논리연구 20-1(2017) pp. 1-19

자기의식 정보와 관찰 선택 효과*

김 명 석

【국문요약】 자기 자신이 지금 깨어나 있다는 사실은 때때로 자기 믿음의 크기를 바꿀 만한 정보로 사용될 수 있다. 현대 우주론에서는 인간 의식의 존재를 가능한 여러 가지 우주들이 현실화되어 있다는 가설을 뒷받침하는 증거로 사용할 수 있는지 논란거리이다. 우리는 이 글에서 자기 자신이 깨어나 있다는 자기의식 정보가 어떤 본성을 갖고 있는지 탐구하고자 한다. 가나와 다라가 참여하는 다음과 같은 실험을 생각해 보자. 일요일에 멀쩡한 동전을 던져 앞면이 나오면 가나와 다라 가운데 마구잡이로 한 명을 뽑아 월요일에 깨우고, 뒷면이 나오면 가나와 다라를 둘 다 월요일에 깨운다. 가나와 다라 가운데 적어도 한 명은 월요일에 깨어날 텐데, 이들은 동전이 앞면이 나왔으리라고 얼마큼 믿어야 할까? 우리는 이 문제의 올바른 답이 1/3이라고 논증한다. 이와 함께 잠에서 깨어난 이가 얻게 되는 “내가 지금 깨어났다”라는 정보는 마구잡이 절차를 거쳐 그 자신에게 알려졌음을 주장한다.

【주요어】 관찰 선택 효과, 마구잡이 절차, 믿음직함, 치우친 절차, 확률

투고일: 2017. 1. 16 심사 및 수정 완료일: 2017. 2. 9 게재확정일: 2017. 2. 10
* 이 글의 논증을 토론해준 김한승 선생과 박일호 선생, 수학동아의 고은영 기자, 논문을 개선하도록 도움을 주신 이 논문의 세 심사자들께 감사드린다.

1. 들어가는 말

현대 과학철학자들은 다양한 인식 상황에서 인식 주체가 자기 믿음의 정도를 베이즈 공리를 따라 바꾸는 것이 합당한지를 검토하고 있다. 특별히 자기가 있는 공간 위치나 시간 위치 정보를 잃었을 때 자기 믿음 정도를 어떻게 바꾸어야 하는지에 큰 관심을 갖는다. 또한 자기 동일성 또는 정체성이 흐려졌을 때도 자기 믿음 정도가 달라질 수 있다. 이를 탐구하기 위해 종말 논증, 미세 조율 논증, 두 딸 문제, 잠자는 미인 등 여러 가지 생각 실험들을 제안한다.[1] 이 글에서 우리는 새로운 생각 실험을 선보이고, 이로부터 인식 상황 변화에 따른 믿음 정도의 변화를 합당하게 가늠하는 인식론 및 과학철학 탐구에 이바지하고자 한다.

새로운 생각 실험은 엘가가 처음 발표한 “잠자는 미인 문제”와 닮아 있다.[2] 그의 실험에 참여하는 사람은 오직 한 사람이다. 우리의 새로운 실험에는 여러 사람들이 참여한다. 가나와 다라는 우리 실험에 스스로 참여한다. 이들은 잠꾸러기이고 누군가 깨워주어야 일어날 수 있다. 우리는 가나와 다라에게 어떻게 실험이 진행되는지 자세하게 알려준다. 먼저 가나와 다라를 일요일 밤에 각기 다른 방에 재운다. 이들은 실험이 끝날 때까지 서로 대화할 수 없으며, 상대방이 어떤 상태에 있는지 전혀 알지 못한다. 이들이 잠든 뒤 멀쩡한 동전을 던진다. 일요일에 던진 이 동전이 앞면이 나올 경우, 월요일 아침에 가나와 다라 가운데 아무나 한 사람을 마구잡이로 뽑아 깨운다. 그 동전이 뒷면이 나올 경우, 월요일 아침에 가나와 다라를 둘 다 깨운다. 깨어난 이에게는 이렇게 묻는다. “일요일에 던진 그 동전이 앞면이 나왔으리라고 얼마큼 믿어야 하는가?”

---

1) Bostrom (2002).

2) 이 문제의 국내 논쟁에 대해서는 송하석 (2012)를 보라.

우리는 제2절에서 가나든 다라든, 이 물음에 1/3이라고 답해야 한다는 손쉬운 셈을 먼저 제안한다. 제3절과 제4절에서 실험에 참여하지 않는 외부자의 관점에서 이 물음의 올바른 답이 1/3이 되는 경우와 1/2이 되는 경우가 있음을 보인다. 외부자가 마구잡이 절차를 거쳐 정보를 얻게 되면 그의 답은 1/3이지만, 치우친 절차를 거쳐 정보를 얻게 되면 그의 답은 1/2이다. 제5절에서 실험에 참여하는 내부자가 얻게 되는 자기의식 정보가 마구잡이 절차를 거쳐 주어진 것인지 치우친 절차를 거쳐 주어진 것인지 따져 보겠다. 우리 실험의 경우 자기의식 정보는 마구잡이 절차를 거쳐 주어진 것이라고 결론내리고 그가 답해야 하는 올바른 답이 1/3이라고 주장한다. 마침내 제6절에서 우리 생각 실험의 교훈을 일부 드러내고 글을 마무리한다.

2. 깨어난 사람들 문제

우리가 선보인 실험을 "잠자는 미인들 문제" 또는 "깨어난 사람들 문제"라고 부르겠다. 엘가의 잠자는 미인 문제에는 오직 한 사람의 미인이 나오지만 우리 문제에는 두 미인들이 나온다. 엘가의 문제에서는 동전을 던져 한 미인을 한 번 깨울지 두 번 깨울지를 결정한다. 우리의 문제에서는 동전을 던져 오직 한 미인만을 깨울지 두 미인을 모두 깨울지를 결정한다.[3)]

문장들의 믿음직함을 셈하기 위해 각 문장들을 다음과 같이 약칭하겠다.

---

3) 엘가의 문제를 일반화하여, 앞면이 나올 가능성이 h인 동전을 던져, 한 미인을 한 번 깨울지, 그를 N번 깨울지를 결정할 수 있듯이, 우리의 문제도 일반화하여, 앞면이 나올 가능성이 h인 동전을 던져, 한 미인을 깨울지, 한꺼번에 N명의 미인들을 모두 깨울지 결정할 수 있다.

앞 := 일요일에 던진 그 동전이 앞면이 나왔다.
뒤 := 일요일에 던진 그 동전이 뒷면이 나왔다.
가 := 가나가 월요일에 깨어났다.
다 := 다라가 월요일에 깨어났다.

일요일 저녁에 가나와 다라 및 우리가 놓인 인식 상황을 S라고 하고, 월요일 아침에 가나와 다라 가운데 누군가 깨어났을 때 이들이 놓일 인식 상황을 M이라고 하자. 우리는 인식 상황 M에서 가나 또는 다라의 믿음직함 C_M(앞)을 셈하고자 한다. 여기서 믿음직함 함수 $C_M(X)$는 인식 상황 M에 놓인 인식 주체가 X를 믿는 정도이다. 인식 상황 S에 놓인 인식 주체가 X를 믿는 정도는 $C_S(X)$라고 쓰겠다.

한 인식 주체가 인식 상황 o에 있을 때 문장들 X, Y, X&Y 등을 믿는 정도가 각각 $C_o(X)$, $C_o(Y)$, $C_o(X\&Y)$였다고 하자. 이 인식 주체가 다른 것을 새로 알지 못한 채 문장 Y만을 새로 알게 된 상황 t에서, 문장 X를 믿는 정도는 베이즈 공리 또는 조건화 규칙에 따라

(C) $C_t(X) = C_o(X|Y) = C_o(X\&Y)/C_o(Y)$

로 바뀐다. C_M(앞)을 셈하기 위해 우리가 당연히 받아들여야 하는 논제를 먼저 말해 놓겠다.

(E1) C_S(가|앞) = C_S(다|앞) = 1/2
(E2) C_S(가|뒤) = C_S(다|뒤) = 1
(E3) C_S(앞) = C_S(뒤) = 1/2

(E1), (E2), (E3)으로부터

(1) C_S(가) = C_S(가|앞)C_S(앞) + C_S(가|뒤)C_S(뒤) = 3/4

이다. 마찬가지로

(2) C_S(다) = C_S(다|앞)C_S(앞) + C_S(다|뒤)C_S(뒤) = 3/4

이다.

만일 가나가 월요일에 깨어났다면 그가 놓인 인식 상황은 S에서 M으로 바뀐다. 그는 다른 것은 새로 알지 못한 채 정보 “가나가 월요일에 깨어났다”만을 새로 알게 된다. 그에게 C_M(앞)은 조건화 규칙 (C)를 써서

(3) C_M(앞) = C_S(앞|가) = C_S(앞&가)/C_S(가)
= C_S(가|앞)C_S(앞)/C_S(가) = 1/3

이다. 만일 다라가 월요일에 깨어났다면 그에게도 C_M(앞)은 1/3이다. 따라서 가나든 다라든, 그가 인식 상황 M에 있다면 그에게 C_M(앞)은 1/3이다. 우리는 이러한 믿음직함 셈을 받아들일 수 있는지 다음 절에서 따져 보겠다. 실험에 참여하지 않은 제3자인 마바가 가나와 다라 가운데 어느 한 사람이 월요일에 깨어났다는 정보를 듣게 된다면, 마바에게도 C(앞)은 1/3일까?

3. 관찰 선택 효과

우리 실험에 참여하지 않은 마바는 월요일에 누가 깨어났는지

모른 채 제3의 방에 앉아 있다. 그는 우리 실험의 설정을 잘 알고 있지만, 일요일에 던진 동전이 앞면이 나왔는지 뒷면이 나왔는지 모르고 있다. 월요일에 마바가 놓인 이러한 인식 상황을 O라고 하겠다. 이 상황에서 그에게 C_O(앞) = 1/2이다.[4] 우리는 마바에게 월요일에 누가 깨어났는지 한 사람만을 알려 줄 것이다.

만일 마바가 정보 "가나가 월요일에 깨어났다"를 얻게 된다면 그는 새로운 인식 상황 A에 놓인다. 이 상황에서 마바에게 C_A(앞)은 조건화 규칙 (C)를 써서 아래와 같이 셈해야 할 것 같다.[5]

(4) C_A(앞) = C_O(앞|가) = C_O(가|앞)C_O(앞)/C_O(가) = 1/3.

만일 마바가 정보 "다라가 월요일에 깨어났다"를 얻게 된다면 그는 새로운 인식 상황 B에 놓인다. 이 상황에서도 마바에게 C_B(앞)은 1/3일 것 같다. 이 셈이 옳다면, 마바가 인식 상황 A에 있든 B에 있든, 그에게 C(앞)은 1/3이다.

월요일 제3의 방에 있는 마바에게 인식 상황 A든 B든 이것은 의미 있는 상황이 아닌 것 같다. 그는 가나와 다라 가운데 적어도 한 사람은 월요일에 깨어난다는 것을 이미 알고 있다. 마바에게 "가나가 월요일에 깨어났다"나 "다라가 월요일에 깨어났다"는 사소한 정보처럼 보인다. 하지만 그가 "가나가 월요일에 깨어났다"를 알게 되면, 식 (4)가 말해주듯이, C(앞)이 1/2에서 1/3로 바뀐다. 달

---

4) 믿음직함 함수 C_O(X)는 인식 상황 O에 놓인 인식 주체가 X를 믿는 정도이다. 실험에 참여하지 않는 외부자가 월요일에 인식 상황 O에 있을 때 논제 (E1), (E2), (E3)과 똑같은 꼴의 논제들이 성립한다. (E1*) C_O(가|앞) = C_O(다|앞) = 1/2; (E2*) C_O(가|뒤) = C_O(다|뒤) = 1; (E3) C_O(앞) = C_O(뒤) = 1/2.

5) C_A(X)와 C_B(X)는 인식 주체가 각각 인식 상황 A와 B에 있을 때 믿음직함 함수이다. 제3의 방에 있는 마바에게 C_O(가) = C_O(가|앞)C_O(앞) + C_O(가|뒤)C_O(뒤) = 3/4이다. 마찬가지로 마바에게 C_O(다) = 3/4.

리 말해 C(뒤)가 1/2에게 2/3로 바뀐다. 또한 그가 “다라가 월요일에 깨어났다”를 알게 되면, 똑같은 까닭에서 C(뒤)가 1/2에게 2/3로 바뀐다. 하찮은 정보처럼 보이는 것이 자기의 믿음직함을 바꿀 만한 정보로 쓰이고 있는 셈이다.

우리 실험을 극단의 실험으로 몰고 가서 우리에게 닥친 야릇함을 확연히 드러내도록 하자. 실험에 스스로 참여할 N명의 사람들을 모아 놓고, C_O(앞) 곧 C_S(앞)이 h인 동전을 던진다.[6] 만일 이 동전이 뒷면이 나오면 월요일에 N명 모두를 한꺼번에 깨우고, 앞면이 나오면 이들 가운데 오직 한 명만 마구잡이로 뽑아 그를 월요일에 깨운다. 이들은 각기 다른 방에 있기 때문에, 다른 이들이 자고 있는지 깨어 있는지 전혀 알지 못한다. N명의 미인들을 $미_1$, $미_2$, … $미_N$ 등으로 부르고, 이들 가운데 어느 한 사람 $미인_n$이 월요일에 깨어났다는 정보를 듣게 되는 인식 상황을 N이라 하자.

딴 방에 있던 마바가 상황 N에서 앞면이 나왔으리라 믿는 정도는 아래와 같이 셈할 수 있겠다.

(5) $C_N(\text{앞}) = C_O(\text{앞}|\text{미}_n) = \dfrac{h}{h+tN}$.[7]

이 식에 따르면, t가 거의 0에 가까워 뒷면이 거의 나오지 않는 매우 치우친 동전을 던지는 실험에서도, N이 매우 크기만 하다면, 마

---

6) h가 1/2일 때 이 동전은 멀쩡한 동전이지만 다른 값일 때는 치우친 동전이다. h가 멀쩡한 동전이냐 치우친 동전이냐는 우리 생각 실험의 핵심 요소는 아니다. h가 0과 1 사이의 아무 값인 동전을 던지는 실험에서 우리가 일반해를 구할 수 있다면, h가 1/2인 동전을 던지는 특수한 경우의 답도 손쉽게 구할 수 있다.

7) $C_O(\text{미}_n|\text{앞})$은 1/N이고, $C_O(\text{미}_n|\text{뒤})$는 1이기 때문에, $C_O(\text{미}_n) = C_O(\text{미}_n|\text{앞})C_O(\text{앞}) + C_O(\text{미}_n|\text{뒤})C_O(\text{뒤}) = h/N + t$. 여기서 $t = C_O(\text{뒤}) = C_S(\text{뒤}) = 1 - h$. $C_O(\text{앞}|\text{미}_n) = C_O(\text{앞}\&\text{미}_n)/C_O(\text{미}_n) = C_O(\text{미}_n|\text{앞})C_O(\text{앞})/C_O(\text{미}_n) = h/(h + tN)$.

바에게 C_N(앞)은 거의 0에 가까이 가고, C_N(뒤)는 거의 1에 가까이 간다.

월요일에는 한 사람이든 여러 사람들이든 반드시 누군가 깨어나게 되어 있다. 식 (5)가 옳다면, 딴 방에 있는 마바가 뒷면이 나오기 몹시 힘든 동전을 던졌다는 것을 알고 있다고 해도, 누군가 깨어났다는 정보를 듣게 된 다음 그는 뒷면이 나왔다고 매우 굳게 믿어야 한다. 마바는 그냥 기다리고 있으면 어차피 누군가 깨어났다는 것을 듣게 될 텐데, 그 사람이 누구인지 이름을 듣기 전에는 C(뒤)가 거의 0에 가깝다고 생각하다가 그 사람의 이름을 듣자마자 C(뒤)가 거의 1에 가깝다고 생각하게 된다. N명의 사람들 가운데 누가 깨어나더라도, 마바에게 이 정보는 일요일에 던진 동전이 뒷면이 나왔다는 것을 거의 확실히 뒷받침하는 입증 증거로 쓰일 수 있다는 말이다. 이 야릇한 상황은 우리 직관과 크게 어긋난다.

식 (4)와 식 (5)에는 잘못된 셈이 들어 있다. 마바가 C_A(앞) = C_O(앞|가) = C_O(앞&가)/C_O(가)라고 셈하기 위해, 정보 “가나가 월요일에 깨어났다”는 그에게 마구잡이 절차를 거쳐 주어져야 한다. 정보, 증거, 증언, 관찰 등이 우리에게 주어지는 절차는 크게 두 가지로 나눌 수 있다.

- 마구잡이 절차: 정보 채널이 특정 정보를 각별하게 고르지 않은 채, 그 정보가 인식 주체에게 어쩌다, 우연히 주어진다.
- 치우친 절차: 정보 채널 자체가 특정 정보를 각별하게 골라 그 정보가 인식 주체에게 주어진다.

정보가 주어지는 절차가 달라지면, 정보의 내용도 달라지고 그에 따른 믿음직함 변화 방식도 달라진다.[8)]

---

8) 이러한 현상을 “관찰 선택 효과” 또는 “선택 효과”라고 한다. Bostrom

한 정보가 마구잡이 절차를 거쳐 주어졌는지 치우친 절차를 거쳐 주어졌는지 판단하는 가장 중요한 단서는 무엇인가? 정보를 주는 사람, 체계, 현상이 관련 정보를 모두 다 아는 상태에서 그 정보의 일부를 우리에게 주었다면, 그 정보는 치우친 절차를 거쳐 주어진 셈이다. 치우친 절차를 거쳐 정보가 주어졌을 때 조건화 규칙을 쓰는 것을 삼가야 한다. 관련 정보가 어쩌다 주어지지 않고 그 정보를 이미 알고 있는 이가 그 정보를 각별히 골라 우리에게 준 것이라면, 그 정보는 때때로 우리의 믿음직함을 바꿀 만한 정보로 쓸 수 없다. 한 정보가 우연히 또는 자연히 우리에게 주어지지 않는다면, 그 정보는 관련 가설을 뒷받침하지 못할 수 있다.

한 정보 또는 증거 E가 마구잡이 절차를 거쳐 주어졌을 때 E^{RP}라고 쓰고, 치우친 절차를 거쳐 주어졌을 때 E^{BP}라고 쓰기로 하자.[9] E를 얻기 전 인식 상황 O에서 정보 E만 새로 주어진 인식 상황 E로 바뀔 때, 주체의 믿음직함은 $C_O(X)$에서 $C_E(X) = C_O(X|E)$로 바뀐다. 이 때 주체에게 주어진 정보가 E^{RP}냐 E^{BP}냐에 따라 $C_O(X|E)$는 다르게 셈해야 한다.[10]

(2002); Bradley (2012).

9) 치우침에도 정도가 있는데 여기서는 완전히 치우친 경우로 한정하겠다. 완전히 치우친 절차를 거쳐 정보 E가 주체에게 주어진다는 것은 E가 틀림없이 참이 되는 방식으로 그 주체에게 정보 E가 주어진다는 것을 뜻한다. 따라서 완전히 치우친 절차를 거쳐 정보 E가 주체에게 주어질 경우 그에게 $C_O(E^{RP})$는 1이다. 그물코 크기가 5센티미터인 그물로 한강에서 물고기 100마리를 잡았는데 이로부터 정보 “한강에서 잡은 100마리 물고기는 모두 5센티미터보다 컸다”를 얻었다. 그물 자체가 완전히 치우친 정보 채널 역할을 하기 때문에, 이렇게 얻은 정보는 치우친 절차를 거쳐 얻은 셈이다. 이 정보 “한강에서 잡은 100마리 물고기는 모두 5센티미터보다 컸다”는 “한강에는 5센티미터보다 큰 물고기만 살고 있다”를 뒷받침하는 증거로 쓰기 어렵다.

10) 김명석 (2016). 이와 비슷하지만 보다 앞선 연구로는 Schrödinger (1947), Hutchison (1999) 등이 있다.

$$(\text{RP})\ C_O(X|E^{RP}) = C_O(X\&E^{RP})/C_O(E^{RP})$$
$$(\text{BP})\ C_O(X|E^{BP}) = C_O(X)$$

따라서 우리는 딴 방에 있던 마바에게 주어진 정보가 어떤 절차를 거쳐 그에게 주어졌는지 따져 보아야 한다.

실험을 꾸리는 사람이 다음과 같은 방식으로 마바에게 정보를 알려준다고 해보자. 월요일에 깨어난 사람이 가나뿐일 때 마바에게 “가나가 월요일에 깨어났다”고 말해준다. 월요일에 깨어난 사람이 다라뿐일 때 마바에게 “다라가 월요일에 깨어났다”고 말해준다. 월요일에 깨어난 사람이 가나와 다라일 때 마바에게 “가나가 월요일에 깨어났다” 또는 “다라가 월요일에 깨어났다”라고 말해준다. 마바는 “가나가 월요일에 깨어났는가?”라고 먼저 물어볼 수 없었으며, 또한 “다라가 월요일에 깨어났는가?”라고 먼저 물어볼 수도 없었다. 정보 “가나가 월요일에 깨어났다”나 정보 “다라가 월요일에 깨어났다”가 마바에게 주어질 때 이 정보가 거짓이 될 위험이 애초에 차단된 채 주어졌다. 이 정보는 마바에게 어쩌다 주어진 정보가 아니라, 특수한 정보 채널을 거쳐 주어졌다. 이런 방식으로 마바에게 주어진 정보는 치우친 절차를 거쳤다고 보아야 한다.

치우친 절차를 거쳐 마바에게 정보가 주어졌기 때문에, 마바는 규칙 (BP)에 따라 자기 믿음직함을 바꾸어야 한다. 따라서 식 (4)의 셈은 다음과 같이 고쳐야 한다.

$$(6)\ C_A(\text{앞}) = C_O(\text{앞}|\text{가}^{BP}) = C_O(\text{앞}) = 1/2$$

마찬가지로 정보 “다라가 월요일에 깨어났다”도 완전히 치우친 절차를 거쳐 마바에게 주어졌기 때문에, 마바는 규칙 (BP)를 따라야 한다. 인식 상황 B에 있는 그에게 $C_B(\text{앞}) = C_O(\text{앞}|\text{나}^{BP}) = C_O(\text{앞})$

= 1/2이다.

딴 방에 있는 마바에게 정보 “가나가 월요일에 깨어났다”나 “다라가 월요일에 깨어났다”는, 일요일에 던진 동전이 뒷면이 나왔다는 것을 뒷받침하는 입증 증거로 쓰일 수 없다. 이 통찰은 완전히 일반화된 실험에서도 똑같이 적용할 수 있다. 딴 방에 있던 마바는 정보 “미인$_n$이 월요일에 깨어났다”를 치우친 절차를 거쳐 받게 된다. 그가 인식 상황 N에 놓이게 될 때 C_N(앞)은 C_O(앞|미$_n^{BP}$)이고 이 값은 C_O(앞) 곧 h이다.

4. 외부자의 마구잡이 절차

실험에 참여하지 않는 외부자 마바의 관점에서, 만일 정보 “가나가 월요일에 깨어났다”나 정보 “다라가 월요일에 깨어났다”가 치우친 절차를 거쳐 자신에게 주어졌다면, 이 정보는 동전이 뒷면이 나왔음을 뒷받침하는 증거로 쓰일 수 없다. 물론 관련 정보가 마구잡이 절차를 거쳐 외부자 마바에게 주어질 수 있다. 월요일에 누가 깨어났는지 이미 알고 있는 사람이 관련 정보를 골라 마바에게 알려주는 것은 치우친 절차이다. 하지만 마바가 가나와 다라 가운데 한 명을 마구잡이로 뽑아 그가 월요일에 깨어났는지를 묻는 방식을 취하는 것은 마구잡이 절차이다.

이런 식으로 정보를 얻는 것이 왜 마구잡이 절차로 얻는 것인가? 완전히 치우친 절차를 거쳐 정보 E가 주체에게 주어진다는 것은 E가 틀림없이 참이 되는 방식으로 그 주체에게 정보 E가 주어진다는 것을 뜻한다. 마바가 가나와 다라 가운데 한 명을 마구잡이로 뽑아 그가 월요일에 깨어났는지를 묻는 방식을 취할 경우, 정보 “가나가 월요일에 깨어났다”나 정보 “다라가 월요일에 깨어났다”는 틀릴 위험을 가진 채 마바에게 주어진다. 이것은 마구잡이 절차에

해당하며, 이와 같은 방식으로 정보가 주어질 경우 마바에게 C_O(가)와 C_O(다)는 1이 아니라 3/4이다. 이 경우 규칙 (RP)에 따라 믿음직함 변화를 셈해야 하기 때문에 C_A(앞)과 C_B(앞)은 1/3이다.[11)]

실험에 참여하는 사람들이 아주 많은 실험 설정에서, 마바가 누가 깨어났는지를 이미 아는 우리에게 미인$_3$이 깨어났는지 묻는다고 생각해 보라. 우리가 미인$_3$이 깨어나지 않았다고 말한다면, 마바는 일요일에 던진 동전이 앞면이 나왔다는 것을 알게 된다. 하지만 우리가 미인$_3$이 깨어났다고 말한다면, 마바는 동전이 뒷면이 나왔으리라고 믿는 쪽으로 기울어야 한다. 일반해 식 (5)가 말해주는 것은 바로 이 점이다. 1억 명의 미인들이 이 실험에 참여하고, C_O(앞) 곧 C_S(앞)이 9999/10000나 되는, 앞면으로 매우 치우친 동전을 던진다고 생각해 보자. 마바가 미인$_{84399876}$이 깨어났는지 물었는데 우연의 일치로 미인$_{84399876}$이 정말로 깨어났다는 답을 얻게 된다면, 그는 앞면이 나왔으리라고

$$C_N(\text{앞}) = C_O(\text{앞}|\text{미}_{84399876}) = \frac{0.9999}{0.9999+10000} \doteq 1/10002$$

만큼 믿어야 한다. 다시 말해 동전이 뒷면이 나왔을 것이라고 매우 굳게 믿어야 한다. 마바 자신이 마구잡이로 뽑은 미인$_{84399876}$이 깨어났다는 사실을 듣는다면, 그는 월요일에 1억 명 가운데 한 명이 깨어났는데 자신이 바로 그 사람을 어쩌다 맞히었을 리가 없다고 생

11) 가나와 다라 가운데 적어도 한 사람은 깨어나는 것이 분명한 상황에서, “누가 깨어났는가?”라는 물음의 답변으로 얻게 된 정보는 치우친 절차를 거쳐 얻은 정보이기 쉽다. 하지만 “가나가 깨어났는가?”라는 물음의 답변으로 얻게 된 정보는 마구잡이 절차를 거쳐 얻은 정보이다. 내가 실험에 직접 참여하는 상황에서 “나는 깨어났다”라는 자기의식 정보는 “누가 깨어났는가?”라는 물음의 답변으로 얻게 된 정보인가, “나는 깨어났는가?”라는 물음의 답변으로 얻게 된 정보인가? 이것을 따지는 것은 제5절의 주제이다.

각해야 한다. 왜냐하면 그것을 어쩌다 맞히리라고 믿는 정도는 1억 분의 1밖에 되지 않기 때문이다. 차라리 마바는 미인들이 모두 깨어났기 때문에 미인$_{84399876}$도 깨어났을 것이라고 생각하는 쪽이 낫다. 모든 사람들이 한꺼번에 깨어나는 일의 일어남직함은 1만 분의 1이고, 이 크기는 1억 분의 1에 비해 만 배나 크다.

5. 내부자의 자기의식 정보

내부자의 관점에 있는 사람은 자신이 받은 정보가 마구잡이 절차를 거친 정보라고 생각해야 할까 치우친 절차를 거친 정보라고 생각해야 할까? 이제 나 자신이 이 실험에 참여하는 경우를 생각해 보겠다. 실험에 참여하는 나는 월요일에 깨어날 수도 있고 깨어나지 않을 수도 있다. 일요일 저녁에 내가 깨어날까 궁금해 하며 잠든 뒤 월요일 아침에 내가 깨어났다. 깨어나자마자 나는 자기의식 정보 “내가 지금 깨어났다”를 얻는다. 실험의 설정 때문에 “지금”은 곧 “월요일”이다. 정보 “내가 월요일에 깨어났다”를 “내”라고 약칭하겠다. 분명 일요일 저녁의 인식 상황 S에 있을 때 C_S(내) = 3/4이다. 월요일에 내가 깨어나서 인식 상황 M에 놓이기 될 때 C_M(앞)은

(7) C_M(앞) = C_S(앞|내) = C_S(내|앞)C_S(앞)/C_S(내) = 1/3

로 셈해야 할 것 같다. 여기서 C_S(내|앞)이 1/2이라는 사실을 썼다.

내부자의 관점과 외부자의 관점에 중요한 차이가 있다.[12] 동전을

12) 관점의 차이가 믿음직함의 차이를 낳는다는 주장은 김한승 (2009), 김한승 (2011)을 보라. 나는 김한승이 말하는 ‘관점’이 일종의 정보이며, 관점의 차이란 곧 정보의 차이라고 생각한다. 또한 정보 또는 증거가 주어지는 절차 자체가 해당 정보의 일부 내용을 구성하기 때문에, 인식 주체는 자기가 받

이미 던졌고 누군가 깨어난 다음에도, 외부자는 인식 상황 O에 놓일 수 있다. 외부자가 인식 상황 O에 있을 때 그에게 C_O(앞)은 1/2이다. 만일 외부자가 누가 깨어났는지에 관한 정보 E를 치우친 절차를 거쳐 얻는다면, 그는 인식 상황 E^{BP}에 놓이게 되고, 이 때 그에게 C_E(앞)은 여전히 1/2이다. 만일 외부자가 누가 깨어났는지에 관한 정보 E를 마구잡이 절차를 거쳐 얻는다면, 그는 인식 상황 E^{RP}에 놓이게 되고, 이 때 그에게 C_E(앞)은 1/3로 바뀐다.

동전을 이미 던졌고 누군가 깨어난 다음에, 내부자는 더 이상 인식 상황 O에 놓일 수 없다. 내부자는 월요일에 의식을 갖자마자 자기의식 정보 “내가 지금 깨어났다”를 얻게 되고 그는 곧바로 인식 상황 M에 놓인다. 월요일 나에게 주어진 정보 “내가 지금 깨어났다”는 마구잡이 절차를 거친 정보인가 치우친 절차를 거친 정보인가? 나의 자기의식 정보가 치우친 절차를 거쳐 나에게 주어졌다고 가정해 보겠다. 만일 월요일에 내가 치우친 절차를 거쳐 정보 “내”를 얻었다면, 내가 얻은 정보는 “내”^{BP}이기 때문에, 규칙 (BP)에 따라 나에게 C_M(앞)은 1/2이다.

일요일에 멀쩡한 동전을 던지고, 예컨대 70억 명의 사람들이 실험에 참여하고 있다고 생각해 보라. 뒷면이 나오면 이들 모두를 월요일에 한꺼번에 깨운다. 앞면이 나오면 이들 가운데 한 명만 마구

은 정보를 제대로 파악하기 위해 그는 정보가 주어지는 절차도 감안해야 한다. 이 절차에 대한 판단을 김한승은 ‘관점’이라고 쓴 것이 아닌가 생각한다. 우리 실험에 참여하는 내부자는 자신에게 주어진 정보가 치우친 절차로 주어진 것인지, 마구잡이 절차로 주어진 것인지 판단해야 하는데, 김한승은 두 개의 판단을 각기 다른 관점이라고 부를 것 같다. 용어의 차이를 빼면 그와 견해를 같이하지만, 나는 바로 이 관점 또는 절차가 정보의 차이를 낳으며, 특정 상황에서 두 관점들 또는 두 절차들 가운데 오직 하나만이 올바른 관점이며 올바른 절차라고 생각한다. 다만 한 주체는 때때로 자신이 취할 올바른 관점 또는 상황에 맞는 절차가 무엇인지 헷갈리는 상황에 처할 수 있다는 데 반대하지 않는다.

잡이로 뽑아 깨운다. 설사 일요일에 던진 동전이 앞면이 나온다고 해도, 70억 명 가운데 월요일에 내가 깨어날 가능성은 고작 70억분의 1밖에 되지 않는다. 희한하게도 월요일에 내가 깨어났을 때 내가 "오직 나 자신만 깨어났다"고 믿는 정도가 1/2로 유지되어야 한다는 점은 야릇한 일이다.

월요일에 내가 얻은 정보 "내"는 치우친 절차가 아니라 마구잡이 절차를 거쳐 나에게 주어진 것이 아닌가 생각된다. 가령 내부자 김명석이 월요일에 깨어날지를 외부자 윤보석이 일요일부터 묻고 있었다고 생각해 보라. 그런데 정말로 "김명석이 월요일에 깨어났다"를 외부자 윤보석이 월요일에 알게 된다면, 그는 이 정보를 마구잡이 절차를 거쳐 알게 된 셈이다. 내부자 김명석도 자신이 월요일에 깨어날지를 일요일부터 주의집중해서 묻고 있었다고 생각해 보라. 정말로 "김명석이 월요일에 깨어났다"를 내부자 김명석이 월요일에 알게 된다면 그 또한 이 정보를 마구잡이 절차를 거쳐 알게 된 셈이라고 말해야 하지 않을까?

실험에 내가 참여할 때 월요일에 얻게 되는 자기의식 정보 "내"가 "내"^{RP}라는 우리의 추정은 야릇한 귀결을 갖지 않는가? 이제 제2절의 처음 문제로 돌아가 보자. 실험에 참여하는 가나와 다라는 둘 다 실험의 내부자들이다. 월요일에 깨어난 사람은 인식 상황 M에 놓이게 되고, 그는 마구잡이 절차를 거쳐 자기의식 정보 "내가 월요일에 깨어났다"를 얻게 되었다고 생각할 것이다. 가나든 다라든, 월요일에 깨어난 미인은 C_M(앞)은 1/3이라고 생각해야 한다. 나아가 매우 일반화된 실험에서, 자기의식 정보 "내"가 마구잡이 절차를 거쳐 나에게 주어졌다면, 식 (5)의 셈과 비슷하게, 실험에 참여하는 사람 누구에게도 C_M(앞)은 $h/(h+tN)$이다. 설사 h가 거의 1에 가깝다 하더라도 N이 엄청나게 크다면 C_M(앞)은 0에 가깝다.

인구가 제한되어 있다고 생각하지 말고 어마어마하게 많은 사람

들 예컨대 $10^{10,000,000}$명의 사람들을 모아 실험한다고 생각해 보라. 뒷면이 나오면 이들 모두를 월요일에 한꺼번에 깨운다. 앞면이 나오면 이들 가운데 한 명만 마구잡이로 뽑아 깨운다. 우리가 일요일에 던질 동전은 뒷면이 나오기 몹시도 힘들다. 예컨대 이 동전이 뒷면이 나오는 일은 멀쩡한 동전 1000만 개를 한꺼번에 던져 모두 뒷면이 나오는 것만큼 힘들다. 이런 동전을 던지면 거의 틀림없이 앞면이 나온다. $10^{10,000,000}$명의 사람들 가운데 적어도 한 명은 월요일에 깨어날 것이다. 월요일에 깨어난 사람이 누구든 그는 자기의식 정보를 얻게 되고, 그는 그 정보를 마구잡이 절차를 거쳐 얻게 되었다고 생각할 것이다. 결국 그가 C_M(앞)이 h/(h+tN)라고 생각한다면,[13] 그는 앞면이 나왔을 리 없고, 뒷면이 나왔음이 거의 틀림없다고 말해야 한다. 하지만 외부자의 관점에서 보면, 깨어난 그 사람은 결국 정답을 맞히는 데 성공하지 못한다.

이런 식의 실험을 1000만 번 반복해도 이 실험에 참여하여 깨어난 거의 모든 이들은 정답을 맞히는 데 실패한다. 이 실험을 $2^{10,000,000}$번 반복했을 때 거의 한 번쯤 뒷면이 나오는데 그 때 비로소 실험에 참여한 사람들이 정답을 맞히게 될 것이다. 이처럼 정답을 맞히는 실험 사례는 매우 드물게 나타나고, 거의 모든 실험 사례들에서 실험 참여자는 정답을 맞히는 데 실패한다. 이것은 무엇을 뜻하는가? 실험 참여자들이 C_M(앞)이 h/(h+tN)라고 생각하는 것이 옳지 않다는 것을 말해주는가? 나아가 이것은 월요일 깨어난 이들이 얻는 자기의식 정보가 마구잡이 절차를 거친 정보라는 우리의 결론이 의심스럽다는 것을 뜻하는가?

실험에 참여한 내부자가 월요일에 깨어나자마자 받은 자기의식 정보가 치우친 절차를 거쳐 주어진 정보라고 섣불리 결론내려서는

13) h는 거의 1이고, tN = $10^{10,000,000}/2^{10,000,000}$ = $5^{10,000,000}$이다. C_M(앞) = C_S(앞|내) = h/(h+tN) ≐ $1/5^{10,000,000}$이고 이 값은 거의 0이다.

안 된다. 우리가 실험에 참여하는 사람들 전체에 관심을 가질 때 앞의 귀결이 야릇해 보일 뿐이다. 우리는 오히려 실험에 참여하는 특정 한 사람에 주목해야 한다. C_M(앞)을 셈하는 일은 실험에 참여하는 김명석 본인이며, 그는 일요일과 월요일에 주의집중해서 “나는 깨어났는가?”를 묻는다. 일요일에 던진 동전은 뻔히 앞면이 나온다. 하지만 $10^{10,000,000}$명의 사람들 가운데 하필이면 나 김명석이 월요일에 깨어나는 일은, 뒷면이 나오는 일보다 훨씬 어려운 일이다. 그 일의 일어남직함보다 뒷면이 나오는 일의 일어남직함이 무려 $5^{10,000,000}$배나 더 크다.

6. 나오는 말

자기 자신이 지금 깨어나 있다는 사실은 때때로 자기 믿음의 크기를 바꿀 만한 정보로 쓸 수 있다. 이 글에서 선보인 깨어난 사람들의 문제에서, 가나든 다라든, 그가 깨어나자마자 자기의식 정보 “나는 지금 깨어났다”를 얻게 된다. 우리는 깨어난 사람들의 문제를 통해 자기 자신이 지금 깨어나 있다는 자기의식 정보가 어떤 본성을 갖고 있는지를 성찰하고 싶었다. 이러한 자기의식 정보가 치우친 절차를 거쳐 주체에게 주어지는지 마구잡이 절차를 거쳐 주체에게 주어지는지 따져 보았다. 이 정보가 치우친 절차를 거쳐 주어졌다면 이 정보는 관련 가설을 입증하는 증거로 쓰기 어렵다.

실험에 참여하는 사람은 “나는 지금 깨어나지 않았다”를 입수할 수도 들을 수도 관찰할 수도 없다. 달리 말해 정보 “나는 지금 깨어났다”가 거짓이 되는 일이 처음부터 실험 내부자에게 아예 차단되어 있었다. 실험의 외부자와 달리, 내부자는 “내가 월요일에 깨어났는가?”를 아예 물을 수 없다. 이 점에서 자기의식 정보는 치우친 절차를 거쳐 그 자신에게 주어진 것으로 착각할 만하다.[14] 하지만

우리 실험에서 자기의식 정보는 어쩌다 또는 자연히 내부자에게 주어진 것으로 보는 것이 보다 합당하다.

내가 일요일 실험에 참여하게 될 때 월요일에 “내”가 참일 것이라는 것이 틀림없이 보장된 것은 아니었다. 월요일에 깨어난 사람이 반드시 내가 아니라는 점에서 “내”는, 실험하는 사람에게든 나에게든, 마구잡이 절차를 거쳐 주어진 것이다. 우리는 여기서 매우 중요한 점을 발견한다. 실험에 참여하는 사람이 일요일부터 주의집중하여 자신이 월요일에 깨어날까 나지 않을까 심려하다가 월요일에 자신이 깨어났다는 정보를 얻게 된다면, 이러한 자기의식 정보는 마구잡이 절차를 거쳐 주어졌다고 보아야 한다는 점이다. 주의집중하던 사람에게 월요일에 자신이 깨어났다는 사실은 뜻밖에 정보이다.

한편 동전을 던져 앞면이 나오면 1명이 생겨나고, 뒷면이 나오면 N명이 생겨나는 이른바 “태어난 사람들의 문제”에서 새로 태어난 사람들은 동전을 던지기 전에 의식 자체가 없었다.[15] 별도 논의가 있어야 하겠지만, 이처럼 실험 과정에서 새로 태어난 사람들의 자기의식 정보는, 현재의 짧은 내 판단으로는, 치우친 절차를 거쳐 의식 주체에게 주어졌다. 그들은 자신이 깨어났다는 또는 의식을 갖게 되었다는 사실에 놀랄 까닭이 별로 없다. 하지만 우리 실험에서 월요일에 깨어난 이들은 하필 자신이 깨어났다는 점에 놀랄 까닭이 있다.[16]

---

14) 한 심사자께서 지적했듯이, 마구잡이 절차와 치우친 절차 사이의 구분이 우리 실험에 논란의 여지없이 깔끔하게 적용될 수 있는지 분명하지 않다.

15) Bostrom (2003); 김한승 (2011).

16) 일요일에 주의집중 없이 흐릿하게 또는 멍청하게 있다가, 의식을 차려보니 월요일이었다고 한다면, 이 경우 이 자기의식 정보가 마구잡이 절차를 거쳐 주체에게 주어진 것인지 치우친 절차를 거쳐 주체에게 주어진 것인지는 하나의 새로운 물음이다. 이 물음의 답을 여기서 답하지는 않겠다.

참고문헌

김한승 (2009), “비개념적 내용으로서의 지표적 내용”, 『철학적 분석』, 20호, pp. 119-140.

김한승 (2011), “확률에 대한 관점주의”, 『논리연구』, 14집 3호, pp. 59-83.

김명석 (2016), “두 딸 문제와 선택 효과”, 『논리연구』, 19집 3호, pp. 369-400.

송하석 (2012), “잠자는 미녀 문제, 다시 보기”, 『철학적 분석』, 25호, pp. 1-21.

Bostrom, N. (2002), *Antropic Bias: Observation Selection Effects in Science and Philosophy*, New York: Routledge.

Bostrom, N. (2003), “The Mysteries of Self-Locating Belief and Anthropic Reasoning”, *The Harvard Review of Philosophy,* 11(1), pp. 59-73.

Bradley, D. J. (2012). “Four Problems About Self-Locating Belief”, *Philosophical Review,* 121(2), pp. 149-177.

Hutchison, K. (1999), “What Are Conditional Probabilities Conditional Upon?”, *British Journal for the Philosophy of Science,* 50(4), pp. 665-695.

Schrödinger, E. (1947), “The Foundation of the Theory of Probability”, *Proceedings of the Royal Irish Academy,* A51, pp. 51-66.

국민대학교 교양대학
College of General Education, Kookmin University
myeongseok@gmail.com

ARTICLE ABSTRACTS

Self-Consciousness Information and Selection Effect

Myeongseok Kim

In modern cosmology, it is controversial whether the existence of human consciousness can be used as evidence to support the hypothesis that many parallel universes are actualized. In this paper, we want to explore the nature of self-consciousness information that I am awake now. Consider the following experiment participating Al and Bob. We throw a fair coin on Sunday. If the coin lands heads we wake up just one of Al and Bob on Monday. If the coin lands tails we wake up both of Al and Bob. On Monday, at least one of Al and Bob will wake up, to what degree ought they believe that the outcome of the coin toss is heads? We will argue that the correct answer to this question is 1/3. To this end, we will argue the awakened person's information that "I am awake" is given to himself through a random procedure.

Key Words: biased procedure, selection effect, sleeping beauties problem, random procedure