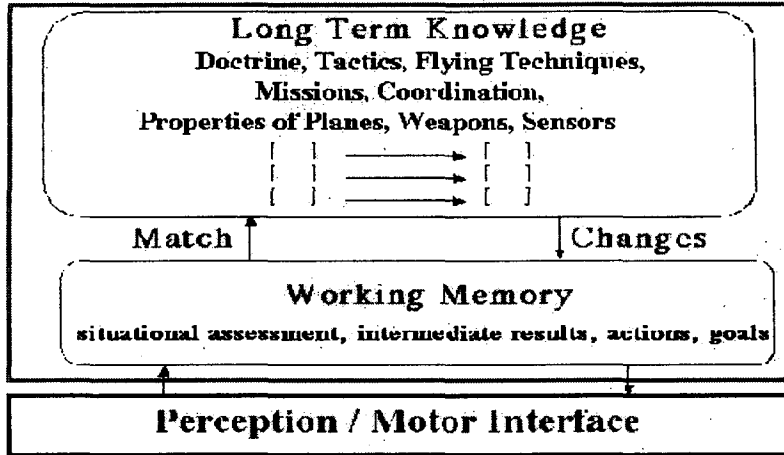
더 느슨한 형태의 규칙이라고 말할 수 있는 발견법(heuristics)에 따른다. 인지 현상을 탐구하는데 있어서 적절한 분석 수준에 대해서 마(Marr, 1982)는 세 가지 수준, 즉 계산 수준, 표상 및 알고리즘 수준, 그리고 구현 수준을 제시했다. 좌뇌 모형에 따르면, 사고언어 가설과 물리적 기호체계 가설에 대한 논의로부터 드러나듯이, 인지 현상은 마가 구분한 세 가지 수준에서 최상위 수준인 계산 수준을 분석함으로써 적절하게 설명될 수 있다.³⁾

사고언어 가설과 물리적 기호체계 가설은 구성(architecture) 개념을 통하여 물리적으로 구현될 수 있는 구체적 가설이 된다. 좌뇌 모형에서 구성 개념은 다음과 같이 설명한다. 즉, “인지 체계의 구성은 기본 작동, 자원, 기능, 원리이며, 그것들의 정의역과 치역은 해당 유기체의 표상 상태이다”(Fodor and Pylyshyn, 1988: 10). 여기서 우리는 서론에서 제시된 모든 항목들이 구성 개념에 관련되어 있음을 알 수 있다. 포도와 필리슨이 제시한 구성 개념은 좌뇌 모형을 위한 것이지만 그것을 약간 수정하면 인지 구성에 대한 일반적 정의가 될 수 있다.⁴⁾ 가장 쉬운 방법은 그들의 설명에서 “표상적 상태”라는 말을 제거하고 “인지 체계의 구성은 기본 작동, 자원, 기능, 원리이며, 그것들의 정의역과 치역은 해당 유기체의 이리저리한 상태이다”로 수정하면 일반적 규정이 된다. 좌뇌 모형을 위한 표준적 인지 구성으로서 ACT* (Anderson, 1995)와 SOAR(Newell, Rosenbloom, and Laird, 1989)를 그 예로 제시할 수 있는데 우리는 여기서 쓰어(SOAR)를 검토하기로 한다. <그림 1>에서 나타나듯이 쓰어의 특징은 인지를 지식과 지각이라는 두 가지 요소들로 구분하고, 다시 지식을 장기 지식(long term knowledge)과 단기 기억(short term memory) 또는 작업 기억(working memory)으로 구분하는데 있다.

3) 이러한 이유 때문에 AI는 종종 계산주의(computationalism)라고도 불린다.

4) 예를 들어, 공란을 “신경세포 상태”로 대체하면 위의 정의는 연결주의가 제시하는 우뇌 모형의 구성 개념이 된다.

<그림 1> Soar의 인지 구성



위의 에서 나타나듯이 쏘어의 장기 지식은 “[] → []”로 표현된 산출 규칙(production rule)들로 구성된다. 개별적 산출 규칙들은, 예를 들어 “if P, then Q (0.7)”로 표현되는데 여기서 P와 Q는 각각 산출을 위한 조건과 그 행동이고 숫자 0.7은 행동이 나타날 확률을 의미한다. 조건은 작업 지식과 대조되어 양자가 일치하면 행동이 산출되고, 이어서 더하기, 지우기, 수정하기와 같은 과정을 통하여 작업 기억의 내용이 변화된다. 물론 양자가 일치하지 않을 경우 어떠한 행동도 산출되지 않는다. 산출 규칙은 논리학의 기본적인 추론규칙 중의 하나인 전건긍정 규칙(modus ponens), “P → Q and P, therefore Q”, 과 형식적으로는 동일하지만 그 운용 방식에서 본질적으로 다르다. 즉, 전건긍정 규칙에서는 “P → Q”와 “P”가 전제인데 비하여 산출 체계에서는 “P → Q”는 규칙이고 “P”는 지각의 내용이라는 점이다. 따라서 “P → Q”라는 규칙이 작동하기 위해서는 작업 기억에 P가 있는가를 확인하는 대조가 요구되는데, 경험으로부터 유래한 P가 어떻게 명제적으로 암호화되는가 그리고 어떤 방식으로 P가 장기 지식을 구성하는 산출 규칙과 비교되는가를 설명하는 것이 좌뇌 모형이 해결해야 할 중요한 과제에 속한다.

이상의 논의를 종합하면 우리는 인지 이론으로서의 좌뇌 모형의 특징을 <표 1>로 정리할 수 있다. 심성 상태는 기호적 표상이고, 심성 과정은 그러한 표상들을 순차적으로 계산하는 과정이다. 표상을 구성하는 근본 요소는 기호이며, 기호적 표상들은 산출 규칙과 같은 명시적으로 규정된 규칙에 의해 처리된다. 한편 인지에 대한 적절한 분석 수준은 표상들이 처리되는 수준인 계산적 수준이다. 쓰어에 대한 논의에서 드러났듯이 좌뇌 모형이 갖는 특징은 인지를 전적으로 기호적 관점에서 분석하고 설명한다. 우리의 기억, 지식, 그리고 사고 과정 전체를 기호라는 '원자'로 구성되어 있다고 보고, 의미가 부여될 수 있는 단위인 '분자' 수준을 그러한 기호들로 구성된 표상으로 봄으로써 좌뇌 모형은 수학, 논리학, 그리고 철학에서 이용되어 왔던 명제 중심적인 사고 방식을 계승했다.

<표 1> 좌뇌 모형의 기본 특징

항 목	좌뇌 모형에 의한 설명
심성 상태	기호적 표상
심성 과정	순차적 계산
기본 요소	기호
규 칙	명시적 규칙 (예, 산출 규칙)
분석 수준	계산적 수준
인지 구성	기호적 구성 (예, SOAR)

3. 연결주의의 우뇌 모형

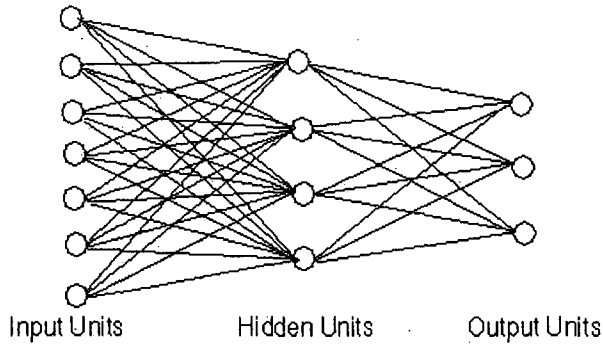
인지과학의 역사에서 볼 수 있듯이 AI는 태동 이후로부터 많은 도전을 받아왔다. AI에 대한 도전은 대체로 이론적 (또는 개념적) 도전과 경험적 (또는 기술적) 도전으로 양분된다. 우리의 관심을 이론적 도전으로 한정하더라도 프레임 문제(frame problem)로부터 시작하여 쉘의 중국어 방 논변

(Chinese room argument)에 이르기까지 진정한 인공지능의 가능성에 대한 많은 비판들이 제기되어 왔다. 그러한 비판들은 서로 다른 이론적 배경으로부터 제시되었기 때문에 그것들로부터 하나의 공통성을 찾는 것은 어려운 일이지만, AI에 대한 거의 모든 이론적 비판들은 물리적 뿌리내림의 문제(physical grounding problem)로 귀결될 수 있다. 앞 절에서 보았듯이 좌뇌 모형이 제시하는 쏘어 모형의 경우 인지의 기본 단위는 기호이고 장기 지식과 작업 기억의 내용은 기호로 이루어지는 표상들(또는 표현들)이기 때문에 그러한 표상들이 실제로 외부 세계에 있는 대상과 사건에 대한 진정한 표현인가에 의문점이 발생한다. 다시 말하면 그러한 표현들은 물리적으로 뿌리를 내리지 못하고 있다는 점에서 물리적 뿌리내림의 문제가 발생한다.

연결주의는 1950년대에 맥컬로치와 피츠(McCulloch and Pitts, 1943)에 의해 인지 이론으로서 제안되었고 그들의 이론은 로젠브라트(Rosenblatt, 1958)에 의해 퍼셉트론(perceptron) 이론으로 발전했다.⁵⁾ 한편 민스키와 페퍼트(Minsky and Papert)는 두 개의 층을 가진 퍼셉트론은 특정한 유형의 문제, 예를 들어 배타적 선언(exclusive or)과 같은 논리적 연산을 할 수 없다는 점을 지적함으로써 새롭게 형성되고 있었던 연결주의에 치명타를 가하였고, 그 결과 연결주의는 거의 20년 동안 인지 이론으로서 포기되었다. 그러나 1980년대에 들어서 연결주의는 숨겨진 층(hidden-layer)의 개념을 이용하여 배타적 선언의 문제를 해결할 수 있다는 점이 드러나면서 재등장하게 된다. 연결주의가 AI와 구별되는 일차적인 이유는 인간의 뇌는 AI가 가정하듯이 폰 노이만 방식의 컴퓨터처럼 작동하지 않는다는 점을 강조하는데 있다. 새롭게 등장한 연결주의의 주창자 중의 한 사람인 루멜하트에 따르면, 연결주의의 기본 전략은 정보 처리의 기본 단위로서 추상적인 신경 세포에 해당하는 기본 단위(unit)를 이용하는데 있다(Rumelhart, 1989: 134).

5) 로젠브라트의 퍼셉트론은 맥컬로치와 피츠가 제시한 형식적 신경세포(formal neurons) 개념에 후일에 가중치(weight) 개념으로 발전되는 강도(strength) 개념을 도입하여 발전시킨 것이다. 로젠브라트는 강도들의 변화에 의하여 퍼셉트론의 반응을 조정할 수 있음을 보였다.

<그림 2> 표준적 신경망의 구성



따라서 연결주의가 제시하는 인지 구성은 신경 세포에 대응하는 많은 단위들이 모여 하나의 신경망(neural network)을 구성한다. 신경망은 여러 가지 방식으로 구성될 수 있는데, 표준적인 신경망은 <그림 2>에서 볼 수 있듯이 세 부분, 즉 입력 단위(input units), 출력 단위(output units), 그리고 숨겨진 단위(hidden-units)로 구성되고, 그러한 단위들은 세 가지 층(입력층, 출력층, 숨겨진 층)을 형성한다. 개개의 단위들은 활성화 값(activation value)을 갖고 있으며 그 값은 함수를 통하여 출력 값을 산출한다. 또한 단위들간의 연결에서 강도는 가중치(weight)를 통하여 결정되고, 학습 규칙(learning rule)은 가중치의 변경을 통하여 망의 전체 행위를 결정한다. 신경망의 상태는 전체 단위들에 걸쳐있는 활성화 패턴을 표상하는 벡터에 의해 규정된다. 신경망은 의미론적 관점에서 다음과 같이 세 가지 종류로 구분된다. 첫 번째 종류는 국소망(localist network)인데 망의 단위들은 개념을 표상하는 것으로 해석된다. 1970년대에 주로 제시된 의미론적 망들과 최근에 제시된 타가드(Thagard, 1989)의 에코(ECHO) 프로그램은 국소망에 해당한다. 국소망은 개발자가 개개의 단위에 개념을 대응시킴으로써 망의 행위를 추적하고 해석할 수 있는 장점이 있지만 때로는 바로 그러한 점이 단점으로 작용하기도 있다. 즉 매우 많은 개별 단위들로 구성된 신경망의 경우 개발자가 개별 단위들이

표상하는 의미를 놓치는 경우가 발생하기도 하고 개별적 표상으로부터 전체적 의미를 추출하는 것이 어려운 경우가 있다. 국소망은 분산 표상이 필요하지 않은 경우에 성공적으로 적용될 수 있다. 예를 들어, 에코 망은 개별적 단위에 의해서 표상되는 개념들간의 정합성을 검사하여 입력층에 주어진 경험적이고 역사적 자료들로부터 산소의 발견과 같은 역사적인 발견 사례들을 출력한다. 정합성은 단위들의 긍정적 연결과 부정적 연결에 의해 판단되는데, 긍정적 연결은 두 단위가 같은 상태에 있는 경우이고 부정적 연결은 서로 다른 상태에 있는 경우이다.

두 번째 신경망은 분산망(distributed network)이다. 분산망에서는 개념은 개별 단위들 전체와 관련된 활성화 패턴을 나타내는 하나의 벡터에 의해 표상된다. 분산망을 구성하는데 있어서 중요한 문제는 전체 단위에 걸쳐서 표상되는 개념을 해석하는 것이다. 분산망에서의 표상에 대한 표준적인 해석은 명제적이 아니라 연상적이다. 즉, 연결주의자들은 분산망의 표상을 “그림 같은 실재”(Rumelhart, 1989: 138)이거나 그림 표상(featural representation)이라고 생각되고 있다. 분산망이 갖는 장점 중의 하나는 표상을 구성하는 일부분이 훼손되더라도 일반적으로는 망의 전체 수행에 큰 장애가 되지 않는 데 있다. 반면에 분산망은 계산적 효율성의 측면에서는 한계를 보이고 있다. 분산망은 일반적으로 역전파(back-propagation)라는 학습 규칙을 이용하는데 이 방법은 원하는 출력 값을 얻을 때까지 출력층에 도달한 표상을 감쳐진 층과 입력층으로 되돌려 보내는 과정을 되풀이한다. 그 결과 역전파에 의한 학습은 인간의 학습에 비하여 엄청나게 많은 시간을 필요로 하고, 바로 이 점에서 좌뇌 모형과 비교해볼 때 계산적으로는 매우 비효율적이다.

분산망이 갖는 이러한 문제점과 관련하여 스몰렌스키(Smolensky)는 연결주의의 기본을 유지하면서도 좌뇌 모형의 기호적 구성의 장점을 이용하여 분산망의 약점을 보완할 수 있는 융화된 연결 기호적 구성(integrated connectionists/symbolic architecture, 이하 ICSA)을 제안했다. 스몰렌스키의 ICSA는 분산망의 표상을 표현하는데 있어서 벡터 표상과 더불어 텐서곱(tensor product) 표상을 이용한다. 스몰렌스키에 따르면, 텐서곱 표상은 분

산된 활성화 패턴과 회귀 함수를 포함한 대부분의 기호적 표현을 표상할 수 있고, 비교적 간단한 신경망에 의해서도 구현이 가능하다는 장점을 갖는다. ICSA에서 표상은 좌뇌 모형에서의 기호와 분산망에서의 그림 사이에 위치하는 중간적 표상(sub-symbolic)으로 해석된다. 이상의 논의를 바탕으로 우리는 세 가지 유형의 신경망에서 나타나는 우뇌 모형의 특징을 <표 2>와 같이 정리할 수 있다.

<표 2> 우뇌 모형의 기본 특징

항 목	국소망 모형	분산망 모형	텐서곱망 모형
심성 상태	의미론적 표상	분산적 표상 (그림)	중간적 표상
심성 과정	병렬적 계산	병렬적 계산	병렬적 계산
기본 요소	개념	활성화 패턴	문장
규 칙	개별 가중치	전체 가중치	개별 가중치
분석 수준	전산 수준	신경세포 수준	중간 수준
인지 구성	의미론적 구성	병렬 분산 구성	융합된 연결기호적 구성

이제 <표 1>과 <표 2>를 이용하여 좌뇌 모형과 우뇌 모형의 이론적인 양립 가능성에 대해 생각해보자. “AI와 연결주의는 이론적으로 양립 가능한가?”라는 질문에 대한 대답은 국소망 모형, 분산망 모형, 그리고 텐서곱망 모형의 경우에 약간씩 달라진다. 국소망의 경우는 두 가지 항목에서 좌뇌 모형과 양립 가능한 대답을 제시한다. 즉 심성 상태를 의미론적 표상으로 파악함으로써 좌뇌 모형의 사고언어 가설에 기반을 둔 표상 상태와 부분적으로 일치하고, 기본 요소로서 파악된 개념은 좌뇌 모형이 제시하는 기호에 의한 표상과 일치한다. 이러한 점에서 국소망 모형은 앤더슨이 제안한 액트 스타 (ACT*)의 구성과 거의 일치한다. 반면에 분산망 모형은 모든 항목에서 좌뇌 모형과 일치하지 않기 때문에 두 모형은 양립 불가능하다고 말할 수 있으며,

이러한 결과는 일반적으로 연결주의자들의 입장을 대변한다. 또한 좌뇌 모형과 분산망 모형간의 양립 불가능성은 분산망 모형이 처칠랜드(Churchland)가 주장하는 제거적 유물론과 같은 급진적인 철학적 해석을 허용한다는 사실로부터 잘 나타난다. 처칠랜드는 역전과 학습 규칙을 갖는 분산망을 이용한 사례 연구를 통하여 통속 심리학을 제거할 것을 주장함으로써, AI의 근본 가설들, 즉 사고언어 가설, 물리적 기호체계 가설, 그리고 인지에 대한 기호적 구성 모두를 부정한다. 따라서 분산망 모형은 AI가 제시하는 좌뇌 모형과 이론적으로 양립 불가능하다. 그러나 텐서곱 모형의 경우 사정은 달라진다. 텐서곱 모형은 다음의 논의에서 나타나듯이 사고언어 가설을 부정하면서도 다른 한편으로는 표상의 생산성, 체계성, 조합성이 텐서곱망 모형에서도 가능하다고 주장하기 때문에 좌뇌 모형과 이론적으로 양립 가능할 수 있다.

좌뇌 모형과 우뇌 모형의 양립 가능성의 문제와 관련하여 포도와 필리슨(1988)은 연결주의자들에게 하나의 딜레마를 제기했다. 그들의 주장에 따르면, 연결주의가 제공하는 우뇌 모형이 좌뇌 모형과 전적으로 다르다면 그것은 인지에 대한 부적절한 이론이고 그렇지 않다면 단지 좌뇌 모형을 구현한 모형에 불과하다. 포도와 필리슨은 자신들의 주장에 대해서 두 가지 논거를 제시했다. 즉, 우뇌 모형에서의 표상은 조합적이 아니며 처리는 통계적일 뿐 구조에 민감하지 않은 그림들의 연합에 불과하다는 것이다. 일부 연결주의자들은 포도와 필리슨의 논변을 대수롭지 않게 생각하는 경향이 있지만 스몰렌스키는 그들의 주장을 연결주의에 대한 심각한 도전으로 받아들인다. 스몰렌스키가 제시한 대응책의 핵심은 분산 표상을 가진 우뇌 모형에서도 조합성이 가능하다는 점을 보이는 것이다. 스몰렌스키(1990)는 먼저 벡터합과 텐서곱이라는 수학적 개념들을 이용하여 텐서곱 변항묶음(variable binding) 개념을 제시한다.⁶⁾ n 항 벡터 v 와 w 에 대해서, 그것들의 합 ($v + w$)은 각각의

6) 스몰렌스키(1991)는 위에서 제시된 예 이외에도 “커피가 들어있는 컵”의 예를 제시했다. 스몰렌스키의 논변에 대한 비판은 포도와 맥로린(Fodor and McLaughlin, 1990)과 맥로린(1997)을 참조.

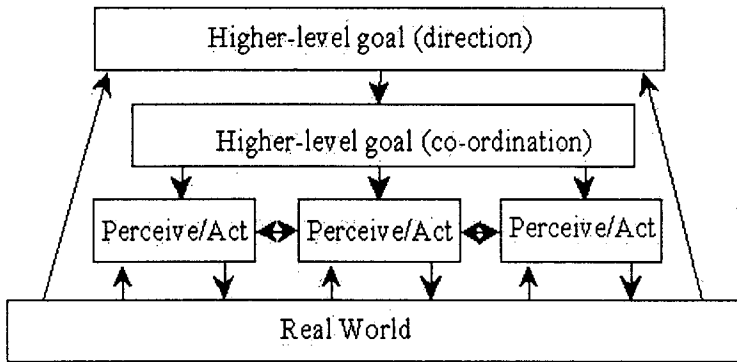
항을 더한 값을 항으로 갖는 벡터이다. 한편 n 항 벡터 v 와 m 항 벡터 w 의 텐서곱 ($n \times m$)의 요소들은 v 와 w 의 곱에 의해서 생성된다. 텐서곱 변항 묶음을 이용한 텐서곱 표상의 특징은 텐서곱을 이용하여 역할이 변화하지 않은 표상을 생성하고 벡터합을 이용하여 단순한 표상들로부터 복잡한 표상을 생성하는 것이다. F 를 충원자라고 불리는 원소들을 갖는 n 항 벡터들의 집합이라고 하고, r 은 충원자들의 역할들이 원소인 m 항 벡터라고 가정된다. 예를 들어, “샌디는 킴을 사랑한다”는 문장을 생각해 보자. 벡터 f_1, f_2, f_3 은 각각 “샌디”, “사랑한다”, “킴”을 표상하고 벡터 r_1, r_2, r_3 는 각각 그것들의 역할, 즉 주어, 동사, 목적어를 표상한다. 이 경우 세 가지 텐서곱이 있다. 즉 “ $f_1 \times r_1$ ”은 “주어의 역할을 하는 샌디”를 표상하고, “ $f_2 \times r_2$ ”는 “동사의 역할을 하는 사랑한다”를 표상하고, “ $f_3 \times r_3$ ”는 “목적어의 역할을 하는 킴”을 표상한다. 이러한 벡터들을 더하면 벡터들의 중첩 “ $(f_1 \times r_1) + (f_2 \times r_2) + (f_3 \times r_3)$ ”이 되는데, 그 중첩이 “샌디는 킴을 사랑한다”는 문장을 표상한다. 스몰렌스키는 이러한 예를 통하여 텐서곱 표상은 좌뇌 모형에서 가정되는 기호적 구성에 의존하지 않고서도 조합성을 갖는다고 주장한다. 우뇌 모형도 좌뇌 모형이 갖는 구성적 특징을 가질 수 있다는 스몰렌스키의 논변은 한 가지 중요한 점을 함축한다. 스몰렌스키의 주장이 사실이라면 우뇌 모형의 일종인 텐서곱 모형과 좌뇌 모형에서 한 가지 중요한 공통점, 즉 사고의 조합성에 대한 일치가 발생한다는 점이다. 환언하면, 두 모형은 사고의 조합성이 인지의 중요한 특징이라는 점을 전제하고 있다. 이러한 경우 우리는 두 가지 모형들을 사고의 조합성에 대한 상호 경쟁하는 이론으로 간주할 수 있고 경험적으로 그것들을 비교할 수 있을 것이다.

4. 상황화된 행동 이론의 로봇 모형

상황화된 행동 이론은 브룩스(Brooks)에 의해 제안되었다. 브룩스의 이론

의 핵심은 인지는 특정한 상황에서 나타나는 행동에 의해서 설명되어야 한다는 것이다. 브록스의 이론은 물리적 뿌리내림 가설과 비표상 가설(non-representation hypothesis)에 근거를 두고 있다. 물리적 뿌리내림 가설에 따르면, 지능적인 체계를 만들기 위해서는 표상들이 물리적 세계에 뿌리를 내려야 한다(Brooks, 1990: 5). 물리적 뿌리내림의 개념은 지능적 체계에 대한 모든 입력이 그 체계가 속한 환경으로부터 직접적으로 유래해야 한다는 것을 주장한다. 물리적 뿌리내림의 기준에 따르면, 좌뇌 모형에 기반을 둔 체계들은 모든 입력이 세계로부터 유래된 감각들이 아니라 감각에 대한 언어적 표상이기 때문에 물리적으로 세계에 뿌리를 내리지 못하고 있고 따라서 적절한 인지모형이라고 볼 수 없게 된다.

<그림 3> 인지에 대한 포섭 구성



브록스는 물리적 뿌리내림이라는 개념을 상황화된 행동 이론에 수용하기 위해서 포섭 구성(subsumption architecture)을 제안한다. <그림 3>에서 나타나는 포섭 구성의 목표는 실제 세계에서 발생하는 지각과 행동을 인지 구성의 기본 단위로 설정하여 인지를 효과적으로 설명하는 것이다. 포섭 구성에서 모든 지각과 행동은 목표 지향적이라고 가정되며 목표는 항상 실제 세

계에 의해서 규정된다. 지각과 행동의 관계는 포섭 프로그램에 의해서 설명되는데, 포섭 프로그램은 부가된 유한한 상태 기계들(augmented finite state machines, 이하 AFSM)로 구성된 망이다. 개개의 AFSM들은 등록기와 시계를 갖고 있고, 정보는 기계들을 연결하는 선에 의해서 전달된다. 매우 흥미롭게도 선에는 입력을 억제하거나 출력을 억제하는 역할을 하는 조절기가 있다. 여기서 우리는 포섭 구성과 우뇌 모형의 구성간의 유사점들을 볼 수 있다. 앞에서 보았듯이 우뇌 모형이 제시하는 신경망은 많은 단위들로 구성되고 포섭망 역시 많은 AFSM들로 구성된다. 또한 신경망의 경우 학습 규칙에 의하여 가중치를 조절함으로써 정보의 흐름을 조절하고 포섭망에서는 조절기에 의해서 정보의 흐름이 통제된다. 그러나 브룩스는 이러한 유사점들은 피상적이며 두 가지 모형은 근본적으로 다르다고 주장한다. 브룩스가 제시하는 첫 번째 차이점은 포섭망을 구성하는 AFSM들은 유한한 상태를 갖고 있으며, 특히 그것들은 독자적으로 등록기와 시계를 완벽하게 암호화한다.(Brooks, 1990: 7) 이에 비하여 신경망에서는 표상은 전체 단위들에 걸쳐서 분산되어 있기 때문에 단위들간에는 질적인 차이가 없다.

한편, 비표상 가설에 따르면 표상은 지능적 체계들을 대량 생산하는데 있어서 잘못된 단위이다. (Brooks, 1991: 140) 인지를 설명하는데 있어서 전통적인 의미에서의 표상은 쓸모가 없다는 것이다. 비표상 가설은 인지에 대한 설명에서 전통적으로 가정되어 왔던 표상 개념을 부정한다는 점에서 매우 혁명적인 사상이다. 브룩스가 주장하듯이 개개의 AFSM들이 독자적이기 때문에 우리는 그것들이 라이프니츠의 '창이 없는 단자'처럼 독자적인 표상을 갖는다고 해석할 수도 있다. 그러나 브룩스는 AFSM의 상태를 '표상'이라고 보는 관점을 거부하고, 설사 우리가 그것을 표상이라고 부르더라도 AFSM의 '표상'에는 변형이나 규칙의 선택이 없기 때문에 전통적 의미에서의 표상은 아니라고 주장한다.(Brooks, 1991: 149) 개별적 AFSM의 상태에 대한 브룩스의 설명에는 분명하지 않은 점이 많지만, 적어도 포섭 구성에는 중앙 표상(central representation)이 없다는 것은 분명하다. 이 점을 반영하여 우리는 비표상 가설을 "인지를 설명하는데 있어서 중앙 표상을 가정할 필요가 없

다”라고 표현할 수 있다. 브룩스가 주장하는 두 번째 차이점은 감각이 주어지면 그것에 적절한 AFSM들이 자동적으로 집중된다는 점이다. 예를 들어, 시각적 감각이 감각기라고 불리는 입력 모듈에 주어지면, 그것에 대한 AFSM의 그룹이 생성되고, 행동기라고 불리는 출력 모듈에 메시지를 전달하고, 행동기는 특정한 행동을 나타낸다. 마찬가지로, 포섭 체계가 다른 유형의 감각을 갖게되면, 다른 그룹의 AFSM이 생성되고 결과적으로 다른 행동이 나타나게 된다. 이상적인 포섭 구성에서는 우리의 감각기관들에 대응하는 그룹의 모듈들이 있게 된다.

브룩스의 상황화된 행동 이론은 “분산된, 구현된, 상황화된 인지”(distributed, embodied, situated cognition)에 대한 인지과학의 최근 조류를 반영한다.⁷⁾ 이러한 최근 동향은 전통적인지 이론, 특히 AI가 인지를 적절히 설명하지 못한다는 인식으로부터 출발하여 기존의 이론들이 간과한 인지 체계와 환경간의 관계를 강조한다. 브룩스의 이론에서 볼 수 있듯이 최근의 조류는 인지를 창발적이고 분산적이며 자기 조직적인 현상으로 간주하기 때문에 AI의 좌뇌 모형에서 나타나는 중앙 집중적인 구성이나 표상 개념을 채택하지 않는다. 이러한 점에서 브룩스의 이론은 AI와 연결주의의 문제점과 한계를 심각하게 고려하는 인지과학자들에게는 매우 매력적인 대안이다. 그러나 상황화된 행동 이론이 적절한 인지 모형을 제시하기에는 본질적인 문제가 있는 것 같다. 비(중앙) 표상 가설과 포섭 구성에 대한 논의로부터 우리는 상황화된 행동 이론이 정상적 인간이 아니라 동물이나 또는 유아기의 인간에 적절한 이론이라는 강한 인상을 받게 된다. 브룩스의 이론은 동물들이 갖고 있는 일차적 감각과 본능에 의한 자동적 행동에 의하여 인지를 설명하기 때문에 그것을 인지 이론이라고 받아들이더라도 상위 차원의 인지가 아니라 하위 차원의 인지를 설명하는데 적절한 이론이라고 생각된다. 실제로 브룩스의 이론은 여러 가지 종류의 로봇들을 제작하는 모형으로서 이용되었다. 인간의 인지가 브룩스가 가정하듯이 로봇과 마찬가지로 전적으로 감각에

7) 상황화된 행동 이론 이외에 역동적 체계(dynamic system)와 구현된 인지(embodied cognition)에 대한 이론이 최근의 조류를 반영한다.

의한 자동적 행동에 기반을 두고 있다면 상황화된 행동 이론은 적절한 인지 모형을 제시할 수 있지만, 인지가 감각에 기반을 둔 자동적 행동 체계라는 주장은 많은 문제점을 갖고 있다. 더구나 행동을 강조하고 표상 개념을 부정함으로써 상황화된 행동 이론은 “새로운 행동주의”(new behaviorism)가 될 가능성이 많다. 이러한 문제점들을 해결하지 못하는 한 상황화된 행동 이론은 현재로서는 인지 모형이 아니라 로봇 모형을 제시한다.

인지 현상의 중추인 인간의 뇌를 현재의 과학, 특히 신경과학의 발전 단계에서 충분히 설명하기 어렵고, 설사 그러한 지식이 가능하더라도 인지에 대한 완벽한 설명이 가능한가라는 문제가 남아있지만, 인지 현상 자체를 설명하는 것을 포기하는 이론은 인지 이론이라고 볼 수 없다. 우리는 브룩스와 마찬가지로 좌뇌 모형이 근거를 두고 있는 기호적 표상 이론을 거부할 수는 있지만 그것이 인지현상을 설명하는 것을 거부하는 것을 의미하지는 않는다. 그러므로 상황화된 행동 이론이 전통적으로 중요하게 다루어져왔던 표상 개념을 부정한다고 해서 인지 이론이 될 수 없다고 말할 수는 없다. 문제는 위에서 지적되었듯이 표상 개념을 포기함으로써 나타나는 이론적 득실 계산에서 실이 많을 것 같다는 데 있다.

<표 3> 로봇 모형의 특징

항 목	좌뇌 모형	로봇 모형
심성 상태	기호적 표상	(설명이 없음)
심성 과정	순차적 계산	분산적 계산
기본 요소	기호	행동
규 칙	산출규칙	(부정)
분석 수준	전산적 수준	구현 수준
인지 구성	기호적 구성	포섭 구성

이상의 논의를 바탕으로 우리는 상황화된 행동 이론이 제안하는 로봇 모

형의 특징을 <표 3>과 같이 정리할 수 있다. <표 3>에서 나타나듯이 로봇 모형은 좌뇌 모형과 모든 항목에서 일치하지 않기 때문에 좌뇌 모형과 양립 불가능하다. 로봇 모형은 심성 상태와 규칙을 가정하지 않기 때문에 두 모형 간의 불일치 정도는 좌뇌 모형과 우뇌 모형간의 경우 보다 더 심하다. 그러나 좌뇌 모형의 지지자들은 우뇌 모형에 대해 구현 논변을 제시했던 것과 마찬가지로 로봇 모형이 좌뇌 모형의 구현 사례라고 주장하기 때문에 우리의 판단이 결정적이지 않을 수도 있다. AI의 지지자들이 제시하는 구현 논변은 베라와 싸이먼(Vera and Simon, 1993)에 의해서 제시되었다. 베라와 싸이먼은 브룩스가 제시하는 움직임은 로봇들이 공학적인 관점에서 보면 성공적이라고 평가하면서도 브룩스의 접근이 보다 더 복잡한 문제 해결 상황에도 적용될 수 있는가에 대해서는 의문을 표시한다. 그들의 논변의 핵심은 비중앙 표상 개념은 곤충과 같은 생물체에는 적절하게 적용될 수 있겠지만 상위차원의 문제 해결에는 적용될 수 없다는 것이다. 또한 베라와 싸이먼은 중앙 표상이 없는 기호적 구성을 이용한 인공지능 체계들이 있다는 점을 지적하면서 그러한 사례들 역시 복잡한 상황에 대한 체계가 아니라는 점을 강조한다. “‘상황화된 행동’이라는 개념은 복잡한 환경에서 실시간에 적응적으로 작동하도록 특수하게 설계된 기호 체계들에 대한 이름이다”(Vera and Simon, 1993: 47). 따라서 베라와 싸이먼에 따르면, 브룩스의 이론은 AI의 좌뇌 모형의 구현 사례에 해당한다. 이러한 구현 논쟁에서 핵심은 AI는 상위차원의 인지에 관한 이론이고 상황화된 행동 이론은 하위 차원의 이론에 관한 이론인가, 만약 그렇다면 어떤 부분이 보다 더 일차적이고 본질적인 차원인가이다. 만약 그렇지 않다면, 브룩스의 이론은 하위 차원뿐만 아니라 상위차원의 인지 현상도 설명 가능하다는 입장이 된다. 이러한 논점들에 대한 논의는 아직도 진행 중이기 때문에 AI와 연결주의간의 구현 논쟁과는 달리 어떤 결론을 내리는 것은 시기상조일 것이다.

5. 결론

지금까지의 우리는 좌뇌 모형, 우뇌 모형, 로봇 모형간의 양립 가능성에 대해 제기된 세 가지 구현 논변들을 파악했다.

- ① 우뇌 모형은 좌뇌 모형의 구현 사례라는 논변(Fodor and Pylyshyn)
- ② 우뇌 모형(텐서망 모형)이 독립적인 조합성을 갖는다는 논변(Smolensky)
- ③ 로봇 모형은 좌뇌 모형의 구현 사례라는 논변(Vera and Simon)

이상의 논변들은 좌뇌 모형을 중심으로 제시되었고 좌뇌 모형이 겨냥하는 분석 수준이 최상위 수준이기 때문에 다른 모형들을 그것의 구현 사례로서 보는 구현 논변으로 전개되었다. 좌뇌 모형의 지지자들이 제시한 논변은 만약 그들의 주장이 참이라면 구현 수준의 이론으로 지목된 이론의 독립성은 보장되지 않는다는 점에서 부정적이다. 다른 한편으로 ①과 ②의 경우에서 나타나듯이 좌뇌 모형의 지지자들이 제시한 논변은 논변의 본래 의도와는 달리 구현 수준의 이론이, 구현 여부와는 관계없이, 상위 수준의 이론과 중요한 공통점을 갖는다는 점을 드러내는데 기여했다는 점에서 긍정적이다. 우뇌 모형이 사고의 조합성을 갖는다는 주장은 그 모형이 인지에 대한 “일반” 모형이 아니라 “우뇌” 모형으로 불리게 되는 중요한 이유 중 한 가지를 제거함으로써 좌뇌 모형이 인지에 대한 적절한 모형으로 발전하는데 기여한다. 인지에 대한 적절한 모형을 모색하는 작업은 이처럼 경쟁하는 이론들간의 공통성을 파악하는데서 시작될 수 있다. 그러나 ③의 경우에서는 아직 그러한 긍정적 결과가 나타나고 있지 않다.

인지에 대한 적절한 이론의 모색에서 두 번째 단계는 모형들의 양립 가능성을 바탕으로 하는 모형들간의 혼합이다. 모형들의 혼합은 다양한 방식으로 진행될 수 있다. 예를 들어, 이 글에서는 다루어지지 않았지만, 하나드(Harnad, 1990)의 경우에서 나타나듯이 좌뇌 모형과 우뇌 모형의 단순한 혼

합이 있다. 또는 로봇 모형과 좌뇌 모형, 로봇 모형과 우뇌 모형과의 양립 가능성을 바탕으로 로봇모형과 좌뇌 모형, 로봇 모형과 우뇌모형의 혼합도 가능하다. 마지막으로 모형들의 전체적인 혼합도 예상된다. 이러한 전체적 혼합은 단순한 가능성에 머물지 않고 이미 뉴웰(Newell, 1990)에 의하여 “통합적 인지 이론들”(unified theories of cognition)로서 시도되었다.

□ 참고 문헌 □

- Anderson, J. R. (1995) *The architecture of cognition*, Lawrence Erlbaum.
- Boden, M. (1990) *The philosophy of artificial intelligence*, Oxford University Press.
- Brooks, R. A. (1990) 'Elephants don't play chess', *Robotics and Autonomous Systems*, 6, pp. 3-15.
- Brooks, R. A. (1991) 'Intelligence without representation', *Artificial Intelligence*, 47, pp. 139-159.
- Christiansen, M. H. and Chater, N. (1993) 'Symbol grounding-the emperor's new theory of meaning?', *In Proceedings of the 15th annual conference of the Cognitive Science Society*. Lawrence Erlbaum: 155-160.
- Fodor, J. (1975) *The language of thought*, MIT Press.
- Fodor, J. and P. Pylyshyn (1988) 'Connectionism and cognitive architecture: A critical analysis', *Cognition* 28, pp. 3-71.
- Fodor, J. and B. P. McLaughlin(1990) 'Connectionism and the problem of systemacity: Why Smolensky's solution does not work', *Cognition* 35, pp. 183-204.

- Harnad, S. (1990) 'The symbol grounding problem', *Physica D*42: 335-346.
- Marr, D. (1982) *Vision*, Freeman.
- McCulloch, W. and W. Pitts (1965) 'A logical calculus for the idea as immanent in nervous activity', Reprinted in M. Boden(1990): 22-39.
- McLaughlin, B. P. (1997) 'Classical constituents in Smolensky's ICS architecture', in D. Chiara et al.(eds.), *Structures and norms in science*, Kluwer: 331-343.
- Newell, A. (1990) *Unified theories of cognition*, Harvard University Press.
- Newell, A., Rosenbloom, P. and Laird, J. (1989) 'Symbolic architecture for cognition', in M. Posner (1989), pp. 93-132.
- Newell, A. and H. Simon (1976) 'Computer science as empirical enquiry: symbols and search', reprinted in M. Boden.
- Rosenblatt, F. (1958) 'The perceptron: A probabilistic model for information storage and organization in the brain', *Psychological review* 65, pp. 368-408.
- Posner, M. ed. (1989) *Foundations of cognitive science*, MIT Press.
- Rumelhart, D. (1989) 'The architecture of ind: A connectionist approach', in M. Posner (1989), pp. 133-160.
- Searle, J. (1980) 'Minds, brains, and programs', reprinted in M. Boden (1990), pp. 67-88.
- Smolensky, P. (1988) 'On the proper treatment of connectionism', *Behavioral and Brain Sciences* 11, pp. 1-74.
- Smolensky, P. (1990) 'Tensor product variable and the representation of symbolic structure in connectionism', *Artificial Intelligence* 46, pp. 161-216.
- Smolensky, P. (1991) 'The constituent structure of mental states: A reply to Fodor and Pylyshyn', reprinted in T. Horgan and J. Tienson

- eds., *Connectionism and the philosophy of mind*, Kluwer, pp. 281-308.
- Smolensky, P. (1995) 'Constituent structure and an integrated connectionist /symbolic cognitive architecture', in C. Macdonald and G. Macdonald eds., *Connectionism: Debates on psychology explanation* Blackwell, pp. 223-290.
- Thagard, P. (1989) 'Explanatory coherence', *Behavioral and Brain Science* 12, pp. 435-467.
- Touretzky, D. (1993) 'The hearts of symbols: Why symbol grounding is irrelevant', in *Proceedings of the 15th annual conference of the Cognitive Science Society*, Lawrence Erlbaum, pp. 165-168.
- Vera, A. and H. Simon (1993) 'Situated action: A symbolic interpretation', *Cognitive Science* 17, pp. 7-48.

improve pro-environmental technologies. In this vein, we have the necessity of making environmental decisions and solving environmental problems by information technologies.

Since the second half of the last century, "environment" is the key word because we have the consciousness of environment strongly. As we solve human problems by making decisions of actions, we must face with environmental decisions in order to solve our environmental problems. If we have the better understanding of the nature of information and the role of information technology, and the relation of information technology and decision-making, we are able to design environmental systems and implement their optimal interfaces of environmental components. For this purpose, we are obliged to combine several useful technologies including GIS, DSS, Knowledge-based system, and artificial neural networks. Therefore the developments and cooperations of these fields in environmental decision making enables us to live in the better and comfortable surroundings in the near future.

Toward a Possibility of the Unified Model of Cognition

Rhee, Young Eui

ABSTRACT :

Models for human cognition currently discussed in cognitive science cannot be appropriate ones. The symbolic model of the traditional

artificial intelligence works for reasoning and problem-solving tasks, but doesn't fit for pattern recognition such as letter/sound cognition. Connectionism shows the contrary phenomena to those of the traditional artificial intelligence. Connectionist systems has been shown to be very strong in the tasks of pattern recognition but weak in most of logical tasks. Brooks' situated action theory denies the notion of representation which is presupposed in both the traditional artificial intelligence and connectionism and suggests a subsumption model which is based on perceptions coming from real world. However, situated action theory hasn't also been well applied to human cognition so far. In emphasizing those characteristics of models I refer those models "left-brain model", "right-brain model", and "robot model" respectively. After I examine those models in terms of substantial items of cognitions- mental state, mental procedure, basic element of cognition, rule of cognition, appropriate level of analysis, architecture of cognition, I draw three arguments of embodiment. I suggest a way of unifying those existing models by examining their theoretical compatability which is found in those arguments.

Human Genome Research and ELSI Program

Yoon, Jeong-Ro

ABSTRACT :

Social issues surrounding biotechnology are the new frontier of