UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO

MARCELO ALVES PEREIRA

**Dilema do prisioneiro contínuo com agentes
racionais e classificadores de cooperação**

Ribeirão Preto
2012

MARCELO ALVES PEREIRA

**Dilema do prisioneiro contínuo com agentes
racionais e classificadores de cooperação**

Tese apresentada à Faculdade de Filosofia,
Ciências e Letras de Ribeirão Preto da
Universidade de São Paulo como parte das
exigências para a obtenção do título de
Doutor em Ciências.

Área de Concentração:

Física Aplicada à Medicina e Biologia.

Orientador:

Prof. Dr. Alexandre Souto Martinez.

Ribeirão Preto

2012

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

FICHA CATALOGRÁFICA

Pereira, Marcelo Alves

Dilema do prisioneiro contínuo com agentes racionais e classificadores de cooperação / Marcelo Alves Pereira; orientador: Prof. Dr. Alexandre Souto Martinez. - - Ribeirão Preto, 2012.

94 f.

Tese (Doutorado) - - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, 2012.

Inclui Bibliografia.

1. dilema do prisioneiro. 2. dilema do prisioneiro contínuo. 3. teoria dos jogos. 4. cooperação. 5. estratégia darwiniana. 6. estratégia pavloviana. 7. entropia de agrupamento. 8. coeficiente de Gini. 9. aversão ao risco. 10. econofísica.

Nome: PEREIRA, MARCELO ALVES

Título: Dilema do prisioneiro contínuo com agentes racionais e classificadores de cooperação

Tese apresentada à Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo como parte das exigências para a obtenção do título de Doutor em Ciências.

Aprovado em: ____/____/____.

Banca Examinadora

Prof. Dr. : _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. : _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. : _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. : _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. : _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Dedicatória

Dedico esta tese aos meus amados pais, Sidnei Alves Pereira e Rosa Maria Cacciolari Alves Pereira, por sempre me darem apoio e forças para que eu possa lutar por meus objetivos e sonhos.

À minha irmã Mônica e meus sobrinhos Leonardo e Rafael.

Agradecimentos

Meus mais profundos agradecimentos ao meu orientador, Prof. Dr. Alexandre Souto Martinez, por sua orientação, críticas e discussões acerca deste trabalho.

Ao Prof. Matheus Grasselli pela orientação durante meu estágio junto ao Department of Mathematics and Statistics (Phimac) na McMaster University, Canadá.

Aos amigos e companheiros diários de laboratório, pelas oportunidades de trocas de experiências e de conhecimento e pelas colaborações.

Aos funcionários do Departamento de Física, em especial Adriano de Jesus Holanda, Fábio José Moretti e Nilza Marina Leone Marino pela presteza e profissionalismo.

Ao Departamento de Física pelo suporte ao trabalho.

Às agências de fomento à pesquisa CAPES e CNPq pelo apoio financeiro.

À Foreign Affairs and International Trade Canada (DFAIT) pelo *grant* concedido através do *Emerging Leaders in the Americas Program* (ELAP).

... [Jacob Bronowski said] “You mean, theory of games like chess.” “No, no” he [von Neumann] said “Chess is not a game. Chess is a well defined form of computation. You may not be able to work out the answers, but in theory there must be a solution, a right procedure in any position. Now real game,” he said, “are not like that at all. Real life is not like that. Real life consists of bluffing, of little tactics of deception, of asking yourself what is the other man going to think I mean to do. And that is what games are about in my theory.” **Em uma conversa de táxi em Londres entre Jacob Bronowski e John von Neumann, do livro *The Ascent of Man*.**

Resumo

PEREIRA, M. A. **Dilema do prisioneiro contínuo com agentes racionais e classificadores de cooperação**. 2012. 94 f. Tese (Doutorado - Programa de Pós-Graduação em Física Aplicada à Medicina e Biologia) - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2012.

O *dilema do prisioneiro* (DP) é um dos principais jogos da teoria dos jogos. No dilema do prisioneiro discreto (DPD), dois prisioneiros têm as opções de cooperar ou desertar. Um jogador cooperador não delata seu comparsa, já um desertor delata. Se um cooperar e o outro desertar, o cooperador fica preso por cinco anos e o desertor fica livre. Se ambos cooperarem, ficam presos por um ano e, se ambos desertarem, ficam presos por três anos. Quando o DP é repetido, a cooperação pode emergir entre agentes egoístas. Realizamos um estudo analítico para o DPD, que produziu uma formulação da evolução do nível médio de cooperação e da tentação crítica (valor de tentação que causa mudança abrupta do nível de cooperação). No dilema do prisioneiro contínuo (DPC), cada jogador apresenta um nível de cooperação que define o grau de cooperação. Utilizamos o DPC para estudar o efeito da personalidade dos jogadores sobre a emergência da cooperação. Para isso, propusemos novas estratégias: uma baseada na personalidade dos jogadores e outras duas baseadas na comparação entre o ganho obtido e a aspiração do jogador. Todas as estratégias apresentavam algum mecanismo de cópia do estado do vizinho com maior ganho na vizinhança, mecanismo este, herdado da estratégia darwiniana. Os resultados mostraram que o DPC aumenta o nível médio de cooperação do sistema, quando comparado ao DPD. No entanto, as diferentes estratégias não aumentaram a cooperação comparado à cooperação obtida com a estratégia darwiniana. Então propusemos o uso do coeficiente de agrupamentos, coeficiente de Gini e entropias de Shannon, Tsallis e Kullback-Leibler para classificar os sistemas, em que os agentes jogam o DPD com a estratégia darwiniana, quanto ao nível de cooperação. Como analisamos valores de médias configuracionais, tais classificadores não foram eficientes ao classificar os sistemas. Isso é consequência da existência de distribuições de extremos nos resultados que compõem as médias. As distribuições de extremos suscitaram uma discussão acerca da definição do regime de cooperação no dilema do prisioneiro. Discutimos também as consequências de utilizar apenas valores médios nos resultados ignorando seus desvios e as distribuições.

Palavras-chave: 1. dilema do prisioneiro. 2. dilema do prisioneiro contínuo. 3. teoria dos jogos. 4. cooperação. 5. estratégia darwiniana. 6. estratégia pavloviana. 7. entropia de agrupamento. 8. coeficiente de Gini. 9. aversão ao risco. 10. econofísica.

Abstract

PEREIRA, M. A. **Continuous prisoner's dilemma with rational agents and cooperation classifiers**. 2012. 94 f. Thesis (Ph.D. - Postgraduate program in Physics Applied to Medicine and Biology) - Faculty of Philosophy, Sciences and Letters, University of São Paulo, Ribeirão Preto, 2012.

Prisoner's dilemma (PD) is one of the main games of game theory. In discrete prisoner's dilemma (DPD), two prisoners have the options to cooperate or to defect. A cooperator player does not defect his accomplice, while a defector does. If one player cooperates and the other defects, the cooperator gets jailed for five years and the defector goes free. If both cooperate, they get jailed during one year and if both defect, they get jailed during three years. When this game is repeated, cooperation may emerge among selfish individuals. We perform an analytical study for the DPD, that produced a formulation for the evolution of the mean cooperation level and for the critical temptation values (temptation values that promote abrupt modifications in the cooperation level). In continuous prisoner's dilemma (CPD), each player has a level of cooperation that defines his/her degree of cooperation. We used the CPD to study the effect of the players' personality on the emergence of cooperation. For this, we propose new strategies: one based on the players' personality and two others based on the comparison between the player's obtained payoff and the desire one. All strategies present some mechanism that copies the state of the neighbor with the highest payoff in the neighborhood, mechanism inherited from the Darwinian strategy. The results showed that the CPD increases the average cooperation level of the system when compared to DPD. However, different strategies do not increased the cooperation compared to cooperation obtained with the Darwinian strategy. So, we propose the use of cluster coefficient, Gini coefficient and entropy of Shannon, Tsallis and Kullback-Leibler as classifiers to classify systems, in which the individuals play DPD with Darwinian strategy, by the cooperation level. As configurational averages were analyzed, such classifiers were not efficient in classifying the systems. This is due to the existence of distributions with extreme values of the results that compose the means. Distributions with extremes values emerged a discussion about the definition of the cooperation state in the prisoner's dilemma. We also discussed the consequences of using only average results in the analysis ignoring their deviations and distributions.

Key-words: 1. prisoner's dilemma. 2. continuous prisoner's dilemma. 3. game theory. 4. cooperation. 5. darwinian strategy. 6. pavlovian strategy. 7. cluster entropy. 8. Gini coefficient. 9. risk aversion. 10. econophysics.

Lista de Figuras

- 2.1 Ganho do jogador A ao confrontar o jogador B no gráfico cartesiano tridimensional. Eixo das abscissas: nível de cooperação do jogador A . Eixo das ordenadas: nível de cooperação do jogador B . Eixo das cotas: ganho do jogador A . Deserção é igual ao valor zero e cooperação igual a um. 13
- 3.1 Autômato celular na rede unidimensional com $N = 11$ jogadores. (a) dilema do prisioneiro discreto (DPD) - jogador coopera ou deserta; (b) dilema do prisioneiro contínuo (DPC) - jogador tem um nível de cooperação. Esquema de cores: branco representa cooperação, preto deserção e tons de cinza os níveis intermediários de cooperação. 31
- 3.2 Vizinhaça do jogador central (de onde partem as setas). Em sistemas que utilizam vizinhaça regular, todos os jogadores utilizam o esquema representado pelas setas para definir seus vizinhos. (a) $k = 8$ (sem auto-interação). (b) $k = 9$ (com auto-interação). 33
- 3.3 Representação de sistemas com condições periódicas de contorno (CPC). (a) Anel representa o sistema unidimensional com suas extremidades conectadas. (b) Toróide representa um sistema bidimensional com suas bordas conectadas. 34
- 3.4 Ganho por interação do jogador A , $g(\alpha_A; \alpha_B)$, em função do seu nível de cooperação, α_A , e do nível de cooperação do oponente, α_B , dado pela Eq. 3.2. Os níveis de cooperação $\alpha_{A,B}$ estão no intervalo $0 \leq \alpha_{A,B} \leq 1$ e os valores para a matriz de ganhos são: $T = 5$, $R = 3$, $P = 1$ e $S = 0$. Os pontos destacados correspondem aos ganhos do dilema do prisioneiro discreto (dilema de prisioneiro de duas-escolhas). . . . 36

3.5	Formas da função subjetividade (Eq. 3.33) utilizada pelos jogadores para avaliar o seu ganho no momento da comparação com o ganho dos oponentes na estratégia subjetiva.	46
3.6	Ganho do jogador A para diferentes personalidades.	47
3.7	Fluxograma das principais etapas do programa utilizado nas simulações. Cada retângulo representa uma operação e cada losango uma decisão. Começamos inicializando os parâmetros da simulação e escolhendo qual será o jogo (Inicializa parâmetros). Em seguida, definimos os vizinhos de cada jogador (Vizinhança) e tem início as repetições para as diferentes configurações ($m \leq M$), onde são criados cada um dos M autômatos distintos (Autômato), que irão interagir durante as rodadas ($t \leq EPC$). Em cada rodada, os jogadores interagem, recebem e somam os ganhos destas interações (Interações). Calculamos o coeficiente de Gini (Gini) utilizando as riquezas dos jogadores. Medimos a distribuição dos grupos em função da configuração do autômato (Agrupamentos) e com isso encontramos o coeficiente de agrupamento (Coeficiente de agrupamento) e esta distribuição também é utilizada no cálculo da entropia do sistema (Entropia). Feito isso, determinamos o como serão feitas as mudanças dos níveis de cooperação dos jogadores e realizamos essa atualização (Estratégia). Todo esse processo é repetido por EPC rodadas e então uma nova configuração é gerada e reinicia-se a sequência. Isso é feito para M configurações diferentes e então salvamos os resultados obtidos.	59
3.8	Representação na rede quadrada das vizinhanças de: (a) von Neumann, $k = 4$ (sem auto-interação) e (b) Moore, $k = 8$ (sem auto-interação). E suas respectivas representações na rede unidimensional para: (c) $k = 4$ (sem auto-interação), e (d) $k = 8$ (sem auto-interação). Cores: preto: jogador central; cinza escuro: primeiros vizinhos; e cinza claro: segundos vizinhos.	61

- 3.9 Fluxograma da definição dos vizinhos de cada jogador. A vizinhança de um jogador depende do tipo escolhido: regular, aleatória fixa ou aleatória dinâmica. Exceto pela aleatória dinâmica, a vizinhança de cada jogador se mantém a mesma durante todo o jogo. Para a vizinhança regular, observa-se se há ou não a auto-interação (**auto**), k ímpar ou par, respectivamente. Os vizinhos do jogador n serão os jogadores que vão de $n + (-k/2)$ até $n + (k/2)$, inclusive n , caso haja auto-interação, e respeitando as condições periódicas de contorno. Para a vizinhança aleatória fixa, são escolhidos, uma única vez, no início do jogo, k vizinhos aleatoriamente pertencentes ao sistema. A variável **aleat** é um número pseudoaleatório uniformemente distribuído entre zero e um. Na vizinhança aleatória dinâmica, a cada rodada, cada jogador escolhe k novos vizinhos, como veremos no fluxograma das interações (Fig. 3.12). O vetor **viz**[**N*****k**] traz os vizinhos de todos os jogadores e é utilizado durante as interações (Fig. 3.12) e para comparar os ganhos dos jogadores com seus vizinhos (Fig. 3.13). 62
- 3.10 Fluxograma da definição do nível de cooperação de cada jogador para cada configuração do automato. No caso discreto, são definidos exatamente $N\rho_0$ jogadores como cooperadores (**coops**). Primeiro, definimos todos os jogadores como desertores e sorteamos aleatoriamente os que serão cooperadores. Para isso, utilizamos um vetor auxiliar (**eleg**[**N**]) com os jogadores que são elegíveis (desertores) para se tornarem cooperadores, e vamos atualizando este vetor para que não sejam sorteados cooperadores. Para os casos contínuos, para cada jogador é definido um número aleatório, entre zero e um, seguindo a distribuição desejada: uniforme, crescente, decrescente, triangular central ou triangular de extremos, que é obtida a partir do método da transformação inversa de uma variável aleatória uniformemente distribuída entre zero e um (**aleat**). 64
- 3.11 Fluxograma da definição da personalidade de cada jogador para cada configuração do automato **q**[**N**]. 65

- 3.12 Fluxograma das interações dos jogadores com seus vizinhos. Cada jogador joga com seus k vizinhos e recebe o ganho das interações e soma estes totalizando o ganho na rodada (`rod[N]`). Caso a vizinhança seja do tipo aleatória dinâmica, os jogadores escolhem seus vizinhos aleatoriamente no momento em que irão interagir com o j -ésimo vizinho. A final da rodada os jogadores somam o ganho da rodada ao ganho acumulado no jogo (`accum[N]`), o ganho acumulado é a riqueza do jogador que será utilizada para calcular o coeficiente de Gini do sistema (Fig. 3.15). 69

- 3.13 Fluxograma que mostra como encontrar o jogador que tem o maior ganho na vizinhança. Para encontrar o jogador com o maior ganho basta encontrar o valor máximo do conjunto: ganho do jogador mais ganhos dos vizinhos. Caso os agentes do sistema adotem a estratégia baseada na personalidade (EBP), o jogador ao invés de utilizar seu ganho real na comparação, ele compara seu ganho subjetivo com os ganhos reais dos vizinhos. Estes jogadores estão agrupados no vetor `venc[N]` que é utilizado pelas estratégias (Fig. 3.14) na comparação de ganhos. 70

- 3.14 Fluxograma da atualização do nível de cooperação dos jogadores de acordo com a estratégia adotada. A estratégia pavloviana é válida somente para o dilema do prisioneiro discreto. Nela, o jogador compara seu ganho com a aspiração e caso tenha um ganho maior que a aspiração mantém seu estado atual, caso contrário, muda de cooperador para desertor ou vice-versa. A estratégia darwiniana é válida tanto para o dilema do prisioneiro discreto quanto para o contínuo. As estratégias baseadas na personalidade e no ganho global e local são válidas para o dilema do prisioneiro contínuo. Tanto para a estratégia darwiniana quanto para a baseada na personalidade, o jogador simplesmente copia o nível de cooperação do jogador da vizinhança com o maior ganho (`venc[N]`). Para as estratégias baseadas no ganho, cada jogador observa a diferença entre o seu ganho e o ganho máximo global ou local, e determina como será sua mudança do nível de cooperação. O novo nível de cooperação é armazenado no vetor `temp[N]`. Para atualizar os estados todos os jogadores modificam seus estados de acordo com o que foi definido pelas estratégias copiando os valores de `temp[N]`. 71
- 3.15 Fluxograma do cálculo do coeficiente de Gini. Criamos um vetor com o ganho acumulado e outro do ganho na rodada dos jogadores que é ordenado em ordem crescente. Calculamos os somatórios da riqueza e se esse ganho for não nulo, calcula-se o coeficiente de Gini, para o ganho acumulado no decorrer das rodadas (`gini`) e da rodada atual (`gini_rod`). 73
- 3.16 Fluxograma que faz o histograma do tamanho dos agrupamentos no sistema. Contamos quantos agrupamentos de cada tamanho há no sistema, considerando condições periódicas de contorno. Esses histogramas são os vetores (`coop`) e (`dsrt`) que serão utilizados no cálculo do coeficiente de agrupamento e das entropias. 74

3.17	Fluxograma da determinação do coeficiente de agrupamento a partir do tamanho do maior agrupamento fazendo a varredura dos histogramas de cooperadores e desertores do fim para o início. Desta forma são determinados os coeficiente de agrupamento cooperador <code>coef-coop[epc]</code> e desertor do sistema <code>coefdsrt[epc]</code>	75
3.18	Fluxograma entropias. (continua)	80
3.18	Fluxograma entropias.	81
4.1	Evolução temporal de sistemas com $N = 1000$ jogadores e nível médio inicial de cooperação $\rho_0 = 1/2$, que jogam o dilema do prisioneiro contínuo durante $EPC = 1000$ rodadas. O tempo evolui da esquerda para direita e no eixo vertical os rótulos dos jogadores. A escala de cores em tons de cinza vai do branco (cooperação) ao preto (deserção). Os gráficos abaixo das imagens de evolução temporal, mostram o nível médio de cooperação em função do tempo (rodadas). Todos esses sistemas possuem a mesma configuração inicial, alterando-se somente os parâmetros e estratégias adotados para cada situação.	87
4.2	Instantes iniciais dos sistemas apresentados na Fig. 4.1 até a rodada $epc = 250$	88
4.3	Ganho de um cooperador ($\alpha_i = 1$) e um desertor ($\alpha_i = 0$) ao jogar contra um oponente com nível de cooperação α_j com $T = 2$, $R = 1$ e $P = S = 0$ (dilema do prisioneiro fraco).	90
4.4	(a) Funções densidade de probabilidade utilizadas para a definição do nível de cooperação dos jogadores no início do jogo. (b)-(f) Histogramas do nível inicial de cooperação dos jogadores, em uma simulação, para cada uma das PDFs utilizadas.	95

4.5	Estudo do efeito das diferentes funções densidade de probabilidade. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas têm agentes que adotam a estratégia darwiniana e não realizam auto-interação (k par).	99
4.6	Desvio padrão de cada um dos gráficos apresentados na Fig. 4.5. Cada conectividade é apresentada em separado.	100
4.7	Estudo do efeito das diferentes funções densidade de probabilidade. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas têm agentes que adotam a estratégia darwiniana e realizam auto-interação (k ímpar).	101
4.8	Desvio padrão de cada um dos gráficos apresentados na Fig. 4.7. Cada conectividade é apresentada em separado.	102
4.9	Estudo do efeito dos tipos de vizinhança. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas adotam a estratégia darwiniana e k assume valores pares.	105
4.10	Desvio padrão de cada um dos gráficos apresentados na Fig. 4.9 Cada conectividade é apresentada em separado, de acordo com o texto acima do gráfico, para facilitar a leitura das barras de desvio padrão. Omitimos a vizinhança regular para o dilema do prisioneiro discreto e contínuo, pois estas foram apresentadas nas Figs. 4.6a e 4.6b.	106

- 4.11 Estudo do efeito dos tipos de vizinhança. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas adotam a estratégia darwiniana e k assume valores ímpares. 107
- 4.12 Desvio padrão de cada um dos gráficos apresentados na Fig. 4.11. Cada conectividade é apresentada em separado, de acordo com o texto acima do gráfico, para facilitar a leitura das barras de desvio padrão. Omitimos a vizinhança regular para o dilema do prisioneiro discreto e contínuo, pois estas foram apresentadas nas Figs. 4.8a e 4.8b. . . . 108
- 4.13 Estudo do efeito das estratégias. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas adotam a estratégia darwiniana e realizam não auto-interação (k par). 110
- 4.14 Desvio padrão de cada um dos gráficos apresentados na Fig. 4.13. Cada conectividade é apresentada em separado, de acordo com o texto acima do gráfico, para facilitar a leitura das barras de desvio padrão. Omitimos a estratégia darwiniana para os casos discreto e contínuo, pois estes já foram apresentados nas Fig. 4.6a e Fig. 4.6b, respectivamente. 111
- 4.15 Estudo do efeito das estratégias. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas adotam a estratégia darwiniana e realizam auto-interação (k ímpar). 113

- 4.16 Desvio padrão de cada um dos gráficos apresentados na Fig. 4.15. Cada conectividade é apresentada em separado, de acordo com o texto acima do gráfico, para facilitar a leitura das barras de desvio padrão. Omitimos a estratégia darwiniana para os casos discreto e contínuo, pois estes já foram apresentados nas Fig. 4.8a e Fig. 4.8b, respectivamente. 114
- 4.17 Comparação entre as diferente estratégias para os casos (a) $k = 8$ e (b) $k = 9$. DPD: dilema do prisioneiro discreto e estratégia darwiniana; DPC: dilema do prisioneiro contínuo e estratégia darwiniana; EBP: dilema do prisioneiro contínuo e estratégia baseada na personalidade; EBGG: dilema do prisioneiro contínuo e estratégia baseada no ganho global; e EBGL: dilema do prisioneiro contínuo e estratégia baseada no ganho local. 115
- 4.18 Histogramas do nível médio assintótico de cooperação para diversas estratégias em sistemas com conectividade $k = 8$ (sem auto-interação). A tentação T de cada histograma é mencionada na primeira linha do gráfico. DPD: dilema do prisioneiro discreto. DPC: dilema do prisioneiro contínuo com PDF uniforme. 118
- 4.19 Histogramas do nível médio assintótico de cooperação para diversas estratégias em sistemas com conectividade $k = 9$ (com auto-interação). A tentação T de cada histograma é mencionada na primeira linha do gráfico. DPD: dilema do prisioneiro discreto. DPC: dilema do prisioneiro contínuo com PDF uniforme. 121
- 4.20 Resultados dos classificadores para um sistema com agentes sem auto-interação, com $k = 8$, utilizando estratégia darwiniana e vizinhança regular. Todos os resultados são as médias configuracionais de $M = 1000$ configurações distintas. Os valores de entropia estão escalonados entre zero e um (u.e., unidade escalonada, adimensional). (continua) 125
- 4.20 Continuação. Os valores de entropia estão escalonados entre zero e um (u.e., unidade escalonada, adimensional). 126
- 4.21 Resultados do classificador para sistemas sem auto-interação ($k = 8$). 131

4.22	Resultados dos classificadores para um sistema com agentes com auto-interação, com $k = 9$, utilizando estratégia darwiniana e vizinhança regular. Todos os resultados são as médias configuracionais de $M = 1000$ configurações distintas. Os valores de entropia estão escalonados entre zero e um (u.e., unidade escalonada, adimensional). (continua)	133
4.22	Continuação. Os valores de entropia estão escalonados entre zero e um (u.e., unidade escalonada, adimensional).	134
4.23	Resultados dos classificadores para sistemas com auto-interação ($k = 9$).	135

Lista de Tabelas

2.1	Ganhos no dilema do prisioneiro. No par ordenado o primeiro valor é o ganho do jogador A e o segundo o do jogador B . Os parâmetros são: T (tentação), R (recompensa), P (punição), S (tolice).	12
2.2	Ganho do jogador A no dilema do prisioneiro de <i>quatro</i> -escolhas com $T = 5$, $R = 4$, $P = 1$ e $S = 0$	21
3.1	Ganhos de um jogador $G_i^{\kappa_i}$ para o estado desertor ($i = D$, $\alpha_D = 0$) e cooperador ($i = C$, $\alpha_C = 1$), de acordo com sua vizinhança, em um sistema que utiliza estratégia darwiniana. Os parâmetros utilizados são: $N = 1000$, $\rho = 1/2$, $k = 4$, $T = 1, 5$, $R = 1$ e $P = S = 0$	53
3.2	Tabela verdade para as desigualdades $G^{\kappa_D} > G^{\kappa_C}$ e $G^{\kappa_C} > G^{\kappa_D}$	54
4.1	Possíveis resultados de $\langle \rho_\infty \rangle$	116

Lista de abreviaturas

C cooperador, p. 11.

CAC coeficiente de agrupamento cooperador, p. 128.

CAD coeficiente de agrupamento desertor, p. 128.

CPC condições periódicas de contorno, p. 34.

D desertor, p. 11.

DP dilema do prisioneiro, p. 1.

DPC dilema do prisioneiro contínuo, p. 2.

DPD dilema do prisioneiro discreto, p. 2.

DPE dilema do prisioneiro espacial, p. 4.

DPI dilema do prisioneiro iterado, p. 3.

EBG estratégia baseada no ganho, p. 40.

EBGG estratégia baseada no ganho global, p. 44.

EBGL estratégia baseada no ganho local, p. 43.

ED estratégia darwiniana, p. 4.

EP estratégia pavloviana, p. 4.

ES estratégia subjetiva, p. 40.

EUT teoria da utilidade esperada (do inglês, *expected utility theory*), p. 26.

LMSC Laboratório de Modelagem em Sistemas Complexos, p. 57.

PDF função de densidade de probabilidade (do inglês, *probability density function*),
p. 94.

PT teoria do prospecto (do inglês, *prospect theory*), p. 26.

TFT estratégia *tit-for-tat*, ou lei do talião (olho-por-olho, dente-por-dente), p. 17.

Lista de símbolos

- α_i nível de cooperação do i -ésimo jogador, p. 21.
- δ_i diferença entre o ganho máximo por interação e o ganho médio obtido por interação, p. 43.
- δ_i^L diferença entre o ganho máximo (tangível ao i -ésimo jogador) e o ganho médio obtido pelo i -ésimo jogador por interação, p. 44.
- δ_i^U diferença entre o ganho máximo por interação no jogo e o ganho médio obtido pelo i -ésimo jogador por interação, p. 44.
- $\Delta_L(\alpha)$ intervalo de ganho local, diferença entre os ganhos máximo e mínimo por interação tangível a um jogador com nível de cooperação α , p. 39.
- Δ_U diferença entre os ganhos máximo e mínimo no jogo, p. 39.
- ϵ_i diferença entre nível de cooperação do i -ésimo jogador e o nível de cooperação do vizinho com maior ganho, p. 42.
- γ probabilidade de um cooperador mudar seu estado para desertor, p. 50.
- $\hat{q}$ personalidade de um jogador, p. 26.
- $\hat{x}$ valores escalonados entre zero e um, p. 123.
- κ conjunto com o nível de cooperação e personalidade dos jogadores de uma vizinhança, p. 32.
- λ probabilidade de um desertor mudar seu estado para cooperador, p. 50.
- $\langle \rho_\infty \rangle$ média configuracional do nível médio assintótico de cooperação do sistema, p. 93.

$\bar{x}$ média, p. 115.

$\bar{f}_\chi$ média espacial, p. 119.

$\bar{f}_\tau$ média temporal, p. 119.

$\Pi(p_i; \hat{q})$ função subjetividade na avaliação do ganho, p. 26.

ρ_0 nível médio inicial de cooperação do sistema, p. 18.

ρ_∞ nível médio assintótico de cooperação do sistema, p. 19.

$\tilde{G}$ ganho médio por interação normalizado, p. 124.

$\tilde{x}$ média configuracional, p. 123.

φ ambição ou aspiração de um jogador, p. 41.

b benefício, p. 23.

B_k estratégia do indivíduo k , p. 17.

c custo, p. 23.

c número de cooperadores, p. 35.

d número de desertores, p. 35.

$E[g(\alpha_i; \rho)]$ ganho esperado por interação do i -ésimo jogador em um sistema com nível médio de cooperação ρ , p. 37.

$E[G_i]$ ganho total esperado do i -ésimo jogador, p. 37.

EPC número de rodadas, p. 57.

$F(x)$ função distribuição acumulada, p. 63.

$f(x)$ função de densidade de probabilidade, p. 63.

$g(\alpha_i; \alpha_j)$ função generalizada de ganho do dilema do prisioneiro, p. 21.

$g_{\max_L}(\alpha_i)$ ganho máximo local, p. 38.

- $g_{\max_U}$ ganho máximo global, p. 38.
- $g_{\min_L}(\alpha_i)$ ganho mínimo local, p. 38.
- $g_{\min_U}$ ganho mínimo global, p. 38.
- G_i ganho total do i -ésimo jogador, p. 35.
- Gini* coeficiente de Gini, p. 72.
- GS ganho subjetivo, p. 45.
- $H(x)$ função de Heaviside ou função degrau, p. 56.
- k conectividade (“vizinhança” ou número de contatos) dos jogadores em um sistema, p. 32.
- M número de configurações do autômato, p. 57.
- N número de jogadores, p. 32.
- P punição, p. 12.
- $p_{i,s}$ probabilidade de um agrupamento da classe i ter o tamanho s , p. 78.
- p_i probabilidade, p. 26.
- p_s probabilidade de um agrupamento ter o tamanho s , p. 76.
- R recompensa, p. 12.
- S “tolo”, do inglês, *sucker*, p. 12.
- Sh entropia de Shannon de agrupamentos, p. 77.
- Sh_C entropia de Shannon de agrupamentos cooperadores, p. 77.
- Sh_D entropia de Shannon de agrupamentos desertores, p. 77.
- SK_C entropia de Kullback-Leibler para a relação entre as distribuições de probabilidade C e D , p. 79.

SK_D entropia de Kullback-Leibler para a relação entre as distribuições de probabilidade D e C , p. 79.

$Sq(q)$ entropia de Tsallis para agrupamentos, p. 78.

$Sq_C(q)$ entropia de Tsallis para agrupamentos cooperadores, p. 78.

$Sq_D(q)$ entropia de Tsallis para agrupamentos desertores, p. 78.

std desvio padrão, p. 115.

T tentação, p. 12.

T_{EDH}^{**} tentação crítica para a estratégia darwiniana considerando o sistema homogêneo, p. 49.

T^* tentação crítica, p. 48.

T_{ED}^* tentação crítica para a estratégia darwiniana, p. 49.

T_{EP}^* tentação crítica para estratégia pavloviana, p. 50.

U utilidade esperada, p. 26.

$u(x_i)$ função utilidade, p. 26.

w_s probabilidade de que um sítio ocupado pertencer a um agrupamento com s sítios, p. 76.

X variável aleatória uniformemente distribuída entre 0 e 1, p. 66.

y_n riqueza do n -ésimo jogador, p. 72.

$\mathcal{P}$ função prospecto, p. 26.

$\text{acum}[N]$ vetor com o ganho acumulado dos jogadores, p. 68.

$\text{coop}[N]$ vetor com o número de agrupamentos cooperadores de tamanho $1 \leq n \leq N$, p. 72.

$\text{dsrt}[N]$ vetor com o número de agrupamentos desertores de tamanho $1 \leq n \leq N$, p. 72.

`nc[N]` vetor com o nível de cooperação dos jogadores, p. 63.

`q[N]` vetor com a personalidade dos jogadores, p. 63.

`rod[N]` vetor com o ganho dos jogadores na rodada, p. 68.

`temp[N]` vetor auxiliar de tamanho N , p. 63.

`venc[N]` vetor com o vizinho vencedor em cada vizinhança, p. 68.

`viz[Nk]` vetor com os vizinhos de cada jogador, p. 61.

Sumário

Lista de Figuras	xv
-------------------------	-----------

Lista de Tabelas	xxv
-------------------------	------------

1 Introdução	1
1.1 Dilema do prisioneiro (DP)	2
1.2 Simulações numéricas de modelos teóricos	5
1.3 Resultados no dilema do prisioneiro	6
1.4 Organização da tese	8
2 Dilema do prisioneiro	11
2.1 Variantes	14
2.1.1 Discreto rodada única (DPD)	15
2.1.2 Discreto iterado (DPI)	16
2.1.3 Espacial (DPE)	18
2.1.4 Iterado unidimensional	19
2.1.5 Contínuo (DPC)	20
2.2 Exemplos	22
2.2.1 Mimercofilia	23
2.2.2 Inspeção de predador em cardumes	23
2.2.3 Troca de ovos em hermafroditas simultâneos	24
2.2.4 Morcegos hematófagos	24
2.3 Aversão ao risco	25
3 Modelos	29
3.1 Unidimensional	30

3.1.1	Autômato celular	30
3.1.2	Conectividade	32
3.1.3	Ganhos	34
3.1.4	Dinâmica do sistema	39
3.1.5	Estratégias	40
3.1.5.1	Pavloviana (EP)	41
3.1.5.2	Darwiniana (ED)	42
3.1.5.3	Baseada no ganho (EBG)	43
3.1.5.4	Subjetiva (ES)	45
3.2	Modelo analítico	47
3.2.1	Tentação crítica (T^*)	48
3.2.1.1	T^* - estratégia darwiniana	49
3.2.1.2	T^* - estratégia pavloviana	50
3.2.2	Evolução temporal	50
3.3	Modelo computacional	57
3.3.1	Conectividade ou vizinhança	60
3.3.2	Autômato celular	61
3.3.3	Interações - dinâmica do sistema	67
3.3.4	Estratégias e atualização dos estados	68
3.3.5	Coefficiente de Gini	70
3.3.6	Agrupamentos de jogadores	72
3.3.7	Entropia da informação	75
3.3.7.1	Shannon	77
3.3.7.2	Tsallis	77
3.3.7.3	Kullback-Leibler generalizada	78
4	Resultados numéricos	83
4.1	Transiente - dilema do prisioneiro contínuo	86
4.2	Regime estacionário	93
4.2.1	Função densidade de probabilidade (PDF)	94
4.2.2	Conectividade	103
4.2.3	Estratégia	109

4.2.4 Médias configuracionais	115
4.3 Novos classificadores	120
5 Conclusões e perspectivas	137
Referências	143

Introdução

“...precisamos estudar principalmente os eventos raros e extremos para que seja possível decifrar os eventos comuns...”

Nassim Nicholas Taleb no livro *A lógica do cisne negro* - 2007.

O *dilema do prisioneiro* (DP) se destacou dentre os jogos estudados pela teoria dos jogos por ter sido o primeiro onde a *cooperação* emergiu como estratégia estável. A resolução do dilema do prisioneiro na teoria dos jogos [1] é baseada em tomadores de decisão racionais¹ e leva à caracterização dos jogadores como egoístas e competitivos. Os jogadores consideram a deserção como melhor estratégia individual, mas a emergência da cooperação como uma estratégia coletiva estável surpreendeu os pesquisadores e transformou o dilema no prisioneiro em referência nos estudos sobre cooperação.

Neste capítulo, situamos o problema do dilema do prisioneiro e mostramos sua relevância. Em seguida, apresentamos algumas considerações sobre os estudos existentes na área e por fim um resumo dos resultados obtidos nesta tese e a organização da mesma.

¹Em economia, teoria dos jogos e teoria de decisão, o tomador de decisão é “racional” [2] no sentido que ele:

- conhece todas suas alternativas;
- tem expectativas quanto ao desconhecido;
- tem preferências claras e
- escolhe suas ações deliberadamente depois de algum processo de otimização.

1.1 Dilema do prisioneiro (DP)

O dilema do prisioneiro (DP) apresenta uma situação na qual dois prisioneiros têm que escolher entre cooperar - ficando em silêncio - ou desertar - delatando seu comparsa. Se um ficar em silêncio e o outro delatar, o que ficou em silêncio fica preso por cinco anos e o delator ganha a liberdade. Caso ambos permaneçam em silêncio, ficarão presos por um ano; se ambos se delatarem, ficarão presos por três anos.

Observando as possíveis combinações de escolhas dos jogadores e seus respectivos resultados, vemos que, tanto enfrentando um cooperador quanto um desertor, a deserção produz um ganho² maior. Então parece óbvio que a melhor escolha seja a deserção. No entanto, se ambos os prisioneiros pensarem desta forma e escolherem a deserção, ficarão presos por três anos. Este resultado é pior do que a cooperação mútua, pois, neste caso, ficariam presos por um ano e não três, como no caso da deserção mútua. Pensando nessa possibilidade, parece razoável que ambos cooperem para aumentar seus ganhos. Entretanto, se qualquer um deles supor que o seu oponente irá cooperar, ele poderá aumentar ainda mais seu ganho desertando. Essa é a essência do conflito no dilema do prisioneiro.

O dilema do prisioneiro apresenta uma variante conhecida como *dilema do prisioneiro contínuo* (DPC) [3, 4, 5, 6, 7, 8, 9, 10, 11]. Nela os jogadores têm um nível de cooperação que varia entre a cooperação total e a deserção total. Isso pode acontecer, por exemplo, se os prisioneiros puderem fornecer somente uma certa quantia de provas para incriminar seu comparsa. Em nosso estudo, consideramos essa possibilidade para compreender seu efeito sobre o sistema. Para diferenciar o DPC do dilema do prisioneiro de duas escolhas, chamamos este de *dilema do prisioneiro discreto* (DPD).

Alguns conflitos da vida real podem ser descritos por situações em que alcançar os objetivos pessoais, é incompatível com alcançar objetivos mútuos, ou de grupo. Cada jogador é tentado a procurar aumentar seu ganho próprio por competição, mas se ambos competem, eles perdem, se eles estiverem dispostos a confiar um no outro e cooperarem, ambos lucram. Entre 78 tipos de jogos de 2 indivíduos,

²Chamamos de ganho a diminuição da pena, ou seja, mais tempo em liberdade.

o DP é o único que tem um resultado fortemente estável que é deficiente, isto é, o ganho dos jogadores é menor que o ganho ótimo para os dois jogadores [12].

Como o DP representa uma situação de conflito entre duas pessoas, ele se aplica em muitas situações de conflito da vida real, como por exemplo, na política (sociofísica) [9, 13] e economia [14, 15, 16]. O DP também representa situações nas quais a disputa entre agentes econômicos pode acabar em tragédia para ambas as partes envolvidas, como a tragédia dos comuns, que foi apresentada por Hardin [17].

Na biologia, algumas situações que podem ser modeladas usando o DP são [18]: mirmecofilia [19], inspeção de predador em peixes [20, 21, 22, 23, 24, 25], troca de ovos em hermafroditas simultâneos [26, 27, 28] e compartilhamento de sangue entre morcegos hematófagos [29]. Estes exemplos serão apresentados em detalhes na Seção 2.2.

Em teoria dos jogos utilizam-se agentes ditos racionais que buscam maximizar seus ganhos. Portanto, a solução racional para o dilema do prisioneiro seria a deserção. Para verificar se a solução racional teórica seria utilizada na prática, a empresa RAND fez um experimento onde pessoas jogaram o dilema do prisioneiro repetidas vezes, situação conhecida como *dilema do prisioneiro iterado* (DPI). No entanto, os resultados obtidos não confirmaram o resultado esperado pela teoria dos jogos.

Buscando compreender e explicar esse resultado inusitado, Axelrod propôs um torneio de computadores [30, 31, 32, 33] para confrontar estratégias propostas por pesquisadores de diversas áreas, em busca de uma estratégia que prevalecesse quando confrontada contra as outras estratégias e com a estratégia aleatória, ou seja, que obtivesse o maior ganho médio no jogo.

A estratégia vencedora neste torneio foi o *tit-for-tat* (olho por olho, dente por dente), uma estratégia simples com memória de uma única jogada, que começa cooperando e passa a copiar a ação do oponente a partir da segunda rodada. Com esta estratégia, a cooperação apareceu como estratégia estável e fez do dilema do prisioneiro o principal arquétipo para explicar a cooperação nas mais diversas situações.

Nowak e May [34, 35] demonstraram a emergência de cooperação entre jogadores racionais, egoístas e sem memória. Os prisioneiros são dispostos nos sítios

de uma rede quadrada e jogam o dilema do prisioneiro contra seus oponentes adjacentes na rede. Assim, eles estabeleceram o chamado *dilema do prisioneiro espacial* (DPE). O ganho obtido por cada jogador é a soma de todas suas interações. Para modificar seu estado, o jogador toma uma decisão baseada em sua estratégia. No DPE, os jogadores copiam o estado do vizinho que obtém o maior ganho, esta é a *estratégia darwiniana* (ED).

Outras estratégias também podem ser utilizadas. Uma que também foi bastante explorada é a *estratégia pavloviana* (EP) [35, 36, 37, 38, 39, 40, 41, 42]. Nela, o jogador compara seu ganho com seu nível de ambição/aspiração e se, o ganho for maior que a ambição, ele mantém o estado atual, e caso contrário, muda de cooperação para deserção ou vice-versa.

O estudo desta tese consiste em:

- resolução analítica da função de tentação crítica (valores de tentação que causam variações significativas no nível médio assintótico de cooperação);
- uma abordagem para a resolução analítica da equação do nível médio de cooperação em função do tempo;
- simulações numéricas do dilema do prisioneiro contínuo com novas estratégias e
- simulações numéricas do dilema do prisioneiro discreto com a estratégia darwiniana para os novos classificadores de cooperação.

Para as simulações numéricas foi utilizado um modelo baseado em um autômato celular unidimensional, onde cada célula é um jogador que joga o dilema do prisioneiro (contínuo ou discreto) com as células adjacentes. No sistema que utiliza uma rede unidimensional também ocorre a emergência de cooperação, com agentes sem memória, assim como observado no trabalho de Nowak e May [34]. Quanto ao modelo analítico, partimos das premissas do modelo numérico e desenvolvemos as equações.

Para explorar o nível de cooperação contínuo, propusemos novas estratégias: a baseada no ganho local ou global e a baseada na personalidade. As estratégias baseadas no ganho comparam um ganho desejado pelo jogador com os ganhos obtidos

pelos vizinhos. Elas não copiam diretamente o estado do vizinho que obteve o maior ganho na vizinhança. Há uma modificação no mecanismo de cópia do estado que a diferencia da ED. A estratégia baseada na personalidade traz uma modificação na comparação dos ganhos, mas a mudança de estado do jogador é copiar o vizinho vencedor, como ocorre na ED.

Além do estudo dos efeitos das estratégias para o DPC, fizemos um estudo de propriedades do sistema que estão relacionadas à configuração espacial dos jogadores ou à distribuição de ganhos no sistema para buscar uma definição do sistema como cooperativo ou desertor. Para isso, utilizamos os conceitos de entropia de informação, coeficiente de agrupamento dos jogadores e distribuição da riqueza entre os jogadores (coeficiente de Gini).

Deste modo, os objetivos desta tese são:

- explorar as oportunidades trazidas pelo nível de cooperação contínuo, como o uso de estratégias com mudança dos estados dos jogadores mais elaboradas, e com isso compreender o papel do nível de cooperação contínuo e de diferentes estratégias na emergência da cooperação;
- dentre as estratégias, estudar o efeito da aspiração por um determinado ganho, que pode ser pessoal (local) ou global, e da personalidade nos jogadores (aversão ou afinidade ao risco);
- resolver analiticamente a função do nível médio de cooperação do sistema em função do tempo e os valores de tentação crítica (valores de tentação que modificam o nível de cooperação abruptamente) e
- buscar métodos que sirvam para classificar o sistema como cooperador ou desertor.

1.2 Simulações numéricas de modelos teóricos

Quando soluções analíticas das funções que regem um problema não são possíveis, experimentos reais e simulações numéricas podem ser as alternativas tangíveis. Experimentos numéricos são a melhor escolha quando experimentos reais apresentam muitas variáveis que estão fora do objetivo do estudo e não há um modo de

controlar essas variáveis, ou quantificar sua influência nos resultados. No caso do dilema do prisioneiro, seriam necessários experimentos com humanos ou animais, mas ao fazer estes estudos não poderíamos desconsiderar alguns aspectos inerentes aos mesmos, como a personalidade, por exemplo.

As simulações numéricas são frequentemente utilizadas no estudo de fenômenos que apresentam caos, afim de aprofundar a análise sobre as características destes sistemas. Em sistemas com caos determinístico, uma pequena modificação nas condições iniciais implica em resultados completamente distintos. Para medir essa falta de previsibilidade é utilizado o coeficiente de Lyapunov, que mede a distância entre os resultados que é causada pela variação da condição inicial. O comportamento caótico é caracterizado pela divergência exponencial de trajetórias que tem condições iniciais próximas. Essas medidas de caos são numéricas e o estudo analítico não mostraria estas características do problema.

Além disso, simulações numéricas permitem o domínio do cenário a ser estudado, permitindo a definição dos parâmetros de controle e das condições iniciais, podendo ser reproduzidas com facilidade. Em experimentos reais, muitas vezes, não há essa possibilidade, como nos experimentos com humanos, onde não podemos remover vieses intrínsecos dos indivíduos. Por estas razões, optamos por utilizar simulações numéricas e mostramos que em nosso estudo existem cenários onde a resolução analítica é muito complicada, ou não é válida em todo o domínio do problema.

1.3 Resultados no dilema do prisioneiro

A maior parte dos estudos sobre dilema do prisioneiro consideram o caso discreto, onde os jogadores, ou cooperam ou desertam. Nesta tese, iniciamos os estudos considerando o caso contínuo, onde cada jogador pode assumir vários graus de cooperação. O nível de cooperação contínuo aliado a uma função generalizada de ganhos, que é uma combinação linear dos possíveis ganhos obtidos no dilema do prisioneiro discreto, nos permitiram explorar os efeitos da continuidade do nível de cooperação e o efeito das novas estratégias.

Estudamos o DPC com a estratégia darwiniana, bastante conhecida na li-

teratura. Também propusemos novas estratégias, uma baseada no ganho e outra, baseada na personalidade do jogador, na expectativa de compreender quais aspectos são fundamentais na emergência, manutenção ou extinção da cooperação. Além das estratégias, analisamos o efeito da função densidade de probabilidade da distribuição inicial do nível médio inicial de cooperação dos jogadores e o efeito do tipo de vizinhança.

A distribuição inicial dos jogadores influencia no nível médio assintótico de cooperação, portanto optamos por utilizar a distribuição uniforme para não inserirmos nenhuma tendência por parte da distribuição adotada. O tipo de vizinhança é muito importante na manutenção da cooperação, pois a vizinhança aleatória extinguiu a cooperação em quase todos os casos em que foi utilizada. Verificamos que o nível contínuo de cooperação aumenta o nível médio assintótico de cooperação quando comparado aos resultados do dilema do prisioneiro discreto. No entanto, as estratégias propostas não tiveram diferenças significativas em relação à estratégia darwiniana, portanto, não pudemos definir a interferência de cada característica das estratégias sobre a cooperação.

Durante a análise, observamos que a distribuição dos resultados que formam um valor médio assintótico de cooperação não é estudada na literatura científica, assim como seu desvio padrão. Então, fizemos um estudo a respeito das médias e suas distribuições e o resultado para o dilema do prisioneiro discreto foi bastante interessante, pois surgem distribuições de extremos.

Na literatura, há muitos resultados que mostram que determinada topologia (tipos e conexões de rede), estratégia, uso de informação, reputação, ruído na comunicação, etc. são responsáveis pela emergência/manutenção/extinção da cooperação no sistema. No entanto, não há um claro consenso sobre a definição de cooperação. Alguns estudos tratam cooperação como a não extinção de cooperadores no sistema. Outros consideram a cooperação como a maioria da população no estado cooperador.

Na busca de indícios se um sistema terá como regime estacionário uma maioria de cooperadores ou então que não haverá a extinção dos cooperadores, propusemos algumas técnicas para classificar o sistema como cooperador ou não. Utilizamos para isso o coeficiente de agrupamentos, o coeficiente de Gini e as entropias de agru-

pamento de Shannon, Tsallis e Kullback-Leibler. Para realizar este estudo, utilizamos o dilema do prisioneiro discreto. Pois, além da maioria dos estudos na literatura sobre o tema tratarem deste caso, o estudo baseado nos agrupamentos é mais simples quando há apenas cooperadores e desertores para a análise dos agrupamentos.

As novas técnicas propostas falharam em seu objetivo de classificar o sistema como cooperador ou desertor, pois esbarramos novamente no problema das distribuições de extremos do nível médio assintótico de cooperação. A impossibilidade de aplicar os classificadores é consequência das distribuições de extremos que aparecem nos resultados. Essa característica não havia sido observada nos estudos presentes na literatura, mas não deve ser desconsiderado em virtude de suas implicações.

1.4 Organização da tese

Começamos com uma breve revisão da literatura e apresentação dos conceitos necessários à compreensão do estudo. Em seguida, apresentamos o modelo adotado, sua abordagem analítica e a implementação numérica. Então, passamos aos resultados numéricos obtidos e finalmente às conclusões. A tese está estruturada em cinco capítulos, alguns divididos em seções conforme descrevemos abaixo.

No Capítulo 2, apresentamos uma revisão da evolução do estudo sobre o dilema do prisioneiro, na Seção 2.1, suas variantes e na Seção 2.2, exemplos de como o DP é utilizado para modelar problemas sociais e biológicos. A Seção 2.3 versa sobre aversão ao risco, que será utilizada na estratégia baseada na personalidade do jogador.

O Capítulo 3 trata do modelo utilizado. Iniciamos na Seção 3.1 abordando suas peculiaridades como a construção do autômato, sua conectividade, a apresentação das estratégias com as regras de interação e toda a dinâmica de evolução do sistema. Depois surgem os desdobramentos do tratamento analítico, Seção 3.2, que mostra a evolução temporal do nível de cooperação do sistema e os valores de tentação crítica. Na Seção 3.3, apresentamos os detalhes dos algoritmos utilizados no cálculo numérico, perfazendo a parte computacional do estudo.

O Capítulo 4 traz os resultados das simulações numéricas. Na Seção 4.1, mostramos a evolução temporal de alguns sistemas selecionados para ilustrar o re-

gime transiente do dilema do prisioneiro contínuo. Na Seção 4.2, ainda utilizando o DPC, mostramos os resultados para as novas estratégias, diferentes distribuições iniciais e diferentes vizinhanças. A Seção 4.3 apresenta os métodos que utilizamos para definir o sistema como cooperador e os resultados obtidos com estes métodos para o DPD com a estratégia darwiniana.

No Capítulo 5, são apresentadas as conclusões e as sugestões para estudos posteriores.

Dilema do prisioneiro

“Nós não aprendemos espontaneamente que *não aprendemos que não aprendemos*.”

Nassim Nicholas Taleb no livro *A lógica do cisne negro* - 2007.

Neste capítulo apresentamos o dilema do prisioneiro, algumas de suas variantes e exemplos de situações reais encontradas na natureza, que podem ser modeladas utilizando este jogo. O *dilema do prisioneiro* é a situação vivida por dois prisioneiros sob interrogatório, que precisam fazer uma escolha: desertar (*D* - desertor) ou cooperar (*C* - cooperador). Os prisioneiros são suspeitos de um crime grave, mas foram presos por um crime leve. Eles ficam separados e não podem se comunicar. Quando são interrogados, individualmente, o investigador oferece duas opções: ou entregar o comparsa (desertar), dando indícios para incriminá-lo pelo crime grave, ou ficar calado e não delatar o comparsa (cooperar). No caso de um dos prisioneiros denunciar o comparsa (desertor) e este ficar calado (cooperador), o denunciante obterá a liberdade, por colaborar com a justiça (delação premiada), enquanto seu comparsa será condenado com a pena máxima devido ao crime grave, pois a justiça conseguiria os indícios necessários para a condenação do cooperador, a partir do depoimento do desertor. Se ambos permanecerem em silêncio, receberão a sentença mais branda, devido ao crime leve que os levou à prisão em flagrante. Contudo, se os dois se delatarem mutuamente serão punidos com uma mesma condenação, menor do que a máxima, por terem colaborado com a justiça, no entanto, maior do que a pena leve.

A formalização do jogo com os tempos de prisão e o nome “dilema do prisioneiro” foi feita por Albert W. Tucker em 1950, na tentativa de tornar as ideias de Flood e Dresher [43] mais acessíveis em um seminário na Universidade de Stanford [44]. De maneira mais formal, com as penas para as possíveis combinações de escolhas dos prisioneiros, surge:

- cooperação mútua, ambos recebem como pena R (recompensa);
- deserção mútua, recebem a pena P (punição);
- atitudes antagônicas, o cooperador recebe S (“tolice”, do inglês, *sucker*) e o desertor recebe T (tentação).

Como a prisão é uma privação da liberdade, naturalmente pensamos nisto como algo negativo. Chamamos de ganho do jogador o aumento do seu tempo em liberdade. Para termos valores positivos, somamos a todos os valores de pena a pena máxima. Desta forma, a pena máxima tem um ganho nulo no aumento da liberdade e as demais penas tem um aumento no tempo em liberdade. A Tab. 2.1 resume os ganhos do DP, as linhas contém as escolhas dos jogador A e as colunas as escolhas do jogador B . Cada célula traz o resultado, na forma de um par ordenado que corresponde ao ganho do jogador A e B , respectivamente, em função das escolhas dos jogadores. Estes ganhos também podem ser representados em um gráfico cartesiano tridimensional, onde o eixo das abscissas representa a escolha do jogador A , o eixo das ordenadas a escolha do jogador B , sendo deserção igual ao valor zero e cooperação igual a um, e o eixo das cotas o ganho do jogador A , que varia de 0 à T , como mostra a Fig. 2.1.

Tabela 2.1 – Ganhos no dilema do prisioneiro. No par ordenado o primeiro valor é o ganho do jogador A e o segundo o do jogador B . Os parâmetros são: T (tentação), R (recompensa), P (punição), S (tolice).

		Jogador B	
		Coopera	Deserta
Jogador A	Coopera	$(R,R)=(3,3)$	$(S,T)=(0,5)$
	Deserta	$(T,S)=(5,0)$	$(P,P)=(1,1)$

Figura 2.1 – Ganho do jogador A ao confrontar o jogador B no gráfico cartesiano tridimensional. Eixo das abscissas: nível de cooperação do jogador A . Eixo das ordenadas: nível de cooperação do jogador B . Eixo das cotas: ganho do jogador A . Deserção é igual ao valor zero e cooperação igual a um.

Para que exista o conflito, os ganhos do DP estão sujeitos às condições $T > R > P > S$ e $2R > T + S$ [32], pois:

- i. $T > R$, a deserção proporciona um ganho maior que a cooperação ao jogar contra um cooperador (menor tempo de prisão);
- ii. $R > P$, o ganho para a cooperação mútua é superior ao ganho da deserção mútua;
- iii. $P > S$, a deserção proporciona um ganho maior que a cooperação ao jogar contra um desertor e
- iv. $2R > T + S$, dois jogadores que alternam entre cooperação e deserção tem um ganho menor do que se ambos cooperarem sucessivamente.

Adotando $T > R$, $R = 1$, $P = S = 0$, os ganhos utilizados por Nowak & May [34], o jogo é conhecido como *dilema do prisioneiro fraco*. Quando essas condições não são satisfeitas surgem outros jogos, como o jogo da galinha (chicken game ou hawk-dove game) ou jogo da caça da lebre ou cervo (stag hunt game).

Analisando os resultados do DP nota-se que a melhor opção para um tomador de decisão racional é a deserção, pois quando faz esta opção ele sempre aumenta seu ganho. Na Tab. 2.1 vemos que se passarmos da cooperação para a deserção (indo da

primeira linha para a segunda, no caso do jogador A , ou da primeira para a segunda coluna, no caso do jogador B) os ganhos aumentam. Ao jogar contra um cooperador, mudando de cooperador para desertor o ganho passa de R para T e $T > R$ e ao jogar contra um desertor, passa de S para P e $P > S$. Também podemos observar este aumento de ganho pela Fig. 2.1. Quando seguimos o eixo das abscissas em direção à origem do gráfico, indo da cooperação (um) para a deserção (zero), o ganho sempre aumenta. Se ambos prisioneiros utilizarem estas informações no processo de tomada de decisão, o resultado estável será a deserção mútua. No entanto, a deserção mútua é desvantajosa para ambos se comparada à cooperação mútua.

Então, como seria possível obter a cooperação se considerarmos tomadores de decisão racionais? Eles querem otimizar seus ganhos, mas não podem fazer acordos para garantir seus ganhos e se proteger de eventuais prejuízos. Assim ambos escolheriam racionalmente pela deserção. Apesar de uma concepção simples o DP gera um dilema de difícil solução.

A tomada de decisão dos jogadores depende da estratégia que eles utilizam. Uma tentativa de resolver o conflito gerado pelo dilema do prisioneiro é utilizar diferentes estratégias, observar o efeito de cada uma delas no resultado e então explicar quais foram as características da estratégia que determinaram aquele resultado. Juntado-se aos predicados de cada estratégia a maximização dos ganhos dos jogadores e assuntos similares podemos buscar compreender situações equivalentes da vida real [45].

2.1 Variantes do dilema do prisioneiro

Agora, apresentamos as principais variantes do dilema do prisioneiro. Cada uma destas variantes agregam novas características que servem para explicar a emergência/manutenção da cooperação no sistema. A citar:

1. dilema do prisioneiro (DP): traz a concepção do problema do dilema do prisioneiro e dois jogadores que jogam o DP uma única rodada.
2. dilema do prisioneiro iterado (DPI): um par de jogadores, que podem adotar diversas estratégias, jogam o DP repetidas vezes.

3. dilema do prisioneiro espacial (DPE): os jogadores estão localizados em uma estrutura espacial bidimensional e jogam com mais de um oponente durante várias rodadas e todos adotam uma única estratégia.
4. dilema do prisioneiro unidimensional - caso particular do DPE: utilizamos uma estrutura unidimensional [46, 47, 48], que apesar de mais simples, obtém resultados para o nível médio assintótico de cooperação qualitativamente condizentes com os observadas no DPE bidimensional e as transições de fase [49, 50] são as mesmas.
5. dilema do prisioneiro contínuo (DPC): os jogadores apresentam um grau de cooperação ao invés de somente cooperar ou desertar.

2.1.1 Dilema do prisioneiro (DP): dois jogadores e uma única rodada

Na versão original de Tucker para o problema do dilema do prisioneiro, há um único confronto, no qual o prisioneiro precisa decidir se vai cooperar ou desertar. Desta maneira, a resolução racional formal tem por resultado a deserção mútua como estratégia dominante. Isso ocorre pois, independente da ação do oponente, a única forma de aumentar seu ganho é desertar, além de ser a única forma de obter o ganho máximo do jogo T e evitar o pior resultado possível S . Mas, para ambos seria melhor que optassem pela cooperação mútua.

Humanos, que vislumbram todas essas nuances do problema, sabem que a cooperação mútua é preferível à deserção mútua. No entanto, como conseguir que ambos os prisioneiros cooperem? Eles não podem se comunicar e fazer um acordo. Mesmo que façam um acordo, não há garantias de que irão cumprir com suas partes no trato, tornando a cooperação pouco atrativa e até perigosa. Com isso, fica evidente a dificuldade em explicar a cooperação entre agentes racionais. Assim, o DP se tornou uma ferramenta de grande valia para explicar sob quais condições a cooperação poderia emergir entre agentes considerados naturalmente competitivos [44].

2.1.2 Dilema do prisioneiro discreto iterado (DPI): dois jogadores e muitas rodadas

No *dilema do prisioneiro iterado* (DPI) os jogadores se encontram e jogam o DP repetidamente. Estabelecido o dilema do prisioneiro e sua resolução racional, surgiu a questão: pessoas fariam as escolhas que a teoria apresentou como mais adequadas? Para responder a esta questão foi feito um experimento com a participação do matemático John Williams, chefe de departamento da corporação estadunidense RAND, e do economista Armen A. Alchian, da Universidade da Califórnia, Los Angeles [51]. Eles jogaram o DP cem vezes. O resultado esperado seria que eles desertassem todas as cem rodadas se agissem conforme o previsto pela teoria para agentes racionais. No entanto, o resultado obtido contrariou a previsão teórica e eles optaram por cooperar mais vezes ao invés de optarem pela estratégia dominante, a deserção, e trair o oponente. Assim não atingiram o ponto de equilíbrio - deserção mútua - previsto na teoria. Isto mostrou a existência de algum problema na concepção teórica de racionalidade, pois, ao invés de buscar resultados mais favoráveis para si imediatamente, os jogadores tendem a escolher a cooperação para efetivamente maximizar seus ganhos a longo prazo.

A existência do reencontro (ou probabilidade de reencontro não desprezível) foi apontada como responsável pela quebra da racionalidade observada no DPI. Neste caso, a melhor escolha não seria necessariamente a deserção, pois existe a possibilidade de retaliação, a qual acaba por inibir a deserção. É conveniente retaliar apenas à uma jogada não-cooperativa anterior. Se os jogadores tentarem aumentar seus ganhos utilizando as cooperações mútuas, eles poderiam se arriscar jogando cooperativamente. Isso poderia levar à acordos tácitos do tipo: jogar como cooperador desde o início do jogo e retaliar respostas desertoras.

Jogadores racionais pensariam que podem aumentar ainda mais seus ganhos ao jogar como desertores na última rodada, pois não seriam retaliados na rodada seguinte. Considerando a mesma racionalidade para ambos os jogadores, eles se utilizarão desta tática, mas surgirá o pensamento: “Meu oponente irá desertar na última rodada, então devo jogar como desertor na penúltima rodada para ter um desempenho melhor que o dele.”. Baseado na lógica de indução reversa, o jogador nota que o oponente também pensará assim e se antecipa para a antepenúltima

rodada e assim sucessivamente. Este procedimento se repete até chegar a primeira rodada. O resultado final deste processo é que ambos devem jogar como desertor desde a primeira rodada. Por isso, o DPI é jogado um número indefinido de vezes, para que os jogadores não possam utilizar a informação sobre o término do jogo.

Axelrod [30, 31, 32, 33] propôs um torneio de computadores para buscar uma estratégia que fosse a mais eficiente quando confrontada com outras estratégias e a estratégia completamente aleatória. Ele convidou pesquisadores de diversas áreas a enviarem suas estratégia para jogar o dilema do prisioneiro contra as demais estratégias e a estratégia aleatória. Aquela que tivesse o maior ganho médio ao final do torneio seria a vencedora. Quinze estratégias participaram do torneio, as estratégias se enfrentavam por 200 rodadas, mas nenhuma das estratégias tinha conhecimento do número de rodadas para que não pudessem usar essa informação. A estratégia vencedora foi a *tit-for-tat* (TFT), conhecida também como a lei do talião (olho-por-olho, dente-por-dente).

O TFT é muito simples: tem memória de uma única iteração e sua dinâmica é cooperar na primeira jogada e repetir a ação do oponente nas próximas rodadas. Sua memória é a ação do oponente na rodada anterior. Se dois jogadores que utilizam TFT se enfrentam, a partir do momento em que um deles coopera, o outro também o fará e a cooperação emerge como comportamento estável.

Neste torneio a estratégia TFT ficou conhecida por ser uma estratégia extremamente estável e robusta. Numa segunda edição do torneio, muitas das estratégias inscritas tinham por objetivo vencer o TFT. Algumas até foram eficientes em sua tarefa, mas ao se confrontarem com as demais estratégias não se saíram tão bem e novamente a TFT foi vencedora.

Porém, Boyd & Lorberbaum [52] demonstraram que nenhuma estratégia determinista é estável no DPI. Segundo eles, uma estratégia comum na população B_e é robusta se, para qualquer estratégia invasora B_i dada, têm-se $E[g(B_e|B_e)] \geq E[g(B_i|B_e)]$, onde $E[g(B_m|B_n)]$ é o ganho esperado de jogadores que utilizam a estratégia B_m ao confrontar os que utilizam B_n . Neste trabalho, eles mostraram que a estabilidade coletiva não implica em estabilidade evolucionária.

Por exemplo, considere dois jogadores com a estratégia TFT jogando DPI. O ganho de ambos os jogadores é maior enquanto eles optam pela cooperação. Se por

um erro um dos jogadores desertar, o outro imediatamente o retalia e ambos passam a jogar como desertores, gerando um ganho menor do que quando eles estavam cooperando. A deserção seria a estratégia invasora que gera a instabilidade no TFT. Existem estratégias mais robustas que o TFT - por exemplo, a estratégia pavloviana - para ambientes menos estáveis, como na presença de ruído de comunicação, isto é, há uma probabilidade não nula de uma cooperação ser considerada como deserção no momento do cálculo do ganho do jogador.

2.1.3 Dilema do prisioneiro espacial (DPE): muitos jogadores, rede bidimensional ou complexa e muitas rodadas

Em 1992, Nowak & May [34, 35] utilizaram uma rede bidimensional quadrada, *dilema do prisioneiro espacial* (DPE), onde cada sítio da rede é um jogador que pode cooperar ou desertar ao jogar o dilema do prisioneiro com os vizinhos. Neste trabalho eles mostraram que a cooperação pode emergir entre jogadores competitivos, sem memória das jogadas anteriores e que utilizam a estratégia darwiniana.

Uma proporção ρ_0 de jogadores são definidos como cooperadores no início do jogo. Eles são escolhidos aleatoriamente na rede (todos os sítios tem a mesma chance de serem escolhidos), os demais, $1 - \rho_0$, são definidos como desertores. Em seguida, cada jogador joga o DP com seus primeiros e segundos vizinhos e com ele mesmo (auto-interação). O ganho do jogador na rodada é a soma dos ganhos que ele obtém de todas suas interações com os vizinhos e da auto-interação. Após todos os jogadores terem interagido e somado seus ganhos decorrentes destas interações, cada jogador compara seu ganho com os ganhos dos vizinhos com quem jogou. Se o ganho for menor do que o de algum vizinho, ele copia o estado deste vizinho, caso contrário mantém seu estado atual (estratégia darwiniana). Este processo se repete em todas as rodadas e assim a proporção de cooperadores no sistema se altera no decorrer do tempo.

Nowak & May utilizaram o chamado *dilema do prisioneiro fraco* no qual se define a matriz de ganhos como $T \geq R$, $R = 1$ e $P = S = 0$. Com as regras simples do dilema do prisioneiro e os jogadores na rede, durante a evolução temporal surgem padrões espaciais complexos com cooperadores e desertores coexistindo [34, 35]. Nestes trabalhos, Nowak & May obtiveram a cooperação facilitada pela auto-interação

que sempre gera um resultado vantajoso dentro de agrupamentos cooperativos, estabilizando as ilhas de cooperadores na população. Em 2006, Soares e Martinez [53] demonstraram que na ausência da auto-interação o transiente do sistema se altera, mas que o regime estacionário não apresenta mudanças significativas.

Quando consideramos a proporção assintótica de cooperadores ρ_∞ observa-se que surgem três regimes qualitativos de ρ_∞ em função do valor de T [54]:

- i. $1 < T \leq 4/3$, é caracterizado por ρ_∞ estacionário ou levemente periódico e ocorre uma maioria de cooperadores $\rho_\infty > 1/2$;
- ii. $4/3 < T \leq 3/2$, é caracterizado por ρ_∞ não estacionário e ocorrência de caos espaço-temporal;
- iii. $3/2 < T \leq 2$, semelhante aos valores baixos de T , é estacionário mas agora com minoria de cooperadores $\rho_\infty < 1/2$.

Além da estratégia darwiniana que foi utilizada nos primeiros trabalhos do DPE, também se destacou a *estratégia pavloviana* (EP) por sua robustez em sistemas com ruído ou com jogadores sujeitos a erros. Na EP, o jogador considera apenas o próprio ganho e o compara a uma ambição ou aspiração pessoal e só mudará seu estado caso seu ganho seja menor que essa aspiração. Em sistemas com ruído, uma mudança de estado aleatória pode ser considerada uma mutação e a cooperação pode emergir em sistemas completamente desertores e vice-versa.

Utilizando a EP na estrutura unidimensional, observou-se um novo regime, o *quasi*-regular, com nível médio de cooperação $\rho_\infty \sim 1/2$ e a cada rodada vários jogadores mudam de estado [48]. Além disso, a extinção da cooperação não foi observada pois quando todos os jogadores jogavam como desertores recebiam ganhos que não eram maiores que suas aspirações e mudavam seus estados para cooperadores. Assim, dependendo da aspiração dos jogadores podemos garantir a existência de cooperação no sistema.

2.1.4 Dilema do prisioneiro unidimensional: muitos jogadores, rede unidimensional e muitas rodadas

Nas Refs. [46, 47, 48], apresentamos uma nova abordagem para o DPI: utilizar uma rede unidimensional com conectividade (número de vizinhos) variável.

Colocamos os jogadores nos sítios da rede unidimensional e eles jogam o dilema do prisioneiro contra seus vizinhos. Utilizando a topologia unidimensional fizemos a exploração exaustiva do espaço de parâmetros. As motivações para a utilização da rede unidimensional foram a possibilidade de representar o histórico da evolução do sistema em uma imagem estática, onde um eixo representa o tempo e o outro os jogadores do sistema; e a possibilidade de variar a conectividade dos jogadores.

Na Ref. [46], mostramos a dinâmica de formação dos agrupamentos de cooperadores/desertores durante a evolução temporal e a partir desta dinâmica pudemos explicar o aspecto oscilatório do nível médio de cooperação no regime estacionário para sistemas onde os jogadores adotam a estratégia darwiniana. Na Ref. [47], ainda utilizando a ED, apresentamos os resultados da exploração exaustiva do espaço de parâmetros e mostramos a dependência do nível médio assintótico de cooperação em função do número de vizinhos com que cada jogador interage e se ele faz ou não a auto-interação (joga contra si mesmo). Na Ref. [48] tratamos das questões sobre a formação de padrões e a exploração exaustiva do espaço de parâmetros para o sistema em que os jogadores adotam a estratégia pavloviana e também o cálculo analítico dos valores das tentações críticas. Além disso, mostramos que o caso particular $T = 1$, que era até então considerado como um caso trivial¹ não é exatamente tão simples quando os jogadores jogam com mais de um oponente.

Nossos estudos mostraram que as transições de fase² [49] ocorrem nos mesmos valores de tentação crítica para as redes uni ou bidimensionais. No entanto, o nível médio assintótico de cooperação é levemente menor nos sistemas unidimensionais comparado ao obtido em sistemas bidimensionais, adotando-se a mesma estratégia e o mesmo nível médio inicial de cooperação em ambas as dimensionalidades.

2.1.5 Dilema do prisioneiro contínuo (DPC)

Quando cada jogador tem um nível de cooperação, ao invés de ter somente as opções de cooperar ou desertar, as quais são os extremos do nível de cooperação, temos o *dilema do prisioneiro contínuo* (DPC) [3, 4, 5, 6, 7, 8, 9, 10, 11, 55, 56, 57].

¹Considerava-se que os jogadores não mudariam de estado para $T = 1$, pois para este valor de tentação jogar com um cooperador ou desertor gera o mesmo ganho para um cooperador.

²No dilema do prisioneiro, considera-se uma transição de fase uma mudança significativa do nível médio assintótico de cooperação do sistema.

Os níveis de cooperação no DPC são baseados na interpolação linear do DP clássico com *duas*-escolhas. Chong *et al.* [58], adotou $T = 5$, $R = 4$, $P = 1$, $S = 0$ e a seguinte função para o ganho: $g_A = 5/2 - (1/2)\alpha_A + 2\alpha_B$, $-1 \leq \alpha_k \leq 1$, $\alpha_k = -1$ é um desertor e $\alpha_k = 1$ um cooperador, onde g_A é o ganho do jogador A , dado que α_k são os níveis de cooperação dos jogadores com $k = \{A; B\}$.

Ao elaborar a tabela de ganhos para o DP com n -escolhas, para $\alpha'_k \neq \alpha_k$, as seguintes condições devem ser satisfeitas:

1. para $\alpha_A < \alpha'_A$ e α_B constante: $g_A(\alpha_A, \alpha_B) > g_A(\alpha'_A, \alpha_B)$;
2. para $\alpha_A \leq \alpha'_A$ e $\alpha_B < \alpha'_B$: $g_A(\alpha_A, \alpha_B) < g_A(\alpha'_A, \alpha'_B)$;
3. para $\alpha_A < \alpha'_A$ e $\alpha_B < \alpha'_B$: $g_A(\alpha'_A, \alpha'_B) > (1/2)(g_A(\alpha_A, \alpha'_B) + g_A(\alpha'_A, \alpha_B))$.

Podemos observar que as condições acima são análogas às aquelas do DP de *duas*-escolhas ($T > R > P > S$ e $2R > T + S$).

Dadas a equação de ganhos e as três condições acima, definimos uma tabela de ganhos para o DP com n -escolhas. Por exemplo, a Tab. 2.2 mostra os ganhos para o DP de *quatro*-escolhas considerando $T = 5$, $R = 4$, $P = 1$ e $S = 0$. Esta tabela também ilustra dois pontos importantes: primeiro, os ganhos dos quatro cantos da tabela de ganhos de um DP com n -escolhas são os mesmos daqueles do DP com *duas*-escolhas; segundo, qualquer sub-tabela 2×2 da tabela $n \times n$ do DP com n -escolhas é por si só, um DP com *duas*-escolhas.

Tabela 2.2 – Ganho do jogador A no dilema do prisioneiro de *quatro*-escolhas com $T = 5$, $R = 4$, $P = 1$ e $S = 0$.

Ganho do jogador A		Nível de cooperação do jogador B			
		1	(1/3)	-(1/3)	-1
Nível de cooperação do jogador A	1	4	2(2/3)	1(1/3)	0
	(1/3)	4(1/3)	3	1(2/3)	(1/3)
	-(1/3)	4(2/3)	3(1/3)	2	(2/3)
	-1	5	3(2/3)	2(2/3)	1

A função generalizada de ganho

$$g(\alpha_i; \alpha_j) = (R + P - T - S)\alpha_i\alpha_j + (S - P)\alpha_i + (T - P)\alpha_j + P, \quad (2.1)$$

com $0 \leq \alpha_i, \alpha_j \leq 1$ onde $\alpha_i = 0$ é um desertor e $\alpha_i = 1$ um cooperador, que apresentamos na Ref. [48] generaliza os ganhos para o DP de n -escolhas para qualquer valor de n e para todo o intervalo do nível de cooperação, além de servir para qualquer matriz de ganhos. Essas propriedades da função ganho foram fundamentais neste trabalho, pois permitiram a criação de novas estratégias, que agregaram características intrínsecas do ser humano, como a liberdade de escolher a maneira que mais lhe agrade numa tomada de decisão ante a situações de risco.

O DPC é mais adequado à modelagem de sistemas reais do que o DPD, pois na natureza não é esperado um comportamento dramaticamente dicotômico, como os atos de cooperar ou desertar [7]. A formulação contínua permite um estudo mais apropriado em sistemas que buscam modelar situações reais e buscam responder duas questões principais sobre a evolução da cooperação:

1. Pode o altruísmo recíproco com níveis variáveis de cooperação resultar em comportamentos cooperativos surgindo gradualmente de um comportamento inicialmente egoísta? Algo que seria mais plausível em um cenário evolutivo.
2. O comportamento cooperativo pode ser mantido indefinidamente? Ou a peculiar natureza contínua resulta em estado cooperativo sendo gradualmente minado?

Portanto, o DPC permite estudar a emergência de comportamentos cooperativos e sua estabilidade, por exemplo. Killingback e Doebeli [7] introduzem um dilema do prisioneiro “interespecífico” baseado em investimento, que pode assumir valores contínuos. Esse investimento, do ponto de vista biológico, é comumente considerado como o custo para o cooperador prover um ganho maior para o desertor. A seguir, trazemos alguns exemplos biológicos que inspiram as pesquisas sobre o dilema do prisioneiro e mostram que o modelo contínuo parece ser o mais adequado em muitas situações.

2.2 Exemplos do DP

Apresentamos alguns exemplos de onde podemos encontrar o DP na natureza [18]. Os exemplos em sociedades biológicas são escassos, não apenas devido às

dificuldades em medir os parâmetros fundamentais, mas também devido ao alcance limitado do DPI.

2.2.1 Mimercofilia

Mimercofilia refere-se às associações de mutualismo com formigas. Leimar e Axén [19] apresentaram um caso de mutualismo entre formigas e lagartas da espécie *lycaenid* que pode ser modelado usando-se o dilema do prisioneiro. As lagartas possuem um “órgão dorsal de néctar” que secreta um líquido contendo carboidratos e aminoácidos livres. As formigas coletam o líquido e em troca oferecem proteção às lagartas. A produção do líquido tem um custo para as lagartas, então elas prefeririam receber a proteção das formigas sem ter que secretar o líquido. No entanto, as formigas para protegerem as lagartas, deixam de coletar comida, o que também implica em um custo para elas. Ambas teriam vantagens se desertassem, a lagarta não secretando o líquido enquanto recebe a proteção das formigas ou as formigas recolhendo o líquido das lagartas sem ter que protegê-las.

O problema das lagartas e das formigas tem algumas diferenças do dilema do prisioneiro iterado clássico: as recompensas são assimétricas; os movimentos não são necessariamente simultâneos; as decisões são influenciadas por terceiros - predadores e parasitas atacando a lagarta; o investimento em cada interação pode ser variado gradualmente (dilema do prisioneiro contínuo). Assim podemos utilizar o dilema do prisioneiro como modelo, mas vemos que as interações entre lagartas e formigas são mais complexas.

2.2.2 Inspeção de predador em cardumes

Evidências da existência de um regime do dilema do prisioneiro iterado (inclusive com a estratégia do TFT) são encontradas em peixes no caso da inspeção de predador, dos gêneros *Phoxinus*, *Poecilia* e *Gasterosteus* [20, 22]. Quando um cardume está sendo perseguido por um predador, alguns membros do cardume irão se separar do bando, nadar em direção ao predador até uma pequena distância dele, esperar lá por alguns segundos e então lentamente voltar ao cardume. Acredita-se que com essas excursões os peixes podem colher informações precisas sobre a identidade, localização e as atuais intenções do predador [20, 21, 22, 23, 24, 25].

Dois peixes podem se aproximar mais do predador do que um único peixe sozinho, portanto ambos recebem o mesmo benefício b e dividem o custo c , logo $R = b - c/2$. Se um dos peixes fica mais afastado, ele irá receber a mesma informação sem o custo, assim $T = b$, enquanto que o outro arcará com todo o custo, $S = b - c$. Se nenhum dos dois se aproxima, o custo e o benefício são nulos $P = 0$. Se os valores seguem às condições do DP: $T > R > P > S$, podemos modelar o problema usando o DP. O problema da inspeção de predadores ainda é controverso, pois não há um estudo de caso que prove que todas as condições do DP são satisfeitas [59]. São levantadas outras questões: os peixes pagam o custo da inspeção e depois compartilham as informações com o cardume? Qual a relação existente entre os membros do cardume?

2.2.3 Troca de ovos em hermafroditas simultâneos

Em estudos sobre peixes da família *Serranidae* (pertencem a esta família, garoupas e robalos, por exemplo), observou-se que o hermafroditismo simultâneo é uma característica comum [26, 27, 28]. Sua fertilização é externa com ovos planctônicos, isto é, ficam à deriva em áreas de mar aberto.

Durante a fertilização, eles dividem seus ovos em parcelas. Após a corte de acasalamento, eles consentem em fornecer parcelas de ovos para a fertilização. Provavelmente a produção de ovos é mais custosa que a produção de esperma. Portanto a tentação de desertar, isto é, fertilizar os ovos do companheiro sem fornecer ovos para a fertilização, é provavelmente alta. Então surge a pergunta, porque os serranídeos cooperam? Fischer [27] sugeriu que os peixes estão jogando usando a estratégia *tit-for-tat*.

2.2.4 Compartilhamento de sangue entre morcegos hematófagos

Aproximadamente 8% dos morcegos hematófagos adultos falham em conseguir comida em uma noite qualquer, mas nessas noites eles são frequentemente alimentados pelos companheiros que obtiveram sucesso. Essa cooperação independe do nível de proximidade de parentesco [60], mas depende da probabilidade de reciprocidade. As habilidades cognitivas dos morcegos podem ser suficientes para garantir

o reconhecimento de desertores e cooperadores [29].

O comportamento de compartilhar o alimento é fundamental na sobrevivência do morcego que não conseguiu sangue durante a noite, pois seu estoque de energia lhe fornece forças para uma sobrevivência de 48 a 72 horas sem uma nova refeição. A perda de peso pós alimentação tem um comportamento exponencial. Por isso, a doação de sangue significa uma diminuição no tempo de sobrevivência do doador menor que o ganho obtido pelo receptor. Então a matriz de ganhos $T > R > P > S$ seria plausível para representar compartilhamento de sangue entre os morcegos hematófagos. Logo, esta situação poderia ser modelada pelo dilema do prisioneiro. No entanto, os pesquisadores que trabalham com morcegos hematófagos, tem dificuldades em mensurar esses parâmetros essenciais ao modelo.

2.3 Aversão ao risco

Para fazer o estudo da aversão ao risco que será utilizada na estratégia baseada na personalidade, utilizamos como referência o trabalho de Anteneodo *et al* [15], no qual foram estudadas as consequências de atitudes avessas ao risco na dinâmica de operações econômicas. Isso foi feito introduzindo um autômato que simula transações monetárias ante escolhas arriscadas, entre operadores com diferentes atitudes.

A preferência por uma escolha com ganho menor, porém com grande probabilidade de ocorrência, em contraste à escolha com ganho maior, mas com probabilidade menor é a chamada *aversão ao risco*. De fato, a maioria das pessoas sente aversão a arriscar quando esperam ganhar (com probabilidade moderada ou alta). Por outro lado, têm tendência a arriscar quando esperam perder. Estas tendências são invertidas para probabilidades muito baixas. Certamente, este comportamento afeta a maioria das decisões humanas uma vez que fatores de chance (probabilidades) sempre estão presentes em, por exemplo, estratégias médicas, jogos de aposta ou transações econômicas.

As estratégias utilizadas no dilema do prisioneiro contínuo são pouco eficientes para determinar quais são as intenções das escolhas intermediárias, as quais podem ser um sinal para firmar uma cooperação ou uma exploração sutil. Ao não conseguirem determinar a intenção do oponente, as estratégias optam por jogadas

menos cooperativas que oferecem ganhos maiores e menor risco de serem exploradas a curto prazo, ou seja, uma aversão ao risco.

Em economia, a análise de decisão sob risco foi tratada pela *teoria de utilidade esperada* (EUT, do inglês *expected utility theory*), assumindo que os jogadores fazem escolhas racionais. O valor esperado U é dado por $U(\mathcal{P}) = \sum_{i=1}^N u(x_i)p_i$, onde a função utilidade $u(x_i)$ aumenta monotonicamente com x_i e $\mathcal{P} \equiv \{x_1, p_1; \dots; x_n, p_n\}$ corresponde ao prospecto, onde o resultado x_i (ganho se positivo; perda se negativo) acontece com uma probabilidade p_i .

Há, porém, aspectos de sensibilidade ao risco que não são contemplados adequadamente dentro da EUT. Experiências com problemas hipotéticos de escolha propostos por Kahneman e Tversky [61] exibiram estas características. Os autores propuseram uma generalização à equação da EUT com a “*teoria de prospecto*” (PT, do inglês, *prospect theory*): $U(\mathcal{P}) = \sum_{i=1}^N u(x_i)\Pi(p_i)$, onde a função peso $\Pi(p_i)$ aumenta monotonicamente com p_i , onde $\Pi(p) = p^{\hat{q}}$ com $\hat{q} \in \mathbb{R}$ e $\Pi(p; \hat{q}) = p^{\hat{q}}/[p^{\hat{q}} + (1 - p)^{\hat{q}}]$.

A função peso $\Pi(p_i)$, para $\hat{q} \leq 1$, é, basicamente, côncava para probabilidades baixas e convexa para altas, com $\Pi(0) = 0$ e $\Pi(1) = 1$. Para $\hat{q} > 1$ temos o inverso em relação a concavidade da função peso. Além disso, são satisfeitas às condições:

1. $\Pi(p) + \Pi(1 - p) \leq 1$ (a igualdade se mantém para $p = \{0; 1\}$) - sub-certeza;
2. $\Pi(ps)/\Pi(p) \leq \Pi(psr)/\Pi(pr)$ para $0 < p, r, s \leq 1$ - sub-proporcionalidade;
3. $\Pi < p$ para altas probabilidades - sobrepeso;
4. $\Pi > p$ para baixas probabilidades - sub-peso; e
5. para probabilidades muito baixas $\Pi(pr) > r\Pi(p)$ para $0 < r < 1$ - sub-aditividade.

Em todos estes casos, cada jogador é caracterizado por um conjunto de parâmetros que rendem uma particular $\Pi_{\hat{q}}(p)$, representando o processo subjetivo que o jogador faz para probabilidades conhecidas p em um jogo de chance.

Nesta seção 2.3, apresentamos uma abordagem para modelar a tomada de decisão frente ao risco. Tal abordagem se baseia nas ideias de utilidade esperada e

teoria do prospecto, que permeiam o meio acadêmico econômico e Otiveram suas origens nas ideias dos matemáticos *Blaise Pascal* (1623-1662) e *Pierre de Fermat* (1601-1655) com as primeiras teorias a tratarem o comportamento humano através de suas teorias de probabilidade. Assim, encerramos a introdução e partimos para o modelo de estudo.

Modelos

“Mas a prova lógica depende sempre de premissas, e a questão decisiva na investigação científica não está portanto em tirar logicamente as conclusões, mas sim em descobrir as premissas.”

Olavo Carvalho no livro *Como vencer um debate sem ter a razão* - 2007.

O dilema do prisioneiro iterado (DPI) já foi estudado em diversas topologias tais como: redes regulares [34, 35], grafos [37], grafos aleatórios [54], redes *livres de escala* [62] e redes do tipo *mundo pequeno* [63, 64]. É interessante observar que, em sistemas unidimensionais, o DPI não havia sido abordado devido à sua aparente simplicidade [54]. No entanto, redes unidimensionais possuem as mesmas transições de fase de redes bidimensionais, isto é, apresentam os mesmos valores de tentação crítica onde ocorrem mudanças abruptas na proporção assintótica de cooperadores [46, 47, 48]. Neste capítulo, apresentamos o modelo adotado nesta tese em três seções: a primeira seção traz a descrição do modelo em detalhes; a segunda, o modelo analítico e a última, o modelo computacional.

Os destaques deste capítulo ficam por conta da função generalizada de ganho apresentada na subseção 3.1.3 e pelas novas estratégias que propomos na subseção 3.1.5 que são novidades no estudo do dilema do prisioneiro. As novas estratégias são consequência do uso do dilema do prisioneiro contínuo e da função generalizada de ganho.

3.1 Caso unidimensional

Nesta seção detalhamos o modo em que os jogadores estão dispostos na rede (autômato), como são formadas suas vizinhanças (conectividade), como interagem e atualizam seus estados (dinâmica do sistema) e as regras que regem as mudanças de estados (estratégias). Utilizamos um autômato celular unidimensional, onde todas as células estão conectadas com o mesmo número de vizinhos e utilizam a mesma estratégia.

3.1.1 Autômato celular

Autômatos celulares [65, 66, 67] são idealizações matemáticas de sistemas físicos nos quais espaço e tempo são discretos e as grandezas físicas assumem valores pertencentes a um conjunto discreto de valores. Um autômato celular consiste em uma rede regular uniforme, normalmente infinita em extensão, com uma variável discreta¹ (estado) em cada sítio (célula).

Nas Refs. [46, 47, 48] apresentamos uma nova abordagem para o DPI, a qual foi utilizada na presente tese. Esta abordagem utiliza um autômato celular em uma rede unidimensional regular, onde cada célula do autômato celular representa um jogador ou uma família (vários membros no mesmo sítio). A partir deste ponto, sempre que utilizarmos o termo jogador estamos nos referindo a uma célula do autômato, que pode representar um único jogador ou então uma família ou grupo.

O *estado* de cada jogador é representado pelo *nível de cooperação*, α_i , com $\alpha_i \in \mathbb{R} : 0 \leq \alpha_i \leq 1$. No *dilema do prisioneiro discreto* (DPD), α_i assume os valores discretos $\alpha_i = 0$, quando o jogador é desertor e $\alpha_i = 1$, quando cooperador. No *dilema do prisioneiro contínuo* (DPC), $0 \leq \alpha_i \leq 1$ assume valores contínuos, sendo $\alpha_i = 0$ deserção e $\alpha_i = 1$ cooperação. O esquema da Fig. 3.1 ilustra exemplos de autômatos para o dilema do prisioneiro discreto (DPD) e para o dilema do prisioneiro contínuo (DPC).

O nível de cooperação contínuo pode ser interpretado das seguintes maneiras:

¹Na definição inicial de autômatos, proposta por Wolfram, considerou-se componentes básicos discretos. No entanto, em alguns casos, o comportamento agregado de grandes números desses componentes podem ser efetivamente contínuos [68]. Então, se consideramos uma célula como um agregado de componentes básicos, podemos atribuir valores contínuos para o nível de cooperação da célula.

Figura 3.1 – Autômato celular na rede unidimensional com $N = 11$ jogadores. (a) dilema do prisioneiro discreto (DPD) - jogador coopera ou deserta; (b) dilema do prisioneiro contínuo (DPC) - jogador tem um nível de cooperação. Esquema de cores: branco representa cooperação, preto deserção e tons de cinza os níveis intermediários de cooperação.

- (i) cooperação parcial: existe a possibilidade de uma cooperação ou deserção parcial, algo equivalente à situação onde o prisioneiro pode fornecer diferentes quantidades de indícios para incriminar seu comparsa. Neste caso, o prisioneiro pode evitar os riscos que corre com uma cooperação plena, ou ser benevolente com seu comparsa e não o delatar completamente, e assim seu ganho resulta em valores intermediários entre os ganhos máximo e mínimo do jogo.
- (ii) dilema do prisioneiro discreto com nível de cooperação estocástico: cada jogador coopera com probabilidade α_i . Desta maneira, jogando n vezes, ele joga $\alpha_i n$ vezes, em média, como cooperador e as demais vezes como desertor. O ganho obtido é igual ao ganho médio de um conjunto de $\alpha_i n$ cooperações e $(1 - \alpha_i)n$ deserções.
- (iii) nível médio de cooperação do grupo: se a célula representar um grupo de jogadores (família), α_i é a proporção de cooperadores entre os membros do grupo (DPD) ou o nível médio de cooperação (DPC).

Utilizamos no decorrer do texto a nomenclatura *nível médio de cooperação*, mesmo quando estivermos falando sobre dilema do prisioneiro discreto, sendo que neste caso o nível médio de cooperação é equivalente à proporção de cooperadores. Sendo N , o tamanho do autômato (número de jogadores, grupos, famílias ou células),

o nível médio de cooperação, ρ , do sistema é definido como

$$\rho = \frac{1}{N} \sum_{j=1}^N \alpha_j, \quad (3.1)$$

onde α_j é o nível de cooperação do j -ésimo jogador.

Além do nível de cooperação, podemos considerar que cada jogador possui uma personalidade $\hat{q}$, com $\hat{q} \in \mathbb{R} : 0 \leq \hat{q} \leq \infty$ que caracteriza os jogadores como afins ou avessos ao risco, em função da probabilidade de ganhar ou perder, que eles consideram no momento de sua tomada de decisão. A personalidade será utilizada na estratégia em que os jogadores consideram um ganho subjetivo, de acordo com suas personalidades, para avaliar seus resultados. A inclusão da personalidade está baseada no estudo de aversão ao risco de Anteneodo *et al* [15] e será explicada em detalhes na Sec. 3.1.5.

Todos os jogadores de um dado sistema adotam a mesma estratégia, tem a mesma conectividade (número de vizinhos) e recebem a mesma quantidade e qualidade de informação fornecida aos demais jogadores, ou seja, há uma equidade entre os jogadores nas condições de jogo. Em uma dada rodada, cada jogador utiliza o mesmo nível de cooperação com todos os seus vizinhos.

3.1.2 Conectividade

Os jogadores interagem jogando o dilema do prisioneiro contra seus “vizinhos”. A conectividade (“vizinhança” ou número de contatos) é dado por k , com $k \in \mathbb{N} : 1 \leq k \leq N$. A vizinhança do i -ésimo jogador é um conjunto $\kappa_i = \{\alpha_j, \hat{q}_j\}$, $1 \leq j \leq k$, que contém os níveis de cooperação e personalidades de todos os seus k vizinhos.

Utilizamos três tipos de vizinhança: *regular*, *aleatória fixa* e *aleatória dinâmica*. A *vizinhança regular* de um jogador são os jogadores adjacentes na rede que estão situados à sua esquerda e direita. Sendo k um valor par, metade dos vizinhos ($k/2$) fica à esquerda e a outra metade à direita (vide Fig. 3.2a). Sendo k ímpar, o jogador realiza a *auto-interação* - interage consigo mesmo - e cada lado tem $(k-1)/2$ vizinhos (vide Fig. 3.2b). Por simplicidade, a vizinhança regular que consideramos é sempre simétrica. Nowak & May [35] explicam a auto-interação considerando que uma célula é ocupada por vários animais (uma família) ou moléculas, em modelos

biológicos, por exemplo. Assim, a *auto-interação* é uma interação *intra-grupo* e as demais são interações *inter-grupos*.

Figura 3.2 – Vizinhança do jogador central (de onde partem as setas). Em sistemas que utilizam vizinhança regular, todos os jogadores utilizam o esquema representado pelas setas para definir seus vizinhos. (a) $k = 8$ (sem auto-interação). (b) $k = 9$ (com auto-interação).

Na *vizinhança aleatória fixa*, no início do jogo, cada jogador escolhe uma vizinhança de tamanho k , aleatoriamente distribuída. A vizinhança que foi escolhida por cada jogador se mantém fixa durante todo o jogo. No momento da escolha da vizinhança de cada jogador, todos os jogadores da rede (inclusive o próprio jogador - auto-interação) são elegíveis como vizinhos com a mesma probabilidade.

Na *vizinhança aleatória dinâmica*, no início de cada rodada, cada jogador escolhe seus k vizinhos aleatoriamente, algo semelhante aos jogadores apresentarem mobilidade [69, 70]. Da mesma forma que na vizinhança aleatória fixa, todos os jogadores estão igualmente disponíveis no momento da escolha da vizinhança de cada jogador a cada rodada.

Nos casos de vizinhança aleatória, não há necessidade de simetria de interação, ou seja, se o jogador A escolhe o jogador B como vizinho, B não é obrigado a escolher A como vizinho, isto é, $B \in \kappa_A \nRightarrow A \in \kappa_B$. Note que na vizinhança regular o “alcance” dos jogadores é $k/2$, com simetria de interação ($B \in \kappa_A \Rightarrow A \in \kappa_B$), enquanto que nas aleatórias o alcance é N , pois todos os jogadores do sistema são

igualmente acessíveis como vizinhos, inclusive o próprio jogador (auto-interação).

Podemos fazer uma analogia entre física de sistemas desordenados [71] e vizinhanças aleatórias. Comparamos o caso da vizinhança aleatória fixa a um sistema com *desordem congelada* (*quenched disorder*), isto é, aquele onde os parâmetros que definem seu comportamento, são variáveis aleatórias que não evoluem no tempo. Enquanto que, a vizinhança aleatória dinâmica se compara a um sistema com *desordem recozida* (*annealed disorder*), pois os parâmetros podem variar durante a evolução temporal. Em nosso problema, a variável aleatória é a vizinhança de cada jogador.

Para que os jogadores dos extremos da rede tenham a mesma conectividade dos demais, utilizamos condições periódicas de contorno (CPC), ligando a “fita” como um “anel” como vemos na Fig. 3.3a. No caso bidimensional, as condições periódicas de contorno são representadas por um toróide da Fig. 3.3b.

Figura 3.3 – Representação de sistemas com condições periódicas de contorno (CPC). (a) Anel representa o sistema unidimensional com suas extremidades conectadas. (b) Toróide representa um sistema bidimensional com suas bordas conectadas.

Definida a vizinhança, apresentamos o cálculo do ganho de cada jogador.

3.1.3 Ganhos

Quando o i -ésimo jogador interage com o oponente j (interação por pares), ele recebe um ganho $g(\alpha_i; \alpha_j)$. Esta função generalizada de ganho foi primeiramente apresentada na Ref. [48].

$$g(\alpha_i; \alpha_j) = (R + P - T - S)\alpha_i\alpha_j + (S - P)\alpha_i + (T - P)\alpha_j + P. \quad (3.2)$$

O ganho depende dos valores da matriz de ganhos (T (tentação), P (punição), R (recompensa) e S (tolice), com $T > R > P > S$, $2R > T + S$), do nível de cooperação do próprio jogador α_i e do nível de cooperação do oponente α_j . Tanto no DPD quanto no DPC, $0 \leq \alpha_k \leq 1$ e $\alpha_k = 1$ representa cooperação e $\alpha_k = 0$ deserção.

A função $g(\alpha_i; \alpha_j)$ é válida tanto para o DPD quanto para o DPC e, a partir dela, também podemos obter os valores de ganho para o DP de *quatro*-escolhas apresentado no estudo de Chong & Yao [58]. Por exemplo, um jogador, com $\alpha_i = 0,3$, joga contra um cooperador, $\alpha_j = 1$, e o ganho é $g(\alpha_i = 0,3; \alpha_j = 1) = 0,7T + 0,3R$. Na literatura, não há outra função de ganho para o dilema do prisioneiro que seja válida, tanto para o dilema do prisioneiro discreto, quanto para o contínuo, e ainda permita o uso de qualquer matriz de ganhos que respeite as condições de conflito do DP.

Podemos obter uma superfície com os possíveis ganhos por interação de um jogador no DPC, a partir da Eq. 3.2, como mostra a Fig. 3.4, onde também destacamos os ganhos para o DPD. Para obter os ganhos do DPD, basta substituir $\alpha_i = 1$, para cooperação, e $\alpha_i = 0$, para deserção, que são os valores extremos do nível de cooperação, na função generalizada de ganho.

Agora, apresentamos os possíveis ganhos que um jogador pode obter ao jogar como cooperador ou desertor confrontando cooperadores e desertores. O jogador i interage com sua vizinhança κ_i de tamanho k . No caso discreto, a vizinhança é constituída somente de cooperadores e desertores, portanto $k = c + d$, onde c é o número de cooperadores e d o número de desertores.

O ganho total G_i é dado pela soma dos ganhos das interações com todos seus vizinhos (κ_i):

$$G_i = G^{\kappa_i}(\alpha_i) = \sum_{j=1}^k g(\alpha_i; \alpha_j). \quad (3.3)$$

Lembramos que devido à forma como definimos nossa conectividade, apenas no caso de k ímpar ocorre a auto-interação, e $j = i \Rightarrow \alpha_i \in \kappa_i$, e que o ganho total depende da configuração de κ_i (níveis de cooperação dos vizinhos). Nas estratégias adotadas neste estudo, os jogadores consideram o ganho total na tomada de decisão quanto à mudança do nível de cooperação. A forma como o ganho é considerado na tomada

Figura 3.4 – Ganho por interação do jogador A, $g(\alpha_A; \alpha_B)$, em função do seu nível de cooperação, α_A , e do nível de cooperação do oponente, α_B , dado pela Eq. 3.2. Os níveis de cooperação $\alpha_{A,B}$ estão no intervalo $0 \leq \alpha_{A,B} \leq 1$ e os valores para a matriz de ganhos são: $T = 5$, $R = 3$, $P = 1$ e $S = 0$. Os pontos destacados correspondem aos ganhos do dilema do prisioneiro discreto (dilema de prisioneiro de duas-escolhas).

de decisão varia de acordo com a estratégia adotada e será explicada em detalhes na apresentação das estratégias (Seção 3.1.5). Basicamente, ou o jogador considera seu ganho como um ganho relativo - quanto ganhou comparado aos seus oponentes; ou como um ganho absoluto - obteve o ganho desejado ou não, independentemente do ganho dos oponentes.

O ganho obtido pelo i -ésimo jogador, caso ele seja cooperador, é:

$$g(\alpha_i = 1; \alpha_j) = (R - S)\alpha_j + S \quad (3.4)$$

e, caso seja desertor:

$$g(\alpha_i = 0; \alpha_j) = (T - P)\alpha_j + P. \quad (3.5)$$

O ganho obtido pelo i -ésimo jogador, devido à sua interação com um único cooperador, é:

$$g(\alpha_i; \alpha_j = 1) = (R - T)\alpha_i + T \quad (3.6)$$

e, contra um desertor:

$$g(\alpha_i; \alpha_j = 0) = (S - P)\alpha_i + P. \quad (3.7)$$

No caso discreto, o ganho do i -ésimo jogador devido às interações com os c cooperadores de sua vizinhança κ_i é: $c[g(\alpha_i; \alpha_j = 1)] = c[(R - T)\alpha_i + T]$ e com os d desertores: $d[g(\alpha_i; \alpha_j = 0)] = d[(S - P)\alpha_i + P]$. O ganho total G_i é dado por:

$$\begin{aligned} G_i &= [(R - T)\alpha_i + T]c + [(S - P)\alpha_i + P]d, \\ &= [(R - T)c + (S - P)d]\alpha_i + Tc + Pd, \end{aligned}$$

fazendo $d = k - c$:

$$G_i = [(R + P - T - S)c + (S - P)k]\alpha_i + (T - P)c + Pk. \quad (3.8)$$

O ganho total do jogador i jogando como desertor é:

$$G(\alpha_i = 0) = (T - P)c + Pk \quad (3.9)$$

e como cooperador:

$$G(\alpha_i = 1) = (R - S)c + Sk. \quad (3.10)$$

Retomando o caso contínuo, com a Eq. 3.2 podemos calcular, também, o ganho esperado $E[g(\alpha_i; \rho)]$. Isto é, o ganho que um jogador, com nível de cooperação α_i , espera obter em média por interação, quando joga em um sistema com nível médio de cooperação ρ . Assim, se $\{\alpha_i; \rho\} \in \mathbb{R} : 0 \leq \{\alpha_i; \rho\} \leq 1$ e logo, $g(\alpha_i; \rho) = E[g(\alpha_i; \rho)]$. O ganho esperado por interação é dado por:

$$E[g(\alpha_i; \rho)] = \frac{1}{N} \sum_{j=1}^N g(\alpha_i, \alpha_j). \quad (3.11)$$

Substituindo a eq. 3.2, obtemos:

$$E[g(\alpha_i; \rho)] = (R + P - T - S)\alpha_i\rho + (S - P)\alpha_i + (T - P)\rho + P. \quad (3.12)$$

Se o jogador for um cooperador, o ganho esperado por interação é

$$E[g(\alpha_i = 1; \rho)] = (R - S)\rho + S \quad (3.13)$$

e se for desertor:

$$E[g(\alpha_i = 0; \rho)] = (T - P)\rho + P. \quad (3.14)$$

O ganho total esperado $E[G_i]$ devido às k interações é $kE[g(\alpha_i; \rho)]$:

$$E[G_i] = k[(R + P - T - S)\alpha_i\rho + (S - P)\alpha_i + (T - P)\rho + P]. \quad (3.15)$$

Logo, um cooperador tem um ganho total esperado igual a:

$$E[G(\alpha_i = 1)] = k[(R - S)\rho + S], \quad (3.16)$$

e um desertor:

$$E[G(\alpha_i = 0)] = k[(T - P)\rho + P]. \quad (3.17)$$

Note que o ganho total também poderia ter sido obtido multiplicando-se o ganho médio $g(\alpha_i; \rho)$ pelo número de interações k e fazendo $\rho k = c$. No caso discreto é fácil ver que o nível médio de cooperação é o número total de cooperadores no sistema dividido pelo tamanho do sistema. Então, em uma amostra de tamanho r de um sistema com nível médio de cooperação ρ , esperamos ter $r\rho$ cooperadores.

O ganho máximo por interação do i -ésimo jogador ocorre quando ele joga contra um cooperador ($\alpha_j = 1$), e o mínimo, quando ele joga contra um desertor ($\alpha_j = 0$). Se o jogador em questão for um desertor ($\alpha_i = 0$) e jogar contra um cooperador ($\alpha_j = 1$), ele obtém o ganho máximo, que é também, o ganho máximo no jogo, ou *máximo global*. Por outro lado, se ele for cooperador ($\alpha_i = 1$) e jogar contra um desertor ($\alpha_j = 0$), ele obtém o ganho mínimo no jogo, ou *mínimo global*. Se um jogador tiver um nível intermediário de cooperação $0 < \alpha_i < 1$ e jogar contra um cooperador ($\alpha_j = 1$), ele obterá o seu ganho máximo, mas este é um *máximo local*. Jogando contra um desertor ($\alpha_j = 0$), receberá o ganho *mínimo local*. Os ganhos máximo e mínimo locais são diferentes para cada jogador, pois dependem do nível de cooperação α_i do jogador.

O ganho máximo local, $g_{\max_L}(\alpha_i)$, é dado por:

$$g_{\max_L}(\alpha_i) = g(\alpha_i; \alpha_j = 1) = (R - T)\alpha_i + T, \quad (3.18)$$

e o ganho mínimo local, $g_{\min_L}(\alpha_i)$, por:

$$g_{\min_L}(\alpha_i) = g(\alpha_i; \alpha_j = 0) = (S - P)\alpha_i + P. \quad (3.19)$$

Na superfície da Fig. 3.4, os valores de ganhos máximos locais formam a linha sobre o nível de cooperação do jogador B igual a um, $\alpha_B = 1$ (cooperador), indo de R a T e os valores de ganhos mínimos locais estão sobre a linha $\alpha_B = 0$ (desertor), indo de S a P .

O ganho máximo global, $g_{\max_U}$, é dado por:

$$g_{\max_U} = g(\alpha_i = 0; \alpha_j = 1) = T, \quad (3.20)$$

e o ganho mínimo global, $g_{\min_U}$, por:

$$g_{\min_U} = g(\alpha_i = 1; \alpha_j = 0) = S. \quad (3.21)$$

Estes valores são os extremos da superfície de ganho por interação (vide Fig. 3.4).

O ganho por interação do i -ésimo jogador está contido no intervalo $g_{\min}(\alpha_i) \leq g(\alpha_i; \alpha_j) \leq g_{\max}(\alpha_i)$. Para um jogador com nível de cooperação qualquer, o intervalo de ganho local $\Delta_L(\alpha_i)$ vale:

$$\Delta_L(\alpha_i) = g_{\max_L}(\alpha_i) - g_{\min_L}(\alpha_i) = (R + P - T - S)\alpha_i + T - P, \quad (3.22)$$

e no jogo, o intervalo de ganho global Δ_U é igual a:

$$\Delta_U = g_{\max_U} - g_{\min_U} = T - S. \quad (3.23)$$

As equações de ganho por interação, ganhos máximo e mínimo e intervalos de ganhos, serão utilizadas para definir as mudanças de níveis de cooperação pelas estratégias.

3.1.4 Dinâmica do sistema

A evolução temporal do sistema é determinista. Todos os N jogadores jogam com suas respectivas vizinhanças κ_i , recebem seus ganhos e atualizam seus níveis de cooperação de forma síncrona, de acordo com a estratégia que estão adotando.

Fazendo alusão à nomenclatura de teoria dos jogos, chamamos de *jogada* a interação do jogador com todos seus vizinhos e de *rodada* a jogada mais a atualização dos níveis de cooperação de todos os jogadores. As etapas de uma rodada para um jogador são: a soma dos ganhos de sua jogada, a comparação do seu ganho com os ganhos dos vizinhos (que já devem saber seu ganho total), a tomada de decisão quanto à sua mudança de nível de cooperação e a atualização do nível de cooperação de acordo com o que foi determinado por sua estratégia. A duração do jogo é dada em rodadas.

Definimos *mudança de estado* de um jogador como a alteração do seu estado de cooperador para desertor ou vice-versa, no caso do DPD, e como o aumento ou diminuição do nível de cooperação atual, no caso do DPC. No decorrer do jogo o *nível médio de cooperação*, ρ varia em função das mudanças dos níveis de cooperação dos jogadores. No primeiro estágio da *evolução temporal* do sistema, existe um *regime transiente*, onde ocorrem as mudanças mais efetivas no sistema, no sentido de delinear o comportamento do *regime estacionário*.

O regime estacionário pode apresentar um nível médio de cooperação *estável* ou *oscilatório*. Em um regime estacionário estável, podemos supor duas situações para explicar esta condição de estabilidade. Uma primeira, na qual todos os jogadores não mudam de estado. Ou então, há vários jogadores que mudam seus níveis de cooperação, mas as mudanças ocorrem simultaneamente e não modificam o nível médio de cooperação do sistema, pois, enquanto alguns jogadores aumentam seus níveis de cooperação, outros diminuem.

Agora que definimos a dinâmica de evolução dos sistemas, apresentamos as estratégias que são as responsáveis pelas mudanças de estado dos jogadores, e portanto, fundamentais na definição das principais características observadas durante os regimes transiente e estacionário.

3.1.5 Estratégias

Em teoria dos jogos, *estratégia* se refere à uma das opções de ação que um jogador pode escolher em um jogo. No caso do dilema do prisioneiro, essas opções seriam cooperar ou desertar [72]. O termo estratégia também é utilizado para definir o plano de ação de um jogador ou plano de contingência, que leva em conta as possíveis ações do oponente. Estes dois usos do termo estratégia podem causar confusão. Então, definimos que utilizamos o termo estratégia, como o conjunto de regras que definem a tomada de decisão dos jogadores.

Em um dado sistema, todos os jogadores adotam a mesma estratégia, ou seja, não há coexistência de estratégias. Neste estudo, foram utilizadas quatro estratégias: uma darwiniana (ED) [34], duas baseadas no ganho (EBG) e uma subjetiva (ES). Além destas, apresentamos também a estratégia pavloviana (EP) [38], que apesar de não ter sido utilizada nas simulações, apresenta alguns princípios que fo-

ram utilizados nas estratégias baseadas no ganho. O papel da personalidade é pouco explorado na explicação da heterogeneidade nas tomadas de decisão no DP [73]. Por isso, propusemos a estratégia subjetiva, baseada na personalidade dos jogadores.

Começamos apresentando a estratégia pavloviana que é a mais simples, em seguida a darwiniana, depois a baseada no ganho que pode ser considerada como uma combinação da darwiniana e da pavloviana e a subjetiva que é uma variante da darwiniana.

3.1.5.1 Estratégia pavloviana (EP)

Na EP [36, 37, 38, 39, 40, 41, 42], o jogador compara seu ganho com sua ambição ou aspiração, φ . Se o ganho for maior que φ , ele mantém o estado atual, caso contrário, muda de estado. Esta estratégia é baseada no modelo pavloviano de reforço positivo para “acertos”. Reforço positivo é premiar, recompensar ou incentivar de alguma forma um sujeito, quando ele realiza uma ação que se pretende reforçar em seu comportamento ou uma ação que se pretende “ensinar” ao sujeito. Outra forma de reforçar a ação desejada é punir ações que não condizem com o desejado. Nesta variante, utiliza-se um reforço negativo, que é retirar uma punição que era aplicada anteriormente. Em nosso problema, utilizamos um reforço positivo como sendo obter um ganho maior que a aspiração. Devido à dificuldade em definir o que seria uma “mudança de estado”, no DPC, utilizando a ideia de reforço positivo ou negativo do modelo pavloviano, não utilizamos esta estratégia nas simulações do dilema do prisioneiro contínuo.

A EP se mostra bastante robusta em ambientes na presença de [74, 41]:

- ruído, isto é, existe uma probabilidade p , não nula, de um jogador mudar seu estado, independentemente da estratégia que ele utiliza;
- estratégias agressivas, enganadoras ou aproveitadoras;
- competição pela sobrevivência, em jogos co-evolucionários.

Essa robustez é decorrente do fato que a estratégia pavloviana [36]:

- é implacável, não se preocupa em perdoar uma deserção;

- explora estratégias altruístas enquanto não é punida com deserções, que façam com que o ganho obtido seja menor que a aspiração;
- pode corrigir erros ocasionais, o que não acontece na estratégia *tit-for-tat*, por exemplo.

No entanto, sua principal fraqueza é que ela pode ser explorada por estratégias desertoras sutis.

Em sistemas que adotam a EP, normalmente as simulações são realizadas sob as seguintes condições: em um dado sistema, todos os jogadores possuem a mesma aspiração φ ; a EP possui uma matriz de ganhos modificada para apresentar valores negativos que fariam o papel de uma punição, assim os ganhos seriam T , R , $P = -R$ e $S = -T$; e a aspiração dos jogadores seria obter um ganho não nulo: $\varphi > 0$. Mas, para evitarmos ganhos negativos² trasladamos³ a matriz de ganhos e a aspiração. Assim, a aspiração passa a ser: $\varphi > T/2$.

Uma vez apresentada a EP, que nos servirá como inspiração nas estratégias baseadas no ganho, vamos à estratégia darwiniana, que é a mais utilizada na literatura para as simulações do dilema do prisioneiro.

3.1.5.2 Estratégia darwiniana (ED)

Na estratégia darwiniana (ED), o jogador compara o seu ganho total com os ganhos totais obtidos pelos vizinhos. O jogador que obtém o maior ganho na vizinhança é considerado o melhor adaptado. Se o jogador não obtiver um ganho maior que o do melhor adaptado de sua vizinhança, ele copia o nível de cooperação do jogador melhor adaptado na vizinhança. Caso contrário, ele mantém o nível de cooperação atual. A ação de copiar o melhor adaptado é equivalente à replicação do mais apto no ambiente, o que caracteriza a estratégia darwiniana.

Quando o i -ésimo jogador não tem um ganho maior que o do melhor adaptado de sua vizinhança κ_i , ele copia o nível de cooperação do melhor adaptado da seguinte forma: observa a diferença ϵ_i entre seu nível de cooperação atual, α_i^{atual} , e o do melhor adaptado, $\alpha_{vencedor}$:

$$\epsilon_i = \alpha_{vencedor} - \alpha_i^{atual}. \quad (3.24)$$

²O cálculo do coeficiente de Gini assume riquezas positivas e não nulas.

³A transformação aplicada foi $x' = (x + T)/2$.

Então, modifica α_i^{atual} de acordo com a função⁴:

$$\alpha_i^{novo} = \alpha_i^{atual} + \epsilon_i. \quad (3.25)$$

Assim, o jogador copiaria o estado do melhor adaptado. No caso do DPD, ele poderia passar diretamente de cooperador para desertor ou vice-versa se o melhor adaptado estivesse no estado oposto ao seu estado atual. Note que a ação de mudar de estado pode causar mudanças abruptas no nível médio de cooperação do sistema.

No entanto, no DPC há uma possibilidade de se fazer mudança de estado mais gradativa. Para explorar esta oportunidade trazida pelo nível de cooperação contínuo, utilizamos uma mudança do nível de cooperação proporcional à diferença entre o ganho obtido e os possíveis ganhos máximos, que são as estratégias baseadas no ganho [75], que apresentamos na subseção seguinte.

3.1.5.3 Estratégia baseada no ganho (EBG)

Na estratégia baseada no ganho, a mudança em α_i é proporcional à diferença entre o ganho máximo por interação e o ganho médio por interação obtido pelo jogador. Essa proporção, δ_i , é calculada como:

$$\delta_i = \frac{g_i^{\max} - \frac{G_i}{k}}{\Delta}, \quad (3.26)$$

onde $g_i^{\max}$ é um ganho máximo de referência para o jogador, G_i é a soma dos ganhos na rodada, k é número de vizinhos com quem ele interagiu e Δ é o intervalo de ganhos que mantém δ_i entre zero e um. Todos os jogadores têm informação plena sobre os ganhos máximo e mínimo - globais e locais - que podem ser obtidos em cada jogada.

Assim na estratégia baseada no ganho, o jogador atualiza seu estado da seguinte maneira:

$$\alpha_i^{novo} = \alpha_i^{atual} + \epsilon_i \delta_i. \quad (3.27)$$

Lembrando que $\epsilon_i = \alpha_{vencedor} - \alpha_i^{atual}$. Se o ganho do jogador está próximo ao ganho máximo, $\delta \rightarrow 0$, e ele muda pouco seu nível de cooperação atual e quanto mais próximo do ganho mínimo, $\delta \rightarrow 1$, maior será a mudança em α_i .

⁴A função α_i^{novo} é apresentada desta forma intencionalmente para posterior modificação com outras estratégias. No caso, como a estratégia é copiar o jogador melhor adaptado, a função poderia ter sido escrita como: $\alpha_i^{novo} = \alpha_{vencedor}$.

A estratégia baseada no ganho local utiliza para a comparação o ganho máximo local $g_{\max_L}(\alpha_i)$ (Eq. 3.18) e a baseada no ganho global utiliza o ganho máximo global $g_{\max_U}$ (Eq. 3.20).

Ganho máximo local

Na estratégia baseada no ganho local (EBGL), o jogador considera os ganhos máximo e mínimo locais, isto é, aqueles tangíveis ao seu α_i . O intervalo de ganhos considerado na EGBL é $\Delta_L(\alpha_i) = (R + P - T - S)\alpha_i + T - P$. A diferença δ_i^L entre o ganho máximo (tangível) e o ganho médio obtido por interação, normalizado por $\Delta_L(\alpha_i)$ é

$$\begin{aligned}\delta_i^L &= \frac{g_{\max_L}(\alpha_i) - \frac{G_i}{k}}{\Delta_L(\alpha_i)}, \\ \delta_i^L &= \frac{(R - T)\alpha_i + T - \frac{G_i}{k}}{(R + P - T - S)\alpha_i + T - P}.\end{aligned}\tag{3.28}$$

O estado modificado do jogador é igual a

$$\alpha_i^{novo} = \alpha_i^{atual} + \epsilon_i \delta_i^L.\tag{3.29}$$

Ganho máximo global

A estratégia baseada no ganho global (EBGG) ocorre quando um jogador considera os ganhos globais do jogo. O intervalo de ganhos considerado é $\Delta_U = T - S$. A diferença δ_i^U entre o ganho máximo por interação e o ganho médio obtido por interação, normalizado por Δ_U é

$$\begin{aligned}\delta_i^U &= \frac{g_{\max_U} - \frac{G_i}{k}}{\Delta_U}, \\ \delta_i^U &= \frac{T - \frac{G_i}{k}}{T - S}.\end{aligned}\tag{3.30}$$

Da mesma forma, o estado modificado do jogador é igual a

$$\alpha_i^{novo} = \alpha_i^{atual} + \epsilon_i \delta_i^U.\tag{3.31}$$

Pela Eq. 3.28, notamos que adotando a estratégia baseada no ganho local, um jogador considera o seu nível de cooperação atual α_i e utiliza como aspiração o ganho máximo que ele pode atingir com α_i . Ao passo que na estratégia baseada no ganho

global (Eq. 3.30), o jogador está desconsiderando o próprio nível de cooperação e considerando como aspiração o ganho máximo no jogo. Nas EBG, o jogador não compara seu ganho com o ganho do vizinho melhor adaptado, compara o seu ganho com o máximo do jogo, uma aspiração.

No entanto, o jogador, além de observar a diferença entre o ganho obtido e o ganho máximo - local ou global - observa também a diferença entre o seu nível de cooperação e o do vizinho melhor adaptado, $\alpha_{vencedor}$, e supõe que $\alpha_{vencedor}$ deve ser o nível de cooperação capaz de o levar a atingir o ganho máximo. Assim, ele muda seu estado em direção ao estado do vizinho vencedor.

Nas EBG, um jogador muda seu estado copiando o nível de cooperação do vizinho melhor adaptado, equivalente à estratégia darwiniana, mas observa o ganho máximo - local ou global - e os toma como uma aspiração, equivalente à estratégia pavloviana. Por isso, consideramos as EBG como uma combinação das estratégias darwiniana e pavloviana. Finalmente, apresentamos a terceira nova estratégia proposta neste estudo, a estratégia subjetiva.

3.1.5.4 Estratégia subjetiva (ES)

Na estratégia subjetiva (ES), um jogador avalia seu ganho total por um valor subjetivo pessoal representado por aquele ganho, que depende do seu nível de cooperação e de sua personalidade. Ele compara seu ganho subjetivo com o ganho real dos vizinhos e reage da mesma forma que na estratégia darwiniana: caso seu ganho subjetivo seja menor que o ganho do jogador melhor adaptado na vizinhança, ele copia o nível de cooperação deste jogador.

O ganho subjetivo GS da ES é dado por:

$$GS_i = G_i \Pi(\alpha_i; \hat{q}_i), \quad (3.32)$$

onde G_i é o ganho total do jogador (Eq. 3.3) e $\Pi(\alpha_i; \hat{q})$ é a função subjetividade na avaliação do ganho sendo igual a:

$$\Pi(\alpha_i; \hat{q}) = \frac{\alpha_i^{\hat{q}}}{\alpha_i^{\hat{q}} + (1 - \alpha_i)^{\hat{q}}}, \quad (3.33)$$

onde $\hat{q}$ é a personalidade do jogador e α_i o seu nível de cooperação. A função subjetividade está baseada nos moldes do estudo realizado sobre o ganho esperado e a busca ou aversão ao risco em transações econômicas [15].

Originalmente, a função utilizada para a subjetividade é uma função prospecto, ou seja, ela traz qual a sua esperança de probabilidade de ganhar uma determinada quantia, dada a real probabilidade de ganhar. O jogador utiliza α_i e $\hat{q}$ para “avaliar” seu ganho antes de compará-lo ao ganho dos oponentes. Na Fig. 3.5 apresentamos gráficos da função $\Pi(\alpha_i; \hat{q})$ para alguns valores de $\hat{q}$.

Figura 3.5 – Formas da função subjetividade (Eq. 3.33) utilizada pelos jogadores para avaliar o seu ganho no momento da comparação com o ganho dos oponentes na estratégia subjetiva.

Observe que, para $\hat{q} = 1$, a função subjetividade é linear e reflete um ganho proporcional ao nível de cooperação, algo esperado para jogadores puramente racionais. Inserimos a personalidade nos jogadores para caracterizá-los como tomadores de risco ($0 < \hat{q} < 1$) ou avessos ao risco ($\hat{q} > 1$). Chamamos de avessos ao risco os jogadores que consideram seu ganho menor do que ele realmente é (subestimado) - $GS_i < G_i$, enquanto o jogador apresenta $\alpha_i < 1/2$, e considerando seu ganho maior do que o real (sobreestimado) - $GS_i > G_i$, quando o jogador apresenta $\alpha_i > 1/2$. E o inverso ocorre para o jogador tomador de risco, ele sobrestima seu ganho enquanto $\alpha_i < 1/2$ e subestima para $\alpha_i > 1/2$. Na Fig. 3.6 apresentamos o como um jogador percebe o seu ganho, dada a sua personalidade para diferentes valores de personalidade. Da maneira como foi considerada a função de subjetividade, um desertor total ($\alpha_i = 0$) sempre verá seu ganho subjetivo como nulo, $\Pi(0; \hat{q}) = 0$. Já um cooperador total ($\alpha_i = 1$) faz uma avaliação real do seu ganho, sem distorções, pois $\Pi(1; \hat{q}) = 1$.

Apresentadas todas as características que norteiam nosso problema, partimos para as resoluções analíticas, que se utilizam das premissas características do dilema

Figura 3.6 – Ganho do jogador A para diferentes personalidades.

do prisioneiro.

3.2 Modelo analítico

No âmbito da modelagem analítica, temos a vantagem de partir das premissas e observar suas implicações. Tais premissas podem ser experiências sensíveis (empíricas), ou serem oriundas de primeiros princípios universais. No entanto, por vezes, a escolha das premissas adequadas pode se mostrar uma tarefa tão árdua quanto encontrá-las empiricamente em um dado problema.

Nesta seção, abordamos o tratamento analítico do ganho esperado, da tentação crítica e da evolução temporal⁵ do nível médio de cooperação. Consideramos um sistema com N jogadores baseado no modelo apresentado na Seção 3.1. A proporção de cooperadores no sistema é ρ . Definido o sistema, vamos ao cálculo da tentação crítica para as estratégias darwiniana e pavloviana.

⁵Para as grandezas que variam no tempo, apresentaremos o índice t em sua primeira ocorrência e, em seguida, passaremos a omitir este índice, para simplificar a notação se não houver perda de informação. O mesmo será feito para a dependência com a configuração do vizinhos (κ) e demais dependências.

3.2.1 Tentação crítica (T^*)

Tentação crítica (T^*) é um valor de tentação na qual ocorre uma *transição de fase* no sistema. Transição de fase em termodinâmica é uma transformação em um sistema termodinâmico de uma fase (estado da matéria) para outra. A transição é caracterizada pela mudança abrupta de uma ou mais propriedades físicas do sistema, por exemplo, parâmetro de ordem e calor específico [50]. Em sistemas com interações que obedecem às regras do dilema do prisioneiro, dizemos que ocorre uma transição de fase quando o nível médio de cooperação estacionário se altera abruptamente.

Quando definimos um determinado sistema e variamos somente a tentação, ela é o parâmetro de controle que será responsável pelas transições de fase. A dependência das transições de fase com a tentação é devida ao fato dos jogadores utilizarem o valor de seus ganhos como parâmetro no processo de tomada de decisão quanto a mudança de estado. Assim, os ganhos obtidos, quando o sistema está sujeito à tentação crítica, serão valores limítrofes dentro da estratégia e farão com que os jogadores decidam mudar seus estados.

No caso da estratégia darwiniana, uma tentação crítica gera ganhos de cooperadores e desertores iguais, sob determinadas configurações de vizinhança. Quando isto acontece, um jogador ao comparar seu ganho com o do oponente não sabe definir quem teve o melhor desempenho e mantém o estado atual, mesmo se o oponente tiver um estado diferente do seu. Entretanto, o jogador irá manter seu estado atual, sem estar seguro de que esta ação é realmente a mais vantajosa.

Para a estratégia darwiniana, a tentação crítica ocorre quando o ganho do desertor é igual ao ganho do cooperador, $G^{\kappa_i}(\alpha_i = 0) = G^{\kappa_j}(\alpha_j = 1)$. Para definirmos os valores de tentação crítica T^* , consideramos os jogadores $i = D$ (desertor) e $j = C$ (cooperador) interagindo com suas respectivas vizinhanças κ_D e κ_C de tamanho k . Consideramos um desertor e um cooperador, pois na estratégia darwiniana para que um jogador mude de estado é necessário que haja interação com um oponente de estado diferente do seu.

Na estratégia pavloviana, a tentação crítica ocorre quando o ganho obtido é igual ao nível de aspiração, φ . Quando o jogador obtém um ganho menor ou igual à sua aspiração ele decide mudar seu estado, na esperança de obter resultados que sejam satisfatórios comparados à sua aspiração φ . Logo, $G^{\kappa_C} = \varphi$, $G^{\kappa_D} = \varphi$ e,

portanto, $G^{\kappa_C} = G^{\kappa_D}$.

Agora, apresentamos o cálculo de tentação crítica, T^* , para as estratégias darwiniana e pavloviana.

3.2.1.1 T^* - estratégia darwiniana

Para a estratégia darwiniana, igualando o ganho de desertores e cooperadores $G^{\kappa_D} = G^{\kappa_C}$ e isolando T_{ED}^* :

$$\begin{aligned} (T - P)c_D + Pk &= (R - S)c_C + Sk, \\ T_{ED}^* &= P + \frac{(R - S)c_C + (S - P)k}{c_D} \end{aligned} \quad (3.34)$$

para $0 \leq c_C \leq k$ e $0 < c_D \leq k$. Por exemplo, os valores de T_{ED}^* para $k = 6$, $R = 1$ e $P = S = 0$ são $T_{ED}^* = \{1; 6/5; 5/4; 4/3; 3/2; 5/3; 2; 5/2; 3; 4; 5; 6\}$.

A Eq. 3.34, que deduzimos para a tentação crítica, é válida para qualquer matriz de ganhos que seja utilizada no dilema do prisioneiro. Durán e Mulet, na Ref. [54], trazem uma função para o cálculo da tentação crítica, $T^*(n, m) = (z - n)/(z - n - m)$, que é válida somente para o dilema do prisioneiro fraco, e as variáveis mudas utilizadas, não estão claramente relacionadas com os elementos do problema.

Se o sistema for homogeneamente distribuído, podemos dizer que $\kappa_D = \kappa_C$, logo $c_D = c_C = c$ e

$$T_{EDH}^{**} = R + (S - P) \left(\frac{k}{c} - 1 \right) \quad (3.35)$$

para $0 < c \leq k$. Nesse caso, com $R = 1$ e $P = S = 0$, para qualquer valor de k , temos $T_{EDH}^{**} = \{1\}$.

Observando a Eq. 3.34, com o número de cooperadores na vizinhança, levanta-se a questão: “Quais os valores de tentação crítica no caso contínuo?” Para resolver essa questão, consideramos, ao invés do número de cooperadores na vizinhança, o nível médio de cooperação, que pode ser interpretado como o número de cooperadores em uma dada célula, caso ela represente um grupo. Assim, a tentação crítica assume todos os valores contínuos menores ou iguais a R , $T^* \leq R : T^* \in \mathbb{R}$. Portanto, pelas condições do DP, a única tentação crítica é $T = R$, e não se observa as mudanças abruptas existentes no DPD, ou seja, não há transição de fase.

A seguir, calculamos o valor de tentação crítica para a estratégia pavloviana.

3.2.1.2 T^* - estratégia pavloviana

Adotando a estratégia pavloviana, um jogador compara seu ganho à sua aspiração e, caso seu ganho seja menor ou igual à aspiração, muda seu estado. Portanto, o ganho de um cooperador ou desertor igual ao nível de aspiração é um valor limite, a partir do qual o jogador muda de estado. Tomando a relação $G^{\kappa_D} = G^{\kappa_C}$ e isolando T^* temos:

$$\begin{aligned} (T - P)c + Pk &= (R - S)c + Sk, \\ T_{EP}^* &= (R + P - S)c + (S - P)k \end{aligned} \quad (3.36)$$

para $0 \leq c \leq k$.

Não é possível calcular a tentação crítica para as estratégias baseadas no ganho e na personalidade, pois a tomada de decisão dos jogadores depende do ganho obtido em cada rodada, o qual está condicionado ao valor de T no sistema. A tentação crítica se torna uma função da tentação, $T^*(T)$, e não faz sentido isolar o valor de T para achar seu valor crítico, pois para cada valor de T , teríamos um conjunto diferente de valores de T^* .

3.2.2 Evolução temporal

Considere um sistema onde N jogadores jogam o dilema do prisioneiro discreto, dos quais c_t são cooperadores, no instante t . A evolução temporal do número de cooperadores é:

$$\begin{aligned} c_{t+1} &= c_t - c_t\gamma + (N - c_t)\lambda, \\ c_{t+1} &= c_t(1 - \gamma - \lambda) + N\lambda, \end{aligned} \quad (3.37)$$

onde γ é a probabilidade de um cooperador mudar seu estado para desertor e λ a probabilidade de um desertor mudar seu estado para cooperador.

Se considerarmos o nível médio de cooperação ao invés do número de cooperadores, reescrevemos a Eq. 3.37 a dividindo por N :

$$\rho_{t+1} = \rho_t(1 - \gamma - \lambda) + \lambda. \quad (3.38)$$

As probabilidades de mudança de estado γ e λ dependem da estratégia adotada. Se for a darwiniana, o jogador irá comparar o seu ganho com os ganhos de

sua vizinhança κ . Já na pavloviana, não há a comparação relativa entre os ganhos, mas sim uma comparação do ganho obtido pelo jogador e seu nível de aspiração φ . Na seção seguinte, explicamos a dinâmica de mudança de estado no sistema.

Mudança de estado na estratégia darwiniana

Considere um jogador totalmente cooperador, $\alpha_C = 1$, e sua vizinhança κ_C . Ele joga o dilema do prisioneiro com seus k vizinhos e obtém o ganho G^{κ_C} . Para que este jogador mude seu estado, de cooperador para desertor, é necessário que:

1. em sua vizinhança κ , haja pelo menos um desertor ($\rho_t < 1$);
2. um desertor tenha o maior ganho da vizinhança κ , isto é,

$$G_{\alpha_{vencedor}=0} = \max(G^{\kappa}). \quad (3.39)$$

3. haja em $\kappa_{vencedor}$, pelo menos c cooperadores, que sejam suficientes para satisfazer a condição anterior.

O mesmo vale na mudança de desertor para cooperador, observando-se as devidas inversões. Da primeira premissa podemos notar que $\rho = 0$ ou $\rho = 1$ são bacias de atração ou estados absorventes. Isto é, o sistema atinge o estado estacionário e não sofre mais modificações no decorrer do tempo. Nas seções seguintes apresentamos alguns cenários e resolvemos analiticamente o nível médio de cooperação do sistema.

Cenário I: ganho esperado

Vamos considerar o caso mais simples: no instante $t = 0$ cooperadores e desertores obtêm o ganho esperado dado pelas equações (3.13) e (3.14), respectivamente. Se $0 < \rho < 1$, ou seja, o sistema não está nos estados absorventes ($\rho = 0$ ou $\rho = 1$), os jogadores podem mudar seus estados e jogadores cooperadores passam a ser desertores se $E[g(\alpha_i = 0; \rho)] > E[g(\alpha_i = 1; \rho)]$ e vice-versa, caso contrário.

Como vemos na superfície de ganhos da Fig. 3.4, o ganho esperado de um desertor é sempre maior que o de um cooperador ao jogarem em um sistema com nível médio de cooperação ρ . Assim, $E[g(\alpha_i = 0; \rho)] > E[g(\alpha_i = 1; \rho)]$ o que implica que a probabilidade de um cooperador passar a ser desertor é unitária, $\gamma = 1$, e que a probabilidade de um desertor passar a ser cooperador é nula, $\lambda = 0$. Da Eq. 3.38,

com $\gamma = 1$ e $\lambda = 0$, temos $\rho_{t+1} = 0$. Logo, o sistema vai para a deserção, que é um estado absorvente, em uma única rodada e não sofre mais alterações.

Portanto, supor que todos os jogadores obtém exatamente o ganho esperado não é verdade, pois sabemos que a extinção da cooperação não é regra. Isso nos leva a acreditar que os jogadores obtém ganhos diferentes do ganho esperado, o que é decorrente de uma distribuição não homogênea dos jogadores. Por isso, propomos um segundo cenário, no qual utilizamos a probabilidade de ocorrência dos possíveis ganhos dos jogadores.

Cenário II: probabilidades de ganhos

Consideremos os ganhos de um jogador, que joga o dilema do prisioneiro discreto com $k = 4$ vizinhos, e a probabilidade de ocorrência de tais ganhos para um sistema com nível médio de cooperação ρ . Para $k = 4$, um jogador pode ter entre zero e quatro cooperadores na vizinhança e, uma vez conhecido o número de cooperadores, os demais são desertores. Vamos supor que uma vizinhança κ , corresponde a uma amostra representativa da população. Logo, consideramos que a probabilidade de um jogador escolhido ao acaso ser cooperador é igual a ρ .

Para sabermos qual a probabilidade do jogador receber um determinado ganho, precisamos responder à pergunta: “Qual a probabilidade de que em uma vizinhança de k jogadores, c sejam cooperadores?”. Da teoria de probabilidades e estatística, temos a distribuição de probabilidade discreta hipergeométrica. Esta distribuição descreve a probabilidade de retirar x elementos do tipo A , numa sequência de n extrações, sem reposição, de uma população finita de tamanho N , com K elementos do tipo A e $N - K$ elementos do tipo B . A distribuição hipergeométrica é:

$$P(X = x|N, K, n) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}}, \quad (3.40)$$

em que $\binom{y}{z} = \frac{y!}{z!(y-z)!}$.

Em nosso problema, os x elementos do tipo A são os c vizinhos cooperadores, as n extrações são os k vizinhos e a população, finita, tem tamanho N , com $N\rho$ cooperadores e $N(1 - \rho)$ desertores. Portanto, a probabilidade $P(c|N, N\rho, k)$ de que

haja c cooperadores, em uma vizinhança de tamanho k é:

$$P(c|N, N\rho, k) = \frac{\binom{N\rho}{c} \binom{N(1-\rho)}{k-c}}{\binom{N}{k}}. \quad (3.41)$$

Na Tab. 3.1, temos os ganhos dos jogadores para cada configuração e a probabilidade de ocorrência de cada configuração para $N = 1000$, $\rho = 1/2$, $k = 4$, $T = 1, 5$, $R = 1$ e $P = S = 0$.

	0C, 4D	1C, 3D	2C, 2D	3C, 1D	4C, 0D
G^{κ_D}	4P	T+3P	2T+2P	3T+P	4T
G^{κ_D}	0	1,5	3	4,5	6
G^{κ_C}	4S	R+3S	2R+2S	3R+S	4R
G^{κ_C}	0	1	2	3	4
$P(c 1000, 500, 4)$	0,0621	0,2500	0,3758	0,2500	0,0621

Tabela 3.1 – Ganhos de um jogador $G_i^{\kappa_i}$ para o estado desertor ($i = D$, $\alpha_D = 0$) e cooperador ($i = C$, $\alpha_C = 1$), de acordo com sua vizinhança, em um sistema que utiliza estratégia darwiniana. Os parâmetros utilizados são: $N = 1000$, $\rho = 1/2$, $k = 4$, $T = 1, 5$, $R = 1$ e $P = S = 0$.

Buscamos encontrar a situação em que o ganho do cooperador é maior que o do desertor $G^{\kappa_C} > G^{\kappa_D}$. O ganho de um jogador, que joga com k vizinhos, é dado pela Eq. 3.9 se ele for desertor e Eq. 3.10 se cooperador. Isolando o termo c em $G^{\kappa_C} > G^{\kappa_D}$ teríamos qual a condição de vizinhança que nos forneceria o ganho do cooperador maior que o do desertor.

$$\begin{aligned}
 G^{\kappa_C} &> G^{\kappa_D} \\
 (R - S)c + Sk &> (T - P)c + Pk \\
 c &> \frac{(P - S)k}{R + P - T - S}
 \end{aligned} \quad (3.42)$$

Como $0 \leq c \leq k$, temos que satisfazer $0 \leq (P - S)/(R + P - (T + S)) \leq 1$. Essa desigualdade só apresenta soluções para $T \leq R$, no entanto, o dilema do prisioneiro tem como condição $T > R$, para que haja a emergência de conflito. Portanto, se o sistema for homogêneo, não há solução válida para essa desigualdade e não é possível um cooperador ganhar mais que um desertor. Assim como mostramos no

cenário baseado no ganho esperado, confirmamos novamente que se o sistema for homogêneo não há emergência de cooperação para $T > R$. No dilema do prisioneiro fraco, admite-se $T = R$ que é um caso particular onde a desigualdade é satisfeita.

Mas, ao observamos a Tab. 3.1, vemos que se nos atermos a algumas combinações de resultados, seria possível que um cooperador obtivesse um ganho maior que um desertor. Vejamos, para $k = 4$, $T = 1, 5$, $R = 1$ e $P = S = 0$, em quais condições isso seria possível. Para isso vamos usar uma tabela verdade que resolva as desigualdades $G^{\kappa_C} > G^{\kappa_D}$ e $G^{\kappa_D} > G^{\kappa_C}$, ou seja, valor um quando a condição é satisfeita e zero caso contrário. Para simplificar o problema, vamos desconsiderar o fato de que, para um cooperador passar a ser desertor adotando a estratégia darwiniana é necessário que ele jogue contra um desertor. Levando em consideração essa condição, deveríamos considerar $k - 1$ interações mais uma interação com um desertor. O mesmo é válido para um desertor passar a desertor, observando-se as devidas inversões.

Tabela 3.2 – Tabela verdade para as desigualdades $G^{\kappa_D} > G^{\kappa_C}$ e $G^{\kappa_C} > G^{\kappa_D}$.

κ	0C, 4D	1C, 3D	2C, 2D	3C, 1D	4C, 0D
G^{κ_C}	$4S = 0$	$R + 3S = 1$	$2R + 2S = 2$	$3R + S = 3$	$4R = 4$
$G^{\kappa_D} = 4P = 0$	0	0	0	0	0
$G^{\kappa_D} = T + 3P = 1, 5$	1	1	0	0	0
$G^{\kappa_D} = 2T + 2P = 3$	1	1	1	0	0
$G^{\kappa_D} = 3T + P = 4, 5$	1	1	1	1	1
$G^{\kappa_D} = 4T = 6$	1	1	1	1	1
G^{κ_D}	$4P = 0$	$T + 3P = 1, 5$	$2T + 2P = 3$	$3T + P = 4, 5$	$4T = 6$
$G^{\kappa_C} = 4S = 0$	0	0	0	0	0
$G^{\kappa_C} = R + 3S = 1$	1	0	0	0	0
$G^{\kappa_C} = 2R + 2S = 2$	1	1	0	0	0
$G^{\kappa_C} = 3R + S = 3$	1	1	0	0	0
$G^{\kappa_C} = 4R = 4$	1	1	1	0	0

Da Tab. 3.2, vemos que as possibilidades de um cooperador passar a desertor, são dadas pelas seguintes combinações:

1. cooperador ganhar $4R = 4$ e o desertor ganhar $2T + 2P = 3$; ou $T + 3P = 1, 5$; ou $4P = 0$; ou
2. cooperador ganhar $3R + S = 3$ e o desertor ganhar $T + 3P = 1, 5$; ou $4P = 0$; ou
3. cooperador ganhar $2R + 2S = 2$ e o desertor ganhar $T + 3P = 1, 5$; ou $4P = 0$; ou
4. cooperador ganhar $R + 3S = 1$ e o desertor ganhar $4P = 0$.

Usando os valores de probabilidade de ocorrência de cada vizinhança, a probabilidade de um cooperador passar a ser desertor seria $\lambda = 0,0621(0,3758 + 0,2500 + 0,0621) + 0,2500(0,2500 + 0,0621) + 0,3758(0,2500 + 0,0621) + 0,2500(0,0621) = 0,2536$. Se adotarmos o mesmo procedimento para o cálculo da probabilidade de um desertor passar a ser cooperador, teríamos $\gamma = 0,0621(0,0621 + 0,2500 + 0,3758 + 0,2500 + 0,0621) + 0,2500(0,0621 + 0,2500 + 0,3758 + 0,2500 + 0,0621) + 0,3758(0,3758 + 0,2500 + 0,0621) + 0,2500(0,2500 + 0,0621) = 0,6486$. Como $\gamma + \lambda \neq 1$ existe a probabilidade de que os jogadores não mudem de estado. Isso é consequência da situação: $G^{\kappa C} = G^{\kappa D}$, por exemplo, para as configurações $G(\alpha_i = 0; \kappa = \{0C; 4D\}) = G(\alpha_i = 1; \kappa = \{0C; 4D\})$ e $G(\alpha_i = 0; \kappa = \{2C; 2D\}) = G(\alpha_i = 1; \kappa = \{3C; 1D\})$.

Pela indução lógica mostrada acima, fica claro que, para se obter todas essas configurações, devemos considerar um sistema não homogêneo, da mesma forma como fizemos no cálculo da tentativa crítica. Logo, para $G^{\kappa D} > G^{\kappa C}$, teríamos que satisfazer à condição:

$$c_D > \frac{(S - R)c_C + (S - P)k}{P - T} \quad (3.43)$$

e para $G^{\kappa C} > G^{\kappa D}$:

$$c_C > \frac{(T - P)c_D + (P - S)k}{R - S}. \quad (3.44)$$

Para obtermos como resultado a Tab. 3.2, podemos utilizar a função de Heaviside⁶ ou função degrau para resolver as desigualdades das eqs. 3.43 e 3.44. Logo,

6

$$H(x) = \begin{cases} 0, & \text{se } x < 0 \text{ e} \\ 1, & \text{caso contrário.} \end{cases} \quad (3.45)$$

é possível calcular γ e λ e temos uma função analítica para o nível médio de cooperação em função do tempo. A probabilidade de um cooperador mudar seu estado para desertor, γ , é o somatório de zero até k , das probabilidades de ocorrência de c_D cooperadores, ponderada pelo somatório de zero até k , das probabilidades de ocorrência das configurações de vizinhança c_C , que satisfazem a desigualdade da Eq. 3.43. Logo, γ é dado por:

$$\gamma = \sum_{c_D=0}^k \left\{ P(c_D|N, N\rho, k) \left[\sum_{c_C=0}^k P(c_C|N, N\rho, k) H(G^{\kappa_D}(c_D) - G^{\kappa_C}(c_C)) \right] \right\}, \quad (3.46)$$

e λ por:

$$\lambda = \sum_{c_C=0}^k \left\{ P(c_C|N, N\rho, k) \left[\sum_{c_D=0}^k P(c_D|N, N\rho, k) H(G^{\kappa_C}(c_C) - G^{\kappa_D}(c_D)) \right] \right\}, \quad (3.47)$$

onde $H(x)$ é a função de Heaviside ou função degrau.

No entanto, quando consideramos o mecanismo de cópia que a estratégia darwiniana adota, percebemos que a função hipergeométrica, adotada para o cálculo da probabilidade de ocorrência de cada vizinhança, não se mantém válida após as primeiras rodadas. Pois, ao copiar os vizinhos, surge um fenômeno de agrupamento que tem aspecto fundamental na dinâmica do sistema. Este agrupamento dos jogadores invalida a premissa estatística de que cada amostra do sistema seria uma amostra representativa que conserva os momentos estatísticos da população.

Poderíamos buscar resolver analiticamente ρ_t para as estratégias: pavloviana, baseada na personalidade e no ganho global ou local, que utilizamos nesta tese. Mas, fica patente que enfrentaríamos o mesmo problema encontrado na estratégia darwiniana, ou seja, inferir qual a real distribuição dos jogadores para calcularmos a probabilidade de ocorrência de cada vizinhança. Então, finalizamos a parte analítica do trabalho, onde confirmamos a dificuldade de resolver analiticamente o problema o que nos levou a uma abordagem numérica do mesmo. Na próxima seção apresentamos o modelo computacional que foi utilizado nas simulações.

3.3 Modelo computacional

Nesta seção descrevemos as principais características do modelo computacional e do programa desenvolvido para as simulações deste trabalho. O programa utilizado nas simulações foi desenvolvido em linguagem *C++* e compilado com o compilador *g++* (*Linux*). As simulações foram realizadas nos computadores do Laboratório de Modelagem em Sistemas Complexos (LMSC), sendo:

- 3 com processador AMD® Phenom™ II X6 1100T (12Mb Cache, 3,30GHz) com 8Gb DDR3 de memória RAM;
- 2 com processador Intel® Core™2 Quad Processor Q9550 (12Mb Cache, 2,83GHz, 1333MHz FSB) com 4Gb DDR3 de memória RAM e
- 2 com processador Intel® Core™2 Duo Processor E8400 (6Mb Cache, 3,00GHz, 1333MHz FSB) com 4Gb DDR3 de memória RAM.

Todos os computadores possuem o sistema operacional *Linux*, distribuição *Ubuntu*. Para a análise gráfica foi utilizado o software *Matlab*®.

Para descrever o programa, construímos um fluxograma (Fig. 3.7), com a lógica do algoritmo e as etapas fundamentais de sua execução, e fluxogramas com detalhes de cada etapa. De maneira sucinta a execução do programa acontece da seguinte maneira:

1. [**inicializa parâmetros**]: define os parâmetros da simulação;
2. [**vizinhança**]: define quem são os vizinhos de cada jogador;
3. [**configurações**]: para M configurações distintas, repete:
 - (a) [**autômato**]: cria o autômato;
 - (b) [**rodadas**]: durante *EPC* rodadas, repete:
 - i. [**agrupamentos**]: encontra a distribuição dos agrupamentos dos jogadores e calcula os coeficientes de agrupamento do sistema;
 - ii. [**entropia**]: calcula a entropia em função da distribuição dos agrupamentos de jogadores;

- iii. **[interações]**: os jogadores jogam o dilema do prisioneiro com seus k vizinhos, recebem e somam seus ganhos;
 - iv. **[Gini]**: calcula o coeficiente de Gini do sistema (distribuição da riqueza);
 - v. **[estratégia]**: de acordo com a estratégia, define a mudança do nível de cooperação de cada jogador;
 - vi. **[atualiza estados]**: todos os jogadores, de forma síncrona, atualizam seus níveis de cooperação, segundo o que foi definido pela estratégia;
 - vii. **[cálculo das médias]**: atualiza os contadores necessários para o cálculo dos valores médios das grandezas de interesse;
4. **[salva resultados]**: salva os resultados em arquivos de texto.

Durante a inicialização dos parâmetros, definimos os seguintes parâmetros das simulações:

- $[N = 1000]$ - tamanho da rede;
- $[EPC = 1000]$ - número de rodadas;
- $[M = 1000]$ - número de configurações distintas do autômato;
- $[\rho_0 = 1/2]$ - nível médio de cooperação inicial;
- $[1 \leq T \leq 2, R = 1 \text{ e } P = S = 0]$ - matriz de ganhos do dilema do prisioneiro (fraco), onde T é tentação, R - recompensa, P - punição e S - tolo;
- $[2 \leq k \leq 27]$ - conectividade;
- **[vizinhança]** - tipo de vizinhança: regular, aleatória fixa ou aleatória dinâmica;
- **[distribuição]** - função densidade de probabilidade utilizada para a distribuição inicial do nível de cooperação dos jogadores: discreto, uniforme, crescente, decrescente, triangular central ou triangular de extremos;

Figura 3.7 – Fluxograma das principais etapas do programa utilizado nas simulações. Cada retângulo representa uma operação e cada losango uma decisão. Começamos inicializando os parâmetros da simulação e escolhendo qual será o jogo (*Inicializa parâmetros*). Em seguida, definimos os vizinhos de cada jogador (*Vizinhança*) e tem início as repetições para as diferentes configurações ($m \leq M$), onde são criados cada um dos M autômatos distintos (*Autômato*), que irão interagir durante as rodadas ($t \leq EPC$). Em cada rodada, os jogadores interagem, recebem e somam os ganhos destas interações (*Interações*). Calculamos o coeficiente de Gini (*Gini*) utilizando as riquezas dos jogadores. Medimos a distribuição dos grupos em função da configuração do autômato (*Agrupamentos*) e com isso encontramos o coeficiente de agrupamento (*Coeficiente de agrupamento*) e esta distribuição também é utilizada no cálculo da entropia do sistema (*Entropia*). Feito isso, determinamos o como serão feitas as mudanças dos níveis de cooperação dos jogadores e realizamos essa atualização (*Estratégia*). Todo esse processo é repetido por EPC rodadas e então uma nova configuração é gerada e reinicia-se a sequência. Isso é feito para M configurações diferentes e então salvamos os resultados obtidos.

- [estratégia] - estratégia adotada pelos jogadores: darwiniana, baseada no ganho local, baseada no ganho global ou baseada na personalidade.

Na próxima seção, descrevemos cada um dos processos do fluxograma principal, com o auxílio de outros fluxogramas. Estas partes estão separadas nas subseções com os pormenores dos processos.

3.3.1 Conectividade ou vizinhança

Como adotamos uma rede unidimensional, podemos variar a conectividade k para qualquer valor inteiro entre um e N . Em nossas simulações, foi utilizada a vizinhança regular, no entanto, testamos os efeitos da utilização de vizinhanças aleatórias: fixa e dinâmica. Na vizinhança regular, os vizinhos são distribuídos de forma simétrica ao redor de cada jogador e o jogador realiza a auto-interação, quando a conectividade é um valor ímpar.

Utilizamos uma rede unidimensional, pois ela apresenta as mesmas transições de fase das bidimensionais. Além disso, as condições periódicas de contorno são mais simples e permitem variar a conectividade livremente, enquanto que em uma rede quadrada normalmente se consideram as *vizinhanças de von Neumann e Moore*. A vizinhança de von Neumann ($k = 4$) são os quatro primeiros vizinhos, sendo composta pelos jogadores ao norte, sul, leste e oeste, como mostrado na Fig. 3.8a. A vizinhança de Moore ($k = 8$) é formada pela vizinhança de von Neumann, mais os segundos vizinhos, isto é, os jogadores das diagonais (nordeste, sudeste, sudoeste e noroeste), como na Fig. 3.8b. Na rede unidimensional, consideramos como sendo equivalentes da vizinhança de von Neumann a conectividade $k = 4$ e equivalente à vizinhança de Moore, $k = 8$, como mostra a Fig. 3.8.

Em nossas simulações, após definirmos os parâmetros da simulação, definimos os vizinhos de cada jogador, que se mantém os mesmos durante todas as *EPC* rodadas e as M configurações. O fluxograma da Fig. 3.9 mostra como são definidos os vizinhos de cada jogador. No caso da vizinhança aleatória dinâmica, os vizinhos se alteram a cada rodada e por isso são escolhidos no momento que o jogador vai fazer suas interações.

Para representar a vizinhança foi utilizado um vetor de tamanho Nk representado como `viz[Nk]`, que armazena os vizinhos de cada um dos jogadores

Figura 3.8 – Representação na rede quadrada das vizinhanças de: (a) von Neumann, $k = 4$ (sem auto-interação) e (b) Moore, $k = 8$ (sem auto-interação). E suas respectivas representações na rede unidimensional para: (c) $k = 4$ (sem auto-interação), e (d) $k = 8$ (sem auto-interação). Cores: preto: jogador central; cinza escuro: primeiros vizinhos; e cinza claro: segundos vizinhos.

segundo seus rótulos. Por exemplo, para $k = 4$, os vizinhos do jogador 42 são $viz[42] = \{40; 41; 43; 44\}$, que estão armazenados nas posições $nk + j$, sendo n o rótulo do jogador, k a conectividade do sistema e j uma variável auxiliar $1 < j \leq k$. Na sequência, apresentamos a construção dos autômatos.

3.3.2 Autômato celular

Consideramos um autômato celular com $N = 1000$ jogadores. O autômato não tem células vazias. Utilizamos a condição periódica de contorno (CPC), ligando as extremidades da rede. A rede pode ser representada por uma fita fechada em forma de anel. As condições periódicas de contorno são necessárias, para que os jogadores das extremidades da rede, tenham a mesma conectividade dos jogadores internos da rede.

Em um autômato, temos os estados de um jogador, de acordo com o conceito de estados de um autômato celular. Em nosso modelo, cada jogador possui seu nível de cooperação e uma personalidade (utilizada apenas na estratégia baseada

Figura 3.9 – Fluxograma da definição dos vizinhos de cada jogador. A vizinhança de um jogador depende do tipo escolhido: regular, aleatória fixa ou aleatória dinâmica. Exceto pela aleatória dinâmica, a vizinhança de cada jogador se mantém a mesma durante todo o jogo. Para a vizinhança regular, observa-se se há ou não a auto-interação (*auto*), *k* ímpar ou par, respectivamente. Os vizinhos do jogador *n* serão os jogadores que vão de $n + (-k/2)$ até $n + (k/2)$, inclusive *n*, caso haja auto-interação, e respeitando as condições periódicas de contorno. Para a vizinhança aleatória fixa, são escolhidos, uma única vez, no início do jogo, *k* vizinhos aleatoriamente pertencentes ao sistema. A variável *aleat* é um número pseudoaleatório uniformemente distribuído entre zero e um. Na vizinhança aleatória dinâmica, a cada rodada, cada jogador escolhe *k* novos vizinhos, como veremos no fluxograma das interações (Fig. 3.12). O vetor *viz*[*N***k*] traz os vizinhos de todos os jogadores e é utilizado durante as interações (Fig. 3.12) e para comparar os ganhos dos jogadores com seus vizinhos (Fig. 3.13).

na personalidade). Para representar o autômato foram utilizados três vetores de tamanho N , um com o nível de cooperação dos jogadores ($\text{nc}[N]$), que é atualizado a cada rodada, de acordo com a tomada de decisão, um segundo, com a personalidade do jogador ($\text{q}[N]$), que se mantém constante durante todo o jogo e um terceiro, auxiliar ($\text{temp}[N]$) para a atualização do nível de cooperação dos jogadores. O rótulo dos jogadores varia de 1 a N e equivale à sua posição nos vetores do nível de cooperação, personalidade e auxiliar.

Na Fig. 3.10, temos o fluxograma com a definição do nível de cooperação dos jogadores de acordo com a função densidade de probabilidade escolhida, sendo possíveis as distribuições: discreta (DPD), uniforme, crescente, decrescente, triangular central e triangular de extremos. No fluxograma da Fig. 3.11, temos a formação do vetor com a personalidade $\text{q}[N]$ dos jogadores. Definimos a personalidade, de forma que, em média, metade dos jogadores tenha um comportamento avesso ao risco $\hat{q} \leq 1$ e a outra metade um comportamento afeito ao risco $\hat{q} > 1$.

Nível médio inicial de cooperação

O nível médio inicial de cooperação é $\rho(t = 0) = \rho_0 = 1/2$. Quando do DPD, no instante inicial da simulação são escolhidos aleatoriamente a partir de uma distribuição uniforme $N\rho_0$ jogadores, que são definidos como cooperadores e os demais como desertores. A configuração inicial do sistema é a única estocasticidade presente no modelo com vizinhança regular. Além do caso discreto, testamos o efeito do uso das distribuições uniforme, crescente, decrescente, triangular central e triangular de extremos no DPC.

Para obtermos estas distribuições, utilizamos o método da transformação inversa [76]. Por este método, podemos obter uma variável aleatória que siga uma distribuição probabilística desejada, com função de densidade de probabilidade $f(x)$ e função de distribuição acumulada igual a $F(x)$, a partir de uma variável aleatória com distribuição uniforme. Este método baseia-se no fato de que uma distribuição acumulada, $F(x)$, tem valor entre 0 e 1, que é o mesmo intervalo da variável aleatória com distribuição uniforme, X , produzida por um gerador de números pseudoaleatórios⁷. Desta forma, consideramos X como um valor de $F(x)$ e para encontrar o

⁷Um computador não é capaz de gerar números aleatórios propriamente ditos, pois os progra-

Figura 3.10 – Fluxograma da definição do nível de cooperação de cada jogador para cada configuração do automato. No caso discreto, são definidos exatamente $N\rho_0$ jogadores como cooperadores (coops). Primeiro, definimos todos os jogadores como desertores e sorteamos aleatoriamente os que serão cooperadores. Para isso, utilizamos um vetor auxiliar (eleg[N]) com os jogadores que são elegíveis (desertores) para se tornarem cooperadores, e vamos atualizando este vetor para que não sejam sorteados cooperadores. Para os casos contínuos, para cada jogador é definido um número aleatório, entre zero e um, seguindo a distribuição desejada: uniforme, crescente, decrescente, triangular central ou triangular de extremos, que é obtida a partir do método da transformação inversa de uma variável aleatória uniformemente distribuída entre zero e um (aleat).

Figura 3.11 – Fluxograma da definição da personalidade de cada jogador para cada configuração do automato $q[N]$.

valor de x , basta achar a função inversa de $F(x)$. Assim, se $y_0 = F(x_0)$, podemos escrever:

$$x_0 = F^{-1}(y_0) \quad (3.48)$$

fazendo $y_0 = X$:

$$x_0 = F^{-1}(X). \quad (3.49)$$

Logo, para as distribuições crescente, decrescente, triangular central e de extremos que desejamos temos as funções de transformação apresentadas à seguir.

Crescente

Primeiro, definimos a função densidade de probabilidade desejada e seu domínio. No caso da distribuição crescente:

$$f(x) = \begin{cases} 2x & \text{se } 0 \leq x \leq 1, \\ 0 & \text{caso contrário.} \end{cases} \quad (3.50)$$

Em seguida, calculamos a função de distribuição acumulada desta distribuição:

$$y = F(x) = \int_0^x dt(2t) = x^2, \text{ se } 0 \leq x \leq 1. \quad (3.51)$$

mas de computador são determinísticos e geram sempre o mesmo resultado quando são executados sob as mesmas condições (variáveis de entrada). No entanto, é possível gerar uma sequência de números com as mesmas propriedades estatísticas de uma sequência oriunda de um processo aleatório natural. Por isso, os números aleatórios gerados por computadores são ditos pseudoaleatórios.

E depois a inversa:

$$F^{-1}(x^2) = x = \sqrt{y}, \text{ se } 0 \leq y \leq 1. \quad (3.52)$$

que pode ser escrita como

$$x_i = \sqrt{X_i}, \quad (3.53)$$

onde X_i é uma variável aleatória uniformemente distribuída entre 0 e 1.

Decrescente

Para a distribuição decrescente:

$$f(x) = \begin{cases} -2x + 2 & \text{se } 0 \leq x \leq 1, \\ 0 & \text{caso contrário.} \end{cases} \quad (3.54)$$

Distribuição acumulada:

$$y = F(x) = \int_0^x dt(-2t + 2) = -x^2 + 2x, \text{ se } 0 \leq x \leq 1. \quad (3.55)$$

Função inversa:

$$F^{-1} = x = 1 \mp \sqrt{1 - y}, \text{ se } 0 \leq y \leq 1. \quad (3.56)$$

Triangular central

Para a distribuição triangular central:

$$f(x) = \begin{cases} 4x, & \text{se } 0 \leq x < 1/2, \\ -4x + 4, & \text{se } 1/2 < x \leq 1, \\ 0 & \text{caso contrário.} \end{cases} \quad (3.57)$$

Distribuição acumulada:

$$y = F(x) = \begin{cases} \int_0^x dt(4t) = 2x^2, & \text{se } 0 \leq x < 1/2, \\ 1/2 + \int_x^1 dt(-4t + 4) = -2x^2 + 4x + 5/2, & \text{se } 1/2 < x \leq 1. \end{cases} \quad (3.58)$$

Função inversa:

$$F^{-1} = x = \begin{cases} \sqrt{y/2}, & \text{se } 0 \leq y < 1/2, \\ 1 \pm (1/2)\sqrt{2y-1}, & \text{se } 1/2 < y \leq 1. \end{cases} \quad (3.59)$$

Triangular extremos

Função de distribuição triangular de extremos.

$$f(x) = \begin{cases} 2 - 4x, & \text{se } 0 \leq x < 1/2, \\ -2 + 4x, & \text{se } 1/2 < x \leq 1, \\ 0 & \text{caso contrário.} \end{cases} \quad (3.60)$$

Distribuição acumulada:

$$y = F(x) = \begin{cases} \int_0^x dt(2 - 4t) = 2x - 2x^2, & \text{se } 0 \leq x < 1/2, \\ 1/2 + \int_x^1 dt(-2 + 4t) = 1/2 + 2x - 2x^2, & \text{se } 1/2 < x \leq 1. \end{cases} \quad (3.61)$$

Função inversa:

$$F^{-1} = x = \begin{cases} 1/2 \pm (1/2)\sqrt{1-2y}, & \text{se } 0 \leq y < 1/2, \\ 1/2 \pm (1/2)\sqrt{2y-1}, & \text{se } 1/2 < y \leq 1. \end{cases} \quad (3.62)$$

Estas foram as funções utilizadas para gerar as distribuições desejadas.

3.3.3 Interações - dinâmica do sistema

Após termos criado o autômato, começam as interações entre os jogadores. A dinâmica do sistema consiste em cada jogador jogar o dilema do prisioneiro com seus vizinhos e somar os ganhos obtidos. A Fig. 3.12 contém o fluxograma que mostra como os jogadores interagem com seus vizinhos. O jogador começa com um ganho da rodada nulo e vai somando a esta variável o ganho de cada interação.

Como já havíamos dito, no caso da vizinhança aleatória dinâmica, a cada uma das k interações do jogador, ele escolhe aleatoriamente um jogador da rede para ser seu oponente.

Para calcular o ganho de uma interação, utilizamos a Eq. 3.2 com $1 \leq T \leq 2$, $R = 1$, $P = S = 0$ e o nível de cooperação do par de jogadores que estão interagindo. Depois de todos os jogadores terem interagido com suas respectivas vizinhanças, temos o vetor `rod[N]` com o ganho de todos os jogadores na rodada. Calculado o ganho da rodada, os jogadores somam o ganho da rodada ao seu ganho acumulado no jogo (`acum[N]`). Este ganho é a riqueza do jogador que será utilizada para calcular o coeficiente de Gini do sistema.

Agora que já sabemos os ganhos de todos os jogadores na rodada, vamos determinar em cada vizinhança qual foi o jogador com o maior ganho. Para isso, cada jogador compara o seu ganho com os ganhos de seus vizinhos, como é mostrado no fluxograma da Fig. 3.13, e obtemos um vetor `venc[N]` com os jogadores com o maior ganho em cada vizinhança. Caso o sistema esteja adotando a estratégia subjetiva (ES), ao invés do jogador utilizar seu ganho real na comparação, ele compara seu ganho subjetivo com os ganhos reais dos vizinhos.

O vetor `venc[N]` é utilizado pelas estratégias nas tomadas de decisão, como veremos na próxima seção.

3.3.4 Estratégias e atualização dos estados

Nesta etapa da simulação, definimos como serão feitas as alterações no nível de cooperação dos jogadores em função da estratégia adotada, como mostra o fluxograma da Fig. 3.14.

A estratégia pavloviana é válida apenas para o dilema do prisioneiro discreto. Nela, o jogador compara seu ganho com sua aspiração: se seu ganho for menor, ele muda seu estado de cooperador para desertor e *vice-versa* e mantém o estado atual, caso contrário.

A estratégia darwiniana serve tanto para o dilema do prisioneiro discreto quanto para o contínuo. Nela, os jogadores copiam o nível de cooperação do jogador com o maior ganho na vizinhança (`venc`).

As estratégias baseadas na personalidade e no ganho (global ou local) são

Figura 3.12 – Fluxograma das interações dos jogadores com seus vizinhos. Cada jogador joga com seus k vizinhos e recebe o ganho das interações e soma estes totalizando o ganho na rodada ($rod[N]$). Caso a vizinhança seja do tipo aleatória dinâmica, os jogadores escolhem seus vizinhos aleatoriamente no momento em que irão interagir com o j -ésimo vizinho. A final da rodada os jogadores somam o ganho da rodada ao ganho acumulado no jogo ($acum[N]$), o ganho acumulado é a riqueza do jogador que será utilizada para calcular o coeficiente de Gini do sistema (Fig. 3.15).

válidas somente para o dilema do prisioneiro contínuo. Na estratégia baseada na personalidade o jogador compara seu ganho subjetivo, um valor que depende de sua personalidade, com o ganho real dos oponentes e copia o estado do oponente com o maior ganho em sua vizinhança.

Na estratégia baseada no ganho, o jogador modifica seu estado atual na direção do estado do jogador com maior ganho, mas essa modificação depende da diferença entre seu estado e o estado deste vizinho e da diferença entre o ganho obtido e o ganho desejado (global ou local). O novo nível de cooperação gerado na

Figura 3.13 – Fluxograma que mostra como encontrar o jogador que tem o maior ganho na vizinhança. Para encontrar o jogador com o maior ganho basta encontrar o valor máximo do conjunto: ganho do jogador mais ganhos dos vizinhos. Caso os agentes do sistema adotem a estratégia baseada na personalidade (EBP), o jogador ao invés de utilizar seu ganho real na comparação, ele compara seu ganho subjetivo com os ganhos reais dos vizinhos. Estes jogadores estão agrupados no vetor **venc[N]** que é utilizado pelas estratégias (Fig. 3.14) na comparação de ganhos.

tomada de decisão é armazenado no vetor auxiliar (**temp[N]**) até o momento em que todos os jogadores atualizam seus estados de forma síncrona.

Antes de atualizarmos os estados dos jogadores, vamos medir o coeficiente de Gini do sistema, que depende somente da riqueza acumulada pelos jogadores, e definir o coeficiente de agrupamento e entropia do sistema.

3.3.5 Coeficiente de Gini

Uma medida amplamente utilizada para representar a extensão da desigualdade da distribuição de riqueza é o coeficiente de Gini [77, 78, 79, 57]. Ela é atribuída ao estatístico e sociólogo italiano Corrado Gini, pelo seu trabalho de 1912, intitulado

Figura 3.14 – Fluxograma da atualização do nível de cooperação dos jogadores de acordo com a estratégia adotada. A estratégia pavloviana é válida somente para o dilema do prisioneiro discreto. Nela, o jogador compara seu ganho com a aspiração e caso tenha um ganho maior que a aspiração mantém seu estado atual, caso contrário, muda de cooperador para desertor ou vice-versa. A estratégia darwiniana é válida tanto para o dilema do prisioneiro discreto quanto para o contínuo. As estratégias baseadas na personalidade e no ganho global e local são válidas para o dilema do prisioneiro contínuo. Tanto para a estratégia darwiniana quanto para a baseada na personalidade, o jogador simplesmente copia o nível de cooperação do jogador da vizinhança com o maior ganho ($venc[N]$). Para as estratégias baseadas no ganho, cada jogador observa a diferença entre o seu ganho e o ganho máximo global ou local, e determina como será sua mudança do nível de cooperação. O novo nível de cooperação é armazenado no vetor $temp[N]$. Para atualizar os estados todos os jogadores modificam seus estados de acordo com o que foi definido pelas estratégias copiando os valores de $temp[N]$.

“Variabilità e mutabilità” (Variabilidade e mutabilidade). É comumente utilizado no cálculo da desigualdade de distribuição de renda, mas pode ser usada para qualquer distribuição [80, 81].

Um estimador para o coeficiente de Gini é dado por [82]:

$$Gini(t) = \frac{2 \sum_{n=1}^N n y_n(t)}{N \sum_{n=1}^N y_n(t)} - \frac{N+1}{N} \quad (3.63)$$

onde y_n é a riqueza do n -ésimo jogador sendo y ordenado em ordem crescente e N é o tamanho da população. A axiomatização e outras formas de calcular este coeficiente são apresentadas nas Refs. [83, 84, 85]. O valor do coeficiente de Gini varia entre zero e um. $Gini = 0$ corresponde à completa igualdade, todos tem a mesma riqueza, e $Gini = 1$ corresponde à completa desigualdade, uma única pessoa possui toda a riqueza e as demais nada tem.

O fluxograma da Fig. 3.15 mostra como é calculado o coeficiente de Gini em nosso sistema. Note que calculamos seu valor para o ganho acumulado e também para o ganho por rodada que equivalem à distribuição de riqueza e a distribuição de “renda”, respectivamente.

3.3.6 Agrupamentos de jogadores

Para o cálculo do coeficiente de agrupamento e da entropia de agrupamento do sistema, é necessário que se conheça o tamanho dos agrupamentos no sistema e sua frequência de ocorrência. O coeficiente de agrupamento é o tamanho do maior agrupamento presente no sistema, dividido pelo tamanho do sistema.

No fluxograma da Fig. 3.16, apresentamos como é feita a contagem dos agrupamentos de cooperadores e desertores, que são armazenadas nos vetores `coop[N]` e `dsrt[N]`, respectivamente. Neste vetor, o índice n dá o tamanho do agrupamento e seu valor é a quantidade de agrupamentos daquele tamanho, um histograma dos tamanhos de agrupamento.

Para o cálculo do coeficiente é feita uma varredura na ordem decrescente do vetor com os tamanhos dos agrupamentos. O primeiro valor não nulo que for encontrado é o tamanho do maior agrupamento presente no sistema. Ao dividirmos o tamanho do maior agrupamento por N , temos o coeficiente de agrupamento, como mostra o fluxograma da Fig 3.17.

Figura 3.15 – Fluxograma do cálculo do coeficiente de Gini. Criamos um vetor com o ganho acumulado e outro do ganho na rodada dos jogadores que é ordenado em ordem crescente. Calculamos os somatórios da riqueza e se esse ganho for não nulo, calcula-se o coeficiente de Gini, para o ganho acumulado no decorrer das rodadas (*gini*) e da rodada atual (*gini_rod*).

Figura 3.16 – Fluxograma que faz o histograma do tamanho dos agrupamentos no sistema. Contamos quantos agrupamentos de cada tamanho há no sistema, considerando condições periódicas de contorno. Esses histogramas são os vetores (*coop*) e (*dsrt*) que serão utilizados no cálculo do coeficiente de agrupamento e das entropias.

Figura 3.17 – Fluxograma da determinação do coeficiente de agrupamento a partir do tamanho do maior agrupamento fazendo a varredura dos histogramas de cooperadores e desertores do fim para o início. Desta forma são determinados os coeficiente de agrupamento cooperador `coefcoop[epc]` e desertor do sistema `coefdsrt[epc]`.

Os vetores `coop[N]` e `dsrt[N]` que servem para o cálculo do coeficiente de agrupamento cooperador e desertor, respectivamente, são utilizados no cálculo da entropia de agrupamentos do sistema.

3.3.7 Entropia da informação

Nesta seção, apresentamos a base teórica sobre entropia de agrupamentos, entre elas as entropias⁸ de Shannon, Tsallis e Kullback-Leibler. Elas serão utilizadas como uma medida de complexidade da rede, baseada na estatística da organização dos agrupamentos de cooperadores e desertores.

A diversidade do tamanho de agrupamentos tem sido utilizada como uma

⁸Jaynes [86] faz um importante comentário sobre o uso do termo entropia, que devemos ter em mente: “A função H é chamada de *entropia*; ou melhor *entropia da informação* da distribuição $\{p_i\}$. Essa é uma infeliz terminologia que agora parece impossível de se corrigir. Devemos advertir, desde o início, que o principal mal ocupacional deste campo é a persistente incapacidade de distinguir entre a *entropia da informação*, que é uma propriedade de qualquer distribuição de probabilidade, e a *entropia experimental da termodinâmica*, que, por sua vez, é uma propriedade de um estado termodinâmico definido, tais como, por exemplo, as quantidades observadas de algum sistema físico: pressão, volume, temperatura e magnetização. Elas nunca deveriam ter sido chamados pelo mesmo nome, visto que, a entropia experimental não faz nenhuma referência a qualquer distribuição de probabilidade, e a entropia da informação não faz referência à termodinâmica. Muitos livros e artigos de pesquisa são fatalmente falhos por falha do autor em distinguir entre estas coisas completamente diferentes, e em consequência provando teoremas sem sentido.”

medida de complexidade em vários sistemas que geram uma distribuição estatística de agrupamentos [87]. Para medir essa complexidade é utilizada a entropia de agrupamentos, que pode ser definida como uma função da probabilidade de que um sítio ocupado pertença ao agrupamento s . Em um sistema, o número de agrupamentos s por sítios na rede ou o número de agrupamento normalizado n_s é definido como:

$$n_s = \frac{\text{número de agrupamentos de tamanho } s}{\text{número total de sítios na rede}}. \quad (3.64)$$

Portanto, a probabilidade de que um sítio ocupado, escolhido arbitrariamente, pertença a um agrupamento contendo s sítios é dado por:

$$w_s = \frac{sn_s}{\sum_s sn_s}. \quad (3.65)$$

Consequentemente, a entropia do sistema é definida como:

$$H = - \sum_s w_s \ln w_s. \quad (3.66)$$

Essa métrica de complexidade, havia sido apresentada anteriormente por Gould e Tobochnik [88], no estudo sobre sistemas percolantes, inclusive com a mesma notação.

Ao invés de utilizarmos a probabilidade w_s de um sítio pertencer à um agrupamento de tamanho s , utilizamos a probabilidade p_s de um agrupamento ter o tamanho s :

$$p_s = \frac{\text{número de agrupamentos de tamanho } s}{\text{número total de agrupamentos}}. \quad (3.67)$$

Consideramos duas classes de agrupamentos: cooperadores ou desertores. Quando não há distinção sobre a classe e utilizamos simplesmente o termo agrupamentos, estamos nos referindo a todos os agrupamentos existentes no sistema, sejam eles cooperadores ou desertores.

De acordo com as observações de Jaynes, calcularemos a entropia de Shannon e a de Kullback-Leibler⁹ que são entropias da informação. Além destas, utilizamos a entropia de Tsallis, que é uma generalização da entropia de Boltzmann-Gibbs, apesar das controvérsias a respeito de sua natureza termodinâmica. Temos a entropia como

⁹A medida representada pela entropia de Kullback-Leibler recebe vários nomes [89]: discriminação (do inglês, *discrimination*), discriminação de informação (do inglês, *discrimination information*), número de Kullback-Leibler e divergência dirigida (do inglês, *directed divergence*).

uma noção de informação intrínseca do sistema, ou seja, a informação fornecida por um processo aleatório sobre si mesmo. Neste caso, a aleatoriedade está na distribuição dos agrupamentos. Assim, iremos considerar a entropia de Tsallis como sendo uma entropia da informação, pois estamos interessados na organização dos agrupamentos e sua relação com a cooperação no sistema. Nas seções seguintes, mostramos a formulação destas entropias.

3.3.7.1 Entropia de Shannon

Shannon introduziu os conceitos de entropia e informação mútua no contexto da teoria da informação [90]. A teoria da informação pode ser entendida como a teoria matemática da teoria da comunicação, que lida com sistemas que transmitem informações ou dados de um lugar para outro. A entropia H é definida como uma medida de incerteza ou aleatoriedade de um fenômeno aleatório e é apresentada como [91, 92, 93]:

$$H(X) = -k \sum_{i=1}^L p(x_i) \log_2[p(x_i)], \quad (3.68)$$

sendo X uma variável aleatória com L componentes e cada um desses x_i componentes tem uma probabilidade $p(x_i)$ de ocorrência e k é uma constante positiva.

No DP, a variável aleatória que consideramos foi a frequência de ocorrência dos diferentes tamanhos de agrupamento. Como temos duas classes de agrupamentos teremos três resultados para a entropia de Shannon. Ao invés de H , adotaremos a nomenclatura Sh , quando considerarmos todos os agrupamentos do sistema, Sh_C , quando considerarmos somente os agrupamentos cooperadores e Sh_D , quando considerarmos somente os desertores.

3.3.7.2 Entropia de Tsallis

A entropia de Tsallis é uma generalização da entropia de Boltzmann-Gibbs, que foi apresentada pelo físico Constantino Tsallis [94] no ano de 1988. Tsallis utiliza uma grandeza q e postula a entropia como:

$$Sq(X)(q) = \frac{k}{(q-1)} \left[1 - \sum_{i=1}^L p(x_i)^q \right], \quad (3.69)$$

onde $q \in \mathbb{R}$ caracteriza a generalização, $\{p_i\}$ são as probabilidades associadas às configurações macroscópicas do sistema e k é uma constante positiva. A entropia de Boltzmann-Gibbs é recuperada no limite $q \rightarrow 1$.

Em consequência das classes dos agrupamentos, definimos a entropia de Tsallis para agrupamentos, $Sq(q)$, a entropia de Tsallis para agrupamentos cooperadores, $Sq_C(q)$, e a entropia de Tsallis para agrupamentos desertores $Sq_D(q)$.

Relembramos que, apesar da entropia de Tsallis estar relacionada às características termodinâmicas do sistema, iremos considerá-la como se ela fosse uma entropia da informação.

3.3.7.3 Entropia de Kullback-Leibler generalizada

Uma primeira ideia de generalização de entropia, chamada de discriminação, foi desenvolvida por Kullback. Atualmente, esta forma de medida da informação é mais chamada de entropia relativa, número de Kullback-Leibler e interpretada mais como uma medida de similaridade entre distribuições de probabilidade do que como uma medida de informação entre variáveis aleatórias [89].

A entropia relativa, ou divergência de Kullback-Leibler, entre duas distribuições de probabilidade $P(x)$ e $Q(x)$ é [95, 96]:

$$D_{KL}(P||Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}. \quad (3.70)$$

Observe que, em geral, a entropia relativa não é simétrica em relação à troca das distribuições P e Q , isto é, $D_{KL}(P||Q) \neq D_{KL}(Q||P)$. Apesar de, por vezes, a entropia relativa ser chamada de distância de Kullback-Leibler, ela não é propriamente uma distância. A grandeza D_{KL} pode divergir se $Q(x) \rightarrow 0$ e para evitar essa divergência, utiliza-se uma constante ξ na definição das probabilidades, onde $0 < \xi \leq 1$. Tratamos a probabilidade de ocorrência dos tamanhos de agrupamento, da seguinte forma [97]:

$$p_{i,s} = \frac{n_{i,s} + \xi}{\sum_s^L n_{i,s} + L\xi}, \quad (3.71)$$

onde $p_{i,s}$ (equivalente à P_s ou Q_s) é a probabilidade de um agrupamento da classe i ter o tamanho s , $n_{i,s}$ é a contagem de agrupamentos da classe i (cooperador ou desertor em nosso problema) de tamanho s e L é o tamanho máximo que um agrupamento pode atingir.

Podemos utilizar a entropia de Kullback-Leibler porque temos as duas classes de agrupamentos. Seguindo nossa notação (S) para a entropia teremos $SK_C = D_{KL}(C||D)$ e $SK_D = D_{KL}(D||C)$, onde C é a distribuição dos agrupamentos cooperadores e D a distribuição dos agrupamentos desertores.

A entropia de Kullback-Leibler pode ser generalizada utilizando a definição do q -logaritmo [98, 99] e temos [100]:

$$SK(q) = \frac{k}{(1-q)} \sum_{i=1}^L P_i^q \left(P_i^{(1-q)} - Q_i^{(1-q)} \right). \quad (3.72)$$

Para obter a Eq. 3.70 da função generalizada, devemos tomar o limite $q \rightarrow 0$ com $k = 1/\ln 2$.

A entropia SK pode divergir em função do valor de q . Com $q < 0$, $P_i^q \rightarrow \infty$ para $P_i = 0$. Se $q > 1$, $P_i^{(1-q)} \rightarrow \infty$ e $Q_i^{(1-q)} \rightarrow \infty$, quando $P_i = 0$ e $Q_i = 0$, respectivamente. Nestes casos utilizamos a probabilidade dada pela Eq. 3.71. Para o caso $0 < q \leq 1$, a entropia SK nunca diverge, assim definimos [100]:

$$p_{ij} = \frac{n_{ij}}{n_i}. \quad (3.73)$$

Nos casos limites, em que $\sum_{i=1}^L p(x_i)^q$ for igual a zero, ou seja, todo o sistema for cooperador ou desertor, consideramos a entropia como sendo nula, tanto para os cooperadores quanto para os desertores. Por exemplo, se tivermos um sistema completamente cooperador, teremos todos os cooperadores organizados em um agrupamento único e podemos pensar que os desertores estão organizados em grupo vazio.

No fluxograma da Fig. 3.18 mostramos o cálculo das entropias mencionadas acima, levando em consideração todas as condições que dizem respeito às divergências em relação a variável q e aos casos limites com o sistema totalmente cooperador ou desertor.

Figura 3.18 – Fluxograma entropias. (continua)

Figura 3.18 – Fluxograma entropias.

Resultados numéricos - dilema do prisioneiro: contínuo, vizinhança e estratégias

“...de um ponto de vista lógico, a negação de uma negação é uma afirmação: ‘A é igual a A’ é o mesmo que ‘A é não igual a não-A’.”

Nassim Nicholas Taleb no livro *A lógica do cisne negro* - 2007.

Neste capítulo, apresentamos os resultados das simulações numéricas. Exploramos os efeitos das seguintes características:

- função densidade de probabilidade utilizada para a distribuição do nível de cooperação inicial dos jogadores,
- vizinhança utilizada pelos jogadores e
- diferentes estratégias.

No dilema do prisioneiro (DP), estamos interessados em determinar quais são as condições necessárias para que ocorra a emergência de cooperação entre os jogadores. Para explorarmos o dilema do prisioneiro iterado contínuo (DPC) e discreto (DPD) utilizamos uma rede unidimensional.

O sistema possui $N = 1000$ jogadores. Em todos os sistemas utilizamos o nível médio de cooperação inicial $\rho_0 \cong 1/2$. No caso discreto, são escolhidos $\rho_0 N$ jogadores aleatoriamente, a partir de uma distribuição uniforme, e definidos como

cooperadores e os demais como desertores. No caso contínuo, cada jogador tem o seu nível de cooperação inicial definido no intervalo entre zero e um, a partir do sorteio de um valor de uma função de densidade de probabilidade com média $1/2$.

As personalidades dos jogadores são definidas da seguinte maneira: metade dos jogadores tem $0 \leq \hat{q} \leq 1$ e a outra metade $1 < \hat{q} < \infty$. Impusemos essa condição para que metade dos jogadores tivesse um comportamento avesso ao risco e a outra metade afim ao risco, lembramos que a aversão ou afinidade ao risco para o dilema do prisioneiro foi definida na Sec. 3.1.5.4.

Utilizamos a matriz de ganhos adotada por Nowak [34], portanto temos o *dilema do prisioneiro fraco*, com $1 \leq T \leq 2$, com passos $\Delta T = 0,01$, $R = 1$ e $P = S = 0$. Os jogadores fazem suas interações com seus k vizinhos, com $k = \{2; 3; 4; 5; 6; 7; 8; 9; 24; 25; 26; 27\}$, recebem os ganhos dessas interações, somam os ganhos das interações, decidem-se quanto à mudança de estado de acordo com a estratégia adotada e realizam a mudança de estado. Essa sequência se repete sucessivamente por $EPC = 1000$ rodadas. Assim, o sistema passa por um regime transiente e então atinge o regime estacionário¹ que pode ser estável ou oscilatório.

Considerando sistemas com as propriedades apresentadas acima, variamos cada uma das características de interesse, para estudar se elas favorecem ou dificultam a emergência da cooperação. Testamos os efeitos causados pela variação das seguintes características² do sistema:

- i. função densidade de probabilidade adotada para a distribuição do nível de cooperação inicial dos jogadores (no caso contínuo);
- ii. tipo de vizinhança dos jogadores e
- iii. estratégias dos jogadores.

Para o estudo das funções de densidade de probabilidade da distribuição inicial do nível de cooperação dos jogadores, foram utilizadas as seguintes distribuições:

1. uniforme;

¹A escolha $EPC = 1000$ foi baseada em observações do regime transiente do dilema do prisioneiro discreto, onde, normalmente, os sistemas atingem o regime estacionário antes da rodada 300.

²Os detalhes destas características estão explicados na primeira seção do Cap. 3 que traz o modelo do problema.

2. triangular central;
3. triangular de extremos;
4. crescente e
5. decrescente.

Para verificar se a estrutura de conexões dos jogadores a seus vizinhos tem alguma importância na cooperação, foram consideradas as vizinhanças:

1. regular;
2. aleatória fixa e
3. aleatória dinâmica (mobilidade).

Além da estratégia darwiniana, que será utilizada como referência, exploramos as possibilidades advindas do nível de cooperação contínuo, propondo novas estratégias. Assim adotamos as estratégias:

1. darwiniana;
2. baseada no ganho global;
3. baseada no ganho local e
4. subjetiva.

A análise do transiente do sistema nos permite compreender a dinâmica de evolução do sistema e a emergência da cooperação. A análise do regime estacionário serve para estudar o efeito de cada característica em relação ao nível de cooperação estacionário resultante.

O presente capítulo possui três seções. Na primeira, apresentamos alguns casos ilustrativos do regime transiente. Nestes exemplos, aparecem fenômenos que explicam a emergência e manutenção da cooperação. Se não surgir nenhum destes fenômenos ocorre a extinção da cooperação. Na segunda seção, analisamos o regime estacionário dos sistemas em função da densidade de probabilidade da distribuição inicial, vizinhança e estratégias. Na terceira, mostramos os resultados dos classificadores que propusemos para analisar se o sistema é cooperador ou não.

4.1 Transiente - dilema do prisioneiro contínuo

Para analisar o regime transiente, apresentamos na Fig. 4.1 imagens da evolução temporal de alguns sistemas selecionados. Em todos os sistemas apresentados foi empregado o dilema do prisioneiro iterado contínuo com nível médio de cooperação inicial $\rho_0 \sim 1/2$. A configuração inicial do nível de cooperação dos jogadores é a mesma em todas os sistemas. O que os diferencia é o conjunto de parâmetros utilizado. Os parâmetros são: tentação T , conectividade k e estratégia adotada pelos jogadores. Com os conjuntos de parâmetros escolhidos, os sistemas produzem regimes transientes que exemplificam comportamentos de interesse para a compreensão da emergência e manutenção da cooperação.

É possível representar a evolução temporal em uma única imagem estática, pois foi utilizada uma rede unidimensional. No caso bidimensional, seria necessário um vídeo para apresentar a evolução temporal. Nas imagens de evolução temporal dos sistemas unidimensionais, temos no eixo das abscissas o tempo, que transcorre da esquerda para a direita, e no eixo das ordenadas os rótulos dos jogadores. As cores representam o nível de cooperação dos jogadores. A escala de cores, em tons de cinza, vai do branco (cooperador) ao preto (desertor). Então, se acompanharmos um jogador da esquerda para a direita, temos toda sua dinâmica de mudanças do nível de cooperação no decorrer do tempo.

Lembramos que os gráficos da Fig. 4.1 foram escolhidos de forma a exemplificar os processos que podem gerar ou manter a cooperação no sistema. Eles não servem para um estudo detalhado no que diz respeito aos efeitos da tentação, conectividade e estratégia. Este estudo será realizado na seção seguinte, quando analisaremos o nível médio de cooperação estacionário, ρ_∞ .

A Fig. 4.2 traz uma ampliação dos instantes iniciais dos mesmos sistemas apresentados na Fig. 4.1 indo somente até a rodada 250 ao invés da rodada 1000. Esta ampliação serve para mostrar em detalhes o que acontece na região do transiente.

Apesar de apresentarmos somente sistemas nos quais o nível médio de cooperação estacionário é maior que $1/2$, isto é $\rho_\infty > 1/2$, sistemas com $0 \leq \rho_\infty \leq 1$ podem ocorrer, indo da deserção total $\rho_\infty = 0$ à cooperação total $\rho_\infty = 1$. Como

Figura 4.1 – Evolução temporal de sistemas com $N = 1000$ jogadores e nível médio inicial de cooperação $\rho_0 = 1/2$, que jogam o dilema do prisioneiro contínuo durante $EPC = 1000$ rodadas. O tempo evolui da esquerda para direita e no eixo vertical os rótulos dos jogadores. A escala de cores em tons de cinza vai do branco (cooperação) ao preto (deserção). Os gráficos abaixo das imagens de evolução temporal, mostram o nível médio de cooperação em função do tempo (rodadas). Todos esses sistemas possuem a mesma configuração inicial, alterando-se somente os parâmetros e estratégias adotados para cada situação.

Figura 4.2 – Instantes iniciais dos sistemas apresentados na Fig. 4.1 até a rodada $epc = 250$.

comportamento geral dos sistemas, temos dois casos no regime estacionário. Um é o que chamaremos de sistemas cooperativos, que apresentam um nível médio de cooperação estacionário acima de meio ($\rho_\infty > 1/2$) e o outro os sistemas desertores, com nível médio de cooperação estacionário igual ou menor à meio ($\rho_\infty \leq 1/2$).

Nos casos cooperativos, geralmente, o transiente começa com uma diminuição do nível médio de cooperação, $\rho(t)$, em relação ao nível médio de cooperação inicial, ρ_0 , seguida de uma recuperação da cooperação que reverte essa diminuição e faz aumentar o nível médio de cooperação até atingir um valor maior que meio e, por vezes, alcançam a cooperação total no regime estacionário. A diminuição do nível de cooperação logo no início da evolução temporal não é regra. Em alguns sistemas, o nível médio de cooperação pode simplesmente aumentar até atingir um patamar com maioria de cooperadores no regime estacionário.

Os sistemas desertores surgem quando a cooperação não consegue se reerguer após a “invasão” inicial dos desertores e o nível médio de cooperação estacionário fica abaixo de meio. A invasão do sistema pela deserção pode ser compreendida observando-se os ganhos obtidos pelos jogadores na primeira rodada do jogo e a dinâmica das estratégias adotadas. As estratégias privilegiam o maior ganho por diversos mecanismos, tais como copiar os estados dos jogadores que obtiveram maior ganho, ou diminuir o nível de cooperação para aumentar o ganho individual. Ao observar o ganho por interação de um desertor e um cooperador, apresentados no gráfico da Fig. 4.3, vemos que o ganho do desertor é sempre maior que o do cooperador ao confrontarem um mesmo oponente com o nível de cooperação α_j . Além disso, ao se confrontarem, os desertores obtêm um ganho maior que os cooperadores, então pode-se dizer que os desertores exploram a cooperação alheia.

Considerando que desertores e cooperadores começam o jogo homogeneamente distribuídos na rede, então ambos terão vizinhanças com $\rho_0 \cong 1/2$ e os desertores terão um ganho médio maior que os cooperadores. Se a estratégia utilizada possuir um mecanismo de cópia, os desertores serão copiados e teremos uma diminuição do nível médio de cooperação. Alternativamente, se o jogador simplesmente buscar aumentar seus ganhos sem observar o nível de cooperação e ganhos dos jogadores de sua vizinhança, sua ação será diminuir seu nível de cooperação. Por esses motivos a deserção começa a invadir o sistema.

Figura 4.3 – Ganho de um cooperador ($\alpha_i = 1$) e um desertor ($\alpha_i = 0$) ao jogar contra um oponente com nível de cooperação α_j com $T = 2$, $R = 1$ e $P = S = 0$ (dilema do prisioneiro fraco).

No entanto, quando a maioria dos jogadores adota ações pouco cooperativas, tem-se o efeito conhecido como tragédia dos comuns. Nesse caso, a deserção, que era vantajosa para os desertores enquanto eles estavam entre os cooperadores, deixa de trazer os mesmos benefícios, quando todos os jogadores se tornam desertores. Se alguns grupos de cooperadores, que se auxiliam ante a exploração dos desertores, resistirem à invasão inicial, eles podem se tornar os focos de resistência cooperadora dentre a maioria desertora, que podem, eventualmente, fazer prosperar a cooperação. Isso acontece quando cooperadores agrupados têm um ganho maior que os desertores exploradores e eles passam a copiar os cooperadores ou aumentar seu nível de cooperação na tentativa de melhorar seus ganhos. Mas esses efeitos ficam condicionados à um cenário onde existam agrupamentos de cooperadores, que sustentem a cooperação, dentro de uma maioria desertora.

Após termos visto em linhas gerais os fenômenos responsáveis pela emergência ou extinção da cooperação, vamos apresentar algumas estruturas que explicam esses efeitos. Nos sistemas das Figs. 4.1a, 4.1b, 4.1e, e 4.1f surgem *fingers* e *gliders*, que são padrões observados comumente em autômatos celulares e explicam as oscilações periódicas presentes no regime estacionário. *Fingers* são estruturas que se propagam no sistema em linha reta e possuem bordas lisas ou rugosas, como dentes-de-serra.

Gliders se propagam atravessando o sistema. Os *fingers* de borda rugosa e os *gliders* podem apresentar comportamento oscilatório fazendo com que o nível médio de cooperação oscile no tempo.

As estruturas de *fingers* e *gliders* são próprias de sistemas com dois estados como os autômatos de Wolfram. No entanto, as estruturas presentes nos sistemas com nível contínuo de cooperação apresentam a mesma forma e propriedades observadas nos sistemas binários, por isso adotamos a mesma nomenclatura. Isso também mostra que, o comportamento qualitativo de sistemas com o DPC é muito similar ao observado em sistemas com o DPD (vide a Ref. [46]).

Nos gráficos da Fig. 4.1 há uma rápida convergência para o regime estacionário. Os sistemas atingem o estado estacionário ou estão muito próximos do valor estacionário a partir da rodada 200. Existem focos de cooperadores que invadem o sistema levando à maioria cooperadora. O transiente tem dependência com a configuração inicial, apesar de não ser caótico.

Na Fig. 4.1a, tem-se um sistema que adota estratégia evolucionária darwiniana, conectividade $k = 2$ e tentação $T = 1,00$. A baixa conectividade do sistema leva à coexistência de vários agrupamentos com diferentes níveis de cooperação. Surgem agrupamentos mais cooperativos em relação à vizinhança a partir dos jogadores 438-439, 570, 618, 652, 734 e 938. Estes agrupamentos são focos de resistência e emergência da cooperação. A colisão que ocorre na rodada 359, entre o *finger* dos jogadores 748-749 e o *glider* oriundo no jogador 936, ilustra um *finger* que absorve um *glider*. Observe que o *glider* oriundo nos jogadores 438-439 vai absorvendo todos os *fingers* com os quais ele colide. Outra colisão importante neste sistema ocorrerá provavelmente após a rodada 1000 na colisão do *finger* dos jogadores 446 com o *glider* oriundo nos jogadores 438-439.

A Fig. 4.1b, apresenta um sistema que adota estratégia evolucionária darwiniana, conectividade $k = 6$ e tentação $T = 2,00$. O regime estacionário está bem definido a partir da rodada 125 e a última modificação no nível médio de cooperação de relevância ocorre na rodada 213, quando da colisão de dois *gliders*, que formam uma figura triangular. À partir deste momento, o sistema apresenta um comportamento oscilatório devido aos padrões dos *fingers* e *gliders* presentes no sistema e poucos agrupamentos de diferente nível médio de cooperação ainda

coexistem. Provavelmente, ocorrerá outra alteração do nível médio de cooperação depois da rodada 1000 devido à colisão do *glider* originado no jogador 443 com a fronteira originada pelos jogadores 434 e 440. Os principais focos de cooperação são os jogadores 342, 440 e 683.

Na Fig. 4.1c, mostramos um sistema que adota estratégia baseada no ganho global, conectividade $k = 6$ e tentação $T = 1,30$. O regime estacionário está bem estabelecido a partir da rodada 200. Ocorrem focos de resistência e emergência da cooperação nos jogadores 438-439 e 745. O *finger* desertor originado no jogador 736 é absorvido pelo *glider* oriundo do jogador 443.

A Fig. 4.1d, traz um sistema que adota estratégia baseada no ganho local, conectividade $k = 8$ e tentação $T = 1,30$. O regime estacionário é bem definido a partir da rodada 400, apresentando um transiente mais longo do que o sistema da Fig. 4.1c. Há um foco principal de resistência e emergência da cooperação nos jogadores 438-439. O *finger* originado no jogador 637 é absorvido na colisão com o *glider* originado no jogador 443. Há uma persistência da deserção dos jogadores 100 a 200, 880 a 970 e de 990 a 60 (condições periódicas de contorno) que conglomeram desertores que vão sendo invadidos pela deserção e acabam por criar formas triangulares.

Na Fig. 4.1e, o sistema adota a estratégia subjetiva, conectividade $k = 2$ e tentação $T = 1,10$. Um regime completamente estacionário é atingido na rodada 358, onde existem vários jogadores que mudam seus estados a cada rodada, mas o nível médio de cooperação se mantém inalterado, evidenciado pela coexistência de grandes agrupamentos com diferentes níveis de cooperação. Algumas colisões de *gliders* produzem *fingers*, por exemplo a colisão na rodada 120 no jogador 317, ou então elas simplesmente se cruzam sem mudarem suas direções, como na rodada 66 jogador 505. Também pode-se notar que a colisão de *gliders* desertores pode bloquear a invasão desertora, como a colisão na rodada 162, do jogador 174 ou do jogador 420 na rodada 358.

Na Fig. 4.1f, o sistema adota a estratégia subjetiva, conectividade $k = 2$ e tentação $T = 1,20$. Como no caso anterior, o regime estacionário atingido na rodada 533 é completamente estacionário, mas aqui há um número menor de agrupamentos com diferentes níveis de cooperação e alguns *fingers*. Há uma série de colisões de

gliders desertores, que acabam por extinguir a deserção: rodada 292 jogador 469, rodada 363 jogador 74, rodada 427 jogador 911, rodada 533 jogador 234.

O núcleo básico de cooperação no jogador 443 é o responsável pela resistência e emergência da cooperação nos sistemas que adotam as estratégias darwiniana ou as baseadas no ganho.

Nesta seção, apresentamos a aparência do regime transiente e verificamos que ele depende da estratégia adotada. A estratégia determina quais fenômenos estarão presentes no transiente e, por conseguinte, sua persistência e forma. Após o transiente, o sistema atinge o regime estacionário, que será abordado na seção seguinte.

4.2 Regime estacionário

A grandeza de interesse do sistema é o nível médio assintótico de cooperação, ρ_∞ (Eq. 3.1), que depende do conjunto de parâmetros (T, k, ρ_0) e da estratégia adotada. Considera-se o regime estacionário a última rodada, $EPC = 1000$, pois, a partir da análise do transiente mostrada na seção anterior, pode-se inferir que a maioria dos sistemas já está no regime estacionário na milésima rodada. Se o número de jogadores, que mudam seus estados a cada rodada for nulo, o sistema atingiu o regime estacionário. Assim, é possível evitar que os jogadores façam suas interações e comparações, quando o sistema já tiver atingido o regime estacionário, verificando se o número de sítios ativos é nulo e, conseqüentemente, reduzir o tempo de simulação computacional. Isso somente é possível pois são utilizadas estratégias deterministas e o sistema não está sujeito a ruídos.

A grandeza ρ_∞ , além da dependência com o conjunto de parâmetros e estratégias, pode depender também da configuração inicial do sistema. Então, considera-se a média configuracional $\langle \rho_\infty \rangle = (1/M) \sum_{m=1}^M \rho_\infty(T; k; \rho_0 \cong 1/2; epc = 1000; m)$, obtida de $M = 1000$ configurações iniciais diferentes, onde m é cada uma das diferentes configurações.

Os resultados estão divididos nas seções que tratam:

- 4.2.1: efeito da função densidade de probabilidade (PDF) adotada no nível inicial de cooperação dos jogadores;

- 4.2.2: efeito do tipo de vizinhança dos jogadores;
- 4.2.3: efeito das estratégias e
- 4.2.4: análise dos histogramas dos resultados que geram $\langle \rho_\infty \rangle$ e seu desvio padrão³.

As três seções iniciais começam com uma comparação entre os resultados obtidos com cada uma das PDFs/vizinhanças/estratégias nos sistemas sem auto-interação, mostrando suas principais propriedades, similaridades e diferenças. Em seguida, apresentamos os sistemas com auto-interação comparando-os com os sem auto-interação e explicando comportamentos distintos que possam aparecer.

4.2.1 Função densidade de probabilidade (PDF)

Iniciamos pelo estudo do efeito da função de densidade de probabilidade (PDF, do inglês, *probability density function*) utilizada para o nível de cooperação inicial dos jogadores sobre $\langle \rho_\infty \rangle$ em sistemas que adotam a estratégia darwiniana.

No DPD, há exatamente $N\rho_0$ cooperadores no instante inicial da simulação. No DPC é necessário utilizar distribuições do nível inicial de cooperação dos jogadores que sigam funções densidade de probabilidade conhecidas. Na Fig. 4.4, apresentamos as formas das PDFs utilizadas e seus histogramas de uma simulação, para exemplificar uma distribuição inicial de cada PDF. Na Fig. 4.4a, temos as PDFs que são geradas pelo método da transformação inversa. As demais figuras mostram histogramas do nível de cooperação inicial de uma simulação, para as PDFs: uniforme - Fig. 4.4b, triangular central - Fig. 4.4c, triangular de extremos - Fig. 4.4d, crescente - Fig. 4.4e e decrescente - Fig. 4.4f.

Apresentamos os gráficos da média configuracional do nível médio assintótico de cooperação em função da tentação e conectividade, $\langle \rho_\infty \rangle \sim \langle \rho_\infty(T, k) \rangle$, para sistemas sem auto-interação na Fig. 4.5 e com auto-interação na Fig. 4.7. Cada curva representa uma conectividade diferente. Testamos as seguintes PDFs, na ausência e presença de auto-interação, respectivamente: discreto - Figs. 4.5a e 4.7a, uniforme

³Utilizaremos o desvio padrão ao invés do erro padrão para não perdermos a amplitude da variação, pois em um mesmo gráfico há regiões com um desvio muito grande e outras com desvio praticamente nulo e o uso do erro padrão tornaria menos perceptível a diferença entre as regiões intermediárias e as de desvio praticamente nulo.

Figura 4.4 – (a) Funções densidade de probabilidade utilizadas para a definição do nível de cooperação dos jogadores no início do jogo. (b)-(f) Histogramas do nível inicial de cooperação dos jogadores, em uma simulação, para cada uma das PDFs utilizadas.

- Figs. 4.5b e 4.7b, triangular central - Figs. 4.5c e 4.7c, triangular de extremos - Figs. 4.5d e 4.7d, crescente - Figs. 4.5e e 4.7e e decrescente - Figs. 4.5f e 4.7f. As Figs. 4.6 e 4.8 mostram o desvio padrão de $\langle \rho_\infty \rangle$ das Figs. 4.5 e 4.7, respectivamente.

A Fig. 4.5a apresenta os resultados obtidos no DPD. Este sistema será nossa referência, pois grande parte dos trabalhos sobre o tema utilizam essa configuração. Para $k = 2$ e $T > 1,00$, $\langle \rho_\infty \rangle < 0,2$ e não surgem agrupamentos de cooperadores, devido à baixa conectividade, que mantém a cooperação pequena. O desvio padrão para $k = 2$ também é pequeno (vide Fig. 4.6a). Para conectividade $k = 4$, o nível médio de cooperação tem duas regiões bem distintas, que são delimitadas pela transição de fase. Para $T \leq 1,50$, $\langle \rho_\infty \rangle \sim 1$ e para $T > 1,50$, $\langle \rho_\infty \rangle \sim 0$ e o desvio padrão se mantém pequeno para $k = 4$ (vide Fig. 4.6a).

Com $k = \{6; 8\}$, os resultados numéricos confirmam um número maior de transições de fase, esperado pela função de tentação crítica da Eq. 3.34. Para $T \sim 1$, $\langle \rho_\infty \rangle \sim 1$ e diminui a cada transição de fase até atingir zero, $\langle \rho_\infty \rangle \sim 0$.

Conforme a Fig. 4.6a, para $k = \{6; 8\}$ o desvio padrão é aproximadamente $1/2$ em grande parte do intervalo de T e passa a ser nulo quando o nível médio de cooperação também é nulo. Para $k = \{24; 26\}$, as curvas têm comportamentos muito parecidos entre si, o que mostra uma convergência do nível médio de cooperação para $k \gg 1$. Apesar de ser esperado um número ainda maior de transições de fase, vemos somente algumas no início do intervalo de T , com um grande desvio padrão, e depois o nível médio de cooperação se anula rapidamente e o desvio padrão também se anula quando a cooperação se extingue, $\langle \rho_\infty \rangle \sim 0$.

Analizamos o efeito das diferentes PDFs através da comparação entre elas. De maneira geral, no DPC, não há transições de fase abruptas, como observado no DPD, e $\langle \rho_\infty \rangle$ é maior do que no DPD, para $T > 1,50$. Logo, as simulações numéricas confirmaram a inexistência de transições de fase, que havia sido constatada no cálculo analítico da tentação crítica mostrado na Seção 3.2.1.

A Fig. 4.5b considera a PDF uniforme. Com $k = 2$, em $T = 1,00$, o nível médio de cooperação é o maior em todo intervalo de T . Enquanto no DPD, para $T > 1$, temos $\langle \rho_\infty \rangle \sim 0,2$ e constante, no DPC uniforme, para $T = 1,01$, $\langle \rho_\infty \rangle \sim 0,5$ e $\langle \rho_\infty \rangle$ decai até $\langle \rho_\infty \rangle \sim 0,3$ em $T = 2,00$. Para $k = 4$, $\langle \rho_\infty \rangle \sim 0,9$ para $T \sim 1$ e diminui de forma mais acentuada até $\langle \rho_\infty \rangle \sim 0,8$, quando em $T \sim 1,10$, surge um

platô que se estende até $T \sim 1,40$ e então volta a diminuir, atingindo $\langle \rho_\infty \rangle \sim 0,5$ para $T = 2,00$. Para $k = \{6; 8\}$, não surge o platô observado com $k = 4$ e o nível médio de cooperação é inferior ao de $k = 4$ até $T \sim 1,80$, quando ocorre uma inversão, mas após a inversão os valores se mantêm próximos. Para $k = \{24; 26\}$, o comportamento de $\langle \rho_\infty \rangle$ é o mesmo para ambas as conectividades os valores são praticamente iguais. O comportamento destas curvas está marcado por uma queda mais acentuada de $\langle \rho_\infty \rangle$ nos valores iniciais de tentação e em $T = 1,70$, atinge um valor que se mantém praticamente constantes até $T \sim 2$. O valor de $\langle \rho_\infty \rangle$ para altas conectividades no DPC é maior do que o obtido no DPD em todo o intervalo de tentação. Para baixa conectividade, enquanto $T < 1,25$ o DPD tem os maiores valores de $\langle \rho_\infty \rangle$ e a partir desta tentação o DPC passa a apresentar valores maiores de $\langle \rho_\infty \rangle$ e para $k \gg 1$ o DPC sempre tem maior nível de cooperação.

Nas Figs. 4.5c-4.5f o comportamento das curvas é o mesmo do sistema com PDF uniforme da Fig. 4.5b. Apesar da forma ser semelhante, os valores apresentam algumas diferenças que podem ser explicados pelas formas das diferentes PDFs. Nas PDFs triangular de extremos (Fig. 4.5d) e crescente (Fig. 4.5e), para $k = 4$, observa-se ao invés de um platô, uma corcova, definida por um aumento do nível médio assintótico de cooperação seguido de sua redução de forma monotônica. A PDF que mais se diferenciou foi a decrescente (Fig. 4.5f), pois o platô, presente para $k = 4$, observado nas demais PDFs não está presente. A PDF crescente (Fig. 4.5e) apresentou os maiores valores para o nível de cooperação em função da tentação para altas conectividades.

A Fig. 4.5c, mostra o nível de cooperação para sistemas com PDF triangular central. Nesta PDF há uma incidência maior dos valores próximos a $1/2$ nos instantes iniciais da simulação e conseqüentemente há uma redução no número de desertores extremos ($\alpha_i \sim 0$), que poderiam obter ganhos maiores nas jogadas iniciais e serem copiados pelos demais jogadores. Por isso, ocorre uma concentração dos resultados de $\langle \rho_\infty \rangle$ próximos a $1/2$. Por outro lado, a PDF triangular de extremos da Fig. 4.5d nos leva ao cenário oposto e surge um maior distanciamento entre os resultados das diferentes conectividades. Seguindo esta lógica, vemos que, para a PDF crescente (Fig. 4.5e), onde a incidência de cooperadores é maior, há um aumento de $\langle \rho_\infty \rangle$, enquanto que a PDF decrescente (Fig. 4.5f), com mais desertores,

há uma redução no valor de $\langle \rho_\infty \rangle$.

Na Fig. 4.6 são mostradas as curvas da Fig. 4.5 em separado com o respectivo desvio padrão. Em geral, o desvio padrão de $\langle \rho_\infty \rangle$ no DPD apresenta regiões com valores bastante elevados na região de tentação intermediária (Fig. 4.6a). No entanto, o DPC apresenta valores de desvio padrão baixo. O resultado que tem o maior desvio padrão é obtido com $k = 8$ na PDF triangular de extremos (Fig 4.6d), e corresponde a curva com a maior rugosidade no nível médio de cooperação.

Na Fig. 4.7, os resultados são equivalentes aos da Fig. 4.5, mas para sistemas com auto-interação e as curvas referentes as altas conectividades, para as PDFs contínuas, têm o mesmo comportamento. Com a auto-interação, o platô no nível de cooperação, que estava associado à conectividade $k = 4$, ficou a cargo de $k = 3$ e está presente até mesmo na PDF decrescente.

As relações de maior ou menor nível de cooperação entre as PDFs utilizadas são equivalentes ao observado no caso com auto-interação. Os menores níveis de cooperação média assintóticos são observados com $k = \{25; 27\}$ para as PDF triangular de extremos (Fig. 4.7d) e a decrescente (Fig. 4.7f). Nestas PDFs, há um grande número de cooperadores no instante inicial. Estes cooperadores são explorados pelos vizinhos desertores e lhes provêem ganhos elevados. Como a conectividade é grande e os jogadores adotam a estratégia darwiniana, muitos jogadores copiam esses desertores que obtiveram ganhos elevados, o que leva o sistema a um nível médio assintótico de cooperação baixo.

Por outro lado, os maiores níveis de cooperação surgem quando há a possibilidade de se formarem agrupamentos de cooperadores que conseguem se manter, apesar da exploração dos desertores. A situação em que isso é mais marcante é para a distribuição crescente (Fig. 4.7e), que apresenta o maior nível de cooperação dentre as PDFs, até mesmo para altas conectividades.

A Fig. 4.8, mostra o desvio padrão dos gráficos da Fig. 4.7. De maneira análoga ao observado para o desvio padrão de sistemas sem auto-interação, o DPD (Fig. 4.8a) apresenta os maiores valores de desvio padrão, em especial, quando temos as transições de fase nos valores centrais de tentação. Já, no DPC, o desvio padrão é pequeno apresentando seus maiores valores para $k = 9$ na PDF triangular de extremos (Fig 4.8d).

Figura 4.5 – Estudo do efeito das diferentes funções densidade de probabilidade. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas têm agentes que adotam a estratégia darwiniana e não realizam auto-interação (k par).

Figura 4.6 – Desvio padrão de cada um dos gráficos apresentados na Fig. 4.5. Cada conectividade é apresentada em separado.

Figura 4.7 – Estudo do efeito das diferentes funções densidade de probabilidade. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas têm agentes que adotam a estratégia darwiniana e realizam auto-interação (k ímpar).

Figura 4.8 – Desvio padrão de cada um dos gráficos apresentados na Fig. 4.7. Cada conectividade é apresentada em separado.

Comparando sistemas com e sem auto-interação para as altas conectividades, os valores de $\langle \rho_\infty \rangle$ são muito próximos, a menos de um deslocamento em T . No entanto, para um dado valor de T , sistemas com auto-interação tem um nível de cooperação maior do que os sem auto-interação. Para baixas conectividades, as curvas são similares nos sistemas com e sem auto-interação, a exceção do caso $k = 2$, que apresenta um comportamento bastante particular, que não é observado para $k = 3$.

Afim de não adicionarmos nenhum tipo de tendência nos resultados, devido à PDF inicial, optamos por utilizar somente a PDF uniforme nas simulações seguintes. Note que a PDF uniforme apresenta os valores intermediários dentre as PDFs contínuas e o comportamento das curvas é semelhante às demais PDFs, apesar de não apresentar a corcova mais marcante para $T \sim 1,40$ com $k = \{3; 4\}$, como nas PDFs triangular de extremos e crescente, ou os menores níveis de cooperação para altas conectividades da PDF decrescente.

4.2.2 Conectividade

O trabalho de Nowak e May [34, 101] credita a emergência da cooperação ao uso da estrutura bidimensional. No entanto, quando utilizamos a estrutura unidimensional, observamos as mesmas transições de fase, mas com uma redução no valor do nível médio de cooperação assintótico [46, 47, 48]. Então, para verificar a importância da forma como os jogadores estão conectados, utilizamos além da vizinhança regular as vizinhanças aleatórias fixa e dinâmica. A vizinhança aleatória fixa tem inspiração nos sistemas congelados (*quenched*) de física de sistemas desordenados e a aleatória dinâmica nos sistemas recozidos (*annealed*).

Não podemos falar em auto-interação para as vizinhanças aleatória fixa e dinâmica, pois os vizinhos são escolhidos aleatoriamente. No entanto, mantivemos o padrão de apresentação dos resultados, separando as conectividades com valores pares na Fig. 4.9 e ímpares na Fig. 4.11. Nestas figuras, apresentamos sistemas que utilizam a estratégia darwiniana. Começamos apresentando os sistemas que nos servirão de referência: dilema do prisioneiro com estratégia darwiniana com PDF discreta (Fig. 4.9a) e contínua uniforme (Fig. 4.9b). Em seguida, apresentamos os resultados para a conectividade aleatória fixa no dilema do prisioneiro com

estratégia darwiniana e PDF discreta (Fig. 4.9c) e contínua com PDF uniforme (Fig. 4.9d). Finalmente, apresentamos a conectividade aleatória dinâmica nos casos discreto (Fig. 4.9e) e contínuo com PDF uniforme (Fig. 4.9f).

As vizinhanças aleatórias reduzem significativamente os valores de $\langle \rho_\infty \rangle$. No DPD com vizinhança aleatória fixa (Fig. 4.9c) não há a completa extinção da cooperação para $k = 4$, mas para $k > 4$ a cooperação é praticamente nula para todo o intervalo de T , isto é, $\langle \rho_\infty \rangle \sim 0$. Neste sistema, para $k = 2$, surge um resultado que merece destaque: o comportamento de $\langle \rho_\infty \rangle$ é o mesmo tanto para conectividade aleatória fixa quanto para a regular (Fig. 4.9a). Para jogadores com vizinhança aleatória fixa no caso do DPD (Fig. 4.9d), a cooperação não se extingue, como no caso do DPD, enquanto $T < 1,20$, mas é muito menor do que o observado no caso contínuo para a vizinhança regular (Fig. 4.9b).

A vizinhança aleatória dinâmica, na qual a cada rodada os jogadores elegem novos vizinhos, é ainda pior no que diz respeito à emergência e manutenção da cooperação. Tanto para jogadores do DPD quanto do DPC, ocorre a extinção da cooperação para quaisquer valores de conectividade, como ilustram as Figs. 4.9e e 4.9f, respectivamente.

Na Fig 4.10, as curvas com o desvio padrão são apresentadas, no entanto, seus valores não agregam informações relevantes. Como poderíamos supor, os maiores valores de desvio padrão ocorrem para a vizinhança aleatória fixa (Fig. 4.10b), que possuem valores de desvio diferentes de zero para $T < 1,20$, porque nas demais vizinhanças, assim como $\langle \rho_\infty \rangle \sim 0$, o desvio padrão também é praticamente nulo.

Na Fig. 4.11, não estamos tratando a auto-interação, mas sim um aumento na conectividade. O efeito do aumento da conectividade é uma diminuição do nível médio de cooperação nos casos em que $\langle \rho_\infty \rangle$ ainda não eram nulos, ou seja, para a vizinha aleatória fixa no caso contínuo (Fig. 4.11d) e para $k = 3$ no caso discreto (Fig. 4.11c). Sobre o desvio padrão, mostrado na Fig. 4.12, podemos dizer o mesmo que para a conectividade par, ou seja, surgem pequenos desvios para os valores não nulos de $\langle \rho_\infty \rangle$ na vizinhança aleatória fixa no caso contínuo (Fig. 4.12b).

Os resultados para os diferentes tipos de conectividade, mostram que o número de vizinhos não é a grandeza mais relevante na explicação da emergência da cooperação. No entanto, a estrutura que define quem são os vizinhos é a responsável

(a) vizinhança regular (discreto) - referência, idem Fig. 4.5a

(b) vizinhança regular (contínuo) - referência, idem Fig. 4.5b

(c) vizinhança aleatória fixa (discreto)

(d) vizinhança aleatória fixa (contínuo)

(e) vizinhança aleatória dinâmica (discreto)

(f) vizinhança aleatória dinâmica (contínuo)

Figura 4.9 – Estudo do efeito dos tipos de vizinhança. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas adotam a estratégia darwiniana e k assume valores pares.

Figura 4.10 – Desvio padrão de cada um dos gráficos apresentados na Fig. 4.9 Cada conectividade é apresentada em separado, de acordo com o texto acima do gráfico, para facilitar a leitura das barras de desvio padrão. Omitimos a vizinhança regular para o dilema do prisioneiro discreto e contínuo, pois estas foram apresentadas nas Figs. 4.6a e 4.6b.

pela emergência e manutenção da cooperação. O nível de cooperação assintótico é muito baixo no caso das vizinhanças aleatórias, pois um jogador A interage com sua vizinhança κ_A e os jogadores pertencentes à κ_A não precisam escolher o jogador A como oponente em sua jogada. Isto, dificulta a retaliação de uma deserção ou a reciprocidade de uma cooperação.

Nas Refs. [46, 47, 48], mostramos que as transições de fase em sistemas uni ou bidimensionais ocorrem com os mesmos valores de tentação crítica. No entanto, o valor assintótico no caso unidimensional era menor que no bidimensional. Agora, com as vizinhanças aleatórias, vimos que as conexões indiretas (vizinhos dos vizi-

(a) vizinhança regular (discreto) - referência, idem Fig. 4.7a

(b) vizinhança regular (contínuo) - referência, idem Fig. 4.7b

(c) vizinhança aleatória fixa (discreto)

(d) vizinhança aleatória fixa (contínuo)

(e) vizinhança aleatória dinâmica (discreto)

(f) vizinhança aleatória dinâmica (contínuo)

Figura 4.11 – Estudo do efeito dos tipos de vizinhança. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas adotam a estratégia darwiniana e k assume valores ímpares.

Figura 4.12 – Desvio padrão de cada um dos gráficos apresentados na Fig. 4.11. Cada conectividade é apresentada em separado, de acordo com o texto acima do gráfico, para facilitar a leitura das barras de desvio padrão. Omitimos a vizinhança regular para o dilema do prisioneiro discreto e contínuo, pois estas foram apresentadas nas Figs. 4.8a e 4.8b.

nhos) influenciam decisivamente na emergência da cooperação. Pudemos notar que no caso aleatório fixo, a cooperação ainda tem uma leve persistência para baixas conectividades $k < 5$, mas não resiste ao aumento da conectividade e da tentação. A persistência da cooperação para a vizinhança aleatória fixa reforça o fato de que manter os mesmos vizinhos tem influência na emergência da cooperação. No outro extremo, está a conectividade $k = 2$ para o caso discreto, que é indiferente ao tipo de conectividade utilizada, ou seja, com poucos vizinhos pouco importa onde eles estão e como interagem entre si.

4.2.3 Estratégia

Optamos pela PDF uniforme para evitar tendências causadas pelas formas da distribuição inicial do nível inicial de cooperação dos jogadores. A vizinhança aleatória leva à extinção da cooperação, então prosseguiremos nossos estudos utilizando apenas a vizinhança regular. Neste momento, vamos estudar a influência das diferentes estratégias na emergência da cooperação. Utilizamos a estratégia darwiniana (referência), a baseada na personalidade e a baseada no ganho global ou local, como ilustrado na Fig. 4.13.

Os casos de um sistema com agentes que jogam o DPD (Fig. 4.13a) e o DPC com PDF uniforme (Fig. 4.13b), em que ambos adotam a estratégia darwiniana foram analisados na Seção 4.2.1.

Na Fig. 4.13c, apresentamos os resultados da estratégia baseada na personalidade. As curvas de $\langle \rho_\infty \rangle$ têm a mesma forma observada na estratégia darwiniana. Para $k = 2$, ocorre um aumento de $\langle \rho_\infty \rangle$ em comparação com os resultados da estratégia darwiniana. No entanto, para as demais conectividades a estratégia baseada na personalidade tem $\langle \rho_\infty \rangle$ sutilmente menor que o observado na estratégia darwiniana.

Os resultados obtidos com a estratégia baseada no ganho global (Fig. 4.13d) e a estratégia baseada no ganho local (Fig. 4.13e) mostram que $\langle \rho_\infty \rangle$ pouco se diferencia do obtido com a estratégia darwiniana, exceto para $k = 2$, onde seu valor aumenta para todos os valores de T . As estratégias baseadas no ganho têm $\langle \rho_\infty \rangle$ um pouco maior do que a da darwiniana, sendo a baseada no ganho local a com maior aumento na cooperação.

No entanto, se considerarmos o desvio padrão dos gráficos da Fig. 4.13, apresentados na Fig. 4.14, podemos considerar os resultados das novas estratégias indistinguíveis dos obtidos com a estratégia darwiniana no caso contínuo. O desvio padrão dos sistemas é sistematicamente pequeno, ainda mais se comparado ao observado no DPD. Seus maiores valores ocorrem para a estratégia baseada na personalidade.

Para estudar o efeito das estratégias em sistemas com auto-interação considere os gráficos da Fig. 4.15. Os resultados da estratégia darwiniana (DPD, Fig. 4.15a e DPC, Fig. 4.15b) foram analisados na Seção 4.2.1.

Para a estratégia baseada na personalidade com auto-interação, os resultados são apresentados na Fig. 4.15c. Assim, como aconteceu na ausência de auto-

(a) darwiniana (discreto) - referência,
idem Fig. 4.5a

(b) darwiniana - idem Fig. 4.5b

(c) subjetiva

(d) baseada no ganho global

(e) baseada no ganho local

Figura 4.13 – Estudo do efeito das estratégias. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas adotam a estratégia darwiniana e realizam não auto-interação (k par).

Figura 4.14 – Desvio padrão de cada um dos gráficos apresentados na Fig. 4.13. Cada conectividade é apresentada em separado, de acordo com o texto acima do gráfico, para facilitar a leitura das barras de desvio padrão. Omitimos a estratégia darwiniana para os casos discreto e contínuo, pois estes já foram apresentados nas Fig. 4.6a e Fig. 4.6b, respectivamente.

interação, comparando com a estratégia darwiniana no caso contínuo, ocorre um aumento do nível de cooperação para $k = 2$ e, para as demais conectividades, o nível de cooperação é levemente menor que o obtido com a estratégia darwiniana.

Os resultados da estratégia darwiniana (Fig. 4.15b) e a baseada na personalidade (Fig. 4.15c) com auto-interação são praticamente idênticos, consequência direta da auto-interação. Se um jogador tivesse obtido o maior ganho real em sua vizinhança e, ao modificar seu ganho segundo sua função subjetividade, seu ganho subjetivo fosse menor que o de algum vizinho, ele iria copiar este vizinho. Mas, se ele realiza a auto-interação, ele copiaria seu próprio estado. Mesmo a personalidade modificando a percepção de ganho, não alteraria a tomada de decisão do jogador. Este efeito é dominante para conectividades mais elevadas.

As estratégias baseadas no ganho global (Fig. 4.15d) e local (Fig. 4.15e) com auto-interação tem o mesmo comportamento do caso sem auto-interação em relação

a estratégia darwiniana. Elas apresentam um nível médio de cooperação um pouco maior do que o da estratégia darwiniana, quando comparada ao sistema no qual os jogadores fazem auto-interação.

O desvio padrão dos gráficos da Fig. 4.15 são apresentados na Fig. 4.16 e, assim como acontece no caso sem auto-interação, seu valores são pequenos e não agregam informações relevantes.

Para facilitar a comparação entre as estratégias, agrupamos todas em um mesmo gráfico, para os casos $k = \{8; 9\}$, ilustrado na Fig. 4.17. Sem a auto-interação (vide Fig. 4.17a), há um pequeno distanciamento entre as curvas no caso contínuo mas, quando se considera o desvio associado a estas curvas, não é possível diferenciá-las sem um teste estatístico específico. Com a auto-interação (vide Fig. 4.17b), o distanciamento fica ainda menor do que o observado sem a auto-interação.

Para todas as estratégias, a auto-interação promove um aumento no nível médio assintótico de cooperação. Relembramos que na estratégia baseada na personalidade, os jogadores consideram seus ganhos segundo sua personalidade. Alguns exemplos dessa percepção do ganho foram representados na Fig. 3.6. Da maneira como foi considerada a função de subjetividade para modificar o ganho, um desertor total nunca está satisfeito com seu ganho, sempre o considerando como nulo ($\Pi(0, \hat{q}) = 0$). Já um cooperador total não distorce seu ganho e faz uma avaliação real do mesmo ($\Pi(1, \hat{q}) = 1$). Por isso, seria de se esperar que a personalidade dos jogadores tivesse papel de destaque na tomada de decisão e consequente emergência ou não da cooperação. Devido às características da função de subjetividade Π , os desertores tendem a serem mais insatisfeitos com seus ganhos e copiam jogadores mais cooperativos que podem ter ganhos reais menores que os seus. Esse comportamento serviria para explicar o não surgimento da deserção $\langle \rho_\infty \rangle \sim 0$.

No entanto, na estratégia darwiniana com PDF contínua não ocorre a extinção da cooperação. Os níveis de cooperação são muito similares entre as estratégias darwiniana, baseada na personalidade e baseadas no ganho. As estratégias propostas só podem ser aplicadas no dilema do prisioneiro contínuo, então não é possível compará-las ao caso discreto. Portanto, as estratégias propostas não se mostraram mais efetivas que a darwiniana em promover a cooperação no sistema.

Visto que, os resultados obtidos com as estratégias propostas não apresen-

(a) darwiniana (discreto)

(b) darwiniana

(c) subjetiva

(d) baseada no ganho (global)

(e) baseada no ganho (local)

Figura 4.15 – Estudo do efeito das estratégias. Gráficos da média configuracional do nível médio assintótico de cooperação do sistema em função da tentação T . Cada curva do gráfico representa um sistema em que os agentes tem uma conectividade diferente, conforme a legenda. Nesta figura todos os sistemas adotam a estratégia darwiniana e realizam auto-interação (k ímpar).

Figura 4.16 – Desvio padrão de cada um dos gráficos apresentados na Fig. 4.15. Cada conectividade é apresentada em separado, de acordo com o texto acima do gráfico, para facilitar a leitura das barras de desvio padrão. Omitimos a estratégia darwiniana para os casos discreto e contínuo, pois estes já foram apresentados nas Fig. 4.8a e Fig. 4.8b, respectivamente.

taram grandes diferenças em relação a estratégia darwiniana. Isso sugere que, ou devem ser feitas modificações na função subjetividade no caso da estratégia baseada na personalidade, ou que o mecanismo de cópia advindo da estratégia darwiniana tem um peso maior na evolução do sistema do que a percepção de ganhos. As estratégias modificaram de maneira significativa somente o transiente, como mostrado na Seção 4.1. Neste caso, possivelmente, o estudo do transiente seja mais importante do que o do regime estacionário para a compreensão das características das estratégias.

A comparação dos casos contínuo e discreto, adotando a estratégia darwiniana, revelou um aspecto importante sobre a distribuição dos resultados das configurações, médias configuracionais e desvio padrão. Exploramos essas características na próxima seção.

Figura 4.17 – Comparação entre as diferentes estratégias para os casos (a) $k = 8$ e (b) $k = 9$. DPD: dilema do prisioneiro discreto e estratégia darwiniana; DPC: dilema do prisioneiro contínuo e estratégia darwiniana; EBP: dilema do prisioneiro contínuo e estratégia baseada na personalidade; EBGG: dilema do prisioneiro contínuo e estratégia baseada no ganho global; e EBGL: dilema do prisioneiro contínuo e estratégia baseada no ganho local.

4.2.4 Médias configuracionais

Para a análise da cooperação em um sistema, quando calcula-se um valor de $\langle \rho_\infty \rangle$ e seu desvio padrão⁴ estamos analisando apropriadamente $\langle \rho_\infty \rangle$ e a cooperação do sistema?

Vamos analisar ρ_∞ a partir de três possíveis resultados que podem ocorrer tanto no DPD, quanto no DPC, e um exclusivo do DPC:

1. todos os jogadores cooperadores ($\alpha_i = 1$), logo $\rho_\infty^A = 1$;
2. todos os jogadores desertores ($\alpha_i = 0$), logo $\rho_\infty^B = 0$;
3. metade dos jogadores $n/2$ cooperadores ($\alpha_i = 1$) e a outra metade desertores ($\alpha_i = 0$), logo $\rho_\infty^C = 1/2$;

⁴O desvio padrão é dado por

$$std = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \right)^{1/2}$$

onde a média vale:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i.$$

4. todos os jogadores com nível de cooperação ($\alpha_i = 1/2$) (caso válido somente para o DPC), logo $\rho_\infty^D = 1/2$.

Usando a proporção de cooperadores, diríamos que 3 e 4 tem o mesmo nível médio de cooperação. No entanto, pode-se dizer que os sistemas são equivalentes ou diferenciá-los de alguma maneira?

Se, ao comparar dois sistemas, já nos defrontamos com esta dificuldade, a situação fica ainda mais complicada ao considerarmos as médias configuracionais do nível médio assintótico de cooperação das diferentes configurações do mesmo sistema. Quando analisamos o nível médio de cooperação de um sistema, $\langle \rho_\infty \rangle = 1/M \sum_{i=1}^M \rho_{\infty_i}$, estamos tratando de seu valor médio sobre todas as configurações. A análise fica ainda mais nebulosa se pensarmos em como podem ser compostos os valores de $\langle \rho_\infty \rangle$ de um sistema⁵. Apresentamos algumas possibilidades de como compor $\langle \rho_\infty \rangle$ na Tab. 4.1.

Tabela 4.1 – Possíveis resultados de $\langle \rho_\infty \rangle$.

$$\langle \rho_\infty \rangle_1 = (\rho_\infty^A + \rho_\infty^B)/2 = (1 + 0)/2 = 1/2$$

$$\langle \rho_\infty \rangle_2 = (\rho_\infty^C + \rho_\infty^C)/2 = (1/2 + 1/2)/2 = 1/2$$

$$\langle \rho_\infty \rangle_3 = (\rho_\infty^D + \rho_\infty^D)/2 = (1/2 + 1/2)/2 = 1/2$$

Ao analisar a Tab. 4.1, surgem questões do tipo: Como dizer que um sistema apresenta mais cooperação que outro? Podemos dizer que os sistemas $\{1,2,3\}$ são equivalentes, uma vez que, $\langle \rho_\infty \rangle_1 = \langle \rho_\infty \rangle_2 = \langle \rho_\infty \rangle_3$? Além da dificuldade inicial em definir o que seria um sistema cooperativo, a medida de proporção de cooperadores ou nível médio de cooperação $\langle \rho_\infty \rangle$ apresenta uma degenerescência muito grande, que torna sua análise pouco esclarecedora sobre as questões mencionadas. Existem várias configurações que são ortogonalmente distintas, mas apresentam o mesmo valor de $\langle \rho_\infty \rangle$.

A presença de eventos extremos, isto é, poucas configurações que geram cooperação total dentro de uma grande maioria de configurações que geram deserção

⁵Note que, para de $\rho_\infty^{(A,B,C,D)}$, não consideramos o desvio padrão. Caso levássemos em conta o desvio padrão teríamos: $\rho_\infty^A = 1 \pm 0$; $\rho_\infty^B = 0 \pm 0$; $\rho_\infty^C = 1/2 \pm 1/2$ (o desvio padrão é igual a $std = N/(2N - 1) \sim 1/2$ para $N \gg 1$) e $\rho_\infty^D = 1/2 \pm 0$ e deveríamos utilizar a propagação de erros para calcular o desvio de $\langle \rho_\infty \rangle$.

total ou vice-versa, não podem ser percebidos na análise do desvio padrão. A única forma de certificar-se que um dado sistema atingirá apenas um estado assintótico plenamente cooperador ou desertor, seria simular *todas* as configurações iniciais possíveis. Para $N = 1000$ e $\rho_0 = 1/2$, há um número muito grande de configurações distintas possíveis⁶ e definir quais são elas é muito complicado devido ao uso das condições periódicas de contorno.

Para ilustrar alguns exemplos de como são constituídos os resultados apresentados nas seções anteriores, mostramos nas Figs. 4.18 e 4.19 os histogramas das $M = 1000$ configurações que produzem $\langle \rho_\infty \rangle$ para alguns valores de tentação $T = \{1,00; 1,01; 1,50; 2,00\}$. Apresentamos apenas as conectividades $k = \{8; 9\}$. Na Fig. 4.18, a conectividade é $k = 8$ (sem auto-interação). Nela mostramos o DPD com estratégia darwiniana (Fig. 4.18a) e o DPC com as estratégias darwiniana (Fig. 4.18b), baseada na personalidade (Fig. 4.18c), baseada no ganho global (Fig. 4.18d) e baseada no ganho local (Fig. 4.18e).

Na Fig. 4.18a, está o dilema do prisioneiro discreto. Nela vemos que, com $T = 1,00$, o nível médio assintótico de cooperação dos sistemas se distribui entre $0,98 \leq \rho_\infty \leq 1$, ou seja, quase totalidade de cooperação. No entanto, um pequeno aumento de tentação para $T = 1,01$, faz surgir sistemas com nível médio de cooperação quase nulos ou nulos ($\rho_\infty \cong 0$). Enquanto que os demais sistemas atingem a quase totalidade ou totalidade ($\rho_\infty \cong 1$) de cooperadores.

A medida que a tentação aumenta, também aumenta a quantidade de configurações que tem maioria de desertores ($\rho_\infty \sim 0$), como vemos para $T = 1,50$, ocorre uma inversão do que era observado para $T = 1,01$. Neste caso, há mais configurações desertoras do que cooperadoras. Com valores ainda mais elevados de tentação, os sistemas ou são totalmente desertores, ou alguns poucos cooperadores resistem, aqueles que não foram dizimados durante a evolução, mas também não conseguem promover a cooperação no sistema. Isso é mostrado pelas poucas configurações que atingem $\rho_\infty \sim 0,01$, para $T = 2,00$. Nestas configurações, existem dez cooperadores em mil jogadores.

Na literatura, normalmente, os resultados apresentados são médias configu-

⁶Essa afirmação é válida somente para o caso discreto, pois, no caso contínuo, há infinitas configurações.

Figura 4.18 – Histogramas do nível médio assintótico de cooperação para diversas estratégias em sistemas com conectividade $k = 8$ (sem auto-interação). A tentação T de cada histograma é mencionada na primeira linha do gráfico. DPD: dilema do prisioneiro discreto. DPC: dilema do prisioneiro contínuo com PDF uniforme.

racionais do nível médio assintótico de cooperação. O que imaginamos, quando nos deparamos com a seguinte frase: “Um sistema apresenta nível médio de cooperação de 0,8 (80%).”? Se estivermos tratando de uma única configuração, seria que 80% dos jogadores são cooperadores. Se estivermos tratando do nível médio de cooperação de diferentes configurações, seria algo do tipo: em média 80% dos jogadores são cooperadores nestas configurações. Mas, imaginariamos que 80% das configurações apresentariam $\sim 100\%$ cooperação e 20%, com $\sim 100\%$ deserção?

A resposta para esta pergunta provavelmente deve ter sido *não*, pois estamos habituados a lidar com sistemas ergódicos [102, 103]. Sistema ergódico é aquele no qual a média temporal é igual à média espacial. Uma média temporal de uma realização, $\bar{f}_\tau$, é definida sobre as interações no regime estacionário, decorrentes de uma posição inicial χ .

$$\bar{f}_\tau = \frac{1}{\tau_f - \tau_i} \sum_{t=\tau_i}^{\tau_f} f(x_0 = \chi, t), \quad (4.1)$$

onde t é um instante em que o sistema já está no regime estacionário e a média é calculada no intervalo $\tau_i \leq t \leq \tau_f$. Na média espacial, $\bar{f}_\chi$, temos uma média do valor estacionário no instante τ decorrentes de diferentes posições iniciais x_0 .

$$\bar{f}_\chi = \frac{1}{\Psi} \sum_{x_0=\chi_i}^{\chi_f} f(x_0, t = \tau). \quad (4.2)$$

Em nosso problema, a média temporal é ρ_∞ e a média espacial $\langle \rho_\infty \rangle$.

Os resultados mostram que o dilema do prisioneiro discreto não é ergódico. Portanto, a média configuracional do nível médio de cooperação no caso discreto deve ser entendida como o número médio de sistemas com totalidade/maioria de cooperação. Os sistemas, em que os agentes jogam o DPC, são ergódicos, ou seja, a média temporal é a mesma que a média espacial.

Na Fig. 4.18b, na qual os agentes utilizam a estratégia darwiniana, surgem distribuições de ρ_∞ distintas para cada valor de tentação. O mesmo ocorre nas estratégias baseadas na: personalidade (Fig. 4.18c), ganho global (Fig. 4.18d) e ganho local (Fig. 4.18e), mas há diferença na forma da distribuição; na dispersão e média dos valores. No entanto, essas diferenças não são tão significativas quanto a observada ao mudar do DPD para o DPC.

Na Fig. 4.19, estão os histogramas dos sistemas com $k = 9$, com auto-interação, sendo: o DPD com estratégia darwiniana (Fig. 4.19a); e o DPC com as estratégias: darwiniana (Fig. 4.19b); baseada na personalidade (Fig. 4.19c); baseada no ganho global (Fig. 4.19d) e baseada no ganho local (Fig. 4.19e). A principal diferença entre os sistemas sem e com auto-interação é o aumento da média configuracional do nível médio cooperação. Além disso, em alguns casos houve também mudanças na forma das distribuição destes valores.

No DPD (Fig. 4.19a), outra diferença proveniente da auto-interação foi um aumento no intervalo da tentação em que só ocorre a cooperação, antes de começar a coexistirem configurações cooperadoras e desertoras. Observando-se as Fig. 4.18a e 4.19a, poderíamos ser levados a dizer que ocorrem dois regimes em relação à ergodicidade no DPD: um ergódico, onde todos as configurações de um sistema são cooperadoras, ou desertoras, e as médias temporais são iguais as médias espaciais, e outro não ergódico. Se fizermos isso, diremos, por exemplo, que o sistema seria não ergódico para $k = 8$ e $T > 1$. Mas, não podemos fazer essa afirmação, pois, como mencionamos, existe a possibilidade de surgirem configurações que geram os eventos extremos. A única forma de assegurar a não ocorrência de tais eventos seria a exploração exaustiva de todas as configurações possíveis de um sistema, algo inviável para $N \gg 10$.

Verificamos que a média e o desvio padrão não são capazes de definir se um sistema é totalmente cooperador/desertor. Então, haveria alguma técnica que poderia definir o sistema como cooperador ou desertor? Ou, confirmar que há uma maioria de configurações cooperadoras ou desertoras para um dado sistema? Para tentar responder a estas perguntas, analisamos na próxima seção os resultados dos novos qualificadores que propusemos.

4.3 Novos classificadores

Iniciamos os estudos desta tese vislumbrando as oportunidades oferecidas pela função generalizada de ganho para valores contínuos do nível de cooperação. Dentre estas oportunidades estavam o uso de novas estratégias e o estudo do papel da personalidade do jogador, afim ou avessa ao risco, pois cada jogador pode

Figura 4.19 – Histogramas do nível médio assintótico de cooperação para diversas estratégias em sistemas com conectividade $k = 9$ (com auto-interação). A tentação T de cada histograma é mencionada na primeira linha do gráfico. DPD: dilema do prisioneiro discreto. DPC: dilema do prisioneiro contínuo com PDF uniforme.

tomar decisões que modificam seu nível de cooperação de forma contínua. Após as primeiras simulações, surgiram as dificuldades em definir meios de comparar as estratégias utilizando o nível de cooperação do sistema e como definir os sistemas como cooperadores ou desertores.

Encontramos na literatura, artigos cujos títulos contém expressões como: “...*determinada estratégia/estrutura/propriedade do sistema/dinâmica de evolução/... aumenta/mantém/emerge/evolui a cooperação.*”, por exemplo, as Refs. [5, 70, 104, 105, 106, 107, 108, 109, 110, 111]. Mas afinal, o que é a cooperação? Convém lembrar que a cooperação é objeto de estudo de áreas como biologia, ecologia, ecossistemas, sociologia, economia, teoria dos jogos, etc. Em cada uma, o enfoque e abordagem adotados se diferenciam de acordo com o interesse da área, por isso, a definição de cooperação não é única. Alguns exemplos de definição para cooperação:

- Cooperação é um processo conjunto pelo qual os agentes envolvidos trabalham juntos para atingir resultados globais.
- “Um cooperador é alguém que paga um custo, c , para outro indivíduo receber um benefício, b .” [107].

No DPD, cooperação seria a maioria de jogadores do sistema jogando como cooperadores? Nesse caso, no DPC, onde existem níveis intermediários de cooperação, o que seria uma maioria de cooperadores? A proporção de cooperadores no DPI ou o nível médio de cooperação no DPC é dado por $\rho = (1/n) \sum_{i=1}^n \alpha_i$. Logo, cooperação seria $\rho > 1/2$? Ou seria a não extinção da cooperação? Ou ainda, um aumento do nível médio de cooperação em relação ao estado inicial do sistema?

A maioria dos trabalhos na literatura não define o que é cooperação e também não esclarece o que será considerado como cooperação no decorrer do trabalho. Normalmente, estes trabalhos consideram o dilema do prisioneiro com N jogadores e utilizam o nível médio de cooperação, ρ , que é, na verdade, uma média configuracional. Muitas vezes, sequer citam o nível de cooperação inicial do sistema e quando tratam de ρ , não fazem qualquer menção sobre seu desvio padrão ou sobre a distribuição dos resultados que compõem esses valores médios. Como vimos na análise das médias configuracionais e suas distribuições, essas informações mudam

drasticamente a interpretação e compreensão dos resultados. Então, como dizer que a cooperação (que não foi definida) aumenta/emerge/se mantém por esse motivo ou aquele outro?

Devido a estas questões, buscamos uma maneira de definir o que seria a cooperação no dilema do prisioneiro. Para isso, propusemos novos classificadores que fossem além da proporção de cooperadores:

- (i) coeficiente de Gini;
- (ii) coeficiente do tamanho do maior agrupamento no sistema;
- (iii) entropia de agrupamentos de Shannon;
- (iv) entropia de agrupamentos de Tsallis; e
- (v) entropia de agrupamentos generalizada de Kullback-Leibler.

Assim, o objetivo inicial de explorar os efeitos das estratégias na cooperação, foi ampliado e passou a incluir a tarefa de definir o que seria um sistema cooperativo e buscar uma forma clara de classificar um sistema como cooperativo ou desertor. Resolvida essa questão sobre a cooperação, poderíamos voltar ao objetivo inicial e comparar a efetividade das estratégias em promover a cooperação. Quando sugerimos estes classificadores, não estávamos cientes da problemática acerca da distribuição dos resultados que compõem a média.

A seguir, apresentamos os resultados dos classificadores, com o desvio padrão, para um sistema em que os agentes jogam o dilema do prisioneiro discreto com a estratégia darwiniana. Trataremos duas conectividades: $k = 9$ (sem auto-interação), ilustrados na Fig. 4.20 e $k = 9$ (com auto-interação), na Fig. 4.22. Os resultados apresentados nos gráficos destas figuras são valores de média configuracional: $\tilde{x} = (1/M) \sum_{m=1}^M x_m$ e os valores das entropias estão escalonados entre zero e um: $\hat{x} = \frac{\tilde{x} - \min(\tilde{x})}{\max(\tilde{x} - \min(\tilde{x}))}$. Fazemos uma análise descritiva do comportamento destes classificadores e apontamos a relação entre eles.

As análises são baseadas em valores médios, como se os sistemas fossem ergódicos. No entanto, o desvio padrão elevado nos remete ao estudo sobre a média

configuracional de $\langle \rho_\infty \rangle$ e pensamos se os valores médios dos classificadores são obtidos de valores extremos ou de valores distribuídos ao redor da média.

Analisando os histogramas dos classificadores (omitidos) verificou-se que aparecem distribuições de extremos e ao redor da média. No entanto, não podemos dizer que tais distribuições são análogas ao que se observa em $\langle \rho_\infty \rangle$, ou consequência da distribuição de $\langle \rho_\infty \rangle$. Portanto, a análise não descritiva são especulações que ilustram como esses classificadores poderiam ser utilizados na análise do sistema.

Para resolver esta questão, seria necessário analisar os resultados dos classificadores em grupos com os resultados correspondentes a cada extremo de $\langle \rho_\infty \rangle$, quando esta tiver uma distribuição de extremos (não ergódica). Por isso, a análise dos resultados dos classificadores fica comprometida, pois não é possível pensar nos resultados como se o sistema fosse ergódico. Se considerássemos cada extremo da distribuição, em separado, as conclusões a que chegaríamos só se aplicariam aquele extremo, em particular, e não ao sistema como todo.

Na Fig. 4.20, analisamos os sistemas sem auto-interação ($k = 8$). A Fig. 4.20a mostra o nível médio assintótico de cooperação $\langle \rho_\infty \rangle$, com seu desvio padrão. Relembrando, há uma diminuição de $\langle \rho_\infty \rangle$ em função do aumento de T , com várias transições de fase e um desvio padrão elevado.

A Fig. 4.20b mostra o ganho médio por interação normalizado: $\tilde{G} = \frac{1}{M} \sum_{m=1}^M \left(\frac{1}{NkT} \sum_{n=1}^N G_n \right)$. Seu valor acompanha a diminuição observada no nível de cooperação e também apresenta um grande desvio padrão.

Quando o valor do coeficiente de Gini é nulo, ocorre uma total igualdade entre os indivíduos e todos tem a mesma riqueza. Este resultado só aparece se todos os jogadores têm o mesmo ganho no decorrer do jogo, o que ocorre quando todos os jogadores têm o mesmo estado ou em uma configuração de cooperadores e desertores (dada uma tentação T) muito específica e pouco provável. Portanto, o coeficiente de Gini nulo não distingue o sistema como cooperador ou desertor.

Já o coeficiente de Gini unitário, significa que um único indivíduo possui toda a riqueza existente, e se for próximo de um, poucos indivíduos dividem toda a riqueza. Essa situação ocorre na presença de poucos cooperadores cercados de desertores. Desta forma, os desertores que exploram os cooperadores tem um ganho

Figura 4.20 – Resultados dos classificadores para um sistema com agentes sem auto-interação, com $k = 8$, utilizando estratégia darwiniana e vizinhança regular. Todos os resultados são as médias configuracionais de $M = 1000$ configurações distintas. Os valores de entropia estão escalonados entre zero e um (u.e., unidade escalonada, adimensional). (continua)

(j) entropia de Tsallis de agrupamentos (Sq) (k) entropia de Tsallis de agrupamentos cooperadores (Sqc) (l) entropia de Tsallis de agrupamentos desertores (Sqd) (m) entropia de Kullback-Leibler generalizada - cooperadores/desertores (SKc) (n) entropia de Kullback-Leibler generalizada - desertores/cooperadores (SKd)

Figura 4.20 – Continuação. Os valores de entropia estão escalonados entre zero e um (u.e., unidade escalonada, adimensional).

maior, acumulando a riqueza entre eles. A primeira utilização do coeficiente de Gini para mensurar a distribuição de riqueza/renda foi apresentada na Ref. [57].

A Fig. 4.20c mostra o coeficiente de Gini do ganho acumulado durante todo o jogo, ou seja, a riqueza de cada jogador. O valor do coeficiente de Gini é próximo de zero para todo o intervalo de tentação. Em $T \sim 1,18$, há uma diminuição/aumento na igualdade/desigualdade na distribuição da riqueza, mas este cenário se mantém apenas até $T \sim 1,20$, quando seu valor diminui para um patamar levemente superior ao observado antes do aumento. Em $T \sim 1,22$ e $T \sim 1,35$ aparecem novos degraus que reduzem o valor do coeficiente de Gini e o levam para valores praticamente nulos, que voltam a oscilar em $T \sim 1,4$ e $T \sim 1,5$, mas essas oscilações são menos significantes que as anteriores. O desvio padrão tem diferentes valores em cada região. Seus maiores valores ocorrem na região $1 \leq T \leq 1,18$.

A Fig. 4.20d traz o coeficiente de Gini do ganho obtido na rodada (renda ao invés da riqueza). Na região $1,00 \leq T \leq 1,75$ o coeficiente de Gini se anula, significando que a distribuição da renda é igual entre todos os jogadores. Acima de $T = 1,75$, os valores do coeficiente de Gini são elevados, o que indica que há alguns poucos jogadores que concentram toda a renda do sistema. Pela presença dos cooperadores, conforme mostrou a Fig. 4.18a com $T = 2,00$, aliado ao fato de que a maioria das configurações é desertora (vide Fig. 4.20a) podemos dizer que os jogadores que acumulam a renda na rodada são os desertores que exploram os cooperadores que resistiram no sistema. Esses desertores, que têm um ganho maior que os demais jogadores, não modificam o ganho médio dos jogadores (vide Fig. 4.20b), porque são em pequeno número. Também não modificam o coeficiente de Gini da riqueza acumulada, pois eles devem formar um *glider* (ver Seção 4.1) dentro do sistema, ou seja, a cada rodada, um jogador diferente joga como cooperador e provê um ganho mais elevado aos seus vizinhos, mas na rodada seguinte volta a ser desertor. Assim, vai se modificando o jogador que teve o maior ganho na rodada e não há um acúmulo de riqueza.

Como vimos nos histogramas de $\langle \rho_\infty \rangle$ para o DPD (Figs. 4.18 e 4.19), há configurações cooperadoras ou desertoras, tornando infrutífera qualquer especulação acerca do regime estacionário baseada no coeficiente de Gini, quando ele for nulo. A única observação que poderíamos fazer sobre o sistema em que os agentes jogam

o DP, utilizando o coeficiente de Gini para a distribuição da renda, seria quando ele for igual ou próximo de um e tivermos uma única configuração do sistema. Neste caso, podemos dizer que existem muitos desertores e alguns cooperadores.

O coeficiente de agrupamento cooperador (CAC) e desertor (CAD), Figs. 4.20e e 4.20f, estão diretamente relacionados com $\langle \rho_\infty \rangle$ e entre si. Para o CAC a região $1 \leq T \leq 1,20$ merece maior destaque, pois, em $T = 1,00$, o maior agrupamento não é igual à proporção de cooperadores do sistema. Isso sugere a presença de alguns desertores que permeiam agrupamentos cooperadores menores, além do principal agrupamento, que é formado por mais de 50% dos jogadores do sistema, mas estes desertores não se agrupam em único grande grupo, segundo mostra o CAD, pois CAC e CAD não são complementares.

No intervalo $1,01 \leq T \leq 1,18$, o CAC ainda apresenta uma diferença com $\langle \rho_\infty \rangle$, revelando que os cooperadores não estão reunidos em um único agrupamento. Entre $1,18 \leq T \leq 1,20$, ocorre um aumento simultâneo do CAC e CAD, mas os valores do CAC e CAD são não complementares, sugerindo a presença de mais de um agrupamento desertor. Para $T > 1,75$, o sistema pode ser considerado plenamente desertor, onde inexistem os cooperadores e os desertores controlam todo o sistema. O desvio padrão é grande em toda a região na qual os valores de CAC não são nulos. Para o CAD, enquanto $T < 1,75$, o desvio padrão é grande e para $T > 1,75$, $CAD \sim 1$ e o desvio padrão é praticamente nulo.

Antes de analisar a entropia, lembramos que foi considerada entropia nula no caso da ausência de grupos cooperadores/desertores quando estivermos analisando a distribuição dos agrupamentos cooperadores/desertores ou ambos. Os valores de entropia foram escalonados para permitir uma comparação qualitativa entre eles.

Iniciamos pela entropia de Shannon de agrupamentos (vide Fig. 4.20g). Seu valor máximo ocorre com $T = 1,00$ e, logo em seguida, em $T = 1,01$ sofre uma redução brusca atingindo um patamar, que se estende até $T = 1,15$, quando sofre uma nova redução formando um mínimo local entre $1,17 \leq T \leq 1,20$. Em $T = 1,20$, a entropia aumenta e entre $1,20 \leq T \leq 1,60$ existem patamares, que se diferenciam pela diminuição do valor da entropia na forma de degraus. Entre $1,60 \leq T \leq 2,00$ a entropia tem seu valor mínimo. O desvio padrão tem seus valores mais elevados entre $1,00 \leq T \leq 1,17$ e diminui a medida que aumenta a tentação, junto com a

diminuição do valor da entropia. A entropia de Shannon de agrupamentos cooperadores (Sc) e seu desvio padrão têm o mesmo comportamento observado na entropia de Shannon de agrupamentos (S), como mostra a comparação das Fig. 4.20h e 4.20g, respectivamente. Como os valores apresentados são escalonados, não é correto dizer que os valores de Sc são menores que os de S .

A Fig. 4.20i, mostra os resultados da entropia de Shannon de agrupamentos desertores, que parece se diferenciar bastante das entropias de agrupamentos e agrupamentos cooperadores. No entanto, isso ocorre porque há um mínimo mais pronunciado entre $1,17 \leq T \leq 1,20$ com desvio padrão nulo, mas o comportamento da curva é similar, mas, nas regiões fora do mínimo local, o desvio padrão é maior do que o das outras entropias de Shannon.

A região de mínimo local coincide com o aumento da coeficiente de Gini para o ganho acumulado (Fig. 4.20c), que é explicado pela coexistência de agrupamentos de desertores e cooperadores, o que pode ser confirmado pelos CAC e CAD dessa região. Nesse caso o desertor da borda tem metade da vizinhança cooperadora e a outra metade desertora, assim ele explora esses cooperadores e tem um ganho elevado e o cooperador da borda tem um ganho reduzido. A medida que nos afastamos da borda, forma-se um gradiente nos valores de ganhos obtidos pelos jogadores. Se essa configuração se mantiver por mais tempo, ocorre uma diferenciação nos ganhos acumulados pelos jogadores que leva ao aumento do coeficiente de Gini, conforme observado. A entropia de Shannon para agrupamentos nula indica que há um único agrupamento ou então, todos os grupos têm o mesmo tamanho.

As Figs. 4.20j, 4.20k e 4.20l mostram a entropia de Tsallis com $q = \{-2; -1/2; 1/2; 2\}$. A Fig. 4.20j, mostra a entropia de Tsallis para agrupamentos cooperadores e desertores. O escalonamento feito nos dados, comprime o gráfico, mas dos resultados sem o escalonamento (omitido) verificamos que as curvas têm o mesmo comportamento presente na entropia de Shannon, tanto para os valores da entropia em si, quanto para seu desvio padrão.

Na Fig. 4.20k, mostramos a entropia de Tsallis de agrupamentos cooperadores, nela ocorre o problema da compressão pelo escalonamento para $q = -2$ e mesmo assim não aparecem diferenças significativas. Para os demais valores de q , os resultados não se diferenciam do observado com a entropia de Shannon (vide Fig. 4.20g).

Para $q = \{-2; 2\}$ não surge o mínimo local entre $1, 17 \leq T \leq 1, 20$. A Fig. 4.20l mostra a entropia de Tsallis de agrupamentos desertores, que tem algumas peculiaridades. Para $q = -2$, o que aparentemente seria uma maior entropia em quase todo o intervalo, pode ser explicada pelo mínimo observado em $T = 1, 20$, um valor negativo com grande valor absoluto, que pelo escalonamento, manteve toda a curva em um nível mais elevado. Para $q = \{-1/2; 2\}$ ocorre o mesmo comportamento da entropia de Shannon (vide Fig. 4.20g). Somente para $q = 1/2$, houve uma inversão do comportamento da curva após o mínimo local. Neste caso, a entropia aumenta em degraus em função do aumento da tentação, enquanto que, nas demais ocorria um breve aumento seguido da diminuição em degraus.

As Figs. 4.20m e 4.20n trazem a entropia de Kullback-Leibler generalizada para $q = \{-2; -1/2; 1/2; 2\}$. A Fig. 4.20m apresenta a entropia de Kullback-Leibler generalizada da relação cooperadores/desertores. Com $q = \{-2; -1/2\}$, ocorre a compressão devida ao escalonamento, mas em todos os casos o comportamento se assemelha ao caso peculiar da entropia de Tsallis para agrupamentos desertores com $q = 1/2$ (vide Fig. 4.20l). Isto é, há um crescimento da entropia a medida que aumentamos a tentação e os acréscimos acontecem na forma de degraus. Mas, aqui a entropia começa em seu valor mínimo e aumenta durante todo intervalo de tentação.

Na Fig. 4.20n, apresentamos a entropia de Kullback-Leibler generalizada da relação desertores/cooperadores. As curvas para $q = \{-2; 1/2; 2\}$ tem o comportamento da entropia de Shannon (Fig. 4.20g), enquanto que, para $q = -1/2$ o comportamento lembra o da Fig. 4.20l para $q = 1/2$. Entretanto, o aumento da entropia após o mínimo local é maior que o observado com a entropia de Shannon e leva o valor da entropia de Kullback-Leibler ao máximo se mantendo constante até o final do intervalo de tentação com desvio padrão nulo.

Nas Figs. 4.21 e 4.23, mostramos $\langle \rho_\infty \rangle$ e os classificadores agrupados de acordo com suas propriedades para facilitar a visualização dos classificadores e compará-los entre si. Nestes gráficos, podemos observar que os classificadores estão relacionados com as oscilações de $\langle \rho_\infty \rangle$. Pretendíamos interpretar o que cada classificador poderia dizer a respeito da cooperação ou deserção no sistema. Esta análise parece inadequada quando observamos a distribuição dos resultados utilizados no cálculo

Figura 4.21 – Resultados do classificador para sistemas sem auto-interação ($k = 8$).

das médias.

Na Fig. 4.22, apresentamos os gráficos equivalentes aos da Fig. 4.20, mas para sistemas com auto-interação ($k = 9$). A análise descritiva destes resultados não acrescenta novas informações, além das obtidas no caso sem auto-interação, pois a diferença entre ambas seria o deslocamento em T de onde ocorre cada região. As diferenças que podem ser vistas nos valores dos gráficos equivalentes estão relacionadas, principalmente, às mudanças de $\langle \rho_\infty \rangle$ causadas pelo maior número de

transições que existem para $k = 9$, do que para $k = 8$.

A Fig. 4.23 é equivalente à Fig. 4.21 mostrando $\langle \rho_\infty \rangle$ e os classificadores agrupados de acordo com suas propriedades, para os sistemas com auto-interação ($k = 9$). Por causa das distribuições de extremos que surgem nos resultados dos classificadores, não é possível nos aprofundarmos na análise dos classificadores.

Assim, finalizamos a análise dos resultados e partimos para o capítulo final da tese com as conclusões.

Figura 4.22 – Resultados dos classificadores para um sistema com agentes com auto-interação, com $k = 9$, utilizando estratégia darwiniana e vizinhança regular. Todos os resultados são as médias configuracionais de $M = 1000$ configurações distintas. Os valores de entropia estão escalonados entre zero e um (u.e., unidade escalonada, adimensional). (continua)

(j) entropia de Tsallis de agrupamentos (S_q) (k) entropia de Tsallis de agrupamentos cooperadores (S_{qc}) (l) entropia de Tsallis de agrupamentos desertores (S_{qd}) (m) entropia de Kullback-Leibler generalizada - cooperadores/desertores (SK_c) (n) entropia de Kullback-Leibler generalizada - desertores/cooperadores (SK_d)

Figura 4.22 – Continuação. Os valores de entropia estão escalonados entre zero e um (u.e., unidade escalonada, adimensional).

Figura 4.23 – Resultados dos classificadores para sistemas com auto-interação ($k = 9$).

Conclusões e perspectivas

“ [Audience in a conference]: ‘They will never make a machine to replace the human mind - it does many things which no machine could ever do. But of course, a mere machine can’t really think, can it?’

[Neumann]: ‘You insist that there is something a machine cannot do. If you will tell me precisely what it is that a machine cannot do, then I can always make a machine which will do just that!’ ”¹

Resposta de J. von Neumann em uma conferência sobre computadores em Princeton no ano de 1948.

E. T. Jaynes no livro *Probability Theory: The Logic of Science* - 1996.

Finalizando, fazemos uma compilação dos resultados desta tese e suas implicações. Iniciamos o estudo focados nas perspectivas abertas pelo dilema do prisioneiro contínuo. Esperávamos compreender o efeito da personalidade dos jogadores e agregar elementos como a inconsistência intertemporal [112], a assimetria na tomada de decisão frente a ganhos e perdas [15]. Para isso, utilizamos o dilema do prisioneiro contínuo na estrutura que propusemos [46, 47, 48] com novas estratégias.

Porém, os resultados iniciais utilizando as novas estratégias eram muitos semelhantes entre si e nos obrigaram a buscar alternativas que pudessem diferenciá-los

¹“ [Platéia em uma conferência]: ‘Nunca será feita uma máquina para substituir a mente humana - ela faz muitas coisas as quais nenhuma máquina jamais poderia fazer. Mas, claro, uma mera máquina não pode pensar, não é?’

[Neumann]: ‘Vocês insistem que há algo que uma máquina não possa fazer. Se você me contar precisamente o que é que uma máquina não pode fazer, então eu sempre posso fazer uma máquina que fará exatamente isso!’ ”

antes de pensar em mudanças nas estratégias propostas. Como o nível médio de cooperação parecia insuficiente para esta tarefa, realizamos uma análise visando a formação dos valores de cooperação e a organização dos jogadores.

Além da análise numérica, obtivemos uma formulação e soluções analíticas importantes. Para o dilema do prisioneiro discreto adotando a estratégia darwiniana, obtivemos a função analítica do nível médio de cooperação em função da tentação. No entanto, esta função é válida somente nos instantes iniciais de uma simulação numérica, pois os vizinhos são fixos e a estratégia é baseada em um mecanismo de cópia. Este processo leva à formação de agrupamentos logo após as primeiras rodadas, o que invalida a premissa de homogeneidade do sistema.

Também definimos uma função que calcula valores de tentação crítica para o dilema do prisioneiro discreto, tanto para estratégia darwiniana quanto para a pavloviana. Estas transições são suprimidas no DPC, portanto não há variações abruptas de ρ . Os valores onde ocorrem as transições foram confirmados numericamente para a estratégia darwiniana no DPD e também mostramos que elas não aparecem no DPC.

Quanto à função de densidade de probabilidade utilizada na distribuição do nível médio inicial de cooperação, observamos que, distribuições com uma maior densidade nos extremos influenciam os resultados sendo tendenciosas. Por isso, optamos por adotar a distribuição uniforme que, apesar de simples, não inibiu qualquer comportamento observado com as demais distribuições. Contudo, a utilização do nível de cooperação contínuo foi responsável pelo aumento da cooperação quando comparado ao dilema do prisioneiro discreto.

As conexões dos jogadores tem um importante papel na emergência/manutenção da cooperação. Quando utilizamos a vizinhança aleatória, a cooperação se extinguiu e conectividades elevadas levam à diminuição do nível médio de cooperação. Por isso, utilizamos a vizinhança fixa como padrão nas simulações.

A busca por uma característica essencial na formação da cooperação, através do uso de diferentes estratégias, não produziu resultados conclusivos nesse aspecto. Isso porque, não foi possível inferir qual o papel da personalidade e/ou da aspiração de ganho do jogador na emergência da cooperação com as estratégias utilizadas. Apesar das estratégias terem apresentado diferentes regimes transientes, o estado

assintótico de todas as estratégias é muito similar. Os resultados das estratégias propostas não tiveram qualquer destaque em relação à estratégia darwiniana. A explicação mais plausível para isso é o mecanismo de cópia do estado de outro jogador.

Uma alternativa nessa direção seria o uso de outro paradigma, diferente da cópia, na mudança do estado do jogador. Algo em que os processos de tomada de decisão e mudança de estado sejam intrínsecos de cada jogador, sem cópias. Talvez assim, a personalidade do jogador tivesse mais destaque. Vimos que apesar da forma de subjetividade adotada ser pouco realista e privilegiar a cooperação, não houve um aumento da mesma nos resultados com esta estratégia. Uma possível correção seria alterar a subjetividade, afim de sanar a anomalia de que um desertor está sempre insatisfeito com seu ganho, pois desta maneira a deserção está subdimensionada.

As estratégias baseadas no ganho consideraram o ganho total do jogador comparado a uma aspiração. Estas estratégias e o uso do coeficiente de Gini, tiveram inspiração na Ref. [57], que apresenta algumas estratégias para o jogo do ultimato.

Considerando um único sistema e analisando conjuntamente os coeficientes de agrupamento cooperador e desertor e o nível médio de cooperação, podemos ter alguns indícios da fragmentação dos agrupamentos. Quando o valor do coeficiente de agrupamento, cooperador ou desertor, é aproximadamente unitário, há no sistema um grande agrupamento de cooperadores ou desertores, respectivamente, e podemos afirmar isso sem sabermos o nível médio de cooperação. Se a soma dos coeficientes de agrupamento cooperador e desertor for aproximadamente unitário, o sistema tem dois agrupamentos principais de cooperadores e desertores e o nível médio de cooperação será equivalente ao maior destes grupos. Se um dos coeficientes for grande e outro aproximadamente nulo, há um grande agrupamento deste tipo e os demais grupos estão fragmentados.

Apesar da análise feita pelo coeficiente de agrupamento, quando pensamos em seus valores médios, surge o problema da distribuição de valores extremos que compõem as médias. Por exemplo, não sabemos se um valor médio de coeficiente de agrupamento igual a 0,8 é equivalente a todos os sistemas com um maior grupo concentrando aproximadamente 80% dos jogadores em um dado estado, ou se existe 80% dos sistemas totalmente em um estado e 20% no outro. Se a mediana for con-

siderada, os efeitos dos eventos extremos nos resultados são fortemente atenuados, pois a mediana não é tão sensível quanto a média à presença dos valores extremos.

O coeficiente de Gini igual à zero acontece em sistemas totalmente desertores ou cooperadores. Logo, não serve como qualificador de cooperação, quando utilizado individualmente. Entretanto, se o coeficiente de Gini for aproximadamente unitário, significa que se trata de um caso em que poucos jogadores possuem praticamente toda a riqueza existente.

Se estivermos calculando o coeficiente de Gini da riqueza acumulada, um valor próximo ao unitário pode ocorrer na presença de *fingers* desertores, ou seja, alguns desertores exploram os cooperadores repetidamente, acumulando um ganho maior que os demais jogadores. No caso do coeficiente de Gini para a distribuição de renda, seu valor é aproximadamente unitário, quando o sistema é de maioria desertora, mas um *glider* cooperador persiste no sistema, ele produz um ganho elevado para os poucos desertores que exploram esse cooperador, ocasionando a distribuição irregular de renda. Novamente, lembramos que tais interpretações do coeficiente de Gini se aplicam para um sistema em separado e não para seus valores médios.

Dentre as entropias utilizadas, vimos que o aumento da complexidade no cálculo da entropia não aperfeiçoou seus resultados. Da mesma forma que o coeficiente de Gini, a entropia também não distingue sistemas com maioria cooperadora ou desertora, pois a entropia é nula em ambos os extremos. Assim, a entropia nula conciliada ao desvio padrão elevado de ρ pode ser um indicativo de que os resultados que compõem a média têm uma distribuição de extremos.

No entanto, configurações regulares também podem gerar entropia nula. Para se assegurar sobre os resultados de entropia, seria de grande valia um estudo mais detalhado sobre qual o valor de entropia de cada configuração, ou padrões de configuração. No entanto, tais conclusões só surgiram no final do trabalho, não tendo sido possível realizar tal estudo.

Em sistemas que apresentam caos, pequenas variações nas condições iniciais provocam grandes alterações nos resultados finais. Assim, conhecer as condições iniciais é fundamental para uma boa análise de sistemas caóticos. Nas simulações numéricas, todos os sistemas tem o mesmo nível médio inicial de cooperação, mas a configuração é muito importante para o resultado final. Seria possível criar algum

tipo de classificação de similaridade de configurações iniciais? Talvez a entropia inicial do sistema se preste a esta função de agrupar os sistemas, e talvez, sirvam para separar aqueles que apresentarão maioria de cooperadores daqueles com maioria de desertores. Isso poderia evitar a distribuição de extremos que obtivemos.

Quanto ao nível médio de cooperação dos sistemas, os resultados numéricos nos alertaram para o fato de que o dilema do prisioneiro carece de uma definição precisa sobre o que é cooperação em sistemas de muitos agentes. Em especial, no caso do dilema do prisioneiro discreto, a análise do nível médio de cooperação deve ser mais cuidadosa para fornecer resultados realmente confiáveis.

Para ilustrar, imagine um jogo em que um sujeito ganha uma unidade monetária cada vez que um dado for lançado e apresentar as faces de um a três e não ganha nada quando as faces de quatro a seis aparecerem. Se esse jogo se repetir cem vezes, seu ganho esperado é 50 unidades monetárias. Um segundo jogo também está disponível, mas, ao invés de um dado, é lançada uma moeda e se a face aparente for cara o sujeito ganha uma unidade monetária e não ganha nada caso contrário. Neste segundo jogo, se jogado por cem vezes, seu ganho esperado também é de 50 unidades monetárias. Ambos os jogos tem o mesmo ganho esperado, mas qual deles parece mais atraente?

Agora, suponha cem sistemas de mil jogadores jogando o dilema do prisioneiro. Em uma situação, os cem sistemas tem 80% dos seus jogadores como cooperadores e 20% desertores. Na outra, 80% dos sistemas são totalmente cooperadores e 20% totalmente desertores. Qual seria o seu ganho esperado em cada uma das situações? Elas são equivalentes? Se o sujeito jogar com todos os jogadores de todos os sistemas, então, seu ganho esperado será o mesmo. Mas, se ele fizer apenas k jogadas em cada um dos cem sistemas, qual cenário seria mais vantajoso?

Essas situações nos levaram a questionarmos a maneira como o dilema do prisioneiro em sistemas com muitos agentes tem sido abordado. A falta de qualquer menção quanto à distribuição dos resultados ou seu desvio padrão parece ser uma falha que os estudos vêm apresentando reiteradamente e não há qualquer sinal que aponte uma mudança desse cenário. Talvez fosse conveniente uma nova abordagem dos resultados que levasse em conta eventos extremos, ou uma demonstração de que eles não são importantes.

Esse estudo foi bastante extenso e suscitou várias questões a respeito dos fundamentos de um dos principais jogos da teoria dos jogos, o dilema do prisioneiro. Os objetivos traçados para a tese foram contemplados. Apesar de alguns resultados não apresentarem diferenças significativas, eles serviram para evidenciar fundamentos que ainda não foram completamente esclarecidos na literatura. Talvez essas questões sobre o nível médio de cooperação, sugiram que grande parte dos estudos atuais na área poderia ser enquadrada como casos particulares de um caso mais geral, que ainda não foi bem posto.

Outro aspecto que é motivo de discussão é o uso da rede unidimensional. Entretanto, o uso de redes bidimensionais torna a análise de agrupamentos muito mais difícil. Seria necessário utilizar algoritmos de percolação para definir o tamanho dos agrupamentos (Hoshen-Kopelman). Como não sabíamos a efetividade dos classificadores propostos, exploramos somente a rede unidimensional.

Como sugestões de estudos futuros na área que poderiam derivar diretamente desta tese, apresentamos a coevolução de estratégias no DPC, novas estratégias em que os jogadores estimam a probabilidade de obter cooperação do oponente numa dada jogada, estratégias que contenham inconsistências na avaliação de risco na hora de ganhar e perder (assimetria para valores positivos e negativos) e inconsistência intertemporal (avaliação de ganhos futuros). Agregando informação ao sistema, poderiam ser utilizados jogadores com reputação, a qual pode ser informada via histórico (com ou sem ruído). A questão da informação poderia ser encarada com naturalidade em sistemas onde os jogadores tem mobilidade, pois eles receberiam informações a respeito dos novos vizinhos pelo histórico das jogadas ou por fofocas.

Referências²

- [1] J. Neumann e O. Morgenstern, *Theory of games and economic behavior*. New York: Princeton University Press, 1947. 1
- [2] M. J. Osborne e A. Rubinstein, *A course in game theory*. MIT Press, 1994. 1
- [3] L. M. Wahl e M. A. Nowak, “The continuous prisoner’s dilemma: I. linear reactive strategies”, *J. Theor. Biol.*, vol. 200, pp. 307–321, 1999. 2, 20
- [4] L. M. Wahl e M. A. Nowak, “The continuous prisoner’s dilemma: II. linear reactive strategies with noise”, *J. Theor. Biol.*, vol. 200, pp. 323–338, 1999. 2, 20
- [5] M. Ifti, T. Killingback e M. Doebeli, “Effects of neighbourhood size and connectivity on the spatial continuous prisoner’s dilemma”, *J. Theor. Biol.*, vol. 231, pp. 97–106, 2004. 2, 20, 122
- [6] T. Killingback, M. Doebeli e N. Knowlton, “Variable investment, the continuous prisoner’s dilemma, and the origin of cooperation”, *P. Roy. Soc. Lond. B Bio.*, vol. 266, no. 1430, pp. 1723–1728, 1999. 2, 20
- [7] T. Killingback e M. Doebeli, “The continuous prisoner’s dilemma and the evolution of cooperation through reciprocal altruism with variable investment”, *Am. Nat.*, vol. 160, no. 4, pp. 421–438, 2002. 2, 20, 22
- [8] S. Le e R. Boyd, “Evolutionary dynamics of the continuous iterated prisoner’s dilemma”, *J. Theor. Biol.*, vol. 245, pp. 258–267, 2007. 2, 20

²De acordo com o estilo LMSC (baseado no estilo IEEE Transactions - Electrical and Electronics Engineers).

-
- [9] P. Darew e X. Yao, “Co-evolution in iterated prisoner’s dilemma with intermediate levels of cooperation: application to missile defense”, *Int. J. Knowledge-Based Intell. Eng. Syst.*, vol. 4, pp. 83–107, December 2002. 2, 3, 20
- [10] R. Alonso-Sanz, M. C. Martín e M. Martín, “The effect of memory in the spatial continuous-valued prisoner’s dilemma”, *Int. J. Bifurcat. Chaos*, vol. 11, pp. 2061–2083, 2000. 2, 20
- [11] R. D. Sommerfeld, H. J. Krambeck, D. Semmann e M. Milinski, “Multiple choices and reputation in multiagent interactions”, *PNAS*, vol. 104, pp. 17435–17440, 2007. 2, 20
- [12] S. Oskamp e C. Kleinke, “Amount of reward as a variable in the prisoner’s dilemma game”, *J. Pers. Soc. Psychol.*, vol. 16, pp. 133–40, 1970. 3
- [13] D. Stauffer, “Introduction to statistical physics outside physics”, *Physica. A*, vol. 336, pp. 1–5, 2004. 3
- [14] J. P. Bouchaud, “An introduction to statistical finance”, *Physica. A*, vol. 313, pp. 238–251, 2002. 3
- [15] C. Anteneodo, C. Tsallis e A. S. Martinez, “Risk aversion in economic transactions”, *Europhys. Lett.*, vol. 59, no. 5, pp. 635–641, 2002. 3, 25, 32, 45, 137
- [16] P. Bak, S. F. Norrelykke e M. Shubik, “Dynamics of money”, *Phys. Rev. E*, vol. 60, p. 2528, 1999. 3
- [17] G. Hardin, “The tragedy of the commons”, *Science*, vol. 162, pp. 1243–1248, 1968. 3
- [18] B. Brembs, “Chaos, cheating and cooperation: potential solutions to the prisoner’s dilemma.”, *Oikos*, vol. 76, pp. 14–24, 1996. 3, 22
- [19] O. Leimar e A. H. Axén, “Strategic behavior in an interespecific mutualism: interactions between lycaenid larvae and ants”, *Anim. Behav.*, vol. 46, pp. 1177–1182, 1993. 3, 23

-
- [20] M. Milinski, “On cooperation in sticklebacks”, *Anim. Behav.*, vol. 40, pp. 1190–1191, 1990. 3, 23
- [21] J. C. Reboreda e A. Kacelnik, “On cooperation, tit-for-tat and mirrors”, *Anim. Behav.*, vol. 40, pp. 1188–1189, 1990. 3, 23
- [22] L. A. Dugatkin, “Dynamics of the tit-for-tat strategy during predator inspection in the guppy (*Poecilia reticulata*)”, *Behav. Ecol. Sociobiol.*, vol. 29, pp. 127–132, 1991. 3, 23
- [23] L. A. Dugatkin, “Predator inspection, tit-for-tat and shoaling: a comment on Masters & Waite.”, *Anim. Behav.*, vol. 41, pp. 898–899, 1991. 3, 23
- [24] L. A. Dugatkin e M. Alfieri, “Guppies and the tit-for-tat strategy: preference based on past interaction”, *Behav. Ecol. Sociobiol.*, vol. 28, pp. 243–246, 1991. 3, 23
- [25] G. F. Turner e R. L. Robinson, “Milinski’s tit-for-tat hypothesis: do fish preferentially inspect in pairs?”, *Anim. Behav.*, vol. 43, pp. 677–678, 1992. 3, 23
- [26] E. A. Fischer, “Egg trading in the Chalk Bass, *Serranus-Tortugarum*, a simultaneous hermaphrodite”, *J. Comp. Ethol.*, vol. 66, no. 2, pp. 143–151, 1984. 3, 24
- [27] E. A. Fischer, “Simultaneous hermaphroditism, tit-for-tat, and the evolutionary stability of social systems”, *Ethol. Sociobiol.*, vol. 9, pp. 119–136, 1988. 3, 24
- [28] R. C. Connor, “Egg-trading in simultaneous hermaphrodites: an alternative to tit-for-tat”, *J. Evol. Biol.*, vol. 5, pp. 523–528, 1992. 3, 24
- [29] G. S. Wilkinson, “Reciprocal food sharing in the vampire bat”, *Nature*, vol. 308, pp. 181–184, 1984. 3, 25
- [30] R. Axelrod, “Effective choice in the prisoner’s dilemma”, *J. Conflict Resolut.*, vol. 24, pp. 3–25, 1980. 3, 17

-
- [31] R. Axelrod, “More effective choice in the prisoner’s dilemma”, *J. Conflict Resolut.*, vol. 24, pp. 379–403, 1980. 3, 17
- [32] R. Axelrod e W. D. Hamilton, “The evolution of cooperation”, *Science*, vol. 211, pp. 1390–1396, 1981. 3, 13, 17
- [33] R. Axelrod, *The evolution of cooperation*. New York: Basic Books, 1984. 3, 17
- [34] M. A. Nowak e R. May, “Evolutionary games and spatial chaos”, *Nature*, vol. 359, pp. 826–829, 1992. 3, 4, 13, 18, 29, 40, 84, 103
- [35] M. A. Nowak e R. M. May, “The spatial dilemmas of evolution”, *Int. J. Bifurcat. Chaos*, vol. 3, pp. 35–78, 1993. 3, 4, 18, 29, 32
- [36] M. A. Nowak e K. Sigmund, “A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner’s dilemma game”, *Nature*, vol. 364, pp. 56–58, 1993. 4, 41
- [37] E. Lieberman, C. Hauert e M. A. Nowak, “A strategy of win-stay and lose-shift that outperforms tit-for-tat in the prisoner’s dilemma game”, *Nature*, vol. 433, pp. 312–316, 2005. 4, 29, 41
- [38] H. Fort e S. Viola, “Spatial patterns and scale freedom in prisoner’s dilemma cellular automata with pavlovian strategies”, *J. Stat. Mech.-Theory Exp.*, vol. 1, p. P01010, 2005. 4, 40, 41
- [39] M. Posch, “Win-stay, lose-shift strategies for repeated games - memory lenght, aspiration levels and noise”, *J. Theor. Biol.*, vol. 198, pp. 183–195, 1999. 4, 41
- [40] L. A. Imhof, D. Fudenberg e M. A. Nowak, “Tit-for-tat or win-stay, lose-shift?”, *J. Theor. Biol.*, vol. 247, pp. 574–580, 2007. 4, 41
- [41] D. Kraines e V. Kraines, “Pavlov and the prisoner’s dilemma”, *Theor. Decis.*, vol. 26, pp. 47–79, 1989. 4, 41
- [42] T. Platkowski, “Enhanced cooperation in prisoner’s dilemma with aspiration”, *Appl. Math. Lett.*, vol. 22, no. 12, pp. 1161–1165, 2009. 4, 41

-
- [43] M. Dresher, *The mathematics of games of strategy: theory and applications*. Englewood Cliffs and NJ: Prentice-Hall, 1961. 12
- [44] W. Poundstone, *Prisoner's dilemma*. New York: Doubleday, 1992. 12, 15
- [45] S. Oskamp, "Effects of programmed strategies on cooperation in the prisoner's dilemma and other mixed-motive games", *J. Conflict Resolut.*, vol. 15, pp. 225–259, 1971. 14
- [46] M. A. Pereira, A. S. Martinez e A. L. Espíndola, "Prisoner's dilemma in one-dimensional cellular automata: visualization of evolutionary patterns", *Int. J. of Modern Phys. C*, vol. 19, pp. 187–201, 2008. 15, 19, 20, 29, 30, 91, 103, 106, 137
- [47] M. A. Pereira, A. S. Martinez e A. L. Espíndola, "Exhaustive exploration of prisoner's dilemma parameter space in one-dimensional cellular automata", *Braz. J. of Phys.*, vol. 38, pp. 65–69, 2008. 15, 19, 20, 29, 30, 103, 106, 137
- [48] M. A. Pereira e A. S. Martinez, "Pavlovian prisoner's dilemma: Analytical results, the quasi-regular phase and spatio-temporal patterns", *J. Theor. Biol.*, vol. 265, no. 3, pp. 346 – 358, 2010. 15, 19, 20, 22, 29, 30, 34, 103, 106, 137
- [49] D. Helbing e S. Lozano, "Phase transitions to cooperation in the prisoner's dilemma", *Phys. Rev. E*, vol. 81, p. 057102(4), 2010. 15, 20
- [50] H. E. Stanley, *An introduction to phase transitions and critical phenomena*. Oxford: Oxford University Press, 1971. 15, 48
- [51] L. B. Lave, "An empirical description of the prisoner's dilemma game", *Support RAND Research*, vol. 2091, pp. 1–21, 1960. 16
- [52] R. Boyd e J. P. Lorberbaum, "No pure strategy is evolutionary stable in the prisoner's dilemma game", *Nature*, vol. 327, pp. 58–59, 1987. 17
- [53] R. O. S. Soares e A. S. Martinez, "The geometrical patterns of cooperation evolution in the spatial prisoner's dilemma: an intra-group model", *Physica. A*, vol. 369, pp. 823–829, 2006. 19

-
- [54] O. Durán e R. Mulet, “Evolutionary prisoner’s dilemma in random graphs”, *Physica. A*, vol. 208, pp. 257–265, 2005. 19, 29, 49
- [55] R. da Silva, A. L. C. Bazzan, A. T. Baraviera e S. R. Dahmen, “Emerging collective behavior and local properties of financial dynamics in a public investment game”, *Physica. A*, vol. 371, pp. 610–626, 2006. 20
- [56] R. da Silva, “The public good game on graphs: Can the pro-social behavior persist?”, *Braz. J. of Phys.*, vol. 38, pp. 74–80, 2008. 20
- [57] R. da Silva, G. A. Kellermann e L. d. C. Lamb, “Statistical fluctuations in population bargaining in the ultimatum game: Static and evolutionary aspects”, *J. Theor. Biol.*, vol. 258, pp. 208–218, 2009. 20, 70, 127, 139
- [58] S. Y. Chong e X. Yao, “Multiple choices and reputation in multiagent interactions”, *IEEE Trans. on Evol. Comput.*, vol. 11, pp. 689–711, December 2007. 21, 35
- [59] M. Milinski, “Predator inspection: cooperation or safety numbers?”, *Anim. Behav.*, vol. 43, pp. 679–680, 1992. 24
- [60] G. S. Wilkinson, “Reciprocal altruism in bats and other mammals”, *Ethol. Sociobiol.*, vol. 9, pp. 85–100, 1988. 24
- [61] D. Kahneman e A. Tversky, “Prospect theory: an analysis of decision under risk”, *Econometrica*, vol. 47, p. 263, 1979. 26
- [62] Z. Wu, J. Guan, X. Xu e Y. Wang, “Evolutionary prisoner’s dilemma game on Barabási-Albert scale-free networks”, *Physica. D*, vol. 379, pp. 672–680, 2007. 29
- [63] G. Abramson e M. Kuperman, “Social games in a social network”, *Phys. Rev. E*, vol. 63, no. 030901, pp. 1–4, 2001. 29
- [64] R. da Silva, A. T. Baraviera, S. R. Dahmen e A. L. C. Bazzan, “Dynamics of a public investment game: from nearest-neighbor lattices to small-world network”, *Lect. Notes Econ. Math.*, vol. 584, pp. 221–233, 2006. 29

-
- [65] S. Wolfram, “Statistical mechanics of cellular automata”, *Rev. Mod. Phys.*, vol. 55, pp. 601–644, 1983. 30
- [66] S. Wolfram, “Cellular automata”, *Science*, vol. 9, pp. 2–21, 1983. 30
- [67] R. da Silva e N. Alves Júnior, “Dynamic exponents of a probabilistic cellular automaton”, *Physica. A*, vol. 350, pp. 263–276, 2005. 30
- [68] S. Wolfram, “Minimal cellular automaton approximations to continuum systems”. Internet, 1986. 30
- [69] M. H. Vainstein, A. T. C. Silva e J. J. Arenzon, “Does mobility decrease cooperation?”, *J. Theor. Biol.*, vol. 244, no. 4, pp. 722–728, 2007. 33
- [70] W. Yu, “Mobility enhances cooperation in the presence of decision-making mistakes on complex networks”, *Phys. Rev. E*, vol. 83, p. 026105(6), 2011. 33, 122
- [71] J.-M. Luck, *Systèmes désordonnés unidimensionnels*. Aléa Saclay, 1992. 34
- [72] R. D. Luce e H. Raiffa, *Games and Decisions*. Wiley, 2nd ed., 1957. 40
- [73] E. M. Pothos, G. Perry, P. J. Corr, M. R. Matthew e J. R. Busemeyer, “Understanding cooperation in the prisoner’s dilemma game”, *Person. Individ. Diff.*, vol. 51, pp. 210–215, 2011. 41
- [74] A. Rapoport e A. M. Chammah, *Prisoner’s dilemma*. University of Michigan Press, 1985. 41
- [75] R. da Silva e G. A. Kellermann, “Analysing the payoff of a heterogeneous population in the ultimatum game”, *Braz. J. of Phys.*, vol. 37, pp. 1206–1211, 2007. 43
- [76] M. P. dos Santos, “Pesquisa operacional”. <http://www.mpsantos.com.br/>, 2003. p. 174. 63
- [77] H. Dalton, “The measurement of the inequality of incomes”, *Econ. J.*, vol. 30, pp. 348–361, 1920. 70

-
- [78] C. Gini, “Measurement of inequality of incomes”, *Econ. J.*, vol. 31, pp. 124–126, 1921. 70
- [79] A. B. Atkinson, “On the measurement of inequality”, *J. Econ. Theory*, vol. 2, pp. 244–263, 1970. 70
- [80] J. Weiner, “Size hierarchies in experimental populations of annual plants”, *Ecology*, vol. 66, pp. 743–752, 1985. 72
- [81] V. Sadras e R. Bongiovanni, “Use of Lorenz curves and Gini coefficients to assess yield inequality within paddocks”, *Field Crop. Res.*, vol. 90, pp. 303–310, 2004. 72
- [82] A. Sen e J. E. Foster, *On economic inequality*. Oxford University Press Inc., 1997. 72
- [83] D. Thon, “An axiomatization of the Gini coefficient”, *Math. Soc. Sci.*, vol. 2, pp. 131–143, 1982. 72
- [84] C. R. Barrett e M. Salles, “On a generalisation of the Gini coefficient”, *Math. Soc. Sci.*, vol. 30, pp. 235–244, 1995. 72
- [85] B. Milanovic, “A simple way to calculate the Gini coefficient, and some implications”, *Econ. Lett.*, vol. 56, pp. 45–49, 1997. 72
- [86] E. T. Jaynes, “Probability theory: The logic of science”. 1995. 75
- [87] I. R. Tsang e I. J. Tsang, “Cluster size diversity, percolation and complex systems”, *Phys. Rev. E*, vol. 60, pp. 2684–2698, 1999. 76
- [88] J. Tobochnik e H. Gould, *An Introduction to computer simulation methods*. Addison-Wesley Publishing Company, 1996. 76
- [89] R. M. Gray, *Entropy and Information Theory*. Springer-Verlag, 1990. 76, 78
- [90] S. Ihara, *Information theory for continuous systems*. World Scientific Publishing Co. Pie. Ltd., 1993. 77

-
- [91] C. E. Shannon, “A mathematical theory of communication”, *The Bell System Technical Journal*, vol. 27, pp. 379–423, 1948. 77
- [92] C. E. Shannon, *The mathematical theory of communication*. Univ of Illinois Press, 1949. 77
- [93] J. R. Iglesias e R. M. C. de Almeida, “Entropy and equilibrium state of free market models”, *Eur. Phys. J. B*, vol. 85, p. 10, 2012. 77
- [94] C. Tsallis, “Possible generalization of Boltzmann-Gibbs statistics”, *J. Stat. Phys.*, vol. 52, pp. 479–487, 1988. 77
- [95] S. Kullback e R. A. Leibler, “On information and sufficiency”, *Ann. Math. Statist*, vol. 22, pp. 79–86, 1951. 78
- [96] D. J. MacKay, *Information Theory, Inference, and Learning Algorithms*. 2002. 78
- [97] B. C. T. Cabella, M. J. Sturzbecher, W. Tedeschi, O. B. Filho, D. B. de Araújo e U. P. da Costa Neves, “A numerical study of the kullback-leibler distance in functional magnetic resonance imaging”, *Braz. J. of Phys.*, vol. 38, pp. 20–25, 2008. 78
- [98] C. Tsallis, “What are the numbers that experiments provide?”, *Quim. Nova*, vol. 17, pp. 468–471, 1994. 79
- [99] E. P. Borges, “On a q -generalizations of circular and hyperbolic functions”, *J Phys. A-Math. Gen.*, vol. 31, pp. 5281–5288, 1998. 79
- [100] B. C. T. Cabella, M. J. Sturzbecher, D. B. Araújo e U. P. C. Neves, “Generalized relative entropy in functional magnetic resonance imaging”, *Physica. A*, vol. 388, pp. 41–50, 2009. 79
- [101] A. Mukherji, V. Rajan e J. R. S. JR, “Robustness of cooperation”, *Nature*, vol. 379, pp. 125–126, 1996. 103
- [102] G. D. Birkhoff, “Proof of the ergodic theorem”, *Proc. Natl. Acad. Sci.*, vol. 17, pp. 656–660, 1931. 119

-
- [103] G. D. Birkhoff, “What is the ergodic theorem?”, *Am Math. Mon.*, vol. 49, pp. 222–226, 1942. 119
- [104] G. Rezaei e M. Kirley, “Dynamic social networks facilitate cooperation in the n -player prisoner’s dilemma”, *Physica. A*, vol. 391, pp. 6199–6211, 2012. 122
- [105] Z. Wang, W.-B. Du, X.-B. Cao e L.-Z. Zhang, “Integrating neighborhoods in the evaluation of fitness promotes cooperation in the spatial prisoner’s dilemma game”, *Physica. A*, vol. 390, pp. 1234–1239, 2011. 122
- [106] Y. Yoshino e N. Masuda, “Evolution of cooperation is a robust outcome in the prisoner’s dilemma on dynamic networks”, *ARXiv*, vol. arXiv:1006.5791v2 [physics.soc-ph], 2011. 122
- [107] M. A. Nowak, “Five rules for the evolution of cooperation”, *Science*, vol. 314, pp. 1560–1563, 2006. 122
- [108] F. C. Santos e J. M. Pacheco, “Scale-free networks provide a unifying framework for the emergence of cooperation”, *Phys. Rev. Lett.*, vol. 95, no. 098104, 2005. 122
- [109] A. Eriksson e K. Lindgren, “Cooperation driven by mutations in multi-person prisoner’s dilemma”, *J. Theor. Biol.*, vol. 232, pp. 399–409, 2005. 122
- [110] J. M. McNamara, Z. Barta e A. I. Houston, “Variation in behaviour promotes cooperation in the prisoner’s dilemma game”, *Nature*, vol. 428, pp. 745–748, 2004. 122
- [111] J. C. Oh, “Promoting cooperation using ‘kin’ biased conditional strategy in the iterated prisoner’s dilemma game”, *Inform. Sciences*, vol. 133, pp. 149–164, 2001. 122
- [112] N. Destefano e A. S. Martinez, “The additive property of the inconsistency degree in intertemporal decision making through the generalization of psychophysical laws”, *Physica. A*, vol. 390, pp. 1763–1772, 2011. 137