

인지과학과 인공지능

존 씨얼

오늘은 철학에서 가장 오래된 문제들 중의 하나에 대해 이야기 하겠습니다. 그 문제란 바로 인간이 세계와 어떻게 관계하는가 하는 것입니다. 오늘 날에 있어서 이 문제는 다음과 같은 형태로 제기되었습니다: 과학은 이 세계가 전적으로 의식도 없고 의미도 없는(mindless, meaningless) 물리적 미립자들로 이루어져 있음을 말해주는데, 그러한 세계 안에 있는 우리들은 스스로를 의식을 가지고 있고 자유로우며 이성적인 존재로 간주한다. 이러한 상식적이고 전통적인 생각 안에 있는 불일치를 어떻게 해소할 수 있을 것인가?

이 문제는 20세기 후반에 우리가 갖게 된 당혹감으로 인해 특히 어려운 문제인듯 싶은데, 그 당혹감은 자연과학의 방법들이 인간행위와 마음에 관한 탐구에 있어서는 별다른 효과를 내지 못하는 데에서 연유한 것입니다. 어쨌거나, 인간행위에 대한 설명 중에서 우리가 실제로 받아들이고 있는 것은 적어도 두가지 종류입니다. 그 하나는 일상적이고 상식적인 의미에서의 심리학적 설명인 바, 인간의 행위를 인간의 심리상태와 관련해서 설명합니다. 또다른 하나는 인간행위에 대한 신경생리학적인 설명입니다. 이와 같이 우리는 인간행위에 대해 두가지 차원의 설명을 합니다. 그런가 하면, 인간행위를 일상적인 의미의 심리현상과 관련해서 설명하는 경우에도 그 근처에 신경생리학적인 차원이 있으리라고 생각합니다. 이렇게 볼 때, 우리는 궁극적으로 이 두 차원을 관련지어 설명할 수 있어야 할 것 같습니다.

* 편집자: 이 논문은 씨얼(John R. Searle, University of California, Berkeley) 교수가 1987년 8월 7일 서울대학교 철학과 주최의 발표회에서 행한 강연의 번역이다. 강연자와 주최자의 허락 하에 여기에 게재한다. 번역은 씨얼의 심리철학에 관한 논문을 쓴 바 있는 梁寶賢(이화여대 대학원) 씨에게 위촉하였다. 게재를 가능케 한 여러분께 감사드립니다.

20세기 후반인 현재의 난점은 바로 우리가 심리학적 설명에 관해서 충분히 알지 못할 뿐만 아니라 두뇌의 활동에 관해서도 충분히 알지 못함에도 불구하고 현재의 두뇌과학이 인간을 설명하려는 데에 있습니다. 이러한 경향은 결과적으로 우리에게 의견상의 틈새, 즉 심리학적 차원의 설명과 신경생리학적 차원의 설명 사이의 틈새를 남겨주었습니다.

이 틈새를 메우기 위해서 많은 시도가 이루어졌지만 모두 실패했습니다. 나중에 말하게 되겠지만, 저는 그러한 틈새 같은 것이 있다고 믿지 않으며 틈새가 있다고 믿는 것은 그야말로 환상이라고 생각합니다. 지금 여기에서 우리는 이 틈새를 메우려는 시도 중의 하나로서 인지과학(Cognitive Science)에 대해 알아보려고 하는데, 인공지능(Artificial Intelligence)은 인지과학의 선두주자이면서 주류라고 할 수 있습니다. 인공지능에 대해서는 여러가지 다른 해석들이 있지만, 미국에서 가장 영향력 있고 강력한 해석은 바로 다음과 같은 것입니다. 즉, 심리적인 차원을 가지고 있거나 마음을 가지고 있다는 것은 잘 짜여진 컴퓨터 프로그램을 가지고 있음을 의미한다는 것입니다. 저는 이 견해를 ‘강한 인공지능’으로 명명하고 이에 대해서 제가 제시한 가장 최근의 반박을 설명함으로써 오늘의 토의를 시작하려고 합니다.

설명 강한 인공지능을 반박했다고 해도 아마 체계는 여전히 하나의 문제가 남아 있을 것입니다. 다시 말해서, 만약 강한 인공지능이 옳지 않다는 주장을 했다고 한다면, 심리학적 차원의 설명과 신경생리학적 차원의 설명 사이에 있다고 했던 틈새가 없는 이유를 또한 제시해야만 하리라는 것입니다. 이는 매우 절박한 문제입니다. 왜냐하면, 강한 인공지능에 따르기만 하면 아주 간결한 설명이 주어지기 때문입니다. 즉, 컴퓨터 프로그램이 그 하드웨어에 대해 가지고 있는 관계를 마음은 두뇌에 대해 가지고 있다는 것입니다. 더 나아가, 만약 이 등식을 거부하려면 마음과 두뇌의 관계가 어떤 것인지를 설명해내야만 합니다. 마음이 컴퓨터 프로그램과 같다면 마음은 두뇌라는 하드웨어 안에 구현되어 있다고 말하면 그만이지만, 그외의 경우에는 그렇지 않기 때문입니다. 따라서, 강한 인공지능을 반박하는 데에 덧붙여서, 마음과 두뇌의 관계는 무엇인가의 문제, 즉 여러분이 심신관계 문제로 알고 있는 문제에 대해 답변하려고 합니다.

이제 이 비오는 오후의 과제는 첫째, 인공지능을 반박하는 것이고 둘째, 심신관계문제를 살펴보는 것입니다. 그리고 만약 시간이 있다면 인공지능의 최근의 발전 결과에 대해 말하고 싶은데, 특히 인공지능에 있어서 가장 최신행태인 연결주의(Connectionism)에 관해서입니다.

지금부터 여러분에게 제시하려고 하는 주장은 이미 여러 해 전에 발표되었

던 것입니다. 그 이후로 엄청난 수의 반응들이 있었는데, 아마도 100 개 이상일 것이고 그 중의 50 개 정도는 간행되기까지 했습니다. 저의 주장은 날이 갈수록 점점 더 심한 공격을 받았습니다. 저는 어안이 병병해졌습니다. 왜냐하면, 제 생각으로는 저의 주장이 매우 간단하고 정말 결정적인 반면에 그 공격들은 상식적이고 부적합하며 그리 좋지 않기 때문입니다.

저의 논거는 이것입니다. 여러분이 컴퓨터라고 상상해 보십시오. 그리고 지금과 같은 인지능력을 위해 스스로 작동하는 프로그램이라고 상상해 보십시오. 그런 후에, 만일 스스로 작동하는 프로그램이라면 과연 적절한 인지적 이해를 가질 수 있을 것인지 자신에게 물어보십시오. 저는 이 문제에 대해 항상 스스로에게 묻는 방식을 취해 왔습니다. 제게는 이런 방식이 효과적이라고 생각되므로, 여러분들도 각자 저처럼 하신다면 좋을 것이라고 봅니다.

지금 저는 중국어를 이해하지 못합니다. 여러분이 만약 중국어를 이해한다면 아랍어나 기타 다른 언어에 대해서 생각하십시오. 제가 상상하는 것은 이것입니다. 어떤 컴퓨터 프로그래머가 중국어로 된 질문에 대답할 수 있도록 컴퓨터 프로그램을 짰습니다. 그래서 당신이 중국어로 물으면 그 컴퓨터는 옳은 대답을 합니다. 물론 당신이 중국어 질문의 형식을 취해서 입력을 하면, 입력받은 프로그램은 특정한 출력을 하도록 조작되어 있고, 그 출력은 중국어로 한 질문에 옳은 대답을 한 것입니다.

자, 이제 우리의 질문은 적절하게 프로그램된 컴퓨터는 과연 중국어를 이해하는가 하는 것입니다. 다시 말하자면, 인간의 마음이 두뇌피질 안에 있다고 말하는 것과 마찬가지로 컴퓨터의 하드웨어 안에 하나의 마음이 있다고 말할 수 있을까 하는 것입니다. 강한 인공지능은 이에 대해서 “그렇다”고 하는 반면, 저는 “결코 그렇지 않다”고 대답하겠습니다. 저의 답변을 가장 손쉽게 이해하는 방법은 당신이 컴퓨터이고 중국어 기호가 가득차 있는 바구니를 가지고 잠겨진 방안에 있다고 상상하는 것입니다. 그리하여 당신이 중국어 기호들을 입력받았고 규칙이라고 생각되는 것을 찾아내서 중국어 기호들을 출력해 내보냈다고 칩시다. 잠시 후에, 우리는 중국어 질문에 대한 당신의 대답이 중국인이 한 것처럼 훌륭하다고 생각합니다. 하지만, 저 뿐만 아니라 중국어 해독을 위해 작동되는 컴퓨터 프로그램에 의해서만은 그 누구도 단 한마디의 중국어도 이해하지 못할 것입니다. 요컨대, 만약 우리가 중국어 해독을 위해 고안된 컴퓨터 프로그램이 작동하는 것과 동일한 방식으로 중국어를 이해하지 않는다면, 그 어떤 디지털 컴퓨터도 작동하는 컴퓨터 프로그램에 의해서만은 중국어를 이해할 수 없으리라는 것입니다. 그것은 그 어떤 컴퓨터도 작동하는 그 프로그램만 가지고서는 우리가 가지지 않은 그 어떤 능력도 가지지 않을 것이기 때문입니다. 일련의 형식기호들로

☆ 존 씨 열

된 모든 프로그램은 컴퓨터의 부호지시를 조작합니다. 그 중에서도 특히 디지털 컴퓨터는 기호들의 조작장치입니다. 만약 여러분이 중국어를 이해하지 못할 경우에, 중국어 해독을 위해 프로그램을 작동하는 것과 한국어나 영어처럼 현재 여러분이 이해하고 있는 언어로 질문하거나 대답하는 것을 대조시킨다면, 우리의 논거에 대해 더 잘 알 수 있을 것입니다. 만약 지금 여러분이 이해하는 언어로 질문한다면, 여러분은 그 질문을 이해하여 대답을 할 것입니다. 주목할 점은, 설령 한국어로 한 여러분의 답변이 훌륭한 것처럼 중국어로 한 여러분의 답변 역시 훌륭하다고 해도 여기에는 분명한 차이가 있다는 사실이라고 하겠읍니다. 여러분이 저와 마찬가지로 중국어를 이해하지 못한다면, 중국어로 된 질문이나 대답을 전혀 이해하지 못했을 것이고 따라서 중국어에 관한 한 여러분은 디지털 컴퓨터와 같기 때문에 그렇습니다.

흔히 중국어 방 논의(the Chinese room argument)로 불리우는 이 논거가 가지고 있는 핵심은 우리를 어떤 낯설지 않은 결론에 이르게 하는 듯합니다. 즉, 컴퓨터는 형식기호들을 조작하는 하나의 장치이고, 의미론이 아니라 통사론에 의해 조작된다는 사실입니다. 의미론은 우리가 이해하고 있는 어떤 언어로 사고할 때에 외부로부터 들어오는 어떤 것입니다. 우리는 통사론과 더불어 의미론에 의해서라야 언어를 이해합니다.

이제 비로소 논거들이 형식적인 구조를 가지게 되었습니다. 이 논거의 구조는 매우 단순합니다.

1. 프로그램은 형식적이거나 통사론적이다.
2. 실제의 인간의 마음은 심리적 내용을 가지며 이 내용은 의미론을 가진다. 이는 언어이해의 문제이다. 그리고 중국어 방 논의가 남겨준 교훈은 통사론과 의미론은 동일하지 않다는 것이다.
3. 이로써 우리는 다음과 같은 강력한 결론을 얻게 되었다. 즉 스스로 작동하는 프로그램은 마음과 동일하지 않다. 왜냐하면 그 프로그램은 의미론을 결여하고 있기 때문이다. 요컨대, 강한 인공지능은 거짓이다. 이것이 과거 30년동안 인공지능학자들 편에 유리하도록 과장되어 만들어져서 반박되어 왔던 저의 논거입니다. 이제 저의 논거를 반박하려고 애쓰고 있는 많은 인공지능 전문가들에 대해 답변하겠습니다.

두가지 답변만 언급하겠습니다. 우선, 인공지능전문가 중의 어떤 사람들은 “중국어 기호더미가 들어있는 방안에 갇혀 있을 때에, 물론 당신은 중국어를 이해하지 못한다. 왜냐하면, 당신은 컴퓨터 본체(the central processing unit of the computer)일 뿐이고 중국어를 이해하는 것은 그 방 전체이기 때문이다” 라고 말합니다. 저는 이러한 주장에 처음 접했을 때

하도 의아해서 “당신은 그 방이 중국어를 이해한다고 말하는 것이냐?” 고 물었고 그 답변자는 “그렇다. 중국어를 이해하는 것은 그 방이다” 라고 확인했습니다. 저는 이로 인해 저의 논거가 쓸모 없어지는 것이 아닌가 두려워 했었는데, 그 까닭은 그 반론이 신비한 여운을 주었기 때문입니다. 그러나 그 반론은 저의 논거를 정말 반박하지는 못했습니다. 그 방은 결코 통사론에서 의미론으로 가는데에 저보다 더 유리하지는 않습니다. 왜냐하면, 작동하는 프로그램일 때의 저에게 있을 수 없었던 것이라면 그 무엇도 그 방에 있을 수는 없기 때문입니다.

이를 잘 이해하기 위해서, 앞서 했던 사유실험을 다음과 같이 변형시켜 봅시다. 저는 잡겨진 중국어 방에 있습니다. 그런데 문득 저는 머리 속에서 모든 조작을 다 하리라고 마음먹고는 그 프로그램을 다 외웠습니다. 그래서 규칙이 적혀 있는 책을 볼 필요도 없어졌습니다. 더 나아가서는 그 방을 제거해도 무방하게 되었습니다. 저는 문 밖으로 나와서 들관 한 가운데에 서 있습니다. 물론 이제 그 방에는 들관에 있는 저에게 없는 것 중의 어떤 것도 있지 않습니다. 그런데도 저는 중국어를 이해하지 못합니다. 그 이유는 아주 분명한데, 저는 그 기호들의 의미를 전혀 알지 못하기 때문입니다. 따라서 그 방이 중국어를 이해한다고 말하는 것은 좋은 대답이 아닙니다. 왜냐하면 그 방은 결코 통사론에서 의미론으로 가는데에 있어서 저보다 더 유리한 조건을 가지고 있지 않기 때문입니다. 이것이 저의 첫번째 답변입니다.

또 하나의 답변이 있습니다. 개인적으로는, 이 답변이 첫번째 답변보다 약하지만 우리의 논의를 계속해 나가는 데에 있어서 우리의 지적봉토에 의미심장한 그 무엇을 드러내 준다고 생각합니다. 그 답변은 이것입니다. 이제 당신이 프로그래밍 되어 있어서 “당신은 중국어를 이해하는가?”라는 질문에 대답할 수 있다고 합시다. 그리고 그 프로그램에 접근하면, “물론 나는 중국어를 이해하고 말고. 왜 그런 어리석은 질문을 하지? 내가 중국어를 이해한다는 사실을 모르겠어? 나는 중국어로 된 질문에 한주일 내내 대답해 왔어.”라고 댓구한다고 합시다. 사실 저는 앞에서 제시했던 논거를 잘 개선했는지 모르겠습니다. 하지만 제가 만들고자 하는 핵심적인 주장을 잘 표현하고 있으므로 진지하게 다룹시다. 제가 말하고 싶은 것은, 단지 기호들만 가지고는 여러분이 그 기호들의 의미나 해석을 가질 수 없다는 것입니다. 만약 여러분이 위의 예에 익숙치 않다면, 다른 예를 고안해내도 좋습니다. 어쨌거나 그 기호 역시 “당신은 중국어를 이해하느냐?”를 의미합니다. 앞의 경우와 유사한 절차에 따라서, 제가 모르는 그 기호들은 “물론 나는 중국어를 이해해. 제발 그렇게 어리석은 질문을 하지 말아줘.”라고 대답할 것입니다. 하지만, 제가 이 일을 하루종일 계속한다고 해도 중국어를 이

해할 수 있게 되지는 않겠지요. 중국어방 논의가 주는 교훈과 인간의 인식을 재현하려는 모든 연산적 시도의 실패를 받아들여야 하면, 디지털 컴퓨터의 강점은 그것이 완전히 형식적이라는 데에 있음을 알아야만 합니다. 바로 이 때문에 동일한 하나의 컴퓨터가 여러 다른 프로그램에 의해 작동될 수 있습니다. 요컨대, 이러한 특징이 컴퓨터가 가진 강점인 것입니다. 그런데 이 특징은 또 다른 특징, 즉 스스로 움직이는 기호체계가 결코 본유적 해석 (intrinsic interpretation)을 수행할 수 없다는 사실과 연관되어 있습니다. 이제 강한 인공지능을 반박하기에 충분하게 되었습니다.

다음은 심신관계 문제에 대해 답변하겠습니다. 여러분은 많은 철학자들이 가장 어렵다고 생각한 문제, 즉 정신과 육체의 관계 또는 심리적 과정과 신경 생리학적 상태나 과정 사이의 관계문제로 들어선 것입니다. 지금 저는 심신문제가 오히려 쉬운 해답을 가지고 있다는 제안을 하려고 합니다. 수수께끼는 왜 철학에서 그토록 오랫동안 심신관계 문제에 대해 명확한 해답을 찾기 위해 애썼는가 하는 것입니다. 서양철학적 전통에서 보자면, 심신관계 문제는 17세기 데카르트로부터 거슬러 올라감은 물론이고 플라톤에서까지 그 연원을 찾을 수 있습니다. 저는 유럽철학에서와 마찬가지로의 방식으로 동양철학에서도 심신관계 문제가 제기된 사실에 대해 흥미롭게 생각하고 있습니다. 언젠가 인도 봄베이에서 달라이 라마(Dalai Lama)와 함께 동일한 문제로 강의한 적이 있었습니다. 그 때, 그는 데카르트와 매우 유사한 맥락에서 심신관계를 논했습니다. 적어도 그 논의에 관한 한, 그는 데카르트주의자라고 생각됩니다. 이렇게 볼 때, 심신관계 문제에 관해서는 보편적인 어떤 것이 있는 것 같습니다.

이제 문제를 말하겠습니다. 왜 우리는 위장과 소화의 관계문제를 다루는 방식으로 마음-두뇌 관계문제를 다루지 않을까요? 그 누구도 위장과 소화의 관계문제가 심오한 형이상학적인 문제라고 생각하지 않습니다. 그 누구도 위장과 소화에 관하여 당신이 이원론자인지 일원론자인지 따지지 않습니다. 그것은 단지 신체활동에 대한 생리학적 문제이기 때문입니다. 왜 소화-위장 관계문제는 심신관계문제와 달라야 할까요? 왜 우리는 소화가 위장에 속하고 걷는 것은 다리에 속한다고 말하는 것과 마찬가지로 마음은 두뇌에 속한다고 말해서는 안되는 것일까요? 마음은 두뇌의 기능일 뿐입니다. 그런데 이 답변에는 심리현상이 몹시 수수께끼 같이 보이는 여러 특징들을 가지고 있다는 사실에 대한 해명도 포함되어 있어야 합니다. 우리는 의식을 가지고 있고 이성적입니다. 우리는 스스로를 자유롭다고 믿으며 우리의 심리현상이 주관성(subjectivity)을 가진다고 믿습니다. 또한 심리현상

들은 인과적으로 기능하는 것 같고 모두가 수수께끼로 보입니다. 그래서 우리는 심신관계문제에 대한 답은 반드시 이러한 현상들에 대한 설명을 포함하고 있어야 한다고 생각합니다. 그 수수께끼 같은 현상들에 대해 모두 말할 수는 없지만, 심신관계문제를 어렵게 만드는 그러한 현상들 중의 몇가지만 이야기하겠습니다.

문제를 어렵게 만드는 첫째 특징은 ‘의식(consciousness)’입니다. 우리는 자주 의식의 존재 때문에 놀랍니다. 누구나 우주가 의식을 가지지 않았으리라는 상상은 쉽게 할 수 있을 것입니다. 그런데 그 우주가 의식적인 두뇌(제 것, 여러분의 것, 아마도 고등동물들의 것)를 포함하고 있다는 사실은 놀라운 일입니다.

수수께끼는 극히 간단합니다. 만약 여러분이 저의 두개골 안을 들여다 본다면, 거기에서 여러분이 볼 수 있는 전부는 약 1.5킬로그램의 내용물일 것입니다. 그리고 만약 그 내용물을 검사해 보면 그것이 다수의 신경세포임을 알게 될 것입니다. 그것은 마치 오트밀처럼 생겼습니다. 이제 수수께끼는 이 1.5킬로그램의 오트밀 같은 것이 어떻게 의식적으로 될 수 있는가 하는 것입니다. 이는 참으로 경이로운 사실입니다.

마음이 가지고 있는 또 하나의 특징은 주관성(subjectivity)입니다. 저는 제가 여러분의 심리현상과 관계하는 방식과는 전혀 다른 방식으로 제 자신의 심리현상과 관계되어 있습니다. 저는 여러분이 느낄 수 있는 방법으로는 여러분의 고통을 느낄 수 없고, 여러분은 제가 느낄 수 있는 방법으로는 저의 고통을 느낄 수 없으리라는 사실은 명백합니다. 그런데 이러한 사실은 17세기 이후의 과학이 가진 신념-실재(reality)는 객관적이어야만 한다, 즉 존재하는 것은 유능하기만 하면 그 어떤 관찰자에게라도 접근가능 해야만 한다는 신념 때문에 곤란한 상황에 처했습니다. 우리가 알고 있는 것은 실제적인 현상으로 추측되기는 하지만 얼핏보기에는 고유하게 주관적인 현상의 존재입니다. 우리는 이 사실을 과학에 대한 우리의 생각과 어떻게 부합시킬 수 있겠습니까?

이제, 셋째로, 마음의 속성은 다음과 같은 것입니다. 즉 마음이 신체를 움직일 수 있다는 것입니다. 우리는 종종 이 사실을 이상하게 생각합니다. 내가 팔을 들어야겠다고 결정하면 나의 팔이 올라가는 놀라운 일이 생깁니다. 그런데 많은 철학자들과 물리학자들이 그러한 일은 불가능하다고 합니다. 아침에 먹는 콘푸로스트 같이 무게없는 나의 결정이 어떻게 물체를 움직일 수 있겠느냐는 것이지요. 하지만 저는 그런 경우가 있다고 말하겠습니다. 그 어떤 철학자도, 그 어떤 생리학자도, 그 어떤 물리학자도 나의 결정, 나의 욕구, 나의 의도가 나의 팔이 올라가는 원인이 아니라는 것을 설득할 수는 없습니다. 왜

☆ 존 씨얼

냐하면 제가 그 사실을 증명하려고 마음먹었을 때마다 팔을 올릴 수 있었기 때문입니다. 심리적인 것이 물리적인 것을 움직일 수 있다는 것이 어떻게 가능한가 하는 것은 정말 묘한 일입니다. 이렇게 볼 때, 심신관계문제에 대한 우리의 답변은 적어도 의식, 주관성, 그리고 심성적 인과성(mental causation)이라는 세가지 특성에 대해 설명하고 있어야 할 듯합니다. 물론 그 밖에도 많은 특성을 설명해내야 하겠지요. 시간관계상 그 특성들이 어떤 것인지는 다 말할 수 없지만, 이 논의 가운데에서 제기된 것 만으로도 그 특성들이 야기하는 난점들의 심자성을 알아주십시오.

자, 하나의 심리적 상태가 어떻게 위의 세가지 특성들을 가질 수 있는지를 설명하는 것이 심신관계문제에 대한 답변이 갖추어야 할 적합조건임이 드러났습니다. 이제 주목해야 할 것은 이것입니다; 모든 심리적 상태는 두뇌 안의 생리학적 과정들에서 일어나는 바, 두뇌 안에서 신경생리학적으로 진행되는 것들이 심리적 상태를 일으키기에 충분하다는 사실입니다. 물론 우리는 신경생리학에 대해서 잘 모르겠습니다만, 다소 아는 것을 이용해서 두뇌가 실제로 어떻게 심리적 상태를 일으키는가 하는 데에 대한 예를 들겠습니다.

지금 저는 목이 마릅니다. 현재의 두뇌 상태는 어떠할까요? 아마도 박약한 지식으로는 이려할 것입니다: 만약 제 몸안에 액체의 양이 부족하다면, 그리고 제 몸의 수분비율이 낮고 염분비율이 지나치게 높다면, 저의 신장에서 분비된 레닌(renin)이라고 불리는 물질이 엔자이오텐신(Enzyotensin)Ⅱ로 불리는 또 다른 물질과 합성됩니다. 엔자이오텐신Ⅱ는 단백질(peptide)의 일종입니다. 그리고 엔자이오텐신Ⅱ는 두뇌의 한 부분인 시상하부(Hypothalamus)로 들어가서 여러가지 신경 자극을 일으킵니다. 이 신경 자극들이 저에게 현재의 목마름이라는 감각을 불러일으킵니다. 요컨대, 나의 시상하부에 있는 신경들이 자극되고 있기 때문에 목마름을 느끼는 것입니다. 이제 저는 이 목마름의 경우를 모든 경우에 확대하려고 합니다. 즉 사고나 느낌과 같은 가장 놀라운 일들이 다른 신경구조와 관련되어 있는 다양한 신경 자극에 의해 일어난다고 말입니다.

지금까지의 이야기를 한 문장으로 정리해 봅시다. 기억해야 할 첫째 요점은 두뇌과정이 심리상태를 일으킨다, 즉 두뇌가 마음을 일으킨다는 사실입니다. 이제 우리에게 하나의 수수께끼가 남습니다: 그러한 심리적 과정은 무엇인가? 목마름의 상태란 무엇인가? 이것들을 어떻게 세계에 관한 포괄적인 설명 안으로 들여올 수 있을 것인가?

여기에서 저는 두번째 주장을 소개하고자 하는 바, 이제 철학적인 문제는

이러한 두개의 주장이 어떻게 둘 다 참일 수 있는가를 살피는 것이 됩니다. 첫째 주장은 ‘두뇌는 마음을 일으킨다’는 것이었고 두번째 주장은 ‘마음이나 심리적 과정은 두뇌의 상위 차원적 특성이다’는 것입니다. 다시 말하지만, 이제 문제는 이러한 두가지 주장이 둘 다 참일 수 있는 방법을 찾는 것이 되었습니다. 즉, 신경생리학적 차원의 설명과 심리적 차원의 설명 사이에 인과관계가 있다는 사실이 도대체 어떻게 참일 수 있겠는가 하는 것이지요. 또한 목마름이나 고통을 느끼는 것, 소득세를 걱정하는 것, 집에 오다가 비를 만나서 당황하는 것과 같은 심리적 상태와 두뇌과정 사이에는 일종의 동일성 관계가 성립하고 있기도 합니다. 바꿔 말해서, 모든 심리적 상태는 여러분의 신경생리학적 구조가 가진 상위차원의 특성일 뿐입니다. 그래서 만약 여러분이 인과성에 대해 어떤 특정한 개념을 가지고 있다면, 이런 종류의 인과관계와 저런 종류의 인과관계를 둘 다 받아들일 수는 없을 것 같습니다.

제가 말하고 싶은 것은 여러분이 자연의 움직임에 대해 생각한다면 그 공통점을 발견하리라는 것입니다. 하위차원의 요소들의 움직임에 의해 일어난 한 체계가 상위차원의 특성을 갖는 것과 공통점이 있다는 것이죠. 예를 들어 우리가 이 연단의 딱딱함(solidity)이나 이 물의 유동성(liquidity)에 대해 생각해 본다면, 그 외관상의 특성에 대해 설명할 수 있습니다. 또한 물의 유동성은 H_2O 분자의 차원의 운동이라는 특성에 의해 설명할 수 있고, 연단의 딱딱함은 격자구조(lattice structure)를 가진 분자들의 진동운동에 의해 설명할 수 있습니다. 물의 분자들은 다소 무질서하게 불규칙적으로 움직이는 반면, 연단의 분자들은 격자구조를 가지고 직선적으로 움직입니다. 그리고 이러한 하위차원의 것들이 유동성이나 딱딱함과 같은 상위차원의 특성들을 발생하게 합니다. 저는 여러분이 자연의 활동에서 배우기를 바랍니다.

마찬가지로, 두뇌의 하위차원의 특성들이 심리현상이라는 상위차원의 특성들을 일으키는 것입니다. 그리고 그 심리현상이라는 상위차원의 특성들은 하위차원의 특성을 가진 요소들이 구성한 하나의 체계의 특성일 뿐입니다. 물론 여기에서의 체계는 전적으로 하위차원의 요소들로 구성되어 있고 그 요소들의 활동이 상위차원의 특성을 일으키는 특정한 종류의 체계일 경우에만 한한 것이겠지요. 그 체계의 특성은 단지 상위차원의 특성들에만 존재한다는 사실을 주목해 주십시오. 그래서 저는 “이 분자가 유동성을 가졌습니다.” 또는 “이 분자들은 유동성을 가지지도 않았고 건조하지도 않습니다.”라고 말할 수 없습니다. 분자들은 젖지도 않고 건조하지도 않습니다. 요컨대 유동성 딱딱함, 촉촉함, 건조함등은 하위차원이 아닌 상위차원에서만 존재합니다. 이와 유사하게, 저는 어떤 신경세포를 가리키면서 “이것이 의식을 가지고

☆ 존 씨 열

있습니다”라고 말할 수 없습니다. 의식은 개별적인 신경세포들 보다 상위차원에 존재하기 때문입니다.

이제, 우리의 세 가지 문제를 풀어 봅시다. 저는 심신관계문제에 대해 답변을 제시했고 그 답변은 심리적인 것과 물리적인 것의 관계가 자연세계에서 보았던 관계와 공통적인 패턴이라는 사실을 아는 것이었습니다. 앞의 논의를 기억해 주십시오. 그런데, 어떤 것이 어떻게 가능한가를 아는 가장 좋은 방법은 그것이 실제로 어떻게 나타나는가, 즉 그것이 현실적으로 존재하는 방식을 아는 것이라고 하겠습니다. 지금 저의 두뇌가 의식을 가지고 있다고 말하는 것은 저의 두뇌에 관한 하나의 사실일 뿐입니다. 그래서 몇 가지 실제의 경우를 살펴봄으로써, 두뇌가 어떻게 의식을 가질 수 있는가를 설명하겠습니다.

그 대답은 여러분에게 이미 말씀드렸습니다. 목마름이라는 나의 의식적인 경험에 대해 생각해 보십시오. 목마름이라는 나의 의식적인 경험은 시상하부 안에 있는 다양한 신경자극들의 연쇄(sequences)에 의해 일어납니다. 고통이라는 나의 감각은 시상(Thalamus)과 두뇌의 또 다른 하부조직, 또는 신체 감각피질(the somatic sensory cortex)안에 있는 신경자극들의 연쇄에 의해 일어납니다. 시각적 경험과 청각적 경험 등도 마찬가지입니다. 대부분의 사고활동은 전두엽(the frontal lobes)에서 일어납니다. 물론 우리는 그것이 어떤 과정을 통해 이루어지는지 잘 알지는 못합니다. 그러나 그것이 다양한 신경자극들과 특정한 신경전달물질(neuro transmitter)이라는 분비물에 반응하는 또 다른 신경구조의 관계문제라는 것을 말할 수는 있습니다. 따라서 의식에 관한 구체적인 상세한 설명을 하는 데에는 수수께끼가 있음에도 불구하고, 그러한 수수께끼들은 우리가 사실을 잘 알고 있지 못할 경우에 생기는 것으로 과학에서는 흔히 볼 수 있는 것들입니다. 제가 말하고 싶은 것은 우리에게는 단지 과학적 무지가 있을 뿐이지 결코 철학적이거나 형이상학적인 수수께끼는 없다는 사실입니다. 마음과 신체 사이의 틈새를 연결하는 문제에 관한 수수께끼 또는 그 사이의 틈새를 극복하는 문제에 관한 수수께끼가 없는 까닭은 그 문제를 발생시킨 바로 그 틈새라는 것이 없기 때문입니다.

이제는 주관성의 문제로 넘어갑시다. 여태껏 저는 우리자신을 의식적인 존재로 간주하는 상식적인 생각이 세계가 어떻게 움직이는가에 대해 과학에서 말해주는 사실 모두와 완전히 일관된다는 것을 주장해왔습니다. 하지만 지금 우리는 하나의 의견상의 상충을 가지고 있는 듯 합니다. 과학은 실재는 궁극적으로 객관적이라고 말하고 우리는 각자 주관적인 경험을 가지고 있다고 말합니다. 경험을 가진 우리는 주관적입니다. 모든 의식상태는 어떤 사

람의 의식상태이고 하나의 주관적인 심리적 상태입니다. 그렇다면 우리들의 경험의 주관성과 17세기 이래로 과학이 가르쳐준 사실, 즉 실재의 객관성에 대한 특정가치의 개입(ideological commitment) 사이에 일어나는 상충을 어떻게 해소할 수 있겠습니까?

여기에서 저는 오늘 낱의 과학이 가진 신앙으로부터의 이탈을 제안하겠습니다. 만약 과학이 하나의 체계적인 설명을 가지고 싶다면, 이 때에 과학이 인정해야 할 사실 중의 하나는 주관성의 존재입니다. 다시 말해서 만약 주관성이 배제된 채로 과학에 정의가 내려졌다면, 그 정의는 포기되어야만 하리라는 것입니다. 하나의 사실과 하나의 정의가 상충할 때에 정의를 포기하는 것은 당연한 일이라 하겠습니다. 왜냐하면, 주관적인 심리적 경험이란 어떤 종류의 신경체계가 진화론적으로 발전된 결과이기 때문입니다. 우리의 신경체계는 주관적인 심리상태를 생산하는 바, 만약 이를 불가능하다고 주장하는 과학에 대한 정의를 가졌다면, 경험을 소유한 인간들이 존재한다는 사실로 인해 그 정의를 바꾸어야 할 것입니다. 우리는 주관성의 존재도 다른 것과 마찬가지로 하나의 객관적인 사실로서 인정해야 합니다. 이는 빠를수록 좋습니다.

이제, 마지막 수수께끼인 심성적 인과성(mental causation)의 문제로 넘어 갑시다. 저의 결정이 팔을 올라가게 함이 분명한 경우에, 하나의 심리적 사건인 저의 결정은 어떻게 저의 팔을 올라가게 할 수 있을까? 특히 우리가 데카르트적인 맥락을 따라 저의 결정을 비물질적이고 무게를 가지지 않는(weight less) 어떤 것으로 생각한다면, 어떻게 그것이 신체적 움직임을 일어나게 할 수 있는 걸까요? 이미 살펴본 답변은 “나의 결정에서 비물질적이고 무게를 가지지 않은 어떤 것이란 도대체 없다”는 것이었습니다. 그 결정은 모두 저의 두뇌에서 진행된 것이고 거기에는 두 종류의 설명이 가능했습니다. 즉 심리적 사건이 뒤이은 물리적 사건을 어떻게 일으킬까에 대한 상위차원의 설명이 있는 반면, 신경생리학적인 차원에 있는 미시적과정들이 생리학적 차원에 있는 또 다른 미시적 과정들을 어떻게 일으킬까에 대한 하위차원의 설명이 있는 것입니다. 마찬가지로, 심성적 인과성에 대해서도 또한 두 종류의 설명을 가지고 있습니다. 저의 결정 혹은 의도가 어떻게 신체적인 동작을 일어나게 할까요? 이 결정 혹은 의도는 저의 신체의 동작을 일으킵니다. 그러나 물론, 이 의도자체는 신경자극의 연쇄에 의해 일어난 것입니다. 그리고 우리는 이 의도가 두뇌 안에 실현되어(is realized) 있음을 압니다. 의도는 여러 신경세포들로 구성된 한 체계의 특성입니다. 이 신경자극들은 또 다른 생리학적 사건들을 일으키는데, 저는 이 과정에 대해서 거의 알고 있지 못

합니다. 박약한 지식에 의하면, 운동피질(the motor cortex)로부터 운동신경세포(the motor neuron)의 축삭돌기 전단부(the axon hillock)로 메시지를 전달하는 특정의 신경전달물질(neuro-transmitter)은 하나의 안착호출(a settle calling)입니다. 그 안착 호출이 근육섬유에 있는 칼슘이온을 자극하면 무슨 일이 일어날까요? 이러한 신경자극은 일련의 생리학적 변화를 일으킵니다. 그리고 이 일련의 변화는 신체적인 움직임을 일으키는데, 이때 우리는 하위차원으로 기술될 수 있을 뿐만 아니라 상위차원으로도 기술될 수 있는 동일한 하나의 사건들의 연쇄를 가지고 있습니다. 이러한 사실은 형이상학적으로 신비로운 것 같지는 않습니다. 왜냐하면, 자연세계는 그러한 종류의 관계들로 가득차 있기 때문입니다.

자동차에 대해 생각해 보세요. 여러분은 자동차의 움직임에 대해 피스톤의 차원에서 기술할 수 있을 것입니다. 탄화수소의 점화, 가솔린의 점화, 실린더에서의 공기 혼합이 피스톤에 압력을 주는 폭발을 일으킵니다. 이것이 상위차원에서의 하나의 기술입니다. 그러나 완전한 설명은 미시적차원에서도 가능합니다. 점화플러그들의 전극사이에 흐르는 전자들이 탄화수소입자들을 발화지점에까지 끌어 올리면 당신은 탄화수소입자들의 급속한 산화를 얻게 됩니다. 이렇게 보면, 동일한 하나의 현상이 상위차원의 기술과 하위차원의 기술을 둘 다 가진다는 사실에서 신비로운 점은 하나도 없습니다.

여기에서 제가 내리고자 하는 결론은 술어적이고 개념적인 영역 밖에서 한번 답변을 찾아보자는 것입니다. 이걸 무슨 뜻인가 하면, 우리는 자연세계에 대해서 생각하는 것과 마찬가지로 마음에 대해서 생각하는 버릇이 있는데 한데 여기에서 벗어나기만 하면, 심신관계문제는 위장의 소화문제가 신비로운 것 만큼만 신비로운 것이 되리라는 것입니다. 만약 우리가 전통적인 단어들을 받아들인다면, 제가 제시한 해답은 소박한 심리주의(naive mentalism, “우리의 모든 심리적 상태들은 실재하고 모두 인과적으로 기능하고 때로는 의식을 가지고 주관성을 가진다”)와 이에 대조되는 소박한 물리주의(naive physicalism, “세계는 궁극적으로 물리적이며 세계 안의 모든 것은 궁극적으로 물리적이다”)가 상충되지 않고 오히려 완벽하게 조화될 수 있음을 보여줄 것입니다. 우리가 세계가 어떻게 움직이는가에 대해 알기만 하면 심리주의와 물리주의는 둘 다 참일 수 있을 것입니다. 더 정확하게 말해서, 우리 자신에 대한 심리학적 생각과 세계에 대한 과학적·물리학적 생각은 부합하리라는 것입니다.

간단히 요약하기로 하겠습니다. 우리가 이 논의를 시작할 때에, 심리학적 설명의 차원과 신경생리학적 설명의 차원 사이에 어떤 틈새가 있다는 하나의 문제 제기에 출발점을 두었음을 기억해 주십시오. 확실히, 심리적이지

도 않고 신경생리학적이지도 않은 모종의 인간과학이 교량역할을 해야 할 것 같은 틈새가 있는 것처럼 보입니다. 그래서 여태까지는 묵시적으로 밖에 표현하지 않았지만, 이제는 그 틈새가 없다는 사실을 드러내어 말하고 싶습니다. 우리는 두뇌를 가지고 있고 실제로 어떤 의미에서는 그 두뇌에 구체화되어(embodied) 있습니다. 그 두뇌는 심리현상을 만들어내고 이 심리현상들은 두뇌의 물리적 특성 중에서 상위차원의 것들입니다. 이러한 사실은 인간에 대한 학문이 여러 다른 차원의 것으로 있을 수 있음을 말해 줍니다. 더 나아가서는, 그러한 여러 다른 차원의 학문이 연산적 모델보다 덜 풍부하리라고는 말할 수 없음을 시사합니다. 제 생각은 오히려 더 풍부하지 않겠느냐는 것입니다. 두뇌에 대해서 여러 다른 차원의 탐구가 이루어질 수 있는 것과 마찬가지로 우리의 정신적인 삶에 대해서도 여러 다른 차원의 탐구—예컨대, 개별적인 감각의 차원, 사회적이고 정치적인 활동 전반의 차원등이 이루어질 수 있기 때문입니다. 따라서 우리들의 과제는 믿음·욕구·신념·희망·전통 그리고 정열과의 연관 하에서 인간을 탐구하는 것에 대해 비과학적이라고 몰아부쳐버리는 모종의 장애물을 제거하는 데에 있습니다. 저는 이러한 방식의 탐구들이 모두 완전하게 그리고 과학적으로 나무랄데 없이 훌륭하다고 생각합니다. 모두 인간에 대한 체계적인 학문들인 것입니다. 그들 사이에는 틈새라고는 있지 않습니다. 존재하지도 않는 그 틈새를 메우기 위해서 하나의 학문을 찾아야 한다고 생각하는 것은 20세기의 환상입니다. 그 환상은 우리로 하여금 컴퓨터에 관심을 갖게 합니다.

묘하게도 자주, 저는 여태껏 마치 발전된 기술문명에 대해 반대되는 입장에서 컴퓨터를 공격하고 있는 것으로 비난 받아왔습니다. 그러나 사실, 저는 컴퓨터를 몹시 좋아합니다. 개인적인 이야기이지만, 저는 모든 작업을 컴퓨터로 해내고 있습니다. 단지 타자기나 필기도구만 가지고 있다면 얼마나 불편할까요? 저는 여러분에게 가능하다면 컴퓨터를 사용하라고 권하겠습니니다. 저는 어떠한 종류의 반전산화(anti-computational) 주장도 하지 않습니다. 저는 제 자동차를 좋아하는 것과 마찬가지로 컴퓨터를 좋아할 뿐입니다. 그러나 이러한 이유 때문에, 제 자동차가 의식을 가진다고 말할 수는 없는 일입니다. 정말 어리석은 일이겠지요. 저는 똑딱거리며 시간을 나타내는 제 손목시계를 좋아하지만, 그 시계가 영혼이나 느낌을 가지고 있다고는 생각하지 않습니다. 그것들은 단지 하나의 기계일 뿐입니다. 컴퓨터 역시 하나의 놀라운 기계일 뿐이므로, 컴퓨터에 관해 년센스를 범하는 논의들은 빨리 중지할수록 좋으리라고 생각합니다.

그러면 컴퓨터과학과 인공지능의 최근의 발전은 어떤지 살펴보겠습니다. 미국에서는 현재 매우 흥미진진한 발전들이 계속되고 있는 바, 때때로 연결주

의 또는 병행적 자료처리 (connectionism or parallel-distributed processing) 라고 불리워지는 것들입니다. 연결주의의 아이디어는 일련의 기호조합을 직선적 구조에서 했던 인공지능의 전통적인 모형과는 다릅니다. 연결주의의 모델은 상호작용하고 또 병행하는 각 마디 (node) 들의 전체 네트워크를 포함하고 있습니다. 간단히 말해서 연결주의의 모델은 그들이 입력을 가지고 출력을 할 뿐만 아니라 그들 다룰 이해할 수 있고 특정한 문제들에 대해서 전통적인 인공지능 보다 훨씬 잘 해결할 수 있다는 점에서 놀랍습니다.

이제, 연결주의의 모델에 대한 저의 평가를 말하겠는데, 두 가지로 요인될 수 있습니다. 첫째, 연결주의의 모델은 다른 종류의 모델보다 강한 인공지능을 더 잘 옹호하고 있지는 못하다는 점입니다. 즉 중국어방 논의가 다른 형태의 인공 지능에게 제기했던 반박은 연결주의에도 마찬가지로 적용된다는 것입니다. 둘째로, 연결주의의 모형이 전통적 인공지능에 비해서 과연 얼마나 다른 종류의 결과를 우리에게 주었겠는가 하는 문제입니다. 적어도 저에게 분명한 사실은 전통적 인공지능에 비해 개량되었다는 점 이외에 어떤 다른 종류의 결과도 던져주지 않았다는 것입니다. 연결주의자들의 네트워크가 하려고 하는 작업은 그들 스스로를 모형화하려는 시도로 해석될 수 있습니다. 그들 자신들도 말했듯이, 연결주의자들의 모델은 커다란 병형체계들 안에서 신경이 상호연결되는 방식에 의해 고안되었기 때문입니다. 따라서, 구성된 연결주의의 모델로 된 컴퓨터의 훌륭함은 단지 컴퓨터가 유용하다는 사실을 증명했을 뿐이라고 생각합니다. 물론 자연세계의 특정한 모습들은 흉내내고 있으므로, 여러분은 컴퓨터를 여러가지 다양한 경우에 사용할 수 있을 것입니다. 아마 어쩌면 연결주의의 모델은 두뇌과정의 모델을 설명하는 적절한 방식일지도 모릅니다. 단지 여기에서 저지르지 않으려는 실수는 뭔가 하면, 한 현상에 대해서 한 종류의 설명을 부여하는 하나의 모델 그 자체를 그 현상과 동일시하는 것입니다. 그래서 저는 여러분께 한 가지 사실을 환기시킴으로써 결론을 맺고자 합니다. 즉, 우리는 모두 진정으로 마음을 가지고 있으며 의식을 가지고 있고 또 우리의 심리상태는 의식과 주관성과 심성적인 과정을 가지고 있습니다. 그리고 컴퓨터는 우리가 마음과 두뇌 혹은 다른 것들을 연구하는 데에 사용할 수 있는 놀라운 기계장치일 뿐입니다. 감사합니다.