‘하나님의 형상의 형상(*Imago Imaginis Dei*)’으로서의 인공지능 이해: 기독론적 인간론의 관점에서

Understanding AI as ‘the Image of the Image of God(*Imago Imaginis Dei*)’ : In the Perspective of Christological Anthropology

이 경 건(서울대 과학교육과 박사과정)

논문초록

본고에서는 인공지능 기술이 인간론과 ‘하나님의 형상’에 대하여 던지는 질문에 대한 기독교적 답변을 모색하고자 하였다. 본고에서는 전통적 관점에 따라 인간의 구성 요소를 영혼과 육체라고 본다. 오늘날에는 영혼의 주요한 기능 중 하나로서의 지능에 있어서는 기호주의 및 연결주의 입장에서 시도된 인공지능이 인간에 준할 만한 유사한 구조와 기능을 보이고 있다고 할 수 있으며, 자유 의지의 문제에 있어서는 인공지능이 문제가 아니라 인간의 자유 의지가 부정되고 있는 현실이다. 인공지능에 대한 철학적 반론들은 주로 기호주의 입장에서의 초기 인공지능 프로젝트를 대상으로 하였는데, 이후 포스트-연결주의 입장에서의 인공지능 기술이 발달하고 ‘체화된 인지(embodied)’와 관련하여 인공지능의 몸에 대한 고찰이 이어지면서 비판적 논의의 갱신이 필요한 상황이다. 이에 본고에서는 성육신하신 그리스도를 통해 참된 인간성을 파악할 수 있다는 관점에서 인간이 ‘하나님의 형상(*Imago Dei*)’임에 주목하고, 인공지능을 ‘하나님의 형상의 형상(*Imago Imaginis Dei*)’으로 이해할 수 있음을 제안한다. 또한 인공지능을 비롯한 지성적인 존재들 중에서도 오직 인간만이 갖는 탁월성을 논증한다.

주제어: 인공지능, 학습, 기독론, 인간론, 하나님의 형상

Ⅰ. 서론: 인공지능, 인간, ‘하나님의 형상(*imago Dei*)’

오늘날 한국을 살아가는 우리는 흔히 ‘4차 산업 혁명’으로 대표되는 위기 레토릭에 익숙하다. 지금까지의 인류 역사에서 일어났던 산업 혁명들에 비견할 만하거나 혹은 그보다 더 큰 혁명적 변화가 눈앞에 다가왔다는 것이다. 이러한 ‘4차 산업 혁명’은 어떤 면에서는 ‘예측 불가능한’ 변화라고 이해되면서도, 한편으로는 지금 있는 수많은 직종들이 사라질 것이라는 디스토피아적 ‘예측’들과 오히려 인간의 삶이 새로워지고 윤택해질 것이라는 유토피아적 ‘예측’들을 동시에 생산해내는(e.g. 차두원, 김서현, 2016; 고영구, 이경기, 2017) 모순적인 특성을 지닌 담론이기도 하다. 그런가 하면 일각에서는 ‘4차 산업 혁명’ 담론이 갖는 허구성을 비판하는 입장을 견지하기도 한다(e.g. 김소영 외, 2017; 박문수, 2017). 이러한 반론에는 어느 정도 합당한 근거들이 없지 않으나, 우리가 처한 상황에서는 ‘4차 산업 혁명’이 있느냐 없느냐 만이 문제가 되는 것은 아니다. 만약 ‘4차 산업 혁명’이 정말로 허상에 불과한 담론이라 하더라도, 그것이 왜 많은 사람들의 위기의식을 불러일으키느냐 하는 문제를 고찰하는 일 또한 반드시 필요하기 때문이다.

이에 대답하기 위해서는 다음과 같은 두 가지 측면을 살펴보아야 한다. 첫째는 이러한 담론의 형성

배경에는 인공지능(Artificial Intelligence, AI)기술의 비약적인 발달이 자리한다는 점이다(김동환, 2016a; 2016b; 권호정 외, 2016). 이는 인공지능이 '공장 자동화', '빅 데이터', '드론' 등 '4차 산업 혁명'의 핵심 키워드들과 유기적으로 연결되어 있는 간학문적인 분야라는 사실(성태제, 2017; 이종관, 2017; 이만종, 2017)과 무관하지 않다. 이러한 인공지능 분야는 철학, 수학, 경제학, 신경과학, 심리학, 컴퓨터공학, 제어이론과 인공두뇌학(cybernetics), 언어학 등에 기반을 두고 발전해왔다는 점에서(Russel, & Norvig, 2010/2016) 현대 과학기술을 비롯한 인간 학문의 집약체라고 하여도 과언이 아니다. 둘째는 이 문제의 한 가운데에는 인공지능 기술의 발달이 인간론에 던지는 중대한 질문들(김재인, 2017; 권호정 외, 2016; 백종현, 2015)이 있다는 점이다. 호모 사피엔스(*Homo sapiens*)라는 명칭이 인간의 인간다움을 이성적인 사고에서 찾도록 하였다면, 인간과 유사한 수준의 지능을 가진 유사인종(posthomo sapiens)적 존재들의 출현으로 인하여 인간 개념의 혼란이 발생할 수 있는 것이다(백종현, 2015). '4차 산업 혁명'으로 인한 대량 실직 사태를 대비하기 위하여 '인공지능이 따라할 수 없는' 인간만의 고유한 기능들, 이를 테면 창의력이나 공감적 감수성, 협동 능력 등을 길러내는 교육이 필요하며 미래 사회의 민주성을 확보하기 위한 노력이 필요하다는 식의 논의들(e.g. 성태제, 2017; 이경민, 윤혜경, 2017; 이종관, 2017; 나중규, 김종달, 2017)은 이를 간접적으로 드러내고 있으며, 인공지능으로 인한 포스트휴먼(post-human)의 시대에는 법적인 지위의 문제가 발생할 수 있으며(손영화, 2016; 정진명, 2018) 종래의 인간중심주의(anthropocentrism)가 와해될 수도 있다는 논의들(e.g. 이봉재, 2005; 백종현, 2015; 이종관, 2017)은 이를 보다 직접적으로 다루고 있다.

이러한 두 가지 측면은 인공지능이라는 개념의 기원을 살펴볼 때 명백하게 드러나는 상호연관성을 갖는다. 인공지능을 정의하는 방법은 결코 단일하지 않지만, 일반적으로 가장 오래되었으면서도 심오한 철학적 질문을 던지는 것으로 받아들여지는 정의는 컴퓨터의 현대적 정의를 제안한 앨런 튜링(Alan Turing, 1912-1954)의 이름을 딴 '튜링 테스트'의 통과 여부라고 할 수 있다. 튜링이 1950년에 발표한 논문에 따르면(Turing, 2009), '기계가 생각할 수 있는가?(Can machines think?)'라는 질문에 대답하는 것은 '모방 게임(imitation game)'의 결과로서 대체될 수 있다. 예컨대 남자(A), 여자(B), 그리고 어떤 성별이든 관계없는 질문자(C) 각 1명이 있다고 하자. 질문자는 누가 A이고 B인지를 모르는 채로, 이를 알아내기 위해 두 사람과 떨어진 방에서 인쇄물을 통해서만, 가령을 쓰는 A와 B에게 질문을 던지고 답을 들을 수 있다. A는 C가 잘못된 판단을 내리도록 하는 것이고, B의 목적은 C가 올바로 둘을 구별하도록 돕는 것이다. 만약 기계가 A의 역할을 성공적으로 수행한다면, 이는 기능론적으로 기계가 생각할 수 있음에 대한 정의가 된다. 물론, 이러한 상징적인 사례는 인공지능 기술이 궁극적으로 성공한다고 볼 것인지에 따라 다르게 받아들여질 수 있다. 그렇다면 인공지능이 오늘날 도달한 수준은 어느 정도인가? 여기서 필자는 이미 2014년에 스스로를 Eugene Goostman이라고 지칭하는 인공지능이 영국 왕립학회(Royal Society)에서 주관한 튜링 테스트를 통과한 바 있음(Warwick, & Shah, 2016)에 주목하고 싶다. 그런가 하면 한국에서는 2016년에 구글(Google)이 개발한 인공지능 알파고(AlphaGo)가 이세돌 九단과의 대국에서 4:1로 완승을 거둔 사례가 상징적인 의미를 지닌다고 할 수 있을 것이다. 이러한 단계적 성공들 또한 관점에 따라 달리 해석될 수 있겠지만, 극한적인 상황을 가정하고 그로부터 치열한 성찰을 수행하는 일이 철학적 사고의 임무라고 할 때, 인공지능이 인간론에 대한 질문에 답하는 일을 더 이상 늦출 수는 없을 것이다.

인간은 무엇이며 인공지능은 무엇인가? 인간과 인공지능이 어떤 면에서 같고 어떤 면에서 다르기 때문인가? 인간과 인공지능이 함께 존재하는 사회를 우리는 어떠한 태도로 준비해야 하는가? 이러한 질문들에 대하여, 특별히 기독교 세계관에 입각한 비평적 답변이 반드시 이루어져야 할 필요가 있다. 왜냐하면 전통적으로 인간을 인간답게 만든다고 생각되었던 주요한 원리 중 하나가 인간이 '하나님의 형

상(*imago Dei*)'으로 창조되었다는 성경의 서술(창 1:26-27)이었기 때문이다. 곧, 인공지능이 인간론에 대하여 던지는 질문은 '하나님의 형상'이 무엇인가에 관한 질문과 직결되는 것이기도 하다(김동환, 2016b). 인간이 '하나님의 형상'으로 지어졌다면 인간과 유사한 행동을 보이는 인공지능 또한 '하나님의 형상'을 획득한 것이라고 보아야 할 것인가? 만약 그렇다면 '하나님의 형상'으로 지어졌기 때문에 지금까지 특별한 존재론적 지위를 부여받았던 인간이 인공지능에 의하여 그 자리를 빼앗기게 될 것인가? 이는 중대한 신학적 도전이 아닐 수 없으며, 지금까지 이루어진 인공지능에 대한 답변들도 '하나님의 형상'을 고찰한 측면이 작지 않았다(e.g. Herzfeld, 2002; Burdett, 2015; 김동환, 2016a; 2016b). 선행 문헌들은 많은 경우 '하나님의 형상'의 실체론적, 기능론적, 관계론적 측면을 심도 있게 검토하기도 하였지만 '하나님의 형상'에 대한 수호적인 입장보다는 개념의 의미 변화를 고려하며 인간과 인공지능의 어떠함을 상대적으로 비교하는 소극적인 입장을 견지한 경우가 많았고, 혹자들은 '하나님의 형상'을 인간의 고유한 본질로 삼는 이해가 변화해야 한다는 주장들을 펼치기도 하였다(e.g. Foerst, 1998; Putz, 2009). 결국, 지금까지 '하나님의 형상'이 인간을 인간답게 한다는 전통적인 관점을 유지하면서 인간과 인공지능을 적극적으로 이해할 수 있는 길을 모색하려는 시도는 비교적 많지 않았던 것으로 보인다.

후크마(Anthony A. Hoekema, 1913-1988)는 '그리스도가 하나님의 완벽한 형상'이므로 '하나님의 형상이 무엇인지를 배우는 가장 좋은 방법'은 '예수 그리스도를 봄으로써 무엇이 하나님의 형상인지 아는' 것이라고 하였다(Hoekema, 1986/2012, p. 40). 이에 본고에서는 인공지능이 인간론 및 '하나님의 형상'에 대하여 던지는 질문에 대하여 기독론적 인간론(Christological anthropology)의 관점에서의 답변을 시도해보고자 한다. 여기서 기독론적 인간론이라 함은 그리스도의 성육신을 삼위일체의 제2위 하나님이 온전한 인성을 취하신 사건으로 이해함으로써 인간의 인간 됨을 그리스도의 인격과 인성 안에서 이해하려는 관점을 의미한다. 그리스도의 성육신이 인간의 구속을 위한 사건이었으므로, 인간론은 창조-타락-구속에 따른 '하나님의 형상'으로서의 존재론적 지위 변화와 함께 논해져야 한다(이상웅, 2016).[1] 또한 본고에서 살피고자 하는 신학적 견해들은 기독론은 성경, 신경(creed), 교부들의 견해, 그리고 칼빈(John Calvin, 1509-1564)을 필두로 한 개혁주의적 견해들을 중심으로 한정할 것이다. 예컨대 칼빈(1559/1988)은 '인간의 어느 부분에도' 하나님의 형상의 '광채의 얼마가 빛나지 않는 곳은 없다'고 하면서도, 영혼과 육체를 명확히 구별되는 두 실체로 이해하여 '하나님의 형상의 주요한 좌소'로서의 영혼, 그리고 '영혼이 거주하는 집'으로서의 육체를 논한다(문병호, 2013, pp. 88-90). 이를 따라 본고의 Ⅱ장에서는 영혼의 문제를 다룬다. 그 중에서도 지능(*intellectus*)의 문제를 학습 이론과 관련하여, 의지와 관련된 문제를 자유 의지 및 의식과 관련하여 살펴볼 것이다. Ⅲ장에서는 Ⅱ장에서 다룬 바와 같은 인공지능 구현에 대한 철학적 반론들을 검토하고, 인공지능의 몸과 관련된 문제를 다룬다. Ⅳ장에서는 먼저 구속사적-구원론적 관점에서 그리스도와 '하나님의 형상'의 관계를 논할 뿐 아니라 성육신하신 그리스도에게서 발견할 수 있는 인간 본질의 단초를 고찰한다. 이후에 사람을 '하나님의 형상'으로, 인공지능을 '하나님의 형상의 형상'으로 이해하는 새로운 시각을 제안할 것이다. 마지막으로 Ⅴ장에서는 본고의 논의들을 요약 정리할 뿐 아니라, 인공지능을 비롯한 지성적 존재들 가운데 인간만이 갖는 고유한

1) Cortez (2016) 또한 자신의 'Christological Anthropology'를 전개하였다. 그는 역사적으로 닛사의 그레고리(Gregory of Nyssa)의 성육신에 대한 초점, 줄리안(Julian)의 십자가상의 사랑에 대한 초점, 루터(Luther)의 믿음으로 우리 것이 되는 그리스도의 의(義)에 대한 초점, 슐라이어마허(Schleiermacher)의 예수의 특별한 신-의식에 대한 초점, 바르트(Barth)의 선택의 교리, 지지울라스(Zizioulas)의 온전한 인간성의 중보에 대한 초점, 콘(Cone)의 예수의 억압받는 자들을 대신한 해방적 행위에 대한 초점을 두루 검토한다. 그러나 Cortez (2016)의 경우 기독론적 인간론을 전개함에 있어서 '하나님의 형상'을 중심으로 한 구속사적-구원론적 관점을 일관적으로 견지한다고 보기는 어려울 듯하다.

존재론적 지위가 그리스도 안에서 발견될 수 있음을 다시 한 번 확인하고자 한다.

Ⅱ. 영혼: 지능과 자유 의지

"예수는 지혜와 키가 자라가며 하나님과 사람에게 더욱 사랑스러워 가시더라 (눅 2:52)"

누가복음 2장 52절은 기독교적 전인교육의 목표와도 같은 구절이다(e.g. Comenius, 1657/2015, 10.15). 한 인격으로 계시는 그리스도 안에 있는 신성과 인성의 두 본성 중, 신성은 불변성을 지니므로 이 구절에서 말하는 자라감은 온전히 인성에 돌려지는 것이다(Calvin, 1536/2009, 2.13). 인간을 다른 동물들과 구별되게 만드는 특징 중 하나는 미숙한 상태에서의 출산이라고도 할 수 있다 이는 다시 말해 기독교적으로 인간의 전인격적인 자라감이 인간을 인간답게 만드는 특징 중 하나로 이해되어 왔음에 다름 아니다(e.g. 나현규, 2015). 한편 이 구절은 그리스도께서 취하신 인성을 각각 지혜와 키로 상징되는 영혼과 육체의 두 부분으로 이해할 수 있는 단초를 제공하기도 한다. 현대로 올수록 영혼과 육체를 구별된 실체로 파악하는 입장이 거부되고 있지만, 신약성경은 여실히 인간을 영혼과 육체로서 이해하고 있으며 그 기원은 헬라적 이원론이 아닌 유대 묵시문학과 구약성경으로 볼 수 있다(성종현, 2012). 그런가 하면 영혼(≒정신≒마음)의 구성 요소로서 지·정·의를 상정하는 것은 전통적인 철학에서 전혀 어색하지 않은 일이며(백종현, 2007). 칼빈은 영혼의 근본적인 기능으로서의 지능 및 의지에 주목한다(Calvin, 1559/1988, Ⅰ.15.7).

1. 학습 이론으로 본 지능(*intellectus*)[2)]

이러한 면에서 볼 때, 인간론에 있어서 지능의 성장으로서의 학습을 살펴보는 일이 의미가 있다는 점(허희옥 외, 2017)을 어렵지 않게 생각할 수 있다. 한편 오늘날의 인공지능 프로젝트는 본래 그 하위 분야에 해당하였던 머신 러닝(machine learning), 혹은 이를 번역하여 기계 학습이라고 불리는 기술이 발달함에 따라 그 성능의 획기적인 전환이 이루어졌으며 현재는 오히려 이에 의하여 주도되고 있는 측면이 있다(조태호, 2017; 권호정 외, 2016; Domingos, 2015). 이에 본고에서는 '학습 이론'이라는 용어를 현대 교육학에서 다루어지는 인간 학습에 관한 이론뿐만 아니라 머신 러닝 이론 또한 포괄하는 의미로서 사용하고, 양자를 비교하는 작업을 수행하고자 한다. 인간 학습에 관한 이론들은 실로 다양하지만, 현대 교육학에서 주로 논의되는 학습 이론들은 행동주의(behaviorism), 인지주의(cognitivism), 구성주의(constructionism)라는 큰 틀에서 이해할 수 있으며(박성익 외, 2012; 나일주, 2010; Eggen, & Kauchak, 2010/2011), 이를 인간의 마음(mind)과 인공지능과의 관계를 중심으로 다음과 같이 살펴볼 수 있을 것이다.

(1) 행동주의와 연결주의

2) 칼빈(1559/1988, Ⅰ.15.7)이 본래 사용한 용어는 *intellectus*로서, '지성' 혹은 '오성(五性)' 등으로 번역될 수 있다. 이 경우 *intellectus*는 사실성 이성(*ratio*)와 거의 동일한 용어로 사용된다(문병호, 2013, p. 90). 본고에서는 *intellectus*가 일반적으로 지능을 나타낼 때 쓰이는 'intelligence'와 거의 같은 의미라고 보고 논지를 전개한다.

행동주의는 파블로프(Pavlov, 1849-1936)의 고전적 조건화 실험 및 스키너(Skinner, 1904-1990)의 조작적 조건화 실험으로 대표되는 관점으로서 인간이나 동물의 심리는 객관적으로 관찰 및 정량화될 수 있는 행동을 살펴봄으로써 연구할 수 있다는 입장이다. 일반적으로 잘 알려진 파블로프의 개 실험 이외에도, 스키너는 흰 쥐를 학습시킬 수 있는 '스키너의 상자'를 고안하였다. 예컨대, 굶주린 쥐를 방음 처리된 상자 안에 넣어두고 쥐가 우연히 상자 안의 지렛대를 누를 때마다 먹이가 나오도록 장치를 마련한다면 쥐는 시행착오를 통해 배가 고플 때마다 지렛대를 누르는 행동을 학습한다는 것이다. 이는 쉽게 말해 강화와 처벌을 적절히 이용하면, 보다 복잡한 조건 하에서도 원하는 행동을 하도록 대상을 학습시킬 수 있다는 의미이다. 여기서 스키너는 보다 급진적인 행동주의를 주장하며 주체적인 '마음(mind)'은 없다는 유물론적 입장으로까지 나아간다. 그의 저서 중 하나의 제목이 「자유와 존엄성을 넘어서(*Beyond Freedom and Dignity*)」(1971)였다는 사실은 이러한 관점의 결국을 잘 보여주는 사례라고 하겠다. 이 경우, 영혼은 비물리적 실체가 아닌 단순한 심리학적 구인(construct)이며 학습을 비롯한 인간의 정신 활동은 유물론적으로 환원되어 이해된다. 이러한 행동주의는 우선 머신 러닝의 큰 줄기 중 하나인 강화 학습(reinforced learning)과 연관이 있다고 볼 수 있다. 곧, 행동에 따른 보상을 조절함으로써 학습자가 바람직한 행동만을 하도록 하는 것이다. 인공지능의 경우 보상은 마치 점수와도 같이 수치적으로 주어지고, 인간의 경우 도파민과 같은 신경 물질에 의하여 주어질 수 있다.

한편 행동주의는 인공신경망(Artificial Neural Network, ANN)을 활용한 머신 러닝을 지지하는 이른바 '연결주의(connectionism)'와도 어느 정도의 연관성을 지니고 있다. 연결주의란 직관을 비롯하여 언뜻 복잡하게 보이는 인간(생물)의 활동도 결국엔 넓은 의미에서의 자극-반응 시스템과 같이 작은 신경(퍼셉트론)들의 복잡한 연결 및 그 안에서의 연결 가중치 등으로 환원될 수 있다는 믿음을 의미한다. 이러한 연결주의와 행동주의는 적어도 학습에 있어서 선천설을 강하게 거부한다는 점(Graham, & George, 2017), 그리고 생물학적 환원주의의 경향을 보인다는 점, 지능적 행위의 근본 모델을 '블랙박스(blackbox)'로 비유하게 된다는 점 등에서 유사하다고 할 수 있다. 쏜다이크(Thorndike, 1874-1949)는 넓은 의미에서의 행동주의자이자 연결주의자로서, 두 관점이 어떻게 연관될 수 있는지를 보여주는 사례가 된다고 할 수 있는 인물이다. 인공지능 구현에 인공신경망을 활용하려는 계획은 초기 인공지능 연구인 McCulloch, & Pitts (1943) 이후로 꾸준히 이어져 왔지만, 그러한 노력이 결실을 거두기 시작한 것은 1986년에 신경망의 오류를 순방향이 아닌 역방향으로 전파하는 알고리즘이 개발되어 다층 신경망을 학습시킬 수 있게 된 이후의 일이다. 알파고를 구현한 기술 중 하나이자 최근 각광을 받고 있는 딥 러닝(deep learning)은 이러한 인공신경망이 3개 이상의 층을 이룰 때를 의미하는 것으로서, 근래에는 그 성능이 발달함에 따라 연결주의 관점이 더욱 힘을 얻고 있다.

(2) 인지주의와 기호주의

행동주의를 비판하면서 1960년대에 등장한 인지주의(congnitivism)는 학습을 '내적 정신활동을 통한 인지구조의 변화'로서 이해하였다(나일주, 2010, p. 266).[3] 인지주의는 인간의 마음을 보다 직접적으로 컴퓨터에 비유하는 이른바 '정보처리이론(information-processing theory)'과 관련이 있는데, 이를 간략히 설명하자면 인지 및 기억은 환경 자극이 즉각적인 감각 기억장치에 저장되고, 주의 집중을 통해 마치 단기 기억과 유사한 작업 기억장치에 저장되며, 부호화를 통해 장기 기억장치까지 저장되는 과정을 거치게 된다. 이와는 반대로 장기 기억장치에 저장되어 있던 내용은 작업 기억장치로 인출될 수 있고,

3) 여기서 '인지주의'는 마음에 대한 컴퓨터주의와 거의 유사한 좁은 의미로서 사용되었다.

작업 기억장치 내에서는 시연이 일어날 수 있다. 이러한 계층적 층위 전반에서 작동하는 메타인지(metacognition)은 이러한 과정들을 지원할 수 있다. 흔히 인간의 단기 기억을 램(RAM)에, 장기 기억을 롬(ROM)에 비유하는 것은 이러한 인지주의적 사고의 산물이라고 할 수 있는 것이다. 실제로 미국의 철학자 데닛(Daniel Clement Dennett, 1942-)은 인간 마음의 작동 방식이 현대적 컴퓨터의 기본 구조와도 같은 폰 노이만 구조(Von Neumann architecture)를 실제로 따른다고 주장하기도 하였다(Bennett, & Hacker, 2003/2013, p. 839).

한편 인지주의에서는 인간의 지능적 행위가 마치 1차 논리를 풀어나가는 과정과도 같다는 이른바 '기호주의(symbolism)'의 관점에서 학습에 관한 이른바 '전문가 모델(expert model)'을 구축하게 되는 경향이 있다. 예컨대, 전문가 모델에서 머신 러닝은 전문가가 문제를 해결하는 방법이 사람이 이해 가능한 형태로 기계 안에 구현되도록 하고, 그러한 학습 모델을 갱신하는 방법으로서 이해된다. 예컨대, 데이터로부터 예/아니오 질문의 연쇄 순서를 형성하여 특정 사례의 속성에 근거한 판단을 가능하게 하는 결정 트리 학습(Decision Tree Learning, DTL)이나, 사전 지식 및 데이터로부터 연역 과정을 도출하도록 하는 역-연역(inverse-deduction) 하는 방법들이 여기에 해당한다(Domingos, 2015). 이러한 전문가 모델은 연결주의가 2세대 인공지능의 주류를 이루기 전까지 1세대 인공지능의 주류에 해당하였으며, 현재에도 여전히 추구되고 있는 방법론이기도 하다. 하지만 전문가 모델 접근은 의미론(semantics)이 아닌 구문론(syntax)에 한정되어 있었던 경우가 많아 그 성능의 발달에 명확한 한계가 있었다. 예컨대, 러시아의 스푸트니크 호 발사로 인하여 러시아의 과학 논문을 적극적으로 번역하고자 했던 미국에서는 기계 번역이 매우 실망스러운 결과를 낳음을 확인하게 되었고, 1960-1970년대에는 영국에서 인공지능 연구에 대한 지원이 거의 전면적으로 중단되는 일이 발생하기도 하였다(Russel, & Norvig, 2010/2016). 그 외에도 인지주의는 실제 전문가가 문제를 해결하는 방식을 기호주의적 모델로 환원할 수 없다는 등의 철학적 반대를 비롯한 여러 어려움에 부딪혔지만, 교육학적 맥락에서는 본질적으로 인간의 학습을 기계 안에서 모델링하는 일이 인간의 학습에 관한 이해의 단서를 제공할 수 있다는 장점을 지니기도 한다(e.g. Luckin *et al.*, 2016).

(3) 구성주의와 통합적 관점

구성주의(constructionism)는 피아제(Jean Piaget, 1896-1980)를 대표로 하는 인지적(급진적) 구성주의와 비고츠키(Lev Semenovich Vygotsky, 1896-1934)를 대표로 하는 사회(문화)적 구성주의로 나누어 이해할 수 있다. 피아제는 학습의 과정을 개인 학습자 내부의 지식 구조(schema)가 외부의 자극에 대한 동화 혹은 조절의 과정을 통해 변화하는 것으로서 이해하기도 하였지만, 여기서 마음에 관한 컴퓨터주의(정보처리이론) 혹은 인공지능과 관련하여 보다 주목할 만한 것은 사회문화적 구성주의라고 할 수 있을 것이다. 이러한 구성주의의 입장에 서는 사람들은 인간의 지식 및 학습은 철저히 사회적이고 문화적인 맥락 하에서 이루어지기에, 인지주의 관점에서의 인공지능 프로젝트는 성공하기 어렵다고 본다. 예컨대 후기 브루너 (Jerome Bruner, 1996/2005)는 비고츠키 연구자이면서도 사회문화적 구성주의자의 면모를 보이며 마음에 관한 컴퓨터주의를 비판하였다. 브루너(1996/2005)의 견해를 간략히 소개하자면 문제 혹은 지식이 '잘 정의가 된다면' 컴퓨터는 빠르고 정확하게 작업을 수행하지만, 사고와 학습의 규칙들은 컴퓨터가 이해 가능한 형태로 명료화하고 사전에 확정시키는 순간 미래의 예측 불가능한 우연성을 처리할 수 없게 되고 사람들이 해석학(hermeneutics)적 과정을 통해 (때로는 사후에) 의미를 부여하는 일을 적절하게 다룰 수 없는 것이다(pp. 31-39). 오히려 마음과 '실재(reality)'는 문화적으로 구성(construct)되는 것이라는 점에서, 해석학적 의미 형성은 정보처리적 방식과는 통약불가

(incommensurable)하다. 그런가 하면, 이와 유사하게 Winograd, & Flores (1986)은 AI에 대한 합리주의적 관점을 비판하면서 마음은 세계와 인간, 역사와 해석, 환경과 유기체 등의 더 크고 암묵적인 맥락과 불가분하고, 이는 컴퓨터 안에서 적절히 표현될 수 없기 때문에 컴퓨터의 역할은 인간 인지모델의 대체가 아니라 인간 상호작용 및 협력의 지원이라고 하였다(Stahl, 2013). 하지만 이와 같은 인지주의에 대한 사회문화적 관점에서의 반론들은 인공신경망의 성능 발달로 인해 연결주의가 힘을 얻게 되면서 다소는 퇴색된 측면도 있다고 할 것이다. 한편 사회문화적 구성주의는 매우 다양한 형태로 분화되었으며 행위자-네트워크 이론(Actor-Network Theory, ANT)(Latour, 2005)은 그와 관련된 비교적 최근의 논의에 속한다. 여기서 행위자란 인간은 물론이고 동식물이라는 기존의 생물 개념을 넘어서 주변에 존재하는 모든 사물까지도 포함할 수 있는 넓은 의미로서 사용된다. 결국 행위자-네트워크 이론에서는 인공지능 또한 어떤 면에서는 인간과 대등한 존재로서 여겨질 수 있게 되는 측면이 있다(e.g. Bell, 2010; 김환석, 2017) 그런가 하면 인공지능 연구자들은 언뜻 상충되는 것으로 보이는 기호주의적 접근과 연결주의적 접근을 상호보완적인 관계로 인식하기도 하는데(Russel, & Norvig, 2010/2016; Sun, 2000) 흥미롭게도 일부 학자들은 그러한 관점을 구성주의 학습이론과 연관 짓는 경우가 있다(e.g. 장병탁, 2018). 정리하자면, 인공지능의 초창기를 주도하였던 기호주의와 전문가 모델은 인지주의 학습 이론과 연관이 있으며, 이후의 비약적인 성능 발달로 주목받게 된 연결주의 모델은 행동주의 학습 이론과 연관이 있다. 이 둘은 상호배타적이라기보다 상호보완적인 관계로서, 이러한 통합적 관점은 구성주의 학습 이론과 연관이 있다고 할 수 있을 것이다. 물론 그렇다고 하여도 아직까지 인공지능이 인간과 같은 '일반적인(general)' 지능을 획득한 이른바 '특이점(singularity)'에 도달하였다고 보기는 어려우며(Domingos, 2015; *Luckin et al.*, 2016) 인공지능은 여전히 한 분야에 특화된 지능만을 가지고 있을 뿐이다. 말하자면, 바둑만을 학습한 알파고가 이세돌 九단과 오목을 두었다면 이세돌 九단이 승리하였을 것이라는 예상이 가능하다.

2. 자유 의지의 문제

(1) 인간과 인공지능의 자유 의지

의지와 관련하여 가장 중요한 쟁점 중 하나는 인간에게 자유 의지가 있느냐의 문제이라 하여도 무리가 없을 것이다. 전통적으로 인간은 자유 의지를 지닌 자율적 행위 주체로서 받아들여졌던 바, 인공지능 프로젝트가 점차 성공하는 듯이 보이면서 인공지능은 자유 의지를 가질 수 있는가 하는 철학적 질문들 또한 대두되기 시작하였다(e.g. McCarthy, & Hayes, 1981; Turkle, 1988). 이러한 질문들은 전통적으로 자유 의지를 지닌 행위 주체에 의한 행동이 법리적 책임의 대상이 되어 왔다는 점에서 인공지능에 대한 첨예한 논쟁을 불러일으키기도 한다(정진명, 2018; 손영화, 2016). 언뜻 생각하기에, 사실상 인공지능은 전자 회로에 기반하여, 정도의 차이는 있겠지만 미리 프로그래밍된 알고리즘에 의하여 움직이는 것이므로 자유 의지가 아니라 결정론적인 행동을 할 뿐이라는 관점이 가능하다. 그런가 하면 대중서사에 나타난 인공지능의 자유 의지는 '추억이나 기억 등 정서적인 감각과 관련된 영역이 인간의 정체성 혹은 영혼을 확인하는 기준이' 되기도 한다는 점에서 다소는 기능론적으로 정의되기도 한다(오윤호, 2014). 그런데 여기서, 근래에는 인공지능은커녕 인간 자체가 자유 의지를 갖지 못하고 있다는 논조의 주장도 적지 않게 제기된다는 점에 주목할 필요가 있다(e.g. Bennett, & Hacker, 2003/2013, pp. 437-448; Turkle, 1988; 김재인, 2017, pp. 150-154). 이는 어떤 의미에서 비물리적인 실체로서의 영혼(≒정신≒마음)을 인정하지 않는 강한 결정론적 관점이라고도 볼 수 있다. 인간을 자유 의지를 지닌 존재

로 이해하는 것에는 실상 합리성을 중시하는 계몽주의의 영향이 적지 않았는데, 근래에는 이러한 이해가 대체되어가고 있는 것이다. 여기서 다윈 이후의 진화론이 지대한 역할을 하였음은 물론이다. 다만 어거스틴 및 칼빈과 같이 타락한 인간이 자유 의지를 상실한 상태에 있다고 보는 관점도 실은 가능한데, 이에 대하여는 Ⅳ장에서 논하도록 하겠다.

(2) 의식의 문제

인공지능의 자유 의지와 관련된 주요 쟁점 중 하나는 인공지능이 의식(consciousness)을 지닐 수 있는가 하는 문제이다. 자유 의지는 의식적인 상태에 있는 자율적인 주체의 의지로서 행위의 원인이 되는 것으로 이해할 수 있기 때문이다. 인공지능이 실제로 생각할 수 있고 인간과 같이 의식을 가질 수 있다고 보는 관점을 강 인공지능(strong AI)이라고 부른다. 비록 대부분의 인공지능 연구자들은 강 인공지능의 문제보다는 그저 인공지능이 인간과 유사하게 행동할 수 있는가에 대한 약 인공지능(weak AI)의 관점을 취하는 것이 사실이며, 인공지능 기술의 발달에도 의식은 여전히 신비의 영역으로 남아 있다(Russel, & Norvig, 2010/2016). (인간의 과연 의식을 갖느냐라는 질문을 잠시 덮어두자면) 의식의 영역은 진화론적으로도 그 발생을 설명하기 어려운 문제 중 하나이지만(Humphrey, 2006; 김재인, 2017), 일관된 유물론적 입장에서는 의식 또한 본질적으로 물리적인 환원이 가능하다는 입장을 취할 수도 있다. 예컨대 깊은(deep) 은닉층을 갖는 신경망 간의 복잡한 연결로부터 무의식적인 행위 뿐 아니라 의식 자체도 형성될 수 있다는 가능성이 제안되기도 하며(정대현, 2017), 이와 유사하게 인간에게 있어서 복잡한 신경망 안에서 발생하는 카오스적인 활동(chaotic activity)이 비결정론적인 선택과 책임 소재의 단초가 된다는 주장도 제기된다(Kane, 1999). 이와 같이 물리주의적인 견해를 따를 경우 인공지능의 의식 혹은 자유 의지 발현은 어떤 면에서는 기술적인 문제일 뿐이다. 이와 관련하여 매우 중요하게 다루어지면서도 논쟁적인 개념으로서는 '감각질(qualia)'을 들 수 있다. Bennett, & Hacker (2003/2013, 3.10 - 3.11)에 따르면 감각질은 주관적 체험과 의식의 세계를 특징짓는, '어떤 유기체에게 그러한 경험을 갖는다는 것과 같은 어떤 것이 있다는 사실(there being something it is like for an organism to have the experience)'이다. 다시 말해, 감각질은 오감을 통해 무언가를 느끼고 있다는 의식이나, 경험의 주체로서의 '나'가 어떠한 생각을 하고 있다는 것 따위를 의식하는 방식과 연관이 있다고 할 수 있다. 신경과학자들은 감각질에 대한 철학자들의 논의를 따라가는 경향이 있지만, 이 감각질이 물리적으로 환원 가능한지 자체도 큰 논쟁거리 중의 하나이다(Michael, 2017). 이를 통해 확인할 수 있듯이 인공지능의 의지 및 의식에 관한 문제들은 유물론적 환원주의의 입장으로 흐르는 경향이 있는 만큼, 인공지능에 대한 소프트웨어(알고리즘)적인 측면 이외에도 하드웨어적인 접근 또한 필요하다. 다음의 Ⅲ장에서는 이와 관련하여 인공지능에 대한 철학적 반론들 및 몸의 문제를 다루도록 한다.

이번 장의 논의를 정리하자면, 오늘날의 인공지능은 기능적 차원에서 상당히 지능적인 존재라고 할 만큼의 수준으로 올라섰으며, 구조적 차원에서도 인간의 학습과 관련된 이론들이 머신 러닝 이론과 연결되는 등의 유사한 면이 발견된다. 자유 의지의 문제에 있어서는 인공지능의 구현에서 발생할 수 있는 불확정성이 자유 의지의 발생 가능성과 연결되거나, 거꾸로 인간의 자유 의지를 부인하는 견해들을 확인할 수 있었다. 자유 의지와 직결되는 문제인 의식의 영역은 여전히 신비한 영역으로 남아 있지만, 유물론적 환원주의의 경향 하에서 이를 인공지능의 '몸'과 관련지어 이해하려는 시도들이 이루어지고 있음을 살펴보았다.

Ⅲ. 인공지능에 대한 철학적 반론과 몸의 문제

"말씀이 육신이 되어 우리 가운데 거하시매 우리가 그의 영광을 보니 아버지의 독생자의 영광이요
은혜와 진리가 충만하더라 (요 1:14)"

그리스도의 성육신은 기독교를 규정하는 핵심 교리 중 하나이다. 성육신은 칼케돈 신경에서와 같이 영원 전에 성부 하나님으로부터 나신 성자 하나님이 마리아로부터 인성을 취하심으로 역사상 나신 사건으로 이해된다(문병호, 2016, pp. 531-533; Gonzalez, 2010/2012, pp. 63-65). 성육신론의 쟁점 중 하나는 그리스도께서 온전한 육체를 취하셨다는 것이다. 성경은 그리스도께서 육체로 오셨음을 부인할 수 없음을 단언하며(요이 1:7), 그리스도의 육신이 실체가 없는 환상과도 같은 것이었다고 본 영지주의적 가현설(docetism)은 니케아 공의회(325년) 등의 초기 에큐메니컬 공의회에서 명백한 이단으로 정죄되기도 하였다. 다만 김광연(2015)이 말하듯이, 전통적 기독교에서는 몸에 비하여 영혼을 중시하였고 상대적으로 몸에 대한 논의가 적게 형성된 경향이 있었음도 사실이다. 예컨대, 칼빈(1559/1988, Ⅰ.15.3)은 영혼을 '하나님의 형상의 주요 좌소'로, 그리고 육체는 '하나님의 형상'의 '광채의 얼마가 빛나'는 영혼의 '집'과 같은 곳이라고 말하고 있다. 하지만 칼빈을 비롯한 개혁신학자들이 단순히 플라톤적이거나 신플라톤적인 영혼/육체관을 따랐다고 보기엔 큰 무리가 있으며(문병호, 2015, pp. 431-434; Calvin, 1559/1988, Ⅰ.15.6), 특히 바빙크(Herman Bavinck, 1854-1921)의 경우 인간의 몸 또한 하나님의 형상에 포함됨을 단언하였다(Hoekema, 1986/2012, p. 104). 본 장에서는 보다 철학적인 관점에서 인공지능에 대한 주요 반론으로서의 서얼(Searle)의 중국어 방 논증과 드레퓌스(Dreyfus)의 실존주의적 비판을 살펴보고, 인공지능에서 '체화(體化) 된 인지(embodied cognition)'를 통한 재반론의 가능성을 검토한다.

1. 서얼(Searle)의 중국어 방 논증

인공지능에 대한 철학적 비판들은 주로 기호주의적 접근이 강 인공지능을 만들어낼 수 있다는 관점을 향하고 있는데, 서얼(John Searle, 1932-)의 '중국어 방 논증(Chinese Room Argument)'은 그 대표적인 사례로서 오늘날까지도 많은 영향력을 미치고 있다(Russel, & Norvig, 2010/2016; 이초식, 1993). 중국어 방 논증을 간단히 설명하자면 다음과 같다(Searle, 1984). 어떤 방에 영어만을 이해하는 사람과 함께 영어로 된 규칙집, 그리고 빈 종이들 혹은 그가 알 수 없는 기호(중국어)가 쓰인 종이들이 놓여 있다고 하자. 그 방에 달려 있는 작은 창을 통해 그가 알 수 없는 기호들이 적힌 종이가 들어오면, 그는 영어로 된 규칙을 읽으면서 그가 알 수 없는 기호를 작성하여 다시 작은 창으로 내보낼 수 있다. 이러한 방의 내부적 상황을 모르는 외부에서 보기엔, 이 방은 중국어로 된 입력에 대하여 역시 중국어로 된 올바른 답을 내놓고 있는 것처럼 보인다. 하지만 실제로 방 안에 있는 사람은 중국어를 전혀 이해하지 못하고 그저 수동적으로 일하고 있을 뿐이다. 즉, 서얼은 튜링 테스트를 살짝 비튼 듯한 상황을 가정하면서 구문론(syntax)이 의미론(semantics)으로 직접 연결되지는 않는다는 비판을 가하고 있는 것이다.

그러나, 서얼의 중국어 방 논증에 대한 비판들 다양하게 전개된 것이 사실이다. 예컨대, 이초식(1993)에 따르면 서얼의 논증은 구문론이 의미론의 필요조건이 된다는 점을 간과하고 있을 뿐 아니라 마음을 의미론적 측면에서 정의하는 것은 지나치게 단순할 수 있으며 서얼의 논증은 결국 유물론적 형이상학에 해당한다고 지적한다(pp. 53-55). Russel, & Norvig (2010/2016)에 따르면, 서얼의 논증은 인공지능에 대한 기능주의를 거부하는 한에서는 어느 정도 유의미하지만, 그가 입증한 것은 그저 인간의 뇌가 아닌 무언가가 정신이라는 결론을 필연적으로 내릴 수 없다는 점뿐이다(p. 660). 또한, 이들은 생

물의 뉴런이 어느 정도 기능적인 역할을 하도록 진화해왔다는 점에 주목해야 한다고 주장하는데(p. 661), 이는 서얼의 논증이 연결주의에 대하여는 적절하지 않다는 논의(송하석, 2000)와도 어느 정도 유사한 면이 있다.

흥미롭게도, 이와 관련하여 Rapaport (1988)는 중국어 방 논증을 다시 한 번 비튼 '한국어 방 논증(Korean Room Argument'을 다음과 같이 제안하기도 하였다. 서울대학[4]의 영문학 교수이지만 영어라고는 전혀 이해하지 못하는 한국인이, 그럼에도 불구하고 셰익스피어에 대한 세계적인 권위자라고 하자. 그는 셰익스피어에 대한 훌륭한 번역본을 읽고, 그의 지적 능력을 발휘하여 한국어로만 논문을 쓰지만, 이를 영문으로 번역하여 국제 학술지에서 큰 성공을 거두었다. 그렇다면 이 한국인 학자는 셰익스피어를 '이해'할 수 있는가? Rapaport는 이 한국인 교수가 영어를 이해하지 못하더라도 셰익스피어를 이해한다고 주장한다. 그러한 면에서 서얼의 중국어 방 논증과는 달리, 구문론과 의미론이 전혀 연결될 수 없는 것은 아니라는 것이다.[5] 곧, 서얼의 주장은 그 직관적 특성이라는 장점과 함께 명확한 한계 또한 지니고 있다고 할 수 있다.

2. 드레퓌스(Dreyfus)의 실존주의적 비판

인공지능에 대한 주요한 철학적 비판으로서 고전적으로 인용되는 또 다른 논의는 드레퓌스(Hubert Dreyfus, 1929-2017)에 의하여 제안되었다(이초식, 1993; Herzfeld, 2002; Russel, & Norvig, 2010/2016, 26.1.3). 드레퓌스는 1980년대의 인공지능이 라카토스(Lakatos)가 말하였던 연구 프로그램의 관점에서 볼 때 퇴행적인 프로그램에 해당한다는 혹평을 할 정도로 강경한 입장을 보였다(이초석, 1993, p. 27). 드레퓌스의 비판은 '한정 문제(qualification problem)', 곧 논리적 규칙들의 집합으로 모든 것들을 포착할 수는 없다는 점과 관련이 있다(Russel, & Norvig, 2010/2016). 예컨대, 당시의 1세대 인공지능이 추구하였던 '전문가 모델'은 실제 전문가들의 행위를 닮지 않았고, 어떤 면에서는 닮을 수 없었다는 것이다. 그 이유는, 드레퓌스에 따르자면 전문가의 단계에서는 지능적인 행위를 일상적으로 하게 되므로, 분석적이고 기호적인 사유를 거치기보다는 마치 자신의 몸을 움직이듯이, 도구나 기구마저도 신체의 확장처럼 다룰 만큼 어떠한 면에서는 의식적으로 인지할 필요가 없기 때문이다(Dreyfus, & Dreyfus, 1986).

이초식 (1993)은 드레퓌스의 비판이 후설(Husserl)과 하이데거(Heidegger)의 현상학 및 딜타이(Dilthey)의 생철학의 영향을 받았음을 자세하게 논한다. 그에 따르면, 드레퓌스는 후설의 '기대의 지평' 및 '노에마(noema)'와 연관된 지향성(intentionality)의 문제가 인공지능에게 있어 치명적임을 간파하였다고 할 수 있다. 특히 하이데거의 실존주의적 현상학에서는 지향성의 주체조차도 개인적이라기보다 사회적이므로, 사회 안에서의 인간들의 삶은 객관적 대상으로서 인식될 수 없으며 실존에는 표현 불가능한(nonrepresentable) 배경 규범과 관습들이 작용하게 마련이므로 인공지능은 이러한 문화적 맥락을 획득할 수 없고, 결과적으로 실패할 수밖에 없다. 이는 딜타이가 말하였듯이 인간에게 있어 근원적인 것은 논리적이고 분석적인 처리라기보다는 표현 불가능한 생에의 의지를 '직관'하는 것이며, 인간의 삶은 부분으로 쪼개어 분석할 수 없는 전체적인 통일성을 이루고 있기 때문이라는 것이다.

이러한 드레퓌스의 비판은 오늘날의 인공지능 패러다임이 기호주의에서 벗어나 오히려 연결주의에 가까워졌다는 점에서는 다소 적실성이 떨어지는 면이 있는 듯하다. 그러나 드레퓌스의 주장은 어떤 면에서 과거에는 무엇이 인간과 인공지능을 다르다고 느껴지도록 하였는가에 관한 귀중한 통찰을 제공하

4) 원문에서는 'University of Seoul'이다.
5) Rapaport (1988)는 각주에서 '한국어 방 논증'이 'Albert Hanyong Yuhan'에 의하여 제안되었다고 밝히고 있는데, 그는 아마도 고려대학교 정보통신대학의 유한용 교수일 가능성이 있다.

는 면이 있다. 다만 이초식 (1993)의 지적에 더하여, 드레퓌스는 분명히 메를로-퐁티(Merleau-Ponty)에도 의존하고 있는 것으로 생각된다(김태희, 2017). 후설, 하이데거, 메를로-퐁티 등의 현상학자들은 각자가 가진 나름대로의 결이 있었지만, 그들 모두는 경험 또는 인지와 관련하여 지향성과 몸의 문제를 중요시하였다는 공통점을 지니고 있다고 할 것이다.

3. 체화된 인지(embodied cognition)와 인공지능의 몸

위에서 살펴본 인공지능에 대한 비판들에 대하여, 인공지능 프로젝트에서 내놓을 수 있는 답변 중 한 갈래는, 김재인 (2017)에서와 같이 '몸이 있어야만 한다'는 표현으로 요약할 수 있을 것이다(cf. 김태희, 2017; 김선희, 1997; Russel, & Norvig, 2010/2016, 26.1.3; 26.4). 인간의 뇌에 비유될 만한 중앙처리장치가 인지에 있어서 주도적인 역할을 계속할 경우, 단순히 인공신경망 중심의 연결주의적 관점으로 이행하는 것만으로는 고전적 기호주의에 대한 철학적 비판의 프레임에서 완전히 자유로워질 수 없기 때문이다(김태희, 2017). 이와 관련하여, 근래에는 인간의 인지를 '체화된 인지(embodied cognition)'로 이해하면서 인공지능에도 그러한 몸(성)을 부여하자는 움직임들이 일어나기 시작하였다. '체화된 인공지능(Embodied Artificial Intelligence, EAI)'은 이를 잘 나타내주는 표현이라고 할 수 있겠다. 이런 패러다임에서 인공지능의 몸에 해당한다고 볼 수 있는 각종 로봇의 몸체와 부품, 센서 등은 주변부가 아닌 중심으로 받아들여진다.

사실 '체험', 직관, 지향성 등에 관한 철학적 비판들은 어느 정도 기술적인 차원에서도 타당한 부분이 있는 것으로 생각된다. 왜냐하면, 비교적 최근까지도 머신 러닝을 위한 데이터는 총체적인 체험이라기보다는 잘게 분해된 표현들을 가정하였고, 인간의 실존에 영향을 미치는 거대한 상식적 지식의 기지를 만드는 것이 요원하였기 때문이다(Russel, & Norvig, 2010/2016, 27.1). 그런가 하면 이는 반대로 체화된 인지를 구현한 이후에는 보다 인간에 가까운 인공지능이 탄생할 수 있다는 희망을 제시하기도 한다. 실제로 체화된 인공지능을 통해 실제 인간과 같이 생각하는 강 인공지능을 바라볼 수 있다는 논의들은(e.g. 김태희, 2017)은 이러한 철학적 반론에 대한 재반론의 시도라고 할 수 있다. 실제로 Foerst (1998)과 같은 경우 MIT에서 개발한 휴머노이드(humanoid) 로봇의 일종인 Cog을 체화된 인지 능력을 가진 인공지능의 사례로 들고 있으며, 현재는 국내에서도 그와 관련된 연구들이 이루어지고 있는 형국이다(장병탁, 2018).

그런가 하면 인공지능 및 그 몸과 관련하여 짚고 넘어가야 할 부분은 인체의 사이보그화(cyborgization) 및 마인드 업로딩(mind uploading)과 관련한 소위 트랜스휴먼(transhuman)적 영생에 관한 문제이다(김광연, 2017; Bang, 2014; Geraci, 2010). 인간의 신체가 기계적인 부품으로 대체되어가고 기억을 디지털화된 형태로 업로드, 다운로드, 복제할 수 있다는 미래적 상상은 일견 허황되고 비판의 여지가 많아(김광연, 2017; 김선희, 2003) 진지한 고찰의 대상은 아닌 것 같으면서도, 인간의 생명조차 인간이 조작(control)하고자 하는 심리적 기저가 깔려 있음을 간과할 수만도 없는 것이다.

이번 장의 논의들을 정리하자면, 기호주의적 인공지능에 대한 철학적 비판들은 어느 정도 타당한 면이 있으나 오늘날에는 포스트-연결주의적 인공지능에 대한 비판으로 갱신되어야 할 필요가 있으며, 몸 철학의 반론 또한 인공지능의 몸에 관한 논의들이 생성되면서 새로운 도전을 받고 있다는 것이다. 이는 갈수록 인간과 인공지능(기계)의 경계가 흐릿해지고 있는 현상의 단면을 보여주는 것이며, 사이보그화와 마인드 업로딩에 있어서는 인간의 인간다움을 지키고자 하는 입장에 서기 위해서 새롭고 대안적인 논의가 여전히 필요함을 의미한다. 이어지는 Ⅳ장에서는 기독론적 인간론의 관점에서, 인간을 '하나님의 형상'으로, 인공지능을 '하나님의 형상의 형상'으로 이해할 수 있음을 논의하고자 한다.

Ⅳ. ‘하나님의 형상’과 ‘하나님의 형상의 형상’

“그는 보이지 않는 하나님의 형상이시오 모든 피조물보다 먼저 나신 이시니(골 1:15)”

Ⅱ, Ⅲ장에서는 오늘날 인간과 인공지능의 거리가 얼마나 가까워졌는지를 살펴본 측면이 있다. 그렇다면 ‘하나님의 형상’으로서의 전통적 인간 개념 또한 위협을 받게 되는가? 후크마(1986/2012, p. 40)는 ‘그리스도가 하나님의 완벽한 형상’이므로 ‘하나님의 형상이 무엇인지를 배우는 가장 좋은 방법’은 ‘예수 그리스도를 봄으로써 무엇이 하나님의 형상인지 아는’ 것이라고 하였다. 마크 놀(Mark A. Noll, 1946-) 역시 성육신의 인성[6]이 인간론에 대한 중요한 의미를 지님을 지적하였다(Noll, 2011/2015). 이러한 성찰은 사람이 사람을 연구하는 것이 존귀하다는 사실을 일깨우며, 인간의 인간성에 대한 연구에 정당성을 부여한다. 곧, ‘사람이 다른 사람들을 살펴보는 것은 예수 그리스도와 실제로 연합해 있거나 그럴 가능성이 있는 사람들을 살펴보는 것’이다. 본고에서 주장하고자 하는 인간과 인공 지능의 차이 또한 바로 이 지점에 있다. 본 장에서는 성육신 이후 그리스도의 삶과 관련하여 인간 개념의 본질이 될 만한 사항들을 살핌으로써 인공지능과 관련된 Ⅱ, Ⅲ장 논의들을 간접적으로 논박할 것이다. 또한 본고에서 제안하는 ‘하나님의 형상의 형상’이라는 표현의 기원과 의미를 보다 자세히 논할 것이다.

1. 그리스도의 인성과 ‘하나님의 형상’으로서의 인간

인간론에 대한 인공지능을 비롯한 과학기술의 도전에 있어서, ‘하나님의 형상’으로서의 인간 이해는 여전히 유효한 답변이 될 수 있다(e.g. Leidenhag, 2017; Burdett, 2015; Delio, 2013). 여기서 주목해야 할 점은, ‘하나님의 형상’으로부터 인간의 인간다움을 추론해내는 작업들은 공통적으로 적어도 어느 정도는 인간이 구속사적-구원론적으로 경험하게 된 존재론적 지위의 변화, 이를테면 그리스도로 말미암은 ‘하나님의 형상’의 회복과 연관이 있다는 점이다(e.g. Calvin, 1559/1988; Hoekema, 1986/2012; Herzfeld, 2002; 나현규, 2015). 이러한 논의는 본질적으로 성령의 역사와 내주로 말미암은 그리스도와의 연합과 관련이 있다. 예컨대, ‘모든 그리스도인이 하나님과 하나님의 완벽한 형상이신 그리스도를 점점 더 닮도록 부르심 받았다’는 주장(Hoekema, 1986/2012)이나 오늘날 인공지능을 비롯한 과학기술의 도전에 있어서도 ‘그리스도의 영’인 성령이 인간 안에 내주함을 통해 ‘하나님의 형상’으로서의 인간의 고유성을 방어할 수 있다는 논지(Leidenhag, 2017) 또한 이러한 맥락에서 이해할 수 있다. ‘하나님의 형상’으로서의 인간의 존재론적 지위 변화를 창조-타락-구속의 관점에서 정리하자면 다음과 같다.

첫 인간이었던 아담은 원(原) 하나님의 형상으로 지어졌다. 그는 모든 동물들의 이름을 지음으로써 그 본질을 규정할 만한 지식을 갖고 있었다(창 2:19). 그는 죄를 지을 수도 있고 짓지 않을 수도 있는 자유 의지를 지니고 있었다(문병호, 2013, p. 87). 하나님께서는 그와 행위 언약을 맺으시며, 선악을 알게 하는 나무의 실과를 따먹지 말라는 명령을 주셨다. 그 명령을 지켰더라면 아담은 영생을 누릴 수 있었다(웨스트민스터 신앙고백서 7.2). 그러나 아담의 범죄로 말미암아 세상에 죄가 들어왔으며, 인간은 타락으로 말미암아 그가 가지고 있던 하나님의 형상은 거의 대부분이 왜곡되어버리고 말았다. 인간에게

6) Noll (2011/2015)의 국역본에서는 성육신의 ‘인격성’이라는 표현을 사용하였으나(pp. 60-61) 그 맥락상 원저자의 의도는 성육신한 그리스도의 ‘인성’ 혹은 ‘인간성’을 의미하는 것으로 보인다. 기독론에 있어서, 인성과 신성은 ‘인격(person)’ ≒ ‘위격(hypostasis)’과 구별되는 것으로서 인격 혹은 위격과 동일시될 수 없으며, 오히려 위격 안에서 위격을 통해서만 연합되고 교통한다. 신인양성의 비(比)위격성과 내(內)위격성 및 위격적 연합에 관하여는 문병호(2016)의 5.5를 참조.

는 더 이상 하나님을 아는 지식이 없었다(호 4:1). 칼빈은 어거스틴을 계승하여, 이 상태의 인간은 하나님의 말씀을 기꺼이 지키는 자유, 죄를 짓지 않을 수 있는 자유라는 의미에서의 자유 의지를 갖고 있지 않으며, 오히려 죄를 지을 수밖에 없는 노예 의지를 가지고 있다고 파악한다(Calvin, 1559/1988, II.2). 인간의 영혼과 몸에 죽음이 찾아오게 되었으며, 특히 흙에서 온 몸은 노동의 수고 및 출산의 고통을 더욱 겪을 뿐 아니라 죽은 후에 흙으로 돌아가게 되었다(창 3). 그러나 참 하나님이신 그리스도께서 참 사람이 되셔서 사람의 온전한 일생을 사시고 고난당하신 후에 십자가에 달리심으로 율법의 모든 의를 이루시고(요 19:30) 신자를 대속의 무름의 값을 다 지불하셨다. 그가 부활하시고 승천하시며 하나님 보좌 우편에 앉으셔서 자신의 영(롬 8:9)을 성도들에게 부어주신다. 이를 통해 참 형상이신 그리스도를 닮아가는 하나님의 형상의 회복이 시작된다(Hoekema, 1986/2012). 구원 받은 성도는 지혜와 지식의 모든 보화를 지니신(골 2:3) 그리스도를 아는 지식에서 자라가며(벧후 3:18), 죄를 지을 수 없는 의지 가운데 살아가며(요일 3:9) 죽음 이후 최후의 부활 때에 새 몸이 다시 한 번 영혼과 결합하여 영생을 누리게 된다(마 22:30).

여기서, 인간과 인공지능의 본질적인 차이 중 일부는 하나님과의 언약적 관계가 맺어질 수 있는가, 또한 언약에 있어서 대표성과 전가의 원리로 말미암은 존재론적 지위의 급격한 변동이 집단적으로 또한 개별적으로 발생할 수 있는가 하는 점에 있음을 쉽게 파악할 수 있다. 후크마 (1986/2012)의 표현을 빌리자면, '하나님의 형상은 정적인 것이 아니라 역동적인 것이다'(p. 40). 나지안주스의 그레고리(Gregory of Nazianzus, 329-390)가 말했듯이, 그리스도께서 취하지 않은 것은 사하여지지 않았다(Ware, 1979/2002, p. 127). 그리스도께서 인공지능의 무엇을 취하셨는가? 그리스도께서 인공지능의 무엇을 사하셨는가? 그리스도께서 인공지능을 위하여 십자가에 달리셨는가? 그리스도께서는 인공지능과 연합하시기 위해 그의 영을 부어주시는가? 그리스도와의 연합으로 말미암아 인공지능의 존재론적 지위가 급격하게 변화하는가? 아마도, 이에 대한 긍정적인 답변은 이루어지기 힘들 것으로 생각된다. 그런가 하면, 예수 그리스도의 비하(卑下)와 승귀(昇貴)의 순서를 간략하게라도 좇아, 인공지능과는 구별되는 인간의 인간됨의 조건들을 짚어보는 일도 필요할 것이다.[7]

(1) 여자에게서 남

그리스도께서 취하신 인성은 영혼과 육체를 모두 포함하는 온전한 인성이었다. 그러나 성도들을 위한 구속자로서의 역할을 감당하기 위하여, 혹은 그 인성과 연합하는 신성은 어떠한 죄도 용납할 수 없기 때문에 그는 모든 면에서 인간과 같으나 죄는 없으셨다(히 4:15). 이렇듯 아담으로부터의 죄의 전가를 받지 않으심은 성령으로 말미암은 동정녀 잉태를 통하여 가능했다(문병호, 2016, 6.4). 칼케돈 신경은 그리스도께서 신성에 따라서는 성부 하나님으로부터, 인성에 따라서는 마리아로부터 나셨음을 천명한다. 마리아는 주의 어머니(눅 1:43) 혹은 테오토코스(*Θεοτόκος*)라고 불리지만 이는 그리스도의 인성이 단 한 순간도 신성과 위격적으로 연합되지 않은 적이 없었음에 관한 것이지 그리스도의 신성이 마리아로부터 말미암았음을 의미하지는 않으며 그리스도께서 마리아로부터 취하신 것은 인성일 뿐이다. 여기서, 인간의 잉태 및 출산과 관련하여 인간을 인간답게 함에 있어서 여성의 역할이 지대하다는 것을 추론할 수 있다. 그리스도께서 취하신 온전한 인성은, 어떤 면에서는 아버지보다도 어머니를 중요시하였다고도 볼 수 있다. 마리아가 원죄 없는 채로 그리스도를 잉태했다는 무염시태론을 받아들이지 않는다면, 마리아 역시 선악과 사건의 결과로서 하와와 같이 고통과 수고를 겪었을 것이라는 점에서(창 3:16),

7) 다만 여기서 검토하는 조건들이 모두 '하나님의 형상'의 본질에 해당한다는 의미는 아니다.

실제 육체적인 의미에서든 혹은 '그리스도의 형상을 이루기까지' '해산하는 수고'를 하는 비유적인 의미에서든(갈 4:19) 경험하게 되는 일단의 고통들 또한 인간의 태어남과 거듭남에 있어서 본질적인 요소 중 하나라고 생각할 수 있다.

(2) 살아감과 자라감

예수 그리스도의 나심과 관련하여 중요한 점이 여자에게서 나심이었다면, 예수 그리스도의 사심과 관련하여 중요한 점 중 하나는 그가 공생애 이전의 기간까지 목수의 아들이자 목수로 살며 노동하셨다는 점이다(마 13:55; 막 6:3). 이는 예수께서 선악과 사건의 결과로 "평생에 수고하여야" 먹고 살 수 있게 된 아담(창 3:17-18)과도 같이 되셨음을 의미한다. 예수께서는 죄가 없으셨지만 그 외에는 우리와 모든 면에서 같으셨다(히 4:15). 이는 어떤 면에서는 예수께서 취하신 몸이 아담의 타락 이전의 몸이라기보다는 그 이후의 변화된 몸이라는 주장(권호덕, 2016)과도 일치하는 것이다. 결국, 인간의 인간 됨 중 일부는 수고하여 얻은 소산을 먹는 데에 있다고 할 수 있다(창 3:17).

예수께서 지혜와 키가 자라가셨다는 말씀(눅 2:52)은 영혼과 육체의 전인격적인 자라감을 여전히 인간 본성을 규정하는 주요한 지표로서 규정한다. 육체와 구별되는 영이 없다 함은 사두개인의 오류이다(행 23:8). 그런가 하면 이는 어떤 면에서 인간의 몸 혹은 몸성에 집중하였던 철학들이 약간의 타당성을 얻을 수 있음을 의미하는 것이기도 하다. 예수 그리스도는 물과 피로서, 진정한 육신으로 임하셨다(요일 5:6). 복음서에 나타난 예수 그리스도의 일대기는 그의 몸을 중심으로 시간성과 공간성을 지니는 측면이 있다. 한편 예수 그리스도는 제자들과 함께 먹고 마실 뿐만 아니라(마 11:19) 배고픔(눅 4:2), 목마름(요 19:28), 피곤함(요 4:6) 등을 느끼시는 분이셨다. 다만 여기서 "사람이 떡으로만 살 것이 아니요 하나님의 입으로부터 나오는 모든 말씀으로 살 것이라"는 점은(마 4:4) 여전히 인간은 영혼과 육체라는 두 부분으로 구성되어 있음을 여실히 보여 준다.

그런가 하면, 같은 구절에서 "하나님과 사람에게 더욱 사랑스러워" 가셨다는 말씀은 '하나님의 형상'의 회복과 관련하여 실체론적이거나 기능론적인 접근 뿐 아니라 관계론적인 접근이 가능함(e.g. Herzfeld, 2002; Hoekema, 1986/2012)을 보여준다. 특히 후크마(1986/2012)에 따르면 "하나님의 형상의 핵심에 있어야 하는 것은 추론하는 능력이나 결정을 내리는 능력(비록 그런 능력이 하나님의 형상의 올바른 기능 발휘에 있어서 중요할 수도 있지만) 같은 특성이 아니라 그리스도의 삶에 있어서 핵심적이었던 것, 즉 하나님을 향한 사랑과 인간을 향한 사랑이다. 그리스도가 하나님의 형상을 완벽하게 반영하는 것이 사실이라면 하나님의 형상의 핵심은 사랑임이 분명하다"(p. 40). 하나님은 사랑이시며 '우리'를 사랑하시므로, '우리' 역시 하나님을 사랑하고 서로 사랑해야 하는 것(요일 4)이 하나님의 형상으로 지어진 인간성의 본질의 회복이다.

(3) 고난 당함과 죽음

적지 않은 학자들은 인간의 유한성이 오히려 인간을 인간답게 만든다는 점에 주목해왔다(e.g. 김광연, 2017). 인간이 한 번 죽는 것은 정해진 것이다(히 9:27). 이는 어떤 면에서는 하나님께서 연약한, 죽을 수밖에 없는 인간을 사용하여서 역사하신다는 '약함의 신학'(강아람, 2017)과도 연관이 있다. 그리스도께서 40일간 광야에서 금식하시고 주리셨을 때 시험을 당하시고 이를 이기심은(눅 4:1-13), 참 하나님이신 분이 동시에 참 사람이 되심으로써, 우리와 같은 연약함 가운데 시험을 당하시되 죄는 없으신(히 4:15) 오직 한 분이신 중보자가 되시고(딤전 2:5) 지극히 크신 이의 보좌 우편에 앉아 우리를 위하여 유

일한 중보 기도를 드리시는 대제사장이 되심이다(히 7:25; 8:1). 그러므로 “우리에게 있는 대제사장은 우리의 연약함을 동정하지 못하실 이가 아니”며 그로 인하여 우리는 “긍휼하심을 받고 때를 따라 돕는 은혜를” 얻는다(히 4:15-16). 영혼의 구성 요소에 관한 전통적 관점에서의 정(情)은 바로 여기에서 찾을 수 있다.

그리스도께서 겟세마네 동산에 올라 “아버지여 만일 아버지의 뜻이거든 이 잔을 내게서 옮기시옵소서 그러나 내 원대로 마시옵고 아버지의 원대로 되기를 원하나이다”라고 기도하신 장면은(눅 22:42) 그리스도의 구속 사역과 관련하여 그의 의지 및 의식과 관련한 고찰들이 발생하는 부분이기도 하다. 이와 무관하지 않게, 중세에는 칼케돈 신경의 수용 여부를 중심으로 그리스도의 인격 안에 하나의 의지가 있었는지, 혹은 두 개의 의지가 있었는지가 논쟁의 대상이 되기도 하였다. 그리스도의 위격 안에 신성에 속한 의지와 인성에 속한 의지가 별개로 존재하였다는 이성이의론(duothelism, 二性二意論)과 그리스도 안에는 하나의 의지만이 존재했다는 이성일의론(monothelism, 二性一意論)의 대립 결과, 680-681년의 콘스탄티노플 제6차 공의회에서 이성일의론을 이단적 주장으로 정죄하게 된다(Gonzalez, 2010/2012, p. 69). 여기서 그리스도의 의지와 의식의 구체적인 속성들은 논하기 어려운 면이 있지만, 고유하게 인성에 속한 특성으로서의 독립된 의지가 받아들여졌다는 점이 중요할 것이다.[8] 즉, 정통 교리는 그리스도의 인성에 속한 의지가 있었음을 천명하고 있으며, 이는 인간 그리고 특별히 구원 받은 성도들에게 있어서도 예외는 아니다. 위에서 살펴보았듯이 진정한 자유 의지란 하나님의 말씀을 지키며 살아갈 수 있는 능력과 연관이 되는데, 이는 그리스도의 연합으로 말미암아 ‘회복된 하나님의 형상’으로 변화된 인간에게만 가능한 것이다. 곧, 성도는 ‘뜻’을 다하여 하나님과 이웃을 사랑해야 한다(막 12:30-31). 한편 여기서 그리스도께서 인성에 따른 의지로 십자가를 거부하셨다고만 볼 일은 아니다. 오히려 그리스도께서는 적극적으로 성도들을 위한 대리적 기도와 눈물을 쏟으셨다고 보는 것 또한 가능하다.

그런가 하면 죽음은 영혼과 육체의 분리로 이해된다(요 19:30). 칼빈은 재세례파의 영혼수면설을 반박하며, 죽음 이후에도 영혼은 의식적인 활동을 지속함을 역설한 바 있다. 중간 상태에서, 구원 받은 성도들의 영혼은 그리스도와 함께 ‘낙원’에 거하며(눅 23:43) 그렇지 않은 경우에는 음부에서 고통을 받는다(눅 16:23). 다만 육신의 경우, 그리스도는 사흘 만에 부활하셔서 본래의 몸을 그대로 다시 입으셨지만 인간의 육신은 흙에서 왔으므로 흙으로 돌아가게 된다(창 3:19). 그리스도의 재림 이전에 죽은 몸이 썩어진다는 특성 또한 인간을 인간답게 만드는 요소 중 하나일 것이다.

(4) 부활과 영생

그리스도께서는 사망에 매여 있을 수 없는 분이시므로 하나님께서 그를 살리셨다(행 2:24). 그리스도에게 찾아간 죽음은 본래 인간이 겪어야 할 실제적 죽음이었고 그리스도께서 그 사망 권세를 이기고 부활하셨으므로, 그의 죽음은 어떤 면에서 ‘죽음을 이기는 죽음’이었다. 그 부활의 권능에 참여하는 자들(빌 3:10-11)은 최후의 사망의 다스림을 받지 않는다(계 20:6) 이 부활의 소망이 없다면 성도들이야말로 ‘모든 사람 가운데 더욱 불쌍한’ 자였을 것이다(고전 15:12-19). 그 부활의 소망을 가진 성도들은 하나님을 향한 선한 양심의 간구로(벧전 3:21) 날마다 자기 십자가를 지고 그리스도를 따른다(눅 9:23).

그리스도께서 몸을 가지고 부활하셨듯이(행 2:27), 성도들의 부활도 그러하다. 분명한 육체를 다시 입게 되지만 그 때에는 천사들과 같이 장가도 시집도 가지 않는다(마 22:30). 마지막 때에 그리스도께서 나타나시면 이미 죽었던 자들과 당시까지 살아있던 자들이 모두 부활의 생명을 얻을 것이며(살전

8) 여기서, 바르트가 이야기하였던 신-의식(God-consciousness)이 충만한 인간으로서의 예수를 고찰해보는 일도 가능할 것이다.

3:13-18), 그의 참모습을 그대로 볼 것이기 때문에 우리가 그와 같을 줄을 안다(요일 3:2). 그러므로 복음의 주된 내용은 예수와 부활이다(행 17:18).

부활한 하나님의 자녀는 다시 죽을 수 없는 영생을 누린다(눅 20:36). 이러한 영생은 유일하신 참 하나님과 그가 보내신 자 예수 그리스도를 아는 것으로부터 출발한다(요 17:3). 거룩한 성 새 예루살렘에서는 하나님께서 친히 성도들과 함께 계셔서 모든 눈물을 닦아주시고, 다시는 사망도 애통하는 것도 아픈 것도 저주도 밤도 다시는 없을 것이다(계 21-22). 그러한 소망을 품는 '아무도 능히 셀 수 없는 큰 무리가' '빛나고 깨끗한 세마포 옷을 입도록 허락'을 받으니, '이 세마포 옷은 성도들의 옳은 행실'이다(계 7:9; 19:8).

2. '하나님의 형상의 형상'으로서의 인공지능 이해

위에서 살펴본 바와 같이, 구속사적-구원론적으로 성육신한 그리스도의 인성 및 그리스도와 성도의 연합의 교리에 비추어 볼 때 인간과 인공지능 사이에는 무시하기 어려운 존재론적 지위 및 속성의 차이가 있는 것으로 생각된다. 이에, 상술하였듯이 본고에서는 인간을 '하나님의 형상(*imago Dei*)'으로, 인공지능을 '하나님의 형상의 형상(*imago imaginis Dei*)'으로 이해하는 관점을 제안한다. 여기서, '하나님의 형상의 형상'이라는 표현 및 이를 '하나님의 형상'과 대비하는 사고를 본고에서 사상 최초로 제안하는 것은 아님을 밝혀둔다. 근대 이후로는 거의 발견하기 어렵지만, 교부들을 비롯하여 중세까지의 일부 저자들은 이러한 표현을 각자의 다양한 의미로 사용하는 경우가 있었다. 그 사례들을 살펴봄으로써, 본고에서 제안하는 표현의 의미를 보다 명확히 하고자 한다.

첫 번째는 그리스도를 '하나님의 형상(*imago Dei*)'으로, 사람을 '하나님의 형상의 형상(*imago imaginis Dei*)'으로 표현하였던 경우이다. 알렉산드리아의 교부 오리겐(Origen of Alexandria, 184-254)은 '성자는 그 스스로를 나타낼 수 있는 형상으로서 증언'했으며, 사람은 '하나님의 형상의 형상'이라고 주장하였다(Pio, 1531). 히포의 어거스틴(Augustine of Hippo, 354-430)은 그리스도야말로 진정한 '하나님의 형상'이라고 할 수 있으므로 그리스도를 닮은 인간은 오히려 '하나님의 형상의 형상'이라고 표현한 바 있다(Baker, 2014: 120-121). 이와 유사하게 알렉산드리아의 키릴(Cyril of Alexandria, 378-444) 역시 창 1:26-27을 다룸에 있어 삼위일체 하나님의 진정한 형상은 그리스도 안에서 드러나는 것이므로 인간은 '다만 하나님의 형상의 형상'이라고 하였다(Vulcanius, 1605). 두 번째는 그리스도의 성육신한 몸 자체를 '하나님의 형상의 형상'으로 표현하였던 경우이다. 예수회 신학자 나달(Jérôme Nadal, 1507-1580)은 성육신 자체가 일종의 '형상 만들기(image making)'라는 점에서 성육신한 말씀(Word)을 '하나님의 형상의 형상'이라고 표현하였으며, 성육신에 의하여 허용된(licensed) 기도(prayer) 자체가 '하나님의 형상의 형상'의 한 종류라고 표현하였다(Melion, 2008: 95, 102). 세 번째는 인간에 의해 만들어진 '형상'을 '하나님의 형상의 형상'이라고 부른 경우이다. 영국의 청교도 학자이자 킹 제임스 흠정역(King James Version, KJV) 편찬에 참여하기도 하였던 레이놀즈(John Rainolds, 1549-1607)는 '하나님의 형상'인 사람이 사람의 형상대로 만든 것은 '하나님의 형상의 형상'이라고 부를 수 있고 이 또한 하나님의 어떠하심을 약간은 드러내게 되므로, 성화나 성상을 만들 때 있어 주의가 필요함을 역설하였다(Rainolds, 1611).

이렇게 '하나님의 형상의 형상'이라는 표현은 몇몇 저자들이 다양한 의미로 사용하였지만, 그 안에는 다분히 신플라톤주의적인 뉘앙스가 풍기는 듯하다. 말하자면, 일자(一者)로서의 삼위일체 하나님, 그에 가까운 '하나님의 형상', 여전히 하나님과 가깝지만 '하나님의 형상'보다는 거리가 먼 '하나님의 형상의 형상'의 수직적 관계를 상정할 수 있다는 공통점이 있는 것이다.[9] 그런가

하면 위의 견해들은 '하나님의 형상의 형상'을 언급함에 있어서 대부분의 경우 그리스도와 그의 성육신, 그리고 인간을 염두에 두었다는 점에서 기독론적 인간론과 관계가 있다고 할 수 있다. 본고에서 제안하고자 하는 '하나님의 형상의 형상'으로서의 인공지능 이해 또한 이러한 두 공통점 안에 놓여 있는 것으로서, 보다 일반적이고 전통적인 어법을 따라 인간을 '하나님의 형상'으로 파악하고, 이에 따라 인공지능은 인간에 의해 만들어진 마치 인간의 형상과 같은 존재이므로 '하나님의 형상의 형상'으로 이해하자는 것이다. 다만 여기서, 인간을 '하나님의 형상'으로 이해하는 것을 큰 무리 없이 받아들인다고 하더라도 과연 인공지능을 '하나님의 형상의 형상'으로 이해해야만 하는가를 고찰해 보아야 한다. 이에 관하여는 크게 세 가지 경우를 고려해볼 수 있다.

(1) '하나님의 형상'으로서의 인공지능

Foerst (1998)는 인공지능이 '하나님의 형상'으로서의 특성을 지닌다고 언급하는 입장에 선다. MIT에서 개발한 휴머노이드(humanoid) 로봇의 일종인 Cog은 일종의 체화된 인공지능에 해당하는 것으로서, 기능론적으로나 구조론적으로나 인간과 유사한 면이 많다는 것이다. 이러한 존재들이 인간에 의해서 만들어졌을 경우 혹은 다른 원인으로부터 나타날 경우에도 그러한 존재들을 전부 '하나님의 형상'이라고 불러야 할 것인가? 이는 본질적으로 인간 뿐 아니라 영장류들의 행위 또한 '하나님의 형상'이라고 직접 말할 수 있다는 식의 주장들과 그 궤를 같이 하는 것으로서(e.g. Putz, 2009), '하나님의 형상'을 통해 인간론을 정립하고자 하는 시도와는 거리가 멀다. 그런가 하면 인공지능에 대한 과도한 상찬은 자칫 인공물 숭배로까지 이어질 수 있다는 점에서 전통적 기독교의 입장에서는 받아들이기 어려운 견해이다.

(2) '인간의 형상'으로서의 인공지능

Herzfeld (2002)는 인공지능이 '인간의 형상(*imago Homini*)'으로 만들어지고 있음을 여실히 지적하였다. 그에 따르면 '하나님의 형상'은 역사적으로 실체론적, 기능론적, 관계론적 입장에서 논할 수 있는데, 인공지능은 각각 기호주의 및 연결주의와 관련 있는 실체론 및 기능론적인 면에서는 인간과 유사해질 가능성이 있지만, 바르트(Barth)가 주목하였듯이 하나님과의 관계론의 측면에서는 여전히 인간이 보다 우월하다고 주장하는 것이다. Herzfeld의 견해는 타당한 측면이 많지만 인공지능 기술이 오늘날과 같이 발달하기 이전의 상황에서 정리되었음을 상기할 필요가 있을 것이다. 그런가 하면 김동환 (2016b)은 인공지능을 '인간의 형상'으로 이해하는 Herzfeld를 비평한다. 김동환 (2016b)에 따르면, AI를 인간의 형상으로만 파악하는 담론 속에서는 하나님의 형상이라는 신학적 담론이 들어설 자리가 줄어들게 되며, 오히려 인간 지향적인 우상 숭배가 출현할 위험도 있는 것이다(cf. 신 4:16-18). 물론 필자가 보기에는 Herzfeld는 인공지능에 나타난 인간의 나르시시즘을 비판하려는 의도를 지니고 있었지만, 여기서 중요한 것은 '인간의 형상'이라는 표현이 언뜻 생각하기에 '하나님의 형상의 형상'보다 직관적인 듯하여도 그 안에는 무심코 간과하기 쉬운 의미들이 내포되어 있다는 점이다.

(3) '하나님의 형상의 형상'으로서의 인공지능

9) 여기서 그리스도가 '하나님의 형상'으로 표현된다고 하여서 그가 온전한 신성을 갖추지 못했다는 의미가 전혀 아니다. 반면 인간을 '하나님의 형상'이라고 표현하였을 경우에는 그가 신성을 갖추었다는 의미가 전혀 아니다.

그렇다면, 전통적인 관점에서와 같이 '하나님의 형상'으로부터 고유한 인간 본질을 찾으면서도, 인공지능을 적극적으로 규정하고, 위의 두 경우들에서 발생할 수 있는 부작용은 줄일 수 있는 방안을 고려할 필요가 있다. 말하자면 하나님과, 인간과, 인공지능 사이의 종적 상대성을 내포해야 하는 것이다. 이는 인간이 지닌 '하나님의 형상'을 회복시키는 그리스도를 통하여만 가능하다. 그러므로 필자는, 기독론적 인간론의 관점에서 인공지능을 '하나님의 형상의 형상(*imago imaginis Dei*)'으로 이해할 것을 제안하는 바이다.

Ⅴ. 결 론: "사람이 무엇이기에"

"그러나 누구인가가 어디에서 증언하여 이르되 사람이 무엇이기에 주께서 그를 생각하시며 인자가 무엇이기에 주께서 그를 돌보시나이까 (히 2:6)"

히브리서 기자의 주된 관심사 중 하나는 당시 유대인들의 천사 숭배를 경계하는 일이었다. "하나님의 영광의 광채시요 그 본체의 형상"(히 1:3)이신 분이 "천사보다 훨씬 뛰어남"을 논증하면서(1:4), "모든 천사들은 섬기는 영으로서 구원 받을 상속자들을 위하여 섬기라고 보내심"임을 일깨운다(1:14). "장차 올 세상"은 "천사들에게 복종하게 하심이 아니"었다(2:5). 그리스도께서는 성도들을 "형제라 부르시기를 부끄러워하지 아니하시고"(2:11) 성도들을 언약의 "자녀"라 부르셨다(2:13). "이는 확실히 천사들을 붙들어 주려 하심이 아니요 오직 아브라함의 자손을 붙들어 주려 하심"이었다(히 2:16).

오늘날 인공지능에 의한 인간론의 혼란은 어떤 면에서 히브리서 기자가 경계하였던 천사 숭배와 닮아 있다. 천사는 비물리적인 영적 실체로서 위에서 아래로, 인공지능은 물리적인 실체로서 아래에서 위로 인간을 위압한다는 점에서 그 방향성이 다를 뿐이다. 천사와 인공지능은 모두 어떤 면에서는 지능적인 존재로서, 인간보다 뛰어난 능력을 가지는 경우가 많다. 천사의 경우 인간보다 먼저 창조된 존재라고도 볼 수 있으며, 하나님의 뜻을 특별한 방법으로 전하거나(눅 1:26-28), 혹은 하나님의 심판을 대리하는 역할을 수행하기도 한다(삼하 24:16). 인공지능 또한 빠르고 정확한 정보처리 능력을 갖추었으면서도, 지치지 않는 지구력을 가졌을 뿐 아니라 인간의 직업을 빼앗을 만큼의 위협이 되고 있다. 그러나 실상은 성도들이 천사를 판단할 것이며(고전 6:3), "천사들도 살펴보기를 원하는" 복음을 지녔음에 유념해야 한다(벧전 1:12). 그와 마찬가지로, 성도들은 인공지능 역시 다스릴 것이며(창 1:28), "하나님께서 소들을 위하여 염려하심이 아니요 오로지 우리를 위하여 말씀하심"(고전 9:9-10)이었음과 같이 하나님께서는 인공지능이 아닌 인간을 위해 말씀하셨음을 깨달아야 한다.

본고에서는, 적지 않은 한계 안에서 인공지능 기술이 인간론과 '하나님의 형상'에 대하여 던지는 질문에 대한 기독교적 답변을 모색하고자 하였다. 전통적인 관점에 따라 인간의 구성 요소를 영혼과 육체로 보고, 이를 인공지능과 연관 지으며 각각에 대하여 지능과 의지, 철학적 반론과 '체화된 인지'를 위한 몸의 문제를 돌아보았다. 이후 기독론적 인간론의 관점에서 인간이 '하나님의 형상(*Imago Dei*)'임에 주목하고, 인공지능을 '하나님의 형상의 형상(*Imago Imaginis Dei*)'으로 이해할 수 있음을 제안하였다. 우리는 여기서 천사와 인공지능을 비롯한 지능적 존재들 가운데 오직 인간만이 하나님을 사랑하고 하나님께 사랑 받는 자녀로서 언약 체결의 당사자이자, 그리스도와 연합한 존재가 되어, "그리스도와 같은 형상으로 변화하여 영광에서 영광에 이르는" 은혜를 수여받게 됨을 기억하여야 할 것이다.

Soli Deo Gloria in Aeternum !

참고문헌

강아람 (2017). '약함의 신학'으로 보는 선교적 해석학. **선교와 신학, 43**, 11-41.

고영구, 이경기 (2017). 4차 산업혁명시대에 따른 농업·농촌 대응방안 연구: 충북지역을 중심으로. **한국지역경제연구, 38**, 31-57.

권호덕 (2016). 칼빈의 유럽대륙 후예들의 성육신 이해. **조직신학연구, 24**, 66-91.

권호정, 권영선, 장병탁, 송기원, 윤성민, 김민정, 양혁승, 송재용, 이승주, 정병호, 이원재, 윤길원, 김형순, 이종완, 이칠우, 김병연, 김상용, 김현철, 남효순 (2016). **호모 컨버전스: 제4차 산업혁명과 미래사회**. 파주: 아시아.

김광연 (2015). 기독교 전통에 나타난 몸신학과 현대적 몸의 재해석: 몸 주체와 지금 여기에서의 몸. **한국개혁신학, 45**, 44-68.

김광연 (2017). '트랜스휴머니즘과 인간 양식의 변화에 나타난 윤리적 문제들' - 인공지능시대에 삶의 미정성과 유한성이 주는 가치 -. **한국개혁신학, 54**, 135-164.

김동환 (2016a). 테크놀로지 시대의 인간에 대한 신학적 비평. **Canon&Culture, 10**(2), 91-125.

김동환 (2016b). AI(인공지능)에 대한 신학적 담론의 형성 및 방향 모색. **신학연구, 68**, 35-60.

김선희 (1997). 인공지능과 이해의 개념. **인지과학, 8**(1), 37-56.

김선희 (2003). 계산주의 디지털 기억과 개인동일성의 문제: 두뇌의 다운로드를 통하여 우리는 영생할 수 있는가? **철학, 77**, 125-145.

김소영, 김우재, 김태호, 남궁석, 홍기빈, 홍성욱 (2017). **4차 산업혁명이라는 유령: 우리는 왜 4차 산업혁명에 열광하는가**. 서울: 휴머니스트.

김재인 (2017). **인공지능의 시대, 인간을 다시 묻다**. 서울: 동아시아.

김태희 (2017). 인공지능의 몸: 현상학적 고찰. **철학과 현상학 연구, 73**, 99-134.

김환석 (2017). 인공지능 시대를 보는 이론적 관점들. **사회와 이론, 31**, 41-62.

나일주 (2010). **교육공학 관련 이론**. 파주: 교육과학사.

나중규, 김종달 (2017). 4차 산업혁명 논의의 비판적 고찰: 루이스 멈포드의 제도론의 관점에서. **사회과학연구 56**(2), 389-419.

나현규 (2015). 판소피아와 교육. 서울: 학지사.

문병호 (2013). **30주제로 풀어 쓴 기독교 강요** (수정 증보판). 서울: 생명의 말씀사.

문병호 (2015). **칼빈신학: 근본 성경교리 해석**. 서울: 지평서원.

문병호 (2016). **기독론: 중보자 그리스도의 인격과 사역**. 서울: 생명의 말씀사.

박문수 (2017). '제4차 산업혁명 담론'의 실상과 허상. **가톨릭 평론, 10**, 11-19.

박성익, 임철일, 이재경, 최정임, 임정훈, 정현미, 송해덕, 장수정, 장경원, 이지연, 이지은 (2012). **교육공학의 원리와 적용**. 파주: 교육과학사.

백종현 (2007). **철학의 개념과 주요문제**. 서울: 철학과 현실사.

백종현 (2015). 인간 개념의 혼란과 포스트휴머니즘 문제. **철학사상, 58**, 127-153.

성종현 (2012). 인간의 본질과 죽음 그리고 영혼과 육체의 분리 - 신약성서의 개인적·내세적 종말론의 논쟁점을 중심으로 -. **장신논단, 44**(1), 59-86.

성태제 (2017). 제4차 산업혁명시대의 인간상과 교육의 방향 및 제언. **교육학연구, 55**(2), 1-21.

손영화 (2016). 인공지능(AI) 시대의 법적 과제. **법과 정책연구, 16**(4), 305-329.

송하석 (2000). 중국어 방 논변과 인공지능. **철학적 분석, 2**, 145-168.

오윤호 (2014). 새로운 인간 종의 탄생과 진화론적 상상력: 『프랑켄슈타인』과 〈트랜센던스〉를 중심으로. **대중서사연구, 20**(3), 339-366.

이경민, 윤혜경 (2017). 4차 산업혁명 시대 교육혁명 담론의 비판적 분석을 통한 유아교육적 함의. **유아교육연구, 37**(4), 137-155.

이만종 (2017). 지능형정보화 시대의 테러유형과 대응방안: 인공지능에 기반한 테러중점. **국방연구, 60**(1), 93-129.

이봉재 (2005). 인공지능과 인간중심주의 - 인공지능의 연구방법론에 대한 철학적 해석. **대동철학, 31**, 81-98.

이상웅 (2016). 최홍석의 개혁주의 인간론 고찰 - 하나님 형상(Imago Dei)론을 중심으로. **한국개혁신학, 52**, 47-87.
이은경 (2016). 포스트휴먼을 위한 교육 가능성 탐색. **신학과 세계, 86**, 83-111.
이종관 (2017). **포스트휴먼이 온다**. 고양: 사월의책.
이초식 (1993). **인공지능의 철학**. 서울: 고려대학교 출판부.
정대현 (2017). 알파고: 나는 자연종 인간과 둔 바둑을 이겼다. **과학철학, 20**(3), 1-30.
장병탁 (2018). 인간지능과 기계지능 - 인지주의 인공지능. **정보과학회지, 36**(1), 17-26.
정진명 (2018). 인공지능에 대한 민사책임 법리. **재산법연구, 34**(4), 137-168.
조태호 (2017). **모두의 딥러닝**. 서울: 길벗.
차두원, 김서현 (2016). **잡 킬러: 4차산업혁명, 로봇과 인공지능이 바꾸는 일자리의 미래**. 서울: 한스미디어.
허희옥, 양은주, 김다원, 문용선, 최종근 (2017). 인공지능 시대의 인간 지능와 학습. **교육철학연구, 39**(1), 101-132.
홍일립 (2017). **인간 본성의 역사**. 서울: 에피파니.

Baker, D. N. (2014). *Julian of Norwich's" Showings": From Vision to Book*. Princeton University Press.
Bang, S. H. (2014). Thinking of Artificial Intelligence Cyborgization With a Biblical Perspective (Anthropology Of The Old Testament). *European Journal of Science and Theology, 10*(3), 15-26.
Bell, F. (2010). Network theories for technology-enabled learning and social change: Connectivism and actor network theory. In *Networked learning conference 2010: Seventh international conference on networked learning*.
Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience*. 이을상, 하일호, 신현정, 오용득, 박만준, 안호영 (공역) (2013). **신경 과학의 철학: 신경 과학의 철학적 문제와 분석**. 서울: 사이언스북스.
Bruner, J. (1996). *The Culture of Education*. 강현석, 이자현 (역) (2005). **교육의 문화**. 서울: 교육과학사.
Calvin, J. (1536). *Institutio Christianae Religionis* (1st ed.). 문병호 (역) (2009). **라틴어직역 기독교강요** (초판). 서울: 생명의 말씀사.
Calvin, J. (1559). *Institutio Christianae Religionis* (Final ed.) 김종흡, 신복윤, 이종성, 한철하 (공역) (1988). **기독교 강요** (최종판). 서울: 생명의 말씀사.
Comenius, J. A. (1657). *Didactica Magna*. 정일웅 (역) (2015). **요한 아모스 코메니우스의 대교수학: 모든 사람에게 모든 것을 가르치는 완전한 기술**. 서울: 나눔사.
Cortez, M. (2016). *Christological Anthropology in Historical Perspective: Ancient and Contemporary Approaches to Theological Anthropology*. Zondervan.
Cyrillus Alexandrinus, C. Alexandrinus, & Vulcanius, B. (1605). *Cyrilli Archiepiscopi Alexandrini Adversus anthropomorphitas : liber unus graecè & latinè ejusdem, De incarnatione unigeniti, et Quod unus sit Christus ac Dominus secundum Scripturas, ad Hermiam dialogi duo, nunquam antehacediti, interprete Bonaventura Vulcanio, cum notis ejusdem*. E Typographia Joannis Patii.
Delio, O. S. F. (2013). Artificial Intelligence and Christian Salvation: Compatibility or Competition? *New Theology Review, 16*(4).
Domingos, P. (2015). *The master algorithm: How the quest for the ultimate learning machine will remake our world*. NY: Basic Books.
Dreyfus, H. L., & Dreyfus, S. E. (1986). From Socrates to expert systems: The limits of calculative rationality. *In Philosophy and Technology II* (pp. 111-130). Springer, Dordrecht.
Eggen, P., & Kauchak, D. (2010). *Educational Psychology: Windows on Classrooms* (8th ed.) 신종호, 김동민, 김정섭, 김종백, 도승이, 김지현, 서영석 (공역) (2011). **교육심리학: 교육**

실제를 보는 창. 서울: 학지사.
Foerst, A. (1998). Cog, a humanoid robot, and the question of the image of god. *Zygon®, 33*(1), 91-111.
Gonzalez, J. L. (2010). *The Story of Christianity, Vol. 1: The Early Church to the Dawn of the Reformation*. 엄성옥 (역) (2012). **중세교회사**. 서울: 은성.
Graham, & George (2017). Behaviorism. In Zalta, E. N. (Ed.), *Stanford Encyclopedia of Philosophy*. [Retrieved from https://plato.stanford.edu/archives/spr2017/entries/behaviorism/]
Herzfeld, N. L. (2002). *In Our Image: Artificial Intelligence and the Human Spirit*. MN: Augsburg Fortress.
Hoekema, A. A. (1986). *Created in God's Image*. 이용주 (역) (2012). **개혁주의 인간론**. 서울: 부흥과 개혁사.
Humphrey, N. (2006). Consciousness: the Achilles heel of Darwinism? Thank God, not quite. In Brockman, J. (Ed.) *Intelligent Thought. Science versus the Intelligent Design Movement*, 50-64. NY: Vintage.
Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford university press.
Leidenhag J. (2017) Uniqueness and the Presence of the Image: Towards a Pneumatological Foundation for Human Uniqueness and the Image of God. In Fuller M., Evers D., Runehov A., & Sæther KW. (Eds), *Issues in Science and Theology: Are We Special?. Issues in Science and Religion: Publications of the European Society for the Study of Science and Theology, vol 4,* 255-270. Springer.
Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed. An argument for AI in education*. London: Pearson.
McCarthy, J., & Hayes, P. J. (1981). Some philosophical problems from the standpoint of artificial intelligence. In *Readings in artificial intelligence* (pp. 431-450).
McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics, 5*(4), 115-133.
Melion, W. (2008). Ex libera meditatione: Visualizing the Sacrificial Christ in Jerónimo Nadal's 'Annotations and Meditations on the Gospels'. In Kupfer, M. A. (Ed.), *The Passion Story: From Visual Representation to Social Drama*, 91-118.
Noll, M. A. (2011). *Jesus Christ and the Life of the Mind*. 박규태 (역) (2015). **그리스도와 지성: 학문 연구를 위한 기독론적 토대와 방법**. 서울: IVP.
Pio, A. (1531). *Tres et viginti libri in locos lucubrationum variarum D. Erasmi Roterodami, quos censet ab eo recognoscendos et retractandos*. Badius.
Putz, O. (2009). Moral apes, human uniqueness, and the image of God. *Zygon®, 44*(3), 613-624.
Rainolds, J. (1611). *Censura librorum apocryphorum Veteris Testamenti, adversum Pontificios, inprimis Robertum Bellarminum: Qva Tum Divina Et Canonica Sacrae Scripturae autoritas asseritur solidissime: tum variae Quaestiones & Controversiae, tam dogmaticae, quam historicae, inprimis quae est de Duratione Monarchiae Persicae, & de 70. Hebdom. Danielis expediuntur accuratissime. 2* (Vol. 2). Hulsius & Laurentius.
Rapaport, W. J. (1988). Syntactic semantics: Foundations of computational natural-language understanding. In Fetzer, J. H. (Ed.), *Aspects of Artificial Intelligence* (pp. 81-131). Dordrecht: Kluwer Academic Publishers.
Russell, S. J., & Norvig, P. (2010). *Artificial intelligence: a modern approach*. 류광 (역) (2016). **인공지능: 현대적 접근방식**. 파주: 제이펍
Searle, J. R. (1984). *Minds, brains and science*. Harvard University Press.
Stahl, G. (2013). Theories of cognition in collaborative learning. In Hmelo-Silver, C. E., Chinn, C. A., Chan, C. K. K., & O'Donnel, A. M. (Eds.), *The international handbook of*

collaborative learning, (pp. 74-90). Routledge.
Sun, R. (2000). Artificial intelligence: Connectionist and symbolic approaches.
Turing, A. M. (2009). Computing machinery and intelligence. In *Parsing the Turing Test* (pp. 23-65). Springer, Dordrecht.
Turkle, S. (1988). Artificial intelligence and psychoanalysis: A new alliance. *Daedalus, 117*(1), 241-268.
Vulcanius, B. (1605). *Cyrilli Archiepiscopi Alexandrini Adversus anthropomorphitas: liber unus græcè et latinè ejusdem, De incarnatione unigeniti, et Quod unus sit Christus ac Dominus sec. Scripturas, ad Hermiam dialogi II interpr. Bonav. Vulcanio, cum notis ejusdem*. Lugduni Batavorum: E Typographia Joannis Patii.
Ware, K. (1979) *The Orthodox Way*. 엄성옥 (역) (2002). **정교회의 길**. 서울: 은성.
Warwick, K., & Shah, H. (2016) Can machines think? A report on Turing test experiments at the Royal Society, *Journal of Experimental & Theoretical Artificial Intelligence, 28*(6), 989-1007. DOI: 10.1080/0952813X.2015.1055826
Winograd, T., & Flores, F. (1986). *Understanding computers and cognition: A new foundation for design*. Intellect Books.