JOGOS OCULTOS DE MARKOV

Rafael Souza Nader

Dissertação de Mestrado apresentada ao Programa de Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Sistemas e Computação.

Orientador: Mario Roberto Folhadela Benevides

Rio de Janeiro
Setembro de 2013

JOGOS OCULTOS DE MARKOV

Rafael Souza Nader

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:

Prof. Mario Roberto Folhadela Benevides, Ph.D.

Prof. Carla Amor Divino Moreira Delgado, D.Sc.

Prof. Felipe Maia Galvão França, Ph.D.

RIO DE JANEIRO, RJ - BRASIL

SETEMBRO DE 2013

Nader, Rafael Souza

Jogos Ocultos de Markov/Rafael Souza Nader. – Rio de Janeiro: UFRJ/COPPE, 2013.

XI, 48 p.: il.; 29,7 cm.

Orientador: Mario Roberto Folhadela Benevides

Dissertação (mestrado) – UFRJ/ COPPE/ Programa de Engenharia de Sistemas e Computação, 2013.

Referências Bibliográficas: p. 47-48.

1. Teoria dos Jogos. 2. Jogos Repetidos. 3. Jogos Estocásticos. 4. Modelo Oculto de Markov. I. Benevides, Mario Roberto Folhadela. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

À minha família

Agradeço ao meu orientador, professor e amigo Mario Benevides, que esteve junto comigo em boa parte da minha formação acadêmica, tanto na graduação como no mestrado, e possibilitou que eu chegasse aonde eu cheguei.

Agradeço aos meus amigos da faculdade com quem tenho contato até hoje, em especial o Isaque, que esteve junto comigo nessa jornada acadêmica até o final.

Agradeço à minha namorada Luisa e sua mãe Marise por terem me apoiado e, principalmente, cobrado a dedicação que este trabalho exigiu. A paciência de vocês comigo foi essencial.

Agradeço à minha prima Bia, sua dedicação e empolgação com o mestrado serviram de inspiração para chegar até aqui.

Agradeço principalmente à minha mãe Claudia e ao meu pai Zeca pelos conselhos e conversas, por todo apoio e paciência que tiveram, e aos meus avós, que sempre me incentivaram, cobraram e apoiaram. Todos, mesmo com as eventuais dificuldades no caminho, me deram mais que o necessário para poder realizar essa formação acadêmica.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

JOGOS OCULTOS DE MARKOV

Rafael Souza Nader

Setembro/2013

Orientador: Mario Roberto Folhadela Benevides

Programa: Engenharia de Sistemas e Computação

Este trabalho analisa elementos da teoria dos jogos, em especial jogos repetidos, com ênfase em jogos repetidos com falta de informação. Apresentamos um modelo para uma nova classe de jogos, chamada de Jogo Oculto de Markov, onde temos um jogo estocástico com falta de informação em um dos lados. Neste jogo o jogador assume diferentes tipos ao longo do tempo, onde a transição entre os tipos ocorre de forma estocástica e independente do histórico. O tipo de um jogador define a forma como ele avalia o resultado das decisões dos jogadores em cada rodada.

Esse jogo difere dos trabalhos relacionados pelas características da falta de informação: os jogadores não conhecem como ocorrem as transições do tipo do oponente. Apresentamos um modelo formal para o jogo oculto de Markov e propomos uma solução heurística para ajudar na escolha de qual estratégia jogar. Essa solução envolve uma hipótese sobre jogos estocásticos com as características do jogo tratado aqui. Utilizamos o Modelo Oculto de Markov na solução para auxiliar na falta de informação. Alguns exemplos são utilizados para demonstrar que a solução apresenta bons resultados.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

HIDDEN MARKOV GAME

Rafael Souza Nader

September/2013

Advisor: Mario Roberto Folhadela Benevides

Department: Systems Engineering and Computer

This work analyses game theory elements, specially repeated games, with focus in repeated games with lack of information. We present a model to a new class of game, called here hidden Markov game, which is a stochastic game with lack of information on one side. In this game each player takes different types through the time in a stochastic way and memoryless. The type of a player defines how he evaluates the result of the decisions of each player each round.

This game differs from related works by the lack of information particularity: the players do not know how opponent's type transition occurs. We present a formal model to the hidden Markov game and suggest a heuristic solution to help player chooses which strategy to play. This solution uses a hypothesis about stochastic games with the game features discussed here. We use hidden Markov model to help with the lack of information. Some examples are used to demonstrate the good results of the heuristic and the solution.

Sumário

1 Introdução.. 1

2 Teoria dos Jogos 4

2.1. Conceitos Básicos de Teoria dos Jogos 5

2.1.1. Jogo estratégico ou Jogo normal 6

2.1.2. Exemplo de jogo estratégico: Dilema do Prisioneiro 7

2.1.3. Estratégia dominante 8

2.1.4. Equilíbrio de Nash 9

2.1.5. Equilíbrio misto 11

2.2. Aplicação 12

3 Trabalhos Relacionados 15

3.1. Jogo Repetido com Informação Incompleta 15

3.2. Jogo de cadeia de Markov com falta de informação em um lado 17

4 Jogo Oculto de Markov 19

4.1. Descrição do problema 19

4.2. Exemplo: jogo de tênis 20

4.3. Modelo 20

4.4. Solução 23

4.4.1. Estratégia do jogador informado 26

4.4.2. Racionalidade limitada 27

4.4.3. Cálculo do vetor B 28

4.4.4. Construção da solução 29

4.4.4.1. Geração do Modelo Oculto de Markov 30

4.4.4.2. Gerar observações 30

4.4.4.3. Atualizar modelo oculto de Markov 31

4.4.4.4. Inferir o estado atual 31

4.4.4.5. Jogar contra o estado inferido 31

4.5. Exemplo aplicado: testes 32

4.6. Resultados 35

4.6.1. Resultados do cenário agressivo 35

4.6.2. Resultados do cenário defensivo 40

5 Conclusão e Trabalhos Futuros 45

6 Bibliografia 47

Lista de Figuras

Figura 1 - Modelo do Jogo Oculto de Markov 23

Figura 2 – Estados observáveis 25

Figura 3 - Probabilidade das observações: vetor B 29

Figura 4 - Modelo oculto de Markov 30

Figura 5 – Exemplo Jogo de Tênis: cenário agressivo 33

Figura 6 - Exemplo Jogo de Tênis: cenário defensivo 33

Figura 7 - Cadeia estacionária original - Agressivo 36

Figura 8 - Cadeia treinada – Agressivo 37

Figura 9 - Distância entre as duas cadeias - Agressivo 38

Figura 10 - Recompensa média - Agressivo 39

Figura 11 - Recompensa ao longo das rodadas – Agressivo 39

Figura 12 - Cadeia estacionária original - Defensivo 41

Figura 13 - Cadeia treinada – Defensivo 42

Figura 14 - Distância entre as duas cadeias – Defensivo 43

Figura 15 - Recompensa média – Defensivo 43

Figura 16 - Recompensa ao longo das rodadas - Defensivo 44

Lista de Tabelas

Tabela 1 – Jogos estratégicos 7

Tabela 2 - Dilema do Prisioneiro 8

Tabela 3 - Eliminação das estratégias dominadas 9

Tabela 4 – Par-ou-ímpar 11

Capítulo 1

Introdução

A teoria dos jogos é amplamente utilizada para modelar diferentes problemas na área da economia, campo de atuação no qual se originou. Embora tenha papel fundamental nesta área, não se restringe à ela e tem sido cada vez mais utilizada em diferentes aplicações, podendo destacar sua importância nas relações políticas, relações diplomáticas, biologia, entre outros.

Na computação sua importância e utilização se dá pela possibilidade de modelar e reproduzir fenômenos sociais. Uma vez que a teoria dos jogos apresenta uma formalização matemática de comportamentos humanos, sua relação com o campo da lógica computacional e da inteligência artificial ocorre de forma direta, se tornando um campo pertencente à computação quase tanto quanto à economia.

O estudo da teoria dos jogos se baseia em descrever, modelar e resolver problemas relacionados a interações entre agentes racionais tomadores de decisão. Embora não esteja restrito somente a isso, a origem e a motivação na pesquisa em teorias dos jogos envolve a disputa entre jogadores que escolhem suas ações da melhor forma possível, e sabem que seus oponentes jogam da melhor forma possível também.

Embora a origem filosófica da Teoria dos Jogos date da Grécia antiga, a sua formalização matemática só ocorreu no século XX com von Neumann & Morgenstern (1944). Essa formalização foi um marco na teoria da economia, que sempre foi considerada uma teoria difícil de ser modelada matematicamente, por apresentar fatores indefinidos e interações humanas. Desde então diversos modelos foram criados, visando especializar a teoria existente para fenômenos específicos, melhorando a qualidade dos resultados.

Dentre as classes de jogos existentes, destacamos neste trabalho os jogos repetidos, que refletem as interações sucessivas entre os mesmo agentes. Essa divisão entre jogos de uma única rodada e jogos de múltiplas ou infinitas rodadas não é tão direta, com diferenças importantes entre os dois.

Nesta dissertação apresentamos uma nova classe de jogos repetidos e propomos uma solução heurística para a escolha da estratégia neste jogo, combinando elementos da teoria dos jogos com modelo oculto de Markov. No capítulo 2 faremos uma revisão dos principais conceitos de teoria dos jogos utilizados neste trabalho, com exemplos clássicos e reais de jogos simples.

No capítulo 3 apresentamos trabalhos relacionados, que nos levaram ao modelo proposto. Neste capítulo faremos uma revisão de jogos repetidos, apresentaremos o jogo repetido com falta de informação, introduzido por Aumman e Maschler em 1967. Em seguida apresentamos o jogo de Markov com falta de informação de um lado, trabalho de Renault em que ele utiliza os resultados de Aumman e Maschler em uma classe específica do problema.

No capítulo 4 definimos o jogo oculto de Markov, com a apresentação formal do problema que queremos resolver, juntamente com um exemplo para ilustrar. Neste capítulo apresentamos uma definição formal do modelo proposto para essa classe de jogos, seguida da solução proposta para a escolha de estratégias, que utiliza uma heurística baseada no princípio da racionalidade limitada, e faz um paralelo com modelos ocultos de Markov. Finalizamos com testes que indicam que a heurística utilizada na solução apresenta bons resultados, e comparamos esses resultados com outras possíveis estratégias para o problema.

No capítulo 5 apresentamos uma conclusão e indicações de pesquisa para trabalhos futuros. Avaliamos a contribuição do modelo proposto para a área de teoria dos jogos, as limitações encontradas com os resultados da solução proposta, e analisamos pontos de interesse para pesquisa futura na verificação e validação analítica do modelo, como a demonstração da existência ou não de um valor para o jogo, de estratégias ótimas e de um equilíbrio uniforme.

Capítulo 2

Teoria dos Jogos

O conceito de teoria dos jogos se refere ao estudo de problemas que envolvem a interação entre tomadores de decisão, cada um com o objetivo de maximizar seu retorno com a decisão escolhida. A palavra jogo corresponde a uma atividade competitiva onde os jogadores interagem seguindo regras (Osborne, 2004). O estudo dessas relações pode ser encontrado em textos de filósofos gregos como Platão e Sócrates, mas artigos relacionando o tema à assuntos econômicos tiveram início com Cournot (1838) em seu modelo de duopólio, onde empresas deveriam decidir como seriam suas produções. Embora Cournot tenha feito uma análise de um problema de teoria econômica utilizando elementos relativos à teoria dos jogos, não é apresentado um modelo geral para problemas da área de economia.

A formação da teoria dos jogos como um campo da economia se deu início mais tarde com o trabalho de von Neumann & Morgenstern (1944), responsáveis pela criação do modelo matemático das relações competitivas entre agentes racionais. Este trabalho é responsável pela criação de diversos modelos matemáticos de fenômenos econômicos. Não existia até então nenhum sistema para a teoria econômica, considerada uma ciência de grande dificuldade em ser modelada, pela falta de conhecimento e descrição dos fatos envolvidos (von Neumann & Morgenstern, 1944).

Outro ponto importante na história da teoria dos jogos é o trabalho de Nash (1950), provando a existência de um equilíbrio em um jogo, ou seja, em todo jogo existe pelo menos uma estratégia (pura ou mista, conceitos apresentados a seguir) em que um jogador pode garantir um valor mínimo. Iremos explicar o conceito de equilíbrio de Nash nas seções a seguir.

Desde então inúmeros estudos relacionados à teoria dos jogos surgiram, complementando o trabalho de von Neumann e Morgenstern, visando aumentar a quantidade de fenômenos econômicos modelados pela teoria dos jogos. A capacidade e os ótimos resultados em explicar fenômenos sociais fez com que diversas áreas se interessassem pelo assunto, expandido o número de aplicações relacionadas.

Na área da computação temos importantes aplicações em redes e sistemas multi agentes, e como instrumento para modelagem de problemas reais, sua aplicação está relacionada diretamente com a inteligência artificial, na modelagem de aprendizado e automação de decisões.

Nesta seção definimos os principais conceitos de teoria dos jogos necessários para o entendimento desta dissertação. Apresentaremos a definição de jogos estratégicos, equilíbrio de Nash puro e misto, e alguns exemplos.

2.1. Conceitos Básicos de Teoria dos Jogos

Como dito anteriormente, a teoria dos jogos é a ciência que estuda a interação social entre dois agentes tomadores de decisão. Nesse contexto, um agente pode ser uma pessoa, uma indústria, uma empresa, um país, um político, um partido, etc. O termo tomadores de decisão define que esses agentes possuem ações que devem ser escolhidas. Esses agentes são chamados de jogadores.

Cada jogador possui um conjunto de ações ou decisões que podem ser tomadas, chamadas de estratégias. A escolha de uma estratégia, combinada com as estratégias escolhidas pelos outros jogadores resulta em uma recompensa, chamada de utilidade. O objetivo de cada jogador é escolher uma estratégia que maximize o valor utilidade no jogo.

Esse objetivo, por sua vez, decorre do princípio da escolha racional, base da teoria da economia. Esse princípio enumera que cada jogador age de forma racional, visando maximizar sua satisfação, e sabe que o jogador oponente é racional, e sabe que o jogador oponente sabe que ele é racional, assim sucessivamente. Embora muito se discuta sobre esse princípio em (Simon, 1955), (Arrow, 1986) e (Rubinstein, 1998), a teoria da economia de baseia nele para conseguir modelar as interações sociais. Iremos discutir o significado e as alternativas à escolha racional na seção 4.1.2.

2.1.1. Jogo estratégico ou Jogo normal

Um jogo pode ser definido de diferentes formas, dependendo da sua natureza e das suas características. Neste tópico iremos apresentar o jogo em sua forma mais simples, chamado jogo estratégico ou jogo normal.

Com os conceitos apresentados anteriormente, podemos utilizar a definição formal de jogo estratégico G apresentada em Osborne (2004):

- Conjunto finito de N jogadores
- Conjunto de estratégias S_i para cada jogador
- Função de recompensa $u_i(s)$ para cada perfil $s = (s_1, ..., s_n)$ de estratégias, sendo u_i a função do jogador i e s_i a estratégia do jogador i

Um jogo é a composição das regras acima, e uma jogada é a escolha de uma estratégia feita por cada jogador. No jogo estratégico não temos o conceito de tempo, e a escolha de cada jogador ocorre simultaneamente, no sentido de que os jogadores não conhecem as estratégias escolhidas pelos outros jogadores. Ao final da jogada, cada jogador conhece as estratégias escolhidas e avalia a recompensa associada pela função $u_i(s)$. Um jogo estratégico é um jogo de uma rodada só, ou seja, ocorre somente uma jogada.

Os jogos estratégicos com 2 jogadores podem ser representados por uma tabela com duas dimensões, com uma linha para cada estratégia do jogador 1 e uma coluna para cada estratégia do jogador 2. O valor de cada entrada da tabela é dados pelo par de valores da função $u_i(s)$ para cada jogador.

Jogador 2

		s_3	s_4
Jogador 1	s_1	$(u_1(s_1, s_3),$ $u_2(s_1, s_3))$	$(u_1(s_1, s_4),$ $u_2(s_1, s_4))$
	s_2	$(u_1(s_2, s_3),$ $u_2(s_2, s_3))$	$(u_1(s_2, s_4),$ $u_2(s_2, s_4))$

Tabela 1 – Jogos estratégicos

Como convenção, nos próximos exemplos iremos omitir o rótulo dos jogadores, sendo a linha sempre referente ao jogador 1 e a coluna ao jogador 2. O conjunto de estratégias S pode ser o mesmo para cada jogador, o que geralmente ocorre.

2.1.2. Exemplo de jogo estratégico: Dilema do Prisioneiro

Para ilustrar melhor a definição de jogo estratégico, iremos utilizar um problema clássico de teoria dos jogos, chamado dilema do prisioneiro. O objetivo é mostrar como um problema pode ser modelado, dando significado aos termos apresentados até o momento.

O problema do dilema do prisioneiro envolve dois suspeitos de terem cometido um crime, que serão nossos jogadores. Ambos estão presos em celas separadas e é oferecido a cada um a opção de delatar o outro preso ou negar ter cometido o crime. Essas serão nossas estratégias. Caso os dois neguem o crime, ambos ficam presos por pouco tempo. Caso somente um preso opte pela delação, ele será liberado e o outro ficará muito tempo na prisão. Caso os dois optem pela delação, ambos ficarão um tempo médio na prisão.

Pela descrição do problema, podemos modelar uma função de utilidade que represente a consequência da escolha de cada estratégia. Iremos utilizar os seguintes valores: 0 para muito tempo na prisão, 1 para tempo médio na prisão, 2 para pouco tempo na prisão e 3 para nenhum tempo na prisão. Poderíamos escolher o tempo na prisão como função de utilidade, mas como por convenção queremos maximizar o valor do jogo, utilizamos essa conversão. A tabela a seguir modela o exemplo descrito.

	Negar	Delatar
Negar	(2, 2)	(0, 3)
Delatar	(3 ,0)	(1, 1)

Tabela 2 - Dilema do Prisioneiro

O objetivo de cada jogador é decidir qual estratégia escolher. O princípio da escolha racional indica que cada jogador quer maximizar seu valor e sabe que o oponente vai querer o mesmo.

2.1.3. Estratégia dominante

O objetivo da teoria dos jogos é modelar o problema e resolvê-lo da melhor forma possível. Por resolver entenda-se encontrar a melhor estratégia para cada jogador. Em um modo geral, a teoria formaliza o comportamento humano diante de uma situação como essa. A intuição na escolha da estratégia no problema anterior é simular uma escolha, analisar o que o outro jogador está pensando e em como ele jogaria, e tentar jogar da melhor maneira possível de acordo com o que imagina que ele irá jogar.

Uma estratégia dominante é uma estratégia que, independente de qual estratégia o outro jogador escolher, ela apresenta um valor maior na função utilidade. Isso significa que as outras estratégias são dominadas por essa, e nunca deverão ser escolhidas.

Uma maneira de encontrar a melhor estratégia para um jogador é através da eliminação iterativa de estratégias dominadas, ou seja, analisar cada estratégia e verificar se ela é dominada por alguma outra. Caso seja, retira ela do jogo, e continua o processo. No final irão restar somente as estratégias dominantes.

	~~Negar~~	Delatar
~~Negar~~	~~(2, 2)~~	~~(0, 3)~~
Delatar	~~(3 ,0)~~	(1, 1)

Tabela 3 - Eliminação das estratégias dominadas

Na tabela anterior, podemos verificar que para o jogador 1 a estratégia Negar é dominada pela estratégia Delatar, uma vez que para qualquer estratégia do jogador 2, o valor utilidade para o jogador 1 será maior escolhendo Delatar (3 > 2 e 1 > 0). O mesmo ocorre com o jogador 2. Podemos observar que restou apenas a estratégia Delatar para cada um dos jogadores, com recompensa (1, 1), e essa deve ser a estratégia escolhida.

2.1.4. Equilíbrio de Nash

A estratégia encontrada anteriormente produz um valor inferior à estratégia (Negar, Negar) com utilidade (2, 2) para ambos os jogadores, como pode ser observado na Tabela 2. O motivo para os jogadores não escolherem essa estratégia é porque ela não está em equilíbrio, ou seja, dada a estratégia escolhida pelo outro jogador, existe uma estratégia diferente da atual que produz um valor melhor. Isso significa que se o jogador 1 acha que o jogador 2 irá jogador Negar, para ele é melhor jogar Delatar. Pelo princípio da racionalidade o jogador 2 sabe disso, e dado que o jogador 1 irá escolher Delatar, ele prefere escolher Delatar também.

Uma estratégia em que nenhum jogador queira mudar de estratégia unilateralmente, ou seja, sem que o outro mude de estratégia, é chamada de equilíbrio de Nash. Segue abaixo a definição formal (Gibbons, 1992).

Definição: Em um jogo estratégico G de n jogadores, com um conjunto de estratégias S_i para cada jogador, função de utilidade u_i para cada jogador, as estratégias (s_1*, s_2*, ..., s_n*) são equilíbrio de Nash se, para cada jogador i, s_i* é a melhor resposta para as estratégias dos outros jogadores n - i, (s_1*, ..., s_{i-1}*, s_{i+1}*, ..., s_n*). Ou seja: $u_i(s_1^*, ..., s_{i-1}^*, s_i^*, s_{i+1}^*, ..., s_n^*) \geq u_i(s_1^*, ..., s_{i-1}^*, s_i, s_{i+1}^*, ..., s_n^*)$ para cada $s_i \in S_i$.

O equilíbrio de Nash representa a solução para o jogo, uma vez que nenhum jogador terá interesse em mudar de estratégia e que ela representa o melhor valor dada as estratégias dos outros jogadores. Em (Nash, 1950) é apresentada a prova de que todo jogo estratégico tem pelo menos um equilíbrio de Nash, ou seja, tem uma solução que pode ser calculada utilizando teoria dos jogos. Essa estratégia, no entanto, pode ser pura, sendo representada por uma estratégia do conjunto de estratégias do jogador, ou mista, que será explicado a seguir.

Uma boa descrição do significado do equilibro de Nash é indicada em (Osborne, 2004), apresentando como sendo a situação definida com uma população de indivíduos representando cada jogador, e em cada jogada um indivíduo é selecionado aleatoriamente de cada população. A ação escolhida por cada jogador vai depender das ações escolhidas anteriormente. Em um determinado momento, cada jogador que venha a ser selecionado irá escolher a estratégia escolhida anteriormente, pois a mudança de estratégia não irá melhorar a sua recompensa. Esse estado é um estado estacionário do jogo, assim como um equilíbrio de Nash.

Uma discussão que ocorre sobre a escolha do equilíbrio é que Nash não apresenta uma forma de escolher um equilíbrio caso haja mais de um. Pela descrição anterior, podemos pensar na solução como sendo a escolha que se estabilizou primeiro. Mesmo podendo não ser a melhor opção, nenhum dos jogadores terá incentivo para desviar da sua escolha. Dentre os possíveis equilíbrios, podemos ter um equilíbrio considerado Ponto Focal[1] no jogo. A escolha do Ponto Focal pode

[1] Ponto Focal é o equilíbrio com maior chance de ser escolhido em um jogo com diversos equilíbrios. É expectativa do indivíduo sobre o que os outros esperam que ele espera ser esperado ser feito (Schelling, 1960).

ocorrer por diversos fatores, pode ser um fator estético, temporal, aleatório, cultural. A teoria do equilíbrio de Nash é neutra em relação à escolha feita, ou seja, não diferencia os diferentes equilíbrios (Osborne, 2004).

2.1.5. Equilíbrio misto

No exemplo anterior, nós encontramos um equilíbrio de Nash utilizando eliminação de estratégias dominadas. Mas existem os casos em que esse método não resulta em nenhuma estratégia, como no exemplo abaixo, chamado de par-ou-ímpar, onde o jogador 1 ganha caso os dois escolham estratégias diferentes, e o jogador 2 ganha caso escolham estratégias iguais.

	Par	Ímpar
Par	(-1, 1)	(1, -1)
Ímpar	(1, -1)	(-1, 1)

Tabela 4 – Par-ou-ímpar

Neste exemplo não existe um equilíbrio de Nash composto por estratégias puras, uma vez que cada jogador irá sempre querer trocar de estratégia. Para resolver esse tipo de jogo, utilizamos uma estratégia mista, que é uma distribuição de probabilidade sobre as estratégias puras.

O equilíbrio misto é dado pela distribuição de probabilidade sobre as estratégias que faz com que o jogador não dependa da escolha dos outros jogadores. O objetivo é que a escolha do jogador sendo independente faz com que a escolha dos outros jogadores também fique independente da sua, não permitindo a oportunidade de explorar comportamentos ou falhas que façam a recompensa ser menor para ele. A forma de calcular o equilíbrio misto é a seguinte, utilizando o exemplo anterior:

- Para o jogador 1, atribuímos probabilidade p de jogar Par e 1-p de jogar Ímpar
- A recompensa média do jogador 1 para cada estratégia do jogador 2
  - Par = $\big((p)*(-1)\big)+\big((1-p)*(1)\big)$
  - Ímpar = $\big((p)*(1)\big)+\big((1-p)*(-1)\big)$
- Como queremos que a escolha da estratégia independa da escolha da estratégia adversária, igualamos as duas equações
  - $\big((p)*(-1)\big)+\big((1-p)*(1)\big)=\big((p)*(1)\big)+\big((1-p)*(-1)\big)$
  - $-p+1-p=p-1+p$
  - $-4p=-2$
  - $p=0.5$

Com p = 0.5, o jogador 1 não depende da estratégia do oponente, e o oponente não conseguirá uma estratégia que faça com que o jogador 1 ganhe menos do que a média, dada por:

- $m(p=0.5)=\ \big((0.5)*(-1)\big)+\big((1-0.5)*(1)\big)=0.$

Como provado por Nash (1950), todo jogo estratégico finito apresenta pelo menos um equilíbrio misto.

2.2. Aplicação

São inúmeras as aplicações de teoria dos jogos na área da teoria econômica. Considerada a formalização matemática mais importante de fenômenos sociais, ela é utilizada para modelar situações entre empresas, estados, governantes, diplomatas, jogadores, sistemas inteligentes, etc.

Como dito no início do capítulo, a criação[2] da teoria dos jogos por von Neumann & Morgenstern (1944) possibilitou o estudo de interações entre tomadores de decisões em níveis não atingidos antes, tendo como resultado a análise de fenômenos na economia como a entrada de uma nova empresa no mercado, a criação de monopólios, duopólios e oligopólios, cálculo de preço de mercadorias, controle da produção em industrias.

Não demorou muito e sua aplicação se estendeu para a área da política, analisando a geração de partidos, coligações, resultados de votações, escolha de representantes, etc. Na área diplomática atua na interação entre países, escolha de aliados, cumprimento de exigências e de acordos.

Em (Aumann & Maschler, 1995) é discutida a importância da Teoria dos Jogos no mundo real, mais especificamente no mundo da diplomacia política, e serviu de motivação para o desenvolvimento deste trabalho. O material apresentado no livro foi desenvolvido entre 1966 e 1968, com a reunião de diversos matemáticos, entre eles os autores do livro, para pensar o problema do desarmamento. De diversas reuniões do grupo Mathematica Inc, foram escritos relatórios encomendados pela agência americana de controle de armas e desarmamento (USDA em inglês). O governo estava pedindo consultoria aos matemáticos para o problema do desarmamento durante a Guerra Fria, problema este que foi modelado como sendo de Teoria dos Jogos, e será discutido mais na frente.

É interessante observar a posição dos próprios autores em relação às limitações do uso de Teoria dos Jogos para resolução do problema: “No mathematical treatment can possibly take account of all the implications involved in a political decision. (…) The analysis of such a highly simplified abstraction can very seldom lead to any specific recommendations in a specific situation. But it can lead to insights of general nature.

[2] Embora o assunto já fosse discutido antes, consideramos a sua criação com a

These insights can then be used by policy makers in making specific decisions or in formulating general policies.” (Aumann & Maschler, 1995)

---

formalização matemática apresentada em (von Neumann & Morgenstern, 1944).

Capítulo 3

Trabalhos Relacionados

Neste capítulo iremos apresentar dois trabalhos que direcionaram a nossa pesquisa e foram importantes na geração do modelo proposto. O primeiro é o trabalho de Aumann & Maschler (1995), considerado o pioneiro na classe de problemas de jogos repetidos com falta de informação. Embora a publicação tenha ocorrido em 1995, os trabalhos foram desenvolvidos entre 1966 e 1968, período de guerra fria.

O outro trabalho é a extensão do trabalho anterior para uma classes específica de problemas. Renault (2006) apresenta uma abordagem para o problema de jogos repetidos com falta de informação aplicada em jogos estocásticos. É discutida a existência de um valor para o jogo e de estratégias ótimas.

3.1. Jogo Repetido com Informação Incompleta

Jogos repetidos e falta de informação sempre foram elementos importantes na modelagem de problemas em teoria dos jogos. O trecho apresentado abaixo é retirado de Aumann & Maschler (1995), e serve de motivação para o estudo de problemas do gênero:

"*Whenever one wants to apply game theory to real life situations, one has to bear in mind, among other things, two aspects in which classical game theory differs from real life:*

I - unlike a game situation, which is essentially a one-shot affair, a conflict situation in real life usually leads to another conflict situation. Thus, in real life, when one takes actions one should consider not only the immediate payoffs but also the effect that the actions may have on the other conflict situations that will occur in the future.

II - unlike in a game situation, in which it is assumed that the players know all the available strategies as well as the payoff functions, in real life each participant usually has only partial knowledge of the strategies available to him and to the other participants; moreover, the actual payoff that results from the possible actions taken by the participants is impossible to determine because of a lack of knowledge of relevant facts."

O conflito em questão tratado na série de artigos encomendada pelo departamento de controle de armas e desarmamento dos Estados Unidos (ACDA em inglês) é em relação ao cumprimento ou não dos acordos de desarmamento. A descrição do problema envolve dois países que têm que decidir entre cumprir ou não termos de desarmamento impostos. Ambos são incertos em relação à alguma informação relevante nesta decisão, tal como arsenal atual ou capacidade de produção de armas do outro país.

Dada a motivação para o estudo, definimos abaixo o modelo apresentado por Aumann & Maschler (1995), que chamaremos neste trabalho de jogo repetido com informação incompleta em um lado:

- Dois jogadores, J_1 e J_2
- Conjunto finito K de estados
- Distribuição de probabilidade $p = p(k)$, $k \in K$, em K
- Conjunto de estratégias S_i para cada jogador
- Conjunto de jogos estratégicos $G = \{G_1, ..., G_K\}$, com o mesmo conjunto S_i de estratégias para cada jogador

No início do jogo um estado k em K é sorteado de acordo com p e anunciado ao jogador J_1. Este jogo tem falta de informação em um lado, no caso o jogador J_2, que desconhece o estado sorteado. A partir de então o jogo G_k é jogado repetidamente, e a cada rodada é informado aos dois jogadores somente a estratégia s_i de cada jogador. A recompensa de cada jogo não é informada, mas é dada por $G_k(s_1, s_2)$.

Em (Aumann & Maschler, 1995) é apresentada uma prova da existência de um valor v para esse jogo, ou seja, um valor mínimo que cada jogador pode adquirir no final jogo, além da prova da existência de uma estratégia ótima para cada jogador que garante v.

É discutido ainda se o jogador J_1 deve ou não jogar de forma reveladora, ou seja, utilizando a informação do estado k informada para ele no início do jogo. A ideia é que se o jogador utilizar essa informação, ao longo do jogo o jogador J_2 pode descobrir o estado inicial k. A estratégia ótima apresentada envolve uma escolha aleatória por parte do jogador J_1 de qual jogo G_k jogar em cada rodada, fazendo com que o jogador J_2 não consiga criar uma relação entre as jogadas dele e o estado inicial.

3.2. Jogo de cadeia de Markov com falta de informação em um lado

No trabalho de Renault (2006), temos um jogo repetido com algumas características diferentes do apresentado anteriormente. Neste trabalho ele apresenta um jogo estocástico, também conhecido como jogo de Markov ou jogo de cadeia de Markov, que em resumo é um jogo repetido com uma série de estados associados, com um jogo diferente em cada estado, mas que o estado atual muda com o tempo, através de uma distribuição de probabilidade independente do tempo, conforme uma cadeia de Markov.

A definição formal deste jogo é dada por:

- Dois jogadores, J_1 e J_2
- Conjunto finito de K de estados
- Distribuição de probabilidade inicial $\pi = \pi$ (k), k $\in$ K, em K
- Matriz de transição M = M(k, k') de probabilidade de transição de cada estado para os outros estados
- Conjunto de estratégias S_i para cada jogador

- Conjunto de jogos estratégicos G = {G_1, ..., G_K}, com o mesmo conjunto S_i de estratégias para cada jogador

O jogo inicia com um estado inicial k_0 dado pela loteria em π e informado apenas ao jogador J_1. Os dois jogadores escolhem suas estratégias, e são informados da escolha do outro jogador, e o próximo estado será escolhido pela matriz de transição M e novamente será informado somente ao jogador J_1.

Renault apresenta formas de calcular o valor *v* para algumas variações desse jogo, como o caso do jogo repetido infinitamente, jogos de soma-zero, e estratégias ótimas para cada jogador. Segundo ele, o jogador informado não deve jogar uma estratégia reveladora, ou seja, não deve utilizar a informação sobre o estado atual para maximizar sua estratégia. A estratégia ótima se baseia em uma série de escolhas ponderadas pela distribuição de probabilidade dos estados e as ações escolhidas pelos jogadores.

Este trabalho serviu como motivação para o estudo de classes de jogos repetidos, que nos levou a geração do modelo apresentado a seguir, juntamente com o trabalho de Aumann & Maschler, onde fica clara a importância de jogos repetidos com falta de informação para modelagem de fenômenos utilizando teoria dos jogos.

Capítulo 4

Jogo Oculto de Markov

O problema da existência de um equilíbrio uniforme em jogos repetidos com falta de informação em um lado, assim como o cálculo do valor do jogo e a existência de uma estratégia ótima para os dois jogadores ainda é um problema em aberto em algumas classes de jogos (Renault, 2012). O estudo de classes de jogos deste tipo é motivada por uma tentativa de encontrar melhores formas de se jogar um jogo com características distintas de informação.

O que propomos neste trabalho é um modelo para uma nova classe específica de jogos repetidos com falta de informação em um lado, chamada de Jogos Ocultos de Markov, apresentando uma solução heurística para o mesmo. Apresentamos neste capítulo a discussão sobre escolha racional presente em (Simon, 1955), (Rubinstein, 1998) e (Arrow, 1986) como embasamento para a nossa solução. Concluímos com alguns testes e comparações de resultados.

4.1. Descrição do problema

O problema apresentado por Renault (2006) ainda exige que o jogador com falta de informação tenha conhecimento de como o estado do jogador informado muda de acordo com o tempo, ou seja, conhece a distribuição de probabilidade de cada estado. Neste trabalho estamos interessados nos problemas em que isso não é conhecido.

Jogo oculto de Markov é um jogo de cadeia de Markov com falta de informação de um lado, onde essa falta de informação é ainda maior. Nas outras classes de jogos com falta de informação, sempre se assumiu um mínimo de conhecimento, como a distribuição de probabilidade dos estados nos jogos Bayesianos, o vetor de transição entre os estados e o estado atual nos jogos estocásticos ou a cadeia de Markov dos estados no jogo de Markov com falta de informação.

4.2. Exemplo: jogo de tênis

O exemplo utilizado para ilustrar o problema descrito anteriormente simula uma partida de tênis, onde estamos interessados somente no momento do saque. Os dois jogadores são o sacador e o receptor. As estratégias possíveis para o sacador são jogar no centro, de forma mais segura, ou jogar aberto, mais arriscado. Para o jogador receptor, as estratégias são tentar prever e se preparar para um saque no centro ou para um saque aberto. Para efeito de nomenclatura, os dois conjuntos de estratégias serão iguais.

O jogador sacador apresenta a característica de mudar de tipo, que pode ser entendido como o seu comportamento ao longo do jogo. Como resultado dessa mudança, sua eficiência com cada estratégia escolhida muda. Os tipos apresentados pelo jogador sacador são agressivo, moderado ou defensivo. A mudança de tipo do jogador ocorre de forma estocástica, mas independente do tempo e do histórico de tipos, importando somente o tipo atual. Tal mudança pode ser modelada como uma cadeia de Markov.

O jogador receptor sabe que o oponente apresenta essa variação de tipos e conhece cada comportamento dele, mas não sabe qual o tipo atual, nem conhece a forma como a mudança de tipo ocorre.

O jogo é composto por várias rodadas repetidas entre os dois jogadores, um número finito, mas desconhecido, de vezes. Ao final de cada rodada, os dois jogadores são informados das ações de cada um. A recompensa de cada jogador pode ser vista como a performance de cada um até o próximo saque. Essa recompensa não é informada imediatamente aos jogadores.

4.3. Modelo

“Um modelo é uma abstração que usamos para entender nossas observações e experiências. (...) Se ele melhora nosso entendimento do mundo, então serve ao seu propósito.” (Osborne, 2004). Esse trecho mostra a importância dos modelos para a compreensão e estudo de diversos fenômenos. O objetivo do modelo deve estar bem estabelecido para que as variáveis mais importantes sejam levadas em consideração, enquanto as variáveis menos relevantes sejam ignoradas.

O problema que queremos resolver, descrito anteriormente, é um caso particular do jogo apresentado em (Aumann & Maschler, 1995) e (Renault, 2006), e o nosso modelo é baseado nesses dois trabalhos. O principal diferença na definição é a mudança na falta de informação que o jogador tem.

A descrição formal do problema segue abaixo:

- Um conjunto finito de N jogadores
- Um conjunto de estratégias S_i para cada jogador
- Um conjunto de tipos T_i com K_i tipos para cada jogador
- Uma função de transição $\mathcal{T}_i : T_i \times S_1 \times ... \times S_N \rightarrow \mathcal{PD}(T_i)$ para cada jogador
    - Onde $\mathcal{PD}(Ti)$ é uma distribuição de probabilidade sobre $T_{i.}$
- Um conjunto de jogos estratégicos $G(t_1, t_2, ..., t_N)$ para cada combinação de tipos $t_i \in T_i$ de cada jogador, com o mesmo conjunto S_i de estratégias para cada jogador
- Funções de recompensa u_i em cada jogo G
- Distribuição de probabilidade $\pi_{i,\ j}$, representando a crença inicial de cada jogador i sobre o tipo do oponente j, para todos os jogadores
- Um conjunto finito de estados observáveis $\mathbf{O}(S_1, ..., S_N)$, representando a informação divulgada em cada rodada associada à estratégia escolhida por cada jogador

As informações disponíveis para cada jogador são:

- Os jogadores conhecem o conjunto de tipos T_i de cada jogador
- Os jogadores conhecem o conjunto de estratégias S_i de cada jogador

- Os jogadores conhecem os jogos estratégicos e suas respectivas tabelas de recompensa de cada combinação de tipo $T_1 \times T_2 \times ... \times T_N$
- Os jogadores conhecem o conjunto de estados observáveis **O**
- Os jogadores não conhecem a função de transição dos outros jogadores
- Os jogadores não são informados do estado atual dos outros jogadores
- Os jogadores não são informados da recompensa ao final de cada rodada

Para melhor entendimento do modelo, neste trabalho estamos interessados em jogos com dois jogadores e com falta de informação em um lado, ou seja, somente um jogador apresenta mais de um tipo. Iremos utilizar N = 2 jogadores, os conjuntos de estratégias $S_1 = S_2 = \{A, B\}$, os conjuntos de tipos $T_1 = \{\text{Tipo } 1\}$ e $T_2 = \{\text{Tipo } 2, \text{Tipo } 3\}$, vetor de transição entre os tipos $\mathcal{T}_1$ = {Tipo 1 -> Tipo 1 = 1}, $\mathcal{T}_2$ = {Tipo 2 -> Tipo 2 = a_2, Tipo 2 -> Tipo 3 = 1- a_2, Tipo 3 -> Tipo 3 -> a_3, Tipo 3 -> Tipo 2 -> 1- a_3}.

Na figura abaixo temos a representação do modelo. Temos os dois jogos $G(t_1, t_2)$ e $G(t_1, t_3)$, com as respectivas tabelas de recompensa. A distribuição de probabilidade inicial $\pi_{1,2} = \{t_2 = x, t_3 = y\}$ e $\pi_{2,1} = \{t_1 = 1\}$.

O conjunto de estados observáveis **O** representa a informação divulgada ao final de cada rodada, associados aqui à estratégia escolhida por cada jogador.

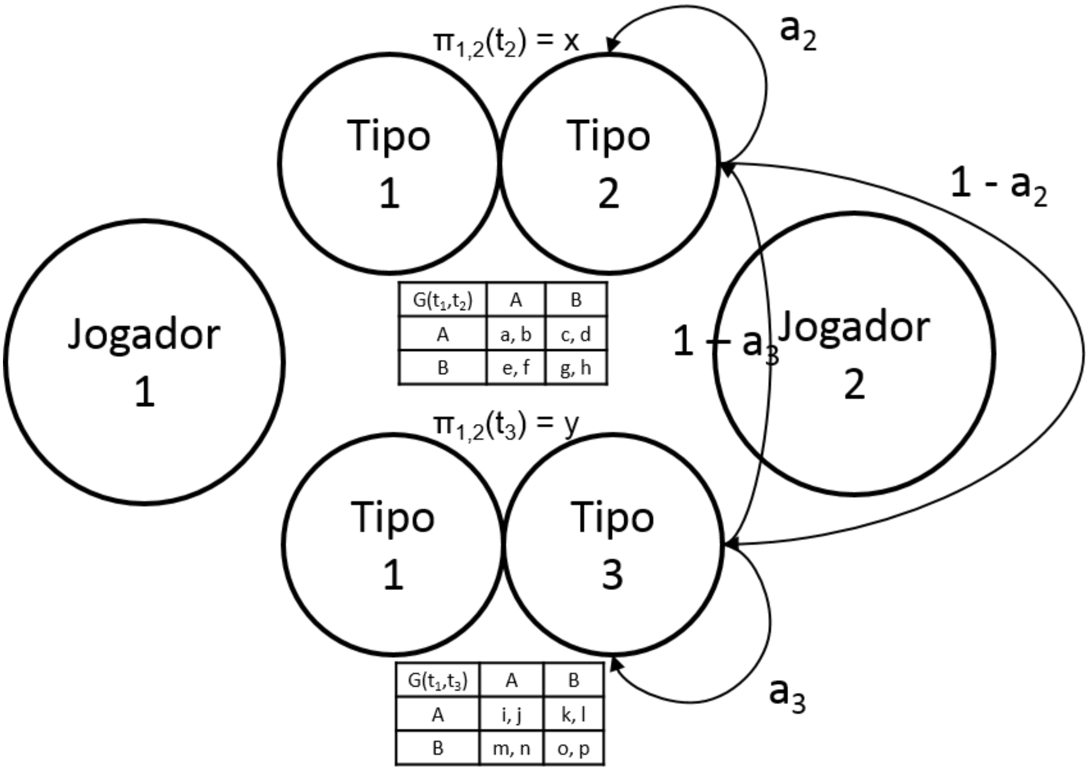

Figura 1 - Modelo do Jogo Oculto de Markov

O jogador 1 é o jogador com falta de informação, logo desconhece o estado atual em cada rodada do jogador 2, assim com a função de transição $\mathcal{T}_2$. Ele é informado somente da estratégia escolhida pelo jogador 2, através do estado observável associado. O jogador 2 é completamente informado, uma vez que o jogador 1 só tem um tipo.

Com o modelo definido, estamos interessados em resolvê-lo, ou seja, apresentar uma forma de cada jogador escolher suas estratégias a cada rodada. Na próxima seção apresentamos a solução proposta.

4.4. Solução

Como o problema do cálculo do equilíbrio, da geração de um valor para o jogo ou da geração de uma estratégia ótima para os jogadores é um problema em aberto, o que propomos para a resolução do jogo apresentado é uma forma prática de escolher a estratégia a ser jogada. Uma heurística é apresentada e comparada com outras possíveis soluções para o problema, que não necessariamente utilizam teoria dos jogos.

Como dito em (Osborne, 2004), “a teoria dos jogos estratégicos se presta ao estudo experimental: providenciar para indivíduos jogarem e observarem suas escolhas é uma relação direta.”, se referindo à necessidade na geração de um modelo de se relacionar com comportamentos reais, que possam ser observados empiricamente.

A solução proposta parte da relação direta que existe entre o modelo apresentado e Modelo Oculto de Markov. A ideia, apresentada primeiramente em (Waghabi, 2009), utiliza a natureza estocástica dos dois modelos para acrescentar informações ao jogador com falta de informação. O que queremos é, a partir das informações apresentadas, inferir outras informações que auxiliem na escolha de uma boa estratégia para o jogador. A intenção é utilizar o ferramental disponível para modelos ocultos de Markov no nosso modelo.

A existência dos estados observáveis no modelo relacionados com as estratégias escolhidas por cada jogador resulta na relação mostrada na figura abaixo. A figura representa a visão do jogador 1 sobre o jogo, sendo a linha tracejada a divisão entre a informação disponível (abaixo) e a informação omitida (acima).

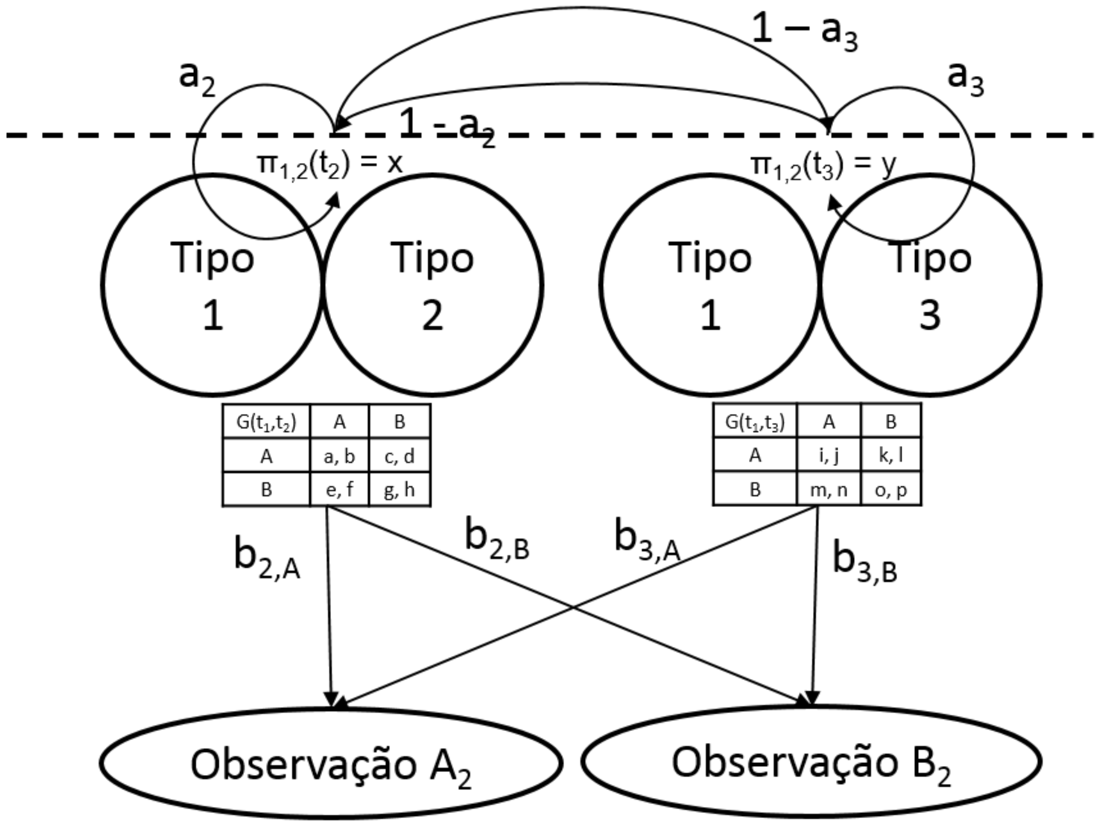

Figura 2 – Estados observáveis

O valores $b_{2,A}$, $b_{2,B}$, $b_{3,A}$, $b_{3,B}$ representam o que chamaremos de vetor B. A cada rodada, a escolha da estratégia pelo jogador 2 resulta em uma observação para o jogador 1. Essa observação representa a estratégia escolhida. O vetor B representa a relação entre o jogo $G(T_1, T_2)$ e a estratégia escolhida.

A representação acima do jogo facilita a visualização da relação entre o modelo apresentado e modelos ocultos de Markov. Temos na definição do nosso jogo mostrada acima todos os elementos necessários para a construção do modelo oculto:

- Conjunto de estados ocultos, representado por {(Tipo 1, Tipo 2), (Tipo 1, Tipo 3)}
- Vetor de transição A entre os estados ocultos, representado pela transição $\mathcal{T}_2$
- Conjunto de estados observáveis **O**, representado pelas estratégias escolhidas {Observação A_2, Observação B_2}

- Vetor de transição B entre os estados ocultos e os estados observáveis

O nosso objetivo é apresentar uma solução para a escolha das estratégias. Primeiramente estamos interessados no caso específico do jogador com falta de informação, onde queremos, a partir das observações, apresentar uma heurística razoável de escolha. Por isso estamos interessados na construção do modelo oculto de Markov. Através dele é possível inferir valores sobre as informações não disponíveis, possibilitando uma escolha menos arbitrária.

Para a construção do modelo oculto de Markov temos todas as variáveis disponíveis, com exceção do vetor de transição A dos estados ocultos e o vetor B de transição entre os estados ocultos e os estados observáveis. Como mostrado anteriormente, o valor do vetor B está relacionado diretamente com a estratégia adota pelo jogador informado. Surge então a questão de qual estratégia o jogador 2 deverá seguir.

4.4.1. Estratégia do jogador informado

Nesta seção apresentamos uma hipótese para a escolha da estratégia por parte do jogador informado. Como o problema de um valor para o jogo está em aberto, assim como a existência de uma estratégia ótima e um equilíbrio para o jogo, discutiremos aqui uma possibilidade de estratégia que garante um valor mínimo para o jogador.

Hipótese: a estratégia S_2 do jogador informado em cada rodada será a estratégia do equilíbrio misto do jogo $G(t_2)$ da rodada atual.

A escolha da estratégia no equilíbrio misto de cada jogo faz com que a recompensa do jogador não dependa da estratégia escolhida pelo outro jogador. Como mostrado na seção 2.1.5, o equilíbrio misto representa uma distribuição de probabilidade entre as estratégias, de forma que a recompensa tenda a ser a mesma.

Não conseguimos provar que esta estratégia é um equilíbrio no jogo repetido como um todo, uma vez que não temos ainda uma forma de calcular as estratégias ótimas de cada jogador. No entanto, baseamos essa escolha heurística na discussão apresentada em (Simon, 1955), (Arrow, 1986) e (Rubinstein, 1998), onde é questionado o princípio da escolha racional.

4.4.2. Racionalidade limitada

Arrow (1986) questiona o uso da racionalidade como princípio básico no estudo da teoria econômica. Para ele, a racionalidade varia de acordo com as situações e não é uma coisa estática e característica do indivíduo, depende do meio. A racionalidade utilizada na teoria é plausível em condições ideais, mas se as condições se degeneram, se degenera também o conceito de racionalidade. Contudo, Arrow coloca que o princípio da escolha racional é essencial para a teoria econômica, sem o qual não haveria qualquer teoria, e mesmo com muitos autores discordando do princípio, não veem um alternativa viável.

Rubinstein (1998) fala sobre a racionalidade limitada, introduzida por Simon (1955), que se refere à tentativa de tornar mais reais os modelos de teoria dos jogos. A ideia de que cada jogador tem controle e conhecimento completo das suas estratégias, das expectativas que cada combinação de estratégias resulta e raciocinam de forma perfeita a ponto de conhecer exatamente os equilíbrios de determinado problema é colocada em discussão.

O conceito de racionalidade limitada é o fato de que um indivíduo, na hora de tomar uma decisão, pode em muitos casos não ter conhecimento de tudo, não ter toda a informação.

Simon (1955) propõe que as escolhas devem seguir estratégias que façam sentido em um contexto humano da situação, onde a falta de informação e o comportamento de cada indivíduo deve ser levado em consideração. Utiliza de elementos empíricos para justificar seu ponto de vista, e apresenta como opção de escolha de estratégia o uso de heurísticas ao invés das estratégias ótimas, representando o conceito de “estratégias satisfatórias”, onde a recompensa não é máxima, mas é um valor suficiente.

No nosso trabalho apresentamos um solução que se encaixa nessa definição. A escolha do equilíbrio misto em cada rodada, garantindo um valor mínimo em cada jogo, é uma estratégia satisfatória dada a falta de informação apresentada no problema.

4.4.3. Cálculo do vetor B

Com a escolha da estratégia do jogador informado como sendo o equilíbrio misto no jogo do tipo atual, podemos voltar a definição do vetor B de transição entre os estados ocultos e os estados observáveis.

Como dito anteriormente, o vetor B representa a probabilidade de cada observação ocorrer em cada estado oculto. Como a observação em questão é a estratégia escolhida pelo jogador, a probabilidade $b_{2,A}$ da observação A_2 é a probabilidade do jogador jogar a estratégia A no jogo $G(t_1,t_2)$. Essa probabilidade é dada calculando a estratégia do equilíbrio misto, dada pela equação:

- Ganho quando o adversário utiliza a estratégia A = $\big((p) * (b)\big) + \big((1 - p) * (d)\big)$
- Ganho quando o adversário utiliza a estratégia B = $\big((p) * (f)\big) + \big((1 - p) * (h)\big)$

Igualamos os ganhos:

- $\big((p) * (b)\big) + \big((1 - p) * (d)\big) = \big((p) * (f)\big) + ((1 - p) * (h))$
- $p = \frac{(h-d)}{(b-d-f+h)}$

O cálculo da probabilidade $b_{2,B}$ da observação B_2 é análogo e dado por (1-p).

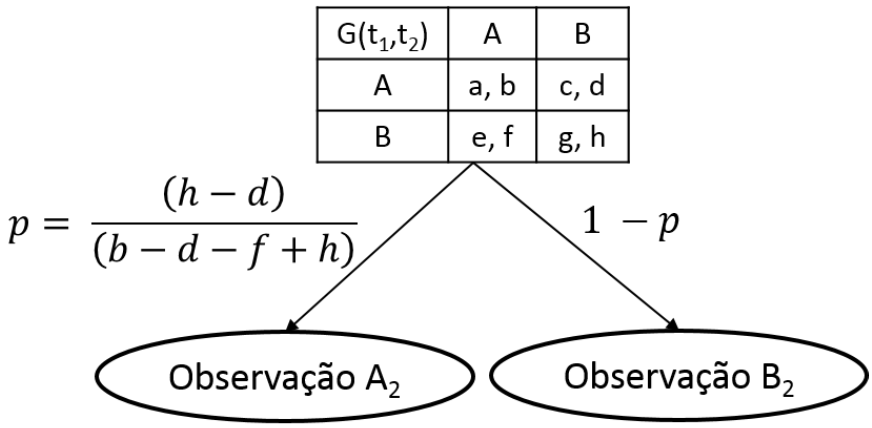

Figura 3 - Probabilidade das observações: vetor B

4.4.4. Construção da solução

Com a definição da estratégia do jogador informado como sendo jogar no equilíbrio misto em cada jogo G, resta definirmos qual estratégia o jogador não informado deverá jogar. A escolha de uma estratégia apresenta a dificuldade do desconhecimento do tipo atual e da forma como o tipo muda com o tempo, dificultando a inferência do tipo atual utilizando métodos probabilísticos como em (Renault, 2006).

Como dito anteriormente, utilizaremos um modelo oculto de Markov, relacionado com o modelo do jogo oculto de Markov apresentado, e os algoritmos disponíveis para o mesmo para o cálculo das transições dos estados ocultos e a inferência do estado atual.

Apresentamos a seguir os passos para a construção da solução:

1. Geração do Modelo Oculto de Markov relacionado ao Jogo Oculto de Markov
2. Gerar observações
3. Atualizar modelo oculto de Markov
4. Inferir estado atual
5. Jogar contra o estado inferido

4.4.4.1. Geração do Modelo Oculto de Markov

A solução proposta parte do mapeamento dos tipos do jogador oponente com a cadeia oculta no modelo oculto de Markov. Como mostrado anteriormente na seção 4.4, e com o cálculo do vetor B em 4.4.3, temos todos os elementos necessários para isso, ficando em aberto somente o cálculo do vetor A de transição dos estados ocultos.

O modelo oculto de Markov a ser utilizado será:

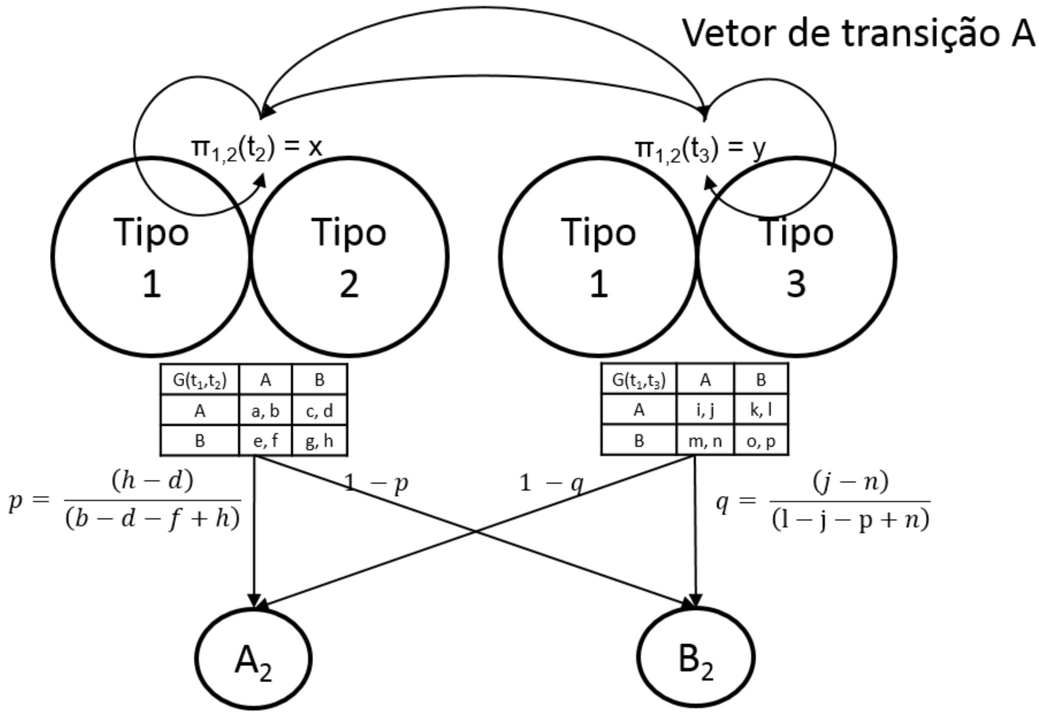

Figura 4 - Modelo oculto de Markov

4.4.4.2. Gerar observações

Uma vez que temos o modelo oculto de Markov, podemos iniciar o nosso jogo. A cada rodada, cada jogador escolhe uma estratégia, que é observada pelos jogadores. A estratégia do jogador informado será a estratégia no equilíbrio misto do jogo atual. A estratégia do jogador com falta de informação será descrita a seguir. Após escolher a estratégia, o tipo do jogador 2 pode mudar, de acordo com o vetor de transição A.

4.4.4.3. Atualizar modelo oculto de Markov

Utilizando o ferramental de modelo oculto de Markov, é possível através das observações obtidas a cada rodada atualizar o modelo de forma que o vetor A da cadeia oculta de Markov tenda a se aproximar do original (Rabiner & Juang, 1986).

Utilizamos o algoritmo Baum-Welch (Welch, 2003), (Baum & Eagon, 1967), para estimar o valor das transições dos estados ocultos. A ideia básica do algoritmo é tentar encontrar a melhor cadeia oculta que geraria a sequência de observações passadas como parâmetro.

Como a única variável que queremos estimar é o vetor A, a eficiência do algoritmo é maior do que a apresentada em (Waghabi, 2009), onde era calculado também o vetor B de transições. Iremos mostrar no exemplo que a cadeia gerada estimada se aproxima bastante da cadeia original.

4.4.4.4. Inferir o estado atual

Uma vez estimada a cadeia oculta do modelo oculto de Markov, precisamos inferir qual o estado mais provável que o jogador está. Utilizamos neste trabalho o algoritmo de Viterbi (Viterbi, 1967), que dada uma sequência de observações, infere qual a sequência de estados mais provável que ocorreu até o momento, e qual a probabilidade de cada estado ser o próximo.

4.4.4.5. Jogar contra o estado inferido

A heurística escolhida nesta solução para a escolha da estratégia do jogador com falta de informação é utilizar o tipo atual mais provável do oponente, inferido anteriormente, e jogar contra ele, ou seja, utilizar a estratégia no equilíbrio misto de Nash do jogo com a tabela do tipo inferido.

O princípio para a escolha dessa estratégia é o mesmo utilizado para o jogador informado, apresentado em 4.4.1. Na seção de resultados, é apresentada outra opção de escolha de estratégia, utilizando as probabilidades resultantes do algoritmo de

Viterbi sobre os tipos como ponderação para a escolha da estratégia, mas que apresentou resultados inferiores.

Com isso, repetimos a partir do passo 4.4.4.2 até o término do jogo.

4.5. Exemplo aplicado: testes

Iremos utilizar o exemplo do jogo de tênis apresentado em 4.2 para a aplicação do modelo proposto. Foram criados dois cenários diferentes, para representar comportamentos diversificados do jogador sacador. Teremos o cenário onde o jogador sacador é mais agressivo, ou seja, a cadeia oculta faz com que ele fiquei mais tempo no tipo agressivo. O outro cenário será defensivo, onde o jogador sacador fica mais tempo no tipo defensivo.

Para construção do modelo, precisamos dos valores das transições dos dois cenários explicados anteriormente, do valor π inicial de cada tipo do jogador sacador, dos valores da matriz de recompensa de cada jogo relacionado aos três tipos: agressivo (A), moderado (M) e defensivo (D); as observações são as duas estratégias aberto (Ab) e centro (Ce), e o vetor B será calculado em cada tabela.

As figuras abaixo mostram os dois jogos, um para cada cenário, com o cálculo do equilíbrio misto já realizado em cada jogo e aplicado no vetor B.

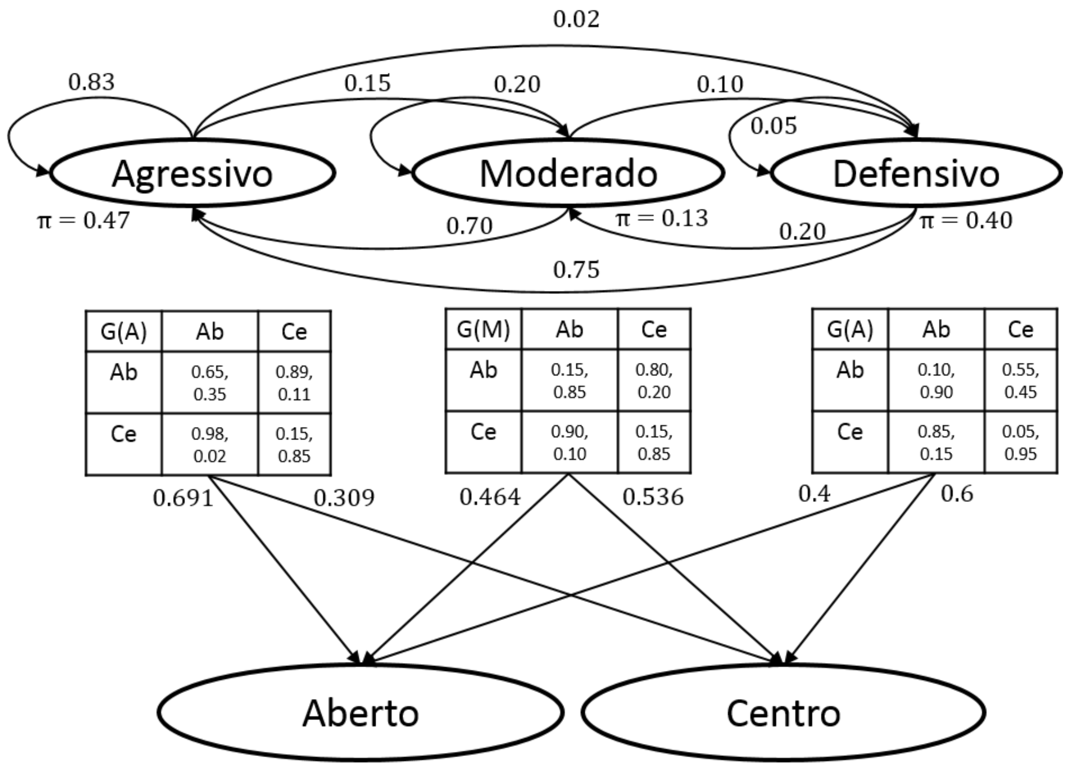

Figura 5 – Exemplo Jogo de Tênis: cenário agressivo

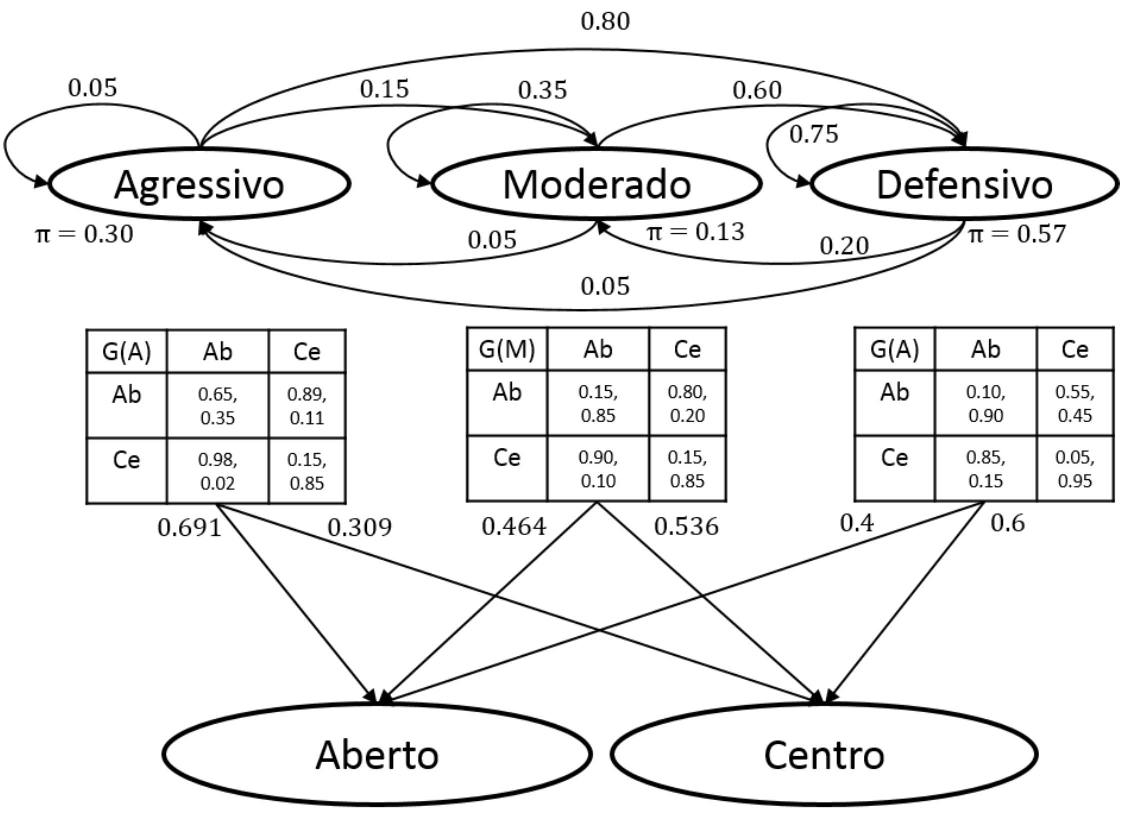

Figura 6 - Exemplo Jogo de Tênis: cenário defensivo

Os valores do vetor A das transições da cadeia oculta são utilizados na simulação do jogo, mas não são conhecidos pelo jogador 1.

Para a realização dos testes em cada cenário, simulamos as rodadas do jogo oculto de Markov, começando pelo passo apresentado em 4.4.4.1, gerando o modelo apresentado na Figura 5 e Figura 6. Um tipo inicial é sorteado para o jogador 2 (jogador informado) de acordo com π inicial.

Na visão do jogador 2, a simulação ocorre da seguinte forma:

- É escolhida uma estratégia de acordo com o equilíbrio misto do tipo atual
- O tipo atual muda de acordo com a cadeia oculta

Na visão do jogador 1 (não informado), ocorre da seguinte forma:

- É gerada uma cadeia oculta com valores iniciais arbitrários. No nosso caso, utilizamos uma divisão igual sempre que possível resultando em:
    - $\pi = {}^{1}/_{3}$ para cada tipo
    - $a = {}^{1}/_{3}$ para toda transição do vetor A
- É sorteado um tipo inicial aleatório do jogador 2 ponderado por π.
- É escolhida uma estratégia de acordo com o equilíbrio misto do tipo sorteado
- O estado observável é revelado
- A cadeia oculta é atualizada utilizando Baum-Welch
- O próximo tipo do jogador 2 é estimado utilizando Viterbi

Para a simulação desenvolvemos programa em Java baseado no apresentado em (Waghabi, 2009), o mesmo utilizado em (Benevides, Lima, Nader, & Rougemont, 2013) com uma estrutura para a geração das variáveis aleatórias de acordo com o nosso modelo. Neste programa, utilizamos o pacote JAHMM disponível em (François, 2006) para a implementação dos algoritmos de modelo oculto de Markov utilizados, Baum-Welch e Viterbi.

4.6. Resultados

Nesta seção iremos apresentar os resultados obtidos resultantes da simulação do exemplo descrito. Como apresentado na descrição do modelo, a solução é empírica e tem como objetivo apresentar uma opção de estratégia em problemas com pouca informação.

O teste realizado é composto de duas simulações, a primeira no cenário agressivo, a segunda no cenário defensivo. Foram executadas 200 rodadas com 100 observações em cada rodada, com o cálculo da recompensa média em cada rodada.

4.6.1. Resultados do cenário agressivo

No cenário agressivo, abaixo temos a cadeia de Markov estacionária original e a cadeia de Markov ao final da simulação. As imagens foram geradas primeiramente através da exportação das cadeias para um arquivo ".dot" através do framework JAHMM, e transformadas em imagem pela aplicação Graphviz (Gansner, Koutsofios, & North, 2010).

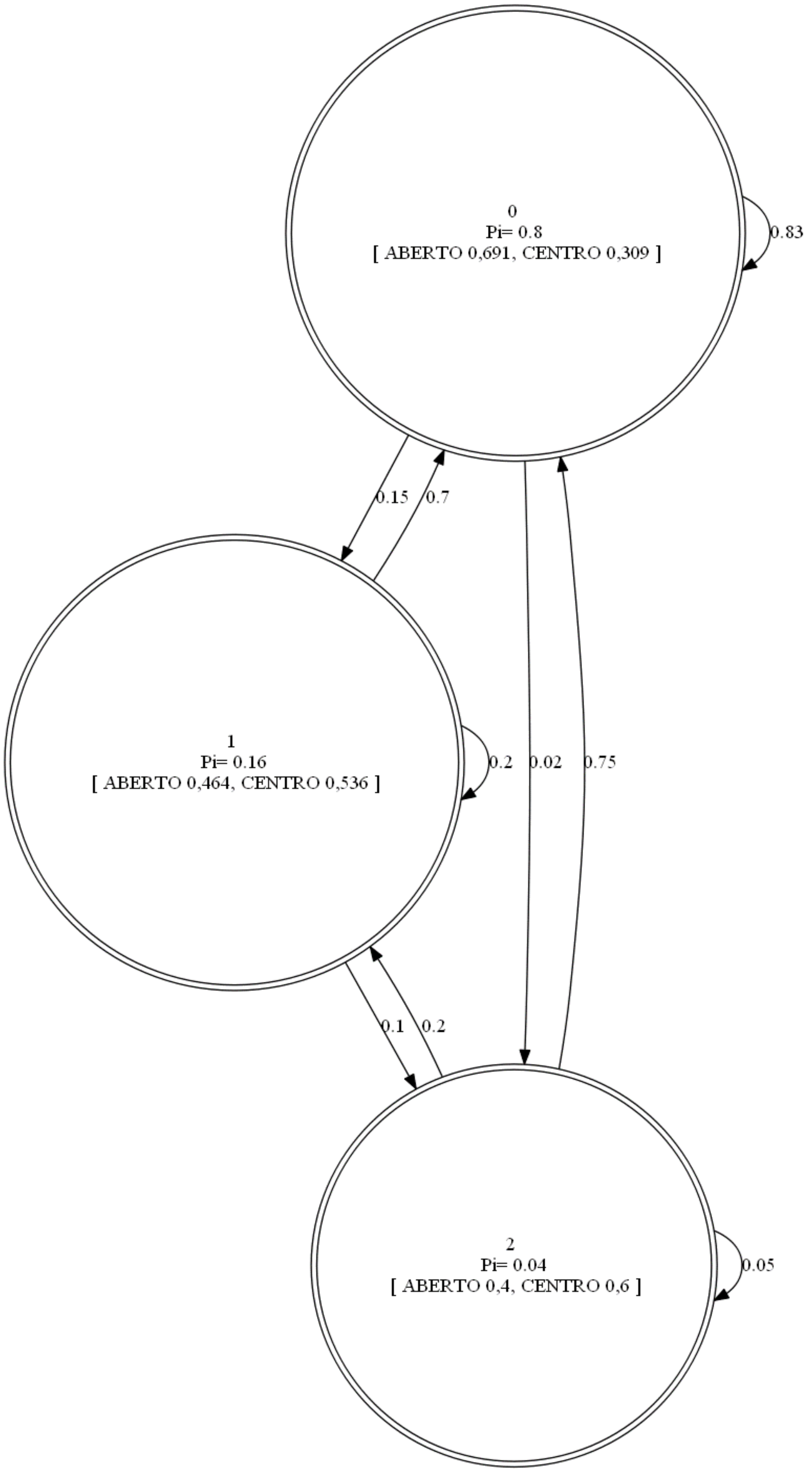

Figura 7 - Cadeia estacionária original - Agressivo

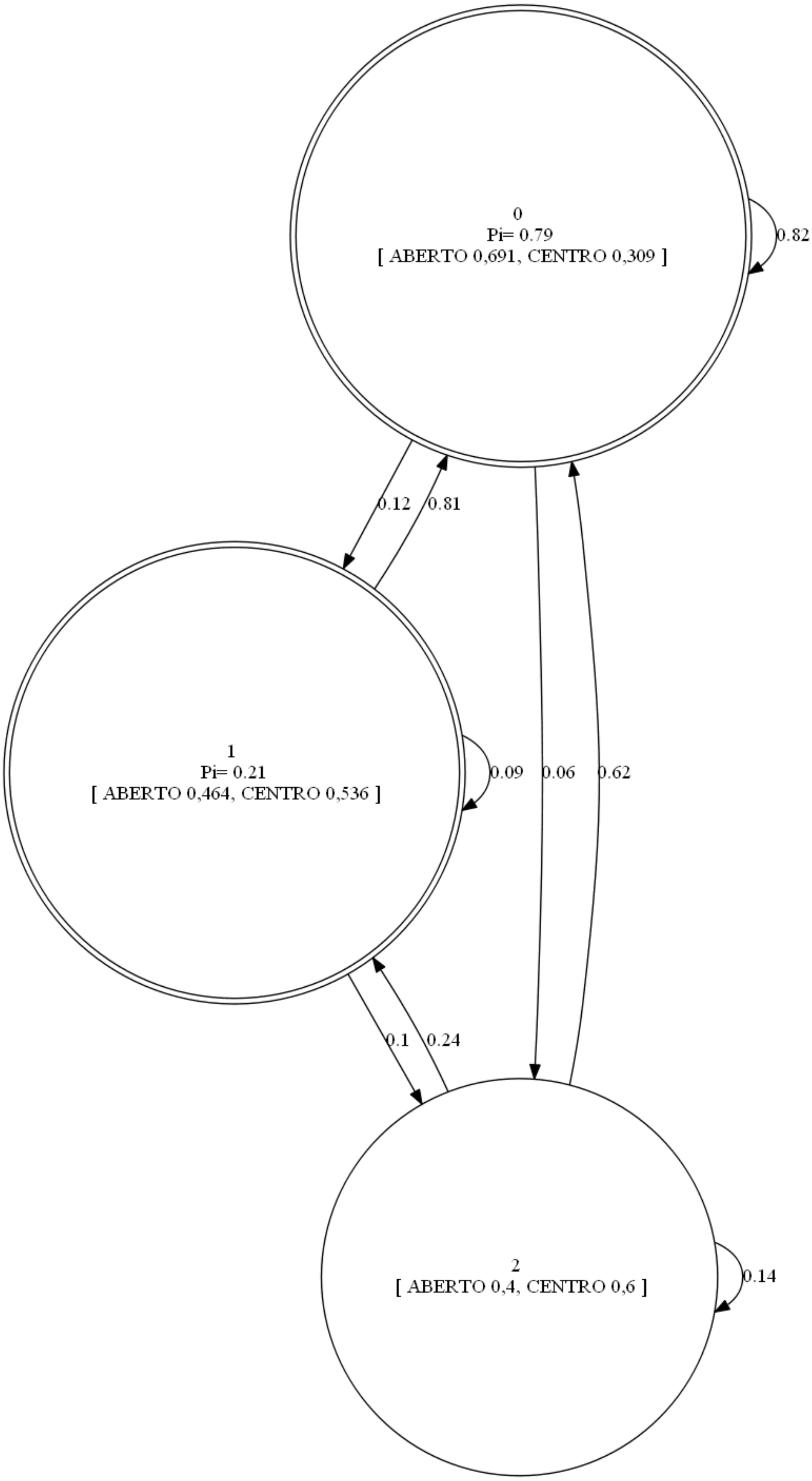

Figura 8 - Cadeia treinada – Agressivo

Na imagem, o estado 0 é o tipo Agressivo, o estado 1 é o tipo Moderado e o estado 2 é o tipo Defensivo. Podemos verificar pelas imagens que as duas cadeias se aproximam. O gráfico abaixo apresenta a distância entre os dois modelos ocultos de Markov ao longo das rodadas, utilizando o algoritmo Kullback-Leibler implementado no framework JAHMM.

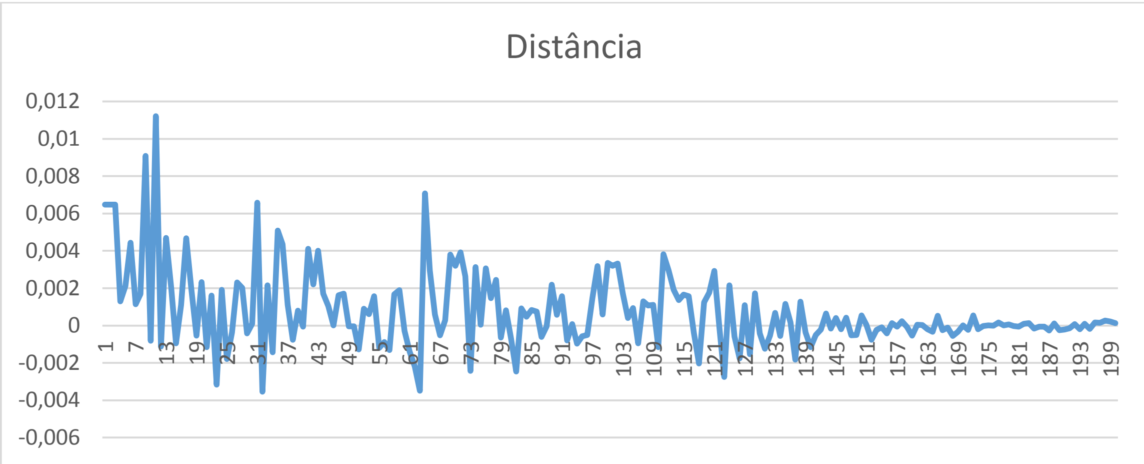

Figura 9 - Distância entre as duas cadeias - Agressivo

Os resultados obtidos utilizando a solução proposta em 4.4.4.5, chamada aqui de solução Viterbi, é comparada com a solução onde o estado atual do adversário é conhecido, e é jogado no equilibro misto dele, com a intenção de apresentar a eficiência do algoritmo de Viterbi em inferir o estado atual. Essa segunda solução foi chamada de “Real” nos gráficos a seguir. Como comparativo, apresentamos também o resultado caso a estratégia do jogador fosse escolhida de forma aleatória.

Abaixo um gráfico com a média das três soluções.

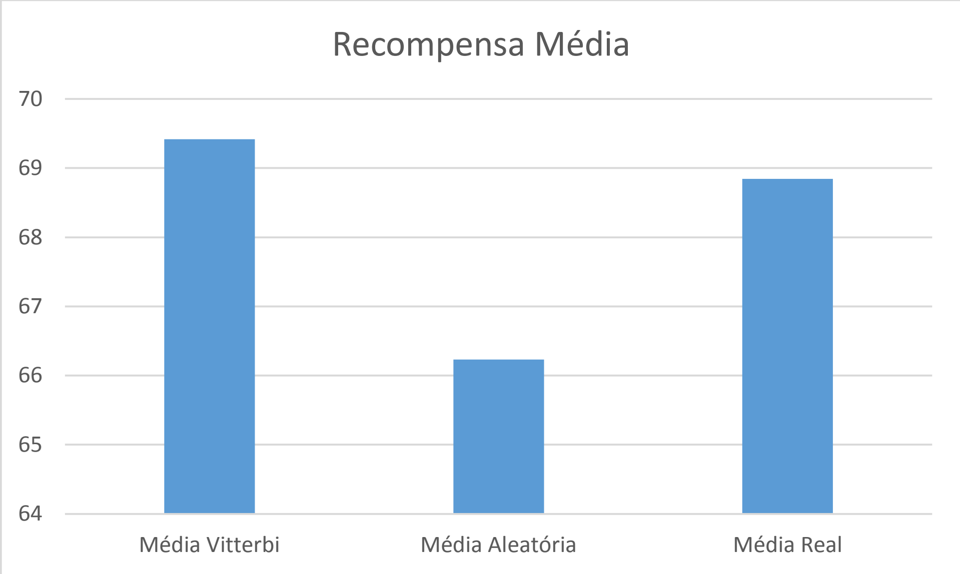

Figura 10 - Recompensa média - Agressivo

Podemos verificar que a recompensa média obtida utilizando a solução proposta é praticamente igual a recompensa média utilizando o estado real do jogador adversário, indicando que a inferência do estado atual é satisfatória. Embora a diferença absoluta observada seja pequena em relação à média aleatória, isso depende da natureza do problema e dos valores apresentados na tabela de recompensa.

Abaixo o gráfico com média ao longo das rodadas das duas soluções.

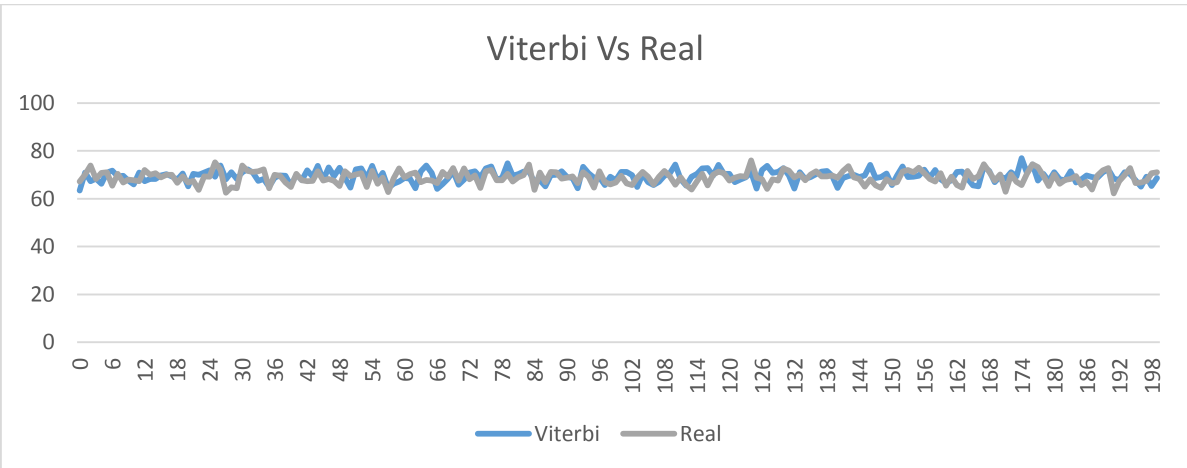

Figura 11 - Recompensa ao longo das rodadas – Agressivo

A média das recompensas em cada rodada varia dentro de um intervalo pequeno, devido à natureza do problema, que apresenta valor mínimo e valor máximo de recompensa muito próximos.

4.6.2. Resultados do cenário defensivo

Os mesmo testes foram realizados no cenário defensivo, e os resultados obtidos foram semelhantes, conforme gráficos abaixo.

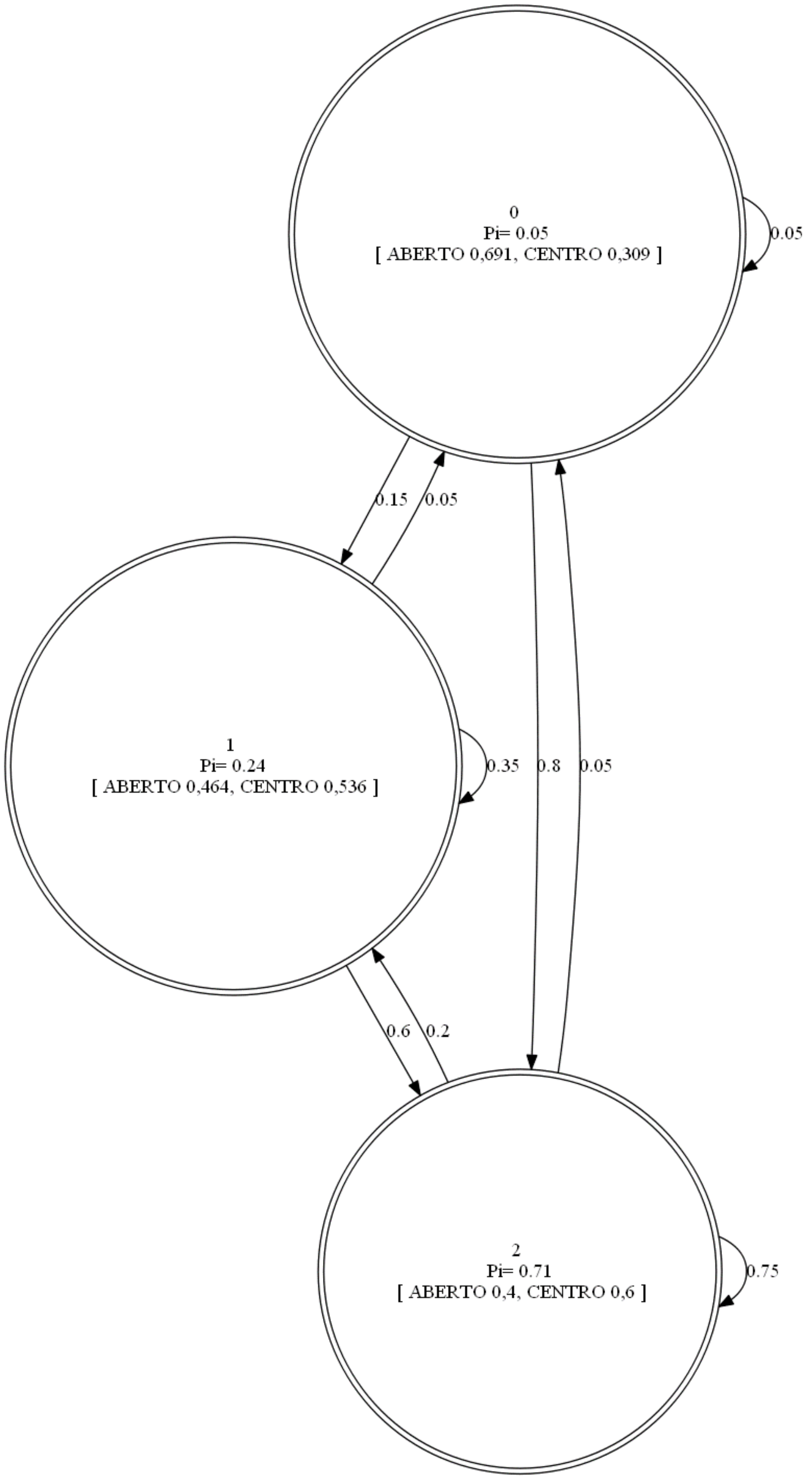

Figura 12 - Cadeia estacionária original - Defensivo

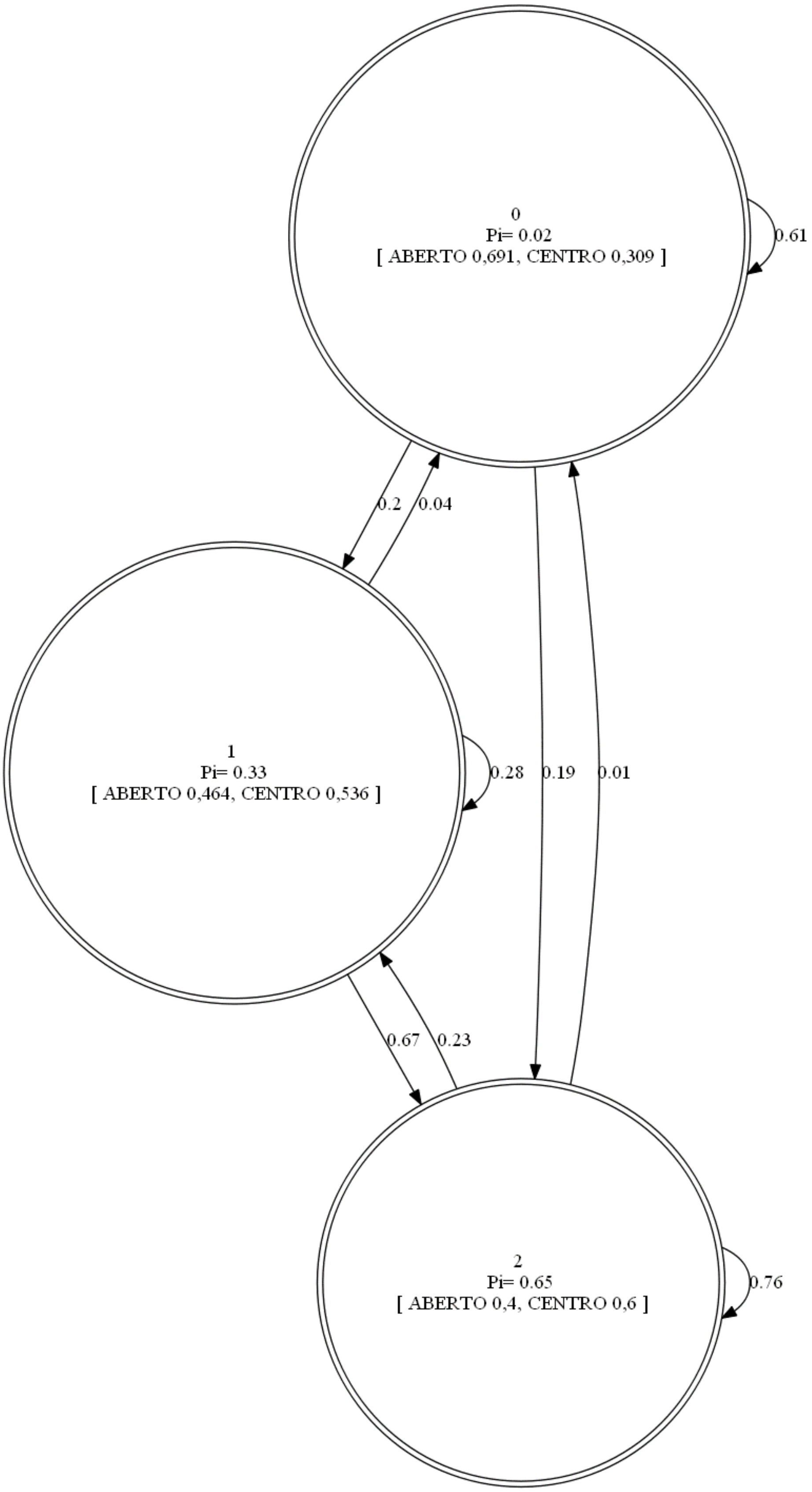

Figura 13 - Cadeia treinada – Defensivo

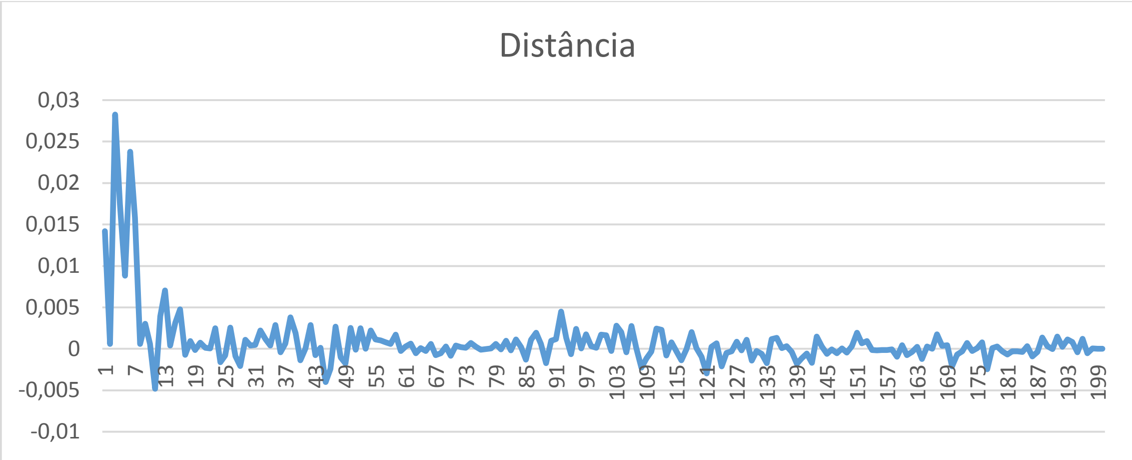

Figura 14 - Distância entre as duas cadeias – Defensivo

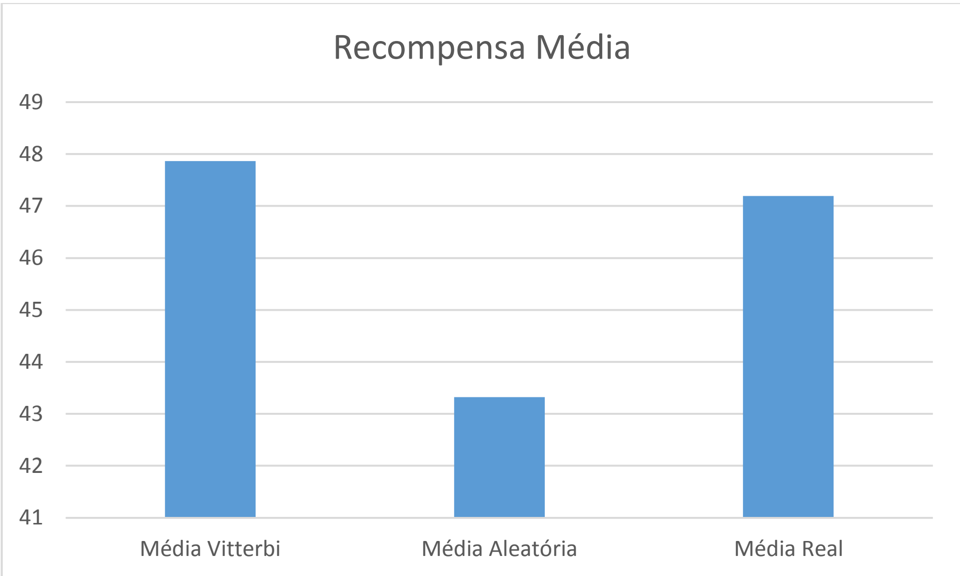

Figura 15 - Recompensa média – Defensivo

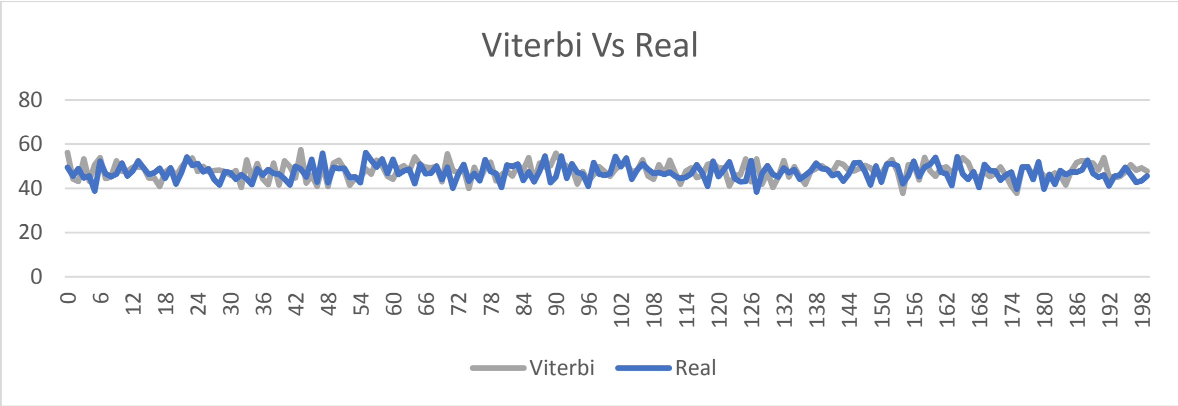

Figura 16 - Recompensa ao longo das rodadas - Defensivo

Capítulo 5

Conclusão e Trabalhos Futuros

Nessa dissertação apresentamos uma classe nova de problemas em teoria dos jogos e propomos um modelo de jogo para este problema, chamado de jogo oculto de Markov. O problema apresentado se assemelha com outros problemas conhecidos da área, mas com uma característica que não permite a utilização de soluções já existentes.

Assim como alguns outros problemas de jogos repetidos, o jogo proposto não apresenta uma prova de existência de um valor nem de estratégias ótimas, o que é proposto nesse trabalho é uma solução heurística baseada na natureza do problema.

Percebemos uma correlação entre os valores do equilíbrio de cada tipo com a performance da solução proposta. Quanto menos definido cada tipo, menos informação conseguimos obter com as observações, e mais próximo de uma abordagem randômica ficará nossa solução. Isso não representa um problema exatamente, mas indica uma classe de problemas onde a nossa solução alcança melhores resultados.

Como questionado em (Simon, 1955), (Arrow, 1986) e (Rubinstein, 1998), o princípio da escolha racional pode não ser a melhor representação da escolha de estratégias por parte do jogador. Utilizamos isso para corroborar a escolha de algumas hipóteses e restrições no desenvolvimento do trabalho.

Os resultados obtidos, conforme apresentado na seção anterior, indicam que a modelagem proposta apresenta bons resultados na inferência das informações ausentes no problema. O treinamento do modelo oculto e sua proximidade com a cadeia original resultaram em uma boa inferência do estado atual pelo algoritmo de Viterbi.

É importante observar que os resultados apresentados estão diretamente relacionados com a natureza do exemplo utilizado. Tanto as cadeias de Markov originais quanto as tabelas de recompensa de cada jogo em cada tipo interferem nos resultados de outras estratégias. O objetivo deste trabalho era apresentar uma solução que garantisse o valor do equilíbrio misto utilizando o estado inferido.

Para trabalhos futuros podemos analisar os tipos de jogos em cada estado em que a solução apresenta melhor resultado, mostrando a relação entre o vetor de probabilidade B e a eficiência da solução. Podemos analisar também que tipo de cadeia oculta apresenta melhor capacidade de treinamento utilizando os algoritmos de modelo oculto de Markov. Ficou em aberto também a verificação da existência de um valor no modelo de jogo proposto, e caso exista esse valor, qual estratégia ótima de cada jogador para atingir esse valor.

Bibliografia

Arrow, K. J. (Outubro de 1986). Rationality of Self and Others in an Economic System. *The Journal of Business*, pp. S385-S399.

Aumann, R. J., & Maschler, M. (1995). *Repeated games with incomplete information.* Cambridge, MA: MIT Press.

Baum, L. E., & Eagon, J. A. (1967). An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. *Bulletin of American Mathematical Society*, pp. 360-363.

Benevides, M., Lima, I., Nader, R., & Rougemont, P. (2013). Using HMM in Strategic Games. *International Workshop on Developments in Computational Models*.

Cournot, A. A. (1838). *Researches Into the Mathematical Principles of the Theory of Wealth.*

François, J.-M. (Abril de 2006). Fonte: Jahmm - Hidden Markov Model (HMM): http://www.run.montefiore.ulg.ac.be/~francois/software/jahmm/

Gansner, E. R., Koutsofios, E., & North, S. (Novembro de 2010). Drawing graphs with dot.

Gibbons, R. (1992). *Game theory for applied economists.* Princeton: Princeton University Press.

Nash, J. F. (1950). Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences of the United States of America*, pp. 48-49.

Osborne, M. J. (2004). *An Introduction to Game Theory.* Oxford: Oxford University Press.

Rabiner, L. R., & Juang, B. H. (1986). An Introduction to Hidden Markov Models. *IEEE ASSP MAg, 3*, 4-16.

Renault, J. (2006). The Value of Markov Chain Game With Lack of Information on One Side. *Mathematics of Operations Research, 31*, 490-512.

Renault, J. (2012). Repeated Games With Incomplete Information. Em R. E. Meyers, *Computational Complexity* (pp. 2635-2655). Springer.

Rubinstein, A. (1998). *Modeling Bounded Rationality.* Cambridge: The MIT Press.

Schelling, T. C. (1960). *The strategy of conflict.* Cambridge, MA: Harvard University.

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics*(69), pp. 99–118.

Viterbi, A. J. (Abril de 1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory, XIII*, pp. 260-269.

von Neumann, J., & Morgenstern, O. (1944). *Theoy of Games and Economic Behavior.* Princeton, NJ: Princeton University Press.

Waghabi, E. R. (2009). Aplicação de Modelos Ocultos de Markov em Jogos de Estratégia Mista.

Welch, R. L. (2003). Hidden Markov Models and the Baum-Welch Algorithm. *IEEE Information Theory Society Newsletter, 53*, 1, 10-13.