



Douglas Sad Silveira

**O Jogo *Snowdrift* Evolucionário com Presença
de Agentes Punidores num Arranjo Espacial de
Duas Populações**

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-Graduação
em Engenharia de Produção da Pontifícia Universidade
Católica (PUC-Rio) como requisito parcial para obtenção
do título de Mestre em Engenharia de Produção.

Orientador: Prof. André Barreira da Silva Rocha

Rio de Janeiro

Abril de 2016



Douglas Sad Silveira

**O Jogo *Snowdrift* Evolucionário com Presença de
Agentes Punidores num Arranjo Espacial de Duas
Populações**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia de Produção da PUC -Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. André Barreira da Silva Rocha

Orientador

Departamento de Engenharia Industrial - PUC-Rio

Prof. Davi Michel Valladão

Departamento de Engenharia Industrial - PUC-Rio

Prof. Antônio Carlos Figueiredo Pinto

Instituto de Administração e Gerência - IAG - PUC-Rio

Prof. Márcio da Silveira Carvalho

Coordenador Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 19 de Abril de 2016

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Douglas Sad Silveira

Possui Pós graduação *lato sensu* em Métodos Estatísticos Computacionais pela Universidade Federal de Juiz de Fora (UFJF) em 2014. Possui graduação em Economia pelas Faculdades Integradas Vianna Júnior com certificação de qualidade da Fundação Getúlio Vargas (FIVJ – FGV) em 2013.

Ficha Catalográfica

Silveira, Douglas Sad

O jogo *snowdrift* evolucionário com presença de agentes punidores num arranjo espacial de duas populações / Douglas Sad Silveira ; orientador: André Barreira da Silva Rocha. – 2016.

87 f. : il. color. ; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2016.

Inclui bibliografia

1. Engenharia Industrial – Teses. 2. Teoria dos jogos. 3. Estratégia evolucionariamente estável. 4. Simulação por agentes. I. Rocha, André Barreira da Silva. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. III. Título.

CDD: 658.5

Ao meu pai, Sérgio (*in
memorian*), e ao saudoso
Professor Carlos Patrício
Samanez (*in memorian*).

Agradecimentos

Ao meu orientador, Professor Dr. André Barreira da Silva Rocha, pela paciência, cooperação, dedicação e pelo conhecimento transmitido.

À CNPq e a PUC-Rio, pelos auxílios concedidos, sem os quais este trabalho não poderia ter sido realizado.

Aos professores e funcionários da PUC-Rio, pela instrução e confiança; bases para este trabalho.

À secretaria do DEI, em especial à Cláudia, Fernanda, Isabel, Gilvan e Graça por toda amizade, carinho e por todo café que me manteve acordado ao longo desses anos.

Aos amigos do mestrado, Igor, Leandro, Ali, Erick, Anna, Maurício, Raffael pela troca de experiências e pela ajuda recebida na execução deste trabalho.

Aos amigos da UFF, Eduardo Ferioli Gomes e Wilson Calmon, pelo compartilhamento do saber e pelas grandes contribuições que deram neste trabalho.

Aos amigos Eric Perim, Ricardo e Daniel pelas bem humoradas e precisas contribuições matemáticas e computacionais.

À Fernanda e minha família, por todo amor, carinho e apoio. Sem vocês, jamais conseguiria concluir este ciclo.

Resumo

Silveira, Douglas Sad; Rocha, André Barreira da Silva (Orientador). **O Jogo *Snowdrift* Evolucionário com Presença de Agentes Punidores num Arranjo Espacial de Duas Populações**. Rio de Janeiro, 2016. 87p. MSc. Dissertação - Departamento de Engenharia Industrial. Pontifícia Universidade Católica do Rio de Janeiro.

O papel da punição e os efeitos de uma população distribuída espacialmente na promoção da cooperação são questões relevantes. Num modelo de *Snowdrift Game* (SG) que incorpora um agente Punidor (P), Chan et al. (2013) estudaram os efeitos de uma população ligada por uma rede espacial. Os punidores, que carregam um caráter cooperativo (C), estão dispostos a pagar um custo α , de modo a punir um desertor (D), isto é, um oponente com comportamento não-cooperativo, em um custo β . Dependendo dos valores atribuídos para α , β , a razão de custo-benefício r , e da alocação das frequências iniciais, o sistema apresenta um comportamento evolutivo que pode ser homogêneo, com apenas uma única estratégia se perpetuando no longo prazo, ou não-homogêneo, com mais de uma estratégia correspondendo ao equilíbrio assintótico. A estrutura espacial impõe uma restrição geométrica sobre o ambiente competitivo. Neste trabalho, estendeu-se a modelagem de Chan et al. (2013) para duas populações com arranjo espacial. São apresentados os resultados de simulações numéricas. As dinâmicas observadas no mecanismo de replicação por imitação em duas redes espaciais são analisadas e discutidas. Em contraste com uma única população estruturada, a estratégia P sobrevive apenas em um dos casos analisados. Além do equilíbrio evolucionário de longo prazo, a dinâmica competitiva do jogo disposto em duas populações espaciais se difere daquela apresentada por Chan et al. (2013), na medida em que podemos observar a formação de populações inteiras de C e de D ou com a coexistência de C e P numa mesma população no caso em que o custo pelo benefício de cooperar é baixo.

Palavras-chave

Teoria dos Jogos; Estratégia Evolucionariamente Estável; Simulação por Agentes.

Abstract

Douglas Sad Silveira; Rocha, André Barreira da Silva (Advisor). **Evolutionary Snowdrift Game Incorporating Costly Punishment in Two Structured Populations**. Rio de Janeiro, 2016. 87p. MSc. Dissertation - Departamento de Engenharia Industrial. Pontifícia Universidade Católica do Rio de Janeiro.

The role of punishment and its effects in a spatial structured population in promoting cooperation are relevant issues. In the Snowdrift Game (SG) incorporating a Punisher agent (P), Chan et al. (2013) studied the effects of a population linked by a spatial network. Punishers, which carry a cooperative character (C), are willing to pay a cost α in order to punish a defector (D), i.e., a non-cooperative opponent, in a β cost. Depending on the values assigned to α , β , the ratio of cost-benefit r , and the initial distribution of the frequency of the strategies, the system displays an evolutionary behavior that can be homogeneous, with only a single strategy perpetuating in the long run, or non-homogeneous, with over a strategy corresponding to the asymptotic equilibrium. The spatial structure imposes a geometric constraint on the competitive environment. In this work, we extended the model proposed by Chan et al. (2013) for two populations with spatial arrangement. The results of numerical simulations are presented. The dynamic observed in the Agent Based Simulation approach in two spatial networks is discussed. In contrast to a single structured population, punishing strategy (P) survive in only one of the cases we shown. In the evolutionary equilibrium long-term, the competitive dynamics set in two spatial populations differs from that one presented by Chan et al. (2013). In this sense, we can observe the formation of entire populations of cooperators and defectors or the coexistence of C and P in each of the two populations for the case that the cost to benefit of cooperation is extremely low.

Keywords

Game Theory; Evolutionary Stable Strategy; Agent Based Simulation.

Sumário

| | |
|--|----|
| 1 Introdução | 16 |
| 2 Teoria dos Jogos | 19 |
| 2.1. Estratégias Mistas | 20 |
| 2.2. Perfil de Estratégia Mista | 23 |
| 2.3. O <i>payoff</i> de uma estratégia mista | 24 |
| 2.4. Correspondência de Melhor Resposta | 27 |
| 3 Teoria dos Jogos Evolucionários | 32 |
| 3.1. Jogos Unidimensionais | 38 |
| 3.2. Jogos Bidimensionais | 41 |
| 4 Replicação por Imitação – simulação computacional | 44 |
| 4.1. O Jogo Espacial | 46 |
| 4.2. A vizinhança de <i>von-Neumann</i> | 47 |
| 4.3. Os Métodos Síncrono e Assíncrono. | 48 |
| 5 Modelo | 50 |
| 5.1. O <i>Snowdrift Game</i> | 50 |
| 5.2. Mecanismos de cooperação no SG – a presença do punidor | 51 |
| 5.3. Projeto Piloto – Comparativo com Chan et al. (2013) | 52 |
| (i) Caso em que $r = 0,02$ | 52 |
| (ii) Caso em que $r = 0,20$ | 53 |
| (iii) Caso em que $r = 0,60$ | 54 |
| 5.4. O <i>Snowdrift</i> com punidor em duas populações espaciais | 55 |
| 6 Resultados | 59 |
| Caso I) $r = 0,02$ | 59 |
| Caso II) $r = 0,20$ | 66 |
| Caso II A) | 67 |

| | |
|------------------------------|----|
| Caso II B) | 73 |
| Caso III) $r = 0,6$ | 74 |
| Caso III A) | 75 |
| Caso III B) | 80 |
| 7 Conclusão | 83 |
| 8 Referências bibliográficas | 85 |

Lista de figuras

| | |
|---|----|
| Figura 1 - Espaço de estratégias mistas - 2 estratégias. | 21 |
| Figura 2 - Espaço de estratégias mistas $R^2 \rightarrow R$. | 21 |
| Figura 3 - Espaço de estratégias mistas - 3 estratégias. | 22 |
| Figura 4 - Espaço de estratégias mistas $R^3 \rightarrow R^2$. | 22 |
| Figura 5 - Espaço de perfis de estratégias mistas. | 23 |
| Figura 6 - Matriz do jogo <i>Hawk – Dove</i> . | 29 |
| Figura 7 - Correspondência de melhor resposta no jogo <i>Hawk – Dove</i> . | 30 |
| Figura 8 - Análise da estabilidade e tipos de jogo. | 39 |
| Figura 9 - Ilustração da dinâmica de um jogo de anti-coordenação. | 39 |
| Figura 10 - Dinâmica do jogo de coordenação. | 40 |
| Figura 11 - Dinâmica de estratégia dominante. | 40 |
| Figura 12 - Forma matricial de um jogo bidimensional. | 41 |
| Figura 13 - Casos possíveis para jogos bidimensionais. | 41 |
| Figura 14 - Diagrama de fase do caso I. | 42 |
| Figura 15 - Diagrama de fase do caso IX. | 43 |
| Figura 16 - População wellmixed e espacial. | 45 |
| Figura 17 - Pagamento do Focal e Referência. | 45 |
| Figura 18 - Vizinhança de <i>von-Neumann</i> . | 47 |
| Figura 19 - Vizinhança de <i>von-Neumann</i> para indivíduos situados nas fronteiras. | 48 |
| Figura 20 - Equilíbrios de Nash em estratégias puras no jogo SG. | 51 |
| Figura 21 - Matriz do jogo <i>Snowdrift</i> com o Punidor. | 51 |

| | |
|---|----|
| Figura 22 - Matriz de <i>payoff</i> do SG com o Punidor. | 52 |
| Figura 23 - Comparativo da evolução das estratégias em Chan et al. (2013) e projeto piloto para $r = 0,02$. | 53 |
| Figura 24 - Comparativo da evolução das estratégias em Chan et al. (2013) e projeto piloto para $r = 0,20$. | 54 |
| Figura 25 - Comparativo da evolução das estratégias em Chan et al. (2013) e projeto piloto para $r = 0,60$. | 55 |
| Figura 26 – Inicialização das Populações 1 e 2. | 56 |
| Figura 27 - Processo de escolha dos agentes Focal e Referência. | 56 |
| Figura 28 - Cálculo do pagamento para 2 populações. | 57 |
| Figura 29 - Exemplo do procedimento computacional para atualização de estratégias. | 57 |
| Figura 30 - <i>Snowdrift</i> com agente punidor em 2 populações com distribuição espacial. Parâmetro $r = 0,02$. | 59 |
| Figura 31 - Equilíbrio assintótico obtido por Chan et al. (2013) para $r = 0,02$. | 60 |
| Figura 32 - Passo de Monte Carlo 1; $r = 0,02$. | 60 |
| Figura 33 – Passo de Monte Carlo 10; $r = 0,02$. | 61 |
| Figura 34 – Exemplo do pagamento obtido pelos jogadores Focal e Referência; $r = 0,02$. | 61 |
| Figura 35 - Matriz de pagamento para o caso em que $r = 0,02$, $\alpha = 0,01$ e $\beta = 0,05$. | 62 |
| Figura 36 - Passo de Monte Carlo 150; $r = 0,02$. | 63 |
| Figura 37 - Passo de Monte Carlo 850; $r = 0,02$. | 64 |
| Figura 38 - Passo de Monte Carlo 1000; $r = 0,02$. | 64 |

| | |
|---|----|
| Figura 39 - Passo de Monte Carlo 6000; $r = 0,02$. | 65 |
| Figura 40 - Passo de Monte Carlo 20000; $r = 0,02$. | 65 |
| Figura 41- Passo de Monte Carlo 100.000; $r = 0,02$. | 66 |
| Figura 42 - Equilíbrio assintótico obtido por Chan et al. (2013) para $r = 0,2$. | 67 |
| Figura 43 - Equilíbrio Assintótico com o comportamento típico para $r = 0,2$. População 1 <i>All C</i> e População 2 <i>All D</i> . | 68 |
| Figura 44 - Matriz de pagamento do jogo; $r = 0,20$, $\alpha = 0,01$ e $\beta = 0,05$. | 68 |
| Figura 45 – Passo de Monte Carlo 10; $r = 0,2$. | 69 |
| Figura 46 - Passo de Monte Carlo 100; $r = 0,2$. | 69 |
| Figura 47 - Passo de Monte Carlo 1000; $r = 0,2$. | 70 |
| Figura 48 – Focal D e Referência C com 3 vizinhos desertores; $r = 0,2$. | 70 |
| Figura 49 - Focal D e Referência C com 4 vizinhos desertores; $r = 0,2$. | 71 |
| Figura 50 - Dinâmica dos passos de Monte Carlo 5.000, 10.000 e 20.000; $r = 0,2$. | 72 |
| Figura 51 - Dinâmica evolutiva verificada no caso II B; $r = 0,2$. | 73 |
| Figura 52 - Evolução espacial do jogo. À esquerda estão representados os passos de Monte Carlo 1.000, 30.000 e 60.000. À direita 1.000.000, 1.500.000 e 2.000.000. | 73 |
| Figura 53 - Estado All C e All D com 2.400.000 Passos de Monte Carlo; $r = 0,2$. | 74 |

| | |
|--|----|
| Figura 54 - Equilíbrio assintótico obtido por Chan et al. (2013) para $r = 0,6$. | 75 |
| Figura 55 - Comportamento típico da dinâmica evolutiva do jogo; $r=0,6$. | 75 |
| Figura 56 - Matriz de pagamentos para o caso em que $r = 0,6$, $\alpha = 0,01$ e $\beta = 0,05$. | 76 |
| Figura 57 - Evolução das populações – passos Monte Carlo 1, 10 e 20; $r = 0,6$. | 76 |
| Figura 58 - Evolução das interações estratégicas - passos 50, 100 e 1.000; $r = 0,6$. | 77 |
| Figura 59 - Focal e Referência competindo com 4 desertores; $r = 0,6$. | 78 |
| Figura 60 - Focal e Referência competindo com 4 cooperadores; $r = 0,6$. | 79 |
| Figura 61 - Distribuição espacial dos passos 2.000, 5.000 e 10.000; $r = 0,6$. | 79 |
| Figura 62 - Equilíbrio assintótico do jogo no PMC 15.000; $r = 0,6$. | 80 |
| Figura 63 - Dinâmica evolutiva verificada no caso III B; $r = 0,6$. | 80 |
| Figura 64 - Evolução espacial do jogo. | 81 |
| Figura 65 - Estados <i>All C</i> e <i>All D</i> no PMC 2.000.000; $r = 0,6$. | 82 |

Lista de abreviaturas e siglas

| | |
|-------|---|
| ESS | <i>Evolutionary Stable Strategy</i> |
| EEE | Equilíbrio Evolucionariamente Estável |
| EM | Equilíbrio de <i>Nash</i> |
| EDO | Equações Diferenciais Ordinárias |
| INT | Interior |
| SG | <i>Snowdrift Game</i> |
| DP | Dilema do Prisioneiro |
| C | Cooperar |
| D | Desertar |
| P | Punir |
| N x N | Dimensão da matriz com N linhas e N colunas |

“A dúvida é o preço da pureza.”

Jean-Paul Sartre

1

Introdução

Em 1950, o matemático John Nash popularizou a aplicação de teoria dos jogos em pesquisas, ao oferecer uma série de soluções a questões até então insolúveis. A premissa de racionalidade completa para os agentes (jogadores) e as várias extensões sobre o conceito de equilíbrio de Nash nem sempre refletiam bem situações reais. Em outras palavras, as pessoas não agiam da forma como se postulava nos modelos. Maynard Smith & Price (1973) definiram o que é uma estratégia evolucionariamente estável (EEE), introduzindo explicitamente o conceito de seleção evolutiva num jogo, fornecendo, também, um refinamento acerca do conceito dos equilíbrios de Nash. Em outras palavras, é possível inferir sobre qual equilíbrio de Nash corresponde a uma EEE. Essa noção da evolução em jogos repetidos encontra respaldo em certos modelos nos quais os jogadores aprendem através das estratégias adotadas no passado, podendo imitar o comportamento mais bem sucedido num dado ambiente competitivo. Essa premissa, associada a pressuposição de que os indivíduos possuem racionalidade limitada (Nowak et al., 2010), contribuiu em larga escala para a aceitação dos modelos evolutivos, fazendo com que estes fossem largamente utilizados a partir da década de 90.

Modelos teóricos de jogos fornecem uma poderosa ferramenta para o estudo da evolução do comportamento cooperativo. A emergência da cooperação é uma questão de extrema relevância e fundamental importância no estudo científico do comportamento estratégico em populações (Nowak e Highfield, 2011), (Hauert e Doebeli, 2004) e (Rocha, 2013). Este tópico tem atraído a atenção de muitos pesquisadores de diversos ramos da ciência, incluindo ecologistas, economistas, engenheiros industriais, físicos e matemáticos. Existe uma similaridade desta questão da cooperação com tópicos que envolvem o tema de finanças, por exemplo, Hirth (2014) e Bolton et al. (2012) tratam da interação estratégica entre as agências de rating de crédito com investidores. Em se tratando de teoria dos jogos, é possível fazer distinção entre dois ramos.

- 1) Jogos cooperativos: agentes competem com o objetivo de obter ganhos mútuos;
- 2) Jogos não cooperativos: agentes competem com o objetivo de conseguir ganhos pessoais.

Na abordagem evolucionária do Dilema dos Prisioneiros e do *Snowdrift game* o jogo é jogado ao longo de um dado intervalo de tempo, repetidas vezes, e os agentes podem escolher entre cooperar ou desertar, evoluindo seu comportamento a partir da avaliação daquela estratégia que melhor performa no ambiente em que estão inseridos. O Dilema do Prisioneiro (DP), de acordo com Xu et al. (2011), é o modelo padrão representativo de um dilema social e tem sido estudado exaustivamente. Porém, devido à dificuldade em se determinar precisamente a ordem de grandeza dos *payoffs* associados a cada estratégia, o *Snowdrift Game* (SG) é considerado uma importante modelagem alternativa ao DP.

No SG, dois motoristas (jogador 1, jogador 2) estão guiando em direções opostas e se deparam com uma estrada bloqueada por um desmoronamento de neve. Eles possuem duas possíveis ações a serem escolhidas simultaneamente: cavar a neve (cooperar, C) ou não cavar (não cooperar, D).

Dessas ações, surgem os seguintes possíveis resultados:

- i) (C,C) – ambos recebem o benefício b . Porém, cavar tem um custo c , que no caso em que ambos cooperam, será dividido. Cada um dos jogadores recebe o pagamento $R = b - \frac{c}{2}$.
- ii) (D,D) – ambos ficam presos e não recebem nada por isso. $P = 0$.
- iii) (C,D) ou (D,C) – ambos vão para casa. Porém, o que jogou D vai para casa sem fazer qualquer esforço, recebendo um pagamento $T = b$, enquanto o seu oponente, que jogou C, recebe $S = b - c$, absorvendo todo o custo de cavar.

No SG $b > c > 0$, o que leva a um ranking em que $T > R > S > P$. No jogo estático, isto é, jogado uma única vez, a melhor resposta a um oponente que joga C é jogar a estratégia D. Quando o oponente joga D com certeza, a melhor resposta é adotar a estratégia C. No âmbito de jogos evolucionários, que são jogados dinamicamente ao longo de um intervalo de tempo, os jogadores visam obter as maiores recompensas e, para tal, podem mudar de estratégia ao longo do jogo, imitando o comportamento de um adversário que melhor performar dentro do ambiente competitivo.

Um mecanismo para a promoção do comportamento cooperativo no jogo SG evolucionário, conforme Chan et al. (2013), é a incorporação de uma estratégia de punição (P) numa estrutura de jogo espacial para uma população, em que um agente punidor - que pode ser visto como um cooperador altruísta - está disposto a abrir mão de uma quantia de seu *payoff*, ganhando um pouco menos, para que os desertores (D) sejam penalizados. O agente punidor (P) é parcialmente racional, dado que a punição leva a uma perda maior no pagamento obtido pelo jogador desertor. A depender do custo pelo benefício da cooperação, fornecido pelo parâmetro r , os resultados apresentados por Chan et al. (2013), indicam que a presença do punidor contribuiu positivamente para a evolução da cooperação num jogo de uma população com estrutura espacial. Sendo assim, a maior contribuição desta dissertação foi estender a abordagem proposta por Chan et al. (2013) para um contexto em que exista interação estratégica entre duas populações distribuídas numa estrutura de jogo espacial.

O objetivo principal dessa dissertação é analisar o efeito da extensão da modelagem do jogo *Snowdrift* evolucionário em uma estrutura espacial aplicado num ambiente competitivo formado por duas populações. Deste modo, pretende-se inferir sobre a dinâmica de longo prazo de cada uma das estratégias e, por conseguinte, qual delas corresponde ao equilíbrio assintótico do jogo em cada uma das populações. Para a extensão da modelagem de Chan et al. (2013), criou-se um programa computacional em linguagem JAVA com o propósito de apresentar e expandir os conceitos e aplicações da Teoria de Jogos Evolucionários para uma estrutura que envolve simulações de jogos espaciais envolvendo duas populações, de modo a confrontar e comparar os resultados obtidos por Chan et al. (2013) nos tradicionais jogos de uma população do *Snowdrift* com incorporação do agente punidor (P). Cabe ressaltar que Hauert (2004) quebrou paradigmas ao mostrar que nem sempre a introdução da estrutura de jogo espacial favorece a cooperação.

2 Teoria dos Jogos

A teoria dos jogos propõe uma abordagem interdisciplinar no que tange ao comportamento do ser humano, provendo fundamentos necessários para a compreensão da interação entre os agentes econômicos. Relacionam-se diretamente com teoria dos jogos as disciplinas ligadas à matemática, física, biologia, economia e outras ciências sociais e do comportamento.

Este capítulo inicia-se com uma revisão da base teórica da teoria dos jogos, revisando a literatura exposta por von-Neumann e Morgenstern (1944) e, principalmente uma revisão sobre o conceito de equilíbrio de Nash (1950). À teoria de jogos clássica, acrescentar-se-á, no capítulo 3, a noção de Estratégias Evolucionariamente Estáveis (EEE). Mais adiante, o conjunto de equações diferenciais ordinárias (EDO) que regem um processo evolutivo biológico através do replicador dinâmico, também será apresentado. Discute-se, ainda, a dinâmica malthusiana do replicador, mostrando que uma EEE também constitui um refinamento do equilíbrio de Nash, uma vez que a solução analítica do sistema de EDO que determina os equilíbrios evolutivos estáveis pode ser vista como um subconjunto dos equilíbrios de Nash.

Ao modelarmos situações conflitantes entre agentes assumidos como racionais, por meio da teoria dos jogos, podemos descrever um conjunto de estratégias de cada agente e seu respectivo *payoff*. O conceito do equilíbrio de Nash é fundamental na previsão do resultado de um jogo e, de acordo com Mas-Collel, Whinston e Green (1995), pode ser definido como uma situação em que nenhum jogador possui incentivo para desviar unilateralmente daquela estratégia que escolheu inicialmente para jogar. Logo, num equilíbrio de Nash, a estratégia escolhida por cada jogador é uma melhor resposta ao perfil de estratégias escolhido por seus oponentes.

Em muitas situações não há um equilíbrio de Nash único, existindo mais de um perfil de estratégia que corresponde a uma melhor resposta ao adversário. Refinando-se o conceito de equilíbrio de Nash, pode-se inferir melhor sobre o

resultado de um jogo, descartando soluções de equilíbrio que não condizem com determinados comportamentos dos agentes.

O modelo usado como base em jogos não cooperativos é conhecido como um jogo de n agentes e em sua forma normal é caracterizado por $\Gamma = (S_1, S_2, \dots, S_n; I_1, I_2, \dots, I_n; \pi_1, \pi_2, \dots, \pi_n)$.

Onde I é o conjunto de jogadores;

$S_i = (1, 2, \dots, m_i)$ é o conjunto das m_i estratégias puras do jogador i ;

$\pi_i(s)$ é o *payoff* do jogador i diante do perfil de estratégias puras $s = \{s_1, s_2, \dots, s_n\}$.

$s_i \in S_i$, isto é, cada estratégia pura s_i , jogada pelo jogador i pertence ao conjunto de estratégias puras S_i deste mesmo jogador i .

O produto cartesiano $S = \times S_i$ é chamado de espaço de estratégias puras, em que $s \in S$.

O *payoff* de cada perfil de estratégia é definido como $\pi(s) = [\pi_1, \pi_2, \dots, \pi_n]$.

Um jogo definido em termos de estratégias puras, em sua forma normal, pode ser escrito como:

$G = (I, S, \pi)$, onde I é o conjunto de jogadores.

Em se tratando de um jogo com dois jogadores, ambos possuindo as mesmas estratégias puras, P e R , os *payoffs* podem ser postos na seguinte forma matricial:

$$A = \begin{bmatrix} \pi_{1PP}; \pi_{2PP} & \pi_{1PR}; \pi_{2PR} \\ \pi_{1RP}; \pi_{2RP} & \pi_{1RR}; \pi_{2RR} \end{bmatrix}$$

Em que

π_{iPR} corresponde ao *payoff* obtido pelo jogador i ao jogar sua estratégia pura P contra a estratégia pura R do seu oponente.

2.1. Estratégias Mistas

Além da possibilidade de representarmos as ações de um jogador através de suas estratégias puras, também é possível que tais ações sejam representadas por meio de estratégias mistas. Considerando uma dada distribuição de probabilidade sobre o conjunto de estratégias puras, obtém-se as chamadas estratégias mistas.

Estas, por sua vez, são representadas por um vetor $x_j \in \mathcal{R}^{m_j}$, onde cada elemento deste vetor x_{jP} é a probabilidade do jogador j jogar a sua estratégia pura P e m_j é o número de estratégias puras de um jogador. O simplex unitário de dimensão $m_j - 1$, definido a seguir, representa o conjunto de todas as estratégias mistas do jogador j :

$$\Delta_j = \left\{ x_j \in \mathcal{R}_+^{m_j}; \sum_{P=1}^{m_j} x_{jP} = 1 \right\} \quad (1)$$

A dimensão do simplex Δ_j é $m_j - 1$, pois é possível escrevermos qualquer probabilidade como 1 menos a soma das outras probabilidades, desde que o somatório de todas estas probabilidades seja igual a um. Deste modo, sem perda de informação, analisaremos o simplex $\Delta_j \in \mathcal{R}_+^{m_j}$ por sua projeção no espaço com dimensão $m_j - 1$.

Para uma melhor visualização, suponha o caso que $m_j = 2$, ou seja, o jogador j possui apenas duas estratégias no jogo, o espaço de estratégias mistas como definido na equação (1) é representado graficamente pela Figura 1 e sua projeção num espaço $m_j - 1$, pode ser vista na Figura 2.

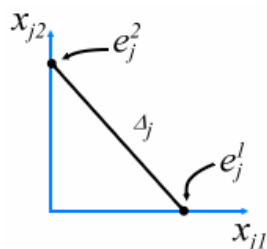


Figura 1- Espaço de estratégias mistas para 2 estratégias.

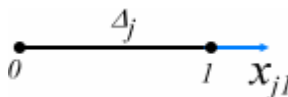


Figura 2- Projeção do espaço de estratégias mistas $\mathcal{R}^2 \rightarrow \mathcal{R}$.

Para o caso em que um jogador possui três estratégias puras, o espaço de estratégias mistas pode ser visto como na figura 3 e sua projeção no espaço bidimensional será mostrado na figura 4 a seguir:

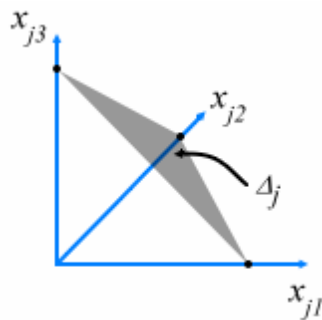


Figura 3 - Espaço de estratégias mistas para o caso de três estratégias.

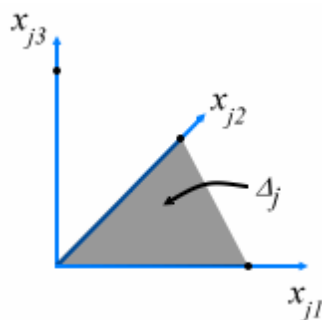


Figura 4 - Projeção do espaço de estratégias mistas $\mathbb{R}^3 \rightarrow \mathbb{R}^2$.

Sendo que

$$e_j^1 = \{1, 0, \dots, 0\}, e_j^2 = \{0, 1, \dots, 0\}, \dots, e_j^{m_1} = \{0, 0, \dots, 1\} \quad (2)$$

denotam os vértices do simplex unitário Δ_j que, por sua vez, são vetores unitários no espaço m_j .

Cada vértice e_j^P representa uma estratégia mista para o jogador j que associa probabilidade 1 à sua P -ésima estratégia pura e probabilidades nulas às demais estratégias disponíveis. Deste modo, cada estratégia mista $x_j \in \Delta_j$ é uma combinação convexa de estratégias puras. Logo, podemos escrever:

$$x_j = \sum_{P=1}^{m_j} x_{jP} e_j^P \quad (3)$$

2.2. Perfil de Estratégia Mista

Seja $x = \{x_1, x_2, \dots, x_n\}$ um perfil de estratégias mistas onde x_j é a estratégia mista do jogador j . Um perfil de estratégias mistas é um ponto do espaço de perfis de estratégias mistas definido como o produto cartesiano dos n simplex Δ_j de dimensão $m_j - 1$.

$$\Theta = \times_{j \in I} \Delta_j \quad (4)$$

O espaço Θ possui dimensão m , sendo que $m = \sum_{j=1}^n m_j$ corresponde ao total de estratégias puras no jogo e sua projeção feita no espaço de dimensão $m - n = m_1 - 1 + m_2 - 1 + \dots + m_n - 1$ é o produto cartesiano das projeções de Δ_j .

Quando o jogo é composto por dois indivíduos ($n = 2$) e cada um possui duas estratégias $m_1 = m_2 = 2$, o espaço é dado por

$\Theta = \{[x_{11}, x_{21}], [x_{11}, x_{22}], [x_{12}, x_{21}], [x_{12}, x_{22}]\} \in \mathcal{R}^4$ corresponde ao produto cartesiano de Δ_1 e Δ_2 .

O produto cartesiano pode ser visualizado através das projeções de Δ_1 e Δ_2 na figura abaixo. A ver:

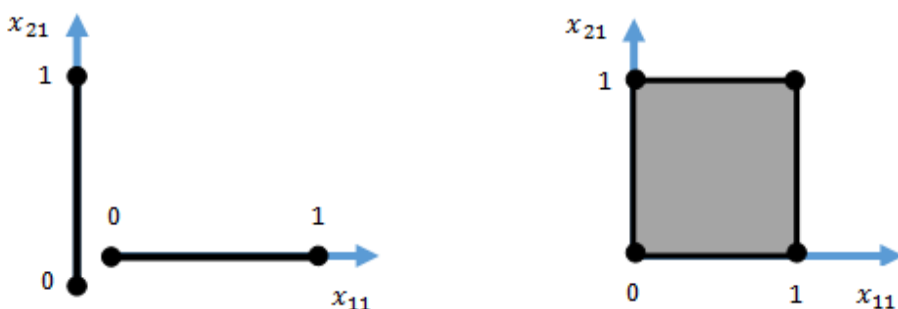


Figura 5- Espaço de perfis de estratégias mistas.

Um perfil de estratégias mistas x é dito interior se cada uma de suas estratégias mistas x_j for interior e, além disso, ser definida por:

$$\text{INT}(\Theta) = \times_{j \in I} \text{INT}(\Delta_j) \quad (5)$$

Considerando (x_j, y_{-j}) como o perfil de estratégias no qual o jogador j joga a sua estratégia $x_j \in \Delta_j$ contra o perfil $y \in \Theta$ jogado pelos demais jogadores, de

modo mais preciso, é possível definir um perfil de estratégia $Z = (x_j, y_{-j}) \in \theta$ como $z_j = x_j$ e $z_i = y_i \forall j \neq i$.

2.3.

O *payoff* de uma estratégia mista

Em termos probabilísticos, a chance de um perfil de estratégias puras $s = \{s_1, s_2, \dots, s_n\} \in S$ ser utilizado quando o perfil de estratégias mistas $x \in \theta$ é jogado pode ser definido como:

$$x(s) = \prod_{j=1}^n x_{js_j} \quad (6)$$

x_{js_j} é a probabilidade do jogador j jogar sua estratégia s_j .

Suponha, agora, um contexto em que existam dois jogadores e, cada um dispõe de duas estratégias. O jogador A possui as estratégias a_1 e a_2 e o jogador B joga com as estratégias b_1 e b_2 . O caso em que $s = (a_1, b_1)$ é um exemplo de um perfil formado por estratégias puras de ambos os agentes. Uma estratégia mista do jogador A , por exemplo, poderia ser $x_A = (x_{A1}, x_{A2})$ onde x_{A1} corresponde à probabilidade do jogador A jogar a sua estratégia a_1 . Agora, fica mais claro enxergar que um perfil de estratégias mistas $x = (x_A, x_B)$, em que x_B é uma estratégia mista de B , pode ser analogamente escrita da seguinte forma:

$$x = (x_A, x_B) = \left(\begin{bmatrix} x_{A1} \\ x_{A2} \end{bmatrix}, \begin{bmatrix} x_{B1} \\ x_{B2} \end{bmatrix} \right) \in \theta \quad (7)$$

O *payoff* esperado para o jogador j associado com o perfil de estratégias mistas $x \in \theta$ é dado por:

$$\mu_j(x) = \sum_{s \in S} x(s) \pi_j(s) \quad (8)$$

$x(s)$ é a probabilidade associada a um perfil de estratégias puras e $\pi_j(s)$ é o *payoff* que o jogador j obtém no caso em o perfil de estratégia pura s for jogado.

Quando um agente opta por jogar a estratégia pura $s_j = k$, diz-se que esta ação é probabilisticamente equivalente a jogar a estratégia mista $e_{jk} \in \Delta_j$, assim

podemos escrever $u_j(e_j^k, x_{-j})$ como o valor esperado do *payoff* do jogador j quando o perfil (e_j^k, x_{-j}) for jogado. Em outras palavras, $u_j(e_j^k, x_{-j})$ é o pagamento obtido pelo jogador j quando ele joga sua k -ésima estratégia pura.

Assim, temos que $\forall x \in \Theta$:

$$\mu_j(x) = \sum_{k=1}^{m_j} \mu_j(e_j^k, x_{-j}) x_{jk} \quad (9)$$

Onde

x_{jk} corresponde à probabilidade do jogador j jogar sua k -ésima estratégia pura.

$\mu_j(x) = (\mu_1(x), \mu_2(x), \dots, \mu_n(x))$ também é chamada de função *payoff* das estratégias mistas do jogo.

O valor esperado do *payoff* do jogador “1” associado com a estratégia mista x é dado por:

$\mu_1(x) = \sum_{h=1}^{m_1} \sum_{k=1}^{m_2} x_{1h} a_{hk} x_{2k} = x_1^T A x_2$, onde \mathbf{A} é a matriz de *payoff* do jogador “1”.

O valor esperado do *payoff* do jogador “2” associado com a estratégia mista x é dado por:

$\mu_2(x) = \sum_{h=1}^{m_1} \sum_{k=1}^{m_2} x_{1h} b_{hk} x_{2k} = x_1^T B x_2$, onde \mathbf{B} é a matriz de *payoff* do jogador “2”.

Para o jogador “1”, a estrutura de *payoff* fica:

$$\mathbf{A} = \begin{matrix} & \mathbf{b}_1 & \mathbf{b}_2 \\ \mathbf{a}_1 & \pi_1(\mathbf{a}_1, \mathbf{b}_1) & \pi_1(\mathbf{a}_1, \mathbf{b}_2) \\ \mathbf{a}_2 & \pi_1(\mathbf{a}_2, \mathbf{b}_1) & \pi_1(\mathbf{a}_2, \mathbf{b}_2) \end{matrix}$$

Sendo $\pi_j(a_h, b_k)$ o *payoff* do jogador j ao jogar sua estratégia pura h contra a estratégia pura k de seu oponente i . Ressaltando que $i \neq j$.

O valor esperado do *payoff* do jogador “1” será dado pela seguinte fórmula:

$$\begin{aligned} \mu_1(s) &= \sum_{s \in S} x(s) \pi_1(s) = \\ &= Prob(a_1, b_1) * \pi_1(a_1, b_1) + Prob(a_1, b_2) * \pi_1(a_1, b_2) + Prob(a_2, b_1) \\ &\quad * \pi_1(a_2, b_1) + Prob(a_2, b_2) * \pi_1(a_2, b_2) \end{aligned}$$

Em que $Prob(a_h, b_k) = Probabilidade(a_h) \times Probabilidade(b_k)$. É muito importante reparar que esta igualdade só será válida para quando a estratégia adotada pelo primeiro jogador for independente daquela adotada pelo segundo jogador.

Fazendo com que $\pi_j(a_h, b_k) = a_{hk}$, temos que:

$$\begin{aligned} & a_{11} * x(a_1, b_1) + a_{12} * x(a_1, b_2) + a_{21} * x(a_2, b_1) + a_{22} * x(a_2, b_2) = \\ & = a_{11} * x_{11} * x_{21} + a_{12} * x_{11} * x_{22} + a_{21} * x_{12} * x_{21} + a_{22} * x_{12} * x_{22} = \\ & = \sum_{h=1}^2 \sum_{k=1}^2 a_{hk} x_{1h} x_{2k} \end{aligned}$$

De maneira análoga, encontra-se:

$$\mu_2(x) = \sum_{h=1}^2 \sum_{k=1}^2 b_{hk} x_{1h} x_{2k} = x_1^T \mathbf{B} x_2 = x_2^T \mathbf{B}^T x_1$$

O *payoff* combinado das estratégias mistas é representado pela função $\mu(x) = (\mu_1(x), \mu_2(x)): IR^4 \rightarrow IR^2$. Em palavras, $IR^m \rightarrow IR^n$, onde $m = m_1 + m_2$ corresponde ao total de estratégias puras disponíveis dos jogadores. No caso em análise, existe um total de quatro estratégias puras e dois jogadores ($n = 2$). Para uma melhor compreensão do que foi exposto anteriormente, apresenta-se um exemplo a seguir.

Considerando uma matriz de *payoff* para um jogador A e para um jogador B:

$$\mathbf{A} = \begin{bmatrix} 4 & 0 \\ 5 & 3 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 4 & 5 \\ 0 & 3 \end{bmatrix}$$

$$\begin{aligned} \mu_1(x) &= x_1 \mathbf{A} x_2 = [x_{11} \quad x_{12}] \begin{bmatrix} 4 & 0 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} \\ &= [4x_{11} + 5x_{12} \quad 3x_{12}] \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} \\ &= 4x_{11}x_{21} + 5x_{12}x_{21} + 3x_{12}x_{22}. \end{aligned}$$

$$\begin{aligned} \mu_2(x) &= x_1 \mathbf{B} x_2 = [x_{11} \quad x_{12}] \begin{bmatrix} 4 & 5 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} \\ &= [4x_{11} \quad 5x_{11} + 3x_{12}] \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} \\ &= 4x_{11}x_{21} + 5x_{11}x_{22} + 3x_{12}x_{22}. \end{aligned}$$

2.4. Correspondência de Melhor Resposta

Jehle e Reny (2011) definem uma melhor resposta em estratégias puras para o jogador j , no perfil de estratégias $y \in \Theta$, como uma estratégia pura $s_i \in S_j$, tal que nenhuma outra estratégia pura que esteja ao alcance do jogador resulte num *payoff* maior.

A correspondência de melhor resposta para o caso de uma estratégia pura de um jogador $\beta_j : \Theta \rightarrow S_j$ mapeia cada perfil de estratégia mista $y \in \Theta$ em um conjunto de melhores respostas puras do jogador j ante o perfil y . Este conjunto é finito e não vazio, podendo ser definido como:

$$\beta_j(y) = \{h \in S_j : \mu_j(e_j^h, y_{-j}) \geq \mu_j(e_j^k, y_{-j}) \forall (k \neq h) \in S_j\}$$

Isto é, $\beta_j(y)$ corresponde ao conjunto de estratégias puras que obtém os melhores *payoffs* quando confrontados com um perfil de estratégias concorrente y . Se a estratégia for uma melhor resposta, podem existir outras estratégias tão boas quanto ela, mas não melhores.

Sendo uma estratégia mista $x_j \in \Delta_j$ uma combinação convexa de estratégias puras e $\mu_j(x_j, y_{-j})$ linear em relação a x_j , nenhuma estratégia mista $x_j \in \Delta_j$ pode dar ao jogador j um *payoff* maior contra o perfil $y \in \Theta$ do que qualquer uma de suas melhores respostas puras diante de um mesmo y .

Suponha que as duas estratégias puras de um jogador j sejam representadas por $e_{j1} = (1,0)$ e $e_{j2} = (0,1)$. Se $\mu_j(e_j^1, y_{-j}) > \mu_j(e_j^2, y_{-j})$, então qualquer estratégia mista x_j que seja formada a partir de uma combinação convexa das estratégias puras disponíveis terá um *payoff* dado por $\mu_j(e_j^1, y_{-j}) > \mu_j(x_j, y_{-j}) > \mu_j(e_j^2, y_{-j})$.

Portanto, podemos escrever de maneira formal que para qualquer $y \in \Theta$, $x_j \in \Delta_j$ e $h \in \beta_j(y)$ o *payoff*

$$\mu_j(x_j, y_{-j}) = \sum_{k=1}^{m_j} \mu_j(e_j^k, y_{-j}) x_{jk} \leq \sum_{k=1}^{m_j} \mu_j(e_j^h, y_{-j}) x_{jk} = \mu_j(e_j^h, y_{-j})$$

Como a estratégia mista x_j pode ser representada por uma combinação convexa dos vértices de Δ_j , têm-se que:

$$\begin{aligned}
 x_j &= \sum_{k=1}^{m_j} e_j^k x_{jk} \\
 \mu_j(x_j, y_{-j}) &= \sum_{k=1}^{m_j} \mu_j(e_j^k, y_{-j}) x_{jk}
 \end{aligned} \tag{10}$$

Assim, definimos o conjunto $\beta_j(y)$ como

$$\beta_j(y) = \{h \in S_j : \mu_j(e_j^h, y_{-j}) \geq \mu_j(e_j^k, y_{-j}) \forall x_j \in \Delta_j\}$$

Em palavras, o que podemos concluir é que a correspondência de melhor resposta é um conjunto formado por todas as estratégias do jogador que levam à maximização não estrita de seu pagamento.

Cabe, ainda, ressaltar que uma melhor resposta em estratégia pura pode ser vista como uma estratégia mista e_{jh} quando atribuímos probabilidade 1 para a estratégia pura s_h . Portanto, cada melhor resposta em estratégia pura é também uma melhor resposta em estratégia mista. O conceito de melhor resposta é importante, uma vez que o equilíbrio de Nash, a ser apresentado na próxima subseção, corresponde a um ponto fixo de estratégias de melhor resposta simultânea num jogo.

2.5. Equilíbrio de Nash

De acordo com Mas-Collel, Whinston e Green (1995), o perfil de estratégias $s = \{s_1, s_2, \dots, s_n\}$ é um equilíbrio de Nash (E.N.) em estratégias puras de um jogo se, para todo jogador i , vale a desigualdade

$$V_i(s_i, s_{-i}) \geq V_i(s_i', s_{-i}) \quad \forall s_i' \in S_i$$

Em que $V_i(s_i, s_{-i})$ é o valor da estratégia s_i escolhida pelo jogador i para jogar o jogo e s_{-i} é o vetor $(I - 1)$ que contém as estratégias de todos os oponentes do jogador i . Num equilíbrio de Nash, a estratégia escolhida por cada jogador é uma melhor resposta ao perfil de estratégias escolhido por seus oponentes e nenhum jogador possui incentivo para desviar unilateralmente da sua estratégia escolhida para jogar.

O equilíbrio de Nash contribuiu muito para deixar bastante clara a distinção entre as classes de jogos cooperativos (coordenação) e não cooperativos (anti-coordenação). Nos cooperativos há existência de acordos que podem ser estabelecidos (em tribunais, contratos, etc.). Já nos jogos não-cooperativos não existem tais mecanismos.

O conceito de E.N. associado à premissa do comportamento racional dos agentes motivaram muitos estudos nas classes de jogos não cooperativos. Em tais contextos, o equilíbrio de Nash pode ser uma ferramenta limitada, uma vez que em nos jogos de anti-coordenação há mais de um equilíbrio no sentido de Nash. Uma escolha arbitrária entre os possíveis equilíbrios de Nash pode não fazer muito sentido como predição do comportamento dos agentes. Além disso, nem todos os resultados são consistentes com a intuição sobre qual deveria ser o resultado do jogo.

Concluindo este capítulo, apresenta-se o jogo *Hawk-Dove*, exemplo clássico de um jogo de anti-coordenação, que auxiliará num entendimento prévio sobre um jogo evolucionário bem como no conceito de Equilíbrio de Nash e Estratégia Evolucionariamente Estável. Considere a seguinte matriz:

| | | JOGADOR B | |
|-----------|----------|--------------------------------|----------------------------|
| | | $H(q)$ | $D(1-q)$ |
| JOGADOR A | $H(p)$ | $\frac{v-c}{2}; \frac{v-c}{2}$ | $v; 0$ |
| | $D(1-p)$ | $0; v$ | $\frac{v}{2}; \frac{v}{2}$ |

Figura 6 - Matriz do jogo *Hawk – Dove*.

A estratégia H faz referência a um comportamento agressivo, do termo em inglês *Hawkish*, ao passo que a estratégia D faz referência a um comportamento pacífico, advindo do termo *Dovish*. Considerando v como benefício, e c como custo, tal que $v < c$, busca-se responder a seguinte pergunta: qual a solução desse jogo em Equilíbrio de Nash (E.N)?

Para a resposta inicialmente calcula-se o valor esperado da estratégia H e da estratégia D para o jogador A :

$$E_A(\pi_H) = \frac{v-c}{2}(q) + v(1-q)$$

$$E_A(\pi_D) = 0(q) + \frac{v}{2}(1-q)$$
(11)

De posse desses valores, e dada uma estratégia mista para o jogador **B**, é possível encontrarmos as condições em que o jogador **A** irá optar por jogar sua estratégia pura *H* ou sua estratégia pura *D*. Para tal, faz-se a correspondência entre os valores esperados de cada uma das estratégias, de modo que se $E_A(\pi_H) \geq E_A(\pi_D) \rightarrow q \leq \frac{v}{c}$. Esta relação se aplica no caso de um jogo de anti-coordenação, e como a matriz é simétrica, o mesmo vale para encontrarmos as correspondências de melhor resposta para o jogador **B**, ou seja, para **B**, temos que $E_B(\pi_H) \geq E_B(\pi_D) \rightarrow p \leq \frac{v}{c}$.

Graficamente, o mapeamento da correspondência de melhor resposta, que nos permitirá encontrar os equilíbrios de Nash, é dado por:

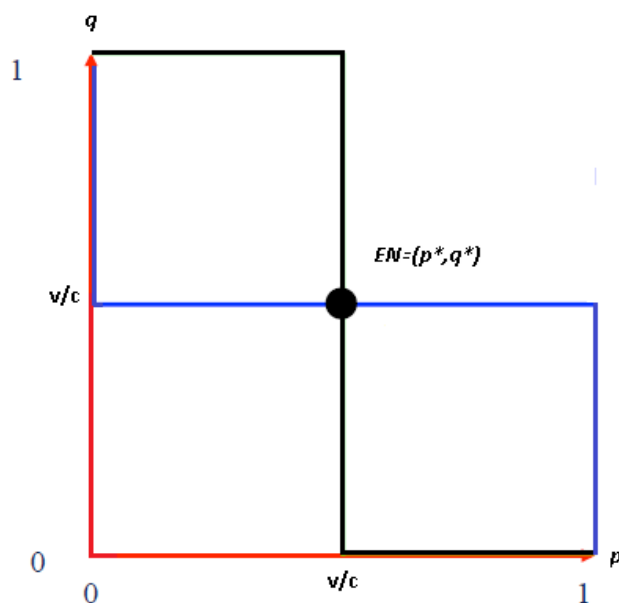


Figura 7 - Correspondência de melhor resposta no jogo Hawk – Dove.

A linha em azul representa a correspondência de melhor resposta do jogador **A** e a linha em preto a do jogador **B**. Como a matriz de pagamento do jogo é simétrica, temos que $p = q$.

O jogador **A** adota sua estratégia *H* quando $E_A(\pi_H) > E_A(\pi_D)$, isto implica em $q < \frac{v}{c}$, e o jogador **A** adota a estratégia *H* com probabilidade igual a 1. Em outras palavras, para qualquer probabilidade $q \in [0, \frac{v}{c}]$ do jogador **B** adotar a sua

estratégia H , a melhor resposta do jogador **A** é jogar sua primeira estratégia H com $p = 1$.

Se $q > \frac{v}{c}$, a melhor alternativa para o jogador **A** é jogar a estratégia D . Ou seja, ele irá alocar probabilidade zero em sua primeira estratégia, fazendo com que $p = 0$. Para quando $q = \frac{v}{c}$, o jogador **A** é indiferente entre suas estratégias, o que significa que para qualquer combinação convexa entre H e D ele estará jogando uma melhor resposta. Fica intuitivo, então, que para quando $q = \frac{v}{c}$, existe um conjunto infinito de melhores respostas e, portanto, qualquer estratégia mista garante uma melhor resposta ao que o oponente jogar.

Finalmente, encontram-se os E.N.:

- I) Dois E.N, em estratégias puras: $S = (0,1); (1,0)$, isto é, $S = (D, H), (H, D)$.
- II) Um E.N. em estratégia mista: $S = (\frac{v}{c}, \frac{v}{c})$, isto é, $S = (p^*, q^*)$.

Que são os três perfis de estratégias onde ambos os jogadores jogam uma melhor resposta à estratégia adotada pelo oponente.

3

Teoria dos Jogos Evolucionários

Muitos foram os refinamentos desenvolvidos a partir do equilíbrio de Nash, sempre pautados na premissa de que os indivíduos são dotados de um alto nível de racionalidade. Teoricamente, tais refinamentos provaram ser bastante úteis para eliminar resultados inadequados do jogo. Por outro lado, qualquer equilíbrio de Nash poderia ser justificado a partir de algum de seus refinamentos.

A Teoria dos Jogos Evolucionários, segundo apresentado por Friedman (1998), advém da biologia teórica, sendo que o jogo envolve uma ou várias populações com número muito grande de indivíduos. A teoria evolucionária assume racionalidade limitada, isto é, cada jogador está programado para escolher uma dada estratégia pura, não necessariamente a estratégia ótima, para jogar contra seu oponente. Durante um dado intervalo de tempo, pares de jogadores são selecionados de modo aleatório para competir num jogo cuja representação é da forma normal (*stage game*).

Os pagamentos representam o número esperado de descendentes de uma população e não dependem somente da função de pagamento do *stage game*, mas também do estado da população, isto é, da proporção de indivíduos adotando cada estratégia pura. Adiante, neste capítulo, será revisitado o jogo conhecido na literatura como *Hawk-Dove*, e estes conceitos serão devidamente aprofundados.

Relaxando a condição de racionalidade, existem dois grandes campos de aplicação da teoria evolucionária no âmbito de populações muito grandes ($n \rightarrow \infty$):

- (i) Jogos Evolucionários Estáticos: usam o conceito de Estratégia Evolucionariamente Estável (EEE) e permite, em alguns casos, refinar a seleção do equilíbrio de Nash. Normalmente são jogos simultâneos, jogados uma única vez. O ponto de equilíbrio é obtido nesta única interação.
- (ii) Jogos Evolucionários Dinâmicos: modelagem feita através de um sistema de equações diferenciais não lineares para inferir sobre a dinâmica evolutiva das estratégias dentro de uma população de jogadores. O replicador

dinâmico, proposto por Taylor & Jonker (1978), geralmente permite selecionar o único equilíbrio de Nash dadas as condições iniciais do jogo.

Maynard Smith & Price (1973), cujo trabalho envolve aplicação da teoria dos jogos à biologia, introduziram o conceito estático e direcionaram o foco das interações estratégicas ao longo do tempo para a dinâmica comportamental e de traços biológicos. Em outras palavras, a superioridade do comportamento da população incumbente em relação à população mutante depende do número de fenótipos (descendentes) que aquela é capaz de gerar no ambiente.

Consideremos uma população incumbente no estado $\sigma = [\sigma_1, \dots, \sigma_n]$, sendo n o número de estratégias. Se substituirmos numa fração $\varepsilon > 0$ o número de indivíduos incumbentes por mutantes que adotam as n estratégias disponíveis no jogo com probabilidade $m = [m_1, \dots, m_n]$, o estado da população fica $\tau = (1 - \varepsilon)\sigma + \varepsilon m$. O pagamento esperado para um indivíduo médio e representativo dessa população dita incumbente quando competindo num *stage game* contra um indivíduo médio da população mutante é dado por:

$$u(\sigma, \tau) = \sigma M \tau^T = (1 - \varepsilon)u(\sigma, \sigma) + \varepsilon u(\sigma, m) \quad (12)$$

O pagamento esperado para um indivíduo mutante é:

$$u(m, \tau) = m M \tau^T = (1 - \varepsilon)u(m, \sigma) + \varepsilon u(m, m) \quad (13)$$

Um mutante invade a população incumbente se para algum $\varepsilon > 0$ suficientemente pequeno, $u(m, \tau) \geq u(\sigma, \tau)$, isto é, o mutante tem uma performance ao menos tão boa quanto um indivíduo da população incumbente no ambiente competitivo da nova população em estado τ .

Seja a matriz de um jogo genérico dada por:

$$\begin{array}{cc} & \begin{array}{c} M_{(\varepsilon)} \\ I_{(1-\varepsilon)} \end{array} \\ \begin{array}{c} M_{(\varepsilon)} \\ I_{(1-\varepsilon)} \end{array} & \begin{bmatrix} \alpha; \alpha & \beta; \gamma \\ \gamma; \beta & \Delta; \Delta \end{bmatrix} \end{array}$$

Imagine que num momento inicial, a população joga a estratégia pura I e, num instante seguinte, uma fração ε dessa população passa a adotar uma estratégia alternativa M . A ideia dos EEE é assegurar que a estratégia mutante M não será capaz de invadir ou dominar a população que adota a estratégia incumbente I . Por esse motivo, deve-se comparar os *payoffs* esperados dos indivíduos que adotam as estratégias M e I .

$$i) \quad E[\pi_M] = \alpha\varepsilon + \beta(1 - \varepsilon)$$

$$ii) \quad E[\pi_I] = \gamma\varepsilon + \Delta(1 - \varepsilon)$$

É preciso garantir que $E[\pi_I] > E[\pi_M]$ quando $\lim \varepsilon \rightarrow 0$. Em palavras, o pagamento da população de indivíduos incumbentes deve ser estritamente maior que o pagamento esperado da população de mutantes no limite em que ε tende a zero.

Se $\varepsilon \rightarrow 0$, então $\Delta > \beta$ garante que $E[\pi_I] > E[\pi_M]$. Para o caso em que $\Delta = \beta$, é preciso que $\gamma > \alpha$, ou seja é necessário analisar infinitesimalmente a matriz de pagamentos para garantir que a estratégia M é EEE. Essas condições estando satisfeitas, pode-se garantir que uma estratégia é evolucionariamente estável.

Finalmente, podemos concluir que o estado da população incumbente é dito EEE se e somente:

$$a) \quad u(\sigma, \sigma) > u(m, \sigma); \forall m$$

$$b) \quad \text{Se } u(\sigma, \sigma) = u(m, \sigma) \text{ então } u(\sigma, m) > u(m, m); \forall m \neq \sigma$$

A EEE resiste às invasões de estratégias mutantes uma vez que os invasores são menos ou igualmente aptos no momento em que ocorre a interação (Rocha et al., 2011). Note, ainda, que uma EEE é um equilíbrio de Nash, mas um equilíbrio de Nash (EN) não é necessariamente uma EEE. Uma forma alternativa de definir uma EEE é dada por:

$$\Delta^{EEE} = \{\sigma \in \Delta^{EN} : u(\sigma, m) > u(m, m); \forall m \in b(\sigma), m \neq \sigma\}.$$

Anastasopoulos (2012) mostra que apesar de uma EEE ser mais robusta que o conceito de equilíbrio de Nash, trata-se de um conceito estático, não provendo uma resposta precisa sobre como essa EEE é alcançada ao longo do tempo nem qual EEE é selecionada quando mais de uma EEE persiste no jogo.

Ainda de acordo com o exposto em Anastasopoulos (2012), os processos evolucionários combinam dois mecanismos distintos:

- 1) Mecanismo de mutação, responsável pela variedade;
- 2) Mecanismo de seleção, responsável pelo favorecimento de um comportamento estratégico sobre o outro.

O replicador dinâmico, introduzido por Taylor e Jonker (1978), consiste em modelar como uma população evolui dinamicamente ao longo do tempo através de um sistema não linear composto por equações diferenciais, fornecendo a estratégia que corresponde a um equilíbrio assintótico no longo prazo.

Menasché (2005) define um conjunto de equações que compõem um sistema dinâmico que ao considerarmos $x(t) = (x_1(t), \dots, x_n(t))$ um ponto no \mathcal{R}^n , tem-se que:

$$\begin{aligned}\frac{dx_1}{dt} &= f_1(x_1, \dots, x_n) \\ \frac{dx_2}{dt} &= f_2(x_1, \dots, x_n) \\ &\vdots \\ \frac{dx_n}{dt} &= f_n(x_1, \dots, x_n)\end{aligned}$$

O ponto x^* é denominado de ponto crítico ou ponto estacionário para o caso em que $\{f_1(x^*), \dots, f_n(x^*) = f(x^*) = (0, \dots, 0)\}$, podendo ser um ponto neutramente estável, assintoticamente estável ou instável. Os pontos estacionários neutramente estáveis são suscetíveis à perturbações no sistema, enquanto os pontos com estabilidade assintótica correspondem a equilíbrios robustos e, por isso, são caracterizados como equilíbrio evolucionário.

Friedman (1991) e Hirth (2014) apresentam o replicador dinâmico como uma ferramenta analítica bastante eficiente, capaz de inferir sobre o comportamento evolutivo de uma determinada população. Considerando a função de utilidade $u(s_i, \sigma)$ o valor esperado do número de descendentes de um indivíduo que adota a estratégia s_i contra uma população cujo estado médio é $\sigma(t)$ e seja no instante de tempo t :

$N(t)$: o tamanho da população;

$N_i(t)$: o número de indivíduos que adota a estratégia s_i ;

$p_i(t)$: a proporção de indivíduos que adotam s_i , i.e. $p_i(t) = \frac{N_i(t)}{N(t)}$;

d : a taxa de mortalidade da população (constante);

$u(s_i, \sigma)$: o número esperado de descendentes de um indivíduo que adota s_i quando compete numa população em estado $\sigma(t)$.

O número de indivíduos adotando s_i em $t + 1$ é:

$$N_i(t + 1) = N_i(t) \left[1 + \sum_{j=1}^n p_j u(s_i, s_j) - d \right] \quad (14)$$

Generalizando para tempo contínuo:

$$N_i(t + \Delta t) = N_i(t) \left[1 + \Delta t \left(\sum_{j=1}^n p_j u(s_i, s_j) - d \right) \right] \quad (14.1)$$

Portanto, no limite quando $\Delta t \rightarrow 0$:

$$\begin{aligned} \dot{N}_i(t) &= N_i(t) \left[\sum_{j=1}^n p_j u(s_i, s_j) - d \right] \\ &= N_i(t) [u(s_i, \sigma) - d] \end{aligned} \quad (14.2)$$

Para a população toda:

$$N(t + \Delta t) = \sum_{i=1}^n \left\{ N_i(t) \left[1 + \Delta t \left(\sum_{j=1}^n p_j u(s_i, s_j) - d \right) \right] \right\} \quad (14.3)$$

Portanto, no limite quando $\Delta t \rightarrow 0$:

$$\begin{aligned} \dot{N}(t) &= N(t) \left\{ \sum_{i=1}^n p_i \left[\sum_{j=1}^n p_j u(s_i, s_j) - d \right] \right\} \\ &= N(t) [u(\sigma, \sigma) - d] \end{aligned} \quad (14.4)$$

Finalmente, chega-se à expressão do replicador dinâmico:

$$\dot{p}_i(t) = \frac{d}{dt} \frac{N_i(t)}{N(t)} = \frac{\dot{N}_i(t)N(t) - N_i(t)\dot{N}(t)}{[N(t)]^2} = p_i(t)[u(s_i, \sigma) - u(\sigma, \sigma)] \quad (15)$$

Uma forma de interpretar os resultados obtidos através da equação do replicador dinâmico, segundo Binmore (1992), é que quanto mais bem sucedida for uma estratégia que compete numa população em estado $\sigma(t)$, maior é o seu *payoff* e, conseqüentemente, maior é seu crescimento relativo. Em outras palavras, o replicador dinâmico fornece a taxa de crescimento da proporção de agentes que adotam determinada estratégia dentro de uma população.

Para encontrar os pontos assintoticamente estáveis, isto é, pontos de equilíbrio, partindo de um sistema de equações diferenciais não lineares, deve-se:

- i) Obter os pontos estacionários fazendo $\dot{p}_i = 0$. Esta condição é necessária, mas não suficiente, pois o ponto em análise deve satisfazer a condição proposta a seguir;
- ii) Verificar a estabilidade dos pontos estacionários encontrados, de modo que a estabilidade assintótica num sistema unidimensional, formado por uma população e duas estratégias, é obtida quando $\frac{\partial \dot{p}_i}{\partial p_i} < 0$.

No caso em que duas populações competem e cada uma delas dispõem de duas estratégias, formar-se um sistema de equações diferenciais não lineares. Por exemplo:

$$\begin{aligned}\dot{p} &= p[u(s_i, \sigma) - u(\sigma, \sigma)] \\ \dot{q} &= q[u(s_i, \sigma) - u(\sigma, \sigma)]\end{aligned}\tag{15.1}$$

Nestes sistemas, a estabilidade nos pontos candidatos ao equilíbrio evolucionário é analisada por meio da Matriz Jacobiana (Ω). Para o cálculo dos autovalores:

$$\Omega(p, q) = \begin{bmatrix} \frac{\partial \dot{p}}{\partial p} & \frac{\partial \dot{p}}{\partial q} \\ \frac{\partial \dot{q}}{\partial p} & \frac{\partial \dot{q}}{\partial q} \end{bmatrix}$$

Fazendo o determinante do Jacobiano:

$$\begin{aligned}\det(\Omega - \lambda I) &= \begin{vmatrix} \frac{\partial \dot{p}}{\partial p} - \lambda & \frac{\partial \dot{p}}{\partial q} \\ \frac{\partial \dot{q}}{\partial p} & \frac{\partial \dot{q}}{\partial q} - \lambda \end{vmatrix} = 0 \\ \lambda_{1,2} &= \frac{\text{tr}\Omega \pm \sqrt{(\text{tr}\Omega)^2 - 4\det\Omega}}{2}\end{aligned}$$

Para que o ponto estacionário seja assintoticamente estável é necessário que os autovalores $\lambda_{1,2}$ da matriz Ω avaliada no ponto estacionário tenham parte real negativa. Nas subseções a seguir, apresentamos os jogos unidimensionais e bidimensionais, cujas soluções são obtidas via replicador dinâmico.

3.1. Jogos Unidimensionais

Segundo Friedman (1991), para jogos de apenas uma população e duas estratégias disponíveis, sendo p a proporção (probabilidade) de indivíduos que adotam a primeira estratégia e^1 e $1 - p$ a proporção de indivíduos que adotam a segunda estratégia e^2 , pode-se representar a tendência de comportamento desta população ao longo do tempo através do replicador dinâmico:

$$\dot{p} = p(u(e^1, S) - u(S, S)) \quad (16)$$

Considerando ainda que a matriz normalizada $A = \begin{bmatrix} 0 & a_1 \\ a_2 & 0 \end{bmatrix}$ corresponde aos valores dos pagamentos obtidos pelas interações estratégicas no jogo e que o estado da população S é dado por $S = \{p, 1 - p\} \therefore p \in [0, 1]$, temos:

$$\begin{aligned} u(e^1, S) &= e^1 * A * S \\ u(e^1, S) &= [1 \quad 0] \begin{bmatrix} 0 & a_1 \\ a_2 & 0 \end{bmatrix} \begin{bmatrix} p \\ 1 - p \end{bmatrix} \\ u(e^1, S) &= a_1(1 - p) \end{aligned} \quad (16.1)$$

De modo análogo, calcula-se $u(S, S)$:

$$\begin{aligned} u(S, S) &= S * A * S \\ u(S, S) &= [p \quad 1 - p] \begin{bmatrix} 0 & a_1 \\ a_2 & 0 \end{bmatrix} \begin{bmatrix} p \\ 1 - p \end{bmatrix} \\ u(S, S) &= a_2 p(1 - p) + a_1 p(1 - p) \end{aligned} \quad (16.2)$$

Após algumas manipulações algébricas, o replicador dinâmico é dado por:

$$\dot{p} = p(1 - p)[a_1(1 - p) - a_2 p] \quad (16.3)$$

De posse do replicador, para encontrar o(s) equilíbrio(s) evolucionário(s), faz-se:

- 1) Pontos estacionários - $\dot{p} = 0$. Pela equação do replicador, encontra-se os seguintes pontos: $p = \left\{0, 1, \frac{a_1}{a_1 + a_2}\right\}$.
- 2) Verificar a estabilidade dos pontos estacionários, isto é, $\frac{\partial \dot{p}}{\partial p} < 0$. A ver:

$$\frac{\partial \dot{p}}{\partial p} = (1 - 2p)[a_1(1 - p) - a_2p] - p(1 - p)[a_1 + a_2]$$

$$\left. \frac{\partial \dot{p}}{\partial p} \right|_{p=0} = a_1$$

$$\left. \frac{\partial \dot{p}}{\partial p} \right|_{p=1} = a_2$$

$$\left. \frac{\partial \dot{p}}{\partial p} \right|_{p=\frac{a_1}{a_1+a_2}} = -\frac{a_1 a_2}{a_1 + a_2}$$

Como os valores de a_1 e a_2 não foram previamente definidos, é preciso analisar a estabilidade para os casos em que $a_1 \geq 0$ e $a_2 \geq 0$. Assim, temos os seguintes casos e os respectivos tipos de jogo, de acordo com o resultado da derivação do replicador dinâmico:

| Caso | Condição | Ponto de Estabilidade | Tipo de Jogo |
|------|-----------------------|-----------------------|----------------------|
| I | $a_1, a_2 > 0$ | $a_1/(a_1 + a_2)$ | Anti-Coordenação |
| II | $a_1, a_2 < 0$ | 0, 1 | Coordenação |
| III | $a_1 > 0$ e $a_2 < 0$ | 1 | Estratégia Dominante |
| IV | $a_1 < 0$ e $a_2 > 0$ | 0 | Estratégia Dominante |

Figura 8 - Análise da estabilidade e tipos de jogo.

Para o caso I, a figura abaixo ilustra a dinâmica do jogo. O equilíbrio evolucionário se dá no ponto $\frac{a_1}{a_1 + a_2}$, que representa uma estratégia mista. No longo prazo ($t \rightarrow \infty$) a proporção $\bar{p} = \frac{a_1}{a_1 + a_2}$ dos indivíduos adotará a estratégia e^1 do jogo e proporção $(1 - \bar{p})$ adotará a estratégia e^2 .

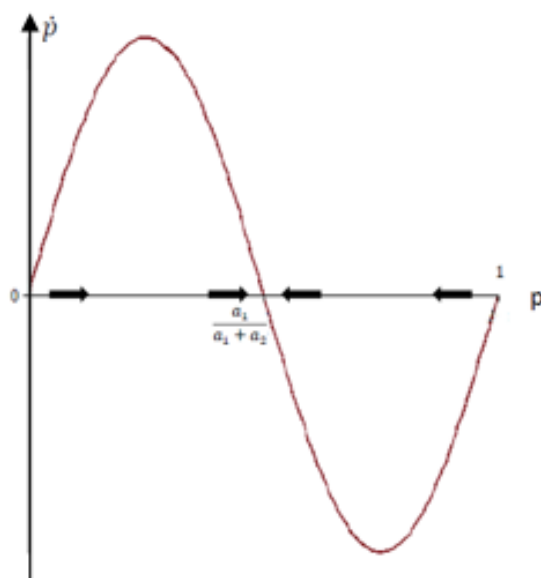


Figura 9 - Ilustração da dinâmica de um jogo de anti-coordenação.

Para o caso II, que corresponde a um jogo de coordenação, surgem dois equilíbrios evolucionariamente estáveis, conforme pode ser visto na figura 10. As derivadas do replicador são negativas nos pontos 0 e 1, ou seja, os equilíbrios se

dão em estratégias puras. O equilíbrio que ocorrerá no longo prazo ($t \rightarrow \infty$) vai depender das condições iniciais da população, isto é, se $p_0 \geq \frac{a_1}{a_1 + a_2}$.

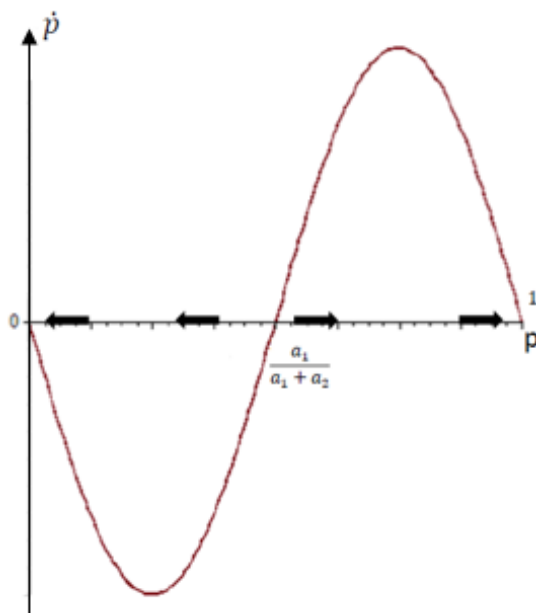


Figura 10 - Ilustração da dinâmica de um jogo de coordenação.

Já para o caso III, o jogo corresponde a um caso conhecido na literatura como jogo em estratégia dominante, isto é, independentemente do que o adversário irá jogar, existe uma estratégia dominante que irá performar sempre melhor em termos de *payoff*. É importante enfatizar que neste caso, quando $t \rightarrow \infty$, todos os indivíduos adotarão a estratégia e^1 para qualquer condição inicial $p_0 \in (0,1]$.

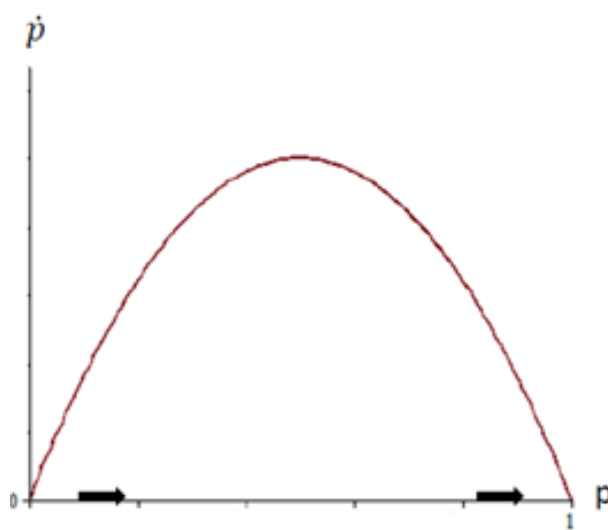


Figura 11 - Ilustração da dinâmica de um jogo com estratégia dominante.

3.2. Jogos Bidimensionais

Ainda de acordo com o exposto em Friedman (1991), em um jogo bidimensional, isto é, composto por duas populações, uma jogando contra a outra, e com duas estratégias disponíveis, S^1 corresponde ao estado da primeira população, dado por $S^1 = \{p, 1 - p\} \therefore p \in [0,1]$, e S^2 ao estado da população 2, de modo que $S^2 = \{q, 1 - q\} \therefore q \in [0,1]$. Considerando a mesma matriz A apresentada na seção de jogos unidimensionais:

$$A = \begin{bmatrix} 0 & a_1 \\ a_2 & 0 \end{bmatrix} \text{ para representar os } \textit{payoffs} \text{ da população um (jogador linha),}$$

e uma matriz

$$B = \begin{bmatrix} 0 & b_1 \\ b_2 & 0 \end{bmatrix} \text{ para representar os } \textit{payoffs} \text{ da população dois (jogador}$$

coluna), temos que;

| | | População 2 | |
|-------------|-----|-------------|------------|
| | | q | 1-q |
| População 1 | p | 0,0 | a_1, b_2 |
| | 1-p | a_2, b_1 | 0,0 |

Figura 12 - Representação matricial de um jogo bidimensional.

De acordo com os possíveis valores de a_1 , a_2 , b_1 e b_2 existem ao todo nove casos a serem analisados. Para exemplificar, apresentam-se apenas os casos I e IX.

| Caso | Condição |
|------|---|
| I | a_1, a_2 e $b_1, b_2 > 0$ |
| II | a_1, a_2 e $b_1, b_2 < 0$ |
| III | $a_1, a_2 > 0$ e $b_1, b_2 < 0$ |
| IV | $a_1, a_2 < 0$ e $b_1, b_2 > 0$ |
| V | $a_1 > 0, a_2 < 0$ e $b_1, b_2 < 0$ |
| VI | $a_1 > 0, a_2 < 0$ e $b_1, b_2 > 0$ |
| VII | $a_1, a_2 > 0$ e $b_1 < 0, b_2 > 0$ |
| VIII | $a_1, a_2 < 0$ e $b_1 < 0, b_2 > 0$ |
| IX | $a_1 > 0, a_2 < 0$ e $b_1 > 0, b_2 < 0$ |

Figura 13 - Casos possíveis para jogos bidimensionais.

Para encontrar o replicador dinâmico de cada população, faz-se primeiramente:

$$\dot{p} = p(u(e^1, S^2) - u(S^1, S^2)) \quad (17)$$

Em palavras, a frequência (p) com que os indivíduos da população 1 irão adotar a estratégia e^1 depende de quão bem ela performa contra o estado da população 2 (S^2). Logo, se e^1 for uma melhor resposta contra S^2 do que S^1 contra S^2 , a proporção de indivíduos adotando a estratégia e^1 no longo prazo aumentará. Obtém-se $u(e^1, S^2)$ e $u(S^1, S^2)$ fazendo:

$$u(e^1, S^2) = e^1 * A * S^2 = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & a_1 \\ a_2 & 0 \end{bmatrix} \begin{bmatrix} q \\ 1-q \end{bmatrix} = a_1(1-q)$$

$$u(S^1, S^2) = S^1 * A * S^2 = \begin{bmatrix} p & 1-p \end{bmatrix} \begin{bmatrix} 0 & a_1 \\ a_2 & 0 \end{bmatrix} \begin{bmatrix} q \\ 1-q \end{bmatrix} = a_2(1-p)q + a_1p(1-q)$$

Finalmente, encontra-se a equação do replicador da população 1:

$$\dot{p} = p(1-p)[a_1(1-q) - a_2q] \quad (18)$$

Pode-se escrever o replicador da população 2 do seguinte modo:

$$\dot{q} = q(1-q)[b_1(1-p) - b_2p] \quad (19)$$

Dado que \dot{p} e \dot{q} formam um sistema de equações não lineares, através da linearização do sistema, pode-se usar a Matriz Jacobiana (Ω) para estudar a vizinhança dos pontos candidatos a equilíbrio assintótico através do traço e do determinante de (Ω).

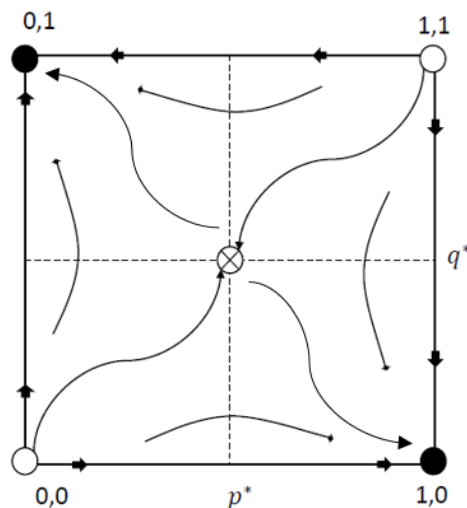


Figura 14 - Diagrama de fase do caso I.

Pelo diagrama de fase do caso I, pode-se ver que existem dois pontos assintoticamente estáveis - $(0,1)$ e $(1,0)$ – e dois pontos instáveis – $(0,0)$ e $(1,1)$. O ponto (p^*, q^*) é um ponto de sela. Este é um jogo conhecido como de anticondição, em que a melhor ação de um jogador é jogar exatamente a estratégia oposta que seu oponente escolher.

Na figura 15, através do diagrama, é possível ver que o único equilíbrio assintótico corresponde à coordenada $(1,1)$. Os pontos $(1,0)$ e $(0,1)$ são pontos de sela e o ponto $(0,0)$ é instável.

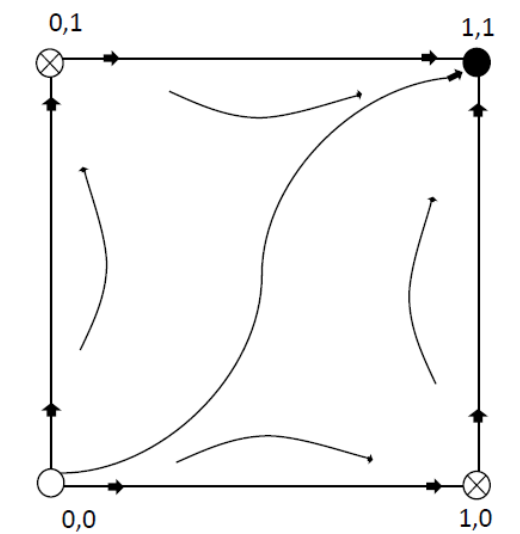


Figura 15 - Diagrama de fase do caso IX.

4

Replicação por imitação – simulação computacional

Para complementar a solução analítica fornecida pelo replicador dinâmico, faz-se uso de uma simulação computacional conhecida na literatura como *Agent Based Simulation* – replicação por imitação - que implementa estocasticidade na análise sobre o equilíbrio evolucionário. Este método de simulação computacional se mostra bastante eficiente e, conforme pode ser visto nos trabalhos de Hauert e Doebeli (2004), Xu et al. (2011) e Chan et al. (2013), apresentaram resultados para o jogo *snowdrift* que contrariam o senso comum.

Para implementar o mecanismo computacional de simulação, num instante de tempo $t = 0$, considera-se uma dada distribuição inicial da proporção dos agentes de cada população que adotam uma das estratégias puras disponíveis no jogo. A dinâmica evolutiva é introduzida em sequência e a cada Passo de Monte Carlo, um agente denominado Focal i , que pode mudar sua estratégia, é escolhido de forma aleatória dentro de uma das populações. Este agente i , por sua vez, joga contra um oponente aleatório e dá início à competição. Assim, o agente i obterá um *payoff* V_i . Ao mesmo tempo, um agente j é aleatoriamente escolhido como agente Referência j , que novamente de forma aleatória joga contra um oponente e obtém um *payoff* V_j .

O agente focal i compara V_i e V_j e analisa a possibilidade de evoluir sua estratégia em duas etapas enunciadas abaixo:

- i) Se $V_i \geq V_j$, o agente i mantém sua estratégia;
- ii) Se $V_i < V_j$ o agente i muda sua estratégia para aquela adotada pelo agente j com uma probabilidade:

$$w = \frac{V_j - V_i}{\text{payoff máximo} - \text{payoff mínimo}}; \quad (20)$$

Quando se aplica o método de simulação, o número de adversários com o que os indivíduos interagem depende do arranjo espacial do jogo. Na forma *wellmixed*, todos os jogadores são dispostos num vetor e interagem com toda a população, sendo a probabilidade de interação entre eles a mesma e independente

da sua localização no vetor. Já numa rede espacial, que será a estrutura utilizada como inspiração neste trabalho, surge o conceito de vizinhança local, e a evolução estratégica se dá conforme descrito em (Hauert e Doebeli, 2004).

A figura 16 apresenta uma ilustração da dinâmica de um jogo no formato *wellmixed* (vetor de dimensão N) e disposto em forma espacial (matrizes $N \times N$), onde também é possível identificar a vizinhança de cada jogador focal e de cada jogador referência, de acordo com sua localização na matriz.

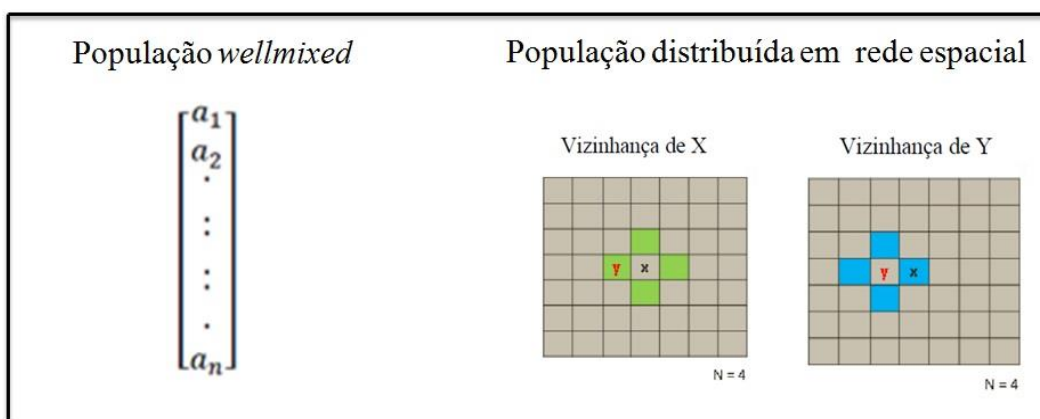


Figura 16 - Distribuição da população *wellmixed* e em rede espacial.

O exemplo a seguir ilustra como funciona o algoritmo computacional para a troca de estratégias:

| | | Adversário do Agente Focal | |
|--------------|---|----------------------------|-----|
| | | C | D |
| Agente Focal | C | 1 | 0,4 |
| | D | 1,6 | 0 |

| | | Adversário do Agente Referência | |
|-------------------|---|---------------------------------|-----|
| | | C | D |
| Agente Referência | C | 1 | 0,4 |
| | D | 1,6 | 0 |

Figura 17 - Matriz de pagamento dos agentes Focal e Referência.

A figura 17 apresenta os pagamentos de cada jogador. É através dessa matriz que se calcula a probabilidade w , que servirá de critério de decisão para saber se um indivíduo troca ou não sua estratégia. Suponha o jogo de anti-coordenação *snowdrift*, cujas estratégias são cooperar (C) e desertar (D), os pagamentos são dados conforme a figura acima. Assim, dado que o Agente Focal está programado para jogar D e seu oponente jogar C, o *payoff* V_i recebido por ele será 1,6. O agente

Referência, por sua vez, estando programado para jogar D contra um adversário que também joga D, recebe um *payoff* V_j , que é igual a zero.

Neste exemplo genérico, o jogador Focal não trocará de estratégia, pois seu *payoff* é maior que o do jogador Referência. Suponhamos que numa outra interação estratégica desse jogo, o agente Referência possa vir a imitar o Focal com uma probabilidade $w = 0,25$. Para estabelecer um critério de decisão sobre a troca ou não da estratégia por parte do agente Focal, faz-se uso de um gerador de número aleatório, aqui convenientemente chamado de *rnd*, também dentro do intervalo entre zero e um e compara-se este número gerado com a probabilidade w , de modo que:

- a) Se $w \geq rnd$, o agente Focal imita o agente Referência;
- b) Se $w < rnd$, o agente Focal não troca sua estratégia.

Deste modo, cria-se um componente estocástico sobre a dinâmica do jogo. A simulação por agente vem sendo muito empregada em jogos clássicos da literatura – como no Dilema do Prisioneiro e no *Snowdrift Game* – e tem contribuído para o debate acadêmico sobre Teoria dos Jogos, pois vêm apresentando resultados que desafiam a intuição e o senso comum, conforme abordado por Chan et al. (2013), Xu et al. (2011) e Hauert e Doebeli (2004).

A próxima subseção apresenta a metodologia proposta na replicação por imitação de uma população distribuída espacialmente (Hauert e Doebeli, 2004) e, em seguida, cita-se a extensão para duas populações conforme realizado por (Arruda, 2015).

4.1. O Jogo Espacial

Neste tipo de jogo, é considerada uma população que, em seu estágio inicial, é composta por indivíduos que estão dispostos a jogar cada uma das estratégias puras disponíveis. O jogo espacial possui um número muito grande de indivíduos, sendo que, conforme apresentado em (Hauert e Doebeli, 2004), a população de agentes é representada por uma matriz de tamanho $N \times N$, só havendo interação entre um número restrito de jogadores, de acordo com uma vizinhança pré-estabelecida.

A cada passo de Monte Carlo os indivíduos da população, selecionados aleatoriamente, têm a oportunidade de, em média, trocar de estratégia pelo menos uma vez, comparando seu *payoff* com o de seus vizinhos. Dizemos que, na média, cada indivíduo pode trocar de estratégia uma vez, pois dentro de um passo de Monte Carlo o mesmo indivíduo pode ser convidado a jogar mais de uma vez, e outros indivíduos podem não ser selecionados, já que o processo de escolha dos jogadores é aleatório. Sendo assim, quando todos os indivíduos da matriz $N \times N$, em média, tiverem a oportunidade de atualizar suas estratégias, dá-se por encerrado um passo de Monte Carlo e um novo passo se inicia de forma a repetir a dinâmica do jogo. Este procedimento é conhecido como replicação por imitação ou *Agent Based Simulation*. A seguir será explicado mais detalhadamente o processo de composição da vizinhança local e o raio de interação estratégica dos jogadores.

4.2. A vizinhança de *von-Neumann*

A vizinhança de *von-Neumann* advém de um modelo discreto estudado na matemática aplicada às Ciências Biológicas e à Teoria da Recursão. Tal vizinhança consiste de quatro células dispostas ortogonalmente ao redor da chamada célula central, conforme figura abaixo:

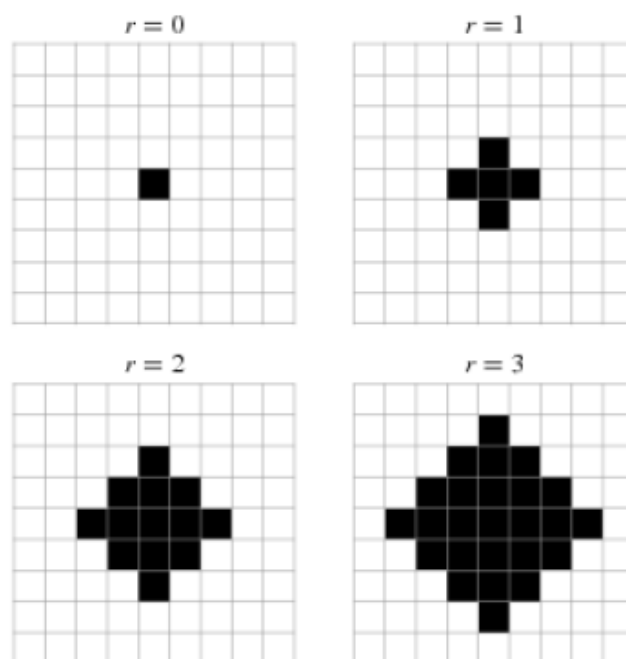


Figura 18 - Vizinhança de *von-Neumann*.
Fonte: Gray (2003).

Nos trabalhos de Hauert e Doebeli (2004) e Chan et al. (2013), que serviram de inspiração para esta dissertação, também são utilizadas a vizinhança de *von-Neumann* de raio $r = 1$ para a execução das simulações computacionais da competição entre indivíduos de uma única população. Em outras palavras, podemos explicar a vizinhança em que $r = 1$ como aquela em que o agente focal interage com seus vizinhos imediatos.

Deve-se ter uma atenção especial ao considerar as interações estratégicas daqueles jogadores situados na fronteira (ou nas bordas) do arranjo espacial em que se aloca a população. Um indivíduo Focal representado na cor azul na figura abaixo terá como vizinho os jogadores do tipo Referência que estão representados pela cor vermelha, admitindo que o jogo se dá numa estrutura espacial que possui “*periodic boundary conditions*”.

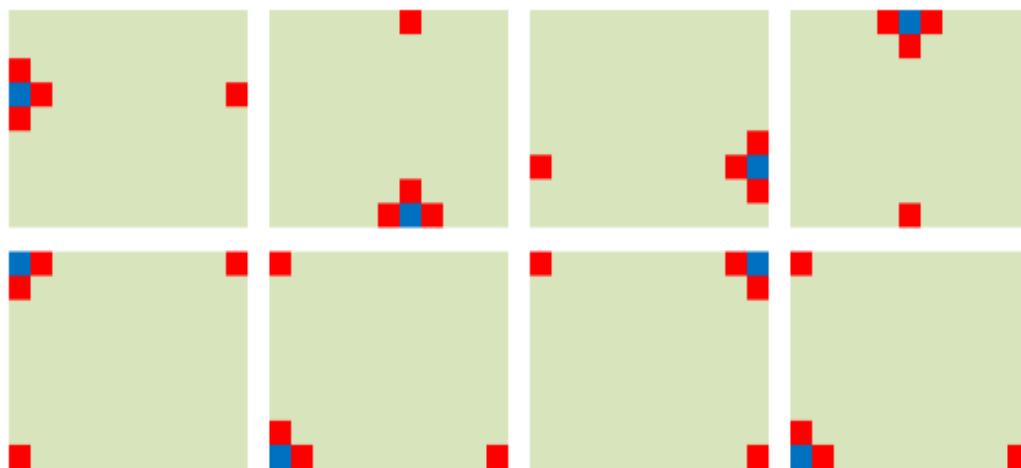


Figura 19 - Vizinhança de von-Neumann para indivíduos situados nas fronteiras.
Fonte: Arruda (2015).

4.3. Os Métodos Síncrono e Assíncrono.

Existe na literatura, dois mecanismos de atualização da população no que diz respeito às interações estratégicas. De acordo com Hauert (2001), na forma de atualização síncrona, a população é atualizada simultaneamente no final do passo de Monte Carlo, isto é, após todos os indivíduos terem, em média, a oportunidade de jogar com seus vizinhos locais uma vez.

Na forma de atualização assíncrona, a qual será utilizada neste trabalho, existe a possibilidade de se imitar a estratégia do vizinho local imediatamente após a interação. Fala-se, na literatura, que a atualização assíncrona é aquela que permite

a intersecção entre gerações de indivíduos, conhecida como *Overlapping Generations*, pois dentro de um mesmo passo de Monte Carlo existirão jogadores que já atualizaram sua estratégia e outros que ainda não tiveram a oportunidade de imitar o comportamento daquele vizinho - a ser utilizado como referência – e que performa melhor dentro do ambiente competitivo.

Como já introduzido no início deste capítulo, o critério de atualização das estratégias passa pelo cálculo do *payoff* médio de cada indivíduo, que é encontrado pela média aritmética dos pagamentos obtidos em cada interação com os n jogadores que compõem a vizinhança local:

$$V_i = \frac{\sum_{i=1}^n v_i}{n} \quad (21)$$

Logo, o jogador Focal compara o seu pagamento V_i com o pagamento médio do jogador Referência V_j , que é aleatoriamente selecionado dentre os vizinhos locais do Focal. O valor de V_j é obtido como resultado das interações estratégicas que o jogador referência faz com seus vizinhos locais – incluindo nesta vizinhança o jogador Focal. O pagamento médio V_j é dado por:

$$V_j = \frac{\sum_{j=1}^n v_j}{n} \quad (22)$$

De posse dos valores dos pagamentos médios de que cada um dos agentes, deve-se realizar o procedimento de comparação entre eles, sendo que caso $V_i \geq V_j$, o jogador Focal não imita o jogador Referência.

Caso $V_j > V_i$, o jogador Focal pode vir a imitar o jogador Referência com a probabilidade w . Note que $w \in [0,1]$, e quanto pior for a performance do jogador Focal em relação ao Referência, maior é a probabilidade daquele imitar este.

5 Modelo

Com o objetivo de verificar a dinâmica espacial e o comportamento dos agentes quando inseridos no *Snowdrift Game* (SG) com a presença de punidores em duas populações com arranjo espacial que são concorrentes, de modo que não existe competição entre indivíduos da mesma população, baseou-se o modelo a ser apresentado nesta seção naquele proposto por Chan et al. (2013), em que foi exposta a dinâmica evolutiva do jogo SG com punidores para apenas uma população. No que se refere ao delineamento do espaço competitivo entre duas populações distribuídas em arranjos espaciais, adotou-se o critério iterativo desenvolvido por Arruda (2015), a ser melhor explicado na subseção 5.4.

5.1. O *Snowdrift Game*

No *Snowdrift Game* (SG), dois motoristas (jogador 1, jogador 2) estão guiando em direções opostas e se deparam com uma estrada bloqueada por um desmoronamento de neve. Eles possuem duas possíveis ações a serem escolhidas simultaneamente: cavar a neve (cooperar, C) ou não cavar (não cooperar, desertar D). Dessas ações, surgem os possíveis resultados:

- iv) (C, C) – ambos recebem o benefício b . Porém, cavar tem um custo c , que no caso em que ambos cooperam, será dividido. Cada um dos jogadores recebe o pagamento: $R = b - \frac{c}{2}$;
- v) (D, D) – ambos ficam presos e não recebem nada por isso. $P = 0$;
- vi) (C, D) ou (D, C) – ambos vão para casa. Porém, o que jogou D vai para casa sem fazer qualquer esforço, recebendo um pagamento $T = b$, enquanto o seu oponente, que jogou C, recebe $S = b - c$, absorvendo todo o custo de cavar. No SG $b > c > 0$, o que leva a um ranking de pagamentos: $T > R > S > P$.

| SG | C | D |
|----|-------------|-------------|
| C | R, R | <u>S, T</u> |
| D | <u>T, S</u> | P, P |

Figura 20 - Equilíbrios de Nash em estratégias puras no jogo SG.

Trata-se de um jogo de anti-coordenação, em que a melhor resposta de um jogador é adotar a estratégia oposta àquela escolhida por seu adversário. Os equilíbrios de Nash em estratégias puras são, portanto, (C, D) e (D, C).

5.2.

Mecanismos de cooperação no SG – a presença do punidor

Para incorporar a punição no SG, o trabalho de Chan et al. (2013) introduziu uma 3ª estratégia (P) no jogo, que imita a presença dos agentes (jogadores) que jogam C. A diferença entre eles é que os que adotam a estratégia (P) estão dispostos a abrir mão de parte de seu *payoff*, ganhando um pouco menos, para que os agentes desertores (D) sejam penalizados, tendo um desconto ainda maior em seus pagamentos e, conseqüentemente, tenham um menor incentivo a não cooperar. A matriz do SG com a presença do Punidor (P) é dada por:

$$\begin{array}{c}
 \begin{array}{c}
 C \\
 D \\
 P
 \end{array}
 \begin{pmatrix}
 \begin{array}{ccc}
 C & D & P \\
 \begin{array}{c} R \\ T \\ R \end{array} & \begin{array}{c} S \\ P \\ S - \alpha \end{array} & \begin{array}{c} R \\ T - \beta \\ R \end{array}
 \end{array}
 \end{pmatrix}
 \end{array}$$

Figura 21 - Matriz do jogo *Snowdrift* com o Punidor.
Fonte: Chan et al. (2013).

O parâmetro α é a quantia que o *punisher* (P) está disposto a abrir mão em seu pagamento para que um oponente desertor (D) tenha um desconto (uma penalização) de β em seu *payoff*. A condição $\alpha < \beta$ implica que o *punisher* (P) permanece sendo parcialmente racional no jogo evolucionário, ao passo que a punição causa uma perda maior no *payoff* dos desertores (D). Quando um jogador que adota a estratégia C encontra outro que esteja jogando P, o resultado é o mesmo para o caso em que dois jogadores utilizando a estratégia C se encontram. Ambos recebem o mesmo *payoff* R . Podemos simplificar a matriz de pagamentos assumindo $R = 1$ e $P = 0$. Outra simplificação é feita ao introduzir um parâmetro de

custo benefício r ($0 < r < 1$), de modo que $S = 1 - r$ e $T = 1 + r$, conforme definido na literatura por Doebeli e Hauert (2005). A matriz passa a ser representada por:

$$\begin{array}{c} C \\ D \\ P \end{array} \begin{array}{ccc} C & D & P \\ \left(\begin{array}{ccc} 1 & 1 - r & 1 \\ 1 + r & 0 & 1 + r - \beta \\ 1 & 1 - r - \alpha & 1 \end{array} \right) \end{array}$$

Figura 22 - Matriz de payoff do SG com o Punidor.
Fonte: Chan et al. (2013).

Os autores consideraram um sistema com N ($N \gg 1$) agentes. Há, ainda, uma distribuição uniforme inicial de proporção das três estratégias na população, ou seja, a frequência de desertores, cooperadores e punidores é a mesma e igual a $1/3$.

5.3.

Projeto Piloto – Comparativo com resultados de Chan et al. (2013)

Antes de iniciar a apresentação do modelo desenvolvido neste capítulo, fez-se necessária a execução de um projeto piloto, com a finalidade de estudar os efeitos do jogo SG espacial em uma população com a presença do agente punidor e replicar o resultado já consagrado na literatura. Isto visou, dentre outras coisas, a validar o código de programação em linguagem JAVA desenvolvido especialmente para este trabalho. Utilizou-se a vizinhança de *von-Neumann* e uma população alocada numa matriz de dimensão 90×90 . Os valores de α e β foram definidos em 0,01 e 0,05, do mesmo modo que Chan et al. (2013). Considerou-se os três casos apresentados pelo autor, a serem divididos pelo valor do parâmetro r , que é a razão entre o custo pelo benefício de cooperar.

(i) Caso em que $r = 0,02$

É possível perceber que nos estágios iniciais do jogo, quando o custo pelo benefício da cooperação é extremamente baixo, a frequência de desertores diminui enquanto as frequências de punidores e cooperadores são praticamente as mesmas e aumentam na mesma proporção até o passo de Monte Carlo (ou passo Evolucionário) de número 100. Isso acontece por conta da existência de desertores numa proporção suficiente para que estes estejam distribuídos ao redor de outros

desertores, o que resulta num *payoff* baixo quando um jogador programado para jogar D encontra com outro que também esteja programado para jogar D. Logo, os jogadores atualizam suas estratégias de D para C ou para P.

Após o centésimo passo de Monte Carlo, a frequência de desertores continua a cair e, além das trocas diretas dos desertores para o perfil cooperador, também ocorrem trocas indiretas para cooperadores, isto é, através da sequência $D \rightarrow P \rightarrow C$. De acordo com Chan et al. (2013), o que justifica a ascensão da frequência de cooperadores e a estabilização do número de punidores, é o fato de que num ambiente competitivo em que haja o mesmo número de desertores, a estratégia P performa pior contra D do que C contra D. Logo, quando um indivíduo P compara seu desempenho com um vizinho C, este último poderá ter o seu comportamento imitado por aquele.

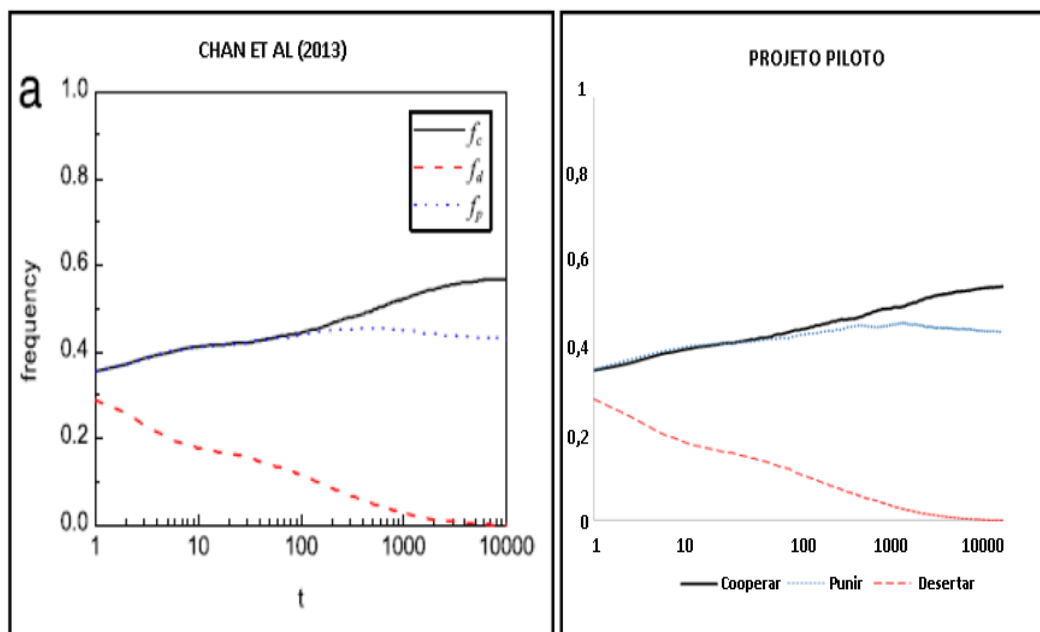


Figura 23 - Comparativo entre evolução das estratégias em Chan et al. (2013) e o projeto piloto para $r = 0,02$.

Este mecanismo fica ainda mais aparente após o milésimo passo evolucionário, cessando somente quando a estratégia D é extinta nessa população. Pela figura 23, é possível ver que o resultado obtido pelo projeto piloto está de acordo com obtido por Chan et al. (2013).

(ii) Caso em que $r = 0,20$

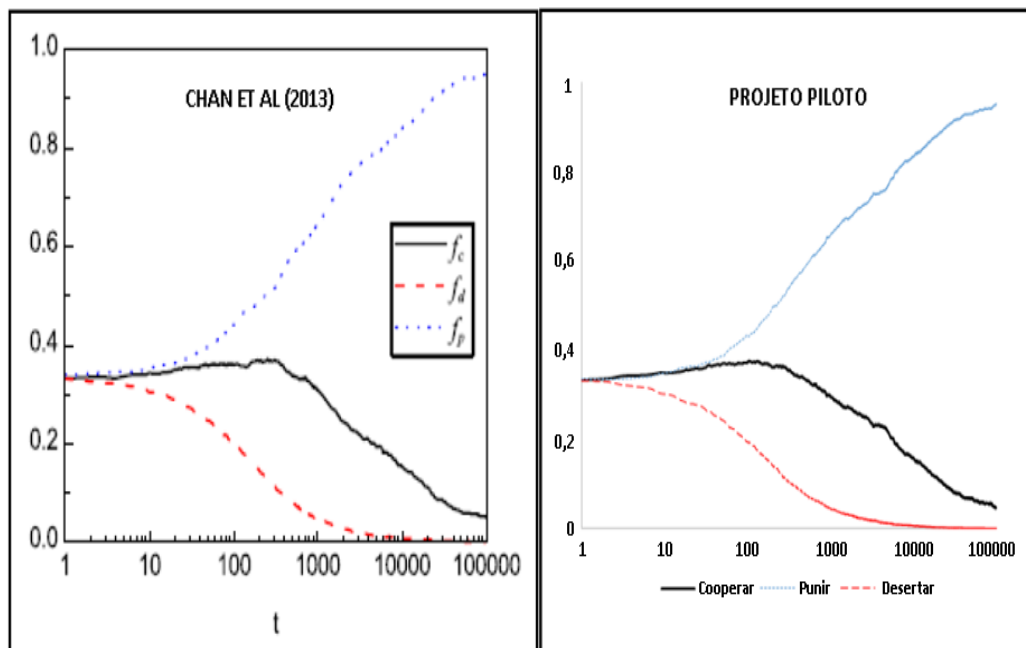


Figura 24 - Comparativo entre evolução das estratégias em Chan et al. (2013) e o projeto piloto para $r = 0,20$.

Neste caso, a frequência de indivíduos cooperadores perde força na medida em que seus *payoffs* se tornam menores quando confrontam com indivíduos desertores. Há um rápido aumento da frequência de punidores acompanhado por um rápido decréscimo da frequência de cooperadores, enquanto a frequência de desertores permanece baixa. As primeiras trocas envolvendo os desertores ocorrem nas regiões em que eles possuem outros desertores como vizinhos e utilizam como agente Referência um cooperador ou punidor. Logo, haverá trocas do tipo desertor para cooperador ($D \rightarrow C$) e do tipo desertor para punidor ($D \rightarrow P$).

Os cooperadores existentes no sistema e que estão situados próximo aos desertores – o que piora o *payoff* daqueles programados para jogar C – poderão imitar tanto os desertores ou punidores que pertencem ao seu ambiente competitivo e, também, ocorrem trocas do tipo cooperador, desertor e punidor, isto é, $C \rightarrow D \rightarrow P$, o que resulta no domínio dos punidores ao longo do tempo. A população é formada no longo prazo por Cooperadores e Punidores, uma dinâmica que se distingue da modelagem realizada por Xu et al. (2011) que trata o mesmo jogo com um arranjo do tipo *wellmixed*.

(iii) Caso em que $r = 0,60$

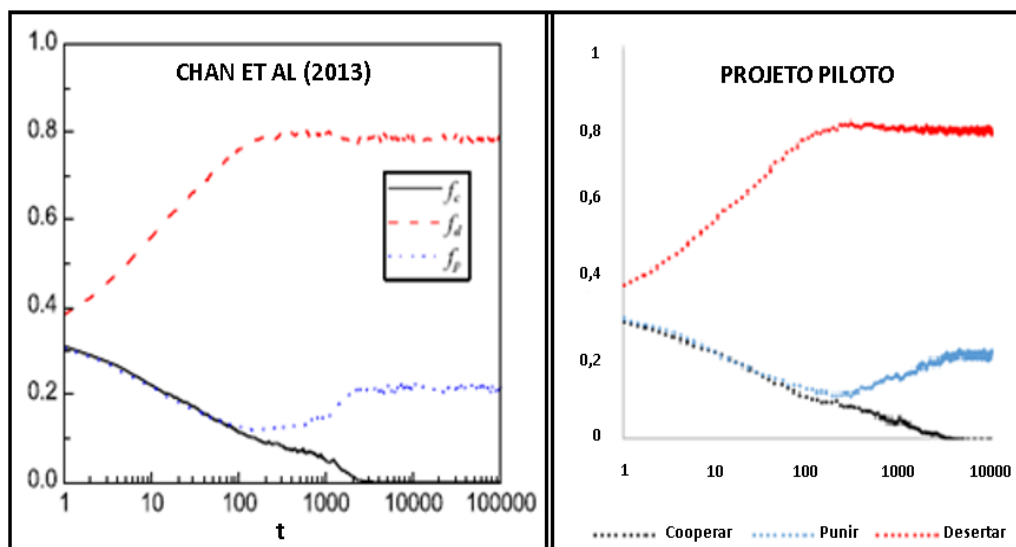


Figura 25 - Comparativo entre evolução das estratégias em Chan et al. (2013) e o projeto piloto para $r = 0,60$.

No primeiro momento, o alto valor do custo pelo benefício da cooperação leva a uma dinâmica intensa de troca entre cooperadores e punidores para desertores. Por volta do centésimo passo de Monte Carlo, onde a frequência de desertores é bastante elevada, a competição entre um desertor com outro desertor é também mais frequente – o que gera o menor *payoff* do jogo.

Então, a partir do passo de Monte Carlo 100, o efeito líquido observado é uma pequena proporção de punidores e o desaparecimento dos cooperadores. Chan et al. (2013) mostra que para $r \gtrsim 0,67$, o efeito da punição se torna desprezível e há formação de uma população homogeneamente formada por desertores.

5.4

O *Snowdrift* com punidor em duas populações espaciais

Inspirado no modelo de Arruda (2015), em que o autor analisou um jogo entre uma população de agências de rating de crédito que interage com uma população de investidores, com distribuição espacial, aqui apresenta-se uma extensão para o jogo *Snowdrift* com punidores, inicialmente proposto por Chan et al. (2013), num ambiente em que os indivíduos competem numa única população.

Assim, duas matrizes de tamanho $N \times N$ são criadas contendo as estratégias jogadas, isto é, alocam-se os jogadores programados para jogar C, D e P, de forma aleatória, com as frequências uniformes e iguais a $1/3$.

| População 1 | | | | | População 2 | | | | |
|-------------|---|---|---|---|-------------|---|---|---|---|
| D | P | P | C | P | P | C | P | C | D |
| D | C | C | C | P | P | D | D | C | C |
| P | D | C | C | D | P | D | P | C | P |
| C | D | D | D | C | D | C | P | C | D |
| D | P | P | P | D | D | P | P | D | C |

Figura 26 – Ilustração da inicialização das Populações 1 e 2.

Após inicializar a população, o programa percorre as matrizes aleatoriamente, selecionando os agentes Focal (em azul) e Referência (em cinza) em cada população. Posteriormente, são calculados o pagamento obtido por eles ao competirem com seus vizinhos. Note que os jogadores Focal e Referência pertencem à mesma população, sendo que o Referência é um dos quatro elementos que compõe a vizinhança de *von Neumann* (apresentada na subseção 4.2) do jogador Focal.

| População 1 | | | | | População 2 | | | | |
|-------------|---|---|---|---|-------------|---|---|---|---|
| D | P | P | C | P | P | C | P | C | D |
| D | C | C | C | P | P | D | D | C | C |
| P | D | C | C | D | P | D | P | C | P |
| C | D | D | D | C | D | C | P | C | D |
| D | P | P | P | D | D | P | P | D | C |

Figura 27 - Ilustração do processo de escolha dos agentes Focal e Referência.

O indivíduo de uma população irá competir com o indivíduo que está situado espacialmente na mesma posição que ele, porém na população rival, mais os indivíduos que compõem a vizinhança de *von Neumann* do oponente, isto é, aqueles jogadores localizados acima, abaixo, à direita e à esquerda. Em outras palavras, ocorre simultaneamente interações da população 1 com a população 2 e, também, da população 2 com a população 1. Além de competir com os 4 vizinhos locais, também haverá competição com o indivíduo central, que ocupa a mesma posição geográfica que o agente focal na população adversária. A ilustração da

figura 28 ajuda a visualizar a dinâmica do jogo. Sendo possível ver que cada agente jogará com cinco adversários.

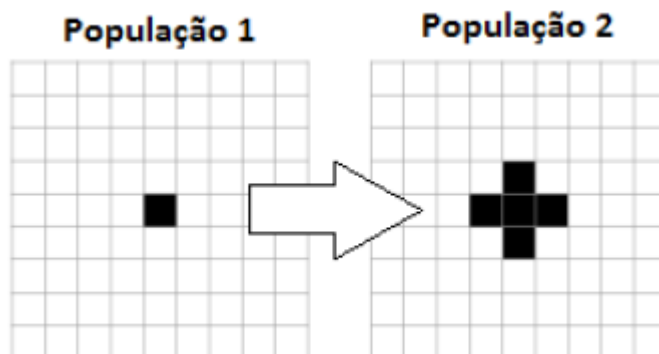


Figura 28 - Cálculo do pagamento para 2 populações.
Fonte: Arruda (2015).

Após o cálculo do pagamento de cada jogador, há uma comparação entre os indivíduos. O procedimento de comparação é idêntico ao explicado para o jogo de uma população. Note que o jogador Focal se compara com o Referência, podendo imitá-lo com uma dada probabilidade w . Na figura 29, através de um exemplo pontual, é apresentado o procedimento computacional para a atualização de estratégias. A matriz do jogo é dada abaixo, considerando $r = 0,6$.

| População 1 | População 2 | |
|--|-----------------|---|
| [p][p][c][c][c] | [d][c][p]*d*[p] | |
| [d][c][d][p][d] | [d][c][d][d][d] | |
| >d<[d][p][p]*p* | [c][c][p][p][p] | |
| [c][p][d][p][p] | [p][c][d][d][p] | |
| [c][d][d][c][c] | [p][p][c]>c<[p] | |
| | | $ \begin{matrix} & C & D & P \\ C & \begin{pmatrix} 1 & 0,4 & 1 \end{pmatrix} \\ D & \begin{pmatrix} 1,6 & 0 & 1,55 \end{pmatrix} \\ P & \begin{pmatrix} 1 & 0,39 & 1 \end{pmatrix} \end{matrix} $ |
| ----- | | |
| Focal (População 1) | : [3][5] p; | payoff: 0.878 |
| Referencia (População 1) | : [3][1] d; | payoff: 1.26 |
| w | : 0.2387500 | |
| numero aleatorio pra comparar com w: | 0.8348215 | |
| trocou de estrategia | : false | |
| ----- | | |
| Focal (População 2) | : [1][4] d; | payoff: 1.59 |
| Referencia (População 2) | : [5][4] c; | payoff: 0.88 |
| retornando antes de calcular 'w' por causa dos valores de payoff | | |

Figura 29 - Exemplo do procedimento computacional para atualização de estratégias.

Um agente Focal escolhido aleatoriamente na população 1, situado na 3ª linha e na 5ª coluna, programado para punir, irá competir com seus vizinhos locais da população 2 e, também, com aquele jogador que ocupa a 3ª linha e 5ª coluna na população 2. Assim, o seu *payoff* é dado por:

$$Payoff\ Focal\ (P) = \frac{4(1) + 0,39}{5} = 0,878.$$

Este agente Focal irá se comparar com o agente Referência, que no exemplo da figura 29 está situado na 3ª linha e na 1ª coluna, programado para desertar. O *payoff* do Referência é dado por:

$$\text{Payoff Referência (D)} = \frac{2(1,55) + 2(1,6)}{5} = 1,26.$$

Como o Focal tem um *payoff* menor que o Referência, há probabilidade de troca de estratégia entre eles, que pode ser calculada por:

$$w = \frac{1,26 - 0,878}{1,6} = 0,23875.$$

Note que o programa faz o mesmo procedimento para a população 2, computando o pagamento dos agentes e, sempre que o Referência performar melhor que o Focal, este poderá imitar aquele com uma probabilidade w . Compara-se, portanto, o valor de w com rnd . Se $w < rnd$, o agente Focal não imita o agente Referência. Na próxima seção serão apresentados os resultados obtidos através da extensão deste modelo. Como será visto, o equilíbrio assintótico depende do valor do parâmetro r , sugerindo que as estratégias que sobrevivem no longo prazo estão correlacionadas com a razão entre o custo pelo benefício da cooperação.

6 Resultados

Os resultados estão divididos em três casos e foram obtidos a partir de uma matriz 90 x 90. A frequência inicial de C, D e P foi de 1/3, o parâmetro $\alpha = 0,01$ e $\beta = 0,05$. No caso I, em que o parâmetro $r = 0,02$, observou-se a coexistência de cooperadores e punidores em cada uma das populações. Este resultado só ocorreu para simulação feita com este parâmetro. Para os casos em que $r = 0,2$ e $r = 0,6$, o que foi observado é a formação de uma população toda de cooperadores (*All C*) e outra toda de desertores (*All D*), não havendo existência dos agentes punidores.

Caso I) $r = 0,02$

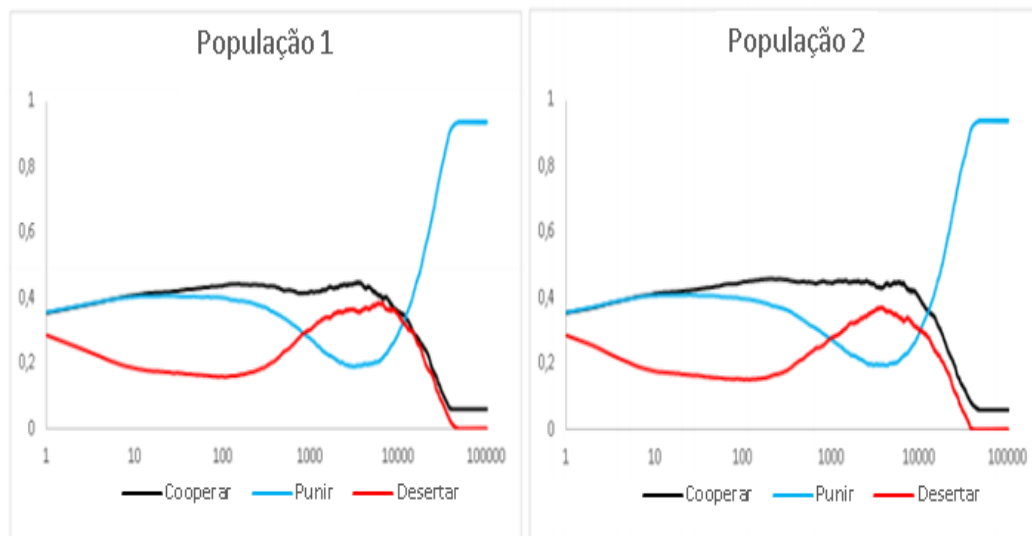
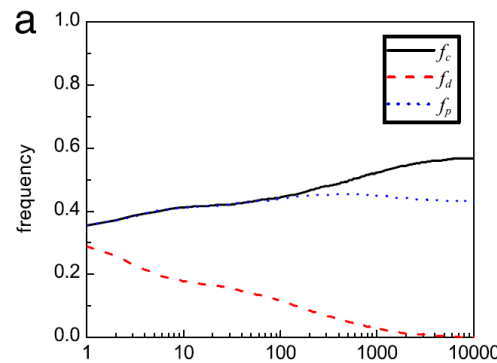


Figura 30 - *Snowdrift* com agente punidor em 2 populações com distribuição espacial. Parâmetro $r = 0,02$.

A figura acima corresponde a uma realização feita com o parâmetro $r = 0,02$. A frequência final de punidores (f_P) e cooperadores (f_C) foi, aproximadamente, de 0,90 e 0,10. Foram realizadas 30 simulações deste mesmo cenário e o comportamento evolutivo se deu de forma bastante similar, isto é, o que se verificou para este caso é um equilíbrio evolucionário composto por uma

população majoritariamente de punidores e por poucos cooperadores. Este resultado se difere daquele observado em Chan et al. (2013), que encontrou para este caso um equilíbrio formado por uma maioria de cooperadores e, aproximadamente, 40% de punidores.



**Figura 31 - Equilíbrio assintótico obtido por Chan et al. (2013) para $r = 0,02$.
Fonte: Chan et al. (2013).**

Os desertores desaparecem do jogo, o que significa que a presença dos punidores foi capaz de inibir a proliferação deste comportamento em ambas as populações. Para a compreensão do processo evolutivo, apresentam-se os *clusters* (subpopulações ou ilhas) com a distribuição espacial do jogo. Sendo que as cores seguem a legenda da figura 30, isto é, os cooperadores são representados pela cor preta, os desertores pela cor vermelha e os punidores pela cor azul.

Nos passos de Monte Carlo iniciais, isto é, $t \lesssim 10$, tanto na população 1 quanto na população 2, a frequência de desertores diminui e a frequência de punidores e cooperadores aumentam em proporções bastante similares.

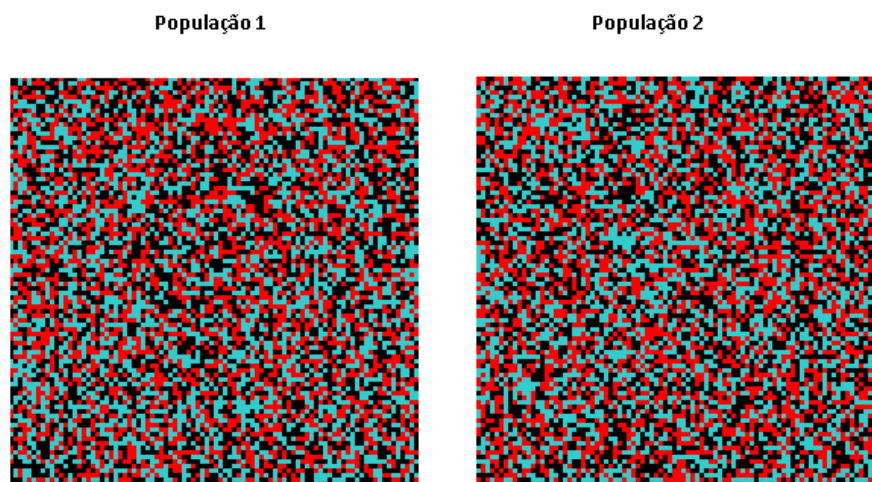


Figura 32 - Passo de Monte Carlo 1; $r = 0,02$.

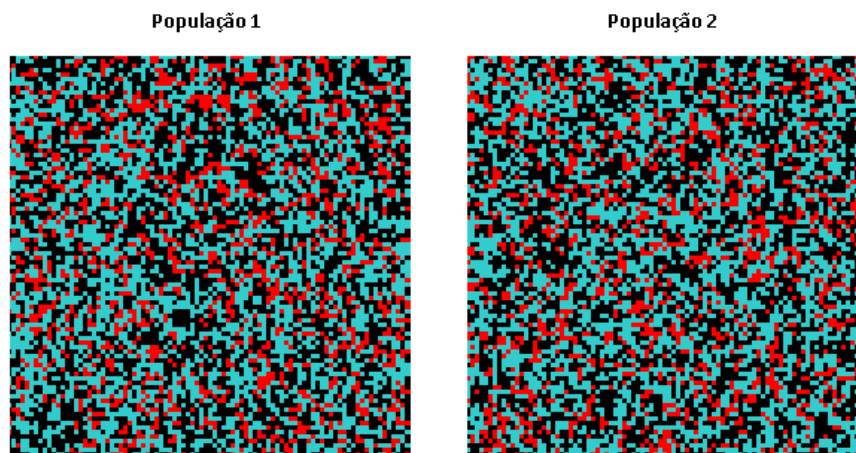


Figura 33 – Passo de Monte Carlo 10; $r = 0,02$.

Nas duas populações há um número suficiente de desertores de modo que eles podem pertencer à vizinhança de outros indivíduos desertores, o que impacta negativamente o *payoff* esperado do jogador que está programado para jogar a estratégia D.

Para exemplificar o que se passa neste estágio do jogo, suponha que na população 1 o jogador D é o Focal e C é o Referência e, na população 2, estão os adversários desses jogadores.

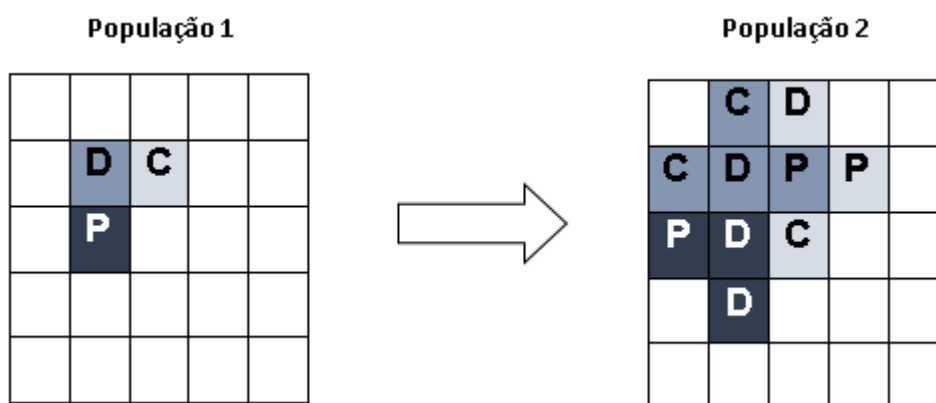


Figura 34 – Exemplo do pagamento obtido pelos jogadores Focal e Referência; $r = 0,02$.

Note que no exemplo a vizinhança competitiva do jogador Focal D é composta por dois cooperadores, dois desertores e um punidor. Já a do jogador Referência C é composta por dois punidores, dois desertores e um cooperador. De acordo com a matriz de pagamento formada a partir das interações estratégicas entre os jogadores, temos:

| | | | |
|---|------|------|------|
| | C | D | P |
| C | 1 | 0,98 | 1 |
| D | 1,02 | 0 | 0,97 |
| P | 1 | 0,97 | 1 |

Figura 35 - Matriz de pagamento para o caso em que $r = 0,02$, $\alpha = 0,01$ e $\beta = 0,05$.

É possível calcularmos os pagamentos dos jogadores Focal e Referência e avaliar se haverá a possibilidade de o agente desertor imitar o comportamento do agente cooperador:

$$\text{payoff do Focal (D)} = \frac{2(1,02) + 0,97}{5} = 0,602$$

$$\text{payoff do Referência (C)} = \frac{3(1) + 2(0,98)}{5} = 0,992$$

$$w = \frac{0,992 - 0,602}{1,02} \cong 0,38$$

Note que a probabilidade w de troca de estratégia é de aproximadamente 0,38. Agora, para concluir o exemplo, suponhamos que um agente D é jogador Focal e P é o Referência.

$$\text{payoff do Focal (D)} = \frac{2(1,02) + 0,97}{5} = 0,602$$

$$\text{payoff do Referência (P)} = \frac{3(0,97) + 2(1)}{5} = 0,988$$

$$w = \frac{0,988 - 0,602}{1,02} \cong 0,38$$

Como visto, a probabilidade w de um jogador Focal (D) imitar um Referência programado para jogar (C) ou (P) é praticamente a mesma nas circunstâncias descritas acima. Isso implica que a troca de estratégias ocorre praticamente com a mesma frequência.

Entre $10 \lesssim t \lesssim 150$, a frequência de desertores continua numa trajetória decrescente, acompanhada por um aumento do número de cooperadores nas duas populações. Neste intervalo de tempo, ocorrem trocas de desertores para cooperadores e punidores, e de punidores para cooperadores.

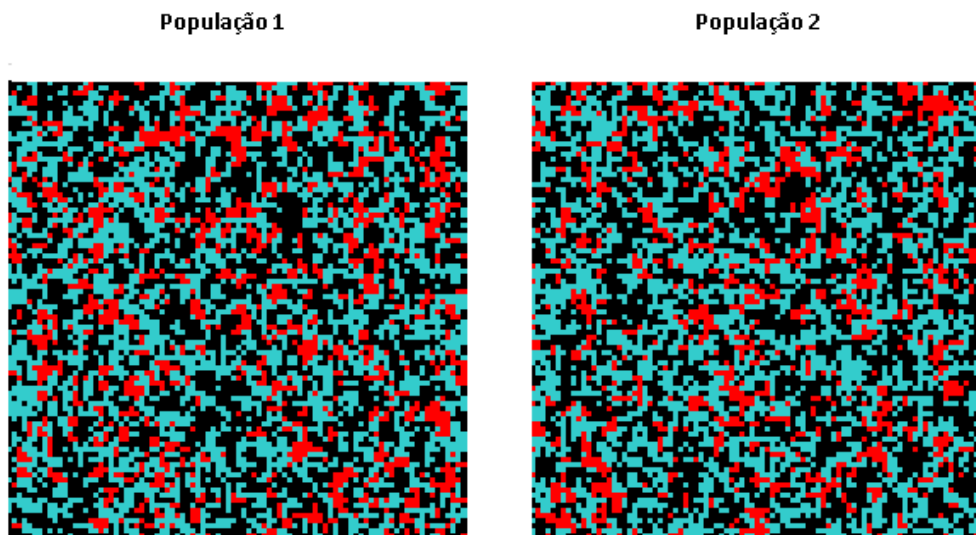


Figura 36 - Passo de Monte Carlo 150; $r = 0,02$.

Trocas do tipo $D \rightarrow C$ se tornam mais frequentes e, nas duas populações, a frequência de D é relativamente menor (aproximadamente 20%), e a distribuição de P e C é aproximadamente a mesma em ambas as populações. Isto é, temos 40% de punidores e 40% de cooperadores. Logo, em aproximadamente 80% dos casos:

$$E(\pi_P) = E(\pi_C) = \frac{1}{2} * (1 + 1) = 1$$

$$E(\pi_D) = \frac{1}{2} * (1,02 + 0,97) < 1$$

Em palavras, na situação descrita acima, o pagamento esperado dos punidores, denotado por $E(\pi_P)$, e o pagamento esperado dos cooperadores, denotado por $E(\pi_C)$, é maior que o pagamento esperado dos desertores, denotado por $E(\pi_D)$.

Há, também, um ligeiro favorecimento de C ante D, dado que através da matriz de pagamentos, é possível inferirmos que quando os vizinhos C e P estão num ambiente competitivo similar, com o mesmo número de desertores D em suas vizinhanças, a estratégia P performa pior contra D do que a estratégia C contra D. Isso acontece pela existência do custo de punição. Logo, os indivíduos que adotam P contra D passam a imitar os cooperadores e trocam de estratégia $P \rightarrow C$.

Entre $150 \lesssim t \lesssim 850$, a dinâmica do jogo passa a ser guiada pelo aumento da frequência de desertores (D), que performam melhor quando são relativamente mais raros no ambiente competitivo. É possível observar pela figura 37 a formação

de pequenas “ilhas” ou subpopulações compactas de desertores em ambas populações.

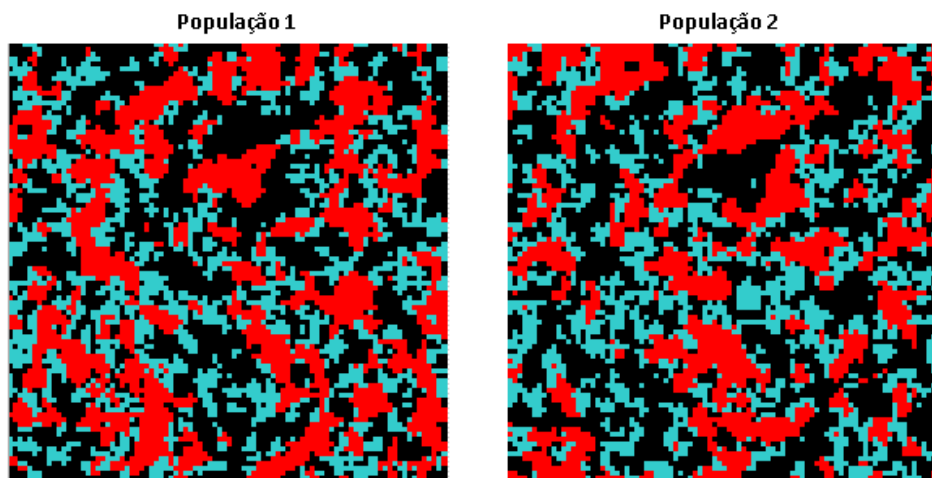


Figura 37 - Passo de Monte Carlo 850; $r = 0,02$.

Os desertores obtêm o maior *payoff* do jogo quando competem num ambiente com vizinhos programados para jogar C. Por volta do passo de Monte Carlo 850 a frequência de desertores se torna maior do que a frequência de punidores. Isso acontece nas duas populações e também pode ser visto através da figura 30 da página 56.

Entre $1000 \lesssim t \lesssim 2500$, a frequência de desertores continua a crescer às custas dos punidores, que observam seus vizinhos desertores obterem melhores *payoffs* num ambiente cuja a predominância é de cooperadores. Logo, os punidores imitam o comportamento dos desertores e ocorrem trocas do tipo $P \rightarrow D$.

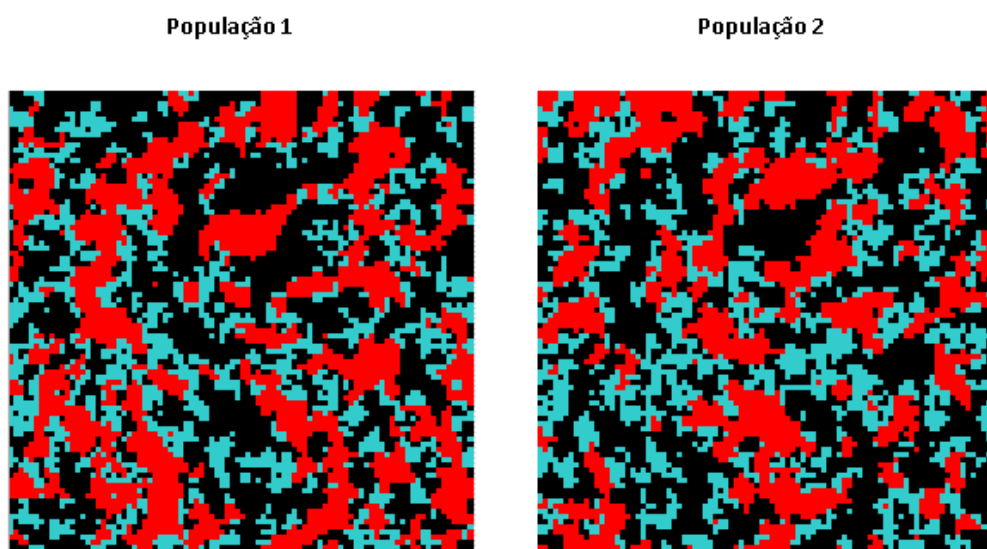


Figura 38 - Passo de Monte Carlo 1000; $r = 0,02$.

Entre $2500 \lesssim t \lesssim 6000$, a frequência de punidores se estabiliza nas duas populações com aproximadamente 20% dos indivíduos adotando a estratégia P. As interações então passam a ocorrer substancialmente entre cooperadores e desertores, o que abre espaço essencialmente para um aumento da frequência de desertores nesse instante do jogo. O efeito líquido observado é uma compactação ainda maior de ilhas de desertores e cooperadores.

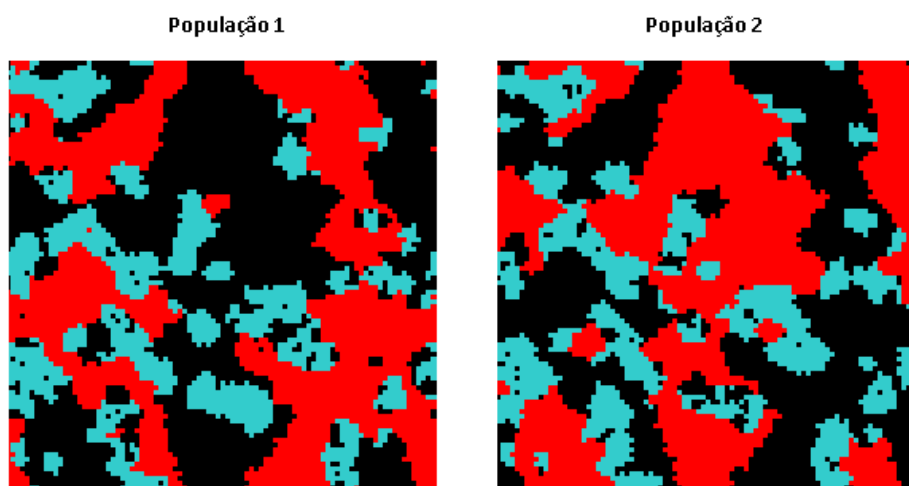


Figura 39 - Passo de Monte Carlo 6000; $r = 0,02$.

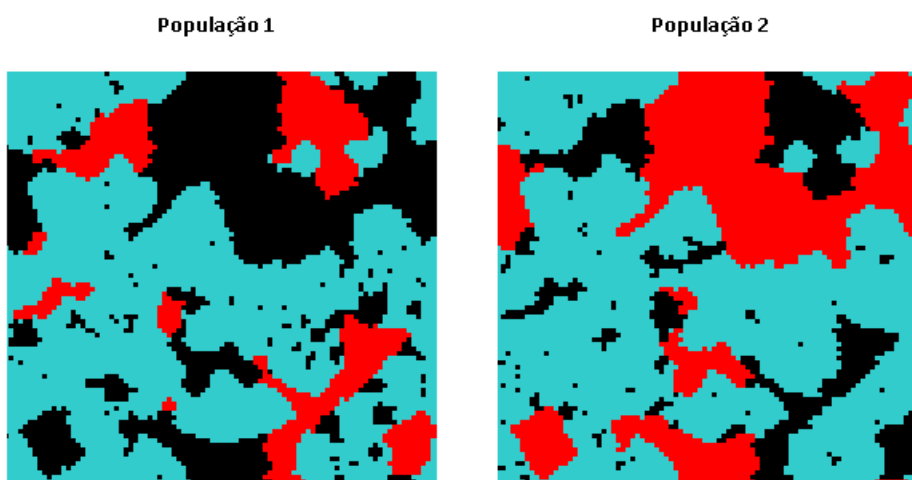


Figura 40 - Passo de Monte Carlo 20000; $r = 0,02$.

Por volta do Passo de Monte Carlo 20.000, as ilhas formadas por cooperadores e desertores vão sendo ocupadas por punidores. Nesse estágio do jogo, ocorrem trocas indiretas, isto é, $C \rightarrow D \rightarrow P$ e, na medida em que se avança no processo de simulação, observa-se a coexistência de punidores e cooperadores como equilíbrio assintótico em ambas as populações. As trocas indiretas se justificam, pois os indivíduos C observam seus vizinhos D performarem melhor

contra os adversários C, favorecendo trocas do tipo $C \rightarrow D$. Por outro lado, quando estes jogadores C que mudaram de estratégia para D competem num ambiente cuja predominância é de punidores, as trocas passam a ser do tipo $D \rightarrow P$, até que a frequência de desertores chegue a zero em ambas as populações. A extinção completa dos desertores pode ser explicada pela própria matriz de pagamentos do jogo. Recordemos que, neste caso em que $r = 0,02$, num ambiente em que há predominância de indivíduos desertores competindo com punidores, a estratégia que performa pior contra P é justamente D. Logo, o equilíbrio se dá com ambas populações compostas por uma maioria de punidores e alguns cooperadores espalhados pelos *clusters*.

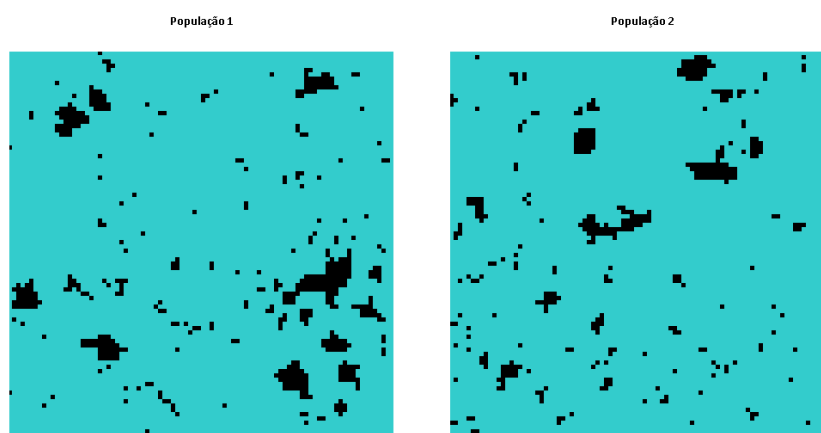


Figura 41- Passo de Monte Carlo 100.000; $r = 0,02$.

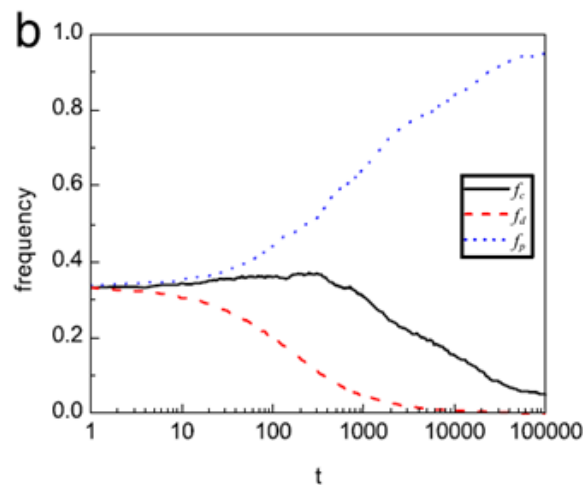
Alguns cooperadores sobrevivem, pois quando o último desertor troca de estratégia para P ou C, estes comparam seus *payoffs* com outros punidores e cooperadores, e não há incentivo para replicarem o comportamento deles, uma vez que recebem o mesmo pagamento caso cooperem ou caso punam quando competem com um jogador programado para jogar P ou C.

Caso II) $r = 0,20$

Tipicamente, o comportamento evolutivo verificado para o caso em que $r = 0,2$ é o demonstrado na figura 43. Com aproximadamente 20.000 passos de Monte Carlo há convergência de uma população para o estado *All C* e da população rival para o estado *All D*, não sendo possível verificar a existência de punidores em nenhum dos cenários observados. Em função da aleatoriedade com que se aloca inicialmente as populações, foi possível encontrar casos em que a população 1

convergiu para o estado *All C* e a população 2 para *All D*. O contrário também foi observado, isto é, a população 1 convergiu para *All D* e a população 2 para *All C*. Há, também, para este parâmetro, uma trajetória até o equilíbrio bastante lenta em que se verifica a convergência das estratégias para os estados *All C* e *All D* por volta dos 2.000.000 passos de Monte Carlo.

Quando se compara os resultados obtidos nesta dissertação com o modelo de Chan et al. (2013) para o caso em que $r = 0,2$, vê-se que o fato de uma população competir com outra população rival impacta o equilíbrio assintótico do jogo. Em um jogo espacial de uma população, no longo prazo, encontra-se uma população composta essencialmente por punidores, conforme figura abaixo.



**Figura 42 - Equilíbrio assintótico obtido por Chan et al. (2013) para $r = 0,2$.
Fonte: Chan et al. (2013).**

Apresenta-se inicialmente o caso II A, que corresponde ao padrão de convergência observado na maioria das realizações – com 20.000 passos de Monte Carlo.

Caso II A)

Este é o caso registrado em aproximadamente 75% das realizações. Nos instantes iniciais, isto é, $t \lesssim 10$, observa-se, que as frequências de cooperadores e punidores são praticamente idênticas. Neste estágio do jogo, as trocas ocorrem entre $D \rightarrow P$ e $D \rightarrow C$, principalmente nas regiões em que desertores interagem com outros desertores.

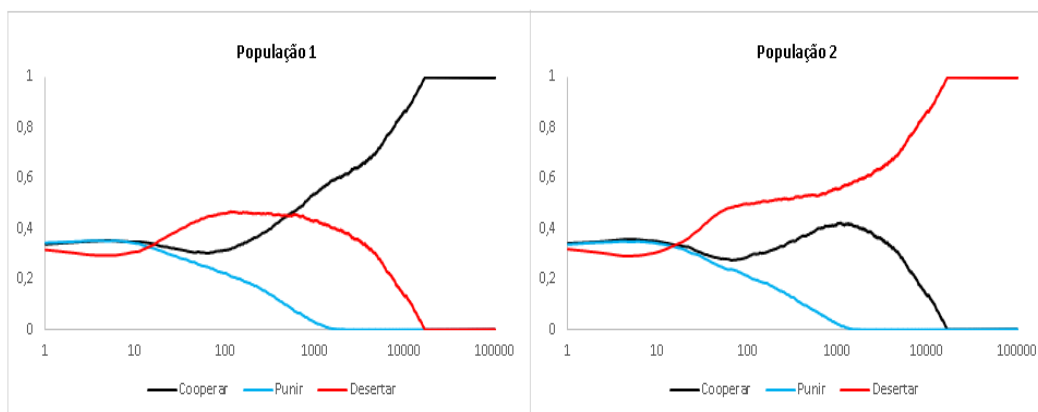


Figura 43 - Equilíbrio Assintótico com o comportamento típico para $r = 0,2$. População 1 All C e População 2 All D.

Com o auxílio da matriz de pagamento, note que jogar C ou P contra D performa melhor que D contra D, que resulta no pior *payoff* do jogo. Por outro lado, diferentemente do caso anterior, em que $r = 0,02$, a estratégia D, agora, performa melhor contra P do que as estratégias C e P contra P. Isso nos permite inferir melhor qualitativamente sobre o equilíbrio do jogo.

| | C | D | P |
|---|-----|------|------|
| C | 1 | 0,8 | 1 |
| D | 1,2 | 0 | 1,15 |
| P | 1 | 0,79 | 1 |

Figura 44 - Matriz de pagamento do jogo; $r=0,20$, $\alpha = 0,01$ e $\beta = 0,05$.

Entre os instantes $10 \lesssim t \lesssim 100$ é possível notar um aumento na frequência relativa de desertores em ambas as populações, incentivada por uma melhor performance desses jogadores programados para jogar D quando inseridos num ambiente competitivo composto basicamente por C e P. Haverá troca de $C \rightarrow D$ e troca de $P \rightarrow D$. A evolução espacial do jogo mostrado nas figuras 45 e 46 proporciona uma melhor percepção das trocas de estratégia até o 100º passo de Monte Carlo. Em vermelho representa-se os desertores, em preto os cooperadores e em azul estão representados os punidores.

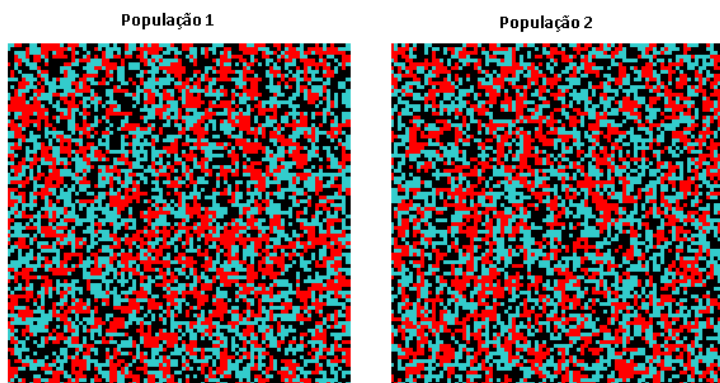


Figura 45 – Passo de Monte Carlo 10; $r = 0,2$.

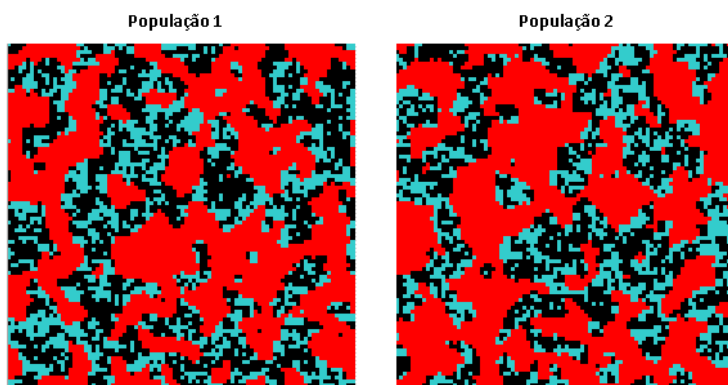


Figura 46 - Passo de Monte Carlo 100; $r = 0,2$.

Por volta do passo de Monte Carlo 100, em ambas as populações há, aproximadamente, 50% de desertores. A partir deste momento é que começa a se desenhar o equilíbrio assintótico do jogo. Entre $100 \lesssim t \lesssim 1.000$, na população 1 ocorrem muitas trocas de punidores para cooperadores, dado que numa competição contra desertores, a estratégia C performa melhor que P. Simultaneamente, na população 2 também ocorre troca de estratégias do tipo $P \rightarrow C$, o que resulta numa frequência de punidores próxima de 2% em ambas as populações por volta do passo de Monte Carlo 1000. Há, ainda, algumas trocas de desertores para cooperadores, principalmente nas regiões em que a competição se dá contra adversários desertores, pois C performa melhor contra D do que D contra D. O inverso também ocorre, isto é, num ambiente de competição formado basicamente por adversários cooperadores, haverá incentivo para troca de estratégia de $C \rightarrow D$, pois contra indivíduos programados para jogar C, os desertores possuem o maior *payoff* do jogo.

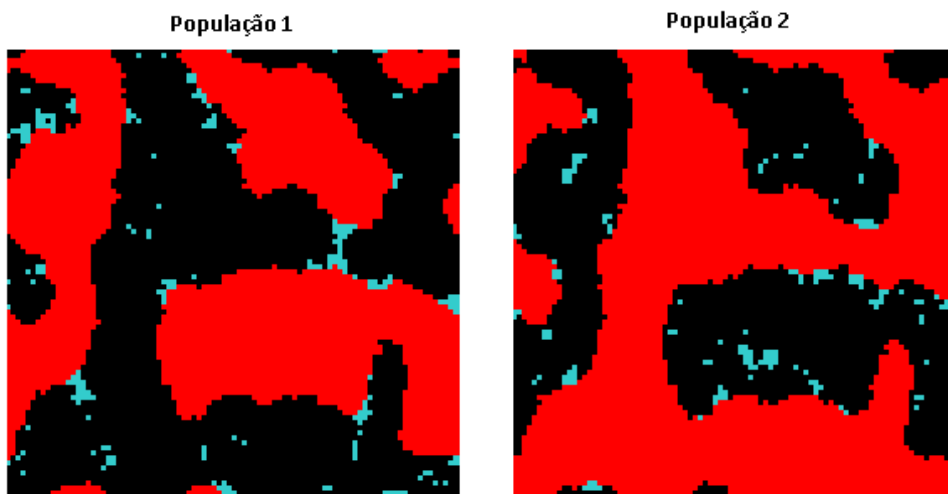


Figura 47 - Passo de Monte Carlo 1000; $r = 0,2$.

Este processo se intensifica na medida em que se avança no número de passos de Monte Carlo simulados. É possível ver a formação de “ilhas” de desertores cercadas por cooperadores e alguns punidores espalhadas na população 1 e, também, a formação de “ilhas” de cooperadores e punidores sendo envolvidas por desertores espalhadas na população 2. Entre $1.000 \lesssim t \lesssim 5.000$, os punidores são eliminados do jogo. As trocas de estratégia ocorrem, então, entre cooperadores e desertores que estão localizados nas fronteiras, exatamente onde agentes Focal e Referência possuem estratégias distintas e estão inseridos num ambiente competitivo bastante similar.

Esse tipo de geometria observada na figura 47 acelera o processo de troca de estratégias e sugere evidências do porquê a convergência se dá de forma mais rápida, uma vez que as ilhas mais dispersas espacialmente levam a maiores pontos de competição, o que contribui para uma troca de estratégia mais intensa.

A população 1, como pode ser visto no exemplo abaixo, terá maior incentivo para fazer trocas de estratégia do tipo $D \rightarrow C$. A ver:

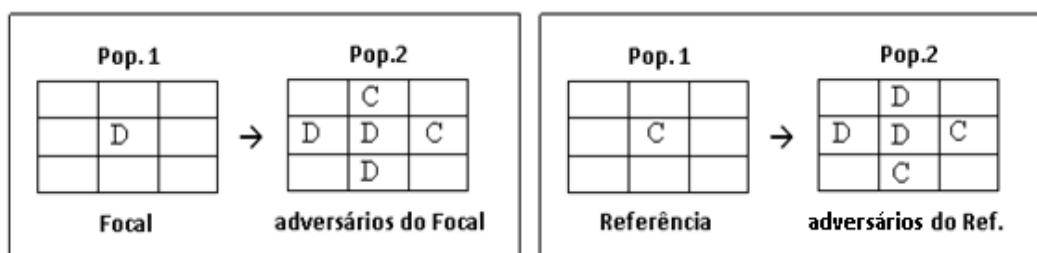


Figura 48 – Focal D e Referência C com 3 vizinhos desertores; $r = 0,2$.

$$\text{payoff do Focal (D)} = \frac{2(1,2)}{5} = 0,48$$

$$\text{payoff do Referência (C)} = \frac{3(0,8) + 2(1)}{5} = 0,88$$

$w \cong 0,33 \rightarrow$ Probabilidade de troca de estratégia de D \rightarrow C.

Nas fronteiras em que existe uma vizinhança formada por 4 desertores, a troca D \rightarrow C na população 1 ocorrerá com a seguinte probabilidade:

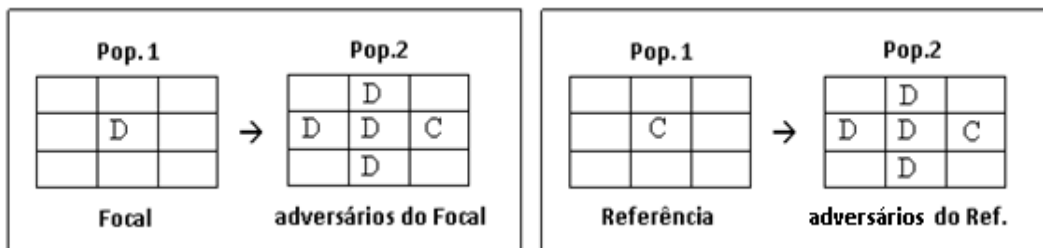


Figura 49 - Focal D e Referência C com 4 vizinhos desertores; $r=0,2$.

$$\text{payoff do focal (D)} = \frac{1,2}{5} = 0,24$$

$$\text{payoff do referência (C)} = \frac{4(0,8) + 1}{5} = 0,84$$

$w = 0,5 \rightarrow$ Probabilidade de troca de estratégia de D \rightarrow C.

Na população 2, as trocas ocorrerão entre C \rightarrow D. Pois, dado que na população 1 há regiões compactas compostas puramente por cooperadores, a estratégia D performa melhor contra C do que C contra C. Essa relação entre o ganho obtido por indivíduos desertores na população 2 é que leva ao equilíbrio assintótico em que a população 1 é *All C* e a população 2 é *All D*. As dinâmicas evolutivas dos Passos de Monte Carlo 5.000, 10.000, 15.000 e 20.000 são apresentadas a seguir. Nelas, é possível visualizarmos que as trocas de estratégias em ambas as populações se iniciam nas fronteiras, onde desertores e cooperadores são vizinhos e a probabilidade de troca de estratégias é maior.

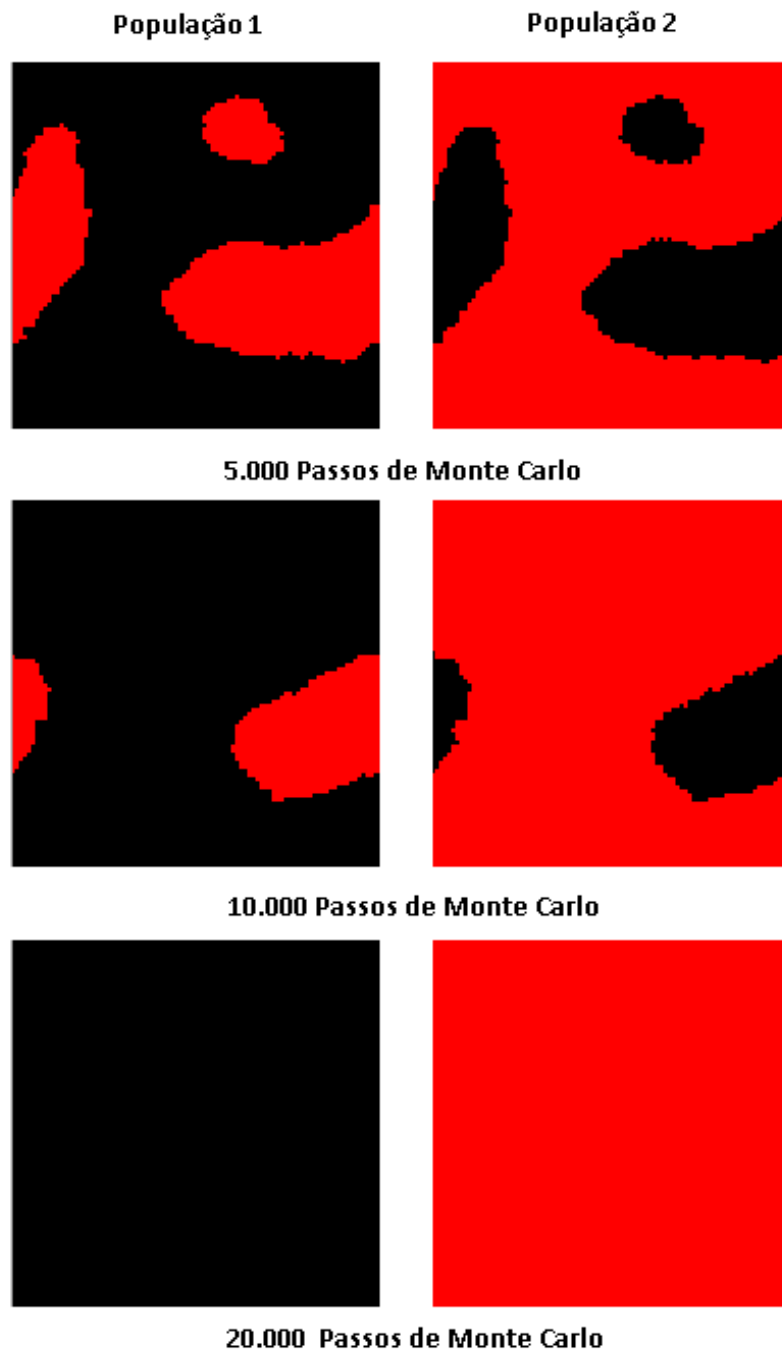


Figura 50 - Dinâmica dos passos de Monte Carlo 5.000, 10.000 e 20.000; $r = 0,2$.

Entre $15.000 \lesssim t \lesssim 20.000$, persiste a troca de $D \rightarrow C$ na população 1 e $C \rightarrow D$ na população 2. Isso ocorre até que a população 1 seja toda formada por cooperadores e a população 2 toda formada por desertores. No caso II B, o equilíbrio de longo prazo é o mesmo, porém a trajetória de convergência possui uma dinâmica diferente.

Caso II B)

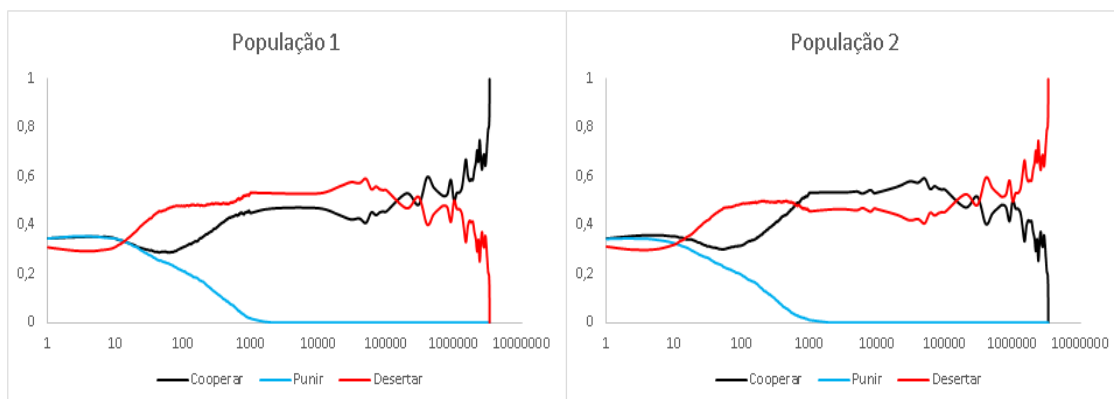


Figura 51 - Dinâmica evolutiva verificada no caso II B; $r = 0,2$.

A partir do 1000º passo de Monte Carlo é visível a diferença nas trajetórias que levam ao equilíbrio do jogo no caso II B, uma vez que a figura geométrica formada pela estrutura espacial muito se difere da mostrada no caso II A.

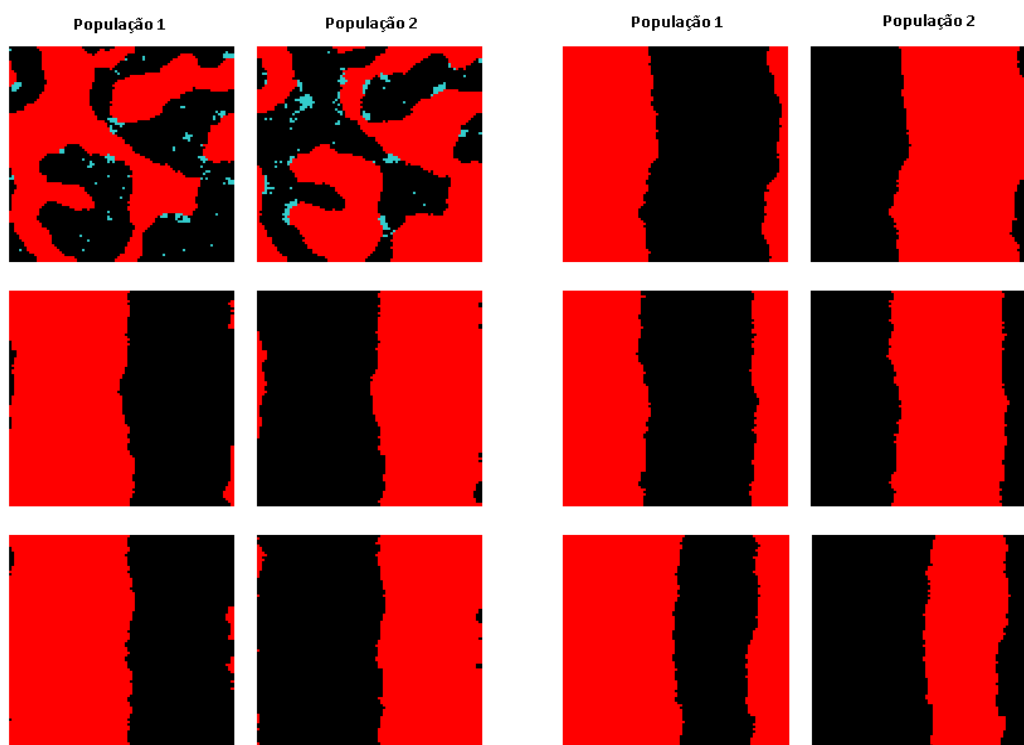


Figura 52 - Evolução espacial do jogo. À esquerda estão representados os passos de Monte Carlo 1.000, 30.000 e 60.000. À direita 1.000.000, 1.500.000 e 2.000.000.

A partir do passo de Monte Carlo 30.000, observa-se a formação de *clusters* mais concentrados espacialmente. Esse desenho espacial, faz com que as trocas $C \rightarrow D$ na população 1 e de $D \rightarrow C$ na população 2 ocorram apenas nas fronteiras. Em

outras palavras, os elementos no interior dos *clusters* ficam protegidos por uma malha, o que torna a convergência mais lenta.

No passo de Monte Carlo 60.000, a frequência de desertores na população 1 é de aproximadamente 55%. Assim, temos que, para a população 2, sendo o jogador Referência um cooperador e o Focal um desertor, na medida em que estes estão inseridos num ambiente competitivo com mais desertores, maior é a probabilidade de troca $D \rightarrow C$ na população 2. Por outro lado, como há um número maior de indivíduos cooperadores na população 2, os jogadores da população 1 adotam a estratégia D, dado que D contra C representa o maior pagamento possível do jogo. Essa dinâmica persiste ao longo do tempo e, lentamente, vai se observando a formação dos estágios *All D* e *All C* nas populações 1 e 2, respectivamente.

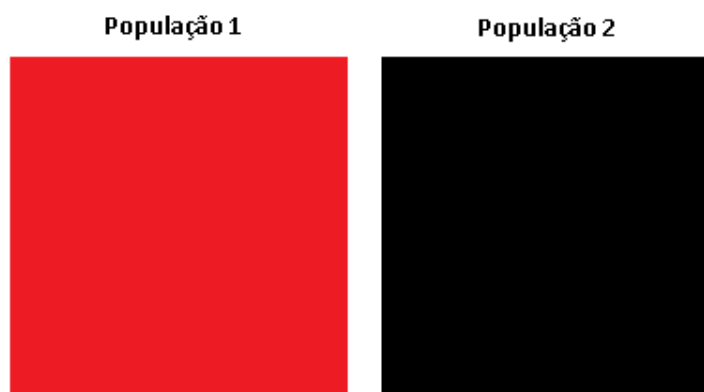
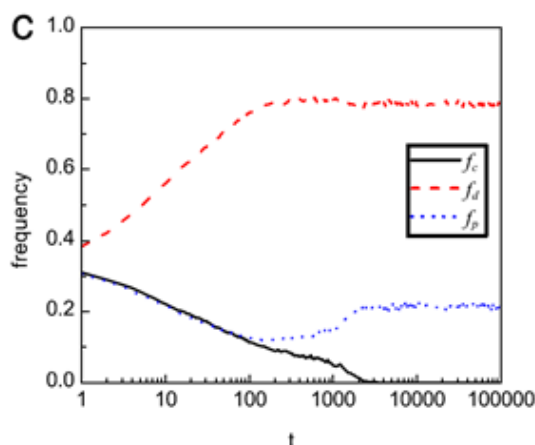


Figura 53 - Estado *All C* e *All D* com 2.400.000 Passos de Monte Carlo; $r = 0,2$.

Caso III) $r = 0,6$

Tipicamente, o comportamento evolutivo verificado para o caso em que $r = 0,6$ é o demonstrado na figura 55. Com aproximadamente 15.000 passos de Monte Carlo, há convergência de uma população para o estado *All C* e da população rival para o estado *All D*, não sendo possível verificar a existência de punidores em nenhum dos cenários observados. Em função da aleatoriedade com que se aloca inicialmente as populações, foi possível encontrar casos em que a população 1 foi composta puramente por desertores e a população 2 totalmente por cooperadores. Há, também, para este parâmetro, uma trajetória bastante lenta em que se verifica a convergência das estratégias para os estados *All C* e *All D* após os 2.000.000 passos de Monte Carlo.

Quando se compara os resultados obtidos nesta dissertação com o modelo de Chan et al (2013) para o caso em que $r = 0,6$, vê-se que o fato de uma população competir com outra rival também impacta o equilíbrio assintótico do jogo. Em um jogo espacial de uma população, no longo prazo, encontra-se uma população composta essencialmente por desertores e por uma frequência baixa de punidores, conforme figura abaixo.



**Figura 54 - Equilíbrio assintótico obtido por Chan et al. (2013) para $r = 0,6$.
Fonte: Chan et al. (2013).**

Apresenta-se inicialmente o caso III A, que corresponde ao padrão de convergência observado na maioria das realizações. Note que, ao contrário do que ocorre no jogo de uma população espacial proposto por Chan et al (2013), os punidores são eliminados no jogo espacial de duas populações, e a frequência relativa de desertores numa das populações converge para 1 e para 0 na população adversária.

Caso III A)

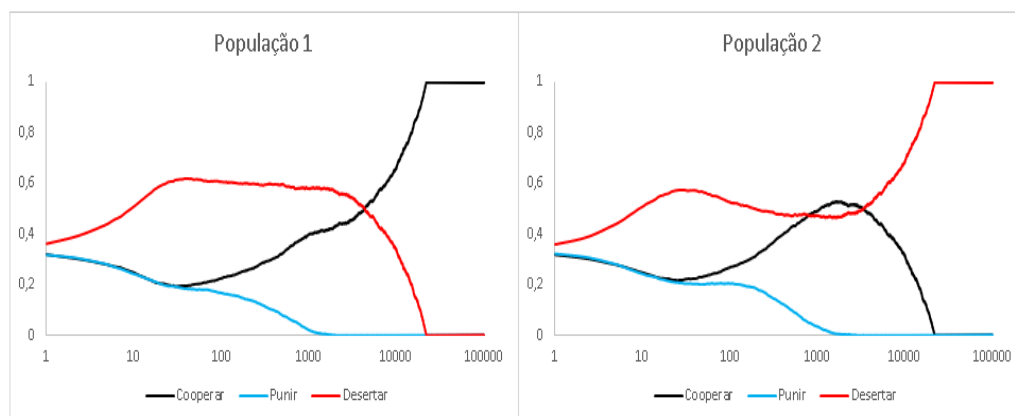


Figura 55 - Comportamento típico da dinâmica evolutiva do jogo; $r=0,6$.

Para o caso em que $r = 0,6$, inicialmente há um favorecimento aos desertores em ambas as populações e a diminuição da importância do papel do custo de punição β aplicado pelos punidores (P) – que são banidos permanentemente do jogo – aos indivíduos programados para jogar D, de modo que o equilíbrio se dará com uma população formada por indivíduos que desertam, jogando D (*All D*) e pela população concorrente adotando a estratégia C (*All C*). Pela relação entre os pagamentos apresentada abaixo podemos ter uma intuição melhor da dinâmica do jogo ao longo do processo simulado. A ver:

$$\begin{array}{c} C \quad D \quad P \\ \begin{array}{c} C \\ D \\ P \end{array} \begin{pmatrix} 1 & 0,4 & 1 \\ 1,6 & 0 & 1,55 \\ 1 & 0,39 & 1 \end{pmatrix} \end{array}$$

Figura 56 - Matriz de pagamentos para o caso em que $r = 0,6$, $\alpha = 0,01$ e $\beta = 0,05$.

Entre $0 \lesssim t \lesssim 50$, é possível visualizarmos nas duas populações a troca maciça de estratégias C e P para D, de modo que a frequência de desertores é de aproximadamente 56% na população 1 e 62% na população 2.

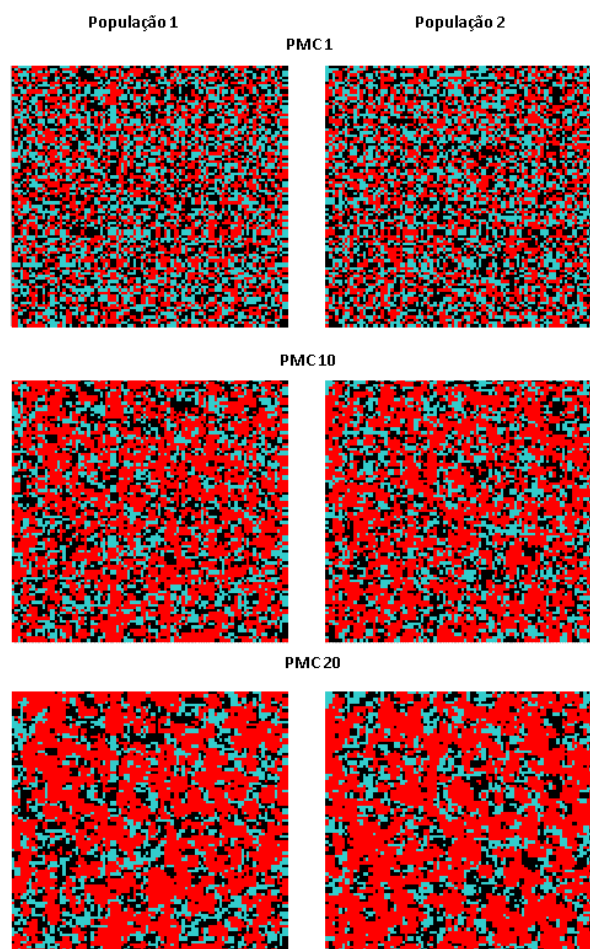


Figura 57 - Evolução das populações – passos Monte Carlo 1, 10 e 20; $r = 0,6$.

Estas trocas ocorrem na medida em que indivíduos programados para jogar D estão num ambiente competitivo formado por cooperadores e punidores. Logo, ao obterem melhores *payoffs*, pela lógica implementada, terão seu comportamento replicado por aqueles jogadores programados para jogar C e P. Porém, na medida em que cresce o número de desertores em ambas as populações, e levando em consideração que D contra D corresponde ao menor pagamento do jogo, é possível inferir que nos passos seguintes, há uma temporária estabilização da frequência de desertores nas duas populações.

Entre $50 \lesssim t \lesssim 100$, é possível observar a estabilização da frequência de desertores. A partir deste intervalo de tempo, ocorrem trocas de estratégias de punidores para cooperadores, uma vez que C performa melhor contra D do que P contra D e, também, ocorrem algumas trocas de D para C, pois nesse estágio do jogo, para uma vizinhança com maior número de desertores, C performa melhor contra D do que D contra D. A dinâmica é bastante parecida nas duas populações.

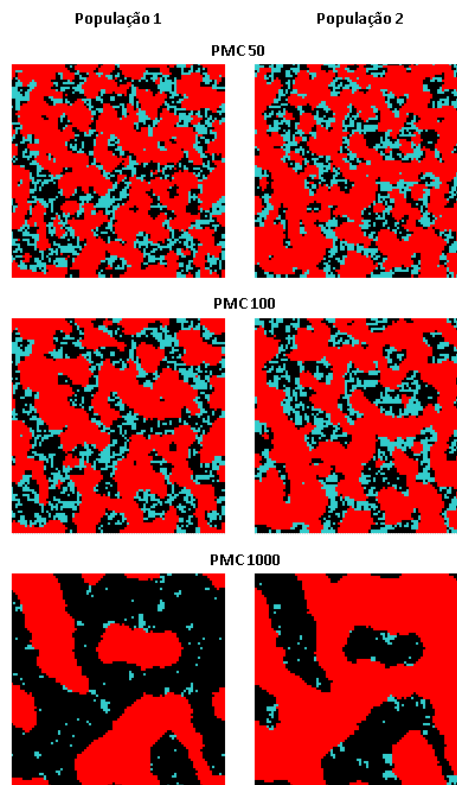


Figura 58 - Evolução das interações estratégicas - passos 50, 100 e 1.000; $r=0,6$.

Entre $100 \lesssim t \lesssim 1.000$, é possível verificar a formação de *clusters* (ou subpopulações) de cooperadores e desertores distribuídos em ambas as populações. Estes *clusters* mais dispersos espacialmente, observados na figura 58, levam a

maiores pontos de competição, contribuindo para uma troca de estratégia mais intensa e, conseqüentemente, a uma convergência mais rápida para os estados *All C* e *All D*.

Na população 1, têm-se maior frequência de cooperadores e na população 2, observa-se um maior número de desertores. Já o número de punidores em ambas as populações é próximo de 2%. A dominância de cooperadores na população 1 pode ser explicado pelo melhor pagamento que estes obtêm em relação aos punidores e desertores quando competem com indivíduos desertores da população rival. Por outro lado, a formação de desertores na população 2 advém da existência de indivíduos cooperadores na população 1, ou seja, o indivíduo da população 2 maximiza seu *payoff* adotando a estratégia D, contra um indivíduo que joga C.

Entre $2.000 \lesssim t \lesssim 5.000$, intensificam-se as trocas de $D \rightarrow C$ na população 1 e de $C \rightarrow D$ na população 2. Em ambas populações a figura do punidor (P) é eliminada dentro deste intervalo de tempo. As subpopulações de desertores na população 1 e de cooperadores na população 2 vão perdendo consistência ao passo que as interações e comparações entre *payoffs* são demonstradas a seguir. Na População 1, existem cooperadores e desertores competindo numa vizinhança formada por uma maioria de desertores, de modo que:

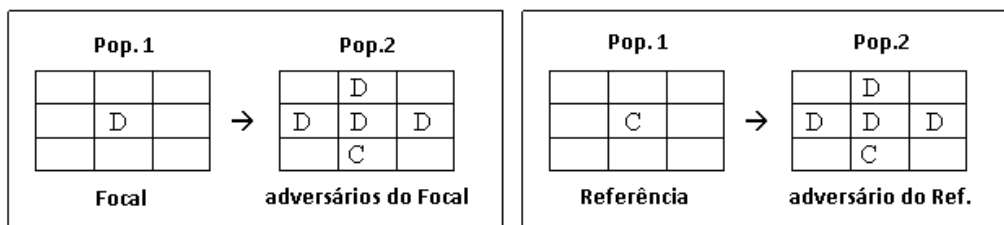


Figura 59 - Focal e Referência competindo com 4 desertores; $r = 0,6$.

$$\text{payoff do Focal (D)} = \frac{1,6}{5} = 0,32$$

$$\text{payoff do Referência(C)} = \frac{4(0,4) + 1}{5} = 0,52$$

$$w = 0,125 \quad \rightarrow \quad \text{Probabilidade de troca de estratégia de } D \rightarrow C.$$

Para uma vizinhança composta apenas por desertores, a probabilidade de troca é ainda maior, chegando a 0,25. Já para a População 2, temos a seguinte situação:

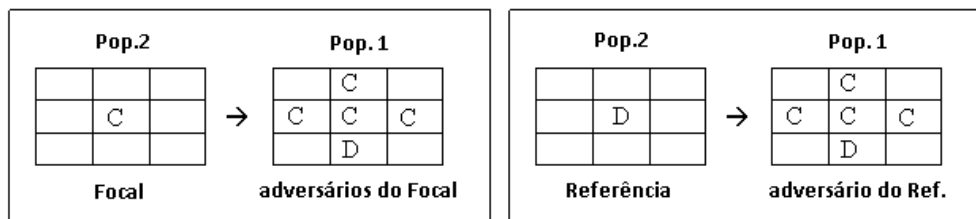


Figura 60 - Focal e Referência competindo com 4 cooperadores; $r = 0,6$.

$$\text{payoff do Referência (D)} = \frac{4(1,6)}{5} = 1,28$$

$$\text{payoff do Focal (C)} = \frac{0,4 + 4(1)}{5} = 0,88$$

$w = 0,25 \rightarrow$ Probabilidade de troca de estratégia de $C \rightarrow D$.

Para uma vizinhança composta apenas por cooperadores, a probabilidade de troca de $C \rightarrow D$ é ainda maior, chegando a 0,75.

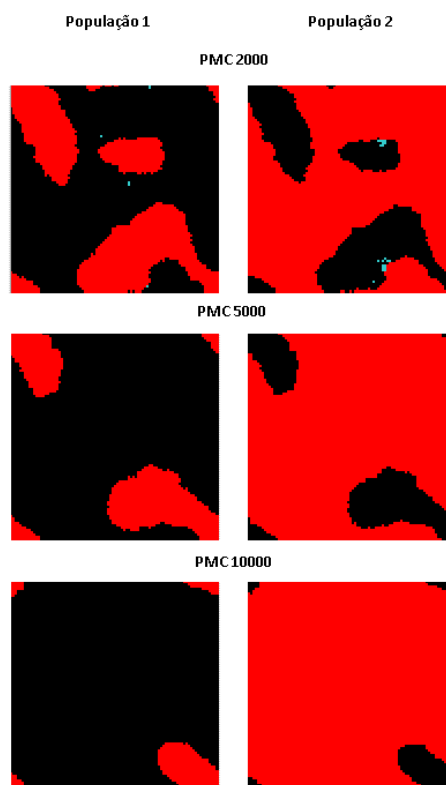


Figura 61 - Distribuição espacial dos passos 2.000, 5.000 e 10.000; $r = 0,6$.

Por volta do passo 15.000, o que pode ser observado é que, de fato, há formação de uma população cujo estado é *All C* e de outra população cujo estado é *All D*. A dinâmica da evolução até se atingir o estado *All C* na população 1 se dá a partir da troca de estratégia de $D \rightarrow C$ dos indivíduos situados nas fronteiras dos

clusters (ou ilhas) de desertores, uma vez que após o milésimo passo é cada vez mais frequente a competição destes com os desertores da população 2. Ao comparar o seu pagamento com um agente Referência programado para cooperar, existe a possibilidade de troca $D \rightarrow C$, conforme mostrado nos exemplos acima.

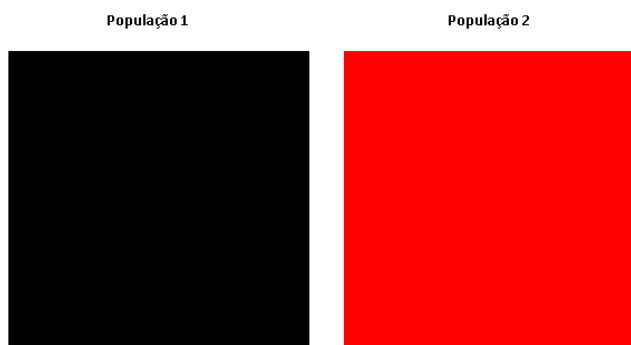


Figura 62 - Equilíbrio assintótico do jogo no PMC 15.000; $r = 0,6$.

Para se atingir o estado *All D* na população 2 as troca de estratégia são de $C \rightarrow D$, também ocorrendo substancialmente nas fronteiras dos *clusters* de cooperadores. Ao comparar o seu pagamento com um vizinho selecionado para ser o agente Referência - programado para desertar - existe uma probabilidade de troca $C \rightarrow D$, conforme mostrado nos exemplos anteriores. Isso faz com que aumente a frequência de desertores na população 2 até que apenas estes indivíduos programados para jogar D sobrevivam no longo prazo.

Caso III B)

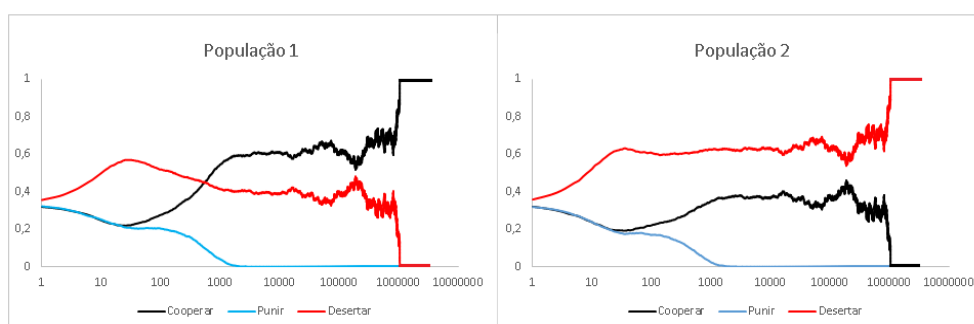


Figura 63 - Dinâmica evolutiva verificada no caso III B; $r=0,6$.

Até por volta do 50º passo de Monte Carlo, a frequência de desertores cresce às custas dos cooperadores e punidores. A partir de então, como a frequência de desertores em ambas as populações se aproxima de 60% é comum que exista num

mesmo ambiente competitivo outros desertores, o que faz o pagamento esperado de desertor ser menor, dado que desertor contra desertor gera um *payoff* igual a zero, possibilitando algumas trocas do tipo $D \rightarrow C$. Nesse estágio do jogo, ocorrem ainda trocas de punidores para cooperadores, dado que os cooperadores performam melhor que os punidores quando interagem com desertores. Essa dinâmica de troca de $P \rightarrow C$ ocorre até por volta do passo de Monte Carlo 2.000, onde já é possível visualizar um *cluster* de cooperadores e de desertores em ambas as populações.

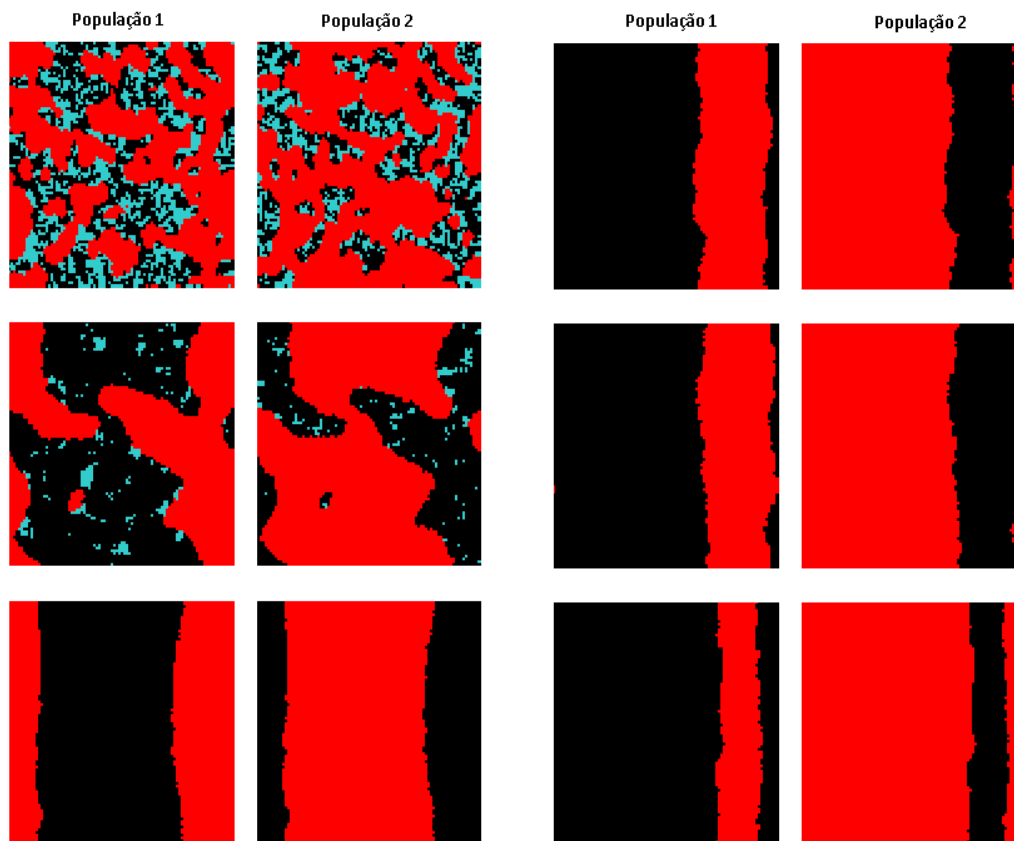


Figura 64 - Evolução espacial do jogo. À esquerda estão representados os passos de Monte Carlo 100, 1.000 e 30.000. À direita estão os passos 500.000, 750.000 e 1.000.000.

Na figura 64, note que a geometria formada pelas duas populações a partir do passo de Monte Carlo 1.000 se distingue daquela apresenta no caso III A. Esse tipo de distribuição espacial impacta a velocidade de convergência (e também a trajetória) para os estados *All C* e *All D*. As trocas de estratégia ficam, portanto, restritas às regiões de fronteira entre cooperadores e desertores.

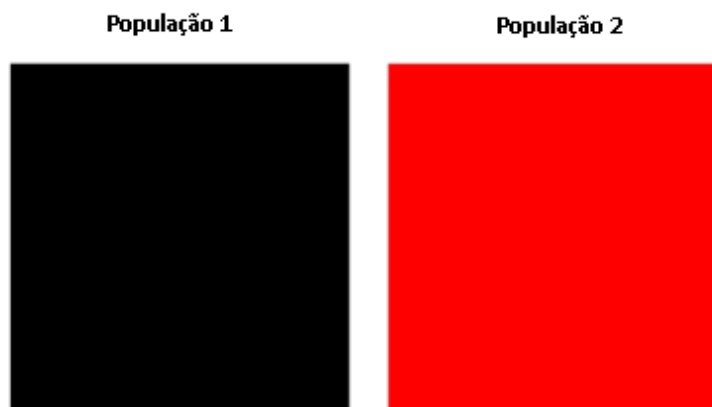


Figura 65 - Estados *All C* e *All D* no PMC 2.000.000; $r = 0,6$.

Por volta do passo de Monte Carlo 30.000, os *clusters* (ou ilhas) se forma de modo bastante compacto, compondo uma malha, o que faz com que os indivíduos cooperadores na população 1 e os desertores na população 2 fiquem protegidos no interior dessas malhas. O processo evolutivo se dá de forma bastante lenta nessas circunstâncias, e após 2.000.000 de passos de Monte Carlo, podemos observar a composição dos estados *All C* e *All D*, conforme mostrado na figura 65.

7 Conclusão

Nesta dissertação foi estudado, através de um modelo baseado em teoria dos jogos evolucionários, a incorporação de um agente punidor no jogo *Snowdrift* em duas populações com distribuição espacial, que impõe restrições geométricas na forma como um jogador pode afetar o seu oponente.

Com a finalidade de analisar o problema proposto, foram utilizados métodos de simulação baseados nos trabalhos de Hauert e Doebeli (2004), Xu et al. (2011) e Chan et al. (2013) para complementar a solução analítica encontrada pelo replicador dinâmico, além de implementar uma componente estocástica na análise.

A partir das simulações numéricas, os resultados encontrados neste trabalho apresentam-se diferentes daqueles alcançados por Chan et al. (2013). Foi possível concluir que numa competição com arranjo espacial e duas populações, a presença de punidores só ocorre quando a relação entre o custo pelo benefício da cooperação é bastante baixo, isto é, $r = 0,02$. Neste caso, em ambas as populações, os desertores foram eliminados, restando apenas cooperadores e uma expressiva maioria de punidores. A extinção completa dos desertores pode ser explicada pela própria matriz de pagamentos do jogo. Recordemos que para uma razão de custo benefício $r = 0,02$, a estratégia D performa pior contra P do que C e P contra P.

Nos demais casos estudados, o punidor não sobrevive na competição, sendo o equilíbrio formado por uma população inteira de cooperadores e a outra totalmente composta por desertores. Os equilíbrios assintóticos do tipo *All C* e *All D* podem ser novamente explicados com o auxílio da matriz de pagamento do jogo associada a cada um dos parâmetros de custo pelo benefício líquido da cooperação (r) analisados.

É possível inferirmos que, diante de uma população totalmente formada por indivíduos desertores, a estratégia que representa maior benefício em termos de pagamento é cooperar. Por outro lado, diante de uma população toda composta por cooperadores, a melhor estratégia a se adotar, neste caso, é desertar. As simulações

elaboradas captaram justamente esta dominância entre as estratégias disponíveis no jogo.

Em função da aleatoriedade, foi possível verificar diferentes trajetórias que levaram aos estágios *All C* e *All D* nos casos II e III. O espalhamento dos *clusters* (ou ilhas), quando mais concentrados espacialmente, faz com que a competição e as trocas de estratégia ocorram somente nas fronteiras, deixando os elementos do interior destes *clusters* protegidos por uma malha – o que leva a uma convergência mais lenta. Já no caso em que as ilhas estão mais dispersas, existem mais pontos de competição, contribuindo para uma troca de estratégias mais intensa – o que leva a uma convergência mais rápida.

Por fim, existem algumas direções para as quais esta pesquisa pode ser estendida. Primeiro, uma aplicação num contexto de Organização Industrial e mercados em competição, inspirando-se nos trabalhos de Young (1993) e Xiao e Yu (2006). Posteriormente, a proposição de um jogo em que haja competição envolvendo populações heterogêneas (Rocha et al, 2015). Outro caminho, no que diz respeito aos métodos de simulação, poderia ser seguido com a implementação de Grafos como alternativa de estabelecer critérios de vizinhança espacial de cada agente, e analisar sua contribuição para a evolução do comportamento cooperativo.

ANASTASOPOULOS, N.P., ANASTASOPOULOS, M.P. **The evolutionary dynamics of audit. European Journal of Operational Research.** Volume 216, pages 469-476, 2012.

ARRUDA, R.C. **Ensaio em jogos evolucionários com aplicação ao Estudo da Interação entre agências de *rating* e investidores.** 2015. 95 f. Dissertação – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2015.

BINMORE, K.G. **Fun and Games: A Text on Game Theory.** D.C. Heath, 1992.

BOLTON, P.; FREIXAS, X.; SHAPIRO, J. **The Credit Ratings Game.** The Journal of Finance. Volume 67, Issue 1, pages 85–111, 2012.

CHAN N. W. H. et al. **Evolutionary snowdrift game incorporating costly punishment in structured populations.** Department of Physics, The Chinese University of Hong Kong, Shatin, Hong Kong. Physica A 392, pages 168 – 176, 2013.

DOEBELI, M.; HAUERT, C. **Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game.** Ecology Letters. Pages 748-766, 2005.

FRIEDMAN, D. **Evolutionary Games in economics.** Econometrica. Volume 59, pages 637-666, 1991.

FRIEDMAN, D. **On economic applications of evolutionary game theory.** Department of Economics, University of California at Santa Cruz, 212 Social Sciences I, Santa Cruz, CA 95064, USA, 1998.

HAUERT, C; DOEBELI, M. **Spatial structure often inhibits the evolution of cooperation in the snowdrift game.** Departments of Zoology and Mathematics. University of British Columbia, Vancouver, British Columbia. Nature 428, 643- 646, 2004.

HAUERT, C. **Effects of Space in 2 x 2 Games.** Institut für Mathematik, Universität Wien, Strudlhofgasse 4, A-1090 Vienna, Austria, 2001.

HIRTH, S. Credit Rating Dynamics and Competition. Aarhus University, School of Business and Social Sciences, Department of Economics and Business,

JEHLE & RENY: Advanced Microeconomic Theory, 3rd edition. Prentice Hall, 2011.

MAS-COLELL, A.; WHINSTON, M.D.; GREEN, J.R. **Microeconomic Theory.** Oxford University Press, 1995.

MAYNARD-SMITH, J., e Price, G. R. **The logic of animal conflict.** School of Biological Sciences. University of Sussex. Nature 246, 15 – 18, 1973.

MENASCHÉ, D. S. **Controle de congestionamento: Uma abordagem dinâmico-populacional via teoria dos jogos evolucionários.** Dissertação (Mestrado em Engenharia de Sistema de Computação - COPPE/UFRJ, 2005.

NASH, J. F. **Equilibrium Points in N-person Games.** Proceedings of the National Academy of Sciences, 1950.

NASH, J.F. **The bargaining problem.** Econometrica 18 155-162, 1950.

NOWAK, M.A.; CORINA E. T.; ANTAL, T. **Evolutionary dynamics in structure populations. Program for Evolutionary Dynamics.** Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, USA, 2010.

NOWAK, M.A.; HIGHFIELD, R. **Super Cooperators: altruism, evolution and why we need each other to succeed**. Free Press, New York, 2011.

ROCHA, A.B.S.;LARUELLE, A.; ZUAZO, P. **Replicator Dynamics and Evolutionary Stable Strategies in Heterogeneous Games**. IKERLANAK 2011-57, Universidad del País Vasco – Departamento de Fundamentos del Análisis Económico I, 2011.

ROCHA, A. B. S. **Evolutionary dynamics of nationalism and migration**. Physica A. Volume 392, pages 3183-3197, 2013.

ROCHA, A. B. S., ESCOBEDO, R., LARUELLE, A. **Emergence of cooperation in phenotypically heterogeneous populations: a replicator dynamic analysis**. Journal of Statistical Mechanics, p.06003, 2015.

TAYLOR, P. D.; JONKER L. **Evolutionarily Stable Strategies and Game Dynamics**. Mathematical Biosciences 40, 145-156, 1978.

VON-NEUMANN, J.; MORGENSTERN, O. **Theory of Games and Economic Behavior**. Princeton University Press, 1944.

XIAO, T., YU, G. **Supply chain disruption management and evolutionary stable strategies of retailers in the quantity setting duopoly situation with homogeneous goods**. European Journal of Operational Research. Volume 173, Pages 648-668, 2006.

XU, C. et al. **Costly punishment and cooperation in the evolutionary snowdrift game**. Physica A: Statistical Mechanics and its Applications. Volume 390, Pages 1607–1614, 2011.

YOUNG, H.P. **The evolution of conventions**. Econometrica 61, 57-84, 1993.