

자연언어 이해의 전산적 가능성

이초식, 이영의
고려대학교 철학과

Computational Possibility of Natural-Language Understanding

Cho-sik Lee and Young-Eui Rhee
Philosophy Dept., Korea University

컴퓨터를 이용하여 자연언어를 처리하려는 연구가 진행되고 있다. 언어가 사고와 밀접한 관계에 있다는 점에서 이러한 연구가 성공한다면 인공지능의 발전과 더불어 인간의 마음에 대한 이해의 폭을 넓히게 될 것이다. 이 글은 이러한 연구와 관련하여 컴퓨터에 의한 자연언어 이해의 가능성을 다루고 있다. 먼저 그러한 이해가 불가능하다는 써얼의 비판을 시발로 해서, 써얼에 대한 라파포트의 재반박을 검토할 것이다. 라파포트는 자신의 인공지능 프로그램과 사고실험 등을 통해서 가능성을 인정한다. 그의 주장의 핵심은 컴퓨터가 자연언어를 이해하는데 있어 구문론적 이해 만으로도 충분하다는 것이다. 이러한 주장은 기호학적 관점에서 볼 때 성립될 수 없다고 비판된다. 인간이나 컴퓨터가 자연언어를 이해하기 위해서는 언어와 그 지시 대상, 그리고 언어의 사용자 간의 관계를 고려하는 기호학적 관점이 요구된다. 그결과 컴퓨터에 의한 자연언어 이해의 가능성에는 한계가 있다는 결론에 이르게 된다.

I. 자연언어 이해

인공지능의 발달과 더불어 컴퓨터를 이용하여 자연언어를 처리하려는 연구가 진행 중이다. 이러한 시도가 성공한다면 그로 인하여 발생하는 유용성은 지대할 것이다. 그러나 한편으로는 전산적 방식에 의한 자연언어 처리는 개념적으로 컴퓨터가 자연언어를 이해한다는 것을 전제로 한다는 주장이 있다. 이러한 주장에 의하면 컴퓨터가 처리 대상인 자연언어를 이해하지 못한다면 "이해 없는 처리"는 올바른 처리가 아니다. 컴퓨터에 의한 자연언어 처리와 자연언어 이해의 동일성에 관한 문제가 제기된다.

컴퓨터에 의한 자연언어 이해(natural-language understanding: NLU)의 가능성을 자연언어 이해의 전산적 가능성이라고 하자. NLU의 전산적 가능성에 대한 강력한 지지자 중의 한사람은 쉥크(Schank)이다. 그는 스크립트(script)를 이용하

여 만든 프로그램 팸(PAM: Plan Applier Mechanism)은 컴퓨터의 “이해” 가능성에 대한 좋은 증거라고 주장했다[Schank and Abelson, 1977]. 쉥크에 의하면 NLU의 전산적 가능성을 논하기 위해서는 “이해”라는 용어를 적절히 이해해야 할 필요가 있다. 그 문제와 관련하여 쉥크[Schank, 1984: 44-48]는 이해의 차원을 구분하고 있는데, 그 중에서 타인의 경험과 태도를 공유할 수 있어야만 가능한 완전한 감정이입(complete empathy) 차원의 이해는 컴퓨터의 경우 불가능하다고 지적한다. 그러나 NLU의 전산적 가능성에 대한 쉥크의 긍정적인 주장은 인지적 이해(cognitive understanding)의 차원에서 제시된다. 인지적 이해의 차원에서 작동하는 컴퓨터는 경험으로부터 학습이 가능하고 또한 추론 과정을 설명할 수 있다.

이러한 쉥크의 주장은 실제로 그러한 인지적 이해가 가능한가에 대한 문제를 제기하는데, 이는 곧 NLU의 전산적 가능성에 대한 문제에 해당한다. 써얼[Searle, 1980]은 쉥크의 주장에 대해 강력한 비판을 제기했다. 그는 중국어방(Chinese room) 논변을 통하여 컴퓨터는 구문론 만이 있을 뿐 의미론을 결여하고 있으므로 자연언어를 이해할 수 없다고 주장했다. 그러나 “강한” 인공지능을 주장하는 라파포트[Rapaport, 1988]는 이러한 써얼의 비판에 대해 NLU가 구문적으로 처리 가능하다고 논박한다. 그는 프로그램 스냅스(SNePS)와 한국어방(Korean Room) 사고 실험을 통하여 자신의 주장을 뒷받침하고 있다.

본 논문이 다루고자 하는 것은 NLU의 전산적 가능성을 중심으로 라파포트의 견해를 비판하고 하나의 대안을 제시하는 것이다. 여기서 제시되는 대안은 퍼스(Peirce)의 기호학(semiotics)을 이용한 페저[Fetzer, 1990]의 기호학적 관점인데, 그것은 써얼의 생물학적 관점이나 라파포트의 전산주의적 관점과 근본적으로 구별된다. 기호학적 관점에 따르면 자연언어에 관한 올바른 이론은 언어와 그 지시 대상, 그리고 사용자 간의 3항 관계를 포함해야 한다. 기호학적 관점에서 보면 이러한 3항 관계를 간과하는 이론은 잘못임이 드러나고 그 결과 라파포트가 주장하는 NLU의 전산적 가능성은 한계가 있다는 점이 드러난다.(본문에서 컴퓨터는 특정 프로그램을 가동하는 상태에 있는 기계를 지칭할 것이다. 이러한 용어법은 라파포트의 soft-being-run-by-the-hardware [Rapaport, 1988:81-82]와 일치한다.)

II. 라파포트와 SNePS

써얼[Searle, 1980:417-418]은 NLU의 전산적 가능성에 대한 쉥크의 주장을 비판하기 위해 중국어방이라고 불리는 다음과 같은 사고 실험을 제시했다.

(가) 중국어방

한 사람이 중국어 문자들이 담긴 바구니가 있는 방안에 갇혀 있다. 그는 영어는 알지만 중국어는 한마디도 모른다. 이제 그에게 중국어 된 질문에 완벽하게 대답할 수 있는 영어로 된 규정집이 주어진다. 이러한 상황에서 규정집에 따라서 밖에서 들어온 중국어로 된 질문에 완벽한 대답을 방밖으로 내보내면 방밖의 중국인은 방안에 있는 그가

중국어를 이해한다고 생각할 것이다.

중국어방 논변은 NLU의 전산적 가능성을 부정하기 위한 논증으로서 제시되었다. (가)의 사고실험에서 방안에 있는 사람은 외부 관찰자로 하여금 자신이 중국어를 이해하고 있다고 생각하도록 속일 수 있다는 점에서 튜링검사(Turing test)를 통과했다. 그러나 써얼이 그가 중국어를 이해하지 못한다고 주장하는 이유는 의미와 관련되어 있다. 써얼에 따르면 의미나 의미론은 단순한 기호 조작, 즉 구문론 이상의 것이다. 언어를 이해한다는 것은 기호의 의미에 대한 해석을 포함한다. 의미는 기호와 그 기호가 지향하는 대상 간의 관계이며, 이러한 지향성은 인간의 마음만이 갖고 있는 성질이다. 뇌와 똑같은 인과적 힘을 지닌 것만이 지향성을 가질 수 있으며, 인과적 힘은 뇌의 생물학적 구조에 연유한다[Searle, 1980:423]. 써얼은 인과적 힘에 대하여 정확히 밝히고 있지 않지만 컴퓨터는 생물학적 존재가 아님을 강조한다.

컴퓨터가 비생물적이라는 것을 기본 전제로 하여 NLU의 전산적 가능성을 부정하는 써얼의 입장은 컴퓨터가 지향성을 가질 수 있는 가능성을 처음부터 부정한다는 점에서 지나친 “단백질주의”라는 비판을 면하기 어렵다. 우리가 검토해야 할 문제는 비록 비생물적인 매체를 통해서라도 특정한 인지 체계가 자연언어를 이해할 수 있는 가능성이다. 따라서 고려해야 할 써얼의 논점은 다음의 명제로 요약된다.

(나) 컴퓨터의 작동은 전적으로 구문적이다.

써얼은 의미론이 기호와 대상 간의 관계에 대한 이론이며, 그러한 관계는 프로그램될 수 없다고 주장했다. 그러나 써얼[Searle, 1980:286, 298]은 의미론적 규칙을 따르는 기계는 자신의 비판을 벗어날 수 있다는 점을 인정했다. 이와 관련하여 무어[Moore, 1988: 43-44]는 써얼의 비판을 반박하기 위하여 하위차원의 구문론으로부터 상위차원의 의미론을 구성할 수 있는 가능성을 제시했다. 라파포트[Rapaport, 1988:87] 역시 구문론적 의미론(syntactic semantics)이 가능하다는 점을 보임으로써 써얼의 비판을 벗어나려고 한다. 라파포트의 전략은 다음과 같다. 그는 먼저 컴퓨터가 외부 대상에 대한 내적 표상(internal representation)을 갖는다는 점을 지적한다. 의미를 구성하는 것은 바로 이러한 내적 표상과 다른 기호와의 관계이다. 의미 문제에 대한 라파포트의 입장은 방법론적 유아론(methodological solisism)을 전제하고 있고 있다. 포더[Fodor, 1980]가 주장하는 방법론적 유아론에 의하면 사고 과정은 사고 언어(language of thought)을 이용하여 구문론적 처리를 하는 과정이다. 그러므로 인지심리학자는 방법론적 측면에서 마음이 조작하는 기호들이 표상하는 것에 대해 고려할 필요가 없고 단지 그것의 구문론적 특징만을 연구해야 한다. 이러한 입장이 현재의 인공지능 연구자들이 대부분 수용하는 입장이지만 철학적으로 논쟁의 여지가 많은 견해라는 점을 지적하고자 한다.

라파포트는 방법론적 유아론을 전제로 해서 NLU는 구문론적 처리 만으로도 충분하다는 결론을 유도한다. 그는 물론 컴퓨터와 외부 세계 사이에 인과적 연결성

이 있으며, 그러한 연결성은 일종의 “비구문론적” 의미론을 요구한다는 점을 인정한다. 그러나 비구문론적 의미론은 계산적인 측면에서는 중요성이 없는 의미론이라고 보고 있다. 의미론적으로 적절한 인과적 관계는 외부 세계와 컴퓨터의 “마음”(computer’s “mind”) 간의 관계이다. 따라서 NLU의 전산적 가능성과 관련된 라파포트의 입장은 다음의 명제로 요약된다.

(다) NLU는 순수한 구문론적 처리로 충분하다.

라파포트[Rapaport, 1986b:393-410, 1988:105-111]는 (나)의 입장을 프로그램 스냅스(SNePS: Semantic Network Processing System)을 이용하여 구현한다. 스냅스는 캐시(CASSIE: Cognitive Agent of the SNePS System) 라고 불리는 (인지자의) 마음을 모델화한 것인데, 가장 큰 특징은 컴퓨터가 입력되는 정보로부터 추론을 유도할 수 있는 변경가능한 지식 베이스(knowledge base)를 갖는다는 점이다. 스냅스는 새로운 입력이 주어지면 변경가능한 지식 베이스에 의해 이미 저장된 정보를 수정할 수 있다. 즉 스냅스는 경험으로부터 배울 수 있다. 스냅스는 기호와 대상을 다음의 방식으로 연결한다. 스냅스의 마음, 즉 캐시는 세계로부터 입력되는 감각자료에 직접적으로 반응하면서 형성한 내적 표상에 접근한다. 그러므로 스냅스의 경우 단어와 대상간의 의미론적 연결은 직접적이 아니라 표상에 의해 중재된다.

스냅스에서 단어와 개념, 그리고 신념들은 내적 의미망(internal semantical network)에 의해 표현되는데, 그 망은 내적 표상에 연결되어 있다. 내적 의미망의 구조는 명제적이다. 즉 망의 노드들이 내적 표상이나 다른 노드들에 대한 주장을 표현한다. 따라서 단어의 의미는 전체망에서의 위치에 의해 결정된다. 여기서 라파포트가 주장하는 것은 모든 개념과 단어들이 궁극적으로 내적 표상에 연결되어 있으므로 의미론적 해석은 실제로 구문론적 기호 조작에 해당한다는 것과 의미망은 순수한 구문론적 과정에 의해 구성되므로 그것을 컴퓨터에 구현할 수 있다는 것이다.

그러나 순수한 구문론적 체계에 의미망을 구현할 수 있다는 라파포트의 주장을 인정하더라도 그 결과가 NLU에 해당하는가를 질문할 수 있다. 이러한 질문에 답하기 위해 라파포트[Rapaport, 1988:114-115]는 써얼의 중국어방을 모방한 한국어방이라는 사고 실험을 제안한다.

(라) 한국어방

한국인 영문학자가 한사람 있는데, 그는 영어는 단 한마디도 알지 못하지만 그러나 셰익스피어에 대한 세계적인 권위자로서 인정받고 있다. 그는 다음과 같은 방식으로 명성을 얻었다. 즉 그는 훌륭한 셰익스피어 한국어 번역만을 읽었지만, 많은 독서와 깊은 통찰력에 의해 한국어로 셰익스피어에 대해 많은 논문을 썼다. 그의 논문은 다시 영어로 번역되어 전문적인 학술지에 실렸고 그결과 많은 명성을 얻었다.

위의 사고실험에서 초점은 한국인 교수가 셰익스피어를 “이해”했는가이다. 중국

어방 논변과 달리 그 교수의 언어 이해에 대한 문제는 발생하지 않는다. 그 교수는 분명히 영어를 이해하지 못한다. 라파포트가 주장하려는 것은 이러한 사실에도 불구하고 한국인 교수는 학자들의 공동체가 보증하듯이 어떤 것을 이해하고 있다고 보아야 한다는 점이다. 마찬가지로 중국어방에 있는 사람은 중국어는 아니지만 무엇인가를 이해하고 있다고 보아야 한다는 것이다.

썬얼과 라파포트의 사고 실험에서 중국어방의 사람과 한국어방의 교수는 각각 중국어와 영어를 이해하지 못한다. 그렇다면 그들이 이해하고 있는 것은 정확히 무엇인가? 라파포트는 한국인 교수는 셰크피어를 “이해”하고 있다고 주장하고자 한다. 그의 주장의 근거는 의미론적 이해와 구문론적 이해를 구별하는데 있다. 특정한 형식 언어를 이해하는 경우를 살펴보자[Rapaport, 1986a: 273-274]. 의미론적 이해는 의미론적 해석을 통한 이해이고, 구문론적 이해는 구문론적 규칙에 따라 형식 언어의 기호들을 직접적으로 조작함으로써 얻어지는 이해이다. 라파포트는 의미론적 해석이 이해에 별로 도움이 되지 않다는 점을 보여주기 위하여 콰인[Quine, 1969:44-45]의 주장을 예로 들면서 의미론적으로 해석하는 유일한 방식이 없다는 점을 지적한다. 다양한 해석 중 단 하나만이 의도된 해석이지만 그것을 확인할 수 없다는 것이다. 또한 라파포트[Rapaport, 1988:103]는 특정한 형식 언어에 대한 해석은 본질적으로 또다른 형식 언어에서의 모의라는 점을 지적한다. 따라서 그것은 단순한 기호 조작 이상의 것이다. 이러한 라파포트의 구분에 대해 썬얼은 구문론적 이해와 의미론적 이해는 질적으로 다르다고 주장할 것이다. 그러나 라파포트는 후자가 전자 보다 더좋은 이해이지만 그 차이는 정도의 차이에 불과하다고 보는 입장이며, 그러한 정도의 차이는 전산적으로 별로 중요한 것이 아니라고 본다.

위에서 살펴본 라파포트의 한국어방 논변은 NLU의 가능성에 대한 썬얼의 비판에 대한 반박이 될 수 없다. 그러므로 NLU의 전산적 가능성에 대한 논증이 되지 못한다. 첫째, 라파포트는 한국인 교수가 구문론적 이해를 넘어서는 이해를 하고 있다는 점을 간과하고 있다. 그의 주장과는 달리 질적인 차이가 있다. 이점은 자렌[Jahren, 1990:325]이 제시하는 다음의 예에서 분명히 나타난다. 한국인 교수가 셰크피어의 「십이야」(Twelfth Night)의 한국어 번역을 읽고 비올라의 대사 “I am not what I am”(IIIi, 141)에 대한 훌륭한 해석을 했다고 하자. 그 교수가 이것을 할 수 있는 본질적인 이유는 “I am not what I am”이라는 대사가 영어 뿐만 아니라 한국어로 표현될 수 있기 때문이라고 보아야 한다. 이는 한국어와 영어라는 구문론적인 차이를 넘어선 이해이다. 따라서 한국인 교수가 셰크피어에 대한 한국어 번역이 아니라 셰크피어를 이해한다는 라파포트의 주장은 옳다. 그러나 그 교수의 이해가 전적으로 구문론적 이해에 해당한다는 주장은 받아들이기 어렵다.

둘째, 라파포트의 주장은 인간의 자연언어 이해가 컴퓨터의 자연언어 이해와 동일하다는 점을 전제로 한다. 이러한 전제가 인정될 경우에만 그는 컴퓨터는 자연언어를 이해한다고 말할 수 있다. 그의 주장을 다음과 같이 재구성해보자.

- (1) 전 제 : 자연언어 이해와 자연언어 처리는 동일하다.
- (2) 한국어방 : 한국어방의 교수는 셰크피어를 이해한다.

(3) 결 론 : 컴퓨터는 자연언어를 이해한다.

위에서 전제 (1)은 논란의 대상이 되는 주장이며 라파포트가 증명하려고 하는 것이다. 그는 유비추리를 통해 (2)에서 (3)을 유도한다. 그런데 (3)이 성립되기 위해서는 (1)을 전제해야 한다. 위의 추리에서 라파포트는 자신의 주장을 논증함에 있어서 바로 그 논증하는 주장과 동의어에 불과한 명제를 논거로 삼고 있다. 따라서 그는 선결문제 요구의 오류(fallacy of begging the question)를 범하고 있다.

III. 기호학적 관점

앞장에서 라파포트가 의미의 문제를 구문적으로 다루려는 입장을 검토했다. 그의 방법은 스티치[Stich, 1983:186]가 지적하고 있듯이 의미론적 속성이나 그 내용에 대해 불가지론적 입장을 취하므로 이해에 대한 올바른 접근이라고 볼 수 없다. 자연언어 이해를 올바르게 다루기 위하여는 구문론과 의미론 뿐만 아니라 화용론에 대해서도 고려해야 한다.

이제 퍼스(Peirce)와 모리스(Morris)에 의해서 개발된 기호학(semiotics)을 이용하여 심성 이론(theory of mentality)을 제시하고 그것에 따라 NLU를 검토해 보기로 한다. 기호학은 기호(sign)에 관한 이론이다. 퍼스는 사고의 요소로서 단어와 문장 뿐만 아니라 시각과 청각 내용도 포함하는 이론을 개발했다. 퍼스 [Peirce, 1931-58, vol. 2:228]에 따르면, 기호는 어떤 점에서 누군가에 대해 무엇을 나타내는 것이다. 퍼스는 기호학의 세부분(pure grammar, logic proper, pure rhetoric)을 제시했는데, 그것은 모리스의 일반화된 용어법에 따르면 구문론과 의미론, 화용론에 해당한다 (아래 그림 1 참조).

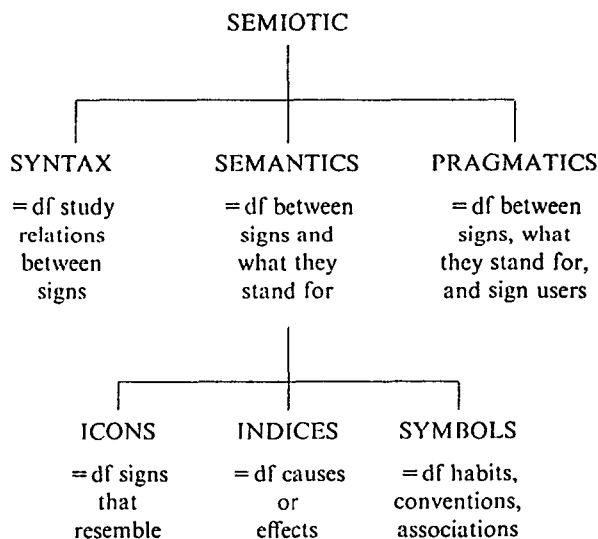


그림 1. 퍼스의 기호 이론

퍼스는 이러한 세가지 영역에 각각 해당하는 기호를 3종류로 구분하여 도합 9 종류의 기호를 구별했다. 여기서는 이 글의 목적상 의미론에 해당하는 기호만을 다루기로 한다. 퍼스는 의미론의 영역에서 기호가 대상을 나타내는 방식에 따라 도상기호(icon)와 지표기호(index), 그리고 상징기호(symbol)로 구분했다. 도상기호는 사진의 예에서 볼 수 있듯이 기호와 그 대상 간의 형태에 있어 유사성에 의해 대상에 관련된다. 지표기호는 실제적 또는 외연적 인과 관계에 의해 그 대상과 관련을 맺는다. 예를 들어 삽엽충은 고생대 캄브리아기의 지표이다. 상징기호는 임의의 규약이나 합의 또는 일반 법칙에 의해 그 대상에 관련된다. 신호등이나 자연언어의 거의 모든 단어들은 모두 상징기호에 속한다.

자연언어의 단어들이 상징기호로서 갖는 특징을 도상기호와 지표기호는 결여하고 있다는 사실을 지적함으로써 사고가 언어를 요구한다는 주장을 반박할 하나의 근거를 얻게 된다. 우리의 사고가 부분적으로 도상기호와 지표기호에 의해 이루어진다면 언어화되지 않은 사고의 가능성을 배제할 수 없다. 그러므로 라파포트가 시도하는 것과 같이 우리의 심성 내용이 모두 명제적인 의미망으로 구성되어 있다고 보는 견해는 잘못이다.

퍼스의 기호 이론은 기호(S)와 기호의 사용자(z), 그리고 그 대상(x)간의 연결성을 고려한다. 이러한 3항 관계가 아래 그림 2에 제시되어 있다.

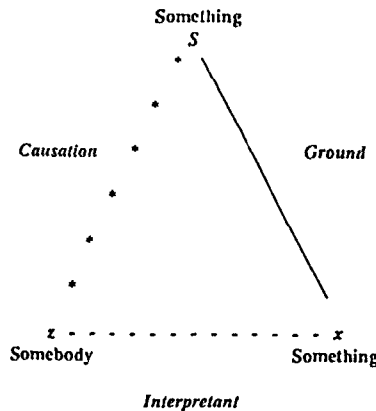


그림 2. 기호 관계

그림 2에서 기호 관계를 사용자(z)의 입장에서 보면 마음은 기호를 만들고 사용하는 체계라고 볼 수 있다. 그러므로 마음은 기호학 체계이다.

(마) 마음은 기호학 체계이다.

마음을 기호학 체계로서 간주하는 접근은 페찌[Fetzer, 1988;1989;1990]에 의해 제안되었다. 그는 퍼스의 기호학적 개념들을 수용하여 다음의 방식으로 확장한다. 즉 z에서 누군가에 대해(for somebody)를 어떤 것에 대해(for something)로 대체하는 것이다[Fetzer, 1991:65]. 이러한 대체 결과 z는 인간이 될 수도 있고 또는 다른 동물이나 기계가 될 수도 있다. 따라서 기호학 체계는 반드시 인간

이어야 한다는 주장은 성립될 수 없으며, 컴퓨터도 포함될 수 있는 가능성을 남기게 된다.

여기서 기호학 체계(semiotic system)와 물리적 기호체계(physical symbol system)의 차이에 주목할 필요가 있다. 뉴웰과 사이몬[Newell and Simon, 1976: 40]에 따르면 물리적 기호체계는 기호들(symbols)의 집합으로 구성되는데, 기호는 표현(expression) 또는 기호 구조(symbol structure)의 요소로서 발생하는 물리적 패턴이다. 물리적 기호체계는 시간의 흐름에 따라 기호 구조들의 집합이 변하는 과정을 산출하는 기계이다. 여기에서 뉴웰-사이몬과 퍼스는 “symbol”이라는 동일한 용어를 사용하고 있지만 그 의미는 전혀 다르다는 것을 알 수 있다. 뉴웰과 사이몬의 경우 기호는 임의의 방식대로 만들어지고 수정될 수 있는 것이다. 반면 퍼스의 상징기호는 규약이나 일반 법칙에 의한 것이므로 그러한 임의의 수정이 거의 불가능하다.

두가지 체계에서 나타나는 가장 큰 차이점은 물리적 기호체계는 기호들의 집합이 기계의 사용자에게 의미가 있는 경우와 기계에 의해서 사용되는데 의미가 있는 경우를 구별하지 못한다는 것이다[Fetzer, 1990: 35]. 기호학 체계와 기호체계의 결정적인 차이점이 바로 여기에 있다. 이러한 차이가 발생하는 이유는 그림 2에서 S와 x간의 관계에서 발생한다. 기호학 체계로서의 마음의 경우 S와 x간에는 근거가 있다. 앞에서 살펴 본 세가지 기호 S, 즉 도상기호와 지표기호 그리고 상징기호는 각각 그 근거에 의해서 대상 x와 관련된다. 그러나 뉴웰과 사이몬의 물리적 기호체계의 경우 임의의 수정이 가능한 것에서 알 수 있듯이 S와 x간에는 아무런 근거가 없다.

이상의 분석을 이용하여 페찌[Fetzer, 1990: 41-60]는 세가지 기호에 대응하는 세가지 심성의 유형이 있다고 주장한다. 즉 첫째, 도상기호를 사용할 수 있는 체계로서 유형 I, 둘째 도상기호와 지표기호를 사용할 수 있는 체계로서의 유형 II, 셋째 도상기호와 지표기호 그리고 상징기호를 모두 사용할 수 있는 체계로서의 유형 III이 그것이다. 이러한 유형들은 점차적으로 더 강력한 마음의 차원을 의미한다. 페찌가 제시하는 이러한 심성 이론은 실수할 수 있는 능력의 형태로서 심성 기준을 제시했다는 점에서 특이성을 지닌다. z가 실수할 수 있는 능력을 지닌다는 것은 S가 지시하는 것 x_1 이 아닌 다른 것 x_2 를 지시하는 것으로 오인하는 것이다. 이러한 오인이 가능한 이유는 z가 x_2 를 지시하기 위해 S를 취할 수 있어야 한다. 물리적 기호체계의 경우 버그(bug)는 있지만 실수는 불가능하다. 따라서 페찌는 인공지능의 목표가 인간의 마음을 모의 또는 복제하는 것이라면, 즉 기호학 체계를 만드는 것이라면 그것의 궁극적 목표는 실수할 수 있는 체계를 개발하는 것에 해당한다는 역설적인 결론을 유도한다.

마음은 위에서 제시된 것 이외에도 고차원의 유형을 갖는다. 마음은 논증을 구성하는데 있어 상징기호를 재배열하는 능력을 갖고 있는데, 이 경우 상징기호는 결론으로서의 다른 상징기호에 대한 이유나 근거를 제시하는 전제로서 작용한다. 이러한 변형적(transformational) 추론 능력을 지닌 마음은 유형 IV의 체계이다. 또한 마음은 기호를 지시하기 위하여 기호를 사용하기도 한다. 이러한 유형의 마음은 기호를 메타 기호(meta-sign)로서 사용할 수 있는 메타심성(metamentality)이며 유형 V의 체계이다. 특정 언어에 대해 언급하기 위해 사용되는 메타언어는

이러한 메타기호에 속한다.

이상에서 검토된 페저의 기호학적 관점은 인간의 마음을 기호학적 체계로 보고 다섯가지의 유형으로 구분했다. 그의 견해가 절대적인 것은 아니지만 현재 NLU에 대한 전산적 접근이 이러한 유형 중 어느 단계에 해당하는가를 검토한다면 그 답은 분명하다. 즉 물리적 기호체계 이론에 기반을 둔 이상 어느 단계에도 해당되지 않는다. 뉴웰과 사이몬의 체계는 유형 III과 유사한 체계로서 보이지만, S와 x 간의 관계에 아무런 근거가 없다는 점에서 근본적 차이가 난다. 더구나 라파포트가 주장하듯이 z와 x간의 관계를 방법론적으로 무시하려는 접근 방식은 NLU의 전산적 가능성을 어렵게 한다.

IV. 결론

이 글은 라파포트의 견해와 그에 대한 비판을 중심으로 NLU의 전산적 가능성에 한계가 있다는 점을 지적했다. 그러나 이러한 결론은 현재 연구 중이거나 이미 개발된 NLU 프로그램이 아무런 쓸모가 없다는 것을 의미하는 것은 아니다. 그러한 프로그램은 일정한 한계, 즉 명제적이라는 점 또는 구문론적이라는 점에서 유용성을 갖는다. 그러므로 “사고언어”로서의 도상기호와 지표기호등을 포함할 수 있는 접근이 요구된다고 말할 수 있을 것이다. 이 글에서는 NLU와 관련하여 연결주의(connectionism)에 대해 다루지 않았지만, 연결주의는 기호학적 관점에서 보면 학습 가능성이나 여러가지 유형의 기호 처리에 있어서 NLU에 대한 장점을 지니고 있다고 생각된다.

참고문헌

- [1] 소홍렬, 1992. 자연주의적 유신론. 서광사.
- [2] 이인식, 1992. 사람과 컴퓨터. 까치.
- [3] 이익환, 1983. 현대의미론. 민음사.
- [4] 이정민, 1992. “츄스키와 언어를 생각한다”. 정보기술 7, 184-189.
- [5] 이초식, 1991. “인공 지능과 카르납의 인공 언어”. 신일철외, 현대철학과 사회. 서광사. 1992, 41-51.
- [6] Cole, D. and Fetzer, J., 1990. Philosophy, Mind, and Cognitive Inquiry. Kluwer.
- [7] Cole, D., 1991. Artificial Intelligence and Personal Identity. Synthese 88, 399-417.
- [8] Fetzer, J., 1988. Aspects of Artificial Intelligence. Kluwer.
- [9] Fetzer, J.H., 1990. Artificial Intelligence: Its Scope and Limits. Kluwer.
- [10] Fetzer, J.H. 1991. Philosophy and Cognitive Science. Paragon House.
- [11] Fodor, J.A., Methodological Solipsism Considered as a Research

- Strategy in Cognitive Psychology. Behavioral and Brain Sciences 3, 63-109.
- [12] Jahren, N., 1990. Can Semantics be Syntactic?. Synthese 82, 309-328.
- [13] Kulas, J., 1988. Philosophy and Natural-Language Processing. In J. Kulas, J. Fetzer, and T. Rantkin. Philosophy, Language, and Artificial Intelligence. Kluwer, 1-45.
- [14] Maloney, J. C., 1985. Methodological Solipsism Reconsidered as a Research Strategy in Cognitive Psychology. Philosophy of Science 52, 451-469.
- [15] Moor, J. 1988. The Pseudorealization Fallacy and the Chinese Room Argument. in J. Fetzer (1988), 35-53.
- [16] Newell, A. and Simon, H., 1976. Computer Science as Empirical Inquiry: Symbols and Search. Reprinted in J. Haugeland (1981), Mind Design. MIT Press, 35-66.
- [17] Pearson, C., 1982. The Semiotic Science: A Semiotic Paradigm. In T. Simon, R. Scholes. Language, Mind, and Brain. Lawrence Erlbaum, 225-240.
- [18] Peirce, C. S. 1931-58. The Collected Papers of Charles Sanders Peirce: Vols. 1-6. Harvard Univ. Press.
- [19] Quine, W. 1969. Ontological Relativity. in Ontological Relativity and Other Essays. Columbia Univ. Press, 26-68.
- [20] Rapaport, W., 1986a. Searle's Experiments with Thought. Philosophy of Science 53, 271-279.
- [21] Rapaport, W., 1986b. Logical Foundations for Belief Representation. Cognitive Science 10, 371-422.
- [22] Rapaport, W., 1988. Syntactic Semantics: Foundations of Computational Natural-Language Understanding. In J. Fetzer (1988a), 81-131.
- [23] Schank, R. 1977. Scripts, Plans, Goals and Understanding. Lawrence Erlbaum
- [24] Schank, R. 1984. The Cognitive Computer. Addison-Wesley.
- [25] Searle, J., 1980. Minds, Brains, and Programs. Behavioral and Brain Sciences 3, 417-457.
- [26] Searle, J., 1990. Consciousness, Explanatory Invention, and Cognitive Science. Behavioral and Brain Sciences 13, 585-596.
- [27] Shapiro, S. and Rapaport, W., 1992. A Fully Intensional Propositional Semantic Network. in L. Burkholder, Philosophy and the Computer. Westview Press, 75-91.
- [28] Stich, S. 1983. From Folk Psychology to Cognitive Science. MIT Press.
- [29] Winograd, T., and Flores, F., 1986. Understanding Computers and Cognition. Addison-Wesley.