

# Dilema do Prisioneiro Evolutivo: Uma Análise Inicial

Marcelo Cargnelutti Rossato<sup>1</sup>, João Frederico da Costa Azevedo Meyer<sup>2</sup>  
IMECC/Unicamp, Campinas, SP

**Resumo** A teoria dos jogos possibilita a modelagem matemática de diversas situações com interações entre indivíduos racionais, tendo aplicações nas ciências políticas, econômicas, biológicas e sociais, por exemplo. Assim, este trabalho tem como objetivos apresentar de forma didática e construtiva o dilema do prisioneiro nas suas versões original, iterativa e evolutiva e fazer uma análise de diferentes estratégias que podem ser adotadas nesses cenários. Foi possível concluir que os melhores resultados nos jogos com repetição são geralmente obtidos com estratégias bondosas e retaliatórias, que incentivam a cooperação com os outros jogadores, mas também punem a falta de cooperação alheia. Além disso, notou-se na versão evolutiva que nenhum jogador sozinho consegue prosperar com outra estratégia em um ambiente dominado por jogadores que não cooperam, mas um grupo de tamanho suficiente de jogadores é capaz de alterar a estratégia dominante.

**Palavras-chave.** Teoria dos Jogos, Dilema do Prisioneiro, Jogos Evolutivos, Estratégia Evolutivamente Estável.

## 1 Introdução

A teoria dos jogos é uma ferramenta que permite a modelagem matemática de situações envolvendo interações entre indivíduos racionais, encontrando aplicações em diversas áreas, tais como ciências políticas, econômicas, biológicas e sociais [1, 3, 6, 7].

A inspiração para esse trabalho veio principalmente de [2], onde há simulações e explicações sobre o dilema do prisioneiro, relacionando uma estratégia do jogo com um movimento espontâneo de cooperação não agressiva dos soldados nas trincheiras da Primeira Guerra Mundial denominado “viva e deixe viver”.

Dessa forma, o objetivo deste trabalho é apresentar de maneira didática e construtiva as versões original, iterada e evolutiva do dilema do prisioneiro, bem como analisar algumas estratégias que podem ser adotadas em cada cenário.

## 2 Dilema do Prisioneiro

O dilema do prisioneiro foi proposto inicialmente por Merrill Flood e Melvin Dresher em 1950 [5], sendo um jogo em que cada indivíduo parece ter maiores vantagens traindo o outro jogador ao mesmo tempo que o resultado mais favorável para ambos é obtido por meio da cooperação mútua. A nomenclatura “dilema do prisioneiro” foi dada por Albert Tucker, que também contribuiu para a popularização desse problema com um enunciado semelhante ao seguinte:

*Dois suspeitos são presos, cada um em uma cela, sem qualquer meio de comunicação com o outro. Como não há provas suficientes para condená-los por todos seus crimes, é oferecido o mesmo acordo a ambos com as seguintes condições:*

---

<sup>1</sup>marcelocrossato@gmail.com

<sup>2</sup>jmeyer@unicamp.br

- Se um dos suspeitos confessar os crimes, traíndo seu companheiro, e o outro permanecer em silêncio, o primeiro fica livre e o último é preso por 10 anos.
- Se ambos traírem seus companheiros confessando, os dois são presos por 5 anos.
- Se eles cooperarem um com o outro permanecendo em silêncio, ambos são presos por 6 meses.

Observando todas as possibilidades de forma coletiva, fica claro que a melhor alternativa para os dois é que ambos permaneçam em silêncio, ficando presos por apenas 6 meses cada um. Porém, analisando de forma individual, podemos considerar as duas possibilidades de ação do outro jogador: se o primeiro jogador escolher trair seu companheiro, o segundo jogador ficaria preso por 10 anos ao cooperar ficando em silêncio e apenas 5 anos ao trair, portanto trair é a melhor opção nesse caso. Já se o primeiro jogador escolhesse ficar em silêncio, o segundo jogador ficaria preso por 6 meses ao cooperar e sairia livre se traísse seu companheiro, ou seja, em ambos os casos trair é a melhor opção quando analisada de forma individual.

Daí segue o dilema, uma vez que a estratégia ótima para o coletivo é que ambos cooperem, enquanto a ação mais favorável analisada de forma individual leva à traição mútua. Além disso, ao analisar a situação em que ambos traem, notamos que nenhum dos indivíduos pode obter uma vantagem alterando sua estratégia sem que o outro indivíduo também altere a sua estratégia, então essa situação é denominada um equilíbrio de Nash [4].

Na Tabela 1, é apresentada a forma geral da matriz de ganhos de uma rodada no dilema do prisioneiro, ou seja, cada elemento  $(i, j)$  é a recompensa ou punição que o jogador que faz a ação da linha  $i$  recebe ao competir com alguém que realiza a ação da coluna  $j$ . Embora também seja possível usar os termos “delatar” e “desertar” (e até podem ser mais adequados em algumas situações) como sinônimos da ação de não cooperar, utilizaremos apenas o termo “trair” para manter um padrão.

Tabela 1: Matriz de ganhos do dilema do prisioneiro.

Ação	Cooperar	Trair
Cooperar	$R$	$I$
Trair	$T$	$P$

Para o texto usado como exemplo do dilema, esses valores representam, em módulo, o tempo que eles perderiam de sua vida presos. Assim, a recompensa  $R$  de ambos cooperarem é ambos ficarem presos por apenas meio ano, ou seja,  $R = -\frac{1}{2}$ , enquanto a punição pela traição mútua é  $P = -5$ . Quando um coopera e outro trai, o traidor recebe o valor da tentação  $T = 0$  e o que cooperou recebe o valor  $I = -10$  por ter sido ingênuo.

Para que tenhamos de fato um dilema, devemos ter a relação  $I < P < R < T$  [3], onde a recompensa pela colaboração mútua é maior do que a punição pela traição mútua, mas ainda é menor do que a tentação de se obter uma vantagem traíndo um oponente que cooperou. Outra condição, que será particularmente importante na versão iterada do dilema, é  $T + I < 2R$ , para que o ganho total quando os dois cooperam seja maior do que quando um trai o outro, ou seja, para que a cooperação mútua seja a melhor situação analisando de forma coletiva.

### 3 Dilema do Prisioneiro Iterado

Percebe-se que uma única jogada do dilema do prisioneiro já pode gerar uma grande reflexão sobre qual melhor ação a se realizar (cooperar ou trair), mas o jogo fica ainda mais interessante se forem realizados torneios com vários jogadores jogando algumas rodadas uns contra os outros. Há inúmeras estratégias que podem ser adotadas nesses torneios, a seguir serão apresentadas algumas que foram utilizadas em [2]:

- Cooperador: coopera em todas as rodadas;
- Trapaceiro: trai em todas as rodadas;
- Copiador: inicia cooperando e depois repete a última jogada do oponente. Também é chamado de "olho por olho";
- "Olho por dois olhos": inicia cooperando e trai somente se o oponente traiu consecutivamente nas duas últimas rodadas. Será referido como OP2O por praticidade;
- Rancoroso: coopera até que o oponente traia. A partir disso, trai em todas as rodadas.
- Aleatório: possui 50% de chance de cooperar e 50% de chance de trair;
- Detetive: coopera nas rodadas 1, 3 e 4 e trai na rodada 2. Se o oponente cooperou em todas as 4 primeiras rodadas, ele começa a atuar como o Trapaceiro. Caso contrário, ele passa a atuar como o Copiador;
- Simplista: inicia cooperando e repete sua própria última jogada sempre que o oponente cooperar. Caso o oponente tenha traído na última jogada, o simplista age de forma oposta ao seu último movimento.

Para introduzir o sistema de torneio, realizamos uma competição com 8 jogadores, cada um com uma das estratégias apresentadas. Em cada etapa do torneio, cada jogador joga 10 rodadas consecutivas do dilema do prisioneiro com cada um dos seus 7 oponentes, calculando o ganho obtido em cada rodada por meio da Tabela 1, com  $R = 2$ ,  $I = -1$ ,  $T = 3$  e  $P = 0$ . Cabe salientar que os jogadores “reiniciam” suas estratégias depois de realizar as 10 rodadas contra um oponente. Se um Rancoroso enfrentar um Trapaceiro, por exemplo, ele iria cooperar na primeira rodada e depois ambos trairiam até a décima rodada, mas após isso o Rancoroso voltaria a cooperar até que seu novo oponente, ou até o mesmo oponente em outra etapa, o traísse. A classificação e as pontuações totais obtidas por cada jogador após 1, 10 e 100 etapas são mostradas na Figura 1.

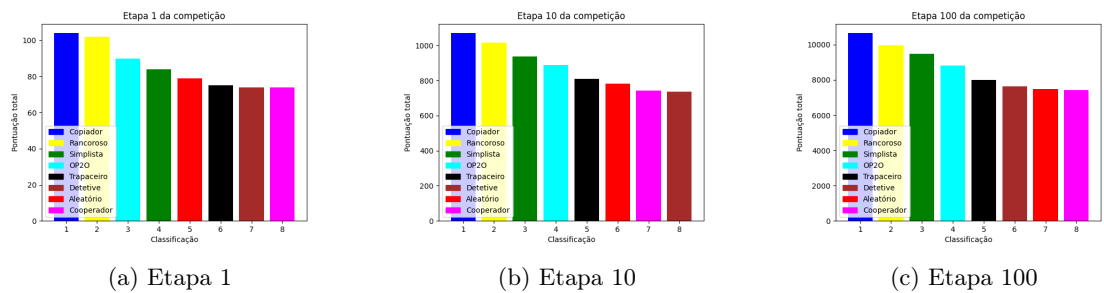


Figura 1: Classificações após 1, 10 e 100 etapas. Fonte: elaborada pelo autor.

Nesse primeiro campeonato, observamos que o Aleatório teve uma boa pontuação na primeira etapa e começou em quinto lugar, mas a ausência de uma estratégia definida fez com que ele terminasse o torneio em sétimo lugar, ganhando somente do Cooperador, cuja bondade excessiva em sempre cooperar fez com que os outros jogadores se aproveitassem dele.

Algumas características foram observadas por [1] em estratégias que se destacaram no Dilema do Prisioneiro Iterado, sendo elas: ser bondosa, ou seja, nunca ser o primeiro jogador a trair; ser retaliatória, isto é, punir de alguma forma as traições do oponente; ser indulgente, ou seja, ter a

capacidade de perdoar e voltar a cooperar mesmo depois de ter sido traído e não ser invejosa, isto é, não tentar obter vantagem sobre seu oponente.

Para o primeiro torneio apresentado em [1], foram submetidas 14 estratégias e todos os jogadores jogavam 200 rodadas contra cada uma das outras estratégias, contra uma cópia de si mesmo e contra um Aleatório, que coopera ou trai com 50% de chance. Os 8 maiores pontuadores foram estratégias bondosas, sendo o Copiador o campeão. Ao observar que as estratégias bondosas estavam se destacando, muitas submissões de estratégias invejosas foram realizadas para o segundo torneio, mas não obtiveram resultados muito bons porque a retaliação dos outros jogadores geralmente fazia com que ambos os jogadores obtivessem menos pontos do que com a simples cooperação.

De forma semelhante, no nosso pequeno torneio, haviam 5 estratégias bondosas: as que eram retaliatórias ficaram nos 4 primeiros lugares e a estratégia Cooperador ficou em último lugar, mostrando a importância de punir as traições dos oponentes. Além disso, o Trapaceiro e o Detetive, que tentam trair para pontuar mais que o oponente, acabaram ficando nas posições 5 e 6, ganhando apenas do Aleatório e do ingênuo Cooperador.

## 4 Dilema do Prisioneiro Evolutivo

Para considerar que os indivíduos podem repensar suas estratégias e se adaptar, introduziremos a ideia de um jogo evolutivo. Assim, o próximo torneio iniciará com 80 jogadores, sendo 10 com cada uma das 8 estratégias. Ao final de cada etapa, os últimos 5 colocados naquela etapa alterarão sua estratégia para as estratégias dos 5 maiores pontuadores da etapa. Na Figura 2 é apresentada a classificação dos jogadores em cada uma das etapas.

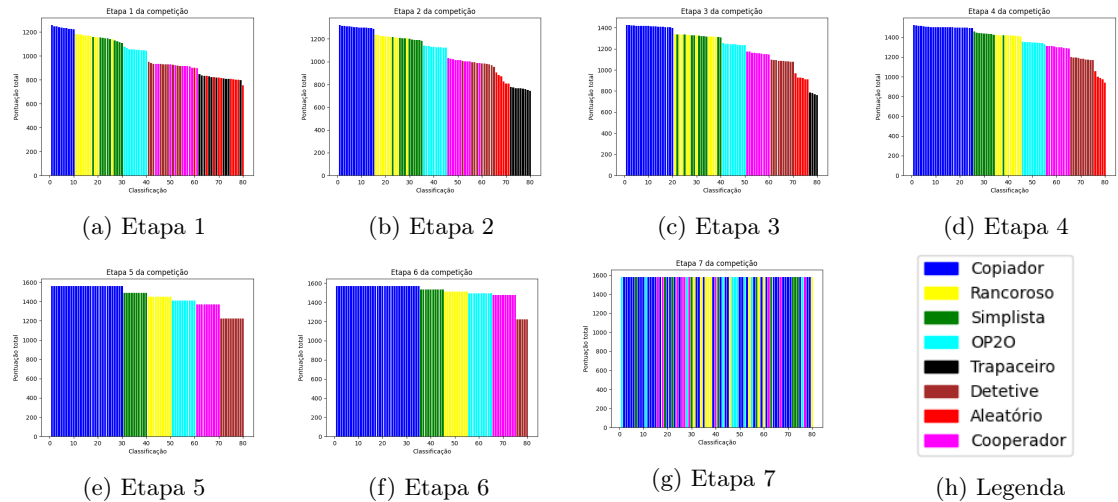


Figura 2: Classificações e estratégias nas 7 primeiras etapas. Fonte: elaborada pelo autor.

Ao analisar a Figura 2, observamos que a estratégia Copiador obteve as maiores pontuações em todas as etapas, enquanto os Trapaceiros e Aleatórios foram os primeiros eliminados, o que mostra que a traição não traz bons resultados em um ambiente com muitas pessoas cooperando e retaliando os traidores. Cabe salientar que todos os jogadores com a mesma estratégia começam a receber as mesmas pontuações quando não há mais Aleatórios.

Além disso, depois que a última estratégia não bondosa dos Detetives é eliminada, todos os jogadores agem de forma igual e obtêm as mesmas pontuações por meio da cooperação mútua em

todas as rodadas. A partir desse ponto, todos vão possuir a mesma pontuação, então a classificação é aleatória, como é observado na Figura 2g. Nesse caso, quaisquer 5 jogadores poderão ser sorteados para substituírem suas estratégias pelas de outros 5 jogadores quaisquer. Assim, mesmo os Copiadores obtendo os maiores resultados nas outras etapas de todos os torneios aqui apresentados, isso não garante que eles conseguirão fazer todos os jogadores adotarem a sua estratégia. Essa aleatoriedade na atualização de estratégias é apresentada na Figura 3.

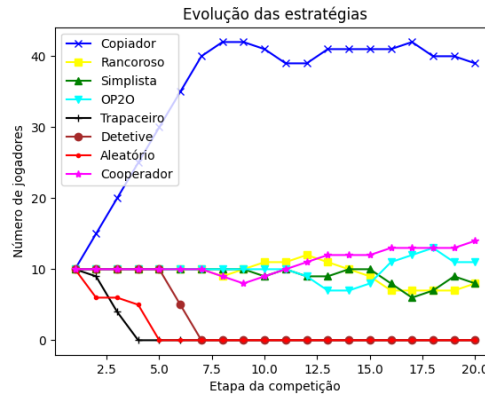


Figura 3: Evolução do número de jogadores por estratégia. Fonte: elaborada pelo autor.

#### 4.1 Estratégias Evolutivamente Estáveis

Com o intuito de definir condições para que os indivíduos possam garantir a dominância de sua estratégia, desde que estejam inicialmente em grande número, introduziremos o conceito de Estratégia Evolutivamente Estável (EEE). Uma EEE é uma estratégia que, se for adotada por toda a população, conseguirá se manter estável por seleção natural mesmo que um jogador “mutante” altere sua estratégia [6].

Seja  $E(T, S)$  a pontuação que um jogador com estratégia  $T$  obtém em uma etapa, que consideramos ser de 10 rodadas, ao jogar contra um jogador com estratégia  $S$ . Definimos que  $S$  é uma EEE se para qualquer estratégia  $T \neq S$  vale que  $E(S, S) > E(T, S)$  ou que  $E(S, S) = E(T, S)$  e  $E(S, T) > E(T, T)$ .

Para entender melhor esse conceito, cada elemento  $(i, j)$  da Tabela 2 é a pontuação que o jogador com a estratégia da linha  $i$  recebe em uma etapa ao competir com um jogador com a estratégia da coluna  $j$ . Nessa análise será desconsiderado o Aleatório, uma vez que seus ganhos não são determinísticos.

Tabela 2: Matriz de ganhos das estratégias por etapa.

Estratégia	Cooperador	Trapaceiro	Copiador	OP2O	Rancoroso	Detetive	Simplista
Cooperador	20	-10	20	20	20	-1	20
Trapaceiro	30	0	3	6	3	9	15
Copiador	20	-1	20	20	20	18	20
OP2O	20	-2	20	20	20	3	20
Rancoroso	20	-1	20	20	20	7	20
Detetive	27	-3	18	15	3	18	7
Simplista	20	-5	20	20	20	11	20

A única EEE é do Trapaceiro, pois  $E(Trapaceiro, Trapaceiro) > E(T, Trapaceiro)$  para qualquer outra estratégia  $T$ . De fato, notamos que qualquer outra estratégia que jogue contra um Trapaceiro em uma etapa recebe uma pontuação negativa enquanto o Trapaceiro recebe uma pontuação positiva quando interage com qualquer uma das outras estratégias. Além disso, na interação entre Trapaceiros, a pontuação obtida é 0. Assim, se considerarmos um ambiente com 100 Trapaceiros no qual um deles altera sua estratégia para  $T$ , teremos, após uma etapa, os 99 Trapaceiros com uma pontuação positiva e o jogador da estratégia  $T$  com uma pontuação negativa. Assim, o jogador “mutante” terá sua estratégia alterada para a do maior pontuador, resultando novamente em uma população inteira de Trapaceiros.

Dessa forma, notamos que se  $E(S, S) > E(T, S)$  ou se  $E(S, S) = E(T, S)$  e  $E(S, T) > E(T, T)$ , então a estratégia  $T$  não consegue invadir e se expandir em uma população de jogadores com estratégia  $S$ . Porém, ao perceber que  $E(Cooperador, Cooperador) < E(Trapaceiro, Cooperador)$  e  $E(Cooperador, Trapaceiro) < E(Trapaceiro, Trapaceiro)$ , podemos fazer uma simulação para descobrir o que acontece quando ocorre uma mutação num ambiente de 100 Cooperadores, resultando em 99 Cooperadores e 1 Trapaceiro. Na Figura 4, observamos que os Trapaceiros conseguem invadir e transformam toda a população de Cooperadores em Trapaceiros. O mesmo ocorre quando surge um Detetive na população de Cooperadores.

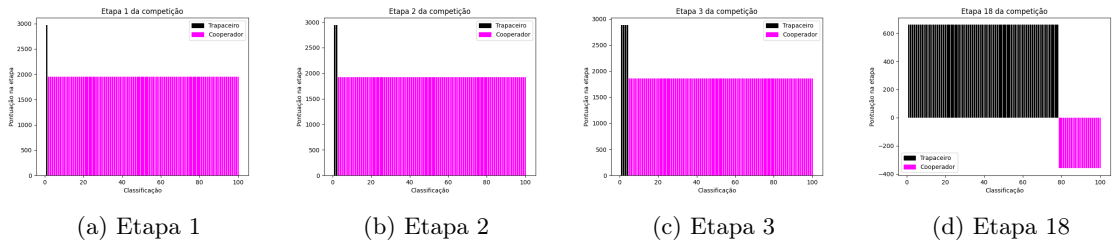


Figura 4: Invasão de Trapaceiros em uma população de Cooperadores. Fonte: elaborada pelo autor.

Outra simulação que poderia ser interessante é a de um Copiador em uma população de 99 Detetives, uma vez que  $E(Detetive, Detetive) = E(Copiador, Detetive)$ , mas  $E(Detetive, Copiador) < E(Copiador, Copiador)$ , ou seja, o Copiador ganha mais pontos do que um Detetive ao jogar contra um Copiador. Na Figura 5, observamos que inicialmente a pontuação é a mesma para todos os jogadores, mas a situação muda assim que o Copiador é sorteado entre os 5 primeiros colocados e surge um outro Copiador. A partir daí, os Copiadores sempre pontuam mais e conseguem transformar toda a população em Copiadores.

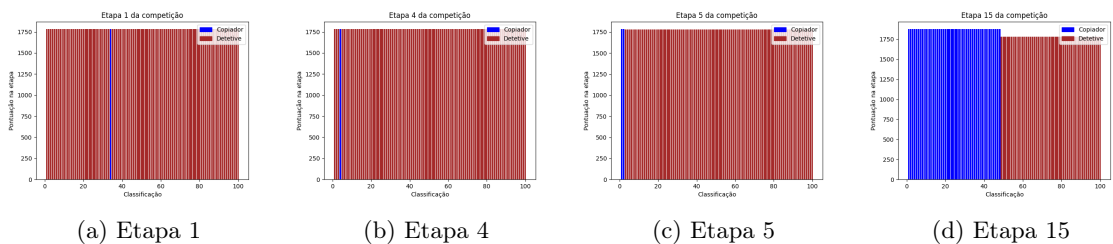


Figura 5: Invasão da estratégia Copiador em um ambiente de Detetives. Fonte: elaborada pelo autor.

Para finalizar, mostraremos que, mesmo o Trapaceiro sendo uma EEE, é possível que uma estratégia  $T$  invada um ambiente com essa estratégia desde que estejam em número suficiente e

tenhamos  $E(\text{Trapaceiro}, T) < E(T, T)$ . Suponhamos que haja uma mutação de  $x$  jogadores para a estratégia  $T$  em um ambiente com 100 Trapaceiros, resultando assim em  $(100 - x)$  Trapaceiros. Interações entre Trapaceiros rendem 0 pontos para ambos, então a pontuação de cada Trapaceiro será  $P(\text{Trapaceiro}) = xE(\text{Trapaceiro}, T)$ , enquanto a pontuação de cada um dos outros  $x$  jogadores será  $P(T) = (x - 1)E(T, T) + (100 - x)E(T, \text{Trapaceiro})$ .

Dessa forma, podemos calcular o menor natural que satisfaz a relação  $P(T) > P(\text{Trapaceiro})$  para obter a quantidade mínima de jogadores  $T$  para garantir a completa invasão do ambiente dominado por Trapaceiros. Realizando os cálculos e confirmando com simulações, chegamos a conclusão de que para invadir um ambiente de Trapaceiros seriam necessários 7 Copiadores ou Rancorosos, 14 OP2Os, 27 Detetives ou 53 Simplistas, ou seja, mesmo em um ambiente dominado pela traição, é possível alterar a estratégia dominante se muitos jogadores cooperarem entre si e retaliarem as traições.

## 5 Considerações Finais

Analisando os resultados das simulações apresentadas nesse trabalho, notamos que o Copiador se destacou positivamente por buscar a cooperação ao mesmo tempo que pune traições de maneira justa, podendo voltar a cooperar se o oponente quiser cooperar também. Essas conclusões nos levam a refletir sobre atitudes cotidianas e a importância de agir de forma cooperativa e justa em interações sociais.

Além disso, o fato de uma estratégia ser uma EEE e ainda assim poder ser invadida por um grupo de jogadores que adotam outra estratégia (como é o caso dos Trapaceiros) possibilita uma reflexão sobre a importância de trabalhar em equipe e de se unir por um objetivo em comum.

A partir desse trabalho, há duas ideias principais para pesquisas futuras. A primeira é considerar uma distribuição espacial dos indivíduos e observar a evolução espacial das estratégias adotadas quando os jogadores competem com os seus vizinhos, podendo então adaptar sua estratégia para corresponder à do vizinho que mais pontuou. A segunda ideia é assumir que há uma probabilidade de erro nas jogadas dos indivíduos, o que pode fazer com que um Cooperador tenha uma pequena chance de trair, por exemplo, e fazer uma análise detalhada de como essas estratégias se sairiam nesse novo sistema, também acrescentando algumas estratégias probabilísticas para a análise e avaliando a importância de as estratégias serem indulgentes.

## Referências

- [1] Robert Axelrod. **The Evolution of Cooperation**. New York: Basic, 1984.
- [2] Nicky Case. **The evolution of trust**. Online. Acessado em 10/03/2023, <https://ncase.me/trust/>.
- [3] Richard Dawkins. **The selfish gene**. Oxford university press, 1976.
- [4] John Nash Jr. “Non-cooperative games”. Em: **Essays on Game Theory**. Edward Elgar Publishing, 1996, pp. 22–33.
- [5] William Poundstone. **Prisoner’s dilemma: John von Neumann, game theory, and the puzzle of the bomb**. Anchor, 1993.
- [6] J Maynard Smith e George R Price. “The logic of animal conflict”. Em: **Nature** 246.5427 (1973), pp. 15–18.
- [7] John Von Neumann e Oskar Morgenstern. “Theory of games and economic behavior”. Em: **Theory of games and economic behavior**. Princeton university press, 1944.