# 動的意思決定と人間行動
## —アレのパラドックス (Allais paradox) の解法—

吉村　仁，伊東　啓，岡部　拓也

期待効用理論は，以前よりアレのパラドックスのために欠陥があると考えられてきた．その後，プロスペクト理論が提案されたことで，この欠陥は人間行動の心理的な問題であると考えられている．吉村らは 2013 年に心理的解釈を必要としない合理的行動理論からパラドックスを説明する解法を提案した．この解法では，ベルマンの最適性原理を一般化して動的意思決定を導き，資産の最大化と破産回避という二つの目的関数を組み合わせることで，アレのパラドックスが資産の中間層の最適化行動として現出することを説明する．この解法から，従来の von Neumann and Morgenstern (1944) の公理体系を基本とした効用理論が本質的に静的モデルであり，動的モデルの効用関数には現資産という状態変数が必然的に入ること，さらに人間には破産回避の行動が本質的に内在していることがわかる．最後に，動的最適化モデルの意義と将来の展望について論じる．

キーワード：アレのパラドックス，動的計画法，動的意思決定，破産回避

## 1. はじめに

意思決定理論 [1] は，1980 年代初頭に動物の採餌行動におけるリスク回避の説明に定性的に使われてきた [2–4]．採餌行動とは動物が餌を採る行動のことである．採餌時間を長くとれば多くの食べ物を得ることができるが，一方で被食リスク（自分が他の動物に食べられて死亡するリスク）も上がる．このことから，利益とリスクのどちらを重視するかという動物行動の研究が数多く行われてきた．動物実験も盛んに行われ，動物には利益よりもリスクに対して敏感に反応するというリスク回避行動があることが実証されてきた [3, 4]．続いて，ベルマンの動的計画法 [5] が動物の生涯行動の最適化問題として本格的に導入された [6, 7]．動的計画法のコンセプトは，連鎖する最適行動をゴール（未来）からスタート（過去）へ向かって各時点における利益が最大化するような戦略を繋げていけば，それが生涯の利益の最大化になるというものであり，動物の生涯の適応度を考えるうえで画期的な試みであった [8, 9]．なぜなら，本来，適応度とはその個体が生涯に残した子供の数であり，動的計画法では直接それを扱えるためである．難点は数値解しか得られないことで，動的計画法が数値計算アルゴリズムであることによるやむ得ない欠点であった．このように，動物生態学では，意思決定理論と動的計画法が，ほぼ同時期に行動適応の説明に使われるようになった [3, 4]．意思決定理論にはアレのパラドックス [10] があり，合理的行動に基づく説明ができなかった．その一方で，動的計画法も数値計算アルゴリズムのために，数値解しか得られないという問題があった．

よしむら じん
静岡大学（名誉教授）
〒 432–8561 浜松市中区城北 3–5–1
yoshimura.jin@shizuoka.ac.jp
いとう ひろむ
長崎大学熱帯医学研究所国際保健学分野
〒 852–8523 長崎県長崎市坂本 1–12–4
ito.hiromu@nagasaki-u.ac.jp
おかべ たくや
静岡大学大学院工学研究科電子物質科学専攻
〒 432–8561 浜松市中区城北 3–5–1
okabe.takuya@shizuoka.ac.jp

Yoshimura et al. [11, 12] では，動的計画法の動的最適化を資産（動物ではエネルギー保存量）の最大化として考え，簡略化により動的最適化理論の基本を導き，アレのパラドックスの合理的解法を導いた．本稿では，2013 年に Journal of Ethology に発表したこの 2 報の論文の概要をまとめて紹介する．まず，動的計画の原理を簡略化して，動物の体サイズ（資源保有量・人間では現資産）のみの長期最適化を定式化し，最終資産の最大化の動的意思決定理論を導出する．これを基本にして資産の最大化と同時に破産の回避の目的関数を組込み複合最適化の目的関数を導出する．得られた結果にもとづき，アレのパラドックスの解法を説明する．最後に，解法の意義や未来の展望では，従来の理論の問題点や，本理論の幅広い応用について言及する．

## 2. 動的意思決定の導出

時間軸上の資産の変動を考えよう．まず，初年度 $t = 0$，最終年度 $t = T$ として，$T$ 年間の資産変動を考える．$t$ 年度の資産を $w_t$ としその年の資産の積

  

算成長率を $r_t$ と表すと次年度の資産 $w_{t+1}$ は $t$ 年度の資産と成長率の積になるので，$w_{t+1} = w_t r_t$ となる ($t = 0 \sim T-1$)．よって，最終資産 $w_T$ は以下の式となる．

$$w_T = w_0 r_0 r_1 \cdots r_{T-1} = w_0 \prod_{j=0}^{T-1} r_j \tag{1}$$

ここで，積算成長率 $r_t$ は独立で同形の分布をもつランダム変数とする．積算成長率 $r$ の確率を $p = p(r)$（$\sum p(r) = 1$ とする）と表すと，幾何平均成長率 $G(r)$ は

$$G(r) = \prod r^{p(r)} \tag{2}$$

となる．最終資産の最大化を考える．

$$w_T \longrightarrow \max \tag{3}$$

資産の最大化は成長率の積算の最大化なので，幾何平均の最大化と同等となる．

$$G(r) \longrightarrow \max \tag{4}$$

さらに対数の最大化とも同等であるので

$$\log G(r) = E\{\log(r)\} \longrightarrow \max \tag{5}$$

つまり，資産の最大化は，対数成長率 $(\log(r))$ の期待値の最大化と同等となる．これを別の形で表すこともできる．すなわち，効用関数

$$u(r) = \log(r) \tag{6}$$

に対して，期待効用の最大化は

$$E\{u(r)\} \longrightarrow \max \tag{7}$$

と表現される．さらにここで，成長率の表現を変更していく．もともと成長率は現資産 $w$ と次年度の資産の違いなので，次年度の獲得資産を $g$ で表すと

$$r = \frac{g+w}{w} \tag{8}$$

で表現できる．式 (6)，(7) の期待効用理論の表現は以下のようになり，動的効用の最大化の原理が導かれる．

$$u(g; w) = \log\left(\frac{g+w}{w}\right) \tag{9}$$

$$E\{u(g; w)\} \longrightarrow \max \tag{10}$$

このように，資産の最大化を目的関数として最大化をすることで動的効用が導かれる．この動的効用関数 $u(g; w)$ には，収入 $g$ の決定変数と現資産 $w$ の状態変数の二つの変数（図 1）があることから，明白に，von Neumann and Morgenstern [1] の期待効用論と異なることがわかる．ここにはベルマンの動的計画法 [5] の状態変数が反映されている．

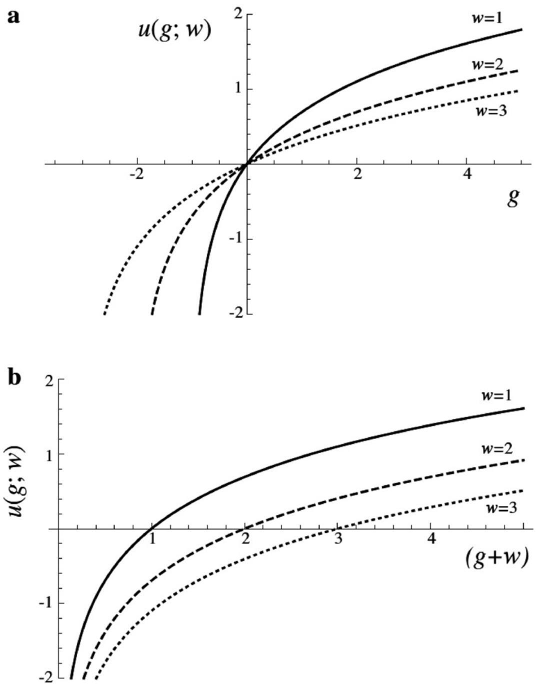


**図 1**　動的効用関数
a. 資産 $w$ を一定にしたときの収入 $g$ の動的効用．b. 意思決定する時点での総資産 $(g+w)$ の評価関数．どちらの場合も状態変数である現資産 $w$ によって異なる関数となる [11]．

## 3. 資産最大化と破産回避の同時最適化

式 (9)，(10) で表した最終資産の最大化では，特に破産回避戦略を考慮に入れていない．そこで破産を避ける戦略として，成長率がマイナスの場合をより重く評価するよう，以下のペナルティ関数を導入する（図 2a）．

$$f(r) = a\left(\frac{1}{r}\right)^b = a\left(\frac{w}{g+w}\right)^b \quad \text{for} \quad -w < g < 0 \tag{11}$$

ここで，ペナルティの重み $a$ と指数 $b$ は定数とする（数値計算では，$a = 2$，$b = 2$ とする）．ペナルティを組込んだ動的効用関数は以下のようになる（図 2b）．

$$u(g; w) = \log\left(\frac{g+w}{w}\right) \quad \text{for} \quad g \geq 0$$
$$u(g; w) = \left(1 + a\left(\frac{w}{g+w}\right)^b\right)\log\left(\frac{g+w}{w}\right) \quad \text{for} \quad -w < g < 0 \tag{12}$$

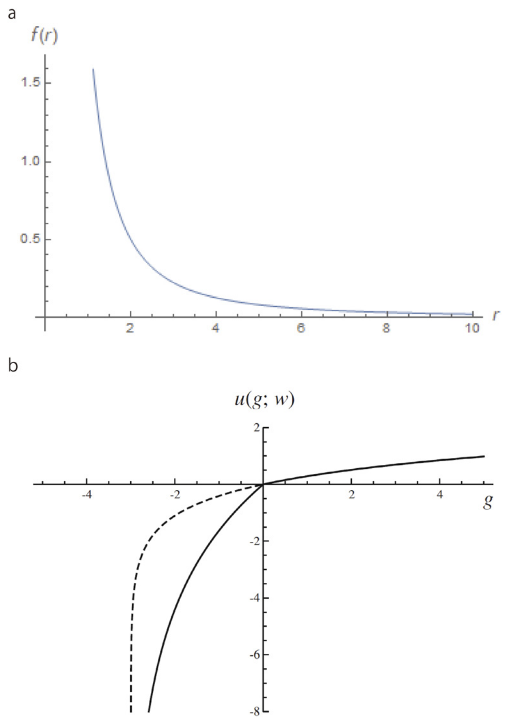


**図 2**　a. ペナルティ関数．b. ペナルティを組込んだ動的効用関数 [12]

上記の式 (12) を アレのパラドックスに使う効用関数とする．

## 4. アレのパラドックス

アレのパラドックスはいろいろな形で提示可能だが，ここでは Machina [13] に沿って，下記の例について数値解析を行う．このパラドックスでは，選択は $(a, b)$ と $(c, d)$ の 2 組あり，それぞれ独立にいずれか一方を選ぶ．

$$
\begin{array}{lll}
a: & \$1\text{million} & \text{with Prob} = 1.00 \\
b: & \$5\text{million} & \text{with Prob} = 0.10 \\
 & \$1\text{million} & \text{with Prob} = 0.89 \\
 & \$0 & \text{with Prob} = 0.01 \\
c: & \$1\text{million} & \text{with Prob} = 0.11 \\
 & \$0 & \text{with Prob} = 0.89 \\
d: & \$5\text{million} & \text{with Prob} = 0.10 \\
 & \$0 & \text{with Prob} = 0.90
\end{array}
$$

パラドックスの本質は，選択 $(c, d)$ では，ほぼ全員が選択肢 $d$ を選ぶ．ところが，選択 $(a, b)$ では，多数の人が選択肢 $a$ を選んでしまう．これに対して，期待効用理論の公理からは，選択肢 $a$ を選んだ人は選択肢 $c$ を選び，選択肢 $b$ を選んだ人は選択肢 $d$ を選ぶと予測される．理由を簡単に説明する．選択 $(a, b)$ の組から確率 0.89 で賞金\$1million を得る期待効用を差し引く．選択 $(c, d)$ では同様に，確率 0.89 の賞金\$0 を差し引く．残りの分の選択 $(ac, bd)$ は以下になる．独立の公理により，以下の選択肢は元の選択肢と同等であることが保証される．

$$
\begin{array}{lll}
ac: & \$1\text{million} & \text{with Prob} = 0.11 \\
bd: & \$5\text{million} & \text{with Prob} = 0.10 \\
 & \$0 & \text{with Prob} = 0.01
\end{array}
$$

このことから，von Neumann and Morgenstern [1] の期待効用論では選択肢 $a$ と $c$ は同等，選択肢 $b$ と $d$ は同等となり，人々はいずれか一方の組を選ぶはずだということになる．実際には，相当数の人々が選択肢 $a$ を選びつつ選択肢 $d$ を選ぶので，期待効用理論では説明がつかない．

## 5. アレのパラドックスの解法

本節では，3 節で導出した式 (12) の破産回避付き動的効用関数を使ってアレのパラドックスの説明をしていく．まず数値解析の設定として，現資産 $w = 10{,}000$ と設定する．まず選択 $(c, d)$ を考えよう．選択肢 $c$, $d$ の期待効用をそれぞれ，$E(u_c)$，$E(u_d)$ と表すと式 (12) に導入して，以下のように計算できる，

$$
E(u_c) = 0.11 \log\left(\frac{1M + w}{w}\right) = 0.507663 \qquad (13)
$$

$$
E(u_d) = 0.10 \log\left(\frac{5M + w}{w}\right) = 0.621661 \qquad (14)
$$

ここで，$M$ は百万の単位で現資産 $w = 10{,}000$ から $E(u_d) > E(u_c)$ となり，選択肢 $c$ よりも選択肢 $d$ が好ましいことがわかる．

つぎに選択 $(a, b)$ を考えよう．選択肢 $a$ は不確定性がなく必ず受け取れるので，意思決定者はすでに受け取ったものと想定できる．この想定を加味して，選択 $(a, b)$ を別の表現に変更することを考えよう．新しく受け取ったことを想定した現資産を $w' = 1M + w$ とすると新しい選択 $(a', b')$ は以下ように書き換わる．

$$
\begin{array}{lll}
a': & \$0 & \text{with Prob} = 1.00 \\
b': & \$4\text{million} & \text{with Prob} = 0.10 \\
 & \$0 & \text{with Prob} = 0.89 \\
 & -\$1\text{million} & \text{with Prob} = 0.01
\end{array}
$$

この選択では，実質的に選択肢 $a'$ は意味がなく，選択

  

| Wealth | $w_1$=0.102400 $w_2$=671350 | | |
|---|---|---|---|
| Prefrerence | a>b c>d | a>b c<d | a<b c<d |
| Wealth Class | poor | middle | rich |

**図 3** 資産レベルとアレのパラドックスの関係

肢 $b'$ をとるかどうかという選択に変わっている．さらに重要なのは，1%の確率で\$1$M$ を失うが 10%の確率で\$4$M$ を得る賭けに挑戦するかである．後でわかるように，この点が破産回避が重要となることを示している．選択肢 $a'$，$b'$ の期待効用を $E(u_{a'})$，$E(u_{b'})$ と表すと式 (12) に導入して，以下のように計算できる．

$$E(u_{a'}) = u(0; w') = 0 \tag{15}$$

$$E(u_{b'}) = 0.10u(4M; w') + 0.89u(0; w') + 0.01u(-1M; w') = -941.463 \tag{16}$$

$E(u_{a'}) > E(u_{b'})$ なので，最適性は選択肢 $a'$ が選択肢 $b'$ より好ましい．つまり，最初の選択 $(a, b)$ では，選択肢 $a$ が選択肢 $b$ より優先される．こうして動的意思決定ではアレのパラドックスが最適化の必然性から起こる．

今回の解析ではパラドックスが起こるかは現資産の数値に依存することがわかる（図 3）．変換点は $w_1 = 0.102400$ と $w_2 = 671{,}350$ である．現資産が $w_1$ より低い場合，つまり貧乏なとき，選択肢 $a$ と $c$ を選びパラドックスは起こらない．また，$w_2$ より高い場合，つまり裕福な場合，選択肢 $b$ と $d$ を選び，これもパラドックスにはならない．つまり，パラドックスは中間層が破産回避するためたまに起こる現象であることが理解できる．もちろん，このアレのパラドックスの証明は定性的なもので，今後定量的にするにはさらなる解析が必要と思われる．

## 6. 解法の意義

動的最適化の基本は，最終時間での資産の最大化（式 (3)）である．時間軸上（式 (1)）の最適化となるため，時系列上の成長率の積算の最適化となる，すなわち，その平均である幾何平均の最大化に一致する．このため，従来の算術平均は使えないが，対数をとることにより，対数成長率の算術平均に変換できる．成長率の変換式（式 (8)）が重要で，これにより，現資産 $w$ の状態変数が必然的に入ってくる．状態変数なしには，最終資産の最大化を規定できないことがわかる．導かれた動的効用では，決定変数は収支（収入）$g$ になるので $x$ 軸は原点ゼロを起点とした単調増加関数となる．さらに状態変数 $w$ を $z$ 軸上にプロットすると曲面となることがわかる．従来の期待効用が全資産の 1 軸しかなかったのは静的モデルだったためで，動的計画とは違い，時間依存性を考えていない．より正確に時間依存性をみるためには動的計画が必要である．

もう一つの重要な要件は破産回避の導入である（式 (11)，(12)）．ここでは負の成長率を加重して破産をより避けるようにしている．実は，これを導入しないとアレのパラドックスは起こらない．つまり，人間はある程度資産が貯まると資産を保全しようとする破産回避の行動を示すことを意味する．破産回避の関数は実際にどのような関数になるかは未知であり，今後の研究に期待したい．

では，従来の von Neumann and Morgenstern [1] の期待効用論の公理体系はどのような数学理論であったのか？ 実は，この体系は，好き（嫌い）の尺度を起点とした公理の上に構築されている．そのため最適化に必須の目的関数が存在しない．つまり，この公理体系は，比較理論であり，最適化理論ではなかったことがわかる．好き嫌いの尺度を効用関数で定義して，平均の効用（期待効用と呼ぶ）が最大であれば最適だと仮定して，公理に沿って数学理論を構築したのだ．最適化理論である限り，最適化の尺度，つまり，目的関数がなければならない．最適化は今回の動的最適化を含め，数理工学の問題であり，純粋な公理体系で表現できると考えるのは無理があると思われる．

アレのパラドックスの存在は，人間の行動は合理的行動 (rational behavior) ではないのではないか，という議論を巻き起こした [14]．今回の解法は，従来の期待効用論が不完全な比較理論であったがためにアレのパラドックスを説明できなかったことを示しており，合理性を排除する必要はないことが理解できる．その点からアレのパラドックスを心理学上の問題として考えたプロスペクト理論 [15] は間違いであることが自明である．半世紀以上解明されずにいたため，心理の問題と考えて，解法を心理に求めたのはやむ得ない歴史であろう．ただ，心理学上の問題はまだほとんど未解明であるが実際に現実に存在しており，プロスペクト理論で提案されている考えをすべて否定できるわけではない．今後の行動の実証研究を通して明らかになると思われる．

## 7. 未来の展望

アレのパラドックスの解法に沿って動的最適化理論を説明してきたが，実は効用理論にはまだ多数の問題が存在する [16]．動的効用関数は実際どのような形な

のか？ 最終資産の最適化（式 (1)–(10)）までは純粋に解析解であるが，破産回避は数値解のための任意の関数（式 (11)）である．本来解析的に求めたいところだが，解析法が不明である．従来から問題になっている保険とギャンブルの問題 [17, 18] を考えると，動的効用関数は式 (10) のような簡単な関数でないことは明確である．これらさまざまな問題の解法が出そろえば，動的最適化理論の全貌が見えてくるかもしれない．

ゲーム理論の展開も気になるところである．これまでの進化ゲームや公共財ゲームなどについても，動的効用を用いた新しい理論展開が必要になってくる．利得表の意味づけにも変更が必要となってくる [19]．たとえば，現資産を組込むとタカハトゲームですら，解には現資産依存性がでてくる [20]．ゲーム理論の本質にかかわるジレンマ問題の本質も動的理論では従来と異なる様相をとる [21]．ナッシュ均衡の概念 [22] には変更は想定されないが，実際の動的ゲームではその計算法も全く未解明である．現代のゲーム理論が，動的効用を基本とした最適化理論の枠組みとした新動的ゲーム理論となるにはまだしばらく時間がかかると思われる．

本稿で扱った動的モデルは従来から関係する意思決定理論を含むオペレーションズ・リサーチ，金融工学，ミクロ経済学の各分野だけでなく，期待値を基準に展開している経営工学，ミクロ経済学にも大きな変革を求めるであろう．動的最適化の確立に本稿が少しでも寄与できればという思いで本稿を上梓した．

**謝辞**　この記事を発表する機会を作っていただいた日本オペレーションズ・リサーチ学会と東京理科大学の朝日弓未教授に心よりお礼を申し上げる．この問題は，吉村が 1989 年にカナダ，ブリティッシュ・コロンビア大学の Colin W. Clark 教授に教わり着手し，1994 年頃に解法をみつけたが，当時はフォン・ノイマンとモルゲンシュテルンの公理体系を絶対視している潮流のために出版を諦めていた．約 20 年の年月を経て，動物行動学会英文誌 J. Ethology の当時の編集長安井行雄氏の招待を受けて，伊東啓らも共著に加わり 2 報セットの論文として 2013 年に公表に至った．Clark 教授は，期待効用理論はアレのパラドックスの存在から本質的に間違っており，動的計画法が唯一の正しい方法だと主張していた．吉村は，期待効用理論の正しさを示すために解法を探したが，解法を実際にみると，Clark 教授の考えが正しかったことがわかった次第である．問題を教えてくださった Clark 教授，論文発表の機会をいただいた安井行雄氏に改めてお礼を申し上げる．

## 参考文献

[1] J. von Neumann and O. Morgenstern, *The Theory of Games and Economic Behavior*, Princeton University Press, 1944.
[2] T. Caraco, “On foraging time allocation in a stochastic environment,” *Ecology*, **61**, pp. 119–128, 1980.
[3] D. W. Stephens and J. R. Krebs, *Foraging Theory*, Princeton University Press, 1986.
[4] L. A. Real and T. Caraco, “Risk and foraging in stochastic environments,” *Annual Review of Ecology, Evolution and Systematics*, **17**, pp. 371–390, 1986.
[5] R. E. Bellman, *Dynamic Programming*, Princeton University Press, 1957.
[6] A. I. Houston and J. McNamara, “A sequential approach to risktaking,” *Animal Behavior*, **13**, pp. 1260–1261, 1982.
[7] M. Mangel and C. W. Clark, “Towards a unified foraging theory,” *Ecology*, **67**, pp. 1127–1138, 1986.
[8] A. Houston, C. Clark, J. McNamara and M. Mangel, “Dynamic models in behavioural and evolutionary ecology,” *Nature*, **332**, pp. 29–34, 1988.
[9] M. Mangel and C. W. Clark, *Dynamic Modeling in Behavioral Ecology*, Princeton University Press, 1988.
[10] M. Allais and O. Hagen (eds.), *Expected Utility Hypotheses and the Allais Paradox*, Reidel, 1979.
[11] J. Yoshimura, H. Ito, D. G. Miller III and K. Tainaka, “Dynamic decision-making in uncertain environments I. The principle of dynamic utility,” *Journal of Ethology*, **31**, pp. 101–105, 2013.
[12] J. Yoshimura, H. Ito, D. G. Miller III and K. Tainaka, “Dynamic decision-making in uncertain environments II. Allais paradox in human behavior,” *Journal of Ethology*, **31**, pp. 107–113, 2013.
[13] M. J. Machina, “Decision-making in the presence of risk,” *Science*, **236**, pp. 537–543, 1987.
[14] R. M. Hogarth and M. W. Reder, “Perspectives from economics and psychology,” *Rational Choice the Contrast between Economics and Psychology*, R. Hogarth and M. Reder (eds.), University of Chicago Press, pp. 1–23, 1986.
[15] D. Kahneman and A. Tversky, “Prospect theory: An analysis of decision under risk,” *Econometrica*, **47**, pp. 263–291, 1979.
[16] M. J. Machina, “Choice under uncertainty: Problems solved and unsolved,” *Journal of Economic Perspective*, **1**, pp. 121–154, 1987.
[17] M. Friedman and L. J. Savage, “The utility analysis of choices involving risk,” *Journal of Political Economy*, **56**, pp. 279–304, 1948.
[18] H. M. Markowitz, “The utility of wealth,” *Journal of Political Economy*, **60**, pp. 151–158, 1952.
[19] H. Ito, Y. Katsumata, E. Hasegawa and J. Yoshimura, “What is true halving in the payoff matrix of game theory?” *PLoS ONE*, **11**, e0159670, 2016.
[20] H. Ito, Y. Katsumata, E. Hasegawa and J. Yoshimura, “The promotion of cooperation by the poor in dynamic chicken games,” *Scientific Reports*, **7**, 43377, 2017.

  

[21] H. Ito and J. Tanimoto, "Dynamic utility: The sixth reciprocity mechanism for the evolution of cooperation," *Royal Society Open Science*, **7**, 200891, 2020.

[22] J. F. Nash Jr, "Equilibrium points in n-person games," *Proceedings of the National Academy of Sciences of the United States of America*, **36**, pp. 48–49, 1950.