Рациональность иррационального выбора: логика дилеммы заключенного

Е. В. Каган, А. С. Рыбалов
Ариэльский университет · Лаборатория LAMBDA, Тель-Авивский университет
`evganyk@ariel.ac.il` · `alxndr_r@yahoo.com`

Аннотация. В работе рассматривается наблюдаемая иррациональность решений в ситуациях, описываемых как одношаговые игры двух игроков. Предлагаемое решение учитывает асимметрию в отношении игроков к собственным выигрышам и выигрышам соперников. Формализация принятия решений основана на некоммутативных операторах многозначной логики. Предложенный метод иллюстрируется примером игры дилемма заключенного. DOI: `10.52119/LPHS.2024.63.70.005`.

Ключевые слова: многозначная логика, некоммутативная алгебра, принятие решений, неопределенность, иррациональные решения.

Введение. Методы логического анализа иррациональных рассуждений основаны на различных вариантах нестандартной логики от логики квантовой механики [1] до вероятностной логики [2], нечеткая логики [3; 4] и теории возможностей [5].

Параллельно с этими работами Ламбек [6] положил начало исследованиям некоммутативных логик, которые были применены для описания структур естественных языков, а затем — для моделирования отношений предпочтения [7; 8]. Полученные результаты позволили описать утверждения, истинность которых зависит от порядка терминов, и смоделировать решения с предпочтениями [9; 10].

В данной работе мы применяем недавно разработанные некоммуникативные логические операторы [11] к известной проблеме принятия решений — дилемме заключенного, и демонстрируем, что учет асимметрии в суждениях заключенного приводит к решению игры.

Постановка задачи. Дилемма заключенного — это игра двух игроков a_1 и a_2 со стратегиями s_1 и s_2. Каждый игрок выбирает стратегию, не зная о стратегии, выбранной другим игроком. Выигрыши игроков определяются таблицей, где $u < v < x < y$:

$a_1 \quad a_2$	s_1	s_2
s_1	(v, v)	(u, y)
s_2	(y, u)	(x, x)

Пусть стратегии суть s_1 — молчать и s_2 — давать показания, и выигрыши — это годы $u = 0, v = 1, x = 2, y = 3$, проведенные в тюрьме. Каждый заключенный встает перед дилеммой: либо молчать (s_1), либо давать показания (s_2).

Оптимальной стратегией для обоих заключенных является молчание (s_1, s_2). Но, поскольку каждый из них не знает о выборе другого, лучшим ответом каждого заключенного будет дать показания (s_2, s_2). Таким образом, равновесие Нэша (s_2, s_2) не является оптимальным.

Дилемма заключенного демонстрирует, что, даже если игрок проинформирован об оптимальных стратегиях, выбор может быть иррациональным из-за влияния неизвестного выбора другого игрока.

Требуется найти способ вычисления рационального или иррационального выбора заключенного относительно заданных выигрышей заключенных. Иными словами, задача состоит в том, чтобы определить метод, демонстрирующий рациональность иррационального выбора заключенного.

Предлагаемое решение. Пусть $\oplus_\theta : [0,1] \times [0,1] \to [0,1]$ — юнинорма [12] с нулем $\theta \in [0,1]$ и $\otimes_\vartheta : [0,1] \times [0,1] \to [0,1]$ — поглощающая норма [13] с единицей $\vartheta \in [0,1]$. Юнинорма $\oplus_1$ является многозначным оператором and, и $\oplus_0$ является многозначным оператором or. Абсорбирующая норма $\otimes_\vartheta$ является многозначной версией логического оператора $not\,xor$. Нормы $\oplus_\theta$ и $\otimes_\vartheta$ вместе с интервалом [0,1] образуют алгебру $A_\eta = \langle [0{,}1], \oplus_\theta, \otimes_\vartheta \rangle$ [14, 15].
Существуют порождающие функции $u_\theta : (0{,}1) \to (-\infty, \infty)$ и $v_\vartheta : (0{,}1) \to (-\infty, \infty)$, такие, что для любых $x{,}y \in (0{,}1)$ [16]

$$x \oplus_\theta y = u_\theta^{-1}(u_\theta(x) + u_\theta(y)), \tag{1}$$

$$x \otimes_\vartheta y = v_\vartheta^{-1}(v_\vartheta(x) \times v_\vartheta(y)). \tag{2}$$

Для граничных значений $x, y \in (0{,}1)$ предполагается, что нормы $\oplus_\theta$ и $\otimes_\vartheta$ являются булевыми операторами.
Некоммутативные юнинорма $\oplus_{\theta_l \vee \theta \vee \theta_r} : [0{,}1] \times [0{,}1] \to [0{,}1]$ и поглощающая норма $\otimes_{\vartheta_l \vee \vartheta \vee \vartheta_r} : [0{,}1] \times [0{,}1] \to [0{,}1]$ определяются следующим образом:

$$x \oplus_{\theta_l \vee \theta \vee \theta_r} y = u_\theta^{-1}(u_{\theta_l}(x) + u_{\theta_r}(y)), \tag{3}$$

$$x \otimes_{\vartheta_l \vee \vartheta \vee \vartheta_r} y = v_\vartheta^{-1}(v_{\vartheta_l}(x) \times v_{\vartheta_r}(y)), \tag{4}$$

где $\theta_l \leq \theta \leq \theta_r$ и $\vartheta_l \leq \vartheta \leq \vartheta_r$. Если $\theta_l \neq \theta_r$ и $\vartheta_l \neq \vartheta_r$, то нормы $\oplus_{\theta_l \vee \theta \vee \theta_r}$ и $\otimes_{\vartheta_l \vee \vartheta \vee \vartheta_r}$ некоммутативны, а если $\theta = \theta_l = \theta_r$ и $\vartheta = \vartheta_l = \vartheta_r$, то эти операторы эквивалентны нормам $\oplus_\theta$ и $\otimes_\vartheta$. Алгебра $A_{l \vee \eta \vee r} = \langle [0{,}1], \oplus_{\theta_l,\theta,\theta_r}, \otimes_{\vartheta_l,\vartheta,\vartheta_r} \rangle$ является расширением алгебры A_η.
Рассмотрим дилемму заключенного в виде биматричной игры [17], где матрицы

$$R^1 = (r^1_{ij})_{2\times 2} \quad \text{и} \quad r^2 = (r^2_{ij})_{2\times 2} \tag{5}$$

представляют собой выигрыши первого и второго игрока.
Пусть

$$r^1_{max} = \max\{|r^1_{ij}|, i, j = 1, 2\} \quad \text{и} \quad r^2_{max} = \max\{|r^2_{ij}|, i, j = 1{,}2\}. \tag{6}$$

Максимальный абсолютный выигрыш в игре равен

$$r_{max} = \max\{r^1_{max}, r^2_{max}\}. \tag{7}$$

Матрицы нормализованых выигрышей имеют вид

$$A^1 = (a^1_{ij})_{2\times 2} \quad \text{и} \quad A^2 = (a^2_{ij})_{2\times 2}, \tag{8}$$

где $(i, j = 1, 2)$

$$a^1_{ij} = r^1_{ij}/r_{max} \quad \text{и} \quad a^2_{ij} = r^2_{ij}/r_{max}. \tag{9}$$

Значения r^1_{max} и r^2_{max} отражают максимальный выигрыш и максимальный проигрыш, из которых обычно исходят суждения, направленные на принятие лучших решений.
Преобразуем выигрыши в значения доверия. Матрицы доверий игроков

$$B^1 = (b^1_{ij})_{2\times 2} \quad \text{и} \quad B^2 = (b^2_{ij})_{2\times 2} \tag{10}$$

включают значения $(i, j = 1, 2)$

$$b^1_{ij} = u_\theta^{-1}(a^1_{ij}) \quad \text{и} \quad b^2_{ij} = u_\theta^{-1}(a^2_{ij}) \quad \text{или} \quad b^1_{ij} = v_\theta^{-1}(a^1_{ij}) \quad \text{и} \quad b^2_{ij} = v_\theta^{-1}(a^2_{ij}), \tag{11}$$

где u_θ^{-1} и v_θ^{-1} — обратные порождающие функции.

Следуя вероятностной интерпретации используемых норм [14], значения b_{ij}^1 и b_{ij}^2, $i,j = 1{,}2$, описывают веру игроков в справедливые выигрыши, что соответствует интерпретации вероятностей Рамсея [18]. Поскольку каждый из заключённых является преступником и знает о преступлении, каждый из них полностью верит в то, что максимальное наказание оправдано, и слабее верит в оправданность меньшего наказания.

Определяя выбор стратегий игроков, мы рассматриваем матрицы убежденностей

$$T^1 = (t_{ij}^1)_{2\times 2} \quad \text{и} \quad T^2 = (t_{ij}^2)_{2\times 2}, \tag{12}$$

определяемые абсорбирующей нормой $(i, j = 1, 2)$

$$t_{ij}^1 = b_{ij}^2 \otimes_{\vartheta_l \vee \vartheta \vee \vartheta_r} b_{ij}^2 \quad \text{и} \quad t_{ij}^2 = b_{ij}^2 \otimes_{\vartheta_l \vee \vartheta \vee \vartheta_r} b_{ij}^1. \tag{13}$$

Здесь предполагается, что игроки выступают в роли оппонентов и реализуют свои отношения «око за око».

Для выбора стратегии убежденности игроков агрегируются с помощью юнинормы. Получаем векторы

$$D^1 = (d_1^1, d_2^1) \quad \text{и} \quad D^2(d_1^2, d_2^2), \tag{14}$$

где

$$d_1^1 = t_{11}^1 \oplus_{\vartheta_l \vee \vartheta \vee \vartheta_r} t_{12}^1 \quad \text{и} \quad d_2^1 = t_{21}^1 \oplus_{\vartheta_l \vee \vartheta \vee \vartheta_r} t_{22}^1, \tag{15}$$

$$d_1^2 = t_{11}^2 \oplus_{\vartheta_l \vee \vartheta \vee \vartheta_r} t_{21}^2 \quad \text{и} \quad d_2^2 = t_{12}^2 \oplus_{\vartheta_l \vee \vartheta \vee \vartheta_r} t_{22}^2. \tag{16}$$

Наконец, стратегия, выбранная игроком, — это стратегия, для которой агрегированные убеждения достигают своего максимума

$$s^1 = \arg\max(d_i^1, i = 1{,}2) \quad \text{и} \quad s^2 = \arg\max(d_i^2, i = 1{,}2). \tag{17}$$

Смысл каждой стратегии определяется формулировкой игры, то есть в дилемме заключенного это молчание или свидетельство.

Рассмотрим дилемму заключенных с выигрышами, заданными матрицами

$$R^1 = \begin{pmatrix} -1 & -3 \\ 0 & -2 \end{pmatrix} \quad \text{и} \quad R^2 = \begin{pmatrix} -1 & 0 \\ -3 & -2 \end{pmatrix}.$$

Максимальный абсолютный выигрыш равен $r_{max} = r_{max}^1 = r_{max}^2 = 3$. Следовательно,

$$A^1 = \begin{pmatrix} -1/3 & -1 \\ 0 & -2/3 \end{pmatrix} \quad \text{и} \quad A^2 = \begin{pmatrix} -1/3 & 0 \\ -1 & -2/2 \end{pmatrix}.$$

Предположим, что $w_\eta = u_\eta = v_\eta$ и $w_\eta(x) = -\ln(x^{-1/\eta} - 1)$, $x \in (0{,}1)$, с параметром $\eta = \theta = \vartheta$. Тогда $w_\eta^{-1} = 1/(1 + \exp(-\xi))^\eta, \xi \in (-\infty, \infty)$. Пусть $\theta_l = \eta_l = \eta/2$ и $\eta_r = (\eta + 1)/2$. Если $\eta = 0.5$, то $\eta_l = 0.25$ (субъективная ложь) и $\eta_r = 0.75$ (субъективная истина) [19]. Тогда

$$B^1 = \begin{pmatrix} 0.42 & 0.27 \\ 0.5 & 0.34 \end{pmatrix} \quad \text{и} \quad B^2 = \begin{pmatrix} 0.42 & 0.5 \\ 0.27 & 0.34 \end{pmatrix}.$$

Согласно этим значениям, каждый игрок почти уверен в том, что будет отбывать 3 года в тюрьме, менее уверен в том, что будет отбывать 2 года, почти не уверен в том, что получит срок в 1 год, и не уверен в том, что немедленно выйдет на свободу.
Далее, применяя поглощающую норму, получим

$$T^1 = \begin{pmatrix} 0.22 & 0.06 \\ 0.41 & 0.24 \end{pmatrix} \quad \text{и} \quad T^2 = \begin{pmatrix} 0.22 & 0.41 \\ 0.06 & 0.24 \end{pmatrix}.$$

Наконец, агрегирование убеждений дает

$$D^1 = (0.02, 0.18) \quad \text{и} \quad D^2 = (0.02, 0.18).$$

В итоге каждый игрок выбирает вторую стратегию:

$$s^1 = 2 \quad \text{и} \quad s^2 = 2,$$

что совпадает с указанным выше равновесием Нэша.
Применение предложенного метода к другим биматричным играм двух лиц, например войне полов [20] и произвольной игре с нулевой суммой, также приводит к предсказанному равновесию.

Заключение. В работе предложен метод принятия решений в условиях неопределенности, который разрешает наблюдаемую иррациональность суждений.
Метод использует асимметрию в отношении игрока к собственному вознаграждению и вознаграждению противника, что формализуется с помощью некоммутативных операторов многозначной логической алгебры.
Метод применяется к одношаговым играм двух игроков, где он успешно предсказывает выбор игроков.
Предложенный метод демонстрирует рациональность иррациональных выборов игроков и может быть использован для объяснения и прогнозирования субъективных решений.

Литература

1. Birkhoff G., Neumann, J. von. The logic of quantum mechanics. *Annals of Mathematics*, Series II, 37.4, 1936, p. 823–843.
2. Nilsson N. J. Probabilistic logic. *Artificial Intelligence* 28, 1986, p. 71–87.
3. Zadeh L. A. Fuzzy sets. *Information and Control* 8, 1965, p. 338–353.
4. Zadeh L. A. Fuzzy logic and approximate reasoning. *Synthese* 30, 1975, p. 407–428.
5. Dubois D., Prade H. *Possibility Theory.* New York, NY: Plenum, 1988.
6. Lambek J. The mathematics of sentence structure. *American Mathematical Monthly* 65, 1958, p. 154–170.
7. Yager R., Rybalov A. Non-commutative self-identity aggregation. *Fuzzy Sets and Systems* 85, 1997, p. 73–82.
8. Ciungu L. *Non-Commutative Multi-Valued Logic Algebras.* Springer, 2014.
9. Fodor J., De Baets B., Perny P. (eds.) *Preferences and Decisions under Incomplete Knowledge.* Springer, 2000.
10. Greco S. et al. (eds.). *Preferences and Decisions. Models and Applications.* Springer, 2010.
11. Kagan E., Novoselsky A., Ramon D., Rybalov A. Non-Commutative logic for collective decision-making with perception bias. *Robotics* 12.76, 2023, p. 1–17.

12. Yager R., Rybalov A. Uninorm aggregation operators. *Fuzzy Sets and Systems* 80, 1996, p. 111–120.
13. Batyrshin I., Kaynak O., Rudas I. Fuzzy modeling based on generalized conjunction operations. *IEEE Transactions on Fuzzy Systems 10*, 2002, p. 678–683.
14. Kagan E., Rybalov A., Siegelmann H., Yager R. Probability-generated aggregators. *International Journal of Intelligent Systems* 28, 2013, p. 709–727.
15. Fodor J., Rudas I., Bede B. Uninorms and absorbing norms with applications to image processing. *Proceedings of the Information Conference SISY, 4th Serbian-Hungarian Joint Symposium on Intelligent Systems, Subotica, Serbia, 29–30 September 2006*. 2006, p. 59–72.
16. Fodor J., Yager R., Rybalov A. Structure of uninorms. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 5, 1997, p. 411–427.
17. Owen G. *Game Theory.* San Diego, CA: Academic Press, 1995.
18. Ramsey F. P. Truth and probability. *The Foundations of Mathematics and Other Logical Essays.* London: Routledge & Kegan Paul, 1926, p. 156–198.
19. Kagan E., Rybalov A., Yager R. Subjective Markov process with fuzzy aggregations. *Proceedings of the 12th International Conference on Agents and Artificial Intelligence ICAART 2020, Valetta, Malta, 22–24 February 2020. Vol. 2.* 2020, pp. 386–394.
20. Luce R. D., Raiffa H. *Games and Decisions.* New York: John Wiley & Sons, 1957.