

9. Teoria dos Jogos

Se as pessoas não acreditam que a matemática é simples, é apenas porque não percebem o quão complicada a vida é.
— John von Neumann

Muitos problemas do mundo real não envolvem um único tomador de decisão trabalhando isoladamente, mas sim múltiplos agentes cujos objetivos podem conflitar, alinhar-se ou sobrepor-se parcialmente. Sempre que o resultado de uma situação depende das ações de outros, o pensamento estratégico torna-se essencial. A teoria dos jogos fornece a estrutura matemática para modelar conflito, cooperação e tomada de decisão entre múltiplos agentes — sejam eles indivíduos, empresas, governos ou até sistemas biológicos.

A teoria dos jogos usa otimização, mas não é apenas otimização.

Ao contrário da otimização tradicional, onde um tomador de decisão busca a melhor solução possível para um problema, a teoria dos jogos explora situações em que cada agente otimiza seu próprio objetivo, levando em conta as estratégias dos outros. Essa interdependência leva a resultados que não são simplesmente "melhores" isoladamente, mas sim a equilíbrios, onde nenhum agente pode melhorar seu resultado unilateralmente.

O poder da teoria dos jogos está em sua versatilidade. Ela pode modelar mercados empresariais competitivos, negociações políticas, dinâmicas evolutivas e até escolhas cotidianas, como cooperação ou retaliação. Através de conceitos como o equilíbrio de Nash, estratégias cooperativas e teoria da escolha social, ganhamos ferramentas para entender tanto o comportamento racional quanto, às vezes, paradoxal, em ambientes com múltiplos agentes.

A teoria dos jogos é o estudo matemático das interações estratégicas entre tomadores de decisão, denominados "jogadores". Quer os jogadores sejam indivíduos, firmas ou países, as decisões que tomam impactam não apenas seus próprios resultados, mas também os dos outros. A teoria dos jogos está no cerne da compreensão da competição, cooperação e da dinâmica da tomada de decisão em várias áreas, como economia, biologia, política e além.

Os conceitos-chave da teoria dos jogos são:

- **Jogadores:** Os tomadores de decisão em um jogo. Os jogadores podem representar indivíduos, empresas, países ou até entidades abstratas, como traços genéticos ou estratégias na natureza.
- **Estratégias:** As opções ou ações disponíveis aos jogadores para alcançar seus objetivos. Uma estratégia pode ser simples, como escolher entre cooperar ou desistir, ou complexa, envolvendo uma sequência de movimentos ao longo do tempo.
- **Payoffs (Recompensas/Resultados):** Os resultados ou recompensas que os jogadores recebem com base na combinação de estratégias escolhidas. Os payoffs podem representar lucros, utilidade, recursos ou qualquer medida de sucesso relevante para o contexto.

A essência da teoria dos jogos está em entender como os jogadores tomam decisões com base em estratégias racionais para alcançar um equilíbrio — um estado onde nenhum jogador se sente incentivado a desviar de sua estratégia escolhida. No entanto, a natureza do equilíbrio e da estratégia racional depende muito do jogo e do contexto específicos. Como a "racionalidade" é subjetiva, diferentes jogadores podem perceber o conceito de forma diferente, influenciados por seus objetivos, payoffs ou até pela psicologia.

A seguir enumeramos as principais características que definem um tipo de jogo.

Jogos Cooperativos e Não Cooperativos

Um jogo é cooperativo se os jogadores estão comprometidos com uma estratégia conjunta. Por exemplo, se dois motoristas estão dirigindo em direções opostas em uma rua estreita e veem um buraco à frente, cada um deve desviar para sua direita. Se um deles desviar para a esquerda e o outro para a direita, eles colidirão. Neste caso, os jogadores estão comprometidos com as leis de trânsito. Portanto, se cooperarem, obterão o melhor resultado possível.

Jogos não cooperativos giram em torno de jogadores agindo de forma independente para maximizar seus próprios interesses, sem compromissos previamente combinados de cooperar. A cooperação não é proibida, mas emerge naturalmente apenas quando se alinha com os incentivos individuais dos jogadores. Por exemplo, dois jogadores (A e B) seguram uma moeda cada e ambos revelam suas moedas simultaneamente. Se ambas as moedas mostrarem o mesmo lado (Cara-Cara ou Coroa-Coroa), o Jogador A ganha 1 ponto; se as moedas mostrarem lados diferentes (Cara-Coroa ou Coroa-Cara), o Jogador B ganha 1 ponto. O jogo não é puramente sorte, pois envolve pensamento estratégico, já que os jogadores podem formar expectativas com base em padrões, tais como:

- **Previsibilidade:** Se o Jogador A mostrar consistentemente Cara, o Jogador B pode responder estrategicamente jogando sempre Coroa para maximizar seus próprios pontos.
- **Estratégia Psicológica:** O Jogador B pode antecipar que o Jogador A tentará quebrar seu padrão após mostrar Cara 10 vezes seguidas, introduzindo um elemento de adivinhação e jogos mentais.
- **Estratégia Mista:** Os jogadores podem escolher probabilidades para determinar suas ações (por exemplo, jogar Cara 50% das vezes e Coroa 50% das vezes) para evitar previsibilidade e manter o outro jogador na dúvida.

Este exemplo mostra a tomada de decisão não cooperativa, onde nenhum jogador se compromete com uma estratégia conjunta, focando em vez disso em seus próprios resultados. O equilíbrio está em sua capacidade de antecipar e contra-atacar os movimentos do oponente, tornando-o mais do que apenas sorte aleatória.

Jogos como este ilustram a interação de estratégias racionais, pensamento probabilístico e comportamento psicológico, que são centrais para a teoria dos jogos.

Jogos de Soma Zero e Soma Não Zero

Jogos de soma zero são aqueles em que a riqueza não é gerada nem destruída, ou seja, a soma de todos os ganhos e de todas as perdas é zero, como, por exemplo, em jogos de pôquer ou tênis. Nestes jogos, o que alguns jogadores perdem será ganho pelos outros jogadores. Este jogo é o modelo que corresponde a uma competição pura.

Jogos de soma não zero são aqueles que aumentam ou diminuem a riqueza dos participantes. Por exemplo, em negociações comerciais, ambos os países se beneficiam do aumento do comércio e do crescimento econômico, criando riqueza. A cooperação pode resultar em ganho mútuo em vez de competição pura. Na guerra, os recursos são destruídos, a infraestrutura é danificada e vidas são perdidas. Isso representa uma diminuição na riqueza geral de todas as partes.

Payoffs (Recompensas/Resultados)

Na teoria dos jogos, um payoff representa o resultado ou recompensa que um jogador recebe com base em suas ações e nas ações dos outros. Os payoffs quantificam os benefícios ou custos associados a diferentes escolhas estratégicas em um jogo. Os payoffs podem representar dinheiro, utilidade, pontos ou qualquer benefício mensurável.

Em um jogo de soma zero:

O ganho de um jogador é exatamente igual à perda de outro jogador.

Exemplo: Xadrez, pôquer ou licitação competitiva.

$$U_1 + U_2 = 0$$

- U_1 é o payoff para o Jogador 1
 - U_2 é o payoff para o Jogador 2
- Se o Jogador 1 ganha \$10, o Jogador 2 perde \$10.

Em um jogo de soma não zero:

$$U_1 + U_2 \neq 0$$

A soma de seus payoffs não é zero, o que significa que ambos os jogadores podem ganhar ou perder juntos. Ambos os jogadores podem se beneficiar da cooperação.

Jogos Sequenciais ou Simultâneos

Jogos sequenciais, também chamados de **forma extensiva**, são uma representação de um jogo usando uma estrutura semelhante a uma árvore. Ela ilustra claramente a sequência de decisões, as escolhas disponíveis em cada ponto de decisão e os payoffs associados aos resultados. É particularmente útil para jogos com múltiplos estágios ou movimentos sequenciais.

Por exemplo, dois jogadores, A e B, decidem como dividir \$100. O Jogador A escolhe Dividir ou Manter. Se o Jogador A escolher Dividir, o Jogador B escolhe Aceitar ou Rejeitar.

- Se o Jogador A escolher Manter, A recebe \$100 e B recebe \$0: (100, 0).
- Se o Jogador A escolher Dividir e B Aceitar, A recebe \$50 e B recebe \$50: (50, 50).
- Se o Jogador A escolher Dividir e B Rejeitar, A e B recebem \$0: (0, 0).

Em jogos simultâneos, também chamados de **forma normal**, todos os jogadores tomam suas decisões ou escolhem suas estratégias ao mesmo tempo, sem saber as escolhas dos outros jogadores. Esses jogos destacam a importância de antecipar as ações dos outros e normalmente dependem de pensamento estratégico, probabilidades e análise de payoff. Eles são representados

usando uma **matriz de payoff**. Ela mostra as estratégias dos jogadores e os payoffs para cada combinação de estratégias.

Por exemplo, duas empresas (X e Y) escolhem se anunciam ou não anunciam.

Empresa Y	Anunciar	Não Anunciar
Empresa X: Anunciar	(-10, -10)	(50, -20)
Empresa X: Não Anunciar	(-20, 50)	(30, 30)

- Se ambas as empresas anunciarem, a competição intensa leva a perdas: (-10, -10).
- Se X anuncia e Y não, X ganha 50 enquanto Y perde 20: (50, -20) e vice-versa.
- Se nenhuma anunciar, ambas obtêm lucros moderados: (30, 30).

Informação perfeita ou imperfeita

A informação perfeita existe quando cada jogador conhece todos os movimentos feitos anteriormente por outros jogadores, ou seja, não há informação oculta no jogo. Isso é comum em jogos sequenciais como xadrez, damas ou jogo da velha, onde os jogadores veem o estado completo do jogo em todos os momentos.

Em contraste, jogos de informação imperfeita incluem cenários onde os jogadores carecem de conhecimento completo sobre ações anteriores ou elementos ocultos.

Por exemplo no pôquer, cartas ocultas dos adversários, em leilões em que as valorizações privadas dos licitantes, em decisões estratégicas de negócios, as empresas não sabem completamente as estratégias dos concorrentes e em negociações salariais em que o limite máximo do empregador é desconhecido.

Estas são muito mais comuns em situações do mundo real.

Jogos de informação perfeita se prestam à dedução lógica, enquanto os de informação imperfeita frequentemente envolvem probabilidade, adivinhação ou blefe estratégico.

Estratégias Dominantes e Não Dominantes

Uma estratégia dominante é aquela que fornece a um jogador um payoff melhor, independentemente do que os outros jogadores façam. É a melhor escolha em todos os cenários. Se uma estratégia dominante existe, um jogador racional sempre a escolherá porque ela garante o maior payoff, independentemente das ações do oponente.

Imagine dois vendedores de sorvete concorrentes, Alice e Bob, que devem decidir se vendem sorvete na Praia ou no Parque. Seus lucros (em unidades arbitrárias) dependem de ambas as escolhas:

	Bob: Praia	Bob: Parque
Alice: Praia	(3, 2)	(5, 1)
Alice: Parque	(1, 4)	(2, 3)

O primeiro número em cada célula é o lucro de Alice, o segundo é o lucro de Bob.

Olhe para as escolhas de Alice

- Se Bob for para a Praia, Alice recebe 3 (Praia) ou 1 (Parque). → Melhor ir para a Praia.
- Se Bob for para o Parque, Alice recebe 5 (Praia) ou 2 (Parque). → Melhor ir para a Praia.
- Não importa o que Bob faça, a melhor escolha de Alice é Praia. Isso significa que Praia é uma estratégia dominante para Alice.

Olhe para as escolhas de Bob

- Se Alice for para a Praia, Bob recebe 2 (Praia) ou 1 (Parque). → Melhor ir para a Praia.
- Se Alice for para o Parque, Bob recebe 4 (Praia) ou 3 (Parque). → Melhor ir para a Praia.
- A melhor escolha de Bob também é Praia, não importa o que Alice faça. Então Praia é uma estratégia dominante para Bob.

Como ambos os jogadores têm estratégias dominantes, o resultado do jogo é (Praia, Praia) = (3, 2). Portanto, uma estratégia dominante dá a um jogador o melhor resultado, independentemente do que os outros façam.

Note que neste exemplo, ambos os jogadores por acaso têm a mesma estratégia dominante ("Praia").

Mas isso não é uma regra geral, pois uma estratégia dominante significa apenas a melhor escolha para um jogador, não importa o que os outros façam. Não precisa ser a mesma escolha para todos os jogadores.

Uma estratégia não dominante não fornece consistentemente o maior payoff, dependendo das ações dos outros jogadores. Pode ser ótima em algumas situações, mas não em outras. Os jogadores adotam estratégias não dominantes quando estratégias dominantes não existem, ou quando estratégias mistas fornecem um resultado melhor.

Por exemplo, no jogo "*matching pennies*", dois jogadores escolhem simultaneamente Cara ou Coroa com a seguinte matriz de payoff:

Jogador B	Cara	Coroa
Jogador A: Cara	(1, -1)	(-1, 1)
Jogador A: Coroa	(-1, 1)	(1, -1)

Para o Jogador A (ou B), escolher Cara leva a uma vitória se o Jogador B jogar Cara, mas a uma perda se o Jogador B jogar Coroa. Da mesma forma, escolher Coroa leva a uma vitória contra Coroa, mas a uma perda contra Cara.

Dilema do Prisioneiro

O Dilema do Prisioneiro é um dos exemplos mais icônicos usados na teoria dos jogos e, em particular, de um jogo com uma estratégia dominante.

Dois suspeitos são presos por um crime e interrogados separadamente. Cada um tem duas opções:

- **Cooperar:** Ficar em silêncio e se recusar a trair o outro.
- **Desistir (Trair):** Trair o outro confessando e o incriminando.

A matriz de payoff define as consequências com base em suas escolhas:

- Se ambos cooperarem: Cada um recebe uma punição leve (por exemplo, 1 ano de prisão).
- Se um desistir enquanto o outro coopera: O desistente sai livre, enquanto o cooperador recebe uma sentença pesada (por exemplo, 10 anos).
- Se ambos desistirem: Cada um recebe uma sentença moderada (por exemplo, 5 anos).

Jogador B	Cooperar	Desistir
Jogador A: Cooperar	(-1, -1)	(-10, 0)
Jogador A: Desertar	(0, -10)	(-5, -5)

Desistir sempre dá ao Jogador A um payoff maior comparado a cooperar, independentemente do que o Jogador B faça, porque:

- Se o Jogador B cooperar, desistir dá 0 vs. -1.
- Se o Jogador B desistir, desistir dá -5 vs. -10.

Da mesma forma, desistir também é uma estratégia dominante para o Jogador B.

Ambos os jogadores, agindo inteiramente de forma racional para minimizar sua própria perda, são levados a escolher desistir.

A estratégia dominante do jogo é (desistir, desistir), com um payoff de (-5, -5). O paradoxo surge porque este resultado é significativamente pior para ambos do que se tivessem decidido confiar um no outro e cooperar, onde teriam obtido o payoff de (-1, -1).

O Dilema do Prisioneiro demonstra que a busca individual pelo autointeresse racional pode levar a um resultado coletivamente ineficiente e inferior para todos.

O Dilema do Prisioneiro é um dos exemplos mais icônicos da tragédia do conflito estratégico.

Cada jogador está agindo racionalmente para se proteger. Mas o resultado é dano mútuo — uma tragédia estratégica.

Isso espelha dilemas do mundo real, como:

- Nações correndo para se armar
- Empresas prejudicando umas às outras
- Pessoas acumulando recursos durante crises

Cooperação e Retaliação: Estendendo o Dilema do Prisioneiro

A lógica do Dilema do Prisioneiro também pode aparecer em situações do mundo real envolvendo confiança, repetição e retaliação. Imagine, por exemplo, um acordo comercial repetido entre duas partes — você e um joalheiro. Vocês concordam em trocar dinheiro por diamantes todo mês, secretamente, deixando dois pacotes em um local escondido: um com dinheiro, outro com diamantes. Cada um de vocês chega separadamente — deixando seu pacote e pegando o do outro.

Logo, ambos enfrentam o mesmo problema:

"Posso confiar que a outra pessoa deixará seu pacote?"

Se você deixar o dinheiro e não encontrar diamantes, você perde. O joalheiro enfrenta o mesmo risco.

Logicamente, ambos podem raciocinar:

"Não importa o que o outro faça, estou mais seguro deixando um pacote vazio."

Se ambos pensarem assim, ninguém ganha — mesmo que ambos tivessem lucrado através da cooperação mútua.

Esta é exatamente a dinâmica do Dilema do Prisioneiro:

- Cooperação mútua (C, C): ambos ganham (2, 2).
- Deserção mútua (D, D): nenhum ganha (0, 0).
- Um trai, o outro coopera (C, D): o traidor ganha muito (4), o cooperador perde (-1).

No entanto, a história adiciona um novo elemento crucial: a interação se repete muitas vezes.

Neste jogo iterado, a decisão não é mais um dilema único — torna-se uma questão de estratégia ao longo do tempo.

Agora, a cooperação pode emergir logicamente porque, se ambos os jogadores esperam se encontrar novamente, a confiança e a retaliação começam a moldar o comportamento.

Você pode começar cooperando, mas se o joalheiro trair, você retalia na próxima vez também traindo.

Se a cooperação for retomada, você perdoa e volta a cooperar.

Esta estratégia — conhecida como **Tit-for-Tat** (Olho por Olho) — foi mostrada em simulações de computador como uma das abordagens mais eficazes para o Dilema do Prisioneiro Iterado. É simples, justa e promove benefício mútuo a longo prazo.

Mesmo assim, a cooperação permanece frágil. Se uma das partes souber que a interação está terminando (por exemplo, o joalheiro está prestes a sair ou se aposentar), o incentivo para trair na última rodada ressurgiu. Isso ilustra como o comportamento racional depende do futuro esperado: quando o futuro desaparece, a cooperação colapsa.

O Dilema do Prisioneiro é um conceito clássico na teoria dos jogos que aparece em muitas situações do mundo real. Aqui estão alguns exemplos:

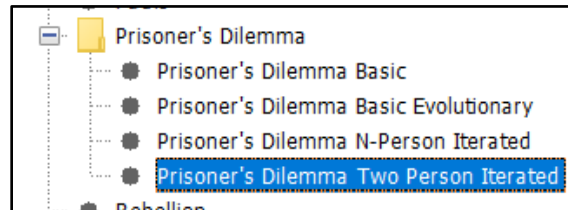
- **Acordos Ambientais:** Países devem decidir se reduzem as emissões de carbono ou continuam poluindo. A cooperação beneficia a todos, mitigando as mudanças climáticas, mas países individuais podem trapacear para evitar custos econômicos enquanto se beneficiam dos esforços dos outros.
- **Corrida Armamentista:** Durante a Guerra Fria, os EUA e a União Soviética enfrentaram o dilema de construir armas nucleares. O desarmamento mútuo teria sido ideal, mas o medo da traição levou ambos a estocar armas, aumentando a tensão global.
- **Greves Trabalhistas:** Trabalhadores e empregadores enfrentam um dilema durante as negociações. Os trabalhadores podem fazer greve por melhores salários, mas arriscam perder renda se a greve falhar. Os empregadores podem atender às demandas ou arriscar perder produtividade durante a greve.

- **Projetos em Grupo:** Na escola ou no trabalho, os indivíduos podem escolher contribuir igualmente ou relaxar. A cooperação leva a um resultado melhor para o grupo, mas os indivíduos podem trapacear, confiando nos esforços dos outros.
- **Cola em Exames:** Os alunos enfrentam o dilema de colar para uma nota melhor ou cooperar estudando honestamente. A honestidade mútua beneficia a todos, mas a cola arrisca punição.

Dilema do Prisioneiro no NetLogo

NetLogo é um poderoso ambiente de modelagem programável usado para simular fenômenos naturais e sociais, que usaremos neste curso, mas explicaremos em detalhes no Volume III.

O Dilema do Prisioneiro Iterado para Duas Pessoas no NetLogo simula o clássico Dilema do Prisioneiro jogado repetidamente entre dois jogadores (agentes) ao longo de múltiplas rodadas. Este tipo de modelo demonstra como estratégias e comportamentos podem evoluir ao longo do tempo à medida que os jogadores interagem repetidamente, revelando dinâmicas como cooperação, deserção, retaliação e construção de confiança.



Dois agentes (humano e computador) interagem repetidamente (jogam uma "rodada") com o Dilema do Prisioneiro. Cada agente tem uma estratégia, decidindo em cada rodada se coopera (C) ou desiste (D).

Os payoffs são determinados pela seguinte matriz. (Os valores absolutos diferem, mas a estrutura de incentivos é idêntica).

- Cooperar-Cooperar: Ambos os jogadores ganham 3 pontos.
- Cooperar-Desertar: O cooperador ganha 0 pontos, e o desistente ganha 5 pontos.
- Desertar-Desertar: Ambos os jogadores ganham 1 ponto.

	partner's action	
your action	C	D
C	3	0
D	5	1

Os jogadores podem usar diferentes estratégias, tais como:

- **Sempre Cooperar:** Sempre escolhe "Cooperar (C)", independentemente das ações do outro jogador.
- **Sempre Desertar:** Sempre escolhe "Desertar (D)".
- **Tit-for-Tat (Olho por Olho):** Começa cooperando e espelha a ação anterior do outro jogador.
- **Aleatório:** Escolhe aleatoriamente entre cooperar e desistir.

O jogo é iterado ao longo de múltiplas rodadas, permitindo que os agentes ajustem suas estratégias dinamicamente com base em interações passadas. Após cada rodada, os payoffs são calculados e acumulados para cada jogador, acompanhando suas pontuações totais ao longo do tempo.

O NetLogo fornece uma plataforma poderosa para simular a dinâmica de interações estratégicas repetidas. No Dilema do Prisioneiro Iterado (DPI) para Duas Pessoas, os agentes repetidamente escolhem cooperar ou desistir, e suas decisões evoluem com base em resultados passados. Aqui estão os principais insights:

- **A interação repetida promove cooperação:** Quando os agentes se enfrentam múltiplas vezes, estratégias como Tit-for-Tat (TFT) podem superar Sempre Desertar. A possibilidade de retaliação ou recompensa incentiva a cooperação mútua.
- **Memória e complexidade da estratégia importam:** Agentes com estratégias de memória-um (que baseiam decisões apenas na última rodada) podem se adaptar efetivamente. Estratégias mais complexas podem superar as simples quando ruído ou incerteza são introduzidos.
- **Fatores ambientais influenciam resultados:** Em simulações onde a probabilidade de encontrar o mesmo agente novamente é baixa, a deserção domina (caso do joalheiro). Mas quando encontros repetidos são prováveis, estratégias cooperativas prosperam.
- **Ruído e erros afetam a confiança:** Introduzir erros de comunicação ou interpretações errôneas pode desestabilizar a cooperação. Estratégias que são perdoadoras (como TFT generosa) tendem a ter melhor desempenho em ambientes ruidosos.
- **Dinâmicas evolutivas revelam dominância estratégica:** Com o tempo, populações de agentes evoluem. Estratégias cooperativas podem dominar se resistirem à exploração e se recuperarem rapidamente de deserções. Estratégias *Grim* (cooperar até ser traído, depois sempre desistir) podem ser eficazes, mas frágeis.

Essas simulações destacam que confiança e memória são cruciais para sustentar a cooperação. O modelo reflete dinâmicas do mundo real em negócios, diplomacia e comportamento social, onde interações repetidas e reputação moldam os resultados.

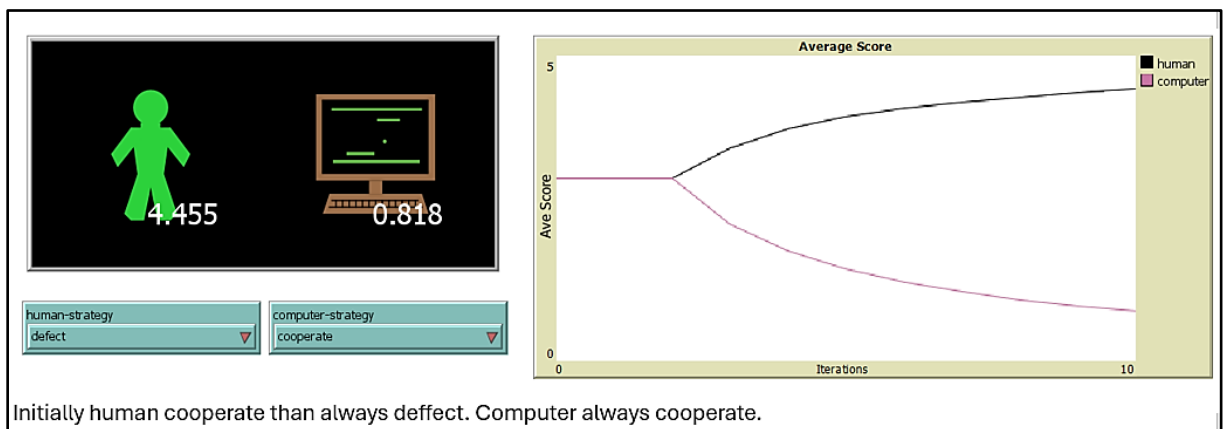
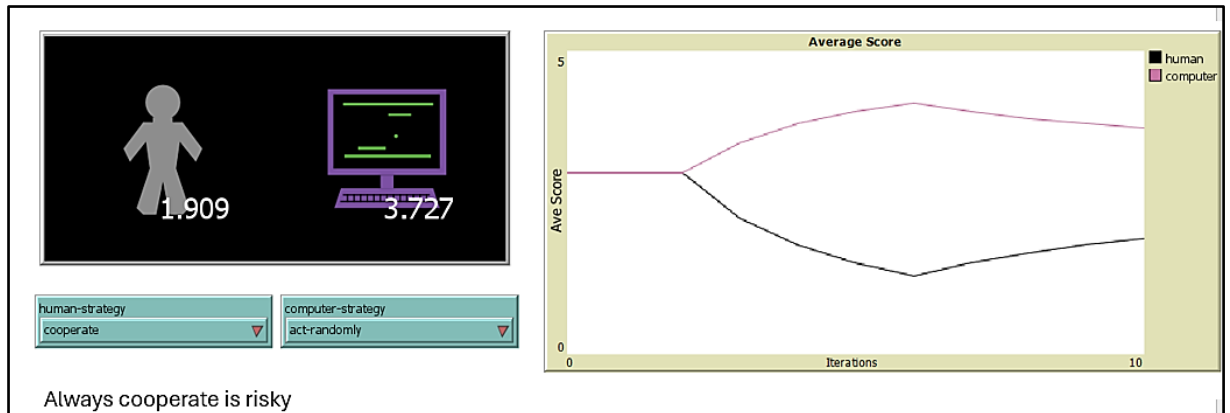
Os jogadores podem começar experimentando a cooperação. Como dissemos, estratégias como Tit-for-Tat incentivam a cooperação mútua se o oponente reciprocara. Se um jogador desistir, o outro pode retaliar na próxima rodada, potencialmente levando a uma deserção prolongada. Ao longo de muitas rodadas, surgem padrões de confiança, traição e perdão, mostrando como a cooperação pode se estabilizar ou colapsar.

Esta simulação ilustra uma lição clássica da teoria dos jogos: A confiança deve ser conquistada e defendida. A cooperação cega em ambientes competitivos pode levar à exploração.

No NetLogo, isso incentiva experimentar com estratégias adaptativas como:

- **Tit-for-Tat (Olho por Olho):** Coopera primeiro, depois espelha o último movimento do oponente.
- **Pavlov (Vitória-Fica, Perda-Muda):** Repete ações bem-sucedidas, muda após perdas.

- **Generous Tit-for-Tat (Olho por Olho Generoso):** Perdoa deserções ocasionais para manter a cooperação em ambientes ruidosos.



Nas figuras acima, vemos que:

- **Primeira simulação:** Humano sempre coopera; computador age aleatoriamente.
 - **Resultado:** Média do Humano $\approx 1,909$; Média do Computador $\approx 3,727$.
- **Segunda simulação:** Humano desiste (após cooperar inicialmente); computador sempre coopera.
 - **Resultado:** Média do Humano $\approx 4,455$; Média do Computador $\approx 0,818$.

Portanto, o DPI no NetLogo demonstra que confiança, memória e interação repetida são essenciais para a cooperação. Ele espelha contextos do mundo real, como negociações comerciais, diplomacia ou redes sociais, onde reputação e relacionamentos de longo prazo moldam o comportamento.

Quando a racionalidade leva a resultados irracionais

Vimos que dois jogadores racionais tentando minimizar suas penas de prisão acabam em uma situação pior do que se tivessem cooperado. O paradoxo é que o interesse próprio racional leva a um resultado coletivamente irracional.

Isso não é apenas um experimento mental — acontece na vida real!

Economistas, biólogos e até cientistas climáticos o usam para explicar por que os países podem falhar em cooperar na redução de emissões, por que os animais às vezes competem excessivamente por recursos e por que a confiança e a reputação são vitais nos mercados.

Mas quando o dilema é repetido (como no Dilema do Prisioneiro Iterado), a cooperação frequentemente emerge naturalmente — **mostrando como o altruísmo pode evoluir a partir da lógica pura.**

John von Neumann, criador da teoria dos jogos, observou que a vida real é justamente esse tipo de jogo — não porque as pessoas são irracionais, mas porque a estrutura dos incentivos às vezes torna a irracionalidade coletiva o único equilíbrio racional individual.

O Teorema Minimax

Desenvolvido por John von Neumann, o Teorema Minimax aplica-se a jogos de soma zero para dois jogadores, onde o ganho de um jogador é exatamente a perda do outro jogador.

Ele afirma que, em tais jogos, existe uma estratégia para cada jogador que minimiza sua possível perda máxima (daí, "minimax"). Esta estratégia garante que nenhum jogador possa se sair melhor desviando-se unilateralmente dela.

Para exemplificar como o jogo funciona, configuramos uma matriz de payoff em alguma moeda (\$), com o objetivo de cada jogador obter o maior payoff possível.

\$3.000,00	\$1.000,00	\$2.000,00
\$6.000,00	\$0,00	-\$3.000,00
-\$5.000,00	-\$1.000,00	\$4.000,00

O jogador da linha (L) joga selecionando uma linha. Simultaneamente, o jogador da coluna (C) seleciona uma coluna. O valor da célula da matriz, que corresponde à linha e coluna selecionadas, é o payoff, que, se positivo, é o valor que L recebe de C. Se este valor for negativo, o jogador L pagará este valor a C. Portanto, o jogador da linha quer obter o valor máximo da matriz, e

o jogador da coluna quer o valor mínimo.

Assim, na matriz abaixo, primeiro L e C escolhem, simultaneamente, um número entre 1, 2 e 3, determinando a linha e a coluna da matriz.

Por exemplo, suponha que L escolheu a linha 3 (L3) e C escolheu a coluna 2 (C2). A posição na matriz é então correspondente a $(L3, C2) = \$ -1.000$. Isso significa que L deve pagar \$1.000 a C. Se a posição fosse $(L3, C3)$, L receberia \$4.000 de C.

A questão então é:

"Existe uma estratégia ótima para L e C, a fim de maximizar o ganho (ou minimizar a perda) dos jogadores? Ou, em outras palavras:

'Existe uma estratégia que torna ambos os jogadores o mais felizes possível (ou o menos insatisfeitos possível)?"

A resposta é afirmativa, como será demonstrado.

Primeiro, considere o raciocínio de L.

Ele vai direto para a posição $(L2, C1)$, onde pode ganhar \$6.000, o que o levaria à linha L2. Mas então ele pensa: "se eu escolher L2, e C tiver escolhido C3, terei que pagar \$3.000", que é o pior que pode acontecer em L2. Então L pensa em $(L3, C3)$, onde ganhará \$4.000. Mas então C pode ter escolhido C1 e L perderá \$5.000, que é o pior que pode acontecer em L3. Finalmente, se L escolher L1, C terá que escolher entre pagar \$1.000, ou pagar \$2.000 ou \$3.000 a L, que são as outras retribuições de L1. Então, a escolha de L é L1.

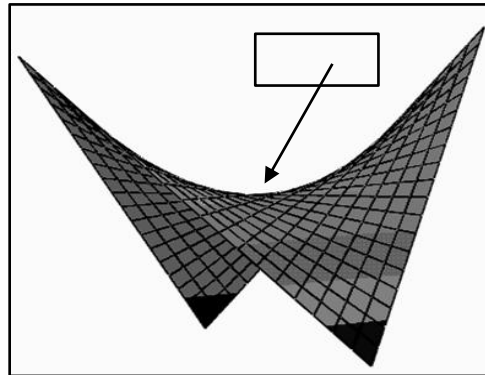
Vimos que o raciocínio de L é baseado no fato de que C pensa exatamente como ele. Também descobrimos que L acabou escolhendo o valor máximo dos mínimos de cada linha, porque os mínimos de cada linha são 1.000, -3.000 e -5.000, respectivamente, então o máximo dos mínimos é 1.000.

Mas esta é a melhor solução para C, que acaba perdendo \$1.000? Vamos ver como a matriz parece para ele. Ele gostaria de cair em $(L3, C1) = \$ -5.000$, mas se escolher C1, L pode ter escolhido $(L2, C1)$ e C perderá \$6.000 (o máximo desta linha). Se ele escolher C3 visando - \$3.000, L pode ter escolhido L3, o que lhe dará uma perda de \$4.000 (o máximo nesta coluna). Finalmente, C se contenta em escolher C2, porque, neste caso, sua perda será um máximo de \$1.000. Neste ponto, C conclui que, do seu ponto de vista, o jogo é injusto, porque o melhor que pode fazer é perder \$1.000. Então, também descobrimos que C escolheu o valor do mínimo dos máximos de cada coluna (6.000, 1.000, 4.000), que é \$1.000.

Como o raciocínio de L e C leva ao mesmo resultado (\$1.000) na matriz de payoff, dizemos que o jogo é determinado ou que tem um valor. Em geral, uma matriz de payoff tem um valor se o máximo dos mínimos de cada linha for igual ao mínimo dos máximos de cada coluna.

$$\begin{array}{cc} \text{Max(min } & \text{Payoff) = Min (max } & \text{Payoff)} \\ \text{Jogador A } & \text{Jogador B} & \text{Jogador B } & \text{Jogador A} \end{array}$$

O valor obtido é conhecido como solução **minimax** ou **maximin** ou mesmo **ponto de equilíbrio** do jogo. O ponto de equilíbrio é às vezes chamado de "**ponto de sela**", como uma analogia topológica ao ponto que é ao mesmo tempo mais alto em uma direção e mais baixo na outra direção, como em uma sela, ou no ponto que é ao mesmo tempo o topo de uma montanha e o fundo de um vale.



Como dissemos, o ponto de equilíbrio de um jogo para duas pessoas é, então, aquele que corresponde à melhor estratégia para os jogadores. Ou seja, considerando estratégias racionais, nenhum dos jogadores pode se sair melhor com qualquer outra estratégia.

Podemos agora resumir o que vimos. Em um jogo de soma zero para dois jogadores:

- Cada jogador tem uma estratégia minimax, onde minimiza a possível perda máxima (ou maximiza o possível ganho mínimo), independentemente da estratégia do oponente.
- O jogo tem um valor específico, representado pelo payoff esperado quando ambos os jogadores usam suas estratégias ótimas.
- O teorema minimax afirma que o valor máximo que um jogador pode alcançar usando sua estratégia ótima é igual ao valor mínimo que seu oponente pode restringi-lo a atingir. Este "valor" é o ponto de equilíbrio do jogo, onde nenhum jogador pode melhorar seu resultado mudando unilateralmente suas estratégias.

Na tabela abaixo, vemos que:

	Coluna 1	Coluna 2	Coluna 3	Coluna 4
Linha 1	-4	3	0	15
Linha 2	5	6	4	5
Linha 3	8	0	2	6

- Mínimos das linhas = (-4, 4, 0); máximo dos mínimos = (4)
 - Máximos das colunas = (8, 6, 4, 15); mínimo dos máximos = (4)
- Portanto, temos um ponto de equilíbrio igual ao valor 4.

Vejamos mais um exemplo.

Qual é a estratégia ótima para ambos os jogadores nesta matriz?

Player A \ Player B	I	II	III	IV	Min
I	-2	0	0	3	-2
II	4	2	1	2	1
III	-4	-3	0	6	-4
IV	5	3	-4	-6	-6
Max	5	3	1	6	

Usamos o princípio minimax para analisar o jogo e descobrimos que 1 é o valor do jogo, o que indica que o jogador A ganhará 1 unidade e o jogador B sacrificará 1 unidade.

Estes são exemplos de **estratégias puras**, ou seja, o jogo no qual há uma estratégia melhor para ambos os jogadores.

Vamos agora ver se o critério minimax nos fornecerá uma solução para o jogo das duas moedas.

Neste jogo, cada jogador tem uma moeda na mão que não é mostrada ao outro jogador. Então, quando os jogadores abrem a mão, se as moedas estão ambas na posição Cara, ou ambas na posição Coroa, L deve um ponto a C. Caso contrário, C deve um ponto a L.

A matriz de payoff é então,

	1	-1
-1	1	

- Mínimos das linhas = (-1, -1); máximo dos mínimos = (-1)
- Máximos das colunas = (1, 1); mínimo dos máximos = (1)

Neste caso, não haverá uma estratégia ótima para um jogador em uma jogada. E também não há ponto de equilíbrio. Qual é a melhor estratégia neste caso?

Estratégias Pura e Mista

Vimos o conceito de estratégia pura. Mas o conceito de equilíbrio evoluiu com a introdução de estratégias mistas nos jogos. Estratégia mista significa uma situação em que um ponto de equilíbrio não existe; o minimax para resolver um problema de jogo se quebra.

Na figura abaixo duas empresas A e B estão competindo pelo mesmo produto. Suas diferentes estratégias são dadas na seguinte matriz de payoff. Primeiro, aplicamos o princípio minimax para analisar o jogo.

Emp. A \ Emp. B	I	II	III	Min
I	-2	14	-2	-2
II	-5	-6	-4	-6
III	-6	20	-8	-8
Max	-2	20	-2	

Vemos que mínimo dos máximos = -2 (linhas) e máximo dos mínimos = -2 (colunas). No entanto, existem dois elementos cujo valor é -2. Portanto, a solução para tal jogo não é única. Nesta situação, ambas as empresas podem recorrer ao que é conhecido como uma estratégia mista.

Em uma estratégia mista, os jogadores atribuem estrategicamente **probabilidades** aos seus movimentos disponíveis e escolhem ações com base nessas probabilidades. Por exemplo, em Pedra-Papel-Tesoura, escolher Pedra, Papel ou Tesoura com probabilidades iguais de 1/3 cada torna impossível para um oponente explorar qualquer viés em suas escolhas. O objetivo de uma estratégia mista é tornar as escolhas do jogador imprevisíveis para seu oponente, impedindo que o oponente explore quaisquer padrões.

O **payoff esperado** refere-se ao payoff médio que um jogador pode esperar receber com base nas probabilidades de vários resultados e nos payoffs correspondentes para cada resultado. É uma medida usada para avaliar e comparar estratégias em jogos onde os resultados são incertos ou dependem das ações de outros. O payoff esperado é calculado como a média ponderada de todos os payoffs possíveis, onde os pesos são as probabilidades dos respectivos resultados. Ele leva em conta a estratégia do jogador, as estratégias de outros jogadores e a estrutura do jogo.

Para 2 jogadores A e B, em que cada jogador tem duas estratégias (S_1, S_2 para A e T_1, T_2 para B), então:

$$E_A = (p_1 \cdot q_1 \cdot \text{Payoff}_A(S_1, T_1)) + (p_1 \cdot q_2 \cdot \text{Payoff}_A(S_1, T_2)) + (p_2 \cdot q_1 \cdot \text{Payoff}_A(S_2, T_1)) + (p_2 \cdot q_2 \cdot \text{Payoff}_A(S_2, T_2))$$

$$E_B = (p_1 \cdot q_1 \cdot \text{Payoff}_B(S_1, T_1)) + (p_1 \cdot q_2 \cdot \text{Payoff}_B(S_1, T_2)) + (p_2 \cdot q_1 \cdot \text{Payoff}_B(S_2, T_1)) + (p_2 \cdot q_2 \cdot \text{Payoff}_B(S_2, T_2))$$

- S_i : Estratégias do Jogador A
- T_j : Estratégias do Jogador B
- p_i : Probabilidade de A jogar S_i
- q_j : Probabilidade de B jogar T_j
- $\text{Payoff}(S_i, T_j)$: Payoff para o jogador para quem estamos calculando

A fórmula diz essencialmente que o payoff esperado do jogador A é a média ponderada de todos os resultados possíveis (payoffs), ponderada pela probabilidade de cada combinação de estratégias ocorrer.

Vamos supor:

- Jogador A tem 2 estratégias (S_1, S_2) com probabilidades $p_1 = 0,4, p_2 = 0,6$.
- Jogador B tem 2 estratégias (T_1, T_2) com probabilidades $q_1 = 0,7, q_2 = 0,3$. Lembrando que em jogos de soma zero, a matriz geralmente mostra payoffs de um jogador (assumindo que o payoff do outro é o negativo), a matriz de payoff para o Jogador A é:

Estrat. A	Jogador B T_1	Jogador B T_2
Jogador A: S_1	3	5
Jogador A: S_2	4	1

Aplicando a fórmula, o payoff esperado E para o Jogador A:

$$E_A = (0,4 \cdot 0,7 \cdot 3) + (0,4 \cdot 0,3 \cdot 5) + (0,6 \cdot 0,7 \cdot 4) + (0,6 \cdot 0,3 \cdot 1) = 3,3$$

O payoff esperado do Jogador A é 3,3.

Como este é um jogo de soma zero, a matriz de payoff do Jogador B é o negativo da do Jogador A, ou seja, -3,3.

Estrat. B	T_1	T_2
S_1	-3	-5
S_2	-4	-1

Como exercício, você pode realizar o cálculo para o jogador B.

Vamos pegar novamente a matriz de payoff para dois jogadores (L e C) jogando matching pennies:

1	-1
-1	1

Para evitar previsibilidade:

- Jogador L randomiza suas escolhas: Joga Cara com probabilidade $x = 0,5$ e joga Coroa com probabilidade $1 - x = 0,5$.
 - Jogador C randomiza suas escolhas: Joga Cara com probabilidade $y = 0,5$ e joga Coroa com probabilidade $1 - y = 0,5$.
- Aplicando a fórmula para o payoff esperado de L ou C,

$$E = 1 \cdot x \cdot y + (-1) \cdot x \cdot (1 - y) + (-1) \cdot (1 - x) \cdot y + 1 \cdot (1 - x) \cdot (1 - y)$$

$$E = (0,25) + (-0,25) + (-0,25) + (0,25) = 0$$

Ambos os jogadores randomizam suas estratégias (50% Cara e 50% Coroa), levando a um payoff esperado de 0 para ambos. Este é o valor ótimo para ambos os jogadores porque nenhum pode melhorar seu resultado esperado mudando unilateralmente sua estratégia. Se o Jogador L se desviar de $x = 0,5$ (por exemplo, jogar Cara com mais frequência), o Jogador C pode explorar isso ajustando suas probabilidades para aumentar suas chances de ganhar. Se o Jogador C se desviar de $y = 0,5$, o Jogador L pode responder de forma semelhante para ganhar uma vantagem. Quando ambos os jogadores mantêm suas estratégias mistas, nenhum pode melhorar seu payoff esperado mudando unilateralmente sua estratégia, criando um **equilíbrio de Nash** (veremos em seguida).

Estratégias mistas são ferramentas poderosas na tomada de decisão estratégica — especialmente quando a previsibilidade é perigosa. Aqui estão alguns exemplos:

Pôquer: Blefe e Equilíbrio

No pôquer, estratégias mistas são essenciais porque:

- Blefar é uma tática randomizada — você às vezes aposta com mãos fracas para manter os oponentes na dúvida.
- O jogo teoricamente ótimo (GTO) envolve misturar entre apostas de valor e blefes em proporções precisas.
- Se você sempre apostar com mãos fortes, os oponentes desistem muito facilmente. Se você blefar com muita frequência, eles pagam sua aposta.
- Estratégias mistas previnem a exploração, tornando seu jogo ilegível e equilibrado.

Cobranças de Pênalti: A Imprevisibilidade é Poder

No futebol, cobranças de pênalti são um jogo clássico de dois jogadores:

- O cobrador escolhe esquerda, centro ou direita.
- O goleiro pula em uma direção.
- Se o cobrador sempre chutar para a direita, o goleiro aprende e se adapta. Então, os cobradores randomizam seus chutes — frequentemente com base em modelos estatísticos ou pistas psicológicas. Estudos mostram que jogadores de elite usam estratégias mistas para manter a imprevisibilidade e maximizar as chances de gol.

Estratégias mistas forçam seu oponente a adivinhar, em vez de explorar. Em qualquer cenário competitivo — esportes, jogos, guerra ou defesa digital — a aleatoriedade é um escudo contra a previsibilidade.