
Uma Abordagem apoiada em Portfólio de Planejadores Probabilísticos aplicada ao Mercado Financeiro

Augusto Afonso Borges Branquinho



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Augusto Afonso Borges Branquinho

**Uma Abordagem apoiada em Portfólio de
Planejadores Probabilísticos aplicada ao
Mercado Financeiro**

Tese de doutorado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Doutor em Ciência da Computação.

Área de concentração: Inteligência Artificial

Orientador: Carlos Roberto Lopes

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

B821a Branquinho, Augusto Afonso Borges, 1984-
2017 Uma abordagem apoiada em portfólio de planejadores
 probabilísticos aplicada ao mercado financeiro / Augusto Afonso Borges
 Branquinho. - 2017.
 163 f. : il.

 Orientador: Carlos Roberto Lopes.
 Tese (Doutorado) - Universidade Federal de Uberlândia, Programa
de Pós-Graduação em Ciência da Computação.
 Disponível em: <http://dx.doi.org/10.14393/ufu.te.2018.765>
 Inclui bibliografia.

 1. Computação - Teses. 2. Mercado financeiro - Planejamento -
Teses. 3. Aprendizado do computador - Teses. I. Lopes, Carlos Roberto,
1962- II. Universidade Federal de Uberlândia. Programa de Pós-
Graduação em Ciência da Computação. III. Título.

CDU: 681.3

Maria Salete de Freitas Pinheiro – CRB6/1262

UNIVERSIDADE FEDERAL DE UBERLÂNDIA FACULDADE DE
COMPUTAÇÃO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO

Os abaixo assinados, por meio deste, certificam que leram e recomendam para a Faculdade de Computação a aceitação do trabalho intitulado "**Uma Abordagem apoiada em Portfólio de Planejadores Probabilísticos aplicada ao Mercado Financeiro**" por **Augusto Afonso Borges Branquinho** como parte dos requisitos exigidos para a obtenção do título de **Doutor em Ciência da Computação**.

Uberlândia, 06 de novembro de 2017

Orientador: _____
Prof. Dr. Carlos Roberto Lopes
Universidade Federal de Uberlândia

Banca Examinadora:

Prof^a. Dr^a. Leliane Nunes de Barros
Universidade de São Paulo

Prof^a. Dr^a. Karina Valdivia Delgado
Universidade de São Paulo

Prof. Dr. Marcelo Keese Albertini
Universidade Federal de Uberlândia

Prof^a. Dr^a. Márcia Aparecida Fernandes
Universidade Federal de Uberlândia

Dedico este trabalho a todos os professores que têm amor pela à educação.

Agradecimentos

À minha esposa Rachell Guedes por me acompanhar durante toda a fase do doutorado.

Aos meus pais Jair Branquinho e Terezinha Carolino por terem sempre incentivado a busca da educação.

Ao meu professor e orientador Carlos Roberto Lopes que sempre acreditou no meu trabalho e esteve sempre a disposição para ajudar neste trabalho.

Ao corpo docente e aos secretários do Programa de Pós-Graduação em Computação da Universidade Federal de Uberlândia. Em especial a Professora Doutora Márcia Aparecida Fernandes e o secretário Erisvaldo Araújo Filho, por serem exemplos de profissionais e terem me ajudado no decorrer do trabalho.

*Se a educação sozinha não transforma a sociedade,
sem ela tampouco a sociedade muda.
(Paulo Freire)*

Resumo

No contexto de planejamento, um dos problemas encontrados é a escolha do planejador mais adequado para um dado problema. Além disso, para certos problemas podem ser usadas combinações de planejadores. Neste trabalho descrevemos uma abordagem que busca, a partir da especificação do problema de planejamento, um planejador que seja adequado para se resolver o problema. Para isto utiliza-se um portfólio de planejadores probabilísticos para problemas especificados como POMDPs (Processos de Decisão Markovianos Parcialmente Observáveis) e técnicas de aprendizado de máquina. Trata-se de uma abordagem de propósito geral, independente do tipo de problema de planejamento. Contudo, para fins de avaliação, a abordagem é aplicada na seleção do planejador mais adequado para o problema de planejamento do mercado financeiro.

O mercado de financeiro pode ser considerado um domínio não-determinístico e parcialmente observável, uma vez que investidores nunca sabem todas as informações que afetam os preços e os resultados de um investimento quase sempre são incertos. Sendo assim, além do portfólio de planejamento, o trabalho propõe diferentes abordagens para o uso de planejamento probabilístico na tarefa de negociação de ativos no mercado financeiro. É apresentada uma nova modelagem para o problema do mercado usando MOMDP (Processos de Decisão Markovianos Observáveis Mistos), o uso de múltiplos ativos, o custo de corretagem e a montagem de carteiras de investimento.

Para o mercado financeiro podem ser usados planejamento *offline* e *online*, sendo que a escolha do tipo de planejamento pode variar segunda a forma de negociação. Em função dos diferentes planejadores existentes, o trabalho explora o uso de ambos os tipos de planejamento, mas em particular planejamento *online* na criação do portfólio.

Palavras-chave: Portfólio de Planejamento; Planejamento Probabilístico; Mercado Financeiro; Análise Técnica.

Abstract

In the context of planning, one of the recurrent problems is choosing the most suitable planner for a given problem. Besides that, for certain problems, several combinations of planners can be used. In this work, we describe an approach that seeks, from the specification of the planning problem, a planner that is suitable to solve the problem. In order to do so, a portfolio of probabilistic planners is used for problems specified as POMDPs (Partially Observable Markov Decision Process) and machine learning techniques, which is a general purpose approach, regardless of the type of planning problem. However, for evaluation purposes, the approach is applied in the selection of the most suitable planner for the financial market planning problem.

The stock market can be considered a domain non-deterministic and partially observable as investors never know all the information that affects prices and the results of an investment are almost always uncertain. Thus, this paper proposes the use of probabilistic planning in order to assist the active of trading in the financial market. It is proposed a new model to the market problem using MOMDP (Mixed Observability Markov Decision Process). Other points are treated, such as the use of multiple assets, the cost of brokerage and mounting investment portfolios.

For the financial market, offline and online planning can both be used, and the choice of planning type can vary according to the form of negotiation. Depending on the different existing planners, the present work explores the use of both types of planning, but in particular online planning in portfolio creation.

Keywords: Planner Portfolio; Probabilistic Planning; Stock Market; Technical Analysis.

Lista de ilustrações

Figura 1 – Preço de fechamento diário e momentos de lucro e prejuízo para o ativo MSFT - Microsoft (2013).	34
Figura 2 – Representação de dois <i>candles</i> , sendo respectivamente, um de baixa e outro de alta.	42
Figura 3 – Gráfico de <i>candles</i> diários para o ativo PETR4.	43
Figura 4 – Padrões engolfo de alta e engolfo de baixa e as respectivas tendências. .	43
Figura 5 – Padrões bebê abandonado de alta e bebê abandonado de baixa e as respectivas tendências.	44
Figura 6 – Exemplos de média móvel simples (SMA) e média móvel exponencial (EMA) para o ativo PETR4.	45
Figura 7 – MACD para o ativo PETR4.	47
Figura 8 – Bandas de Bollinger para o ativo PETR4.	48
Figura 9 – Oscilador de preço absoluto para o ativo PETR4 com 12 e 26 períodos. .	49
Figura 10 – <i>RSI</i> para o ativo PETR4.	50
Figura 11 – Média Móvel Exponencial Dupla para o ativo PETR4 usando 20 e 50 períodos.	51
Figura 12 – Preços de fechamento normalizados dos ativos AELP3, BIOM3, PETR3 e IDVL4 para o período de 02/01/2012 até 30/12/2013.	55
Figura 13 – Fluxo de execução dos AGs.	56
Figura 14 – “ <i>Mundo dos Blocos</i> ” representado em STRIPS.	61
Figura 15 – Estruturas utilizados pela linguagem PDDL.	61
Figura 16 – Representação do problema “Mundo dos Blocos” na linguagem PDDL. .	62
Figura 17 – Histórico da Competição Internacional de Planejamento (IPC) (COLES et al., 2012).	63
Figura 18 – Estrutura do domínio de problemas representados com RDDDL.	64
Figura 19 – Estrutura <i>non-fluent</i> de problemas representados com RDDDL.	65
Figura 20 – Estrutura <i>instance</i> de instâncias de problemas representados com RDDDL. .	65
Figura 21 – Árvore de estados de crença para problemas POMDP.	68

Figura 22 – B é o espaço de crenças, $R(b_0)$ é o espaço alcançável e $R^*(b_0)$ corresponde ao espaço ótimo alcançável. É importante observar que $R^*(b_0) \subseteq R(b_0) \subseteq B$ (KURNIAWATI; HSU; LEE, 2008).	74
Figura 23 – Comparação entre os planejamentos <i>offline</i> e <i>online</i> .	75
Figura 24 – Transição de estados para POMDP (esquerda) e MOMDP (direita). O estado s do MOMDP é decomposto em duas partes: $s = (x, y)$, onde x é completamente observável e y parcialmente observável. t representa o tempo/momento de cada termo, a é a ação, o é a observação e r é a recompensa.	86
Figura 25 – Árvore de estados de crença para problemas MOMDP.	87
Figura 26 – Árvore de estados e transições para o problema do mercado financeiro usando MOMDP.	87
Figura 27 – Separação da parte determinística da modelagem MOMDP.	88
Figura 28 – Modelagem considerando a variável determinística na definição das ações.	89
Figura 29 – Modelagem MOMDP para o mercado financeiro.	90
Figura 30 – Sequência de passos da metodologia usada para a montagem do portfólio.	96
Figura 31 – Sequência de passos da metodologia usada na pesquisa.	98
Figura 32 – Preço e volume do ativo PETR4 durante um período de desdobramento. São mostrados os dados não normalizados e normalizados.	100
Figura 33 – Aplicação responsável pelo processamento de dados e análises técnicas.	104
Figura 34 – Aplicação responsável por realizar a carga e correção dos dados.	105
Figura 35 – Aplicação responsável por realizar as tarefas de planejamento para o mercado financeiro.	106
Figura 36 – Elementos que compõem o ambiente.	108
Figura 37 – Sistema projetado para processar os AGs em paralelo e distribuído.	110
Figura 38 – Experimentos por período e distribuição.	114
Figura 39 – Negócios realizados para PETR4, usando MACD 1, durante o ano de 2014, com o planejador SARSOP e modelagem POMDP.	120
Figura 40 – Melhores rendimentos acumulados para cada ativo nos períodos de 2008 a 2015 e 2011 a 2015 com a cobrança da taxa 2.	123
Figura 41 – Comparação dos melhores rendimentos acumulados para cada ativo no período de 2008 a 2015 usando a taxa 1.	124
Figura 42 – Resumo geral dos rendimentos obtidos utilizando a taxa 2 e todos os planejadores e modelagens de planejamento.	126

Lista de tabelas

Tabela 2 – Exemplos de valores da SMA_7 e EMA_7 para o preço de fechamento do período de 18/07/2014 a 30/07/2014.	46
Tabela 3 – Matriz de correlação dos ativos AELP3, BIOM3, PETR3 e IDVL4 para o período de 02/01/2012 a 30/12/2013.	54
Tabela 4 – Estados na modelagem POMDP para o índice IBOVESPA.	78
Tabela 5 – Alterações da posição do investidor de acordo com as ações/operações de compra e venda.	79
Tabela 6 – Transições de estados para o dia 02/02/2000 da Tabela 4.	79
Tabela 7 – Regras de geração das observações.	80
Tabela 8 – Observações para o índice IBOVESPA utilizando a regra RSI_2 com 14 períodos.	80
Tabela 9 – Recompensa de cada estado.	81
Tabela 10 – Exemplos de estados para PETR4 e VALE5 ao usar múltiplos ativos.	82
Tabela 11 – Ações para PETR4 e VALE5 ao usar múltiplos ativos.	83
Tabela 12 – Observações para PETR4 e VALE5 ao usar múltiplos ativos.	83
Tabela 13 – Exemplos de recompensas de estados formados com os ativos PETR4 e VALE5.	84
Tabela 14 – Distribuição dos domínios por período usando a estratégia <i>Leave One Domain Out</i>	97
Tabela 15 – Proventos em ações da empresa Petrobras.	98
Tabela 16 – Preço e volume do ativo PETR4 durante um período de desdobramento. São mostrados os dados não normalizados e normalizados.	99
Tabela 17 – Lista de análises técnicas.	101
Tabela 18 – Exemplo de taxa de corretagem cobrada sobre cada ordem de compra ou venda.	112
Tabela 19 – Períodos e distribuições dos experimentos A , B , C , D e E	114
Tabela 20 – Atributos dos experimentos A , B , C e D	114

Tabela 21 – Planejadores selecionados através do Portfólio de Planejamento conforme os experimentos.	115
Tabela 22 – Percentual de seleção dos planejadores.	116
Tabela 23 – Tipos de taxas de corretagem.	117
Tabela 24 – Correlação de ativos para cada período de treino.	117
Tabela 25 – Comparação dos resultados obtidos com o SARSOP, modelagem POMDP e taxas 1 e 2 para o ativo CMIG4.	118
Tabela 26 – Resultados para CRUZ3 e ITSA4 com capital inicial de R\$ 25,000.00, utilizando a taxa número 2, o planejador SARSOP e os períodos a partir de 2011.	119
Tabela 27 – Resultados para PETR4, usando o SARSOP, modelagem POMDP, capital inicial de R\$ 25,000.00 e a taxa 2.	119
Tabela 28 – Resultados das simulações com múltiplos ativos para CMIG4.	121
Tabela 29 – Resultados das simulações para CRUZ3, múltiplos ativos e taxa 2.	122
Tabela 30 – Resultados das simulações usando MOMDP para CMIG4.	124
Tabela 31 – Resultados usando MOMDP para CRUZ3.	125

Lista de siglas

ADL	Accumulation Distribution Line
AG	Algoritmo Genético
AMA	Adaptive Moving Average
APO	Absolute Price Oscillator
Chaikin A/D Line	Chaikin Accumulation Distribution Line
EMA	Exponential Moving Average
GP	Genetic Programming
IA	Inteligência Artificial
IPC	International Planning Competition
IPPC	International Probabilistic Planning Competition
KAMA	Kaufman Adaptive Moving Average
LDA	Linear Discriminant Analysis
MA	Moving Average
MACD	Moving Average Convergence/Divergence
MDP	Markov Decision Process
MFM	Money Flow Multiplier
MFV	Money Flow Volume
MLP	Multilayer Perceptron
MOMDP	Mixed Observable Markov Decision Process
OST	Stochastic Oscillator
PBVI	Point-based Value Iteration
POMDP	Partially Observable Markov Decision Processes
QDA	Quadratic Discriminant Analysis
RSI	Relative Strength Index
SARSOP	Successive Approximations of the Reachable Space under Optimal Policies
SMA	Simple Moving Average
Stochastic RSI	Stochastic Relative Strength Index

SVM

Support Vector Machine

UFU

Universidade Federal de Uberlândia

Lista de algoritmos

1	PBVI($B_{Inicial}, \Gamma_0, N, T$)	72
2	BACKUP(B, Γ_{t-1})	72
3	PBVI_EXPANDIR $_{ra}(B, \Gamma)$	73

Sumário

1	INTRODUÇÃO	27
1.1	Objetivos e Desafios da Pesquisa	29
1.2	Contribuições	30
1.3	Organização do Trabalho	30
2	CARACTERIZAÇÃO DO PROBLEMA E TRABALHOS COR- RELATOS	33
2.1	Caracterização do Problema	34
2.2	Trabalhos Correlatos	35
3	FUNDAMENTAÇÃO TEÓRICA	41
3.1	Análise Técnica	41
3.1.1	Candlesticks	42
3.1.2	Médias Móveis	44
3.1.3	Convergência/Divergência das Médias Móveis (MACD)	46
3.1.4	Oscilador Estocástico	46
3.1.5	Bandas de Bollinger	47
3.1.6	Oscilador de Preço Absoluto	48
3.1.7	Índice de Força Relativa	48
3.1.8	Índice de Força Relativa Estocástica	50
3.1.9	Média Móvel Exponencial Dupla	50
3.1.10	Média Móvel Adaptativa	51
3.1.11	Linha de Acumulação Distribuição	52
3.1.12	Oscilador de Chaikin	53
3.1.13	Linha Acumulação/Distribuição de Chaikin	53
3.2	Correlação de Ativos	53
3.3	Algoritmos Genéticos (AG)	55

4	PLANEJAMENTO	59
4.1	Planejamento Clássico	59
4.2	Planejamento Probabilístico	60
4.2.1	Relational Dynamic Influence Diagram Language (RDDL)	62
4.2.2	Processo de Decisão de Markov	65
4.2.3	Processo de Decisão de Markov Parcialmente Observável	67
4.2.4	Algoritmos Exatos	70
4.2.5	Algoritmos Aproximados	71
4.2.6	Planejamento Online com POMDPs	74
5	PLANEJAMENTO PROBABILÍSTICO APLICADO NO MER- CADO FINANCEIRO	77
5.1	Modelagem Observável Parcialmente	77
5.1.1	Modelagem para Múltiplos Ativos	82
5.2	Modelagem Observável Mista	84
5.3	Portfólio de Planejamento	88
5.3.1	Características dos Problemas	91
5.3.2	Planejadores	91
5.3.3	Problemas de Planejamento	92
5.3.4	Classificadores e Seleção de Características	93
6	METODOLOGIA	95
6.1	Metodologia do Portfólio	95
6.2	Metodologia do Planejamento e Negociação	97
7	ARQUITETURA DA SOLUÇÃO	103
7.1	Arquitetura	103
7.2	Sistema Desenvolvido	107
8	EXPERIMENTOS E DISCUSSÃO DOS RESULTADOS . . .	113
8.1	Portfólio de Planejamento	113
8.2	Planejamento	116
9	CONCLUSÃO	127
9.1	Trabalhos Futuros	128
	REFERÊNCIAS	131

APÊNDICES 137

APÊNDICE A	—	EXEMPLO DE MODELAGEM - CASSANDRA	139
-------------------	----------	---	------------

APÊNDICE B	–	EXEMPLO DE MODELAGEM POMDPX . . .	145
APÊNDICE C	–	EXEMPLO DE MODELAGEM RDDDL	151

CAPÍTULO 1

Introdução

No início, o mercado financeiro realizava suas negociações pessoalmente entre corretores. Com a digitalização dos mercados financeiros foram criados sistemas que automatizam as negociações. A partir da virtualização das negociações foram criados robôs virtuais que conseguem atuar de forma autônoma no mercado. Estes robôs utilizam estratégias para analisar o mercado e negociar ativos. O problema neste tipo de negociação consiste em obter lucro através da predição de preços futuros. Sendo assim, existem os desafios de selecionar as técnicas usadas no processo de predição e o de escolher quais ativos negociar.

Técnicas de inteligência artificial (IA) podem ser aplicadas para a negociação de ações do mercado financeiro. Neste caso, o objetivo consiste em buscar padrões de tendências de modo que seja possível prever com certa segurança as variações futuras dos preços. Em tal contexto, uma possibilidade constitui-se no uso de planejamento probabilístico para tratar da automação das negociações de ativos. Uma proposta nesta área é realizada por Baffa (BAFFA; CIARLINI, 2010b). Baffa propõe o uso de processos de decisão de Markov parcialmente observáveis (denotado pelo acrônimo POMDP, da expressão em inglês *Partially Observable Markov Decision Process*) para planejar ações de compra e venda de ativos na bolsa de valores BM&FBOVESPA. Além de (BAFFA; CIARLINI, 2010b), (BRANQUINHO; LOPES; BAFFA, 2016) também explora este tipo de planejamento, contudo, utilizando múltiplos ativos.

O planejamento probabilístico para o mercado financeiro consiste em modelar os estados do mercado a partir das tendências passadas dos preços, tendências futuras dos preços e da posição do investidor. As observações são indicações de pontos em que devem ou não serem realizados negócios e são geradas através de análise técnica. A partir dos dados históricos são calculadas as probabilidades das transições dos estados e das observações. Uma vez realizada a construção do modelo são usados algoritmos POMDP (CASSANDRA; KAEHLING; LITTMAN, 1994) ou processos de decisão de Markov observáveis mistos (denotado pelo acrônimo MOMDP, da expressão em inglês *Mixed Observable Markov Decision Process*) (ONG et al., 2010) para gerar políticas. Tais políticas são usadas

para indicar quais ações devem ser executadas em futuras negociações.

No trabalho de Baffa (BAFFA; CIARLINI, 2010a) foram apresentados diferentes algoritmos de planejamento, contudo, nos experimentos foi utilizado apenas o *Point-based Value Iteration* (PBVI) (PINEAU; GORDON; THRUN, 2003). Já em (BRANQUINHO; LOPES; BAFFA, 2016) foi utilizado o SARSOP (dado pela expressão em inglês *Successive Approximations of the Reachable Space under Optimal Policies*) (KURNIAWATI; HSU; LEE, 2008). Apesar dos bons resultados mostrados em (BAFFA; CIARLINI, 2010a) e (BRANQUINHO; LOPES; BAFFA, 2016), não foram apresentadas justificativas das escolhas, respectivamente, do PBVI (PINEAU; GORDON; THRUN, 2003) e do SARSOP (KURNIAWATI; HSU; LEE, 2008). No contexto de planejamento, existem muitos planejadores e cada um se adapta melhor a certos tipos de problemas (ROBERTS; HOWE, 2009). Desta forma, a escolha do planejador é algo muito importante que impacta nos resultados e precisa ser investigada. Para explorar a escolha do planejador é possível o uso de portfólio de planejadores (SOUSA, 2014). O portfólio de planejamento é uma contribuição ainda não explorada para planejamento não clássico, logo, esta é uma contribuição almejada no trabalho. A proposta fundamenta-se em uma abordagem de escopo geral para a formulação do portfólio de planejamento, posteriormente, é explorado o portfólio para o problema do mercado financeiro.

Normalmente a descrição dos problemas de planejamento é realizada usando uma linguagem de especificação. Uma vez feita a descrição é possível extrair características do problema. Com base nas características do problema entende-se que é possível selecionar o(s) planejador(es) mais adequado(s). Nos trabalhos de (ROBERTS; HOWE, 2009) e (SOUSA, 2014), com base em características, é proposto o uso de aprendizado de máquina, em específico reconhecimento de padrões, para resolver a tarefa do portfólio de planejamento para planejamento clássico. Vale ressaltar que o planejamento não clássico é formado por problemas onde os resultados das ações e os estados não são determinísticos.

Com base nestas informações, este trabalho propõe um conjunto de contribuições na área de planejamento probabilístico e no uso de técnicas de planejamento para o mercado financeiro. A principal contribuição almejada é a criação de um portfólio de planejadores para o planejamento não clássico e sua aplicação no mercado financeiro. O portfólio de planejamento tem o objetivo de auxiliar a escolha do planejador mais adequado para o mercado financeiro.

Outra contribuição consiste em explorar o planejamento usando múltiplos ativos. Para isso novas modelagens são propostas, tanto com a modelagem POMDP, assim como, com modelagens MOMDP. O trabalho também explora o uso de taxas de corretagem e o impacto que tais cobranças têm nas negociações. No trabalho de Baffa (BAFFA; CIARLINI, 2010b) os resultados foram obtidos apenas em relação ao ativo IBOVESPA. Já esta pesquisa realiza experimento com 5 diferentes ativos.

A tarefa de planejamento pode ser caracterizada como sendo *offline* ou *online*. O

planejamento *offline* dispõe de tempo para encontrar uma solução e apenas ao ser encontrada uma ação é executada. Já no *online*, soluções parciais e execução de ações são intercaladas de acordo com um intervalo de tempo.

Os trabalhos de (BAFFA; CIARLINI, 2010b) e (BRANQUINHO; LOPES; BAFFA, 2016) utilizam apenas planejadores *offline*. Contudo, o mercado financeiro apresenta características em que o planejamento *online* é muito indicado. Por exemplo, quando as negociações são realizadas no decorrer do dia as análises devem considerar os últimos dados e é dado um curto período de tempo para indicar qual operação realizar. Isto deve-se ao fato do mercado financeiro ser caracterizado um problema de séries temporais (KWON; MOON, 2007) e com o decorrer do tempo novos dados estão disponíveis. Outro ponto a ser levado em conta é que o portfólio de planejamento proposto no trabalho tem como base dados da competição *International Probabilistic Planning Competition* (IPPC). Na IPPC são utilizados planejadores *online*, logo, as indicações do portfólio de planejamento são realizadas para este tipo de planejamento. Deste modo, é importante que isto seja levado em conta na pesquisa.

Por último, uma atenção foi dada ao desenvolvimento de um ambiente que permita de forma fácil realizar a implementação de novas técnicas de IA destinadas ao mercado financeiro. Este projeto está disponível em <<https://github.com/bbranquinho/wpattern-wtrader>>.

Existem diferentes pontos de motivação no trabalho. A primeira motivação é o desafio de realizar o desenvolvimento do portfólio de planejamento de contexto geral e, posteriormente, seu uso no mercado financeiro. A possibilidade de estender o problema de planejamento usando uma modelagem MOMDP também é investigado nesta pesquisa. Finalmente, existe a motivação de que a pesquisa pode ser usada em um sistema real no mercado financeiro real.

1.1 Objetivos e Desafios da Pesquisa

O principal objetivo desta pesquisa é o desenvolvimento de um portfólio de planejamento aplicado para o planejamento probabilístico e seu respectivo uso como ferramenta de apoio a decisão em investimentos no mercado financeiro. Os objetivos específicos são listados a seguir:

- ❑ Pesquisa e desenvolvimento do portfólio de planejamento.
- ❑ Construção de uma modelagem POMDP para múltiplos ativos.
- ❑ Construção de uma modelagem MOMDP para o problema do mercado financeiro usando apenas um ativo.

- ❑ Construção de uma modelagem MOMDP para o problema do mercado financeiro usando múltiplos ativos.
- ❑ Aplicação do portfólio de planejamento no mercado financeiro.
- ❑ Uso de planejadores *offline* e *online* no mercado financeiro. Tem-se o objetivo de avaliar qual pode ser o melhor tipo de planejamento para o mercado financeiro.
- ❑ Uso de taxas de corretagem nas simulações de negociação. Tem-se o objetivo de avaliar o impacto que estas cobranças exercem nos resultados.
- ❑ Montagem de carteiras de investimento utilizando cálculos de correlação de ativos.

1.2 Contribuições

As principais contribuições alcançadas são:

- ❑ Desenvolvimento de um portfólio de planejamento de contexto geral para problemas de planejamento parcialmente observáveis.
- ❑ Aplicação do portfólio de planejamento para o problema do mercado financeiro.
- ❑ Elaboração de propostas de múltiplos ativos para problemas de planejamento do mercado financeiro.
- ❑ Construção de uma modelagem MOMDP para o mercado financeiro.
- ❑ Desenvolvimento de um ambiente que permita a implementação de novos algoritmos de IA para o mercado financeiro.
- ❑ Obtenção de melhores investimentos no mercado financeiro.
- ❑ Publicações em conferências e revistas relevantes as áreas pesquisadas.

1.3 Organização do Trabalho

O trabalho está estruturado nos seguintes capítulos especificados a seguir. No Capítulo 2 é realizada a apresentação de forma breve do problema do mercado financeiro e uma visão geral dos trabalhos correlatos. O Capítulo 3 descreve os fundamentos teóricos. Em seguida, o Capítulo 4 expõe os fundamentos sobre planejamento e diferentes planejadores. Com base nos fundamentos e nos trabalhos correlatos, o Capítulo 5 descreve as propostas de planejamento. A metodologia usada para realizar o trabalho é detalhada no Capítulo 6. Para avaliar a viabilidade técnica e teórica do trabalho foi realizado o desenvolvimento de um ambiente que suporte a pesquisa. Este sistema segue uma arquitetura brevemente

mostrada no Capítulo 7. Já o Capítulo 8 descreve os resultados obtidos. Por último, no Capítulo 9, é apresentada a conclusão e os trabalhos futuros.

Caracterização do Problema e Trabalhos Correlatos

Quando empresas buscam recursos uma das alternativas é lançar ações no mercado financeiro. O mercado financeiro dispõe de meios para a transferência de fundos entre empresas e investidores. As instituições que recebem investimentos disponibilizam a venda de ativos para os investidores. As ações, também chamadas de ativos, constituem frações das empresas, sendo separadas em ordinárias ou preferenciais. As ações ordinárias, dependendo de sua quantidade, possibilitam o poder de voto aos seus detentores nas decisões sobre o futuro das instituições. Já as ações preferenciais dão prioridade à divisão dos dividendos. Sendo assim, os investidores ao adquirirem ações se tornam sócios das empresas e podem até mesmo influenciar na condução dos negócios.

Para realizar as aplicações, os investidores avaliam criteriosamente cada aplicação, já que existem riscos ao atuar neste tipo de negócio. Apesar de normalmente serem aplicações lucrativas, não existem garantias de que um investimento sempre terá lucro. Sendo assim, os investidores avaliam os resultados das empresas para definir onde aplicar seus recursos. Além dos resultados das empresas, fatores econômicos nacionais e mundiais influenciam os preços dos ativos.

Para auxiliar os investidores, otimizar o investimento e minimizar os riscos são usadas as análises técnicas e análises fundamentalistas. A análise técnica utiliza um conjunto de técnicas e instrumentos analíticos para realizar a projeção de preços futuros das ações. Já a análise fundamentalista avalia a situação financeira da empresa, considerando seu estado econômico e mercadológico. Em síntese, a análise fundamentalista busca avaliar a saúde financeira das empresas, projetando seus resultados futuros e determinando o preço justo de suas ações.

Uma das origens da predição de preços é datada de 1700 no Japão pelo investidor Munehisa Homma. Ele aplicou técnicas para avaliar as variações dos preços no mercado de arroz (NISON, 1991). As técnicas usadas por Munehisa Homma deram origem aos padrões *Candlestick*.

Assim como Munehisa, Charles Henry Dow, um dos co-fundadores da financeira internacional Dow Jones, foi responsável por uma das teorias mais importantes em que se fundamentam o estudo do comportamento dos preços no mercado. Esta teoria, chamada de teoria de *Dow*, deu origem as análises técnicas (SCHANNEP, 2008).

Utilizando análises técnicas são realizados estudos para determinar o comportamento dos preços. Estes comportamentos podem indicar tendências de alta, baixa ou estabilidade. Se os valores das predições dos preços futuros forem superiores aos valores atuais um sinal de alta é gerado. Se as predições têm valores abaixo do preço atual é gerado o sinal de baixa. Caso sejam iguais ou próximos os valores atuais e das predições, temos um sinal de estabilidade. Sinais de alta indicam momentos de compra, sinais de baixa sugerem venda e sinais de estabilidade indicam a manutenção da posição sem que negócios sejam realizados. Com base nestes sinais temos o problema da obtenção do lucro. Este problema é descrito na Seção 2.1.

2.1 Caracterização do Problema

O lucro ou o prejuízo são dados pela diferença dos preços de compra e venda ou venda e compra. Quando é realizada a compra em um valor menor que a venda é obtido o lucro, no entanto, caso o preço na venda tenha sido superior ao da compra temos o prejuízo. Outra possibilidade de negociação é quando um investidor requisita a uma corretora ativos emprestados, que por sua vez são vendidos pelo investidor. Após certo tempo estas ações precisam ser recompradas e devolvidas. Neste caso o lucro é obtido quando os preços estão em queda. A Figura 1 mostra dois momentos, um de lucro e outro de prejuízo.

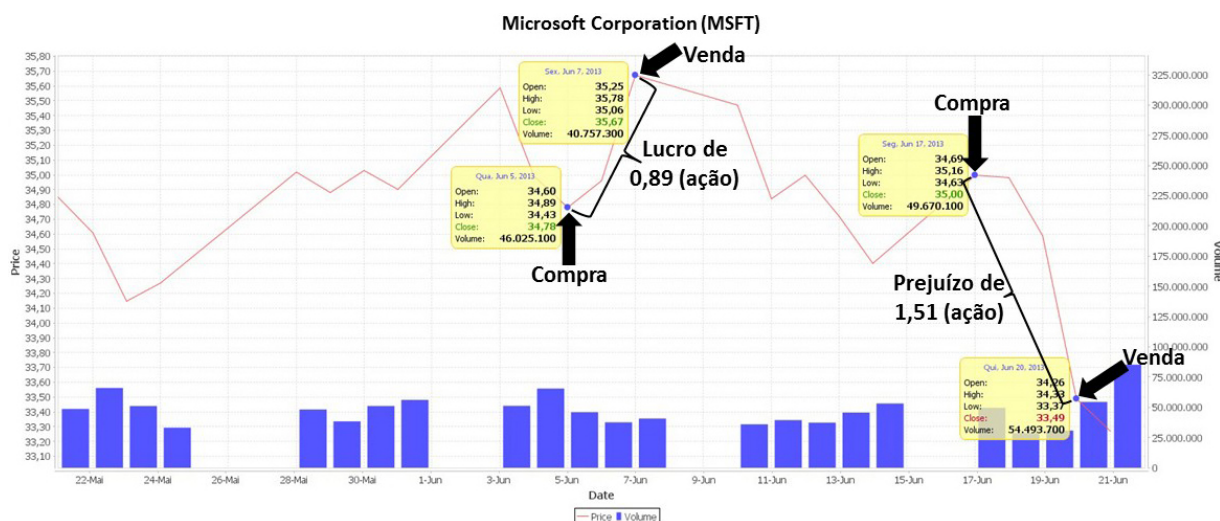


Figura 1 – Preço de fechamento diário e momentos de lucro e prejuízo para o ativo MSFT - Microsoft (2013).

Os dados de negociação do mercado financeiro correspondem a uma série temporal. Séries temporais consistem em um conjunto de observações ordenados por períodos regulares de tempo. Uma característica que norteia as séries temporais é que o tempo exerce influência nos dados de forma mais ou menos constante. Este sistema causal costuma criar padrões não aleatórios que podem ser detectados através de gráficos, métodos estatísticos ou técnicas de IA. No mercado financeiro os dados da série temporal correspondem aos preços das ações (COWPERTWAIT; METCALFE, 2009).

Um instante de tempo t de negociação é normalmente composto pelos dados $x(t)$, $x_h(t)$, $x_l(t)$ e $v(t)$. Estas variáveis representam, respectivamente, o preço de fechamento, o maior valor negociado, o menor valor negociado e o volume negociado. Formalmente, o problema de predição em séries temporais pode ser dada pela Equação 1.

$$x(t+1) = f(x(t), x(t-1), x(t-2), \dots). \quad (1)$$

Logo, o problema de predição consiste em prever x em um tempo $t+1$. Os valores de entrada para f podem ser valores resultantes de outros cálculos, tais como análises técnicas. Simplificações são adotadas de acordo com os modelos adotados em cada trabalho e algoritmo.

Um desafio consiste em determinar os momentos de inversão dos preços. Uma venda tardia ou antecipada pode ocasionar em prejuízos. Igualmente, as compras em momentos indevidos podem causar resultados não desejados. Desta forma, realizar a predição de preços e determinar a inversão de comportamento no mercado financeiro são grandes desafios.

Além da tarefa de predição, existe o problema da escolha dos ativos que serão negociados. Este problema é tratado como a geração de uma carteira de ações. Normalmente para esta tarefa são considerados fatores de risco dos ativos e correlação entre eles. Em geral as carteiras auxiliam a escolha dos investimentos, além de ajudar no processo de predição. Para montar a carteira de ativos uma técnica que geralmente é usada é a correlação cruzada ou também chamada de correlação de ativos.

O problema fundamental do mercado é a maximização do lucro. Este trabalho além de se preocupar com os resultados de lucratividade, tem um esforço para tratar de problemas que envolvem IA, em especial a área de planejamento.

2.2 Trabalhos Correlatos

O mercado financeiro permite o uso de diversas técnicas de IA. Apesar desta ser uma excelente característica o problema de determinar quais métodos são os mais adequados é um desafio. Além disso, em geral os resultados dos algoritmos de IA variam de acordo com as bolsas de valores e os ativos. Portanto, realizar a comparação entre trabalhos é algo desafiador.

Atsalakis realiza um estudo em (ATSALAKIS; VALAVANIS, 2009) e (ATSALAKIS; VALAVANIS, 2013) que compara os resultados de 150 artigos relacionados a predição do comportamento do mercado financeiro. São avaliados por este trabalho os algoritmos estatísticos clássicos: Volatilidade Estocástica, Média Móvel Auto Regressiva, Auto Regressivo, Média Móvel Auto Regressiva Integrada e Heterocedasticidade Condicional Auto Regressiva. Os mesmos autores realizam em (ATSALAKIS; VALAVANIS, 2009) um estudo comparativo com mais de 100 artigos relacionados a predição a partir de algoritmos de aprendizado de máquina supervisionado. Os autores indicam que para a tarefa de predição do comportamento do mercado, algoritmos de aprendizado de máquina supervisionado produzem melhores resultados em relação aos algoritmos estatísticos clássicos.

A predição de alta e queda de preços para o dia seguinte é investigada em (ABRAHAM; PHILIP; SARATCHANDRAN, 2003). Na pesquisa é usado um período de sete anos do índice Nasdaq-100¹ e um período de quatro anos do índice NIFTY 50². Foram usados os algoritmos de aprendizado de máquina supervisionado redes neurais artificiais, *Levenberg-Marquardt*, máquina de vetores de suporte (denotado pelo o acrônimo SVM, da expressão em inglês *Support Vector Machine*), neural-nebuloso e *Difference Boosting Neural Network*. Em geral foram obtidos bons resultados, em especial para o SVM.

É realizado em (HUANG; NAKAMORI; WANG, 2005) a predição semanal de alta e queda do índice NIKKEI-225 para o período de um ano. Foi analisada a capacidade preditiva dos seguintes algoritmos de aprendizado de máquina supervisionado: análise discriminante linear (denotado pelo o acrônimo LDA, da expressão em inglês *Linear Discriminant Analysis*), análise discriminante quadrática (denotado pelo o acrônimo QDA, da expressão em inglês *Quadratic Discriminant Analysis*) e rede neural de Elman. Os resultados indicam uma taxa de 75% de acerto para o SVM tendo sido este o melhor dentre os algoritmos.

Existem diferentes algoritmos de treinamento de redes neurais, tais como *backpropagation* (RUMELHART; HINTON; WILLIAMS, 1988), *simulated annealing backpropagation* (ENGEL, 1988) e *momentum* (SUTSKEVER et al., 2013). Apesar dos diferentes meios de realizar os ajustes dos pesos da rede, os métodos citados podem ficar presos em mínimos ou máximos locais. Para lidar com este problema Kim (KIM; MOON, 2002) propõe o uso de uma abordagem usando Algoritmo Genético (AG) para otimizar o processo de treinamento das redes. É importante ressaltar que a proposta de Kim permite trabalhar juntamente com os diferentes algoritmos de treinamento.

Nos trabalhos (KIM et al., 2001) e (KIM; KIM; LEE, 2002) são apresentados estudos sobre a correlação entre ativos. Nesta pesquisa foram usados os dados do índice S&P 500 durante o período 1993 a 1997. São utilizados os cálculos de correlação simples e correlação relativa. Com base nos cálculos de correlação é gerado um grafo completo.

¹ Corresponde a um índice da bolsa norte-americana NASDAQ que reúne 100 das maiores empresas não financeiras

² Corresponde a um índice da bolsa Indiana que reúne 50 empresas

Cada nó está associado a um ativo e a aresta corresponde a correlação entre dois ativos. Além dos valores de correlação o trabalho avalia a influência absoluta de cada vértice do grafo, ou seja, a influência que cada ativo exerce sobre o cenário global do grupo de empresas que compõem o índice S&P 500. Kim conclui que existe um comportamento compartilhado entre diferentes ativos. Deste modo, o trabalho demonstra que os fatores cooperativos e adaptativos refletem as variações de preços das companhias. Posto isto, os dados de variações dos preços, e consequentemente os valores de correlação, podem ser usados para melhorar resultados de predições de preços dos ativos.

Kwon (KWON; MOON, 2003) utiliza das técnicas apresentados em (KIM; MOON, 2002) e propõe um sistema de predição e simulação de preços de ativos do mercado financeiro. É usado um AG em conjunto de uma rede neural recorrente. Como proposto em (KIM; MOON, 2002), o AG é responsável por ajudar no treinamento da rede. Uma vez encontrada a melhor rede, são realizadas simulações de negociação e o lucro ou prejuízo são calculados. São usados 64 valores como entrada, todos gerados a partir de análise técnica. Os experimentos foram executados para 36 ativos da bolsa NYSE e NASDAQ durante o período de 1992 a 2001, tendo obtido melhores resultados (AG com rede neural) que a rede neural quando usada individualmente. Os resultados obtidos são comparados com a estratégia *buy & hold*. A estratégia *buy & hold* consiste em realizar uma compra no início do período e a venda apenas ao final. O AG conseguiu melhores resultados que a estratégia *buy & hold* em 153 casos, piores em 88 casos e similares em 118 casos. Na média o AG e a rede obtiveram performance melhor que a estratégia *buy & hold*, apesar de em alguns casos terem tido prejuízo.

Kwon apresenta em (KWON; MOON, 2004) uma continuação de seu trabalho (KWON; MOON, 2003), sendo realizada a adição da cobrança de taxas e do uso de *ensemble methods* (MACLIN; OPITZ, 1999) (ZHOU, 2012). Um fator que tem impacto direto nos resultados de negócios no mercado financeiro são as taxas cobradas para realizar as operações de compra e venda. Dependendo da quantidade de operações e das taxas, o valor pago pode tornar inviável algumas técnicas. Uma característica levantada por Kwon (KWON; MOON, 2004) é que não necessariamente uma rede neural com a menor taxa de erro garante bons resultados. Para tratar este problema uma possibilidade é o uso *ensemble methods*, que têm o objetivo de utilizar mais de uma técnica e combina-lás de modo que as decisões sejam tomadas em grupo. Normalmente, em um AG a melhor solução corresponde ao indivíduo com a melhor aptidão. Kwon (KWON; MOON, 2004) realiza o uso de uma abordagem *ensemble* aplicada sobre um AG, de modo que um grupo de indivíduos são selecionados e posteriormente é aplicado o algoritmo de clusterização *k-means* (MACQUEEN, 1967). Logo, tem-se uma solução formada por um grupo de indivíduos, chamada de *cluster*. São usados 75 valores como entrada, todos gerados a partir de análise técnica.

Kwon propõe em (KWON; CHOI; MOON, 2005) a continuação de seu trabalho reali-

zado em (KWON; MOON, 2004). Neste artigo são criados os portfólios de ativos através do cálculo dos coeficientes de correlação, tendo como base as propostas de (KIM et al., 2001) e (KIM; KIM; LEE, 2002). De posse do portfólio contendo k ativos, são realizados os cálculos das análises técnicas, em seguida, estes valores são dados como entrada para as redes neurais. Para cada ativo tem-se o total de 75 variáveis, totalizando $75 * k$ entradas para as redes. Assim sendo, é realizada a seleção de ativos por meio de um vetor de *bits* que representam a presença ou ausência de cada variável de entrada da rede. Consequentemente, para cada execução do AG tem-se um vetor de tamanho $75 * k$. Apesar de não ter obtido lucro para todos os ativos, os resultados foram melhores em relação ao trabalho anterior (KWON; MOON, 2004).

Em (KWON; MOON, 2007) é realizada a continuação dos trabalhos (KWON; MOON, 2003), (KWON; MOON, 2004) e (KWON; CHOI; MOON, 2005). Foi realizada uma comparação do AG juntamente da rede neural recorrente em relação a duas outras abordagens, a programação genética (denotado pelo acrônimo GP, da expressão em inglês *Genetic Programming*) (KABOUDAN, 2000) e SVM (CAO; TAY, 2003). A proposta do trabalho apresentou bons resultados em comparação a GP e ao SVM. No contexto geral todos os pontos dos trabalhos são bem detalhados por Kwon, entretanto, algumas informações não foram descritas. Por exemplo, nem todas os cálculos das variáveis de entrada foram apresentados.

Alvim (ALVIM, 2013) propõe uma metodologia que consiste em dividir o dia de mercado em intervalos de comprimentos distintos e gerar subconjuntos de dados agrupados por comprimento de intervalo e ativo. Desta forma, para cada conjunto de dados é analisada a possibilidade da existência de oportunidades. É dito que no tempo t ocorre uma oportunidade para negociar um ativo quando é realizada uma previsão de aumento significativo no preço no instante de tempo $t + \alpha$. Ao observar essa oportunidade, uma estratégia de negociação emite ordens de compra para o tempo t e ordens de venda para o tempo $t + \alpha$.

Para a tarefa de negociação Alvim (ALVIM; MILIDIÚ, 2013) (ALVIM, 2013) usa o conceito do *trader atômico*. Define-se o conceito de *trader atômico* τ dividindo o dia de mercado em intervalos discretos de tempo $[b, e]$. O *trader atômico* τ é representado por uma 3-upla (s, b, e) , tal que s representa um ativo, b é o tempo inicial e e é o tempo final do intervalo de negociação. Uma vez formulado o conceito do *trader atômico* tem-se a formação de um time de *traders*. Existem duas abordagens para a formação dos times de *traders*. A primeira abordagem adiciona mais de um *trader* para os intervalos e trabalha com o montante total de dinheiro aplicado por ativo em cada intervalo. A segunda abordagem considera aplicações em diferentes ativos usando o fator de risco como meio de decisão. Para isso um conjunto de *traders* compartilham um mesmo intervalo. Este problema é equivalente ao problema de agendamento de intervalos ponderados. Para os experimentos foram selecionados 10 ativos com maior liquidez no período de 2009 para a

bolsa BM&FBOVESPA.

No trabalho de Baffa (BAFFA; CIARLINI, 2010a) tem-se o objetivo de aplicar técnicas de planejamento para a análise e o auxílio na tomada de decisão de compra e venda de ativos no mercado financeiro. Desta forma, propõe-se um ambiente de estudo para modelagem, planejamento e simulação, utilizando POMDP (CASSANDRA; KAEHLING; LITTMAN, 1994). A proposta de (BAFFA; CIARLINI, 2010a) consiste em modelar os estados do mercado a partir das tendências passadas e futuras dos preços e a posição do investidor. As observações são indicações de pontos em que devem ou não serem realizados negócios e são geradas usando análise técnica. A partir dos dados históricos são calculadas as probabilidades das transições dos estados e das observações. Uma vez realizada a construção do modelo são usados algoritmos de POMDP (CASSANDRA; KAEHLING; LITTMAN, 1994) para gerar políticas, de maneira que sejam mapeados estados de crença em ações.

Um dos primeiros problemas encontrados em Baffa (BAFFA; CIARLINI, 2010b) é que o planejador PBVI foi usado sem existir um embasamento para a sua escolha. Conforme resultados das competições *International Planning Competition* (IPC)³ e *International Probabilistic Planning Competition* (IPPC)⁴, o melhor planejador pode variar de acordo com o domínio e com a instância de um problema. Assim sendo, o planejador mais adequado para o mercado financeiro pode variar de acordo com a instância do problema ou da modelagem escolhida. Uma possibilidade para realizar esta escolha consiste em executar todos os planejadores para cada instância e modelagem, contudo, esta solução é inviável quando existem muitos planejadores, inúmeras instâncias e várias formas de modelagem. Desta forma, a proposta do uso de portfólio de planejamento torna-se uma excelente opção para tratar da escolha do planejador.

Outros problemas encontrados nas pesquisas de Baffa (BAFFA; CIARLINI, 2010b) são pontos de estudo neste trabalho. Dentre estes pontos, tem-se a necessidade de tratar o uso de múltiplos ativos. Em outras palavras, utilizar mais de um ativo durante o planejamento e negociações. Outra necessidade é o de considerar o custo das negociações, uma vez que estas cobranças afetam diretamente os resultados. Um característica observada na modelagem de Baffa, é que o estado é composto de múltiplas “variáveis”, sendo cada variável de um tipo (determinística, observável, parcialmente observável). Esta característica abre margem para outra forma de modelagem, sendo ela a MOMDP (ONG et al., 2010).

É realizada em Roberts (ROBERTS; HOWE, 2007) uma pesquisa sobre planejamento clássico com 28 planejadores e 4726 problemas. O trabalho utiliza aprendizado de máquina para avaliar o desempenho de planejadores a partir dos dados das características dos problemas. São levantadas duas questões. A primeira é se é possível obter resultados

³ Detalhes disponíveis em: <<http://www.cs.colostate.edu/~ipc2014/>>

⁴ Detalhes disponíveis em: <http://users.cecs.anu.edu.au/~ssanner/IPPC_2014/>

precisos de desempenho do planejador? Em segundo lugar, os dados e os modelos podem ser usados para melhorar o desempenho do planejador? Com base nestas questões é demonstrado que os modelos podem ajudar no entendimento dos problemas e na avaliação do desempenho dos planejadores. O estudo indica que o portfólio de planejadores é melhor que planejadores individuais. Também é discutido como os modelos podem ser analisados para entender melhor de que maneira as decisões de design do planejador contribuem para o seu desempenho.

Em (ROBERTS; HOWE, 2009), Roberts organiza os problemas e os planejadores clássicos em termos de dificuldade, diversidade e tendência ao longo do tempo. Foi confirmado estatisticamente que os conjuntos de problemas se tornaram mais difíceis e que os novos planejadores são mais capazes de tratar tais problemas. O estudo valida os resultados que ligam a topologia de busca dos planejadores ao seu desempenho. São lacunas encontradas nas complexidades de certos problemas, para isso são construídos problemas mais desafiadores. Estes estudos permitem ajuda pesquisadores na avaliação comparativa de seus planejadores e se existe a necessidade de esforços adicionais.

É apresentando em (SEIPP et al., 2013) uma proposta de portfólio de planejamento clássico e otimização de parâmetros para o planejador *fast downward*. A otimização dos parâmetros é realizada com a biblioteca *ParamILS*⁵. O portfólio de planejamento é aplicado nos domínios da competição IPC e obtiveram pontuações melhores que os planejadores das competições anteriores.

Nos trabalhos de Sousa (SOUSA, 2014)(SOUSA; LOPES, 2014) é colocado o estudo sobre portfólio de planejamento clássico para os problemas da competição IPC. Para montar o portfólio são usados algoritmos de reconhecimento de padrões, também chamados de classificadores. São extraídas características dos problemas e um AG é usado para selecionar as características mais adequadas para cada algoritmo. Com base nas indicações do portfólio é realizado o escalonamento dos planejadores de acordo com o percentual de indicações, este escalonamento consiste apenas em atribuir um percentual de tempo para executar o planejamento.

Um estudo da arte sobre portfólio de planejamento é realizado em (VALLATI; CHRPA; KITCHIN, 2015). São levantados trabalhos correlatos, definições sobre o tema e desafios que estão em aberto e demandam pesquisa.

Concluindo, existem muitos desafios sobre planejamento e mercado financeiro. Há diversos trabalhos com as mais variadas técnicas de IA aplicadas no mercado financeiro. Os próximos capítulos apresentam a pesquisa proposta tomando como base os trabalhos correlatos.

⁵ Informações disponíveis em: <<http://www.cs.ubc.ca/labs/beta/Projects/ParamILS/>>

Fundamentação Teórica

3.1 Análise Técnica

A análise técnica consiste no estudo dos preços de ativos de uma empresa utilizando análise gráfica. A base deste estudo consiste na teoria de que os preços dos ativos seguem padrões que se repetem durante intervalos de tempo. Através destes padrões torna-se possível a realização das previsões dos preços. Com estes estudos investidores realizam negociações, tendo como objetivo melhorar os investimentos.

O conceito e o uso de análise técnica é fundamentada pela “Teoria de Dow”. Esta teoria sugere que o mercado de ações apresenta comportamentos previsíveis. A “Teoria de Dow” tem como base os princípios “a natureza da tendência dos preços”, “os preços descontam tudo”, “o volume confirma a tendência”, “confirmação e divergência” e “suporte/resistência” (SCHANNEP, 2008).

Em 1933 (COWLES, 1933) foi realizada a primeira publicação sobre previsões de preços baseadas em análises técnicas, tendo uma taxa de 55% de métricas de avaliação. Contudo, em (FAMA, 1965) e (JENSEN; BENINGTON, 1970) foram publicadas indicações da ineficácia de certos indicadores para a previsão do mercado. Estas publicações deram origem a conhecida Hipótese do Mercado Eficiente (FAMA, 1970), que “sugere que os preços correntes de mercado já refletem toda a informação histórica de mercado e que não é possível realizar previsões de mercado baseadas em um histórico de mercado” (ALVIM, 2013). Entretanto, estes estudos foram realizados considerando períodos curtos e com poucos ativos. Apesar das evidências levantadas por (FAMA, 1970), muitos trabalhos demonstram que é possível realizar previsões e obter lucro no mercado financeiro através da aplicação de análises técnicas. (BROCK; LAKONISHOK; LEBARON, 1992) e (LO; MAMAYSKY; WANG, 2000) são trabalhos relevantes na área financeira e sustentam a aplicação das análises técnicas para a realização de negócios.

Parte das análises utilizadas na pesquisa são brevemente descritas ou citadas durante o trabalho. É importante observar que foram mostrados detalhes de apenas algumas análises, já que este tópico é extremamente extenso e amplamente difundido na literatura.

A pesquisa apenas utiliza os valores gerados pelas análises, não avaliando efetivamente a eficiência e eficácia de cada uma delas. As próximas seções apresentam brevemente algumas destas análises técnicas.

3.1.1 Candlesticks

Uma das formas mais simples de interpretar gráficos é a partir de padrões de *candles*, também conhecidos como *candlesticks*. Um *candle* é formado por um corpo, sombra/pavio, preço de abertura, preço de fechamento, maior preço e menor preço. Existem dois tipos de *candles*, sendo o de baixa e o de alta. Estes *candles* ocorrem, respectivamente, quando o preço de fechamento é inferior ao de abertura e quando o preço de fechamento é superior ao preço de abertura. A Figura 2 exibe exemplos de *candles*.

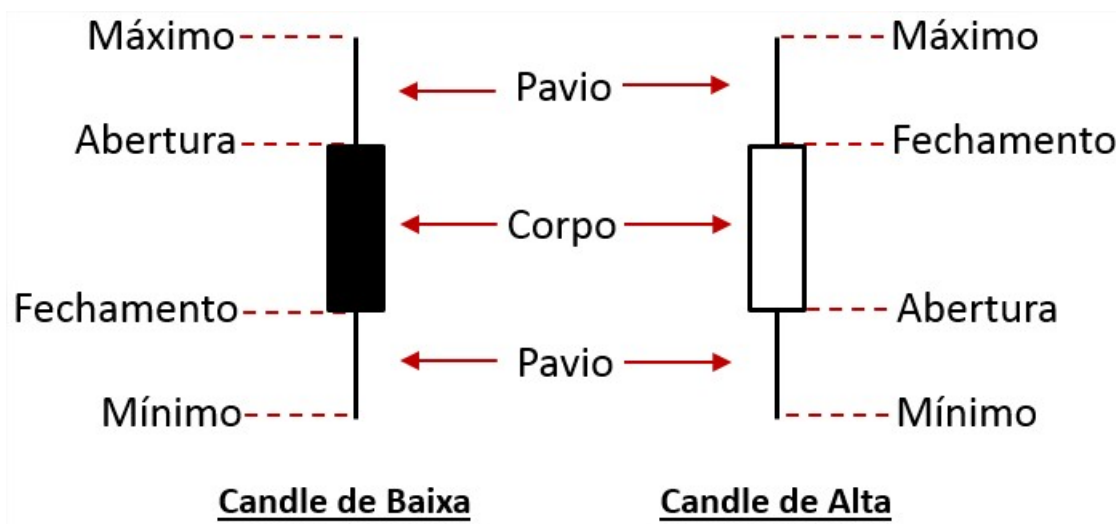


Figura 2 – Representação de dois *candles*, sendo respectivamente, um de baixa e outro de alta.

Para a formação dos *candles* primeiro é necessário determinar o intervalo de tempo, sendo normalmente mensurado em minutos, dias ou semanas. Após determinar o intervalo, os negócios são divididos em janelas respeitando os intervalos. Para cada uma destas janelas os preços são recuperados para a formação dos gráficos. Um exemplo de *candles* diários para o ativo PETR4 é mostrado na Figura 3.

Analisando sequências de *candles* são abstraídos comportamentos dos ativos. Para que seja considerado um padrão é preciso que em geral a taxa de acertos seja superior a 50%. Em (NISON, 1991) são mostrados mais de 50 tipos de padrões *candlesticks*.

Entre os diversos padrões de *candles*, são citados como exemplos o **engolfo de alta**, **engolfo de baixa**, **bebê abandonado de alta** e **bebê abandonado de baixa**.

O **engolfo de alta** ocorre após um momento de queda, o valor do ativo no período analisado possui uma abertura inferior ao anterior e o fechamento ultrapassa a abertura do dia anterior. Essas características indicam que provavelmente o preço irá subir. O



Figura 3 – Gráfico de *candles* diários para o ativo PETR4.

engolfo de baixa segue o mesmo conceito, porém, ocorre após um mercado em alta e indica tendência de queda (BARROSO et al., 2012). A Figura 4 exhibe estes padrões de engolfo.

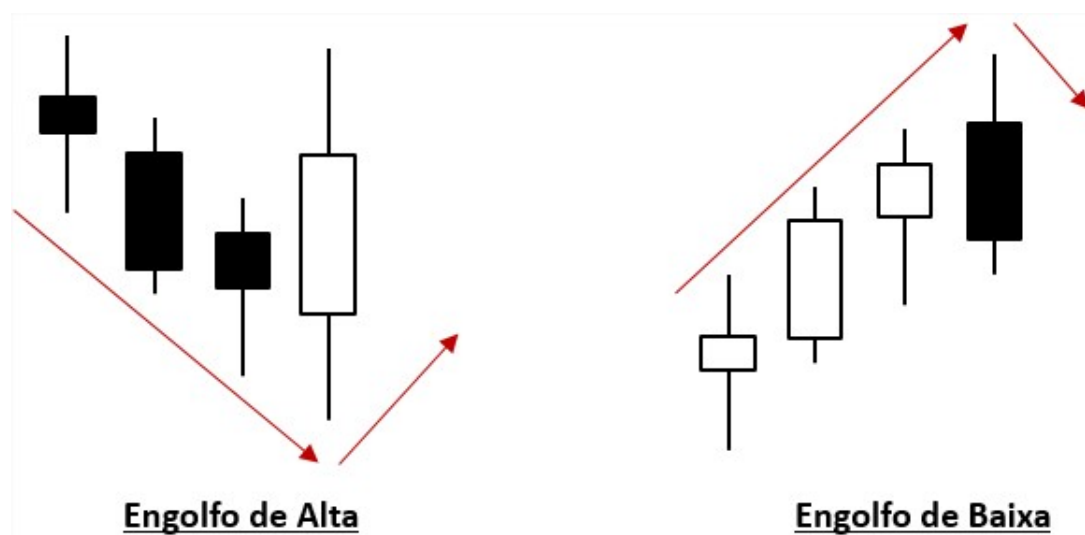


Figura 4 – Padrões engolfo de alta e engolfo de baixa e as respectivas tendências.

O padrão **bebê abandonado** ocorre após uma queda de preço, as compras se equiparam às vendas, estabilizando o preço. Esse equilíbrio entre compras e vendas significa que a abertura e o fechamento possuem valores próximos. Após isso, o mercado começa a subir, e possivelmente continuará assim, pois finalmente as compras conseguiram se

sobressair às vendas. O bebê abandonado de baixa segue a mesma explicação do bebê abandonado, porém, ocorre após um período de alta e indica tendência de baixa (BARROSO et al., 2012). A Figura 5 mostra o padrão bebê abandonado.

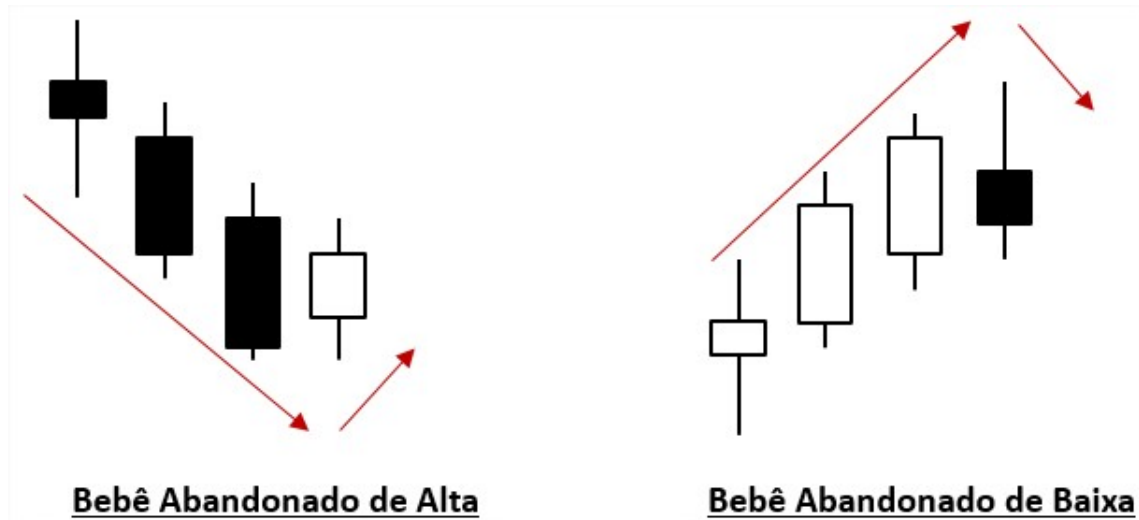


Figura 5 – Padrões bebê abandonado de alta e bebê abandonado de baixa e as respectivas tendências.

Os padrões de *candles* podem ser usados tanto para o treinamento de algoritmos, assim como para a confirmação de oportunidades. Um exemplo de uso é realizado por Barroso (BARROSO et al., 2012), onde os padrões são usados para confirmar tendências geradas por redes neurais.

3.1.2 Médias Móveis

A média móvel (denotado pelo acrônimo MA, da expressão em inglês *Moving Average*) consiste em suavizar conjuntos de dados para formar indicadores de tendências. O objetivo não é de realizar a predição da direção dos preços, mas sim de definir a direção considerando um pequeno atraso. Este atraso consiste em uma suavização das variações considerando os preços passados com a função de filtrar “ruídos”. Estes ruídos são tratados como variações bruscas que podem levar a falsas interpretações das direções dos preços (MURPHY; MURPHY, 1999) .

As duas MAs mais comuns são a média móvel simples (caracterizada pelo acrônimo SMA, da expressão em inglês *Simple Moving Average*) e a média móvel exponencial (caracterizada pelo acrônimo EMA, da expressão em inglês *Exponential Moving Average*). Algumas análises, como *Bandas de Bollinger* e *Convergência/Divergência das Médias Móveis* (caracterizada pelo acrônimo MACD, da expressão em inglês *Moving Average Convergence/Divergence*), têm MAs como base. A Figura 6 exhibe exemplos da SMA e EMA para o ativo PETR4.



Figura 6 – Exemplos de média móvel simples (SMA) e média móvel exponencial (EMA) para o ativo PETR4.

A SMA é calculada a partir da média dos últimos k períodos. O valor de k em geral é dado como sendo de curto, médio ou longo prazo. Os períodos em geral variam, mas normalmente o curto prazo é caracterizado de 5 a 20 períodos, o médio prazo em torno de 50 períodos e o longo prazo 100 períodos. Considerando x_i um valor no período i , a Equação 2 expressa a SMA_k para k períodos.

$$SMA_k = \frac{\sum_{i=1}^k x_i}{k} \quad (2)$$

Para realizar o cálculo de uma SMA_7 são necessários os dados dos últimos 7 períodos. Conforme o exemplo da Tabela 2 a SMA_7 da data 28/07/2014 é igual a $SMA_7 = (11,22 + 11,02 + 11,49 + 11,47 + 11,68 + 11,82 + 12,01)/7 = 11,53$.

A EMA consiste em aplicar um maior peso para os dados mais recentes. A EMA_k possui k períodos e é calculada em três passos. No primeiro passo é calculada a SMA usando o mesmo número de períodos da EMA. Posteriormente, é calculado o peso de acordo com a quantidade de períodos conforme a Equação 3.

$$peso = 2/(k + 1) \quad (3)$$

Ao final é calculada a EMA_k usando a Equação 4.

$$EMA_k = EMA_k^{anterior} + (x - EMA_k^{anterior}) * peso \quad (4)$$

Tabela 2 – Exemplos de valores da SMA_7 e EMA_7 para o preço de fechamento do período de 18/07/2014 a 30/07/2014.

Período	Preço de Fechamento	SMA_7	peso	EMA_7
18/07/2014	11,22			
21/07/2014	11,02			
22/07/2014	11,49			
23/07/2014	11,47			
24/07/2014	11,68			
25/07/2014	11,82			
28/07/2014	12,01	11,53		12,01
29/07/2014	11,92	11,63	0,25	11,98
30/07/2014	12,14	11,79	0,25	12,02

Na Tabela 2 são mostrados exemplos de valores da EMA_7 . Na data 28/07/2014 o EMA_7 é igual a SMA_7 , já que não existe $EMA_7^{anterior}$. No próximo período, 29/07/2014, o valor do EMA_7 é igual a $EMA_7 = (11,92 - 12,01) * 0,25 + 12,01 = 11,98$. Por último, para 30/07/2014 tem-se $EMA_7 = (12,14 - 11,98) * 0,25 + 11,98 = 12,02$.

3.1.3 Convergência/Divergência das Médias Móveis (MACD)

A Convergência/Divergência das Médias Móveis (caracterizada pelo acrônimo MACD, da expressão em inglês *Moving Average Convergence/Divergence*) consiste em um indicador baseado em duas médias móveis (APPEL, 1985). Para isso é realizada a subtração da $MA^{rápida}$ e a MA^{lenta} , que correspondem a dois períodos, sendo respectivamente, o menor período e o maior período. O MACD varia abaixo e acima de 0 de acordo com a convergência, cruzamento e divergência das MAs. Além do indicador, o MACD gera um sinal e um histograma (APPEL, 2005). Os cálculos do MACD são expressos pelas Equações 5a, 5b e 5c.

$$MACD = MA^{rápida} - MA^{lenta} \quad (5a)$$

$$MACD^{sinal} = MA_k \text{ da } MACD \quad (5b)$$

$$MACD^{histograma} = MACD - MACD^{sinal} \quad (5c)$$

A Figura 7 mostra o diagrama do MACD. Neste exemplo são mostrados os 3 atributos do MACD. Foram usadas a EMA_{12} , EMA_{26} e o sinal com $k = 9$ períodos.

De acordo com os valores do MACD e das MAs são realizadas análises de tendências. Em (APPEL, 2005) são descritas diversas interpretações possíveis para a MACD.

3.1.4 Oscilador Estocástico

O Oscilador Estocástico (denotado pelo acrônimo OST, da expressão em inglês *Stochastic Oscillator*) é um indicador que mostra a relação entre o preço de fechamento e



Figura 7 – MACD para o ativo PETR4.

suas máximas e mínimas durante certo período de tempo (ACHELIS, 2006). O cálculo da OST é realizado pelas Equações 6a e 6b.

$$\%K = 100 \left[\frac{Close - Low_{14}}{High_{14} - Low_{14}} \right] \quad (6a)$$

$$\%D = 3\text{-period moving average of } \%K \quad (6b)$$

As variáveis $\%K$ e $\%D$ correspondem a duas linhas do OST, sendo a primeira considerada mais lenta (sensível) e a outra mais rápida (menos sensível).

3.1.5 Bandas de Bollinger

As bandas de bollinger foram desenvolvidas por (BOLLINGER, 2001) e corresponde a regiões de volatilidade que estão acima e abaixo de uma MA. A volatilidade é baseada no desvio padrão, o que faz com que as regiões se alarguem ou se estreitem. As Equações 7a, 7b e 7c, demonstram o cálculo das bandas de bollinger.

$$\text{Banda Média} = MA_k \quad (7a)$$

$$\text{Banda Superior} = MA_k + \sigma_k * 2 \quad (7b)$$

$$\text{Banda Inferior} = MA_k - \sigma_k * 2 \quad (7c)$$

O termo σ_k corresponde ao desvio padrão para os últimos k períodos. A Figura 8 mostra o gráfico das bandas de bollinger para o ativo PETR4.

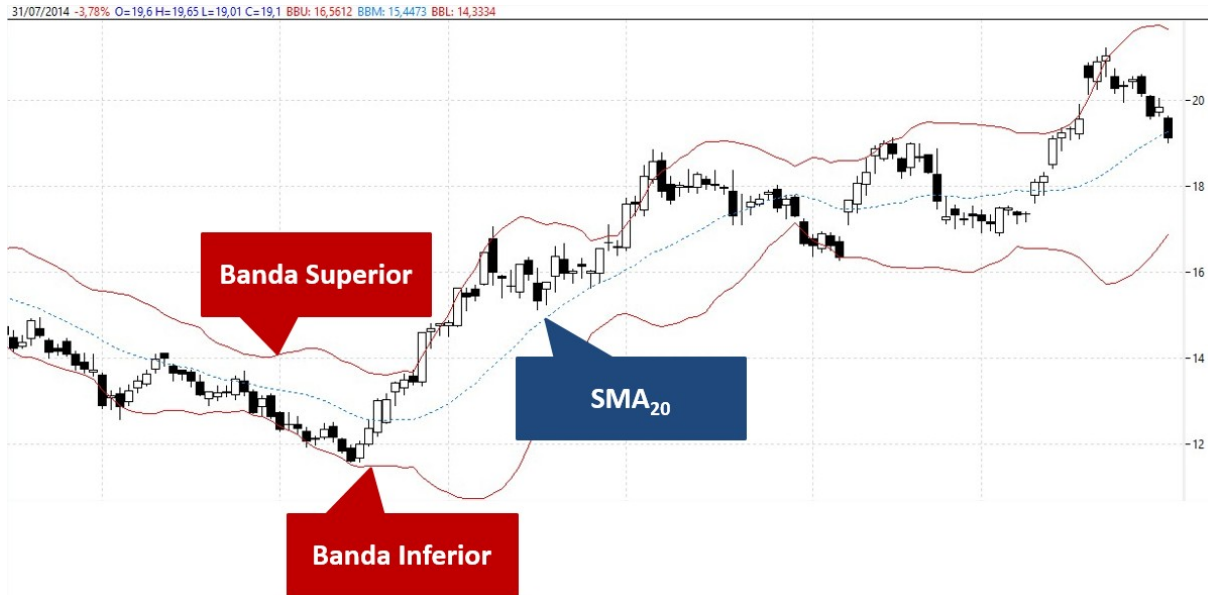


Figura 8 – Bandas de Bollinger para o ativo PETR4.

3.1.6 Oscilador de Preço Absoluto

O oscilador de preço absoluto (denotado pelo acrônimo APO, da expressão em inglês *Absolute Price Oscillator*) é um indicador baseado na diferença de duas EMAs. Esta análise corresponde a uma redução do MACD, sendo este o valor absoluto por ser apenas o sinal, ou seja, a diferença do menor período da EMA ($EMA^{rápida}$) e do maior período da EMA (EMA^{lenta}). O cálculo desta análise é expresso pela Equação 8.

$$APO = EMA^{lenta} - EMA^{rápida} \quad (8)$$

A Figura 9 exibe o gráfico do oscilador de preço absoluto para o ativo PETR4. O sinal com valor positivo indica um movimento de tendência de alta, enquanto um valor negativo representa uma tendência de baixa.

3.1.7 Índice de Força Relativa

O Índice de Força Relativa (denotado pelo acrônimo RSI, da expressão em inglês *Relative Strength Index*) mede a velocidade dos movimentos de mudanças de preços e tem seu valor variando entre 0 e 100. De acordo com (WILDER, 1978) quando o RSI está acima de 70 indica que o ativo está *overbought*¹ e quando está abaixo de 30 está *oversold*². O cálculo do RSI é realizado pela Equação 9.

$$RSI = 100 - (100 / (1 + \text{Ganho Médio} / \text{Perca Média})) \quad (9)$$

¹ Overbought: Preços considerados muito elevados levando a uma tendência de queda.

² Oversold: Preços considerados muito baixos levando a uma tendência de alta.



Figura 9 – Oscilador de preço absoluto para o ativo PETR4 com 12 e 26 períodos.

Ganho Médio e *Perca Média* correspondem, respectivamente, aos ganhos e percas durante 14 períodos. Este número de períodos é o padrão sugerido por (WILDER, 1978). As percas são expressas por valores positivos e não negativos. O primeiro cálculo destes dois atributos são dados pelas Equações 10a e 10b.

$$\text{Ganho Médio} = \frac{\sum_{i=1}^{14} \text{Ganho}_i}{14} \quad (10a)$$

$$\text{Perca Média} = \frac{\sum_{i=1}^{14} \text{Perca}_i}{14} \quad (10b)$$

Após o primeiro cálculo, os próximos valores de *Ganho Médio* e *Perca Média* são baseados nas Equações em 11a e 11b.

$$\text{Ganho Médio} = ((\text{Ganho Médio}_{\text{Anterior}}) * 13 + \text{Ganho}_{\text{Atual}}) / 14 \quad (11a)$$

$$\text{Perca Média} = ((\text{Perca Média}_{\text{Anterior}}) * 13 + \text{Perca}_{\text{Atual}}) / 14 \quad (11b)$$

Os parâmetros *Ganho Médio_{Anterior}* e *Perca Média_{Anterior}* são iguais aos valores que precedem o valor atual. Já os valores atuais do ganho e perca correspondem, respectivamente, a *Ganho_{Atual}* e *Perca_{Atual}*.

A Figura 10 mostra o gráfico do RSI para o ativo PETR4. Ocorrem 4 eventos nesta análise, sendo os dois primeiros *Oversold* e os últimos dois como *Overbought*. Além destas duas tendências, existem outras avaliações que podem ser realizado considerando a RSI (WILDER, 1978).



Figura 10 – *RSI* para o ativo PETR4.

Utilizar os valores posteriores de ganho e perda permitem “suavizar” o cálculo, sendo esta operação similar ao que é feito na EMA. Isso também significa que os valores são mais apurados com o decorrer dos períodos.

3.1.8 Índice de Força Relativa Estocástica

O Índice de Força Relativa Estocástica (denotado pelo acrônimo Stochastic RSI, da expressão em inglês *Stochastic Relative Strength Index*) consiste em comparar o valor relativo do maior e menor RSI de um período. Sendo assim, este é um indicador de um indicador. O resultado é expresso no intervalo de 0 a 1 (BROWN, 1999). O cálculo é realizado pela Equação 12.

$$RSI_{Stoch} = (RSI - RSI_{Menor}) / (RSI_{Maior} - RSI_{Menor}) \quad (12)$$

Stochastic RSI mensura o valor do RSI em relação ao intervalo do maior e do menor RSI do período. Assim como no RSI, por padrão são usados 14 períodos por cálculo da análise.

3.1.9 Média Móvel Exponencial Dupla

MAs proporcionam uma “suavização” dos dados relativos a preços para formar linhas de tendência. Contudo, as MAs não realizam a predição dos preços e sim um pequeno

atraso do valor atual. Murphy (MURPHY; MURPHY, 1999) define a Média Móvel Exponencial Dupla, também chamada de “*double crossover method*”, como a correlação de duas EMA para verificar o cruzamento de sinais e avaliar tendências. A Figura 11 mostra os sinais de suas EMA com 20 e 50 períodos. São mostrados 2 momentos de baixa e 3 de alta. Essas tendências são realizadas pelo cruzamento dos sinais das EMAs. Quando $EMA_{50} > EMA_{20}$ existe a tendência de baixa e se $EMA_{50} < EMA_{20}$ tem-se a tendência de alta.



Figura 11 – Média Móvel Exponencial Dupla para o ativo PETR4 usando 20 e 50 períodos.

3.1.10 Média Móvel Adaptativa

A Média Móvel Adaptativa (caracterizada pelo acrônimo AMA, da expressão em inglês *Adaptive Moving Average*), também chamada de Média Móvel Adaptativa de Kaufman (caracterizada pelo acrônimo KAMA, da expressão em inglês *Kaufman Adaptive Moving Average*), foi proposta inicialmente em (KAUFMAN, 1995). Esta análise trabalha com uma MA adaptativa, visto que seu cálculo tem como base o mesmo cálculo da EMA. Contudo, em vez de ser definido um peso fixo, como na EMA, a AMA possui constantes que se adaptam de acordo com parâmetros. Os parâmetros são determinados de acordo com o desejo do usuário que fará a análise do gráfico. A Equação 13 representa o cálculo da AMA.

$$AMA_k = AMA_{Anterior} + \alpha * (x - AMA_{Anterior}) \quad (13)$$

Em (KAUFMAN, 1995) é sugerido por padrão a análise com 10 períodos. Para o cálculo de α temos a Equação 14.

$$\alpha = ([VI * (FC - SC)] + SC)^2 \quad (14)$$

No cálculo de α a variável VI representa uma medida de volatilidade ou força da tendência. (KAUFMAN, 1995) sugere que VI seja igual ao Índice de Eficiência (denotado pelo acrônimo ER, da expressão em inglês *Efficiency Ratio*), sendo este parâmetro calculado pela Equação 15.

$$ER = \frac{|x_n - x_1|}{\frac{n-1}{n * \sum_{i=1}^{n-1} |x_i - x_n|}} \quad (15)$$

As variáveis SC e FC são dadas pelas Equações 16a e 16b.

$$SC = 2/(SN + 1) \quad (16a)$$

$$FC = 2/(FN + 1) \quad (16b)$$

De modo geral os atributos SN e FN podem variar de acordo com o desejo de quem está realizando a análise. No entanto, (KAUFMAN, 1995) determina que $SN > FN$ e sugere $SN = 30$ e $FN = 2$.

3.1.11 Linha de Acumulação Distribuição

A Linha de Acumulação Distribuição (caracterizada pelo acrônimo ADL, da expressão em inglês *Accumulation Distribution Line*) é um indicador baseado no volume. São necessários três cálculos para encontrar o ADL, como é mostrado nas Equações 17a, 17b e 17c (ACHELIS, 2006).

$$MFM_i = \frac{(x_i^{Fechamento} - x_i^{Menor}) - (x_i^{Maior} - x_i^{Fechamento})}{x_i^{Maior} - x_i^{Menor}} \quad (17a)$$

$$MFV_i = MFM_i * x_i^{Volume} \quad (17b)$$

$$ADL_i = ADL_{i-1} + MFV_i \quad (17c)$$

Para calcular a ADL primeiro tem-se o MFM_i , que corresponde ao Multiplicador do Fluxo de Dinheiro (denotado pelo acrônimo MFM, da expressão em inglês *Money Flow Multiplier*) para o período i . Posteriormente, tem-se o MFV_i que é o Volume do Fluxo de Dinheiro (denotado pelo acrônimo MFV, da expressão em inglês *Money Flow Volume*) e está associado ao volume de negócios no período i . Finalmente, a ADL do período i é dada pelo ADL do período anterior mais o MFV_i .

3.1.12 Oscilador de Chaikin

O Oscilador de Chaikin consiste na diferença entre a EMA_3 de uma ADL e da EMA_{10} de uma ADL. Este indicador é projetado para antecipar momentos de inversão de preços na ADL mensurando o início do movimento. O Oscilador de Chaikin gera sinais acima e abaixo de zero, sendo isso os indicadores de *bullish* e *bearish*³ (ACHELIS, 2006).

Para calcular o Oscilador de Chaikin o primeiro passo é realizar o cálculo da ADL de cada período. Em seguida, é feita a diferença da EMA_3 da ADL e da EMA_{10} da ADL, sendo este o Oscilador de Chaikin. A Equação 18 mostra o cálculo desta análise.

$$\text{Oscilador de Chaikin} = EMA_3 \text{ da ADL} - EMA_{10} \text{ da ADL} \quad (18)$$

A Oscilador de Chaikin se torna positivo quando a EMA_3 se move acima da EMA_{10} . Inversamente, o indicador se torna negativo quando a EMA_3 está abaixo da EMA_{10} .

3.1.13 Linha Acumulação/Distribuição de Chaikin

Este é um indicador com base no volume de negócios e mede o fluxo cumulativo de dinheiro para um ativo. A Linha Acumulação/Distribuição de Chaikin (caracterizada pelo acrônimo *Chaikin A/D Line*, da expressão em inglês *Chaikin Accumulation Distribution Line*) é calculada multiplicando o volume do período específico com um multiplicador que é baseado na relação entre o preço de fechamento de alta e baixa. Este indicador pode ser usado para afirmar uma tendência subjacente ou para prever reversões. As Equações 19a, 19b e 19c, expressam esta análise.

$$MFM_i = \frac{(x_i - x_i^{Low}) - (x_i^{High} - x_i)}{(x_i^{High} - x_i^{Low})} \quad (19a)$$

$$MFV_i = V_i * MFM_i \quad (19b)$$

$$CADL_i = CADL_{i-1} + MFV_i \quad (19c)$$

Nas Equações em 19 o cálculo é feito com base em um período i . A variável MFM_i é um multiplicador usado para calcular a *Chaikin A/D Line* e tem como base o preço de fechamento (x_i), o maior preço (x_i^{High}) e o menor preço (x_i^{Low}). A MFV_i é o produto do volume e da MFM_i . Finalmente, a *Chaikin A/D Line*, representada por $CADL_i$, é calculada pela *Chaikin A/D Line* do período anterior mais a MFV_i .

3.2 Correlação de Ativos

Os preços dos ativos estão diretamente relacionados com os resultados financeiros das empresas. Fatores particulares como resultados nacionais e globais da economia tam-

³ *Bullish* e *Bearish* mensuram condições para a ocorrência de momentos de *overbought* e *oversold*.

bém determinam padrões de variações dos ativos. O estudo da correlação entre ativos possibilita a confirmação de resultados e avaliação das variações dos preços.

A variação dos preços de empresas são influenciadas umas pelas outras. Uma vez que essas variações dependem de vários fatores, é muito difícil determinar equações e prever a evolução dos preços futuros. Deste modo, a correlação entre ativos e variações dos preços são focos de estudos em (KIM; KIM; LEE, 2002), (KIM et al., 2001) e (KWON; CHOI; MOON, 2005).

Um meio para realizar a verificação das correlações é através do cálculo do coeficiente de correlação. O trabalho de (KIM; KIM; LEE, 2002) tem como base um valor de retorno, sendo este valor expresso pela Equação 20.

$$S_i(t) \equiv \ln x_i(t + \Delta t) - \ln x_i(t) \quad (20)$$

Considerando $x_i(t)$ o preço do ativo i no tempo t e Δt um intervalo de tempo, é realizado o cálculo do retorno de um ativo. Para realizar a avaliação da correlação entre os ativos i e j temos a Equação 21.

$$C_{i,j} \equiv \frac{\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle}{\sqrt{(\langle S_i^2 \rangle - \langle S_i \rangle^2)(\langle S_j^2 \rangle - \langle S_j \rangle^2)}} \quad (21)$$

O termo $\langle - \rangle$ correspondem ao valor médio sobre um dado período de tempo. O valor do coeficiente de correção é dado pelo intervalo $-1 \leq C_{i,j} \leq 1$. Valores próximos de -1 indicam uma correlação inversa, próximos de 1 expressam correlação direta e 0 determina que é baixa a correlação.

A partir do cálculo do coeficiente de correlação para um conjunto de ativos é montada a matriz de correlação. Essa matriz permite identificar quais ativos estão mais correlacionados. A Tabela 3 mostra um exemplo de Matriz de Correlação entre os ativos AELP3, BIOM3, PETR3 e IDVL4 para o período de 02/01/2012 a 30/12/2013.

Tabela 3 – Matriz de correlação dos ativos AELP3, BIOM3, PETR3 e IDVL4 para o período de 02/01/2012 a 30/12/2013.

	AELP3	PETR3	IDVL4	BIOM3
AELP3	1	0.8339	0.3869	-0.9469
PETR3	0.8339	1	0.218	-0.8146
IDVL4	0.3869	0.218	1	-0.389
BIOM3	-0.9469	-0.8146	-0.389	1

A Tabela 3 mostra que os ativos AELP3 e PETR3 estão diretamente correlacionados. Para AELP3 e BIOM3 tem-se que estão inversamente relacionados. No caso do par AELP3 e IDVL4 a correlação é muito baixa, se aproximando de 0. Os “comportamentos” dos preços destes ativos são mostrados na Figura 12. Este gráfico foi montado considerando a variação de preços normalizados entre 0 e 1. Neste gráfico é demonstrado

como o coeficiente de correlação tenta representar que os preços dos ativos possuem comportamentos similares e ou divergentes. Os coeficientes de AELP3 em relação a BIOM3, PETR3 e IDVL4 são, respectivamente, 0.8339, 0.3869 e -0.9469. Deste modo, AELP3 acompanha PETR3, já que realizaram um comportamento de queda constante durante os períodos. Entre AELP3 e BIOM3 existe um comportamento inverso, já que ocorre a queda dos preços para AELP3 e alta em BIOM3. Finalmente, em AELP3 e IDVL4 existem momentos de alta e baixa aleatórios e não tendo uma clara correlação.



Figura 12 – Preços de fechamento normalizados dos ativos AELP3, BIOM3, PETR3 e IDVL4 para o período de 02/01/2012 até 30/12/2013.

Outra forma de avaliar a correlação entre ativos apresentada por (KIM; KIM; LEE, 2002) consiste em utilizar o retorno relativo. Este retorno relativo é expresso pela Equação 22.

$$G_i(t) = S_i(t) - \frac{1}{N} \sum_i S_i(t) \quad (22)$$

$G_i(t)$ indica o retorno relativo do ativo i em relação a média dos retornos de um conjunto de N ativos. Utilizando $G_i(t)$ é realizado o cálculo do coeficiente de correlação relativa $w_{i,j}$ entre dois ativos i e j , conforme colocada na Equação 23.

$$w_{i,j} \equiv \frac{\langle G_i G_j \rangle - \langle G_i \rangle \langle G_j \rangle}{\sqrt{(\langle G_i^2 \rangle - \langle G_i \rangle^2)(\langle G_j^2 \rangle - \langle G_j \rangle^2)}} \quad (23)$$

3.3 Algoritmos Genéticos (AG)

Esta teoria idealiza a competição de indivíduos por recursos limitados. Os indivíduos que não obtêm êxito tendem a ter uma prole menor e consequentemente menos descen-

dentes diminuindo dessa forma a probabilidade de propagação genética. A combinação gênica dos indivíduos que sobreviveram por meio da reprodução tende a gerar um novo indivíduo com melhores características para se adaptar ao seu meio ambiente. Portanto, um indivíduo mais adaptado terá mais probabilidade de propagar seus genes criando um ciclo evolutivo contínuo.

Na computação os AGs foram elaborados em 1940 quando alguns cientistas trabalharam em modelos bio inspirados para produzir o ramo da IA. Entretanto começaram a ser utilizados apenas na década de 60, quando John Holland estudou meteticulosamente o processo de seleção natural propondo modelos computacionais heurísticos para procura de soluções.

Inicialmente temos a criação de uma população, que corresponde a um grupo de indivíduos onde cada indivíduo é uma possível solução do problema. Para cada indivíduo é dado um valor que define o quão apto ele é, em outras palavras, a aptidão determina se um indivíduo é melhor ou pior que outros indivíduos. São realizadas evoluções e novos indivíduos são criados a partir da população existente. A criação de novos indivíduos é feita usando os processos de crossover e mutação. O crossover é realizado pela troca de material de dois ou mais indivíduos. A mutação é a alteração aleatório de parte do gene do indivíduo. Ao final os piores indivíduos são eliminados, mantendo assim as melhores soluções. A Figura 13 mostra o fluxo dos AGs.

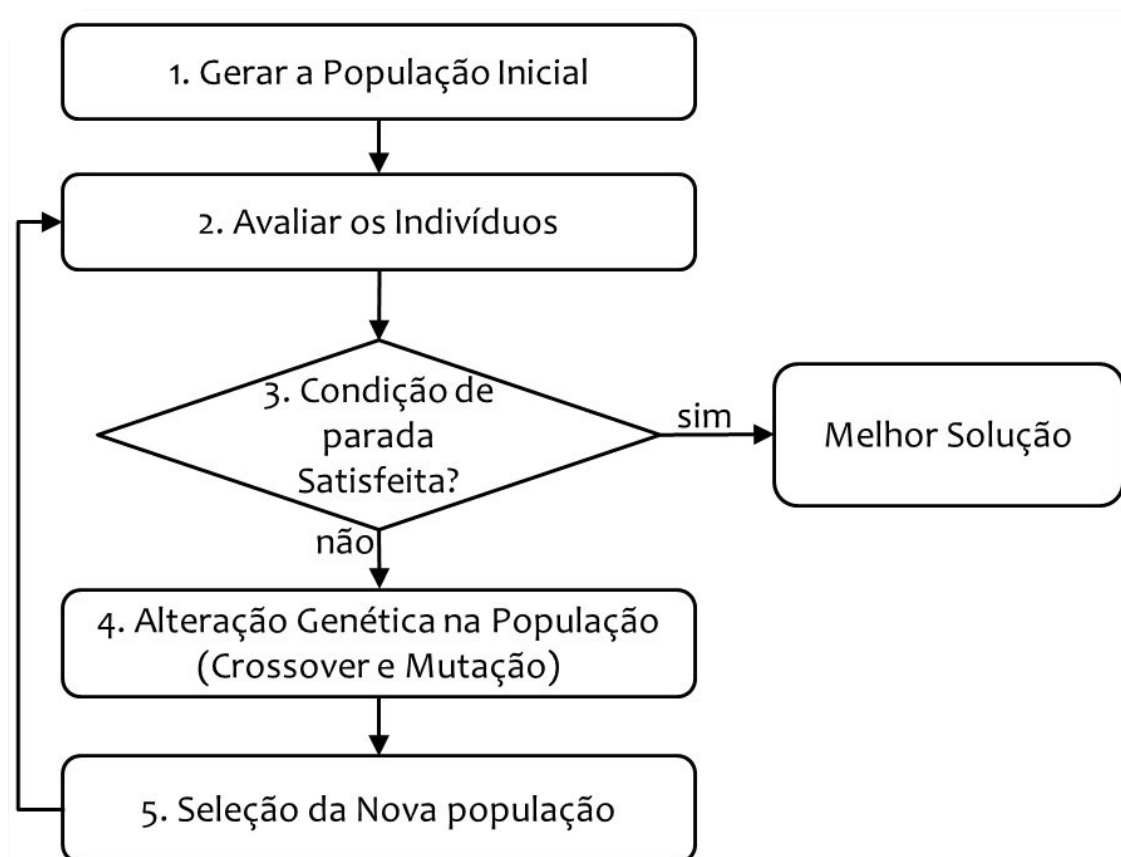


Figura 13 – Fluxo de execução dos AGs.

Basicamente, o objetivo da evolução é que com o decorrer do tempo a população evolua e que o melhor indivíduo seja a solução.

Planejamento

Neste capítulo se descreve conceitos relacionados a planejamento. A abordagem clássica de planejamento é apresentada para em seguida se descrever uma abordagem não clássica que faz uso de probabilidades para caracterizar incertezas na especificação de estados e efeitos não determinísticos das ações. Esta última abordagem é conhecida como planejamento probabilístico e é com base nesta abordagem que esta tese foi desenvolvida.

4.1 Planejamento Clássico

O planejamento clássico pode ser definido por um estado inicial, uma meta a ser atingida e um conjunto de ações. Estes três parâmetros devem ser definidos em uma linguagem formal que possibilite a representação do domínio do problema. A solução consiste em gerar automaticamente um conjunto de ações que ao serem executadas satisfaçam a meta. Este conjunto de ações também é chamado de plano. O planejamento clássico, também chamado de planejamento determinístico, é conhecido quando o plano é executado é o resultado final é exatamente o desejado. O ambiente é estático e as ações têm exatamente o comportamento esperado (BRANQUINHO, 2009) (RUSSELL; NORVIG, 2009) (ALVES; LOPES; BRANQUINHO, 2013).

Os planejadores consistem em ferramentas que resolvem problemas de planejamento. Existem muitos planejadores para o planejamento clássico, tais como o LAMA (RICHTER; WESTPHAL, 2010), *Fast Forward* (HOFFMANN; NEBEL, 2001) e Graphplan (BLUM; FURST, 1997). Formalmente este planejamento é dado pela Definição 4.1.1.

Definição 4.1.1 O problema de planejamento clássico (determinístico) é definido formalmente como uma 5-upla $\langle S, A, s_i, s_g, P \rangle$, onde:

S é o conjunto finito de estados.

A é o conjunto finito de ações.

s_i é o estado inicial, tal que $s_i \in S$.

s_g é um estado meta, tal que $s_g \in S$.

P representa o plano, sendo formado por um conjunto ordenado e finito de ações.

O estado é representado por um conjunto de proposições. Uma característica importante considerada durante o planejamento é a hipótese de mundo fechado. Isso significa que quaisquer proposições não mencionadas em um estado são consideradas falsas (RUSSELL; NORVIG, 2003). Sendo assim, os estados s_i e s_g são representados por proposições. As ações são colocadas em termos de pré-condições que devem ser satisfeitas antes de ser executada e dos efeitos resultantes de sua execução.

Um componente muito importante é a linguagem de representação. Entre as linguagens mais tradicionais temos *STanford Research Institute Problem Solver* (STRIPS) (FIKES; NILSSON, 1971), *Action Description Language* (ADL) (PEDNAULT, 1989) e *Planning Domain Definition Language* (PDDL) (MCDERMOTT et al., 1998).

O “Mundo dos Blocos” é um problema clássico de planejamento e consiste em mover os blocos usando as ações **Mover** e **MoverParaMesa**, tendo como objetivo organizar os blocos de modo a alcançar uma certa configuração. A Figura 14 mostra este problema em STRIPS (BRANQUINHO, 2009). STRIPS é composto por um conjunto de átomos e as ações formadas por operadores com pré-condições e efeitos. As ações removem e adicionam átomos de estados.

PDDL é uma linguagem formada por objetos, predicados ou preposições, estados e ações. Em PDDL o problema é separado em dois elementos, sendo o domínio e o problema. O *Domínio* descreve os predicados, ações e classes de objetos. Já o *Problema* compõe os objetos, o estado inicial e a especificação da meta do problema. Estes dois elementos são exibidos na Figura 15.

Em PDDL o problema “Mundo dos Blocos” é modelado conforme a Figura 16. Normalmente cada um dos elementos da PDDL, o domínio e o problema, são colocados em arquivos distintos.

Uma das principais competições de planejamento, a *International Planning Competition* (IPC), utiliza tais linguagens em suas competições. A evolução das linguagens da IPC, montada por Coles em (COLES et al., 2012), é mostrada na Figura 17. É interessante observar que as linguagens da IPC evoluíram junto das necessidades da competição.

4.2 Planejamento Probabilístico

O planejamento clássico não aborda certos tipos de problemas. Um destes casos ocorre quando as ações têm efeitos não determinísticos e/ou os estados são incertos. Em essência, este tipo de problema é um planejamento sob incertezas, sendo visto como um processo de decisão onde, a cada passo, o agente observa o estado atual do ambiente e escolhe a ação que maximiza a expectativa de se alcançar um estado desejado. Logo, o

Estado Inicial ($Sobre(A, Mesa) \wedge Sobre(B, Mesa) \wedge Sobre(C, A) \wedge$
 $Bloco(A) \wedge Bloco(B) \wedge Bloco(C) \wedge Livre(B) \wedge Livre(C) \wedge Livre(mesa)$)
Estado Meta ($Sobre(A, B) \wedge Sobre(B, C) \wedge Sobre(C, mesa)$)
Ação ($Mover(b, x, y)$,
 pré-condição: $Sobre(b, x) \wedge Livre(b) \wedge Livre(y) \wedge Bloco(b) \wedge Bloco(y) \wedge$
 $(b \neq x) \wedge (b \neq y) \wedge (x \neq y)$
 efeito: $Sobre(b, y) \wedge Livre(x) \wedge \neg Sobre(b, x) \wedge \neg Livre(y)$)
Ação ($MoverParaMesa(b, x)$,
 pré-condição: $Sobre(b, x) \wedge Livre(b) \wedge Bloco(b) \wedge Bloco(x) \wedge (b \neq x)$
 efeito: $Sobre(b, mesa) \wedge Livre(x) \wedge \neg Sobre(b, x)$)

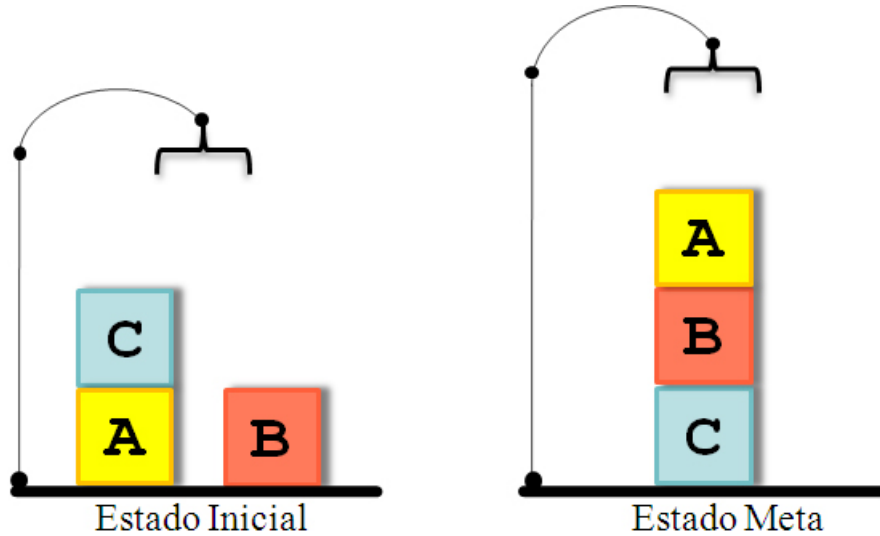


Figura 14 – “Mundo dos Blocos” representado em STRIPS.

ESTRUTURA DO DOMÍNIO

```
(define (domain <nome do domínio>)
  (:requirements <requisitos da representação>)
  (:predicates <código para predicados>)
  (:action <nome da 1ª ação>
    :parameters (<lista de parâmetros>)
    :precondition (<lista de pré-condições>)
    :effect (<lista de efeitos>))
  ...
  (:action <nome da nª ação>
    :parameters (<lista de parâmetros>)
    :precondition (<lista de pré-condições>)
    :effect (<lista de efeitos>))
)
```

ESTRUTURA DO PROBLEMA

```
(define (problem <nome do problema>)
  (:domain <nome do domínio>)
  (:objects <código para objetos>)
  (:init <código para o estado inicial>)
  (:goal <código para a especificação do estado meta>)
)
```

Figura 15 – Estruturas utilizados pela linguagem PDDL.

```

(define (domain MundoDosBlocos)
  (:requirements :strips)
  (:predicates (em ?x ?y) (naMesa ?x) (livre ?x) (bloco ?x))
  (:action Mover
    :parameters (?b ?x ?y)
    :precondition (and (em ?b ?x) (livre ?b) (livre ?y) (bloco ?b) (bloco ?y))
    :effect (and (livre ?b ?y) (livre ?x) (not (em ?b ?x)) (not (livre ?y))))
  (:action MoverParaMesa
    :parameters (?b ?x)
    :precondition (and (em ?b ?x) (livre ?b) (bloco ?b) (bloco ?x))
    :effect (and (naMesa ?b) (livre ?x) (not (em ?b ?x))))
)
(define (problem Blocos-1)
  (:domain MundoDosBlocos)
  (:objects A B C)
  (:init (livre B) (livre C) (naMesa A) (naMesa B) (em A C)
    (bloco A) (bloco B) (bloco C))
  (:goal (and (em C B) (em B A) (naMesa A)))
)

```

Figura 16 – Representação do problema “Mundo dos Blocos” na linguagem PDDL.

planejamento probabilístico é uma extensão do planejamento não determinístico de modo que são associadas probabilidades aos eventos.

A representação dos problemas de planejamento probabilístico pode ser realizada com diferentes linguagens. Este trabalho utiliza a linguagem RDDL, mais detalhes são apresentados na Seção 4.2.1. Nas demais seções são apresentados dois tipos de planejamento probabilístico, a Seção 4.2.2 apresenta o MDP e a Seção 4.2.3 descreve o POMDP.

4.2.1 Relational Dynamic Influence Diagram Language (RDDL)

Para o portfólio de planejamento é usada a *Relational Dynamic Influence Diagram Language* (RDDL) (SANNER, 2010). A RDDL é uma linguagem onde os estados, ações e observações são parametrizados como variáveis, já as probabilidades de transições são expressas como funções. A parametrização em forma de variáveis é simplesmente um modelo para expressar as definições dos objetos de um problema, além disso, esta representação permite definir recompensas as transições de estados e observações. Outra característica relevante é que os estados podem ser expressos como sendo observáveis ou parcialmente observáveis e determinísticos ou não determinísticos. Tais características são importantes para permitir representar diferentes tipos de problemas. Finalmente, o planejamento é especificado via funções estocásticas aplicadas no estado corrente.

A RDDL é composto por três partes, sendo a *domain*, onde é detalhado o domínio; *non-fluent* que apresenta a instanciação de não-fluentes, como estruturas fixas entre os problemas; *instance* que define a descrição do problema, o estado inicial, o número de ações concorrentes e outras características da instância do problema. Geralmente estes

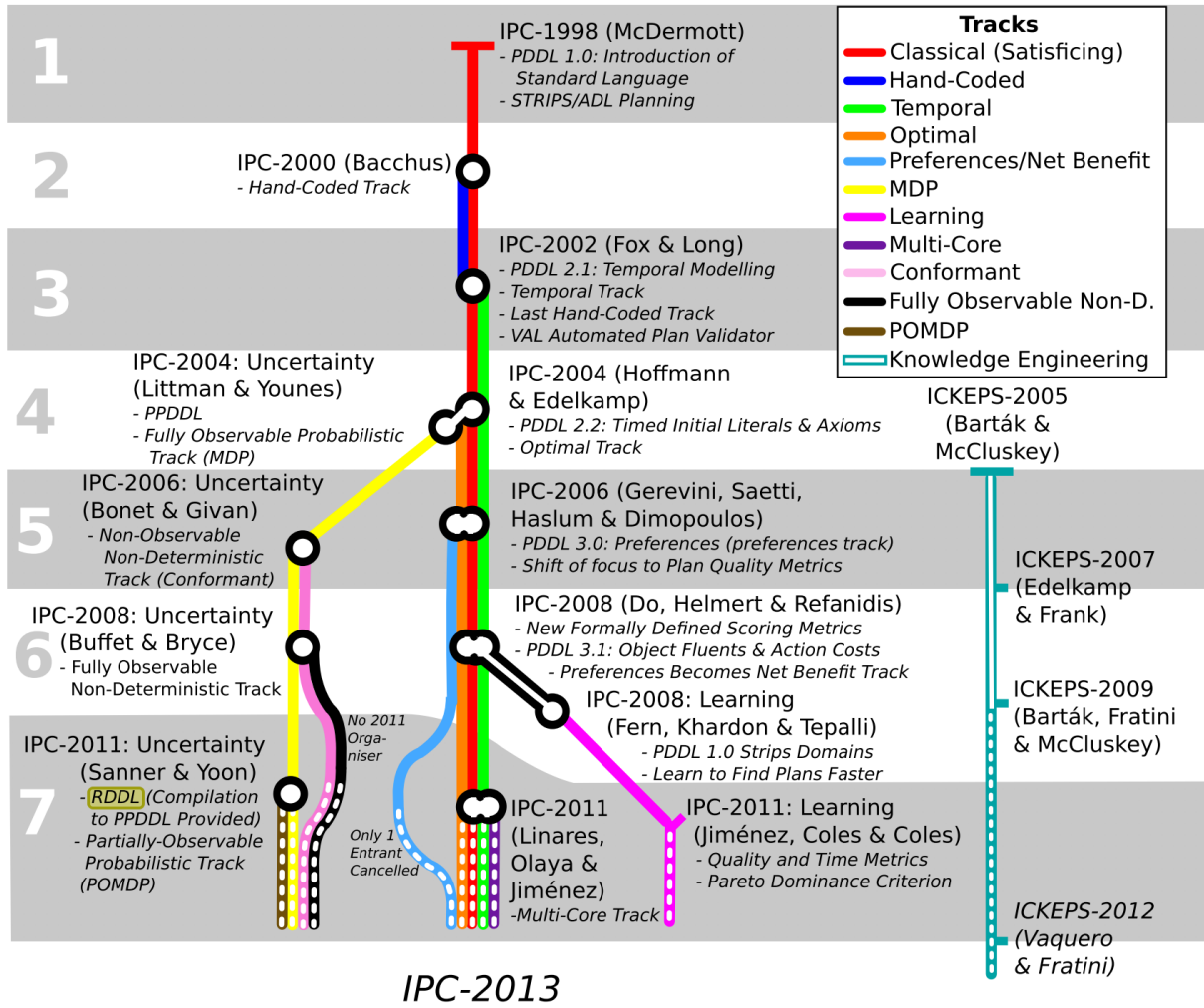


Figura 17 – Histórico da Competição Internacional de Planejamento (IPC) (COLES et al., 2012).

elementos são separados em diferentes arquivos.

A seção *domain* é composta pela estrutura mostrada na Figura 18. O primeiro valor é o nome do domínio. Em seguida, em *requirements*, são colocadas as exigências de linguagem que o parser RDDDL deve possuir para interpretar o arquivo, sendo possível especificar: **reward-deterministic**, **integer-valued**, **continuous**, **multivalued**, **intermediate-nodes**, **constrained-state**, **concurrent**, **cpf-deterministic** e **partially-observed**.

Os *types* estabelecem tipos de dados e variáveis do problema, podendo ser *object* ou *enumerated*. Na seção *pvariables* são listadas as variáveis e o tipo de *fluent* que ela pode ser. As variáveis podem assumir os seguintes *fluents*: **non-fluent**, que apresenta operadores que representam fatos estáticos; **state-fluent**, que são operadores que representam fatos dinâmicos; e **action-fluent**, que são operadores que controlam o comportamento das funções. Além disso, cada variável está associada a um tipo, podem ser **integer**, **bool**, **real**, **object** e **enumerated**. Também é colocado um valor padrão (*default*) para as variáveis. Posteriormente, são apresentadas as CPFs (*conditional probabilistic func-*

```

domain <nome_domínio> {
  requirements = {
    <requisito1>,
    ...,
    <requisiton>
  };
  types {
    <nome_tipo> : <tipo_fluent>;
  };
  pvariables {
    <nome_variável1> : {
      <tipo_variável>, <tipo_dados>, default=<valor_padrão>;
    };
    ...;
    <nome_variávelk> : {
      <tipo_variável>, <tipo_dados>, default=<valor_padrão>;
    };
  };
  cpfs {
    <nome_cpf> (parâmetros) = <função>;
  };
  state-action-constraints {
    <restrição1>;
    ...;
    <restriçãom>;
  };
  reward = <função>;
}

```

Figura 18 – Estrutura do domínio de problemas representados com RDDDL.

tions - funções condicionais probabilísticas), que são as funções a serem aplicadas no problema. Já em *state-action-constraints* restrições que as ações possuem e devem ser respeitadas para que sejam executadas. Por último, em *reward*, são definidas as funções de recompensas.

A parte *non-fluent* apresenta a instanciação de não-fluentes, tais como estruturas fixas entre problemas. A estrutura básica deste elemento é apresentada na Figura 19. Inicialmente é colocado o nome do *non-fluent* e posteriormente o domínio ao qual pertence. Em seguida tem-se os *objects*, onde são listados todos os objetos que serão manipulados. Subsequentemente, na subseção *non-fluents*, é apresentada a lista de fatos estáticos. A seção *non-fluents* é opcional e pode ser representada dentro do *instance*.

Por último, a estrutura da seção *instance* é mostrada na Figura 20. No início é especificado o nome da instância, o nome do domínio e o nome do *non-fluent*. Caso o *non-fluent* não seja especificado, a seção *objects* deve ser inserida antes da *init-state*. Na *init-state* são apresentados os fatos dinâmicos e, caso não tenha sido especificada a

```

non-fluent <nome_non-fluent> {
    domain <nome_domínio>;
    objects {
        <tipo_objeto> : { <nome_objeto> };
    }
    non-fluents {
        // Lista de fatos estáticos.
    }
}

```

Figura 19 – Estrutura *non-fluent* de problemas representados com RDDL.

non-fluent, neste local também são declarados os fatos estáticos. Por fim, são declarados três valores: o primeiro valor é o *max-nondef-actions*, que determina o número de ações concorrentes; o segundo valor é o *horizon* e determina o horizonte de ações; a última variável, a *discount*, pondera a média da função recompensa.

```

instance <nome_instância> {
    domain <nome_domínio>;
    non-fluent <nome_non-fluent>;
    init-state {
        // Lista de fatos dinâmicos.
        // Neste local também são colocados os fatos estáticos,
        // caso não tenha sido declarada a seção non-fluent.
    }
    max-nondef-actions = <valor>;
    horizon = <valor>;
    discount = <valor>;
}

```

Figura 20 – Estrutura *instance* de instâncias de problemas representados com RDDL.

O Apêndice C apresenta um exemplo de modelagem para o problema do mercado financeiro utilizando a linguagem RDDL.

4.2.2 Processo de Decisão de Markov

Em um processo de decisão de Markov (denotado pelo acrônimo MDP, da expressão em inglês *Markov Decision Process*) (HOWARD, 1960) completamente observável o agente possui informações que determinam completamente o estado corrente e probabilidades são associadas as ações. Desta forma, faz-se um raciocínio com ações caracterizadas por efeitos estocásticos para que um agente saia de um estado corrente e alcance um estado desejado. Tal raciocínio caracteriza o chamado planejamento probabilístico.

Um processo estocástico é formado por uma sequência de variáveis aleatórias representando estados e uma função de transição que estabelece probabilidades às transi-

ções de estados (HOLGUIN, 2013). Formalmente, a função de transição é dada por $P[s_{t+1}|s_0, s_1, \dots, s_t]$, tal que s_i corresponde ao estado s no tempo i .

A solução de um problema de planejamento probabilístico pode ser obtida com a teoria de processos de decisão markovianos. O processo estocástico é markoviano se o estado futuro depende apenas do estado corrente. Fundamentalmente, não é necessário armazenar o histórico de ações para tomar novas decisões. Posto isto, para este caso tem-se que a função de transição é expressa por $P[s_{t+1}|s_t]$, tal que s_i corresponde ao estado s no tempo i .

Normalmente, além da função de transição, é dado um reforço/recompensa ao agente em cada transição. Estas recompensas podem estar associadas as ações, estados ou ao par ação mais estado. Considerando a função de recompensa R e o estado s , formalmente, a recompensa é expressa por $R(s_t)$, tal que R é a função de recompensa para o estado s_t no tempo t . Um MDP é definido formalmente pela Definição 4.2.1.

Definição 4.2.1 O MDP é definido formalmente como uma 5-upla $\langle S, A, P, R, C \rangle$, onde:

S é o conjunto finito de estados.

A é o conjunto finito de ações.

P é dado por $P_a(s'|s)$ e corresponde a probabilidade de atingir o estado s' ao se executar a ação a no estado s , tal que $a \in A$, $s \in S$ e $s' \in S$. Para qualquer estado s e ação a , se existe s' tal que $P_a(s', s)$, então $\sum_{s' \in S} P_a(s'|s) = 1$.

R é dado por $S \rightarrow \mathfrak{R}$ e consiste em uma função que atribui recompensa a cada estado.

C é dado por $S \rightarrow \mathfrak{R}$ e corresponde a uma função que atribui um custo a cada ação.

O resultado do planejamento é um plano que especifica a ação que deve ser executada em cada estado. O plano é representado por uma política π que mapeia os estados em ações: $\pi : S \rightarrow A$. A política pode ser estacionária, ou seja, as ações são sempre as mesmas independente do tempo. Já a não estacionária as ações mudam durante o tempo. Para conseguir avaliar uma política é preciso antes definir por quanto tempo ela será seguida, podendo ser de *horizonte finito* ou *horizonte infinito*.

Quando temos um *horizonte finito* normalmente a política é não estacionária. Para uma política π e v_n^π correspondendo a soma dos ganhos esperados, a partir do estado s , seguindo a política π por n passos, temos a Equação 24.

$$v^\pi(s) = R(s) - C(\pi(s)) + \sum_{s' \in S} P(s, \pi(s), s') v^\pi(s') \quad (24)$$

Ao tratar do *horizonte infinito*, como são realizados infinitos passos, é mais adequado considerar uma política estacionária. Para garantir que o valor de ganho esperado seja finito temos uma taxa de desconto $0 < \gamma < 1$, sendo dada pela Equação 25.

$$v^\pi(s) = R(s) - C(\pi(s)) + \gamma \sum_{s' \in S} P(s, \pi(s), s') v^\pi(s') \quad (25)$$

Uma política π^* é ótima se maximiza os ganhos esperados v , podendo ser definida pela Equação 26.

$$\pi^*(s) = \arg \max_{a \in A} \{R(s) - C(a) + \gamma \sum_{s' \in S} p(s, a, s') v^*(s')\} \quad (26)$$

Para obter políticas ótimas existem dois algoritmos comuns na literatura, o Iteração de Valor (denotado pelo acrônimo VI, da expressão em inglês *Value Iteration*) (BELLMAN, 1957) e o Iteração de Política (denotado pelo acrônimo PI, da expressão em inglês *Policy Iteration*) (PUTERMAN, 2009).

4.2.3 Processo de Decisão de Markov Parcialmente Observável

Quando um agente possui informações que não permitem determinar com precisão seu estado tem-se um MDP parcialmente observável (denotado pelo acrônimo POMDP, da expressão em inglês *Partially Observable Markov Decision Process*). Essencialmente, no POMDP os estados em que o agente explora e o resultado das ações não são conhecidos. São utilizados estados de crença para resolver este problema e eles representam a probabilidade de estar em um estado.

Quando se trabalha com ambientes parcialmente observáveis as ações são tomadas considerando o estado atual e as probabilidades de alcançar cada um dos próximos estados. As observações, que levam a avaliação das probabilidades, são capturadas através dos sensores do agente. As probabilidades de transição dependem exclusivamente do estado atual e não são influenciadas pelos estados anteriores. Formalmente, o POMDP tem a Definição 4.2.2.

Definição 4.2.2 O POMDP é definido formalmente como uma 7-upla $\langle S, A, T, O, P, C, R \rangle$, onde:

S é o conjunto finito de estados.

A é o conjunto finito de ações.

T é dado por $T_a(s'|s)$ e corresponde a probabilidade de atingir o estado s' ao executar a ação a no estado s , tal que $a \in A$, $s \in S$ e $s' \in S$. Para qualquer estado s e ação a , se existe s' tal que $T_a(s'|s) \neq 0$, então $\sum_{s' \in S} T_a(s'|s) = 1$.

O é o conjunto finito de observações.

P é dado por $P_a(o|s)$ e corresponde a probabilidade de observar o no estado s após executar a ação a , tal que $a \in A$, $o \in O$ e $s \in S$. Para o estado s e ação a , se existe observação o tal que $P_a(o|s) \neq 0$, então $\sum_{o \in O} P_a(o|s) = 1$.

C é dado por $S \times A \rightarrow \mathbb{R}$ e corresponde a uma função que atribui um custo para cada estado e ação.

R é dado por $S \rightarrow \mathbb{R}$ e consiste em uma função que atribui recompensa a cada estado.

As observações O e distribuições de probabilidades P introduzem o modelo de observações parciais. Sendo assim, as informações de um estado são obtidas através de observações. É preciso calcular a distribuição de probabilidade de estar em cada estado, sendo isso chamado de crença. Considerando B o conjunto de todas as crenças, se $b \in B$ é um estado de crença, a probabilidade de estar no estado s é definida por $b(s)$, tal que $0 \leq b(s) \leq 1$ para todo $s \in S$ e $\sum_{s \in S} b(s) = 1$. A Figura 21 mostra como ocorrem as transições dos estados de crença, ações e observações. A raiz é composta pelos estados de crença, enquanto as transições para novos estados de crença ocorrem de acordo com a escolha das ações e das observações.

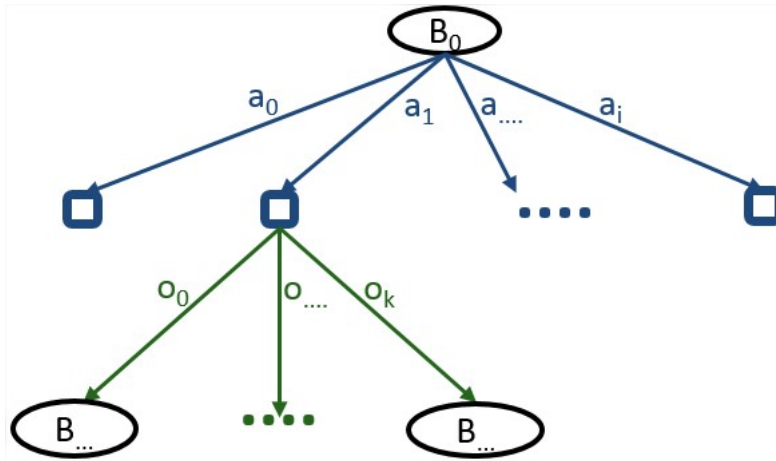


Figura 21 – Árvore de estados de crença para problemas POMDP.

O cálculo do estado de crença $b(s)$ é realizada para uma observação e uma ação, sendo assim expressa pela função $b_a^o(s)$. Portanto, $b_a^o(s)$ corresponde a crença de alcançar o estado s e observar o após executar a ação a . Esta função é representada pela Equação 27.

$$b_a^o(s) = \frac{P_a(o|s)b_a(s)}{b_a(o)} \quad (27)$$

Para calcular $b_a^o(s)$ são requeridas as crenças $b_a(s)$ e $b_a(o)$. A função $b_a(s)$ corresponde a probabilidade de alcançar s ao executar a ação a no estado atual, sendo esta dada pela equação 28.

$$b_a(s) = \sum_{s' \in S} T_a(s|s')b(s') \quad (28)$$

O estado de crença $b_a(o)$ corresponde a probabilidade de observar o após executar a sobre o estado de crença atual. O valor de $b_a(o)$ é obtido pela Equação 29.

$$b_a(o) = \sum_{s \in S} P_a(o|s)b_a(s) \quad (29)$$

O custo de transição do estado de crença b ao executar a ação a é calculada pela Equação 30.

$$C(b, a) = \sum_{s \in S} b(s)C(s, a) \quad (30)$$

A recompensa para um estado de crença b ao executar a ação a é calculada pela Equação 31.

$$\rho(b, a) = \sum_{s \in S} b(s)R(s, a). \quad (31)$$

Com base no custo de transição e na recompensa é calculada a função de utilidade. Esta função corresponde a utilidade de uma crença b ao executar a ação a , sendo assim expressa pela Equação 32.

$$V(b, a) = \sum_{s \in S} b(s)(R(s) - C(s, a)) \quad (32)$$

O problema de planejamento pode ser encarado como um problema de otimização, logo o objetivo consiste em maximizar o benefício esperado para cada estado de crença $E(b)$. Esta otimização pode ser realizada através da equação de Bellman (BAFFA; CIARLINI, 2010a), conforme a Equação 33.

$$E(b) = \max_{a \in A} \{V(b, a) + \gamma \sum_{o \in O} b_a(o)E(b_a^o)\} \quad (33)$$

Uma forma para resolver este problema é dada por algoritmos que convertem o POMDP em MDP com base nos estados de crença em vez do domínio de estados. Uma vez convertido o problema é resolvido através dos algoritmos *Policy Iteration* (PUTERMAN, 2009) ou *Value Iteration* (BELLMAN, 1957). Como o conjunto de estados de crença é infinito e contínuo, computacionalmente esta solução é muito difícil, e consequentemente, é normal recorrer-se a programação dinâmica (PUTERMAN, 1994) para tratar o espaço de estados de crença. Mesmo com o uso de programação dinâmica, políticas ótimas são encontradas apenas para problemas muito pequenos.

4.2.4 Algoritmos Exatos

Para encontrar uma solução exata para POMDP é necessário utilizar uma abordagem como a proposta em (SONDIK, 1973), que utiliza múltiplas iterações de programação dinâmica. Neste caso a função de utilidade E , dada pela Equação 33, representa o benefício esperado a partir de cada estado de crença.

Para resolver este problema de horizonte infinito é dado um fator de desconto e o problema é então resolvido como sendo de horizonte finito até que a contribuição das parcelas futuras convirja para zero. Pode-se calcular o benefício de um estado de crença no caso de um horizonte finito de t transições, utilizando-se os valores dos benefícios para horizontes de $t - 1$ transições conforme a Equação 34.

$$E_t(b) = \begin{cases} \max_{a \in A} \{V(b, a)\} & t = 0 \\ \max_{a \in A} \{V(b, a) + \gamma \sum_{o \in O} b_a(o) E_{t-1}(b_a^o)\} & t > 0 \end{cases} \quad (34)$$

A solução resulta de um problema de programação dinâmica sobre o espaço de estados de crença. Conforme demonstrado em (SONDIK, 1973), a solução decorre do fato de ser uma função linear em trechos e convexa, sendo expressa pela Equação 35.

$$E_t(b) = \max_{\alpha \in \Gamma_t} \left\{ \sum_{s \in S} \alpha(s) b(s) \right\} \quad (35)$$

O valor no horizonte t é obtido a partir de um conjunto de funções $\Gamma_t = \{\alpha_0, \alpha_1, \dots, \alpha_m\}$, sendo que cada α corresponde a função em que está associado um valor de benefício para cada um dos estados. Este também pode ser tratado como um espaço vetorial de $|S|$ dimensões, onde cada posição contém um valor de benefício para um estado.

Para cada vetor α é associada uma ação a que corresponde a uma região convexa do espaço de estados de crença. Tal região é delimitada por hiperplanos e o benefício gerado com a ação a pode ser calculada através do produto escalar de α pelo estado de crença. Deste modo é possível adaptar a função utilidade original para calcular a melhor ação a para um estado de crença b , um horizonte t e o vetor α , sendo expresso pela Equação 36.

$$E_t(b) = \max_{a \in A} \left\{ V(b, a) + \gamma \sum_{o \in O} b_a(o) \max_{\alpha \in \Gamma_{t-1}} \left\{ \sum_{s \in S} \alpha(s) b_a^o(s) \right\} \right\} \quad (36)$$

Com base nesta equação, temos que para cada estado de crença b , Γ_t de conter pelo menos um vetor, associado a uma ação a , que permite calcular o benefício esperado a partir de b . Vários vetores podem estar associados a uma mesma ação, contudo, cada vetor está relacionado a apenas uma ação.

É preciso definir o conjunto de vetores Γ_t^a para cada ação a de modo que seja possível realizar o cálculo do benefício esperado ao executar tal ação em qualquer estado de crença. Para realizar o cálculo de Γ_t^a é necessário criar os conjuntos $\Gamma_t^{a,*}$ para cada par ação e observação.

O conjunto $\Gamma_t^{a,o}$ contém os vetores que contabilizam os benefícios a partir dos estados subsequentes, sendo expresso pela Equação 37.

$$\Gamma_t^{a,o} = \left\{ \left\{ \gamma \sum_{s' \in S} T_a(s'|s) P_a(o|s') \alpha_i(s') \mid \forall s \in S \right\} \mid \forall \alpha_i \in \Gamma_{t-1} \right\} \quad (37)$$

Posteriormente, é preciso criar o conjunto Γ_t^a a partir da soma de $\Gamma_t^{a,*}$ pela soma cruzada (representado por \oplus) de cada $\Gamma_t^{a,o}$ para todas as observações $o \in O$, conforme a Equação 38.

$$\Gamma_t^a = \begin{cases} \Gamma_t^{a,*} & t = 0 \\ \Gamma_t^{a,*} + \Gamma_t^{a,o_1} \oplus \Gamma_t^{a,o_2} \oplus \dots & t > 0 \end{cases} \quad (38)$$

Por último é realizada a soma dos conjuntos Γ_t^a para todas as ações $a \in A$, sendo isto dada pela Equação 39.

$$\Gamma_t = \{\Gamma_t^a \mid \forall a \in A\} \quad (39)$$

Ao final temos a função utilidade para todos os estados de crença.

4.2.5 Algoritmos Aproximados

Mesmo para pequenos problemas a solução da Seção 4.2.4 é inviável, uma vez que explora todos os estados de crença. Por este motivo algoritmos que geram soluções aproximadas são mais utilizados. Estes algoritmos têm os benefícios de consumir menos memória e encontram soluções em menor tempo. Para isso estes algoritmos realizam atualizações em conjuntos pequenos e em específicos estados de crença.

As próximas seções apresentam algoritmos aproximados para resolver POMDP. Na Seção 4.2.5.1 é apresentado o algoritmo PBVI. Já a Seção 4.2.5.2 descreve o SARSOP. Por último, é mostrado o MCVI na Seção 4.2.5.3.

4.2.5.1 PBVI

O PBVI (da expressão em inglês Point-Base Value Iteration) (PINEAU; GORDON; THRUN, 2003) consiste em buscar uma solução aproximada através da definição de um método de backup exato para cada estado de crença, em outras palavras, apenas um vetor α é mantido. O método de backup retorna um vetor que passa a ser válido para a região ao redor da solução do estado de crença b . Sendo assim, os outros estados de crença similares a b (ou seja, que estão na mesma região) passam a assumir que têm características semelhantes. Estas características podem ser a ação escolhida que direciona para a mesma equação E_{t-1} e o estado de crença futuro b' .

É mostrado no Algoritmo 1 o pseudocódigo do PBVI (PINEAU; GORDON; THRUN, 2006). É dado como parâmetros de entrada o estado de crença $B_{Inicial}$, o conjunto inicial

de soluções Γ_0 , o número desejado de expansões N e o horizonte de planejamento T . Durante N iterações são realizadas expansões do conjunto de estados de crença, ele realiza o cálculo das utilidades e é gerada a política ótima para um horizonte T . O algoritmo finaliza após N expansões.

Algoritmo 1 PBVI($B_{Inicial}, \Gamma_0, N, T$)

```

B  $\leftarrow$  Binicial
 $\Gamma \leftarrow \Gamma_0$ 
for  $N$  expansões do
  for  $T$  iterações do
    |  $\Gamma \leftarrow \text{BACKUP}(B, \Gamma)$ 
  end
  BNovo  $\leftarrow$  PBVI_EXPANDIR( $B, \Gamma$ )
  B  $\leftarrow$  B  $\cup$  BNovo
end
return  $\Gamma$ 

```

Neste contexto, o PBVI trabalha com duas funções auxiliares, a BACKUP e a PBVI_EXPANDIR. O Algoritmo 2 mostra a função BACKUP, sendo ele responsável pelo o cálculo da função utilidade para um conjunto de crenças B e um conjunto de soluções atuais Γ . Vale ressaltar que o BACKUP é normalmente usado por outros algoritmos POMDP.

Algoritmo 2 BACKUP(B, Γ_{t-1})

```

foreach  $a \in A$  do
  foreach  $o \in O$  do
    foreach  $\alpha_t \in \Gamma_{t-1}$  do
      |  $\alpha_t^{a,o}(s) \leftarrow \gamma \sum_{s' \in S} T_a(s'|s) P_a(o|s') \alpha_t(s'), \forall s \in S$ 
    end
     $\Gamma_t^{a,o} \leftarrow \Gamma_t^{a,o} \cup \{\alpha_t^{a,o}\}$ 
  end
end
 $\Gamma_t \leftarrow \emptyset$ 
foreach  $b \in B$  do
  |  $\alpha b \leftarrow \max_{a \in A} [\sum_{s \in S} b(s) V(b, a) + \sum_{o \in O} \max_{\alpha \in \Gamma_t^{a,o}} \{\sum_{s \in S} \alpha(s) b(s)\}]$ 
  | if  $\alpha b \notin \Gamma_t$  then
  | |  $\Gamma_t \leftarrow \Gamma_t \cup \{\alpha b\}$ 
  | end
end
return  $\Gamma_t$ 

```

Existem variações do método PBVI_EXPANDIR, sendo as mais comuns o *Random Action*, *Stochastic Simulation with Random Action*, *Stochastic Simulation with Greedy Action* e *Stochastic Simulation with Exploratory Action* (PINEAU; GORDON; THRUN, 2006). No Algoritmo 3 é mostrado o método PBVI_EXPANDIR baseado na implementação *Random Action*. A *Random Action* é uma estratégia simples que seleciona

aleatoriamente estados de crença. Ela consiste em duplicar a quantidade de estados de crença de B e gerar estados de crença aleatoriamente.

Algoritmo 3 PBVI_EXPANDIR_{ra}(B, Γ)

```

 $B_{novo} \leftarrow B$ 
 $\Gamma \leftarrow \Gamma_0$ 
foreach  $b \in B$  do
  |  $S \leftarrow$  Número de Estados
  | for  $i \leftarrow 0$  to  $S$  do
  | |  $b_{temporário}[i] \leftarrow aleatorio_{uniforme}(0, 1)$ 
  | end
  | Ordenar Ascendente  $b_{temporário}$ 
  | for  $i \leftarrow 1$  until  $S$  do
  | |  $b_{novo} \leftarrow b_{temporário}[i + 1] - b_{temporário}[i]$ 
  | end
  |  $B_{novo} \leftarrow B_{novo} \cup b_{novo}$ 
end
return  $B_{novo}$ 

```

4.2.5.2 SARSOP

A ideia do algoritmo SARSOP (da expressão em inglês *Successive Approximation of Reachable Space under Optimal Policy*) (KURNIAWATI; HSU; LEE, 2008) é de se aproximar dos espaços ótimos alcançáveis através de amostragem. Da mesma forma que outros algoritmos baseados em pontos, tal como o PBVI, o SARSOP monta um conjunto de pontos de crenças do espaço de crença, sendo os pontos amostrados dos nós de uma árvore. O SARSOP aproxima as funções de valor dos pontos amostrados na forma de um conjunto linear de vetores. A partir de uma crença inicial b e com base em distribuições de probabilidade, ele realiza a amostragem de uma nova crença ao escolher uma ação $a \in A$ e observação $o \in O$. O principal ponto do algoritmo é de focar em nós promissores em termos de recompensas. São usados os limites superiores e inferiores na função de valor para orientar o processo de aprendizagem. Para amostrar novos pontos de crença, o SARSOP define um tamanho de intervalo entre os limites superiores e inferiores no nó raiz e caminha para baixo na árvore. Em cada nó, uma ação é escolhida com o limite superior mais alto enquanto a observação escolhida é aquela que faz a maior contribuição para o espaço no nó da raiz da árvore. O algoritmo termina quando o intervalo dos limites inferiores é menor que um dado limiar. A Figura 22 ilustra a ideia de busca no espaço de crença do algoritmo SARSOP.

4.2.5.3 Iteração de Valor de Monte Carlo

O algoritmo Iteração de Valor de Monte Carlo (denotado pelo acrônimo MCVI, da expressão em inglês *Monte Carlo Value Iteration*) (BAI et al., 2010) é semelhante ao

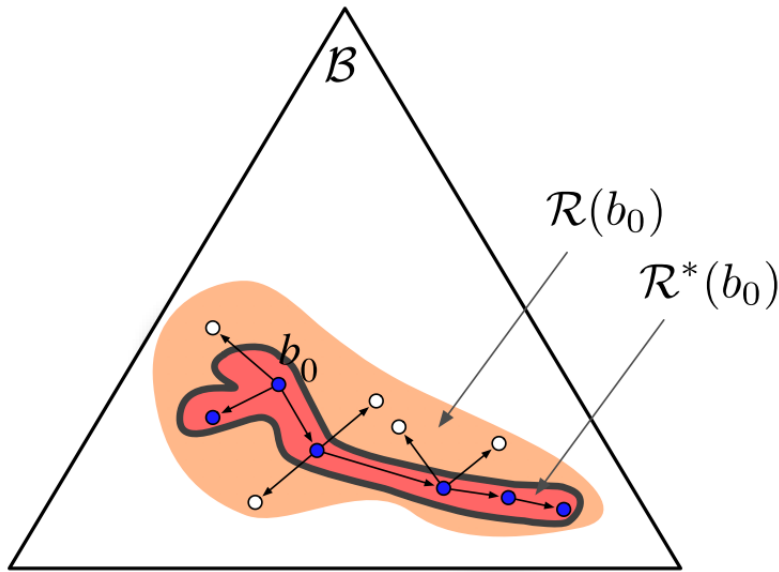


Figura 22 – B é o espaço de crenças, $R(b_0)$ é o espaço alcançável e $R^*(b_0)$ corresponde ao espaço ótimo alcançável. É importante observar que $R^*(b_0) \subseteq R(b_0) \subseteq B$ (KURNIAWATI; HSU; LEE, 2008).

SARSOP, uma vez que também é um algoritmo baseado em pontos. Entretanto, a principal vantagem é que ele usa simulações de Monte Carlo para representar os vetores. O resultado do algoritmo é um grafo de políticas que contém nós que correspondem a ações e arestas a observações. O algoritmo começa com uma amostragem a partir do espaço de crença e obtendo um subconjunto de pontos de crença B . Em seguida, um método de *backup* é executado nos pontos de crença B . O algoritmo é mais eficiente ao considerar apenas os pontos de B que são alcançáveis de acordo com uma política ótima. Ele usa o backup de Monte Carlo nos pontos de crença B e filtragem de partículas para lidar com os espaços contínuos do estado. Estes são os pontos de crença alcançáveis e, ao fazer atualizações nesses pontos, o algoritmo ganha eficiência computacional. Os pontos são amostrados construindo uma árvore de pontos de crença da mesma maneira que no algoritmo SARSOP. A vantagem aqui é a capacidade do MCVI de lidar com espaços de estados contínuos.

4.2.6 Planejamento Online com POMDPs

Quando tratamos de planejamento POMDP *offline*, existe o problema do tempo necessário para encontrar políticas. Em alguns casos existe um limite de tempo para a execução das ações, logo, os algoritmos *offline* podem não conseguir entregar uma solução adequada no momento devido. Os algoritmos baseados em pontos, tais como o PBVI e o SARSOP, ajudam a minimizar este problema mantendo vetores para um conjunto limitado de nós de crença, entretanto, com a restrição de tempo as soluções fornecidas podem ser de baixa qualidade (ROSS et al., 2008).

Em contra partida, temos os algoritmos de planejamento POMDP *online*. A maior vantagem deles é que requerem um custo computacional menor, portanto, encontram uma solução local ao invés de uma solução generalizada considerando todo o espaço de crença. Os algoritmos *online* seguem uma abordagem em que o planejamento e a execução são alternados em cada etapa. Alternar entre estas duas etapas permite que os algoritmos *online* se adaptem mais rápido a um ambiente dinâmico. A Figura 23 ilustra a diferença na estrutura de planejamento entre abordagens *online* e *offline*.

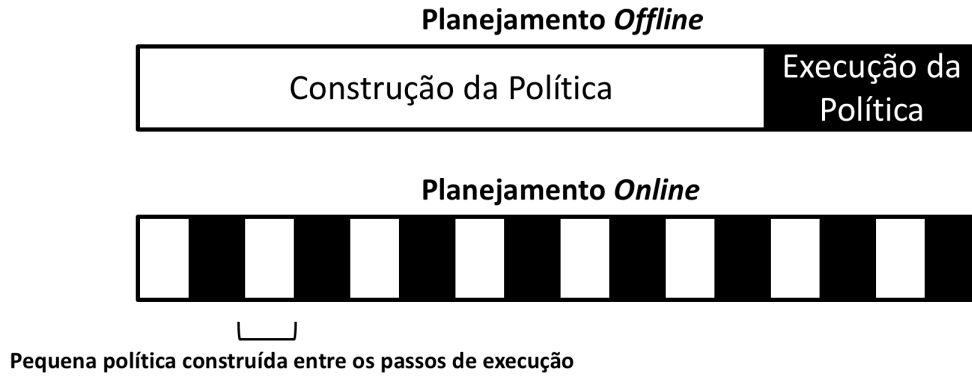


Figura 23 – Comparação entre os planejamentos *offline* e *online*.

Um dos algoritmos *online* mais conhecidos é o POMCP. O POMCP foi introduzido por (SILVER; VENESS, 2010) e consiste em um algoritmo de planejamento online para resolver grandes POMDPs. Ele trabalha com uma árvore de Monte Carlo e executa as atualizações de Monte Carlo sobre o estado de crença. A principal vantagem do POMCP é que ele aborda ambos, a maldição da dimensionalidade e da história. Ele aborda a maldição da dimensionalidade usando o poder da amostragem de Monte Carlo. O estado de crença é representado usando um filtro de partículas contendo k partículas e n estados possíveis (infinito no caso do espaço de estado contínuo). As atualizações do estado de crença são feitas usando amostragem, devida a representação muito compacta, filtro de partículas são computacionalmente atraentes. A maldição da história é abordada focando a busca na parte alcançável do espaço de crença, reduzindo assim o fator de ramificação. Outra característica importante da abordagem do algoritmo POMCP é que ele usa um “simulador de caixa preta” do POMDP para gerir as transições de estado e observações, eliminando assim a necessidade de distribuições explícitas de probabilidade.

Planejamento Probabilístico Aplicado no Mercado Financeiro

Este capítulo descreve as propostas de planejamento para resolver o problema de negociação de ações do mercado financeiro. Tais propostas têm como base o uso de algoritmos de planejamento probabilístico. Inicialmente é apresentada na Seção 5.1 o uso de POMDP para realizar a tarefa de planejamento. Na Seção 5.2 é realizada a decomposição do estado em múltiplas variáveis. Uma vez feita a decomposição, é proposta a modelagem do problema usando MOMDP. Por último, a Seção 5.3 apresenta a principal proposta do trabalho, o portfólio de planejamento. O portfólio tem o objetivo de suportar a escolha de planejadores mais aderentes ao problema do mercado financeiro.

5.1 Modelagem Observável Parcialmente

A abordagem apresentada nesta seção tem o objetivo de aplicar técnicas de planejamento no auxílio a tomada de decisões de compra e venda ações no mercado financeiro. Propõe-se um ambiente de estudo para modelagem, planejamento e simulação, utilizando Processos de Decisão de Markov Parcialmente Observáveis (POMDP) (CASSANDRA; KAEHLING; LITTMAN, 1994).

Nesta modelagem o mercado financeiro é considerado como sendo totalmente e parcialmente observável. Deste modo, é criado um modelo probabilístico que é resolvido por meio de algoritmos de POMDP. Esta modelagem tem como base o trabalho proposto por Baffa (BAFFA; CIARLINI, 2010b).

Seguindo a descrição sobre POMDP da Seção 4.2.3, primeiramente, é preciso definir o metamodelo. Neste contexto o modelo é constituído pela tupla $\langle S, A, O, T, P, R \rangle$, ou seja, uma composição de estados, ações, observações, probabilidades de transições entre estados, probabilidades das observações por estado e recompensas.

Os estados são compostos por três partes. A primeira corresponde à parte observável formada pela tendência (*ALTA*, *BAIXA* ou *ESTÁVEL*) do preço de um período passado.

Para realizar esta avaliação considera-se um intervalo de negociações n entre a data atual e um data do passado. A partir da diferença dos preços calcula-se a variação percentual v . Supondo que l seja uma variável que corresponda a um limite de variação, caso $v \geq l$ a tendência é de *ALTA*, caso $v \leq -l$ tem-se uma *BAIXA* e para todos os outros casos temos a tendência *ESTÁVEL*. A segunda parte do estado é parcialmente observável já que são considerados os dados futuros. Da mesma forma que a primeira parte, os tipos e as variações das tendências são os mesmos, sendo que a única diferença é a avaliação da variação do preço de um período de negociações entre uma data atual e uma data futura. Apesar deste trabalho utilizar apenas as tendências citadas anteriormente, existe a possibilidade do uso de mais regras/tendências, por exemplo, se $v \geq 2l$ a tendência poderia ser *ALTA ACENTUADA*. Por último tem-se a parte determinística, que representa o estado atual da carteira do investidor (*VENDIDO*, *COMPRADO* ou *FORA*). O investidor fica na posição de *VENDIDO* quando pega emprestado papéis de um ativo em uma corretora. Os papéis emprestados são devolvidos em uma data futura e neste caso o lucro é obtido quando o preço da ação esteve em queda. Quando o investidor compra ações sua posição é de *COMPRADO*. Já o caso *FORA* é quando o investidor não possui ações emprestadas e nem compradas. A Tabela 4 mostra exemplos de estados calculados a partir de um conjunto de dados históricos.

Tabela 4 – Estados na modelagem POMDP para o índice IBOVESPA.

Data	Índice	Passado		Futuro		Estado
		Variação	Tendência	Variação	Tendência	
24/01/2000	16972	-	-	-2.65%	BAIXA	-
26/01/2000	17105	-	-	-1.39%	ESTÁVEL	-
27/01/2000	17081	-	-	2.20%	ALTA	-
28/01/2000	16734	-	-	7.16%	ALTA	-
31/01/2000	16388	-	-	10.61%	ALTA	-
01/02/2000	16522	-2.72%	BAIXA	13.10%	ALTA	BAIXA_ALTA
02/02/2000	16868	-1.41%	ESTÁVEL	8.85%	ALTA	ESTÁVEL_ALTA
03/02/2000	17457	2.15%	ALTA	6.57%	ALTA	ALTA_ALTA
04/02/2000	17932	6.68%	ALTA	0.84%	ESTÁVEL	ALTA_ESTÁVEL
07/02/2000	18126	9.59%	ALTA	-1.69%	ESTÁVEL	ALTA_ESTÁVEL
08/02/2000	18686	11.58%	ALTA	-	-	-
09/02/2000	18361	8.13%	ALTA	-	-	-
10/02/2000	18604	6.17%	ALTA	-	-	-
11/02/2000	18083	0.84%	ESTÁVEL	-	-	-
14/02/2000	17819	-1.72%	ESTÁVEL	-	-	-

Na Tabela 4 são mostrados os estados para o índice IBOVESPA utilizando um intervalo de negociações $n = 5$ e o limite de variação $l = 2\%$. A posição do investidor é omitida e usada apenas no planejamento e nas simulações. De acordo com o tipo da tendência, passada ou futura, as $n = 5$ primeiras e últimas negociações não possuem um estado definido.

As ações aplicáveis são *COMPRA*, *COMPRA_DUPLA*, *VENDA*, *VENDA_DUPLA* e *NADA*. Cada ação altera o estado em relação a posição do investidor. Algumas ações não podem ser executadas para certas posições, como por exemplo, se o investidor estiver *COMPRADO* não é possível executar a ação de *COMPRA*. A Tabela 5 exibe as possíveis alterações de estados do investidor de acordo com as ações.

Tabela 5 – Alterações da posição do investidor de acordo com as ações/operações de compra e venda.

Ação	Alteração da Posição do Investidor
COMPRA	VENDIDO → FORA FORA → COMPRADO
COMPRA_DUPLA	VENDIDO → COMPRADO
VENDA	COMPRADO → FORA FORA → VENDIDO
VENDA_DUPLA	COMPRADO → VENDIDO
NADA	VENDIDO → VENDIDO COMPRADO → COMPRADO FORA → FORA

Como existem três valores para cada uma das três variáveis (tendência passada; tendência futura; posição do investidor), existe um total de 27 estados. Além destes 27 estados, foi adicionado o estado *INDESEJÁVEL* que representa um estado em que não se espera que seja alcançado. O *INDESEJÁVEL* é importante pois modela a impossibilidade de aplicar certas ações em algumas posições do investidor. Ao final tem-se o total de 28 estados.

A partir dos dados históricos são avaliadas as probabilidades de transições entre os estados. As probabilidades referem-se as transições entre as partes não determinísticas dos estados de dois instantes consecutivos. A Tabela 6 demonstra as transições entre estados de acordo com cada ação para o dia 02/02/2000 com base nos dados da Tabela 4.

Tabela 6 – Transições de estados para o dia 02/02/2000 da Tabela 4.

Data	Ação	Transições
02/02/2000	COMPRA	BAIXA_ALTA_VENDIDO → ESTÁVEL_ALTA_FORA BAIXA_ALTA_FORA → ESTÁVEL_ALTA_COMPRADO BAIXA_ALTA_COMPRADO → INDESEJÁVEL
	COMPRA_DUPLA	BAIXA_ALTA_VENDIDO → ESTÁVEL_ALTA_COMPRADO BAIXA_ALTA_COMPRADO → INDESEJÁVEL BAIXA_ALTA_FORA → INDESEJÁVEL
	VENDA	BAIXA_ALTA_COMPRADO → ESTÁVEL_ALTA_FORA BAIXA_ALTA_FORA → ESTÁVEL_ALTA_VENDIDO BAIXA_ALTA_VENDIDO → INDESEJÁVEL
	VENDA_DUPLA	BAIXA_ALTA_COMPRADO → ESTÁVEL_ALTA_VENDIDO BAIXA_ALTA_VENDIDO → INDESEJÁVEL BAIXA_ALTA_FORA → INDESEJÁVEL
	NADA	BAIXA_ALTA_COMPRADO → ESTÁVEL_ALTA_COMPRADO BAIXA_ALTA_VENDIDO → ESTÁVEL_ALTA_VENDIDO BAIXA_ALTA_FORA → ESTÁVEL_ALTA_FORA

As observações são definidas pelo usuário e têm como base indicadores gerados a partir de dados de análises técnicas. Foram usadas as análises RSI, MACD e OST. São geradas três observações, sendo o *PONTO_VENDA*, *PONTO_COMPRA* e *PONTO_NADA*. O *PONTO_VENDA* é uma indicação para venda de ações, o *PONTO_COMPRA* um momento para compra e por último o *PONTO_NADA* que sugere que nenhuma negociação seja realizada. A Tabela 7 mostra as observações geradas a partir de regras baseadas em dados de análise técnica.

Tabela 7 – Regras de geração das observações.

Indicador	Parâmetros	Observação
RSI ₁ RSI ₁ RSI ₁	$(rsi_{atual} \geq 0.2) \wedge (rsi_{anterior} \leq 0.2)$ $(rsi_{anterior} \geq 0.8) \wedge (rsi_{atual} \leq 0.8)$	PONTO_COMPRA PONTO_VENDA PONTO_NADA
RSI ₂ RSI ₂ RSI ₂	$(rsi_{atual} \geq 0.3) \wedge (rsi_{anterior} \leq 0.3)$ $(rsi_{anterior} \geq 0.7) \wedge (rsi_{atual} \leq 0.7)$	PONTO_COMPRA PONTO_VENDA PONTO_NADA
MACD ₁ MACD ₁ MACD ₁	$(macd_{atual} \geq 0) \wedge (macd_{anterior} \leq 0)$ $(macd_{anterior} \geq 0) \wedge (macd_{atual} \leq 0)$	PONTO_COMPRA PONTO_VENDA PONTO_NADA
MACD ₂ MACD ₂ MACD ₂	$(histograma_macd_{atual} \geq 0) \wedge (histograma_macd_{anterior} \leq 0)$ $(histograma_macd_{anterior} \geq 0) \wedge (histograma_macd_{atual} \leq 0)$	PONTO_COMPRA PONTO_VENDA PONTO_NADA
OST ₁ OST ₁ OST ₁	$(ost_K_{atual} \geq 0.2) \wedge (ost_K_{anterior} \leq 0.2)$ $(ost_K_{anterior} \geq 0.8) \wedge (ost_K_{atual} \leq 0.8)$	PONTO_COMPRA PONTO_VENDA PONTO_NADA
OST ₂ OST ₂ OST ₂	$(ost_D_{atual} \geq 0.2) \wedge (ost_D_{anterior} \leq 0.2)$ $(ost_D_{anterior} \geq 0.8) \wedge (ost_D_{atual} \leq 0.8)$	PONTO_COMPRA PONTO_VENDA PONTO_NADA

Para cada umas das análises técnicas RSI, MACD e OST, são dadas apenas duas interpretações. Os parâmetros das análises, tais como os períodos, dependem do experimento e são detalhados no Capítulo 8.

As probabilidades das observações são geradas a partir da análise dos dados de séries históricas. As probabilidades das observações referem-se a correspondência entre a parte não-determinística do estado associado a um instante e a observação gerada para aquele instante. A Tabela 8 mostra exemplos de observações obtidas a partir da regra RSI₂ calculada para 14 períodos.

Tabela 8 – Observações para o índice IBOVESPA utilizando a regra RSI₂ com 14 períodos.

Data	Índice	Estado	RSI (14 períodos)	Observações (RSI ₂)
31/01/2000	16388	-	0.290	-
01/02/2000	16522	BAIXA_ALTA	0.386	PONTO_COMPRA
02/02/2000	16868	ESTÁVEL_ALTA	0.564	PONTO_NADA
03/02/2000	17457	ALTA_ALTA	0.704	PONTO_NADA
04/02/2000	17932	ALTA_ESTÁVEL	0.686	PONTO_VENDA
07/02/2000	18126	ALTA_ESTÁVEL	0.693	PONTO_NADA

Por último são estabelecidas as recompensas que estão diretamente relacionadas aos estados. As recompensas são baseadas em fatores qualitativos, ou seja, os valores são determinados de acordo com a avaliação do quão bom ou ruim é um estado em relação aos outros. Por exemplo, comparando os estados *ALTA_ALTA_COMPRADO* e *BAIXA_BAIXA_COMPRADO* a recompensa é melhor no primeiro caso, já que nesta situação o mercado está em alta e o investidor já comprou ações. Deste modo, para o primeiro estado, quando o mercado inverter a tendência e o investidor vender suas ações possivelmente ele terá lucro.

Quantitativamente as recompensas podem ter valores diferentes, ou seja, outros valores podem afetar os resultados. Sendo assim, trabalhos futuros constituem em pesquisar quantitativamente os melhores valores das recompensas. A Tabela 9 mostra as recompensas para cada estado. É dada uma penalidade alta como recompensa ao estado *INDESEJÁVEL*, já que não é desejável que ele esteja na política final.

Tabela 9 – Recompensa de cada estado.

Estado	Recompensa
BAIXA_BAIXA_COMPRADO	-10
BAIXA_ESTÁVEL_COMPRADO	-2
BAIXA_ALTA_COMPRADO	5
ESTÁVEL_BAIXA_COMPRADO	-5
ESTÁVEL_ESTÁVEL_COMPRADO	0
ESTÁVEL_ALTA_COMPRADO	5
ALTA_BAIXA_COMPRADO	0
ALTA_ESTÁVEL_COMPRADO	5
ALTA_ALTA_COMPRADO	10
BAIXA_BAIXA_VENDIDO	10
BAIXA_ESTÁVEL_VENDIDO	5
BAIXA_ALTA_VENDIDO	0
ESTÁVEL_BAIXA_VENDIDO	5
ESTÁVEL_ESTÁVEL_VENDIDO	0
ESTÁVEL_ALTA_VENDIDO	-5
ALTA_BAIXA_VENDIDO	5
ALTA_ESTÁVEL_VENDIDO	-2
ALTA_ALTA_VENDIDO	-10
BAIXA_BAIXA_FORA	-2
BAIXA_ESTÁVEL_FORA	0
BAIXA_ALTA_FORA	0
ESTÁVEL_BAIXA_FORA	0
ESTÁVEL_ESTÁVEL_FORA	2
ESTÁVEL_ALTA_FORA	0
ALTA_BAIXA_FORA	0
ALTA_ESTÁVEL_FORA	0
ALTA_ALTA_FORA	-2
INDESEJÁVEL	-25

A representação para a modelagem POMDP pode ser realizada de diferentes formas, tais como a proposta de Cassandra¹ (CASSANDRA; KAEHLING; LITTMAN, 1994),

¹ Disponível em <<http://www.pomdp.org/code/pomdp-file-spec.html>>

SPUDD (HOEY et al., 1999), POMDPX²(ONG et al., 2010) ou RDDDL³ (SANNER, 2010). O Apêndice A mostra um exemplo de modelagem para o mercado financeiro segundo Cassandra (CASSANDRA; KAEHLING; LITTMAN, 1994).

A modelagem apresentada nesta Seção contempla o uso de apenas um ativo. Logo, para tratar dois ou mais ativos alguns pontos precisam ser ajustados. Neste contexto, a Seção 5.1.1 descreve meios de trabalhar com múltiplos ativos.

5.1.1 Modelagem para Múltiplos Ativos

Para a tarefa de planejamento utilizando múltiplos ativos é levado em conta as diferentes formas de modelagem. Contudo, antes do planejamento, é preciso realizar a seleção de dois ou mais ativos. Tal problema pode ser resolvido através da correlação de ativos, conforme descrito na Seção 3.2.

Para o uso do POMDP com múltiplos ativos é necessário redefinir a modelagem de cada componente. Para melhorar o entendimento, o nome dos ativos são adicionados ao modelo. Conforme as definições, a primeira variável da tupla $\langle S, A, O, T, P, R \rangle$ corresponde aos estados. Os estados são formados pelo produto das tendências do passado, tendências do futuro e da posição do investidor, em relação a cada ativo. Em outras palavras, é realizada a combinação de todas as possibilidades de tendências e posições em relação aos ativos. Conforme a Seção 5.1, existem 27 estados para um único ativo mais o estado *INDESEJÁVEL*. Sendo assim, o número total de estados para n ativos é igual a $27^n + 1$.

A Tabela 10 exibe 3 exemplos de estados para os ativos PETR4 e VALE5, de um total de 729 possíveis estados. Neste exemplo, o primeiro estado indica que para PETR4 a tendência do passado foi de *BAIXA*, a tendência do futuro é de *ESTÁVEL* e o investidor comprou ações de tal ativo. Já para VALE5 a tendência do passado foi de *ALTA*, a tendência futura é de *BAIXA* e o investidor não possui papéis da VALE5. Os demais estados seguem a mesma concepção.

Tabela 10 – Exemplos de estados para PETR4 e VALE5 ao usar múltiplos ativos.

Estados
BAIXA_ESTÁVEL_COMPRADO_PETR4_ALTA_BAIXA_FORA_VALE5
ALTA_BAIXA_FORA_PETR4_ALTA_ALTA_VENDIDO_VALE5
ALTA_BAIXA_COMPRADO_PETR4_ALTA_ALTA_VENDIDO_VALE5

As ações seguem a mesma lógica dos estados, ou seja, é realizada a combinação das possibilidades para sua formação. Como existem 5 ações para um único ativo, para n ativos temos o total de 5^n ações. A Tabela 12 mostra todas as possíveis ações ao combinar PETR4 e VALE5, logo, para dois ativos existem 25 ações. Por exemplo, a

² Disponível em: <http://bigbird.comp.nus.edu.sg/pmwiki/farm/appl/index.php?n=Main.PomdpXDocumentation>

³ Disponível em: <https://github.com/ssanner/rddlsim>

ação COMPRA_PETR4_NADA_VALE5 determina que seja comprado PETR4 e que nenhuma operação seja feita em relação a VALE5.

Tabela 11 – Ações para PETR4 e VALE5 ao usar múltiplos ativos.

Ações	
COMPRA_PETR4_COMPRA_VALE5	COMPRA_PETR4_COMPRA_DUPLA_VALE5
COMPRA_PETR4_VENDA_VALE5	COMPRA_PETR4_VENDA_DUPLA_VALE5
COMPRA_PETR4_NADA_VALE5	COMPRA_DUPLA_PETR4_COMPRA_VALE5
COMPRA_DUPLA_PETR4_COMPRA_DUPLA_VALE5	COMPRA_DUPLA_PETR4_VENDA_VALE5
COMPRA_DUPLA_PETR4_VENDA_DUPLA_VALE5	COMPRA_DUPLA_PETR4_NADA_VALE5
VENDA_PETR4_COMPRA_VALE5	VENDA_PETR4_COMPRA_DUPLA_VALE5
VENDA_PETR4_VENDA_VALE5	VENDA_PETR4_VENDA_DUPLA_VALE5
VENDA_PETR4_NADA_VALE5	VENDA_DUPLA_PETR4_COMPRA_VALE5
VENDA_DUPLA_PETR4_COMPRA_DUPLA_VALE5	VENDA_DUPLA_PETR4_VENDA_VALE5
VENDA_DUPLA_PETR4_VENDA_DUPLA_VALE5	VENDA_DUPLA_PETR4_NADA_VALE5
NADA_PETR4_COMPRA_VALE5	NADA_PETR4_COMPRA_DUPLA_VALE5
NADA_PETR4_VENDA_VALE5	NADA_PETR4_VENDA_DUPLA_VALE5
NADA_PETR4_NADA_VALE5	

De forma análoga, as observações são dadas pelas combinações das observações de cada ativo. Se existem 3 observações para um único ativo, o total de possibilidades para n ativos é igual a 3^n . A Tabela 12 mostra todas as possíveis observações ao combinar PETR4 e VALE5, logo, para dois ativos existem 9 observações.

Tabela 12 – Observações para PETR4 e VALE5 ao usar múltiplos ativos.

Observações
PONTO_COMPRA_PETR4_PONTO_COMPRA_VALE5
PONTO_COMPRA_PETR4_PONTO_VENDA_VALE5
PONTO_COMPRA_PETR4_PONTO_NADA_VALE5
PONTO_VENDA_PETR4_PONTO_COMPRA_VALE5
PONTO_VENDA_PETR4_PONTO_VENDA_VALE5
PONTO_VENDA_PETR4_PONTO_NADA_VALE5
PONTO_NADA_PETR4_PONTO_COMPRA_VALE5
PONTO_NADA_PETR4_PONTO_VENDA_VALE5
PONTO_NADA_PETR4_PONTO_NADA_VALE5

As probabilidades de transições dos estados e observações são calculadas a partir dos dados históricos. Para os múltiplos ativos é feita a contagem das transições entre datas consecutivas para todo um período e ao final são realizados os cálculos das probabilidades.

Por último, as recompensas para múltiplos ativos são calculadas através da soma das recompensas individuais (Tabela 9). A Tabela 13 exhibe três exemplos de recompensas ao combinar os ativos PETR4 e VALE5. No primeiro estado, temos PETR4 com tendência passada de BAIXA, tendência futura de ESTÁVEL e o investidor COMPRADO, logo o valor da recompensa é igual a -2. Para VALE5 com tendência passada de BAIXA, tendência futura de BAIXA e o investidor VENDIDO, o valor da recompensa é igual a -10. Então, a recompensa final é dada por $-2 + 10 = 8$. Os demais estados seguem a mesma ideia de cálculo de recompensa. É importante salientar que, independente do número de ativos, existe apenas um estado *INDESEJÁVEL* e sua recompensa é igual a $-25 * n$, sendo n igual ao número de ativos.

Tabela 13 – Exemplos de recompensas de estados formados com os ativos PETR4 e VALE5.

Estado	Recompensa
BAIXA_ESTÁVEL_COMPRADO_PETR4_BAIXA_BAIXA_VENDIDO_PETR4	8
ALTA_ALTA_COMPRADO_PETR4_BAIXA_ESTÁVEL_VENDIDO_PETR4	15
BAIXA_BAIXA_COMPRADO_PETR4_BAIXA_BAIXA_COMPRADO_PETR4	-20

Existem dois problemas na modelagem desta seção. O primeiro problema é que o número de elementos é exponencial, logo executar o planejamento é uma tarefa difícil, mesmo para poucos ativos.

Para mensurar a complexidade do problema, um recurso a ser avaliado é o espaço para armazenar o modelo. Para n ativos temos 27^n estados, 5^n ações e 4 bytes para cada valor numérico. Considerando apenas os dados das transições entre estados, é necessária uma matriz quadrada em função do número de estados para cada ação, sendo assim, são necessários $27^{2n} * 5^n * 4$ bytes. O espaço para armazenar as probabilidades de transições entre estados para 2 e 3 ativos são respectivamente, $27^4 * 5^2 * 4$ bytes $\cong 50$ megabytes (MB) e $27^6 * 5^3 * 4$ bytes $\cong 180$ gigabytes (GB). Consequentemente, a execução do planejamento para 3 ou mais ativos é uma tarefa muito difícil. Outro problema é que devida a grande quantidade de estados a distribuição de probabilidades pode ser esparsa. Com o objetivo de tratar estes problemas é proposta uma nova abordagem de planejamento na Seção 5.2.

5.2 Modelagem Observável Mista

Uma das características da modelagem da Seção 5.1 é que a formação do estado é feita a partir de 3 tipos de variáveis. Uma variável completamente observável, outra parcialmente observável e uma determinística. Em outras palavras, as tendências do passado são completamente observáveis, as tendências futuras são parcialmente observáveis e a posição do investidor é determinística. Em problemas POMDP normalmente o estado de crença é caracterizado apenas por valores parcialmente observáveis, como é mostrado na Figura 21.

Observando as características apresentadas anteriormente, os estados contêm dados que não são necessários para resolver o problema exclusivamente como um POMDP. Deste modo algumas variáveis impactam na performance e nos resultados do planejamento (ONG et al., 2010). Um modo para resolver este problema consiste em modelar o planejamento como sendo um MOMDP (denotado pelo acrônimo MOMDP, da expressão em inglês *Mixed Observable Markov Decision Process*). Além da modelagem, Ong (ONG et al., 2010) apresenta uma proposta de algoritmo para resolver modelos MOMDP. Ong demonstrou bons resultados de performance e recompensas ao comparar MOMDP e POMDP.

No MOMDP o estado é decomposto em variáveis de acordo com o tipo. Logo, a Definição 4.2.2 é dividida segundo cada tipo de elemento dos estados. Formalmente, a

modelagem MOMDP é especificada pela tupla $\langle X, Y, A, T_x, T_y, O, P, C, R \rangle$. Os termos A, O, P, CeR são os mesmos do POMDP, já X, Y, T_x, T_y correspondem a Definição 5.2.1.

Definição 5.2.1 O MOMDP é definido pela a tupla $\langle X, Y, A, T_x, T_y, O, P, C, R \rangle$. Além dos elementos comuns ao POMDP, tem-se:

X é um conjunto finito de estados completamente observáveis.

Y é um conjunto finito de estados parcialmente observáveis.

T_x corresponde a $T_x(x, y, a, x') = p(x'|x, y, a)$ e representa a probabilidade de $x' \in X$ ser alcançado ao executar a ação $a \in A$ sobre o estado (x, y) .

T_y corresponde a $T_y(x, y, a, x', y') = p(y'|x, y, a, x')$ e representa a probabilidade de $y' \in Y$ ser alcançado ao executar a ação $a \in A$ sobre o estado (x, y) e que o próximo estado completamente observável seja $x' \in X$.

No MOMDP o estado s é decomposto nas variáveis x e y e correspondem, respectivamente, aos estados completamente observável e parcialmente observável. A Figura 24 mostra a diferença entre as transições do POMDP e do MOMDP.

Para o MOMDP o conjunto S é fatorado em $S = X \times Y$. A vantagem computacional decorre do fato de que, uma vez que a variável x é completamente observável é preciso manter apenas o estado de crença b_y e a distribuição de probabilidades do estado y . Agora qualquer estado de crença $b \in B$ para $s = (x, y)$ é então representado por (x, b_y) . O conjunto de estados de crença B é formado por $|X||Y|$ dimensões, onde $|X|$ e $|Y|$ são os números de estados em X e Y . Já $B_y(x)$ possui apenas $|Y|$ dimensões. É associado a cada valor x um espaço de estado de crença em $y|B_y(x) = (x, b_y)|b_y \in B_y$ é um subespaço em B , onde B é a união destes subespaços $\cup_{x \in X} B_y(x)$. Então, com esta representação o espaço B é formado pela união de subespaço de “baixo número de dimensões”. Efetivamente, agora tem-se que o espaço de incertezas é menor, já que $|Y| < |X||Y|$. Com isso o modelo do MOMDP lida com menos estados de crença, fazendo com que o tempo computacional seja menor. Até então, as alterações introduzidas pela modelagem MOMDP evidenciam mais fatores de notação, logo a forma com que é gerada a política se matem com pequenas alterações.

Como apresentada na Seção 4.2.3, uma política POMDP pode ser representada como uma função valor $V(b) = \max_{\alpha \in \Gamma} (\alpha b)$, onde Γ é um conjunto de vetores α . De modo correspondente a função valor para o MOMDP é $V(x, b_y)$ e também é expresso por uma coleção de vetores α , tal que $\{\Gamma_y(x)|x \in X\}$, onde para cada x , $\Gamma_y(x)$ é um conjunto de vetores α definidos em $B_y(x)$. Para avaliar $V(x, b_y)$ procurasse o vetor α usando x que maximiza a Equação 40.

$$V(x, B_y) = \max_{\alpha \in \Gamma_y(x)} (\alpha.b_y) \quad (40)$$

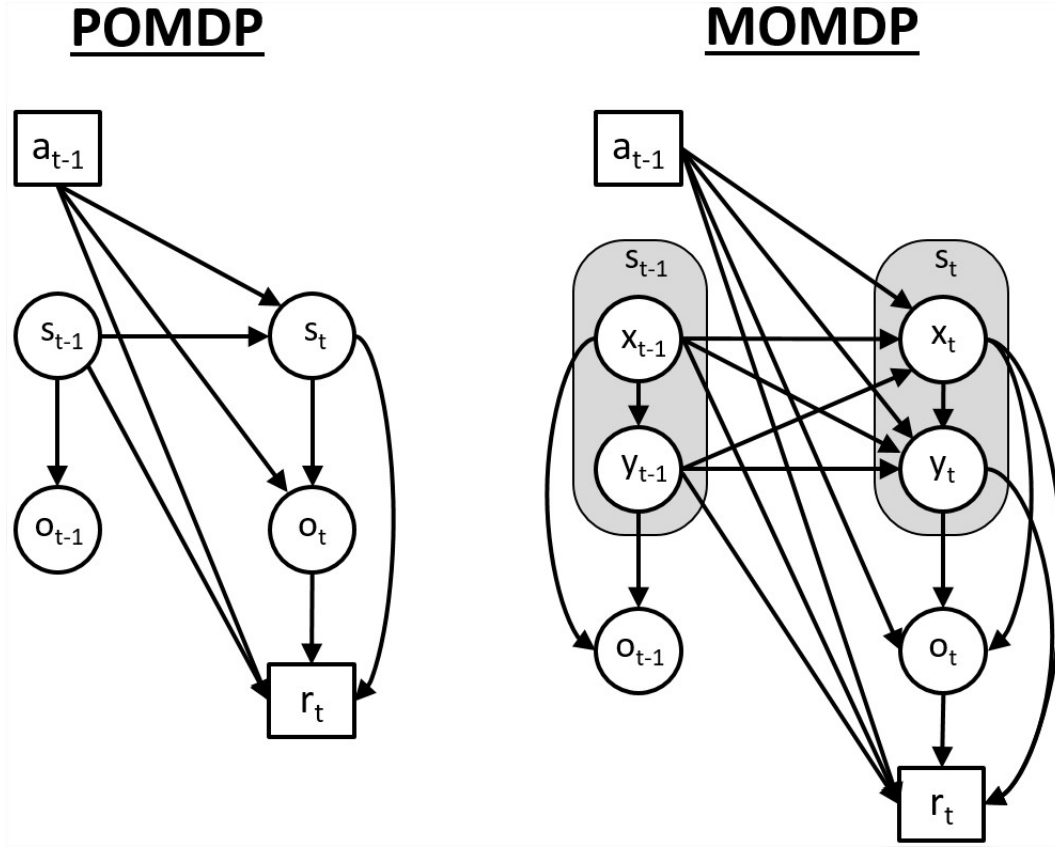


Figura 24 – Transição de estados para POMDP (esquerda) e MOMDP (direita). O estado s do MOMDP é decomposto em duas partes: $s = (x, y)$, onde x é completamente observável e y parcialmente observável. t representa o tempo/momento de cada termo, a é a ação, o é a observação e r é a recompensa.

De modo geral qualquer valor da função $V(b) = \max_{\alpha \in \Gamma} (\alpha.b)$ pode ser representado desta nova forma (ONG et al., 2010). Com a decomposição do estado, além da performance de planejamento ser melhorada, a execução da política também é mais rápida. A Figura 25 mostra a árvore de estados para problemas MOMDP. Nesta árvore é exibida a decomposição dos estados e o modo como ocorre a ramificação de cada componente.

A modelagem MOMDP para o mercado financeiro pode ser realizada de diferentes formas. O primeiro modelo proposto é formado por duas variáveis. A primeira variável é completamente observável e é composta pela união da tendência do passado e a posição do investidor. A segunda é parcialmente observável e é dada pelas tendências futuras, sendo esta parte responsável por constituir o estado de crença. A Figura 26 exhibe a árvore de estados para esta proposta.

Com esta modelagem as crenças em B são compostas por apenas 3 valores. Cada valor corresponde as tendências futuras de *ALTA*, *BAIXA* e *ESTÁVEL*. Efetivamente, tal alteração reduz o número de estados de crença de 28 para 3, fazendo com que B seja formado apenas pela parte parcialmente observável.

Apesar do avanço da primeira modelagem MOMDP, ainda existe uma característica

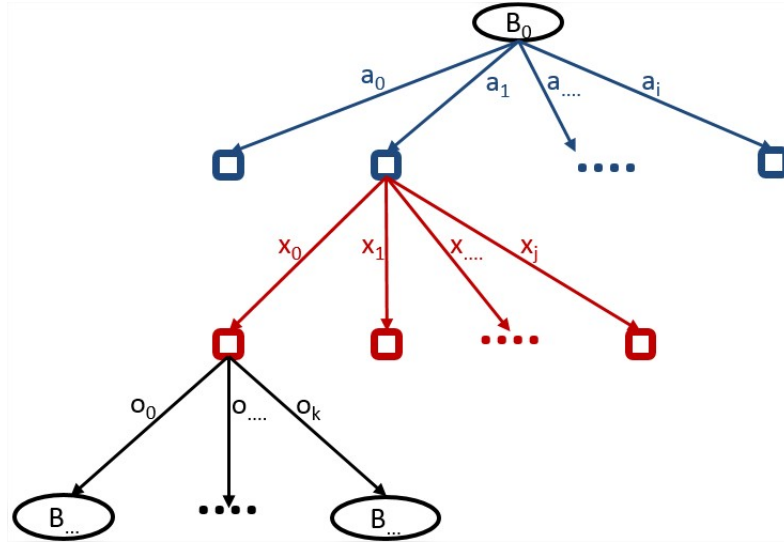


Figura 25 – Árvore de estados de crença para problemas MOMDP.

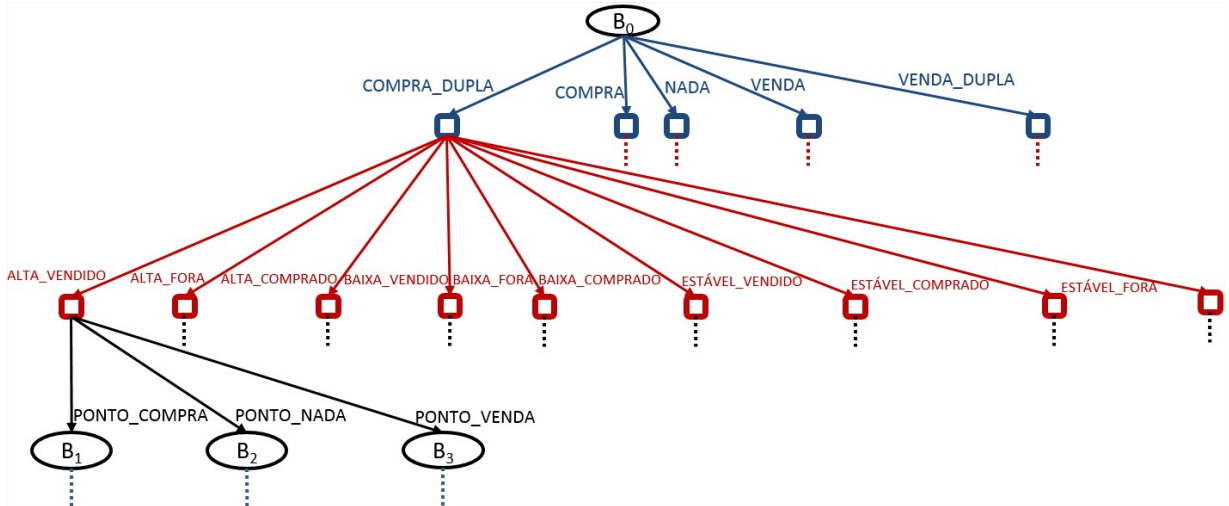


Figura 26 – Árvore de estados e transições para o problema do mercado financeiro usando MOMDP.

que precisa ser melhorada. Desta forma, a segunda modelagem consiste em separar a posição do investidor. Esta alteração é importante pois a posição do investidor é determinística, diferente da tendência do passado que é não determinística. Existem diferentes formas de determinar as transições entre os estados. Com uma representação mais flexível a ação passa a delimitar quais são os estados possíveis de serem alcançados.

A separação da posição do investidor e da tendência do passado é mostrada na Figura 27. Neste exemplo a ação *COMPRA* pode alterar a variável determinística apenas para *COMPRADO* ou *FORA*, as demais ações seguem a mesma lógica.

A terceira forma de melhorar a modelagem consiste em, além de separar a parte determinística, utilizar as informações do estado atual para estabelecer quais transições podem ocorrer. Uma das propriedades de Markov é de que o próximo estado depende exclusiva-

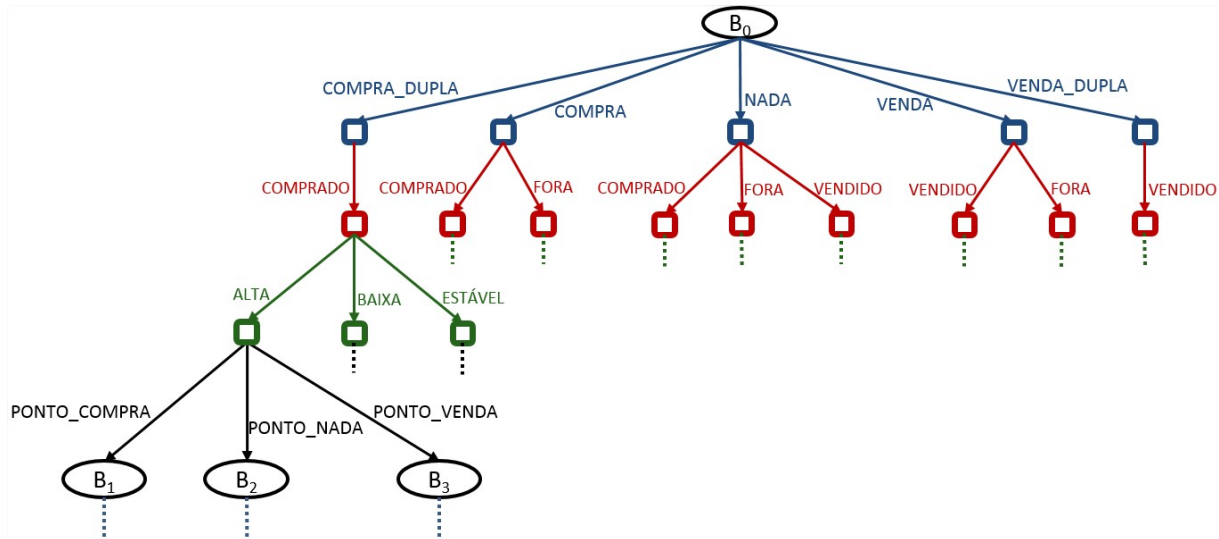


Figura 27 – Separação da parte determinística da modelagem MOMDP.

mente do estado atual, não sendo necessário guardar o histórico de ações e estados. Com base neste fundamento a alteração consiste em guardar a posição do investidor e utilizá-la como orientação para a aplicação das ações. Com isso o número de transições é reduzido e consequentemente a performance tende a ser melhor. É importante salientar que a posição do investidor é conhecida apenas no estado atual e as transições são realizadas com base neste dado. Outro benefício é que o estado *INDESEJADO* pode ser removido, já que o planejador controlaria as possibilidades de execução das ações. A Figura 28 apresenta esta alteração.

Atualmente são usados planejadores de contexto geral, logo algumas peculiaridades da modelagem não são consideradas no planejamento. Sendo assim, apesar da terceira proposta de modelagem MOMDP representar um avanço no que se refere ao planejamento para o mercado financeiro, neste momento ela não é utilizada. Logo, este é um ponto a ser tratado em trabalhos futuros.

Na pesquisa o estado foi separado segundo a característica de cada componente, ou seja, é usada a segunda abordagem. A Figura 29 mostra como foi estruturada a modelagem MOMDP.

O Apêndice B mostra um exemplo de modelagem MOMDP utilizando POMDPX.

5.3 Portfólio de Planejamento

Para a tarefa de planejamento existem muitos planejadores e cada um deles se adapta melhor a certos tipos de problemas (ROBERTS; HOWE, 2009). Uma tarefa difícil consiste em, dado um grupo de planejadores e sem executar todos eles, escolher o melhor planejador para resolver um dado problema. Para realizar esta tarefa é analisado o problema e de

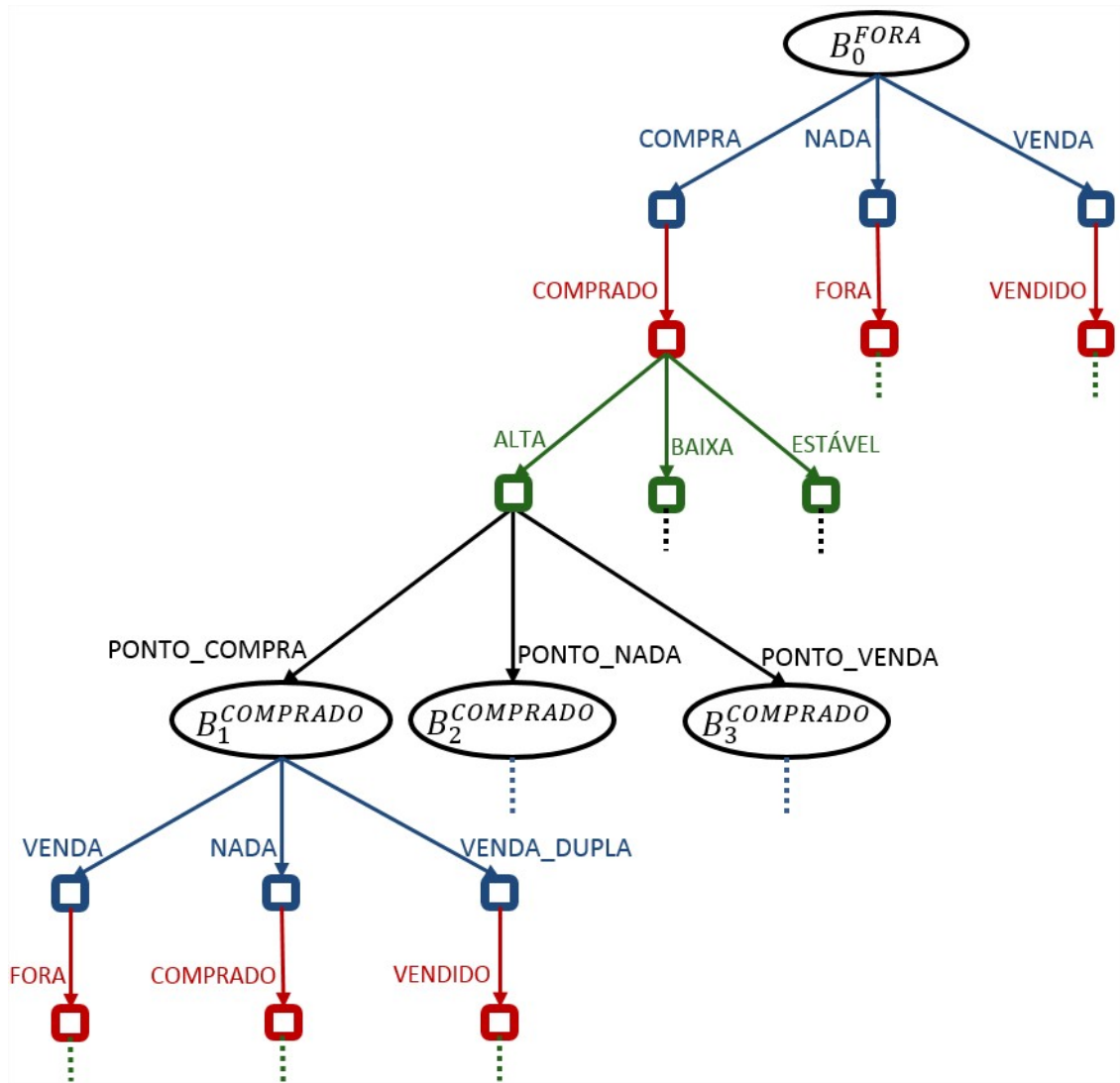


Figura 28 – Modelagem considerando a variável determinística na definição das ações.

acordo com suas características são selecionados planejadores. Este problema de seleção é caracterizado como um portfólio de planejamento (SOUSA, 2014).

Uma vez que múltiplos planejadores podem ser selecionados torna-se necessário utilizar técnicas de escalonamento (SOUSA, 2014) ou *ensemble method* (ZHOU, 2012) para combinar as diferentes estratégias de planejamento.

Para resolver o portfólio de planejamento os trabalhos de (ROBERTS; HOWE, 2009) e (SOUSA, 2014) propõem o uso de aprendizado de máquina, em específico reconhecimento de padrões. Contudo, tais trabalhos tratam apenas do planejamento clássico, deixando de explorar o planejamento não clássico. É válido lembrar que o planejamento não clássico é formado por problemas onde os resultados das ações e os estados não são determinísticos. Neste contexto tem-se os problemas POMDP e MOMDP.

Uma contribuição não explorada até o momento por outras pesquisas é o portfólio de planejamento para o planejamento não clássico. Logo, esta é uma contribuição almejada neste trabalho.

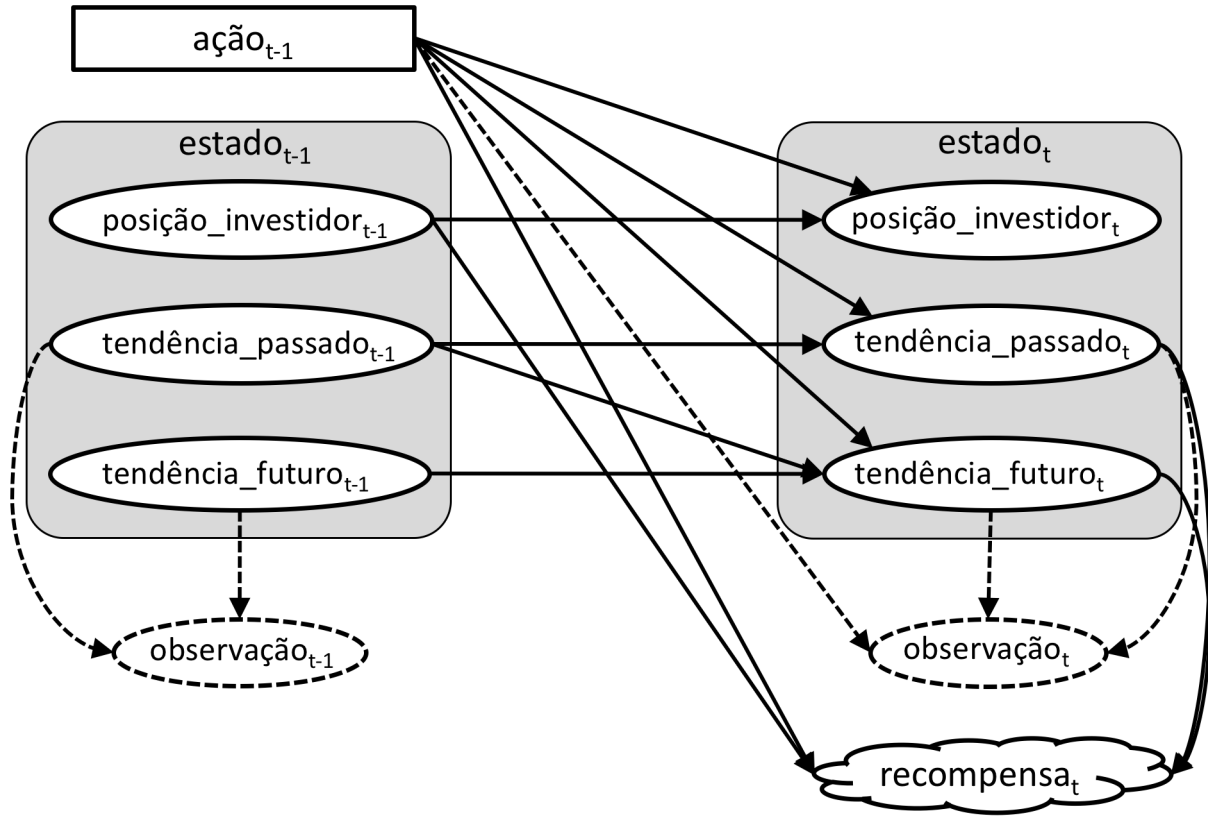


Figura 29 – Modelagem MOMDP para o mercado financeiro.

Para implementar o portfólio de planejamento é necessário determinar os seguintes parâmetros: escopo (geral ou específico), fator a ser otimizado, linguagem de representação dos problemas, características extraídas dos problemas baseada na linguagem de representação, quantidade de planejadores, quantidade de instâncias de problemas, algoritmos de classificação e a metodologia de treinamento e teste.

É utilizado um escopo geral para a formulação do portfólio. Para avaliar os resultados é usado o problema de planejamento para o mercado financeiro. Entre os fatores de otimização temos a qualidade das soluções, a maximização da quantidade de problemas resolvidos e a maximização das recompensas.

A linguagem escolhida para tratar do portfólio de planejamento é a RDDDL, mais detalhes desta linguagem são colocados na Subseção 4.2.1. As características que são extraídas dos problemas são detalhadas na Subseção 5.3.1. Já os planejadores que podem compor o portfólio são colocados na Subseção 5.3.2. Os problemas utilizados são aqueles disponíveis na competição IPPC e são detalhados na Subseção 5.3.3. Finalmente, a Subseção 5.3.4 descreve a lista de classificadores e o processo de seleção de características.

5.3.1 Características dos Problemas

Com base na RDDDL são extraídas as características dos problemas. As características selecionadas são listadas abaixo:

- ☐ Presença ou ausência do requisito *continuous*.
- ☐ Presença ou ausência do requisito *multivalued*.
- ☐ Presença ou ausência do requisito *reward-deterministic*.
- ☐ Presença ou ausência do requisito *intermediate-nodes*.
- ☐ Presença ou ausência do requisito *constrained-state*.
- ☐ Presença ou ausência do requisito *partially-observed*.
- ☐ Presença ou ausência do requisito *concurrent*.
- ☐ Presença ou ausência do requisito *integer-valued*.
- ☐ Presença ou ausência do requisito *cpf-deterministic*.
- ☐ Quantidade de fatos em *non-fluent* da instância.
- ☐ Quantidade de fatos em *init-state* da instância.
- ☐ Quantidade de fatos *types* do domínio.
- ☐ Quantidade de fatos *pvariables* do domínio.
- ☐ Quantidade de *objects* em *non-fluent*.
- ☐ Quantidade de literais *pvariable* do domínio que são do tipo *intermediate-fluents*.
- ☐ Quantidade de literais *pvariable* do domínio que são do tipo *state-fluent*.
- ☐ Quantidade de literais *pvariable* do domínio que são do tipo *action-fluent*.
- ☐ Quantidade de literais *pvariable* do domínio que são do tipo *observation-fluent*.
- ☐ Quantidade de restrições dos estados/ações.
- ☐ Quantidade de objetos definidos nos *non-fluent*.

5.3.2 Planejadores

Os planejadores foram selecionados a partir da competição IPPC (*International Probabilistic Planning Competition*). Esta competição foi realizada nos anos de 2011 e 2014. Em 2011 foram utilizados o *Symbolic Perseus*, *POMDPX NUS*, *KAIST AILAB*, *McGill* e *HyPlan* e em 2014 o *KAIST_AIPR_LAB*, *NUS-POMDPGroup*, *NoopPolicy* e *Random-BoolPolicy*.

5.3.3 Problemas de Planejamento

Assim como os planejadores, os problemas foram selecionados com base na IPPC. Nas competições de 2011 e 2014 foram usados vários problemas, a seguir são descritos brevemente estes problemas:

- ❑ *Academic Advising* (2014): Um aluno pode selecionar cursos a um custo determinado e passa no curso com uma probabilidade determinada pelo número de pré-requisitos que ele passou com sucesso.
- ❑ *Crossing Traffic* (2011 e 2014): Em um gride um robô deve chegar na meta desviando de diversos obstáculos que aparecem aleatoriamente.
- ❑ *Elevator* (2011 e 2014): Um número de elevadores é fornecido e devem coletar todos os passageiros e transportá-los para seu destino. O objetivo é levar os passageiros aos seus destinos sempre que um novo passageiro surgir.
- ❑ *Game of Life* (2011): Este problema representa o “jogo da vida”, um famoso problema da área de autômatos celulares (GARDNER, 1970).
- ❑ *Navigation* (2011): Este problema é semelhante ao *Crossing Traffic*, porém, cada célula tem uma probabilidade que o robô desapareça ao pisar nela.
- ❑ *Reconnaissance* (2011): Em um grid existe um robô, um base, células contaminadas e objetos espalhados. O robô utiliza ferramentas para passar pelo grid e encontrar o maior número de seres vivos e fotografá-los.
- ❑ *Skill Teaching* (2011 e 2014): Um agente tenta ensinar uma série de habilidades a um estudante, usando aulas e questões. O objetivo é aumentar o nível de conhecimento e resolver o máximo de questões.
- ❑ *SysAdmin* (2011): O administrador de um sistema é responsável por manter uma rede de computadores ativa. Cada computador pode estar ou não em funcionamento. O objetivo é manter o máximo de computadores ativos na rede.
- ❑ *Invasive Species (Tamarisk)* (2014): Um sistema é organizado em *slots*. Cada *slot* pode estar vazio, ter uma planta nativa ou uma planta invasora de *Tamarisk*. O objetivo consiste em restaurar as plantas nativas com o menor esforço de trabalho.
- ❑ *Traffic* (2011 e 2014): Um agente controla o tráfego de veículos em cruzamentos de uma cidade. O objetivo é diminuir o engarrafamento de carros em determinadas vias, desafogando ou desviando o trânsito para vias alternativas.
- ❑ *Triangle Tireworld* (2014): O agente deve chegar a um dos vértices em um gráfico plano de forma triangular, mas após cada movimento há a possibilidade de ter um

pneu furado. Se isso acontecer, ele deve obter um pneu sobressalente antes de poder se mover novamente.

- *Wildfire* (2014): Em um cenário de incêndio geral a sua propagação é determinada pelos atributos clima, declive do terreno e tipo de combustível. Neste ambiente, um mapa é representado com grades onde cada grade possui atributos. O objetivo consiste em controlar o fogo.

Para cada domínio existem 10 instâncias. Estas instâncias têm diferentes níveis de complexidade.

5.3.4 Classificadores e Seleção de Características

Para realizar a escolha dos planejadores com base nas características são utilizados algoritmos de classificadores. Para esta tarefa é proposto o uso da ferramenta Weka ⁴. O Weka é um software com diversos algoritmos de aprendizado de máquina. É proposto o uso dos algoritmos *Decision Stump*, *Decision Table*, *Hoeffding Tree*, *J48*, *JRip*, *k-NN* (usando o K igual a 3, 5, 7 e 9), *KStar*, *LMT*, *Logistic*, *LWL*, *MLP*, *Naive Bayes*, *One R*, *Part*, *Random Forest*, *Random Tree*, *REP Tree*, *Simple Logistic*, *SMO* e *ZeroR*.

Para realizar o treinamento é utilizado o método de validação cruzada (*Cross-validation*). São usados os valores de *folds* iguais a 3, 5, 7 e 9. O objetivo consiste em encontrar o valor de *fold* mais adequado para cada classificador.

Para auxiliar os classificadores é proposto o uso de um AG para a seleção das melhores características. Esta técnica foi usada anteriormente por (KWON; MOON, 2007) para auxiliar o treinamento de redes neurais e seleção de atributos. Para realizar o processamento do AG é usado um sistema distribuído (KWON; MOON, 2007).

Na seleção de características do AG é preciso especificar a estrutura dos indivíduos. O indivíduo é representado por um vetor de valores booleanos. Cada posição do vetor indica se uma dada característica está presente (*true*) ou ausente (*false*) no processo de classificação.

O desempenho, também chamado de aptidão, de cada indivíduo corresponde a taxa de acerto do algoritmo de classificação. Existem variações de AGs, no trabalho é utilizado o *Steady State* (LUKE, 2013).

Cada execução tem como base o uso de um classificador. A população inicial é gerada de forma aleatória. A avaliação de cada indivíduo consiste em, seguindo as restrições de características, realizar o processo de treinamento usando o classificador. O desempenho dos indivíduos é avaliado considerando a taxa de acerto da classificação, sendo que, quanto maior a taxa de acerto melhor o indivíduo. Para o processo de seleção tem-se o método de roleta (*roulette*) (LUKE, 2013). Na fase de mutação, levando em conta o vetor de

⁴ Informações disponíveis em: <http://www.cs.waikato.ac.nz/ml/weka>

características dos indivíduos, é realizada a alteração de um valor para uma ou mais posições selecionadas aleatoriamente. Para manter a população sempre com o mesmo número de indivíduos, na reposição dos indivíduos é feito o descarte dos piores indivíduos. O critério de parada do AG é satisfeito quando um número máximo de iterações ocorre ou caso seja encontrada a solução ótima.

Metodologia

Este capítulo tem como objetivo apresentar a metodologia utilizada para a realização do trabalho, descrevendo os passos da pesquisa sobre planejamento e portfólio de planejamento voltados a aplicações de negociação de ativos do mercado financeiro. Os conceitos e técnicas propostos têm como base os trabalhos correlatos e as abordagens de planejamento apresentadas nas seções anteriores. Existem dois grandes momentos que compõem a metodologia. A primeira consiste na montagem do portfólio, conforme descrito na Seção 6.1. Posteriormente, a Seção 6.2 existe a metodologia utilizada para realizar o processo de planejamento e simulação das negociações.

6.1 Metodologia do Portfólio

Os passos utilizados para a geração do portfólio são compostos segundo a Figura 30.

No primeiro momento, em “1. Seleção dos Problemas”, é realizada a seleção dos problemas, os quais são todos aqueles que estão presentes na IPPC. Na IPPC (2011 e 2014) são utilizados 12 domínios e 10 instâncias por domínio, totalizando 120 instâncias/problemas. Mais detalhes sobre os domínios estão disponíveis na Seção 5.3.3. Em seguida, é realizada a “2. Extração das Características”. As características são aquelas apresentadas na Seção 5.3.1 e são extraídas para cada uma das instâncias. Sendo assim, para cada instância existe um grupo de características. Com base nestas características, no passo “3. Classes das Instâncias”, é associada uma classe a cada instância. Esta classe corresponde ao planejador que obteve o melhor resultado na IPPC de acordo com as instâncias.

Para realizar o treino e o teste é preciso organizar as instâncias. No passo “4. Escolha da Distribuição” é realizada a distribuição das instâncias. A primeira possível distribuição consiste em aleatoriamente organizar 2/3 das instâncias para treino e 1/3 para teste, sendo chamada de Aleatória. A segunda forma consiste em realizar a separação por domínio, sendo 2/3 dos domínios para treino e 1/3 para teste, chamando-se de *Leave One Domain Out*. Para a estratégia *Leave One Domain Out* foi realizada a separação conforme a Tabela 14.

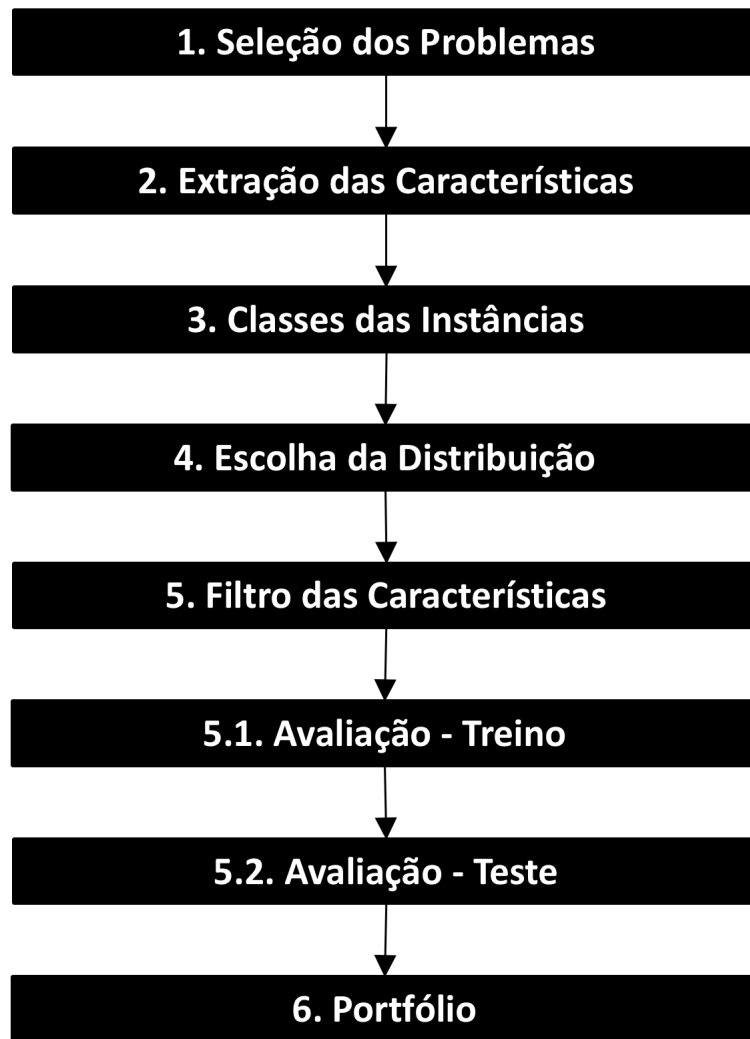


Figura 30 – Sequência de passos da metodologia usada para a montagem do portfólio.

Para cada um dos períodos da estratégia *Leave One Domain Out* foram selecionados aleatoriamente 6 domínios para treino e 2 domínios para teste. É importante salientar que a abordagem *Leave One Domain Out* se aproxima com a condição apresentada na IPPC, onde não se sabe exatamente quais problemas serão enfrentados em uma nova competição.

Tomando como base a Seção 5.3.4, o passo “5. Filtro das Características” usa um AG para selecionar as melhores características de acordo com o algoritmo de classificação. Como está sendo usado o AG *Steady State*, sua execução finaliza apenas se for ultrapassado o limite de iterações sem evolução. São avaliados diferentes números de evoluções, tamanhos da população e taxas de mutação. A taxa de mutação consiste em duas probabilidades, a primeira indicando se o indivíduo está apto para sofrer mutação e a segunda consiste na probabilidade de um gene do indivíduo sofrer mutação. O gene do indivíduo pode sofrer mutação apenas se o indivíduo está apto para sofrer mutações. Os valores para estes atributos são apresentados nos experimentos do trabalho, vide Capítulo 8.

O próximo passo consiste em realizar o cálculo da aptidão de cada indivíduo. O

Tabela 14 – Distribuição dos domínios por período usando a estratégia *Leave One Domain Out*.

	2011	2014
Treino	Crossing Traffic	Academic Advising
	Elevator	Elevator
	Game of Life	Skill Teaching
	Skill Teaching	Invasive Species (Tamarisk)
	SysAdmin	Traffic
	Traffic	Wildfire
Teste	Navigation	Crossing Traffic
	Reconnaissance	Triangle Tireworld

treino, passo “5.1. Avaliação - Treino”, utiliza a base de treino para treinar o classificador e determinar a taxa de acerto desta fase. Por sua vez, em “5.2. Avaliação - Teste”, o classificador treinado é utilizado para classificar a base de testes e determinar a respectiva taxa de acerto. Ao final, a aptidão é igual a média das taxas de acerto do treino e do teste.

Por último, no passo “6. Portfólio”, é realizada a montagem do portfólio. O portfólio consiste em, primeiro, representar o problema do mercado financeiro na linguagem RDDL. Em seguida são extraídas as características. Com base nos resultados do AG, é utilizado o melhor classificador e as características selecionadas para indicar o planejador mais adequado para o mercado financeiro.

6.2 Metodologia do Planejamento e Negociação

Os processos de planejamento e negociação seguem o conjunto de passos apresentados na Figura 31. Inicialmente, no passo “1. Carga dos Dados”, é necessário realizar a carga dos dados. Esta carga consiste em selecionar os dados importantes e permitir que estejam disponíveis de acordo com os requisitos dos algoritmos. Os dados são carregados em um banco de dados.

No momento da carga os dados devem ser “corrigidos”. Para esta tarefa foi desenvolvido um sistema de carga de dados e um web crawler¹. A primeira correção é realizada em função da ausência de negociação para uma determinada data de um ativo. Essa correção é feita copiando os dados do dia anterior. É muito importante salientar que são consideradas apenas datas em que tenham ocorrido negociações para pelo menos um ativo. Além da copia dos dados, é armazenado no banco de dados um valor indicando o percentual de correção de cada ativo e são ignorados aqueles que tenham necessitado de mais de 5% de correções por ausência de operações.

Posteriormente, é preciso corrigir os dados históricos em função dos proventos em

¹ Código fonte disponível em: <<https://github.com/bbranquinho/wpattern-wtrader>>



Figura 31 – Sequência de passos da metodologia usada na pesquisa.

ações. Para a bolsa BM&FBOVESPA os proventos em ações podem ser consultados por empresa na área de eventos corporativos². Um web crawler foi desenvolvido para capturar tais dados e armazenar no banco de dados¹. Os proventos em ações podem ser na forma de *bonificação*, *cisão*, *desdobramento*, *grupamento* e *restituição de capital*. A Tabela 15 mostra alguns proventos para a empresa Petrobras (ativos PETR3 e PETR4).

Tabela 15 – Proventos em ações da empresa Petrobras.

Proventos	Negócios até	% Bon./Desd. ou Fator de Grupamento
Desdobramento	25/04/2008	100
Desdobramento	31/08/2005	300
Grupamento	21/06/2000	100/1
Bonificação	25/03/1994	33.33333333

² Informações das empresa disponíveis em: <<http://www.bmfbovespa.com.br/cias-listadas/empresas-listadas/BuscaEmpresaListada.aspx?Idioma=pt-br>>

Os proventos em ações são colocados como sendo de *split* ou *reverse split*. Os proventos de *bonificação*, *cisão* e *desdobramento* são normalizados na forma de *split*. Já o *grupamento* é normalizado na forma de *reverse split*.

Quando as cotações estão muito elevadas, isso dificulta a entrada de novos investidores. Uma possibilidade consiste em realizar uma operação de *split*. Um dos objetivos da *split* é de dar a um ativo maior liquidez. Já um dos objetivos do *grupamento* de ações é de tentar diminuir a volatilidade dos preços dos ativos. Para o *split* e *reverse split* a normalização é feita, respectivamente, dividindo e multiplicando os dados históricos pelo fator de provento. Em tais casos o capital do investidor é mantido o mesmo. A Tabela 16 demonstra volumes e preços, respectivamente, não normalizados e normalizados do ativo PETR4 no período de 23/04/2008 a 30/04/2008.

Tabela 16 – Preço e volume do ativo PETR4 durante um período de desdobramento. São mostrados os dados não normalizados e normalizados.

Data	Preço		Volume	
	Não Normalizado	Normalizado	Não Normalizado	Normalizado
23/04/2008	R\$ 104,21	R\$ 52,105	50968213,5	101936427
24/04/2008	R\$ 102,00	R\$ 51	117105758	234211516
25/04/2008	R\$ 103,00	R\$ 51,50	108113012,5	216226025
28/04/2008	R\$ 51,85	R\$ 51,85	152439128	152439128
29/04/2008	R\$ 49,74	R\$ 49,74	160996604	160996604
30/04/2008	R\$ 50,60	R\$ 50,60	342878966	342878966

Os gráficos da Figura 32 mostram os dados de volume e preço para o ativo PETR4 não normalizados e normalizados. É importante observar que o capital do investidor continua o mesmo, a diferença é que o preço e o volume são ajustados.

O próximo passo é a “2. Seleção da Carteira de Ativos”. Essa seleção consiste em determinar quais são os ativos que serão alvo de estudos pelos algoritmos de aprendizado, porém, vale ressaltar que não necessariamente todos os ativos serão negociados. Inicialmente são pré-selecionados os ativos das empresas que têm o maior preço de mercado. Em seguida são realizados os cálculos do coeficiente de correlação simples para todos os demais ativos. Usando essa proposta são montadas diversas carteiras. Outra proposta consiste em apenas selecionar alguns ativos das empresas que têm o maior valor de mercado.

Na “3. Seleção de Algoritmos” é feita a escolha dos algoritmos de planejamento que irão compor as fases de treinamento e teste de cada experimento. Além disso, neste passo são escolhidos os parâmetros necessários para o uso de cada algoritmo.

No planejamento probabilístico é proposto o uso dos métodos POMDP e MOMDP, conforme descritos no Capítulo 5. Para ambos os casos, POMDP e MOMDP, o portfólio de algoritmos é usado para selecionar o melhor planejador.

No passo “4. Seleção das Análises Técnicas”, são escolhidas as análises técnicas e os parâmetros necessários para processá-las. As análises técnicas usadas são mostradas

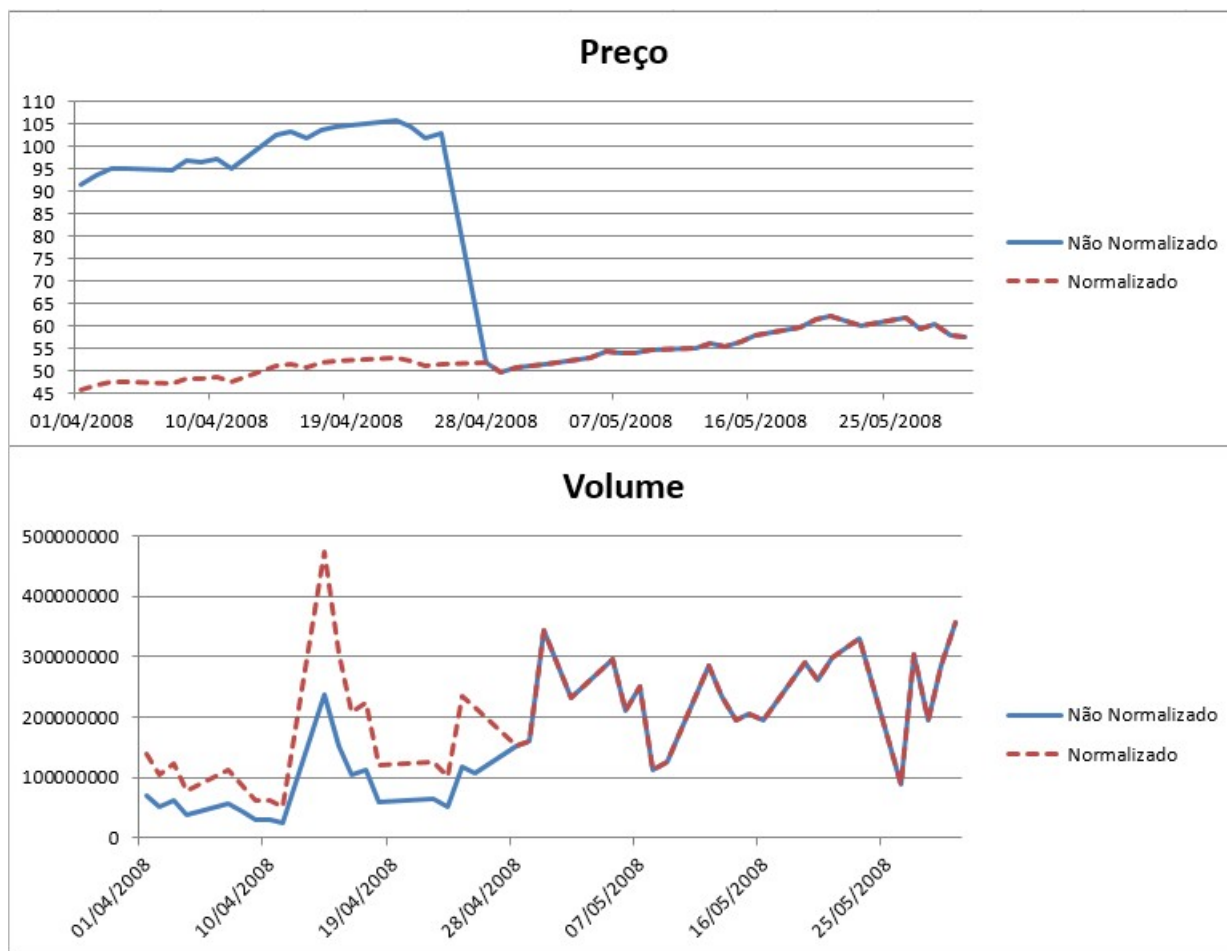


Figura 32 – Preço e volume do ativo PETR4 durante um período de desdobramento. São mostrados os dados não normalizados e normalizados.

na Tabela 17 (ACHELIS, 2006). O uso das análises dependem dos algoritmos e dos experimentos.

Certas análises geram mais de um dado, por exemplo, o MACD produz três dados, sendo eles o MACD, o sinal e o histograma. Detalhes do MACD são mostrados na Seção 3.1.3. Para processar as análises técnicas são utilizados como origem os dados do preço de fechamento e o volume. No total, são produzidos 52 valores.

O “5. Treinamento” é composto pelo treinamento e a validação. O treinamento consiste nos algoritmos realizarem os processos de aprendizado sobre uma base de treinamento. Durante o treinamento algumas abordagens adotam uma validação. Essa validação consiste em selecionar um grupo de dados posteriores ao período dos dados de treinamento. Em seguida estes dados são colocados como entrada para o(s) algoritmo(s) treinado(s). A validação permite confirmar os resultados do treinamento e em alguns casos compõe a avaliação do(s) algoritmo(s). Nesta fase é proposto o uso de AGs para a seleção dos atributos de treinamento. Este processo segue os conceitos do trabalho de Sousa (SOUSA, 2014) e Kwon (KWON; MOON, 2007) e variam de acordo com os

Tabela 17 – Lista de análises técnicas.

Análises Técnicas	
Absolute Price Oscillator	Aroon
Aroon Oscillator	Average Directional Movement Index
Average Directional Movement Index Rating	Average True Range
Balance of Power	Bollinger Bands
Chaikin Line	Chaikin Oscillator
Commodity Channel Index	Directional Movement Index
Double Exponential Moving Average	Exponential Moving Average
Hilbert Transform Dominant Cycle Period	Stochastic
Kaufman Adaptive Moving Average	Linear Regression Slope
Moving Average Convergence/Divergence	Parabolic Parabolic Stop and Reverse
Relative Strength Index (RSI)	Simple Moving Average (SMA)
Hilbert Transform Dominant Cycle Phase	Stochastic Relative Strength Index

algoritmos.

No passo “6. Teste” são realizados testes sobre os algoritmos treinados. Os dados de teste são selecionados e dados como entrada para os algoritmos treinados. Para cada amostra de teste tem-se uma predição de modo que estes valores sejam avaliados na forma de sinais de venda, compra ou estabilidade.

Na “7. Simulação” é realizada uma simulação dos negócios seguindo os sinais e realizando os cálculos de lucro ou prejuízo. Para cada operação de compra e venda existem taxas que atuam sobre o montante negociado. Dependendo do número de negócios realizados as taxas afetam diretamente os resultados de lucratividade. Sendo assim, o projeto desenvolvido permita simular negócios do mercado financeiro independente dos algoritmos e ativos. Para isso, as operações são realizadas seguindo sinais. As taxas normalmente são formadas por um preço fixo independente do valor negociado, uma taxa percentual sobre o lucro e/ou uma taxa percentual independente do custo da operação. Essas taxas podem combinar taxações fixas e percentuais. Os valores cobrados por operação variam de corretora para corretora.

Por último, “8. Resultados” resume todos os lucros e prejuízos. Neste momento são gerados os relatórios com as informações dos valores de cada operação e das taxas. Assim como na simulação, esta fase é independente dos algoritmos e dos ativos. Os resultados são comparados também em relação a estratégia *Buy & Hold* (KWON; MOON, 2003). Esta estratégia consiste em, dado um capital inicial e um período, comprar o maior número de ações possíveis no início do período e vender no final do mesmo.

Arquitetura do Projeto

Foi desenvolvido um ambiente que permite a execução de algoritmos de planejamento e aprendizado de máquina que realizam a predição de preços do mercado financeiro. Este ambiente tem como objetivo possibilitar a fácil adaptação de novos algoritmos e suportar o processamento dos dados do mercado financeiro. Neste momento estão sendo usados os dados provenientes da bolsa de valores BM&FBOVESPA. Para a inclusão de outras bolsas é necessário realizar pequenas alterações na parte de carga dos dados. Sendo assim, os algoritmos de treinamento e simulação não sofrem alterações em função do mercado financeiro.

Este capítulo está estruturado nas seguintes seções especificadas a seguir. Na Seção 7.1 é apresentada a arquitetura do projeto com base em módulos. A Seção 7.2 descreve passo a passo o que o sistema executa para realizar os experimentos, sendo descrita a fase inicial em que ocorre a carga de dados até a fase final, onde tem-se a simulação e coleta de resultados.

7.1 Arquitetura

Para realizar o desenvolvimento os seguintes princípios foram considerados durante a concepção da arquitetura: “desenvolvimento modular”, “baixo acoplamento”, “alta coesão”, “computação paralela” e “computação distribuída”. O “desenvolvimento modular” consiste em criar bibliotecas que tratem de requisitos específicos. Estes requisitos consistem nas funcionalidades necessárias para realizar a execução de experimentos de predição de preços e simulação do mercado financeiro a partir de dados históricos. O “baixo acoplamento” consiste em reduzir ao máximo a dependência entre módulos e classes, tendo o objetivo de permitir a inclusão de novos módulos e algoritmos com o menor impacto entre as partes do sistema. A “alta coesão” é usada para garantir que cada parte do projeto execute apenas uma dada função de forma bem simples e coesa. Já as “computações paralela e distribuída” têm o objetivo de permitir a otimização dos recursos computacionais para a execução dos experimentos. Estes padrões foram adotadas com o objetivo de

permitir incluir novas funcionalidades de modo que seja necessário realizar apenas o desenvolvimento de novos códigos, ou seja, sem alterar o que já está funcionando. Diversos padrões de projeto são usados, contudo, de forma geral o ambiente é fortemente orientado aos padrões inversão de controle e injeção de dependência.

Para suportar a pesquisa foram projetadas 3 diferentes aplicações estruturadas em módulos. Estas aplicação são apresentadas nas Figuras 33, 34 e 35, e correspondem respectivamente, ao processador de dados, carga de dados e planejamento.

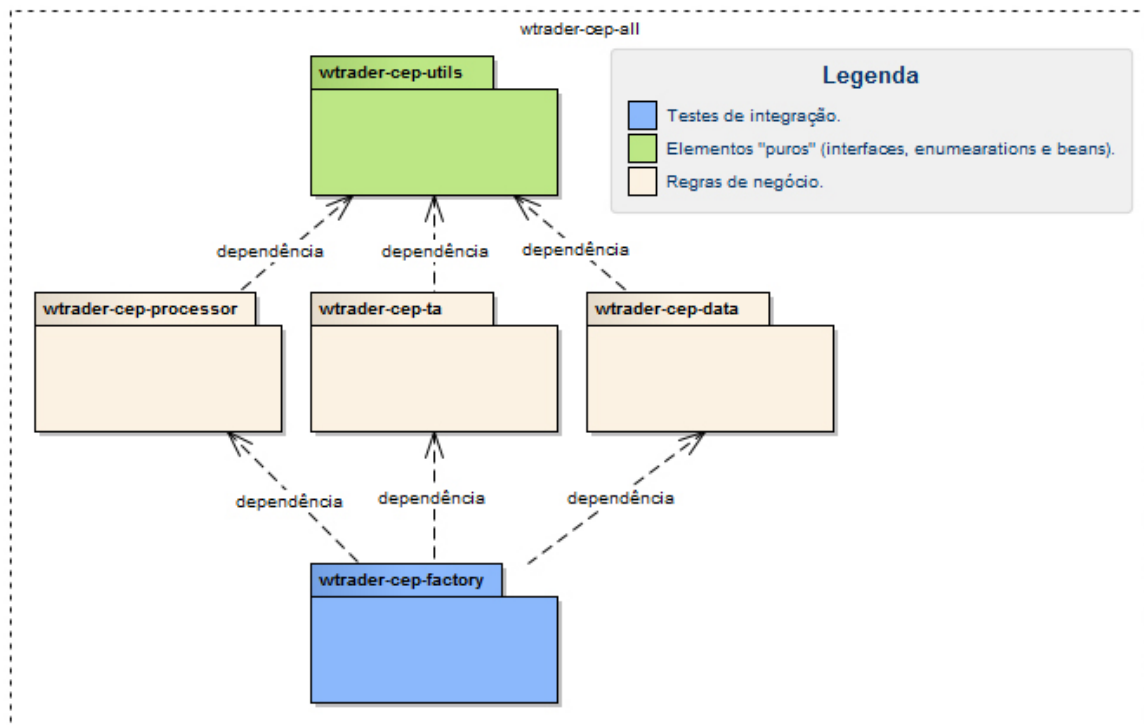


Figura 33 – Aplicação responsável pelo processamento de dados e análises técnicas.

Cada aplicação é formada por 4 tipos de módulos. O primeiro tipo são os módulos compostos por interfaces, enumerações e classes puras. As classes puras não possuem regras de negócio, ou seja, são usadas para transporte de dados e são compostas apenas por variáveis e métodos *get* e *set*. Posteriormente, existem os módulos que possuem a implementação das regras de negócio. Em terceiro estão os módulos responsáveis por testes de integração. Por último, temos os módulos responsáveis por realizar a integração de cada parte do sistema, executar a aplicação e também de realizar testes de integração. O projeto modular permite substituir alguns módulos por novas implementações, reduzindo impactos de desenvolvimento em caso de novos requisitos.

A Figura 33 apresenta o CEP (*Complex Event Processing*). O CEP é responsável por realizar todos os cálculos de análise técnica e operações que possam ser centralizadas, sendo assim usados por diferentes algoritmos. O “wtrader-cep-data” é uma das partes mais importantes, pois é responsável por acessar o banco de dados e é usado por diversos módulos. O “wtrader-cep-ta” contém as implementações das análises técnicas. O

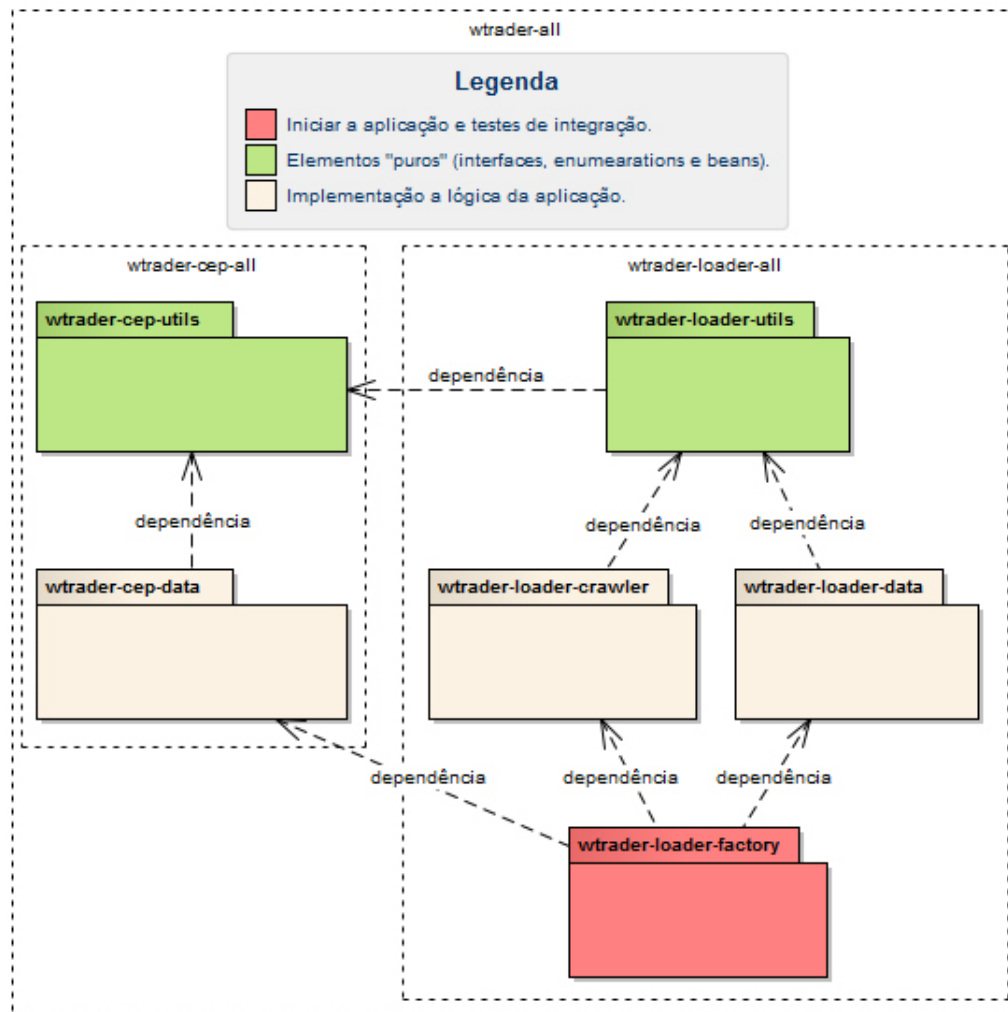


Figura 34 – Aplicação responsável por realizar a carga e correção dos dados.

“wtrader-cep-processor” é responsável por receber as requisições de dados e posteriormente retornar os valores após o processamento. Este módulo é muito importante pois seu desenvolvimento foi realizado com o objetivo de tornar simples o processamento de cálculos complexos. A requisição é composta apenas por parâmetros, ficando o módulo responsável por determinar quais são os dados necessários para os cálculos. Em geral as requisições definem uma data inicial, uma data final, uma análise técnica, os parâmetros da análise, um ativo e o tipo de dado (preço de fechamento, volume e entre outros). Em geral as análises necessitam de dados que excedem as datas. Sendo assim, é responsabilidade do “wtrader-cep-processor” e “wtrader-cep-ta” de tratar estas necessidades. Logo, a complexidade de processamento não é repassada para outras partes do projeto. Por último, o “wtrader-cep-factory” possui apenas testes unitário e de integração.

Em um trabalho futuro é sugerido o desenvolvimento de dois módulos no CEP, o *cache* e o *proxy*. O *cache* tem o objetivo de guardar dados que já tenham sido processados, evitando a repetição de cálculos. O *proxy* possui a função de agrupar as requisições de dados de modo que as porções de dados em comum sejam processadas e recuperadas

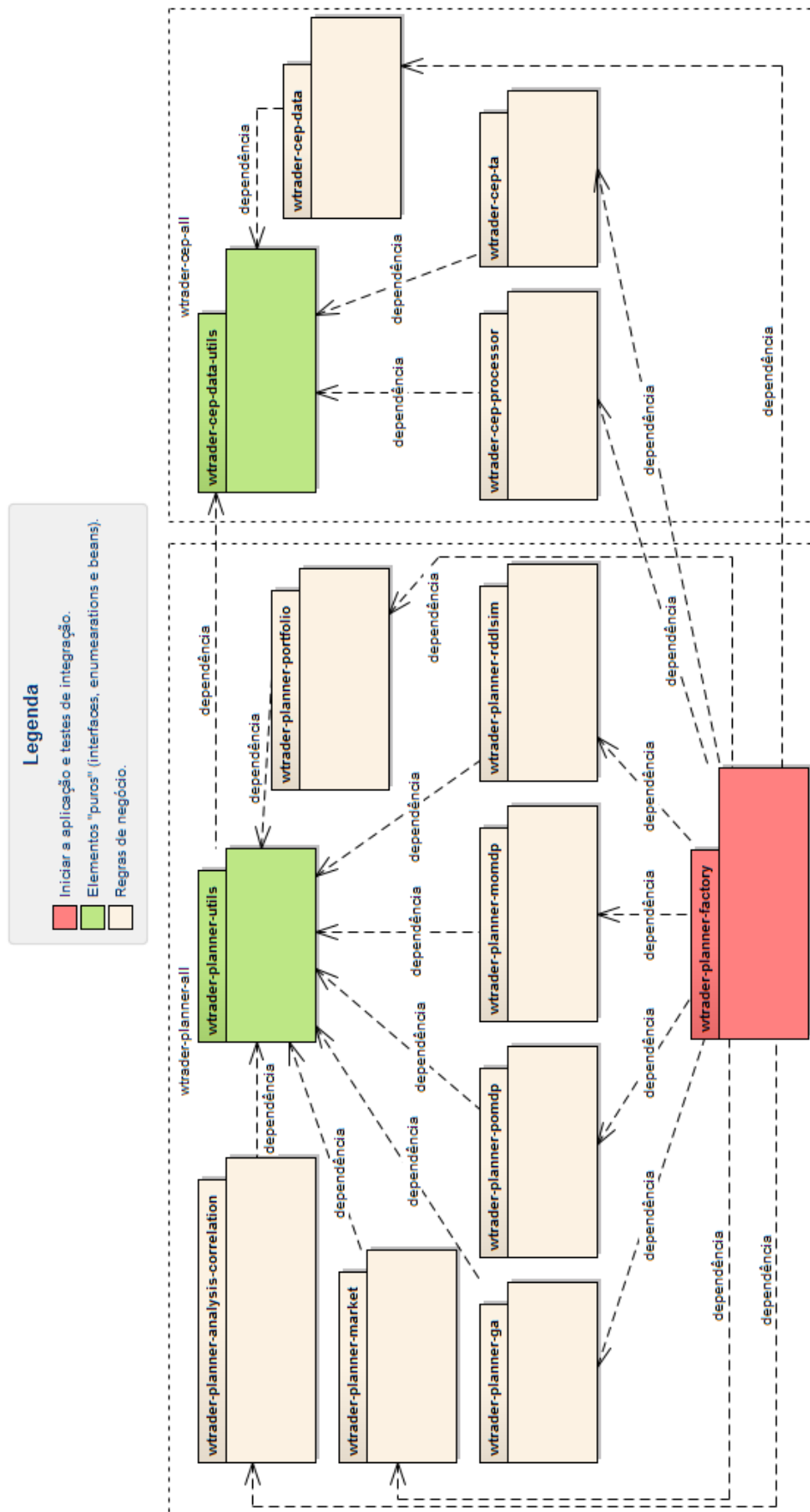


Figura 35 – Aplicação responsável por realizar as tarefas de planejamento para o mercado financeiro.

apenas uma vez. Estas otimizações contribuem para os experimentos onde os cálculos sejam pesados e a massa de dados muito grande.

A aplicação “loader”, Figura 34, é responsável por realizar a carga e correção dos dados históricos. Inicialmente, o “wtrader-loader-data” realiza a carga de arquivos contendo os históricos das cotações. Os arquivos são disponibilizados pela BM&FBOVESPA¹. Já o “wtrader-loader-crawler” acessa o site da BM&FBOVESPA e captura dados das empresas, tais como os proventos em ações. Posteriormente, estes dados são armazenados no banco de dados e usados para normalizar as cotações em função das operações de *split* e *reverse split*. Detalhes das correções são apresentados no Capítulo 6.

A parte mais importante da pesquisa, colocada na Figura 35, é a aplicação “planner”. Nestes módulos estão os algoritmos de planejamento. Inicialmente, o “wtrader-planner-analysis-correlation” possui as técnicas de correlação de ativos e montagem da carteira de investimento. Em “wtrader-planner-market” são realizadas as simulações de negociação e geração de relatórios contendo os resultados. No “wtrader-planner-ga” são colocados os códigos relativos aos AGs. Os métodos referentes aos processos de planejamento estão em “wtrader-planner-pomdp” e “wtrader-planner-momdp”. O portfólio de algoritmos e a extração de características estão presentes no “wtrader-planner-portfolio”. Em “wtrader-planner-rddlsim” está a biblioteca para processar os modelos em RDDDL. Finalmente, para iniciar a execução dos experimentos temos o “wtrader-planner-factory”.

Todo o desenvolvimento foi realizado em Java, C e C++. Para otimizar o desenvolvimento foram usadas diversas bibliotecas. Entre as principais bibliotecas destacam-se a *RDDLSim*², *Encog*³, *Neuroph*⁴, *Weka*⁵, *pomdp-solve*⁶, *APPL*⁷, *libpomdp*⁸, *AI-Toolbox*⁹, *ZMDP*¹⁰, *TA-Lib*¹¹, *Spring*, *Spring Data JPA*, *Log4j*, *JFreeChart*, *Guava*, *Apache Mina*, *Hibernate*, *C3P0*, *Apache Commons IO*, *Joda Time*, *Commons Lang3*, *MySQL* e *Apache POI*. Outras tecnologias e bibliotecas são estudadas de acordo com as necessidades do projeto.

7.2 Sistema Desenvolvido

Além dos conceitos que norteiam a arquitetura, alguns requisitos funcionais são necessário para modelar o sistema. A Figura 36 demonstra cada parte que compõe o sistema

¹ Disponível em: <<http://www.bmfbovespa.com.br/pt-br/cotacoes-historicas/FormSeriesHistoricasArq.asp>>

² Disponível em: <<https://github.com/ssanner/rddlsim>>

³ Disponível em: <<http://www.heatonresearch.com/encog/>>

⁴ Disponível em: <<http://neuroph.sourceforge.net/>>

⁵ Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>

⁶ Disponível em: <<http://www.pomdp.org/>>

⁷ Disponível em: <<http://bigbird.comp.nus.edu.sg/pmwiki/farm/appl/>>

⁸ Disponível em: <<https://github.com/dmaniloff/libpomdp>>

⁹ Disponível em: <<https://github.com/Svalorzen/AI-Toolbox>>

¹⁰ Disponível em: <<https://github.com/trey0/zmdp>>

¹¹ Disponível em: <<http://ta-lib.org/>>

e a sequência de passos executadas para a realização dos experimentos. Os passos foram separados em dois momentos. No primeiro momento é realizada a **Carga e Correção**, que consiste em carregar os dados do mercado em um banco de dados e posteriormente remover inconsistências dos dados. Estas inconsistências são descritas na metodologia apresentada no Capítulo 6, já no **Treinamento e Simulação** são realizados os procedimentos de treinar os algoritmos de aprendizado e ao final simular situações de negociação no mercado financeiro. O que difere **Carga e Correção** e **Treinamento e Simulação** é que a **Carga e Correção** ocorre apenas uma vez e o **Treinamento e Simulação** consiste nos experimentos e simulações.

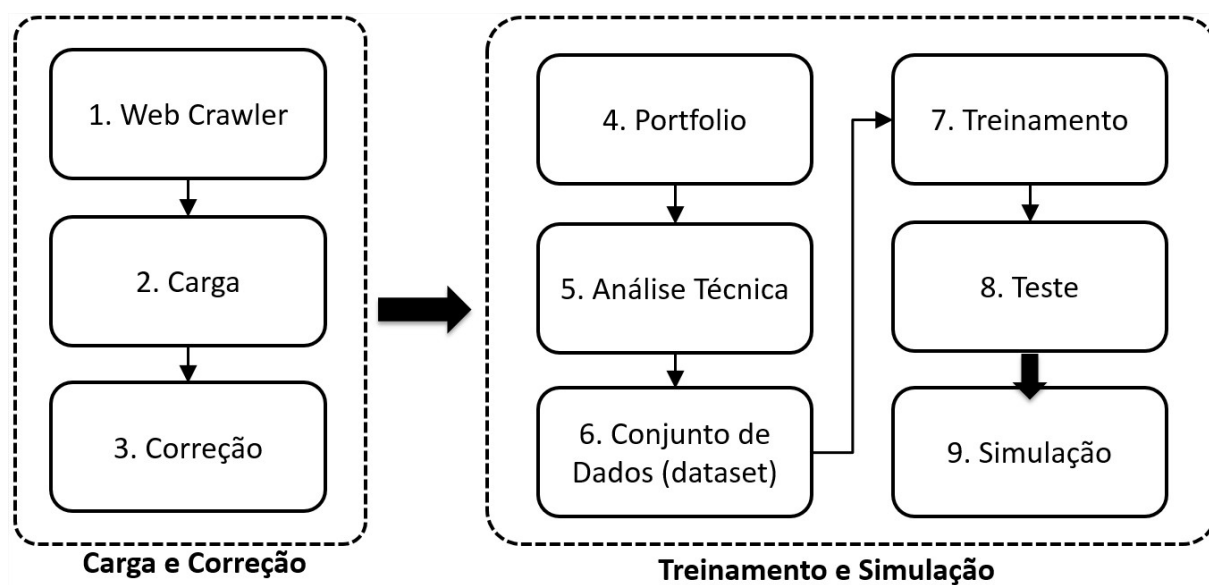


Figura 36 – Elementos que compõem o ambiente.

Primeiramente, antes de realizar treinamentos e simulações é preciso que os dados sejam carregados e corrigidos. Para realizar a carga dos dados inicialmente um *Web Crawler* (1. *Web Crawler*) realiza a captura de dados na web. Os dados capturados pelo *Web Crawler* correspondem a informações sobre as empresas, tais como a lista de ativos, início das atividades na bolsa, data de início das negociações em lotes unitários e proventos em ações.

Uma vez que os dados das empresas foram armazenados em um banco de dados, é realizada a carga dos arquivos (2. *Carga dos Dados*) que contêm as cotações históricas. Estas cotações contemplam apenas os dados diários. Para se ter acesso aos dados de todas as negociações durante todo um dia se faz necessário entrar em contato com a bolsa ou empresas especializadas em prover dados de negociação. O histórico diário normalmente pode ser obtido no site dos mercados¹² ou através de serviços de dados

¹² O histórico de cotações da BM&FBOVESPA e a respectiva formatação dos arquivos históricos estão disponíveis em: <<http://www.bmfbovespa.com.br/pt-br/cotacoes-historicas/FormSeriesHistoricas.asp>>

de negociação¹³. Neste momento apenas os dados diários de fechamento de pregão estão sendo armazenados. Contudo, poucas alterações são necessárias para adaptar o ambiente para trabalhar com intervalos menores, tais como 5, 10 ou 15 minutos entre cada negociação. Neste momento os dados por ativo que são armazenados são o preço de abertura, maior preço negociado, menor preço negociado, média do preço de negociação, preço de fechamento, maior preço de oferta de compra, menor preço de oferta de venda, número de negócios efetuados, quantidade total de ações negociadas, volume total negociado e a data do registro.

Uma vez que todo o processo de carga foi realizado a próxima fase, *3. Correção dos Dados*, consiste na correção da base de dados. Essa correção é descrita no Capítulo 6 (Metodologia). Em resumo, os registros são normalizados em função dos proventos em ações (*split* e *reverse split*) e da ausência de registros de negociação. Um dado que é armazenado no banco de dados para cada ativo é o percentual de correções decorrentes da ausência de registros. A **Carga e Correções** é realizada apenas uma vez no início dos experimentos, sendo executada novamente apenas se for necessário incluir novos dados.

Em seguida temos o **Treinamento e Simulação**. Neste momento é realizado o treinamento, as simulações de negociação e avaliação de resultados.

Na **4. Análise do Portfólio** é montada uma carteira de ativos e o portfólio de planejamento. Para auxiliar na definição de quais ativos usar foram desenvolvidos os cálculos de coeficiente de correlação simples e coeficiente de correlação relativa, explicados na Seção 3.2. O cálculo de correlação pode ser feito sobre um grupo de ativos predeterminados ou considerando todos os ativos que tenham uma certa data de início de negociação. É importante salientar que neste momento é possível limitar o cálculo para aqueles ativos que tenham um certo percentual de correções. Neste momento o portfólio de planejamento seleciona os domínios, instâncias e planejadores que iram compor o portfólio de planejamento.

Posteriormente, na **5. Seleção dos Atributos** são selecionados os elementos que farão parte do treinamento, teste e simulação. Entre eles são determinadas as características dos problemas de planejamento, planejadores, algoritmos de treinamento, os parâmetros de cada algoritmo, período dos dados de treinamento, período dos dados de validação, período dos dados de teste, análises técnicas, parâmetros de cada análise técnica e parâmetros de configuração do ambiente.

Já no passo **6. Conjunto de Dados (Dataset)** todos os dados necessários para o treinamento, validação e teste, são recuperados ou gerados. Em outras palavras, de acordo com os períodos selecionados os dados são recuperados do banco de dados e a partir deles são realizados os cálculos das análises técnicas.

A base do **7. Treinamento** são os algoritmos usados para o treinamento, sendo

¹³ O *Yahoo Finance* fornece dados históricos de negociação de diversos mercados do mundo. Mais informações estão disponíveis em: <<http://finance.yahoo.com/market-overview>>

aqueles escolhidos na **5. Seleção dos Atributos**. No treinamento o AG está sendo usado para a seleção de dados de entrada. Para processar os experimentos foi realizado o desenvolvimento de um sistema que permite realizar o processamento em paralelo e distribuído. A Figura 37 mostra uma abstração de como está estruturado este ambiente.

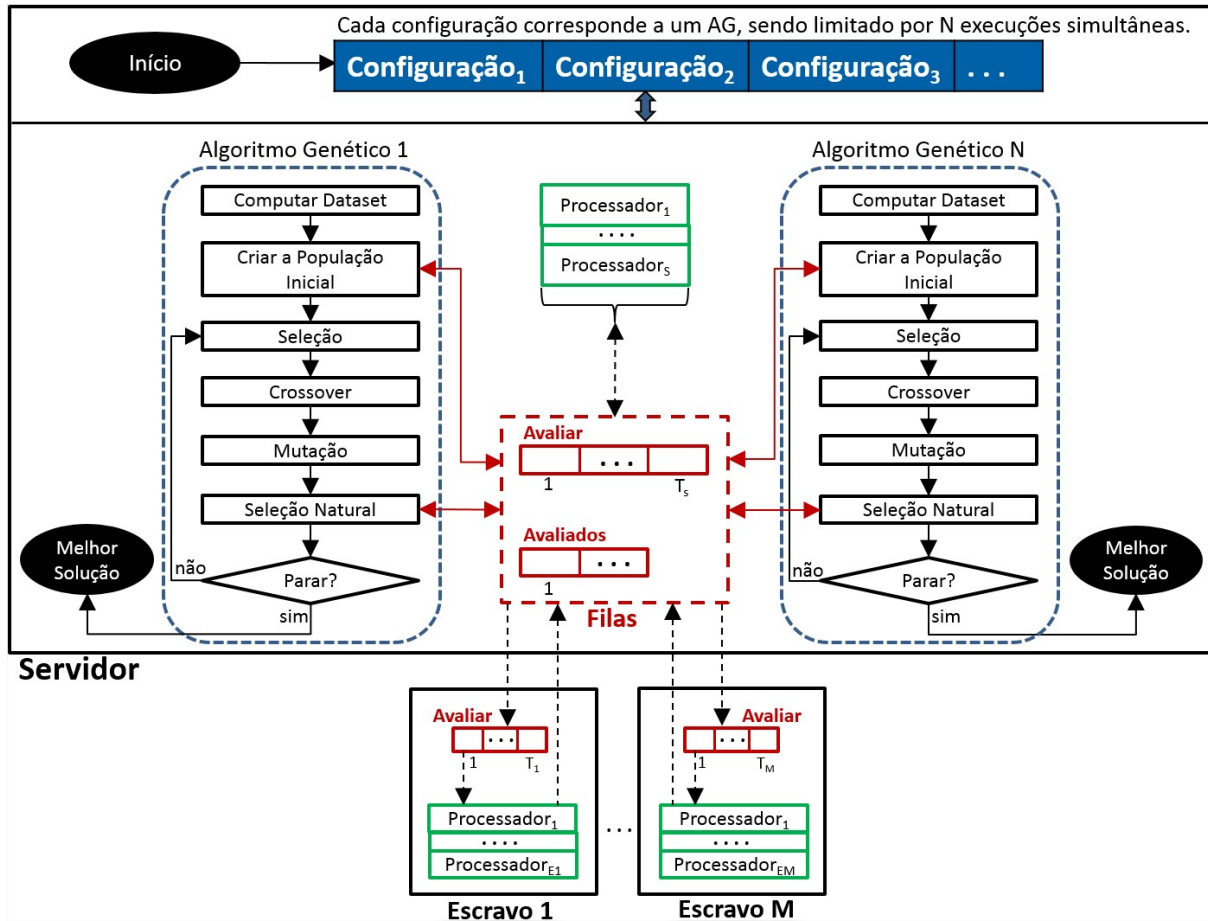


Figura 37 – Sistema projetado para processar os AGs em paralelo e distribuído.

O processamento do AG é realizado de forma distribuída e paralela. Todas as partes que compõem o ambiente foram projetadas para serem facilmente adaptadas para a inclusão de novos algoritmos. O ciclo principal do AG é processado no *servidor* e as aptidões são calculadas nos escravos e/ou no servidor. É usado apenas um servidor e vários escravos. O servidor é responsável por gerenciar o que é processado e o fluxo de dados entre todas as partes. Os escravos realizam apenas o processamento, “não” tendo conhecimento do que é processado e para quem se destinam os resultados. A qualquer momento novos escravos podem se conectar ao servidor e auxiliar na tarefa de processamento. O servidor e os escravos normalmente estão em uma rede local. Sendo assim, em geral são raros os casos de perda de conexão. Contudo, se um escravo perder a conexão com o servidor é dado um intervalo de tempo para ele se reconectar, caso ele não se reconecte durante este período o servidor recupera as mensagens daquele escravo e os repassa para os outras

máquinas.

Uma evolução em relação aos trabalhos correlatos é que o sistema permite paralelizar as instâncias de AGs. Foi criada uma fila que guarda um conjunto de configurações. Cada configuração representa um experimento centrado em um AG. Um parâmetro N limita a quantidade de configurações em execução e pode ser alterada durante a execução.

A troca de dados entre o servidor e os escravos pode ser síncrona ou assíncrona. Isto significa que ao chegar ao final da evolução o AG pode ou não esperar os últimos indivíduos serem processados. Caso seja sincronizado ao chegar ao final o AG espera pelos indivíduos, se o AG receber bons resultados ele pode retomar o processo de evolução.

O sistema é flexível e permite a inclusão de diferentes tipos de AGs. Até o momento foi desenvolvido apenas o *AG Steady State*. O *AG Steady State* é ideal para o processamento em paralelo, já que a cada ciclo são gerados dois novos indivíduos e não é necessário esperar a avaliação deles para que um novo ciclo seja executado. Contudo, para evitar que indivíduos sejam criados discriminadamente é realizada uma configuração que limita o número de indivíduos sendo processados. Este controle é feito através do bloqueio dos AGs enquanto existirem T_s indivíduos na fila “Avaliar”. A fila *Avaliar* armazena todos os indivíduos que precisam ter a aptidão calculada. Toda vez que um “Processador” está disponível ele retira um elemento da fila *Avaliar* para que realize o seu processamento. Uma vez que tenha finalizado um processamento, o “Processador” repassa os resultados para a fila “Avaliados”. Durante as fases “Criar a População Inicial” e “Seleção Natural” são enviados os indivíduos para a fila “Avaliar” e também são recuperados os indivíduos que tenham sido processados. O número de “Processadores” é configurável, tanto no servidor como nos escravos, logo, múltiplas avaliações são realizadas em uma mesma máquina. Por padrão o quantidade de “Processadores” é igual ao número de *threads* da máquina.

Após o treinamento tem-se a fase de **8. Teste**. Esta fase consiste em utilizar os dados de testes para avaliar como são os resultados dos algoritmos treinados. Utilizando os resultados de saída dos testes são gerados sinais de *Alta*, *Baixa* e *Estabilidade*. De posse dos sinais o passo **9. Simulação** realiza simulações de negociações no mercado financeiro. Nestas simulações os sinais são usados para realizar eventos de compra e venda de ações. Ao final destes negócios tem-se como resultado se os investimentos tiveram lucro ou prejuízo.

Uma característica importante já desenvolvida nas simulações são as taxações de corretagem cobradas sobre cada operação. Para cada negócio realizado são descontadas tais taxas. Cada experimento pode usar diversas taxas de corretagem e diferentes valores de investimento. Quanto maior a taxação de corretagem e menor o investimento, mais difícil de se obter lucro.

A *Folhainvest*¹⁴ promove a competição de investimentos na bolsa BM&FBOVESPA

¹⁴ Mais detalhes disponíveis em: <<http://folhainvest.folha.com.br/>>

usando dados reais do mercado em um ambiente simulado de negociação. A Tabela 23 mostra um exemplo de taxa de corretagem, sendo esta à adotada pela *Folhainvest*.

Tabela 18 – Exemplo de taxa de corretagem cobrada sobre cada ordem de compra ou venda.

Valor das Operações	Corretagem
Até R\$ 135,07	R\$ 2,70
Acima de R\$ 135,08 até R\$ 498,62	2%
Acima de R\$ 498,63 até R\$ 1.514,69	1,5% + R\$ 2,49
Acima de R\$ 1.514,70 até R\$ 3.029,38	1% + R\$ 10,06
Acima de R\$ 3.029,39	0,5% + R\$ 25,21

As taxas podem ser compostas por um valor fixo ou percentual conforme o valor da operação. É válido enfatizar que esta taxa incide sobre ambos os lados, ou seja, tanto para quem compra e assim como para quem está vendendo.

Experimentos e Discussão dos Resultados

Os experimentos foram separados em duas partes. Cada uma destas partes têm diferentes pontos de avaliações e objetivos. A primeira parte, presente na Seção 8.1, consiste no portfólio de planejamento. Já a Seção 8.2 corresponde aos experimentos com planejadores aplicados no mercado financeiro. É importante citar que a metodologia da pesquisa é descrita no Capítulo 6 e os experimentos foram executados em um Pentium i7-4930K (6 núcleos, 12 threads, 3.4GHz com turbo de 3.9GHz), 16GB de ram - 2400MHz, 250 GB SSD - RAID 0 e placa mãe ASUS X79-DELUXE.

Cada experimento consiste em duas fases. A primeira consiste em montar o portfólio de planejamento, ou seja, escolher o(s) planejador(es). Na segunda fase o(s) planejador(es) selecionados no portfólio são utilizados para seleção de ações do mercado financeiro, com base nas ações são simuladas negociações e os resultados de lucro ou prejuízo são obtidos. Detalhes dos experimentos são colocados nas Seções 8.1 e 8.2.

8.1 Portfólio de Planejamento

Os experimentos para a criação do portfólio de planejamento levam em conta a avaliação dos resultados em relação ao percentual de acerto dos algoritmos de classificação para as instâncias dos problemas da IPPC. Para criar a base de dados são consideradas as instâncias dos problemas da competição IPPC e da forma da distribuição *Aleatório* ou *Leave one Domain Out*).

Os melhores experimentos para cada período e distribuição, conforme a Tabela 19, foram selecionados e nomeados como *A*, *B*, *C*, *D* e *E*. Os atributos utilizados para a seleção de características de cada experimento são listados na Tabela 20. O experimento *E* corresponde ao classificador que obteve a melhor média de classificação considerando todos os períodos e distribuições.

Tabela 19 – Períodos e distribuições dos experimentos *A*, *B*, *C*, *D* e *E*.

	Período	Distribuição
A	2011	Aleatório
B	2014	Aleatório
C	2011	Leave One Domain Out
D	2014	Leave One Domain Out
E	Média	

Tabela 20 – Atributos dos experimentos *A*, *B*, *C* e *D*.

	Classificador	Tam. População	N° Iterações	N° Folds	% Mutação	
					Indivíduo	Gene
A	LMT	50	200	3	15	25
B	Part	150	200	3	5	25
C	MLP	50	200	3	5	15
D	Logistic	50	200	5	15	5

A Figura 38 apresenta os melhores classificadores para os 4 grupos de dados, sendo mostrados os valores percentuais médio, de treino e de teste. O valor médio corresponde a taxa de acerto média obtida no AG considerando os percentuais de classificação obtidos nas fases treino e de teste.

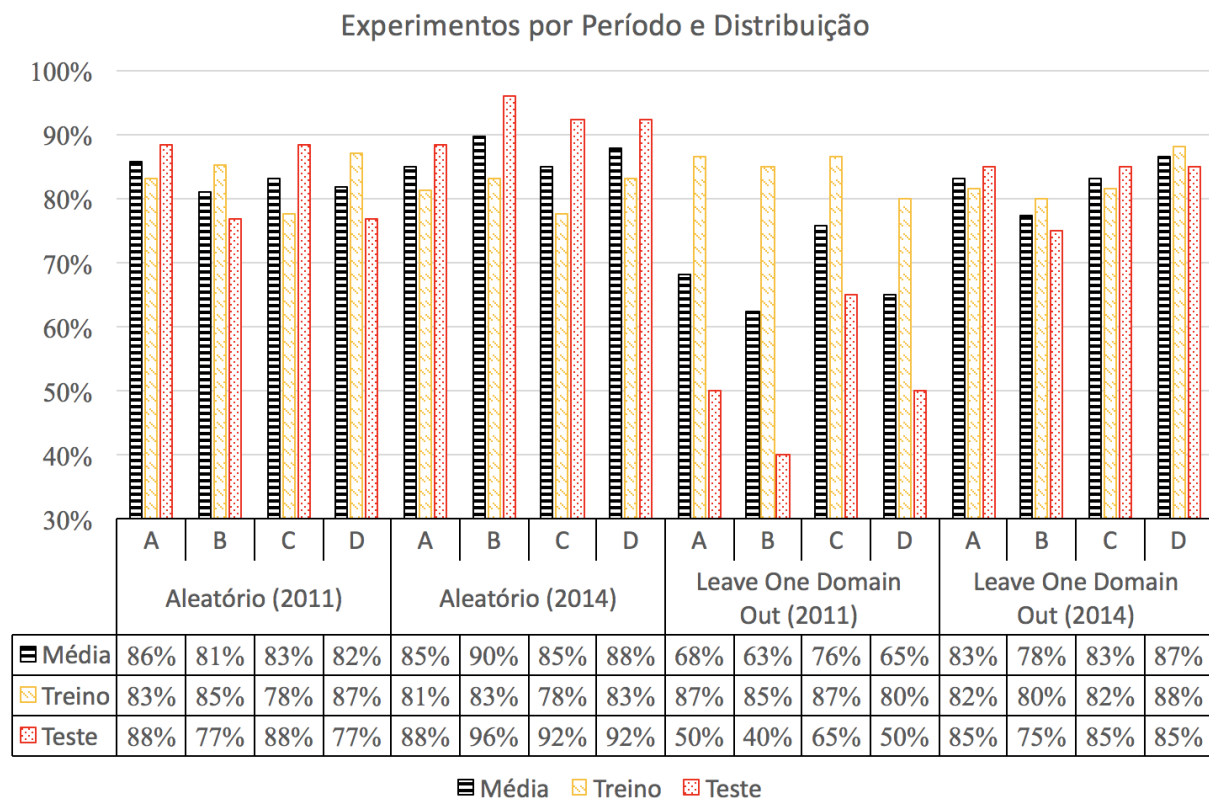


Figura 38 – Experimentos por período e distribuição.

Para *Aleatório (2011)* o melhor experimento foi o *A*, tendo obtido o percentual de

Tabela 21 – Planejadores selecionados através do Portfólio de Planejamento conforme os experimentos.

	Aleatório (2011)		Aleatório (2014)	
	1 Ativo	2 Ativos	1 Ativo	2 Ativos
A	POMDPX_NUS	POMDPX_NUS	NUS-POMDPGroup	NUS-POMDPGroup
B	POMDPX_NUS	POMDPX_NUS	NUS-POMDPGroup	NUS-POMDPGroup
C	POMDPX_NUS	POMDPX_NUS	NUS-POMDPGroup	NUS-POMDPGroup
D	POMDPX_NUS	POMDPX_NUS	NUS-POMDPGroup	NUS-POMDPGroup

	Leave One Domain Out (2011)		Leave One Domain Out (2014)	
	1 Ativo	2 Ativos	1 Ativo	2 Ativos
A	POMDPX_NUS	POMDPX_NUS	NUS-POMDPGroup	NUS-POMDPGroup
B	POMDPX_NUS	POMDPX_NUS	NUS-POMDPGroup	NUS-POMDPGroup
C	POMDPX_NUS	POMDPX_NUS	NUS-POMDPGroup	NUS-POMDPGroup
D	POMDPX_NUS	POMDPX_NUS	NUS-POMDPGroup	NUS-POMDPGroup

acerto de 85,90%. Apesar de *A* ter apresentado bons resultados para *Aleatório (2014)*, o melhor conjunto de atributos neste caso é o *B* com 89,74%. De modo geral, os percentuais de acerto para *Leave one Domain Out (2011)* foram inferiores aos demais, tendo como melhor experimento o *C* com 75,83%. Finalmente, para *Leave One Domain Out (2014)* o *D* obteve um percentual de 85,83%. Por último, o valor percentual de *E* é igual a 81,83%.

Nas fases de treino e teste em alguns casos foram obtidos valores discrepantes de acordo com cada grupo. O caso mais importante é o *Leave one Domain Out (2011)*, já que durante o treino foram encontrados excelentes valores e no teste as classificações não foram tão boas. Isto é o que torna os resultados deste grupo discrepante em relação aos demais. Logo, é possível dizer que isso ocorre quando um domínio de teste é muito diferente daqueles utilizados no treino. Já para os demais grupos os resultados de teste foram muitas vezes melhores que o treino, em especial para o *Aleatório (2014)*.

Foram feitas as modelagens booleanas do problema do mercado financeiro para um e dois ativos utilizando a linguagem RDDDL, tal como o exemplo do Apêndice C. Com base nestas modelagens são extraídas as características. Estas características são dadas como entrada para os classificadores utilizados nos experimentos, ao final é dada a indicação de qual planejador deve ser utilizado. O resultado das indicações de planejadores é mostrado na Tabela 21.

O “*POMDPX*” é indicado como o melhor planejador para ambas as modelagens. É importante ressaltar que o *POMDPX_NUS* e o *NUS-POMDPGroup* correspondem ao mesmo grupo de pesquisa, sendo aqui chamado de *POMDPX*¹. Isto indica que até o momento o planejador mais adequado para o mercado financeiro é o *POMDPX*. Contudo, além do *POMDPX* os demais classificadores e grupos podem ter selecionado outros planejadores. Estes múltiplos planejadores podem ser combinados para gerar uma solução

¹ Detalhes em: <http://bigbird.comp.nus.edu.sg/pmwiki/farm/appl/index.php?n=Main.HomePage>

em conjunto. Uma possibilidade de tratar este cenário consiste em realizar o agendamento com base no número de vezes que o planejador foi selecionado. Esta abordagem tem como base os conceitos apresentados em (SEIPP et al., 2013) e (SOUSA, 2014). A Tabela 22 mostra o portfólio de planejamento conforme o percentual de vezes que cada planejador foi selecionado para o mercado financeiro usando 1 e 2 ações.

Tabela 22 – Percentual de seleção dos planejadores.

	Aleatório (2011)		Leave One Domain Out (2011)	
	1 Ativo	2 Ativos	1 Ativo	2 Ativos
POMDPX_NUS	85%	85%	85%	80%
KAIST-AILAB	10%	10%	15%	20%
HyPlanClient	5%	5%	0%	0%

	Aleatório (2011)		Leave One Domain Out (2011)	
	1 Ativo	2 Ativos	1 Ativo	2 Ativos
NUS-POMDPGroup	90%	90%	90%	80%
NUS-KAIST_AIPR_LAB	10%	10%	10%	20%

É importante lembrar que foram utilizados 20 diferentes tipos de classificadores e estes dados percentuais correspondem as indicações considerando tais algoritmos.

8.2 Planejamento

São descritos nesta Seção os experimentos de negociação para o mercado financeiro utilizando planejamento. São usados diferentes planejadores, parâmetros de configuração, valores de capital, taxas e ativos. Cada experimento é baseado em um período de treinamento e um período de teste.

Os ativos selecionados para os experimentos são o CMIG4, CRUZ3, ITSA4, PETR4 e VALE5. Durante a fase de testes, nas simulações de negociações, são utilizados diferentes valores de capital de investimento e taxas. O capital de investimento é dado pelos valores de R\$ 10.000,00, R\$ 25.000,00, R\$ 50.000,00, R\$ 100.000,00 e R\$ 500.000,00. A taxa de corretagem é cobrada em cada operação de compra e venda. Os valores costumam variar de acordo com o tipo de mercado (à vista, fracionário, ...) e do tipo de operação (*day trade* e *normal trade*). O *day trade* ocorre quando a compra e a venda ocorrem no mesmo dia. Já o *normal trade* é quando a compra e a venda são realizadas em dias diferentes. O mercado fracionário, no qual as ações são compradas separadamente e não em lotes de cem ou mil, a taxa fixa pode sair muito mais cara. Este trabalho opera apenas com lotes de 100. A Tabela 23 mostra os 3 tipos de taxas de corretagem. A variável c define o valor de negociação em uma operação. As taxas são formadas por um valor percentual e/ou um valor fixo de acordo com c . A primeira forma de tributação foi baseada em dados de mercado da simulação Folhainvest².

² Disponível em: <<http://folhainvest.folha.com.br>>

Tabela 23 – Tipos de taxas de corretagem.

Nº Taxa	Valor da Operação	Taxa
1	$c \leq R\$ 135.07$	R\$ 2.70
	$R\$135.08 < c \leq R\498.62	2%
	$R\$498.63 < c \leq R\$1,514.69$	1.5% + R\$ 2.49
	$R\$1,514.70 < c \leq R\$3,029.38$	1% + R\$ 10.06
	$c > R\$3,029.39$	0.5% + R\$ 25.21
2	c	R\$ 9.90
3	c	R\$ 20.00

Além de avaliar os resultados obtidos entre as diferentes técnicas, os experimentos são comparados com a estratégia *Buy & Hold*. Esta estratégia consiste em, dado um capital inicial e um período de tempo, adquirir o maior número de ações possíveis no início do período e ao final do período é realizada a venda do mesmo.

Em relação ao uso de múltiplos ativos, são formados dois grupos de ativos. O primeiro grupo é formado pelos principais ativos, ou seja, CMIG4, CRUZ3, ITSA4, PETR4 e VALE5. O segundo grupo é formado pelos ativos mais correlacionados a cada ativo do primeiro grupo. Para isso é utilizado o cálculo de correlação simples para estabelecer as correlações. Não são realizadas simulações com os ativos do segundo grupo, já que eles usados apenas para confirmar os sinais do primeiro grupo. O cálculo de correlação é realizado para cada experimento e respeita os dados da fase de treino. A Tabela 24 mostra a correlação de ativos. É importante enfatizar que para cada ativo as correlações variam de acordo com os períodos.

Tabela 24 – Correlação de ativos para cada período de treino.

Ano	Ordem de Correlação	CMIG4	CRUZ3	ITSA4	PETR4	VALE5
2007	1º	SHUL4	VALE3	FBMC4	EALT4	GETI3
	2º	TIBR6	GETI3	USIM3	INEP3	USIM3
2008	1º	SHUL4	VALE3	FBMC4	EALT4	USIM3
	2º	TIBR6	USIM3	USIM3	INEP3	ITSA4
2009	1º	SHUL4	VALE3	FBMC4	EALT4	ITSA4
	2º	TIBR5	USIM3	USIM3	INEP3	USIM3
2010	1º	SHUL4	VALE3	FBMC4	EALT4	ITSA4
	2º	TIBR5	PNVL3	TRPL3	INEP4	USIM3
2011	1º	SHUL4	PNVL3	FBMC4	EALT4	ITSA4
	2º	TIBR5	VALE5	TRPL3	INEP4	USIM3
2012	1º	SHUL4	PNVL3	VALE5	EALT4	ITSA4
	2º	TIBR5	SCAR3	FBMC4	INEP4	FBMC4
2013	1º	SHUL4	PNVL3	TRPL3	EALT4	ITSA4
	2º	TIBR5	SCAR3	VALE5	INEP4	FBMC4
2014	1º	SHUL4	PNVL3	TRPL3	EALT4	FBMC4
	2º	TKNO4	SCAR3	VALE5	INEP4	ITSA4

Para o treinamento, a data inicial de todos os experimentos é 01/01/2000 e a data

final corresponde ao dia anterior ao início da fase de testes. A fase de testes é dada pelo período de um ano, cobrindo os anos de 2008 a 2015, totalizando assim 8 experimentos por ativo. Contudo, é preciso enfatizar que para o último experimento a data final é dada por 11/06/2015. No início da fase de treino, um modelo POMDP ou MOMDP é construído a partir de dados históricos. O processo de planejamento é realizado por um período máximo de 6 horas ou até que sua convergência ocorra. O fator de desconto é igual a 0,95. Os estados que compõem a modelagem são formados por tendências com períodos de 5 dias. A simulação é realizada usando os dados de teste e seguindo as indicações de ações da fase de treinamento. Todas as operações são realizadas considerando o preço de abertura. Ao final uma avaliação do resultado é executada comparando capital inicial e o capital final, com base nesta comparação é aferido o lucro ou o prejuízo.

De modo geral, os experimentos em que ocorreram muitas negociações têm uma direta influência das taxas. Quanto maior o número de operações, piores são os resultados, especialmente aqueles sob a taxa 1. De modo geral, a taxa 1 teve um grande impacto nos rendimentos. Um número considerável de casos envolvendo lucro com as taxas 2 e 3, tornaram-se em prejuízo com a taxa 1. Os resultados percentuais entre as taxas 2 e 3 foram muito próximos. Além disso, para os valores de capital superiores a R\$ 25.000,00 os percentuais de rendimento são muito próximos. Já com o valor de R\$ 10.000,00 existe uma diferença considerável dos percentuais de rendimento em relação aos demais capitais.

A Tabela 25 mostra uma comparação dos resultados obtidos entre as taxas de número 1 e 2, com o planejador SARSOP, modelagem POMDP, para o ativo CMIG4 e com um capital inicial de R\$ 25.000,00.

Tabela 25 – Comparação dos resultados obtidos com o SARSOP, modelagem POMDP e taxas 1 e 2 para o ativo CMIG4.

Ano	Anual				Acumulado			
	Buy & Hold		MACD 2		Buy & Hold		MACD 2	
	Taxa 1	Taxa 2	Taxa 1	Taxa 2	Taxa 1	Taxa 2	Taxa 1	Taxa 2
2008	-3%	-2%	-40%	-22%	-3%	-2%	-40%	-22%
2009	20%	21%	-10%	16%	17%	19%	-51%	-5%
2010	-6%	-6%	-12%	11%	10%	13%	-63%	6%
2011	18%	19%	-10%	9%	28%	33%	-73%	15%
2012	-15%	-14%	-10%	10%	14%	19%	-83%	25%
2013	-8%	-7%	-27%	-13%	6%	12%	-110%	11%
2014	-4%	-3%	1%	27%	2%	9%	-109%	38%
2015	5%	6%	-6%	4%	7%	15%	-115%	42%

A estratégia *Buy & Hold* é pouco afetada pelas taxas, uma vez que apenas duas operações são realizadas no período de um ano. Por outro lado, quando utiliza-se o planejamento proposto, os resultados foram diretamente influenciados pelas taxas. Por exemplo, para MACD 2 em 2012 um lucro de 10% com a taxa 2 torna-se em um prejuízo de 10% ao utilizar a taxa 1. No contexto geral os resultados sofreram um impacto direto pela taxa 1, em muitos casos o lucro não foi obtida devido a estes encargos. Os rendimentos

acumulados para um dado ano correspondem a soma dos resultados passados até aquele período. Analisando os valores acumulados para as taxas 2 e 1, respectivamente, 42% de lucro torna-se em um prejuízo de 115%.

Os ativos CMIG4, PETR4 e VALE5 obtiveram melhores resultados com o SARSOP comparado a *Buy & Hold*. Já para CRUZ3 e ITSA4, a estratégia *Buy & Hold* se saiu melhor. Contudo, analisando os casos onde *Buy & Hold* mostram melhores resultados, o SARSOP foi capaz de obter rendimentos próximos ou até melhores ao se considerar apenas os períodos a partir de 2011. A Tabela 26 mostra os resultados para CRUZ3 e ITSA4 no período a partir de 2011.

Tabela 26 – Resultados para CRUZ3 e ITSA4 com capital inicial de R\$ 25,000.00, utilizando a taxa número 2, o planejador SARSOP e os períodos a partir de 2011.

		Anual				Acumulado			
		B. & H.	OST 2	MACD 2	RSI 2	B. & H.	OST 2	MACD 2	RSI 2
CRUZ3	2011	23.75%	-9.23%	-4.99%	-14.87%	23.75%	-9.23%	-4.99%	-14.87%
	2012	29.28%	3.35%	-16.14%	-35.45%	53.03%	-5.88%	-21.13%	-50.32%
	2013	-21.02%	36.04%	36.1%	20.86%	32.01%	30.16%	14.97%	-29.46%
	2014	-17.44%	41.01%	71.01%	6.5%	14.57%	71.17%	85.98%	-22.96%
	2015	31.12%	-25.52%	-35.8%	0.87%	45.69%	45.65%	50.18%	-22.09%
		Anual				Acumulado			
		B. & H.	OST 2	MACD 2	RSI 2	B. & H.	OST 2	MACD 2	RSI 2
ITSA4	2011	-15.57%	-28.06%	7.67%	9.32%	-15.57%	-28.06%	7.67%	9.32%
	2012	-6.02%	-18.14%	-38.67%	11.04%	-21.59%	-46.2%	-31%	20.36%
	2013	-0.7%	-7.4%	2.36%	3.01	-22.29%	-53.6%	-28.64%	23.37%
	2014	17.32%	-15.47%	8.88%	-19.49%	-4.97%	-69.07%	-19.76%	3.88%
	2015	3.18%	7.75%	-18.95%	-3.34%	-1.79%	-61.32%	-38.71%	0.54%

As análises técnicas têm diferentes resultados de acordo com os ativos utilizados. Geralmente, a análise técnica mantém os resultados para os diferentes períodos de um ativo. Em outras palavras, uma análise técnica que mostra bons ou maus resultados para um período de tempo tende a manter esse tipo de comportamento para os demais períodos. A Tabela 27 mostra os resultados para PETR4, usando o SARSOP, modelagem POMDP, capital inicial de R\$ 25.000,00 e a taxa 2.

Tabela 27 – Resultados para PETR4, usando o SARSOP, modelagem POMDP, capital inicial de R\$ 25,000.00 e a taxa 2.

	Buy & Hold	OST 1	OST 2	MACD 1	MACD 2	RSI 1	RSI 2
2008	-42%	42%	7%	38%	15%	26%	41%
2009	53%	-53%	-50%	-51%	-10%	-46%	-46%
2010	-23%	23%	4%	-12%	12%	23%	23%
2011	-21%	21%	11%	41%	6%	21%	24%
2012	-9%	9%	-14%	-13%	-19%	9%	-1%
2013	-15%	15%	-15%	28%	35%	15%	11%
2014	-39%	39%	-2%	83%	33%	39%	-10%
2015	31%	-31%	-34%	9%	-32%	-27%	-9%

Os resultados para PETR4 variam de acordo com o período. Usando MACD 2, de modo geral, os resultados para os períodos foram positivos e mais estáveis, ou seja, obti-

veram ganhos com uma menor variação. No caso de MACD 1 os rendimentos alcançaram extremos, tendo este obtido o maior lucro (83%) e um dos maiores prejuízos (51%). Baseado nestes dados pode-se indicar que MACD 2 é uma boa opção para uma abordagem conservadora de investimento e MACD 1 é apropriada para investidores mais agressivos e que estão dispostos a correrem maiores riscos. O rendimento acumulado para as estratégias MACD 1, OST 1, RSI 1, MACD 2, RSI 2, *Buy & Hold* e OST 2, é igual a, respectivamente, 122%, 63%, 60%, 41%, 33%, -65% e -93%. Quase todas as estratégias obtiveram lucro acumulado, exceto *Buy & Hold* e OST 2.

A Figura 39 mostra os negócios realizados para PETR4 no ano de 2014 usando o SARSOP com a análise técnica MACD 1 e a modelagem POMDP. Este experimento obteve rendimento de 83%, sendo este o melhor dentre todos os resultados para PETR4. Neste gráfico são colocados os preços diários da PETR4 e os momentos de alternância de posição do investidor (*COMPRADO*, *VENDIDO* e *FORA*). Além disso, na parte inferior é mostrada a diferença do preço entre cada momento que o investidor executou uma operação.



Figura 39 – Negócios realizados para PETR4, usando MACD 1, durante o ano de 2014, com o planejador SARSOP e modelagem POMDP.

É apresentada na Seção 5.1.1 o uso de múltiplos ativos com uma modelagem POMDP. Para este problema é utilizado o SARSOP. Nos experimentos esta abordagem é chamada apenas de múltiplos ativos.

A Tabela 28 apresenta os resultados para CMIG4 usando múltiplos ativos. É utilizado o ativo mais correlacionado ao CMIG4. São mostrados os rendimentos percentuais para as taxas 1 e 2. Novamente, a taxa 1 teve um grande impacto nos rendimentos. Muitos resultados de lucro com as taxas 2 e 3 se tornam em prejuízo com a taxa 1. Um ponto importante é que os rendimentos foram melhores com o planejamento usando múltiplos ativos. Além dos resultados anuais, os valores acumulados de modo geral foram melhores.

Enquanto o maior rendimento com um ativo foi de 40,32%, com dois ativos foi obtido 91,76%. A MACD 2 apresentou um desempenho muito bom, exceto no ano de 2014 onde ocorreu um prejuízo de 50,95%.

Tabela 28 – Resultados das simulações com múltiplos ativos para CMIG4.

Resultado Percentual Anual										
	1 ativo (POMDP)						Múltiplos ativos (POMDP)			
	Buy & Hold		OST 2		MACD 2		OST 2		MACD 2	
	Nº da Taxa									
	1º	2º	1º	2º	1º	2º	1º	2º	1º	2º
2008	-3,2%	-2,1%	-40,4%	-21,7%	-17,8%	-1,6%	-16,5%	27,4%	-31,8%	-25,1%
2009	19,8%	20,9%	-10,1%	16,3%	-22,7%	-10,6%	-4,3%	27,0%	-27,3%	-3,2%
2010	-6,3%	-5,6%	-12,4%	11,3%	-28,7%	-17,1%	-7,3%	-4,5%	-19,3%	-5,9%
2011	18,2%	19,4%	-10,0%	8,8%	-20,7%	-10,7%	-15,2%	3,8%	-37,4%	-18,6%
2012	-15,0%	-14,0%	-10,4%	10,1%	-58,3%	-50,3%	-10,7%	16,3%	-57,7%	-41,5%
2013	-7,8%	-6,8%	-27,0%	-13,5%	0,6%	17,1%	-12,3%	21,5%	22,0%	40,6%
2014	-3,9%	-2,8%	1,3%	26,7%	20,6%	40,3%	-56,4%	-51,0%	35,9%	91,8%
2015	4,8%	5,9%	-6,1%	3,6%	-9,0%	-4,2%	-11,1%	0,1%	0,4%	10,0%

Resultado Percentual Acumulado										
	1 ativo (POMDP)						Múltiplos Ativos (POMDP)			
	Buy & Hold		OST 2		MACD 2		OST 2		MACD 2	
	Nº da Taxa									
	1º	2º	1º	2º	1º	2º	1º	2º	1º	2º
2008	-3,2%	-2,1%	-40,4%	-21,7%	-17,8%	-1,6%	-16,5%	27,4%	-31,8%	-25,1%
2009	16,6%	18,8%	-50,5%	-5,5%	-40,5%	-12,2%	-20,8%	54,5%	-59,1%	-28,3%
2010	10,3%	13,2%	-62,9%	5,9%	-69,2%	-29,3%	-28,1%	50,0%	-78,4%	-34,2%
2011	28,5%	32,6%	-72,9%	14,9%	-89,9%	-40,0%	-43,3%	53,8%	-115,8%	-52,9%
2012	13,5%	18,6%	-83,3%	24,8%	-148,3%	-90,3%	-53,9%	70,1%	-173,5%	-94,4%
2013	5,7%	11,8%	-110,2%	11,4%	-147,6%	-73,2%	-66,3%	91,7%	-151,5%	-53,8%
2014	1,8%	9,0%	-108,9%	38,0%	-127,0%	-32,9%	-122,6%	40,7%	-115,5%	38,0%
2015	6,6%	14,9%	-115,0%	41,7%	-136,0%	-37,0%	-133,7%	40,8%	-115,2%	48,0%

Os resultados para CRUZ3 e o ativo mais correlacionado são exibidos na Tabela 29. O uso de múltiplos ativos demonstrou resultados melhores que ao utilizar apenas um ativo. Uma característica interessante é que a análise técnica MACD 2 deixou de realizar negócios para alguns períodos. Quando existem muitas incertezas do comportamento dos preços é mais indicado que operações não sejam realizadas. Deste modo, certos riscos de investimento são evitados e abre-se a possibilidade de explorar outros ativos. Apesar da estratégia *Buy & Hold* ter sido a melhor opção para CRUZ3, caso sejam considerados os experimentos a partir de 2011, múltiplos ativos obtém valores próximos ou até superiores.

Na Figura 40 é mostrado um resumo dos resultados de cada ativo para o período que perfaz os anos de 2008 a 2015. A estratégia de **múltiplos ativos** obteve um desempenho melhor que **1 ativo** em 4 casos (CMIG4, CRUZ3, ITSA4 e VALE5). Apesar da **múltiplos ativos** ter sido pior que a **1 ativo** para PETR4, os resultados ficaram próximos. Ao comparar a **múltiplos ativos** e a **buy & hold**, a **múltiplos ativos** atingiu melhores rendimentos para 3 ativos (CMIG4, PETR4 e VALE5).

Segundo os dados exibidos na Figura 40, outra forma de realizar a análise dos resultados é considerando o período que perfaz os anos de 2011 a 2015. A estratégia **múltiplos ativos** é melhor que a **buy & hold** e melhor que a *1 ativo* em 4 casos (CMIG4, ITSA4,

Tabela 29 – Resultados das simulações para CRUZ3, múltiplos ativos e taxa 2.

Resultado Percentual Anual

	BUY & HOLD	1 ativo			Múltiplos ativos		
		MACD 2	OST 2	RSI 1	MACD 2	OST 2	RSI 1
2008	-8,48%	24,37%	23,04%	8,32%	0,00%	-13,38%	8,28%
2009	31,06%	-38,53%	-41,21%	-31,22%	0,00%	0,00%	0,00%
2010	53,85%	-42,98%	-22,90%	-48,38%	0,00%	0,00%	0,00%
2011	23,75%	-4,99%	-9,23%	-23,91%	0,00%	-28,41%	47,34%
2012	29,28%	-16,14%	3,35%	-24,57%	-7,07%	4,70%	-22,08%
2013	-21,02%	36,10%	36,04%	20,86%	30,23%	46,39%	4,90%
2014	-17,44%	71,01%	41,01%	17,28%	60,65%	43,91%	17,28%
2015	31,12%	-35,80%	-25,52%	-31,28%	-34,98%	-25,52%	-31,28%

Resultado Percentual Acumulado (2008-2015)

	BUY & HOLD	1 ativo			Múltiplos ativos		
		MACD 2	OST 2	RSI 1	MACD 2	OST 2	RSI 1
2008	-8,48%	24,37%	23,04%	8,32%	0,00%	-13,38%	8,32%
2009	22,58%	-14,16%	-18,17%	-22,90%	0,00%	-13,38%	8,32%
2010	76,43%	-57,14%	-41,07%	-71,28%	0,00%	-13,38%	8,32%
2011	100,18%	-62,13%	-50,30%	-95,19%	0,00%	-41,79%	55,66%
2012	129,46%	-78,27%	-46,95%	-119,76%	-7,07%	-37,09%	33,58%
2013	108,44%	-42,17%	-10,91%	-98,90%	23,16%	9,30%	38,48%
2014	91,00%	28,84%	30,10%	-81,62%	83,81%	53,21%	55,76%
2015	122,12%	-6,96%	4,58%	-112,90%	48,83%	27,69%	24,48%

Resultado Percentual Acumulado (2011-2015)

	BUY & HOLD	1 ativo			Múltiplos ativos		
		MACD 2	OST 2	RSI 1	MACD 2	OST 2	RSI 1
2011	23,75%	-4,99%	-9,23%	-23,91%	0,00%	-28,41%	47,34%
2012	53,03%	-21,13%	-5,88%	-48,48%	-7,07%	-23,71%	25,26%
2013	32,01%	14,97%	30,16%	-27,62%	23,16%	22,68%	30,16%
2014	14,57%	85,98%	71,17%	-10,34%	83,81%	66,59%	47,44%
2015	45,69%	50,18%	45,65%	-41,62%	48,83%	41,07%	16,16%

PETR4 e VALE5). Apesar da *1 ativo* de ter se saído melhor para CRUZ3, o resultado ficou muito próximo a estratégia **múltiplos ativos**.

No contexto geral, o uso de múltiplos ativos apresentou melhores resultados que as demais estratégias. Como o modelo POMDP limita o planejamento, devido ao crescimento exponencial do espaço de estados, é dada a abordagem MOMDP, conforme a Seção 5.2. É muito importante lembrar que na abordagem MOMDP as variáveis que compõem os estados são decompostas. Além disso, para trabalhar com MOMDP é utilizada representação em POMDPX e o planejador SARSOP.

Para a MOMDP a taxa 1 não teve um grande impacto. Isso deve-se ao fato do número de negociações ocorrerem poucas vezes. Esta pode ser considerada uma estratégia mais conservadora, uma vez que as operações são realizadas com menor frequência. A Tabela 30 mostra os resultados do MOMDP para CMIG4 usando as taxas de número 1 e 2.

Os resultados com MOMDP são melhores que aqueles obtidos com o POMDP para a taxa 1. Ao final, a RSI 1 teve o menor impacto das taxas, tendo sido até melhor que a *Buy & Hold*.

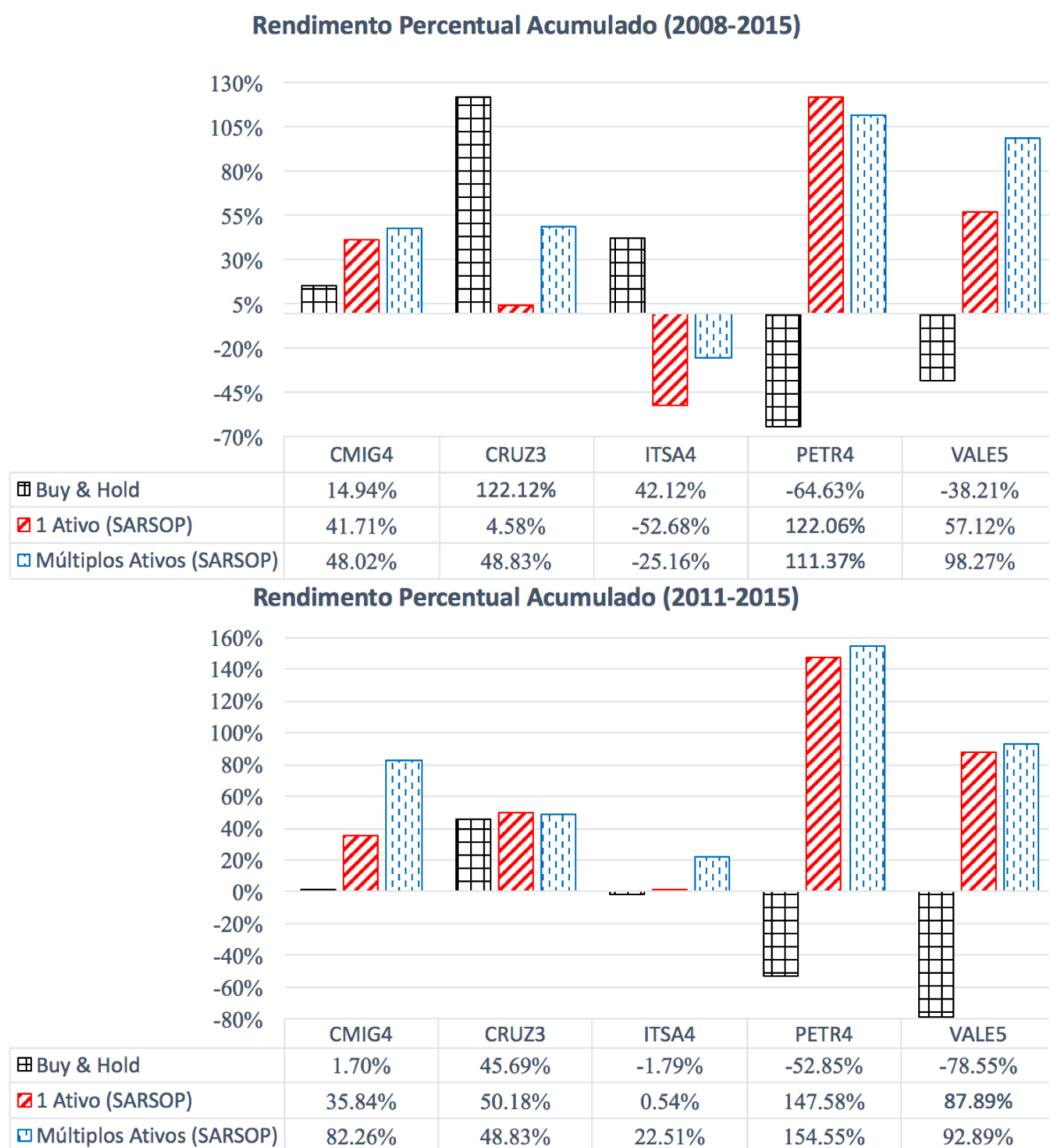


Figura 40 – Melhores rendimentos acumulados para cada ativo nos períodos de 2008 a 2015 e 2011 a 2015 com a cobrança da taxa 2.

A Tabela 31 mostra os resultados para CRUZ3 com a taxa 2 usando POMDP com 1 ativo, múltiplos ativos e MOMDP. Foram encontrados excelentes resultados com MOMDP em relação as demais técnicas. Com o MOMDP o destaque foi a OST 1 com um lucro acumulado de 126,7%. Além do melhor rendimento, também são mostrados os piores casos para cada análise técnica.

Como as análises individuais são muito similares as descrições anteriores, um resumo do valores máximos obtidos considerando todos os anos, cada técnica e ativo, é realizada a partir de gráficos. A Figura 41 mostra os resultados obtidos com cada estratégia ao

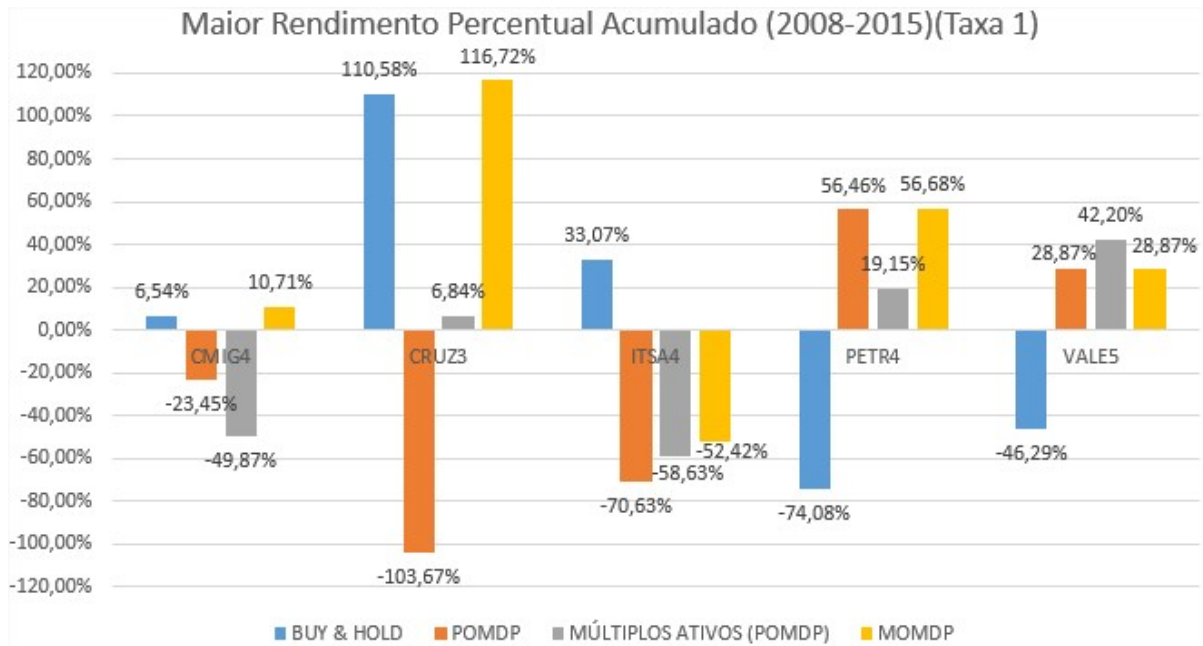


Figura 41 – Comparação dos melhores rendimentos acumulados para cada ativo no período de 2008 a 2015 usando a taxa 1.

Tabela 30 – Resultados das simulações usando MOMDP para CMIG4.

Resultado Percentual Anual								
	Buy & Hold		OST 1		OST 2		RSI 1	
			Nº da Taxa					
	1º	2º	1º	2º	1º	2º	1º	2º
2008	-3,26%	-2,12%	-3,4%	-2,33%	0,89%	1,96%	0,89%	1,96%
2009	19,75%	20,93%	11,62%	13,85%	0%	0%	-22,27%	-21,09%
2010	-6,26%	-5,57%	-6,51%	4,7%	4,04%	5,41%	4,04%	5,41%
2011	18,18%	19,35%	15,92%	18,16%	-20,67%	-19,51%	15,92%	18,16%
2012	-14,98%	-14%	-15,09%	-14,09%	12,86%	13,84%	12,86%	13,84%
2013	-7,83%	-6,75%	-10,57%	-8,43%	5,52%	6,59%	-10,57%	-8,43%
2014	-3,86%	-2,8%	-4,6%	-2,47%	1,58%	2,64%	-4,6%	-2,47%
2015	4,8%	5,9%	14,44%	16,67%	-7,16%	-6,06%	14,44%	16,67%

Resultado Percentual Acumulado								
	Buy & Hold		OST 1		OST 2		RSI 1	
			Nº da Taxa					
	1º	2º	1º	2º	1º	2º	1º	2º
2008	-3,26%	-2,12%	-3,4%	-2,33%	0,89%	1,96%	0,89%	1,96%
2009	16,49%	18,81%	8,22%	11,52%	0,89%	1,96%	-21,38%	-19,13%
2010	10,23%	13,24%	1,71%	16,22%	4,93%	7,37%	-17,34%	-13,72%
2011	28,41%	32,59%	17,63%	34,38%	-15,74%	-12,14%	-1,42%	4,44%
2012	13,43%	18,59%	2,54%	20,29%	-2,88%	1,7%	11,44%	18,28%
2013	5,6%	11,84%	-8,03%	11,86%	2,64%	8,29%	0,87%	9,85%
2014	1,74%	9,04%	-12,63%	9,39%	4,22%	10,93%	-3,73%	7,38%
2015	6,54%	14,94%	1,81%	26,06%	-2,94%	4,87%	10,71%	24,05%

utilizar a taxa 1. O MOMDP foi melhor em 4 casos em relação a *Buy & Hold* e a múltiplos ativos (POMDP). Já comparando com a 1 ativo a MOMDP se saiu melhor em todos os casos.

Apesar dos bons resultados apresentados até o momento, ao considerar a taxa 2 o MOMDP acabou não se saindo tão bem no contexto geral em relação as demais técnicas. A Figura 42 mostra os valores percentuais máximos alcançados por cada estratégia ao

Tabela 31 – Resultados usando MOMDP para CRUZ3.

Resultado Percentual Anual

	BUY & HOLD	1 ativo		Múltiplos ativos		MOMDP	
		MACD 1	OST 2	MACD 1	MACD 2	MACD 1	OST 1
2008	-8,48%	22,24%	23,04%	-15,81%	0,00%	4,33%	0%
2009	31,06%	-43,76%	-41,21%	0,00%	0,00%	-26,67%	25,86%
2010	53,85%	-44,61%	-22,90%	0,00%	0,00%	-47,6%	46,98%
2011	23,75%	-23,36%	-9,23%	0,00%	0,00%	-30,13%	22,9%
2012	29,28%	-24,33%	3,35%	-27,82%	-7,07%	-29,88%	27,12%
2013	-21,02%	1,96%	36,04%	6,01%	30,23%	23,28%	-21,86%
2014	-17,44%	-24,87%	41,01%	-20,20%	60,65%	17,28%	-12,62%
2015	31,12%	-15,93%	-25,52%	-19,79%	-34,98%	-31,28%	38,32%

Resultado Percentual Acumulado (2008-2015)

	BUY & HOLD	1 ativo		Múltiplos ativos		MOMDP	
		MACD 1	OST 2	MACD 1	MACD 2	MACD 1	OST 1
2008	-8,48%	22,24%	23,04%	-15,81%	0,00%	4,33%	0%
2009	22,58%	-21,52%	-18,17%	-15,81%	0,00%	-22,34%	25,86%
2010	76,43%	-66,13%	-41,07%	-15,81%	0,00%	-69,94%	72,84%
2011	100,18%	-89,49%	-50,30%	-15,81%	0,00%	-100,07%	95,74%
2012	129,46%	-113,82%	-46,95%	-43,63%	-7,07%	-129,95%	122,86%
2013	108,44%	-111,86%	-10,91%	-37,62%	23,16%	-106,67%	101%
2014	91,00%	-136,73%	30,10%	-57,82%	83,81%	-89,39%	88,38%
2015	122,12%	-152,66%	4,58%	-77,61%	48,83%	-120,67%	126,7%

usar a taxa 2. Fica evidente que os valores máximos são piores com MOMDP.

Os melhores rendimentos acumulados usando a taxa 2 são colocados na Figura 42. Considerando os valores acumulados, o MOMDP foi superior a *Buy & Hold* para todos os ativos, exceto com o ITSA4. Contudo, o MOMDP não conseguiu superar as abordagens que utilizam POMDP.

A competição IPPC utiliza planejadores *online*, logo, o portfólio de planejamento trabalha com este gênero de planejadores. Sendo assim, além do SARSOP, foram realizados experimentos para o mercado financeiros com o Despot e o POMCP, de modo a avaliar parte dos planejadores do portfólio de planejamento.

Foram realizados diversos experimentos com o Despot e o POMCP. Alguns resultados foram os mesmos obtidos anteriormente, tais como, o impacto da taxa 1 e o comportamento das análises técnicas. Os resultados são apresentados na Figura 42, sendo este um resumo geral considerando todos os planejadores e modelagens de planejamento.

Os planejadores Despot e o POMCP obtiveram ótimos resultados em relação as demais propostas. Estes planejadores alcançaram os melhores resultados. Além disso, os valores acumulados foram muito bons, em especial para o ativo ITSA4 com um lucro acumulado de 134,4%. Estes resultados são devidos a frequência de negociações, uma vez que foram realizadas diversas operações ao usar o Despot e o POMCP.



Figura 42 – Resumo geral dos rendimentos obtidos utilizando a taxa 2 e todos os planejadores e modelagens de planejamento.

Conclusão

Neste trabalho foram propostas diferentes formas de modelagem para o uso de planejamento probabilístico no mercado financeiro. Além das modelagens é apresentado o portfólio de planejamento, sendo ele responsável por escolher o(s) planejador(s) mais apto para cada problema.

No trabalho de Baffa (BAFFA; CIARLINI, 2010b) os experimentos foram executados apenas com o índice IBOVESPA e não foram cobradas taxas de corretagem. Neste trabalho foi colocado o uso de um número maior de ações, aplicação de taxas de corretagem e planejadores.

A primeira modelagem segue o trabalho de Baffa (BAFFA; CIARLINI, 2010b) usando POMDP. Este modelo conta com 28 estados, 5 ações e 3 observações. Os estados são formados pelas tendências do passado, tendências futuras e as posições do investidor. As ações correspondem as operações de negociação que podem ser executadas no mercado. Já as observações são geradas usando análise técnica e indicam pontos de compra, pontos de venda e pontos sem negociação. Para cada estado é dado um valor de recompensa. As probabilidades de transições dos estados e observações são realizadas a partir de um conjunto de dados históricos correspondente ao período de treinamento.

Para tratar de múltiplos ativos com POMDP os elementos da modelagem são organizados de modo a conter cada parte de cada ativo. Entretanto, esta proposta é exponencial para os estados, ações e observações. Quando três ou mais ativos são usados o espaço de estados começa a se tornar intratável. Apesar deste problema, de forma geral os experimentos com dois ativos apresentaram resultados melhores que as demais estratégias. A escolha de quais ações utilizar foi realizada utilizando cálculos de correlação.

Uma característica observada na modelagem POMDP é que os estados são compostos por diferentes tipos de variáveis. Com base nestas variáveis foi proposta uma nova modelagem baseada em MOMDP. Esta modelagem obteve bons resultados em relação a estratégia *Buy & Hold*. Uma característica interessante é que com a abordagem MOMDP o impacto das taxas foram menores. A MOMDP alcançou melhores rendimentos em relação ao POMDP para a taxa 1 e com as taxas 2 e 3 os rendimentos foram melhores com

POMDP.

Apesar de toda a proposta do MOMDP, parte do planejamento depende de alterações do planejador. Isso deve-se ao fato da possibilidade de utilizar variáveis determinísticas para determinar quais ações podem ser aplicadas no estado atual. Sendo assim, um trabalho futuro consiste em realizar o desenvolvimento de um planejador que trate de todos os detalhes da modelagem proposta em MOMDP.

Não existe um planejador que encontra a melhor solução para todos os tipos de problemas. Devida a esta característica se investigou o desenvolvimento de um portfólio de planejamento não clássico. Para montar o portfólio foram utilizados algoritmos de classificação e as características são extraídas dos problemas representados com a linguagem RDDDL. Os domínios e as instâncias foram retirados da competição IPPC dos anos de 2011 e 2014. Dois desafios foram encontrados na pesquisa em relação a IPPC. Primeiro, nem todos os planejadores estão disponíveis. Segundo, os planejadores trabalham com parser de RDDDL para outras linguagens, isto dificulta a compreensão e o parser não necessariamente suporta todo do domínio da RDDDL. Um trabalho futuro é o de realizar o desenvolvimento de planejadores que trabalhem nativamente com RDDDL e suportem completamente a linguagem.

Analisando os resultados dos experimentos foram observados certos comportamentos. Em geral a taxa 1 teve um grande impacto nos rendimentos e apenas a modelagem MOMDP com o planejador SARSOP conseguiu lidar com tal taxação. É indicado em um trabalho futuro a adição do custo de negociação na modelagem, com isso, o planejamento é realizado já considerando os custos e o capital de investimento. Contudo, para isso é indicado o uso planejamento *online* e que as operação sejam *intraday* (obtenção de lucro com a oscilação de preço ao longo do dia). Outro comportamento observado é o de que para os capitais de investimento superiores a R\$ 25.000,00 os percentuais de rendimento foram muito próximos, entretanto, para o valor de R\$ 10.000,00 as taxas impactaram consideravelmente os resultados.

Foram realizados experimentos utilizando planejadores *offline* e *online*. Os planejadores *online* foram escolhidos em função das características do problema do mercado financeiro e devido aos planejadores usados na IPPC. Excelentes resultados foram obtidos com os planejadores Despot e POMCP para a taxa 2.

A pesquisa resultou na publicação dos trabalhos (ALVES; LOPES; BRANQUINHO, 2013), (BRANQUINHO et al., 2015) e (BRANQUINHO; LOPES; BAFFA, 2016). Outras publicações foram submetidas a conferência/revistas e aguardam o processo de revisão.

9.1 Trabalhos Futuros

Existem muitos pontos para serem tratados em trabalhos futuros. É preciso que estudos sejam realizados sobre as características estatísticas dos problemas de planejamento,

levantando as propriedades e o nível de dificuldade dos problemas. Outro ponto é o de que o portfólio de planejamento deve ser melhorado com a combinação dos planejadores.

Em um trabalho futuro é preciso pesquisar o uso de *ensemble methods* e técnicas de aprendizado para combinar mais de uma análise técnica para gerar as observações. A modelagem demanda de vários parâmetros para sua configuração, logo, é necessário pesquisar diferentes combinações de tais parâmetros.

Outro trabalho futuro é o de utilizar um número maior de tipos de tendências, tais como tendência de alta acentuada e tendência de baixa acentuada. Por último, é necessário analisar o cenário onde o mercado está em alta. Durante esta fase obter lucro torna-se um desafio maior, já que em muitas vezes o ideal pode consistir apenas em executar a estratégia *Buy & Hold* ou operar *intraday*.

Referências

- ABRAHAM, A.; PHILIP, N. S.; SARATCHANDRAN, P. Modeling chaotic behavior of stock indices using intelligent paradigms. **Neural, Parallel & Scientific Computations**, Dynamic Publishers, Inc., Atlanta, GA, USA, v. 11, n. 1 & 2, p. 143–160, mar. 2003. ISSN 1061-5369.
- ACHELIS, S. B. **Technical Analysis From A To Z**. India: McGraw-hill Education, 2006. ISBN 9780070636576.
- ALVES, R. M. F.; LOPES, C. R.; BRANQUINHO, A. A. B. Generating plans using lrt. In: **Brazilian Conference on Intelligent Systems (BRACIS)**. Brazil: IEEE, 2013. p. 207–212.
- ALVIM, L. G.; MILIDIÚ, R. L. Trading team composition for the intraday multistock market. **Decision Support Systems**, v. 54, n. 2, p. 838–845, 2013. ISSN 0167-9236.
- ALVIM, L. G. M. **Estratégias de Negociação de Ativos Financeiros utilizando Agendamento por Intervalos Ponderados**. Tese (Doutorado), 2013.
- APPEL, G. **The moving average convergence-divergence trading method: advanced version**. Pennsylvania, USA: Scientific Investment Systems, 1985.
- _____. **Technical Analysis: Power Tools for Active Investors**. First. Upper Saddle River, NJ, USA: Financial Times/Prentice Hall, 2005. (Financial Times Prentice Hall books). ISBN 9780131479029.
- ATSALAKIS, G. S.; VALAVANIS, K. P. Surveying stock market forecasting techniques – part ii: Soft computing methods. **Expert Systems with Applications**, v. 36, n. 3, Part 2, p. 5932 – 5941, 2009. ISSN 0957-4174.
- _____. Surveying stock market forecasting techniques - part i: Conventional methods. In: **Computation Optimization in Economics and Finance Research Compendium**. New York, NY, USA: Nova Science Publishers, 2013. ISBN 978-1-62081-928-9.
- BAFFA, A. C. E.; CIARLINI, A. E. M. **Ambiente para Modelagem, Planejamento e Avaliação de Políticas no Mercado de Ações**. 104 p. Dissertação (Mestrado) — Universidade Federal do Estado do Rio de Janeiro, 2010.
- BAFFA, A. C. E.; CIARLINI, A. E. M. Modeling pomdps for generating and simulating stock investment policies. In: **Proceedings of the 2010 ACM Symposium on**

- Applied Computing - SAC '10**. New York, NY, USA: ACM Press, 2010. p. 2394. ISBN 9781605586397.
- BAI, H. et al. Monte carlo value iteration for continuous-state pomdps. In: **Algorithmic foundations of robotics IX**. [S.l.]: Springer, 2010. p. 175—191.
- BARROSO, B. C. et al. Uma ferramenta computacional eficiente para identificação de melhores gatilhos de compra e venda de ativos financeiros. **BRACIS**, 2012.
- BELLMAN, R. **Dynamic Programming**. 1. ed. Princeton, NJ, USA: Princeton University Press, 1957.
- BLUM, A. L.; FURST, M. L. Fast planning through planning graph analysis. **Artificial Intelligence**, Elsevier Science Publishers Ltd., Essex, UK, UK, v. 90, n. 1-2, p. 281–300, 1997. ISSN 0004-3702.
- BOLLINGER, J. **Bollinger on Bollinger Bands**. New York, NY, USA: McGraw-Hill, 2001. ISBN 9780071373685.
- BRANQUINHO, A. et al. Aquisição automática de competências num ambiente educacional ubíquo. Maceió, Brazil, 2015.
- BRANQUINHO, A. A. B. **Planejamento de recursos para jogos de estratégia em tempo real**. 133 p. Dissertação (Mestrado) — Federal University of Uberlândia, 2009.
- BRANQUINHO, A. A. B.; LOPES, C. R.; BAFFA, A. C. E. Probabilistic planning for multiple stocks of financial markets. In: **IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI)**. San Jose, CA, USA: IEEE, 2016. p. 501–508.
- BROCK, W.; LAKONISHOK, J.; LEBARON, B. Simple technical trading rules and the stochastic properties of stock returns. **Journal of Finance**, v. 47, n. 5, p. 1731–64, December 1992.
- BROWN, C. **Technical Analysis for the Trading Professional**. [S.l.]: Mcgraw-hill, 1999. (Irwin trader's edge series). ISBN 9780070120624.
- CAO, L.; TAY, F. Support vector machine with adaptive parameters in financial time series forecasting. **Neural Networks, IEEE Transactions on**, v. 14, n. 6, p. 1506–1518, Nov 2003. ISSN 1045-9227.
- CASSANDRA, A. R.; KAEHLING, L. P.; LITTMAN, M. L. Acting optimally in partially observable stochastic domains. **AAAI**, Brown University Providence, RI, USA, n. April, 1994.
- COLES, A. J. et al. A survey of the seventh international planning competition. **AI Magazine**, v. 33, n. 1, 2012.
- COWLES, A. Can stock market forecasters forecast? **Econometrica**, v. 1, p. 309–324, 1933.
- COWPERTWAIT, P. S. P.; METCALFE, A. V. **Introductory Time Series with R**. 1st. ed. New York, NY, USA: Springer Publishing Company, Incorporated, 2009. ISBN 0387886974, 9780387886978.

- ENGEL, J. Teaching feed-forward neural networks by simulated annealing. **Complex Systems**, Complex Systems Publications, Inc., Champaign, IL, USA, v. 2, n. 6, p. 641–648, dez. 1988. ISSN 0891-2513.
- FAMA, E. F. The behavior of stock-market prices. **The Journal of Business**, The University of Chicago Press, v. 38, n. 1, p. 34–105, 1965. ISSN 00219398.
- _____. Efficient capital markets: A review of theory and empirical work. **The Journal of Finance**, Blackwell Publishing Ltd, v. 25, n. 2, p. 383–417, 1970. ISSN 1540-6261.
- FIKES, R. E.; NILSSON, N. J. Strips: A new approach to the application of theorem proving to problem solving. **Artificial Intelligence**, v. 2, p. 189, 1971.
- GARDNER, M. The fantastic combinations of john conway’s new solitaire game “life”. **Scientific American**, v. 223, p. 120—123, out. 1970.
- HOEY, J. et al. Spudd: Stochastic planning using decision diagrams. In: **In Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence**. San Francisco, CA, USA: Morgan Kaufmann, 1999. p. 279–288.
- HOFFMANN, J.; NEBEL, B. The FF planning system: Fast plan generation through heuristic search. v. 14, p. 253–302, 2001.
- HOLGUIN, M. G. **Planejamento Probabilístico usando Programação Dinâmica Assíncrona e Fatorada**. 130 p. Dissertação (Mestrado) — Instituto de Matemática e Estatística da Universidade de São Paulo, 2013.
- HOWARD, R. A. **Dynamic Programming and Markov Processes**. Cambridge, MA, USA: MIT Press, 1960.
- HUANG, W.; NAKAMORI, Y.; WANG, S.-Y. Forecasting stock market movement direction with support vector machine. **Comput. Oper. Res.**, Elsevier Science Ltd., Oxford, UK, UK, v. 32, n. 10, p. 2513–2522, out. 2005. ISSN 0305-0548.
- JENSEN, M. C.; BENINGTON, G. A. Random walks and technical theories: Some additional evidence. **The Journal of Finance**, Blackwell Publishing Ltd, v. 25, n. 2, p. 469–482, 1970. ISSN 1540-6261.
- KABOUDAN, M. A. Genetic programming prediction of stock prices. **Computational Economics**, Springer, v. 16, n. 3, p. 207–236, 2000.
- KAUFMAN, P. J. **Smarter Trading: Improving Performance in Changing Markets**. New York, NY, USA: McGraw-Hill, 1995. ISBN 9780070340022.
- KIM, H.; KIM, I.; LEE, Y. Scale-free network in stock markets. v. 40, n. 6, p. 1105–1108, 2002.
- KIM, H. J. et al. **Scale-free network in financial correlations**. 2001.
- KIM, J.-H.; MOON, B.-R. Neuron reordering for better neuro-genetic hybrids. In: **Proceedings of the 4th Annual Conference on Genetic and Evolutionary Computation**. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2002. (GECCO’02), p. 407–414. ISBN 1-55860-878-8.

KURNIAWATI, H.; HSU, D.; LEE, W. S. Sarsop: Efficient point-based pomdp planning by approximating optimally reachable belief spaces. 2008.

KWON, Y.-K.; CHOI, S.-S.; MOON, B.-R. Stock prediction based on financial correlation. **Proceedings of the 2005 conference on Genetic and evolutionary computation - GECCO '05**, p. 2061, 2005.

KWON, Y.-K.; MOON, B.-R. Daily stock prediction using neuro-genetic hybrids. **Genetic and Evolutionary Computation-GECCO**, p. 2203–2214, 2003.

_____. Evolutionary ensemble for stock prediction. **Genetic and Evolutionary Computation-GECCO**, p. 1102–1113, 2004.

_____. A hybrid neurogenetic approach for stock forecasting. **IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council**, v. 18, n. 3, p. 851–64, maio 2007. ISSN 1045-9227.

LO, A. W.; MAMAYSKY, H.; WANG, J. **Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation**. [S.l.], 2000. (Working Paper Series, 7613).

LUKE, S. **Essentials of Metaheuristics**. [S.l.]: Lulu, 2013. Disponível gratuitamente em <http://cs.gmu.edu/~sean/book/metaheuristics>.

MACLIN, R.; OPITZ, D. Popular ensemble methods: An empirical study. **Journal of Artificial Intelligence Research**, v. 11, p. 169–198, 1999.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In: . Berkeley, CA, USA: University of California Press, 1967. v. 233, n. 233, p. 281–297.

MCDERMOTT, D. et al. **PDDL - The Planning Domain Definition Language**. 1998.

MURPHY, J. J.; MURPHY, J. J. **Technical Analysis of the Financial Markets**. Fishkill, NY, USA: New York Institute of Finance, 1999. ISBN 0735200661 9780735200661.

NISON, S. **Japanese Candlestick Charting Techniques**. [S.l.: s.n.], 1991. 315 p. ISBN 0139316507.

ONG, S. C. W. et al. Planning under uncertainty for robotic tasks with mixed observability. **International Journal of Robotics Research**, Sage Publications, Inc., Thousand Oaks, CA, USA, v. 29, n. 8, p. 1053–1068, jul. 2010. ISSN 0278-3649.

PEDNAULT, E. P. D. Adl: exploring the middle ground between strips and the situation calculus. In: **Proceedings of the first international conference on Principles of knowledge representation and reasoning**. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1989. p. 324–332. ISBN 1-55860-032-9.

PINEAU, J.; GORDON, G.; THRUN, S. Point-based value iteration: An anytime algorithm for pomdps. In: **International Joint Conference on Artificial Intelligence (IJCAI)**. [S.l.: s.n.], 2003. p. 1025 – 1032.

- PINEAU, J.; GORDON, G. J.; THRUN, S. Anytime point-based approximations for large pomdps. **Journal of Artificial Intelligence Research**, v. 27, p. 335–380, 2006.
- PUTERMAN, M. L. **Markov Decision Processes: Discrete Stochastic Dynamic Programming**. 1st. ed. New York, NY, USA: John Wiley & Sons, Inc., 1994. ISBN 0471619779.
- _____. **Markov Decision Processes: Discrete Stochastic Dynamic Programming**. [S.l.]: John Wiley & Sons, 2009. v. 414.
- RICHTER, S.; WESTPHAL, M. The lama planner: Guiding cost-based anytime planning with landmarks. **Journal of Artificial Intelligence Research**, v. 39, n. 1, p. 127–177, 2010.
- ROBERTS, M.; HOWE, A. Learned models of performance for many planners. In: **ICAPS 2007, Workshop AI Planning and Learning**. [S.l.: s.n.], 2007.
- _____. Learning from planner performance. **Artificial Intelligence**, v. 173, n. 5–6, p. 536–561, 2009. ISSN 0004-3702. Advances in Automated Plan Generation.
- ROSS, S. et al. Online planning algorithms for pomdps. **Journal of Artificial Intelligence Research**, v. 32, p. 663–704, 2008.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Neurocomputing: Foundations of research. In: ANDERSON, J. A.; ROSENFELD, E. (Ed.). Cambridge, MA, USA: MIT Press, 1988. cap. Learning Representations by Back-propagating Errors, p. 696–699. ISBN 0-262-01097-6.
- RUSSELL, S.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. 3rd. ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2009. ISBN 0136042597, 9780136042594.
- RUSSELL, S. J.; NORVIG. **Artificial Intelligence: A Modern Approach (Second Edition)**. [S.l.]: Prentice Hall, 2003.
- SANNER, S. Relational dynamic influence diagram language (rddl): Language description. 2010. Disponível em: <http://users.cecs.anu.edu.au/~ssanner/IPPC_2011/RDDL.pdf>.
- SCHANNEP, J. **Dow Theory for the 21st Century: Technical Indicators for Improving Your Investment Results**. Hoboken, NJ, USA: John Wiley & Sons, 2008. ISBN 9780470370759.
- SEIPP, J. et al. Learning portfolios of automatically tuned planners. In: **Proceedings of the Twenty-Second International Conference on International Conference on Automated Planning and Scheduling**. Atibaia, SP, Brazil: AAAI Press, 2013. (ICAPS'12), p. 368–372.
- SILVER, D.; VENESS, J. Monte-carlo planning in large pomdps. In: LAFFERTY, J. D. et al. (Ed.). **Advances in Neural Information Processing Systems 23**. Vancouver, BC, Canada: Curran Associates, Inc., 2010. p. 2164–2172.
- SONDIK, E. J. The optimal control of partially observable markov processes over a finite horizon. **Oper. Res.**, INFORMS, v. 21, n. 5, p. 1071–1088, out. 1973. ISSN 0030-364X.

- SOUSA, J. L. d. **Aplicando Técnicas de Aprendizado de Máquina em Planejamento**. 101 p. Dissertação (Mestrado) — Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Uberlândia, 2014.
- SOUSA, J. L. d.; LOPES, C. R. Aplicando técnicas de classificação em planejamento probabilístico. **Encontro Nacional de Inteligência Artificial e Computacional**, 2014.
- SUTSKEVER, I. et al. On the importance of initialization and momentum in deep learning. In: DASGUPTA, S.; MCALLESTER, D. (Ed.). **Proceedings of the 30th International Conference on Machine Learning (ICML-13)**. [S.l.]: JMLR Workshop and Conference Proceedings, 2013. v. 28, n. 3, p. 1139–1147.
- VALLATI, M.; CHRPA, L.; KITCHIN, D. E. Portfolio-based planning: State of the art, common practice and open challenges. **AI Communications**, IOS Press, v. 28, n. 4, p. 717–733, September 2015.
- WILDER, J. **New Concepts in Technical Trading Systems**. [S.l.]: Trend Research, 1978.
- ZHOU, Z. H. **Ensemble Methods: Foundations and Algorithms**. [S.l.: s.n.], 2012. ISBN 9781439830055.

Apêndices

Exemplo de Modelagem - Cassandra

Este apêndice apresenta um exemplo de modelagem POMDP para o mercado financeiro conforme a descrição da Seção 5.1. Foi usada a representação proposta por Cassandra¹ (CASSANDRA; KAEHLING; LITTMAN, 1994) para expressar o problema. Tal exemplo, mostrado a seguir, foi construído a partir dos dados do ativo PETR4 para o período de 2000 a 2007. Para produzir os dados das observações foi utilizada a análise técnica OST₁ conforme as regras mostradas na Tabela 7.

```
# discount: <fator de desconto>
discount: 0.95

# values: [ reward, cost ]
values: reward

# states: [ <lista de estados> ]
states: BAIXA_BAIXA_COMPRADO BAIXA_ESTAVEL_COMPRADO BAIXA_ALTA_COMPRADO ESTAVEL_BAIXA_COMPRADO
ESTAVEL_ESTAVEL_COMPRADO ESTAVEL_ALTA_COMPRADO ALTA_BAIXA_COMPRADO ALTA_ESTAVEL_COMPRADO ALTA_ALTA_COMPRADO
BAIXA_BAIXA_VENDIDO BAIXA_ESTAVEL_VENDIDO BAIXA_ALTA_VENDIDO ESTAVEL_BAIXA_VENDIDO ESTAVEL_ESTAVEL_VENDIDO
ESTAVEL_ALTA_VENDIDO ALTA_BAIXA_VENDIDO ALTA_ESTAVEL_VENDIDO ALTA_ALTA_VENDIDO BAIXA_BAIXA_FORA
BAIXA_ESTAVEL_FORA BAIXA_ALTA_FORA ESTAVEL_BAIXA_FORA ESTAVEL_ESTAVEL_FORA ESTAVEL_ALTA_FORA
ALTA_BAIXA_FORA ALTA_ESTAVEL_FORA ALTA_ALTA_FORA INDESEJAVEL

# actions: [ <lista de ações> ]
actions: COMPRA COMPRA_DUPLA VENDA VENDA_DUPLA NADA

# observations: [ <lista de observações> ]
observations: PONTO_COMPRA PONTO_VENDA PONTO_NADA

# start: [ <distribuição inicial das crenças> ]
start: 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.1111111112 0.1111111111 0.1111111111 0.1111111111
0.1111111111 0.1111111111 0.1111111111 0.1111111111 0.1111111111 0

# Probabilidade de transição entre os estados. A sequência de valores segue a ordem dos estados.
# Cada linha corresponde ao estado de origem e a coluna especifica o próximo estado.
# T: <ação>
# %f %f ... %f
# ...
# %f %f ... %f
```

¹ Detalhes da modelagem disponíveis em: <<http://www.pomdp.org/code/pomdp-file-spec.html>>


```
0.1274900399 0.0199203187 0.0039840637 0.3784860558 0.0836653386 0.0079681275 0.2589641434 0.1035856574
0.0159362550 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.0986547085 0.0896860987 0.0224215247 0.1479820628 0.3094170404 0.1255605381 0.0269058296 0.1076233184
0.0717488789 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.0223463687 0.0893854749 0.2067039106 0 0.1173184358 0.3743016760 0 0.0167597765 0.1731843575 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.0103092784 0 0 0.1408934708 0.0309278351 0 0.5945017182 0.2096219931 0.0137457045 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0
0.0077519380 0.0038759690 0 0.1007751938 0.1162790698 0.0348837209 0.1550387597 0.3837209302
0.1976744186 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.0045248869 0.0180995475 0.0316742081 0.0135746606 0.0904977376 0.1628959276 0.0045248869 0.1719457014
0.5022624434 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0.5358649789 0.0843881857 0.0042194093 0.1983122363 0.1012658228 0.0210970464
0.0253164557 0.0253164557 0.0042194093 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0.2280701755 0.3742690059 0.1228070175 0.0292397661 0.1520467836 0.0760233918
0.0058479532 0.0058479532 0.0058479532 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0.0466666667 0.2733333333 0.5200000000 0.0066666667 0.0200000000 0.1266666667 0 0
0.0066666667 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0.1274900399 0.0199203187 0.0039840637 0.3784860558 0.0836653386 0.0079681275
0.2589641434 0.1035856574 0.0159362550 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0.0986547085 0.0896860987 0.0224215247 0.1479820628 0.3094170404 0.1255605381
0.0269058296 0.1076233184 0.0717488789 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0.0223463687 0.0893854749 0.2067039106 0 0.1173184358 0.3743016760 0 0.0167597765
0.1731843575 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0.0103092784 0 0 0.1408934708 0.0309278351 0 0.5945017182 0.2096219931 0.0137457045 0
0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0.0077519380 0.0038759690 0 0.1007751938 0.1162790698 0.0348837209 0.1550387597
0.3837209302 0.1976744186 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0.0045248869 0.0180995475 0.0316742081 0.0135746606 0.0904977376 0.1628959276
0.0045248869 0.1719457014 0.5022624434 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.5358649789 0.0843881857 0.0042194093 0.1983122363 0.1012658228
0.0210970464 0.0253164557 0.0253164557 0.0042194093 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.2280701755 0.3742690059 0.1228070175 0.0292397661 0.1520467836
0.0760233918 0.0058479532 0.0058479532 0.0058479532 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0466666667 0.2733333333 0.5200000000 0.0066666667 0.0200000000
0.1266666667 0 0 0.0066666667 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.1274900399 0.0199203187 0.0039840637 0.3784860558 0.0836653386
0.0079681275 0.2589641434 0.1035856574 0.0159362550 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0986547085 0.0896860987 0.0224215247 0.1479820628 0.3094170404
0.1255605381 0.0269058296 0.1076233184 0.0717488789 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0223463687 0.0893854749 0.2067039106 0 0.1173184358 0.3743016760
0 0.0167597765 0.1731843575 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0103092784 0 0 0.1408934708 0.0309278351 0 0.5945017182
0.2096219931 0.0137457045 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0077519380 0.0038759690 0 0.1007751938 0.1162790698 0.0348837209
0.1550387597 0.3837209302 0.1976744186 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0045248869 0.0180995475 0.0316742081 0.0135746606 0.0904977376
0.1628959276 0.0045248869 0.1719457014 0.5022624434 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0077519380 0.0038759690 0 0.1007751938 0.1162790698 0.0348837209
0.1550387597 0.3837209302 0.1976744186 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0045248869 0.0180995475 0.0316742081 0.0135746606 0.0904977376
0.1628959276 0.0045248869 0.1719457014 0.5022624434 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0077519380 0.0038759690 0 0.1007751938 0.1162790698 0.0348837209
0.1550387597 0.3837209302 0.1976744186 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0045248869 0.0180995475 0.0316742081 0.0135746606 0.0904977376
0.1628959276 0.0045248869 0.1719457014 0.5022624434 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.0077519380 0.0038759690 0 0.1007751938 0.1162790698 0.0348837209
0.1550387597 0.3837209302 0.1976744186 0
0 0 
```

```
# Probabilidade de observar uma observação ao alcançar um estado.
# A sequência de valores segue a ordem dos estados e das observações.
# Cada linha e cada coluna correspondem, respectivamente, a um estado e uma observação.
# O: <ação>
# %f %f ... %f
# ...
# %f %f ... %f

O:*
0.0379746836 0.0210970464 0.9409282700
0.0350877193 0.0350877193 0.9298245614
```

```

0.0200000000 0.0466666667 0.9333333333
0.0677290837 0.0956175299 0.8366533865
0.0493273543 0.0852017937 0.8654708520
0.0558659218 0.0614525140 0.8826815642
0.0136986302 0.0342465753 0.9520547945
0.0232558140 0.0155038760 0.9612403101
0.0316742082 0.0316742081 0.9366515837
0.0379746836 0.0210970464 0.9409282700
0.0350877193 0.0350877193 0.9298245614
0.0200000000 0.0466666667 0.9333333333
0.0677290837 0.0956175299 0.8366533865
0.0493273543 0.0852017937 0.8654708520
0.0558659218 0.0614525140 0.8826815642
0.0136986302 0.0342465753 0.9520547945
0.0232558140 0.0155038760 0.9612403101
0.0316742082 0.0316742081 0.9366515837
0.0379746836 0.0210970464 0.9409282700
0.0350877193 0.0350877193 0.9298245614
0.0200000000 0.0466666667 0.9333333333
0.0677290837 0.0956175299 0.8366533865
0.0493273543 0.0852017937 0.8654708520
0.0558659218 0.0614525140 0.8826815642
0.0136986302 0.0342465753 0.9520547945
0.0232558140 0.0155038760 0.9612403101
0.0316742082 0.0316742081 0.9366515837
0.3333333334 0.3333333333 0.3333333333

```

```
# R: <ação> : <estado corrente> : <próximo estado> : <observação> %f
```

```

R: * : * : BAIXA_BAIXA_COMPRADO : * -10
R: * : * : BAIXA_ESTAVEL_COMPRADO : * -2
R: * : * : BAIXA_ALTA_COMPRADO : * 5
R: * : * : ESTAVEL_BAIXA_COMPRADO : * -5
R: * : * : ESTAVEL_ESTAVEL_COMPRADO : * 0
R: * : * : ESTAVEL_ALTA_COMPRADO : * 5
R: * : * : ALTA_BAIXA_COMPRADO : * 0
R: * : * : ALTA_ESTAVEL_COMPRADO : * 5
R: * : * : ALTA_ALTA_COMPRADO : * 10
R: * : * : BAIXA_BAIXA_VENDIDO : * 10
R: * : * : BAIXA_ESTAVEL_VENDIDO : * 5
R: * : * : BAIXA_ALTA_VENDIDO : * 0
R: * : * : ESTAVEL_BAIXA_VENDIDO : * 5
R: * : * : ESTAVEL_ESTAVEL_VENDIDO : * 0
R: * : * : ESTAVEL_ALTA_VENDIDO : * -5
R: * : * : ALTA_BAIXA_VENDIDO : * 5
R: * : * : ALTA_ESTAVEL_VENDIDO : * -2
R: * : * : ALTA_ALTA_VENDIDO : * -10
R: * : * : BAIXA_BAIXA_FORA : * -2
R: * : * : BAIXA_ESTAVEL_FORA : * 0
R: * : * : BAIXA_ALTA_FORA : * 0
R: * : * : ESTAVEL_BAIXA_FORA : * 0
R: * : * : ESTAVEL_ESTAVEL_FORA : * 2
R: * : * : ESTAVEL_ALTA_FORA : * 0
R: * : * : ALTA_BAIXA_FORA : * 0
R: * : * : ALTA_ESTAVEL_FORA : * 0
R: * : * : ALTA_ALTA_FORA : * -2
R: * : * : INDESEJAVEL : * -25

```

Exemplo de Modelagem POMDPX

Este apêndice apresenta um exemplo de modelagem MOMDP para o mercado financeiro conforme a descrição da Seção 5.2. Foi usada a proposta de APPL (*Approximate POMDP Planning Toolkit*)¹ (ONG et al., 2010) para expressar este problema, sendo tal representação chamada de POMDPX. Tal exemplo, mostrado a seguir, foi construído a partir dos dados do ativo CMIG4 para o período de 2000 a 2007. Para produzir as observações foi utilizada a análise técnica OST₁ conforme as regras mostradas na Tabela 7.

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <pomdp id="experimento_001" version="1.0" xmlns:xsi="http://www.w3.org/2001/
  XMLElementSchema-instance" xsi:noNamespaceSchemaLocation="pomdp.xsd">
3   <Description>Experimento 001 (ativo: CMIG4; análise técnica: OST 1; período: 2000 a
     2007)</Description>
4   <Discount>0.95</Discount>
5
6   <!-- Definição das variáveis. -->
7   <Variable>
8     <!-- Variável que define as posições do investidor. -->
9     <StateVar vnamePrev="POSICAO_INVESTIDOR_0" vnameCurr="POSICAO_INVESTIDOR_1"
       fullyObs="true">
10      <ValueEnum>COMPRADO VENDIDO FORA</ValueEnum>
11    </StateVar>
12
13    <!-- Variável que define as tendências passadas. -->
14    <StateVar vnamePrev="TENDENCIA_PASSADO_0" vnameCurr="TENDENCIA_PASSADO_1"
       fullyObs="true">
15      <ValueEnum>PASSADO_BULL PASSADO_BEAR PASSADO_FLAT</ValueEnum>
16    </StateVar>
17
18    <!-- Variável que define as tendências futuras. -->
19    <StateVar vnamePrev="TENDENCIA_FUTURO_0" vnameCurr="TENDENCIA_FUTURO_1" fullyObs=
       "false">
20      <ValueEnum>FUTURO_BULL FUTURO_BEAR FUTURO_FLAT</ValueEnum>
21    </StateVar>
22
23    <!-- Variável que define as observações. -->
24    <ObsVar vname="OBSERVACAO">

```

¹ Detalhes da modelagem disponíveis em: <<http://bigbird.comp.nus.edu.sg/pmwiki/farm/appl/index.php?n=Main.PomdpXDocumentation>>

```

25     <ValueEnum>PONTO_COMPRA PONTO_VENDA PONTO_NADA</ValueEnum>
26 </ObsVar>
27
28 <!-- Variável que define as ações. -->
29 <ActionVar vname="ACAO">
30     <ValueEnum>COMPRA COMPRA_DUPLA VENDA VENDA_DUPLA NADA</ValueEnum>
31 </ActionVar>
32
33 <!-- Variável que representa as recompensas. -->
34 <RewardVar vname="RECOMPENSA"/>
35 </Variable>
36
37 <!-- Atribuição dos valores iniciais dos estados. -->
38 <InitialStateBelief>
39     <CondProb>
40         <Var>POSICAO_INVESTIDOR_0</Var>
41         <Parent>null</Parent>
42         <Parameter type="TBL">
43             <Entry>
44                 <Instance>-</Instance>
45                 <ProbTable>0 0 1</ProbTable>
46             </Entry>
47         </Parameter>
48     </CondProb>
49     <CondProb>
50         <Var>TENDENCIA_PASSADO_0</Var>
51         <Parent>null</Parent>
52         <Parameter type="TBL">
53             <Entry>
54                 <Instance>-</Instance>
55                 <ProbTable>uniform</ProbTable>
56             </Entry>
57         </Parameter>
58     </CondProb>
59     <CondProb>
60         <Var>TENDENCIA_FUTURO_0</Var>
61         <Parent>null</Parent>
62         <Parameter type="TBL">
63             <Entry>
64                 <Instance>-</Instance>
65                 <ProbTable>uniform</ProbTable>
66             </Entry>
67         </Parameter>
68     </CondProb>
69 </InitialStateBelief>
70
71 <!-- Funções de transição entre as variáveis. -->
72 <StateTransitionFunction>
73     <CondProb>
74         <Var>POSICAO_INVESTIDOR_1</Var>
75         <Parent>ACAO POSICAO_INVESTIDOR_0</Parent>
76         <Parameter type="TBL">
77             <Entry>
78                 <Instance>COMPRA - </Instance>
79                 <ProbTable>1 0 0 0 0 1 1 0 0</ProbTable>
80             </Entry>
81             <Entry>
82                 <Instance>COMPRA_DUPLA - </Instance>
83                 <ProbTable>1 0 0 1 0 0 1 0 0</ProbTable>
84             </Entry>

```

```

85     <Entry>
86         <Instance>VENDA - -</Instance>
87         <ProbTable>0 0 1 0 1 0 0 1 0</ProbTable>
88     </Entry>
89     <Entry>
90         <Instance>VENDA_DUPLA - -</Instance>
91         <ProbTable>0 1 0 0 1 0 0 1 0</ProbTable>
92     </Entry>
93     <Entry>
94         <Instance>NADA - -</Instance>
95         <ProbTable>1 0 0 0 1 0 0 0 1</ProbTable>
96     </Entry>
97 </Parameter>
98 </CondProb>
99 <CondProb>
100     <Var>TENDENCIA_PASSADO_1</Var>
101     <Parent>ACAO TENDENCIA_PASSADO_0</Parent>
102     <Parameter type="TBL">
103         <Entry>
104             <Instance>* PASSADO_BULL -</Instance>
105             <ProbTable>0.7295081967 0.0601092896 0.2103825137 </ProbTable>
106         </Entry>
107         <Entry>
108             <Instance>* PASSADO_BEAR -</Instance>
109             <ProbTable>0.0586510264 0.7008797654 0.2404692082 </ProbTable>
110         </Entry>
111         <Entry>
112             <Instance>* PASSADO_FLAT -</Instance>
113             <ProbTable>0.2768959436 0.2839506173 0.4391534392 </ProbTable>
114         </Entry>
115     </Parameter>
116 </CondProb>
117 <CondProb>
118     <Var>TENDENCIA_FUTURO_1</Var>
119     <Parent>ACAO TENDENCIA_PASSADO_0 TENDENCIA_FUTURO_0</Parent>
120     <Parameter type="TBL">
121         <Entry>
122             <Instance>* PASSADO_BULL FUTURO_BULL -</Instance>
123             <ProbTable>0.6510067114 0.0570469799 0.2919463087 </ProbTable>
124         </Entry>
125         <Entry>
126             <Instance>* PASSADO_BEAR FUTURO_BULL -</Instance>
127             <ProbTable>0.7052631579 0.0789473684 0.2157894737 </ProbTable>
128         </Entry>
129         <Entry>
130             <Instance>* PASSADO_FLAT FUTURO_BULL -</Instance>
131             <ProbTable>0.7671957672 0.0370370370 0.1957671958 </ProbTable>
132         </Entry>
133         <Entry>
134             <Instance>* PASSADO_BULL FUTURO_BEAR -</Instance>
135             <ProbTable>0.0862068966 0.7370689655 0.1767241379 </ProbTable>
136         </Entry>
137         <Entry>
138             <Instance>* PASSADO_BEAR FUTURO_BEAR -</Instance>
139             <ProbTable>0.0453074434 0.7313915858 0.2233009709 </ProbTable>
140         </Entry>
141         <Entry>
142             <Instance>* PASSADO_FLAT FUTURO_BEAR -</Instance>
143             <ProbTable>0.0497512438 0.7313432836 0.2189054726 </ProbTable>
144         </Entry>

```

```

145     <Entry>
146         <Instance>* PASSADO_BULL FUTURO_FLAT -</Instance>
147         <ProbTable>0.3316831683 0.2524752475 0.4158415842 </ProbTable>
148     </Entry>
149     <Entry>
150         <Instance>* PASSADO_BEAR FUTURO_FLAT -</Instance>
151         <ProbTable>0.2841530055 0.3387978142 0.3770491803 </ProbTable>
152     </Entry>
153     <Entry>
154         <Instance>* PASSADO_FLAT FUTURO_FLAT -</Instance>
155         <ProbTable>0.2259887006 0.2655367232 0.5084745763 </ProbTable>
156     </Entry>
157 </Parameter>
158 </CondProb>
159 </StateTransitionFunction>
160
161 <!-- Funções de transição das observações. -->
162 <ObsFunction>
163     <CondProb>
164         <Var>OBSERVACAO</Var>
165         <Parent>ACAO TENDENCIA_PASSADO_1 TENDENCIA_FUTURO_1</Parent>
166         <Parameter type="TBL">
167             <Entry>
168                 <Instance>* PASSADO_BULL FUTURO_BULL -</Instance>
169                 <ProbTable>0.0202020202 0.0505050505 0.9292929293 </ProbTable>
170             </Entry>
171             <Entry>
172                 <Instance>* PASSADO_BEAR FUTURO_BULL -</Instance>
173                 <ProbTable>0.0526315789 0.0052631579 0.9421052632 </ProbTable>
174             </Entry>
175             <Entry>
176                 <Instance>* PASSADO_FLAT FUTURO_BULL -</Instance>
177                 <ProbTable>0.0317460317 0.0634920635 0.9047619048 </ProbTable>
178             </Entry>
179             <Entry>
180                 <Instance>* PASSADO_BULL FUTURO_BEAR -</Instance>
181                 <ProbTable>0.0301724138 0.0474137931 0.9224137931 </ProbTable>
182             </Entry>
183             <Entry>
184                 <Instance>* PASSADO_BEAR FUTURO_BEAR -</Instance>
185                 <ProbTable>0.0258064516 0.0483870968 0.9258064516 </ProbTable>
186             </Entry>
187             <Entry>
188                 <Instance>* PASSADO_FLAT FUTURO_BEAR -</Instance>
189                 <ProbTable>0.0597014925 0.0746268657 0.8656716418 </ProbTable>
190             </Entry>
191             <Entry>
192                 <Instance>* PASSADO_BULL FUTURO_FLAT -</Instance>
193                 <ProbTable>0.0198019802 0.0396039604 0.9405940594 </ProbTable>
194             </Entry>
195             <Entry>
196                 <Instance>* PASSADO_BEAR FUTURO_FLAT -</Instance>
197                 <ProbTable>0.0382513661 0.0382513661 0.9234972678 </ProbTable>
198             </Entry>
199             <Entry>
200                 <Instance>* PASSADO_FLAT FUTURO_FLAT -</Instance>
201                 <ProbTable>0.0508474576 0.0734463277 0.8757062147 </ProbTable>
202             </Entry>
203         </Parameter>
204     </CondProb>

```

```

205     </ObsFunction>
206
207     <!-- Recompensas atribuídas a transição entre variáveis. -->
208     <RewardFunction>
209         <Func>
210             <Var>RECOMPENSA</Var>
211             <Parent>ACAO POSICAO_INVESTIDOR_0 TENDENCIA_PASSADO_1 TENDENCIA_FUTURO_1</
                Parent>
212             <Parameter type="TBL">
213                 <Entry>
214                     <Instance>COMPRA COMPRADO * *</Instance>
215                     <ValueTable>-25</ValueTable>
216                 </Entry>
217                 <Entry>
218                     <Instance>COMPRA VENDIDO - -</Instance>
219                     <ValueTable>-2 0 0 0 -2 0 0 0 2</ValueTable>
220                 </Entry>
221                 <Entry>
222                     <Instance>COMPRA FORA - -</Instance>
223                     <ValueTable>10 0 3 5 -10 -2 5 -5 0</ValueTable>
224                 </Entry>
225                 <Entry>
226                     <Instance>COMPRA_DUPLA COMPRADO * *</Instance>
227                     <ValueTable>-25</ValueTable>
228                 </Entry>
229                 <Entry>
230                     <Instance>COMPRA_DUPLA VENDIDO - -</Instance>
231                     <ValueTable>10 0 3 5 -10 -2 5 -5 0</ValueTable>
232                 </Entry>
233                 <Entry>
234                     <Instance>COMPRA_DUPLA FORA * *</Instance>
235                     <ValueTable>-25</ValueTable>
236                 </Entry>
237                 <Entry>
238                     <Instance>VENDA COMPRADO - -</Instance>
239                     <ValueTable>-2 0 0 0 -2 0 0 0 2</ValueTable>
240                 </Entry>
241                 <Entry>
242                     <Instance>VENDA VENDIDO * *</Instance>
243                     <ValueTable>-25</ValueTable>
244                 </Entry>
245                 <Entry>
246                     <Instance>VENDA FORA - -</Instance>
247                     <ValueTable>-10 5 -2 0 10 3 -5 5 0</ValueTable>
248                 </Entry>
249                 <Entry>
250                     <Instance>VENDA_DUPLA COMPRADO - -</Instance>
251                     <ValueTable>-10 5 -2 0 10 3 -5 5 0</ValueTable>
252                 </Entry>
253                 <Entry>
254                     <Instance>VENDA_DUPLA VENDIDO * *</Instance>
255                     <ValueTable>-25</ValueTable>
256                 </Entry>
257                 <Entry>
258                     <Instance>VENDA_DUPLA FORA * *</Instance>
259                     <ValueTable>-25</ValueTable>
260                 </Entry>
261                 <Entry>
262                     <Instance>NADA COMPRADO - -</Instance>
263                     <ValueTable>10 0 3 5 -10 -2 5 -5 0</ValueTable>

```

```
264         </Entry>
265     <Entry>
266         <Instance>NADA VENDIDO - -</Instance>
267         <ValueTable>-10 5 -2 0 10 3 -5 5 0</ValueTable>
268     </Entry>
269     <Entry>
270         <Instance>NADA FORA - -</Instance>
271         <ValueTable>-2 0 0 0 -2 0 0 0 2</ValueTable>
272     </Entry>
273 </Parameter>
274 </Func>
275 </RewardFunction>
276 </pomdpX>
```


Exemplo de Modelagem RDDDL

Este apêndice apresenta um exemplo de modelagem MOMDP para o mercado financeiro utilizando a linguagem RDDDL. Detalhes da linguagem são colocadas na Seção 4.2.1. Tal exemplo, mostrado a seguir, foi construído a partir dos dados do índice IBOVESPA para o período ano de 2000.

```

domain market_domain {
  requirements = {
    continuous,
    reward-deterministic,
    partially-observed,
    multivalued
  };
  types {
    ativo: object;
    tendencia: { @alta, @baixa, @estavel };
    posicao: { @vendido, @comprado, @fora };
    observacao_enum: { @ponto_compra, @ponto_venda, @ponto_nada };
    acao: { @compra, @compra_dupla, @venda, @venda_dupla, @nada };
  };
  pvariables {
    TENDENCIA_PASSADO_PROB(tendencia, tendencia, tendencia):
      { non-fluent, real, default = 0.0 };
    TENDENCIA_FUTURO_PROB(tendencia, tendencia, tendencia):
      { non-fluent, real, default = 0.0 };
    OBSERVACAO(tendencia, tendencia, observacao_enum):
      { non-fluent, real, default = 0.0 };
    RECOMPENSA(acao, posicao, tendencia, tendencia):
      { non-fluent, int, default = 0 };
  }
}

```

```

posicao_investidor(posicao):
    { state-fluent, bool, default = false };
tendencia_passado(tendencia):
    { state-fluent, bool, default = false };
tendencia_futuro(tendencia):
    { state-fluent, bool, default = false };

negociar(acao):{ action-fluent, bool, default = false };

observacao_prob(observacao):{ observ-fluent, bool };
};
cpfs {
    posicao_investidor'(?p) =
        if (?p == @vendido) then
            (posicao_investidor(@fora) ∧ negociar(@venda)) |
            (posicao_investidor(@comprado) ∧ negociar(@venda_dupla)) |
            (posicao_investidor(@vendido) ∧ negociar(@venda)) |
            (posicao_investidor(@vendido) ∧ negociar(@venda_dupla)) |
            (posicao_investidor(@vendido) ∧ negociar(@nada)) |
            (posicao_investidor(@vendido) ∧ ~(exists_{?n: acao} negociar(?n)))
        else if (?p == @comprado) then
            (posicao_investidor(@fora) ∧ negociar(@compra)) |
            (posicao_investidor(@vendido) ∧ negociar(@compra_dupla)) |
            (posicao_investidor(@comprado) ∧ negociar(@compra)) |
            (posicao_investidor(@comprado) ∧ negociar(@compra_dupla)) |
            (posicao_investidor(@comprado) ∧ negociar(@nada)) |
            (posicao_investidor(@comprado) ∧ ~(exists_{?n: acao} negociar(?n)))
        else
            (posicao_investidor(@vendido) ∧ negociar(@compra)) |
            (posicao_investidor(@comprado) ∧ negociar(@venda)) |
            (posicao_investidor(@fora) ∧ negociar(@compra_dupla)) |
            (posicao_investidor(@fora) ∧ negociar(@venda_dupla)) |
            (posicao_investidor(@fora) ∧ negociar(@nada)) |
            (posicao_investidor(@fora) ∧ ~(exists_{?n: acao} negociar(?n)));

    tendencia_passado'(?x) =
        if (tendencia_passado(@alta) ∧ tendencia_futuro(@alta))
            then Bernoulli(TENDENCIA_PASSADO_PROB(@alta, @alta, ?x))
        else if (tendencia_passado(@alta) ∧ tendencia_futuro(@baixa))

```

```

        then Bernoulli(TENDENCIA_PASSADO_PROB(@alta, @baixa, ?x))
    else if (tendencia_passado(@alta) ∧ tendencia_futuro(@estavel))
        then Bernoulli(TENDENCIA_PASSADO_PROB(@alta, @estavel, ?x))
    else if (tendencia_passado(@baixa) ∧ tendencia_futuro(@alta))
        then Bernoulli(TENDENCIA_PASSADO_PROB(@baixa, @alta, ?x))
    else if (tendencia_passado(@baixa) ∧ tendencia_futuro(@baixa))
        then Bernoulli(TENDENCIA_PASSADO_PROB(@baixa, @baixa, ?x))
    else if (tendencia_passado(@baixa) ∧ tendencia_futuro(@estavel))
        then Bernoulli(TENDENCIA_PASSADO_PROB(@baixa, @estavel, ?x))
    else if (tendencia_passado(@estavel) ∧ tendencia_futuro(@alta))
        then Bernoulli(TENDENCIA_PASSADO_PROB(@estavel, @alta, ?x))
    else if (tendencia_passado(@estavel) ∧ tendencia_futuro(@baixa))
        then Bernoulli(TENDENCIA_PASSADO_PROB(@estavel, @baixa, ?x))
    else
        Bernoulli(TENDENCIA_PASSADO_PROB(@estavel, @estavel, ?x));

tendencia_futuro'(?x) =
    if (tendencia_passado(@alta) ∧ tendencia_futuro(@alta))
        then Bernoulli(TENDENCIA_FUTURO_PROB(@alta, @alta, ?x))
    else if (tendencia_passado(@alta) ∧ tendencia_futuro(@baixa))
        then Bernoulli(TENDENCIA_FUTURO_PROB(@alta, @baixa, ?x))
    else if (tendencia_passado(@alta) ∧ tendencia_futuro(@estavel))
        then Bernoulli(TENDENCIA_FUTURO_PROB(@alta, @estavel, ?x))
    else if (tendencia_passado(@baixa) ∧ tendencia_futuro(@alta))
        then Bernoulli(TENDENCIA_FUTURO_PROB(@baixa, @alta, ?x))
    else if (tendencia_passado(@baixa) ∧ tendencia_futuro(@baixa))
        then Bernoulli(TENDENCIA_FUTURO_PROB(@baixa, @baixa, ?x))
    else if (tendencia_passado(@baixa) ∧ tendencia_futuro(@estavel))
        then Bernoulli(TENDENCIA_FUTURO_PROB(@baixa, @estavel, ?x))
    else if (tendencia_passado(@estavel) ∧ tendencia_futuro(@alta))
        then Bernoulli(TENDENCIA_FUTURO_PROB(@estavel, @alta, ?x))
    else if (tendencia_passado(@estavel) ∧ tendencia_futuro(@baixa))
        then Bernoulli(TENDENCIA_FUTURO_PROB(@estavel, @baixa, ?x))
    else
        Bernoulli(TENDENCIA_FUTURO_PROB(@estavel, @estavel, ?x));

observacao(?x) =
    if (tendencia_passado(@alta) ∧ tendencia_futuro(@alta))
        then Bernoulli(OBSERVACAO_PROB(@alta, @alta, ?x))

```

```

    else if (tendencia_passado(@baixa) ∧ tendencia_futuro(@alta))
        then Bernoulli(OBSERVACAO_PROB(@baixa, @alta, ?x))
    else if (tendencia_passado(@estavel) ∧ tendencia_futuro(@alta))
        then Bernoulli(OBSERVACAO_PROB(@estavel, @alta, ?x))
    else if (tendencia_passado(@alta) ∧ tendencia_futuro(@baixa))
        then Bernoulli(OBSERVACAO_PROB(@alta, @baixa, ?x))
    else if (tendencia_passado(@baixa) ∧ tendencia_futuro(@baixa))
        then Bernoulli(OBSERVACAO_PROB(@baixa, @baixa, ?x))
    else if (tendencia_passado(@estavel) ∧ tendencia_futuro(@baixa))
        then Bernoulli(OBSERVACAO_PROB(@estavel, @baixa, ?x))
    else if (tendencia_passado(@alta) ∧ tendencia_futuro(@estavel))
        then Bernoulli(OBSERVACAO_PROB(@alta, @estavel, ?x))
    else if (tendencia_passado(@baixa) ∧ tendencia_futuro(@estavel))
        then Bernoulli(OBSERVACAO_PROB(@baixa, @estavel, ?x))
    else
        Bernoulli(OBSERVACAO_PROB(@estavel, @estavel, ?x));
};

reward = sum_{ ?a:acao, ?p:posicao, ?tp:tendencia, ?tf:tendencia } [
    if (negociar(?a) ∧ posicao_investidor(?p) ∧
        tendencia_passado'(?tp) ∧ tendencia_futuro'(?tf)) then
        RECOMPENSA(?a, ?p, ?tp, ?tf)
    else
        0
];
}

```

A seguir é apresentada a parte *non-fluents*.

```

non-fluents market_non_fluent {
    domain = market_domain;

    objects {
        ativo : { bvsp };
    };

    non-fluents {
        TENDENCIA_PASSADO_PROB(@alta, @alta, @alta) = 0.780;
        TENDENCIA_PASSADO_PROB(@alta, @baixa, @alta) = 0.603;
    };
}

```

```

TENDENCIA_PASSADO_PROB(@alta, @estavel, @alta) = 0.737;
TENDENCIA_PASSADO_PROB(@baixa, @alta, @alta) = 0.029;
TENDENCIA_PASSADO_PROB(@baixa, @baixa, @alta) = 0.008;
TENDENCIA_PASSADO_PROB(@baixa, @estavel, @alta) = 0.010;
TENDENCIA_PASSADO_PROB(@estavel, @alta, @alta) = 0.337;
TENDENCIA_PASSADO_PROB(@estavel, @baixa, @alta) = 0.124;
TENDENCIA_PASSADO_PROB(@estavel, @estavel, @alta) = 0.234;

TENDENCIA_PASSADO_PROB(@alta, @alta, @baixa) = 0.010;
TENDENCIA_PASSADO_PROB(@alta, @baixa, @baixa) = 0.021;
TENDENCIA_PASSADO_PROB(@alta, @estavel, @baixa) = 0.017;
TENDENCIA_PASSADO_PROB(@baixa, @alta, @baixa) = 0.615;
TENDENCIA_PASSADO_PROB(@baixa, @baixa, @baixa) = 0.758;
TENDENCIA_PASSADO_PROB(@baixa, @estavel, @baixa) = 0.746;
TENDENCIA_PASSADO_PROB(@estavel, @alta, @baixa) = 0.095;
TENDENCIA_PASSADO_PROB(@estavel, @baixa, @baixa) = 0.316;
TENDENCIA_PASSADO_PROB(@estavel, @estavel, @baixa) = 0.195;

TENDENCIA_PASSADO_PROB(@alta, @alta, @estavel) = 0.210;
TENDENCIA_PASSADO_PROB(@alta, @baixa, @estavel) = 0.376;
TENDENCIA_PASSADO_PROB(@alta, @estavel, @estavel) = 0.246;
TENDENCIA_PASSADO_PROB(@baixa, @alta, @estavel) = 0.356;
TENDENCIA_PASSADO_PROB(@baixa, @baixa, @estavel) = 0.234;
TENDENCIA_PASSADO_PROB(@baixa, @estavel, @estavel) = 0.244;
TENDENCIA_PASSADO_PROB(@estavel, @alta, @estavel) = 0.568;
TENDENCIA_PASSADO_PROB(@estavel, @baixa, @estavel) = 0.560;
TENDENCIA_PASSADO_PROB(@estavel, @estavel, @estavel) = 0.571;

TENDENCIA_FUTURO_PROB(@alta, @alta, @alta) = 0.698;
TENDENCIA_FUTURO_PROB(@alta, @baixa, @alta) = 0.016;
TENDENCIA_FUTURO_PROB(@alta, @estavel, @alta) = 0.233;
TENDENCIA_FUTURO_PROB(@baixa, @alta, @alta) = 0.726;
TENDENCIA_FUTURO_PROB(@baixa, @baixa, @alta) = 0.024;
TENDENCIA_FUTURO_PROB(@baixa, @estavel, @alta) = 0.254;
TENDENCIA_FUTURO_PROB(@estavel, @alta, @alta) = 0.715;
TENDENCIA_FUTURO_PROB(@estavel, @baixa, @alta) = 0.019;
TENDENCIA_FUTURO_PROB(@estavel, @estavel, @alta) = 0.255;

TENDENCIA_FUTURO_PROB(@alta, @alta, @baixa) = 0.029;

```

```

TENDENCIA_FUTURO_PROB(@alta, @baixa, @baixa) = 0.746;
TENDENCIA_FUTURO_PROB(@alta, @estavel, @baixa) = 0.178;
TENDENCIA_FUTURO_PROB(@baixa, @alta, @baixa) = 0.010;
TENDENCIA_FUTURO_PROB(@baixa, @baixa, @baixa) = 0.613;
TENDENCIA_FUTURO_PROB(@baixa, @estavel, @baixa) = 0.218;
TENDENCIA_FUTURO_PROB(@estavel, @alta, @baixa) = 0.011;
TENDENCIA_FUTURO_PROB(@estavel, @baixa, @baixa) = 0.708;
TENDENCIA_FUTURO_PROB(@estavel, @estavel, @baixa) = 0.181;

TENDENCIA_FUTURO_PROB(@alta, @alta, @estavel) = 0.273;
TENDENCIA_FUTURO_PROB(@alta, @baixa, @estavel) = 0.238;
TENDENCIA_FUTURO_PROB(@alta, @estavel, @estavel) = 0.589;
TENDENCIA_FUTURO_PROB(@baixa, @alta, @estavel) = 0.264;
TENDENCIA_FUTURO_PROB(@baixa, @baixa, @estavel) = 0.363;
TENDENCIA_FUTURO_PROB(@baixa, @estavel, @estavel) = 0.528;
TENDENCIA_FUTURO_PROB(@estavel, @alta, @estavel) = 0.274;
TENDENCIA_FUTURO_PROB(@estavel, @baixa, @estavel) = 0.273;
TENDENCIA_FUTURO_PROB(@estavel, @estavel, @estavel) = 0.564;

OBSERVACAO_PROB(@alta, @alta, @ponto_compra) = 0.000;
OBSERVACAO_PROB(@baixa, @alta, @ponto_compra) = 0.043;
OBSERVACAO_PROB(@estavel, @alta, @ponto_compra) = 0.021;
OBSERVACAO_PROB(@alta, @baixa, @ponto_compra) = 0.005;
OBSERVACAO_PROB(@baixa, @baixa, @ponto_compra) = 0.017;
OBSERVACAO_PROB(@estavel, @baixa, @ponto_compra) = 0.029;
OBSERVACAO_PROB(@alta, @estavel, @ponto_compra) = 0.014;
OBSERVACAO_PROB(@baixa, @estavel, @ponto_compra) = 0.036;
OBSERVACAO_PROB(@estavel, @estavel, @ponto_compra) = 0.039;

OBSERVACAO_PROB(@alta, @alta, @ponto_venda) = 0.054;
OBSERVACAO_PROB(@baixa, @alta, @ponto_venda) = 0.010;
OBSERVACAO_PROB(@estavel, @alta, @ponto_venda) = 0.066;
OBSERVACAO_PROB(@alta, @baixa, @ponto_venda) = 0.022;
OBSERVACAO_PROB(@baixa, @baixa, @ponto_venda) = 0.017;
OBSERVACAO_PROB(@estavel, @baixa, @ponto_venda) = 0.063;
OBSERVACAO_PROB(@alta, @estavel, @ponto_venda) = 0.034;
OBSERVACAO_PROB(@baixa, @estavel, @ponto_venda) = 0.031;
OBSERVACAO_PROB(@estavel, @estavel, @ponto_venda) = 0.035;

```

```

OBSERVACAO_PROB(@alta, @alta, @ponto_nada) = 0.946;
OBSERVACAO_PROB(@baixa, @alta, @ponto_nada) = 0.947;
OBSERVACAO_PROB(@estavel, @alta, @ponto_nada) = 0.913;
OBSERVACAO_PROB(@alta, @baixa, @ponto_nada) = 0.973;
OBSERVACAO_PROB(@baixa, @baixa, @ponto_nada) = 0.966;
OBSERVACAO_PROB(@estavel, @baixa, @ponto_nada) = 0.908;
OBSERVACAO_PROB(@alta, @estavel, @ponto_nada) = 0.952;
OBSERVACAO_PROB(@baixa, @estavel, @ponto_nada) = 0.933;
OBSERVACAO_PROB(@estavel, @estavel, @ponto_nada) = 0.926;

```

```

RECOMPENSA(@compra, @comprado, @alta, @alta) = -25;
RECOMPENSA(@compra, @comprado, @alta, @baixa) = -25;
RECOMPENSA(@compra, @comprado, @alta, @estavel) = -25;
RECOMPENSA(@compra, @comprado, @baixa, @alta) = -25;
RECOMPENSA(@compra, @comprado, @baixa, @baixa) = -25;
RECOMPENSA(@compra, @comprado, @baixa, @estavel) = -25;
RECOMPENSA(@compra, @comprado, @estavel, @alta) = -25;
RECOMPENSA(@compra, @comprado, @estavel, @baixa) = -25;
RECOMPENSA(@compra, @comprado, @estavel, @estavel) = -25;
RECOMPENSA(@compra, @vendido, @alta, @alta) = -2;
RECOMPENSA(@compra, @vendido, @alta, @baixa) = 0;
RECOMPENSA(@compra, @vendido, @alta, @estavel) = 0;
RECOMPENSA(@compra, @vendido, @baixa, @alta) = 0;
RECOMPENSA(@compra, @vendido, @baixa, @baixa) = -2;
RECOMPENSA(@compra, @vendido, @baixa, @estavel) = 0;
RECOMPENSA(@compra, @vendido, @estavel, @alta) = 0;
RECOMPENSA(@compra, @vendido, @estavel, @baixa) = 0;
RECOMPENSA(@compra, @vendido, @estavel, @estavel) = 2;
RECOMPENSA(@compra, @fora, @alta, @alta) = 10;
RECOMPENSA(@compra, @fora, @alta, @baixa) = 0;
RECOMPENSA(@compra, @fora, @alta, @estavel) = 3;
RECOMPENSA(@compra, @fora, @baixa, @alta) = 5;
RECOMPENSA(@compra, @fora, @baixa, @baixa) = -10;
RECOMPENSA(@compra, @fora, @baixa, @estavel) = -2;
RECOMPENSA(@compra, @fora, @estavel, @alta) = 5;
RECOMPENSA(@compra, @fora, @estavel, @baixa) = -5;
RECOMPENSA(@compra, @fora, @estavel, @estavel) = 0;

```

```

RECOMPENSA(@compra_dupla, @comprado, @alta, @alta) = -25;

```



```

RECOMPENSA(@compra_dupla, @comprado, @alta, @baixa) = -25;
RECOMPENSA(@compra_dupla, @comprado, @alta, @estavel) = -25;
RECOMPENSA(@compra_dupla, @comprado, @baixa, @alta) = -25;
RECOMPENSA(@compra_dupla, @comprado, @baixa, @baixa) = -25;
RECOMPENSA(@compra_dupla, @comprado, @baixa, @estavel) = -25;
RECOMPENSA(@compra_dupla, @comprado, @estavel, @alta) = -25;
RECOMPENSA(@compra_dupla, @comprado, @estavel, @baixa) = -25;
RECOMPENSA(@compra_dupla, @comprado, @estavel, @estavel) = -25;
RECOMPENSA(@compra_dupla, @vendido, @alta, @alta) = 10;
RECOMPENSA(@compra_dupla, @vendido, @alta, @baixa) = 0;
RECOMPENSA(@compra_dupla, @vendido, @alta, @estavel) = 3;
RECOMPENSA(@compra_dupla, @vendido, @baixa, @alta) = 5;
RECOMPENSA(@compra_dupla, @vendido, @baixa, @baixa) = -10;
RECOMPENSA(@compra_dupla, @vendido, @baixa, @estavel) = -2;
RECOMPENSA(@compra_dupla, @vendido, @estavel, @alta) = 5;
RECOMPENSA(@compra_dupla, @vendido, @estavel, @baixa) = -5;
RECOMPENSA(@compra_dupla, @vendido, @estavel, @estavel) = 0;
RECOMPENSA(@compra_dupla, @fora, @alta, @alta) = -25;
RECOMPENSA(@compra_dupla, @fora, @alta, @baixa) = -25;
RECOMPENSA(@compra_dupla, @fora, @alta, @estavel) = -25;
RECOMPENSA(@compra_dupla, @fora, @baixa, @alta) = -25;
RECOMPENSA(@compra_dupla, @fora, @baixa, @baixa) = -25;
RECOMPENSA(@compra_dupla, @fora, @baixa, @estavel) = -25;
RECOMPENSA(@compra_dupla, @fora, @estavel, @alta) = -25;
RECOMPENSA(@compra_dupla, @fora, @estavel, @baixa) = -25;
RECOMPENSA(@compra_dupla, @fora, @estavel, @estavel) = -25;

RECOMPENSA(@venda, @comprado, @alta, @alta) = -2;
RECOMPENSA(@venda, @comprado, @alta, @baixa) = 0;
RECOMPENSA(@venda, @comprado, @alta, @estavel) = 0;
RECOMPENSA(@venda, @comprado, @baixa, @alta) = 0;
RECOMPENSA(@venda, @comprado, @baixa, @baixa) = -2;
RECOMPENSA(@venda, @comprado, @baixa, @estavel) = 0;
RECOMPENSA(@venda, @comprado, @estavel, @alta) = 0;
RECOMPENSA(@venda, @comprado, @estavel, @baixa) = 0;
RECOMPENSA(@venda, @comprado, @estavel, @estavel) = 2;
RECOMPENSA(@venda, @vendido, @alta, @alta) = -25;
RECOMPENSA(@venda, @vendido, @alta, @baixa) = -25;
RECOMPENSA(@venda, @vendido, @alta, @estavel) = -25;

```

```

RECOMPENSA(@venda, @vendido, @baixa, @alta) = -25;
RECOMPENSA(@venda, @vendido, @baixa, @baixa) = -25;
RECOMPENSA(@venda, @vendido, @baixa, @estavel) = -25;
RECOMPENSA(@venda, @vendido, @estavel, @alta) = -25;
RECOMPENSA(@venda, @vendido, @estavel, @baixa) = -25;
RECOMPENSA(@venda, @vendido, @estavel, @estavel) = -25;
RECOMPENSA(@venda, @fora, @alta, @alta) = -10;
RECOMPENSA(@venda, @fora, @alta, @baixa) = 5;
RECOMPENSA(@venda, @fora, @alta, @estavel) = -2;
RECOMPENSA(@venda, @fora, @baixa, @alta) = 0;
RECOMPENSA(@venda, @fora, @baixa, @baixa) = 10;
RECOMPENSA(@venda, @fora, @baixa, @estavel) = 3;
RECOMPENSA(@venda, @fora, @estavel, @alta) = -5;
RECOMPENSA(@venda, @fora, @estavel, @baixa) = 5;
RECOMPENSA(@venda, @fora, @estavel, @estavel) = 0;

RECOMPENSA(@venda_dupla, @comprado, @alta, @alta) = -10;
RECOMPENSA(@venda_dupla, @comprado, @alta, @baixa) = 5;
RECOMPENSA(@venda_dupla, @comprado, @alta, @estavel) = -2;
RECOMPENSA(@venda_dupla, @comprado, @baixa, @alta) = 0;
RECOMPENSA(@venda_dupla, @comprado, @baixa, @baixa) = 10;
RECOMPENSA(@venda_dupla, @comprado, @baixa, @estavel) = 3;
RECOMPENSA(@venda_dupla, @comprado, @estavel, @alta) = -5;
RECOMPENSA(@venda_dupla, @comprado, @estavel, @baixa) = 5;
RECOMPENSA(@venda_dupla, @comprado, @estavel, @estavel) = 0;
RECOMPENSA(@venda_dupla, @vendido, @alta, @alta) = -25;
RECOMPENSA(@venda_dupla, @vendido, @alta, @baixa) = -25;
RECOMPENSA(@venda_dupla, @vendido, @alta, @estavel) = -25;
RECOMPENSA(@venda_dupla, @vendido, @baixa, @alta) = -25;
RECOMPENSA(@venda_dupla, @vendido, @baixa, @baixa) = -25;
RECOMPENSA(@venda_dupla, @vendido, @baixa, @estavel) = -25;
RECOMPENSA(@venda_dupla, @vendido, @estavel, @alta) = -25;
RECOMPENSA(@venda_dupla, @vendido, @estavel, @baixa) = -25;
RECOMPENSA(@venda_dupla, @vendido, @estavel, @estavel) = -25;
RECOMPENSA(@venda_dupla, @fora, @alta, @alta) = -25;
RECOMPENSA(@venda_dupla, @fora, @alta, @baixa) = -25;
RECOMPENSA(@venda_dupla, @fora, @alta, @estavel) = -25;
RECOMPENSA(@venda_dupla, @fora, @baixa, @alta) = -25;
RECOMPENSA(@venda_dupla, @fora, @baixa, @baixa) = -25;

```

```

RECOMPENSA(@venda_dupla, @fora, @baixa, @estavel) = -25;
RECOMPENSA(@venda_dupla, @fora, @estavel, @alta) = -25;
RECOMPENSA(@venda_dupla, @fora, @estavel, @baixa) = -25;
RECOMPENSA(@venda_dupla, @fora, @estavel, @estavel) = -25;

RECOMPENSA(@nada, @comprado, @alta, @alta) = 10;
RECOMPENSA(@nada, @comprado, @alta, @baixa) = 0;
RECOMPENSA(@nada, @comprado, @alta, @estavel) = 3;
RECOMPENSA(@nada, @comprado, @baixa, @alta) = 5;
RECOMPENSA(@nada, @comprado, @baixa, @baixa) = -10;
RECOMPENSA(@nada, @comprado, @baixa, @estavel) = -2;
RECOMPENSA(@nada, @comprado, @estavel, @alta) = 5;
RECOMPENSA(@nada, @comprado, @estavel, @baixa) = -5;
RECOMPENSA(@nada, @comprado, @estavel, @estavel) = 0;
RECOMPENSA(@nada, @vendido, @alta, @alta) = -10;
RECOMPENSA(@nada, @vendido, @alta, @baixa) = 5;
RECOMPENSA(@nada, @vendido, @alta, @estavel) = -2;
RECOMPENSA(@nada, @vendido, @baixa, @alta) = 0;
RECOMPENSA(@nada, @vendido, @baixa, @baixa) = 10;
RECOMPENSA(@nada, @vendido, @baixa, @estavel) = 3;
RECOMPENSA(@nada, @vendido, @estavel, @alta) = -5;
RECOMPENSA(@nada, @vendido, @estavel, @baixa) = 5;
RECOMPENSA(@nada, @vendido, @estavel, @estavel) = 0;
RECOMPENSA(@nada, @fora, @alta, @alta) = -2;
RECOMPENSA(@nada, @fora, @alta, @baixa) = 0;
RECOMPENSA(@nada, @fora, @alta, @estavel) = 0;
RECOMPENSA(@nada, @fora, @baixa, @alta) = 0;
RECOMPENSA(@nada, @fora, @baixa, @baixa) = -2;
RECOMPENSA(@nada, @fora, @baixa, @estavel) = 0;
RECOMPENSA(@nada, @fora, @estavel, @alta) = 0;
RECOMPENSA(@nada, @fora, @estavel, @baixa) = 0;
RECOMPENSA(@nada, @fora, @estavel, @estavel) = 2;
};
}

```

A seguir é apresentada a parte *instance*.

```

instance market_instance_01 {
    domain = market_domain;

```

```
non-fluents = market_non_fluent;

init-state {
    posicao_investidor(@fora);
};

max-nondef-actions = 2;
horizon = 40;
discount = 0.9;
}
```