

Capítulo 1

Jogos Evolutivos

As interações humanas em sua maioria são comparável com algum tipo de jogo, mais especificamente com o jogo dilema do prisioneiro. Neste capítulo estudaremos apenas este jogo.

Afim de melhorarmos nossa compreensão do jogo, usaremos uma matriz R que mostra apenas o ganho individual do jogador j_i , a notação C para coopera e N para não coopera

$$R = \begin{matrix} & \begin{matrix} C & N \end{matrix} \\ \begin{matrix} C \\ N \end{matrix} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{matrix}$$

Então, se os indivíduos estão interagindo usando a estratégia $(C, C) = (\text{coopera}, \text{coopera})$, então ambos obtém como recompensa o valor a . Se estão interagindo usando a estratégia $(N, N) = (\text{não coopera}, \text{não coopera})$, então ambos obtém como recompensa o valor d . Se um usa a estratégia $C = (\text{coopera})$ e o outro $N = (\text{não coopera})$ na interação, então o primeiro obtém b como recompensa e segundo, c . E para simplificar a notação, representaremos R apenas por

$$R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Os valores das recompensas devem seguir as seguinte regras $b < d < a < c$ para que o jogo continue sendo um dilema do prisioneiro, ou seja, o jogador j_i que cooperar com

quem não coopera tem a pior recompensa, seguido pela recompensa por não cooperar com quem não coopera, que por sua vez é menor que a recompensa de quem coopera com quem coopera e a melhor recompensa é pra quem não coopera com quem coopera.

Usaremos os seguintes valores de recompensa, que chamaremos convenientemente de pontos.

$$R = \begin{pmatrix} 3 & 0 \\ 5 & 1 \end{pmatrix}$$

Vimos no capítulo anterior que o equilíbrio de Nash para esse jogo é a estratégia pura $\mathbf{e} = (N, N) = (\text{não coopera}, \text{não coopera})$, ambos acusam o outro e recebem 1 ponto por isso.

1.1 Jogos Repetidos

Suponhamos que os jogadores repitam o mesmo jogo e somem os pontos obtidos. Vejamos alguns exemplos onde um jogador j_i que nunca coopera joga 10 jogos com um jogador j_n com jogadas aleatórias.

j_i	Jogadas	N	N	N	N	N	N	N	N	N	N	Total
	Pontos	5	1	1	1	1	5	5	1	5	1	26
j_n	Jogadas	C	N	N	N	N	C	C	N	C	N	Total
	Pontos	0	1	1	1	1	0	0	1	0	1	6

Tabela 1.1: Saldo de pontos do jogador j_i jogando com um jogador j_n com jogadas aleatórias.

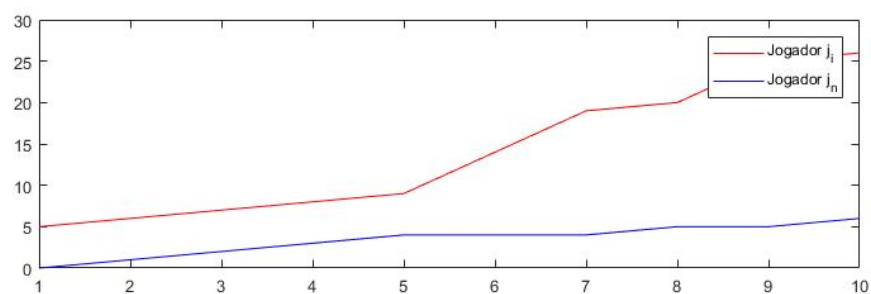


Figura 1.1: Soma de pontos dos jogadores j_i e j_n .

j_i	Jogadas	N	N	N	N	N	N	N	N	N	N	Total
	Pontos	5	1	1	1	1	1	5	1	1	1	
j_n	Jogadas	C	N	N	N	N	N	C	N	N	N	Total
	Pontos	0	1	1	1	1	1	0	1	1	1	

Tabela 1.2: Saldo de pontos do jogador j_i jogando com um jogador j_n com jogadas aleatórias.

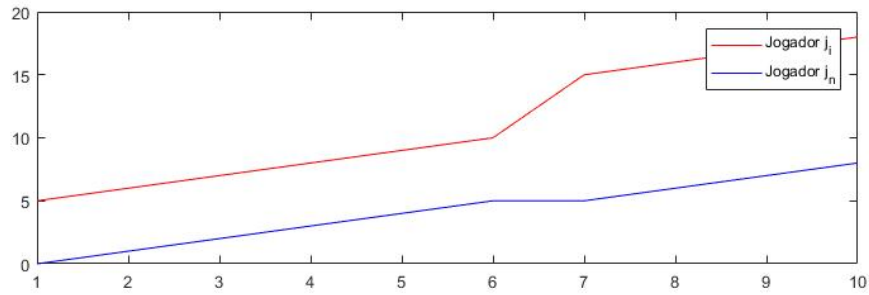


Figura 1.2: Soma de pontos dos jogadores j_i e j_n .

Notemos que, apesar de o jogador j_i sempre vencer, o jogador j_n fica com saldo de pontos mais próximo de j_i quando coopera menos vezes, vejamos agora o mesmo jogador j_i jogando com outro jogador j_n que também não coopera.

j_i	Jogadas	N	N	N	N	N	N	N	N	N	N	Total
	Pontos	1	1	1	1	1	1	1	1	1	1	
j_n	Jogadas	N	N	N	N	N	N	N	N	N	N	Total
	Pontos	1	1	1	1	1	1	1	1	1	1	

Tabela 1.3: Saldo de pontos do jogador j_i jogando com um jogador j_n que nunca coopera.

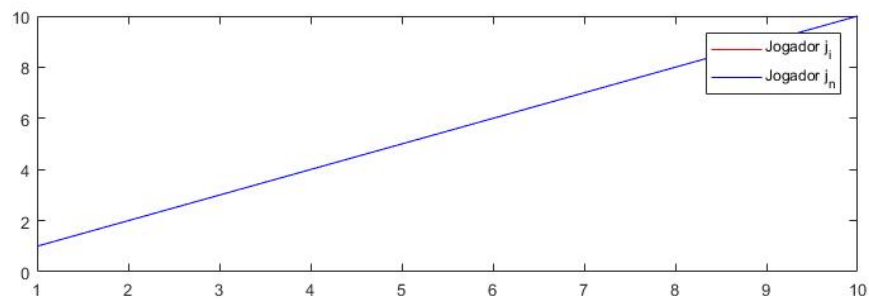


Figura 1.3: Soma de pontos dos jogadores j_i e j_n .

Isso já era esperado porque não cooperar é um equilíbrio de Nash no jogo único. Mas vejamos um jogo curioso, onde dois jogadores cooperam entre si.

j_i	Jogadas	C	C	C	C	C	C	C	C	C	C	Total
	Pontos	3	3	3	3	3	3	3	3	3	3	30
j_n	Jogadas	C	C	C	C	C	C	C	C	C	C	Total
	Pontos	3	3	3	3	3	3	3	3	3	3	30

Tabela 1.4: Saldo de pontos de dois jogadores que sempre cooperam.

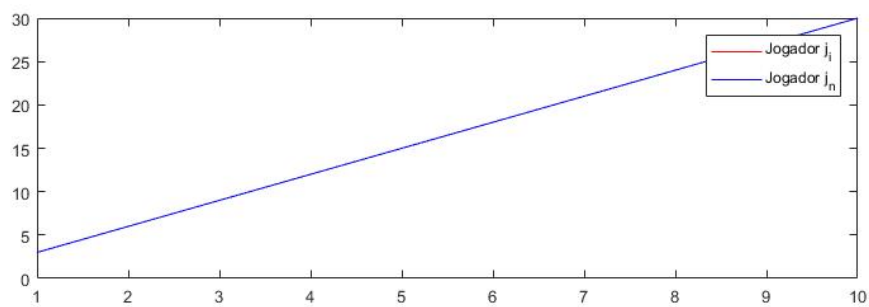


Figura 1.4: Soma de pontos dos jogadores j_i e j_n .

Apesar de os ganhos serem garantidamente maiores quando ambos coopera, assim que um descobre que o outro está cooperando ele é tentado a não cooperar e maximizar sua recompensa, e por mais que um deles insista em continuar cooperando, as perdas o força a não cooperar também.

Uma estratégia que evita essa perda, é a *olho por olho*, onde o jogador começa cooperando e depois repete a jogada do outro jogador no jogo anterior. Essa estratégia deixa o jogador o mais próximo possível do vencedor, mas não o supera, alcança no máximo um empate. Veja alguns exemplos:

j_i	Jogadas	N	C	C	N	N	N	C	N	C	N	Total
	Pontos	5	0	3	5	1	1	0	5	0	5	25
Olho por olho	Jogadas	C	N	C	C	N	N	N	C	N	C	Total
	Pontos	0	5	3	0	1	1	5	0	5	0	20

Tabela 1.5: Saldo de pontos do jogador j_i , com jogadas aleatórias, jogando com um jogador *olho por olho*.

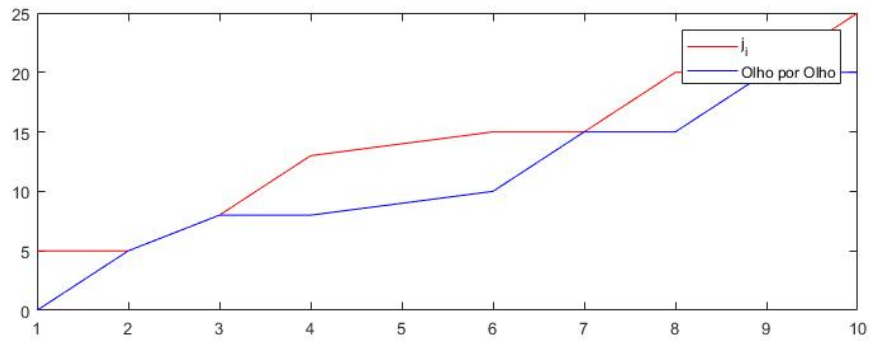


Figura 1.5: Soma de pontos dos jogadores j_i e *olho por olho*.

j_i	Jogadas	N	C	N	C	N	C	N	C	C	C	Total
	Pontos	5	0	5	0	5	0	5	0	3	3	
Olho por olho	Jogadas	C	N	C	N	C	N	C	N	C	C	Total
	Pontos	0	5	0	5	0	5	0	5	3	3	

Tabela 1.6: Saldo de pontos do jogador j_i , com jogadas aleatórias, jogando com um jogador *olho por olho*.

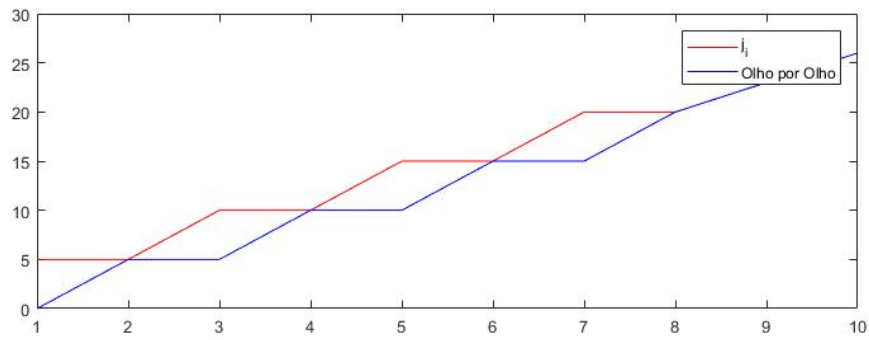


Figura 1.6: Soma de pontos dos jogadores j_i e *olho por olho*.

j_i	Jogadas	N	N	N	N	N	N	N	N	N	N	Total
	Pontos	5	1	1	1	1	1	1	1	1	1	
Olho por olho	Jogadas	C	N	N	N	N	N	N	N	N	N	Total
	Pontos	0	1	1	1	1	1	1	1	1	1	

Tabela 1.7: Saldo de pontos do jogador j_i , com que nunca coopera, jogando com um jogador *olho por olho*.

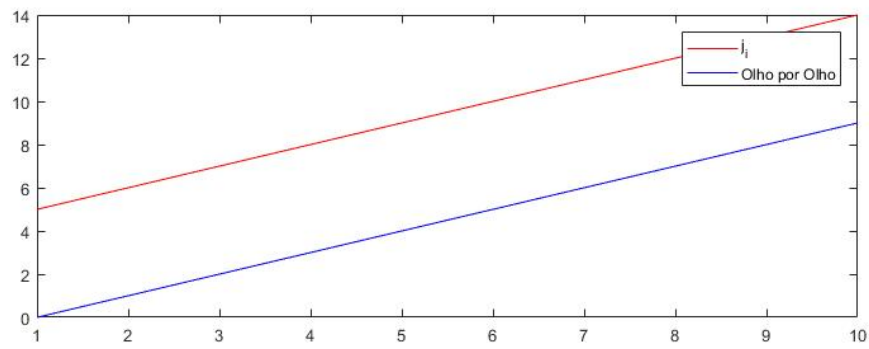


Figura 1.7: Soma de pontos dos jogadores j_i e *olho por olho*.

Olho por olho	Jogadas	C	C	C	C	C	C	C	C	C	Total
	Pontos	3	3	3	3	3	3	3	3	3	
Olho por olho	Jogadas	C	C	C	C	C	C	C	C	C	Total
	Pontos	3	3	3	3	3	3	3	3	3	

Tabela 1.8: Saldo de pontos de dois jogadores *olho por olho* jogando entre si.

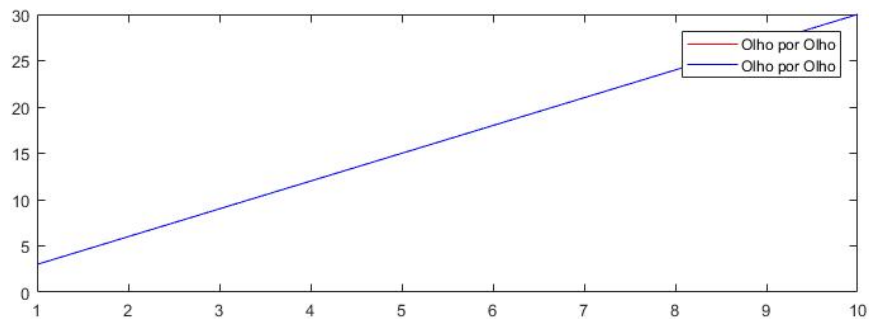


Figura 1.8: Soma de pontos dos jogadores *olho por olho*.

A estratégia *olho por olho* tem uma vantagem sobre a não coopera nunca, que é a de maximizar os ganho quando joga com outro jogador com a mesma estratégia, enquanto o não que coopera nunca, em 10 jogadas ganhou apenas 10 pontos contra um opositor com a mesma estratégia, o olho por olho ganhou 30 pontos.

A limitação da estratégia olho por olho, como veremos a seguir, é que o jogador só pode retaliar o outro jogador na jogada seguinte, se os jogos forem sempre com o mesmo jogador. Numa população de indivíduos onde eles jogam entre si, essa estratégia não é viável.

1.1.1 Coeficiente de Reputação

Em um grupo ou população de jogadores, podemos fazer jogos onde pegamos aleatoriamente dois jogadores e fazemos o jogo entre eles, isso faz com que o próximo jogo de cada jogador, raramente seja com o mesmo jogador anterior, então para guardar um histórico de cada jogador, associaremos a cada um deles um coeficiente $p \in [0, 1]$, que chamaremos de coeficiente de reputação, ou apenas a reputação do jogador j_i . Esse coeficiente será a probabilidade desse indivíduo cooperar no jogo, ou seja, quem tem uma reputação maior, tem maior chance de cooperar.

Os jogadores na hora da partida vão escolher sua estratégia baseada na sua reputação e na do seu oponente. A probabilidade do jogador j_i cooperar vai ser a média aritmética da sua reputação e a reputação do outro jogador. O gráfico 1.9 a seguir, mostra a quantidade de pontos que cada jogador ganha em função da sua reputação, numa simulação¹ onde dois jogadores são selecionados aleatoriamente para jogar 10^5 vezes, de uma população de 100 indivíduos com reputação linearmente distribuída de 0 a 1, e seus pontos são acumulados.

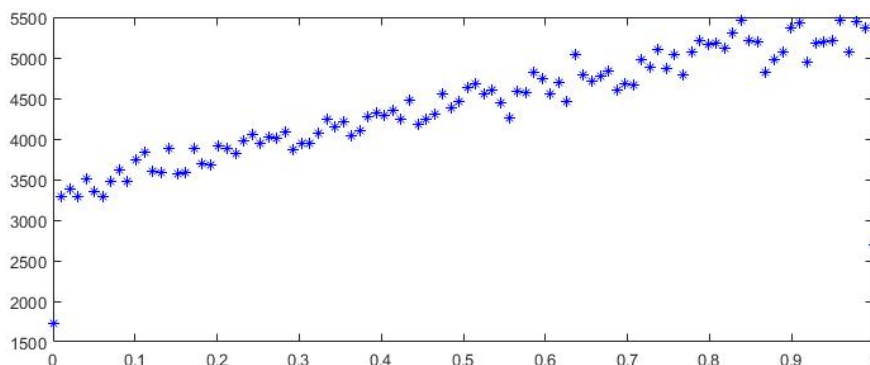


Figura 1.9: Quantidade de pontos em função da reputação dos jogadores.

O jogador com maior reputação acumula mais pontos, e essa diferença fica mais visível quando usamos uma média aritmética ponderada, onde cada jogador dá mais peso para a reputação do seu adversário. Veja os gráficos a seguir onde o jogador dá peso 2 para reputação do adversário e peso 1 pra sua (gráfico 1.10), e peso 3 para seu adversário e 1 para a sua (gráfico 1.11).

¹A simulação foi feita no MATLAB e o algoritmo está anexado ao trabalho, no Apêndice 1

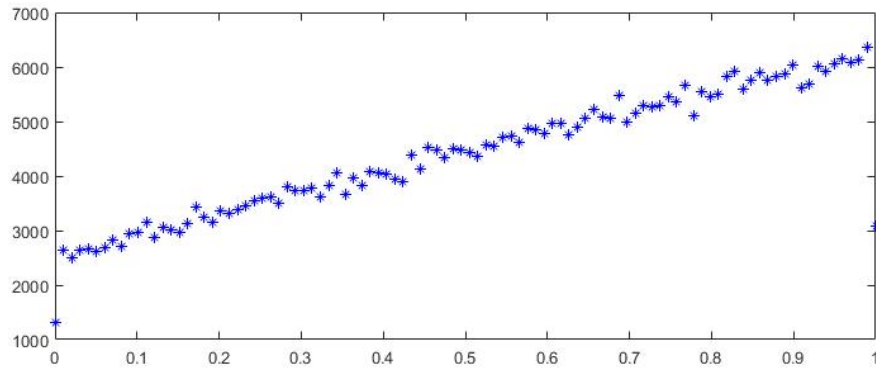


Figura 1.10: Quantidade de pontos em função da reputação dos jogadores.

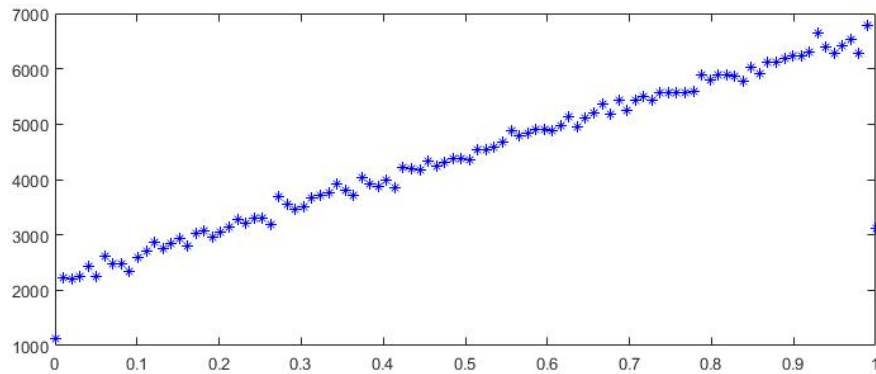


Figura 1.11: Quantidade de pontos em função da reputação dos jogadores.

É possível ver que o indivíduo que tem reputação 0 tem pouco pontos, isso se dá porque os outros jogadores quando feito a média ficam com uma probabilidade baixa de cooperar com ele. O jogador com reputação 1 também fica com pouco pontos, e isso é resultado de ele cooperar sempre.

1.2 Jogos Evolutivos

Suponhamos que os jogadores possa aprender com suas jogadas de jogos anteriores e as jogadas dos outros jogadores, ou seja, que eles possam evoluir suas estratégias. Definiremos uma nova forma de pontuação, a matriz de recompensas continuará a mesma, mas os pontos passarão a ser número de indivíduos gerados na população. Inicialmente vamos dividir a população em dois grupos, o grupo dos que cooperam sempre A e o dos que nunca cooperam B , eles vão jogar entre si, e o resultado de cada jogo, será o numero de indivíduo que aquele

grupo ganhará ou gerará, ou seja, se ambos os indivíduos selecionados forem do grupo A , então o jogo será (C, C) onde cada um ganha 3 pontos que significa mais 6 jogadores para o seu grupo A , se ambos forem do grupo B o jogo será (N, N) e por consequência cada um ganhará 1 ponto que resulta em 2 indivíduos para o grupo B , e por fim se os dois selecionados forem um de cada grupo, o que coopera não ganha ponto portanto seu grupo A não ganha indivíduo e o que não coopera ganha 5 indivíduos para seu grupo B ; e então mediremos qual grupo cresce mais.

O gráfico 1.12 mostra uma simulação² de uma população de 1000 indivíduos em que 50% deles cooperam sempre, grupo A e os outros 50% nunca cooperam, grupo B ; dois indivíduos serão selecionados aleatoriamente para jogar um com o outro, e os indivíduos gerados a cada jogo será somado ao seu respectivo grupo, ao repetirmos esse processo calcularemos o percentual de cada grupo e aplicaremos esse percentual sobre a população inicial, corrigido o percentual de indivíduos de cada grupo pelo resultado do jogo. Chamaremos essas novas população de novas gerações gerações.

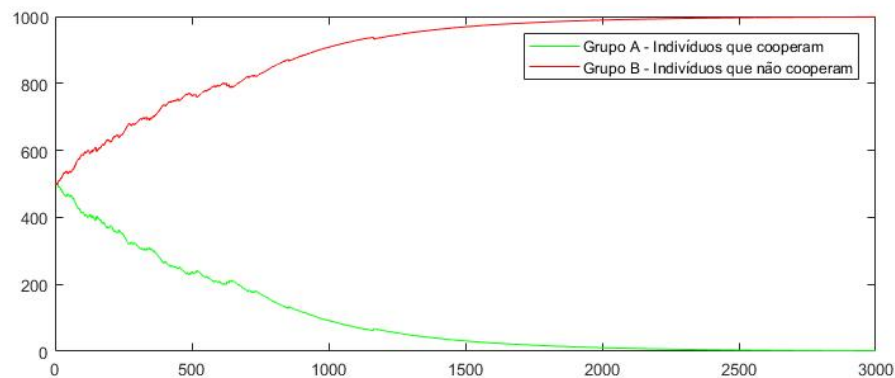


Figura 1.12: Evolução da não cooperação em 3000 gerações.

Não cooperar é a melhor estratégia sempre, esse resultado é obtido independente do percentual inicial da população, como mostra o gráfico 1.13, de uma simulação com uma população inicial com percentual de 90% de indivíduos que cooperam e apenas 10% de não cooperam, e ainda assim a não cooperação prevalece.

²A simulação foi feita no MATLAB e o algoritmo está anexado ao trabalho, no Apêndice 2

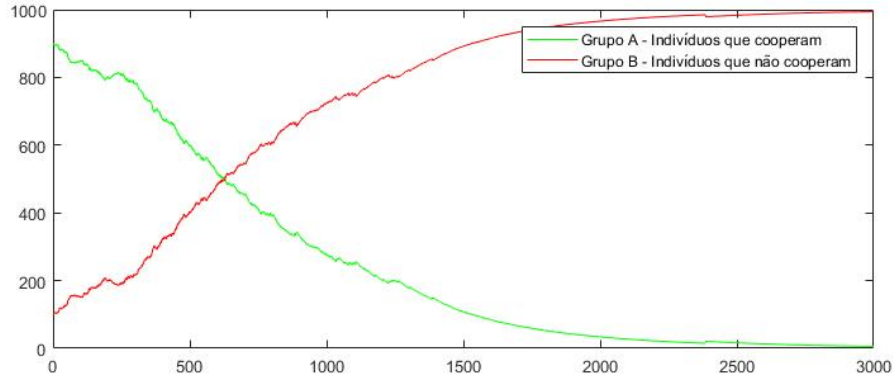


Figura 1.13: Evolução da não cooperação em 3000 gerações.

1.2.1 Evolução da Cooperação

Vamos criar uma população com 100 indivíduos e terá dois coeficientes p e q associados a cada um dele, o coeficiente p será sua reputação, que inicialmente será um número aleatório no intervalo $[0, 1]$ e o coeficiente q , que chamaremos perfil do jogador, será o peso que esse jogador dará a reputação do seu oponente, também no intervalo $[0, 1]$.

Na primeira etapa da simulação³, será ser escolhido aleatoriamente dois jogadores e eles jogarão usando uma probabilidade de cooperar baseada na média aritmética ponderada entre a reputação dele e a do adversário, usando o seu perfil como peso, seja p_i e p_n a reputação do jogador j_i e j_n , e q_i o perfil do jogador j_i , então a probabilidade do jogador j_i cooperar nessa jogada contra o jogador j_n será $(1 - q_i)p_i + q_i p_n$. Vai ser executado 5000 jogos e depois somados seus pontos, através da matriz R e as reputações serão ajustadas obedecendo a seguinte matriz

$$P = \begin{pmatrix} 0.5 & 1 \\ -1 & 0 \end{pmatrix}$$

Onde se o jogador cooperou com outro que cooperou, cada um ganha 0.5 ponto de reputação, se ambos não cooperam um com o outro então ambos não ganham reputação e se um coopera e o outro não, o que cooperou ganha 1 ponto de reputação e o que não cooperou perde 1 ponto.

Na segunda etapa, os perfis da população inicial serão atualizados, os perfis não

³A simulação foi feita no MATLAB e o algoritmo está anexado ao trabalho, no Apêndice 3

serão distribuídos aleatoriamente com distribuição normal entre 0 e 1, e sim obedecendo que mais se aproxima ao gráfico que relaciona pontos com perfil, ou seja, os perfis que geraram mais pontos para seus jogadores, serão mais replicados; como se os jogadores estivessem aprendendo com os que têm melhores resultados. E isso vai se repetir 1000 etapas.

O que podemos observar foi que reputação quando maior a reputação do indivíduo, maior a sua quantidade de pontos, veja o gráfico 1.14 que relaciona a reputação de cada jogador com seus pontos acumulados, e traça uma reta que mais se ajusta aos pontos.

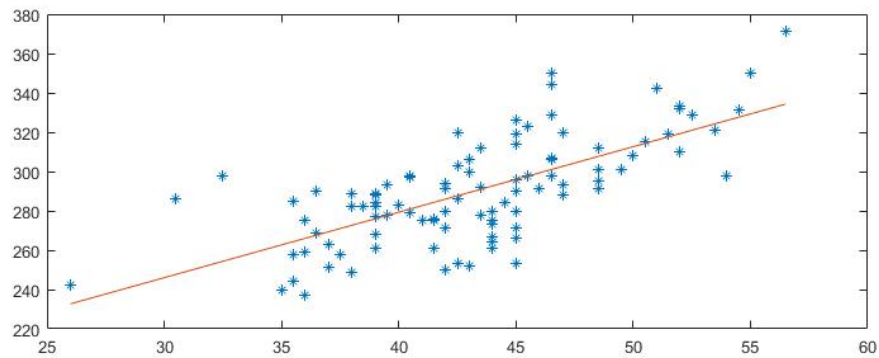


Figura 1.14: Pontos acumulados em relação a reputação e o ajuste dos pontos por uma reta

E a média das somas dos pontos ganhos, pelos dois jogadores, por partida foi se aproximando de 6, que é ganho máximo, ou seja, 3 pontos para cada um, que só é possível quando ambos cooperam. Veja o gráfico 1.15.

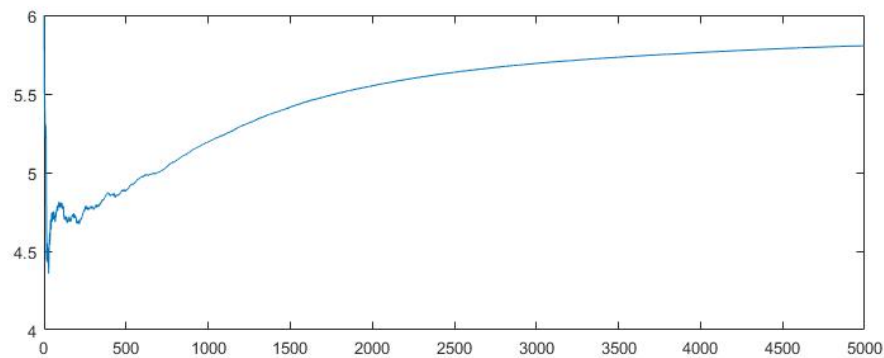


Figura 1.15: Pontos acumulados em relação a reputação e o ajuste dos pontos por uma reta

O gráfico 1.16 mostra que a maioria da população migrou para um perfil próximo de 1, que é significa dar peso máximo a reputação do oponente na hora de escolher a jogada.

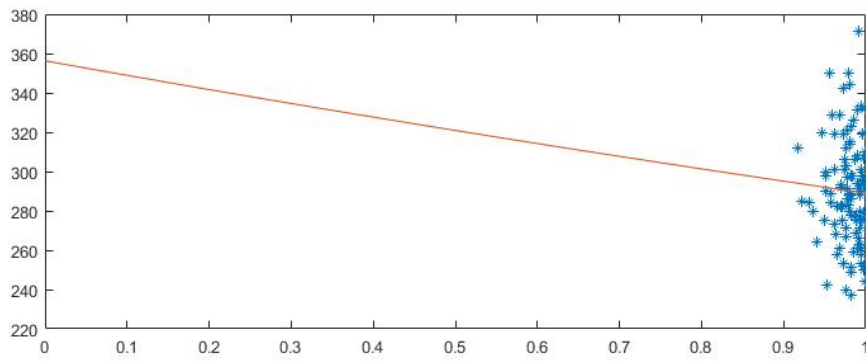


Figura 1.16: Pontos acumulados em relação ao perfil e o ajuste dos pontos por uma reta

Esses resultados foram da última população, a mais aperfeiçoada. Agora veremos abaixo dois gráficos que mostram a evolução da cooperação e do perfil ao longo de todas as populações. No gráfico 1.17 temos o ganho médio por jogo em cada simulação, nas primeiras populações era próximo de 5, e o 5 só pode ser obtido pela matriz de pontos somando $5+0$, ou seja, um cooperando e outro não, mas nas etapas seguintes esse número foi subindo e tendendo a 6, que é obtido por $3+3$, ou seja, quando ambos cooperam. No gráfico 1.18 o perfil médio da população tendendo a um, ou seja, cada jogador dando prioridade máxima a reputação do seu oponente.

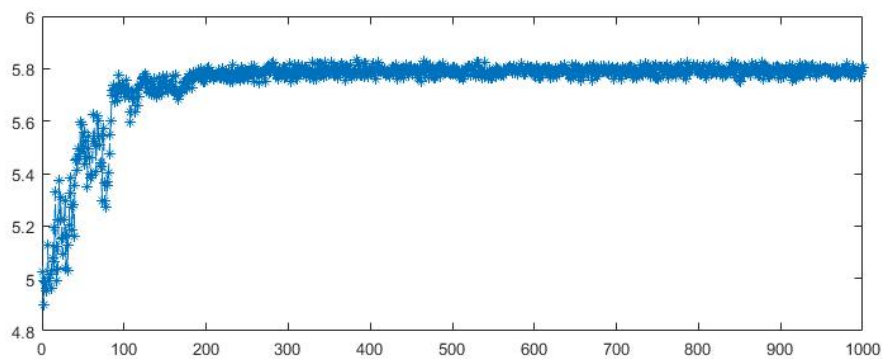


Figura 1.17: Ganho médio por jogada em cada simulação.

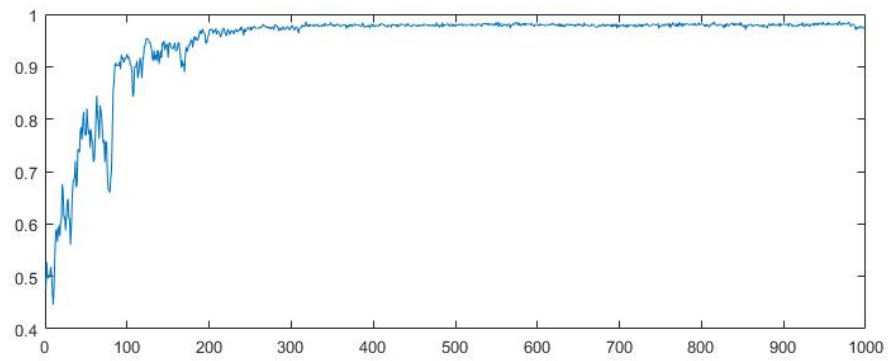


Figura 1.18: Perfil médio por simulação.

1.3 Aplicabilidade

Esse modelo mostra como é possível forçar a cooperação através da reputação. Essa estratégia já é usada desde os primórdios da civilização, sempre que alguém vai negociar com alguma pessoa, primeiro ela se informa pra saber se a pessoa é confiável ou não, isso nada mais é que dar prioridade a reputação da outra pessoa.

Com a globalização e o comércio pela internet, surgiu a necessidade de pensar em métodos de fazer as pessoas cooperarem mesmo sem conhecer seu consumidor ou fornecedor, e a solução foi criar mecanismo que expõe a reputação do indivíduo para consulta. Isso viabilizou e alavancou o comércio digital. Empresas como Uber, Mercado Livre, Submarino e outras empresas de vendas virtuais, controlam seus negócios e mantêm um nível de cooperação reduzindo os golpes através da reputação.