게임 이론으로 보는 분배 정의의 문제

DOI: 10.3938/PhiT.24.005

백 승 기

Game Theory on the Problem of Distributive Justice

Seung Ki BAEK

How to distribute a certain amount of goods between rational players is an important question both in theory and practice. This article is written as a basic and brief introduction to game-theoretic approaches to this issue. Beginning with the ultimatum game, we explain Nash-Rubinstein bargaining and its relation to Rawlsian political theories. We also discuss some methodological aspects of game theory from a physicist's point of view.

서 론

분배의 문제가 우리 사회에서 얼마나 중요한 문제인지는 그리 많은 설명이 필요하지 않다. 샌델의 '정의란 무엇인가'에 끌렸던 관심, 그리고 이어 피케티의 '21세기 자본'을 대하는 반응은 다른 철학이나 경제학 서적의 경우와 분명 달랐다. 사실 분배 문제에 대한 관심은 꼭 오늘날에 국한된 것도 아니며, 어쩌면 경제학 자체를 '희소한 재화의 생산과 분배, 소비를 연구하는 학문'이라 생각할 수도 있겠다. 비록 일반적으로 가장 잘 알려진 게임 이론은 협력의 어려움을 말해주는 '죄수의 딜레마'이지만, 게임 이론에서 다루는 분배의 문제도 그에 못지않게 흥미롭고 깊이 있는 주제이다. 이 글에서는 그 간략한 역사와 흥미로운 논증들을 훑어보고 그 결과를 어떻게 이해해야 할지 생각해보고자 한다. 왜 하필 물리학자가 분배의 문제에 관심을 가지는지 독자들 중 누군가 궁금해 한다면 일단 피케티의 말

로 답을 대신해보도록 하자: "부의 분배는 매우 중요한 문제여서 경제학자, 사회학자, 철학자들에게만 맡겨둘 수 없다."

협상 이론

1. 최후통첩 게임

'최후통첩 게임'에는 두 명의 참가자 A와 B가 있어서, 100이라는 양의 재화를 놓고 얼마를 나누어 가질지를 A가 제안한다. 이 제안을 B가 받아들이면 분배가 이루어진 채로 끝나고, 거부하면 둘 다 아무 것도 얻지 못한 채 끝난다. 순서쌍에서 앞의 숫자가 A가 얻어가는 몫이고 뒤의 숫자가 B의 몫이라 했을 때, A의 분배 제안은 (1,99), (2,98), ..., (99,1) 사이에서 나올 수 있고, A와 B 모두 이 재화를 가능한 한 많이 차지하고 싶어 한다. 분석의 편의를 위해 A의 제안 중 두 가지만 생각해보자. 첫 번째는 (50,50)으로 공평(fair)하게 나누는 제안이고 두 번째는 (99,1)로 A가 더 많이 가져가는 불공평(unfair)한 제안이다. 이 둘 각각의 경우에 대해 B가 수락(Y) 혹은 거부(N)를 할 수 있으므로, B의 전략은 (YY, YN, NY, NN)이라는 네 가지 가능성이 있다. 예컨대 B의 YY의 전략은 무조건 A의 제안을 받는다는 뜻이고 YN은 공평한 제안만을 받아들이고 불공평한 제안이 들어오면 아예 판을 깨버리겠다는 뜻이 된다. 다음의 표 1에서 A의 전략 (fair, unfair)과 B의 전략 (YY, YN, NY, NN)의 조합에 따라 분배되는 재화의 양을 볼 수 있다. A나 B 중 누구든 굳이 제안을 바꿀 이유가 없는 지점들을 빨간색으로 표시해두었다. 예컨대 A가 공평한 제안을 하고 B가 YN이라는 전략을 가지고 있다면 둘 다 50씩을 가져가는데, A가 불공평한 제안으로 바꾸어 B가 판을 깨게끔 할 이유가 없고 B는 다른 전략으로 바꾸어봐야 50보다 더 받을 길이 없기 때문에 그러하다. 이 빨간색 지점들이 말하자면 내쉬 균형(Nash equilibrium)이라고 불리는 점들이다. 그러나 그 중 B가 NY를 하는 제일 마지막 균형은 A가 불공평한 제안을 했을 때에만 수락하겠다는 전략이고, 사실 전혀 말이 되지 않는 것이다. A가 실수로라도 공평한 제안을 해버렸을 때 B가 거부할

저자약력

백승기 교수는 KAIST에서 물리학 박사를 받고 지금은 부경대학교 물리학과 교수로 재직하고 있다. 통계 물리학을 바탕으로 갈등 상황 중 의사 결정의 문제를 연구하고 있다. (seungki@pknu.ac.kr)

Table 1. Normal form of the ultimatum game.

		B			
		YY	YN	NY	NN
A	fair	50,50	50,50	0,0	0,0
	unfair	99,1	0,0	99,1	0,0

이유가 있을까? 마찬가지로 A가 공평하게 나오고 B가 YN을 한다는 (50,50)의 배분도 비록 균형점이지만 신뢰가 가지 않는다. A가 불공평한 제안을 해서 B가 1만 받을 수 있다고 해도 어쨌든 빈손으로 돌아가는 것보다는 나운데 YN을 고수할 수 있을까? 물론 B는 그러겠다고 위협하겠지만 믿기 어렵다. A가 충분히 영리하다면 이 비현실성을 꿰뚫어보기 때문에 불공평한 제안을 하고 B는 언제나 수락할 수밖에 없다. 이런 식으로 A가 뭔가를 해버렸다고 가정해 게임을 축소시킴으로써, 즉 부분 게임(subgame)을 고려함으로써 비현실적인 전략들을 걸러낼 수 있다. 그 결과 얻어지는 균형이 소위 부분 게임 완전균형(subgame perfect equilibrium)이고 표 1에 굵은 글씨로 표시해두었다. 최후통첩 게임의 결과를 설명하면서 내쉬 균형과 부분 게임 완전 균형을 혼동할 수 있는데, 내쉬 균형은 꽤 많을 수 있고 공평한 분배조차도 포함할 수 있다. 그 상당수가 믿을 수 없는 것일 뿐이다.

2. 3단계 협상 게임

최후통첩 게임을 좀 더 확장해보자. 이번에는 A가 먼저 1의 양을 가지는 재화의 분배를 제안하는데, B가 이것을 수락하면 게임이 끝나고 거부하면 B가 역제안을 할 차례가 된다. B의 제안을 A가 수락하면 그렇게 분배가 끝나고, A가 거부하면 다시 A에게 제안할 권리가 넘어오게 된다. 다만 어느 쪽에서든 한번 제안이 거부될 때마다 전체 재화의 양은 $1/3$ 씩 사라진다고 하자. 따라서 거부가 세 번 발생하면 분배는 의미 없다. 그림 1이 이 게임을 나타낸 것이다.

여기에서도 핵심은 마지막 부분부터 생각하는 것이다. A가 세 번째 제안에서 재화 $1/3$ 중 최소 단위 e 만큼을 떼어주면서 $(1/3 - e, e)$ 로 나누자고 제안할 때 B가 거부할 이유가 있을까? 상황은 최후통첩 게임과 같고 따라서 B는 이 제안을 언제나 수락할 것이다. 물론 B는 이 상황이 달갑지 않을 테고 따라서 자기가 역제안할 차례일 때에 미리 A에게 $1/3 - e$ 만을 떼어주면 나머지 $1/3 + e$ 를 취할 수 있다. A는 B의 제안을 거부해야 어차피 마지막 단계에서 더 얻을 것이 없다는 사실을 알기 때문에 이를 수락할 것이다. 그러나 사실 A도 B의 제안까지 기다릴 필요가 없다. 그럴 바에야 제일 처음부터 B에게 $1/3 + e$ 를 주고 나머지 $2/3 - e$ 를 가져가도 B가 수락할 것이기 때문

Fig. 1. Extensive form of the three-stage bargaining game.

이다. 따라서 A는 게임이 시작하자마자 $(2/3 - e, 1/3 + e)$ 의 제안을 하고 B는 이를 즉각 수락하면서 상황을 종료한다는 것이 이론의 해법이다.

3. 루빈슈타인 협상 이론

이번에는 약간 이야기를 바꿔서 $t=1$ 에서 원래 1이었던 재화의 양이, 제안이 거부당할 때마다 $p, p^2, p^3, \dots$ 순으로 줄어든다고 해보자 ($p < 1$). 제안이 두 번 거부당한 $t=3$ 시점이라면 재화가 p^2 만큼 남고 이때 다시 A에게 제안 차례가 돌아올 것이다. A가 기본적으로 이 협상으로부터 가져갈 수 있는 비율이 z 로 유일하게 존재한다고 가정해보자. 가정에 의해 A는 $t=3$ 에서 p^2 중 $a_3 = p^2 z$ 를 요구하고 B에게 그 나머지 $b_3 = p^2(1 - z)$ 를 넘겨주겠다고 제안할 것이다. B는 이 상황까지 오기보다 그 앞 단계인 $t=2$ 에서 제안권을 가질 때에 더 많이 가져갈 수 있다. 즉 B가 A의 첫 제안을 막 거절한 $t=2$ 의 상황이라면 아직 재화가 p 만큼 남아있으니 그는 A에게 $a_2 = p z$ 를 미리 챙겨주고 나머지 $b_2 = p - a_2 = p - p^2 z$ 를 가지겠다고 제안했을 것이다. 이 양이 $t=3$ 에서 A로부터 받는 $b_3 = p^2(1 - z)$ 보다 분명히 많다. 그리고 A도 어차피 같은 양을 얻으니 손해를 보지도 않는다. 하지만 A의 관점에서 보면, A도 애초 $t=1$ 에서 B가 거부할 구실을 줌으로써 제안권을 넘겨줄 필요가 없다. 맨 처음부터 B에게 $b_2 = p - p^2 z$ 를 주고 나머지 $a_1 = 1 - b_2 = 1 - (p - p^2 z)$ 를 가지겠다고 했어도 B가 받아들였을 테니까 말이다. 우리의 가정으로부터, 즉 z 가 유일하다면 일관성을 위해 바로 이 a_1 이 z 와 일치해야 한다. 방정식 $1 - p + p^2 z = z$ 를 풀면 $z = 1/(1 + p)$ 를 얻는다. 즉 A는 게임을 시작하자마자 자신에게 $1/(1 + p)$, B에게 $p/(1 + p)$ 를 주는 분배안을 제시하고 B는 이를 즉각 수락하면서 상황을 종료한다는 것이 답이다. 그리고

이것이 루빈슈타인 협상 이론^[1]의 뼈대이다. 일반적으로 두 사람이 느끼는 시간 감소율이 다르다고 놓고 풀 수도 있는데, 그러면 재화의 양이 빨리 줄어든다고 느끼는 쪽, 그러니까 시간에 쫓겨 초조한 쪽이 작은 양을 얻는다. 이는 상식과 부합한다. 그리고 이로부터 협상은 대체로 기존 강자에게 유리하게 작용할 것이라고 추측할 수도 있겠다. 루빈슈타인은 p 가 1로 가는 극한에서 이렇게 제안과 역제안을 반복하는 협상 절차의 결과가 내쉬 협상의 결과와 일치한다는 것을 증명했다.

4. 내쉬 협상 이론

내쉬 협상 이론^[2]은 지금까지와 완전히 다른 접근법이다. 여기에서는 협상을, 두 사람 모두 결렬을 원하지 않지만 서로 원하는 합의 지점은 다른 상황이라고 개념화한다. 그 다음 복잡한 세부 사항을 과감히 쳐내고 합리적인 협상이 만족할 법한 몇 가지 조건만을 요구한다. 협상에 참여하는 사람들이 재화를 하나씩 더 받을 때마다 얻게 되는 행복감을 측정해서 이를 각자의 효용함수라고 부르자. 내쉬가 제안한 합리적 협상의 네 가지 조건은 이러하다: 첫째, 효용함수가 같은 사람들끼리는 50:50으로 가져가야 한다. 둘째, 하나도 남김없이 모든 재화가 분배되어야 한다. 셋째 조건은, 효용함수를 예컨대 온도처럼 표현한다면 썩씨를 쓰든, 화씨를 쓰든, 아니면 절대온도를 써서 표현하든 협상 결과에 영향을 미치지 않아야 한다. 그러면 내가 50이라 말하는 즐거움과 상대가 50이라 말하는 즐거움이 같은 정도인지 아닌지를 놓고 고민할 필요가 없다. 이를 ‘개인 간 비교(interpersonal comparison)’의 문제라고 한다. 마지막 조건은 이러하다: 예를 들어 A와 B 두 사람이 60:40으로 나누기로 협상한 때에 갑자기 “A가 50 미만을 가져서는 안 된다” 같은 불필요한 구속이 생겨난들 결과가 바뀌지 않는다. 내쉬는 결렬지점을 기준으로 이 네 조건을 만족하는 협상 결과가 유일함을 증명하고 그 답을 제시했다.

내쉬의 답이 앞에서 언급한 ‘내쉬 균형’과는 다른 개념이라는 것을 확실히 해두자. 그리고 이렇게 공리를 세움으로써 해의 존재를 가늠해보는 접근법은 엄청나게 위력적이어서 게임 이론의 다른 논법들에도 차용되고 있다. 예컨대 퇴머의 연구^[3]는 국제기구에서 어떤 병을 치료할 예산을 어떻게 배분해야 하는지 다루었다. 그는 여기에서도 5개의 합리적으로 보이는 공리를 도입함으로써 유일한 해법을 제시했다. 즉 발병률 1위 국가를 2위 국가 수준으로, 다음으로 이 두 나라를 3위 국가 수준으로 개선하며, 예산이 허용하는 한 이 작업을 반복하는 것이다. 퇴머의 예처럼 유일한 해가 발견되어도 흥미롭지만, 해가 존재하지 않는다는 것을 증명하게 되면 ‘불가능성 정리’의 형태가 된다.

다만 내쉬의 공리적 접근법에는 구체적으로 “어떻게”라고 하

는 부분이 미흡했는데, 루빈슈타인이 여기에 구체적인 실행 절차까지 제시해준 것이다. 그러나 루빈슈타인의 버전에서 정말로 해가 유일해야 하는지는 어느 정도 가정의 영역이다. 내쉬가 가정한 조건이 모두 현실적인지도 재고해볼 문제이다. 예컨대 최저임금의 존재는 그보다 높은 임금에 대한 노동자의 협상력도 함께 높여주는데 이는 내쉬가 가정한 마지막 조건을 의문스럽게 만든다. 나아가 내쉬의 가정들이 모두 옳다고 해도, ‘협상 결렬시의 상황’을 어떻게 놓느냐에 따라서도 다양한 답이 존재한다. 예컨대 새플리는 두 사람이 그 경우 맥시민(maximin) 해법을 시도할 것이라고 논한다. 이는 두 사람 사이에 아무 협력의 여지가 없는 소위 제로섬 게임의 답으로, 가능한 최악의 경우들을 모두 생각한 다음 그나마 나은 것을 고르는 방어적인 전략이다. 새플리에 비해 내쉬가 원래 생각했던 협상 결렬의 기준점 계산^[4]은 좀 더 복잡하다.

정의의 이론

어쩌면 답이 유일하지 않기 때문에 분배가 여전히 논란거리인지도 모른다. 현대 정치철학에서 분배적 정의를 논하는 데 있어 가장 중요한 저작 중 하나는 존 롤즈의 ‘정의론’이다. 롤즈는 리버럴의 입장에서 평등을 옹호하고, 불평등은 오직 최악의 상황에 놓인 사람에게 도움이 될 때에만 가능하다고 주장한다. 그는 이 주장을 위해 ‘무지의 장막’이라는 사고실험에 기댄다. 이 장막은 사람들이 모여서 앞으로 만들 사회를 논의할 때에 그 안에서 자신의 처지가 무엇일지 가려놓는 역할을 한다. 그러면 사람들은 자신이 최악의 처지에 놓일 우려를 합리적으로 계산에 넣기 때문에 평등한 사회를 만드는 데 동의할 것이라는 논리이다. 이것은 계층 간 이동이 활발하고 그래서 어느 정도 보편적 관점을 가능케 하는 현대 사회에서 가능한 평등론일 것이다. 다만 장막이 가려야 할 정보가 무엇이여야 하는지는 이견들이 있다.^[5] 롤즈의 이론에서 한 가지 흥미로운 부분은, 이 ‘무지의 장막’ 논리를 전개하는 중에 게임 이론으로부터 빌려온 개념이 협상 이론이 아니라 바로 맥시민이라는 점이다. 앞서 언급했듯, 맥시민은 일체의 협력을 기대할 수 없는 완전한 갈등 상황에서 자신의 몫을 보호하려는, 극단적으로 방어적인 전략이다. 게임 이론에서 일반적으로 가정하는 행위자라면 자신의 보수 기댓값을 높이기 위해 애쓸 뿐이지 안전 위

REFERENCES

- [1] A. Rubinstein, *Econometrica* **50**, 97 (1982).
- [2] J. Nash, *Econometrica* **18**, 155 (1950).
- [3] J. Roemer, *J. Econ. Theory* **45**, 1 (1988).
- [4] A. Rapoport, *Two-Person Game Theory* (Dover, NY, 1966).
- [5] J. Moreno-Torero and J. Roemer, *Econ. Philos.* **24**, 233 (2008).

주의 전략을 사용하지는 않을 것이다.^[6] 다만 칼라이는 내쉬와 다른 공리들에 기반한 자신의 협상 이론을 폴즈의 생각과 연결 짓고 있다.^[7]

더 난해한 문제는 평등의 이념 일반에 합의한다고 해도 '무엇의 평등'이어야 하는지의 부분이다.^[8] 크게 두 개의 노선이 있는데, 하나는 생애 전체에 걸쳐 누리는 복지를 평등하게 해야 한다는 입장이고 다른 하나는 출발점의 자원을 평등하게 한 후 나머지는 개인의 책임으로 남겨두어야 한다는 입장이다. '복지의 평등'이라는 노선에 반대해 드워킨^[9]이 제기한 한 가지 역설은 이러하다. 복지라는 개념에 정치적 선호의 충족까지 포함시키고, 정치적 선호를 복지의 분배에 대한 선호라고 순환적으로 정의해보자. 복지의 평등을 지지하는 사람들로 이루어진 사회에서 이에 반대하는 정치적 선호를 가진 사람은 불만족스러울 텐데, 그러면 이 사람은 자신의 정치적 불만족을 상쇄할 만한 지원을 사회에 요구할 수 있을 것이다. 그런데 사회가 공유하는 가치관에 반대함으로써 오히려 더 많은 것을 얻을 수 있다면 그러한 평등의 개념은 의심스럽다는 것이 드워킨의 주장이다. 사실 여기에서 드워킨은 민주주의 사회가 가지는 본질적인 역설, 즉 의견의 다양성을 반대하는 의견 따위를 어떻게 처리해야 하는가의 고민을 압축적으로 드러내 보이고 있다. 그러나 드워킨의 지적이 복지의 평등에 효과적인 반론이라고 말하기는 어렵다. 자원의 평등과 복지의 평등을 이론적으로 화해시키거나,^[10] 위 역설을 더 진행시킴으로써 오히려 '복지의 평등' 선호가 내쉬 균형일 가능성을 보여주는 연구^[11]도 있다.

N명 사이의 배분을 앞에서처럼 게임 이론으로 바로 다뤄볼 수는 없을까? N명 게임은 맥시민이나 내쉬 균형 같은 해의 개념부터가 의문시되는 매우 복잡한 영역이라, 여기에서 다루기에 지나치게 광범위하다. 다만 연구자들은 내쉬와 루빈슈타인을 따라 N명 사이 협상의 이론을 개발하려 노력해왔고^[12] 개인들이 손잡고 나서는 연합 협상(coalitional bargaining)^[13]의 이론까지도 활발히 연구하고 있다는 점을 언급해두자.

다른 한편, 폴리^[14] 이후로 자원의 배분을 둘러싼 경제학의 흐름이 있다. 한 가지 흥미로운 최근 연구는 디클리펠 등^[15]이 제시한 물음이다. 1 달러를 참가자들의 주장에 기초해 배분한다고 하자. 주장은 편파적일 수 있지만 이를 종합해 공정한 배분에 이르게끔 규칙을 설계하고자 한다. 저자들이 제안하는 첫 번째 규칙은 각자가 자신을 제외하고 다른 사람들이 받아야 할 몫만을 보고하게끔 하는 것이다. 두 번째 규칙은, 한 사람의 몫은 오직 다른 사람들의 보고서에 의해서만 결정된다는 것이다. 이렇게 규칙을 정했을 때 진실하게 평가를 보고하는 것이 거짓말에 비해 불리하지 않은 전략이 된다. 그들은 어떤 배분이 모든 보고와 부합한다면 그것이 답이라고 놓고 4명 이상의 참가자에 대해 이 조건들을 만족할 자연스런 방법들을

제시했다.

또 하나 눈여겨 볼 경제학의 조류는 '타인 고려 선호'^[16]이다. 경제학의 개인은 무자비하게 자신의 선호만을 좇아 이기적 욕망을 추구하는 것처럼 종종 그려지지만, 사실 선호의 개념을 이기주의로 축소할 필요는 없다. 선호란 본래 행동을 통해 측정해야 할 대상이다. 따라서 인간의 행동에서 타인의 상황을 고려하는 요소가 발견된다면 선호 자체에 타인에 대한 부분을 포함할 수 있고, 그렇다면 타인의 삶을 배려하는 자세도 경제학적 합리성과 어긋나지 않는다.

물리학자의 관점에서

물리학자들은 협상 모형으로서 최후통첩 게임 자체의 가치를 찾기보다 '협력의 진화'라는 방향에서 접근하는 경우가 많다. 즉 '죄수의 딜레마'의 대체물로 사용하는 것이다. 따라서 이런 흐름의 연구들에서는 공평한 분배 제안을 협력으로, 그리고 손해에도 불구하고 불공평한 제안을 거부하는 행위를 이타적 처벌로 해석한다. 대개 주제는 공간상에 혹은 어떤 연결망 위에 근시안적인 행위자들을 배치해놓고 적절한 규칙을 통해 제안/수락 전략이 진화해나갈 때에 공정성이 살아남을 수 있는가의 물음이다.^[17] 이런 연구에서 주의해야 할 점은, 진화적 규칙이 내쉬 균형을 찾아간다는 보증만으로는 불충분하다는 것이다. 협력의 진화 문제와 달리 내쉬 균형이 유일하지 않기 때문에, 합리성을 진화로 대체하기 위해서는 그 규칙이 더 좁은 개념의 해, 예컨대 부분 게임 완전균형을 옹계 찾아가는지^[18] 보증해야 한다.

좀 더 근본적인 수준에서 생각해보다면, 물리학자로서는 경

REFERENCES

- [6] S. Hargreaves-Heap and Y. Varoufakis, *Game Theory: A Critical Introduction* (Routledge, NY, 1995).
- [7] E. Kalai, *Econometrica* **45**, 1623 (1977).
- [8] A. Sen, *Equality of What?* (Cambridge Univ. Press, Cambridge, 1980).
- [9] R. Dworkin, *Philos. Public Aff.* **10**, 185 (1981).
- [10] J. Roemer, *Q. J. Econ.* **101**, 751 (1986); J. Moreno-Ternero and J. Roemer, *Econometrica* **74**, 1419 (2006).
- [11] S. K. Baek *et al.*, *PLoS ONE* **7**, e38529 (2012).
- [12] S. Chae and J.-A. Yang, *Econ. Lett.* **28**, 221 (1988); V. Krishna and R. Serrano, *Rev. Econ. Stud.* **63**, 61 (1996).
- [13] F. Gul, *Econometrica* **57**, 81 (1989).
- [14] D. Foley, *Yale Econ. Essays* **7**, 43 (1967).
- [15] De Clippel *et al.*, *J. Econ. Theory* **139**, 176 (2008).
- [16] J. Sobel, *J. Econ. Lit.* **43**, 392 (2005).
- [17] R. Sinatra *et al.*, *J. Stat. Mech.* P09012 (2009); A. Szolnoki *et al.*, *Phys. Rev. Lett.* **109**, 078701 (2012); X. Wang *et al.*, *Sci. Rep.* **4**, 4534 (2014).
- [18] S. Hart, *Games Econ. Behav.* **41**, 227 (2002).

제학에서 가정하는 합리성이 실제 관찰과 다르다는 사실로 인해 이 조건을 포기하는 데 큰 부담감이 없을 수 있다. 또 그것을 무시한 계산으로부터도 실제 현상을 기술할 수 있다면 그것으로 충분하다고 느낄 수도 있다. 비슷한 이유로, 물리학자가 협상 이론을 공부하다가 내쉬의 답안이 사실은 “그래야 한다는 구속력이 있는 것도 아니고 일어날 일의 예측도 아니다”^[19]라는 설명을 만나면 의미를 몰라 당황할 수 있다. 이것은 학문 접근법의 차이를 반영하고 있다. 참고로 ‘tit-for-tat’ 전략의 창안자로 널리 알려진 라포포트는, ‘이론’의 의미를 다음처럼 분류한 바 있다.^[20] 첫째, 일정한 조건으로부터 사건을 예견하는 과정에서 실험을 거쳐 이끌어낸 정리(theorem)의 집합이다. 둘째, 인식의 직관적 조직이다. 셋째, 사회 행위, 제도, 단체, 정치체제, 문화에 대한 직관적 이해이다. 넷째, 규범적으로 당연히 그렇게 되어야 하는 바를 다루는 것이다. 라포포트는 사회과학에서 이론의 목표가 자연과학이 누리는 첫 번째의 수준보다는 낮아야 한다고 주장한다. 그는 심지어 게임 이론의 가장 큰 가치가 “그 자체의 불완전함을 보여주는 능력”이라고까지 말한다. 이는 마치 “경제학적 관점은 일면적이지만 그것이야말로 의도가 되어야 한다”는 막스 베버의 ‘사회과학방법론’과 비슷하다. 베버에 따르면, 사회학의 밑바탕은 ‘사물’ 사이의 관련이 아니라 우리가 문제를 바라보는 ‘사고’ 사이의 연관이고, 이론적 모형은 우리의 사고 틀을 논리의 끝까지 진행시킴으로써 불완전함을 보여주는 데 목적이 있다. 물리학자가 게임 이론을 공부하면서 결국 만나게 되는 가장 어려운 부분은 수학적 논증이라기보다 이렇게 이론을 바라보는 기본 관점의 차이일 것이다.

마치며

우리는 게임 이론의 협상 문제를 중심으로 한정된 재화의 분배를 둘러싼 여러 가지 논점들을 살펴보았다. 상상할 수 있는 거의 모든 사회적 상호작용이 게임 상황으로 묘사될 수 있기 때문에 게임 이론은 오늘날 일종의 ‘통일 이론’ 비슷한 위상을 누리게 되었다. 그러나 다른 한편으로 라포포트처럼 이론의 의미를 소극적으로 해석하면 게임 이론의 연구가 분배 문제를 해결하는 데 직접적 도움이 되지 못하리라는 비관적 전망도 상당 부분 타당하다. 필자는 라포포트의 의견에 기본적으로 동의하지만, 현실 사회 역시 어느 정도 관념의 구성물이기 때문에 게임 이론의 결과들도 그만큼의 현실적 영향력을 가진다고 본다. 예컨대 협상 이론으로 에너지 등분배 정리를 설명하겠다면 터무니없는 일이지만, ‘합리적인 협상이 이 조건들을 만족해야만 한다’고 우리가 믿는다면 그 논리적 귀결은 반드시 이것일 수밖에 없다’는 지적은 사람들로 하여금 자신이 수행하는 협상의 의미에 대해 숙고할 기회를 준다. 이렇게 물리학에서와 다른 ‘게임의 규칙’이 존재한다는 사실을 염두에 둔다면, 모형계의 분석을 위해 물리학에서 개발된 강력한 접근방법들은 게임 이론이라는 틀 안에서도 풍성한 성과를 거둘 수 있을 것이다.

REFERENCES

- [19] M. Davis, *Game Theory: A Nontechnical Introduction* (Dover, NY, 1970).
- [20] A. Rapoport, *Am. Polit. Sci. Rev.* **52**, 972 (1958).