

생겨난 이의 자기의식 정보*

김 명 석

【국문요약】 리처드 고트, 브랜던 카터, 존 레슬리, 닉 보스트롬 등은 데카르트가 그랬던 것처럼 “나는 지금 여기 관찰자로서 있다”를 다른 주요 믿음들의 증거로 삼았다. 인간 원리 또는 관찰선택 효과를 연구했던 보스트롬 및 여러 학자들은 두 가지 점을 놓치고 있다. 첫째, 처음 생겨난 이가 갖는 자기의식 정보는 다시 깨어난 이가 갖는 자기의식 정보와 다르다. 다시 깨어난 ‘나’는 기억을 더듬어 과거로 거슬러 갈 수 있지만 새로 생겨난 ‘나’는 그렇게 할 수 없다. 둘째, 믿음직함을 선택할 때 가능 세계 안에 있는 이들과 실현된 세계 안에 있는 이들을 구별해야 한다. 두 가능 세계들 가운데 하나만 실현되는 상황에서 믿음직함 선택은 두 세계들이 모두 실현되는 상황에서 믿음직함 선택과 근본 차원에서 다르다. 이 점들을 염두에 두면서 우리는 새로 생겨난 이의 자기의식 정보가 어떤 본성을 갖는지 탐구했다. 먼저 보스트롬의 인큐베이터 생각실험을 소개하고 이와 비슷한 두 인간 배아 생각실험을 주의 깊게 살펴보았다. 이로써 우리는 “나는 생겨났다”는 정보와 생겨난 이가 차츰 알게 된 정보는 “누군가 생겨났다”와 다를 바 없는 정보라는 것을 논증했다. 우리 논증이 옳다면 정보 “나는 지금 여기 처음 생겨났다”와 정보 “나는 지금 여기 다시 깨어났다”가 다른 믿음의 믿음직함을 바꾸는 정도는 다르다. 만일 두 정보가 다른 믿음의 믿음직함을 바꾸는 정도가 다르다면 두 정보는 다른 정보인 셈이다.

【주요어】 자기의식 정보, 확률, 믿음직함, 관찰선택 효과, 보스트롬

투고일: 2018. 9. 13 심사 및 수정 완료일: 2018. 11. 24 게재확정일: 2018. 11. 24

* 잘못 읽힐 가능성을 줄이도록 마땅한 조언을 주신 심사자들에게 감사드린다. 지적인 바를 반영하여 논문을 개선했다.

1. 들어가는 말

데카르트에게 ‘나는 생각한다’가 엄청난 정보였듯이 ‘나는 생각한다’는 현대 인식론과 과학철학에서도 엄청난 정보다. 1969년 리처드 고트는 자신이 베를린 장벽이 세워진 뒤 8년 만에 그 현장을 지켜보고 있다는 사실로부터 3년 후 내지 24년 후에 베를린 장벽이 무너질 확률이 50%라고 예측했다.¹⁾ 1974년 브랜턴 카터는 우리 관찰자가 여기에 있다는 정보를 우주론 연구를 돕는 핵심 증거로 여길 수 있다고 제안했다.²⁾ 1989년 존 레슬리는 지구에 존재했던 인구들 가운데 자신이 놓인 위치를 가늠함으로써 인류의 멸종이 예상보다 가까이 다가와 있다고 주장했다.³⁾ 2003년 닉 보스트롬은 우리가 미래 인류가 만든 가상 세계의 인물일지 모른다는 논증을 고안했다.⁴⁾

“내가 지금 여기 관찰자로서 있다”나 “내가 지금 여기 의식을 갖고 있다”가 우리의 다른 믿음들에서 정확히 어떤 역할을 하는지 성찰하는 연구는 인식론 탐구이자 과학철학 탐구이며 나아가 형이상학 탐구이기도 한다. “내가 지금 여기 처음 생겨났다”라는 정보는 “나는 지금 여기 다시 깨어났다”라는 정보와 다를 것 같다는 추정을 바탕으로 “내가 지금 여기 처음 생겨났다”라는 정보의 본성을 성찰하고자 한다. 생겨난 이의 자기의식 정보를 깊이 성찰함으로써 보스트롬 등 이 분야의 관련 연구들이 놓치고 있는 점을 드러내고 그들 연구를 보완하기를 바란다.

1) Gott (1993). 교양 수준의 논의는, 리처드 고트 (2003), 『아인슈타인 우주로의 시간여행』, 박명구 옮김, 한승.

2) Carter (2011). 보다 방대한 논의는, Barrow, J., Tipler, F. (1986), *The Anthropic Cosmological Principle*, Oxford University Press.

3) Leslie (1989).

4) Bostrom (2003a).

우리는 다음 순서에 따라 생겨난 이의 자기의식 정보를 탐구한다. 이 주제를 깊게 연구한 보스트롬의 연구를 살펴볼 텐데 제2절에서 그의 인큐베이터 생각실험을 소개한다. 제3절에서 이 생각실험과 비슷한 두 배아 문제를 풀고 이 결과를 성찰한다. 제4절과 제5절에서 생겨난 이가 차츰 알게 된 정보가 자신이 생기기 전에 무슨 일이 일어났는지를 짐작하게 하는 실마리가 될 수 있는지 따진다. 나오는 말에서 우주론과 관련된 우리 논증의 함축을 살펴본다.

이 글에서 우리가 가정하는 것을 먼저 말해 놓겠다. 처음 생겨난 이가 자신은 다시 깨어난 이라고 착각하거나 다시 깨어난 이가 자신은 처음 생겨난 이라고 착각할 가능성은 당연히 있다. 하지만 우리는 정보 “나는 지금 여기 처음 생겨났다”나 정보 “나는 지금 여기 다시 깨어났다”가 주체에게 직접 주어진다고 가정한다. 주체에게 주어진 정보는 정보의 말뜻 그대로 참인 정보다. 우리는 현실의 주체들이 실제로 이러한 두 자기의식 정보들을 가리고 직접 알아채는지를 탐구하려는 것이 아니다. 우리는 주체가 그러한 정보를 가졌을 때 그의 다른 믿음들을 어떻게 재조정해야 하는지를 탐구하려고 한다. 정보 “나는 지금 여기 처음 생겨났다”와 정보 “나는 지금 여기 다시 깨어났다”가 다른 믿음의 믿음직함을 바꾸는 정도가 다르다면 두 정보는 다른 정보다. 우리는 두 정보가 믿음직함을 바꾸는 정도가 다르다는 점을 이 글에서 논증한다.

2. 보스트롬의 인큐베이터 생각실험

보스트롬은 생겨난 이의 자기의식 정보를 탐구하려고 2002년 책과 2003년 논문에서 인큐베이터 생각실험을 제안했다.⁵⁾ 아무 것도 없

⁵⁾ Bostrom (2002), p. 64. Bostrom (2003b), p. 63. 우리는 Bostrom (2002)에 나온 버전을 다룰 것이고 우리 논의와 연관이 잘 드러나도록 실험 설정도

는 곳에 ‘인큐베이터’라는 자동장치가 동전을 던진다. 앞면이 나오면 방 하나에 검은 수염이 난 사람 하나를 생기게 한다. 뒷면이 나오면 각기 다른 방에 검은 수염이 난 사람과 흰 수염이 난 사람을 각각 생기게 한다. 방은 아예 캄캄해 자기 수염 색깔을 알지 못하지만 자신이 어떤 실험 설정을 거쳐 이 세계에 생기게 되었는지는 잘 알고 있다. 캄캄한 방에서 이들은 동전이 앞면이 나왔으리라고 얼마큼 믿어야 할까? 한 시간 뒤에 방이 밝아져 자기 수염 색깔이 검다는 것을 안 이는 동전이 앞면이 나왔으리라고 얼마큼 믿어야 할까? 보스트롬은 실험에 참여한 이들은 처음에 앞면이 나왔으리라 1/2만큼 믿지만 나중에는 2/3만큼 믿어야 한다고 답한다.⁶⁾

인큐베이터 생각실험에서 동전 결과에 따라 세계는 두 개의 가능 세계로 쪼개진다. 하나는 한 사람만 생겨나는 가능 세계고 다른 하나는 두 사람이 생겨나는 가능 세계다. 하지만 둘 가운데 오직 하나만 실현된다. 이 생각실험이 묻고자 하는 요지는 분명하다. 실현된 세계 안에서 실제로 생겨난 이는 자기 세계가 한 사람만 있는 세계라고 믿어야 할까 여러 사람들이 있는 세계라고 믿어야 할까? 우리는 이 상황을 일반화할 수 있다. 동전이 앞면이 나올 가능성이 h 고 뒷면이 나올 가능성이 t 인데,⁷⁾ 앞면이 나오면 한 명만 생

약간 바꾸겠지만 주요 내용은 바꾸지 않을 것이다. 동전이 앞면이 나올 때 일어나는 일과 뒷면이 나올 때 일어나는 일을 바꾸었다. 이 글에서 다루지는 않지만 다른 주요한 생각실험인 잠자는 미인 실험과 짝을 맞추려는 의도 때 문이다.

⁶⁾ Bostrom (2002), p. 65. 하지만 그는 증거 집단을 좀 더 넓게 잡고 그가 새로 고안한 관찰선택이론을 적용함으로써, 실험 참여자들이 자기 수염이 검다는 것을 안 다음에도 앞면이 나왔으리라고 1/2만큼 믿어야 한다는 결론을 이끌어내기도 했다. 같은 책, p. 174.

⁷⁾ ‘가능성’은 ‘일어남직함’(chance)을 뜻한다. 일어남직함은 믿음직함(credence)과 마찬가지로 콜모고로프 공리들을 따를 것을 요구받는다. 하지만 내 생각에 콜모고로프 공리들을 엄격하게 따라야 하는 것은 믿음직함밖에 없다. 믿음직함이 콜모고로프 공리들을 따라야 한다고 말하는 것은 그 공리들이 믿

기지만 뒷면이 나오면 모두 N 명이 각기 다른 방에 생겨난다. 이 경우 실현된 세계 안에 있는 한 생겨난 이는 자신이 속한 세계가 여러 명이 생겨난 세계라는 데 t 만큼 믿어야 할까 아니면 이보다 더 굳게 또는 더 흔들리게 믿어야 할까? 만약 이것이 t 보다 크다면, 이것은 자신이 생겨나 얻게 된 자기의식 정보 “나는 지금 여기 의식을 갖고 있다”가 많은 의식들이 생겨났음을 짐작하게 하는 증거라는 것을 뜻한다.

보스트롭과 다른 학자들이 놓치고 있는 점이 있다. 그것은 처음 생겨난 이가 갖는 자기의식 정보는 다시 깨어난 이가 갖는 자기의식 정보와 다르다는 점이다. 사안을 더욱 또렷하고 쉽게 파악하려고 두 사람 곧 남자 아담과 여자 이브를 생각해 보겠다. 이들은 이미 생겨났고 잠든 뒤 다시 깨어나면 자신이 누구인지 안다. 우리가 다음처럼 이들을 다시 깨우는 실험을 할 수 있다. 아담과 이브가 잠든 사이에 동전을 던져 앞면이 나오면 아담과 이브 가운데 한 사람만 깨우고 뒷면이 나오면 둘 다 깨운다. 이 때 다시 깨어난 이들은 던진 동전이 앞면이 나왔으리라고 얼마큼 믿어야 하는가? 다시 깨어난 이들은 앞면이 나왔더라면 자신이 줄곧 잠잘 수 있었을 텐데 지금 다시 깨어난 것을 보니 둘 다 깨어나야 하는 상황 곧 뒷면이 나왔으리라고 조금 더 굳게 믿게 될 것이다.⁸⁾ 하지만 이러한 실험 설정은 아담과 이브가 아예 없었던 상황에서, 앞면이 나오면 아담과 이브 가운데 한 사람만 생기게 하고, 뒷면이 나오면 아담과 이브 둘 다를 생기게 하는 설정과 다르다. 이제 생겨난 아담

음 개념의 핵심 구성 요소라고 말하는 것이다. 우리는 일어난직함과 믿음직함의 관계 문제에 깊이 들어가지 않을 것이다. 말이 나온 김에 이 글에서 ‘믿음직함’이라고 쓴 것을 다른 많은 저자들은 ‘신념도’라고 쓴다.

⁸⁾ 이 문제를 “깨어난 이들의 문제”라고 부를 수 있다. 이 실험에 참여하는 이가 N 명이고 앞면이 나올 가능성이 h 인 경우 이 물음의 답은 $h/(h+tN)$ 이다. 여기서 t 는 $1-h$ 다. 김명석 (2017), p. 15.

과 이브는 단지 그들이 생겨났기 때문에 생겨난 자신에 관해 생각할 수 있을 뿐이다.

보스트롬 및 다른 학자들이 놓치고 있는 점이 또 하나 더 있다. 믿음직함을 선택할 때 단지 가능한 세계 안에 있는 이들과 실현된 세계 안에 있는 이들을 구별해야 한다. 인큐베이터 생각실험에서 생겨난 이들의 정보는 다음과 같은 생각실험에서 생겨난 이들의 정보와 매우 다르다. 우리 우주에는 두 세계가 있다. 세계 A에는 한 사람만 생겨나지만 세계 B에는 두 사람이 생겨난다. 다만 세계 B에 사는 이는 다른 이의 존재를 알지 못한다. 이런 우주에 지금 의식을 가진 이는 자신이 세계 A에 살고 있다고 믿어야 하는가 세계 B에 살고 있다고 믿어야 하는가?⁹⁾

생겨난 이들이 모두 3명이라는 것을 아는 이는 이 물음에 아주 쉽게 답할 수 있다. 자신이 세계 A에 살고 있으리라 믿는 정도는 1/3이고 세계 B에 살고 있으리라 믿는 정도는 2/3다. 만일 실제로 존재하는 이들이 실현된 세계들에 존재하는 이가 모두 3명임을 안다면 그들은 자신이 그 셋 가운데 하나이며 셋 가운데 둘은 세계 B에 살고 있다고 생각해야 한다. 보스트롬은 이러한 생각방식을 “자기 표본추출 가정”이라고 불렀다.¹⁰⁾

9) 레슬리 (1998), p. 239.

10) Bostrom (2002), p. 57. “우리는 우리가 우리 준거 집단 안에 있는 모든 관찰자들의 집합에서 마구잡이로 뽑힌 표본인 양 추리해야 한다.” 그는 이 원리가 몇몇 사례들에서 믿음직함을 잘못 선택하게 한다는 점을 지적하고 그 대안이 되는 “강한 자기 표본추출 가정”(strong self-sampling assumption)을 제안한다. 믿음직함을 선택하는 이는 지금의 의식 순간을 준거 집단 안에 있는 모든 의식 순간들의 모임에서 마구잡이로 뽑힌 한 표본으로 여겨야 한다는 가정이다. 같은 책, p. 162. 보스트롬은 믿음직함을 선택하는 이가 처음에 생겨나는 경우와 다시 깨어나는 경우 모두에 자기 표본추출 가정을 똑같이 적용할 수 있다고 본다. Bostrom (2003b), p. 62.

믿음직함을 셈하는 이는 자신을 준거 집단 안에 있는 모든 의식들의 모임에서 마구잡이로 뽑힌 한 표본으로 여겨야 한다.

자신이 준거 집단 안에서 매우 특별한 사람인 양 뽑내서는 안 된다는 말이다. 세 의식들이 존재하고 그들과 나의 다른 점을 아직 알지 못한다면 나는 세 의식들 가운데 마구잡이로 뽑힌 한 사람이라고 생각해야 한다.¹¹⁾

이와 비슷한 실험 설정들을 여러 가지로 생각해볼 수 있다. 세 사람이 잠든 사이에 한 사람은 A동 기숙사에 몰래 옮겨 놓고 두 사람은 B동 기숙사에 각기 다른 방에 옮겨 놓는다. 깨어난 사람은 자신이 기숙사 어느 동에 놓여 있다고 믿어야 하는가? 자신이 A동 기숙사에 있다고 믿는 정도는 1/3이다. 또 비슷한 설정으로는 세 명이 태어날 텐데 한 명은 병원 A에서 태어나고 두 명은 병원 B에서 태어난다. 자신이 이렇게 태어난 셋 가운데 하나라는 것을 안다면, 그는 어느 병원에서 태어났다고 믿어야 하는가? 그가 병원 A에서 태어났다고 믿는 정도는 1/3이어야 한다. 하지만 이러한 설정들은 보스트롬의 인큐베이터 생각실험과는 매우 다르다.

내가 다시 깨어난 곳이 어디든, 내가 처음 태어난 곳이 어디든, 자기의식이 처음 출현한 세계가 어디든, 현재 실제로 존재하는 이들이 모두 셋이라면 나는 이 셋 가운데 하나일 뿐이다. 하지만 우리가 풀고자 하는 물음에서 실제로 존재하는 이들은 셋이 아니라 ‘하나 또는 둘’이다. 한 경우에는 오직 한 사람만 생겨나고 다른

11) 이 유형의 문제에서 보스트롬이나 레슬리 등은 믿음직함을 잘못 셈하고 있지는 않다. 하지만 그들은 실현된 세계들 가운데 우리가 어디에 있는지 가늠하는 물음과 가능 세계들 가운데 무엇이 실현되어 그 안에 우리가 있게 될지를 가늠하는 물음을 매우 다르게 접근해야 한다는 것을 또렷이 인지하지는 못한 것 같다. 둘을 구별해야 한다는 연구로는 김명석 (2018)이 있다. 가능 세계들을 실현된 세계들과 비슷하게 여김으로써 믿음직함을 셈할 때 잘못된 믿음직함 셈에 이를 수 있다. 김명석 (2016a), p. 248.

경우에는 두 사람이 생겨난다. 내가 그렇게 생겨난 사람이라면 나는 생겨난 한 사람 가운데 바로 그 사람이거나 생겨난 두 사람 가운데 하나다.

3. 생겨난 이들은 다시 깨어난 이들이 아니다.

우리가 풀고자 하는 생겨난 이들의 문제를 정식화해 보겠다. 하느님은 일요일 밤늦게 한 사람 또는 두 사람을 생기게 할 것이다. 한 사람을 만들 가능성은 h 고 두 사람을 만들 가능성은 t 다. 생겨난 이들은 서로 멀리 떨어져 있어서 다른 이가 생겨났는지 알지 못한다. 월요일 아침 생겨난 이가 자신이 어떤 과정을 거쳐 생겼는지 알게 되었다. 자기의식 정보를 가진 그는 자기말고 다른 의식도 생겨났으리라고 얼마큼 믿어야 할까? 그것은 t 일까 t 보다 클까? 자기의식 정보는 다른 의식이 어딘가에 있다는 가설을 입증할까? 이야기를 쉽게 하도록 h 와 t 를 모두 1/2로 두되 일반 풀이를 제안할 때는 다시 h 와 t 를 변수로 나타낼 것이다. 멀쩡한 동전을 던져 앞으로 떨어질 일을 정하면 h 와 t 를 1/2로 둘 수 있다.

태어난 사람의 자기의식 정보는 다시 깨어난 이가 아니라 처음 생겨난 이의 자기의식 정보를 닮는다. 두 부부가 있는데 동전을 던져 한 부부만 아이 하나를 낳거나 두 부부 모두 아이 하나씩 낳기로 공동 가족계획을 세웠다고 해보자. 이렇게 만들어진 아이들이 스무 살이 되면 출생에 얽힌 비밀을 듣게 된다. 이런 정보를 가진 그들은 자기말고 다른 사람도 그런 방식으로 태어났으리라고 얼마큼 믿어야 할까? 자기 출생의 비밀은 자신과 같은 방식으로 태어난 다른 이가 어딘가에 있다는 가설을 입증할까? 이 문제를 ‘태어난 이들의 문제’라 부르겠다.

괴짜 유전공학자 지오디는 동전을 던져 한 인간 배아를 사람으

로 자라나게 하거나 두 인간 배아를 사람들로 자라나게 할 작정이다. 이렇게 자라날 이들이 스무 살이 되면 지오디의 실험 설정을 듣게 된다. 이런 정보를 가진 그들은 자기 말고 다른 사람도 그런 방식으로 자라났으리라고 얼마큼 믿어야 할까? 자기 생성의 비밀은 자신과 같은 방식으로 자라난 다른 이가 어딘가에 있다는 가설을 입증할까? 이 문제를 ‘두 배아 문제’라고 부르겠다.

방금 이야기한 생겨난 이들의 세 가지 문제들은 믿음직함에 관한 한 똑같은 문제다. 이 문제들의 대표로 두 배아 문제를 주로 이야기하겠다. 유전공학자 지오디는 동전이 앞면이 나오면 인간 배아 하나를 키우고 뒷면이 나오면 인간 배아 둘을 키운다. 새로 생겨나 실험에 참가하게 된 이들은 동전을 던질 때 아예 없었다. 그들은 동전을 던진 다음에야 비로소 있게 되었다. 하느님의 창조 문제, 태어난 이들의 문제, 두 배아 문제의 공통점은 실험 참여자들이 두 가능 세계로 쪼개지기 전의 세계에 존재한 적이 없었기 때문에 그때의 기억을 갖지 않는다는 점이다.¹²⁾ 그들은 의식을 되찾은 것이 아니라 이제 처음 의식을 얻었다. 반면 다시 깨어나는 실험의 참여자들은 처음 의식을 갖는 것이 아니라 의식을 되찾는다. 우리는 태어난 이들 곧 생겨난 이들의 자기의식 정보가 다시 깨어난 이들의 자기의식 정보와 어떤 점에서 다르고 어떤 점에서 같을지 알아내고 싶다.

지오디가 처음부터 두 개의 배아를 준비하여 각각 ‘이브’와 ‘아담’이라고 이름지었다고 생각해 보자. 동전이 앞면이 나오면 배아 이브와 배아 아담 가운데 하나만 실제 사람으로 자라게 하고, 뒷면이 나오면 두 배아를 모두 사람으로 자라게 한다. 이렇게 자라난 이들이 명제 X에 매기는 믿음직함 값을 $C(X)$ 라 쓰겠다. 그들이 명

12) 실험 참여자에게 가짜 기억을 주입하는 경우와 그가 착각하는 경우는 우리 관심사가 아니다. 가짜와 착각은 오직 진짜 기억과 실제 사실에 비추어서만 제대로 이해될 수 있다.

제 Y만을 새로 알게 된 다음 명제 X에 주는 믿음직함 값은 $C(X|Y)$ 며 베이즈 공리에 따라 이것은 $C(X\text{이고 } Y)/C(Y)$ 와 같다.¹³⁾

여러 명제들의 믿음직함을 셈해야 할 텐데 우리가 셈해야 하는 명제들을 다음과 같이 짧게 쓴다. 지오디가 동전을 던진 것은 아주 오래전이다.

- 앞: 동전이 앞면이 나왔다.
- 뒤: 동전이 뒷면이 나왔다.
- 아: 아담이 의식을 갖고 생겨난다.
- 이: 이브가 의식을 갖고 생겨난다.
- 나_아: 생겨난 나는 아담이다.
- 나_이: 생겨난 나는 이브다.

아담과 이브는 생겨났을 때 자신이 아담인지 이브인지 알지 못한 다. 이 실험을 살펴보는 외부자의 관점에서 먼저 셈해 보면,¹⁴⁾

$$\begin{aligned} C_{외}(아) &= C_{외}(아\text{앞})C_{외}(앞) + C_{외}(아\text{뒤})C_{외}(뒤) \\ &= (1/2)(1/2) + (1)(1/2) = 3/4 \end{aligned}$$

이다. 우리는 외부자가 명제 X에 매기는 믿음직함 값을 $C_{외}(X)$ 라고

13) 명제 Y만을 새로 알게 된 다음 명제 X에 매기는 믿음직함 값을 $C(X|Y)$ 로 쓰는 것은 $C(X|Y)$ 의 정의에 해당하며 우리는 이를 조건부 믿음직함의 정의로 여긴다. 많은 이들이 $C(X|Y)$ 와 $C(X\text{이고 } Y)/C(Y)$ 가 같다는 규칙을 조건부 믿음직함의 정의로 여기지만 우리는 이 규칙을 “베이즈 공리”라고 부른다. 그 규칙이 정의나 공리냐는 우리 결론을 바꾸지 못한다. 참고로 베이즈 공리와 베이즈 정리는 다른 규칙이다.

14) 명제 “A이거나 B”가 참이라는 것이 알려져 있고 또한 A와 B 가운데 오직 하나만 참이라는 것이 알려져 있을 때, 아무 명제 X의 믿음직함 $C(X)$ 는 $C(X|A)C(A) + C(X|B)C(B)$ 다.

썼다. 그에게 $C_{외}(아앞)$ 은 $1/2$ 이고 $C_{외}(아뒤)$ 는 1 이다. 무엇보다 $C_{외}(앞)$ 과 $C_{외}(뒤)$ 는 둘 다 $1/2$ 이다. 똑같이 셈해서 $C_{외}(이)$ 는 $3/4$ 이다.

생겨난 이들은 자신이 아담인지 이브인지 모르지만 다음 사실은 안다고 가정하겠다.¹⁵⁾ 우리 논증을 반박하고자 한다면 이 가정을 부정하면 된다.

$$(E1) C(아) = C(이) = 3/4.$$

의식을 새로 얻은 이가 셈해야 하는 것은 $C(앞)$ 또는 $C(뒤)$ 다. “동전이 뒷면이 나왔다”는 “아담과 이브가 모두 생겨났다”와 논리 동치다. 따라서 $C(뒤) = C(아\&이)$ 다. 한편 아무 명제 X 와 명제 Y 에 대해 $C(X \vee Y) = C(X) + C(Y) - C(X \& Y)$ 가 성립하기 때문에 우리는 다음을 얻는다.¹⁶⁾

$$(1) C(뒤) = C(아\&이) = C(아) + C(이) - C(아 \vee 이) \\ = 3/4 + 3/4 - 1 = 1/2.$$

우리는 (E1)과 $C(아 \vee 이)$ 가 1이라는 것을 썼다. $C(뒤)$ 가 $1/2$ 이니 $C(앞)$ 도 $1/2$ 이다.¹⁷⁾ 이것은 내가 생겨났다는 정보가 여러 사람들이 생겨났음을 입증하는 증거가 되지 못함을 뜻한다.

의식을 갖게 된 이가 “아담이 생겨났다”는 것 곧 “아”를 알게 되었다고 생각해 보자.¹⁸⁾ 이때 그는 동전이 앞면이 나왔으리라고

15) 누군가 명제 X 를 안다면 명제 X 는 참이다. 하지만 명제 X 가 참이라고 해서 누군가 그 명제를 아는 것은 아니다.

16) 말꼴 ‘&’는 ‘이고’를, 말꼴 ‘ \vee ’는 ‘이거나’를 뜻한다.

17) 뒷면이 나오면 새로 생겨나는 이들이 N 명이고 앞면이 나올 가능성이 h 인 일반화된 실험 설정에서 $C(앞)$ 은 h 다.

18) “아담이 생겨났다”가 새로 생겨난 이에게 주어지는 절차는 마구잡이 절차여야 한다. 다음과 같은 절차는 마구잡이 절차 가운데 하나다. 먼저 새로 생겨

얼마큼 믿어야 하는가? 이 값은 C(앞아)를 셈하면 되겠다. C(앞)은 1/2이기 때문에

$$(2) C(\text{앞아}) = C(\text{앞\&아})/C(\text{아}) = C(\text{아\앞})C(\text{앞})/C(\text{아}) \\ = (1/2)(1/2)/(3/4) = 1/3$$

이다. 마찬가지로 셈으로 C(앞이)는 1/3이다. 생겨난 이가 정보 “아담이 생겨났다”나 정보 “이브가 생겨났다”를 새로 얻게 될 때, 그에게 C(앞)은 1/2에서 1/3로 낮아지고, C(뒤)는 1/2에서 2/3로 높아진다. 아담이 생겨났다는 정보는 뒷면이 나왔음을 곧 여러 사람들이 생겨났음을 입증하는 증거다. 또한 이브가 생겨났다는 정보는 여러 사람들이 생겨났음을 입증하는 증거다. 하지만 내가 생겨났다는 정보는 여러 사람들이 생겨났음을 입증하는 증거가 되지 못한다.¹⁹⁾

의식을 새로 얻은 이가 자신이 아담이라는 것을 알게 된 다음에도 C(앞)은 그대로 1/2임을 셈할 수 있다. “생겨난 나는 아담이다” 곧 “나_아”를 안 다음에 동전이 앞면이 나왔으리라 믿는 정도는 C(앞나_아)다. 새로 생겨난 이가 자신이 아담이라고 믿는 정도는 자신이 이브라고 믿는 정도와 똑같은 것이다. 그 정도는 동전이 앞면이

난 이에게 “아담 또는 이브가 생겨났다”는 것을 알려준다. 새로 생겨난 이가 ‘아담’이라는 이름과 ‘이브’라는 이름 가운데 하나를 마구잡이로 뽑는다. 그가 ‘아담’을 뽑았다고 하자. 새로 생겨난 이가 다음과 같이 묻는다. “아담이 생겨났는가?” 이 물음에 대해 “그렇다”라는 대답이 주어졌다면, 새로 생겨난 이는 정보 “아담이 생겨났다”를 마구잡이 절차를 거쳐 얻은 셈이다.

19) 우리 실험 실험실에서 생겨난 이에게 정보 “나는 생겨났다”는 “누군가 생겨났다”나 “적어도 한 사람이 생겨났다”와 같은 정보인 셈이다. 이것을 우주론에 적용하면 우주에 생겨난 이에게 “나는 생겨났다”는 우주의 생성에 관한 한 “의식을 가진 이가 생겨났다”와 비슷한 역할을 한다. 이 글에서는 이 주제를 다루지 않겠다.

나왔음을 안 다음에도 뒷면이 나왔음을 안 다음에도 똑같다. 따라서 우리는 다음을 가정한다.

$$(E2) C(\text{나}_{\text{아}}|\text{앞}) = C(\text{나}_{\text{이}}|\text{앞}),$$

$$(E3) C(\text{나}_{\text{아}}|\text{뒤}) = C(\text{나}_{\text{이}}|\text{뒤}).$$

우리 논증에 오류가 있다면 (E1)이나 (E2)와 (E3)이 미덥지 못하기 때문이다. 하지만 우리 생각에 (E2)와 (E3)은 거의 명백하다. 가정 (E2)와 (E3)에서 $C(\text{나}_{\text{아}}|\text{뒤}) = C(\text{나}_{\text{이}}|\text{뒤}) = 1/2$ 을 얻고²⁰⁾ 이로부터

$$\begin{aligned} (3) C(\text{나}_{\text{아}}) &= C(\text{나}_{\text{아}}|\text{앞})C(\text{앞}) + C(\text{나}_{\text{아}}|\text{뒤})C(\text{뒤}) \\ &= (1/2)\{C(\text{앞}) + C(\text{뒤})\} = 1/2 \end{aligned}$$

을 얻는다. 마찬가지로 $C(\text{나}_{\text{이}})$ 도 1/2다. (E2), (E3), (3)을 써서

$$\begin{aligned} (4) C(\text{앞}|\text{나}_{\text{아}}) &= C(\text{앞}\&\text{나}_{\text{아}})/C(\text{나}_{\text{아}}) = C(\text{나}_{\text{아}}|\text{앞})C(\text{앞})/C(\text{나}_{\text{아}}) \\ &= (1/2)(1/2)/(1/2) = 1/2 \end{aligned}$$

임을 알 수 있다.²¹⁾ 마찬가지로 셈으로 $C(\text{앞}|\text{나}_{\text{이}})$ 도 1/2이다. 이처럼

20) “ $\text{나}_{\text{아}} \vee \text{나}_{\text{이}}$ ”가 참이라는 것이 이미 알려져 있고 “ $\text{나}_{\text{아}}$ ”와 “ $\text{나}_{\text{이}}$ ” 가운데 오직 하나만 참이기 때문에 $1 = C(\text{나}_{\text{아}} \vee \text{나}_{\text{이}}|\text{앞}) = C(\text{나}_{\text{아}}|\text{앞}) + C(\text{나}_{\text{이}}|\text{앞})$ 이다. 이것과 (E2)로부터 $C(\text{나}_{\text{아}}|\text{앞}) = C(\text{나}_{\text{이}}|\text{앞}) = 1/2$ 이라는 것을 알 수 있다. 마찬가지로 $1 = C(\text{나}_{\text{아}} \vee \text{나}_{\text{이}}|\text{뒤}) = C(\text{나}_{\text{아}}|\text{뒤}) + C(\text{나}_{\text{이}}|\text{뒤})$ 기 때문에 (E3)으로부터 $C(\text{나}_{\text{아}}|\text{뒤}) = C(\text{나}_{\text{이}}|\text{뒤}) = 1/2$ 이라는 것을 알 수 있다. 새로 생기는 이들이 N명이고 앞면이 나올 가능성이 h인 일반화된 실험 설정에서 $C(\text{나}_{\text{x}}|\text{앞}) = C(\text{나}_{\text{x}}|\text{뒤}) = 1/N$ 이다. 여기서 ‘x’는 새로 만들어지는 N명의 사람들 가운데 아무 이름이다. 새로 생기는 이는 자기 이름과 관련된 차별화된 정보를 전혀 갖지 않는다.

21) 보스트롬의 원래 인큐베이터 실험에는 비대칭 요소가 있다. 한 사람만 생길 때 검은 수염이 난 사람이 생긴다. 두 사람이 생길 때 검은 수염이 난 사람

새로 생긴 이가 자기 이름을 안다 해도 그에게 C(앞)은 1/2에 머물러 있다.

4. 하필 왜 내가 생겨났는가?

생겨나자마자 자신이 아담인지 이브인지 모르지만 나중에 자기 이름을 알려주는 우리의 실험 설정에서 C(앞)은 1/2이다. 하지만 다시 깨어나는 실험 설정에서 깨어난 이들은 깨어나자마자 자기가 아담인지 이브인지 아는데 이 경우 C(앞)은 1/3이다.²²⁾ 이 다름은 두

과 흰 수염이 난 사람이 각각 생긴다. Bostrom (2002), p. 64. 이것은 사실상 앞면이 나오면 아담이 생기고 뒷면이 나오면 아담과 이브가 생기는 실험 설정과 같다. 보스트롬은 C(앞)과 C(뒤)가 1/2이라고 논증 없이 주장하는데 이를 논증하는 것은 쉽지 않다. C(아)가 1이라는 것은 명백하다. 외부자의 관점에서 섯한 값 $C_{외}(이)$ 가 1/2이라고 해서 자신이 아담인지 이브인지 모르는 실험 참여자의 관점에서 섯한 값 C(이)도 1/2이라 말하는 것은 예단이다. 사실 C(이)가 1/2이라고 말하는 것은 곧 C(뒤)가 1/2이라고 말하는 것이다. 여하튼 보스트롬이 주장했듯이 C(앞)과 C(뒤)가 1/2이라고 일단 가정하기로 하자. C(나_앞앞)은 1이고 C(나_앞뒤)는 1/2이기 때문에 C(나_이)는 $C(나_{앞}앞)C(앞) + C(나_{앞}뒤)C(뒤)$ 이고 이는 3/4이다. 따라서 C(앞_{나_이)는 $C(나_{앞}앞)C(앞)/C(나_{이})$ 이며 이는 2/3다. 보스트롬의 인큐베이터 실험에서 “나_이”가 과거 일을 추정하게 하는 증거 역할을 하는 까닭은 ‘아담’이라는 존재와 ‘이브’라는 존재가 비대칭 역할을 하기 때문이다.}

아무튼 보스트롬은 이와 같은 믿음직함 섯을 받아들이지만 이 방식에 미덥지 못한 측면이 있다면서 ‘사탄의 조언’ 생각실험을 제안한다. Bostrom (2002), p. 142. 하지만 이 생각실험은 그의 애초 인큐베이터 생각실험과 매우 다르다. 특별히 동전을 던지는 시점이 믿음직함을 섯하는 시점보다 늦어질 가능성이 있는 상황에서 믿음직함 섯은 매우 조심해야 한다. 김명석 (2016b), p. 8. 사탄의 조언 생각실험의 올바른 풀이는 여기서 다루지 않겠다.

²²⁾ 김명석 (2017), p. 5. 일반화된 실험 설정에서 C(앞)은 $h/(h+tN)$ 이다. h는 앞면이 나올 애초 가능성이고 t는 뒷면이 나올 애초 가능성이다. N은 뒷면이 나왔을 때 다시 깨어날 사람의 수다. N이 매우 커지면 C(앞)은 매우 작아진다. 반면 생겨난 실험 설정에서 C(앞)은 h다. 뒷면이 나왔을 때 새로 생겨난 사람이 아무리 많다 하더라도 C(앞)은 바뀌지 않는다.

실험 설정에서 자신의 이름이 알려지는 절차의 다름에서 비롯된다. 나아가 이 다름은 새로 생겨난 이들과 다시 깨어난 이들의 다름이기도 하다. 만일 누군가 깨어났다는 사실이 이미 알려져 있다 하더라도, 다시 깨어난 이는 자신이 지금 깨어났다는 사실에 놀랄 까닭이 있다. 자신이 잠들 때 자기가 다시 깨어나지 않을 가능성이 있다는 것을 알고 있었기 때문에 지금 자신이 다시 깨어났다는 사실은 놀랄 일이다. 하지만 만일 누군가 생겨났다는 사실이 이미 알려져 있다면, 새로 처음 생겨난 이는 하필 자신이 지금 생겨났다는 사실에 놀랄 까닭이 없다. 그가 생기기 전에는 자기가 생기지 않을 가능성이 있다는 것을 그가 알 수 없기 때문에 지금 자신이 생겨났다는 사실은 그에게 놀랄 일이 아니다.

이 점은 극단의 경우를 살펴봄으로써 확인할 수 있다. 실험에 참여하는 이들이 100억 명이라고 생각해 보라. 동전이 앞면이 나왔다면 100억 명 가운데 한 명만 다시 깨어난다. 만일 내가 다시 깨어났다면 내가 지금 깨어났다는 사실은 매우 놀랄 일이다. 나는 동전이 앞면이 나와 100억 명 가운데 나 하나가 깨어났다고 믿기보다는 동전이 뒷면이 나와서 100억 명 모두 깨어났다고 믿어야 할 것이다. 이처럼 다시 깨어난 이들의 문제에서 자신이 지금 깨어났다는 정보는 동전이 뒷면이 나왔음을 곧 다른 많은 이들도 지금 깨어나 있음을 강력하게 뒷받침한다. 다시 깨어난 이들은 잠들기 전의 자신과 지금의 자신 사이의 연속성을 파악한다. 이뿐만 아니라 다시 깨어난 그는 100억 명 가운데 구별되는 한 개별자로서 자신을 의식한다.

이제 비로소 생겨난 이의 인식 상황 또는 정보 상황은 다시 깨어난 이의 정보 상황과 매우 다르다. 다시 깨어난 ‘나’는 기억을 더듬어 과거로 거슬러 갈 수 있지만 새로 생겨난 ‘나’는 그렇게 할 수 없다. 처음 생겨난 이 또는 이제 태어난 이는 동전을 던질 때

자신이 없었다는 것을 안다. 동전이 뒷면이 나오면 100억 명이 새로 생기지만 앞면이 나오면 오직 한 명만 생겨난다고 생각해보라. 만일 내가 그렇게 생겨난 이라면 하필이면 왜 내가 생겨났는지 묻는 것은 올바른 물음이 아니다. 왜냐하면 생겨난 다음에야 비로소 내가 나를 나로서 의식했을 뿐이기 때문이다. 내 이름이 ‘아담’이라는 것을 듣게 된다 해도 왜 하필이면 내가 아담인지 묻는 것도 올바른 물음이 아니다. 지금 누군가 생겨났고 그렇게 생겨난 이에게 ‘아담’이라는 이름이 붙여졌을 뿐이기 때문에 “나 아담이 지금 생겨났다”는 정보는 다른 이가 더 존재할 것이라는 증거로 쓰일 수 없다.

내가 생겨났고 그 다음 내 이름을 물어 내 이름이 ‘아담’임을 알게 됨으로써 얻게 된 정보 “아담이 생겨났다”는 “내가 생겨났다”는 정보 그 이상이 될 수 없다. 생겨난 사람이 자기 이름을 물어보면 그때 비로소 이름을 붙여주어도 되며 나 스스로 아무 이름을 붙여주어도 된다. 의식을 갖자마자 내가 나에게 ‘오돔’이라는 이름을 붙여줄 수 있다. 그렇게 붙여진 이름을 두고 “왜 하필이면 내 이름이 오돔인가?”라고 묻는 것은 의미 없는 물음이다. 나아가 “하필이면 왜 오돔이 생겨났는가?”라는 물음은 “생겨난 내가 왜 생겨났는가?”라는 물음과 다를 바 없다. 생겨난 내 이름이 ‘아담’임을 알게 됨으로써 얻게 된 정보 “아담이 생겨났다”는 내 이름이 무엇인지 알지 못한 채 “아담이 생겨났는가?”라는 내 물음의 답변으로 얻게 된 정보 “아담이 생겨났다”와 매우 다르다.

믿음직함을 셈하는 주체가 갖는 자기의식 정보 “나는 지금 여기 의식을 갖고 있다”는 자신이 지금 깨어 있지 않을 수도 있는데 지금 놀랍게도 깨어 있다는 것을 알려줄 때 의미 있는 정보가 된다. 처음 생겨난 이들은 자신이 생겨난 다음에야 자신이 생겨나지 않았을 반사실 상황을 사후에 상정할 수 있다. 이 때문에 그들은 자신

이 생겨났다는 사실에 놀라지 않아도 된다. 왜 하필이면 내가 생겨났는가 하고 놀라려면 그러한 물음을 갖기 전에 내가 생겨나지 않을 가능성을 이미 갖고 있어야 한다. 우리가 명제 X에 놀란다는 것은 우리가 이전에 X가 거짓일 가능성을 갖고 있다가 지금 그 가능성이 사라졌다는 점에 놀란다는 것을 뜻한다. 우리는 “나는 생겨났다”가 거짓일 가능성을 우리가 생기기 전에 가졌던 적이 없다.

5. 생겨난 이들이 차츰 알게 된 정보의 가치

생겨난 이들이 생겨난 다음에 자기 자신과 주변에 대해 알게 되는 정보들 대부분도 앞면 또는 뒷면이 나왔음을 입증하지 못한다. 앞에서 말했듯이 생겨난 이가 자기 이름을 새로 알게 된다 해도 C(앞)은 그대로 1/2이었다. 자신의 성별을 조사해 자신이 남자였다는 것을 알게 되었다 해도 이 사실은 과거에 일어난 일을 추정하는데 도움이 되지 못한다. 새로 생겨난 자기 자신을 조사해 알게 된 정보 “남자가 생겨났다”는 내가 남자인지 여자인지 모른 채 알게 된 정보 “남자가 생겨났다”와 다르다.²³⁾ 믿음직함을 직접 셈함으로써 이를 보여줄 수 있다. 아담은 남자고 이브는 여자라고 가정한다.

자신의 성별을 모르는 상황에서 C(남자가 생겨났다)는 3/4이기 때문에

C(앞남자가 생겨났다)

23) 자신이 남자인지 여자인지 모른 채 “남자가 생겨났는가?”라는 자기 물음의 대답으로서 정보 “남자가 생겨났다”를 얻게 되었다면, 이 정보는 동전이 뒷면이 나왔음을 입증한다. 왜냐하면 “남자가 생겨났다”가 거짓일 가능성을 갖고 있었는데 지금 그 가능성이 사라졌기 때문이다. 그 가능성이 사라진 것을 설명하는 하나의 방식은 남자와 여자가 모두 생겨났다는 가설을 받아들이는 것이다.

$$\begin{aligned}
 &= C(\text{남자가 생겨났다})C(\text{앞})/C(\text{남자가 생겨났다}) \\
 &= (1/2)(1/2)/(3/4) = 1/3
 \end{aligned}$$

이다. 생겨난 이가 정보 “남자가 생겨났다”를 얻는다면 C(앞)은 1/2에서 1/3로 바뀌고 C(뒤)는 1/2에서 2/3로 바뀐다. 정보 “남자가 생겨났다”는 뒷면이 나왔음을 곧 여자도 생겨났음을 입증하는 증거로 쓰일 수 있다.

하지만 정보 “나는 남자다”는 정보 “남자가 생겨났다”와 다르다. 자신의 성별을 모르는 상황에서 C(나는 남자다)는 1/2이기 때문에

$$\begin{aligned}
 C(\text{앞나는 남자다}) &= C(\text{나는 남자대앞})C(\text{앞})/C(\text{나는 남자다}) \\
 &= (1/2)(1/2)/(1/2) = 1/2
 \end{aligned}$$

이다. 이처럼 생겨난 이가 정보 “나는 남자다”를 얻게 된다 해도 C(앞)과 C(뒤)는 바뀌지 않고 그대로 1/2로 남아있다. 정보 “나는 남자다”는 뒷면이 나왔음을 곧 여자도 생겨났음을 입증하지 못한다. 이것은 정보 “나는 여자다”도 마찬가지다. 자신의 성별을 탐구함으로써 알게 된 성별 정보는 자신과 함께 몇 명이 새로 생겨났는지를 추정할 만한 실마리가 되지 못한다.

글을 끝맺기 전에 생겨난 이들의 상황을 보다 현실에 맞게 일반화해 보자. 유전공학자 지오디는 매우 많은 인간 배아들 곧 M개의 배아들을 갖고 있다. 동전을 던져 앞면이 나오면 M개 가운데 한 배아만을 사람으로 자라게 한다. 뒷면이 나오면 M개 가운데 두 배아를 사람으로 자라게 한다. 이렇게 일반화하는 까닭은 생겨난 사람을 아담이나 이브로 미리 정해 놓는 것이 현실에 맞지 않기 때문이다. 나아가 한 사람만 생겨날 때 생겨날 그 사람이 두 사람이 생겨날 때 또 생겨나야 할 까닭이 없다. 지오디는 각 배아의 유전

자에 따라 고유번호를 매겨놓았다. “유전자 m을 가진 배아가 사람으로 자라난다”를 짧게 G_m 이라 쓰겠다. 우리가 $C(G_i) = C_{\text{외}}(G_i)$ 와 $C(G_i \vee G_j) = C_{\text{외}}(G_i \vee G_j)$ 를 가정한다면²⁴⁾ $C(\text{뒤})$ 를 셈할 수 있다. “뒤”는 “ $G_1 \& G_2 \vee G_1 \& G_3 \vee \dots \vee G_M \& G_{M-1}$ ”과 동치고, $G_i \& G_j$ 와 같은 짝은 모두 $M(M-1)/2$ 개 있으며, 이 짝들은 서로 배타 관계에 있기 때문에

$$\begin{aligned} C(\text{뒤}) &= C(G_1 \& G_2 \vee G_1 \& G_3 \vee \dots \vee G_M \& G_{M-1}) \\ &= C(G_1 \& G_2) + C(G_1 \& G_3) + \dots + C(G_M \& G_{M-1}) = 1/2 \end{aligned}$$

이다. 따라서 $C(\text{앞})$ 또한 1/2이다.

이 실험 과정에서 새로 생겨난 내가 우연히 G_i 가 참이라는 것을 알게 되었다면, $C(G_i|\text{앞}) = 1/M$ 이고 $C(G_i) = 3/2M$ 이기 때문에

$$\begin{aligned} C(\text{앞}G_i) &= C(G_i|\text{앞})C(\text{앞})/C(G_i) \\ &= (1/M)(1/2)(2M/3) = 1/3 \end{aligned}$$

이다. 이처럼 특정 유전자를 가진 사람이 생겨났다는 정보는 다른 사람도 생겨났음을 입증한다. 그 특정 유전자를 가진 사람이 ‘레이’라고 불린다면 결국 정보 “레이가 생겨났다”는 다른 누군가도 생겨났음을 알려주는 실마리로 쓰일 수 있다. 하지만 생겨난 이가 자신

24) $C_{\text{외}}(G_j|\text{앞})$ 이 $1/M$ 이라는 사실, $C_{\text{외}}(G_i|\text{뒤})$ 가 $2/M$ 라는 사실, $C_{\text{외}}(G_i|G_j \& \text{뒤})$ 가 $1/(M-1)$ 이라는 사실을 쓴다면, 우리는 쉽게 다음을 셈할 수 있다. (i) $C_{\text{외}}(G_i) = C_{\text{외}}(G_i|\text{앞})C_{\text{외}}(\text{앞}) + C_{\text{외}}(G_i|\text{뒤})C_{\text{외}}(\text{뒤}) = 1/2M + 1/M = 3/2M$. (ii) $C_{\text{외}}(G_i \& G_j) = C_{\text{외}}(G_i \& G_j|\text{앞})C_{\text{외}}(\text{앞}) + C_{\text{외}}(G_i \& G_j|\text{뒤})C_{\text{외}}(\text{뒤}) = C_{\text{외}}(G_i \& G_j|\text{뒤})C_{\text{외}}(\text{뒤}) = C_{\text{외}}(G_i \& G_j \& \text{뒤}) = C_{\text{외}}(G_i|G_j \& \text{뒤})C_{\text{외}}(G_j \& \text{뒤}) = C_{\text{외}}(G_i|G_j \& \text{뒤})C_{\text{외}}(G_j|\text{뒤})C_{\text{외}}(\text{뒤}) = (1/M-1)(2/M)(1/2) = 1/M(M-1)$. (3) $C_{\text{외}}(G_i \vee G_j) = C_{\text{외}}(G_i) + C_{\text{외}}(G_j) - C_{\text{외}}(G_i \& G_j) = 3/M - 1/M(M-1)$. 따라서 $C(G_i)$ 는 $3/2M$ 이고, $C(G_i \vee G_j)$ 는 $3/M - 1/M(M-1)$ 이다. 이로부터 $C(G_i \& G_j)$ 는 $1/M(M-1)$ 임을 알 수 있다.

의 유전자를 조사해 얻은 정보 “내 유전자는 유전자 j 였다”는 그렇지 못하다. 명제 “내 유전자는 유전자 j 였다”를 짧게 “ $나_j$ ”라고 쓴다면, $C(나_j|앞)$ 은 $1/M$ 이고 $C(나_j|뒤)$ 도 $1/M$ 이기 때문에 $C(앞|나_j)$ 는 $1/2$ 일 수밖에 없다. 이처럼 내 유전 정보를 파악함으로써 알게 된 “ $나_j$ ”는 다른 누군가도 생겨났음을 알려주는 실마리로 쓰일 수 없다.²⁵⁾

나아가 새로 생겨난 이가 속성 K 를 가질 가능성이 k 라고 해보자. 내가 K 를 갖는지 모른 채 얻은 정보 “ K 인 이가 생겨났다”는 다른 이도 생겨났음을 추정케 하는 단서로 쓰일 수 있지만 정보 “나는 K 다”는 그렇지 못하다. 이를 살펴보려고 명제 “ K 인 이가 생겨났다”를 짧게 “ $K_{생}$ ”으로, 명제 “나는 K 다”를 짧게 “ $나_k$ ”로 쓰겠다. 우리가 알아내야 하는 것은 $C(앞|K_{생})$ 인데 이를 셈하려면 먼저 $C(K_{생})$ 을 셈해야 한다.²⁶⁾

$$\begin{aligned} C(K_{생}) &= C(K_{생}|앞)C(앞) + C(K_{생}|뒤)C(뒤) \\ &= k/2 + k - k^2/2 = (3k - k^2)/2. \end{aligned}$$

25) 한 심사자는 “유전자 j 를 가진 사람이 생겨났다”는 증거는 가설 “나 말고 다른 사람도 생겨났다”를 입증하지만 “생겨난 내 유전자는 유전자 j 다”는 그 가설을 입증하지 못하는 현상이 매우 흥미롭다고 논평했다. 이것은 마치 정보 A 는 가설 H 를 입증하지만 정보 $A \& B$ 는 가설 H 를 입증하지 못하는 현상이기 때문이다. 이는 정보 B 가 가설 H 를 반입증하는 경우다. 정보 “유전자 j 를 가진 사람이 생겨났다”는 “내 유전자는 유전자 j 다”와 “내 유전자는 유전자 j 가 아니다”를 모두 허용한다. 주어진 정보 상황에서 “내 유전자는 유전자 j 가 아니다”는 “나 말고 다른 사람도 생겨났다”를 100% 입증한다. “내 유전자는 유전자 j 다”가 추가로 주어진다면 “내 유전자는 유전자 j 가 아니다”가 완전히 배제된다. 이로써 “나 말고 다른 사람도 생겨났다”가 참일 가능성이 그만큼 낮아진다. 이 현상이 흥미로운 교훈을 품고 있다면 이를 드러낸 사람은 이 심사자다. 이 밖에 여러 가지를 지적해주신 이 심사자에게 감사드린다.

26) $C(K_{생}|앞)$ 이 k 라고 가정하는 데 무리가 없다. $C(K_{생}|뒤)$ 는 $1 - C(둘 다 K가 아니다|뒤)$ 다. $C(둘 다 K가 아니다|뒤)$ 는 $(1-k)(1-k)$ 라고 가정할 수 있다. 따라서 $C(K_{생}|뒤)$ 는 $2k - k^2$ 이다.

이를 써서 $C(\text{앞}|K_{\text{생}})$ 을 셈하면 결국 우리는

$$\begin{aligned} C(\text{앞}|K_{\text{생}}) &= C(K_{\text{생}}|\text{앞})C(\text{앞})/C(K_{\text{생}}) \\ &= (k)(1/2)\{2/(3k - k^2)\} = 1/(3-k) \end{aligned}$$

이라는 것을 알 수 있다. k 가 1에 가까이 가면 “ $K_{\text{생}}$ ”은 아무 실마리가 되지 못하고 k 가 0에 가까이 가면 “ $K_{\text{생}}$ ”은 상당한 실마리가 된다. 생겨난 사람의 키가 정확히 170.12345cm가 될 가능성은 거의 0에 가깝고 “키가 170.12345cm인 사람이 생겨났다”라는 정보는 다른 이도 생겨났음을 추정케 하는 증거가 된다. 하지만 생겨난 사람이 갖는 자기 신체 정보 “내 키는 정확히 170.12345cm다”는 그렇지 않다. 왜냐하면 $C(\text{나}_k|\text{앞})$ 이 k 고 $C(\text{나}_k|\text{뒤})$ 도 k 라서²⁷⁾ $C(\text{앞}|나_k)$ 를 셈하면 $C(\text{앞})$ 과 같게 나오기 때문이다.

6. 나오는 말

우리는 의식을 처음 갖게 된 이가 갖는 자기의식 정보의 본성을 탐구하고자 했다. 이 정보는 의식을 잃었다가 의식을 되찾은 이가 갖는 자기의식 정보와 다르다. 매우 복잡한 상황에서 자기의식 정보를 더욱 깊이 탐구해야 하겠지만 우리는 씨앗 연구로서 가장 단순한 상황에서 자기의식 정보를 탐구했다. 우리 생각실험의 상황에서 실험 참여자들은 자기의식 정보만으로 자기말고 다른 의식이 생겨났음을 추정할 수 없다는 결론을 얻었다.

두 배아 실험 설정에서 자라난 사람이 자기 유전 정보를 알게 되었다 해도 그 유전 정보는 몇 개의 배아들이 사람으로 자라났는

27) 생겨난 내가 속성 K 를 가질 가능성은 함께 생겨난 이가 있든 없든 달라지지 않는다.

지를 넘지시 알려주지 못한다. 이 실험 설정에서 이십 년 동안 온갖 경험을 한 인식 주체가 가진 자기의식 정보든, 자기 신체 정보든, 유전 정보든 그것은 자신이 배양되기 전에 일어난 일을 말해주지 못한다. 또한 자기말고 다른 이가 같은 배양 과정을 거쳐 자라나 지금 어딘가에 살고 있다는 가설을 입증하지 못한다. 공동 가족 계획을 짰 두 부부의 실험 설정에서도 똑같은 것을 말할 수 있다. 스무 살 때 출생의 비밀을 듣게 된 인식 주체가 갖는 자기의식 정보는 자신이 태어나기 전의 가족계획 내용을 알려주는 실마리가 되지 못한다. 또한 자기 출생의 비밀은 자기말고 다른 이가 같은 계획 아래서 태어나 지금 어딘가에 살고 있다는 가설을 입증하지 못한다.

이 결과들을 우주론에 적용할 수 있다. 우리에게 두 가지 우주론이 있다고 생각해보자. 한 우주론은 오직 한 세계만 실현되며 그 세계를 관찰하는 의식들이 그 세계 안에 반드시 생겨난다. 다른 우주론은 여러 세계들이 실현되며 그 세계들 가운데 적어도 하나는 그 안에 자기 세계를 관찰하는 의식들이 반드시 생겨난다. 우리는 세계를 관찰하는 의식으로서 지금 여기 생겨났다. 이것은 우주론의 ‘코기토’다. 이 정보에 비추어 볼 때 우리는 두 우주론들 가운데 무엇을 더 믿어야 하는가? 이 글의 우리 탐구가 옳다면 이 정보는 두 우주론 가운데 어느 하나를 뒷받침하는 데 도움을 줄 수 없다. 우리 자신을 탐구하여 알게 된 우리 유전자의 특성도 도움을 주지 못하며 바깥 세계를 탐구하여 알게 된 우주 배경복사의 특성도 도움을 주지 못한다. 그런 정보는 “의식을 가진 이가 생겨났다”와 비슷한 정보일 뿐이다. 하지만 우리가 받아들일 만한 우주론들은 매우 많으며 우리의 코기토 정보는 그 가운데 특정 우주론을 선호할 가능성이 여전히 있다.

참고문헌

- 김명석 (2016a), “믿음직함과 가능 세계”, 『범한철학』, 제80집, pp. 239-259.
- 김명석 (2016b), “잠자는 미인에게 ‘아직’과 ‘이미’”, 『과학철학』, 제19권 제3호, pp. 1-20.
- 김명석 (2017), “자기의식 정보와 관찰선택 효과”, 『논리연구』, 제20집 제1호, pp. 1-19.
- 김명석 (2018), “자기 표본추출 가정과 자기 표지 가정”, 『동서철학 연구』 제89집, pp. 353-368.
- 존 레슬리 (1998), 『세계의 종말』, 이충호 옮김, 사람과사람.
- Bostrom, N. (2002), *Anthropic Bias*, Routledge.
- Bostrom, N. (2003a), “Are You Living in a Computer Simulation?”, *Philosophical Quarterly* 53, pp. 243-255.
- Bostrom, N. (2003b), “The Mysteries of Self-Locating Belief and Anthropic Reasoning”, *The Harvard Review of Philosophy* 11, pp. 59-73.
- Carter, B. (2011), “Large Number Coincidences and the Anthropic Principle in Cosmology”, Republication in *General Relativity and Gravitation* 43, pp. 3225-3233.
- Gott, R. (1993), “Implications of the Copernican Principle for our Future Prospects”, *Nature* 363, pp. 315-319.
- Leslie, J. (1988), “Risking the World’s End”, *Bulletin of the Canadian Nuclear Society*, 1989 May, pp. 10-15.

국민대학교 교양대학

Kookmin University, College of General Education

myeongseok@gmail.com

Self-Consciousness Information of the one who just came up

Myeongseok Kim

Gott, Carter, Leslie, Bostrom and so on, as Descartes did, have made the evidence that is “I am here as an observer” to support many other beliefs. Bostrom and others who studied observation selection effects are missing two points. First, the self-consciousness information of the ones who just came up is distinct from that of the ones who have awoken. The awoken ‘I’ can trace back by memory to the past, but the ‘I’ who just came up can not. Second, when calculating credence, we must distinguish the ones in the possible worlds from the ones in the actual worlds. An estimate of credence where only one possible world is actualized among all possible worlds, differs fundamentally from that where all possible worlds are actualized. Keeping these two points in mind, we have explored what is the nature of the self-consciousness information of the one who just came up. We examine in depth the two human embryos thought experiment.

Key Words: Self-Consciousness Information, Probability, Credence, Observational Selection Effect, Bostrom