

UFMG

O Estabelecimento da Cooperação no Contexto das Estratégias Reativas

por

Elton José da Silva Júnior

Dissertação submetida como requisito parcial para
obtenção do grau de mestre em Física.

Orientador: Jafferson Kamphorst Leal da Silva.

Co-orientador: Lucas Lages Wardil.

Instituto de Ciências Exatas
Programa de Pós-Graduação em Física

2 de maio de 2014

“A preguiça é a mãe do progresso. Se o homem não tivesse preguiça de caminhar, não teria inventado a roda.”

Mário Quintana

“Viver no mundo sem tomar consciência do significado do mundo é como vagar por uma imensa biblioteca sem tocar os livros.”

Os Ensinos Secretos de Todos os Tempos

“Here the climax of the darkening is reached. The dark power at first held so high a place that it could wound all who were on the side of good and of the light. But in the end it perishes of its own darkness, for evil must itself fall at the very moment when it has wholly overcome that good, and thus consumed the energy to which it owned its duration.”

Antigo ensinamento chinês, interpretado por R. Wilhelm e C. F. Baynes

UFMG

Resumo

Instituto de Ciências Exatas
Programa de Pós-Graduação em Física

por Elton José da Silva Júnior

Diversas espécies na natureza exibem comportamento cooperativo. O chamado dilema do prisioneiro é um jogo amplamente estudado para modelar o fenômeno da emergência da cooperação. Nesse jogo, os indivíduos tem duas opções: cooperar (C) ou desertar (D), que é o comportamento não cooperativo. Se o jogo é composto de uma única rodada, a deserção é a melhor opção, visto que ela fornece um ganho maior para o jogador. Porém, uma vez que indivíduos se encontram e jogam várias vezes entre si, a cooperação pode emergir. Se p e q são, respectivamente, as probabilidades do jogador cooperar dado que o seu oponente tenha cooperado e desertado na rodada anterior, uma infinidade de estratégias é permitida. A evolução temporal das frequências dos indivíduos que jogam uma estratégia é ditada pela equação do replicador. Como existem versões diferentes para essa equação e formas diferentes de resolvê-la numericamente (usando abordagens computacionais para se resolver equações contínuas e discretas), resultados diferentes podem ser obtidos. Neste trabalho foi mostrado que os resultados da literatura (a vitória da estratégia “tit-for-tat” generosa) é encontrado somente sob condições específicas. Um argumento analítico envolvendo análises de equilíbrio de Nash foi construí a fim de confirmar os resultados. Para investigar o estabelecimento da cooperação, as soluções numéricas foram obtidas usando-se as versões contínua e discreta das duas formas da equação do replicador (forma de Taylor e forma de Maynard Smith). Basicamente, a cooperação só é capaz de se manter se a densidade de estratégias presente no jogo não é muito grande.

UFMG

Abstract

Instituto de Ciências Exatas
Programa de Pós-Graduação em Física

por Elton José da Silva Júnior

Individuals in nature exhibit cooperative behavior. The so called Prisoner's dilemma is a game which is widely used to model this phenomenon. Players in this game have two options: cooperation (C) or desertion (D). If there is only one round, deserting is the best option. But once the individuals meet each other several times, cooperative behavior can emerge. Being p and q the probabilities of cooperating given that the opponent had cooperated and deserted in the last encounter, respectively, an infinity number of strategies is available. The time evolution of the fractions of individuals playing a given strategy is governed by the replicator equation. Since we have distinct versions for this equation and different ways to solve it (using continuous or discrete time approaches) we can obtain discordant outcomes. In this work, it is shown that the usual results which are presented in literature (Generous-tit-for-tat's victory) is found only within some specific conditions. The results were confirmed by using an analytical argument related to Nash equilibrium calculations. In order to investigate the establishment of cooperation, the numerical solutions were obtained by using both discrete and continuous versions of the replicator equation (Taylor's and Maynard Smith's). Basically cooperation is able to survive whether the density of strategies is not too large.

Agradecimentos

Aos meus pais, Maria José e Elton, que NUNCA mediram esforços para que eu pudesse estudar, pelo amor, pela criação, pelo conforto de um lar.

Ao todos os meus familiares: tias e madrinhas (minhas “segundas mães”), tios, primos, irmão. Obrigado pela torcida, pelas orações e pelos desejos de sucesso.

A todos os amigos: os de Janaúba, de Viçosa e de BH. Obrigado por terem estado comigo nos momentos mais importantes da minha vida, por termos dividido experiências sensacionais.

Ao Betão, ao Dalso e ao Gu, pelo companheirismo em todas as horas nessa nova etapa que se iniciava aqui na UFMG.

Aos meus professores, meus mestres, que desde o jardim de infância até o mestrado contribuíram para eu me tornar o aprendiz que ainda sou.

Ao Jafferson e ao Lucas, pela orientação neste trabalho. Sem eles, nada disso teria sido possível.

Às agências financiadoras CAPES, CNPq e FAPEMIG, pelo suporte financeiro.

Aos professores Jeferson Jacob Arenzon, da UFRGS, e Sílvio da Costa Ferreira Júnior, da UFV, que compuseram a banca examinadora e contribuíram nas discussões com sugestões bastantes pertinentes ao trabalho.

Obrigado a todos vocês, que direta ou indiretamente contribuíram para que eu subisse mais esse imenso degrau!

Sumário

Resumo	ii
Abstract	iii
Agradecimentos	iv
Lista de Figuras	vii
Lista de Tabelas	ix
Abbreviations	x
1 Introdução	1
2 Teoria Evolutiva de Jogos	4
2.1 A Teoria da Evolução de Darwin	4
2.1.1 Reprodução	5
2.1.2 Seleção	6
2.1.3 Mutação	9
2.2 Jogos Evolucionários	10
2.2.1 Jogos de dois jogadores	12
2.2.2 Equilíbrio de Nash	15
2.2.3 Estratégia Evolutivamente Estável (ESS)	17
2.2.4 A Equação do Replicador	18
3 O Dilema do Prisioneiro	21
3.1 Definição do jogo	21
3.2 Reciprocidade Direta e Estratégias Determinísticas	24
3.3 O Torneio de Axelrod	27
4 A Cooperação e as Estratégias Reativas	32
4.1 Estratégias Determinísticas e Estocásticas	32
4.2 Definições	33

4.3	O Nível Ótimo de Generosidade: GTFT	37
4.4	O Estabelecimento da Cooperação Quando Existem Poucas Es- tratégias	39
4.4.1	Três Estratégias	39
4.4.2	Quatro Estratégias	45
4.4.3	Cinco Estratégias	47
4.5	O Estabelecimento da Cooperação Quando Existem Muitas Es- tratégias	48
5	Conclusões	56
A	A Estratégia “win-stay, lose-shift”	58
	Referências Bibliográficas	60

Lista de Figuras

2.1	Representação dos simplexes S_2 , S_3 e S_4 . O simplex S_2 , por exemplo, é dado pelo intervalo fechado $[0, 1]$. O simplex S_n é uma estrutura $(n - 1)$ -dimensional contida num espaço euclidiano n -dimensional. O simplex S_n tem n faces que, por sua vez, constituem um simplex S_{n-1}	8
2.2	Análise de estabilidade dos pontos de equilíbrio da equação 2.19. A curva $f_A(x) - f_B(x)$ é a diferença entre aptidões de A e B como uma função de x . No intervalo nos quais essa diferença é positiva, a frequência de A , isto é, x , aumenta, como indicado pelas setas, que apontam na direção da dinâmica de seleção; se a diferença é negativa, x diminui. Os pontos nos quais essa diferença é nula são os pontos de equilíbrio, ou seja, o valor de x não muda. Onde $f_A(x^*) - f_B(x^*)$ é positivo, o ponto de equilíbrio é instável; caso contrário, o ponto de equilíbrio é estável. Os pontos $x = 0$ e $x = 1$ são pontos de equilíbrio, de forma que, se $[f_A(0) - f_B(0)] < 0$, $x = 0$ é estável e, se $[f_A(1) - f_B(1)] > 0$, $x = 1$ é estável [9].	12
2.3	Cinco possibilidades para a dinâmica de seleção entre duas estratégias A e B , de cima para baixo: (i) A domina B , (ii) B domina A , (iii) A e B são bi-estáveis, (iv) A e B coexistem num equilíbrio estável e (v) A e B são estratégias neutras [8].	14
3.1	TFT não pode corrigir mal-entendidos: caso um erro ocorra (asterisco vermelho), o jogo transforma-se numa alternância entre cooperação e deserção. Outro erro cometido leva à deserção mútua. Erros futuros trazem de volta a cooperação e esse ciclo é infundável [9].	29
3.2	Comparação da dinâmica de seleção entre as três estratégias básicas: ALLD, TFT e ALLC [9].	31
4.1	Diagrama de fluxos das frações dos indivíduos jogando as estratégias ALLC, ALLD, TFT e GTFT, tomadas de três a três. A versão de Taylor da equação do replicador foi resolvida pelo método de Runge-Kutta de quarta ordem para várias condições iniciais. O fluxo é mais rápido onde as cores são avermelhadas e mais lento nas cores azuladas. As figuras foram obtidas utilizando-se o <i>software</i> Dynamo [38].	40

4.2	Diagrama de fluxos das frações dos indivíduos jogando as estratégias PALLC, PALLD, PTFT e PGTFT1, tomadas de três a três. A versão de Taylor da equação do replicador foi resolvida pelo método de Runge-Kutta de quarta ordem para várias condições iniciais. O fluxo é mais rápido onde as cores são avermelhadas e mais lento nas cores azuladas. As figuras foram obtidas utilizando-se o <i>software</i> Dynamo [38].	42
4.3	Diagrama de fluxos no simplex com quatro estratégias. A origem dos eixos coordenados representa a estratégia PALLC. Os fluxos em cada plano são mostrados em (a) e em (b) observa-se a evolução da condição inicial homogênea, que converge para o estado final dominado por PGTFT2. A versão de Taylor da equação do replicador foi resolvida pelo método de Runge-Kutta de quarta ordem.	45
4.4	Fração final de PGTFT quando quatro estratégias estão fixas e a quinta varre o quadrado $[0,01; 0,99] \times [0,01; 0,99]$. A evolução temporal é descrita pela equação do replicador de Taylor.	46
4.5	Fração final de PGTFT quando quatro estratégias estão fixas e a quinta varre o quadrado $[0,01; 0,99] \times [0,01; 0,99]$. A evolução temporal é descrita pela equação do replicador de Maynard Smith.	47
4.6	Estratégia vencedora, dependendo do valor de d e da versão da equação do replicador utilizada.	49
4.7	Evolução das principais estratégias para o caso $d = 1$, isto é, 9 estratégias, utilizando-se a equação do replicador de Taylor contínua.	50
4.8	Evolução das principais estratégias para o caso $d = 4$, isto é, 36 estratégias, utilizando-se a equação do replicador de Maynard Smith contínua.	51
4.9	Ilustração da transição de ALLD para TFT ao longo da linha de estratégias (p, ε)	54

Lista de Tabelas

4.1	Pontos fixos do confronto das estratégias ALLC, ALLD, TFT e GTFT, três a três, bem como de suas versões perturbadas.	44
-----	--	----

Abbreviations

ALLC	“always cooperates”, a estratégia que sempre coopera
ALLD	“always defects”, a estratégia que sempre deserta
ESS	“evolutionarily stable strategy”, estratégia evolutivamente estável
GRIM	a estratégia dita “impiedosa”
GTFT	“generous tit for tat”, a estratégia TFT generosa
PALLC	a estratégia ALLC perturbada
PALLD	a estratégia ALLD perturbada
PGTFT	a estratégia GTFT perturbada
PTFT	a estratégia TFT perturbada
TFT	“tit for tat”, (olho por olho, dente por dente)

A Maria José e Elton.

Capítulo 1

Introdução

A cooperação é um fenômeno interessante que é evidenciado em muitos tipos de populações na natureza, e, ao mesmo tempo, intrigante [1, 2]. Em algumas espécies de macaco, por exemplo, um indivíduo grita quando avista um predador a fim de que os outros macacos saibam da situação de perigo e se protejam, o que faz com que aquele que gritou seja um alvo fácil. Mesmo correndo um certo risco, o indivíduo que gritou está cooperando com seus companheiros. Diversos tipos de pássaros abrem mão de ter a sua própria prole para cuidar da prole de outros. Na formação de um organismo multicelular as células cooperaram para formarem tecidos, que por sua vez formam órgãos e indivíduos inteiros através da cooperação. Os seres humanos apresentam formas ainda mais complexas de cooperação, desde o simples sistema de distribuição de tarefas, a ajuda a um desconhecido até a formação de sociedades [3–5]. O que torna a cooperação um fenômeno interessante é o fato de que o ato de cooperar envolve fornecer um bônus a algum indivíduo e um ônus a si próprio. Assim, é complicado de explicar, pela seleção natural de Darwin, o estabelecimento de tal comportamento. No entanto, algumas conclusões podem ser obtidas a partir da teoria evolutiva de jogos [6–8].

Nesse contexto, os indivíduos de uma dada espécie são considerados jogadores que podem escolher entre duas ações: cooperar ou desertar (não cooperar) com o seu oponente. Para cada resultado do confronto entre os jogadores (ambos cooperarem, ambos desertarem, um cooperar e o outro desertar, e vice-versa) associa-se um ganho diferente, dado pela matriz de ganho do conhecido dilema do prisioneiro [7, 9]. O ganho, por sua vez, é traduzido como aptidão (*fitness*), fazendo a ligação com teoria da evolução de Darwin. Devido a forma como o

dilema do prisioneiro é proposto, a melhor opção de um jogador é a deserção, caso o jogo seja constituído de uma única rodada. No entanto, se eles jogam por várias rodadas, o fenómeno chamado de reciprocidade direta pode promover a emergência da cooperação [6, 10].

O cenário no qual o número de rodadas que compõe o jogo é incerto (ou seja, os jogadores não sabem quando o jogo acaba), diversas estratégias determinísticas podem ser formuladas. O número de estratégias aumenta se as chamadas estratégias estocásticas são introduzidas. Em particular, o conjunto de estratégias chamadas de reativas foi criado com a intenção de entender como a cooperação inicialmente surge e como ela se estabelece [9, 11]. Agora um indivíduo pode cooperar com probabilidade p se o seu oponente cooperou na rodada passada e pode ainda cooperar com probabilidade q , caso seu oponente tenha desertado na última rodada. Dessa forma, uma estratégia reativa é definida pelo par (p, q) , sendo ambas as variáveis restritas ao intervalo $[0, 1]$. O ganho médio de um conjunto de indivíduos que joga determinada estratégia pode ser calculado utilizando-se técnicas de campo médio. Esses ganhos médios são inseridos na equação do replicador [7, 9], cuja solução fornece a evolução temporal da fração de cada população que adota uma estratégia específica.

Algumas das estratégias mais importantes são ALLC ($p = q = 1$, ou seja, sempre coopera), ALLD ($p = q = 0$, isto é, nunca coopera). Os valores $p = 1$ e $q = 0$ caracterizam a estratégia “tit-for-tat” (TFT). Nesse caso, o jogador faz na rodada atual exatamente o que o seu oponente fez na rodada anterior. Em dois torneios propostos por Axelrod [6], diversas estratégias foram colocadas para competir entre si, duas a duas; ao final dos embates, foi considerada vencedora a estratégia que obteve o maior ganho acumulado. Surpreendentemente, uma das estratégias mais simples sagrou-se campeã: TFT. No entanto, análises do equilíbrio de Nash das estratégias reativas mostraram que a então chamada “TFT generosa” (GTFT) é a estratégia mais bem sucedida. Tal estratégia é vitoriosa porque, além de cooperar sempre que o oponente cooperou na rodada passada ($p = 1$), existe também a chance dela cooperar mesmo que o oponente tenha desertado ($q = 1/3$ é o valor ótimo), o que justifica seu adjetivo “generosa”.

A descrição apresentada na literatura de como se dá a dinâmica evolutiva das frequências das estratégias reativas é a que se segue: dado um conjunto de estratégias reativas sorteadas aleatoriamente, aquela que mais se assemelha a ALLD cresce inicialmente; em seguida, TFT surge e ALLD tem seu declínio. A ascensão

de TFT logo é interrompida para dar lugar à vitoriosa GTFT [7, 9, 12]. Mas existem alguns problemas com essa descrição. Alguns autores usam a versão da equação do replicador de Taylor [9], enquanto outros usam a versão de Maynard Smith [7], ou mesmo um modelo diferente [12]. A equação do replicador pode ainda ser discretizada [7] a fim de se obter a solução de uma maneira mais simples, ou a equação diferencial pode ser resolvida pelos métodos numéricos tradicionais (neste trabalho foi utilizado o método de Runge-Kutta de quarta ordem). Como as estratégias reativas são definidas no quadrado $[0, 1] \times [0, 1]$, os estudos podem ser feitos utilizando-se estratégias dispostas numa grade ou escolhidas de forma aleatória. Além de tudo isso, pode-se analisar o espaço de fase completo com poucas estratégias ou estudar a evolução temporal de apenas uma condição inicial quando muitas estratégias estão disponíveis. Este trabalho propõe-se a analisar alguns desses aspectos.

Apesar do estabelecimento da cooperação ser essencialmente um problema matemático, os físicos tem voltado sua atenção para essa questão devido ao fato de que fenômenos como o caos, por exemplo, podem emergir de populações heterogêneas [13]. Além disso, o dilema do prisioneiro pode ser estruturado em redes complexas, onde várias ferramentas da Física Estatística podem ser utilizadas [7, 9].

A presente dissertação está organizada da seguinte maneira: no capítulo 2 os principais conceitos da teoria evolutiva de jogos são introduzidos; o capítulo 3 descreve o problema do dilema do prisioneiro; no capítulo 4 o cenário das estratégias reativas é apresentado, bem como os resultados obtidos por nós durante o mesurado. As versões conhecidas da equação do replicador são estudadas para diversas condições iniciais, com muitas e poucas estratégias, e os resultados encontrados são discutidos, sendo que eles mostraram-se diferentes daqueles apresentados na literatura; por fim, as conclusões são mostradas no capítulo 5.

Capítulo 2

Teoria Evolutiva de Jogos

Neste capítulo serão introduzidos os conceitos fundamentais da Teoria Evolutiva de Jogos, nos quais o presente trabalho é fundamentado. O capítulo foi baseado nas referências [9] e [7].

2.1 A Teoria da Evolução de Darwin

A teoria da evolução de Darwin é construída a partir de três fenômenos básicos: reprodução (replicação), seleção e mutação. Para que a evolução ocorra, uma dada população de indivíduos precisa se reproduzir. Sob as condições necessárias, um organismo vivo, seja ele uni ou multicelular, é capaz de fazer cópias de si mesmo. Dessa forma o material genético, na forma de DNA ou RNA, é replicado e transmitido para a prole. Quando indivíduos distintos competem entre si, a seleção atua. Indivíduos diferentes reproduzem-se de formas diferentes, e aquele que o faz de forma mais eficaz sobrepõe-se aos outros. O processo de mutação é o responsável por produzir os tipos diferentes de indivíduos envolvidos na seleção. Tal processo resulta em diversidade, que pode ser ou não favorável. Assim, estatisticamente, a seleção manterá aquelas mutações que beneficiam esses indivíduos frente aos outros e eliminará as desfavoráveis. Essas três etapas serão mais detalhadas nas próximas subseções.

Daqui em diante o conceito de aptidão, ou *fitness*, será bastante utilizado. Como o próprio nome sugere, a aptidão mede o quão adaptado o indivíduo está

em seu ambiente. Do ponto de vista biológico, a aptidão é, geralmente, medida em termos da taxa de reprodução.

2.1.1 Reprodução

Considere uma bactéria num ambiente propício ao seu desenvolvimento, contendo todos os nutrientes necessários à sua reprodução. Admitindo que a cada intervalo de tempo t as bactérias dupliquem-se, a lei de crescimento para esse sistema pode ser escrita de acordo com a seguinte equação recursiva, chamada de equação de diferença,

$$x_{t+1} = 2x_t , \quad (2.1)$$

em que x_t refere-se ao número de bactérias no instante de tempo t . Note que o tempo é medido em número de gerações. Sendo x_0 o número de células em $t = 0$, a solução da equação 2.1 é

$$x_t = x_0 2^t . \quad (2.2)$$

Esse problema também pode ser formulado em termos de uma equação diferencial ao invés de uma equação recursiva, considerando o tempo como uma variável contínua. Sendo $x(t)$ a quantidade de células no instante t e assumindo que elas se reproduzam a uma taxa r , tal equação diferencial é escrita como

$$\dot{x} = \frac{dx}{dt} = rx , \quad (2.3)$$

cujas solução é

$$x(t) = x_0 e^{rt} . \quad (2.4)$$

Outros parâmetros podem ser introduzidos no modelo a fim de descrever de forma melhor o sistema em questão. Pode-se supor, por exemplo, de forma simplificada, que as células morrem a uma taxa d , o que pode ser descrito pela equação

$$\dot{x} = (r - d)x , \quad (2.5)$$

cujas solução também é um crescimento (ou decrescimento, dependendo da diferença $r - d$) exponencial. Modelando o problema de uma maneira mais real, pode-se supor que a população que se reproduz encontra barreiras que se opõem a tal crescimento, como a limitação espacial e a escassez de recursos, por exemplo.

Definindo assim uma capacidade máxima K para o número de indivíduos, um modelo para esse crescimento populacional é dado pela equação logística

$$\dot{x} = rx \left(1 - \frac{x}{K}\right) . \quad (2.6)$$

À medida que x aumenta, a taxa de crescimento diminui. Quando x atinge a capacidade de suporte K , o crescimento da população é interrompido. A solução da equação 2.6 é

$$x(t) = \frac{Kx_0e^{rt}}{K + x_0(e^{rt} - 1)} . \quad (2.7)$$

No limite em que $t \rightarrow \infty$ o tamanho da população converge para o valor de equilíbrio $x^* = K$. Note que os valores de equilíbrio x^* são definidos como os valores em que $\dot{x} = 0$. Uma descrição mais detalhada sobre pontos de equilíbrio se encontra na subseção 2.1.2.

2.1.2 Seleção

A seleção é um fenômeno que ocorre toda vez que indivíduos diferentes se reproduzem a taxas diferentes. A taxa de reprodução reflete o quão adaptado o indivíduo está em relação ao ambiente, isto é, reflete a aptidão darwiniana. Um modelo relativamente simples que ilustra tal fenômeno será agora apresentado. Sejam dois tipos de indivíduos A e B , que se reproduzem a taxas a e b , respectivamente. Sendo $x(t)$ e $y(t)$ o número de indivíduos A e B no instante t , respectivamente, tais subpopulações evoluem segundo as equações

$$\begin{aligned} \dot{x} &= ax , \\ \dot{y} &= by , \end{aligned}$$

cujas soluções são

$$\begin{aligned} x(t) &= x_0e^{at} , \\ y(t) &= y_0e^{bt} . \end{aligned}$$

Se $a > b$, então A reproduz-se mais rapidamente do que B . Isso significa que após um intervalo de tempo haverá mais indivíduos do tipo A do que do tipo B . Denotando-se por $\rho(t)$ a razão entre as populações A e B no instante t e dada a

condição inicial $\rho_0 = x_0/y_0$, tem-se que

$$\rho(t) = \rho_0 e^{(a-b)t} . \quad (2.8)$$

Se $a > b$, $\rho \rightarrow \infty$, ou seja, A irá se sobrepor a B ; por outro lado, se $b > a$, $\rho \rightarrow 0$ e a seleção favorecerá B em detrimento de A . Considere agora uma população composta por duas espécies distintas, e sejam $x(t)$ e $y(t)$ as frações dos indivíduos do tipo A e B no instante t , respectivamente, também denominadas frequências. Como as únicas espécies presentes são A e B , tem-se que $x + y = 1$ para todos os instantes de tempo. As taxas de reprodução de A e B são a e b , respectivamente. Pode-se escrever

$$\begin{aligned} \dot{x} &= x(a - \phi) , \\ \dot{y} &= y(b - \phi) , \end{aligned}$$

onde ϕ é o termo que garante que $x + y = 1$. Somando as duas equações tem-se que

$$\begin{aligned} \dot{x} + \dot{y} &= xa - x\phi + yb - y\phi \\ 0 &= ax + by - \phi(x + y) \\ \phi &= ax + by . \end{aligned} \quad (2.9)$$

Assim sendo, vê-se que ϕ é a aptidão média da população. Usando o fato de que $y = 1 - x$, obtém-se que

$$\dot{x} = x(1 - x)(a - b) . \quad (2.10)$$

A equação diferencial 2.10 é do tipo $\dot{x} = f(x)$. Os valores x^* são chamados pontos fixos, ou pontos de equilíbrio, se $f(x^*) = 0$; eles correspondem aos pontos nos quais o fluxo \dot{x} cessa. Em termos de equações diferenciais, os pontos fixos representam soluções de equilíbrio (uma vez que $x = x^*$, então $x(t) = x^*$ para sempre) [14]. Os pontos de equilíbrio da equação 2.10 são $x^* = 0$ e $x^* = 1$. Se $a > b$, então $\dot{x} > 0$ para qualquer valor de x no intervalo aberto $(0, 1)$. Isso significa que para qualquer condição inicial a fração de indivíduos A aumentará, já que que a aptidão de A é maior. Ou seja, $x \rightarrow 1$ e $y \rightarrow 0$ porque $a > b$, o que exemplifica o conceito de “sobrevivência do mais apto”.

O modelo pode ser estendido para descrever o mecanismo de seleção entre

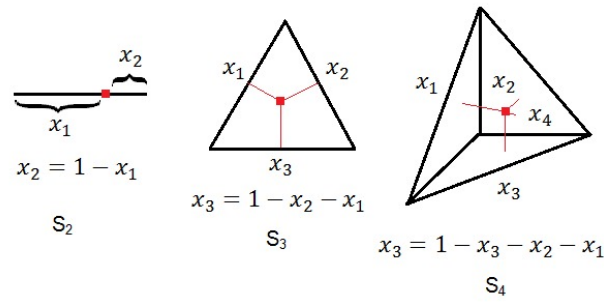


FIGURA 2.1: Representação dos simplexes S_2 , S_3 e S_4 . O simplex S_2 , por exemplo, é dado pelo intervalo fechado $[0, 1]$. O simplex S_n é uma estrutura $(n - 1)$ -dimensional contida num espaço euclidiano n -dimensional. O simplex S_n tem n faces que, por sua vez, constituem um simplex S_{n-1} .

várias espécies diferentes. Denotando por $x_i(t)$ a frequência da espécie i no instante t , $i = 1, \dots, n$, a população pode ser representada pelo vetor $\vec{x} = (x_1, \dots, x_n)$. Sendo f_i a aptidão (o *fitness*¹) da espécie i , a aptidão média da população é dada por

$$\phi = \sum_{i=1}^n x_i f_i . \tag{2.11}$$

Como $\sum_{i=1}^n x_i = 1$ e $\sum_{i=1}^n \dot{x}_i = 0$, a dinâmica de seleção pode ser escrita como

$$\dot{x}_i = x_i(f_i - \phi) . \tag{2.12}$$

A frequência da espécie i aumenta se a sua aptidão é maior que a aptidão média da população.

O conjunto de pontos que obedecem a propriedade $\sum_{i=1}^n x_i = 1$ é chamado de simplex S_n , como exemplificado na figura 2.1. Cada ponto do simplex representa uma configuração particular da população, com determinadas frequências para cada espécie i . O interior do simplex é o conjunto de pontos caracterizados por $x_i > 0$ para todo $i = 1, \dots, n$, enquanto que a face do simplex é o conjunto de pontos em que $x_i = 0$ pelo menos para um i . Os vértices do simplex são os pontos nos quais apenas uma espécie é presente, $x_i = 1$, e, portanto, todas as outras estão ausentes, ou seja, $x_j = 0$ para todo $j \neq i$.

¹Novamente, a aptidão f_i é um número real não-negativo que, nesse caso, é identificado com a taxa de reprodução da espécie em questão.

2.1.3 Mutaçãõ

Nesta subseção será apresentado um modelo simples que inclui o processo de mutação. Novamente, sejam duas espécies distintas A e B . A taxa de mutação é a probabilidade de que a reprodução de A (B) gere um indivíduo do tipo B (A). Sejam u_1 e u_2 as taxas de mutação de A para B e de B para A , respectivamente. Como antes, x e y são as frequências de A e B , respectivamente. Então pode-se escrever que

$$\begin{aligned}\dot{x} &= x(1 - u_1) + yu_2 - \phi x, \\ \dot{y} &= xu_1 + y(1 - u_2) - \phi y.\end{aligned}$$

Por simplicidade, considere que A e B tenham a mesma aptidão ($a = b = 1$), de forma que a aptidão média da população é uma constante, $\phi = 1$ nesse caso. Como $x + y = 1$, segue-se que

$$\dot{x} = u_2 - x(u_1 + u_2). \quad (2.13)$$

A equação 2.13 tem como ponto de equilíbrio

$$x^* = \frac{u_2}{u_1 + u_2}. \quad (2.14)$$

Um ponto de equilíbrio x^* é dito estável se, para toda vizinhança aberta U de x^* , existe uma outra vizinhança aberta $O \subseteq U$ tal que qualquer trajetória inicialmente contida em O permanece em U . Por sua vez, um ponto fixo é dito instável se ele não é estável. De modo geral, seja x^* um ponto fixo da equação diferencial $\dot{x} = f(x)$. O ponto fixo x^* é instável se $f'(x^*)$ é positivo e é estável se $f'(x^*)$ é negativo [14]. Voltando à equação 2.13, nota-se que x^* é um ponto de equilíbrio estável, ou seja, a mutação leva à coexistência de A e B . A proporção de uma espécie em relação à outra depende das taxas de mutação. No equilíbrio, pode-se escrever que

$$\frac{x^*}{y^*} = \frac{x^*}{1 - x^*} = \frac{u_2}{u_1}. \quad (2.15)$$

Se as taxas de mutação são as mesmas, então $x^* = y^*$.

Ocorre que, às vezes, a taxa de mutação em uma dada direção é muito maior do que em outra, de modo que nesses casos é uma boa aproximação ignorar a

mutação na direção desfavorecida. Assim, se $u_2 = 0$, tem-se que

$$\dot{x} = -xu_1, \quad (2.16)$$

cuja solução é

$$x(t) = x_0 e^{-u_1 t}. \quad (2.17)$$

Como $y = 1 - x$, segue-se que

$$y(t) = 1 - (1 - y_0) e^{-u_1 t}. \quad (2.18)$$

Nota-se que a frequência de A diminui com o tempo, enquanto que a de B aumenta. Visto que a mutação ocorre somente no sentido de A para B , os indivíduos do tipo A serão extintos e os do tipo B irão compor toda a população. Assim sendo, a mutação afeta a sobrevivência dos indivíduos de um determinado tipo: taxas de mutações diferentes podem atuar como agentes de seleção, mesmo quando as taxas de reprodução são as mesmas.

2.2 Jogos Evolucionários

Até agora considerou-se que a aptidão de uma população era uma grandeza constante. A Teoria Evolutiva de Jogos, por sua vez, interpreta a aptidão como uma grandeza que não é constante, mas que depende da frequência das espécies presentes na população. Assim sendo, tal teoria é uma aproximação mais geral para a dinâmica evolutiva e tem, como exemplo particular, o caso em que a aptidão é uma constante.

A teoria de jogos foi inicialmente desenvolvida por John von Neumann e Oskar Morgenstern [15], cujo objetivo era desenvolver um modelo para estudar o comportamento humano em estratégias e decisões relacionadas à Economia.

Para a Teoria Evolutiva de Jogos não importa se os jogadores agem de forma racional: ela simplesmente considera uma população de indivíduos interagindo em um jogo. Os jogadores possuem estratégias, fixas ou não, e interagem (aleatoriamente, por exemplo) com outros indivíduos. O ganho relativo dessas interações, também chamado de ganho, é calculado para cada jogador e interpretado como aptidão. Dessa forma, o sucesso no jogo é relacionado com sucesso na reprodução:

estratégias que acumulam mais ganho se reproduzem mais, enquanto que aquelas cujo ganho não é satisfatório são eliminadas. De uma forma simplificada, esse é o processo de seleção natural. Como exemplo ilustrativo, suponha duas espécies de bactérias, A e B . A bactéria do tipo A pode se mover, enquanto que a do tipo B permanece fixa. Por causa de sua mobilidade, A paga um certo custo, mas também possui um ganho associado à sua vantagem de locomoção. Suponha que uma análise de custo-benefício leve a uma aptidão de 1.1 para A e 1.0 para B . Se a aptidão for constante, A com certeza se sobressairá e eliminará B . Admita agora que a vantagem relacionada a se deslocar é grande quando existem poucas bactérias, já que o caminho não estará congestionado. Por outro lado, a vantagem é pequena se existem muitos indivíduos em volta bloqueando o caminho. Nesse caso, a aptidão de A não é constante, mas uma função decrescente com a frequência de A . A aptidão de A é maior que a de B quando A é escassa, e é menor quando A é abundante.

Formalizando o caso geral em que a seleção envolvendo duas estratégias A e B depende das frequências, define-se por x_A e x_B as frequências de A e B respectivamente. O vetor $\vec{x} = (x_A, x_B)$ representa a população. Sendo $f_A(\vec{x})$ a aptidão de A e $f_B(\vec{x})$ a aptidão de B , a dinâmica de seleção pode ser descrita como

$$\begin{aligned}\dot{x}_A &= x_A[f_A(\vec{x}) - \phi], \\ \dot{x}_B &= x_B[f_B(\vec{x}) - \phi].\end{aligned}$$

A aptidão média é dada por $\phi = x_A f_A(\vec{x}) + x_B f_B(\vec{x})$. Como $x_A + x_B = 1$ para todo instante de tempo, pode-se introduzir a variável x de modo que $x_A = x$ e $x_B = 1 - x$. Assim, a dinâmica resume-se a

$$\dot{x} = x(1-x)[f_A(x) - f_B(x)]. \quad (2.19)$$

cujos pontos de equilíbrio são $x^* = 0$, $x^* = 1$ e todos os valores de $x \in (0, 1)$ que satisfazem $f_A(x) = f_B(x)$. O ponto $x^* = 0$ é estável se $f_A(0) < f_B(0)$; $x = 1$, por sua vez, é estável se $f_A(1) > f_B(1)$. O ponto de equilíbrio intermediário, x^* , é estável se as derivadas de f_A e de f_B satisfazem a condição $f'_A(x^*) < f'_B(x^*)$. É importante notar que podem haver vários pontos de equilíbrio no intervalo $(0, 1)$. A figura 2.2 mostra a análise de estabilidade dos pontos fixos.

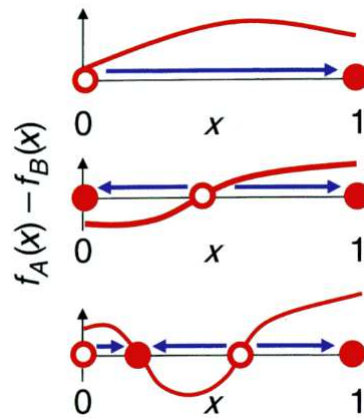


FIGURA 2.2: Análise de estabilidade dos pontos de equilíbrio da equação 2.19. A curva $f_A(x) - f_B(x)$ é a diferença entre aptidões de A e B como uma função de x . No intervalo nos quais essa diferença é positiva, a frequência de A , isto é, x , aumenta, como indicado pelas setas, que apontam na direção da dinâmica de seleção; se a diferença é negativa, x diminui. Os pontos nos quais essa diferença é nula são os pontos de equilíbrio, ou seja, o valor de x não muda. Onde $f_A(x^*) - f_B(x^*)$ é positivo, o ponto de equilíbrio é instável; caso contrário, o ponto de equilíbrio é estável. Os pontos $x = 0$ e $x = 1$ são pontos de equilíbrio, de forma que, se $[f_A(0) - f_B(0)] < 0$, $x = 0$ é estável e, se $[f_A(1) - f_B(1)] > 0$, $x = 1$ é estável [9].

2.2.1 Jogos de dois jogadores

Um jogo com duas estratégias A e B pode ser descrito pela matriz de ganho

$$\begin{array}{c|cc} & A & B \\ \hline A & a & b \\ B & c & d \end{array} . \tag{2.20}$$

A matriz de ganho é lida da seguinte maneira: A ganha a quando joga contra A e ganha b quando joga contra B ; B ganha c quando joga contra A e ganha d quando joga contra B . De forma generalizada, o elemento a_{ij} da matriz de ganho fornece o ganho da estratégia i jogando contra a estratégia j .

As estratégias representadas na matriz de ganho são denominadas estratégias puras. Em muitos jogos, no entanto, os jogadores podem também se valer de estratégias mistas, que são descritas por distribuições de probabilidades sobre as estratégias puras. Pode-se assumir que os jogadores possuem um dispositivo interno aleatório que pode ser usado em situações que exijam tomadas de decisões. Jogar uma estratégia significa que, em cada decisão, o jogador escolhe determinada estratégia com uma probabilidade pré-determinada. Assume-se que em um

jogo os indivíduos jogam independentemente, mas escolhem simultaneamente uma de suas possíveis estratégias (isto é, eles jogam juntos, mas sem saber a escolha do seu oponente); e os jogadores recebem ganhos de acordo com o perfil da ação realizada.

A ideia geral da teoria evolutiva de jogos é considerar uma população de indivíduos que jogam as estratégias A e B e interpretar seus respectivos ganhos como aptidões. Seja x_A a frequência de A e x_B a frequência de B . Para uma população dita bem-misturada, na qual muitos indivíduos interagem aleatoriamente, os ganhos esperados para A e B são, respectivamente,

$$f_A = ax_A + bx_B , \quad (2.21)$$

$$f_B = cx_A + dx_B . \quad (2.22)$$

Note que uma aproximação de campo médio está sendo usada, cuja definição implica nas seguintes suposições:

- (i) o número de indivíduos racionais conectados é muito grande;
- (ii) todos os indivíduos são equivalentes e possuem matrizes de ganhos idênticas;
- (iii) em cada rodada, os indivíduos se encontram aleatoriamente com igual probabilidade;
- (iv) atualização de estratégias são raras quando comparadas à frequência com a qual as rodadas acontecem;
- (v) todos os jogadores usam a mesma regra de atualização das estratégias.

Ao se escrever os ganhos na forma das equações 2.21 e 2.22, assume-se que a probabilidade de que uma interação com um jogador do tipo A ocorra é x_A e a probabilidade de que uma interação ocorra com um jogador do tipo B é x_B . As probabilidades são assim definidas porque os jogadores se encontram de forma aleatória. Colocando-se as equações 2.21 e 2.22 na equação 2.19, e fazendo $x_A = x$, obtém-se

$$\dot{x} = x(1-x)[(a-b-c+d)x + b-d] . \quad (2.23)$$

Dependendo do valores das constantes a , b , c e d , pode-se distinguir cinco casos:

- (i) A domina B . Isso ocorre se $a > c$ e $b > d$. Assim os únicos pontos fixos no intervalo $[0, 1]$ são $x^* = 0$ e $x^* = 1$. A melhor escolha para um jogador é a estratégia A , não importando o que seu oponente escolha. Para uma população de jogadores do tipo A e B , os valores dos ganhos implicam que a aptidão de A sempre será maior que o de B , de forma que a seleção favorecerá aquele ao invés deste, para qualquer configuração inicial da população. A seleção levará a população para a configuração caracterizada por $x^* = 1$;
- (ii) B domina A . Isso ocorre se $c > a$ e $d > b$. Nesse caso, a melhor escolha para um jogador é a estratégia B , não importando o que seu oponente escolha. Novamente, os únicos pontos fixos no intervalo $[0, 1]$ são $x^* = 0$ e $x^* = 1$. A seleção levará a população para a configuração caracterizada por $x^* = 0$, situação oposta ao caso (i);
- (iii) A e B são bi-estáveis. Isso ocorre se $a > c$ e $d > b$. Nesse caso, a melhor estratégia a ser adotada é a estratégia do seu oponente: A é a melhor resposta para A e B é a melhor resposta para B . Além de $x^* = 0$ e $x^* = 1$, no intervalo $[0, 1]$ surge o ponto fixo $x^* = (d - b)/(a - b - c + d)$. Como x^* é um ponto de equilíbrio instável, a configuração final da população vai depender da

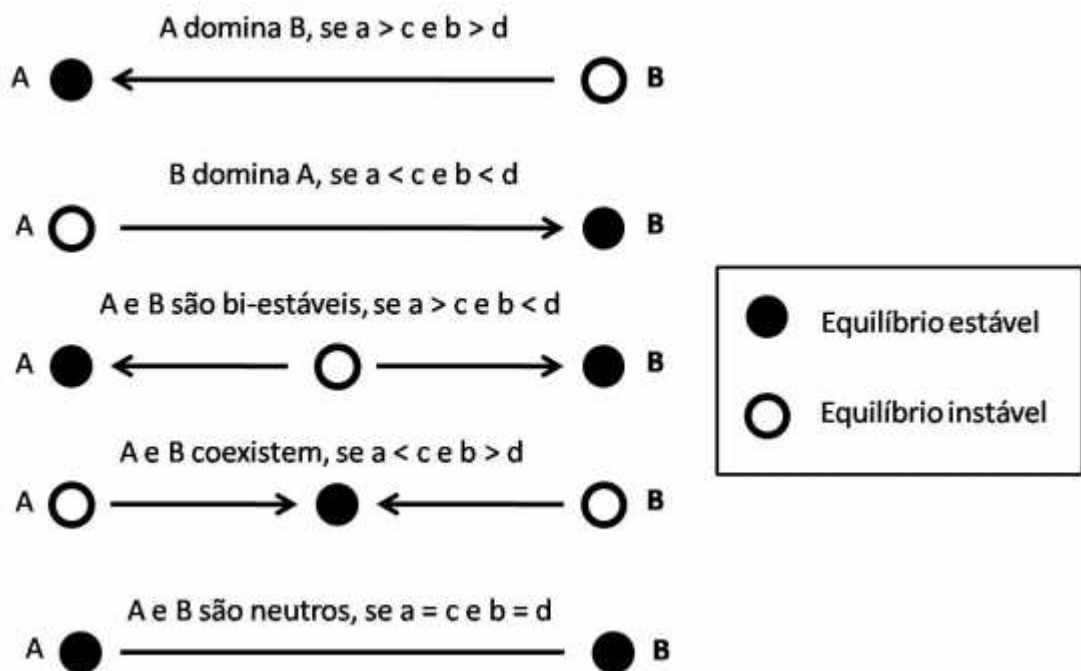


FIGURA 2.3: Cinco possibilidades para a dinâmica de seleção entre duas estratégias A e B , de cima para baixo: (i) A domina B , (ii) B domina A , (iii) A e B são bi-estáveis, (iv) A e B coexistem num equilíbrio estável e (v) A e B são estratégias neutras [8].

configuração inicial: se $x(0) < x^*$, o sistema irá convergir para a completa dominação de B ; por outro lado, se $x(0) > x^*$, o sistema irá convergir para a completa dominação de A ;

- (iv) A e B coexistem. Isso ocorre se $c > a$ e $b > d$. Novamente, a melhor estratégia a ser escolhida é a jogada pelo oponente. Além de $x = 0$ e $x = 1$, no intervalo $[0, 1]$ o ponto fixo $x^* = (d - b)/(a - b - c + d)$ também existe e é estável. Qualquer configuração inicial no intervalo $(0, 1)$ irá convergir para o ponto x^* ;
- (v) A e B são neutros. Isso ocorre se $a = c$ e $b = d$. Todos os pontos são fixos. Não importa qual estratégia um jogador escolha, ele terá o mesmo ganho do seu oponente. A seleção não irá alterar a composição da população, já que qualquer mistura de A e B é um ponto de equilíbrio.

Essa descrição está resumida na figura 2.3.

2.2.2 Equilíbrio de Nash

A teoria de jogos clássica é baseada em duas hipóteses essenciais: a perfeita racionalidade dos jogadores e o fato de que essa racionalidade é um conhecimento dividido por todos, chamado de conhecimento comum. Por perfeita racionalidade entende-se que os jogadores possuem ganhos representados por funções bem definidas e eles estão completamente cientes sobre as estratégias disponíveis para eles e seus oponentes. Não há limitações cognitivas na escolha da melhor maneira possível para se jogar, não importa o quão complicadas sejam as regras. Desse modo, a análise não tem custo e é instantânea. Já o conhecimento comum implica que, além do fato de que todos os jogadores são racionais, eles o sabem.

O conceito de equilíbrio de Nash foi introduzido pelo matemático norte-americano John Forbes Nash, ganhador do prêmio Nobel de Economia em 1994. Suponha um jogo entre dois indivíduos. O equilíbrio de Nash é assim definido: se os dois jogadores usam uma estratégia que é equilíbrio de Nash, então nenhum dos dois pode aumentar seu ganho mudando de estratégia. Um dos resultados mais fundamentais da teoria de jogos clássica é o teorema de Nash [16]: em jogos com um número finito de jogadores e um número finito de estratégias puras, existe, no mínimo, um equilíbrio de Nash, possivelmente envolvendo estratégias mistas.

Utilizando a matriz de ganho generalizada

$$\begin{array}{c|cc} & A & B \\ \hline A & a & b \\ B & c & d \end{array}$$

tem-se as seguintes definições:

- (i) A é um equilíbrio de Nash estrito se $a > c$;
- (ii) A é um equilíbrio de Nash se $a \geq c$;
- (iii) B é um equilíbrio de Nash estrito se $d > b$;
- (iv) B é um equilíbrio de Nash se $d \geq b$.

Como exemplo, considere o seguinte jogo:

$$\begin{array}{c|cc} & A & B \\ \hline A & 3 & 0 \\ B & 5 & 1 \end{array} \quad (2.24)$$

Se os dois jogadores escolhem A , qualquer um deles pode aumentar seu ganho trocando sua estratégia para B . Por outro lado, se os dois jogadores escolhem B , nenhum deles pode aumentar seu ganho mudando para a estratégia A . Então B é um equilíbrio de Nash e A é dominada por ela. Considere agora um outro jogo, definido pela matriz

$$\begin{array}{c|cc} & A & B \\ \hline A & 3 & 1 \\ B & 5 & 0 \end{array} \quad (2.25)$$

Se os dois jogadores escolhem A , então qualquer um deles pode aumentar seu ganho trocando sua estratégia para B . Novamente, se os dois jogadores escolhem B , também é possível aumentar o ganho trocando sua estratégia para A . Então, nesse caso, não há equilíbrio de Nash envolvendo estratégias puras. Finalmente, considere o jogo definido pela matriz

$$\begin{array}{c|cc} & A & B \\ \hline A & 5 & 0 \\ B & 3 & 1 \end{array} \quad (2.26)$$

Se os dois jogadores escolhem a mesma estratégia, A ou B , nenhum deles aumenta seu ganho mudando de estratégia. Portanto, as duas estratégias A e B são equilíbrio de Nash.

2.2.3 Estratégia Evolutivamente Estável (ESS)

Paralelamente ao desenvolvimento da teoria de Nash, John Maynard Smith inventou o conceito de estratégia evolutivamente estável, ou ESS². Suponha uma população muito grande de indivíduos que jogam a estratégia A , e é introduzido um único mutante do tipo B . O jogo entre A e B é dado pela matriz de ganho generalizada 2.20 e as aptidões são funções do tipo 2.21 e 2.22. A pergunta a ser respondida é: qual a condição para que a seleção se oponha à invasão de A por B ?

Assuma que há uma quantidade infinitesimal de invasores B . Assim, a frequência de B é ϵ e a frequência de A é $1 - \epsilon$. Para essa população, a aptidão de A é maior que o de B se

$$a(1 - \epsilon) + b\epsilon > c(1 - \epsilon) + d\epsilon . \quad (2.27)$$

Cancelando os termos com ϵ , a inequação leva a

$$a > c . \quad (2.28)$$

Se, porventura, $a = c$, a inequação 2.27 leva a

$$b > d . \quad (2.29)$$

Portanto, os resultados podem ser resumidos da seguinte maneira: a estratégia A é ESS se (i) $a > c$ ou (ii) $a = c$ e $b > d$ são satisfeitas. Essa definição garante que a seleção irá se opor à invasão de A por B .

Para jogos com mais de duas estratégias, seja $E(S_i, S_j)$ o ganho da estratégia S_i jogando contra a estratégia S_j .

(i) A estratégia S_k é um equilíbrio de Nash estrito se

$$E(S_k, S_k) > E(S_i, S_k) \quad \forall i \neq k . \quad (2.30)$$

²ESS, do inglês *Evolutionarily Stable Strategy*.

(ii) A estratégia S_k é um equilíbrio de Nash se

$$E(S_k, S_k) \geq E(S_i, S_k) \quad \forall i . \quad (2.31)$$

(iii) A estratégia S_k é ESS se, $\forall i \neq k$,

$$E(S_k, S_k) > E(S_i, S_k) \quad (2.32)$$

ou

$$E(S_k, S_k) = E(S_i, S_k) \quad , \quad E(S_k, S_i) > E(S_i, S_i) . \quad (2.33)$$

Note que uma ESS garante que a seleção irá se opor a qualquer invasor em potencial, o que também é garantido pelo equilíbrio de Nash estrito, mas não pelo equilíbrio de Nash. Se $E(S_k, S_k) = E(S_j, S_k)$ e $E(S_k, S_j) < E(S_j, S_j)$, então S_k é ainda equilíbrio de Nash, mas a seleção irá favorecer S_j invadindo S_k , o que torna útil a definição que se segue.

(iv) A estratégia S_k é estável contra invasão por seleção (ESS fraca) se, $\forall i \neq k$,

$$E(S_k, S_k) > E(S_i, S_k) \quad (2.34)$$

ou

$$E(S_k, S_k) = E(S_i, S_k) \quad , \quad E(S_k, S_i) \geq E(S_i, S_i) . \quad (2.35)$$

Se a estratégia é um equilíbrio de Nash estrito, então ela é também uma ESS. Se a estratégia é uma ESS fraca, então ela também é um equilíbrio de Nash. Assim, tem-se que

$$\text{equilíbrio de Nash estrito} \Rightarrow \text{ESS} \Rightarrow \text{ESS fraca} \Rightarrow \text{equilíbrio de Nash}.$$

2.2.4 A Equação do Replicador

Peter Taylor e Leo Jonker foram os primeiros a introduzir uma equação diferencial para a dinâmica de jogos evolucionários em 1978 [17]. O modelo gerador da equação considera uma população infinita, onde cada indivíduo pode adotar uma dentre n estratégias possíveis. Em cada interação entre um tipo i e um tipo j , o ganho para a estratégia i é dado por a_{ij} , e o ganho para cada estratégia j é dado por a_{ji} . A matriz $A = [a_{ij}]_{n \times n}$ é a matriz de ganho. Seja x_i a frequência da

estratégia i . Se as interações são uniformemente aleatórias, o ganho médio do tipo i é dado por

$$f_i = \sum_{j=1}^n x_j a_{ij} , \quad (2.36)$$

sendo que

$$\phi = \sum_{i=1}^n x_i f_i \quad (2.37)$$

é o ganho médio da população. Associando-se o ganho médio com a aptidão, a equação do replicador é

$$\dot{x}_i = x_i(f_i - \phi) , \quad i = 1, \dots, n . \quad (2.38)$$

Note que \dot{x}_i depende do desvio da aptidão. A diferença entre as equações 2.38, usualmente chamada de forma de Taylor da equação do replicador, e 2.12 é que agora a aptidão é uma função linear das frequências, ao invés de ter valores constantes.

Admitindo agora que \dot{x}_i dependa do desvio relativo da aptidão, a equação do replicador pode ser escrita também como

$$\dot{x}_i = x_i \frac{(f_i - \phi)}{\phi} , \quad i = 1, \dots, n , \quad (2.39)$$

chamada de forma de Maynard Smith da equação do replicador, ou equação do replicador ajustada. As duas formas da equação do replicador podem ser deduzidas a partir de modelos microscópicos [7]. Note que, como os pontos fixos são obtidos em $\dot{x} = 0$, ambas as equações apresentam os mesmos pontos fixos. No entanto, devido a presença do termo ϕ no denominador da equação 2.39, os fluxos podem ser diferentes.

Ambas as equações são definidas no simplex S_n dado por $\sum_{i=1}^n x_i = 1$. O interior do simplex, dado pelo conjunto $\{x \in \mathbb{R}^n \mid x_1, \dots, x_n > 0\}$, é invariante: se uma condição inicial não contém, porventura, uma dada estratégia i , ou seja, $x_i(0) = 0$, então $x_i(t) = 0$ para todo instante t . Além disso, as faces do simplex — uma face é um subconjunto do simplex de dimensão maior ou igual a três onde, no mínimo, uma estratégia tem frequência igual a zero — também são invariantes. Note que a dinâmica do replicador não cria novas estratégias, haja visto que, se uma estratégia está ausente na população, a dinâmica permanece sempre dentro da respectiva face do simplex. A dinâmica do replicador enquadra-se numa vasta classe de dinâmicas denominadas dinâmicas não-inovadoras [7].

Os vértices do simplex são pontos fixos da dinâmica do replicador. Dependendo da matriz de ganho podem existir outros pontos fixos no interior e nas faces do simplex. Além de estável ou instável, os pontos fixos x^* podem ainda ser classificados como: (i) atrativo, se existir uma vizinhança aberta U de x^* tal que toda trajetória inicialmente contida em U converge para x^* (o valor máximo de U é chamado de bacia de atração de x^*); e (ii) assintoticamente estável, ou atrator, se ele for estável e atrativo. Um ponto fixo é dito global-assintoticamente estável se a sua bacia de atração cobre todo o espaço. O equilíbrio de Nash e a dinâmica de estabilidade da equação do replicador estão relacionados pelo teorema de Folk [7], cujas implicações são:

- (i) equilíbrios de Nash são pontos fixos;
- (ii) equilíbrios de Nash estritos são atratores;
- (iii) se uma órbita interior converge para x^* , então x^* é um equilíbrio de Nash;
- (iv) se um ponto fixo é estável, então ele é um equilíbrio de Nash.

Capítulo 3

O Dilema do Prisioneiro

Neste capítulo será descrito o jogo conhecido por Dilema do Prisioneiro (ou PD, do inglês *Prisoner's Dilemma*) e será estudada a evolução das estratégias determinísticas. Esse jogo, que se trata de um paradoxo, foi idealizado pelo matemático Albert Tucker em 1950 para ilustrar a dificuldade de analisar certos tipos de jogos estudados previamente por Melvin Dresher e Merrill Flood. O paradoxo de Tucker, como também é conhecido o dilema do prisioneiro, dá margem para uma vasta literatura em diversas áreas, como física, filosofia, biologia, economia, ciências políticas e comportamentais, e teoria de jogos [7].

3.1 Definição do jogo

O nome “dilema do prisioneiro” é apenas figurativo, pois o que interessa são as estratégias e os ganhos representados no jogo. Duas pessoas são suspeitas de terem cometido um crime juntas. Os suspeitos são presos em diferentes celas e não tem contato um com o outro. A polícia não tem evidências suficientes para convencer o júri de que os suspeitos são de fato os responsáveis pelo crime. A promotoria tenta fazer o seguinte acordo com os suspeitos: se eles confessarem o crime, passarão a ser testemunhas de acusação para, assim, evitarem uma sentença de prisão. Se um dos prisioneiros confessar o crime e o outro não o fizer, então aquele que confessar será libertado imediatamente e o que permaneceu em silêncio ficará preso por dez anos. Se ambos confessarem, receberão uma sentença de sete anos de prisão. Se nenhum dos dois confessar, eles serão libertados depois um ano, já que a autoria

do crime não seria comprovada [18]. Esse jogo pode ser representado pela matriz de ganho

$$\begin{array}{c|cc}
 & \textit{silenciar} & \textit{confessar} \\
 \hline
 \textit{silenciar} & -1 & -10 \\
 \textit{confessar} & 0 & -7
 \end{array} . \tag{3.1}$$

Sob o ponto de vista somente de um dos prisioneiros, o comportamento que traria mais benefícios para si próprio seria a deserção, ou seja, confessar o crime. Por outro lado, cooperar com o seu cúmplice (ambos permanecerem em silêncio) representa um benefício mútuo maior. No entanto, quando um dos suspeitos permanece em silêncio, existe a chance de que seu companheiro confesse e saia impune, prejudicando aquele que cooperou. Aí está o paradoxo: deve o indivíduo silenciar-se ou confessar? O que os suspeitos devem então fazer, e o que isso tem a ver com evolução?

Da análise do PD observa-se que a cooperação não é a melhor estratégia a ser seguida, visto que ela não fornece o maior ganho possível para o indivíduo. No entanto, em diversas populações presentes na natureza, o comportamento cooperativo emerge [19–27], o que poderia, a princípio, parecer paradoxal. Do ponto de vista biológico, o problema da cooperação é tão velho quanto o da evolução. O progresso evolutivo e a manifestação de novos fenótipos¹, por exemplo, geralmente requerem a cooperação das partes mais simples já existentes em um organismo. As moléculas replicadoras precisam cooperar para formar as primeiras células, por exemplo. Células, por sua vez, tem que cooperar para formar um organismo multicelular. As células somáticas de um organismo também cooperam para que as células reprodutoras possam fazer seu trabalho e transmitir a herança genética do indivíduo. Os animais cooperam para formar estruturas sociais, como grupos e sociedades. Abelhas operárias cooperam para defender a vida da abelha rainha e garantir a perpetuação da espécie. Algumas espécies de pássaros cooperam para alimentar uma prole que não seja a sua própria. Humanos cooperam em larga escala, criando desde povoados até cidades, estados e países. Vários outros exemplos podem ser citados, evidenciando a importância do estudo da cooperação e o seu estabelecimento [9].

Em suma, o problema pode ser colocado da seguinte forma: suponha dois indivíduos que podem cooperar, C, ou desertar, D. Se ambos cooperam, ganham

¹O fenótipo é a expressão física das características genéticas que determinam a estrutura e o funcionamento de um organismo [28].

3 pontos cada. Se um coopera e o outro deserta, o cooperador não ganha nada e o desertor ganha 5 pontos. Se ambos desertam, cada um ganha um ponto. Assim, a matriz de ganho é

$$\begin{array}{c|cc} & C & D \\ \hline C & 3 & 0 \\ D & 5 & 1 \end{array} . \quad (3.2)$$

Essa matriz de ganho tem a mesma estrutura da matriz 3.1 do dilema do prisioneiro. Novamente, a pergunta a ser respondida é: o que é melhor, cooperar ou desertar?

Partindo-se do pressuposto de que o oponente irá cooperar, o indivíduo receberá 3 pontos se sua escolha for também cooperar e receberá 5 pontos se escolher desertar. Dessa forma, é melhor desertar. Por outro lado, assumindo-se agora que o oponente irá desertar, o jogador terá um ganho 0 caso escolha cooperar e ganhará 1 ponto se resolver também desertar. Novamente, a melhor escolha é desertar. Logo, não importa o que o oponente faça, a deserção sempre é a melhor escolha [9].

Se o oponente faz essa mesma análise lógica, ele chegará à conclusão de que também é melhor que ele não coopere. No final, ambos os indivíduos terão 1 ponto cada, que, por sinal, é menor do que os 3 pontos que ambos receberiam caso tivessem cooperado. Dessa forma, o dilema é posto: jogadores ditos racionais irão desertar a fim de obter um ganho maximizado no PD; a cooperação mútua levaria a um ganho maior do que aquele obtido com a deserção mútua. No entanto, é arriscado cooperar, pois o oponente pode desertar. Sob este ponto de vista, a cooperação é “irracional”. Experimentos envolvendo teoria de jogos mostram, no entanto, que na maioria dos casos os indivíduos se comportam de maneira “irracional”. No dilema do prisioneiro, humanos geralmente tendem a cooperar, e somente quando aprendem que esse comportamento não funciona é que mudam sua estratégia para a deserção [9, 29].

De volta ao dilema inicial, cooperação significa, nesse contexto, não cooperar com a promotoria e cooperar com o seu cúmplice, permanecendo em silêncio. Se ambos silenciam-se, nenhum crime pode ser provado. Por sua vez, a deserção significa confessar. Se ambos confessam, eles ficarão presos por muito tempo. Conclui-se que não importa o que seu parceiro fale, a melhor opção sempre será desertar. Essa análise racional sugere que ambos os prisioneiros confessarão e passarão sete anos na cadeia [9].

Cooperar e desertar são comportamentos que podem ser associados a estratégias em um jogo, cuja matriz de ganho é dada por 3.2. Analisando o problema sob o ponto de vista da seleção, considere uma população de cooperadores e desertores. A frequência de cooperadores é dada por x e a de desertores é $1-x$. O ganho médio dos cooperadores é $f_C = 3x$ e o dos desertores é $f_D = 5x + 1 - x = 4x + 1$. Nota-se que os desertores sempre tem um *fitness* maior que os cooperadores, fazendo com que os cooperadores sejam dominados. A seleção natural faz com que a frequência dos desertores aumente até que os cooperadores estejam extintos. Nesse contexto, a seleção natural favorece a deserção [9].

3.2 Reciprocidade Direta e Estratégias Determinísticas

Considere a seguinte matriz de ganho:

$$\begin{array}{c|cc} & C & D \\ \hline C & R & S \\ D & T & P \end{array} . \quad (3.3)$$

Nessa matriz P é “punição por deserção mútua”, T é a “tentação a desertar”, S é o “ganho do perdedor” e R é a “recompensa por cooperação mútua”². No dilema do prisioneiro tem-se $T > R > P > S$, isto é, a tentação de desertar excede o benefício da cooperação mútua, que por sua vez é maior que a punição da deserção, que é maior que o valor recebido pelo perdedor que coopera com um desertor. Além disso, exige-se que $R > (T + P)/2$ para que a alternância entre cooperar e desertar não leve a um ganho maior do que a cooperação num jogo que se repete várias vezes [7].

O conceito de reciprocidade direta é um fenômeno que ocorre quando se joga várias vezes o mesmo jogo com o mesmo oponente. Nesse caso a cooperação pode se tornar uma estratégia promissora. Como exemplo, considere um jogo de duas estratégias: a estratégia GRIM (impiedosa, em tradução livre) e a estratégia ALLD. A GRIM coopera na primeira rodada e, a partir daí, coopera enquanto o oponente não deserta, quando, então, a GRIM passa a desertar permanentemente. A ALLD não coopera em todas as rodadas. Se o jogo for repetido m vezes, no

²P, T, S e R, do inglês *Punishment*, *Temptation*, *Sucker* e *Reward*, respectivamente.

confronto entre dois jogadores que adotam a estratégia GRIM, eles cooperam um com o outro na primeira rodada e o farão em todas as m rodadas. Assim, o ganho para cada um é o número de rodadas vezes a recompensa pela cooperação, mR . Se um jogador adotando a estratégia ALLD joga com um GRIM, este coopera na primeira rodada e aquele sempre deserta. ALLD ganha o valor T na primeira rodada, pois conseguiu explorar seu oponente, que ganha S , e nas $(m - 1)$ rodadas restantes ambos ganham P , já que GRIM irá desertar sempre. Assim, o ganho de GRIM é $S + (m - 1)P$ e o de ALLD é $T + (m - 1)P$. Por fim, se dois indivíduos do tipo ALLD jogam, ambos sempre desertam e ganham P em todas as m rodadas. A matriz de ganho é

$$\begin{array}{c|cc}
 & GRIM & ALLD \\
 \hline
 GRIM & mR & S + (m - 1)P \\
 ALLD & T + (m - 1)P & mP
 \end{array} . \tag{3.4}$$

Se $mR > T + (m - 1)P$, a GRIM é um equilíbrio de Nash estrito quando está competindo contra ALLD, o que significa que se ambos os jogadores usam a estratégia GRIM, nenhum deles pode aumentar seu ganho mudando para a estratégia ALLD. Em termos de dinâmica evolutiva, uma população inteira de GRIM não pode ser invadida por um mutante ALLD. Dessa forma, GRIM é estável contra invasão de ALLD se o número de rodadas excede o valor crítico

$$\begin{aligned}
 mR(1 - \epsilon) + [S + (m - 1)P]\epsilon &> [T + (m - 1)P](1 - \epsilon) + mP\epsilon \\
 mR &> T + mP - P \\
 m &> \frac{T - P}{R - P} .
 \end{aligned} \tag{3.5}$$

Portanto, esse mecanismo descrito estabiliza a cooperação, uma vez que ela tenha sido estabelecida. Note que ALLD também é um equilíbrio de Nash estrito, pois $mP > S + (m - 1)P$. Então não há, a princípio, um mecanismo evolutivo capaz de explicar a emergência da cooperação [9].

Suponha agora que ambos os jogadores saibam que o jogo é composto por m rodadas. Como o jogo acaba na última rodada e o objetivo é maximizar o ganho, não há motivos para cooperar na última rodada. Assim, justifica-se pensar na estratégia GRIM com a seguinte modificação, denotada por GRIM*: ambos os jogadores certamente irão desertar na última rodada. A matriz de ganho de

GRIM *versus* GRIM* é

$$\begin{array}{c|cc}
 & GRIM & GRIM^* \\
 \hline
 GRIM & mR & (m-1)R + S \\
 GRIM^* & (m-1)R + T & (m-1)R + P
 \end{array} . \quad (3.6)$$

Note que GRIM é dominado por GRIM*, pois $(m-1)R + T > mR$, e uma população de jogadores GRIM pode ser invadida por uma pequena fração de jogadores GRIM*.

Uma vez que alguém joga GRIM*, o mesmo argumento pode ser aplicado para a penúltima rodada, pois não é racional cooperar dessa vez visto que em seguida ambos os jogadores irão desertar. Da mesma forma, o argumento pode ser usado novamente para a antepenúltima rodada, e para a rodada anterior à ela, e assim por diante até chegar à primeira rodada. Pode-se escrever uma sequência de estratégias, começando com GRIM, que é dominada por uma estratégia que deserta na última rodada, que por sua vez é dominada pela estratégia que deserta na penúltima rodada, e assim sucessivamente, até chegar-se a ALLD. Nesse espaço de estratégias, somente ALLD é um equilíbrio de Nash estrito e uma ESS [9].

No entanto, seres humanos não usam esse tipo de raciocínio em situações experimentais [27, 29]. As pessoas normalmente notam que desertar no final é a melhor opção, mas elas não levam essa estratégia até suas últimas consequências, que seria desertar sempre. Uma explicação é que os instintos relacionados a estratégias da espécie humana não são formados por jogos com um número de rodadas pré-definido. Não é certo quando o jogo irá se encerrar, pois sempre pode haver uma outra rodada [9].

Ao invés de fixar o número de rodadas do dilema do prisioneiro repetido, suponha agora que exista uma probabilidade w de que outra rodada aconteça. Então a probabilidade de que exatamente o jogo seja jogado $n-1$ vezes e encerrado

logo após é $w^{n-1}(1-w)$. Portanto, o número médio de rodadas \bar{m} é

$$\begin{aligned} \bar{m} &= \sum_{n=1}^{\infty} n w^{n-1} (1-w) \\ &= (1-w) \frac{\partial}{\partial w} \sum_{n=1}^{\infty} w^n \\ &= (1-w) \frac{\partial}{\partial w} \left(\frac{1}{1-w} \right) \\ &= \frac{1-w}{(1-w)^2} = \frac{1}{1-w} . \end{aligned}$$

A matriz de ganho para as estratégias GRIM e ALLD é

	<i>GRIM</i>	<i>ALLD</i>	
<i>GRIM</i>	$\bar{m}R$	$S + (\bar{m} - 1)P$	(3.7)
<i>ALLD</i>	$T + (\bar{m} - 1)P$	$\bar{m}P$	

GRIM é ESS se

$$\begin{aligned} \bar{m}R &> T + (\bar{m} - 1)P \\ \bar{m} &> \frac{T - P}{R - P} . \end{aligned}$$

Nada muda, exceto pelo fato de que agora não há estratégia que possa desertar na última rodada, pois sempre existe uma probabilidade não-nula de que uma próxima rodada ocorra [9].

3.3 O Torneio de Axelrod

A busca pela melhor estratégia levou o cientista político Robert Axelrod a idealizar em 1978 uma espécie de campeonato do dilema do prisioneiro [6]. Ele convidou pessoas de todas as partes do mundo a submeterem estratégias formuladas em termos de programas de computador para o seu torneio. Todas as estratégias jogaram umas contra as outras e os ganhos calculados eram somados. Para isso, ele utilizou os valores $T = 5$, $R = 3$, $P = 1$ e $S = 0$ na matriz de ganho. Por fim, Axelrod analisou qual estratégia tinha o maior ganho acumulado.

Um total de quatorze estratégias foram inscritas no torneio. Algumas dessas estratégias baseavam-se em mecanismos inteligentes, capazes de enganar o oponente ou mesmo prever seu comportamento. No entanto, a estratégia vencedora foi a mais simples de todas, chamada de *tit-for-tat* (TFT, ou “olho por olho, dente por dente”, em tradução livre). TFT é a estratégia que consiste em começar cooperando e, partir de então, fazer na próxima rodada o que quer que o oponente tenha feito na rodada anterior. Dessa forma, TFT irá cooperar se o oponente cooperou anteriormente e irá desertar se o oponente desertou anteriormente. A vitória foi dada ao estudioso de teoria de jogos Anatol Rapoport, que submeteu a estratégia TFT no torneio [6, 9].

Os resultados e as análises das estratégias do torneio foram publicados por Axelrod. Ele então convidou mais pessoas a submeterem novas estratégias para um segundo campeonato [6]. Dessa vez houveram sessenta e três estratégias, e, novamente, TFT foi a vencedora. Num conjunto de estratégias acessíveis é possível prever qual delas é a melhor, enquanto que essa previsão torna-se difícil num conjunto de estratégias desconhecidas. De qualquer forma, TFT foi consagrada, sem questionamentos, a campeã mundial do torneio.

Axelrod frisou as qualidades importantes que fizeram de TFT a estratégia vencedora. Ela é uma estratégia “bondosa”, no sentido de que nunca é a primeira a desertar; TFT nunca tenta conseguir mais lucro do que seu oponente teria num confronto direto; em cada partida isoladamente, ela recebe, no máximo, o mesmo número de pontos do seu oponente. A soma de pontos de todas as partidas jogadas, no entanto, é maior para TFT do que para seus concorrentes. Conclui-se que TFT não é triunfante sob pareamento direto, mas seu sucesso é devido ao fato de que ela consegue, em média, um ganho maior num confronto com a estratégia X comparado ao ganho obtido pelas outras estratégias num confronto com a mesma estratégia X . Por fim, TFT é muito bem sucedida em induzir o comportamento cooperativo a partir de outras estratégias [6, 9].

Além disso, TFT é estável contra invasão de ALLD se o número médio de rodadas \bar{m} é grande o suficiente. TFT irá cooperar na primeira rodada, mas irá desertar nas rodadas seguintes. A matriz de ganho de TFT *versus* ALLD é

$$\begin{array}{c|cc}
 & TFT & ALLD \\
 \hline
 TFT & \bar{m}R & S + (\bar{m} - 1)P \\
 ALLD & T + (\bar{m} - 1)P & \bar{m}P
 \end{array} . \quad (3.8)$$

TFT: C C C D C D C D D D ...
 TFT: C C C C D C D D D D ...

FIGURA 3.1: TFT não pode corrigir mal-entendidos: caso um erro ocorra (asterisco vermelho), o jogo transforma-se numa alternância entre cooperação e deserção. Outro erro cometido leva à deserção mútua. Erros futuros trazem de volta a cooperação e esse ciclo é infindável [9].

Essa matriz de ganho é a mesma para GRIM *versus* ALLD. TFT pode resistir à invasão de ALLD se $\bar{m} > (T - P)/(R - P)$. A vantagem de TFT em relação a GRIM é que ela volta a cooperar se o oponente cooperar na rodada anterior, enquanto que GRIM sempre irá desertar uma vez que o oponente tenha desertado [9].

O torneio de Axelrod foi conduzido num universo digital imune a erros, mas no mundo real situações onde mal-entendidos ocorram são cabíveis. Verificou-se que, na presença de erro, dois jogadores do tipo TFT acumulam um ganho baixo. A figura 3.1 ilustra esse processo. Um único erro desloca o jogo do estado de cooperação mútua para alternância entre cooperação e deserção. Um segundo erro pode levar o jogo à deserção mútua. Para um jogo com muitas rodadas, dois jogadores do tipo TFT, com uma chance pequena de cometerem erros, obtêm o mesmo ganho obtido por dois jogadores que escolhem aleatoriamente quando desertar e quando cooperar. O ganho de dois jogadores do tipo TFT num mundo onde exista uma pequena possibilidade de se cometer erros é

$$E(TFT, TFT) = \frac{R + T + P + S}{4} . \tag{3.9}$$

Como $R > (T + S)/2$ e $R > P$, então $E(TFT, TFT) < R$. Por isso TFT é considerada uma estratégia fraca na presença de erros.

Mesmo na ausência de erros, TFT possui outra fraqueza: ela não é nem equilíbrio de Nash estrito nem uma ESS [30]. Suponha um embate entre TFT e uma estratégia do tipo “sempre cooperar”, denotada por ALLC. A matriz de ganho é dada por

	<i>TFT</i>	<i>ALLC</i>	
<i>TFT</i>	$\bar{m}R$	$\bar{m}R$.
<i>ALLC</i>	$\bar{m}R$	$\bar{m}R$	

(3.10)

Ambos os jogadores cooperam em todas as rodadas. Então, de fato, TFT não é equilíbrio de Nash estrito e nem uma estratégia evolutivamente estável.

De forma simplificada, como pode ser observado na figura 3.2, tem-se que:

- (i) ALLC é explorado por ALLD. Numa população misturada com jogadores que utilizam essas duas estratégias, ALLD sempre possui uma aptidão maior e domina ALLC;
- (ii) ALLD consegue explorar TFT apenas na primeira rodada, e, partir daí, TFT sempre deserta. Então, num jogo desse tipo, ALLD recebe um ganho um pouco maior do que o de TFT. Numa população de jogadores envolvendo essas duas estratégias, a dinâmica de seleção mostra-se bi-estável. Utilizando-se a matriz de ganho 3.8, nota-se, a partir da equação 2.23, que o ponto fixo é

$$\begin{aligned} x^* &= \frac{\bar{m}P - [S + (\bar{m} - 1)P]}{\bar{m} - [S + (\bar{m} - 1)P] - [T + (\bar{m} - 1)P] + \bar{m}P} \\ &= \frac{P - S}{\bar{m}(R - P) - S - T + 2P} , \end{aligned}$$

onde x é a frequência de ALLD. Usando os valores de Axelrod para T, R, P e S, tem-se que

$$x^* = \frac{1}{2\bar{m} - 3} . \quad (3.11)$$

Quanto maior o número médio de rodadas \bar{m} , mais a seleção favorece TFT. ALLD tem uma aptidão maior somente quando TFT é raro, o que significa que a maior parte das condições iniciais favorecem TFT. É preciso ter, inicialmente, um número muito grande de jogadores do tipo ALLD para que TFT não domine;

- (iii) Dois jogadores do tipo TFT são como um TFT e um ALLC, ou ainda ambos ALLC: sempre cooperam um com o outro. Numa população mista de jogadores usando essas duas estratégias, todos tem a mesma aptidão, e, portanto, TFT não é evolutivamente estável.

Portanto, conclui-se que TFT é uma excelente estratégia que possibilita a emergência da cooperação no contexto das estratégias determinísticas. No entanto, se a ocorrência de erros é considerada, TFT é uma estratégia ruim no que diz respeito à manutenção da cooperação. Considerando que o número de rodadas

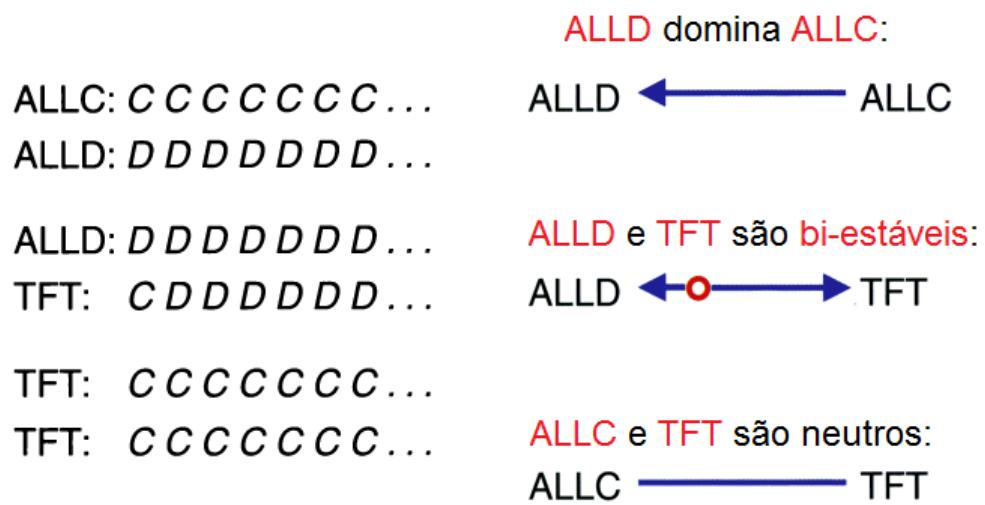


FIGURA 3.2: Comparação da dinâmica de seleção entre as três estratégias básicas: ALLD, TFT e ALLC [9].

que compõem o jogo não é fixo, a questão dos erros na evolução das estratégias determinísticas leva ao estudo das estratégias estocásticas.

Capítulo 4

A Cooperação e as Estratégias Reativas

Neste capítulo o conceito de estratégias reativas será apresentado. Todo o trabalho inédito desenvolvido nesta dissertação se encontra neste capítulo, bem como os seus resultados. Questões como o estabelecimento da cooperação e como as estratégias reativas evoluem no tempo serão estudadas.

4.1 Estratégias Determinísticas e Estocásticas

Uma estratégia determinística é uma regra que se baseia exclusivamente no histórico do jogo para tomar a decisão de cooperar ou não na rodada seguinte. Já uma estratégia estocástica é uma regra que usa o histórico do jogo como base para calcular as probabilidades de cooperar ou desertar na rodada seguinte [9].

Cada rodada do jogo tem quatro resultados possíveis: ambos os jogadores cooperam (CC), o jogador coopera e seu oponente deserta (CD), o jogador deserta e seu oponente coopera (DC), ou ambos desertam (DD). Considerando apenas a rodada anterior, existem $2 + 2 + 2 + 2 = 2 \times 4 = 8$ estratégias determinísticas acessíveis a cada jogador (para cada resultado diferente, o jogador pode cooperar ou desertar, o que justifica a expressão 2×4). Então a estratégia determinística cuja memória guarda apenas a última rodada pode ser discriminada por uma sequência binária de quatro algarismos. Dessa maneira, 0000 significa “sempre desertar”, enquanto que 1000 significa cooperar somente se o resultado da última

rodada foi CC. De forma semelhante, há 32 estratégias determinísticas possíveis que consideram as duas últimas rodadas do jogo. Assim, existem 2×4^m estratégias determinísticas que consideram as m últimas rodadas do jogo [9].

As estratégias estocásticas que consideram as m últimas rodadas do jogo formam um espaço de estratégias de dimensão 4^m ; cada rodada possui quatro resultados possíveis, e a cada um deles associa-se uma probabilidade de cooperar, o que justifica a expressão 4^m . Cada dimensão é restringida ao intervalo $[0, 1]$, para que a probabilidade faça sentido. Um jogo com um número arbitrário de rodadas tem um espaço de possíveis estratégias infinito, de modo que é impossível, mesmo para um computador, considerar todas as estratégias possíveis no dilema do prisioneiro repetido [9].

4.2 Definições

O conjunto de estratégias acessíveis aos jogadores no dilema do prisioneiro repetido é infinito. Além das estratégias ditas determinísticas, que se baseiam no histórico do jogo para tomar a decisão de cooperar ou não com o oponente, existem também as estratégias estocásticas, as quais associam ao histórico do jogo uma distribuição de probabilidades de que na próxima rodada o indivíduo coopere ou deserte.

Dentro do universo das estratégias estocásticas que considera somente as decisões tomadas na rodada anterior define-se o subconjunto das estratégias reativas. As estratégias reativas são caracterizadas por dois parâmetros: p denota a probabilidade de que o indivíduo coopere, dado que seu oponente cooperou na rodada passada, e q é a probabilidade de que o indivíduo coopere caso seu oponente tenha desertado na rodada anterior. É importante ressaltar que o conjunto formado pelas estratégias reativas levam em consideração apenas a última rodada do jogo, e por isso são ditas estratégias de memória curta (pode-se criar estratégias que ainda considerem as duas, três últimas rodadas, e assim por diante [31]). Além disso, as estratégias reativas consideram o que foi feito apenas pelo oponente na última rodada, não importando o que o próprio indivíduo tenha jogado [9].

Como p e q são probabilidades, uma dada estratégia $A(p, q)$ é um ponto contido no quadrado unitário $[0, 1] \times [0, 1]$. O conjunto das estratégias reativas engloba as já conhecidas ALLD, ALLC e TFT, representadas pelos pontos $A(0, 0)$,

$A(1, 1)$ e $A(1, 0)$, respectivamente. Essas estratégias estão localizadas em três dos quatro vértices do quadrado unitário; o outro vértice é a estratégia $A(0, 1)$, dita paradoxal, pois coopera quando o oponente desertou e deserta quando o oponente cooperou [9].

O próximo passo é descrever a dinâmica evolucionária do espaço de fase das estratégias reativas. O dilema do prisioneiro repetido entre duas estratégias pode ser mapeado numa cadeia de Markov, cujos estados são CC, CD, DC e DD, rotulados de 1 a 4, nessa ordem. Assim, o estado 1 é aquele no qual o ambos os jogadores cooperam (CC); o estado 2 é aquele no qual o jogador coopera e seu oponente deserta (CD); o estado 3 é o oposto do estado 2 (DC) e o estado 4 é quando ambos desertam (DD). Denotando por $A(p, q)$ a estratégia do jogador e por $A'(p', q')$ a estratégia do seu oponente, a cadeia de Markov é representada pela matriz de transição $M_{4 \times 4} = [m_{ij}]$. O elemento m_{32} , por exemplo é a probabilidade de transição do estado DC para o estado CD, dada por $p(1 - q')$; p é a probabilidade de que o jogador coopere, já que seu oponente cooperou na rodada passada, e $(1 - q')$ é a probabilidade de que o oponente deserte, já que o jogador desertou na rodada passada [9]. Dessa forma, a matriz de transição é

$$M = \begin{pmatrix} pp' & p(1 - p') & (1 - p)p' & (1 - p)(1 - p') \\ qp' & q(1 - p') & (1 - q)p' & (1 - q)(1 - p') \\ pq' & p(1 - q') & (1 - p)q' & (1 - p)(1 - q') \\ qq' & q(1 - q') & (1 - q)q' & (1 - q)(1 - q') \end{pmatrix}. \quad (4.1)$$

Seja \vec{x}_t a distribuição de probabilidade do jogo após t rodadas. Cada uma das quatro componentes deste vetor fornece a probabilidade do jogo estar em um dos quatro possíveis estados na rodada t . Tal distribuição de probabilidade pode ser obtida multiplicando-se a distribuição da rodada anterior pela matriz de transição:

$$\vec{x}_{t+1} = \vec{x}_t M. \quad (4.2)$$

Se existir um número positivo k tal que todos os elementos da matriz estocástica M^k são positivos, M é dita uma matriz regular. Como as entradas da matriz M são variáveis aleatórias associadas a probabilidades, cujos valores são limitados no intervalo $[0, 1]$, então M é regular. Se M é uma matriz regular, pode ser provado

[32] que existe um único autovetor \vec{x} associado ao autovalor 1, de modo que

$$\vec{x} = \vec{x}M . \quad (4.3)$$

As componentes normalizadas do autovetor \vec{x} são a distribuição estacionária da cadeia de Markov, pois $x_{t+1} = x_t$. Para obter a distribuição estacionária considere que ambos os jogadores comecem cooperando na rodada inicial, chamada rodada 0. A probabilidade s_1 de que o indivíduo que joga a estratégia A coopere na rodada 1 é

$$s_1 = p \quad (4.4)$$

e para o indivíduo que joga a estratégia A' , a probabilidade de que ele coopere na rodada 1 é

$$s'_1 = p' , \quad (4.5)$$

pois ambos os jogadores cooperaram na rodada 0. As probabilidades s_2 e s'_2 dos jogadores cooperarem na rodada 2 é, então,

$$s_2 = ps'_1 + q(1 - s'_1) , \quad (4.6)$$

$$s'_2 = p's_1 + q'(1 - s_1) . \quad (4.7)$$

De forma semelhante, para a rodada 3, as probabilidades de que os indivíduos cooperem são

$$s_3 = ps'_2 + q(1 - s'_2) , \quad (4.8)$$

$$s'_3 = p's_2 + q'(1 - s_2) . \quad (4.9)$$

Substituindo a equação 4.7 em 4.8, tem-se que

$$\begin{aligned} s_3 &= p[p's_1 + q'(1 - s_1)] + q\{1 - [p's_1 + q'(1 - s_1)]\} \\ &= pp's_1 + pq' - pq's_1 + q - p'qs_1 - qq' + qq's_1 \\ &= s_1(p - q)(p' - q') + q'(p - q) + q \end{aligned} \quad (4.10)$$

e, de forma análoga,

$$s'_3 = s'_1(p - q)(p' - q') + q(p' - q') + q' . \quad (4.11)$$

Da mesma forma, para a rodada 4, as probabilidades de que os jogadores cooperem são

$$s_4 = ps'_3 + q(1 - s'_3) , \quad (4.12)$$

$$s'_4 = p's_3 + q'(1 - s_3) . \quad (4.13)$$

Substituindo as equações 4.8 em 4.13, tem-se que

$$\begin{aligned} s_4 &= p[p's_2 + q'(1 - s_2)] + q\{1 - [p's_2 + q'(1 - s_2)]\} \\ &= pp's_2 + pq' - pq's_2 + q - p'qs_2 - qq' + qq's_2 \\ &= s_2(p - q)(p' - q') + q'(p - q) + q \end{aligned} \quad (4.14)$$

e, de forma análoga,

$$s'_4 = s'_2(p - q)(p' - q') + q(p' - q') + q' . \quad (4.15)$$

Então, de maneira geral, pode-se escrever que

$$s_{n+2} = s_n(p - q)(p' - q') + q'(p - q) + q , \quad (4.16)$$

$$s'_{n+2} = s'_n(p - q)(p' - q') + q(p' - q') + q' . \quad (4.17)$$

Como no estado estacionário $s_{n+2} = s_n$, então a probabilidade do indivíduo jogando a estratégia A cooperar no estado estacionário é

$$s = s(p - q)(p' - q') + q'(p - q) + q \quad (4.18)$$

$$s = \frac{q'(p - q) + q}{1 - (p - q)(p' - q')} . \quad (4.19)$$

Analogamente, para a estratégia A' , tem-se que

$$s' = \frac{q(p' - q') + q'}{1 - (p - q)(p' - q')} . \quad (4.20)$$

Logo, a distribuição estacionária da cadeia de Markov é dada pelo vetor

$$\vec{x} = (ss', s(1 - s'), (1 - s)s', (1 - s)(1 - s')) \quad (4.21)$$

onde s e s' são as probabilidades do jogador que adota a estratégia A e do jogador que adota a estratégia A' cooperarem na distribuição estacionária, respectivamente

[7, 33]. Dessa forma, o ganho esperado da estratégia A contra a estratégia A' é

$$E(A, A') = Rss' + Ss(1 - s') + T(1 - s)s' + P(1 - s)(1 - s') . \quad (4.22)$$

A expressão 4.22 é o valor do ganho da estratégia. A partir dela, o ganho médio é calculado (equação 2.36) e interpretado como aptidão, que é inserido na equação do replicador, tanto para a forma de Taylor (equação 2.38) como para a forma de Maynard Smith (equação 2.39). Utilizando-se um método numérico (Runge-Kutta de quarta ordem, como foi o caso), a evolução temporal das frequências das estratégias pode ser obtida.

É importante ressaltar as escalas de tempo diferentes que estão envolvidas nesse processo. A frequência com a qual os jogadores se encontram é muito maior do que a frequência com a qual a reprodução ocorre. Dessa forma, os indivíduos jogam um número de vezes grande o suficiente de forma que as probabilidades de cooperar atinjam o estado estacionário, dadas pelas equações 4.18 e 4.20. Esse número de vezes é muito grande quando comparado à frequência com a qual esses mesmos indivíduos reproduzem-se.

4.3 O Nível Ótimo de Generosidade: GTFT

Num ambiente no qual erros possam ocorrer (figura 3.1), mas não muito frequentemente, já foi mostrado (vide seção 3.3) que jogadores que utilizam a estratégia TFT não acumulam ganhos altos. Um único erro cometido entre dois indivíduos do tipo TFT desloca o jogo da cooperação mútua para a alternância entre cooperação e deserção. Molander [34] propôs que, se a porcentagem de erros ε for baixa ($\varepsilon \ll 1$), a estratégia mais bem sucedida tem um nível ótimo de generosidade. Tal estratégia é

$$(p, q) = \left(1, \min \left\{ 1 - \frac{T - R}{R - S}, \frac{R - P}{T - P} \right\} \right) . \quad (4.23)$$

Ela maximiza o ganho da população e é estável contra a invasão de desertores [35]. Ou seja, a estratégia que possui o melhor desempenho é aquela que coopera sempre que seu oponente coopera e também coopera com uma probabilidade igual a $\min \left\{ 1 - \frac{T - R}{R - S}, \frac{R - P}{T - P} \right\}$ mesmo se seu oponente deserta, o que justifica falar em

nível ótimo de generosidade. Essa estratégia é a chamada “TFT generosa”, ou GTFT (*generous tit-for-tat*).

Em uma simulação guiada por Martin Nowak [9, 33, 35] é possível entender a dinâmica evolutiva das estratégias reativas. Utilizando um gerador de números aleatórios, 100 pares de números contidos no intervalo $[0, 1]$ foram obtidos para compor um conjunto de 100 estratégias reativas. Foi utilizada a matriz 3.2 para calcular o ganho esperado de cada estratégia, dada pela expressão 4.22. A partir desses valores o ganho médio foi calculado e interpretado como aptidão. Admitindo que todas as estratégias eram igualmente abundantes em $t = 0$ e usando a versão de Taylor da equação do replicador 2.38, foi observado como a frequência das estratégias evoluía no tempo. O processo foi repetido várias vezes, com diferentes estratégias.

Na maioria dos casos o cenário era o seguinte: muitas estratégias eram extintas, e aquelas com características mais cooperativas (p e q próximos de 1) eram as primeiras a desaparecer; depois de um período de tempo, somente a estratégia mais próxima da ALLD (p e q próximos de 0) permanecia. Entretanto, em alguns casos, acontece de uma das estratégias presentes ter o par (p, q) próximo a $(1, 0)$, que é a estratégia TFT. Inicialmente a frequência da estratégia mais próxima a ALLD aumentava, até o momento em que quase todas as outras estratégias tinham sido abatidas; é nesse momento que a frequência da estratégia próxima a TFT aumenta rapidamente, enquanto os desertores vão desaparecendo. Com a mesma rapidez com que TFT surge, ela dá lugar a estratégia mais próxima de GTFT (mais próxima de $p = 1$ e $q = 1/3$, nesse caso), encerrando a dinâmica de seleção. Portanto, TFT é necessária para a emergência da cooperação e GTFT é responsável por mantê-la. No entanto, essa descrição ainda encontra problemas no contexto das estratégias reativas: GTFT não é uma ESS, mas a estratégia $(\varepsilon, \varepsilon)$, aquela mais próxima de ALLD, é uma ESS [36]. Mas também é preciso levar em consideração o fato de que a existência de uma ESS não garante que ela seja acessível, o que pode explicar a vitória de GTFT [37].

Levando-se em conta não apenas o espaço de estratégias reativas, mas todo o espaço de estratégias estocásticas de memória curta, onde apenas a última rodada importa, GTFT não se estabelece (vide apêndice A). Esse cenário, embora mais geral, apresenta um número maior de parâmetros, o que dificulta a análise da emergência da cooperação. Então, por simplicidade, será estudado somente o conjunto das estratégias reativas.

4.4 O Estabelecimento da Cooperação Quando Existem Poucas Estratégias

Para compreender como se dá o surgimento e o estabelecimento da cooperação, é conveniente começar a análise utilizando poucas estratégias, por se tratar de um sistema mais simples. Como já foi dito, também por simplicidade, será utilizado o conjunto das estratégias reativas ao invés de todo o conjunto das estratégias estocásticas. As estratégias acessíveis aos jogadores são ALLC, ALLD, TFT e GTFT. A metodologia, de forma resumida, é: calcular os ganhos das estratégias (equação 4.22), os ganhos médios (equação 2.36), definir a equação do replicador a ser utilizada, resolvê-la numericamente, analisar as trajetórias das condições iniciais e os pontos de equilíbrio e, finalmente, interpretar os resultados em termos do nível de cooperação da população.

4.4.1 Três Estratégias

Sejam x_C , x_D , x_T e x_G a frequência de indivíduos jogando as estratégias ALLC, ALLD, TFT e GTFT, respectivamente. Escolhendo-se as estratégias três a três, calculou-se os pontos fixos e a evolução de diversas condições iniciais foi desenhada; em outras palavras, o diagrama de fluxo foi obtido, como mostrado na figura 4.1. O ganho de uma estratégia¹ contra a outra foi calculado usando a equação 4.22 e o ganho médio, interpretado como aptidão, foi calculado usando a equação 2.36. Para cada trio de estratégias a equação do replicador 2.38

$$\dot{x}_i = x_i(f_i - \phi)$$

foi resolvida, para várias condições iniciais. Todas as equações diferenciais deste trabalho foram resolvidas numericamente pelo método de Runge-Kutta de quarta ordem, com um incremento de 0,001. Os pontos críticos foram calculados usando-se a versão 7.0 do *software* Wolfram Mathematica e os fluxos foram desenhados utilizando-se o *software* Dynamo [38].

¹Aqui, e em diversas partes do texto, foi cometido este abuso de linguagem: não é a *estratégia* quem obtém um ganho, mas sim um *indivíduo* que adota determinada estratégia que acumula ganhos. Portanto, todas as vezes em que for dito que uma *estratégia obteve um ganho*, entenda-se que um *jogador adotando tal estratégia* obteve um ganho.

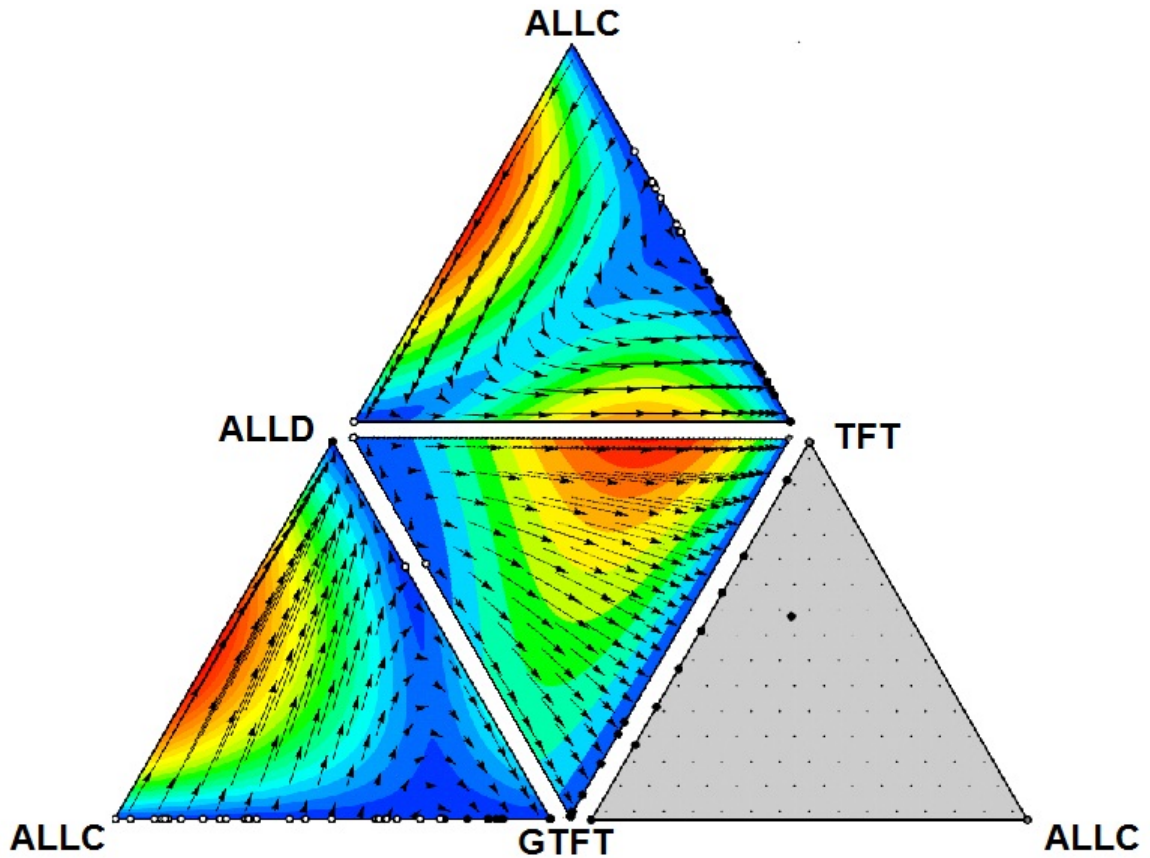


FIGURA 4.1: Diagrama de fluxos das frações dos indivíduos jogando as estratégias ALLC, ALLD, TFT e GTFT, tomadas de três a três. A versão de Taylor da equação do replicador foi resolvida pelo método de Runge-Kutta de quarta ordem para várias condições iniciais. O fluxo é mais rápido onde as cores são avermelhadas e mais lento nas cores azuladas. As figuras foram obtidas utilizando-se o *software* Dynamo [38].

Os vértices de cada diagrama de fluxo representam populações puras, isto é, com a frequência de um dos indivíduos igual a 1. As linhas representam populações cuja frequência de um dos indivíduos é igual a zero. Os pontos interiores tem todas as frequências de indivíduos estritamente positivas. Em todos os casos os vértices dos simplexes são pontos fixos, pois, uma vez que o sistema está nessa condição inicial, ali ele permanecerá para sempre, já que apenas uma estratégia está inicialmente presente e não ocorre mutação no processo de reprodução dos indivíduos.

O caso $GTFT \times TFT \times ALLC$ é ilustrado pelo simplex da direita, de cor cinza. Na primeira rodada, ALLC, TFT e GTFT cooperam², e o farão em todas

²Novamente outro abuso de linguagem recorrente: quando for dito que *ALLC coopera*, por exemplo, entenda-se que *jogadores que adotam a estratégia ALLC cooperam*. De maneira geral, quando uma ação for atribuída a uma estratégia, quer-se dizer que os indivíduos que jogam aquela estratégia são quem praticam a determinada ação.

as outras rodadas. O ganho de cada estratégia é o mesmo, e, por isso, não existe fluxo nesse plano; a equação do replicador para cada estratégia é a mesma e todos os pontos são fixos.

O próximo caso é ALLC \times TFT \times ALLD, ilustrado pelo simplex superior. A linha entre as estratégias TFT e ALLC é uma linha fixa. O fluxo dos pontos interiores vai, inicialmente, em direção a ALLD, e então muda continuamente em direção à linha fixa. Vale ressaltar que apenas parte da linha fixa é atrativa, que é a parte mais próxima de TFT.

O caso GTFT \times TFT \times ALLD é ilustrado no simplex central. Desta vez a linha entre as estratégias GTFT e TFT é fixa. O ponto dado por $x_D = 1/3$, $x_G = 2/3$ e $x_T = 0$ também é um ponto fixo e é instável. Todos os pontos interiores do simplex convergem para a linha fixa.

Por fim, o caso GTFT \times ALLD \times ALLC é ilustrado pelo simplex da esquerda. Existe uma linha fixa entre as estratégias ALLC e GTFT e um ponto fixo instável dado por $x_D = 1/3$, $x_G = 2/3$ e $x_C = 0$. Parte dos pontos interiores converge para o vértice no qual somente ALLD existe e a outra parte converge para o vértice no qual somente GTFT existe.

Até o momento todas as soluções numéricas foram obtidas usando-se estratégias perfeitas, isto é, sem nenhuma perturbação. É também útil conhecer a evolução das frequências dos jogadores utilizando as estratégias mencionadas num universo propenso a erros, pois, no mundo real, ocorrências desse tipo são bastante plausíveis. Para isso foram usados valores de p e q para as estratégias perturbadas que diferem de, aproximadamente, 10% dos valores de p e q das estratégias ditas perfeitas. Portanto, ao invés de usar-se (1, 0) para TFT, utilizou-se (0,9; 0,1) para a sua versão perturbada, denominada PTFT. As estratégias perturbadas PALLC e PALLD são descritas por (0,9; 0,9) e (0,1; 0,1), respectivamente. A estratégia GTFT pode ser perturbada de duas maneiras, dependendo se q é maior ou menor que 1/3: PGTFT1, descrita por (0,9; 0,3), e PGTFT2, descrita por (0,9; 0,4). As frequências dos indivíduos jogando as estratégias PALLC, PALLD, PTFT e PGTFT1 são designadas por x_{PC} , x_{PD} , x_{PT} e x_{PG1} , respectivamente. O diagrama de fluxos das estratégias perturbadas está ilustrado na figura 4.2. Como ocorre no caso das estratégias sem perturbação, os vértices dos simplexes são pontos fixos.

O primeiro caso é PALLC \times PTFT \times PGTFT1 e está ilustrado no simplex da direita. Em comparação com o caso não perturbado (figura 4.1), nota-se que o

plano fixo deixa de existir, permanecendo apenas cinco pontos fixos: três deles são os vértices do simplex, sendo PGTFT1 estável ao longo da linha PGTFT1-PTFT. Os outros dois pontos são $(x_{PC} \simeq 0,72, x_{PT} \simeq 0,28, x_{PG1} = 0)$, denotado por PF1), que é estável ao longo da linha PTFT-PALLC, e $(x_{PC} \simeq 0,13, x_{PT} = 0, x_{PG1} \simeq 0,87)$, denotado por PF2), que é globalmente estável, já que o fluxo de todos os pontos interiores converge para esse ponto. Se substituir-se PGTFT1 por PGTFT2, o ponto fixo globalmente estável passa a ser $(x_{PC} = 0, x_{PT} \simeq 0,06, x_{PG2} \simeq 0,94)$.

O caso PALLC \times PALLD \times PTFT está representado no simplex superior. Fazendo uma comparação com o caso em que as estratégias não estão perturbadas, nota-se que a linha fixa desaparece. PALLD é um ponto fixo atrator; PTFT é estável ao longo da linha PALLD-PTFT; PALLC é instável. Três novos pontos fixos apareceram: (a) o atrator $(x_{PC} \simeq 0,24, x_{PD} \simeq 0,54, x_{PT} \simeq 0,22)$; (b) o

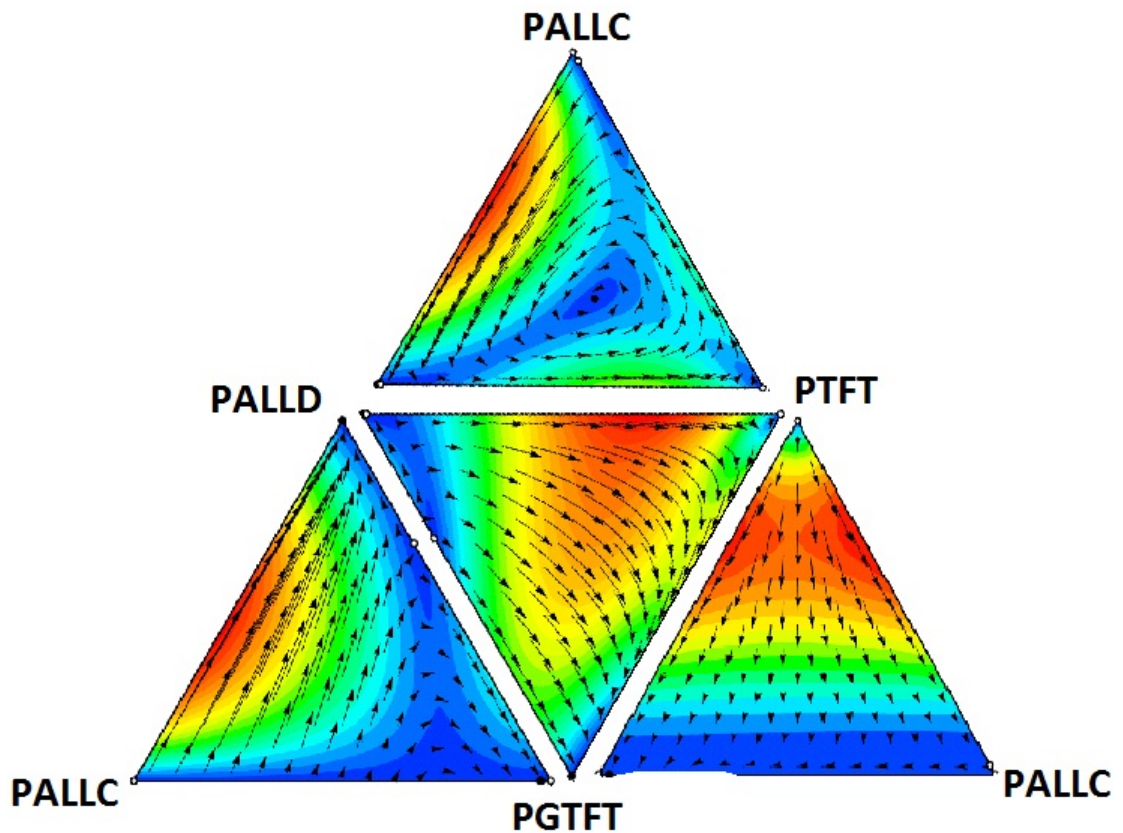


FIGURA 4.2: Diagrama de fluxos das frações dos indivíduos jogando as estratégias PALLC, PALLD, PTFT e PGTFT1, tomadas de três a três. A versão de Taylor da equação do replicador foi resolvida pelo método de Runge-Kutta de quarta ordem para várias condições iniciais. O fluxo é mais rápido onde as cores são avermelhadas e mais lento nas cores azuladas. As figuras foram obtidas utilizando-se o *software* Dynamo [38].

ponto fixo instável ($x_{PC} = 0$, $x_{PD} \simeq 0,11$, $x_{PT} \simeq 0,89$, denotado por PF3); (c) e o ponto fixo PF1. Parte dos pontos interiores tem o fluxo convergindo para um dos atratores, e os pontos interiores restantes convergem para o outro atrator.

O caso $PALLC \times PALLD \times PGTFT1$ é mostrado no simplex da esquerda. Dos pontos fixos que estão nos vértices do simplex, somente $PALLD$ permanece estável (atrator). $PALLC$ é instável e $PGTFT1$ é estável somente ao longo da linha $PALLD$ - $PGTFT1$. O antigo ponto fixo instável é substituído por ($x_{PC} = 0$, $x_{PD} \simeq 0,58$, $x_{PG1} \simeq 0,42$, denotado por PF4) e um novo ponto fixo atrator, PF2, surge. Novamente, parte dos pontos interiores tem o fluxo convergindo para um dos atratores, e os pontos interiores restantes convergem para o outro atrator. Se $PGTFT1$ for substituído por $PGTFT2$, o fluxo praticamente permanece inalterado e o novo ponto fixo atrator que havia surgido passa a ser $(0,3; 0; 0,7)$.

Finalmente, o caso $PALLD \times PTFT \times PGTFT1$ é ilustrado no simplex central. A antiga linha fixa, presente no caso em que não há perturbação, deixa de existir. Um novo ponto fixo instável PF3 surge. O antigo ponto fixo na linha $PGTFT1$ - $PALLD$ é substituído por PF4 e permanece instável. $PGTFT1$ e $PALLD$ são atratores e $PTFT$ é estável somente ao longo da linha $PALLD$ - $PTFT$. Mais uma vez parte dos pontos interiores tem o fluxo convergindo para um dos atratores, e os pontos interiores restantes convergem para o outro atrator. Se $PGTFT1$ for substituído por $PGTFT2$, o fluxo é pouco alterado; o ponto fixo instável é agora $(0; 0,21; 0,79)$ e $PGTFT2$ deixa de ser um atrator, sendo substituído por $(0,0002; 0,0517; 0,9481)$. Todos os resultados apresentados estão resumidos na tabela 4.1.

Utilizando-se a equação do replicador de Maynard Smith (2.39) ao invés da de Taylor (2.38), os resultados são muito similares aos já mostrados: os pontos fixos são exatamente os mesmos e ocorrem pequenas alterações no fluxo.

Observa-se que, na maioria dos casos, o comportamento final da população é caracterizado pela cooperação. Quando a perturbação não está presente, os casos que incluem a estratégia $ALLD$ tem a seguinte característica: todos os pontos interiores evoluem, inicialmente, na direção de $ALLD$, pois no início a sua população aumenta devido à exploração das estratégias cooperadoras. Esse crescimento da população de desertores é freado por eles próprios: a quantidade de indivíduos jogando $ALLD$ é tão grande que a quantidade de jogadores de outras estratégias é pequena demais para ser explorada. Nesse momento as outras estratégias tem um ganho médio maior do que $ALLD$ e fazem com que o fluxo mude sua direção,

Estratégias presentes	PF estáveis	PF instáveis
ALLC × TFT × GTFT	Plano fixo	
PALLC × PTFT × PGTFT1	(0; 0; 1), estável ao longo da linha PGTFT1-PTFT; (0,72; 0,28; 0) (0,13; 0; 0,87), atrator.	(1; 0; 0) (0; 1; 0)
ALLC × ALLD × TFT	(0; 1; 0), estável ao longo da linha ALLC-ALLD; a linha fixa ALLC-TFT; (0; 0; 1), ao longo da linha ALLC-ALLD.	(1; 0; 0)
PALLC × PALLD × PTFT	(0; 0; 1), ao longo da linha PALLD-PTFT; (0; 1; 0); (0,24; 0,54; 0,22), atrator; (0,72; 0; 0,28).	(0; 0,89; 0,11) (1; 0; 0)
ALLD × TFT × GTFT	(1; 0; 0), ao longo da linha ALLD-GTFT; (0; 1; 0), ao longo da linha ALLD-TFT; a linha fixa TFT-GTFT; (0; 0; 1), ao longo da linha ALLD-GTFT.	(1/3; 0; 2/3)
PALLD × PTFT × PGTFT1	(0; 0; 1), atrator; (1; 0; 0), atrator; (0; 1; 0), ao longo da linha PALLD-PTFT.	(0,58; 0; 0,42) (0,89; 0,11; 0)
ALLC × ALLD × GTFT	a linha fixa ALLC-GTFT; (0; 1; 0), atrator; (0; 0; 1), atrator.	(1; 0; 0) (0; 1/3; 2/3)
PALLC × PALLD × PGTFT1	(0; 1; 0), atrator; (0,13; 0; 0,87), atrator; (0; 0; 1), ao longo da linha PALLD-PGTFT1.	(0; 0,58; 0,42) (1; 0; 0)

TABELA 4.1: Pontos fixos do confronto das estratégias ALLC, ALLD, TFT e GTFT, três a três, bem como de suas versões perturbadas.

indo de encontro a uma linha fixa na qual coexistam duas estratégias cooperadoras (ALLC, TFT ou GTFT). A exceção ocorre quando é formada uma população de jogadores que usam as estratégias ALLC, ALLD e GTFT. Apenas algumas condições iniciais levam à vitória de ALLD, já que é fácil para os desertores explorarem cooperadores na presença da estratégia GTFT. A maioria das condições iniciais é levada para a linha fixa (mistura das estratégias puras ALLC e GTFT),

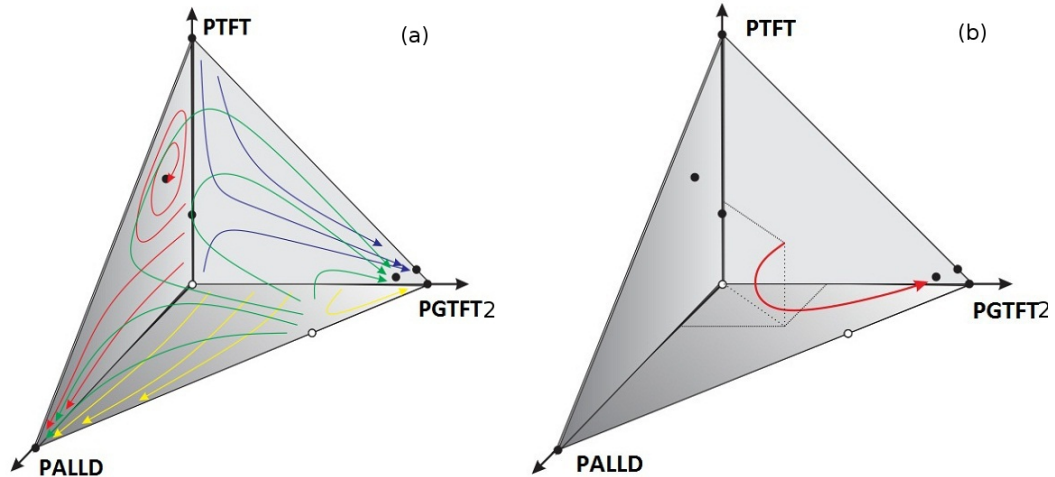


FIGURA 4.3: Diagrama de fluxos no simplex com quatro estratégias. A origem dos eixos coordenados representa a estratégia PALLC. Os fluxos em cada plano são mostrados em (a) e em (b) observa-se a evolução da condição inicial homogênea, que converge para o estado final dominado por PGTFT2. A versão de Taylor da equação do replicador foi resolvida pelo método de Runge-Kutta de quarta ordem.

implicando a vitória da cooperação.

Esse cenário é praticamente o mesmo quando estratégias perturbadas são usadas. Com exceção do caso $PALLC \times PTFT \times PGTFT1$, em todos os outros existem condições iniciais que levam à vitória de PALLD. Essas situações correspondem a uma fração pequena das condições iniciais, e a maioria delas converge para um ponto no qual só há indivíduos que jogam estratégias cooperadoras (mistura das estratégias perturbadas PALLC, PTFT e PGTFT).

4.4.2 Quatro Estratégias

A análise agora será feita utilizando-se as quatro estratégias. Novamente, a equação 4.22 fornece os ganhos de cada estratégia. O ganho médio é calculado e esses valores são inseridos na equação do replicador. Para resolver esse sistema de quatro equações diferenciais ordinárias utilizou-se ainda o método de Runge-Kutta de quarta ordem, com incremento igual a 0,001.

Comparando-se com o caso em que apenas três estratégias são colocadas juntas, nenhum novo ponto fixo surge, para ambas as situações em que as estratégias são perturbadas ou não. O simplex agora tem a forma de um tetraedro, e cada uma de suas faces são os simplexes (planos) analisados na subseção 4.4.1. Como

nenhum ponto interior é ponto de equilíbrio, todas as trajetórias convergem para as faces do simplex.

Nesta subseção, em particular, o termo “condição inicial” fará alusão a um ponto fora das faces do simplex, isto é, um ponto interior. No caso de quatro estratégias, o simplex é um tetraedro. Os infinitos pontos que compõem o tetraedro são possíveis condições iniciais do problema. O volume do tetraedro foi dividido em 166650 partes, isto é, observou-se a evolução temporal de 166650 condições iniciais diferentes. Na ausência de perturbação, todas as condições iniciais são levadas para algum ponto do plano fixo no qual $x_D = 0$. Já no caso perturbado, aproximadamente 82% das condições iniciais são levadas para o ponto $(x_{PC} \simeq 0,13, x_{PD} = 0, x_{PT} = 0, x_{PG1} \simeq 0,87)$ e os 18% restantes são levados para ALLD. Usando-se PGTFT2 ao invés de PGTFT1, 73% das condições iniciais convergem para o ponto $(x_{PC} = 0, x_{PD} \simeq 0,0517, x_{PT} \simeq 0,0002, x_{PG1} \simeq 0,9481)$ e os 27% restantes vão para ALLD. A figura 4.3(a) mostra o fluxo em cada uma das faces do tetraedro e a figura 4.3(b) mostra, como exemplo ilustrativo, a evolução a da condição inicial chamada de homogênea, na qual todas as estratégias são igualmente abundantes no instante inicial. Ambas as figuras são para as estratégias perturbadas.

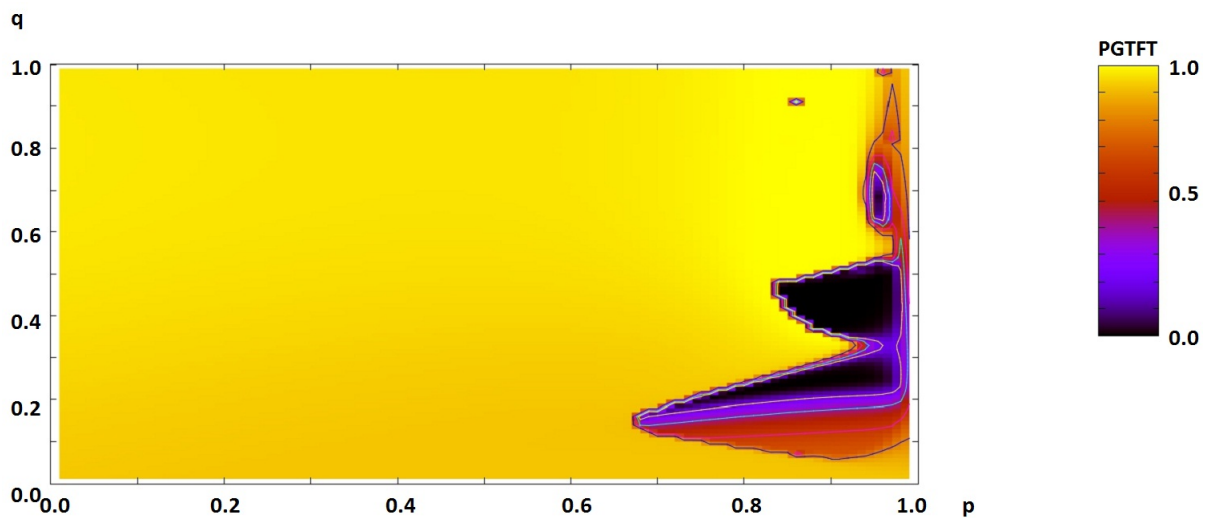


FIGURA 4.4: Fração final de PGTFT quando quatro estratégias estão fixas e a quinta varre o quadrado $[0,01; 0,99] \times [0,01; 0,99]$. A evolução temporal é descrita pela equação do replicador de Taylor.

4.4.3 Cinco Estratégias

De maneira geral, o dilema do prisioneiro com três e quatro estratégias é dominado pela cooperação, exceto para algumas condições iniciais que levam o sistema para um estado no qual a estratégia ALLD se estabelece. O próximo passo é descobrir o que acontece quando novas estratégias são adicionadas ao jogo; especificamente, deseja-se saber como a frequência final de indivíduos semelhantes a GTFT se comporta quando uma quinta estratégia está disponível. Como existem infinitas condições iniciais no problema, a partir de agora as análises estarão focadas na condição inicial homogênea, na qual todas as frequências iniciais são iguais.

O espaço de estratégias está agora restrito ao intervalo $\{(p, q) \in \mathbb{R} \mid \varepsilon \leq p, q \leq 1 - \varepsilon\}$, onde $\varepsilon = 0,01$, de modo que agora PALLC passa a ser $(0,99; 0,99)$, PALLD é $(0,01; 0,01)$, PTFT é $(0,99; 0,01)$ e PGTFT é $(0,99; 0,33)$. Para cada quinta estratégia (p, q) no quadrado $[0,01; 0,99] \times [0,01; 0,99]$ que é adicionada, foi calculada a frequência final da estratégia PGTFT utilizando-se a equação do replicador de Taylor 2.38 e de Maynard 2.39. O método de Runge-Kutta de quarta ordem foi utilizado para resolver numericamente as equações diferenciais ordinárias. Os resultados estão apresentados nas figuras 4.4 e 4.5.

Como pode ser observado nos gráficos, na maioria dos casos PGTFT é a estratégia vencedora; quando isso não ocorre, a quinta estratégia adicionada é

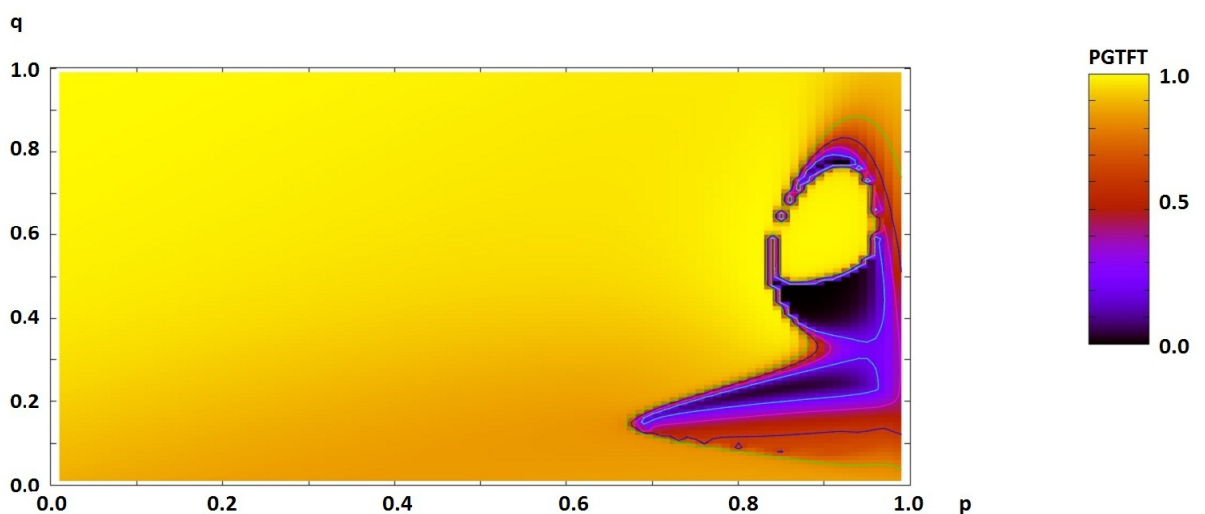


FIGURA 4.5: Fração final de PGTFT quando quatro estratégias estão fixas e a quinta varre o quadrado $[0,01; 0,99] \times [0,01; 0,99]$. A evolução temporal é descrita pela equação do replicador de Maynard Smith.

quem vence. Em alguns poucos casos PALLD é capaz de explorar os cooperadores e se estabelece como a estratégia campeã. Se essas exceções forem descartadas, o cenário inteiro é dominado pela cooperação.

Numa visão geral, quando uma pequena quantidade de estratégias (três, quatro e cinco, tipicamente) competem entre si, é fácil para a cooperação se estabelecer.

4.5 O Estabelecimento da Cooperação Quando Existem Muitas Estratégias

Nesta seção será estudado o que acontece quando muitas estratégias (mais de cinco) estão presentes no jogo. À medida que número de estratégias aumenta, torna-se cada vez mais complicado estudar todas as condições iniciais. Por isso, a partir de agora, o foco será somente na condição inicial homogênea, como já foi feito para o caso de cinco estratégias.

Será estudada a evolução de n estratégias no espaço de estratégias reativas dado por $\{(p, q) \in \mathbb{R} \mid \varepsilon \leq p, q \leq 1 - \varepsilon\}$, sendo $\varepsilon = 0,01$. Cada estratégia (p_i, q_j) é definida numa grade da seguinte forma:

$$(p_i, q_j) = \left(\varepsilon + i \frac{1 - 2\varepsilon}{d}, \varepsilon + j \frac{1 - 2\varepsilon}{d} \right), \quad (4.24)$$

onde $i, j = 0, 1, 2, \dots, d$ e $d \in \mathbb{N}^*$. Para $d = 0$ as estratégias presentes no jogo são $(0,01; 0,01)$, $(0,01; 0,99)$, $(0,99; 0,01)$ e $(0,99; 0,99)$; quando $d = 1$, por exemplo, as estratégias presentes são $(0,01; 0,01)$, $(0,01; 0,50)$, $(0,01; 0,99)$, $(0,50; 0,01)$, $(0,50; 0,50)$, $(0,50; 0,99)$, $(0,99; 0,01)$, $(0,99; 0,50)$ e $(0,99; 0,99)$, e assim sucessivamente. Dessa forma, d é o parâmetro que controla o número de estratégias no jogo, dado por $n = (d + 2)^2$.

A análise da evolução das frequências das estratégias foi feita utilizando-se as equações do replicador de Taylor e Maynard Smith (equações 2.38 e 2.39). Anteriormente, a equação do replicador havia sido resolvida pelo método de Runge-Kutta de quarta ordem. No entanto, a forma mais simples de se obter, de forma aproximada, a evolução temporal de cada x_i é discretizando a equação do replicador. Entretanto, deve-se ter em mente que as versões contínuas e discretas

dessas equações podem apresentar comportamentos completamente distintos. As versões discretas da equação do replicador de Taylor e de Maynard Smith são, respectivamente,

$$x_i(t + 1) = x_i(t) + x_i(t)[f_i(t) - \phi(t)] , \tag{4.25}$$

$$x_i(t + 1) = \frac{x_i(t)f_i(t)}{\phi(t)} , \tag{4.26}$$

onde t é agora uma variável discreta que assume valores inteiros.

Enfim, a trajetória evolutiva das frequências das estratégias foi estudada sob quatro óticas distintas: as equações do replicador de Taylor e Maynard Smith, em suas versões contínuas e discretas. Os resultados estão sintetizados na figura 4.6.

As simulações foram feitas para valores de d indo de 1 até 18. A equação de Taylor discreta apresenta soluções que divergem, sem qualquer significado físico. Os resultados para $d = 1$ não são apresentados na figura 4.6. O que ocorre é que, para todas as versões da equação do replicador (exceto para a versão discreta da equação de Taylor, que diverge), o cenário final é composto por um equilíbrio dinâmico entre as estratégias (0,99; 0,01) e (0,01; 0,50), que oscilam na liderança. Como pode ser verificado na figura 4.7, ALLD começa vencendo, mas logo é substituído por TFT, que alterna com a estratégia (0,01; 0,50).

Dependendo da versão da equação do replicador escolhida, a população pode evoluir na direção de um estado caracterizado pela deserção, com o domínio da estratégia similar a ALLD, ou pode ir em direção a um estado cuja estratégia dominante é aquela mais próxima de GTFT. A cooperação sobrevive na versão contínua da equação do replicador de Taylor somente quando para valores de d tais que $d \leq 6$; caso contrário, ALLD domina. Tanto para a versão contínua da

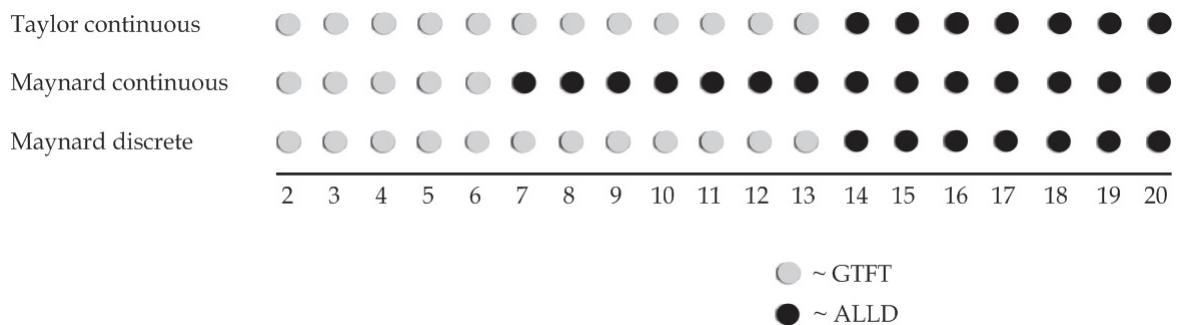


FIGURA 4.6: Estratégia vencedora, dependendo do valor de d e da versão da equação do replicador utilizada.

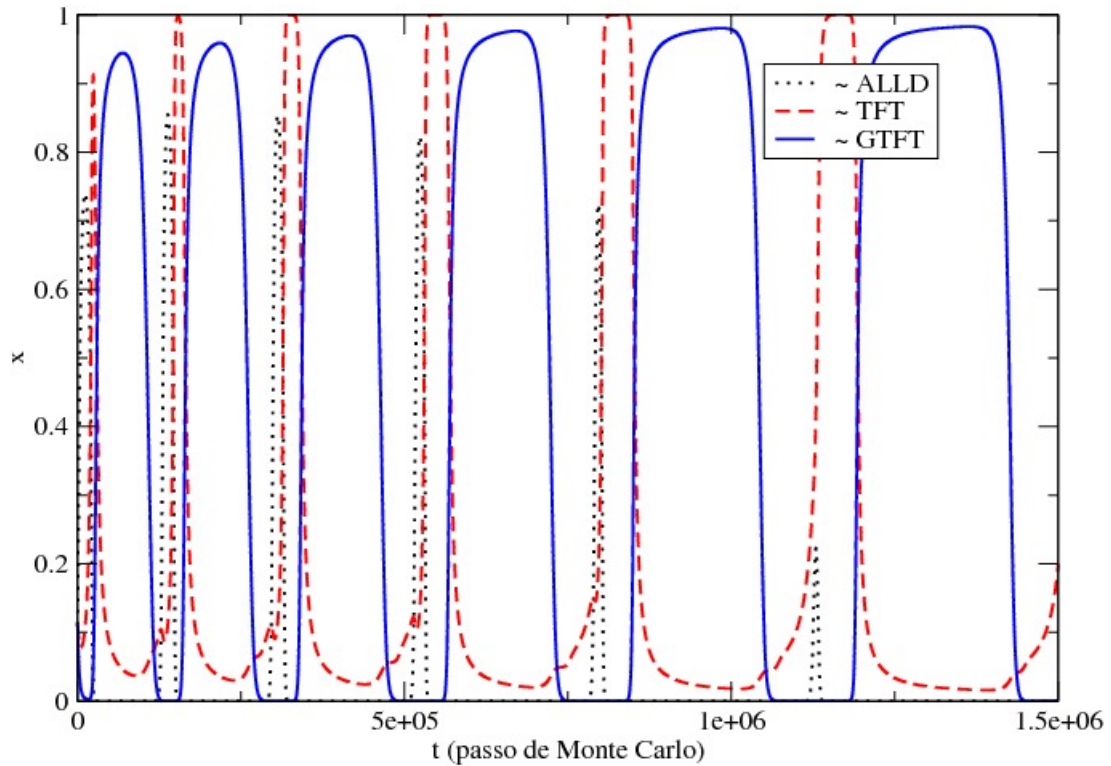


FIGURA 4.7: Evolução das principais estratégias para o caso $d = 1$, isto é, 9 estratégias, utilizando-se a equação do replicador de Taylor contínua.

equação de Taylor quanto para a versão discreta da equação de Maynard Smith, a cooperação sobrevive para valores de d tais que $d \leq 13$.

Quando GTFT vence, a trajetória típica das frequências das estratégias no simplex S_n é como descrita na seção 4.3. Inicialmente a frequência das estratégias similares a ALLD aumenta, seguido pelo aumento das estratégias parecidas com TFT, que, finalmente, são substituídas pela estratégia mais próxima de GTFT. Um exemplo é apresentado na figura 4.8.

A transição crucial é de ALLD para TFT: a estrutura do ganho entre ALLD e TFT é tal que a melhor estratégia é fazer o mesmo que o oponente fez. Numa população de ALLD e TFT, TFT pode se espalhar somente se existem indivíduos suficientes que adotam a estratégia TFT. Quando a população sai do estado no qual a maioria dos indivíduos joga ALLD e passa o estado no qual a maioria adota TFT, a transição ocorre ao longo da linha (p, ε) . A frequência das estratégias cujos valores de p são levemente maiores aumenta, enquanto que a frequência das estratégias cujos valores de p são levemente menores diminui.

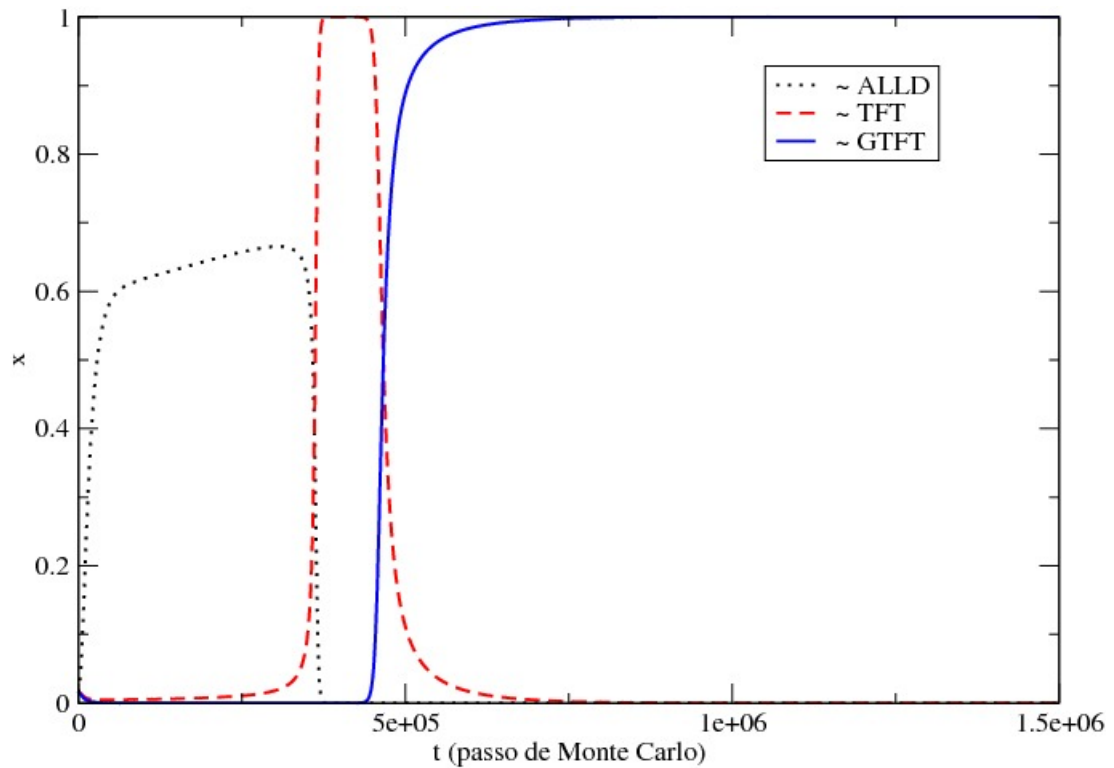


FIGURA 4.8: Evolução das principais estratégias para o caso $d = 4$, isto é, 36 estratégias, utilizando-se a equação do replicador de Maynard Smith contínua.

Para compreender o que ocorre quando a densidade de estratégias no jogo aumenta, será apresentado um argumento simples. Considere o jogo entre as estratégias (p, ε) e $(p + \Delta p, \varepsilon)$. Seja A a estratégia (p, ε) e A' a estratégia $(p + \Delta p, \varepsilon)$. A variável Δp simboliza o quão densa é a grade de estratégias: se Δp é grande, duas estratégias consecutivas estão muito separadas, indicando que existem poucas estratégias no jogo; por outro lado, se Δp é pequeno, as estratégias estão bem próximas e a grade de estratégias é densa. O ganho associado a cada um dos confrontos entre jogadores que adotam essas duas estratégias é, em primeira ordem em ε :

(i) $A \times A$

$$\begin{aligned} s &= \frac{\varepsilon(p - \varepsilon) + \varepsilon}{1 - (p - \varepsilon)^2} \\ &\simeq \frac{\varepsilon p + \varepsilon}{1 - p^2 + 2p\varepsilon} \\ &\simeq \frac{\varepsilon}{1 - p} \left(1 - \frac{2p\varepsilon}{1 - p^2} \right) \simeq \frac{\varepsilon}{1 - p}, \end{aligned} \quad (4.27)$$

$$\begin{aligned} E(A, A) &\simeq Ts + p - Ps - Ps = 5s + 1 - s - s \\ &= 1 + 3s = 1 + \frac{3\varepsilon}{1 - p}. \end{aligned} \quad (4.28)$$

(ii) $A' \times A'$

Analogamente ao item (i):

$$s' = \frac{\varepsilon}{1 - (p + \Delta p)}, \quad (4.29)$$

$$E(A', A') = 1 + \frac{3\varepsilon}{1 - (p + \Delta p)}. \quad (4.30)$$

(iii) $A \times A'$

$$\begin{aligned} s &= \frac{\varepsilon(p - \varepsilon) + \varepsilon}{1 - (p - \varepsilon)(p + \Delta p - \varepsilon)} \\ &\simeq \frac{\varepsilon p + \varepsilon}{1 - p(p + \Delta p) + \varepsilon p + \varepsilon(p + \Delta p)} \\ &\simeq \frac{\varepsilon(1 + p)}{1 - p(p + \Delta p)} \left[1 - \frac{\varepsilon(1 + p + \Delta p)}{1 - p(p + \Delta p)} \right] \\ &\simeq \frac{\varepsilon(1 + p)}{1 - p(p + \Delta p)}, \end{aligned} \quad (4.31)$$

$$\begin{aligned} s' &= \frac{\varepsilon(p + \Delta p - \varepsilon) + \varepsilon}{1 - (p - \varepsilon)(p + \Delta p - \varepsilon)} \\ &= \frac{\varepsilon(1 + p + \Delta p)}{1 - p(p + \Delta p) + \varepsilon(2p + \Delta p)} \\ &\simeq \frac{\varepsilon(1 + p + \Delta p)}{1 - p(p + \Delta p)}, \end{aligned} \quad (4.32)$$

$$\begin{aligned} E(A, A') &\simeq Ts' + P - Ps' - Ps = 5s' + 1 - s' - s \\ &= 1 + 4s' - s \\ &= 1 + \frac{4\varepsilon(1 + p + \Delta p)}{1 - p(p + \Delta p)} - \frac{\varepsilon(1 + p)}{1 - p(p + \Delta p)} \\ &= 1 + \frac{\varepsilon(3 + 3p + 4\Delta p)}{1 - p(p + \Delta p)}. \end{aligned} \quad (4.33)$$

(iv) $A' \times A$

$$\begin{aligned}
 s &= \frac{\varepsilon(1+p)}{1-p(p+\Delta p)}, \\
 s' &= \frac{\varepsilon(1+p+\Delta p)}{1-p(p+\Delta p)}, \\
 E(A', A) &\simeq Ts + p - Ps - Ps' \\
 &= 5s + 1 - s - s' = 1 + 4s - s' \\
 &= 1 + \frac{4\varepsilon(1+p)}{1-p(p+\Delta p)} - \frac{\varepsilon(1+p+\Delta p)}{1-p(p+\Delta p)} \\
 &= 1 + \frac{\varepsilon(3+3p-\Delta p)}{1-p(p+\Delta p)}. \tag{4.34}
 \end{aligned}$$

Como a matriz de ganho das estratégias A e A' é da forma

$$\begin{array}{c|cc} & A & A' \\ \hline A & E(A, A) & E(A, A') \\ A' & E(A', A) & E(A', A') \end{array}, \tag{4.35}$$

Para que a estratégia $(p + \Delta p, \varepsilon)$ domine a estratégia (p, ε) , tem-se que

$$\begin{aligned}
 E(A', A) &> E(A, A) \\
 1 + \frac{\varepsilon(3+3p-\Delta p)}{1-p(p+\Delta p)} &> 1 + \frac{3\varepsilon}{1-p} \\
 p &> \frac{1}{4}
 \end{aligned}$$

e

$$\begin{aligned}
 E(A', A') &> E(A, A') \\
 1 + \frac{3\varepsilon}{1-p-\Delta p} &> 1 + \frac{\varepsilon(3+3p+4\Delta p)}{1-p(p+\Delta p)} \\
 p + \Delta p &> \frac{1}{4}.
 \end{aligned}$$

Ou seja, exige-se simplesmente que $p > 1/4$. Agora, para que a estratégia $(p + \Delta p, \varepsilon)$ seja dominada pela estratégia (p, ε) , tem-se que

$$\begin{aligned}
 E(A', A) &< E(A, A) \\
 p &< \frac{1}{4}
 \end{aligned}$$

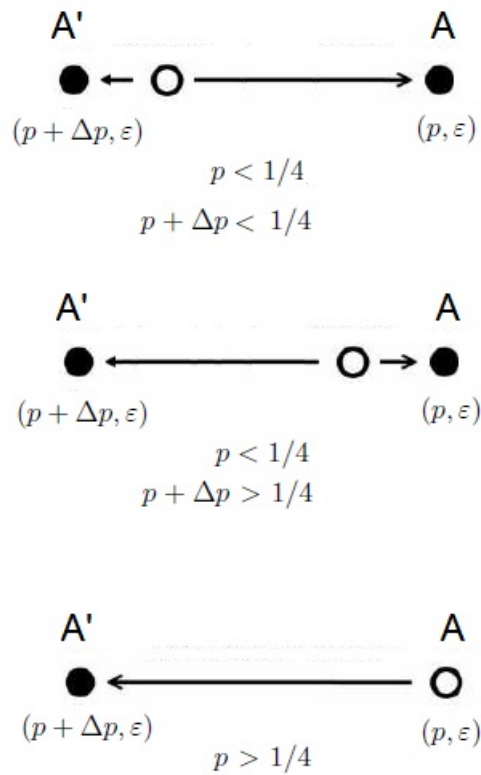


FIGURA 4.9: Ilustração da transição de ALLD para TFT ao longo da linha de estratégias (p, ε) .

e

$$E(A', A') < E(A, A')$$

$$p + \Delta p < \frac{1}{4} .$$

Finalmente, para que ambas as estratégias sejam equilíbrio de Nash, tem-se que

$$E(A', A) < E(A, A)$$

$$p < \frac{1}{4}$$

e

$$E(A', A') > E(A, A')$$

$$p + \Delta p > \frac{1}{4} ,$$

sendo que o ponto de equilíbrio instável (frequência da estratégia A) entre as duas estratégias é, utilizando-se a expressão encontrada na subseção 2.2.1,

$$\begin{aligned} x^* &= \frac{E(A', A') - E(A, A')}{E(A, A) - E(A, A') - E(A', A) + E(A', A')} \\ &= \frac{(1-p)[4(p+\Delta p) - 1]}{3\Delta p} . \end{aligned} \quad (4.36)$$

Note que

$$\frac{\partial x^*}{\partial \Delta p} = \frac{(1-p)(1-4p)}{3(\Delta p)^2} \quad (4.37)$$

é sempre positiva para $p < 1/4$, ou seja, x^* é crescente com Δp .

Considere, inicialmente, que $p < 1/4$. Se $p + \Delta p > 1/4$, ou seja, existem poucas estratégias no jogo, o sistema pode ser atraído tanto para ALLD quanto para TFT, dependendo do valor de Δp . Quanto maior for Δp , maior é x^* , implicando que a bacia de atração de $(p + \Delta p, \varepsilon)$ também aumenta. Esse processo desloca o sistema para a dominância de estratégias cujo p é cada vez maior. Uma vez que p passe a ser maior que $1/4$, o sistema vai em direção a TFT. Mas à medida que a densidade de estratégias aumenta, Δp diminui, aumentando a bacia de atração da estratégia (p, ε) , fazendo com que o sistema convirja para ALLD. A figura 4.9 ilustra a explicação. Logo, conclui-se que a cooperação consegue se estabelecer somente quando a densidade de estratégias no jogo não é muito grande. Se esse é o caso, a estratégia mais próxima de ALLD começa ganhando, mas logo TFT desempenha seu papel como catalisador da cooperação. Como TFT não é bem sucedida num ambiente onde erros ocorrem, a estratégia mais próxima de GTFT firma-se como a vencedora graças à sua capacidade de poder cooperar mesmo quando o oponente tenha desertado.

Também estudou-se a evolução da cooperação num ambiente de estratégias reativas cujos valores de p e q foram sorteados aleatoriamente. Como os resultados são extremamente sensíveis às estratégias sorteadas, nenhuma conclusão definitiva pode ser tirada até o momento.

Capítulo 5

Conclusões

Para entender como a emergência e manutenção da cooperação ocorre numa população de indivíduos que podem cooperar ou desertar, foram analisadas a evolução temporal da frequências de estratégias diferentes, dependendo da quantidade de estratégias presentes.

A evolução da cooperação foi analisada em termos de quatro equações dinâmicas diferentes (equações do replicador de Taylor e Maynard Smith, nas suas versões contínua e discreta), cada uma fornecendo resultados distintos. Quando a quantidade de estratégias é pequena (três, quatro e cinco estratégias), a cooperação pode, de maneira relativamente fácil, dominar o cenário, com exceção de poucos casos — aqueles cujas condições iniciais favorecem a deserção.

Em se tratando do caso no qual as estratégias estão perturbadas, vários fluxos foram desenhados e o estado final é basicamente a cooperação, representada por PGTFT, ou a deserção, representada por PALLD, dependendo de onde a condição inicial está localizada. Nota-se que a bacia de atração de PGTFT é maior do que a bacia de atração de PALLD.

À medida que o número de estratégias aumenta e utilizando-se a condição inicial homogênea, verificou-se que a cooperação consegue se estabelecer apenas para uma quantidade de estratégias abaixo de um valor limite, que depende da versão do replicador que está sendo utilizada ($d \leq 6$, ou $n = 64$, para a versão de Maynard contínua e $d \leq 13$, ou $n = 225$, para as versões de Taylor contínua e de Maynard Smith discreta). A partir daí, a cooperação não mais se mantém e o

cenário é dominado pela deserção. Um argumento analítico, baseado em cálculos de equilíbrio de Nash, foi elaborado para explicar tal dinâmica.

É útil enfatizar que se pode obter resultados completamente distintos do mesmo problema (mesmas estratégias e mesma condição inicial), dependendo de qual versão da equação do replicador é adotada. Além disso, se a versão da equação do replicador escolhida for a de Taylor discreta, as frequências das estratégias divergem. Para estratégias localizadas numa grade, existe um valor máximo de estratégias que podem estar disponíveis no jogo para que os resultados descritos na literatura sejam obtidos. Se o número de estratégias não é muito grande, a cooperação emerge com TFT e é mantida por GTFT. Em suma, a cooperação não consegue se estabelecer em um ambiente no qual a densidade de estratégias é alta, diferentemente do que apontam os resultados da literatura. Para estratégias sorteadas aleatoriamente ao invés de escolhidas numa grade, conclusões gerais não puderam ainda ser obtidas.

Apêndice A

A Estratégia “win-stay, lose-shift”

No universo das estratégias estocásticas a decisão de cooperar ou desertar na rodada seguinte depende do cenário da rodada anterior (não somente do movimento do adversário, mas também da ação do próprio jogador). Cada estratégia pode ser definida em termos de quatro probabilidades condicionais: p_1, p_2, p_3 e p_4 são as probabilidades do jogador cooperar dado que o cenário da rodada passada foi CC, CD, DC e DD, respectivamente. Nesse contexto as conhecidas estratégias ALLC, ALLD, TFT e GTFT são representadas como $(1,1,1,1)$, $(0,0,0,0)$, $(1,0,1,0)$ e $(1,1/3,1,1/3)$, respectivamente. De forma semelhante às estratégias reativas, o confronto entre duas estratégias $A(p_1, p_2, p_3, p_4)$ e $A'(p'_1, p'_2, p'_3, p'_4)$ pode ser mapeado em uma cadeia de Markov, cuja matriz de transição é

$$M = \begin{pmatrix} p_1 p'_1 & p_1(1-p'_1) & (1-p_1)p'_1 & (1-p_1)(1-p'_1) \\ p_2 p'_3 & p_2(1-p'_3) & (1-p_2)p'_3 & (1-p_2)(1-p'_3) \\ p_3 p'_2 & p_3(1-p'_2) & (1-p_3)p'_2 & (1-p_3)(1-p'_2) \\ p_4 p'_4 & p_4(1-p'_4) & (1-p_4)p'_4 & (1-p_4)(1-p'_4) \end{pmatrix}. \quad (\text{A.1})$$

Como M é uma matriz estocástica e regular, ela possui um único autovetor $\vec{x} = (x_1, x_2, x_3, x_4)$ associado ao autovalor 1, que é a distribuição estacionária da cadeia de Markov [32]:

$$\vec{x} = \vec{x}M. \quad (\text{A.2})$$

Portanto, o ganho esperado para a estratégia A jogando contra a estratégia A' é

$$E(A, A') = Rx_1 + Sx_2 + Tx_3 + Px_4 . \quad (\text{A.3})$$

Nowak reproduziu o experimento citado na seção anterior afim de confirmar a vitória da estratégia GTFT. De fato, ela é a vencedora sob certas circunstâncias; mas, de maneira inesperada, e com ocorrência frequente, uma nova estratégia dominava o cenário: (1,0,0,1). Essa estratégia coopera se a última rodada foi CC ou DD, e deserta caso contrário. Isso significa que um jogador adotando tal estratégia repete seu movimento anterior caso ele tenha recebido um ganho alto (T ou R), mas muda sua jogada se o seu ganho for baixo (P ou S). Por isso a estratégia recebe o nome de “win-stay, lose-shift”, WSL (“vence-permanece, perde-troca”, numa tradução livre) [9, 39].

Referências Bibliográficas

- [1] A. Gardner S. A. West, A. S. Griffin. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J. Evolution Biol.*, (20): 415–432, 2007.
- [2] N. Henrich J. Henrich. Culture, evolution and the puzzle of human cooperation. *Cogn. Syst. Res.*, (7):220–245, 2006.
- [3] M. Casari M. Bigoni, G. Camera. Strategies of cooperation and punishment among students and clerical workers. *J. Econ. Behav. Organ.*, (94):172–182, 2013.
- [4] H. Kunreuther M. Gong, J. Baron. Why do groups cooperate more than individuals to reduce risks. *Theory Dec.*, (75):101–116, 2012.
- [5] C. Krumme D. G. Rand T. Pfeiffer, L. Tran. The value of reputation. *J. R. Soc. Interface*, (9):2791–2797, 2012.
- [6] R. Axelrod. *The evolution of cooperation*. Basic Books, New York, 1984.
- [7] G. Fáth G. Szabó. Evolutionary games on graphs. *Physics Reports*, (446): 97–216, 2007.
- [8] L. L. Wardil. *Mecanismos cooperativos: adotando estratégias diferentes contra oponentes distintos no dilema do prisioneiro*. PhD thesis, Universidade Federal de Minas Gerais, 2012.
- [9] M. A. Nowak. *Evolutionary dynamics: exploring equations of life*. The Belknap Press of Harvard University Press, Cambridge, 2006.
- [10] M. Nowak. Five rules for the evolution of cooperation. *Science*, (314):1560–1563, 2006.

-
- [11] K. Sigmund M. Nowak. Game-dynamical aspects of the prisoner's dilemma. *Appl. Math. Comput.*, (30):191–213, 1989.
- [12] K. Sigmund M. Nowak. Tit for tat in heterogeneous populations. *Nature*, (355):250–253, 1992.
- [13] K. Sigmund M. Nowak. Chaos and evolution of cooperation. *Proc. Natl. Acad. Sci. USA*, (90):5091–5094, 1993.
- [14] S. H. Strogatz. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry and engineering*. Perseus Books, New York, 1994.
- [15] O. Morgenstern J. von Neumann. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 2007.
- [16] J. F. Nash. Equilibrium points in n-person games. *P. Natl. Acad. Sci.*, (36):48–49, 1950.
- [17] L. B. Jonker P. D. Taylor. Evolutionary stable strategies and game dynamics. *Math. Biosci.*, (40):145–156, 1978.
- [18] A. M. Chammah A. Rapoport. *Prisoner's dilemma*. University of Michigan Press, Ann Arbor, 1965.
- [19] G. S. Wilkinson. Reciprocal food sharing in the vampire bat. *Nature*, (308):181–184, 1984.
- [20] L. A. Dugatkin. *Cooperation among animals*. Oxford University Press, Oxford, 1997.
- [21] I. J. Molineaux J. J. Bull. Molecular genetics of adaptation in an experimental model of cooperation. *Evolution*, (46):882–895, 1992.
- [22] C. Parker R. Heinsohn. Complex cooperative strategies in group-territorial african lions. *Science*, (269):1260–1262, 1995.
- [23] Y. N. Yu G. J. Velicer. Evolution of novel cooperative swarming in the bacterium *myxococcus xanthus*. *Nature*, (425):75–78, 2003.
- [24] L. Chao P. E. Turner. Prisoner's dilemma in an rna virus. *Nature*, (398):441–443, 1999.

- [25] S. Nee. Mutualism, parasitism, and competition in the evolution of coviruses. *Philos. T. Roy. Soc B*, (355):1607–1613, 2000.
- [26] S. Bonhoeffer T. Pfeiffer, S. Schuster. Cooperation and competition in the evolution of atp-producing pathways. *Science*, (292):504–507, 2001.
- [27] C. Wedekind M. Milinski. Working memory constrains human cooperation in the prisoner’s dilemma. *P. Natl. Acad. Sci. USA*, (95):13755–13758, 1998.
- [28] R. E. Ricklefs. *A Economia da Natureza*. Guanabara Koogan, São Paulo, 2010.
- [29] M. Milinski C. Wedekind. Human cooperation in the simultaneous and the alternating prisoner’s dilemma: Pavlov versus generous tit-for-tat. *Proc. Natl. Acad. Sci. USA*, (93):2686–2689, 1996.
- [30] P. Hammerstein R. Selten. Gaps in harley argument on evolutionarily stable learning rules and in the logic of tit for tat. *Behav. Brain Sci.*, (7):115–116, 1984.
- [31] F. Mengel G. Horváth, J. Kovářík. Limited memory can be beneficial for the evolution of cooperation. *Journal of Theoretical Biology*, (300):193–205, 2012.
- [32] H. M. Taylor S. Karlin. *A first course in stochastic processes*. Academic Press, New York, 2 edition, 1975.
- [33] M. Nowak. Stochastic strategies in the prisoner’s dilemma. *Theor. Popul. Biol.*, (38):93–112, 1990.
- [34] P. Molander. The optimal level of generosity in a selfish, uncertain environment. *J. Conflict Resolut*, (29):611–618, 1985.
- [35] K. Sigmund M. Nowak. The evolution of stochastic strategies in the prisoner’s dilemma. *Acta Appl. Math.*, (20):247–265, 1990.
- [36] A. Shastri L. E. Sine J. P. Lorberbaum, D. E. Bohning. Are there really no evolutionarily stable strategies in the iterated prisoner’s dilemma? *Journal of Theoretical Biology*, (214):155–169, 2002.
- [37] M. A. Nowak. An evolutionarily stable strategy may be inaccessible. *Journal of Theoretical Biology*, (142):237–241, 1990.

-
- [38] F. Franchetti W. H. Sandholm, E. Dokumaci. Dynamo: Diagrams for evolutionary game dynamics. <http://www.ssc.wisc.edu/~whs/dynamo>, 2012.
- [39] K. Sigmund M. Nowak. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, (364):56–58, 1993.