

한국윤리학회

2022년 춘계 워크숍

일시 | 2022년 5월 21일(토)

오후 2시 - 오후 5시 5분

방식 | 코로나 19 로 인한 ZOOM 화상회의

주최 | 한국윤리학회

프로그램

사회 : 강철 (서울시립대)

14:00 - 14:05	개회사 및 회장 인사	김수정 (가톨릭대)
14:05 - 14:10	연구윤리교육	최경석 (이화여대)

14:10 - 15:30 **“호네트의 사회적 자유 개념”**
발표: 강병호 (서울과기대) (40분)
논평: 오근창 (성균관대) (20분)
참석자 토론 (20분)

Break Time (10분)

15:40 - 17:00 **“인공지능의 도덕적 지위와 현상적 의식”**
발표: 정태창 (공주대) (40분)
논평: 김은희 (경인교대) (20분)
참석자 토론 (20분)

17:00 - 17:05 폐회사

호네프의 사회적 자유 개념*

강 병 호**

【주제분류】사회정치철학, 비판이론, 현대유럽철학

【주요어】호네프, 비판이론, 사회주의, 소극적 자유, 성찰적 자유, 사회적 자유

【요약문】『자유와 권리』(2011) 이후 호네프는 “사회적 자유”를 사회비판의 “선도 이념”(Leitidee)으로 제안한다(IS: 154). 그런데 이 개념은 호네프도 인정하다시피 오늘날의 우리에게 “매우 특이하고 이상해 보일 수 있다”(RF: 94). 이 논문은 이러한 사회적 자유 개념을 가능한 한 명료하게 체계적으로 분석·해명하고, 이 개념의 타당성을 검토하는 것을 주된 목적으로 한다. 먼저 이 개념과 호네프가 구상하는 사회비판과의 관계를 밝힌다(1). 이어서 소극적 자유와 성찰적 자유에 대한 호네프의 이해를 살펴본다(2). 이제 이를 배경으로 해서 한편으로는 (일상의 예들도 활용하여) 호네프의 사회적 자유 개념을 가능한 한 우리들의 직관에 친숙하고 그럴 듯하게 제시하고, 다른 한편으로는 그러면서도 이 자유 개념의 특성을 분석적으로 명료하게 드러낸다(3). 마지막 절에서는 두 가지 점에서 호네프의 사회적 자유 개념의 난점을 검토한다. 하나는 연대 개념과의 관계이며, 다른 하나는 롤즈의 “합당한 다원주의의 사실” 논제를 받아들일 때 호네프의 사회적 자유 기획이 마주하게 되는 어려움이다(4).

1. 비판이론과 사회적 자유

비판이론의 주요한 과제는 사회가 규범적으로 개선될 수 있는 전망을 제시하는 것이라고 할 수 있다. 그런데 그것은 단순히 바람직한 이상이나 기준을 제시하는 것 이상이어야 한다. 제시된 이상이나 기준이 그 자체로 바람직하고 고상하기는 하나 정작 비판의 대상이 되는 사회와 접촉면이 너무 없다면, 그런 비판은

* 이 논문은 2020년 대한민국 교육부와 한국연구재단의 지원을 받았다(NRF-2020S1A5B5A16083644).

세 명의 심사자의 논평은 논문을 개선하는 데 큰 도움이 되었다. 진심으로 감사드린다.

** 서울과학기술대학교 강사

그 사회구성원들에 의해 납득되거나 받아들여지기 힘들고, 따라서 실천적으로 공허하거나 때로는 의도와 달리 해로운 결과를 초래할 수도 있기 때문이다. 이러한 “외재적 비판”(external criticism)보다는 해당 사회에서 이미 확고하게 자리 잡고 있는 규범적 원칙이나 가치에 호소하는 “내재적”(immanent) 접근이 규범적으로도 실천적으로도 훨씬 더 효과적일 수 있다.¹⁾ 다만 이때 활용되는 원칙과 가치는 기존의 관행과 제도를 넘어서는 전망을 제시하기에 충분한 규범적 잠재력을 담보하고 있어야 할 것이다. 악셀 호네프는 이러한 “헤겔 좌파적” 전통을 따라 자신의 비판이론을 발전시켜 왔는데, 이전의 “인정이론”에 비해 여러 면에서 새로워진 『자유의 권리』(*Das Recht der Freiheit* [RF], 2011)와²⁾ 그 후속 작업들에서³⁾ “사회적 자유”(soziale Freiheit)를 내재성과 비판적 잠재력을 함께 갖춘, 사회비판의 “선도 이념”(Leitidee)으로 제안한다(IS: 154).

자유가 프랑스 혁명 이후로 현대사회의 정당성의 기반이 되는 규범 원칙임은 부정하기 어렵다. 그렇지만 기본적 자유의 법적 보장이 어느 정도 성취된 사회에서도 여전히 자유가 자본주의적·자유민주주의적 질서의 변혁을 위한 생동하는 규범적 원천이 될 수 있다는 주장은 상당히 논쟁적이다. “신자유주의”라는 표현이 오늘날의 사회현실을 비판적으로 보는 사람들에게 갖는 부정적 어감에서 드러나듯이, 자유 개념은 많은 사람들에게 고삐 풀린 경쟁과 이기심의 논리를 위장하는 의심스러운 가치가 되어 버렸다. 그러나 거꾸로 그렇게 생각하는 경향이야말로 우리 시대 자유 이해의 “빈곤함”의 징표일 수도 있다.⁴⁾ 호네프에 따르면 자유 원칙의 확립 이래로 자유를 단지 소극적·법적 자유로, 기껏해야 도덕적 자유로 “오해”하는 “뿌리 깊은 경향”이 있어 왔다(RF: 223). 그러나 자유에 대한 다

1) “외(재)적 비판”과 “내(재)적 비판”에 대해서는 다음을 참조: Honneth (2007[2000]), 앞짜 2007[1987], 1장과 2장.

2) 호네프의 기존 작업의 연장선에서 『자유의 권리』가 갖는 “혁신적 종합”의 면모에 대해서는 강병호 (2021), pp. 145-149 참조.

3) *Die Idee des Sozialismus* (2015)과 *Die Armut unserer Freiheit* (2020). 앞의 책에서 호네프는 사회적 자유를 재생될 가치가 있는 사회주의의 이념으로 제시하면서 이 개념의 변혁적 잠재력을 보여주려고 한다. 뒤의 책은 사회적 자유라는 주제를 중심으로 묶어진 논문집이다. 나의 현재 작업을 위해서는 앞의 책과 뒤의 책에 실려 있는 한 논문, 「두 가지가 아니라 세 가지 자유 개념」(“Drei, nicht zwei Begriffe der Freiheit”)이 특히 중요하다. 앞으로 본문에서 앞의 책을 ‘IS’로, 뒤의 논문을 ‘Drei’로 표기하겠다.

4) Honneth (2020)의 제목이 “우리들의 자유의 빈곤”이다.

른 이해가 있고, 우리가 제대로 알아채지 못해서 그렇지 그러한 “제3의” 자유가 사실은 우리의 삶과 제도를 지탱하고 있는 실제 가치임을 드러낼 수 있다면 사정은 달라진다. 호네프가 새롭게 상기시키고 싶어하는 그러한 자유 개념이 바로 “사회적 자유”이다.

이 논문은 호네프가 소극적(negativ)·법적 자유와 성찰적(reflexiv)·도덕적 자유 개념 옆에 “제3의” “독자적 범주”로(Drei: 140, 149) 제 자리를 찾아주려고 하는 사회적 자유 개념을 가능한 한 명료하게 체계적으로 분석·해명하고, 그 개념의 타당성을 검토하는 것을 주된 목적으로 한다. 그러나 본격적으로 그러한 분석 작업에 착수하기 전에 어떻게 사회적 자유가 비판적 사회이론의 “선도 이념”의 역할을 할 수 있는지를 살펴보자.

사회적 자유는 한편에서는 푸리에(Charles Fourier)와 푸루동(Pierre-Joseph Proudhon) 등의 초기 사회주의 사상 전통과 다른 한편에서는 헤겔, 맑스로 이어지는 헤겔 좌파 전통에서 발전된 자유 개념으로서, “사회주의의 가장 핵심적인 윤리적 동인”이었다”(Drei: 154). 오늘날에는 거의 잊혀진 전통에 속하고, 그래서 우리에게 “매우 특이하고 이상해 보이기까지 할 수도 있는” 자유 개념이지만(RF: 94), 호네프의 주장에 따르면 실제로 있어서는 소극적, 성찰적 자유 개념보다 우리들의 “선이론적 직관 및 사회적 경험과 훨씬 더 일치하는 바가 많으며”(RF: 113), 우리가 제대로 이해한다면 사랑과 가족, 시장, 민주주의 같은 현대 사회의 제도의 형성에 “가장 큰 영향”을 끼쳐온 자유 개념이다(RF: 222). 이 자유 개념을 새롭게 발굴하고, 이 자유 개념에 비추어 사회를 “재구성”하는 것은⁵⁾ 다음과 같은 이유에서 실천적 함축을 갖는 사회비판의 기획이 될 수 있다.

첫째로 앞에서 언급했듯이 현대사회에는 자유를 소극적 자유로만 이해하는 강한 경향이 있으며, 이런 경향을 뒷받침하고 강화하는 지배적 지식 체계가 있다. 이러한 지적 헤게모니의 지원을 받아 사회는 소극적 자유라는 협소한 개념틀에 맞춰 이해되고 구성되고 재생산되고 있다. 이런 상황에서 사회적 자유 개념을 소극적 자유, 성찰적 자유와 더불어 “독자적인”, 나아가 훨씬 더 풍부하고 추구할 만한 자유 개념으로 제시하는 이론 작업은 기존의 지배적인 지식 체계에 대한 비

5) 호네프가 『자유와 권리』에서 제안하고 실행하는 “재구성”이라는 연구방법에 대해서는 강병호 (2021), pp. 149ff. 참조.

판 작업이면서(IS: 140, Fn. 141 참조), 동시에 현실을 다르게 해석할 수 있는 문화적 자원과 “정치적 해석 유형”을 제공하는 실천이 된다(IS: 17-19 참조).

둘째로 앞으로 제대로 설명되어야 할 것이지만, 사회적 자유 개념은 그 이름이 암시하듯이 개인의 자유 영역이나 능력보다 사람들 사이의 상호협력과 보완의 관계를 강조한다. 그리고 그러한 협력과 보완은 “제도”를 통해서만 안정화될 수 있다고 본다. 그러니까 사회적 자유 개념은 사람들의 상호작용이 보완과 협력이 되도록 매개하는 제도와 관행을 자유 실현의 “필수” 조건으로 포함한다(RF: 123). 따라서 사회적 자유 개념에 비추어 사적 관계와 경제, 정치 영역을 분석하는 것은 자연스럽게 그러한 영역이 사회적 자유를 실현하는 곳으로 변화하기 위해서 필요한 제도적 개선 방향을 제시하는 것으로 이어지고, 나아가 “사회적 자유의 실현 조건을 위한 투쟁”으로 사람들을 동기부여하는 정치적 기획이 될 것이다(IS: 135, 137).

이상에서 드러나듯이 호네프가 『자유와 권리』 이후 추구하는 사회비판이론의 성패는 이 사회적 자유 개념이 얼마나 건실하고 사람들에게 매력적인가에 상당 정도 달려있다. 사회적 자유가 독자적 자유 범주로, 그리고 사회 현실을 평가할 때 사람들이 선호하고, 그것의 실현을 위해 나설 수 있는 매력적인 가치로 받아들여질 때 호네프의 사회적 자유 기획은 성공할 수 있다. 그런데 이렇게 중요한 사회적 자유 개념은 호네프가 시인하다시피 오늘날에는 이상하게 느껴질 정도로 낮은 개념이다. 그래서 이 논문은 한편으로는 (일상의 예들도 활용하여) 호네프의 사회적 자유 개념을 가능한 한 우리들의 직관에 친숙하고 그럴 듯하게 제시하고, 다른 한편으로는 그러면서도 이 자유 개념의 특성을 분석적으로 명료하게 드러내려고 한다(3). 이를 위한 배경으로서 먼저 소극적 자유와 성찰적 자유에 대한 호네프의 이해를 살펴본다(2). 마지막으로 두 가지 점에서 호네프의 사회적 자유 개념의 난점을 검토한다. 하나는 연대 개념과의 관계이며, 다른 하나는 롤즈의 “합당한 다원주의의 사실” 논제를 받아들일 때 호네프의 사회적 자유 기획이 마주하게 되는 어려움이다(4).

본격적인 논의로 들어가기에 앞서 한 가지 유보 사항을 분명히 하는 것이 좋을 것 같다. 호네프가 사회적 자유 개념을 발굴하고 재구성할 때 가장 크게 의존하는 이론적 자원은 헤겔의 『법철학』⁶⁾이다. 여기서 헤겔은 ‘사회적 자유’라는 용어를 직접 사용하지는 않는다. 사회적 자유는 헤겔의 “실체적 자유”(substantielle

Freiheit, §§149, 257) 또는 “객관적 자유” 개념을 재구성한 것이다.⁷⁾ 그런 만큼 호네트의 사회적 자유 개념이 헤겔에 대한 얼마나 적절한 해석인지도 매우 흥미로운 사항이지만, 논의의 분량과 편의를 위해 여기서는 이 점은 전혀 고려되지 않는다.⁸⁾ 헤겔뿐만 아니라 이샤야 벌린이나 맑스 등 호네트가 사회적 자유 개념을 계발(啓發)하기 위해서 활용하는 이론가들에 대한 해석의 적절성은 전혀 따지지 않는다. 다른 이론가들에 대한 해석은 호네트가 자신의 사회적 자유 개념을 제시하기 위한 수단으로 여겨지며, 이 논문이 집중하는 것은 ‘호네트의’ 사회적 자유 개념이다.

2. 소극적 자유와 성찰적 자유

호네트에 따르면 서구 사회에서는 근대 이후 뚜렷이 구별되는 세 가지 자유 모델이 형성되어 왔다. 호네트는 그것들을 각각 “소극적”, “성찰적”, “사회적” 자유라고 부른다.⁹⁾ 여기서는 먼저 호네트가 생각하는 소극적 자유와 성찰적 자유를 설명하고, 이를 배경으로 다음 절에서 본격적으로 사회적 자유 개념의 특성을 밝히려 한다.

일단 오해를 방지하기 위해서 두 가지를 분명히 하는 것이 좋겠다. 우선 ‘사회적’ 자유는 ‘개인적’ 자유의 반대 개념이 아니다. 사회적 자유는, 벌린의 적극적 자유의 집단적 형태처럼¹⁰⁾ 자유의 주체를 개인이 아니라 어떤 집단으로 상정하지 않는다(Drei: 147-148). 소극적, 성찰적 자유처럼 사회적 자유도 ‘개인의’ 자유에 대한 하나의 이해방식이다. 다만 사회적 자유 모델은 자유의 상호주관적 성격을 강조한다. 이런 점에서 사회적 자유에 대립하는 것은 개인의 자유가 아니라

6) Hegel (1986[1821]).

7) Neuhausser (2000), pp. 5, 82ff.

8) 헤겔 『법철학』에 대한 호네트의 “재구성적” 해석으로는 다음을 참조: 호네트 (2017[2001]), Honneth (2010), (2020b), (2020c). 이에 대해 비판적 입장은 Pippin (2014), 서정혁 (2018).

9) 이러한 구분은 즉시 벌린의 유명한 소극적/적극적 자유 구분을 연상시키고 실제로 많은 유사점이 있지만, 그보다는 헤겔이 『법철학』에서 각각 ‘추상법’, ‘도덕’, ‘인륜’으로 펼친 자유를 재구성한 것이다.

10) Berlin (2002[1958]), pp. 191-200.

자유를 “단지 개인주의적으로” 이해하는 방식이다(IS: 38).¹¹⁾ 다음으로 사회적 자유 모델은 소극적 자유와 성찰적 자유를 대체하려고 하지 않는다. 오히려 이 두 모델이 현대사회에서 갖는 정당한 자리와 가치를 인정하고, 그것들 각각의 합리적 핵심을 수용하고 “확장”한다. 이런 의미에서 사회적 자유 개념은 다른 두 자유 개념의 “종합”이라고 할 수 있다(Drei: 150).

1) 소극적·법적 자유

소극적 자유는 현대인에게 가장 친숙한 자유 개념일 것이다. 이 친숙함과 설득력은 한편으로는 이 자유 모델이 인간을 “원자”로 보는 근대의 개인주의 정신에 잘 부합하기 때문이고, 또 부분적으로는 자유에 대한 단순명료한 정의(定義) 덕분이다. 자유에 대한 소극적 개념 규정은 흄으로 소급된다. 흄은 『리바이던』 제21장에서 자유를 “저항의 부재”로 정의한다. “여기서 저항이란 외부적 장애를 말한다.” 자유란 “자기가 하고자 하는 것을 방해받지 않는” 것이다.¹²⁾ 그런데 자유주의 전통에서 자유에 대한 ‘외부의 방해’는 보통 다른 사람의 행위에서 기인하는 것으로 여겨진다. 그래서 소극적 자유에 대한 전형적인 정의는 “다른 사람의 자의적 간섭”이 없는 것이다.¹³⁾ 이 소극적 자유는 현실에서 법적 권리라는 제도를 통해 보장된다(RF: 221).

자유 규정의 이런 단순명료함은 소극적 자유 개념에 직관적 호소력을 더해주지만, 다른 한편 이 자유 모델이 전제하는 인간상의 얇음 또한 드러낸다. 이런 점에서 흄스가 자유를 외적 장애의 부재로 정의한 다음 곧바로 이어서 “이것은 이성 있는 인간에게만 적용되는 것이 아니라 이성이 없는 생물이나 무생물에게도 적용된다”고 덧붙이는 것은 무척 시사적이다.¹⁴⁾ 자유의 전제는 “운동” 능력 뿐이라는 것이다. 자유의 주체에게 전제되는 것이 이렇게 얇기 때문에 쉽게 자유

11) 글자만 보자면 호네프가 “사회적 자유”를 “개인의 자유”(individuelle Freiheit)와 대립시키는 경우도 가끔 있다. 그러나 맥락을 보면 이때 “individuelle Freiheit”는 ‘단지 개인주의적으로 이해된 자유’를 의미한다.

12) 흄스(2008[1651]), pp. 279, 280.

13) Berlin (2002[1958]), p. 169: “the deliberate interference of other human beings”. 소극적 자유 개념의 부족함 점에 대해서는 Petti (2014), ch. 2 참조.

14) 흄스(2008[1651]), p. 279.

에 대한 방해와 간섭은 모두 “외부”에서 오는 것으로 여겨지게 된다.¹⁵⁾ 이것을 인간에 한정해서 말하자면 소극적 자유 모델이 전제하는 것은 무언가를 욕구하고, 욕구하는 바를 행하려는 능력뿐이다. 외부의 방해와 간섭 말고 인간 내부의 (불안, 의지박약, 자신감 부족 같은) 심리적 요인 또한 인간을 자유롭지 못하게 만들 수 있다는 사정은 고려되지 않는다. 마찬가지로 사람들이 진정 자신이 원하는 것이 무엇인지를 알기가 그렇게 쉽지 않다는 것, 무엇을 욕구한다고 해도 그것이 그 사람의 진정한 욕구가 아닐 수 있다는 사실은 무시된다. 다시 말해서 자주 지적되듯이 소극적 자유 모델은 사람이 진정 원하는 것을 알고 그것을 욕구하기 위해서는, 즉 자유롭기 위해서는 단순히 욕구 능력과는 다른 능력, 이성의 성찰 능력 또한 요구된다는 사실을 놓치고 있다.¹⁶⁾

소극적 자유 개념이 자유의 전제에 행위자의 성찰 능력을 포함시키지 않는 것은 자연스럽게 행위자의 의도와 목적의 질(質)을 따지지 않는 것으로 귀결된다. 이 자유 모델에는 행위자가 내세우는 의도와 진정한 의도를 구분할 기반이 없기 때문이다. 의도와 목적의 질을 따지지 않는 것은 다시금 자신의 의도와 목적을 다른 사람들에게 정당화할 필요가 없는 것으로 귀결된다. 이제 자유란 (다른 사람의 자유를 침해하지 않는 한) 행위자의 의도와 목적이 무엇이든 그 사람의 행위가 방해받지 않는 것, 다시 말해서 행위자가 정당화 부담 없이 “자기 마음대로” 할 수 있는 ‘자기만의 공간’을 확보하는 것이다.¹⁷⁾

소극적 자유 개념이 포착해 낸, 정당화 압력에서 벗어난 간섭 없는 이 자기만의 공간은 부정적인 면과 긍정적인 면을 모두 갖고 있다. 부정적으로 보면 이 공간은 개인들이 순전히 자신들의 이익 계산에 따라 “전략적으로” 행동하는 것이 허락된 “사적 이기주의”의 공간이다(RF: 55; IS: 31). 그러나 동시에 이 공간은 개인들이 사회와 여론의 정당화 압박에서 벗어나 그것이 무엇이든 자신이 해보고 싶은 것을 해보고 자기가 살고 싶은 삶을 실험해 수 있는 무중력의 공간이기

15) 홉스(2008[1651]), p. 279: “운동을 가로막는 장애가 물체 그 자체의 구조 내에 있을 경우에는 자유가 없다고 말하는 것이 아니라, 움직일 힘이 없다고 말하는 것이 보통이다.”

16) Taylor (1985), pp. 211-229.

17) 호네프는 자유의 내적 장애 요인을 무시하는 것과 목적의 질적 차이를 인정하지 않는 것 사이의 내적 연관 관계를 홉스의 자유 개념과 사르트르의 자유 개념에 입각해 잘 보여준다(RF: 45-50).

도 하다. 개인이 자신의 “특별함”을 추구할 수 있는 공간(RF: 47, 50). 호네프는 이 후자의 측면에서 소극적 자유에 특별한 가치가 있다고 본다(RF: 50). 소극적 자유를 한낱 “미개한”(primitiv) 자유 개념으로 여기는 것은 “설부른” 것이다. 거기에는 “개인이 외적 제한 없이 그리고 동기 점점의 강제 없이 ‘마음대로’ 행동할 수 있는 권리를 누려야 한다는 근대의 도덕적 자기이해가 표현되어 있다. 소극적 자유는 이러한 자기이해의 독특하고 포기될 수 없는 요소이다”(RF: 58; RF 137-138도 참조).

2) 성찰적·도덕적 자유

소극적 자유가 ‘외부의’ 간섭을 제거하는 데 집중한다면, 성찰적 자유는 주체의 “자기관계”에 주목한다. 주체는 자신의 진정한 의도에 따라 행동할 때만 자유롭다고 여겨진다(RF: 58-59). 이 자유 모델에 따르면 주체가 원하는 것이라고 해서 자동적으로 그의 ‘진정한’ 의도가 되는 것은 아니다. ‘진정한 의도’를 어떻게 파악하느냐에 따라 성찰적 자유 모델은 다시 칸트적 자율성 버전과 헤르더(J. G. Herder)를 필두로 한 낭만주의자들의 자기실현 버전으로 분화된다. 자율성 버전은 인간의 본성을 이성에 의해 인도되는 “의지”와 그렇지 않은 “욕구”로 나누고, 욕구가 아니라 이성과 의지에서 나오는 것을 행위자의 진정한 의도로 파악한다. 이때 의도가 이성의 산물인지 한낱 욕구의 산물인지를 구분할 수 있는 기준은 그 의도를 표현하는 준칙의 보편화 가능성 여부이다. 자기실현 버전은 진정한 의도는 행위자 자신의 “고유하고 진실한 욕구”에서 나오고 그것에 일치하는 것이라고 본다. 따라서 사람은 자유롭기 위해 먼저 자기 자신의 고유한 욕구와 바람을 알아야 하고 그것을 실현할 수 있어야 한다(RF: 67). 이러한 성찰적 자유는 현실에서 도덕이란 제도를 통하여 구현된다(RF: 221).

성찰적 자유의 두 버전 중에서 호네프가 사회적 자유와 집중적으로 대비시키는 것은 칸트의 자율성 버전이다. 호네프는 자율성의 소극적 의미와 적극적 의미를 구분한다. 소극적 의미에서 성찰적·도덕적 자유는 주체에게 사회적 관계와 거기에서 나오는 행위 기대를, 모든 이성적 존재의 동의를 받을 만한 것인가라는 기준에 맞춰 스스로 검토하고, 경우에 따라서는 거부할 수 있는 권한을 부여한다(RF: 179, 180). 일상적 표현을 쓰자면 현대인은 “양심”의 이름으로(RF: 192) 기

성의 지배 질서와 그 질서가 요구하는 것을 비판하고 거부할 자유가 있다. 이러한 도덕적 자유는 기존의 관계와 질서에 대한 “거리두기” 혹은 단절이란 점에서 소극적·법적 자유와 공통점이 있다. 그러나 소극적 자유가 기존의 관계에서 빠져나와 정당화 부담 없이 자기 마음대로 행위할 자유라면, 도덕적 자유는 비판하고 거부하는 이유를 다른 사람들에게 설명하고 정당화할 능력과 자세를 전제한다(RF: 193). 이를 통해 도덕적 주체는 급진적 거리두기와 단절 속에서도 혼자가 아니다. 그는 자신이 제시하는 이유를 이해하고 정당성을 인정해 줄 나머지 인류 공동체가 자신의 편에 서 있다고 믿는다(RF: 190).

성찰적 자유의 적극적 의미는 자율성이, 정당성이 의심스러운 질서에 반대하는 데 머물지 않고 “자기입법”의 이념을 좇아 규범을 새롭게 해석하고, 필요하면 새로운 규범을 만드는 방식으로 “구성적으로” 사용되는 것이다(RF: 190). “법적 자유에서 우리는 방해 없이 우리의 삶을 바꿀 권리만을 갖는다면, 도덕적 자유에서 우리는 도덕 규범에 대한 공적 해석에 영향력을 행사할 권리를 갖는다”. 이렇게 해서 도덕적 자유는 기존의 사회질서를 “변화시키는 힘”이 된다(RF: 205). 호네프는 바로 이러한, “급진적 거리두기‘와’ 일반적 동의의 결합, 모든 기성의 것에서 벗어나‘과’ 동시에 보편주의적 합의의 결합”이, 성찰적·도덕적 자유 개념이 현대사회에서 갖는 “매력”과 “가치”의 실체라고 생각한다(RF: 189, 190).

3. 사회적 자유

소극적 자유가 단지 외적 방해와 간섭만 없으면 자유가 실현되었다고 간주함으로써 행위자의 의도가 진정 자유롭게 형성된 것인지를 놓치고 만다면, 성찰적 자유는 행위자의 성찰적 “능력”에만 몰두한 나머지, 자율적으로 선택된 행위 의도가 현실에서 실제로 실행될 수 있는지의 여부에 전혀 주목하지 않는다. 따라서 성찰적 자유는 그저 의지의 자율적 결정에 만족하고 만다(RF: 83, 38). 그러나 “스스로 결정한 목적의 실현을 위한 전제 조건들이 외부 현실에 마련되어 있지 않는 한, 우리는 스스로를 진정 자유롭다고 경험할 수 없다”(RF: 90). 사회적 자유는 자율적으로 선택된 목적의 실행을 가능하게 하는 “사회적 조건”을 자유 개념의 “구성 부분”, “내적 계기”로 포함한다(RF: 79). 이때 자유 실현의 사회적

“필수적인 조건”이 되는 것은 바로 나의 행위를 “보완”해주고 “완성”해주는 다른 사람들의 행위, 그리고 사람들의 행위를 상호 보완이 되도록 안정적으로 “여어주는” 사회의 “제도”이다(RF: 224). 이런 의미에서 사회적 자유는 “자유 개념을 제도 차원으로 확장한다”(RF: 80).

계속해서 사회적 자유 개념의 주요 특성을 살펴보기 전에, 나의 자유로운 행위의 실행이 다른 사람들의 상응하는 행위를 통해 제도의 매개에 의해 “보완”되고 비로소 “완성”, “실현”된다는 것이 무엇을 의미하는지를 일상의 예를 통해 분명히 해보자. 먼저 직접적 대면 관계에서의 예로 선물을 생각해 보자. 진심에서 우러러 나오는 마음에서 선물을 전달했다. 이것으로 선물이란 행위는 완성되었을까? 개인주의적 모델에 따르면 그렇겠지만 사회적 자유의 관점에서 보면 그렇지 않다. 선물은 받는 사람의 기쁘고 즐거워하는 반응 행위를 통해 드디어 실현되고 완성된다고 할 수 있다. 재즈 밴드도 사회적 자유 수행의 좋은 예가 될 수 있다. 나의 연주는 나의 연주만으로 결코 완성되지 않는다. 나만 잘 한다고 내 연주가 돋보이지도 않는다. 나의 연주를 받쳐주고 이어주는 다른 사람의 연주를 통해서 나의 연주는 완성된다.¹⁸⁾ 매개해 주는 ‘제도’가 필요한 예로, 환경을 생각하는 마음에서 재활용 쓰레기를 분리 배출하는 행위도 마찬가지이다. 깨끗이 씻어 분리 배출해도 그것을 재활용까지 이어주는 다른 사람들의 수고와 그것들을 “맞물리게”(ineinandergreifen) 여어주는 제도가 없다면, 환경을 지키고자 하는 개인의 작은 실천은 결국 실현되지 못하고 좌절된다. 예로 든 세 행위 모두 소극적, 성찰적 자유의 의미에서는 다른 사람의 응답·보완하는 행위 없이도 온전히 자유롭지만, 사회적 자유의 관점에서 보면 다른 사람들의 보완적 응답이 없이는 나의 행위 목적은 실현될 수 없고, 그런 점에서 부자유하다. 소박한 세 가지 예를 들었지만 현실에서 이런 행위유형은 매우 많다. 사회 현실을 구성하는 대부분의 사회적 행위는 이런 유형에 속한다.

1) 사회적 자유의 경험: “타자 속에서 자기 자신 곁에 존재함”

호네트는 헤겔의 『법철학』 「서론」에 나오는 한 표현이 사회적 자유 경험의 이

18) Brudney (1997)에 따르면 음악 협주는 철학자들이 공동체를 묘사할 때 즐겨쓰는 예이다. 396쪽 이하와 414쪽 주 43번 후반부를 참조.

해에 “열쇠”가 될 만하다고 생각한다. 그 표현은 “타자 속에서 자기 자신 곁에 존재함”(Bei-sich-selbst-Sein im Anderen, RF: 85)이다.¹⁹⁾ 헤겔의 이 표현을 이해하는 데에는 “자기 곁에 존재함”(Bei-sich-selbst-Sein)을 어떻게 해석하느냐가 적지 않게 중요하다. Neuhouser는 이것을 “집에 있는 것 같은 편안함”(being at home)으로 해석한다.²⁰⁾ 사회적으로 자유롭다는 것은 사회에서, 다른 사람들과의 관계 속에서, 친한 친구들(“우정”) 또는 친밀한 연인(“사랑”)과 함께 있을 때 처럼 “편안함”을 느끼는 것이다. 반대의 경우, 즉 사회적으로 부자유한 경험은 “꾸어다 놓은 보릿자루 같다”는 한국어 속담이 잘 포착하고 있다고 할 수 있다. 다른 사람들과 함께 있으나 환영받지 못하는 것처럼 느껴지고, 잘 어울리지 못해서 어딘가 어색하고, 있을 자리가 아닌 곳에 와 있는 것같이 불편하고 부자유한 것이다.

사회 안에서 타자와 함께 있으면서 마치 “집에 있는 것 같은 편안함”을 느낄 수 있는 것은 타자와의 관계가 “신뢰로 가득한 공동의 삶”일 때 가능하다(Drei: 159). 이러한 신뢰 관계에서 행위자들은 서로 상대방의 행위를 자신의 행위를 완성시켜주는 “협력”(Kooperation)으로 “보완”(Ergänzung)으로 파악하고, 이런 의미에서 자신의 행위를 타자에 의해 환영받는 것으로, 타자가 “바라는”(gewollt) 것으로 경험한다. 여기서는 “상호작용 상대방에게서 어떤 자의적 간섭도 예상되지 않을 뿐만 아니라, 오히려 한 사람이 기꺼이 하는 행동이, 다른 모든 참여자들이 마찬가지로 기꺼이 원했던(gewünscht) 행동이 되리라고 신뢰할 수 있다”(Drei: 157).

자신의 행동이 상대방이 “원하고 바랬던”(erwünscht oder erstrebt, RF: 222) 것으로 환영받으리라 믿을 수 있는 곳에서, 그런 “편안함” 속에서 사람들은 자유롭다고 느낀다. 이 자유 속에서 사람들은 더 이상 꾸어다 놓은 보릿자루 같이 움츠리지 않고, 자신 안에 갇히지 않고, 오히려 “풀려나는”(Ich-Entschränkung), 자신이 “확장”되는 경험을 한다(Drei: 144). “사람들은 그들을 둘러싼 환경을 자신들의 자아가 확장된 공간으로 지각할 수 있게 된다”(RF: 113).

19) 헤겔 (2016[1821]), p. 279. §7 보유(Zusatz)에 나오는 표현이다.

20) Neuhouser (2020), p. 105. 특별히 305쪽 주 34번의 일상언어적 해설 참조. Rose(2007)도 이 해석을 따르고 있다.

2) 사회적 자유의 표준 행위유형: “상호보완”

사회적 자유 경험이 전달하는 “편안함”과 “신뢰”와 “자아 확장”의 느낌은, 그 사회의 구성원들의 행위가 실제로 서로의 행위를 “가능하게” 하고 “보완”하고 “촉진”하는, 공동의 “협력”과 “협동”이라는 사실에 기반하고 있다. 이런 의미에서 사회적 자유의 “표준” 행위유형은 자유로운 “상호보완”이다.²¹⁾

그런데 호네프트는 사회구성원들이 서로를 “상호보완”하는 “간접적” 방식과 “직접적” 방식을 구분한다. 기준이 되는 것은 행위자들의 “동기”이다(IS: 39-42). 간접적 방식은 자본주의 “시장”에서처럼 각자는 자신의 이윤만을 추구하지만 “보이지 않는 손”이 각자의 교환 행위를 서로의 욕구 충족으로 이어주는 방식이다. 상호보완이 이렇게 간접적으로 이루어지는 곳에서 참여자들의 관계는 “순전히 도구적”일 수 있다.²²⁾ 상호보완의 “직접적이고 의지적인”(direkt und willentlich, IS: 41) 방식은 참여자들이 “상호의존성에 대한 의식”에 기반하여 서로에 대한 “관심”(Anteilnahme)과 “배려”(Rücksichtsnahme)에서, “상대방을 위하려는 의도에서” 상호협력이 이루어지는 것이다. 여기서 상호작용은 “단지 함께 하는 활동이 아니라 ‘서로를 위한’ 활동”이 된다.²³⁾ 1844년 경의 맑스가 마음에 품고 있던 공산주의 사회에서의 삶의 방식이다.

사회적 자유의 행위 도식은 의식적이고 의도적인 상호협력이다. 행위자는 자신의 행위를 계획할 때부터 이미 “‘우리’의 관점”에서(‘Wir’-Perspektive, Drei: 145, 152) 상대방의 목적과 필요를 고려한다(IS: 40). 각자는 의식적으로 서로를 “위하고” “배려”하는 협력 속에서 각자의 목적을 이루면서 동시에 상대방의 목적 실현에 기여한다. 이런 의미에서 사회적 자유는 개인이 고독하게 실행할 수 있는 것이 아니라 “‘우리’의 구성원으로서(als Mitglieder eines ‘Wir’, Drei: 156) 함께 하는 “공동의 실천”, “공동의 기획”이다(Drei: 146, 151, 142). 각자는 서로에게 “자유의 조건”이 된다(IS: 132).

21) “Muster eines Sich-wechselseitig-Ergänzens” (IS: 129).

22) Brudney (1997), p. 390.

23) “nicht bloß miteinander, sondern ‘füreinander’ tätig” (IS: 41, 46)

3) 사회적 자유의 “존재 조건”으로서 제도

소극적 자유가 인간을 고립된 원자로, 성찰적 자유가 독백적 주체로 간주한다면, 사회적 자유는 인간을 “구성적으로” 서로 의존하는 존재로 본다(RF: 145). 인간은 인간으로의 성장을 위해서도, 욕구 충족을 위해서도 항상 다른 사람들의 “보완적 행위”를 필요로 하는 존재이다(IS: 38). 이 보완은 일정 시간 이상 지속되는 관계에서는 제도를 통해 매개될 수밖에 없다. 보통 사회적 자유를 가능하게 하는 제도는 실정법을 통해 형식적으로 성립된 제도보다는, 관행, 관습, 역할기대, 행위자의 적절한 태도와 행위방식 등의 복합체로서 “약한” 의미에서의 제도이다(RF: 125-126, 131). 사회적 자유가 “사회적”이라는 수식어를 갖는 것은 바로 자유 개념이 이렇게 제도를 자유의 “매개체와 실행조건”으로(RF: 86), 자유의 “필수적인” “존재 조건”으로 개념 안에 포함하기 때문이다(RF: 123, 125).

제도는 사회적 자유의 실현을 위해 두 가지 이유에서 필수적이다. 하나는 각자의 행위가 실현되고 완성되기 위해서 필요한 서로 간의 여러 행위를, 한 행위의 실현이 다른 행위의 실현조건이 되는 방식으로 “맞물리게” 안정적으로 “보장”해주는 것이 바로 제도이기 때문이다(RF: 86). 마찬가지로 중요한 다른 이유는, 제도와 관행과 사회적 역할기대 등이 행위자를 “사회화”하는 효과 때문이다. 사회적 자유가 실현되기 위해서는 단지 사람들의 행위가 상호 보완된다는 것으로 부족하다. 앞에서 말했듯이 행위자들이 서로를 소극적·법적 권리를 갖춘 이기적 모나드로 간주하는 자본주의 시장에서도 일정 형태의 보완이 일어난다. ‘사회적’ 자유가 실현되기 위해서는 행위자들이 서로를 상호의존적이면서 상호보완적인 존재로 이해하는 인식과 신념, “‘우리’라는 관점”에서 서로를 “배려”하는 태도와 행위 습관, 그리고 그에 적합한 “욕구 체계”를 갖추고 있어야 한다. 이러한 사회적 자유에 필요한 자질을 사람들은 개별 행위를 맞물리게 엮어주는 제도에 참여함으로써, 그 제도가 요구하는 역할을 수행함으로써 함양하게 된다(RF: 92-93).

4) 세 가지 자유 개념의 분석적 비교

앞에서의 세 가지 자유 개념에 대한 설명은 세 가지 측면에서 다음과 같이 분석적으로 비교, 정리될 수 있다.

(1) 행위자

- ① 소극적 자유의 행위자는 “원자”와 같은 고립된 존재로 표상된다. 이 행위자는 오직 자신의 선호에 입각해 행위한다.
- ② 성찰적 자유의 주체도 고독한 개인이지만, 자신의 행위를 일반화될 수 있는 근거에 입각해 결정할 수 있는 이성적 “능력”을 갖춘 존재로 여겨진다. 단순히 자신의 선호만이 아니라 일반화될 수 있는 이유에 관심을 갖는다는 점에서 원칙적으로 타자를 향하고 있다. 그러나 그 타자는 여전히 추상적이고 독백적으로 표상된 타자일 뿐이다.
- ③ 사회적 자유의 행위자는 개인이지만 “가치관을 공유하는 공동체의 일원”으로 상정된다.²⁴⁾ 그는 “어떤 ‘우리’”의 구성으로서(Drei: 156), 자신의 행위 계획을 세울 때부터 ‘우리’의 관점에서 다른 사람들의 행위 목적을 고려할 뿐만 아니라 “배려”한다.

(2) 행위 수행 방식

- ① 소극적 자유의 행위자는 고독한 개인으로서 자신의 목적을 추구한다. 그는 타자의 행위에 무관심하거나, 관심을 갖는다면 자신의 목적 실현과 관련하여 전략적 관심을 갖는다.
- ② 성찰적 자유의 주체도 마찬가지로 단독자이다. 그러나 전략적으로만 행동하지 않으며 타자에게도 정당화될 수 있는 방식으로 행위하려고 한다.
- ③ 사회적 자유 모델에 따르면 행위의 수행과 실현은 음악 밴드의 협주처럼 공동의 “협업”이다. 이 협업은 한 구성원의 목적 실현이 다른 구성원의 목적 실현에 의식적으로 기여하는 “상호보완”의 방식으로, 특히 그런 상호보완을 매개하는 “제도”에 참여함으로써 이루어진다.

(3) 자유의 실현 상태

- ① 소극적 자유가 실현된 상태는 타자의 “자의적인 간섭”이 없는 상태이다. 주체는 타자의 자유를 침해하지 않는 한에서 타자를 신경 쓸 필요 없이 자

24) “Zugehörigkeit zu einer Gemeinschaft ethisch übereinstimmender Mitglieder” (Drei: 158)

기 “마음대로” 행동할 수 있다.

- ② 성찰적 자유 모델에 따라 자유로운 주체는, 자신의 의지의 주인이라고 할 수 있다. 그는 자신의 행위 이유들에 대해서 투명하게 의식하며, 자신이 승인한 이유들에 따라서만 행동한다.
- ③ 사회적 자유 모델에 따라 자유롭다는 것은, 사회 안에서 개인이 다른 사람들로부터 협력자로 인정과 환영을 받으며, 그들의 관심과 선의를 기대하고 신뢰할 수 있는 상태이다. 이런 상태에서 개인은 “집에 있는 것 같은 편안함”을 느끼며, 자신의 자아가 타자에게로까지 “확장”되는 경험을 한다. 개인들은 서로를 “자기 자신의 타자로”(als Andere ihrer Selbst), 즉 자신의 다른 모습으로, 다른 모습을 한 자신으로 인정하고 경험한다(RF: 85, 86).

호네프는 오해를 방지하기 위하여 사회적 자유가 실현된 “공동체”와 별린의 적극적(positiv) 자유의 집단주의적 버전과의 차이점에 주의를 환기한다. 핵심적 차이는 별린의 집단에서는 개인성과 개인의 권리가 사라지지만, 자유의 공동체에서는 소속감, 가치와 관점의 공유, 서로에 대한 배려와 연대에도 불구하고 구성원 각자는 개인으로서 항상 소극적 자유와 성찰적 자유를 누리며, 공동체의 사안에 대해 동등한 구성원으로서 함께 결정할 권한과 기회를 갖는다는 것이다(Drei: 147-148).²⁵⁾

5) 사회적 자유와 소극적, 성찰적 자유의 관계

사회적 자유 개념을 부각시키고 활성화시키고 싶어하는 호네프에게도 사회적 자유만이 아니라 소극적 자유와 성찰적 자유도 근대 이후 우리의 자유 이해의 “본질적 계기”를 이룬다(RF: 114). 그렇다면 호네프는 이 세 가지 자유 모델의 관계를 어떻게 파악할까?

앞에서 확인했던 소극적 자유와 성찰적 자유의 “기능”과 “가치”를 기억해 보자. 이 두 자유의 본질적 기능은 기존의 관계와 역할기대에 거리를 두기 위해 철수하거나, 그것의 정당성에 의문을 제기하는 것이다. 그런데 그렇다면 이 두 개인주의적 자유가 실행되기 위해서는 거리두기의 대상이 되는 사회적 관계와 상호작용이

25) 이 점에서 호네프는 Petti (1992)에 의존해서 “전체론적 개인주의”(holistischer Individualismus)라는 개념을 도입한다. 그러나 충분히 설명하지는 않는다(IS: 47-48).

이미 있어야 한다. 이러한 상호작용에서 사람들이 살고 있는 자유가 바로 사회적 자유이다. 이로부터 떨어져 나와 ‘혼자만의 공간’으로 들어가는 것은 이미 맺고 있던 관계와 상호작용에서 생긴 문제에 대한 “반응”이다. 호네프는 이러한 사실에서 소극적, 성찰적 자유는 “파생적”(derivativ)이고 “부차적”(sekundär)이라고 본다(RF: 114-115, 221-222).

개인주의적 자유가 “파생적이고 부차적”이라는 것은 발생론적 측면에서 그렇다는 것이지 정당성 측면에서 그렇다는 것은 아니다. 사회적 자유가 실현되기 위해서는 제도가 전제되어야 하는데, 제도는 자유를 증진할 수도 있고 억압할 수도 있다. 참여자들을 동등하게 대우하지 않는다면, 그들의 자율성을 침해한다면, 역할 분담이 불공정하다면, 그런 제도는 자유를 보장하고 촉진하기보다 방해할 것이다. “그 자체로 자유를 구현하고 있는 제도만이 자유를 보장”할 수 있다(RF: 101). 그리고 “자유를 구현하고 있는 제도”는 바로 참여자들에게 그 제도를 “언제든지 자신의 의도와 신념에 비추어 검토하고 경우에 따라서 떠날 수도 있는 기회를 보장하는” 제도이다(RF: 115). 따라서 소극적 자유와 성찰적 자유가, 사회적 자유의 조건이 되는 제도의 정당성, 제도들의 복합체인 사회의 정당성을 평가하는 “정당한 기준”이 된다(RF: 115). 이렇게 소극적 자유와 성찰적 자유가 평가 기준으로 활용되면서, 이 두 개인주의적 자유의 “잠재력”, 즉 “개인화와 자율성의 힘”이 제도에 “활력”(Dynamik)을 불어넣는다. 이제 어떤 제도도 고정된 모습으로 남아있을 수 없게 된다(RF: 116, 117). 사회적 자유를 구현하고 있는 사회 제도는 그것이 보장하는 소극적 자유와 성찰적 자유, “개인화와 자율성의 힘”에 의해 “계속해서 항구적으로 스스로를 넘어서다”(IS: 102, Fn. 107).

물론 호네프는 이렇게 소극적, 성찰적 자유의 온당한 가치와 기능을 인정하면서도 그것이 절대화되거나 “자립화”되는 것에 대해서는 경계한다(RF: 125). 사회적 관계와 상호작용이 성립하고 나서야 그것을 “일시 정지”시키는(Moratorium, RF: 155, 205) 소극적 자유와 성찰적 자유의 사용이 있는 것이다. 스스로를 고립시키는 개인주의적 자유만으로는 사회적 삶이 형성될 수 없다. 개인주의적 자유를 활용하여 사회질서의 정당성에 의문을 표할 수 있지만, 그것만 가지고서는 다른 주체들과 함께 새로운 질서를 세울 수 없다(RF: 110). 따라서 사회적 삶의 이런 상호주관적이고 상호의존적인 성격을 잊고 개인주의적 자유 이해가 절대화되면 그에 따른 “병리 현상들”이 발생하게 된다.²⁶⁾

4. 사회적 자유 개념의 곤란함

지금까지의 설명을 통해 호네트의 사회적 자유 개념의 내용과 구조가 선명히 드러났기를, 그 결과 이 개념이 언뜻 생각하기보다 우리의 일상적 직관과 가까우며, 다른 두 개인주의적 자유 개념만큼 진지하게 논의할 만한 “독자적” 개념일 수 있음이 좀 더 그럴듯해졌기를 희망한다. 마지막으로 호네트의 이 자유 개념이 맞닥뜨리게 되는 곤란함 두 가지를 살펴보고자 한다. 첫 번째 곤란함은 개념적 성격의 것으로, 자주 제기되어 온 의문, 이 자유 개념이 실제로 ‘자유’ 개념인가 하는 것이다.²⁷⁾ 두 번째 곤란함은 호네트가 정치적 실천의 “선도 이념”으로 제안하는 사회적 자유 개념이 “합당한 다원주의의 사실”을 충분히 진지하게 고려하고 있는가 하는 의문이다.

1) 자유와 연대

호네트가 생각하는 자유가 실현된 상태는 서로의 행위와 목적을 “상호보완”해주는 사회 제도를 기반으로, 참여자들이 서로를 “인정”하고 “배려”하고 “신뢰”하는 가운데, 타자와 더불어 “집에 있는 것 같은 편안함”을 느끼며, 서로의 자아가 서로에게 “확장”되어지는, 그래서 타자를 “자신의 타자로” 여기게 되는 상태라는 것을 보았다. 그런데 이러한 묘사는 많은 사람들에게 자유보다는 ‘우애’(Brüderlichkeit)나 ‘연대’(Solidarität)를 떠올리게 한다. 호네트 자신도 자신의 사회적 자유 개념이 연대와 갖는 밀접한 관련성을 잘 알고 있고, 스스로 이를 부각시킨다. 초기 사회주의자들은 자유가 “사적 이기주의”로 이해되면 우애와 양립 불가능하며, 평등과도 불화하게 된다는 것을 인식하고, 자유 개념을 확장하여 자유를 평등 및 연대와 화해시키려고 하였다(IS: 30-33). 이러한 화해 시도의 개념적 “해결책”이 사회적 자유 개념이다.²⁸⁾ 사회적 자유에서 자유, 평등, 우애는 “완전히 일치”하게 된다는 것이다(vollends zusammenfällt, IS: 48). 호네트는 초기 사회주의자들의 이러한 구상을 따르면서, 사회적 자유 개념이 연대에 대한

26) 이러한 병리 현상에 대해서는 RF: 157-172, 206-218 참조.

27) Nys (2013); Zürn (2015), pp. 192ff.; 강병호 (2021), pp. 162ff.

28) “das Lösungswort für die Versöhnung” (IS: 51).

“단지 개선된 표상”이 아니라 “자유에 대한 자립적이고 새로운 이해”라고 강조한다(IS: 43).

사회적 자유와 연대의 관계에 대한 호네트의 주장의 요점은 전자가 후자의 “상위 개념”이라는 것이다(Oberbegriff, Drei: 160). 두 가지 근거를 제시한다. 하나는 자유는 여러 모습으로 실현될 수 있는데, 연대는 자유가 실현되는 모습 중 하나이다. 자유가 “타입”(type)이라면, 연대는 “토큰”(token)이다(Drei: 161). 다른 근거는 자유가 연대가 추구되는 궁극적인 이유, 목적이라는 것이다. 사람들은 왜 연대를 추구하는가? “자유 특별한 형태”(Drei: 160), 즉 “자신의 행위에서 완전한 비강제성을 경험하기 위하여”(Drei: 161).

여기서 이 문제를 본격적으로 다룰 수는 없다. 다만 호네트가 제시한 두 번째 근거에 대해서 간단한 반론을 제시겠다. 에리히 프롬이 “자유로부터의 도피”라고 포착한 현대인의 자유 경험과 관련된 흥미로운 현상은, 자유가 사람들에게 “견딜 수 없이” 부담스러운 것, 피하고 싶은 것이 될 수도 있다는 것이다.²⁹⁾ 현대 사회에서 자유의 증가는 개인화와 함께 일어나고, 개인화는 유대 관계의 해체를 초래하며, 유대 관계의 상실은 개인을 고독하게 만들고, 개인의 삶의 의미 찾기를 어렵게 한다는 것이다. 그래서 현대인은 고독과 불안과 무의미에서 벗어나기 위해서 기꺼이 “도피”와 “복종”을 선택하기도 한다. 물론 이런 문제를 일으키는 것은 자유에 대한 “소극적” 이해이고, 사회적 자유가 실현된다면 이런 문제 현상 자체가 일어나지 않을 것이라고 할 수도 있다. 프롬이 대안으로 제시하는 “적극적 자유”도 호네트의 사회적 자유와 유사한 점이 많다.³⁰⁾

그런데 내가 프롬의 “자유로부터의 도피” 논제에서 도출하고자 하는 함축은, 사람들이 연대와 유대를 추구하는 이유와 관련된다. 그것은 호네트의 주장과 달리 “자유 경험을 위해서”가 아닐 수 있다는 것이다. 외로움과 그것이 초래하는 불안과 의미상실은 보통 사람들에게 감당하기 힘든 심리적 부담이고, 그만큼 유대와 연대는 고독에서 벗어나기 위해서 ‘그 자체로’ 추구된다고 할 수 있다. 그 대가가 부자유와 복종이라고 하더라도. 그렇다면 연대는 자유를 위한 것이고, 자유가 실현되는 하나의 형식일 뿐이라는 주장은 호네트가 생각하는 것처럼 그렇

29) 프롬(2014[1941]), 2장과 4장 참조.

30) 프롬(2014[1941]), p. 149.

게 자명한 것이 아니다. 오히려 연대와 유대는 그 자체로 추구되나, 그러한 유대는 자유로운 결합의 형태가 될 수도 있고, 복종의 형태가 될 수도 있다는 것이 더 경험적 사실에 부합할 것이다.

2) 포괄적 신념체계로서 사회적 자유 개념

사회적 자유는 벌써 그 개념상 “우리라는 관점”에서 수행된다(Drei: 145, 152). 그러니까 사회적 자유는 자신의 “존재 조건”으로서 “공동체”를 전제한다(IS: 141; Drei: 148). 호네트에 따르면 그렇다면 이 공동체가 “그 구성원들이 서로 얼굴을 아는 작은 규모의 공동체”이어야 하는 것은 아니다. 일상언어의 용법이 보여주듯이 “민족”이나 “정치적 운동” 같이 “상당히 규모가 크고 익명인” 집단도 충분히 공동체가 될 수 있다(IS: 53). 호네트는 현재 전망이 아주 밝지는 않지만 유럽 연합도 일종의 공동체로 발전할 가능성이 여전히 있다고 생각한다(RF: 612-624). 결정적인 것은 집단의 크기나 구성원들 간의 친밀성 정도가 아니라 구성원들이 “공동의 목표 설정”과 관련하여 “마음이 같은”(gleichgesinnt) 것이다(IS: 53). “마음이 같다”는 것은 다음과 같이 상술될 수 있다: 사회적 자유의 실현을 위해 해결해야 할 “문제들의 인식”에서 상당히 일치하여 “공동의 행위”로 나아갈 수 있을 정도로 “규범적 감수성과 관심이 일치”하는 것이다. 이런 의미에서 “마음이 같기” 위해서는 구성원들이 상당한 정도로 가치와 목적, 문화적 자원을 공유해야 할 것이다(IS: 159-160).

호네트는 이 정도로 “같은 마음”이 요구된다는 것이 꼭 “가족 같은” 공동체를 전제하는 것은 아님을 보이기 위해 롤즈의 『정의론』의 “질서 정연한 사회”도 “차등의 원칙”의 적용을 위해서 상당한 정도의 연대나 우애가 필요하다고 지적한다(IS: 53-54). 롤즈가 『정의론』(1971)에서 질서 정연한 사회를 일종의 공동체로 구상했음은 분명하다.³¹⁾ 그런데 여기서 호네트가 스스로 선택한 롤즈의 질서 정연한 사회와의 비교는, 호네트의 의도와 다른 방향으로도 시사하는 바가 있다. 롤즈는 후에 『정의론』에서 질서 정연한 사회를 사고한 방식이 “유토피아적”이었 고, “다원주의의 조건을 고려하는 데 실패”했다고 자인하기 때문이다.³²⁾ 롤즈는

31) 롤즈 (2014[1999/1971]), 79절 참조. Brudney는 『정의론』의 질서 정연한 사회에 대한 구상이 『정치적 자유주의』에서보다 훨씬 더 공동체주의적이라고 본다. Brudney (1997), pp. 399ff.

『정의론』에서 공정이라는 정의관이 자유주의라는 “포괄적 신념 체계”에서 도출될 수 있다고 가정했는데(JF: §55.4), 그것은 그가 “합당한 다원주의의 사실”이라고 부르는 현대사회의 조건에 맞지 않는다는 것이다. “합당한 다원주의의 사실”이란 “자유로운 제도의 기본적인 권리들과 자유들에 의해 보장되는 정치적·사회적 조건하에서는 상충하고 화해 불가능하지만 합당한 포괄적 신념 체계들이 다양하게 -만약 아직 존재하지 않는다면- 생겨나고 지속될 것”이라는 사정이다. 롤즈에 따르면 이러한 “근대 민주 사회에서 발견되는 종교적·철학적·도덕적 신념 체계들의 다양성은 곧 사라질 수도 있는 단지 역사적인 조건이 아니다. 그것은 민주주의의 공적 문화가 갖는 영구적인 특징이다”(JF: §11.3). 이런 조건에서 “하나의 포괄적 신념 체계를 지속적으로 공유하고 고수하는 것은 … 국가권력의 억압적 사용에 의해서만 유지될 수 있다”(“억압의 사실”, JF: §11.3). 그래서 롤즈는 『정치적 자유주의』(1993)에서 공정으로서의 정의에 입각한 질서 정연한 사회를 “정치적 관점”으로, “정치적 자유주의”의 가장 합당한 형태로 제시한다. 정치적 관점은 포괄적 신념 체계들 사이의 “중첩적 합의”의 대상이 된다. “시민들은 상충하는 종교적·철학적·도덕적 견해들을 가지며, 따라서 서로 다르고 대립하는 포괄적 신념 체계들 내에서, 그리고 적어도 부분적으로는 서로 다른 이유들로 정치적 관점을 승인한다”(JF: §11.1). 이제 민주적 정치사회는 공동체와 분명히 구별된다. 전자는 “공동체가 아니며, 합당한 다원주의의 사실을 고려할 때 그렇게 될 수도 없다”(JF: §7.3).

나는 호네트도 “합당한 다원주의의 사실”을 인정해야 한다고 생각한다. 앞에서 보았듯이 사회적 자유를 실현하는 제도들은 동시에 그 자체로 소극적, 성찰적 자유를 구현하고 있어야 한다. 그리고 그러한 조건에서는 “상충하고 화해 불가능하지만 합당한 포괄적 신념 체계들이 다양하게 -만약 아직 존재하지 않는다면- 생겨나고 지속될 것”이다. 그렇다면 이제 관건은 호네트의 사회적 자유가 “정치적 관점”인지 아니면 “포괄적 신념 체계”인지 하는 것이다. 나는 다음의 두 가지 이유에서 호네트의 사회적 자유가 포괄적 신념 체계에 가깝다고 본다.³³⁾

32) 롤즈(2016[2001]), §55.5. 이하에서는 이 책을 ‘JF’로 약칭하고, 해당 절을 표시하겠다.

33) 롤즈는 칸트의 의무론, 밀의 공리주의를 모두 철학적·도덕적 포괄적 신념 체계로 본다(JF: §11.2). “공리주의의 한 형태나 칸트 혹은 밀의 도덕적 견해로 통일되어 있는 사회도 또한 통일된 사회로 남아 있기 위해 국가권력의 억압적 제재를 필요로 할 것이다”(JF: §11.3).

첫째, 롤즈에 따르면 정치적 관점은 “헌법”과 같이 “사회의 기본 구조”에서 드러나는 가치와 목적만을 포함한다(§11.2). 이와 달리 호네프의 사회적 자유의 원초적 장소는 우정과 사랑 같은 친밀한 관계이고, 그것을 경제 영역과 정치적 공론장을 비롯해 전 사회로 확장해서 실현하고자 한다. 사회적 자유는 “사회의 기본 구조”에서는 물론이고 사회구성원들의 일상의 삶에서, 그들의 일상적 “관례”와 “관습”과 “행위방식”과 “습관”에서 구현되어야 한다(RF: 125-126). 둘째로 호네프는 사회적 자유가 실현되기 위해서는 행위자들의 목적이 그저 “중첩”되는 것으로는 부족하다고 본다. 그들은 “의도적으로 서로를 위해서”(absichtsvoll füreinander) 활동해야 한다. “상호배려”가 행위자들의 “태도”(Einstellung)와 “동기”(Antrieb)로 요구된다(IS: 39-42, 46, 53). “사회적 자유란 공동체의 사회적 실천에 참여하는 것이다. 이 공동체에서 구성원들은 그들의 정당한 욕구 실현을 위하여 상호적으로 협력하는데, 이때 이것을 ‘서로 상대방을 위하는 마음에서’ 할 만큼 서로를 배려한다.”³⁴⁾

호네프의 자유 이론이 개인주의적으로 쫓고라든 현대사회의 자유 이해의 “빈곤”을 지적하면서, 사회적 자유라는 “파문혀” 잊혀진 전통을 상기시키는, 상대적으로 소박한 목적에 머문다면(Drei: 161), 그것이 포괄적 신념 체계라는 것은 큰 문제가 되지 않을 것이다. 그러나 호네프가 천명하듯이 사회적 자유가 사회 변혁을 위한 정치적 기획의 “선도 이념”이 되어야 한다면, 호네프는 “합당한 다원주의의 사실” 그리고 그로부터 곧장 도출되는 “억압의 사실”을 지금보다 훨씬 더 진지하게 고려할 필요가 있다.³⁵⁾

34) “Dann heißt soziale Freiheit, an der sozialen Praxis einer Gemeinschaft teilzunehmen, in der die Mitglieder sich untereinander so viel Anteilnahme entgegenbringen, daß sie sich *um des jeweils anderen willen* wechselseitig zur Verwirklichung ihrer begründeten Bedürfnisse verhelfen.”(IS: 47, 인용자의 강조)

35) 같은 맥락에서 Ferrara (2019)의 문제 제기도 참조.

참고문헌

- 강병호. 2021. 「악셀 호네트의 “규범적 재구성”」. 『범한철학』 제102집: 145-171.
- 롤즈, 존. ¹²2014[1999/1971]. 『정의론』, 황경식 옮김. 이학사.
- _____. 2016[2001]. 『공정으로서의 정의. 재서술』, 김주희 옮김. 이학사.
- 서정혁. 2018. 「호네트의 헤겔 해석: 『비규정성의 고통』을 중심으로」. 『현상과 인식』 42권 4호: 105-126.
- 왈쩨, 마이클. 2007. 『해석과 사회비판』, 김은희 옮김. 철학과현실사.
- 프롬, 에리히. ³2014[1941]. 『자유로부터의 도피』, 김석희 옮김. 휴머니스트.
- 호네트, 악셀. 2017[2001]. 『비규정성의 고통』, 이행남 옮김. 그린비.
- 홉스, 토마스. 2008[1651]. 『리바이어던 1』, 진석용 옮김. 나남.
- 헤겔, G. W. F. 2016[1821]. 『법철학 강요. 「서문」과 「서론」』, 백훈승 역주. 서광사.
- Berlin, Isaiah. 2002[1958]. “Two Concepts of Liberty.” in: *Liberty. Incorporating ‘Four Essays on Liberty’*. Oxford University Press. (「자유」의 두 개념, 『자유론』, 박동천 옮김. 아카넷)
- Brudey, Daniel. 1997. “Community and Completion.” in: Reath, Herman, and Korsgaard (ed.), *Reclaiming the History of Ethics. Essays for John Rawls*. Cambridge University Press.
- Ferrara, Alessandro. 2019. “Social Freedom and Reasonable Pluralism: Reflections on Freedom’s Right.” *Philosophy & Social Criticism*. 45(6): 635-642.
- Hegel, Georg Wilhelm Friedrich. 1986[1821]. *Grundlinien der Philosophie des Rechts*, in: Werke, Bd. 7, Suhrkamp: Frankfurt am Main.
- Honneth, Axel. 2007. *Pathologien der Vernunft. Geschichte und Gegenwart der Kritischen Theorie*. Frankfurt a. M.: Suhrkamp.
- _____. 2007[2000]. “Rekonstruktive Gesellschaftskritik. Zur Idee der Kritik in der Frankfurter Schule”, in: Honneth (2007), pp. 57-69.
- _____. 2010. “Das Reich der verwirklichten Freiheit. Hegels Idee einer ‘Rechtsphilosophie’.” in: *Das Ich im Wir*. Berlin: Suhrkamp.
- _____. 2011. *Das Recht der Freiheit. Grundriss einer demokratischen Sittlichkeit* [RF], Berlin: Suhrkamp.

- _____. 2015. *Die Idee des Sozialismus* [IS]. Berlin: Suhrkamp. (『사회주의의 재발명』, 문성훈 옮김. 사월의 책)
- _____. 2020. *Die Armut unserer Freiheit. Aufsätze 2012-2019*. Berlin: Suhrkamp.
- _____. 2020a. “Drei, nicht zwei Begriffe der Freiheit. Zur Reaktualisierung einer verschütteten Tradition” [Drei], in: Honneth (2020), pp. 139-161.
- _____. 2020b. “Von der Armut unserer Freiheit. Größe und Grenzen der Hegelschen Sittlichkeitslehre.” in: Honneth (2020), pp. 38-58.
- _____. 2020c. “Die Normativität der Sittlichkeit. Hegels Lehre als Alternative zur Ethik Kants.” in: Honneth (2020), pp. 59-74.
- Neuhouse, Frederick. 2000. *Foundations of Hegel's Social Theory: Actualizing Freedom*. Cambridge (MA): Harvard University Press.
- Nys, Thomas. 2013. “Which Justice, Whose Pathology?” in: *Krisis. Journal of Contemporary Philosophy*. Issue 1: 10-13.
- Pettit, Philipp. 1993. *The Common Mind: An Essay on Psychology, Society, and Politics*. Oxford University Press.
- _____. 2014. *Just Freedom. A Moral Compass for a Complex World*. W.W. Norton & Company: New York. (『왜 다시 자유인가』, 박준혁·윤채영 옮김. 한길사)
- Pippin, Robert. 2014. “Reconstructivism: On Honneth's Hegelianism.” *Philosophy & Social Criticism* 40(8): 725-741.
- Rose, David. 2007. *Hegel's Philosophy of Right. A Reader's Guide*. Continuum International Publishing Group. (『헤겔의 <법철학> 입문』, 이신철 옮김. 서광사)
- Taylor, Charles. 1985. “What's Wrong with Negative Liberty.” in: *Philosophy and The Human Science. Philosophical Papers 2*. Cambridge University Press.
- Zurn, C. F. 2015. *Axel Honneth: A Critical Theory of the Social*, Polity Press.

Abstract

On Honneth's Concept of Social Freedom

Byoung-ho Kang

Since *Freedom's Right*, Honneth has proposed 'social freedom' as the "normative guideline" for social criticism. However, his concept of social freedom may seem to be "downright peculiar, and even eccentric". This essay aims at explaining his concept of social freedom systematically and examining its validity in modern pluralistic society.

Subject Areas: Social Political Philosophy, Critical Theory, Modern European Philosophy

Keywords: Honneth, Negative Freedom, Reflexive Freedom, Social Freedom, Socialism, Critical Theory

제가 논평을 맡게 된 강병호 선생님(이후 발표자)의 논문을 간단히 요약하면서 논평문을 시작해보겠습니다. 발표자는 악셀 호네프가 말하는 세 가지 자유 개념을 해설합니다. 즉 홉스가 정의하듯 “저항의 부재”로 정의되는 자유, 즉 소극적, 법적 자유(헤겔에서 추상법에 해당), 그리고 주체의 자기관계에 주목하는 성찰적, 도덕적 자유(헤겔에서 도덕에 해당)에 대한 개념적 해명 이후 사회적 자유(헤겔에서 인륜에 해당) 개념의 독특성에 대해 명료한 설명을 제공하고 있습니다. 특히 소극적 자유는 발표자에 따르면 단순히 미개하거나 부정적인 것은 아닌데, 이 자유는 “사회와 여론의 정당화 압박에서 벗어나 그것이 무엇이든 자신이 해보고 싶은 것을 해보고 자기가 살고 싶은 삶을 실험해볼 수 있는 무중력의 공간”(229)이라는 점에서 존 스튜어트 밀 같은 사람의 자유론을 떠올리게 하는 것이기 때문입니다. 개인이 외적 제한 없이 마음대로 행동할 수 있는 권리를 누려야 한다는 것은 “근대의 도덕적 자기이해”이고 이러한 근대적 성과는 되돌릴 수 없이 중요한 것이라는 점이 인상 깊습니다. 성찰적 자유 역시 칸트의 도덕철학이 잘 보였듯이 자기입법적인 자유로서 자율성으로 이해될 수 있습니다. 이는 근대의 도덕적 자기이해를 이루는데, 이는 기존의 사회질서를 변화시키는 “급진적 거리두기와 일반적 동의의 결합”(231)으로서 의의가 깊은 것입니다.

이처럼 소극적 자유는 원자적인 개인에 대한 외적 방해와 간섭에만 관련되고 성찰적 자유는 독백적 행위자의 능력에만 몰두하여 그 현실적 실행가능성을 보지 못한다는 점에서, 발표자는 궁극적으로 자유의 사회적 조건으로서 사회적 자유 개념에 주목합니다. 발표자가 재구성하는 호네프의 사회적 자유는 그야말로 “자유 개념을 제도 차원으로 확장”(232)하는 것입니다. 사회적 부자유가 일종의 소외의 경험, 또는 “꾸어다 놓은 보릿자루 같다”라는 속담에서 포착되는 상태라면, 사회적 자유에서 우리는 탈소외의 경험, “집에 있는 것과 같은 편안함”(233)을 느끼게 되어 각자는 상대방의 행위를 신뢰하고 상호보완으로 파악하여 자아 확장의 느낌을 가질 수 있게 됩니다. 여기서 각자는 상호의존성을 의식하여 서로에 대해 관심을 가지고 배려하며, ‘우리’의 관점에서 협업합니다. 발표자가 정확히 지적하듯이 각자가 서로의 자유의 조건이 되는 이러한 이념은 1844년의 마르크스가 생각하고 있던 “공산주의 사회에서의 삶의 방식”인 “의식적이고 의도적인 상호협력”(234)입니다. 발표자는 특히 235쪽 이하의 세 가지 자유 개념의 분석적 비교를 통해 호네프의 이론적 입지점을 명료하게 해명하고 있습니다. 마지막으로 발표자는 사회적 자유 개념이 갖는 두 가지 곤란함에 대해 지적함으로써 이 논문의 의도가 호네프의 이론을 무비판적으로 지지하는 것은 아님을 명확히 합니다.

발표자가 비판적 재구성한 호네프의 사회적 자유 개념은 그 자체로도 매우 흥미로우며 설득력 있으나, 논평자로서의 의무를 다하기 위해 서투르나마 몇 가지 질문을 제기해보겠습니다. 첫째 질문은 사회적 자유 개념이 논문에서 소개된 것처럼 자율적인 것인지에 대한 것입니다. “사회적 자유 모델”이 “소극적 자유와 성찰적 자유를 대체하려고 하지 않”고, “그것들 각각의 합리적 핵심을 수용하고 ‘확장’하여 “사회적 자유 개념은 다른 두 자유 개념의 ‘종합’”(228)이라고 할 때, 자연스러운 질문은 왜 우리가 소극적 자유와 성찰적 자유로 만족할 수는 없고 사회적 자유라는 추가적인 개념을 필요로 하는가라는 것이기 때문입니다. 이는 한편으로 호네프가 말하는 비판이론의 내재적 비판에서 사회적 자유는 사회비판의 선도 이념이지만, 다른 한편으로 소극적 자유 및 성찰적 자유가 “사회의 정당성을 평가하는 정당한 기준”(238)이라는

점을 고려하면 더욱 그렇습니다. 호네프에게 관건은 자유 개념이 실은 (벌린이 생각했던 것처럼) 두 개가 아니라 세 개라는 점이지만, 사회적 자유가 제3의 자유 개념으로 확고하게 자리 잡을 수 있는지를 따져볼 수 있을 것입니다. 아마 호네프의 논의에 비판적인 독자라면, 실은 사회적 자유 개념은 제3의 자유 개념이 아니라 자유가 아닌 다른 무엇으로 불려야 하는 것이 아니냐는 의문을 제기할 것입니다. 호네프에 따르면 사회적 자유 개념은 “사랑과 가족, 시장, 민주주의 같은 현대 사회의 제도의 형성에 ‘가장 큰 영향’을 끼쳐온 자유 개념”(225)이면서도, 동시에 “오늘날에는 이상하게 느껴질 정도로 낮은 개념”(226)이기 때문입니다. 왜 이런 간극이 발생할까요?

사회적 자유의 경험은 헤겔이 <법철학> 서론에서 사용한 용어에 따르자면 “타자 속에서 자기 자신 곁에 존재함”이라고 요약될 수 있습니다. 그런데 이처럼 사회적 자유 개념이 칸트의 소위 ‘형식적’ 자유나 ‘주관적’ 자유와 대비되는바 헤겔의 실체적 자유나 객관적 자유 개념을 재구성한 것이라 할 때, 칸트주의자라면 애초에 이러한 헤겔적 자유 개념(실체적, 객관적 자유) 자체에 대해 의구심을 가질 것입니다. “사회적 자유 모델은 자유의 상호주관적 성격을 강조”(227)하고, 성찰적 자유는 그렇지 못한다고 할 때, 칸트주의자라면 성찰적 자유가 상호주관적이지 못하고 단순히 단독자의 자유라는 가정을 거부할 수도 있을 듯 보입니다. 발표자(그리고 아마도 호네프)도 인정하듯이, “도덕적 자유는 비판하고 거부하는 이유를 다른 사람들에게 설명하고 정당화할 능력과 자세를 전제”하고, “도덕적 주체는 급진적 거리두기와 단절 속에서도 혼자가 아니”(231)기 때문입니다. 이처럼 성찰적 자유는 칸트가 생각하듯이 준칙의 보편화 가능성 여부에 관련되고, “모든 이성적 존재의 동의를 받을 만한 것인가라는 기준에 맞춰 스스로 검토하고, 경우에 따라서는 거부할 수 있는 권한을 부여”(230)하기에 정당화가능성은 항상 이미 타인과의 연관을 가정하는 것처럼 보입니다. 만약 칸트주의자의 이런 생각이 옳다면, 사회적 자유만 상호주관성을 가진다고 할 수 있을까요? 만약 칸트주의자도 상호주관적인 자유를 생각할 수 있다고 한다면, 사회적 자유는 잉여적인 것으로 보일 것입니다.

오히려 사회적 자유 개념은 자유 개념의 일종이 아니라, 자유를 위한 제도 또는 발표자가 논문 말미에서 언급하듯이 연대 같은 개념으로 불려야 하는 것일 수도 있습니다. 관련하여 발표자가 “성찰적 자유는 현실에서 **도덕이란 제도를** 통해서 구현”(230, 논평자의 강조)된다고 말하는데, 칸트주의자라면 도덕을 제도라고 부르는 것이 적절한 일인지, 오히려 도덕을 실현시키기 위한 제도(헤겔이라면 인륜이라고 불렀을)는 자유 그 자체는 아니지 않을까라고 물을 것 같습니다. 그렇게 본다면 자유의 사회적 조건이나 제도는 자유 개념의 구성 부분이나 “내적 계기”가 아니라 외적인 계기일 수도 있습니다. 유사한 맥락에서 선물 주기나 재활용 쓰레기 배출, 재즈 밴드의 예(232)도 제3의 자유로서 사회적 자유에 대한 것이라기보다는 도덕을 매개하는 에티켓이나 제도와 같은 사회적 규범 내지 관행의 문제로 볼 수도 있을 것입니다. 오히려 문제는 사회적 자유를 가능하게 하는 제도가 아니라 소극적, 성찰적 자유를 가능하게 하는 제도, “관행, 관습, 역할기대, 행위자의 적절한 태도와 행위방식 등의 복합체로서 ‘약한’ 의미에서의 제도”(235)일 수 있습니다. 만약 이런 생각이 옳다면, 서로 간의 행위를 협응시키고 보장하는 제도, 사회적 역할 기대를 사회화하는 것을 사회적 규범이나 제도로 부르는 것이 그것을 사회적 자유라고 부르는 것보다 더 자연스러워 보입니다. 그리고 이러한 제도는 연대와 마찬가지로 자유를 증진할 수도 억압할 수도 있는 것입니다.

아마도 논평자의 이런 질문은 발표자가 지적하는 사회적 자유 개념의 난점과도 연관될지도 모르겠습니다. 발표자 역시 호네프가 말하는 사회적 자유가 실상 우애나 연대를 상기시킨다고 인정하기 때문입니다(239). 호네프에 따르면 사회적 자유 개념에서 자유, 평등, 우애는 완전히

일치하게 됩니다. 여기서 그는 자유가 하나의 유형(type)이고 연대가 사례(token)라는 독특한 주장을 하는데, 사람들이 추구하는 연대는 발표자가 에리히 프롬(‘자유로부터의 도피’)에 의거하여 지적하듯이 자유의 경험(보다 정확히는 비강제성의 경험)을 위한 것이 아닐 수도 있을 것입니다. 이것이 발표자가 보는 호네프의 첫 번째 곤란입니다.

발표자가 보기에 보다 중요한 두 번째 곤란은 호네프의 사회적 자유 개념이 루소가 말하는 식으로 구성원들이 “마음이 같은” 상태, 상당한 정도로 가치, 목적, 문화적 자원이 공유되는 그런 이상적인 동질적 공동체를 전제하는 것 같다는 점입니다. 이는 롤즈의 표현을 빌리자면 민주주의 공적 문화의 영구적 특징으로서 “합당한 다원주의의 사실”을 고려하지 못한 것처럼 보입니다. 여기에서도 발표자는 호네프가 말하는 사회적 자유가 정치적 관점이라기보다는 롤즈가 말하는 포괄적 신념 체계에 가까운 것이라고 주장합니다. 다만 호네프 본인이 “사회적 자유의 원초적 장소는 우정과 사랑 같은 친밀한 관계이고, 그것을 경제 영역과 정치적 공론장을 비롯해 전 사회로 확장해서 실현하고자”(243) 한다는 발표자의 지적에 동의할지는 의문입니다. 근대 이후 돌이킬 수 없는 사회의 분화를 고려하면, <법철학>을 현대적으로 재구성하는 호네프의 논의는 사랑, 권리, 연대의 영역을 각기 고유한 영역으로서 구별하지 않을 수 없기 때문입니다. 사랑과 연대의 영역이 닮은 구석이 있을 수는 있으나 호네프가 사회적 자유 또는 연대를 정말로 소박하게 사랑과 같은 친밀성을 모델로 삼고 있는지는 더 자세한 탐구를 요구할 듯 합니다.

호네프의 두 번째 곤란과 관련하여 왜 호네프의 관점을 구태여 롤즈를 따라서 이해해야 할 까라는 질문을 제기하기 전에, 저는 위의 주장을 읽고 발표자 역시 제가 앞에서 제기한 질문처럼 실은 칸트주의적인 의무론적 관점에서 호네프를 이해하는 것이 아닐까라는 생각이 들었습니다. 즉 왜 사회적 자유가 제3의 자유인가? 소극적 자유와 성찰적 자유로도 충분하고 호네프가 사회적 자유라고 부르는 것은 오히려 사회적 관행이나 제도 같은 것이 아닌가? 라는 질문이 바로 칸트주의자가 던질 법한 질문이기 때문입니다.

지금까지 한 이야기가 발표자의 본의에 대한 부정확한 이해일지는 모르겠으나, 적어도 생산적 오해여서 발표자의 추후 작업과 청중들의 토론에 자극이 되길 바랍니다. 많은 것을 공부할 수 있었던 좋은 논문을 읽게 되어 감사하다는 말씀을 드리며 논평문을 마칩니다.

The background is a light gray gradient. On the left side, there is a complex network of thin gray lines connecting various black dots of different sizes, creating a web-like structure. Scattered across the entire background are several thin, light gray triangles of various sizes and orientations. Some of these triangles are solid, while others are just outlines. In the upper right corner, there are small, faint clusters of dots.

인공 지능의 도덕적 지위와 현상적 의식

정태창 (공주대학교)

인공 지능의 도덕적
지위와
현상적 의식

01

속성론적으로 이해된
기능적 환원주의 검토

02

TABLE OF CONTENTS

03

인공 지능의 도덕적
지위에 대한 관계론적
해석 검토

04

인공 지능에 대한
인간중심주의적 관점의
극복

01

인공 지능의 도덕적 지위와 현상적 의식



1. 인공 지능의 도덕적 지위에 대한 물음

2000년대 이후 머신 러닝(machine learning)의 한 분야인 딥 러닝이 비약적인 발전을 이루게 되면서

인간에 의존하지 않고 스스로 규칙 및 해답을 발견한다는 의미에서 자율성(autonomy)을 갖는 인공 지능이 등장하게 되었다.

이와 함께 이전에는 사물 혹은 도구로만 간주되었던 인공 지능에 도덕적 지위를 부여해야 하는가라는 물음이 첨예하게 대두되고 있다.



1. 인공 지능의 도덕적 지위에 대한 물음(계속)

인공 지능이 딥 러닝에 기반하여 내린 결정은 현재로서는 인간의 관점에서 설명될 수 없다.

2016년의 알파고-이세돌 대국은 이 점을 잘 보여주었다.

인간에 대해 독립적으로, 그리고 인간보다 우월한 결정을 내리면서 그 결정의 근거에 대해 인간이 설명조차 할 수 없는 존재자가 등장한 것이다.



2. 도덕적 지위에 대한 입장

어떤 대상의 도덕적 지위(moral status)를 결정하기 위한 접근 방식은 크게 다음의 두 가지가 있다.

- 속성론적 접근
- 관계론적 접근



3. 속성론적 접근의 개요

속성론적 접근에 따르면 어떤 존재자의 도덕적 지위는 그것이 내재적으로 보유하는 **비관계적 속성(non- relational property)**에 의해 결정된다.

이러한 기준이 되는 속성으로는 합리성(rationality), 자유 의지, 자율성, 쾌고감수능력(sentience), 현상적 의식(phenomenal consciousness) 등이 거론된다.

현존하는 인공 지능은 이러한 속성을 갖지 못하기 때문에 **속성론적 접근을 따를 경우 인공 지능의 도덕적 지위는 대체로 부정된다.**

이러한 입장을 취하는 학자로는 드레이퍼스(H. Dreyfus), 쉘(J. Searle), 그리고 최근에는 존슨(D. Johnson) 등이 있다.



3. 속성론적 접근의 개요(계속)

우리는 어떤 존재자가 도덕적 지위를 갖기 위해서는 그것이 의식을 가져야 하고, 쾌락과 고통을 느낄 수 있어야 한다고 생각하는 경향이 있다.

예를 들어 동물은 의식을 가진 존재로서 쾌락과 고통을 느낄 수 있기 때문에 동물에게 불필요한 해악을 가하는 것은 부도덕한 행위로 간주된다.

반면 길가에 있는 돌을 걷어차는 행위는, 그 행위가 다른 사람에게 피해를 주지 않는다면, 부도덕한 행위로 간주되지 않는다.

왜냐하면 돌에는 의식이 없기 때문이다.



3. 속성론적 접근의 개요(계속)

따라서 속성론적 접근에 따라 인공지능의 도덕적 지위를 결정하고자 할 경우, 인공지능이 현상적 의식을 가질 수 있는가가 관건이 된다.

의식을 가질 수 있는 인공지능이 등장한다면, 속성론적 접근을 따르더라도 인공지능은 도덕적 지위를 갖는 것으로 인정될 가능성이 높다.

그런데 속성론적 접근을 지지하는 학자들은 대체로 인공지능이 의식을 가질 가능성 자체를 부정하는 경우가 많다.



4. 관계론적 접근의 개요

관계론적 접근에 따르면 어떤 존재자의 도덕적 지위는 '해당 존재자는 실제로 무엇인가'라는 인식적 물음에 의해서가 아니라 '**우리는 해당 존재자와 어떠한 관계를 맺고 있는가(혹은 맺고자 하는가)**'라는 실천적 물음에 의해서 결정된다.

즉 **도덕적 지위**는 어떤 존재자가 갖고 있는 속성을 근거로 하여 객관적으로 인식되는 것이 아니라, **우리의 일상적인 삶 속에서의 관계 맺음의 방식에 따라 실천적으로 구성(construct)된다는 것이다.**

따라서 이러한 관계론적 접근을 따를 경우 인공지능이 전통적으로 도덕적 지위를 얻기 위한 근거로 간주되어온 현상적 의식 등의 속성을 결여하더라도 도덕적 지위를 부여받게 될 가능성이 열리게 된다.

이러한 입장을 취하는 학자로는 코겔버그(M. Coeckelbergh), 건켈(D. Gunkel), 고든(J.-S. Gordon) 등이 있다.



4. 관계론적 접근의 개요(계속)

개체가 어떤 속성을 가졌는지를 확인하는 것이 어렵다면, 그리고 도덕적 지위의 개념이 실천적 관계를 위한 것이라면, 어떤 존재자가 그러한 지위를 갖는 것으로 인정될 수 있는지를 따져보는 실천적 방식이 더 효과적일 수 있다는 것이다.

이러한 사고 방식은 튜링 테스트의 취지와 비슷하다. 튜링 테스트는 기계가 지능이라는 속성을 내재적으로 가졌는가보다는, 기계가 생각한다고 인정될 수 있는가의 물음을 다루기 때문이다.



4. 관계론적 접근의 개요(계속)

관계론적 접근은 속성론적 접근의 배제의 문제를 극복할 수 있다는 점에서 매력적인 대안으로 간주된다.

그것은 인류의 도덕적 진보의 과정으로서의 '포함(inclusion)'의 역사, 즉 평등한 도덕적 지위를 인종, 나이, 성별, 계층 등의 속성과 상관없이 인간 일반으로 확장하고, 더 나아가 동물을 비롯한 생물들에게 일정 수준의 도덕적 지위를 부여함으로써 우리의 도덕 공동체 안으로 포함시켜온 역사에 잘 부합하는 관점으로 생각된다.

관계론적 접근에 따라 '인격성(personhood)'의 관념을 인공지능으로 확장하고 인공지능을 우리의 도덕 공동체 안으로 포함시키는 것은 그 자체로 도덕적 진보라는 것이다.



5. 본 발표의 핵심 내용

본 발표는 **현상적 의식이 인공지능에 도덕적 지위를 부여하기 위한 필요조건을 이룬다**는 점을 보이기 위한 논증의 개요를 제시하고자 한다.

이를 위해 현상적 의식을 갖지 않는 인공지능에 도덕적 지위를 부여할 수 있다고 주장하는 두 가지 주요 입장, 즉

- 1) **속성론적으로 이해된 기능적 환원주의**
- 2) **인공지능의 도덕적 지위에 대한 관계론적 해석**

의 문제점을 분석하는 방식으로 논의를 전개하고,

본 발표의 주장이 타당할 경우 **인공지능에 대한 관점이 어떻게 변화할 필요가 있는지**를 살펴보고자 한다.





02

속성론적으로 이해된 기능적 환원주의 검토

1. 속성론적으로 이해된 기능적 환원주의

기능적 환원주의(functional reductionism)는 속성론의 관점에서 현상적 의식을 갖지 않는 인공 지능에 도덕적 지위를 부여하기 위한 이론적 기초를 이루고 있다.

차머스(D. Chalmers)에 따르면 심성적 상태는 '행동을 산출할 때 올바른 종류의 인과적 역할을 수행하는' **심리적 상태(psychological state)**와 질적 특성을 갖는 상태인 **현상적(phenomenal state)**로 구분될 수 있다.

이러한 구분은 블록(N. Block)의 **접근 의식(access consciousness)**과 **현상적 의식(phenomenal consciousness)**의 구분으로 이어진다.

전자는 의식의 기능적, 인과적, 인지적 관점에서 설명가능한 측면이고 후자는 그러한 관점에서 설명불가능한 의식의 질적 측면을 나타낸다고 할 수 있다.



1. 속성론적으로 이해된 기능적 환원주의(계속)

인공 지능의 도덕적 지위에 대한 기능적 환원주의에 입각한 논변의 핵심은 다음과 같이 요약될 수 있다.

- 1) 어떤 존재자가 접근 의식을 갖는 경우, 그것은 도덕적 지위를 갖는다.
- 2) 인공 지능은 복수실현의 방식으로 접근 의식을 가질 수 있다.
- 3) 따라서 인공 지능은 도덕적 지위를 가질 수 있다.

이러한 논변은 특정한 속성, 즉 의식의 기능적 구현을 인공 지능이 도덕적 지위를 갖기 위한 조건으로 놓고 있다는 점에서 속성론의 입장을 취하는 것으로 이해될 수 있다.



2. 기능적 환원주의 비판(1): 의식의 본질은 기능인가?

기능(function) 및 그것의 복수실현 가능성(multiple realizability)이라는 개념은 특히 도구에 적용하면 잘 들어맞는 것으로 보인다. 도구는 기능 외에 별다른 본질을 갖지 않는 것으로 생각되기 때문이다.

예를 들어 '사람이 앉을 수 있음'이 의자의 기능이라면, 이 기능을 잘 실현하는 한에서 그것의 물리적 구성이나 메커니즘은 무엇이 됐든 의자의 본질과는 별 상관이 없을 수도 있을 것이다.

하지만 그렇다면 **기능 개념은 본질적으로 도구가 아닌 존재자에게는 적용되기 어려운 것이 아닐까?**

예를 들어 개 혹은 고양이의 기능에 대해 말할 수 있는가? 개의 전형적인 행동 패턴을 묶어 개의 '기능'으로 규정하고 그것을 기능적으로 복수실현하면, 개가 구현된 것인가?



2. 기능적 환원주의 비판(1): 의식의 본질은 기능인가?

도구의 기능은 그것을 사용하는 주체와의 관계 하에서만 의미를 갖는 것으로 보인다.

예를 들어 경비견 대신 방범 장치를 설치함으로써 경비견을 기능적으로 교체했다고 하면, 그러한 교체는 '경비'라는 기능을 필요로 하는 주체와의 관계에서만 등가적으로 기능하는 것이다.

그렇다면 이러한 의미에서 의식을 이루는 일련의 '기능'들에 대해 말하는 것은 어쩌면 범주적 착오가 아닐까?

의식에 대한 기능적 환원주의를 설명하면서 거의 예외없이 엔진 등의 도구를 사례로 들고 있는 것은 우연만은 아닌 것으로 보인다.



3. 기능적 환원주의 비판(2): 방만한 기능주의의 문제

기능주의에 기초한 복수실현에 있어 관건은 기능의 개념을 어떻게 해석해야 하는가이다. 여기에는 다음과 같은 문제가 있다.

관건이 되는 기능의 개념을 너무 구체적으로 규정하면 복수실현이 거의 불가능해진다. 반면 그것을 너무 추상적으로 규정하면 일종의 개념적인 분류에 그치는 방만한 기능주의로 귀결된다.

예를 들어 인류의 역사와 함께 해온 활을 생각해보자. 활의 기능을 추상적으로 규정하면 장난감 활도 활의 기능을 복수실현하게 된다. 반면 활의 기능을 특정 갑옷을 특정 거리에서 관통하는 살상력 등으로 구체적으로 규정하면, 그것은 복수실현이 그만큼 어렵게 된다.

대상의 기능과 그것을 구현하는 재료, 메커니즘 사이에 필연적 연관 관계가 성립하기 시작하는 것이다.



3. 기능적 환원주의 비판(2): 방만한 기능주의의 문제



우리가 복수실현하고자 하는 대상이 도구라면, 기능 해석의 문제는 간단하게 해결될 수 있다. 그것을 사용하는 주체가 구체성의 수준을 결정하면 된다.

반면 의식에 대해서는 어떠한가? 기능주의의 관점에서 접근 의식을 구성하는 것으로 간주되는 수많은 '기능'들을 어느 정도의 구체성의 수준에서 해석해서 구현해야 의식이 복수실현되었다고 할 수 있는가?

3. 기능적 환원주의 비판(2): 방만한 기능주의의 문제(계속)

신상규는 이와 관련된 어려움을 다음과 같이 요약하고 있다(2010, <'기능'과 기능주의>, p.113).

인과적 구조의 척도를 너무 잘게 설정할 경우 인간 중심주의에 빠질 가능성이 커진다. 구조에 대한 기능의 규제를 감안하면, 인과적 역할을 상세히 규정하면 할수록 이는 인간에 유사한 특성의 신경 구조에 대한 요구로 이어지고, 중국에는 심성에 대한 인간중심주의로 회귀하게 될 것이다. 기능주의는 복수실현가능성을 유지하기 위해서 그러한 물리적 상세보다는 추상화된 층위에 자리 잡아야만 한다. 그런데 인과 구조의 척도를 너무 굵게 설정할 경우, 이는 반대로 튜링기계 기능주의가 될 위험성이 커지며, 모든 인과적 역할을 심성의 표지로 간주하게 되는 심성의 편재성, 보편성에 대한 주장으로 향하게 된다.



3. 기능적 환원주의 비판(2): 방만한 기능주의의 문제(계속)

노형래는 신상규의 논의를 이어받아 다음과 같이 지적한다(2021, <진화론적 기능 개념에 따른 복수실현

생물 시스템의 기능이 복잡해질수록 그에 따르는 구조적 제약 조건도 다양해진다. 인간의 마음은 다양한 모듈들의 조직적 협업을 통해 작동하는데, 그러한 작동방식을 복수실현하는(또는 인간의 두뇌와 전혀 다른 구조를 갖는) 실현자의 존재는 알려진 바 없다. 즉, 생선의 눈이나 동물의 날개가 복수실현된 것과 달리 인간의 마음은 복수실현되지 않았다. 이는 신경구조만이 진화적 정상 조건 아래에서 작동 가능한, 즉 인간 마음의 원인론적 기능을 실현하는 기계장치일 수도 있다는 가설에 대한 귀추법적(abductive) 증거가 된다. 반면, 비-신경구조의 로봇도 인간 마음의 원인론적 기능을 실현할 수 있는지에 대한 증거는 없다.



3. 기능적 환원주의 비판(2): 방만한 기능주의의 문제(계속)

지금까지의 논의는 예컨대 현존하는 컴퓨터가 인간의 마음을 복수 실현한다는 식의 널리 퍼져있는 생각이 방만한 기능주의에 기초해 있을 가능성을 강하게 시사한다.

의식의 '기능'들을 어느 정도의 구체성의 수준에서 기능적으로 구현해야할지에 대한 기준을 (어쩌면 원리적으로) 찾기 어렵고, 현존하는 인공지능의 아키텍처를 통해 그것을 구현하는 것이 가능하다고 믿을만한 근거도 제시된 바 없는 것이다.

여기서 한 가지 문제를 더 제기해보자면, 의식의 '기능'에 대해 현상적 의식이 비본질적이라고 믿을만한 근거도 부족하다는 점이다. 어쩌면 현상적 의식은 마음의 잔여물(residue) 같은 것이 아니라, 의식이 올바르게 '기능'하기 위한 본질적인 부분을 이룰지도 모른다. 심적 인과라는 것이 가능하려면 그렇게 되어야 하지 않을까?



4. 기능적 환원주의 비판(3): 도덕적 지위와 기능

의식의 '기능'들에 대한 실현의 문제는 일단 접어두고, 인간과 구별 불가능한 수준까지 접근 의식을 기능적으로 구현한 인공지능이 있다고 상상해보자.

이러한 인공지능에게 도덕적 지위를 부여해야 할까? 만약 그렇다면 그 이유는 무엇인가?

접근 의식과 도덕적 지위 사이의 연관 관계가 무엇인지는 분명하지 않다.



4. 기능적 환원주의 비판(3): 도덕적 지위와 기능(계속)

적어도 일부의 동물은 도덕적 지위를 갖는다. 그러면 그것은 동물들이 인간과 접근 의식의 측면에서 유사성을 갖고 있기 때문일까?

만약 의식의 '기능'이 도덕적 지위를 부여할 근거가 된다면 인간의 언어에 대한 처리능력에 있어 모든 동물들을 능가하는 것으로 보이는 Siri를 비롯한 일련의 현존하는 인공 지능들은 모두 기능적 구현의 정도에 따라 도덕적 지위를 부여받아야 할 것이다. 하지만 이는 반직관적인 것 같다.

개나 고양이는 Siri만큼 인간의 언어를 이해하지는 못하는 것으로 보이지만, 그럼에도 분명 도덕적 지위를 갖는다. 그것은 그들이 현상적 의식을 갖기 때문에, 그리고 인간과 마찬가지로 쾌락과 고통을 느낄 능력을 갖기 때문이 아닐까?





03

인공 지능의 도덕적 지위에 대한 관계론적 해석의 검토

1. 관계론적 해석에 대한 설명

관계론적 해석의 관점에서는 인공지능이 접근 의식을 기능적으로 복수실현하는 것은 그것에 도덕적 지위를 부여하기 위한 근거가 되지 않는다.

이러한 관점에서는 인공지능의 도덕적 지위는 관계 속에서 구성되는 것이므로, 인공지능에 의한 접근 의식의 구현은 인간이 그것을 도덕적 지위를 갖는 존재자로 대우하기 위한 하나의 배경조건에 불과하다.

관계론적 해석에서 의식의 구현 여부는 도덕적 지위의 문제와 적어도 원칙적으로는 무관하다고 할 수 있을 것이다.



2. 관계론적 해석에 대한 일반적인 비판

관계론적 해석에 대해서는 일반적으로 그것이 도덕적 지위에 대한 일종의 상대주의로 귀결된다는 비판이 제기된다.

즉 이러한 방식으로 도덕적 지위를 얻지 못할 존재자는 원칙적으로 존재하지 않는다는 점에서 도덕적 지위를 너무 주관적인 것으로 만든다는 것이다(Müller, 2021; Mosakas, 2021; Gibert & Martin, 2022).

관계론적 해석의 주창자인 건켈이나 코겔버그 등은 관계론적 해석이 '진리의 다수성을 허용'한다거나, '역동적인 윤리적 사유의 방식'에 해당한다거나 등등의 말로 이에 대해 방어하고 있지만, 이러한 대응은 위의 비판을 인정하는 것이 다름없어 보인다.



3. 관계론적 해석과 에우티프론의 문제

관계론적 해석은 도덕적 지위에 대한 일종의 구성주의 (constructivism) 관점에 기초해 있다. 그렇다면 구성주의에 대한 일반적인 비판이 관계론적 해석에도 적용될 수 있을 것이다.

모든 구성주의는 다음과 같은 '에우티프론의 문제'에 직면하게 되는 것으로 보인다.



소크라테스: 그렇다면, 에우티프론! 당신이 말하듯, 신들의 사랑을 받는 것이 경건한 것이 아니고, 경건한 것이 신들의 사랑을 받는 것도 아니니, 둘은 서로 다른 것이오.

에우티프론: 어째서죠, 소크라테스님?

소크라테스: 경건한 것은 이 때문에, 즉 그것이 경건하기 때문에 사랑받게 되지, 그것이 사랑받기 때문에 경건하지는 않다는 데 우리가 동의하고 있기 때문이오. 아니 그렇지소?

3. 관계론적 해석과 에우튀프론의 문제(계속)

에우튀프론의 문제를 관계론적 해석에 적용하면 다음과 같은 결론에 이르게 된다.

즉 우리가 인공지능에게 도덕적 지위를 부여하는 방식으로 관계를 맺고 싶기 때문에 인공지능이 도덕적 지위를 갖는 것이 아니라, 반대로 인공지능이 도덕적 지위를 갖기 때문에 우리가 인공지능에게 도덕적 지위를 부여하게 된다는 것이다.

만약 우리가 인공지능에게 도덕적 지위를 부여하고 싶다면 곧 그것이 도덕적 지위를 갖게 된다는 식으로 주장한다면, 도덕적 지위는 매우 상대적이고 주관적이고 불안정한 것이 되어 객관적 의미를 상실하게 될 것이다.



4. 인공 지능의 도덕적 지위와 의식의 '어려운 문제'

현상적 의식을 갖지 않는 인공 지능에게 도덕적 지위를 부여하려는 입장들은 해결하기 어려운 이론적 난관에 직면하게 되는 것으로 보인다.

그보다는 '**감성주의(sentientism)**', 즉 쾌고감수능력으로서의 감성 (sentience)을 갖는 모든 존재에게는 도덕적 지위가 있다는 입장을 받아들이는 것이 도덕적 지위에 대한 우리의 확신에도 부합하고, 지금까지의 관행에도 잘 들어맞는다.

우리가 곤충 등의 하등동물이나 식물들보다는 개나 고양이, 원숭이 등의 고등동물에 더 높은 도덕적 지위를 부여하는 것은 결국 쾌고감수능력에 대한 고려에 기초해 있는 것이 아닌가?



4. 인공 지능의 도덕적 지위와 의식의 '어려운 문제'(계속)

이러한 생각이 타당하다면, 인공 지능에 도덕적 지위를 부여하기 위해서는 먼저 차머스가 말하는 의식의 '어려운 문제', 즉 두뇌의 물리화학적 작용이 어떻게 주관적 경험을 일으키는가의 문제를 해결할 필요가 있다.

공상과학에서처럼 인공 지능이 '우연히', '저절로' 의식을 갖게 된다는 창발론적 생각에 의존할 것이 아니라면, 인공 지능이 현상적 의식을 가질 수 있도록 만드는 수밖에 없을 것이다.

하지만 그러한 것이 미래에 기술적으로 가능하더라도, 인공 지능을 쾌락과 고통을 느끼는 존재로 만드는 것이 옳은가라는 문제는 여전히 남게 된다.



04

인공 지능에 대한 인간중심주의적 관점의 극복



1. 인공 지능을 어떻게 이해해야 할까?



본 발표의 논의가 타당하다면, 그것은 인공 지능에 대한 관점에 대해 어떠한 귀결을 가질까?

2. 인간중심주의의 문제

관계론적 해석의 지지자들은 도덕적 지위를 인간과 동물 등의 자연종(natural kinds)에만 한정하려는 태도는 일종의 인간중심주의에 불과하다고 비판한다.

그러한 태도는 인공종(artificial kinds)을 처음부터 배제하는 일종의 종차별주의라는 것이다.

하지만 이러한 주장은 인공 지능이 도덕적 지위를 가질 수 있어야만 성립할 수 있다는 점에서 일종의 선결문제 요구의 오류를 범하고 있는 것으로 보인다.



2. 인간중심주의의 문제(계속)

관계론적 해석의 지지자들은 사실 다른 유형의 인간중심주의에 빠져 있는 것으로 생각될 수 있다.

그것은 인공 지능을 '인간의 형상에 따라' 이해하는 것, 인공 지능이 궁극적으로는 인간처럼 되어야 한다는 선입견을 갖는 것, 인공 지능을 가능한 한 인간과 동화하는 방식으로 이해하려는 유형의 인간중심주의이다.

하지만 이러한 선입견에서 벗어나 현존하는 인공 지능들을 실제로 살펴보면 그것을 '인간의 형상에 따라' 이해해야 할 본질적인 이유를 찾기 어려울 수도 있다.



3. 프랑켄슈타인 콤플렉스?

‘인간처럼 사고하고 행동하지만 인간과는 다른 인공물, 그리고 그런 인공물을 편견을 가지고 배척하는 인간’이라는 기본 아이디어는 메리 셸리가 거의 200년전에 소설로 그려낸 바 있고, 이후에도 아이작 아시모프 등을 거쳐 끊임없이 변주되어왔다.

이러한 문학적 상상력이 우리가 인공 지능을 이해하는 기본적인 틀을 이루어야 할까? 인간처럼 사고하고 행동하는 인공 지능이라고 하는, 현실에 존재하지 않는 이 관념은 사실 문학에서 내러티브 구성의 필요성에 의해 가공된 산물이 아닐까?

내러티브 구성에서는 인간처럼 사고하고 행동하지 않으면 ‘등장인물’이 될 수 없기 때문에, 인공물 또한 등장인물이려면 그렇게 그려질 수밖에 없다.



4. 인공 지능에 대한 과장

이와 관련하여 맥클루어(Maclure)는 최근의 인공 지능 낙관주의(AI inflationism)를 경계하면서,

그림 속에서 고양이를 구별해내거나 문장을 번역하는 등의 컴퓨터 알고리즘으로서의 현존하는 인공 지능과

인류의 생존을 위협하거나 인류의 구원을 이끌 압도적인 행위자로서의 인공 초지능이라는 상상된 인공 지능 사이에는 데우스 엑스 마키나가 개입할 수밖에 없다고 지적한다(Maclure, 2020).

이러한 데우스 엑스 마키나는 공상 과학에 기초하여 상상된 인공 지능과 현존하는 실제의 인공 지능 사이의 갭을 메우기 위해 요청되는 것은 아닌가?



5. 인공 지능에 대한 현실적인 이해 방식

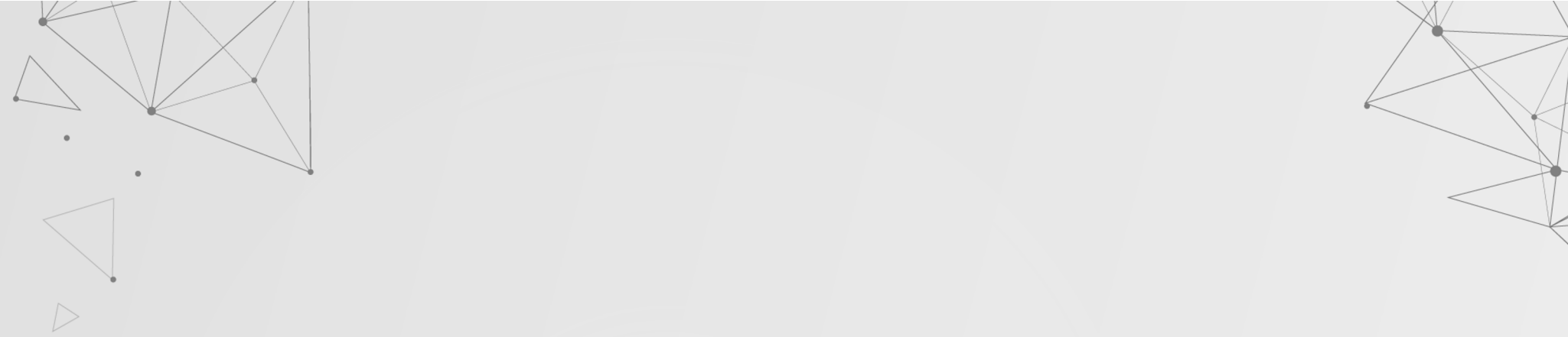
커즈와일의 특이점 논의나 보스트롬의 인공 초지능 종말론 같은 극단적인 논의는 배제하더라도,

그간의 인공 지능에 대한 논의를 살펴보면 인공 지능이 인간처럼 될 수 있고 또한 인간을 궁극적으로 넘어설 수 있다는 일종의 문학적 상상력이 인공 지능에 대한 우리의 이해 방식을 지배해온 것으로 보인다.

예를 들면 범용 인공 지능에 대한 논의가 그러한데, 사실 인공 지능이 도구라는 점을 고려하면 그것이 범용으로 발전해야 할 이유도 확실하지 않고 그것이 그렇게 발전할 거라고 믿을만한 기술적 이유도 별로 없다.

앞으로는 상상된 인공 지능보다는 구체적인 아키텍처에 기반해 있는 현실의 인공 지능에 초점을 맞추면서 그것이 인간과 갖는 차이점



Abstract geometric line art in the top corners, consisting of interconnected lines and dots forming various triangular and polygonal shapes.

감사합니다