

6

OS JOGOS E O COMPORTAMENTO

Ao sugerir um jogo baseado em dinâmicas, alguém poderia argumentar que tais processos não representam jogos do mundo real, onde os agentes não agem exclusivamente de modo racional selecionando as estratégias de melhores réplicas tentando atingir um ponto de equilíbrio. Num mundo real nem sempre existirá um equilíbrio e contempla uma mescla de situações e jogadores que produzem uma diversidade impossível de ser modelada com dinâmicas.

Tais questões são discutidas mais profundamente neste capítulo confrontando toda uma gama de idéias onde a importância do comportamento e do aprendizado será uma conclusão inevitável.

6.1 Aprendizado e Imitação como Dinâmica do Replicador

Até mesmo se nos adaptarmos à idéia básica da dinâmica do replicador e a considerarmos como uma ferramenta útil para selecionar alguém no equilíbrio de Nash, a aplicação para a economia deste conceito biológico de aptidão guiando a dinâmica não é óbvia. Nesta linha de pesquisa [Ref: 3] poderia-se sugerir que o mecanismo genético de seleção natural fosse recolocado por um mecanismo social de aprendizado e imitação. Porém com isto encontramos vários problemas.

Primeiramente, seleção natural por aptidão resulta num processo dinâmico na qual a frequência das estratégias jogadas pelos membros é ajustada através das diferenças nos números de descendentes. Membros com estratégias de maior sucesso gerarão mais filhos. Assim, o ajuste das frequências é um processo autônomo. Os membros da população não fazem escolhas racionais, mas o ajuste do comportamento a nível individual e conseqüentemente o comportamento racional são uma característica essencial para o aprendizado.

Em segundo lugar, na dinâmica do replicador as frequências das estratégias na população são ajustadas de acordo com sua aptidão.

Como consequência, as frequências de todas as estratégias com uma aptidão acima da média crescem

mesmo quando a estratégia não é uma melhor réplica. Dentro de uma estrutura de aprendizado e imitação isto não é muito realístico. Podemos esperar que jogadores racionais, num processo de aprendizado ou imitação, recolocarão suas estratégias pelas melhores réplicas. Deste modo, espera-se que em tal processo somente as frequências das estratégias de melhor réplica cresçam. Mas aí aparece a questão de se o processo de aprendizado ou imitação pode justificar a dinâmica do replicador como definida pela Equação 3-7. As pesquisas neste assunto têm seguido em várias direções. Alguns pesquisadores têm discutido a questão de se é possível formular o processo de aprendizado ou de imitação resultando na dinâmica do replicador como determinado pela Equação 3-7. Os indivíduos poderiam seguir um comportamento imitador e nunca imitariam um indivíduo que possui uma performance pior do que a sua, mas imitariam indivíduos melhores com uma probabilidade proporcional ao número das melhores realizações. Esta regra de comportamento resulta num processo de ajuste que pode ser aproximado pela dinâmica do replicador.

Dinâmicas como a de aprendizado e imitação foram apresentada por Weibull [Ref: 1], porém não adentraremos na análise de tais dinâmicas desde que sua aplicação parece restrita a jogos em que todos os jogadores possuem as mesmas características. Uma estratégia pode ser muito boa para um concorrente com determinada característica, mas pode não ser para outros com características completamente diferentes das do concorrente. Aí aparece uma outra questão que seria a quem imitar. Além do mais, mesmo que tentemos imitar alguém com alta performance e com características iguais às nossas, a estratégia admirada usada por nosso concorrente pode cair no caso de só ser eficiente se for a única no jogo. Caso tentemos utilizar a mesma estratégia, o resultado dessa estratégia contra ela mesma pode não corresponder às nossas expectativas.

Finalmente, é interessante mencionar um terceiro problema que ocorre quando se substitui um mecanismo genético por uma história de imitação e aprendizagem. Este problema pode ser formulado quando se discute o problema da racionalidade limitada dos jogadores.

Jogadores nestes modelos são freqüentemente estúpidos. Por exemplo, nos modelos evolutivos, os jogadores não são capazes de detectar ciclos que podem

ser gerados pela dinâmica. A crítica é que os modelos parecem contar com que os jogadores sejam implausivelmente estúpidos. Porque os jogadores não são capazes de compreender o que o modelador pode fazer? Relembre que na teoria tradicional, presume-se freqüentemente que os jogadores são computacionalmente superiores ao modelador.

Pode parecer não muito realístico adaptar as estratégias conforme um processo não tão sofisticado como dinâmica do replicador se as freqüências são ajustadas na prática, por um processo de aprendizagem.

Por outro lado, a dinâmica do replicador é um mecanismo que, além de encontrar equilíbrios, traça todo um panorama do espaço de estados que deve ser explorado, de modo inteligente pelo jogador e mesmo que estratégias que não sejam uma melhor réplica possam proliferar na população caso possuam um *payoff* acima da média, a convergência da dinâmica atingirá um equilíbrio de estratégias puras e portanto só uma estratégia sobrar na população..

Neste trabalho a dinâmica do replicador, assim como os equilíbrios, serão tratados como entidades matemáticas que nos ajudarão a pincelar a face do jogo.

Questões como racionalidade ou racionalidade limitada, serão tratadas na formação das estratégias. Conseqüentemente cada jogador terá o impacto de suas estratégias medido pela matriz de *payoff*.

Para explicitar a importância do comportamento associado às dinâmicas daremos um exemplo simples a seguir que consiste na livre concorrência de geradores num mercado hipotético pela venda de energia. Digamos que existam dois agentes hidrelétricos concorrendo pela venda de energia e que os seus *payoffs* dependem das energias armazenadas em suas usinas hidrelétricas. Eles competem em nível de igualdade, pois possuem reservatórios capazes de armazenar energia suficiente para atender toda a demanda e assim suas estratégias de venda nesta situação são preços reduzidos. Neste caso, a matriz de *payoff* do agente um, M_1 , possui a característica de $A = \mu(x,x) - \mu(y,x) > 0$, ou seja, que a sua estratégia residente de preço reduzido forme um equilíbrio de Nash com ela própria.

Depois de algum tempo de afluências bem abaixo da MLT (Média de Longo Termo), seus reservatórios começam a atingir níveis críticos de modo que só os dois agentes em conjunto consigam atender a demanda sem ocorrência de déficit. Assim, a nova estratégia para ambos, será a de elevar os preços de modo a desviar a geração de energia para outra fonte mais barata e, portanto, a nova matriz de

payoff não mais conservará a característica de $A > 0$, desde que as estratégias de preços reduzidos não são mais interessantes. Então a condição para que o agente tenha uma boa estratégia ao longo do tempo, seria evitar que seu volume caísse a um nível crítico, ou seja, a melhor forma de lidar com o futuro seria reduzir as incertezas, que no caso de tais agentes seria o aumento de seus reservatórios de modo que, mesmo num período seco os seus volumes fossem suficientes para atender toda a demanda, num determinado período e, portanto, as estratégias de preços reduzidos continuassem sendo a melhor estratégia.

Pode parecer uma conclusão óbvia e cara a de investir em garantias de lastro, mas as técnicas mais aceitas na atualidade apontam para esta solução pela relação entre o custo marginal de operação e o custo marginal de expansão. Por exemplo, os modelos usados para o planejamento energético apontam para a necessidade de investimento em novas usinas termelétricas ou hidrelétricas para atender o crescente mercado de energia elétrica quando o custo marginal de operação se aproxima do custo marginal de expansão. Ambas possuem a utilidade de evitar o déficit no sistema de energia. Tal solução óbvia não requereria o suporte de tais modelos se esta fosse a única solução. Porém, se existir a chance de haver estratégias capazes de operar o sistema sem déficits e sem novos investimentos dentro de um certo período de tempo, será justificado o desenvolvimento de tais modelos que a encontrem e, mesmo que não encontrem, eles podem apontar que investimentos devem ser feitos, no sentido de que venham a proporcionar o aparecimento de tal estratégia.

Para exemplificar tal situação descrita anteriormente, admitamos a situação em que dois agentes hidrelétricos concorrem pela venda de energia sob o suporte de um agente termelétrico emergencial. O agente termelétrico por ter um preço de venda mais caro que os agentes hidrelétricos não deverá entrar em operação a não ser para prevenção de déficit.

Durante um período de um ano, as operações de cada agente serão determinadas de forma que o equilíbrio do sistema seja mantido, considerando-se que a configuração deste sistema seja suficiente para manter o equilíbrio. A cada estágio é computado o perfil de estratégias de equilíbrio de preço mais atrator. Definidos os preços o agente com preço mais barato vende a sua energia e caso esta energia não seja suficiente para atender toda a demanda o agente com preço mais caro completa a energia que falta. Através de uma equação de balanço

hídrico são calculados os níveis de armazenamento ao final do estágio. A simulação implementada em Matlab segue até o final do período de estudo permitindo que observemos quais estratégias serão utilizadas em cada situação.

Para cada agente hidrelétrico foram dadas duas opções de preço, que são preço alto e preço baixo. O preço baixo é utilizado para ganhar a concorrência que privilegia aquele de menor preço, e o preço alto é utilizado para propiciar a entrada da termelétrica por ordem de mérito, já que o preço dela está situado ligeiramente abaixo do preço alto das hidrelétricas. Num sistema puramente competitivo quando existe sobra de natureza estrutural pela lei da oferta e procura o preço tende a cair e neste caso numa concorrência como esta do exemplo, onde não existe outro tipo de comercialização como leilões, os preços podem inviabilizar a atividade geradora de energia, mostrando que existe uma quantidade ótima a ser investida.

Nos primeiros cinco estágios o equilíbrio em que os dois agentes jogam preços baixos é o único equilíbrio do jogo. Isso acontece, porque no início do período de estudo os reservatórios encontram-se com elevados níveis de volume, resultando no excesso de oferta e permitindo ao agente com preço mais baixo a preferência do consumidor. Com isso os agentes competem por preços baixos.

Nas figuras abaixo vemos retratos dos estados do sistema à medida que os agentes operam.

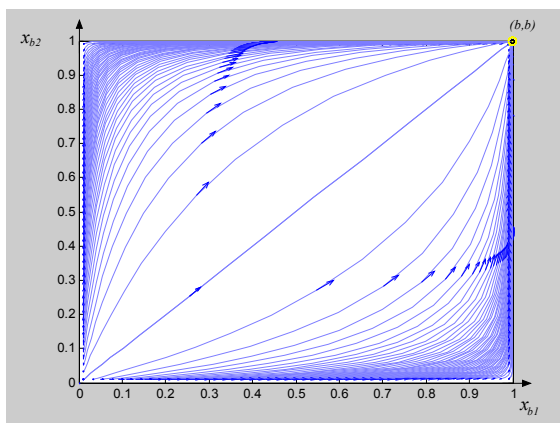


Figura 6-1 – Estágios 1º a 5º

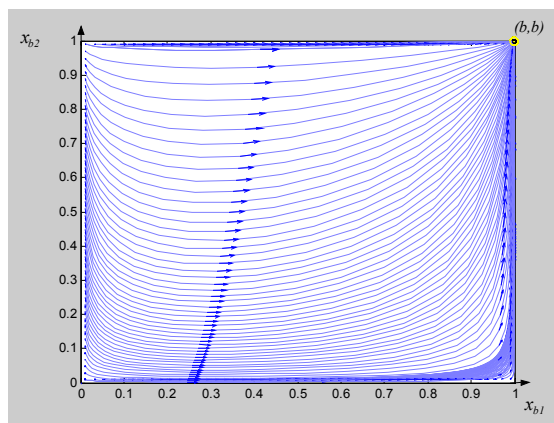


Figura 6-2 – Estágio 6º

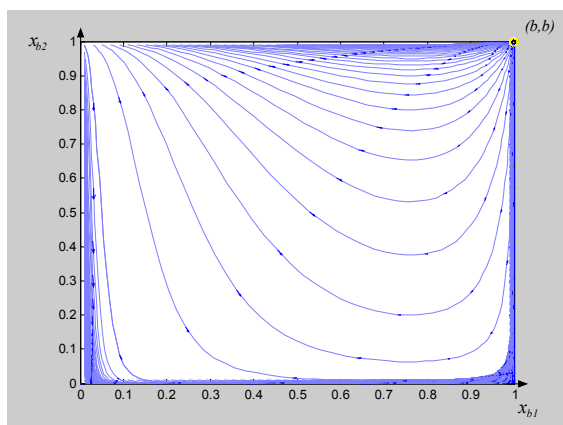


Figura 6-3 – Estágio 7º

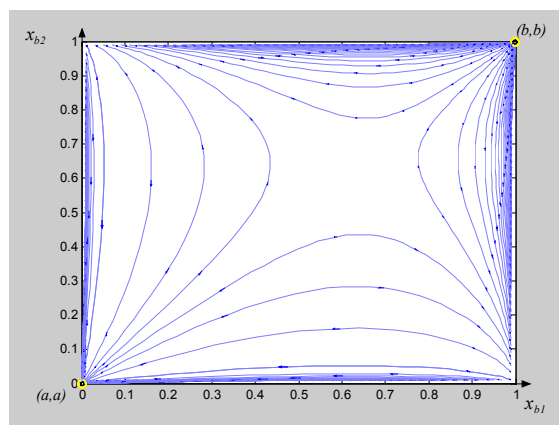


Figura 6-4 – Estágio 8º

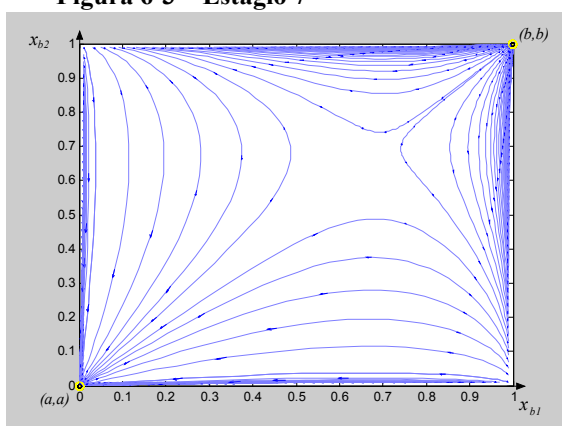


Figura 6-5 – Estágio 9º

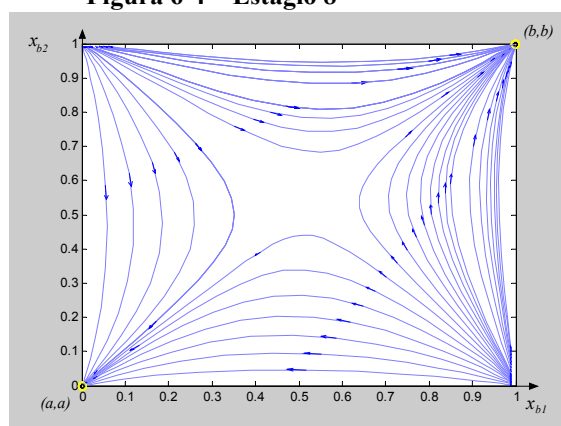


Figura 6-6 Estágio 10º

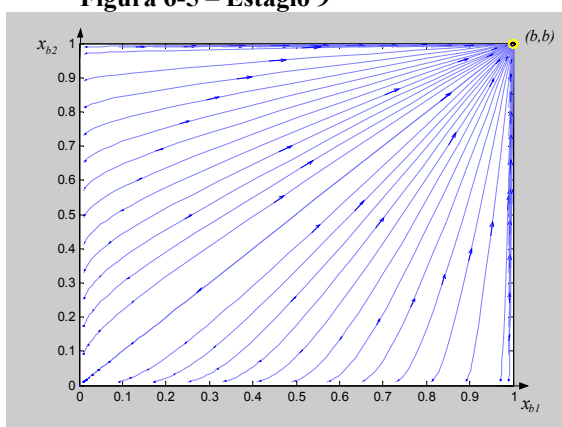


Figura 6-7 – Estágio 11º

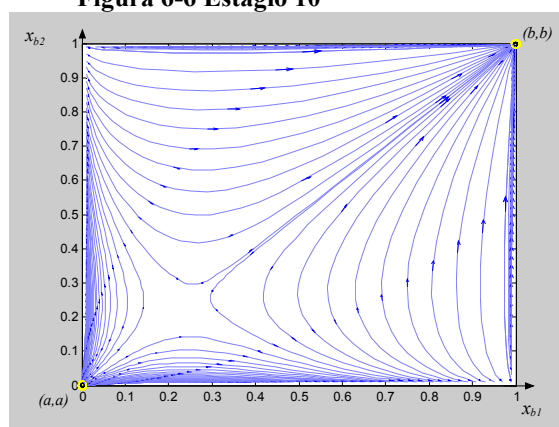


Figura 6-8 – Estágio 12º

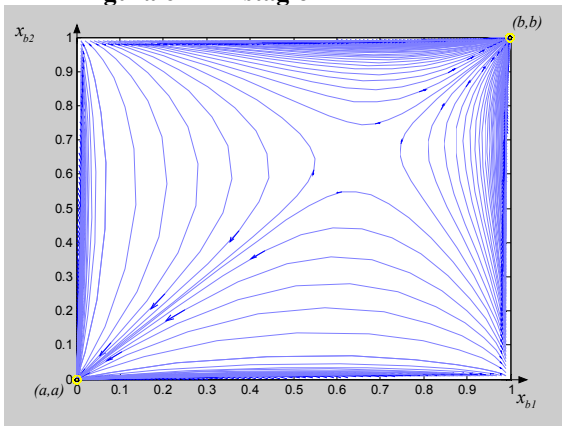


Figura 6-9 – Estágio 13º

A partir do sexto estágio as tensões observadas na bacia de atração já antecipam o aparecimento de um novo equilíbrio em que ambos agentes jogam preços altos. No sexto estágio o perfil de estratégias de preços baixos já perdeu o poder de atração que possuía no início quando os reservatórios estavam cheios.

A partir do oitavo estágio, os níveis individuais dos reservatórios não são mais suficientes para sozinhos suprirem a demanda. Como consequência, no oitavo estágio, já é interessante para os agentes a escolha de preços mais altos.

Este novo equilíbrio de preços altos atrai a maioria dos estados do espaço de estados durante os oitavo e nono estágios. Esta situação é coerente com a escassez dos recursos hidrológicos na aplicação de preços altos. Novamente, no décimo estado, com uma pequena recuperação nos níveis dos reservatórios, as distorções da bacia de atração antecipam a volta do único equilíbrio de preços baixos.

Na Figura 6-10 são mostrados os níveis dos reservatórios que foram operados de modo a estarem sempre em equilíbrio.

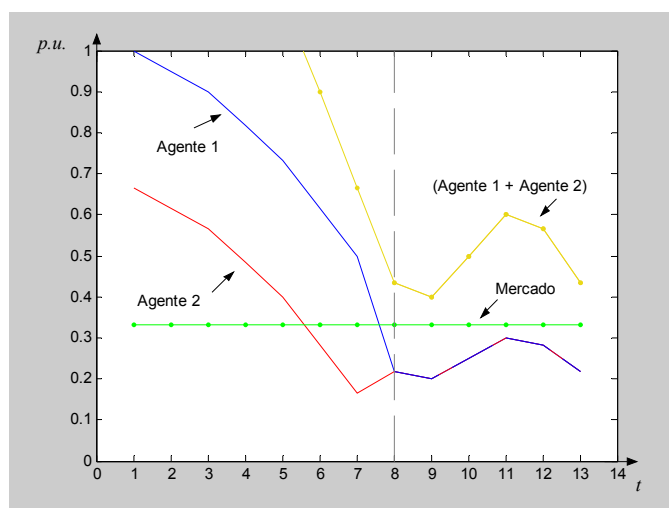


Figura 6-10 – Níveis de armazenamento

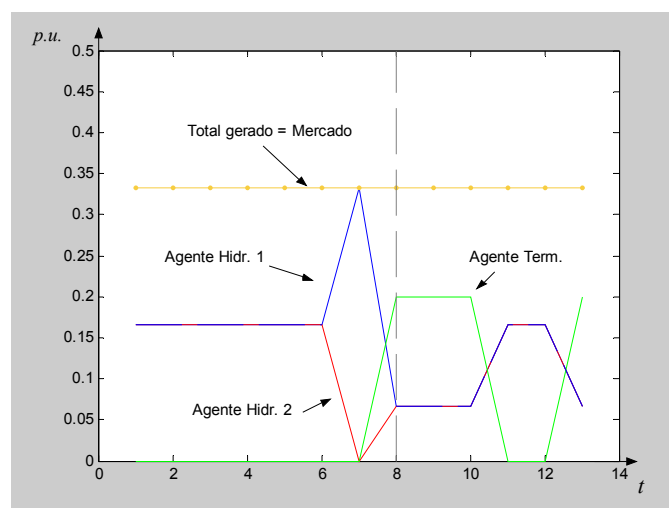


Figura 6-11 – Quantidades geradas pelos agentes

No caso, a configuração do sistema atual conseguiu atender à demanda mesmo em um período seco. Caso isto não fosse possível, não existiria nenhuma estratégia pura que assegurasse um futuro livre de riscos.

O problema neste jogo não é a inexistência de um equilíbrio estrito, mas sim a existência de vários deles. Um modo de escolhê-los, portanto, seria o modo baseado no poder de atração de cada equilíbrio.

Um outro modo seria o de adotar estratégias comportamentais advindas de um especialista ou aprendizado obtido nos próprios cenários da concorrência.

6.2 A estratégia comportamental

Como já abordado neste trabalho, uma importante linha de pesquisa dentro da teoria dos jogos evolutivos é a consideração do comportamento e do aprendizado como impulsionadores da evolução.

Uma estratégia comportamental age conforme a situação. Perante situações de risco alguém poderia ter um comportamento mais parcimonioso, já em situações sem risco este mesmo jogador poderia ser mais agressivo. A consequência de se adotar tais estratégias no jogo seria o aparecimento de estratégias dominantes que não existiriam com o uso de estratégias mais simples.

Vamos utilizar o mesmo exemplo já visto da concorrência entre três geradores de energia. Adotemos de antemão as seguintes estratégias especialistas:

1. Estratégia e_i^1 do jogador i

Se ($Vol\ Tot \geq Niv$) **então**

Preço baixo

Senão

Preço alto

2. Estratégia e_i^2 do jogador i

Se ($Vol\ Tot \geq Niv$) **então**

Preço alto

Senão

Preço baixo

Onde *Vol. Tot.* é o volume total dos dois agentes hidrológicos e *Niv* um determinado nível de volume.

Note que a estratégia comportamental 2 é antagonica à estratégia racional 1. A função desta estratégia é simplesmente propiciar a concorrência entre um estratégia racional mais eficiente e uma estratégia suicida ineficiente. Assim poderemos com facilidade observar o efeito das estratégias comportamentais.

Os resultados entre os primeiro e sétimo estágios foram idênticos aos do exemplo anterior em termos de estados e equilíbrios. Nestes estágios, a estratégia comportamental e_i^1 de ambos os jogadores indica a opção por preços baixos e formam os equilíbrios selecionados pela dinâmica do replicador. Entre os oitavo e nono estágios as mesmas estratégias comportamentais e_i^1 de ambos os jogadores, indicam preços elevados e formam o equilíbrio (e_1^1, e_2^1) . Nos últimos estágios restantes o equilíbrio foi o mesmo (e_1^1, e_2^1) .

Com o aglutinamento de estratégias puras em estratégias comportamentais promove-se o aparecimento de equilíbrios mais robustos às variações do sistema.

Enquanto com estratégias simples os equilíbrios alcançados mudavam conforme a situação, com uma estratégia comportamental um único equilíbrio foi suficiente para operar em todo o período de estudo.

Os estados e equilíbrios resultantes deste exemplo estão dispostos nas Figura 6-12, Figura 6-13, Figura 6-14 e Figura 6-15.

A existência de um equilíbrio evolutivamente estável é interessante para suprir a impossibilidade de uma previsão do futuro que evite qualquer prejuízo. Como o futuro, na grande parte dos casos, não é possível de ser previsto com precisão, é interessante a existência de uma estratégia que seja apta durante pelo menos o período de estudo.

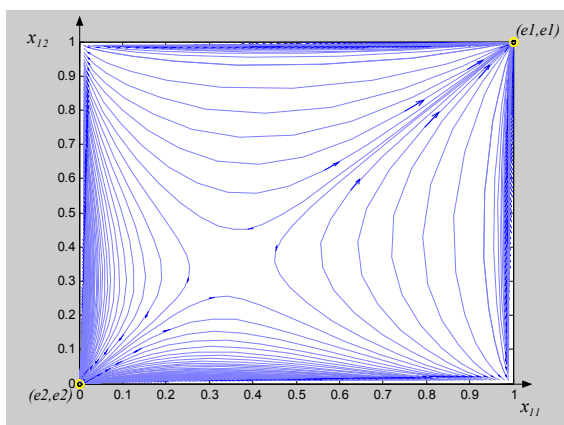


Figura 6-12 – Estágio 8º

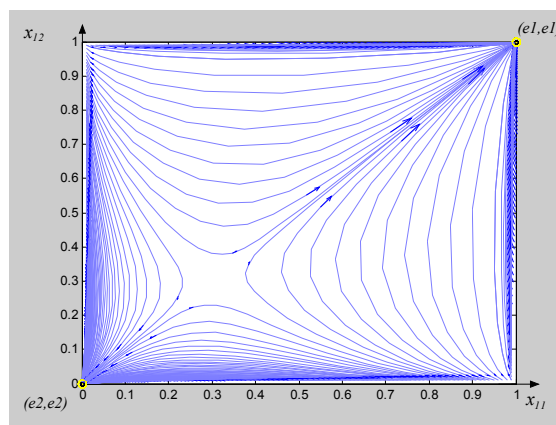


Figura 6-13 – Estágio 9º

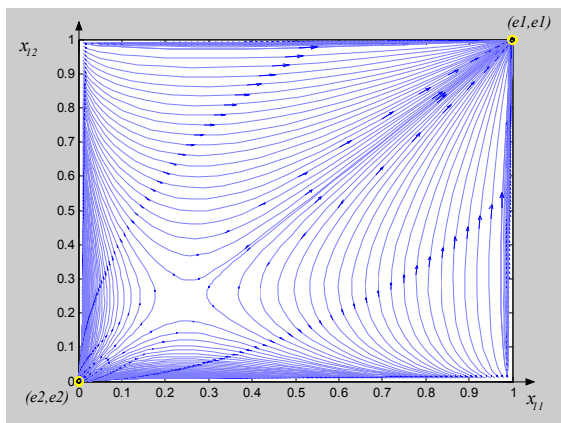


Figura 6-14 – Estágio 12º

Equilíbrios formados por estratégias comportamentais mantêm seu poder de atração durante as várias etapas do jogo

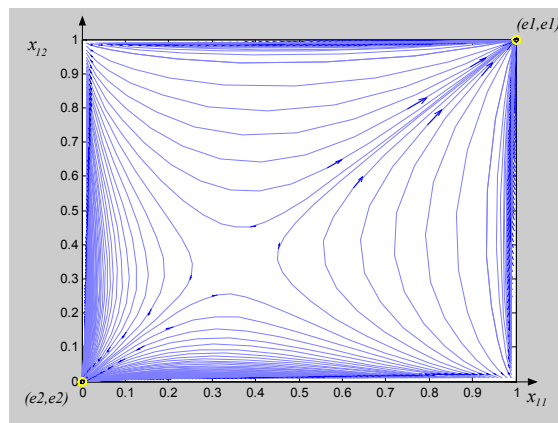


Figura 6-15 – Estágio 13º

As estratégias poderão ser mais sofisticadas imitando o comportamento de especialistas que agem conforme suas experiências sobre a situação atual e crenças futuras. Por exemplo, se o reservatório estiver cheio e o custo futuro obtido dos modelos do planejamento energético for baixo, venda o mais que puder. Se o reservatório estiver vazio e o custo futuro for alto, venda o mínimo possível. Neste caso o cromossomo seria formado por vários genes que se tornariam ativos conforme a situação como no exemplo da estratégia comportamental. Tais estratégias também podem ser criadas com técnicas *neuro fuzzy* em que após um período de aprendizado poderiam tomar a decisão correta, baseadas em algumas variáveis de entrada, como nível atual dos reservatórios e custo futuro. Porém um sistema *neuro fuzzy* não é capaz por si próprio de gerar os cenários de treinamento, nem de saber o custo futuro.

Os algoritmos baseados em jogos evolutivos podem por si próprios gerar os cenários resultantes da concorrência entre os agentes envolvidos e ao mesmo tempo irem treinando um sistema de cromossomos *neurofuzzy*, onde o cromossomo seria uma rede *neurofuzzy* e seus genes regras que representariam as estratégias comportamentais dos agentes a serem treinadas no processo de aprendizado. Ainda existiria a possibilidade do cálculo da variável de custo futuro, tal que a matriz de *payoff* não altere drasticamente suas características de estabilidade evolutiva originais ao longo do período. Assim, ao se percorrer por longo tempo cenários de afluências diversas, deve ser encontrado o melhor custo

futuro para cada um dos estágios do período de estudo, tal que, a quantidade gerada implique nos melhores lucros possíveis sem ocorrência de déficits.

A importância das estratégias comportamentais fica clara pela exploração da nossa própria realidade onde temos que aprender a lidar com as diversas situações de nos comportarmos adequadamente frente a cada uma destas situações.

6.3 Cooperação

A teoria dos jogos evolutivos tem obtido sucesso em jogos nos quais os equilíbrios vêm de um ambiente competitivo, porém alguns pesquisadores alegam que a aplicação da teoria dos jogos evolutivos não funciona como desejado em casos onde a cooperação entre os jogadores traria um benefício social.

Para explicar melhor os argumentos contrários ao uso da teoria dos jogos evolutivos em jogos de cooperação lancemos mão, mais uma vez do dilema dos prisioneiros, onde dois suspeitos são presos, postos em celas separadas e interrogados. Desta vez, exemplificado pelas seguintes matrizes de *payoff*.

$$M = \begin{matrix} & \begin{matrix} e_2^c & e_2^d \end{matrix} \\ \begin{matrix} e_1^c \\ e_1^d \end{matrix} & \begin{bmatrix} (20,20) & (0,40) \\ (40,0) & (5,5) \end{bmatrix} \end{matrix}$$

Onde e_i^c é a estratégia onde o jogador i coopera, ou seja não delata o outro prisioneiro e a estratégia e_i^d é a estratégia onde o prisioneiro i delata o outro.

O único equilíbrio de Nash existente é o perfil (e_1^d, e_2^d) com *payoffs* iguais a cinco para ambos prisioneiros. É justamente este o equilíbrio para o qual a dinâmica do replicador vai convergir.

O que alguns pesquisadores alegam é que, em vez de um lucro de cinco para ambos os jogadores, seria mais vantajoso para ambos se eles ganhassem vinte que seria o caso em que cada um deles jogasse sua estratégia de cooperação.

É importante observar que na situação onde ambos os prisioneiros cooperam existe um risco intrínseco, pois cada um dos prisioneiros perderia totalmente o controle da situação caso seus opositores resolvessem delatar. O único modo de garantir que a cooperação mútua fosse um bom negócio seria um contrato no qual

estivesse embutida uma multa e assim caso um dos dois prisioneiros delatasse, ele acabaria por pagar mais caro.

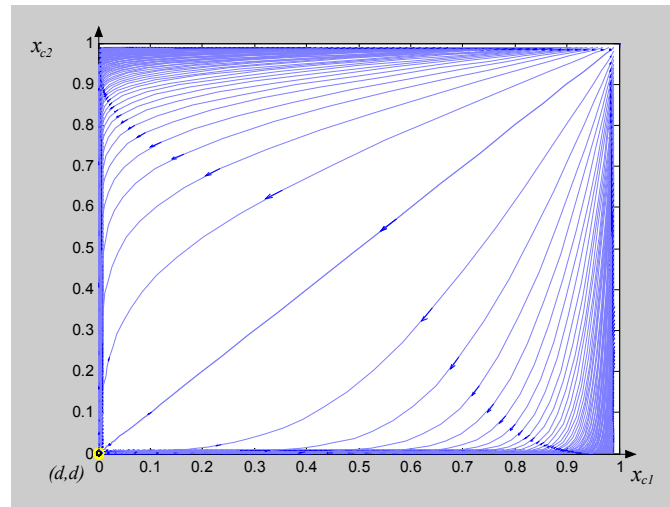


Figura 6-16 – Equilíbrio (delata,delata)

Uma simples transformação na matriz de *payoff* pode levar a dinâmica do replicador a perfis de cooperação. Uma possível transformação seria:

$$\mu'_1 = \mu'_2 = (\mu_1 + \mu_2) + \text{abs}(\mu_1 - \mu_2)$$

Esta função prioriza tanto o bem comum ou, a igualdade de ganhos. Ao aplicarmos esta transformação na matriz de *payoff* M , teremos:

$$M' = \begin{matrix} & e_2^c & e_2^d \\ \begin{matrix} e_1^c \\ e_1^d \end{matrix} & \begin{bmatrix} (40,40) & (0,0) \\ (0,0) & (10,10) \end{bmatrix} \end{matrix}$$

A análise de M' nos mostra agora, que existem dois equilíbrios que são os perfis (e_1^c, e_2^c) e (e_1^d, e_2^d) . Simulando as órbitas que convergem para estes dois equilíbrios pode-se definir as bacias de atração de cada um deles e observar através da Figura 6-17 que o perfil (e_1^c, e_2^c) é mais atrator que o perfil (e_1^d, e_2^d) . É correto admitir que o perfil (e_1^c, e_2^c) seria uma melhor opção para ambos os jogadores. Após a conclusão de que cooperar é interessante para ambos, poderia ser feito um acordo entre a classe dos criminosos, que implicaria em severas penas para quem quebrasse tal contrato.

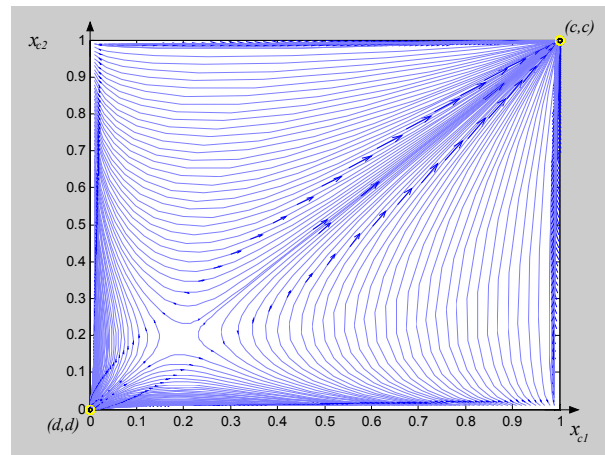


Figura 6-17 – Equilíbrios (delata, delata) e (coopera, coopera)

Esse tipo de acordo ou contrato modificaria a matriz de *payoff* de modo que o novo equilíbrio correspondesse à situação de cooperação na matriz original. Admitamos uma pena de 50 para quem quebrasse o acordo de não cooperar. A nova matriz de *payoff* seria:

$$M' = \begin{matrix} & e_2^c & e_2^d \\ e_1^c & (20, 20) & (0, -10) \\ e_1^d & (-10, 0) & (-45, -45) \end{matrix}$$

Com a inclusão do acordo, passa a existir um único equilíbrio no jogo, que é o de cooperar (e_1^c, e_2^c) . Deste modo a dinâmica do replicador convergirá diretamente para este equilíbrio, mostrando que o tipo de comportamento cooperativo pode ser uma alternativa ao comportamento competitivo sob certas condições.

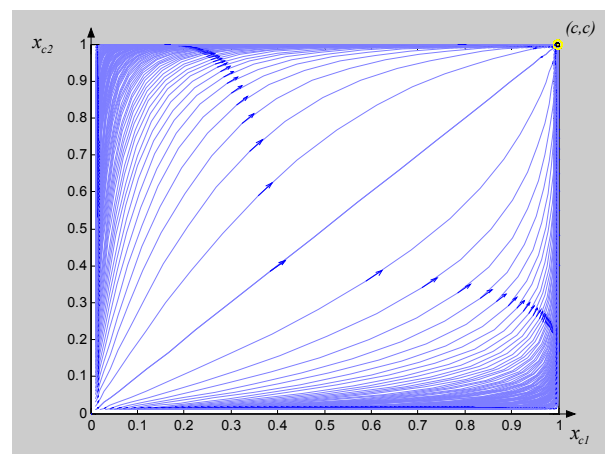


Figura 6-18 – Equilíbrio (coopera, coopera)