



**UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

FELIPE HOUAT DE BRITO

DM 13/2011

**GAIA - UMA PROPOSTA DE AGENTE INTELIGENTE
ACOPLADO A ALGORITMOS EVOLUCIONÁRIOS
PARA AJUSTE DINÂMICO DE PARÂMETROS
ATRAVÉS DE CONCEITOS DE DESIGN
INTELIGENTE E LÓGICA FUZZY**

Belém, Março de 2011.

**UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA**

FELIPE HOUAT DE BRITO

DM 13/2011

**GAIA - UMA PROPOSTA DE AGENTE INTELIGENTE
ACOPLADO A ALGORITMOS EVOLUCIONÁRIOS
PARA AJUSTE DINÂMICO DE PARÂMETROS
ATRAVÉS DE CONCEITOS DE DESIGN
INTELIGENTE E LÓGICA FUZZY**

Dissertação submetida à banca examinadora do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Pará, como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica.

Prof. Ádamo Lima de Santana, Dr.
Orientador

Belém, Março de 2011.

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

FELIPE HOUAT DE BRITO

**GAIA - UMA PROPOSTA DE AGENTE INTELIGENTE
ACOPLADO A ALGORITMOS EVOLUCIONÁRIOS
PARA AJUSTE DINÂMICO DE PARÂMETROS
ATRAVÉS DE CONCEITOS DE DESIGN
INTELIGENTE E LÓGICA FUZZY**

Esta Dissertação foi julgada adequada para obtenção do título de Mestre em Engenharia Elétrica Área de Concentração (Computação Aplicada) e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica.

Aprovada em: 31/03/2011

Banca Examinadora:

Prof. Ádamo Lima de Santana, Dr.
Orientador
Universidade Federal do Pará – UFPA/PPGEE

Prof. Roberto Célio Limão de Oliveira, Dr.
Membro da Banca
Universidade Federal do Pará – UFPA/PPGEE

Prof. Marco Jose de Sousa, Dr.
Membro da Banca
Universidade Federal do Pará – UFPA/PPGEE

Prof. Dionne Cavalcante Monteiro, Dr.
Membro da Banca
Universidade Federal do Pará – UFPA/PPGCC

Visto

Prof. Marcus Vinicius Alves Nunes, Dr.
Coordenador do PPGEE
Universidade Federal do Pará – UFPA

Belém, Março de 2011.

A Deus, meus pais, queridos irmã e irmão, meus avós e meus mestres.

AGRADECIMENTOS

Aos meus pais, que desde pequeno, seguraram minha mão pelas dificuldades da vida até que eu aprendesse a caminhar sozinho, cuidando sempre de mim como maior prioridade em sua vida não importando qual fosse a dificuldade de fazê-lo, oferecendo-me o acesso à boa educação e a possibilidade de realizar um curso superior, fatores que contribuíram para a ótima formação de meu caráter.

À minha “mãezinha”, minha vida, minha luz e meu abrigo, por tudo de que sempre abriu (e abre até hoje) mão pela minha felicidade e sucesso. Nos longos anos passados e em todas as transformações que ocorreram comigo nesse período de mestrado, recebi dela de comida para continuar escrevendo, até a mão e um abraço em cada um de meus tropeços.

À minha irmã Carla, que, apesar de mais nova, durante toda a minha vida sempre procurou cuidar de mim e me ajudar a superar as dificuldades. Seus cuidados, atenção, paciência e até mesmo as brigas (!), foram de fundamental importância para minha vida.

Ao meu irmão caçula Marcelo, futuro médico, que busca sempre seguir todos os bons exemplos de aluno e de vida que lhe dei. Uma das poucas pessoas que conheço, que ainda mais novo e menos experiente, é capaz de entender as minhas mais malucas idéias. Sua ajuda na extensa codificação da função de *fitness*, ainda no meu primeiro trabalho com Computação Evolucionária no ano de 2004, foi fundamental e impressionante para um “menino” (à época), de apenas 14 anos.

Aos meus avós Stephan Houat (*in memorian*), Jacqueline Houat, Thereza Brito e Ruy Nobre de Brito (*in memorian*), que além de acreditarem muito no meu potencial, sempre contribuíram de forma direta para minha formação, servindo sempre como espelho e exemplos a serem seguidos.

À minha princesa, minha noiva e futura esposa, Priscila Alves Noronha, que por tantas vezes, mesmo que a magoasse, abriu mão do convívio comigo, sempre em prol da minha carreira profissional. Uma das poucas pessoas que consegui me amar, compreender-me, perdoar-me e tolerar todos os meus vacilos, por vezes, tudo isso ao mesmo tempo. Sei que ainda venceremos muitas batalhas juntos e logo formaremos nosso “*lar dos sonhos*”.

Ao grande amigo e “irmão por opção” Artur Noura Texeira. Conheci-o por meio de seu irmão, meu amigo e co-orientador, durante o mestrado e, nos muitos desafios impostos por este curso que vencemos, formamos uma grande amizade. Foram ótimos anos de trabalho, conversas, troca de experiências, madrugadas trabalhando, “farras”, companheirismo e até mesmo de colega de trabalho durante minha passagem pelo SERPRO. Tenho certeza de que esse convívio mudou para sempre nossas vidas. Tive a oportunidade de apresentá-lo a sua atual esposa, indiquei a ele que fizesse o concurso do SERPRO (seu atual emprego) e alcançamos muitos objetivos profissionais, os quais, para mim pareciam impossíveis, como as publicações e defesas internacionais na China e em Las Vegas. Enfim, são tantas histórias vividas, que seria preciso uma dissertação inteira apenas para contá-las. Não posso também esquecer de dizer que muito da proposta inicial deste trabalho foi concebida em parceria com ele. Por isso tudo e mais, obrigado amigo!

Ao meu co-orientador e amigo Otávio Noura Texeira, que sempre foi minha inspiração de profissional ideal da área acadêmica. Desde meu TCC, ainda nos tempos de CESUPA, foi responsável por redespertar os meus sonhos de criança de ser pesquisador e cientista, sempre inovando, descobrindo e explorando. Tal qual, o período de meu TCC possibilitou-me realizar este trabalho, sempre evitando me influenciar muito, deixando com que eu pudesse escolher as direções a seguir, até mesmo como forma de me fazer crescer e “evoluir”. Hoje, após nossas inúmeras publicações e outros trabalhos, tenho certeza que a

confiança e a troca de informações é uma via de mão-dupla. Ainda iremos realizar muitos projetos de sucesso juntos.

Ao meu orientador Professor Dr. Roberto Célio Limão, que apesar de seu “ar austero” e disciplinador, mostrou-se um grande parceiro, compreensivo com as dificuldades e grande incentivador de todas as publicações e viagens que fizemos. Sem ele, com certeza, na haveria dissertação, muito menos publicações e boas histórias das viagens. Admiro-o, e desejo ainda mais sucesso para a ótima fase profissional que ele vive. Obrigado, e até o doutorado, quem sabe...

Ao amigo, companheiro de política, colega de profissão e muitas vezes um irmão mais velho, Antonio Carlos Melo (ACM), que sempre acreditou em mim, mesmo nos momentos mais difíceis e, permitiu-me sempre mostrar meu valor através de grandes desafios que me ofereceu. O ACM é uma das poucas pessoas que soube me valorizar profissionalmente desde quando eu era um simples estagiário do SERPRO, até os momentos mais duros de minha carreira. Se estou onde estou profissionalmente, boa parte é graças a ele, obrigado irmão!

Ao meu ex-chefe e grande mestre Ticiano Monteiro, de professor no CESUPA, a chefe no SERPRO durante quase três anos. É, e sempre será meu grande modelo de gestor profissional. Posso dizer que se hoje faço isso bem, devo muito a sua forma inovadora de gestão e de sempre buscar a paz e a competência individual de todos. Obrigado também por me valorizar, confiar em mim e entregar grandes desafios. Conte sempre com um amigo!

À esposa do Ticiano, minha ex-coordenadora do curso de Ciência da Computação no CESUPA, Professora Leda Monteiro, que sempre me apoiou e esteve por perto nos momentos mais importantes de minha carreira acadêmica. Desde minha escolha profissional no momento “Direito x Ciência da Computação”, até sua aparição inesperada no dia da defesa desta dissertação. Mais uma vez, muito obrigado por tudo!

À família SERPRO, que me acolheu como estagiário durante dois anos, oferecendo-me intensas experiências e oportunidades de aprendizado que eu jamais poderia ter em outro lugar. Voltei àquela casa como concursado e, por quatro anos participei lá de inúmeras realizações de pequenos à gigantescos projetos por todo o país. Com muito pesar e dor no coração, outras oportunidades apareceram na vida e precisei agarrá-las, deixando mais uma vez essa família. Saudades e gratidão sempre. A formação como profissional que lá recebi e para sempre, assim como os amigos que lá fiz, são amigos pra vida toda, obrigado a todos.

Aos novos e velhos colegas e amigos do TRE-PA, neste novo momento de minha carreira, por toda uma nova experiência, que já me proporcionou algumas boas aventuras e muitos aprendizados pelo nosso enorme Pará. Viagens de barco pelo rio Amazonas, estradas de terra em encostas de serras, repteis perigosos e muitas histórias, tudo em prol da cidadania e da garantia dos direitos de cada um dos brasileiros. Espero contribuir e aprender ainda mais e, prometo sempre tentar fazer jus ao apelido que recebi, *Bear Grylls* (“À prova de tudo”, *Discovery Channel*). Agradeço a todos da equipe, por me apoiarem neste momento que culmina o encerramento de uma longa jornada de minha vida.

À família MARISTA, que durante minha formação no ensino fundamental e médio, construiu muitos dos alicerces de minha educação moral, civil e religiosa.

Às pessoas, eventos e acontecimentos que tentaram contribuir para que este trabalho não fosse concluído, desestabilizando-me, adoentando-me, invejando-me, entristecendo-me ou tentando tirar minhas forças. Todos esses eventos, que nesses anos foram um grande desafio, serviram para dar maior gosto de vitória e glória a este término de trabalho, fazendo-me sempre crescer e “evoluir”. Citando Darwin: “*the survival of the fittest*”.

Ao ano de 2009, sem dúvida, um ano de luz, mudanças positivas constantes, retomada de auto-estima e crescimento em minha vida pessoal e profissional. Em nossa vida, de tempos em tempos, tal qual na proposta do trabalho, faz-se preciso um grande recomeço e uma

limpeza de tudo que nos acomoda e nos coloca para trás. Obrigado a este ano. Espero que outros como ele surjam novamente.

Aos meus queridos amigos, alguns de longa data e outros mais recentes, mas sem dúvida, do mais alto calibre, Mario Fima (meu compadre), Iuri Fernandes, Juliana Carvalho, Claudionor Vieira, José Vicente (Neto), Ewerton Leite, Adgerson Galeno, Alexandre Maradei e tantos outros que não estão aqui listados, mas que sem dúvida estão no meu coração..Obrigado pelo apoio, ombro nos momentos mais difíceis e por todos os bons momentos que passamos juntos!

A Deus, Jesus e minha mãe Nossa Senhora de Nazaré, e todos os seres de luz, que me acompanham sempre em toda minha vida, sem nunca me abandonar e que, de maneira misteriosa e magnífica, fazem com que todas as coisas dêem certo ao final. Obrigado sempre por me proteger, oportunizar, iluminar e guiar!

“Algo só é impossível até que alguém duvide e resolva provar o contrário.”

Albert Einstein

SUMÁRIO

LISTA DE FIGURAS	xvi
LISTA DE TABELAS	xviii
LISTA DE GRÁFICOS	xix
LISTA DE PUBLICAÇÕES.....	xxii
RESUMO	xxiv
ABSTRACT	xxv

CAPÍTULO I

1. INTRODUÇÃO.....	26
1.1 OBJETIVOS DA DISSERTAÇÃO	28
1.1.1 Objetivo Geral.....	29
1.1.2 Objetivos Específicos	29
1.2 METODOLOGIA E PESQUISA ADOTADA	30
1.3 ESTRUTURA GERAL DA DISSERTAÇÃO.....	31

CAPÍTULO II

2. ASPECTOS BIOLÓGICOS FUNDAMENTAIS: ORIGEM DA VIDA, EVOLUÇÃO E GENÉTICA	34
2.1 ORIGEM E EVOLUÇÃO DO UNIVERSO	34
2.2 EVOLUÇÃO E A “ORIGEM DA ESPECIES”	37
2.2.1 Histórico do Desenvolvimento de Teorias.....	38
2.2.2 Seleção Natural de Charles Darwin.....	41
2.3 NEODARWINISMO: EVOLUÇÃO E GENÉTICA.....	45
2.3.1 Histórico.....	46
2.3.2 Conceitos Fundamentais de Genética.....	47
2.3.3 Leis de Mendel	52
2.4 CONSIDERAÇÕES FINAIS.....	54

CAPÍTULO III

3. ALGORITMOS GENÉTICOS E UMA VISÃO GERAL DE COMPUTAÇÃO EVOLUCIONÁRIA.....	55
3.1 HISTÓRICO E CONTEXTUALIZAÇÃO	59
3.2 IMPLEMENTAÇÃO DE ALGORITMOS GENÉTICOS	60
3.2.1. Operação de Seleção.....	64
3.2.1.1. Roleta (<i>Roulette Wheel</i>).....	65
3.2.1.2. Rank (<i>Ranking Selection</i>).....	67
3.2.1.3. Torneio (<i>Tournament Selection</i>).....	67
3.2.1.4. Estado Fixo (<i>Steady State</i>).....	68
3.2.1.5. Elitismo (<i>Elitism</i>).....	69
3.2.1.6. Outros Métodos para o Operador de Seleção	69
3.2.2. Operação de Recombinação	70
3.2.2.1. Um Ponto de Corte (<i>1-Point Crossover</i>)	71
3.2.2.2. Dois Pontos de Corte (<i>2-Point Crossover</i>)	72
3.2.2.3. Uniforme (<i>Uniform Crossover</i>)	72
3.2.2.4. Parcialmente Mapeado (<i>PMX</i>).....	73
3.2.3 Operação de Mutação	75
3.3 “QUANDO UTILIZAR ALGORITMOS GENÉTICOS?” - APLICAÇÕES ..	76
3.4 O PROBLEMA DAS ESCOLHAS: MÉTODOS E PARÂMETROS	77
3.5 CONSIDERAÇÕES FINAIS	80

CAPÍTULO IV

4. PRINCIPAIS TÉCNICAS E CONCEITOS DE SUPORTE À PROPOSTA..	81
4.1 DESIGN INTELIGENTE.....	82
4.1.1 O Princípio do “ <i>No Free Lunch Theorem</i> ”	86
4.2 TEORIA DA EXTINÇÃO CÍCLICA DE RAUP-SEPKOSKI E VARIAÇÕES	89
4.3 TEORIA DOS CONJUNTOS E SISTEMAS NEBULOSOS – LÓGICA FUZZY	93
4.3.1 Conjuntos Nebulosos	94

4.3.2 Operações e Relações Sobre Conjuntos Nebulosos	96
4.3.2.1. Operações	96
4.3.2.2 Relações.....	100
4.3.3 Variáveis Lingüísticas.....	102
4.3.4 Lógica, Regras “Se-Então” e Inferência Nebulosa.....	106
4.3.5 Sistemas Nebulosos	108
4.3.5.1 Base de Regras Nebulosas.....	109
4.3.5.2 Motor de Inferência Nebuloso.....	110
4.3.5.3 Fuzzificador.....	114
4.3.5.4 Defuzzificador.....	115
4.4 ALGORITMOS GENÉTICOS COM INTERAÇÃO FENOTÍPICA: UMA VISÃO GERAL	116
4.4.1 Teoria dos Jogos e o Jogo Evolucionário Hawk-Dove.....	117
4.4.2 Codificação e Transmissão Genética do Comportamento	120
4.4.3 Estrutura dos Algoritmos Genéticos com Interação Fenotípica.....	120
4.5 CONSIDERAÇÕES FINAIS.....	122

CAPÍTULO V

5. GAIA: “GENETIC ALGORITHM INFERENCE AGENT”.....	123
5.1 VARIÁVEIS: O QUE MONITORAR E CONTROLAR?	126
5.1.2 Variáveis de Entrada.....	128
5.1.2.1 Valor Ótimo.....	128
5.1.2.2 Frequência do Melhor Resultado (ou Indivíduo).....	130
5.1.2.3 Percentual de Indivíduos Repetidos.....	132
5.1.3 Variáveis de Saída	133
5.1.3.1 Percentual de Mutação	133
5.1.3.2 Percentual de Recombinação.....	135
5.1.3.3 Percentual de Sobreviventes.....	136
5.2 BASE DE REGRAS E O MOTOR DE INFERÊNCIA FUZZY	138
5.3 ROLETA DA MORTE	146

5.4 PERIODICIDADE E OS DIFERENTES TIPOS DE GAIA	147
5.5 CONSIDERAÇÕES FINAIS	148

CAPÍTULO VI

6. TESTES E AVALIAÇÃO DA PROPOSTA	149
6.1 ESTRUTURA E ABORDAGEM DE TESTES DEFINIDA.....	150
6.1.1 O Problema Utilizado: PCVS-BR-26.....	153
6.1.1.1 O Problema do Caixeiro Viajante(<i>The Traveling Salesman Problem</i>).....	153
6.1.1.2 PCVS-BR-26.....	155
6.2 PRIMEIROS TESTES – PERÍODO 2005/2006.....	156
6.3 TESTES DEFINITIVOS – PERÍODO 2010/2011	160
6.3.1 Testes-Padrão.....	162
6.3.1.1 GAIA Permanente	165
6.3.1.2 GAIA Estocástico	167
6.3.1.3 GAIA Periódico.....	179
6.3.1.4 Hawk-Dove Roleta e Roleta Simples	187
6.3.2 Super-Testes.....	190
6.3.2.1 GAIA Permanente	192
6.3.2.2 GAIA Estocástico	194
6.3.2.3 GAIA Periódico.....	198
6.3.2.4 Hawk-Dove Roleta e Roleta Simples	202
6.4 QUADRO GERAL E ANÁLISE CONSOLIDADA DOS RESULTADOS	205
6.4.2 Medidas e Fatores de Análise e Comparação Interna	205
6.4.2.1 Variável “Melhores Resultados Alcançados”	206
6.4.2.2 Variáveis de Variabilidade: “Piores”, “Média” e “Dispersão”	210
6.4.2 Comparação com Outros Trabalhos.....	214
6.5 CONSIDERAÇÕES FINAIS	217

CAPÍTULO VII

7. CONCLUSÕES	218
REFERÊNCIAS BIBLIOGRÁFICAS	223
NORMAS ADOTADAS.....	227
ANEXO A.....	228
ANEXO B.....	231
ANEXO C.....	234
ANEXO D.....	244
ANEXO E.....	253
ANEXO F.....	257

LISTA DE FIGURAS

Figura 2.1 – O experimento de Louis Pasteur.	39
Figura 2.2 – A viagem do Beagle.	42
Figura 2.3 – Exemplo da girafa e a comparação: Darwin x Lamarck.	44
Figura 2.4 – Cariótipo humano do sexo masculino, 22 pares de autossomos mais um par de cromossomos alossomos (XY), indicando o sexo masculino.....	49
Figura 2.5 – Pares de cromossomos com genótipos e fenótipos correspondentes para determinação do tipo de vagens de ervilhas.	51
Figura 2.6 – As sete características das ervilhas-de-jardim, estudadas por Mendel.....	53
Figura 3.1 – Classes de Técnicas de Busca.	57
Figura 3.2 – Comparação: AG x processo biológico de evolução.	60
Figura 3.3 – Representação binária de cromossomos de soluções.	62
Figura 3.4 – Cromossomo codificado utilizando a representação por caminho.....	62
Figura 3.5 – Cromossomos codificados utilizando a representação por valores.....	63
Figura 3.6 – Roleta montada para população hipotética.	66
Figura 3.7 – <i>Crossover</i> com um ponto de corte.	71
Figura 3.8 – <i>Crossover</i> com dois pontos de corte.	72
Figura 3.9 – <i>Crossover</i> uniforme.	73
Figura 3.10 – Realizando o PMX.	74
Figura 3.11 – Operação de Mutação.	75
Figura 4.1 – Principais Formas de Representação das Funções de Pertinência.....	96
Figura 4.2 – Exemplo do Complemento de um Conjunto Nebuloso.....	97
Figura 4.3 – Exemplo de União de Conjuntos	97
Figura 4.4 – Exemplo de Interseção de Conjuntos	98
Figura 4.5 – Funções de Pertinência da Variável Velocidade.....	104

Figura 4.6 – Funções de Pertinência para as Restrições “Muito” e “Mais ou menos”.....	105
Figura 4.7 – Sistema Nebuloso de Mamdani com Fuzzificador e Defuzzificador.....	109
Figura 4.8. Estrutura de um AG segundo a abordagem híbrida de (BRITO et al., 2005).	122
Figura 5.1 – Nova Estrutura de Algoritmo Genético com o GAIA.....	125
Figura 5.2 – Funcionamento geral do GAIA.....	128
Figura 5.3 – Variável de Entrada – “Valor Ótimo”.....	130
Figura 5.4 – Variável de Entrada – “Frequência do Melhor Resultado”	132
Figura 5.5 – Variável de Entrada – “Percentual de Indivíduos Repetidos”	133
Figura 5.6 – Variável de Saída – “Percentual de Mutação”	134
Figura 5.7 – Variável de Saída – “Percentual de Recombinação”	136
Figura 5.8 – Variável de Saída – “Percentual de Sobreviventes”	138
Figura 6.1 – Estrutura Geral de Testes e Simulações GAIA.....	153

LISTA DE TABELAS

Tabela 3.1 – População hipotética com valores de <i>fitness</i> individual e proporcional à população.....	66
Tabela 4.1 – Diferentes Formas de Interpretação da Implicação Nebulosa.....	107
Tabela 4.2 – Exemplos de Motores de Inferência Nebulosos	114
Tabela 4.3 – Exemplo de Fuzzificadores	115
Tabela 4.4 – Exemplo de Defuzzificadores	116
Tabela 4.5 – Tabela de Pagamentos (<i>payoffs</i>) do jogo Hawk-Dove.....	119
Tabela 5.1 – Base de Regras Fuzzy - GAIA	139
Tabela 6.1 – Resultados Testes Período 2005/2006 (20 Melhores Resultados)	157
Tabela 6.2 – Testes-Padrão: Execuções e Resultados (Sintético)	162
Tabela 6.3 – Super-Testes: Execuções e Resultados (Sintético).....	191
Tabela 6.4 – Testes-Padrão: Variável “Melhores Resultados Alcançados”	208
Tabela 6.5 – Super-Testes: Variável “Melhores Resultados Alcançados”	209
Tabela 6.6 – Testes-Padrão: Variabilidade – “Piores”, “Média” e “Dispersão”	212
Tabela 6.7 – Testes-Padrão: Variabilidade – “Piores”, “Média” e “Dispersão”	213
Tabela B.1 – População inicial aleatória gerada pelo Framework Beagle XP v.4.5, utilizada em todos os experimentos e testes executados.	231
Tabela C.1 – Conjunto de Configurações de Todas as Simulações Realizadas.....	235
Tabela C.2 – Configurações 2010/2011 x Melhores Rotas Encontradas.....	238

LISTA DE GRÁFICOS

Gráfico 4.1 – Percentual de Extinção x Tempo (250 ma).....	92
Gráfico 6.1 – GAIA Periódico 100 (2005/2006): Melhor, Pior, Média x Geração.....	158
Gráfico 6.2 – GAIA Estocástico 25% (2005/2006): Melhor, Pior, Média x Geração.....	159
Gráfico 6.3 – GAIA Permanente: Melhores Valores x Geração	166
Gráfico 6.4 – GAIA Permanente: Média dos Valores x Geração	166
Gráfico 6.5 – GAIA Estocástico 0,2%: Média dos Valores x Geração	169
Gráfico 6.6 – GAIA Estocástico 0,2%: Melhores Valores x Geração.....	169
Gráfico 6.7 – GAIA Estocástico 01%: Média dos Valores x Geração	170
Gráfico 6.8 – GAIA Estocástico 01%: Melhores Valores x Geração.....	170
Gráfico 6.9 – GAIA Estocástico 05%: Média dos Valores x Geração	171
Gráfico 6.10 – GAIA Estocástico 05%: Melhores Valores x Geração.....	171
Gráfico 6.11 – GAIA Estocástico 25%: Média dos Valores x Geração	172
Gráfico 6.12 – GAIA Estocástico 25%: Melhores Valores x Geração.....	172
Gráfico 6.13 – GAIA Estocástico 50%: Média dos Valores x Geração	173
Gráfico 6.14 – GAIA Estocástico 50%: Melhores Valores x Geração.....	173
Gráfico 6.15 – GAIA Estocástico 75%: Média dos Valores x Geração	174
Gráfico 6.16 – GAIA Estocástico 75%: Melhores Valores x Geração.....	174
Gráfico 6.17 – GAIA Estocástico 95%: Média dos Valores x Geração	175
Gráfico 6.18 – GAIA Estocástico 95%: Melhores Valores x Geração.....	175
Gráfico 6.19 – GAIA Estocástico 98%: Média dos Valores x Geração	176
Gráfico 6.20 – GAIA Estocástico 98%: Melhores Valores x Geração.....	176
Gráfico 6.21 – GAIA Estocástico 99,8%: Média dos Valores x Geração	177
Gráfico 6.22 – GAIA Estocástico 99,8%: Melhores Valores x Geração.....	177
Gráfico 6.23 – GAIA Estocástico 99%: Média dos Valores x Geração	178

Gráfico 6.24 – GAIA Estocástico 99%: Melhores Valores x Geração.....	178
Gráfico 6.25 – GAIA Periódico 20: Média dos Valores x Geração	181
Gráfico 6.26 – GAIA Periódico 20: Melhores Valores x Geração.....	181
Gráfico 6.27 – GAIA Periódico 50: Média dos Valores x Geração	182
Gráfico 6.28 – GAIA Periódico 50: Melhores Valores x Geração.....	182
Gráfico 6.29 – GAIA Periódico 100: Média dos Valores x Geração	183
Gráfico 6.30 – GAIA Periódico 100: Melhores Valores x Geração.....	183
Gráfico 6.31 – GAIA Periódico 500: Média dos Valores x Geração	184
Gráfico 6.32 – GAIA Periódico 500: Melhores Valores x Geração.....	184
Gráfico 6.33 – GAIA Periódico 1000: Média dos Valores x Geração	185
Gráfico 6.34 – GAIA Periódico 1000: Melhores Valores x Geração.....	185
Gráfico 6.35 – GAIA Periódico 2500: Média dos Valores x Geração	186
Gráfico 6.36 – GAIA Periódico 2500: Melhores Valores x Geração.....	186
Gráfico 6.37 – Hawk-Dove Roleta: Média dos Valores x Geração	188
Gráfico 6.38 – Hawk-Dove Roleta: Melhores Valores x Geração	188
Gráfico 6.39 – Roleta Simples: Média dos Valores x Geração.....	189
Gráfico 6.40 – Roleta Simples: Melhores Valores x Geração	189
Gráfico 6.41 – Super-Teste GAIA Permanente: Média dos Valores x Geração.....	193
Gráfico 6.42 – Super-Teste GAIA Permanente: Melhores Valores x Geração.....	193
Gráfico 6.43 – Super-Teste GAIA Estocástico 25%: Média dos Valores x Geração.....	195
Gráfico 6.44 – Super-Teste GAIA Estocástico 25%: Melhores Valores x Geração	195
Gráfico 6.45 – Super-Teste GAIA Estocástico 50%: Média dos Valores x Geração.....	196
Gráfico 6.46 – Super-Teste GAIA Estocástico 50%: Melhores Valores x Geração	196
Gráfico 6.47 – Super-Teste GAIA Estocástico 98%: Média dos Valores x Geração.....	197
Gráfico 6.48 – Super-Teste GAIA Estocástico 98%: Melhores Valores x Geração	197

Gráfico 6.49 – Super-Teste GAIA Periódico 500: Média dos Valores x Geração.....	199
Gráfico 6.50 – Super-Teste GAIA Periódico 500: Melhores Valores x Geração	199
Gráfico 6.51 – Super-Teste GAIA Periódico 1.000: Média dos Valores x Geração.....	200
Gráfico 6.52 – Super-Teste GAIA Periódico 1.000: Melhores Valores x Geração	200
Gráfico 6.53 – Super-Teste GAIA Periódico 10.000: Média dos Valores x Geração.....	201
Gráfico 6.54 – Super-Teste GAIA Periódico 10.000: Melhores Valores x Geração	201
Gráfico 6.55 – Super-Teste Hawk-Dove Roleta: Média dos Valores x Geração.....	203
Gráfico 6.56 – Super-Teste Hawk-Dove Roleta: Melhores Valores x Geração	203
Gráfico 6.57 – Super-Teste Roleta Simples: Média dos Valores x Geração	204
Gráfico 6.58 – Super-Teste Roleta Simples: Melhores Valores x Geração.....	204
Gráfico 6.59 – Hawk-Dove Roleta (LEHRER, 2000): Melhor Valor e Média x Geração....	215
Gráfico 6.60 – Hawk-Dove Roleta (BRITO et. al, 2005): Média, Melhor e Pior x Geração	216
Gráfico 6.61– Hawk-Dove Roleta (TEIXEIRA, 2005): Média, Melhor e Pior x Geração ...	216

LISTA DE PUBLICAÇÕES

BRITO, F. H. ; TEIXEIRA, A. N. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. . **A Fuzzy Approach to Control Genetic Algorithm Parameters**. In: Electronic Journal of SADIO (Sociedad Argentina de Informática e Investigación Operativa), 2006, Argentina, Buenos Aires. (Publicação em Revista Internacional).

BRITO, F. H. ; TEIXEIRA, A. N. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. . **A Fuzzy Intelligent Controller for Genetic Algorithms' Parameters**. In: Second International Conference on Natural Computation (ICNC'06) and the Third International Conference on Fuzzy Systems and Knowledge Discovery (FSKD'06), 2006, Xi'an. Advances in Natural Computation - Second International Conference, ICNC 2006, Xi an, China, September 24-28, 2006. Proceedings, Part I, 2006. v. 4222. p. 633-642. (Capítulo de Livro).

BRITO, F. H. ; TEIXEIRA, A. N. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. . **A Fuzzy Approach to Control Genetic Algorithm Parameters**. In: ASAI 2006 - VIII Argentine Symposium on Artificial Intelligence, 2006, Mendoza. Proceedings of VIII Argentine Symposium on Artificial Intelligence, 2006.

BRITO, F. H. ; TEIXEIRA, A. N. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. . **Aplicando datamining com redes SOM em base de dados de execuções de algoritmos genéticos utilizando os novos métodos de seleção hawk-dove roleta e hawk-dove torneio**. In: CLEI 2006 - Conferencia Latinoamericana de Informática, 2006, Santiago. Proceedings of CLEI 2006 - Conferencia Latinoamericana de Informática, 2006.

BRITO, F. H. ; TEIXEIRA, A. N. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. . **Hawk-Dove Torneio: Um Novo Método de Seleção para os Algoritmos Genéticos**. In: SEPAI/CETIC, 2006, Belém, Pará.

BRITO, F. H. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. **A introdução da interação fenotípica em algoritmos genéticos através dos jogos evolucionários e da codificação e transmissão genética do comportamento** (Presented Conference Paper Style), presented at VII SBAI/II IEEE LARS: Simpósio Brasileiro de Automação Inteligente / II Latin-American IEEE Robotics Symposium, São Luís, MA, September 18-23 2005.

BRITO, Felipe Houat de ; TEIXEIRA, Artur Noura ; TEIXEIRA, Otávio Noura ; OLIVEIRA, Roberto Célio Limão de . **Darwin-Less Evolutionary Algorithm - Less Randomness, More Intelligence**. In: Worldcomp 2009 - World Congress in Computer Science, Computer Engineering and Applied Computing, 2009, Las Vegas, NV, USA. GEM 2009 - Genetic and Evolutionary Methods, 2009.

BRITO, Felipe Houat de ; TEIXEIRA, Artur Noura ; TEIXEIRA, Otávio Noura ; OLIVEIRA, Roberto Célio Limão de . **GAIA: A Fuzzy Intelligent Controller for Genetic Algorithms Parameters and Performance**. In: Worldcomp 2007 - World Congress in Computer Science, Computer Engineering and Applied Computing, 2007, Las Vegas. GEM 2007 - Genetic and Evolutionary Methods, 2007.

BRITO, Felipe Houat de ; TEIXEIRA, Artur Noura ; TEIXEIRA, Otávio Noura ; OLIVEIRA, Roberto Célio Limão de . **Hawk-Dove Tournament: A New Selection Technique for Genetic Algorithms Based on Evolutionary Game Theory**. In: Worldcomp 2009 - World Congress in Computer Science, Computer Engineering and Applied Computing, 2009, Las Vegas, NV, USA. GEM 2009 - Genetic and Evolutionary Methods, 2009.

TEIXEIRA, O. N. ; BRITO, F. H. ; TEIXEIRA, A. N. ; OLIVEIRA, R. C. L. . **Genetic Algorithms with Social Interaction Phase as Phenotype Characterization**. In: Second International Conference on Natural Computation (ICNC'06) and the Third International Conference on Fuzzy Systems and Knowledge Discovery (FSKD'06), 2006, Xi'an. Advances in Natural Computation and Data Mining. Xi'an, China: Xidian University Press, 2006. p. 48-58.

TEIXEIRA, O. N. ; BRITO, F. H. ; TEIXEIRA, A. N. ; OLIVEIRA, R. C. L. . **O Paradigma do Dilema do Prisioneiro Aplicado ao Método de Seleção Roleta dos Algoritmos Genéticos**. In: CLEI 2006 - Conferencia Latinoamericana de Informática, 2006, Santiago. Proceedings of CLEI 2006 - Conferencia Latinoamericana de Informática, 2006.

TEIXEIRA, O. N. ; BRITO, F. H. ; TEIXEIRA, A. N. ; OLIVEIRA, R. C. L. . **Game Theory as a New Paradigm for Phenotype Characterization of Genetic Algorithms**. In: GECCO 2006 - Genetic and Evolutionary Computation Conference, 2006, Seattle-WA. Proceedings of GECCO 2006 - Genetic and Evolutionary Computation Conference. New York : ACM Press, 2006. v. 2. p. 1431-1432.

TEIXEIRA, Otávio Noura ; BRITO, Felipe Houat de ; TEIXEIRA, Artur Noura ; OLIVEIRA, Roberto Célio Limão de . **Genetic Algorithms with Social Interaction: A New Step for Phenotype Characterization Thru Game Theory**. In: Worldcomp 2007 - World Congress in Computer Science, Computer Engineering and Applied Computing, 2007, Las Vegas. GEM 2007 - Genetic and Evolutionary Methods, 2007.

TEIXEIRA, Otávio Noura ; LOBATO, Walter Avelino da Luz ; YASOJIMA, Carlos Takeshi Kudo ; BRITO, Felipe Houat de ; TEIXEIRA, Artur Noura ; OLIVEIRA, Roberto Célio Limão de . **Fuzzy Social Interaction Genetic Algorithm**. In: GECCO 2010 - Genetic and Evolutionary Computation Conference, 2010, Portland, Oregon. Proceedings of the 2010 Genetic and Evolutionary Computation Conference. Alpha, New Jersey: Sheridan Printing Company, Inc, 2010. p. 2113-2114.

RESUMO

Há muitos anos, técnicas de Computação Evolucionária vem sendo aplicadas com sucesso na solução dos mais variados tipos de problemas de otimização. Na constante procura pelo ótimo global e por uma melhor exploração da superfície de busca, as escolhas para ajustar estes métodos podem ser exponencialmente complexas e requerem uma grande quantidade de intervenção humana. Estes modelos tradicionais darwinianos apóiam-se fortemente em aleatoriedade e escolhas heurísticas que se mantêm fixas durante toda a execução, sem que acompanhem a variabilidade dos indivíduos e as eventuais mudanças necessárias. Dadas estas questões, o trabalho introduz a combinação de aspectos da Teoria do Design Inteligente a uma abordagem híbrida de algoritmo evolucionário, através da implementação de um agente inteligente o qual, utilizando lógica fuzzy, monitora e controla dinamicamente a população e seis parâmetros definidos de uma dada execução, ajustando-os para cada situação encontrada durante a busca. Na avaliação das proposições foi construído um protótipo sobre a implementação de um algoritmo genético para o problema do caixeiro viajante simétrico aplicado ao cenário de distância por estradas entre as capitais brasileiras, o que permitiu realizar 580 testes, simulações e comparações entre diferentes configurações apresentadas e resultados de outras técnicas. A intervenção inteligente entrega resultados que, com sucesso em muitos aspectos, superam as implementações tradicionais e abrem um vasto espaço para novas pesquisas e estudos nos aqui chamados: “Algoritmos Evolucionários Híbridos Auto-Adaptáveis”, ou mesmo, “Algoritmos Evolucionários Não-Darwinianos”.

Palavras-chave: Computação Evolucionária, Design Inteligente, Algoritmos Genéticos Híbridos, Lógica Fuzzy, GAIA.

ABSTRACT

For many years Evolutionary Techniques have been successfully applied in several computational optimization problems. For the matter of obtaining “best results” and a wide exploration of the search surface, the choices for tuning those methods can be exponentially complex and require a large human intervention. Those traditional Darwinian models rely only on randomness and heuristic choices, which are kept fix during the entire algorithm execution without following the individuals’ variability and any eventual changes needed. For that matter, the present work introduces the combination of Intelligent Design Theory aspects with a hybrid approach of an evolutionary algorithm, through the implementation of a Fuzzy Intelligent Designer agent, which dynamically monitors and regulates six chosen algorithms’ parameters, adjusting their values for any given situation. In order to testify these proposals, a Brazilian’s capitals symmetric TSP problem genetic algorithm prototype was also built, allowing the execution of 580 tests, plus several simulations and comparisons between many proposed set of configurations, regular techniques and other recent works. The Intelligent Design intervention deliveries over-expectation results, opening a wide new space for research of the: “Darwin-less Evolutionary Algorithms” or “The Hybrid Self-Adaptable Evolutionary Algorithms”.

Keywords: Evolutionary Computation, Intelligent Design, Hybrid Genetic Algorithms, Fuzzy Logic, GAIA.

CAPÍTULO I

Deus disse: "Faça-se a luz!" E a luz foi feita.

Gênesis 1:3.

1. INTRODUÇÃO

O homem, com sua sede de conhecimento e seu desejo profundo de dominar e compreender os eventos que regem o universo e a sua própria existência, inspira-se por muitas vezes em suas observações da natureza na construção de ferramentas e conhecimento.

Esse conhecimento, que é construído de forma cumulativa, não tem limites, e dá origem a ferramentas e processos não antes pensados para a solução dos inúmeros problemas do cotidiano. Até, segundo alguns autores como KURWEIL (1999), chegar ao ponto em que as próprias ferramentas passarão a evoluir por si mesmas.

Deste anseio, surge a área da computação conhecida como Computação Natural, que representa um grande conjunto de técnicas que podem ser subdivididas em três principais áreas: ferramentas computacionais (algoritmos) desenvolvidas com inspiração na natureza; modelos de simulação de sistemas naturais; e, por último, os novos paradigmas de computação que envolvem a utilização de mecanismos naturais para armazenagem, processamento de informação e solução de problemas (EBERHART et al., 1996).

Seguindo este raciocínio mimético dos processos da natureza, HOLLAND (1975) apud MITCHELL (1999), propôs, entre as décadas de 60 e 70, a estruturação de uma técnica de otimização baseada nos conceitos de genética e na Teoria da Evolução de Charles Darwin, chamada de Algoritmos Genéticos (AGs) e que, posteriormente, deu origem a todo um ramo de conhecimento, hoje, conhecido como Computação Evolucionária ou Evolutiva.

Desde então, Algoritmos Genéticos, têm sido aplicados na solução de diversos problemas, sendo até, para alguns autores, considerado “*the ultimate problem solver*”, numa tradução livre, “o solucionador definitivo de problemas” (DEMBSKI, 2007). Exatamente desta forma, HOLLAND (1995) apud HAUPT et al. (2004), imaginou-os como um método eficiente, de fácil uso e aplicável a um vasto conjunto de problemas. Entretanto, um fato que é pouco citado na literatura tradicional dos AGs, é que para que sejam aplicados de forma eficiente, eles requerem um grande conhecimento para adequá-los a uma aplicação particular.

Essa expertise é requerida porque seus usuários geralmente não tem idéia de como se decidir entre o uso dos inúmeros parâmetros, operadores, codificações e métodos envolvidos. Em síntese, um grande paradoxo. Já que, ao invés de serem uma técnica eficiente e prática como originalmente foram pensados, o uso de AGs acaba por se tornar uma busca interminável e complexa por ajustes de parâmetros e configurações.

Na busca pela solução desse problema, como exposto em mais detalhes nos próximos capítulos, uma grande quantidade de proposições são feitas, desde a definição de um conjunto de valores padrão aplicável a qualquer problema, passando pela codificação dos parâmetros dentro das próprias populações, até mesmo técnicas que, em paralelo a execução principal, realizam um meta algoritmo genético para otimizar o conjunto de parâmetros (HAUPT et al., 2004).

No geral, estas soluções nem diminuem a complexidade do uso da técnica, muito menos se adéquam a qualquer tipo de problema.

Neste ponto, o presente trabalho, formula uma nova proposta de Algoritmo Genético, buscando inspiração em um conjunto recente de abordagens (DEMBSKI, 2007, 2004; DANIKEN, 2005; MOURÃO, 2002; RAUP e SEPKOSKI, 1984), as quais tentam explicar algumas lacunas do darwinismo, como o alto estágio que a evolução atingiu apenas “aleatoriamente”, as constantes e cíclicas catástrofes que ocorrem na Terra que, apesar de

severas, após algum tempo, trouxeram um novo caminho para a evolução da vida, os vestígios arqueológicos e antropológicos desconexos, as evidências da chamada assinatura “inteligente” no DNA humano, entre outros fatores.

Esta proposta é baseada substancialmente na construção de um agente inteligente - aqui chamado de GAIA (*Genetic Algorithm Inference Agent*) - que, anexado a um Algoritmo Genético tradicional, monitora sua população e o estágio da evolução, regulando de forma automática parâmetros e garantindo a variabilidade dos indivíduos (soluções). Desta maneira, evita que a execução fique presa em ótimos locais e busca constantemente pelo ótimo global, sem que, para tal, seja necessária qualquer intervenção do pesquisador.

O agente inteligente mimetiza a idéia do *designer* da Teoria do Design Inteligente (DEMBSKI, 2007, 2004), utilizando um motor de inferência baseado em lógica fuzzy, sustentado pelo conhecimento paramétrico apresentado na literatura e pela vasta experiência do autor e seu grupo de pesquisas na resolução de problemas de otimização por meio do uso de técnicas de Computação Evolucionária.

O ateste da eficácia da proposta é feito por meio da elaboração de 48 configurações diferentes para o agente, avaliadas em 580 testes (vide detalhes, Anexo C), executados sobre um problema comum na literatura, porém de alta complexidade e sem o ótimo global conhecido, uma aplicação do Problema do Caixeiro Viajante Simétrico às distâncias entre as capitais brasileiras, o PCVS-BR-26 (LEHRER, 2000).

1.1 OBJETIVOS DA DISSERTAÇÃO

Nos tópicos que seguem, são descritos os objetivos principais e secundários deste trabalho, bem como a delimitação de seu escopo dentro da utilização da Teoria do Design Inteligente e variações à estrutura de um Algoritmo Genético.

1.1.1 Objetivo Geral

O objetivo geral deste trabalho é, utilizando aspectos da Teoria do Design Inteligente (DEMBSKI, 2004, 2007) e da Teoria da Extinção Cíclica de Raup-Sepkoski (RAUP e SEPKOSKI, 1984; MOURÃO, 2002), propor a construção de uma nova estrutura de algoritmo evolucionário em que um agente inteligente, percebendo o estágio de evolução em que se encontra a população, intervenha ajustando automaticamente parâmetros e eliminando indivíduos. Desta forma, visa contornar dois grandes problemas desta técnica: o primeiro, a grande dificuldade na escolha de parâmetros ótimos ao início de cada execução e, o segundo, evitar a todo custo a convergência prematura e “prisão” em ótimos locais (HAUPT et al., 2004).

Além disso, a nova estrutura proposta, deve-se tornar então, o passo inicial na construção de uma técnica de otimização que exima completamente os pesquisadores de diferentes áreas de conhecimento da extensa compreensão das técnicas de computação evolucionária e do grande tempo gasto, geralmente, via “tentativa e erro”, na busca de parâmetros ótimos para cada tipo de problema.

1.1.2 Objetivos Específicos

1. Compreender os mecanismos tradicionais inerentes ao processo evolutivo, da concepção do universo aos primeiros seres vivos, em especial a formalização dada pelo Neodarwinismo, a evolução biológica com os conceitos de Seleção Natural, a Luta pela Existência e a Variabilidade das Espécies. Além disso, confrontá-los com outras Teorias mais recentes, antagônicas e até mesmo polêmicas, como o Design Inteligente e suas diversas correntes;

2. Compreender dentro do contexto da Computação Evolucionária, os Algoritmos Genéticos como um todo, abordando histórico, estrutura básica, principais métodos de implementação e aplicabilidade, assim como, seus dois principais problemas: convergência prematura e escolha de parâmetros;

3. Realizar estudo sobre a proposta da existência de um “designer” inteligente por meio da Teoria do Design Inteligente, bem como, suas possíveis formas de interferência sobre a origem e manutenção da vida, vislumbrando sua correlação e aplicabilidade nas estruturas de Algoritmos Genéticos;

4. Compreender aspectos da Lógica Difusa ou Nebulosa (Lógica Fuzzy) para construção de um motor de inferência capaz de atuar como o “designer” em algoritmos genéticos, percebendo seus aspectos e intervindo quando necessário;

5. Desenvolver programaticamente um protótipo de fácil utilização para depuração das propostas, tendo sempre registro em base de dados para uma melhor análise dos dados e posteriores comparações.

1.2 METODOLOGIA E PESQUISA ADOTADA

Para a consecução deste trabalho foi realizada uma ampla pesquisa bibliográfica nas principais áreas abordadas: Algoritmos Genéticos e o problema da variação e escolha de parâmetros, o Design Inteligente, a Teoria da Extinção Cíclica e Lógica Fuzzy. Esta pesquisa inicial se destinou, principalmente, a obtenção de uma base sólida nos assuntos abordados, possibilitando a definição da estrutura do agente inteligente, suas variáveis de monitoramento e controle, seu motor de inferência e também a forma como seria realizada sua intervenção na população do AG.

Após a proposição dos métodos e estratégias adotados, iniciou-se a busca por um problema de ampla utilidade e facilidade de implementação, ideal para testar as propostas. Neste ponto, foi escolhido o problema do Caixeiro Viajante Simétrico para as capitais do Brasil (PCVS-BR-26), de forma semelhante à implementada por LEHRER (2000) e também em diversos outros trabalhos e propostas do grupo de pesquisa o qual pertence o autor - que conta com um número significativo de publicações nacionais e internacionais, vide Anexos E e F-, para que, desta forma, pudesse estabelecer certa comparação da abordagem e dos resultados obtidos.

Na continuidade do projeto, ocorreu a escolha da plataforma de testes e dos aspectos de implementação da proposta para fins de simulação. Optou-se por evoluir o já existente framework de testes, Beagle XP (BRITO, 2004), que, conforme descrito no Anexo D, possuía uma boa estrutura de AGs construída, necessitando apenas da construção e adaptação do agente fuzzy.

Com o protótipo plenamente desenvolvido, foram realizados testes comparativos para comprovar a eficácia da proposição e analisar de forma criteriosa os impactos das intervenções dentro da população durante as execuções.

Por fim, tabulou-se todos os resultados em gráficos e estatísticas que possibilitaram comprovar o êxito das proposições do trabalho.

1.3 ESTRUTURA GERAL DA DISSERTAÇÃO

No segundo capítulo, são apresentados os aspectos biológicos que fundamentam a maior parte das teorias aplicadas no escopo deste trabalho. No primeiro momento, é apresentado o surgimento do Universo e apontado o processo de evolução que culminou na criação das primeiras formas biológicas conhecidas. Depois, são apresentadas as primeiras

tentativas de explicar a evolução dos seres vivos até chegar a Seleção Natural proposta por Charles Darwin. Na seqüência, são apresentados os conceitos básicos inerentes ao Neodarwinismo e a genética, que partem do histórico, passando pelas Leis de Mendel e, culminam no modelo moderno do DNA de James Watson e Francis Crick.

No terceiro capítulo, são abordados os Algoritmos Genéticos, contextualizados dentro da Computação Evolucionária. É apresentada uma conceituação inicial e histórico, sendo tratados na seqüência aspectos da implementação de AGs. Nesta parte, são descritos os principais Operadores Genéticos usados, bem como os principais métodos utilizados na sua implementação. O capítulo é encerrado abordando quando devem ser aplicados os AGs, seguidos de alguns exemplos de sucesso, e, por último, tratando de apresentar a principal problemática deste trabalho: a dificuldade na escolha dos inúmeros parâmetros e métodos disponíveis em AGs para chegar à execução “ideal” e eficaz na solução de problemas.

No quarto capítulo, são apresentados os principais conceitos e técnicas de suporte à proposta do trabalho. Após uma breve introdução conectora com as devidas justificativas das abordagens, o capítulo se diversifica. Inicia com uma seção destinada a tratar de forma resumida dos conceitos da Lógica Difusa (Lógica Fuzzy), alicerce do motor da proposta GAIA. Depois, trata, também de forma sintética, da abordagem de AGs Híbridos com Interação Social utilizada em todos os testes deste trabalho. Na seqüência, apresenta os aspectos conceituais inspiradores da proposta: a polêmica e controversa teoria do Design Inteligente, e a Teoria da Extinção Cíclica de Raup-Sepkoski.

No quinto capítulo, é apresentada a proposta principal do trabalho na forma da construção de um agente inteligente que, periodicamente, monitora a evolução de uma população de um AG, intervindo e controlando-a quando necessário. Para tal, primeiramente, é descrita toda a conexão com a fundamentação teórica de um agente (ou *designer*) e a justificativa do porquê de um controle da evolução dentro de um AG. Em seguida, são

apresentadas as variáveis de monitoramento e controle escolhidas, sua modelagem fuzzy, a base de regras elaborada e os diferentes tipos de ação do agente. Ao final, são propostas, baseado na teoria da extinção cíclica, as conseqüências e a forma da intervenção na população pelo agente inteligente.

No sexto capítulo, são descritos todos os testes realizados e a metodologia adotada. Primeiramente, é descrito o problema do Caixeiro Viajante de forma geral e sua variação adotada. Posteriormente, é discutida a estratégia de testes, com todas as configurações de parâmetros e métodos utilizados. Ao final, são apresentados e analisados os resultados obtidos, realizando a comparação com as configurações de execuções geradas e também com outros trabalhos.

No sétimo capítulo, as considerações finais e principais conclusões são apresentadas, além de serem sugeridos novos trabalhos e pesquisas na área abordada.

CAPÍTULO II

O homem, com suas nobres qualidades, ainda carrega no corpo a marca indelével de sua origem modesta.

Charles Darwin

2. ASPECTOS BIOLÓGICOS FUNDAMENTAIS: ORIGEM DA VIDA, EVOLUÇÃO E GENÉTICA

Nos itens que seguem, estão descritos alguns dos conceitos biológicos fundamentais à compreensão de Algoritmos Genéticos.

Inicia-se a abordagem com início do universo até a evolução e origem da vida, passando pelas primeiras teorias que buscaram explicar este evento, tratando da Teoria da Evolução de Charles Darwin e o paradigma Neodarwinista, onde também são apresentados os principais conceitos inerentes à genética, em especial as proposições feitas por Mendel.

Em um capítulo posterior será apresentada outra teoria sobre a evolução e origem da vida, o Design Inteligente, que em muitos aspectos se opõe às demais expostas nesta seção, mais comumente aceitas no meio científico. Entretanto, agrega valor à compreensão, já que foi a base inspiratória para a proposta desta nova técnica.

2.1 ORIGEM E EVOLUÇÃO DO UNIVERSO

Como surgiu o universo? O que existia antes do universo? Quando ele deixará de existir? Estas questões perseguem o homem desde que ele começou a pensar e se deu conta de sua existência.

No princípio as únicas explicações que se tinham eram míticas e normalmente associadas às religiões e culturas locais. À medida que a ciência evoluiu, explicações racionais baseadas em astronomia e física começaram a surgir, mas mesmo estas ainda estão longe de serem a verdade absoluta.

A abordagem descrita a seguir é a do “Big-Bang”, proposta por Alexandre Friedmann (1888-1925) em 1922, que atualmente apresenta inúmeras contribuições de outros cientistas da área. Foi escolhida com base na conveniência dos efeitos causados frente à Teoria da Evolução. Outras teorias que fogem ao escopo desse trabalho, as quais também possuem forte embasamento nos novos teoremas da física podem ser encontradas em GOTT (2002).

Em seu livro “O Livro de Ouro do Universo”, MOURÃO (2002), descreve o “Big-Bang” e afirma que graças aos físicos de partículas, toda a história de criação do universo, até o estado que hoje conhecemos, foi um processo que pode ser contado num espaço de tempo muito curto (10^{-35} s), e concluído em aproximadamente três minutos.

Há 15 bilhões de anos, uma enorme quantidade de energia se encontrava localizada numa esfera de diâmetro inferior a 1 cm. Neste estado não existiam as dimensões de espaço e tempo de forma separada, estas estavam intrinsecamente ligadas numa forma semelhante a uma esponja (GOTT, 2002).

Por algum motivo esta singularidade entrou em colapso e iniciou um processo súbito de expansão. A energia espalhada entrou em processo de resfriamento de forma uniforme, dando origem a sucessivas transformações que culminaram em matéria e antimatéria em proporções praticamente semelhantes.

No entanto, no instante entre 10^{-43} s e 10^{-35} s, a antimatéria “desapareceu” (MOURÃO, 2002), graças a uma pequena desvantagem na quantidade em relação à matéria existente. Esta pequena quantidade remanescente de matéria, no instante 10^{-30} s, deu origem a todas as partículas elementares (prótons, nêutrons, elétrons, etc.), as quais compõe o universo, e toda a

matéria aniquilada deu origem aos fótons (partículas de luz) existentes, que entregam a transparência a grande parte do universo.

Depois de 1 segundo, alguns nêutrons continuaram a se transformar em prótons, fato que justifica a razão encontrada atualmente de três prótons para um nêutron. Dez segundos depois, as primeiras reações de fusão ocorrem e surgem os primeiros átomos de Hélio (He). Aos três minutos tudo já estava constituído.

Após uma longa pausa de 1 milhão de anos, a “sopa” de matéria se resfria ainda mais e permite a fusão de átomos em núcleos mais pesados, dando início a era química. Antes de 1 bilhão de anos de idade, surge a força mais fraca de interação, a gravidade, e os aglomerados estelares - galáxias- têm sua origem. Mais 4 bilhões de anos e surge, circundando em torno de uma estrela da Via-Láctea, o até então planetóide, Terra.

Somente há 3,4 bilhões de anos, é que, como citado por KURZWEIL (1999), “da química nasceu a biologia”, quando as primeiras cadeias carbônicas dão origem ao primeiro ser vivo da Terra, um procarionte anaeróbico - organismo unicelular, o qual não necessitava de oxigênio. A evolução é continuada até que somente a menos de 1 milhão de anos, surge o *Homo Sapiens sapiens*.

Toda esta descrição temporal, foi exposta com intuito de apresentar que, à medida que, o caos (desordem, aleatoriedade) foi se instaurando no universo, os eventos mais significativos passaram a ocorrer com intervalos de tempo cada vez maiores, de nano segundos a bilhões de anos.

Esta desaceleração é justificada pela **Segunda Lei da Termodinâmica**, também conhecida como Lei do Aumento da Entropia, cujo enunciado afirma que, num sistema fechado, onde não existe troca de energia (interação) com forças externas, a desordem tende a aumentar, ou seja, leva-se a entropia (KURZWEIL, 1999). Com isso, o surgimento natural da

inteligência, a qual está diretamente associada à ordem, torna-se inviável, já que comportamento inteligente é o oposto de aleatoriedade.

De acordo com KURZWEIL (1999), esta contradição é explicada, devido à insignificância estabelecida pela ordem que levou ao surgimento de vida inteligente, frente ao estado caótico em que se encontra o universo.

O autor ainda propõe a “*Law of Accelerating Returns*”, ao afirmar que, quanto mais evoluído um sistema, menor é o tempo entre o acontecimento de eventos significativos. Exatamente a explicação que foi dada à desaceleração do acontecimento de grandes eventos, com o aumento do caos do universo. Explicação também adotada para prever a evolução da tecnologia e o seu aumento exponencial, até o ponto em que a tecnologia irá evoluir por si só.

A principal intenção do exposto, foi demonstrar que o processo Evolutivo, transcende os limites biológicos, sendo mais macro do que se usualmente acredita. Aplicar a evolução a processos da física e à partículas, é algo ainda mais polêmico do que a aplicação aos seres vivos e, ainda mais complexo de ser explicado.

Na sequência do capítulo, parte-se para a abordagem biológica da evolução, as principais proposições para a origem das espécies, seguidos dos estudos revolucionários realizados por Charles Darwin, Gregor Mendel e outros.

2.2 EVOLUÇÃO E A “ORIGEM DA ESPECIES”

Como citado anteriormente, há 3,4 bilhões de anos, o primeiro organismo surgiu na Terra, o procarionte anaeróbico. O surgimento deste primeiro ser vivo - dentre as teorias mais defendidas - se deu graças ao ambiente propício oferecido pelo planeta, em que constantes cargas elétricas associadas a muitas moléculas de hidrogênio (H_2), amônia (NH_3), gás carbônico (CO_2), metano (CH_4) e vapor d'água (H_2O) criaram as primeiras moléculas

orgânicas, que deram origem a certas cadeias de proteínas que agrupadas, num processo de milhões de anos, geraram o primeiro ser (GOWDAK et al., 1991).

A partir deste ser, a evolução se encarregou de continuamente otimizar a estrutura biológica dos seres vivos, até que estes pudessem se adaptar ao meio. Entretanto, para conseguir esta façanha de busca por soluções, a natureza criou a estrutura de DNA que, segundo TEIXEIRA (2005), “é a base genética que guiou e registrou o processo de evolução, de forma a não permitir encontrar soluções para problemas já solucionados anteriormente”.

O DNA possibilitou a combinação de soluções mais adequadas e o seu rápido incremento. Logo, seres unicelulares deram origem a seres pluricelulares, que de simples plantas, tornaram-se peixes, até conseguirem sair da água e, tornarem-se anfíbios, parcialmente dependentes de água como seu *habitat*, até o surgimento dos répteis que, num intervalo de tempo mais curto, dominaram a terra até a extinção dos dinossauros (700 milhões de anos), e a ascensão dos mamíferos, culminando na supremacia humana, há menos de 1 milhão de anos atrás (AMABIS et al., 1994).

Antes de se chegar ao exposto sobre o surgimento da vida hoje largamente aceito, muitas proposições absurdas míticas e até mesmo de cunho “científico” foram apresentadas. Além disso, para explicar a variabilidade dos seres vivos, outras inúmeras teorias foram propostas, até culminar na defendida por Charles Darwin, em sua célebre obra, “A Origem das Espécies”. Estes fatos serão abordados no item a seguir.

2.2.1 Histórico do Desenvolvimento de Teorias

Até meados do século XIX, acreditava-se que os seres vivos eram gerados de forma espontânea a partir de matéria não-viva, teoria que ficou conhecida como Abiogênese. Esta

idéia era tão defendida, que até mesmo pensadores de renome como Descartes (1596-1650) e Isaac Newton (1642-1727) acreditavam nela (AMABIS et al., 1994).

Um famoso médico da época, chegou até a criar uma receita para a criação da vida, como citado em AMABIS et al. (1994): “(...) coloca-se, num canto sossegado e pouco iluminado, camisas sujas. Sobre elas espalham-se grãos de trigo, e o resultado será que em 21 dias, surgirão ratos (...)”.

Estas hipóteses foram derrubadas principalmente pelos experimentos realizados por Francesco Redi (1626-1697), Lazzaro Spalanzanni (1729-1799) e Louis Pasteur (1822-1895). Pasteur, que realizou o experimento mais significativo, já realizava trabalhos com microorganismos e, provou que estes eram os causadores do “surgimento de vida”, em restos de alimentos (GOWDAK et al., 1991).

Utilizando um frasco com um tubo de oxigenação em formato de pescoço de cisne, Pasteur conseguiu impedir a entrada de microorganismos no seu “caldo nutritivo” por quatro meses, quando então, quebrou a entrada especial do frasco e, em poucos dias, o frasco já estava repleto de microorganismos, como visto na figura 2.1. Desta forma, a Abiogênese cede lugar definitivamente a Biogênese, que afirma que a vida só pode surgir de vida, ou seja, pela reprodução.

Figura 2.1 – O experimento de Louis Pasteur.



Fonte: <http://www.personal.psu.edu/faculty/j/e/jel5/micro/art.htm>

Explicado em parte o surgimento da vida, questões ainda mais complexas surgiram, tais como: Como explicar a variabilidade das espécies? Seriam todas as espécies estáticas? Então, como justificar os fósseis de animais extintos que se assemelham aos atuais?

As explicações religiosas sempre foram tidas como verdade absoluta e, serviram, por muito tempo, para refrear inúmeras teorias que se aproximavam daquilo que hoje se acredita. Predominavam na época, principalmente as correntes Criacionista e Fixista. A primeira apontando para a criação divina de toda a vida, e a segunda, para o número fixo de espécies de forma estática e fixa (LOPES, 1998).

Segundo FAVARETTO et al. (1999), Jean-Baptiste Lamarck (1744-1829), em seu livro “Filosofia Zoológica”, foi o primeiro a afirmar que todas as espécies, inclusive a humana, descendiam de organismos preexistentes, adotando o termo “evolução biológica”. Sua obra foi mal recebida na época, mas suas idéias foram de grande contribuição para as que viriam suceder-lhe.

Lamarck sustentava que os seres vivos apresentavam um “impulso interior”, o qual os permitia se adaptar ao meio e às condições impostas por ele, por meio de alguma necessidade. Desta forma, eles usariam um determinado membro mais que outros, que no caso se desenvolveria mais, em detrimento dos outros que se atrofiariam. Esta hipótese, ficou conhecida como “Lei do Uso e Desuso” (GOWDAK et al., 1991).

Além disso, Lamarck, afirmava que todas as mudanças ocorridas numa geração, graças a “lei” anterior, seriam transmitidas aos descendentes, no que ficou conhecido como “Lei da Transmissão de Características Adquiridas”.

Para FAVARETTO et al. (1999), dois aspectos propostos por Lamarck merecem destaque: o primeiro, é que a lei do uso e desuso é válida para alguns órgãos do corpo, como os músculos estriados, que se desenvolvem com o uso contínuo; o segundo, é que a segunda

lei não é válida de forma alguma, uma vez que, nenhuma das características adquiridas durante a vida é transmitida aos descendentes, os quais recebem somente o material genético contido nos gametas.

Lamarck impulsionou os principais naturalistas da época a exporem suas idéias sobre Evolução e a diversificação das espécies. Dentre esses naturalistas, encontravam-se Charles Darwin e Alfred Russel Wallace (1823 - 1913), que no ano de 1858, percebendo a semelhança entre suas teorias, publicaram um artigo em conjunto proclamando a chamada **“seleção natural”**. Reafirmada no ano seguinte com a publicação de Darwin, “A Origem das Espécies”, assuntos descritos com mais detalhes a seguir.

2.2.2 Seleção Natural de Charles Darwin

No ano de 1831, Charles Darwin integra-se a uma expedição da coroa inglesa, que duraria cinco anos, para uma viagem ao redor do mundo, cujos principais objetivos eram a análise geológica dos continentes.

Durante as inúmeras paradas do Beagle (nome da embarcação ocupada por ele), Darwin, que já possui um grande interesse na biologia, fica fascinado pela grande variedade de espécies observadas, passando a se questionar sobre qual fator estas espécies seriam levadas a tão grande variedade (TEIXEIRA, 2005). A viagem do Beagle encontra-se na figura 2.2, a seguir.

Darwin supôs que os fatores alimento e espaço controlariam o tamanho das populações, uma vez que, durante seus estudos, embora houvesse motivos para as populações selvagens apresentarem um grande número de indivíduos, ele não havia encontrado nenhuma população nessas condições, pois estas sempre se encontravam em equilíbrio.

Sendo assim, Darwin concluiu que, nas circunstâncias expostas, haveria uma grande **luta pela sobrevivência**, onde apenas os mais aptos sobreviveriam e deixariam mais descendentes de sua variabilidade. Esta seleção feita pelas condições impostas pelo ambiente foi chamada de **Seleção Natural** (GOWDAK et al., 1991).

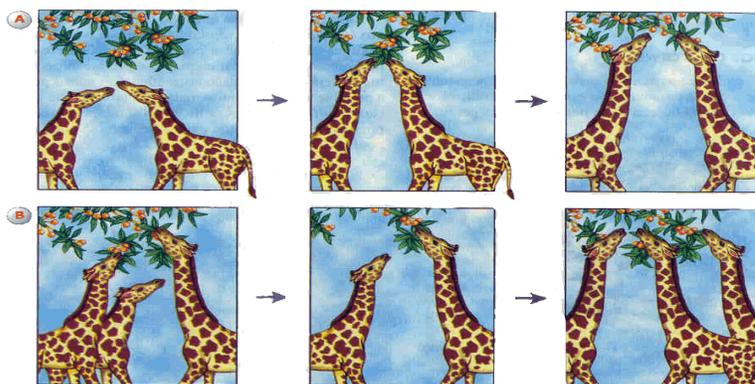
A **Seleção Natural** de Darwin é bem explicada através da reflexão apresentada no início do capítulo quarto de sua célebre obra:

Como o esforço pela existência (...) pode agir no que diz respeito à variação? Pode o princípio de seleção, que como já podemos ver é tão potente nas mãos do homem [neste trecho, ele está se referindo a criação de raças de cavalos e cães de forma artificial] ser aplicado na natureza? Eu penso que veremos que ele age exatamente de forma semelhante e ainda mais eficaz. Deixe pensar que o interminável número de estranhas peculiaridades de nossas produções domésticas, e num grau menor, as provenientes da natureza, variam; e o quão forte é a tendência da hereditariedade. (...) Deixe pensar o quão infinitamente complexo são as relações mutuas entre todos os organismos vivos, nas condições físicas de vida. Pode então, desse modo, ser improvável que, se as variações esperadas pelo homem ocorreram, as variações úteis de alguma forma a cada ser vivo na grande complexa batalha pela vida, pode algumas vezes ocorrer no curso de milhares de gerações? Se isso realmente ocorrer, podemos nós duvidar que indivíduos possuindo alguma vantagem, ainda que superficial, sobre os outros, teriam a melhor chance de sobreviver e perpetuar a sua espécie? E em contrapartida, nós poderemos afirmar que qualquer variação prejudicial em algum grau seria rigorosamente eliminada. Essa preservação de variações favoráveis e a rejeição de variações desfavoráveis é o que eu chamo de Seleção Natural. (Traduzido de DARWIN, 1999).

Um exemplo clássico utilizado para explicar a teoria de Darwin, é o apresentado em FAVARETTO et al. (1999), onde é confrontada a Seleção Natural de Darwin com a Lei do Uso e Desuso de Lamarck, para explicar o pescoço longo das girafas.

Na figura 2.3, os três quadros superiores (seção A), utilizam a explicação lamarckista em que, à medida que os alimentos passaram a se localizar em árvores mais altas, as girafas passaram a esticar os seus pescoços, fazendo com que estes crescessem ao longo do tempo. Na seção B, a explicação darwinista é aplicada, afirmando que entre as girafas, existia a diversidade que possuía pescoços mais longos, com o aumento das árvores que continham os alimentos, as girafas de pescoço menor foram sendo eliminadas e a variação de pescoço maior predominou perpetuando-se através da reprodução (seleção natural).

Figura 2.3 – Exemplo da girafa e a comparação: Darwin x Lamarck.



Fonte: (FAVARETTO, 1999)

Apesar da grande mudança de paradigma e progresso trazida pela teoria de Darwin, este recebeu sérias críticas em alguns pontos em que sua teoria não conseguia explicar ou não tinha como ser comprovada. Uma das principais críticas foi a falta de exemplos concretos e atuais de evolução, uma vez que este é um processo extremamente lento.

Outra crítica foi a falta de explicação convincente para as causas da diversidade intra-específica e os fatores que levavam ao afastamento de espécies que descendiam de um mesmo ancestral. Nesta época os estudos da genética ainda não existiam e Darwin, de acordo com

LOPES (1998), buscou explicação para a hereditariedade no que ele chamou de Teoria da Pangênese.

Segundo essa teoria, todos os órgãos e componentes do corpo produzem suas próprias cópias em miniatura infinitamente pequenas, denominadas gêmulas, que são carregadas pela corrente sanguínea até os gametas para serem utilizadas durante a fecundação.

Com o acréscimo de outras teorias e descobertas de diversos pesquisadores posteriores a Darwin, muitas questões da seleção natural foram explicadas. Esta agregação deu origem ao que hoje é chamada de **Teoria Sintética da Evolução** ou **Neodarwinismo**, exposta no item a seguir.

2.3 NEODARWINISMO: EVOLUÇÃO E GENÉTICA

O Neodarwinismo, como citado anteriormente, é o estado atual da teoria da evolução, onde são mantidas as principais idéias inerentes à seleção, somadas às contribuições propostas pela Genética Clássica e pelas mutações.

A diversidade dos organismos, como será exposto a seguir, é garantida pela variabilidade do seu material genético (genes). Desta forma, a seleção natural estabelece uma taxa diferencial de reprodução, ou seja, os indivíduos com um conjunto de genes mais adaptado ao meio têm mais chances de sobreviver e gerar descendentes, transmitindo assim, seu material genético (LOPES, 1998).

Pelo princípio de Hardy-Weinberg, apresentado em (FAVARETTO et al., 1999), em uma população em equilíbrio a frequência dos alelos (valores que um gene pode assumir), é constante; porém as populações não são estáticas e evoluem constantemente. Este fato, se deve às mutações, à pressão da seleção natural e outros fatores.

Desta forma, a Teoria Sintética da Evolução pode ser definida como o conjunto de alterações ocorridas nas frequências dos alelos com o decorrer do tempo, onde a mutação é responsável pela inclusão de novos alelos no *pool* gênico e, a seleção natural age controlando a frequência daqueles que são mais adaptados às condições do meio (LOPES, 1998).

2.3.1 Histórico

Para alcançar a definição proposta pelo Neodarwinismo descrita acima, além de outras contribuições, foram de fundamental importância as de quatro cientistas: Gregor Mendel (1822-1884), Hugo de Vries (1848-1935), James Watson (1928-) e Francis Crick (1916-2004).

O monge Johann Gregor Mendel realizando experimentações com ervilhas nos jardins do mosteiro de Brün, na República Tcheca, buscou uma explicação matemática para o aparecimento de certas características nas ervilhas-de-jardim (*Pisum sativum*).

Mendel acabou descobrindo uma certa frequência com que algumas características se manifestavam, postulando que a transmissão de caracteres hereditários era feita através de partículas ou fatores que se encontravam nos gametas, que deu origem a suas duas leis (LOPES, 1998).

Mais tarde estes fatores foram chamados de genes e, hoje, as leis de Mendel são tidas como base nas noções atuais de transmissão de características hereditárias, além de terem derrubado qualquer outra teoria que tentara explicar estes eventos.

Os trabalhos de Mendel foram publicados pela primeira vez em 1866, e praticamente passaram despercebidos pelo mundo. Até que em 1900, o holandês Hugo de Vries, em suas pesquisas sobre mutações redescobre os trabalhos de Mendel e os relança na comunidade científica. De Vries proferiu significativas contribuições ao Neodarwinismo, em suas explicações sobre a mutação e a forma com que ocorria (LOPES, 1998).

O ponto culminante para o sucesso da genética atual, veio na década de 50, quando James Watson e Francis Crick, apresentaram o seu modelo helicoidal da composição química do ácido desoxirribonucléico, também conhecido como DNA (sigla em inglês). O DNA é a matéria que constitui os genes e o conhecimento de sua estrutura possibilitou um nível de manipulação genética que tem revolucionado a ciência nos últimos 45 anos (AMABIS et al., 1994).

Um dos impactos diretos dessas descobertas é o Projeto Genoma Humano (PGH), que constitui uma força-tarefa de pesquisadores, centros de pesquisa e governos do mundo todo, para mapear todos os genes humanos e suas funcionalidades, o que irá possibilitar uma melhor compreensão de doenças, aptidões, relógio biológico, inteligência e tudo o que for proveniente dos humanos que não é herdado das interações sociais (GENOME, 2004).

Para que se tenha noção do funcionamento da transmissão das características de forma hereditária é necessária à compreensão de alguns conceitos de genética expostos no item que segue.

2.3.2 Conceitos Fundamentais de Genética

Os conceitos expostos a seguir são de comum definição entre os autores de biologia e foram extraídos de FAVARETTO et al. (1999), GOWDAK et al. (1991), AMABIS et al. (1994) e LOPES (1998).

Cada ser vivo, no interior de suas células, apresenta um número constante de pares de cromossomos ($2n$), de acordo com a sua espécie.

Cromossomos são estruturas que se replicam durante as divisões celulares e, que carregam toda a carga genética de uma dada espécie. Seu número é constante, porque na

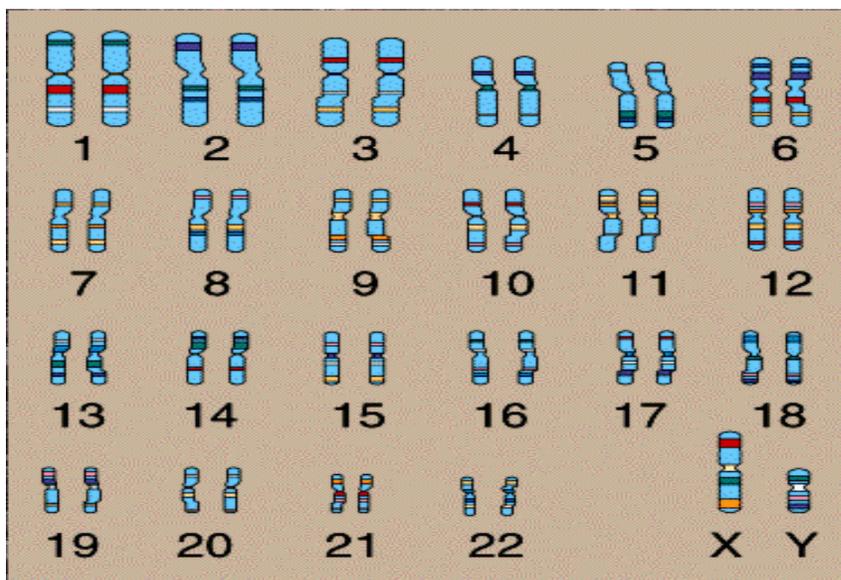
formação dos gametas esse número é reduzido à metade e, durante a fecundação, restabelecido.

As células do corpo, também chamadas de somáticas **diplóides** ($2n$), apresentam pares de cromossomos “parecidos” chamados de **cromossomos homólogos**. Durante a formação dos gametas, como dito anteriormente, o número de cromossomos é reduzido à metade, o que faz dessas células serem **haplóides** (n).

A figura 2.3 apresenta um **cariótipo** humano, extraído de uma célula diplóide com 23 pares de cromossomos. Um cariótipo de uma espécie é o conjunto de todos os cromossomos encontrados nos núcleos de suas células somáticas, com suas características de número, tamanho e forma.

A maioria das espécies apresenta um par de cromossomos sexuais (**alossomos**), que apresentam constituição diferente nos machos e nas fêmeas, os demais pares, são formados por cromossomos **autossomos** e são idênticos em ambos os sexos.

Figura 2.4 – Cariótipo humano do sexo masculino, 22 pares de autossomos mais um par de cromossomos alossomos (XY), indicando o sexo masculino.



Fonte: <http://www.accessexcellence.org/AE/AEPC/NIH/gene03.html>.

Cada par de cromossomos homólogos contém genes para as mesmas características. Por gene, entende-se cada segmento de DNA capaz de transcrever sua mensagem em uma molécula de uma estrutura chamada de **RNA** (ácido ribonucléico, da sigla em inglês), fundamental nas etapas de divisão celular para a replicação do DNA.

Desta forma, um cromossomo possui uma seqüência linear de genes e o local ocupado pelo gene é chamado de **locus gênico** (no plural *loci*). Os valores ou formas de gene que atuam para uma mesma característica são chamados de **alelos**, ocupando o mesmo *locus* em cromossomos homólogos.

O conjunto de todos os alelos que um indivíduo herda de seus genitores é chamado de **genótipo**. As características internas ou externas manifestadas por este genótipo recebem a denominação de fenótipo.

Certas características, como o grupo sanguíneo (exposto a seguir neste trabalho), dependem exclusivamente do genótipo, outras sofrem também influência direta do meio ambiente. Um exemplo é a cor da pele, que apesar de possuir características definidas na carga genética (genótipo) pode sofrer alterações de acordo com o nível de exposição à luz solar. Sendo assim, o conceito de fenótipo consiste em ambiente mais genótipo.

Quando uma característica possui uma ou mais variedades, podemos concluir que estas são representadas por diferentes alelos, que ocupam o mesmo *locus* gênico. Neste ponto, de acordo com a frequência de aparição das características, surgem dois conceitos: **alelo dominante** e **alelo recessivo**.

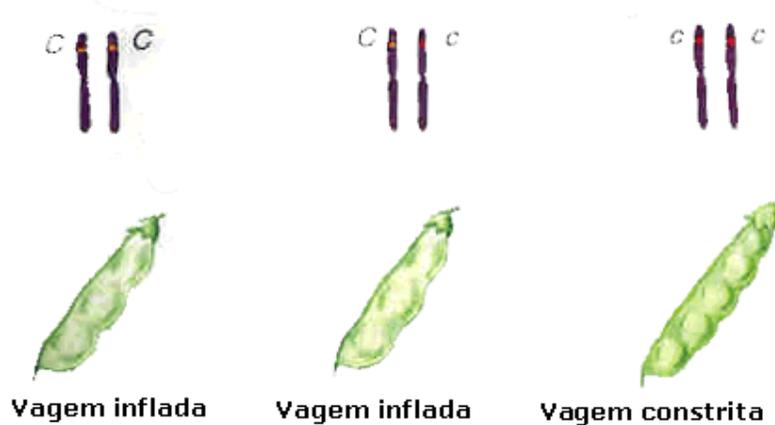
O alelo dominante é aquele que pode manifestar-se tanto em dose dupla quanto em dose simples, já os alelos recessivos, necessitam aparecer em dose dupla para que sejam manifestados. Em geral, os alelos são representados por uma letra, que corresponde, usualmente, à inicial do fenótipo recessivo. Quando um indivíduo apresenta um genótipo com dois alelos iguais, diz-se que ele é **homozigoto**, com dois alelos diferentes ele é chamado de **heterozigoto**.

Os conceitos apresentados podem ser bem visualizados no exemplo exposto por FAVARETTO et al. (1999), que descreve o aspecto das vagens em ervilhas. No *locus* que controla o aspecto das vagens, um alelo determina vagens infladas, enquanto o outro determina vagens constrictas.

Na figura 2.5, são apresentados os respectivos genótipos e fenótipos possíveis para a situação descrita. A representação do alelo é a letra “c”, devido ao fenótipo vagens constricta ser manifestado pelo alelo recessivo, estando a relação de dominância expressa por $C > c$. O primeiro indivíduo apresentado é homozigoto dominante e, apresenta vagem inflada; o segundo é heterozigoto, então a característica imposta pelo alelo dominante é que se

manifesta, sendo assim, ele apresenta vagem inflada; o terceiro indivíduo é heterozigoto recessivo, portanto apresenta vagem constricta.

Figura 2.5 – Pares de cromossomos com genótipos e fenótipos correspondentes para determinação do tipo de vagens de ervilhas.



Fonte: (FAVARETTO et al., 1999)

Embora cada célula contenha o DNA completo do indivíduo, elas se utilizam dos genes de forma seletiva. Alguns genes possibilitam às células produzir proteínas específicas para funções básicas. Outros genes, no entanto, estão ativos a maior parte do tempo. Alguns genes realizam algum papel durante a fase embrionária e são desativados para sempre. De uma forma geral, nas células os genes são ativados sempre que uma dada função é requerida.

Apresentados estes conceitos sobre genética, é possível expor as Leis de Mendel descritas no item que segue.

2.3.3 Leis de Mendel

Gregor Mendel em seu estudo analisando sete características de variações das ervilhas-de-jardim, através do controle do cruzamento e uso de métodos estatísticos, postulou a **Lei da Segregação dos Fatores** ou **Lei da Pureza dos Gametas**, hoje chamada de **Primeira Lei de Mendel**.

Seus testes consistiam em obter uma geração pura de indivíduos para uma determinada característica. Depois, cruzando indivíduos puros de características diferentes, ele obtinha na geração seguinte, indivíduos, os quais ele denominou de “híbridos”. Nestes, ele observou que uma característica sempre permanecia oculta e só se manifestava, na outra geração, quando da fecundação destes híbridos, que sempre geravam uma proporção fenotípica de 3:1.

Mendel chamou o fator que se manifestava em indivíduos híbridos de dominante e, o fator que pulava uma geração resultado do cruzamento de indivíduos puros diferentes, de fator recessivo.

Sendo assim, a Primeira Lei de Mendel, de acordo com LOPES (1998), afirma que: “cada característica é determinada por um par de fatores, que se segregam durante a formação dos gametas, os quais sempre são puros”.

Os conceitos de “fatores”, “indivíduos puros” e “indivíduos híbridos” são, o que hoje se chamam, respectivamente: alelos homozigotos e heterozigotos. A figura 2.6 apresenta uma tabela com os fenótipos das ervilhas que levaram Mendel a formulação de sua primeira Lei.

Figura 2.6 – As sete características das ervilhas-de-jardim, estudadas por Mendel.

Característica	Dominante	Recessiva	Característica	Dominante	Recessiva
Forma da semente	 Lisa	 Enrugada	Posição da flor	 Axial	 Terminal
Cor da semente	 Amarela	 Verde		Altura do caule	 Alta
Cor da flor	 Violeta	 Branca			
Forma da vagem	 Inflada	 Constrita			
Cor da vagem	 Verde	 Amarela	©1998 Sinauer Associates, Inc.		

Fonte (adaptado): <http://www.people.virginia.edu/~rjh9u/monotraits.html>.

Mendel prosseguiu seus estudos e passou a analisar duas características ao mesmo tempo. Testou a combinação de todas as sete características e, chegou à conclusão que estas eram independentes. Além disso, a proporção fenotípica sempre se aproximava da proporção 9:3:3:1, onde nove se referia a ocorrência de dois fenótipos dominantes no mesmo indivíduo, o 3 à de um fenótipo dominante qualquer e 1 à de dois fenótipos recessivos (GOWDAK ET AL., 1991).

Assim era anunciada a **Segunda Lei de Mendel** ou Diíbridismo, Triíbridismo ou Poliíbridismo: “dois ou mais fatores segregam-se de forma independente durante a formação dos gametas” (FAVARETTO et al., 1999).

Ao escolher estas sete características das ervilhas, pode-se dizer que Mendel “teve sorte”, pois os alelos que as representavam se comportavam da mesma maneira, respeitando o monoíbridismo e possibilitando a formulação de sua lei.

2.4 CONSIDERAÇÕES FINAIS

Apresentados e compreendidos os conceitos biológicos de seleção natural, genética e genética do comportamento, é possível demonstrar a sua aplicabilidade computacional na resolução de problemas, o que servirá para compreensão da problemática tratada neste trabalho. Além disso, cria-se suporte para a compreensão de outras teorias suplementares às lacunas deixadas pelos evolucionistas puros e simples, que conforme será demonstrado, têm impacto significativo na eficácia da aplicabilidade dos Algoritmos Genéticos tradicionais.

No capítulo a seguir serão apresentados os algoritmos genéticos, suas principais formas, aplicações e problemas.

CAPÍTULO III

“Eu não tenho muito respeito por talento. Talento é genético. É o que você faz com ele que conta”.

Martin Ritt

3. ALGORITMOS GENÉTICOS E UMA VISÃO GERAL DE COMPUTAÇÃO EVOLUCIONÁRIA

A Computação Evolucionária é um dos paradigmas computacionais da Computação Natural, que se fundamenta nos conceitos apresentados no capítulo anterior - em especial o da Seleção Natural, da Luta pela Sobrevivência e do Neodarwinismo- , para otimizar funções e, de uma forma geral, resolver problemas.

EBERHART et al. (1996) define este paradigma da seguinte forma:

Computação Evolucionária consiste no paradigma de otimização do aprendizado de máquinas e classificação de paradigmas que são baseados nos mecanismos da evolução, como genética biológica e seleção natural. (Traduzido do original)

A Computação Evolucionária compreende quatro modelos principais, ou agrupamentos de técnicas: Algoritmos Genéticos (AG ou GA), Programação Evolucionária (EP), Programação Genética (GP) e Estratégias Evolutivas (ES) (MITCHELL, 1996). Sendo destes, o AG o mais destacado e o objeto de estudo e proposições deste trabalho.

Os AGs são algoritmos de busca bastante eficientes que implementam de uma maneira primitiva alguns dos processos da evolução natural. São idéias para otimizações de funções e classificação de padrões, pois com sua estrutura conseguem realizar uma busca eficiente num espaço muito grande de soluções (EBERHART et al., 1996).

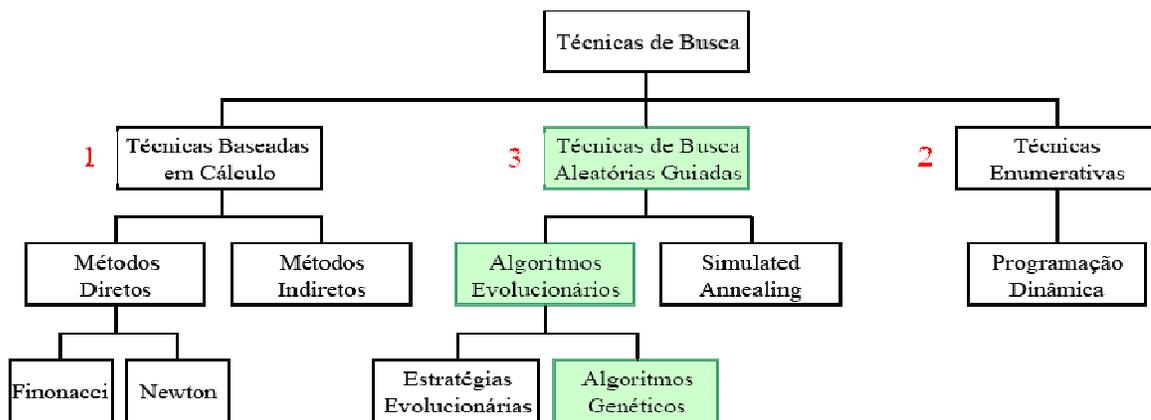
KOZA (1990) apud LEHRER (2000) define os AGs como algoritmos matemáticos altamente paralelos, que utilizando os operadores genéticos, como mutação e reprodução sexual (recombinação), transforma populações de objetos matemáticos (indivíduos) em novas populações, de forma proporcional a adaptabilidade (*fitness*) a função desejada, remetendo ao princípio exposto no capítulo anterior da “sobrevivência do mais apto”.

De forma a representar o escopo de aplicação de AGs, MITCHELL (1996), define três tipos principais de buscas realizadas na ciência da computação:

- **Busca por dados armazenados:** consiste em obter de forma eficiente informações armazenadas numa base de dados de um computador.
- **Busca por caminhos para atingir metas:** neste caso, a busca é por um conjunto de ações que permitam mover de um estado inicial para um estado “meta” final. Este é um dos tipos mais comuns de busca realizadas em IA, onde os caminhos são geralmente representados por estruturas de árvores e grafos.
- **Busca por Soluções:** é uma classe de busca, ainda mais geral, do que a anterior. Consiste em obter uma solução para um problema num espaço de busca muito grande e, muitas vezes, indeterminado. Este é o tipo de busca em que os AGs são utilizados.

Conforme apresentado em TEIXEIRA (2005), os AGs pertencem a classe de **Técnicas de Buscas Aleatórias Guiadas**, que tem obtido grande sucesso face à incompletude e complexidade dos demais tipos de busca (Técnicas Baseadas em Cálculo e Técnicas Enumerativas). A figura 3.1 fornece uma visão do agrupamento das principais técnicas de busca.

Figura 3.1 – Classes de Técnicas de Busca.



Fonte: (TEIXEIRA, 2005).

GOLDBERG (1989) apud LEHRER (2000), define as principais diferenças entre os AGs para os métodos tradicionais de busca:

- AGs, geralmente, trabalham com codificação de parâmetros, ao invés de trabalhar com os parâmetros em si;
- AGs utilizam regras de transição probabilística, no lugar de regras determinísticas;
- AGs não utilizam conhecimentos auxiliares ou derivados¹ na busca pela solução, estando esta condicionada a informação fornecida pela função objetivo;
- AGs trabalham com populações de pontos, em vez de um único ponto;

Além das diferenças já expostas, AGs apresentam certas vantagens que têm estimulado bastante sua utilização, como apresentado em HAUPT et al. (2004):

¹ Segundo DEMBSKI (2007), este é exatamente o principal ponto de falha dos Algoritmos Genéticos e a maior crítica à teoria darwinista, dada tão alta complexidade e especificidade dos seres vivos. É sobre esta lacuna que o presente trabalho agrega um agente inteligente à estrutura convencional dos AGs, evitando a alta convergência e garantindo a constante busca pelo ótimo global do problema.

- Otimizam funções com parâmetros discretos ou contínuos;
- Não requerem informação derivada;
- Simultaneamente explora um amplo espaço da superfície de busca (n soluções de uma população);
- Pode lidar com uma grande quantidade de parâmetros;
- É bem ajustado para computadores com processamento paralelo;
- Otimizam parâmetros com superfícies de busca extremamente complexas;
- Podem sair de um mínimo local e continuar a busca pelo global;
- Oferecem uma lista de inúmeras soluções, não estando restrito a uma única;
- Trabalham com dados numéricos, dados experimentais ou funções analíticas;

Entretanto, assim como as demais técnicas, AGs também apresentam certas desvantagens, já que não estão adaptados a todos os tipos de problemas e, geralmente sua utilização remete a um alto custo computacional e conseqüente longo tempo de processamento. Além da alta dependência na função de avaliação de *fitness*, cuja própria modelagem, por definição, pode apresentar um espaço de busca exponencialmente maior e mais complexo do que o do próprio problema em questão (DEMBSKI, 2007).

Ao final deste capítulo, outros problemas e desvantagens dos Algoritmos Genéticos serão também apresentados.

No item que segue, está explicitada a origem e desenvolvimento das técnicas de Computação Evolucionária, em especial dos Algoritmos Genéticos, para que se possa entender o porquê de certas estruturas adotadas hoje.

3.1 HISTÓRICO E CONTEXTUALIZAÇÃO

Antes de chegar ao paradigma da Computação Evolucionária, diversos fatores biológicos foram descobertos e propostos, como explicados no capítulo anterior.

De posse da compreensão destes trabalhos e observando o sucesso obtido da própria natureza, engenheiros e cientistas da computação, nas décadas de 50 e 60, em trabalhos paralelos passaram a tentar simular os processos biológicos de evolução para otimizar funções e solucionar problemas (MITCHELL, 1999). Estes trabalhos resultaram no paradigma da Computação Evolucionária e de suas subáreas.

Algoritmos Genéticos, especificamente, foram desenvolvidos por John Holland e seus estudantes durante as décadas de 60 e 70. Foram popularizados pelas mãos de um dos seus estudantes, David Goldberg, na década de 80, com a solução de um problema relacionado à implantação de gasodutos (HAUPT et al., 2004).

Segundo MITCHELL (1999), o que tornou a técnica desenvolvida por Holland eficaz e diferente das outras propostas, foi o seu forte estudo na função de adaptabilidade e em como a transcrever para simulação em computadores sem focar na solução de um problema específico.

HOLLAND (1975) apud MITCHELL (1999) define o AG como um método de se movimentar de uma população a outra, utilizando um tipo de “Seleção Natural” em conjunto com outros operadores genéticos, tais como mutação, onde um gene é substituído por um alelo de forma aleatória, e recombinação, em que os cromossomos escolhidos trocam pedaços gerando novas soluções. Este processo é repetido até que a solução ou conjunto de soluções satisfatórias sejam obtidos.

Apesar de, atualmente, o conceito de AGs ter absorvido outras áreas e se aproximado mais ainda da simulação biológica, a estrutura proposta por Holland ainda é válida e defendida por muitos autores, sendo chamada de **AG canônico**.

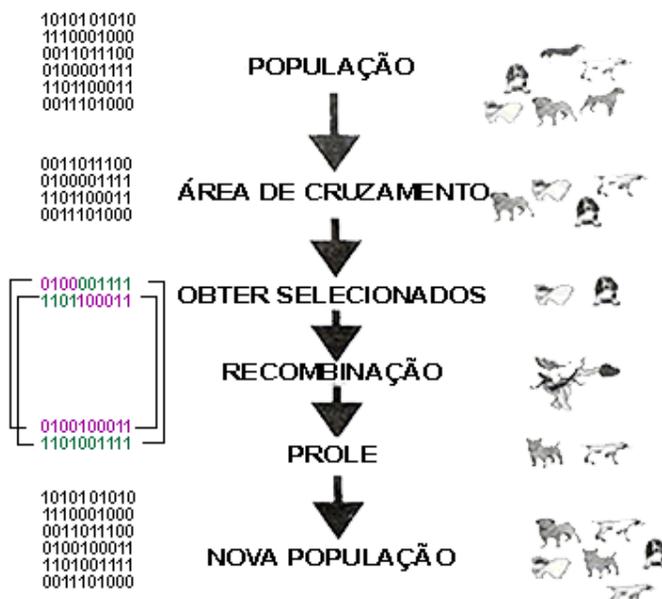
No item a seguir, é apresentada a implementação de AGs e as principais técnicas utilizadas para definição dos operadores genéticos.

3.2 IMPLEMENTAÇÃO DE ALGORITMOS GENÉTICOS

O principal propósito de algoritmos genéticos é, a partir de uma população de soluções candidatas (cromossomos ou indivíduos) para um determinado tipo de problema, evoluir os indivíduos até encontrar a solução ou conjunto de soluções esperadas (MITCHELL, 1999).

A figura 3.2 expõe uma comparação entre o processo evolutivo biológico e o utilizado em AGs.

Figura 3.2 – Comparação: AG x processo biológico de evolução.



Fonte (adaptado): (HAUPT et al., 2004)

De uma forma bem geral, de acordo com HAUPT et al. (2004), a estrutura de um AG consiste em :

1. Definir os parâmetros a serem utilizados: **formato do cromossomo** (normalmente um vetor de x posições) e codificação (valores de alelos);

2. Definir a função de avaliação de adaptabilidade ou **função de fitness**;
3. Gerar uma população de n indivíduos solução, geralmente inicializados com valores aleatórios do **alfabeto de alelos** (todos os parâmetros possíveis);
4. Avaliar os indivíduos, definindo um valor de adaptabilidade (*fitness*) para cada um;
5. Utilizar o **Operador de Seleção** para escolher os indivíduos que deverão gerar novos descendentes, de acordo com a adaptabilidade;
6. Gerar novos descendentes, através do **Operador de Recombinação**;
7. Realizar a **Operação de Mutação** sobre os descendentes gerados;
8. Avaliar convergência da solução;
9. **Repetir** os passos 4 a 8 (**Gerações**) até atingir uma solução;

De acordo com TEIXEIRA (2005), a definição da estrutura de codificação das soluções é um dos aspectos mais importantes na construção de um AG, sendo que um cromossomo solução é codificado geralmente de uma das quatro formas: **Codificação Binária, Código de Permutação, Código de Valores, e Árvores**.

A Codificação Binária é um dos métodos mais utilizados, já que os primeiros estudos realizados por Holland se utilizaram dessas estruturas, tendo diversos métodos para os operadores genéticos bem definidos (MITCHELL, 1999).

É definida por uma cadeia de caracteres (*strings*), que podem ser guardada num vetor, onde seus pedaços (posições) representam os genes responsáveis pela solução. Esta codificação apresenta limitações já que não serve para representar quaisquer tipos de problemas, além de se encontrar, muitas vezes, bastante complexa e de difícil compreensão dado determinado tipo de problema. (HAUPT et al., 2004).

Na figura 3.3, encontra-se uma representação de dois cromossomos binários de soluções candidatas. Ressaltando que, o tamanho das cadeias e suas posições é algo arbitrário e inerente ao problema.

Figura 3.3 – Representação binária de cromossomos de soluções.

Cromossomo A	101100101100101011100101
Cromossomo B	111111100000110000011111

Fonte: (TEIXEIRA, 2005).

Outra representação de cromossomos bastante utilizada é o Código de Permutação ou Representação por Caminho. Como apresentado em TEIXEIRA (2005), este tipo de codificação é utilizada em problemas de sequencialidade, arrumação e disposição, tais como o Problema do Caixeiro Viajante (PCV), que é apresentado neste trabalho.

A figura 3.4 apresenta um cromossomo codificado utilizando a representação por caminho. Este, foi gerado pelo protótipo **Beagle XP**, construído e evoluído pelo autor deste trabalho, para suportar as propostas apresentadas, bem como outras, cuja finalidade é encontrar a solução para um Problema do Caixeiro Viajante Simétrico aplicado entre as capitais do Brasil por meio da utilização de estradas federais disponíveis.

Figura 3.4 – Cromossomo codificado utilizando a representação por caminho.

CROMOSSOMOPROBLEMA
04-01-05-18-06-15-03-07-08-13-17-09-14-20-25-24-23-00-11-12-21-16-19-22-10-02-04

Fonte: Beagle XP.

Uma das codificações que tem passado a ser bastante utilizada, apesar da não garantia probabilística de encontrar o valor ótimo, é a Codificação de Valores, onde são utilizados vetores de parâmetros contínuos (inteiros ou reais), caracteres comuns, palavras e outros valores (HAUPT et. al., 2004).

Esta representação é de fácil compreensão já que pode se aproximar bastante do problema modelado. Entretanto, devem ser observados que os operadores genéticos precisam ser ajustados para atender a este formato. A figura 3.5 apresenta cromossomos codificados com esta estrutura.

Figura 3.5 – Cromossomos codificados utilizando a representação por valores.

Cromossomo A	1.2324 5.3243 0.4556 2.3293 2.4545
Cromossomo B	ABDJEIFJDHDIERJFDLDFLFEGTAZ
Cromossomo C	(voltar), (voltar), (direita), (a frente), (esquerda)

Fonte: (TEIXEIRA, 2005).

Uma representação menos utilizada é o esquema de codificação em árvores que, de acordo com LEHRER (2000), consiste num grafo unidirecional sem ciclos fechados, utilizado bastante na programação genética.

Ainda que existam estas formalizações sobre a modelagem do cromossomo, deve-se ressaltar que estas não são tidas como verdades absolutas e, diversas vezes, precisam ser combinadas ou alteradas para que o problema possa ser representado de forma correta e funcional dentro de uma estrutura de AG.

Em alguns casos até matrizes multidimensionais podem ser utilizadas na modelagem do problema, porém deve-se tomar cuidado para não fazer com que a complexidade de avaliação utilize muitos recursos computacionais.

Com o formato dos cromossomos definidos, é necessário compor a função de avaliação ou função de fitness. Esta função está intrinsecamente relacionada à modelagem do problema e, não possuindo uma fórmula exata, pode ser extremamente complexa de ser desenvolvida (EBERHART et al., 1996).

Uma forma mais comum de avaliar indivíduos consiste em varrer os cromossomos e, de acordo com seus alelos ou combinação de alelos, atribuir pontuações negativas ou

positivas para os resultados esperados. Desta forma, encontra-se ao final, um valor de adaptabilidade próprio do cromossomo.

Segundo LEHRER (2000), a função de adaptabilidade é que transforma o desempenho de cada cromossomo numa distribuição de chances para que seu genótipo seja transmitido através das gerações pelo operador de recombinação.

Depois de modelado o formato dos indivíduos, definidos os alelos que irão compor o alfabeto e criada a função de avaliação, devem ser escolhidos os métodos para os operadores de AG: Seleção, Recombinação e Mutação.

Estes operadores e suas técnicas estão expostos nos itens a seguir.

3.2.1. Operação de Seleção

O operador de seleção é responsável por escolher os genótipos que serão utilizados na geração de novos indivíduos por meio do operador de recombinação, produzindo descendentes mais bem adaptados (MITCHELL, 1999).

Segundo WHITLEY (1994) apud LEHRER (2000), a operação de seleção irá selecionar cromossomos da população, utilizando na maioria das vezes o valor de *fitness*, possibilitando aos indivíduos mais aptos gerar maior número de descendentes, ou seja, passar pelo operador de recombinação mais vezes, fazendo um paralelo direto com a Seleção Natural de Charles Darwin.

Os métodos de seleção não podem ser muito fortes, ou seja, escolher sempre os melhores ou o melhor indivíduo de uma geração, pois isso leva a uma rápida convergência da população para um resultado que pode ser um ótimo local, mas não o ótimo global (HAUPT et al., 2004).

Existem certos parâmetros e técnicas para que a pressão seletiva mantenha um valor sempre equilibrado. A seguir, estão descritos alguns dos métodos mais utilizados para este operador.

3.2.1.1. Roleta (*Roulette Wheel*)

O método da Roleta é um dos métodos mais populares e tradicionais em AGs e é baseado na chamada Seleção Proporcional ao *Fitness*.

Segundo MITCHELL (1999), métodos de seleção proporcionais ao *fitness* existem desde a proposição original de AGs por Holland, onde a média esperada para a participação de um indivíduo em recombinações é igual a proporção do valor do *fitness* do indivíduo em relação ao *fitness* total da população.

A roleta proposta por GOLDBERG (1989) funciona atribuindo a cada indivíduo uma fatia na roleta de tamanho proporcional ao seu valor de adaptabilidade. Criada a roleta, esta é girada n vezes, até que uma seção correspondente a um indivíduo _{x} esteja sobre o “apontador” da roleta. Este, então, é selecionado e levado para a “área de cruzamento” (*mating pool*). O processo é repetido até que o número de pais seja igual ao esperado (EBERHART et al., 1996).

Na tabela 3.1, encontram-se os indivíduos de uma população qualquer, apresentando um identificador numérico, genótipo, valor de adaptabilidade e o respectivo percentual de adaptabilidade dentro da população.

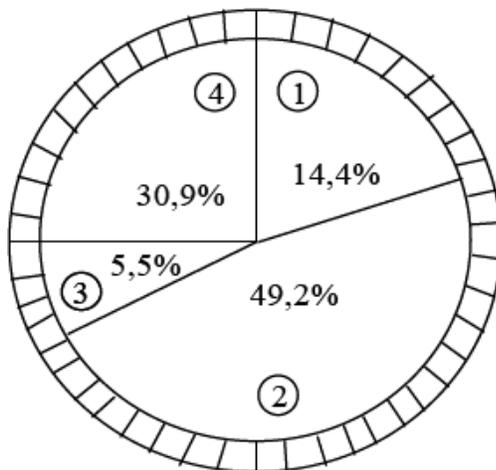
Tabela 3.1 – População hipotética com valores de *fitness* individual e proporcional à população.

Nº.	Indivíduos	(<i>Fitness</i>)	% do Total
1	01101	169	14,4
2	11000	576	49,2
3	01000	64	5,5
4	10011	361	30,9
Total		1170	100,0

Fonte: (TEIXEIRA, 2005).

Na figura 3.6, foi montada a roleta de acordo com os valores de adaptabilidade expostos na tabela 3.1 para esta população hipotética.

Figura 3.6 – Roleta montada para população hipotética.



Fonte: (TEIXEIRA, 2005).

3.2.1.2. Rank (*Ranking Selection*)

Uma alternativa à seleção proporcional ao *fitness* é o método de seleção por ranking. Este método procura evitar a convergência prematura da população para um provável mínimo local (MITCHELL, 1999).

Na versão proposta por BAKER (1985) apud TEIXEIRA (2005), os indivíduos são ordenados de acordo com o seu *fitness* em posições seqüenciais que desprezam o montante de diferença existente entre eles.

Esse desprezo da diferença total de *fitness*, de acordo com MITCHELL (1999), leva às vantagens de evitar a convergência prematura, mantendo a pressão seletiva controlada quando a variação de *fitness* for muito alta. Entretanto, ele não é indicado em alguns casos em que seja importante conhecer o quanto um indivíduo é melhor que o outro (diferença de *fitness*).

Os indivíduos selecionados dependem de sua posição no ranking e da distribuição de probabilidade utilizada pela função adotada. Desta forma, pode-se ter o Rank Linear e Exponencial.

Maiores detalhes deste método, bem como as funções de probabilidade adotadas podem ser encontrados em TEIXEIRA (2005) e MITCHELL (1999).

3.2.1.3. Torneio (*Tournament Selection*)

Segundo HAUPT et. al. (2004), o método de seleção por Torneio é um dos mais próximos de simular o processo natural de competição para obter a chance de reprodução. Na abordagem definida por este autor e por EBERHART et al. (1996), são escolhidos dois indivíduos da população de forma aleatória, o de maior *fitness* “vence” a disputa e é escolhido como um dos possíveis pais. O processo é repetido até atingir o número de pais esperado.

MITCHELL (1999) defende outra abordagem para o método do Torneio. São escolhidos dois indivíduos aleatoriamente da população e, um fator k , variando no intervalo $[0,1]$, é gerado aleatoriamente. Se este fator for maior que um determinado valor (estabelecido pelo usuário, por exemplo, 0,75), o indivíduo de maior *fitness* é selecionado, caso contrário, o de menor é selecionado para a área de cruzamento.

Algumas variações do método do Torneio selecionam um grupo de indivíduos de tamanho t , realizando sucessivas disputas entre eles, até que o melhor seja escolhido (EBERHART et al., 1996).

Pode-se verificar que o método de seleção por Torneio se comporta ainda melhor que o método por Rank. Isto porque, além de controlar a pressão seletiva e evitar a rápida convergência de forma equilibrada, diferente do método de Rank, não necessita de ordenação, possuindo um algoritmo de complexidade $O(n)$.

3.2.1.4. Estado Fixo (*Steady State*)

O Estado Fixo é uma abordagem que propõe que apenas alguns indivíduos da população sejam substituídos pelos novos indivíduos gerados (MITCHELL, 1999). Ao contrário de grande parte das abordagens de AGs, em que todos os indivíduos da geração parental são substituídos pela filial.

Na implementação desse método, geralmente, é escolhido uma parcela da população (indivíduos de menor *fitness*) para serem substituídos pela nova prole (TEIXEIRA, 2005).

Esta abordagem tem larga utilidade quando da aplicação de AGs na otimização de sistemas baseados em regras e outros caso, em que a resposta depende de todo o conjunto da população (MITCHELL, 1999).

3.2.1.5. Elitismo (*Elitism*)

Para EBERHART et al. (1996), é possível que o indivíduo de maior *fitness* de uma geração não sobreviva às operações de seleção, recombinação e mutação. Sendo assim, uma excelente solução que demorou muitas gerações para surgir poderia ser destruída e, em alguns casos, dificilmente seria encontrada novamente.

O Elitismo é um método complementar aos outros e visa estabelecer uma cópia exata do melhor ou dos n melhores indivíduos de uma geração a outra (MITCHELL, 1999). Deve-se tomar cuidado ao escolher a quantidade, pois um mínimo local pode ser fixado e levar a população à estagnação frente à evolução.

3.2.1.6. Outros Métodos para o Operador de Seleção

Os métodos que já foram apresentados acima apresentam inúmeras variações, que acabam atendendo a um ou outro tipo de problema, ou seguem descritos de formas alternativas por autores diferentes.

Em TEIXEIRA (2005), são apresentados os métodos: Seleção por Diversidade, onde os mais diversos indivíduos da população são selecionados; Seleção Bi-classista, onde os “P%” melhores indivíduos são selecionados e os “100 – P%” são selecionados; Seleção Aleatória, todos os indivíduos são selecionados de forma aleatória; e Seleção Aleatória Salvacionista, onde somente o primeiro melhor é escolhido de forma não aleatória.

Nos trabalhos de LEHRER (2000), BRITO (2004), TEIXEIRA (2005) e suas publicações derivadas, bem como em todos os testes realizados neste trabalho, são utilizados operadores de seleção híbridos que envolvem uma fase de disputa (interação) baseada no Jogo Evolucionário Hawk-Dove e em comportamentos geneticamente codificados. Estes métodos

como o Hawk-Dove Roleta² e o Hawk-Dove Torneio, tendem a diminuir a pressão seletiva e evitar a convergência prematura. Mesmo assim, como será exposto a seguir, estes estão longe de impedir completamente que a população fique presa em mínimos locais.

Na seqüência, estão expostos os operadores genéticos de recombinação ou cruzamento (*crossover*) e mutação.

3.2.2. Operação de Recombinação

O operador de recombinação ou *crossover* é a principal característica que distingue os AGs de outros métodos de busca, e é considerado o mais importante entre os operadores genéticos (MITCHELL, 1999).

Conforme, HAUPT et al. (2004), é a primeira maneira que um algoritmo genético usa para explorar a superfície de busca dentro do material genético (alelos) que já se encontra nos indivíduos da população, levando o AG à convergência de resultados.

A recombinação, de um modo geral, consiste na troca de pedaços (partes) de cromossomos entre os indivíduos selecionados como pais, para a criação de novos indivíduos melhores adaptados (FOGEL, 2000).

Diversos métodos de recombinação foram propostos, adequando-se a inúmeras situações, sendo os principais expostos nos itens que seguem.

² Maiores detalhes sobre esta variação em algoritmos genéticos, chamados de Algoritmos Genéticos Híbridos com Interação Fenotípica, podem ser obtidos no Anexo E, que contém uma listagem de publicações e um artigo na íntegra (BRITO et. al., 2005) sobre o assunto. Dentro do quarto capítulo, pode-se encontrar também uma abordagem sintética destes métodos já que, apesar de não ser o foco do trabalho, é utilizada em todos os testes realizados.

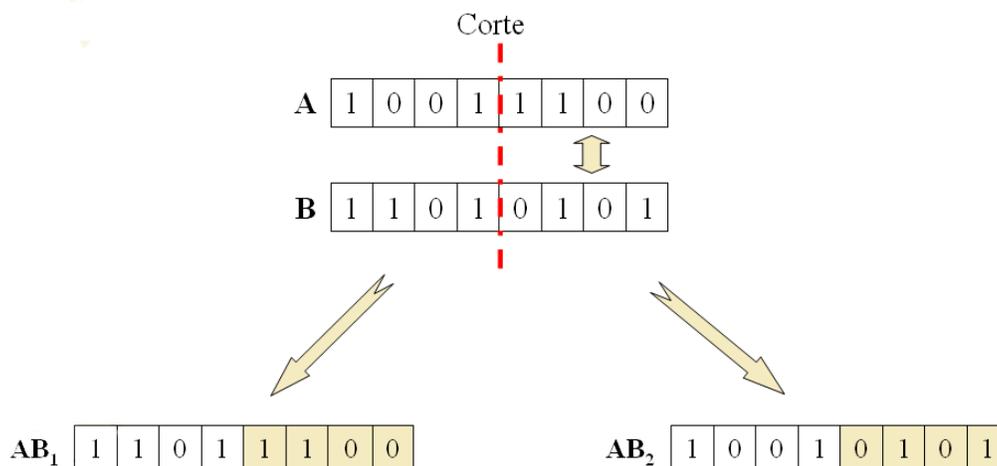
3.2.2.1. Um Ponto de Corte (*1-Point Crossover*)

É apresentado por praticamente todos os autores, sendo considerado o método mais comum de recombinação, com funcionamento mais simples.

Apresentado originalmente por Holland, este método gera dois cromossomos filhos a partir de dois pais, selecionando-se uma posição aleatória (ponto de corte) nos cromossomos pais, e intercalando os pedaços gerados. Um filho é gerado pelo pedaço do cromossomo do primeiro pai até a posição do ponto de corte, mais o pedaço do segundo pai da posição do ponto de corte até o final; o outro filho é gerado pela composição das partes que restaram (FOGEL, 2000).

A figura 3.7 exemplifica o método proposto, onde os cromossomos A e B geram os filhos AB₁ e AB₂.

Figura 3.7 – Crossover com um ponto de corte.

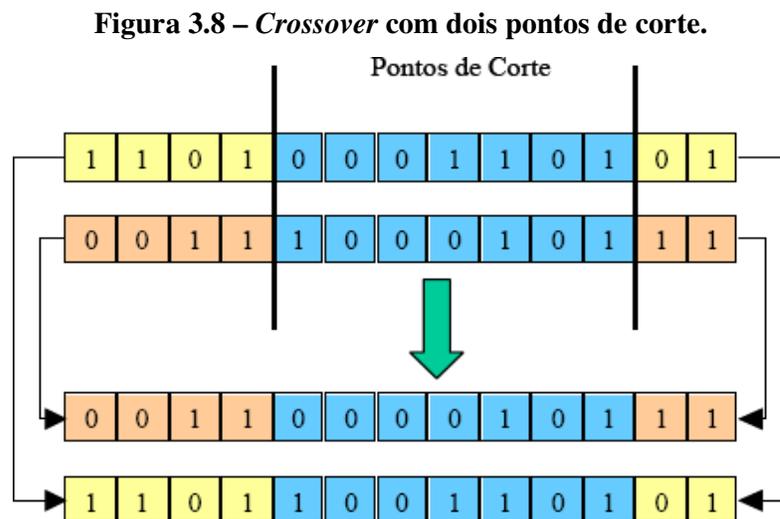


Fonte: (MOREIRA,2004)

3.2.2.2. Dois Pontos de Corte (*2-Point Crossover*)

A operação utilizando dois pontos de corte, ocorre de forma semelhante a anterior, a diferença é que são selecionados dois pontos de corte aleatórios.

Segundo TEIXEIRA (2005), o pedaço remanescente entre os pontos de corte é preservado, sendo trocadas apenas as extremidades, como demonstrado na figura 3.8.



Fonte: (TEIXEIRA, 2005)

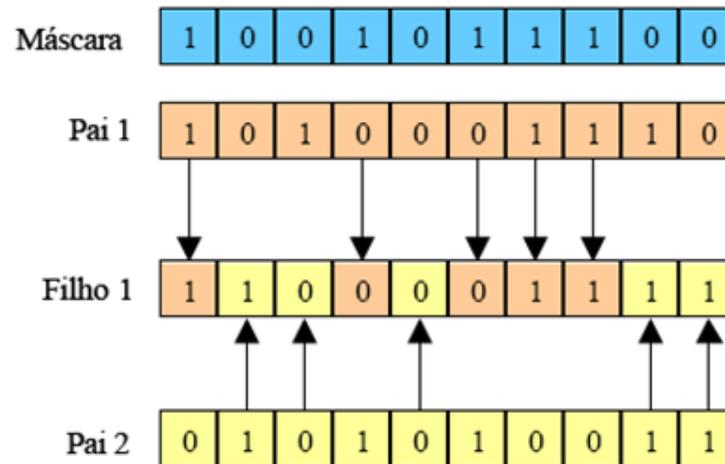
3.2.2.3. Uniforme (*Uniform Crossover*)

Uma abordagem que difere significativamente das demais é o Operador de Recombinação Uniforme. Nele, cada gene herdado pelo filho tem 50% de chances de vir de cada um dos pais (EBERHART et al., 1996).

Para que seja implementado, é gerada uma máscara binária de tamanho igual ao do cromossomo, sendo preenchida com valores aleatórios. Se o valor numa posição x for o bit 1, o filho herda o alelo do pai 1 da mesma posição, caso contrário herda o alelo de mesma

posição do pai 2 (HAUPT et al., 2004). Esta operação é demonstrada em detalhes na figura 3.9.

Figura 3.9 – Crossover uniforme.



Fonte: (TEIXEIRA, 2005).

3.2.2.4. Parcialmente Mapeado (*PMX*)

Para certos tipos de problema, os operadores de recombinação apresentados geram filhos incoerentes com a solução esperada. Um grupo de problemas afetados pelos operadores apresentados, são os problemas de ordenação, onde o exemplo mais famoso é o Problema do Caixeiro Viajante (HAUPT et al., 2004), apresentado neste trabalho nos capítulos a seguir.

Visando atender a estes problemas, Goldberg sugeriu o método Parcialmente Mapeado ou Operador de Comparação Parcial (*PMX*). Este método, segundo HAUPT et al. (2004), funciona mapeando uma parte dos caracteres (alelos) de um dos cromossomos pais em função de uma parte dos caracteres do outro cromossomo, sendo os caracteres remanescentes trocados.

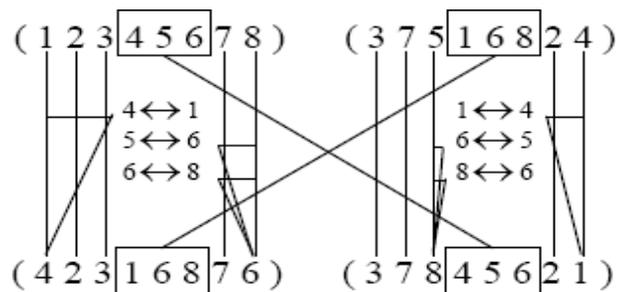
TEIXEIRA (2005) se utiliza o exemplo descrito a seguir para demonstrar o funcionamento do método. São utilizados os dois cromossomos-rotas abaixo:

Pai 1: (1 2 3 4 5 6 7 8)

Pai 2: (3 7 5 1 6 8 2 4)

Selecionam-se dois pontos de corte aleatórios e uniformes sobre os cromossomos pais, supondo que o primeiro ponto seja entre o terceiro e quarto alelo, e o segundo entre o sexto e sétimo alelo. Ao pedaço entre os pontos de corte é dado o nome de “Seção de Mapeamento”. Estes primeiros passos estão representados na figura 3.10.

Figura 3.10 – Realizando o PMX.



Fonte: (TEIXEIRA, 2005)

As seções de mapeamento aparecem envolvidas na figura 3.10, e a partir dela são gerados os seguintes mapeamentos: $4 \leftrightarrow 1$, $5 \leftrightarrow 6$, $6 \leftrightarrow 8$. Depois, o filho 1 recebe a seção de mapeamento do pai 2 e o filho 2 recebe a do pai 1, apresentando a estrutura abaixo:

Filho 1: (x x x | 1 6 8 | x x)

Filho 2: (x x x | 4 5 6 | x x)

No momento seguinte, os filhos i ($i=1, 2$) recebem o elemento do i -ésimo pai, se algum alelo já estiver presente no cromossomo do filho, este é substituído pelo seu correspondente na seção de mapeamento, até que seja encontrado um não repetido.

Desta forma, são gerados, sem repetição de alelos, os filhos 1 e 2, apresentados abaixo:

Filho 1: (4 2 3 | 1 6 8 | 7 5)

Filho 2: (3 7 8 | 4 5 6 | 2 1)

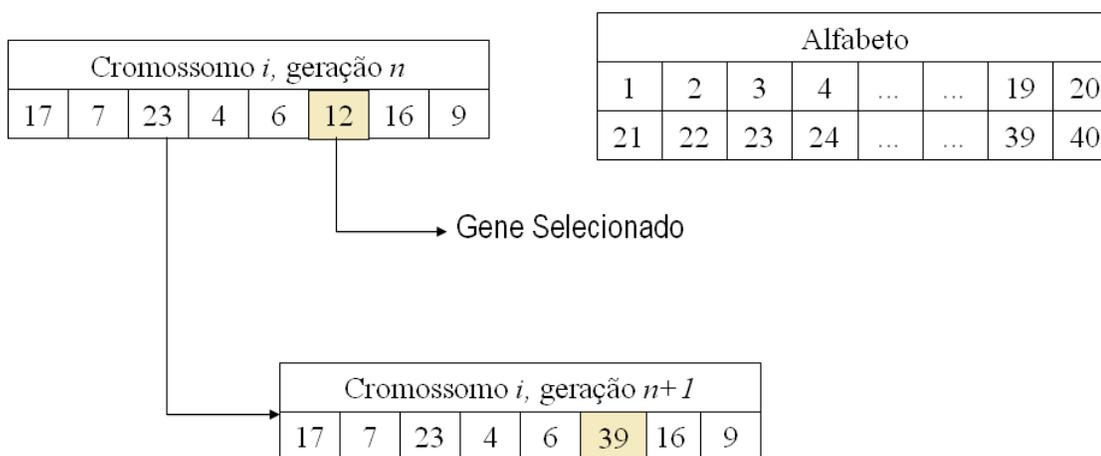
No item que segue será abordado outro importante operador genético, o Operador de Mutação.

3.2.3 Operação de Mutação

Segundo HAUPT et al. (2004), o Operador de Mutação é a segunda maneira que um AG utiliza para explorar a superfície de busca do problema. Ele pode introduzir um alelo que não estava na população e evitar a sua rápida convergência “escapando” de mínimos locais.

A operação de mutação funciona fazendo pequenas alterações dos genes da população substituindo-se de forma aleatória por um elemento do alfabeto de alelos (EBERHART et al., 1996). A figura 3.11 exemplifica esta operação.

Figura 3.11 – Operação de Mutação.



Fonte: (MOREIRA, 2004)

Para MITCHELL (1999), a probabilidade de um gene qualquer ser “mutado” deve ser aproximadamente 0,001 %.

Sobre as mutações, LEHRER (2000) afirma que esta não é a melhor forma de gerar novas estruturas, sendo o Operador de Recombinação muito mais eficiente para tal função.

De posse dos conhecimentos acerca do funcionamento de AGs, é importante compreender quando estes deverão ser aplicados, e conhecer algumas soluções práticas de sucesso que os aplicam. Estes assuntos estão descritos na seção a seguir.

3.3 “QUANDO UTILIZAR ALGORITMOS GENÉTICOS?” - APLICAÇÕES

Como já exposto no início do capítulo, os AGs em alguns problemas apresentam, se comparados com outros métodos de busca, baixo desempenho em relação ao tempo. Nestas condições, como saber quando utilizá-los?

MITCHELL (1999), defende a idéia compartilhada por muitos pesquisadores de que se o espaço de busca for muito grande (inúmeras soluções), a superfície for rugosa ou pouco compreendida, ou a função de *fitness* for ruidosa, e se não for necessário encontrar um ótimo global, então os AGs apresentam um bom desempenho em relação a outros métodos.

Esta definição da utilização não é totalmente certa, haja vista que o desempenho do AG estará sempre diretamente ligado à modelagem do problema, métodos de operadores escolhidos e parâmetros utilizados em sua aplicação.

É importante ressaltar, que esta é ainda uma técnica pouco explorada dentro do ramo da IA, que durante muito tempo voltou praticamente todos os seus esforços para a área conexionista (EBERHART et al.,1996). Além disso, os AGs têm apresentado muitos casos de sucesso aplicados a diversas áreas, da medicina à astrofísica.

KOZA et al.(2003), pesquisadores de renome da computação evolucionária, simularam a construção de circuitos eletrônicos em AGs, modelando à função de adaptação a função esperada do circuito. Com algumas iterações conseguiram alcançar a mesma estrutura

de circuito criada pelos humanos. Depois de muitas gerações conseguiram produzir diversos modelos completamente inovadores que estão sendo patenteados, reforçando que a estrutura de AG não utilizou nenhum conhecimento avançado de eletrônica. A evolução se encarregou do processo.

Outra aplicação que tem economizado milhares de dólares para empresas controladoras de redes de satélite, foi a formação de “constelações” inovadoras de satélites. Apresentada em CROSSLEY (2001), o uso de constelações é necessário para otimizar a presença de poucos satélites de baixa altitude, ideais para aplicações de comunicação móvel, diminuindo ao máximo as janelas (intervalo de tempo em que as antenas não “enxergam” os satélites devido a curvatura da Terra).

AGs também têm sido utilizados para otimizar aditivos de combustíveis, explorando a combinação de milhares de cadeias de moléculas, e produzindo resultados excepcionais (VENKATASUBRAMANIAN, 1999).

Além dessas aplicações AGs também têm sido aplicados em robótica, biologia, otimização de grades de horários, classificação de dados, otimização de pesos para redes neurais e em muitas outras áreas (MITCHELL, 1999).

3.4 O PROBLEMA DAS ESCOLHAS: MÉTODOS E PARÂMETROS

Como descrito em DEMBSKI (2007), Algoritmos Genéticos estão longe de serem “o solucionador definitivo de problemas” (“*the ultimate problem solver*”). Pelo contrário, ainda segundo este autor e as deduções do Teorema do “*No Free Lunch*”, dada a falta de

conhecimento do domínio e o conseqüente mal uso de parâmetros, estas buscas não são nenhum pouco melhores que as buscas cegas tradicionais (“*Blind Searches*”)³.

A já mencionada e notada – até mesmo pela extensão deste capítulo – escolha de parâmetros, métodos, técnicas de codificação e operadores, tornam o uso e a plena difusão de AGs substancialmente complexos. Isto porque a expertise requerida, restringe e muito o possível número de usuários e dá lugar a outras técnicas.

Sem falar na grande quantidade de novas variáveis biológicas (observáveis na natureza) quem vem sendo acrescentadas a estrutura canônica dos AGs e que, apesar de melhorarem em alguns casos suas performance, ainda assim implicam em um número ainda maior de parâmetros a serem configurados, aumentando exponencialmente esta complexidade.

Um exemplo local é a própria abordagem híbrida (LEHRER, 2000; BRITO, 2004) utilizada neste trabalho, que, conforme exposto, trata também do comportamento e da interação entre os indivíduos (soluções), acrescentando parâmetros como número de partidas, tipo de jogo, custos da tabela de pagamentos, distribuição dos comportamentos e memória.

Descartando a escolha dos métodos descritos nas seções anteriores (operadores de recombinação, mutação e seleção), que, naturalmente, estão intrinsecamente ligadas ao tipo do problema que se pretende resolver, recai-se ainda sobre a complexa combinação de parâmetros, como por exemplo: o tamanho da população, o critério de parada e as taxa de recombinação e mutação.

³ Em seu livro “*No Free Lunch: Why Specified Complexity Cannot Be Purchased Without Intelligence*”, William A. Dembski, utiliza-se de um capítulo inteiro para apresentar uma dedução lógica-matemática, derrubando o tradicional exemplo de algoritmo Genético de suporte ao darwinismo de Richard Dawkins, “ME*THINKS*IT*IS*LIKE*A*WEASEL”, argumentando que, sem inteligência ou conhecimento, ainda que parcial sobre o domínio, os Algoritmos Genéticos não passam de buscas cegas.

Em HAUPT (2004), o estudo da escolha dos parâmetros é agrupado em três grandes grupos: **estudos empíricos, estudos teóricos baseados em fatos e a auto-adaptação de parâmetros.**

No primeiro grupo, o dos estudos empíricos, um dos primeiros e principais trabalhos desenvolvidos foi o de DEJONG (1975), que, após exaustivos testes com muitas combinações de parâmetros para um conjunto de cinco problemas, concluiu serem os seguintes valores ideais para uma boa performance em AGs: população de 50 a 100 indivíduos; probabilidade de recombinação de 0,6 (60%); e, probabilidade de mutação em 0,001 (0,1%) por gene.

Os estudos de DEJONG (1975) foram tão difundidos e adotados, que até hoje aparecem na literatura, erroneamente, como padrão para qualquer problema a ser solucionado usando AGs.

O segundo grupo, o dos estudos teóricos, dada a complexidade da correlação entre os valores dos diferentes parâmetros, apresenta inúmeros estudos de parâmetros isolados, que permitiram um aprofundamento na real dinâmica dos AGs. Destes, podem-se destacar: os estudos dos parâmetros de seleção por GOLDBERG e DEB (1991) apud HAUPT et al. (2004); do operador de mutação por MÜHLENBEIN (1992) e BÄCK (1993) apud HAUPT et al. (2004); e, do tamanho de populações por GOLDBERG et al. (1992) apud HAUPT et al. (2004).

Só o último trabalho sobre o tamanho das populações já apresenta um resultado completamente conflitante ao valor “ideal” de 50 a 100 indivíduos apresentado em DEJONG (1975).

O terceiro e último grupo, que diz respeito aos parâmetros codificados e auto-adaptáveis durante o processo de execução do algoritmo evolucionário, são bastante referenciados na literatura e suas abordagens podem ser enquadradas em dois grandes agrupamentos: o primeiro que trata de colocar estes parâmetros dentro da execução normal da

técnica evolucionária, geralmente codificando-os como parte dos cromossomos que buscam retratar as soluções do problema, tal qual fazem as Estratégias Evolucionárias; e, o segundo, que define um meta algoritmo genético de nível superior (ou paralelo), que se auto adapta ao longo da execução (GREFENSTETTE, 1986 apud HAUPT et al. 2004).

De um modo geral, nenhuma das abordagens apresentadas nos três grupos pode ser tida como verdade absoluta e solução definitiva para o problema da definição paramétrica e das escolhas dos Algoritmos Genéticos. De um jeito ou de outro, o “melhor” conjunto sempre estará atrelado ao problema, e, irá requerer esforço do pesquisador e um conseqüente ajuste.

Estar atrelado ao problema que se quer solucionar, não é a principal falha, mas sim a forma como este relacionamento é feito e o esforço requerido do pesquisador.

Neste ponto, é onde o presente trabalho introduz o agente inteligente, o qual, monitora a população durante a execução e ajusta os parâmetros, sem no entanto, requerer qualquer intervenção do pesquisador, como será visto nos capítulos que seguem.

3.5 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados os principais elementos de estrutura de AGs, paralelo com a biologia, nomenclatura básica, principais métodos de seleção, principais métodos para os operadores genéticos, aplicações práticas e o problema das escolhas e da variação paramétrica.

Compreendidos estes assuntos, para que se possa apresentar as propostas do trabalho, ainda se faz necessário a compreensão dos conceitos complementares expostos no capítulo a seguir: Lógica Nebulosa ou Difusa (Fuzzy), utilizada na construção do motor da proposta; a abordagem híbrida de AG utilizada baseada na Teoria dos Jogos Evolucionários (LEHER, 2000; BRITO, 2004; TEIXEIRA, 2005); a Teoria do Design Inteligente; e, a Teoria da Extinção Cíclica (RAUP e SEPKOSKI, 1984).

CAPÍTULO IV

Aquilo que é impenetrável para nós existe de fato. Por trás dos segredos da natureza há algo sutil, intangível e inexplicável. A veneração a essa força que está além de tudo que podemos compreender é a minha religião.

Albert Einstein

4. PRINCIPAIS TÉCNICAS E CONCEITOS DE SUPORTE À PROPOSTA

Realizada a explanação dos assuntos fundamentais à compreensão dos mecanismos da evolução e da genética (Capítulo II), e das técnicas de computação evolucionária, em especial os Algoritmos Genéticos (Capítulo III), faz-se necessária a apresentação de conceitos estritamente ligados à proposta do agente GAIA, tanto em aspectos filosóficos ou suporte teórico, quanto em aspectos práticos de implementação. Neste sentido, o presente capítulo é plural.

Inicia apresentando duas teorias polêmicas e substancialmente controversas, na medida em que sofrem diversas retaliações da maior parte da comunidade científica. A primeira delas é chamada de Teoria do Design Inteligente⁴ e, diferente do que fora apresentado, trata de uma nova perspectiva para a origem e a evolução da vida. Ainda assim, pode-se dizer, é o “cerne teórico” da proposta do agente do trabalho. Já a segunda, intitulada

⁴ De uma maneira geral, além desta teoria, que por si só já é bastante controversa, muitos tópicos polêmicos e de diferentes crenças, religiões e culturas serão expostos neste capítulo. **Nenhum deles expressa de forma alguma a opinião ou crença própria do autor.** Todavia, são necessários para a total compreensão dos aspectos da proposta de melhoria da estrutura tradicional dos Algoritmos Genéticos deste trabalho.

de Teoria da Extinção Cíclica de Raup-Sepkoski, de forma complementar explica como a “intervenção” do agente e suas conseqüências devem ter reflexo sobre a população.

Na seqüência do capítulo, dedica-se uma seção exclusiva a apresentação da Lógica Difusa ou Nebulosa (Fuzzy), seus conceitos e sua aplicação, que serão o motor fundamental para a materialização da “inteligência” do agente proposto.

Por fim, na última seção, é tratada, sem no entanto desviar o foco principal do trabalho, a abordagem híbrida de Algoritmos Genéticos de BRITO (2004) e BRITO et al. (2005)⁵, que, apesar de preservar na íntegra todas as etapas do ciclo de um AG Canônico (vide Capítulo III), inclui uma fase especial de interação social (ou fenotípica).

4.1 DESIGN INTELIGENTE

O renomado matemático Blaise Pascal (1623-1622), em sua obra “A Arte da Persuasão”, afirma que: “[...] as pessoas, quase que invariavelmente, chegam às suas crenças, não com base em provas, mas sim naquilo que lhes é atraente.” É com essa frase que DEMBSKI (2004) abre uma de suas principais obras sobre a Teoria do Design Inteligente (*Intelligent Design*).

Desde que surgiu até os dias de hoje, a evolução nos moldes concebidos por Charles Darwin, é a teoria prevalecente para explicar a origem, manutenção e evolução da vida. Esta até pouco tempo, não encontrou nenhuma outra “rival” à altura, salvo suas próprias falhas, lacunas e falta de explicação, que quase sempre são respondidas com as palavras: “acaso”, “sorte” e “eventualidade”.

⁵ Maiores detalhes descritos no Anexo E.

Nesta lacuna, é que cientistas e pesquisadores renomados das mais diversas áreas do conhecimento, da matemática à biologia, propõe uma nova abordagem chamada de Design Inteligente. Definida como a ciência que estuda os chamados “sinais de inteligência”.

Para facilitar a compreensão do conceito, DEMBSKI (2004) sugere que se imagine a escultura do Monte Rushmore⁶ sendo esculpida de forma completamente aleatória, apenas pela erosão causada pela chuva e pelo vento. Certamente, seria impossível de acontecer, pois existem padrões e complexidades em sua elaboração que identificam claramente o sinal de uma inteligência por trás, ou seja, de um *designer*⁷ inteligente.

Os teóricos do design inteligente não tentam decifrar a forma como o *designer* pensa, nem quem ele é, muito menos seus próximos passos ou algum tipo de adoração. De forma simples e objetiva, a teoria é focada em utilizar métodos científicos para identificar traços de inteligência nos mais variados acontecimentos e elementos físicos, em especial nos aspectos biológicos dos seres e, isso a faz tão controversa e criticada no meio científico.

Nos últimos dois séculos a falta de uma metodologia científica e de evidências concretas foram as maiores dificuldades desta proposição. Além do fato de que muitas vezes esta é confundida ou posta como um “neo-criacionismo”, ou mesmo uma doutrina diretamente ligada a religiosos que se opõe completamente a evolução de Darwin.

Como uma teoria da origem da vida e seu desenvolvimento, o ponto central do Design Inteligente é que somente seria adequado explicar a complexidade e as estruturas biológicas ricas em informação por causas inteligentes, e que estas podem ser detectadas empiricamente. Isto quer dizer que existem métodos bem definidos, que baseados em características

⁶ Escultura artística em formação rochosa homônima dos quatro presidentes norte americanos de maior admiração por aquele povo: George Washington, Thomas Jefferson, Theodore Roosevelt e Abraham Lincoln.

⁷ Em todas as vezes que a palavra “designer” aparecer no texto, entenda-se, numa tradução livre, o significado de arquiteto, planejador, elaborador, ou até mesmo construtor.

observáveis do mundo, podem distinguir causas indiretas e naturais das inteligentes (DEMBSKI, 2007).

Um exemplo científico claro destes métodos, são os procedimentos utilizados pelos astrônomos do Projeto SETI⁸, onde, por meio de algoritmos e proposições matemáticas, tenta diferenciar padrões aleatórios de ondas de rádio vindos do espaço, de padrões que efetivamente demonstrem alguma assinatura de inteligência.

O notório astrônomo Carl Sagan, em seu livro de ficção “Contato” apud DEMBSKI (2004), propõe uma situação em que, para diferenciar-se dos sinais aleatórios vindos do espaço, uma suposta civilização inteligente, transmite uma seqüência de 101 números primos por meio de variações de ausência e presença numa portadora de sinais.

Neste caso, ficam exemplificadas duas características propostas pelos teóricos do Design Inteligente: primeiro, nada nas leis da física obrigada que um sinal adquira uma forma ou outra, além disso o seu tamanho exclui qualquer possibilidade do simples acaso, levando-se ao conceito de “complexidade”; segundo, ela não simplesmente transmitiu uma seqüência qualquer de números, mas sim uma seqüência que segue um padrão definido, a seqüência dos números primos, de onde se infere o segundo conceito, o de “especificidade”.

Segundo DEMBSKI (2004), a combinação deste dois conceitos é a “marca registrada” dos sinais do design inteligente, a chamada *specified complexity* (“complexidade especificada”, numa tradução livre).

Portanto, na tentativa de determinar se organismos biológicos possuem esta característica (*specified complexity*), os teóricos focam nos sistemas identificáveis que são

⁸ SETI (*Search for Extra-Terrestrial Intelligence*, ou Busca por Inteligência Extraterrestre) é um projeto que tem por objetivo analisar, por meio de inúmeros radiotelescópios, sinais de rádio provenientes dos confins do universo na tentativa de descobrir e/ou comunicar-se com alguma forma de vida inteligente no universo (MOURÃO, 2002).

específicos em virtude de sua independência funcional e ao mesmo tempo exibem um alto grau de complexidade, tais como: enzimas, vias metabólicas, “máquinas” moleculares e outros.

DEMBKSI (2004) conclui que se uma parte essencial de um organismo possuir características do *design*, então ela por si só agrega essa característica a todo o organismo. Sem no entanto a exigência de que todas as partes possuam evidências claras de design. De fato, até numa forma flexível e combinável com a evolução de Darwin e com a proposta dos Algoritmos Genéticos, algumas partes serão plenamente geradas por causas naturais (acaso), no modo “tentativa e erro”.

O que se argumenta é que, não é impossível nem proibido que a natureza e o acaso gerem a complexidade especificada, o problema argumentado é que ela faça isso sem nenhum conhecimento, ou partindo de um ponto completamente aleatório do vazio. Em DEMBKSI (2007), prova-se que complexidade especificada pré-existente, quando “misturada” por causas naturais pode gerar novas formas evoluídas que também a apresentem sobre nova forma.

Sobre este aspecto é que se pode comparar com o tema deste trabalho, mesmo os algoritmos genéticos, quando modelados em termos de cromossomos, função de avaliação e outros, expressam o conhecimento do pesquisador sobre o problema e, de uma forma ou de outra, acabam por deixar sua assinatura.

O *designer* não é necessariamente aquele que tem todas as respostas ou soluções, mas aquele que, por algum objetivo, estabeleceu padrões específicos, essenciais para o funcionamento de um sistema.

Outro conceito proposto por um dos teóricos do design, Michael Behe apud DEMBSKI (2007), é o de *irreducible complexity*, ou, numa tradução livre, complexidade irredutível. Este conceito trata de sistemas grandes e complexos em que remover a menor que seja das partes, destrói completamente a funcionalidade do todo. Um exemplo biológico

clássico, é o “motor” do flagelo de bactérias, um sistema complexo baseado em uma série de enzimas, em que a simples inibição de uma proteína é capaz de inviabilizar todo o sistema.

Estes dois principais conceitos tornam causas inteligentes empiricamente detectáveis e fazem do Design Inteligente uma teoria científica completa, distinguindo-a dos argumentos de *design* postos por filósofos e teólogos, ou do que antes era chamado de “Teologia Natural”.

Apresentada a conceituação geral da teoria, nas seções seguintes, serão tratados aspectos do design inteligente mais propriamente relacionados com a proposta deste trabalho.

4.1.1 O Princípio do “*No Free Lunch Theorem*”

William A. Dembski dedica uma obra inteira (DEMBSKI, 2007) ao uso do Teorema “*No Free Lunch for Optimization*” de Wolpert e Macready, para comprovar que o mecanismo darwiniano de busca (“sem objetivo”) para criação e evolução da vida, não faz sentido sem algum conhecimento prévio do que se pretende encontrar, tal qual, na proposta original, comparam-se os métodos de computação evolucionária e suas respectivas performances que, segundo os autores provam, nestas mesmas condições não passam de “buscas cegas”.

O teorema original em si exige uma certa dedução matemática que envolve uma alta complexidade, porém, utilizando o exemplo e a dedução de DEMBSKI (2004), o qual ele chama de “*displacement*” (ou deslocamento, numa tradução livre) está compreensão faz-se de forma facilitada.

DEMBSKI (2004) cita como exemplo um espaço de busca bastante extenso, em que as possibilidades são tantas e tão improváveis de se imaginar, que uma busca exaustiva não é viável e, uma aleatória certamente não irá encontrar a solução. Como consequência, para que esta busca seja realizada, é necessária alguma restrição, qualquer informação sobre o espaço que servirá como guia na busca. Acontece que, qualquer informação útil possui também o seu próprio espaço ou domínio. Desta forma, a busca principal cede o lugar, ou seja, é deslocada,

para a busca de informações que ajudariam ao encontro da solução do problema original. E é esse segundo espaço que, segundo DEMBKI (2007), é no mínimo de mesma ordem e complexidade, se não superior, ao original.

Em linhas gerais, DEMBKI (2007), conclui que conhecer qualquer informação deste espaço superior que leva à solução, remete ao conceito de “*specified complexity*” e, como já abordado, automaticamente o leva para o espaço de ordem inferior, ou original. Sendo assim, não há como se dizer que a natureza, sem conhecer onde quer se chegar, ou ter objetivo, não conseguiria por si só desenvolver tais complexas estruturas.

Pensando em termos de computação evolucionária – algoritmos inspirados no modelo original de Darwin –, a única maneira de resolver este problema e não ficar sempre propondo uma série de soluções híbridas que são, como ditas no teorema do NFL, iguais em performance a uma busca cega, é inserir informação sobre o espaço de busca e, ao longo da evolução, ajustar a função de avaliação de adaptabilidade (*fitness*).

DEMBKI (2007) apresenta esta solução como “*time-varying/coevolving fitness measures*”, ou, numa tradução livre, função de avaliação que evolui/modifica-se em função do tempo. Exatamente o que é feito com a introdução de um agente inteligente para o controle da execução dos algoritmos genéticos proposto neste trabalho.

Ainda que, como será visto no próximo capítulo, este agente não haja diretamente sobre a função de *fitness*, ele indiretamente produz modificações na população e na sobrevivência dos indivíduos, valendo-se para tal do que ele passa a considerar, ao longo da evolução, o melhor caminho de busca.

4.1.2 Intervencionismo

O Design Inteligente não exige que os organismos apareçam repentinamente, tampouco sejam criados do “nada” por meio da simples intervenção de um *designer*

inteligente. Todavia, a teoria não se restringe a estes fatores. Ela é tanto compatível com essa idéia apresentada do criacionismo, quanto com a idéia de que os organismos evoluíram de forma natural em estruturas mais complexas com o tempo, tal qual propõe Darwin. A teoria não se atem a estes fatores, mas sim ao fato do que foi responsável pela evolução.

Os teóricos da evolução tradicional pregam que dispositivos mecânicos e materiais por si sós são responsáveis pela evolução, enquanto isso, o Design Inteligente, propõe que estes eventos materiais são capazes apenas de mudanças evolutivas limitadas. Para DEMBKI (2004), qualquer grande mudança, irá exigir a **intervenção** de uma concepção inteligente.

Não importa como um organismo surgiu ou chegou a um estágio de evolução, o que de fato importa para esses teóricos é que num dado momento a intervenção de uma inteligência foi crucial na formação deste indivíduo.

Dadas essas colocações, surgem os questionamentos de onde, quando, como e com que frequência devem estas intervenções ocorrerem. DEMBKI (2004) apresenta este ponto como o de maior debate entre os propositores desta teoria. E destes, pode-se inferir três principais correntes:

- **Intervenção mínima:** ocorre somente quando se faz estritamente necessária, deixando todo o resto para os processos evolutivos naturais;
- **Intervenção constante:** todas as modificações certas ou não são frutos de testes e acertos realizados pelo *designer*;
- **Intervenção no ambiente:** esta corrente está mais ligada ao como do que propriamente ao tempo em que ocorrem, nela os teóricos afirmam que o designer manipula o ambiente em busca das mudanças necessárias e que estes traços também podem ser identificados por meio de métodos científicos;

Neste trabalho, como será apresentado no próximo capítulo, todas essas abordagens intervencionistas serão modeladas e utilizadas e, naturalmente, chegar-se-á a conclusão de que a situação e o problema são os que melhor definem o momento da intervenção.

Ainda dentro deste capítulo e seu contexto teórico de subsídio da proposta, na seção seguinte será apresentada uma teoria e um conjunto de crenças que definem um pouco do momento e da forma como a intervenção supostamente ocorre.

4.2 TEORIA DA EXTINÇÃO CÍCLICA DE RAUP-SEPKOSKI E VARIAÇÕES

De forma complementar ao que já fora apresentado sobre o conceito de um suposto *designer* e de sua vontade se valer, ainda que de forma eventual, sobre a criação e a evolução da vida, faz-se ainda uma breve síntese sobre como tal “vontade” (ou intervenção) supostamente ocorre na forma de catástrofes de cunho planetário ou até mesmo localizado sobre uma ou outra espécie.

Como apresentado na síntese de pesquisas arqueológicas em DANIKEN (2005), muitas civilizações antigas acreditavam que de tempos em tempos a própria natureza se voltava contra elas em eventos cataclísmicos dizimadores de boa parte da vida na Terra.

O povo judeu, por exemplo, narra em seus textos que por muitas vezes, o *designer*, por meio do que se foi chamado de “juízo”, aniquilou uma grande quantidade de povos e até mesmo espécies que não estavam de acordo com sua “vontade” ou objetivos. Os maias, os quais em sua cultura dividem o tempo em ciclos, narram que ao término de cada um destes uma grande catástrofe reiniciaria toda a vida e muitas mudanças se desencadeariam. Assim como eles, o povo veda, que deu origem aos hindus, também acreditava nessa destruição cíclica.

Egípcios e sumérios em suas escrituras foram até mais específicos, chegando a colocar que essa era a maneira de proporcionar, que o mundo fosse habitado por novas espécies, melhores e mais evoluídas (DANIKEN, 2005).

Para a maioria das pessoas, essas afirmações não passam de lendas, teorias conspiratórias e crenças religiosas. “Quem” ou “o quê” as causou sempre será algo difícil de provar, e talvez até impossível.

O fato é que hoje a ciência, por meio dos achados fósseis e do estudo da geologia, tem evidências concretas de que de tempos em tempos grandes extinções em massa ocorreram e, que estes restos de seres se depositam no solo como camadas para cada um desses eventos, que inclusive são utilizados para separar as “eras” da história da terra e os seres vivos que nela habitavam.

Em MOURÃO (2002) é dito que, registros e achados fósseis mostram classes inteiras de formas de vida desaparecendo em diferentes épocas nos últimos 650 milhões de anos. Um desses maiores eventos de extinção ocorreu há cerca de 65 milhões de anos, nos limites entre as eras geológicas Secundária e Terciária. Nessa época, uma enorme quantidade de plantas e animais - quase a metade de todo o biogênero existente - desapareceu completamente. Este evento de extinção narrado, talvez seja o de maior notoriedade e conhecimento, e diz respeito à Extinção dos Dinossauros⁹.

⁹ Em 1981, Luis Walter Alvarez, prêmio Nobel de Física em 1968 e seu filho, o geólogo Walter Alvarez, apresentaram a hipótese de que, há 65 milhões de anos, um asteroide de uma dezena de quilômetros de diâmetro e massa de quase 13 trilhões de toneladas se teria chocado com a Terra. Tal choque, provocou uma explosão de 100 milhões de *megatons* e aniquilou com boa parte da vida e das espécies dominantes no planeta (MOURÃO, 2002).

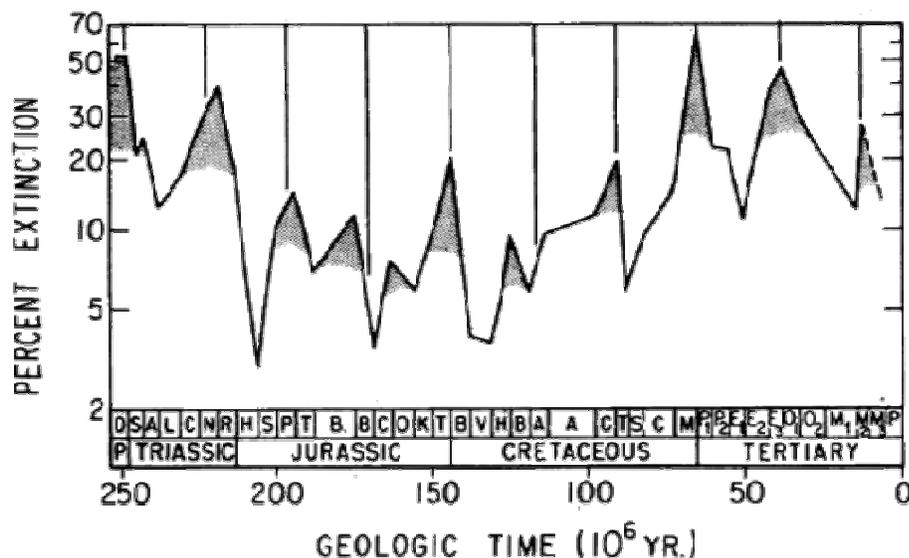
A hipótese dos Alvarez, diferente da explicação darwiniana tradicional de que todos os dinossauros simplesmente evoluíram para outros seres¹⁰, tem grande valor ao explicar o súbito desaparecimento dos dinossauros e outras espécies de uma maneira muito mais lógica que as hipóteses anteriores.

A primeira suspeita científica de que esses eventos de extinção ocorrem regularmente foi levantada em 1977 por dois pesquisadores americanos, David Raup e John Sepkoski Junior, da Universidade da Califórnia. Estes, depois de estudarem a extinção de 600 famílias de vida marinha nos últimos 250 milhões de anos, constataram doze diferentes ocorrências de extinção no período e, sem apontarem a causa, propuseram a chamada Teoria da Extinção Cíclica (RAUP e SEPKOSKI, 1984).

O gráfico 4.1, listado a seguir, foi apresentado originalmente nos trabalhos de RAUP e SEPKOSKI (1984), como síntese dos resultados de seus estudos. Neste, é possível notar, por meio dos picos de extinção de espécies, a repetição de padrões ao longo dos últimos 250 milhões de anos.

¹⁰ Pela semelhança genética, estudos recentes apontam para as aves como a evolução das espécies remanescentes de dinossauros no momento posterior ao cataclismo (MOURÃO, 2002).

Gráfico 4.1 – Percentual de Extinção x Tempo (250 ma)



Fonte: (RAUP e SEPKOSKI, 1984).

A teoria do impacto dos Alvarez pode parecer em contradição com a Teoria da Extinção Cíclica de Raup-Sepkoski. Todavia, estudos em 88 crateras produzidas na superfície terrestre por impacto de corpos celestes mostram uma interessante periodicidade entre elas - algo em torno de 28 a 31 milhões de anos (MOURÃO, 2002).

Como dito, não se pode afirmar o que ou quando esses eventos ocorreram, apenas o que se pode concluir é o que os fatos apresentam. Estes apontam para uma extinção de mais 90% das espécies já existentes na Terra e, diferente do motor proposto por Darwin, boa parte dessas mudanças não aconteceu de forma gradual, mas sim, como afirmado em NATGEO (2010), de forma repentina e fora dos padrões convencionais associados às taxas de evolução.

De certo, essas catástrofes parecem ter servido como catalisadores do processo tradicional da evolução, pois sempre após sua ocorrência, novas espécies mais evoluídas surgem e, rapidamente, tornam-se dominantes.

Baseado nesta idéia principal é que o agente proposto, ao intervir, dependendo do estágio de estagnação da evolução, causa grandes extinções em massa e outras conseqüências

para a população, as quais irão permeá-la por muitas gerações, na tentativa de trazê-la novamente para a busca do ótimo global. Maiores detalhes desta representação na estrutura de algoritmos genéticos serão descritos no próximo capítulo.

4.3 TEORIA DOS CONJUNTOS E SISTEMAS NEBULOSOS – LÓGICA FUZZY

No início dos anos 60, Lotfi A. Zadeh afirmou que a teoria de controle clássica havia posto muito ênfase na precisão e, assim, não conseguiria lidar com sistemas complexos. Dessa forma, ele disse que seria necessário um tipo de matemática radicalmente diferente, a matemática de quantidades nebulosas, que não fosse descrita em razão de distribuições de probabilidade. Mais tarde, em 1965, com o artigo intitulado *Fuzzy Sets*, Zadeh inspirou os matemáticos a “fuzzificarem” as estruturas matemáticas e, conseqüentemente, nasceu a Lógica Fuzzy ou Nebulosa (FUZZY, 1999).

Apesar de um nascimento conturbado, onde alguns acadêmicos apoiaram as idéias e, então, passaram a trabalhar no desenvolvimento da área, enquanto outros a rejeitaram, pois a viam na contramão dos princípios científicos. Dentre estes, os matemáticos, especializados em estatística e probabilidade, acreditavam que: (1) a Teoria das Probabilidades era suficiente para caracterizar incerteza; (2) todo e qualquer problema resolvido pela teoria nebulosa poderia ser resolvido igualmente ou melhor através da Teoria das Probabilidades. E, esta visão foi difícil de ser alterada, pois não haviam aplicações práticas reais, somente seus fundamentos filosóficos (WANG, 1997).

Lotfi A. Zadeh, em seu trabalho *Fuzzy Sets*, de 1965, apresentou as noções dos conjuntos nebulosos como uma extensão dos conceitos da teoria dos conjuntos clássica, onde uma afirmação só pode ser verdadeira ou falsa (0 ou 1). A proposta da lógica nebulosa, base

para compreensão da teoria dos conjuntos nebulosos, é permitir a qualquer afirmação a atribuição de um valor verdade variando continuamente entre os valores 0 e 1.

A Lógica Nebulosa, dessa forma, representa uma classe distinta de incerteza, que não é em essência probabilística, ao lidar qualitativamente com as características de um objeto, que podem ser vagas ou imprecisas, tais como: “alto”, “baixo”, “grande”, “pequeno”, entre outros.

Nas seções seguintes são apresentados os principais aspectos para a compreensão e, posterior, aplicação dos conceitos pertinentes à Teoria dos Conjuntos e Sistemas Nebulosos aos propósitos deste trabalho.

4.3.1 Conjuntos Nebulosos

O conceito de conjunto nebuloso é muito simples e intuitivamente interessante de ser aplicado na modelagem de situações e problemas do mundo real. Além disso, as suas vantagens são ainda maiores a partir do momento que se tenta imitar o comportamento humano no processo de tomada de decisão. Como exemplo, ao decidir sobre a compra de um produto, uma pessoa qualifica as variáveis numéricas presentes no processo e, assim, verifica se o produto é caro, confiável, bom ou ruim, de forma a tomar a sua decisão.

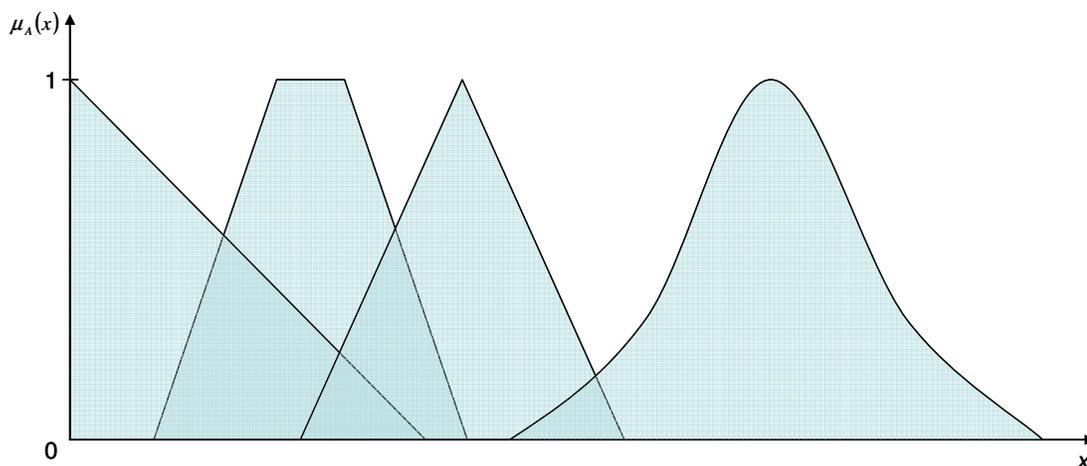
Dessa forma, pode-se definir que um conjunto nebuloso A no universo de discurso U é caracterizado pela função de pertinência $\mu_A(x)$, que pode assumir valores contínuos no intervalo $[0,1]$. Assim, um conjunto nebuloso nada mais é do que uma generalização de um conjunto clássico.

Alguns aspectos dos conjuntos nebulosos são importantes de se destacar, entre eles (FUZZY, 1999):

- as propriedades caracterizadas por eles são nebulosas e, assim, diferentes funções de pertinência podem representá-las. Essas funções são funções matemáticas precisas e, a partir do momento que representam propriedades nebulosas, nada mais é nebuloso e, nesse sentido, o que ocorre na verdade é a “defuzzificação” dessa propriedade;
- como determinar as funções de pertinência? Existem duas abordagens: (1) usar o conhecimento do especialista humano. Normalmente se faz necessária uma sintonia fina para ajustar as funções de pertinência, pois esta abordagem fornece o conhecimento bruto, não refinado; (2) usar dados coletados por sensores como sintonia fina para funções de pertinência previamente estabelecidas;
- um conjunto nebuloso tem uma referência de um para um com sua função de pertinência, ou seja, ela é única para cada conjunto.

Como dito anteriormente, um conjunto nebuloso pode ser representado através de sua função de pertinência e, nesse sentido, a figura 4.1, a seguir, apresenta algumas das principais formas de se representar essas funções.

Figura 4.1 – Principais Formas de Representação das Funções de Pertinência



Fonte: Adaptado de FUZZY (1999).

4.3.2 Operações e Relações Sobre Conjuntos Nebulosos

Os conceitos apresentados anteriormente estão relacionados a um único conjunto nebuloso. E, assim, esta seção objetiva abordar as operações e relações entre conjuntos nebulosos, conforme pode ser visto a seguir.

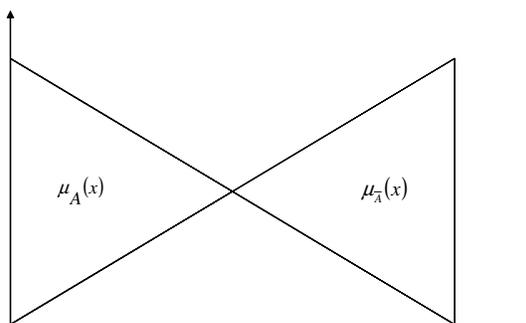
4.3.2.1. Operações

As operações para manipulação de conjuntos nebulosos foram definidas a partir das operações existentes na Teoria Clássica dos Conjuntos e, assim, há uma série de operações envolvendo-os. Como forma de apresentá-las, considere-se os dois conjuntos nebulosos A e B , e a partir deles pode-se definir as seguintes definições e operações.

- **Igualdade:** A e B são *iguais*, se e somente se, $\mu_A(x) = \mu_B(x)$, $\forall x \in U$;
- **Subconjuntos:** $(A \subset B)$, se e somente se, $\mu_A(x) \leq \mu_B(x)$, $\forall x \in U$;

- **Complemento:** o conjunto A tem como complemento \bar{A} e a sua função de pertinência é definido por $\mu_{\bar{A}} = 1 - \mu_A$, conforme a figura 4.2 abaixo.

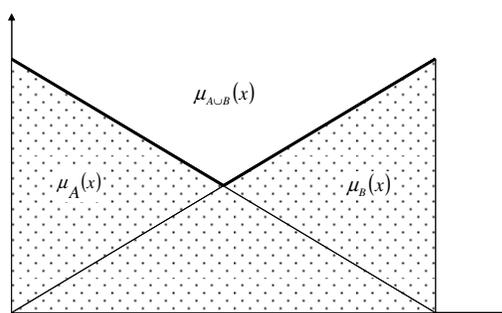
Figura 4.2 – Exemplo do Complemento de um Conjunto Nebuloso



Fonte: Adaptado de WANG (1997).

- **União:** a união de A e B ($A \cup B$) tem a função de pertinência definida por $\mu_{A \cup B} = \max[\mu_A(x), \mu_B(x)]$, ou seja, é o menor conjunto nebuloso que contém ambos os conjuntos A e B , conforme a figura 4.3 a seguir apresenta.

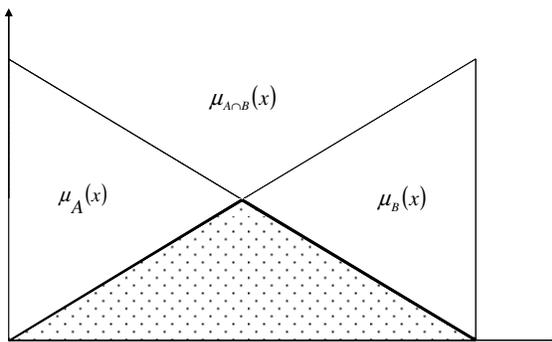
Figura 4.3 – Exemplo de União de Conjuntos



Fonte: Adaptado de WANG (1997).

- **Interseção:** a interseção de A e B ($A \cap B$) tem a função de pertinência definida por $\mu_{A \cap B} = \min[\mu_A(x), \mu_B(x)]$, ou seja, é o maior conjunto nebuloso que contém ambos os conjuntos A e B , conforme pode ser visto na figura 4.4 a seguir.

Figura 4.4 – Exemplo de Interseção de Conjuntos



Fonte: Adaptado de WANG (1997).

As operações de complemento, união e interseção apresentadas até aqui são consideradas operações básicas e, assim, não se adequam satisfatoriamente a toda e qualquer situação existente. Dessa forma, existem outros tipos de operações que são utilizados em situações onde, por exemplo, a união de A e B não contenha necessariamente o menor conjunto. Assim tem-se (WANG, 1997):

- **Complemento Nebuloso** (*Fuzzy Complement*): seja $c: [0,1] \rightarrow [0,1]$ a função de mapeamento que transforma a função de pertinência de um conjunto A na função de pertinência do seu conjunto complemento \bar{A} . ou seja:

$$c[\mu_A(x)] = \mu_{\bar{A}}(x) = 1 - \mu_A(x) \quad (01)$$

Para c poder se qualificado como complemento de A , duas condições devem ser consideradas:

1. Limite (*Boundary Condition*): $c(0) = 1$ e $c(1) = 0$;
2. Não-Crescimento (*Nonincreasing Condition*): um aumento no grau de pertinência de um conjunto, deve resultar no decréscimo ou nenhuma mudança no grau de pertinência de seu complemento e, assim, para todo $a, b \in [0,1]$, se $a < b$ então

$c(a) \geq c(b)$. Caso contrário, qualquer violação dessas condições, fará com que c não seja uma função aceitável.

- **União Nebulosa: Norma-S** (*Fuzzy Union: The S-Norms*): seja $s: [0,1] \times [0,1] \rightarrow [0,1]$ a função de mapeamento que transforma as funções de pertinência de dois conjuntos nebulosos A e B , na função de pertinência da união desses conjuntos, ou seja:

$$s[\mu_A(x), \mu_B(x)] = \mu_{A \cup B} = \max[\mu_A(x), \mu_B(x)] \quad (02)$$

Para s poder se qualificado como a união de A e B , tem-se quatro condições que devem ser consideradas:

1. Limite (*Boundary Condition*): $s(1,1) = 1$, $s(0,a) = s(a,0) = a$;
2. Comutativa (*Comutative Condition*): a ordem dos conjuntos não influencia no resultado, ou seja, $s(a,b) = s(b,a)$;
3. Não-Decrescimento (*Nondecreasing Condition*): o aumento nos graus de pertinência dos dois conjuntos resulta no crescimento no grau de pertinência da união deles., ou seja, se $a \leq a'$ e $b \leq b'$, então $s(a,b) \leq s(a',b')$;
4. Associativa (*Associative Condition*): permite estender o uso da união para mais de dois conjuntos nebulosos. Assim, $s(s(a,b),c) = s(a,s(b,c))$. Caso contrário, qualquer violação dessas condições, fará com que s não seja uma função aceitável.
5. Existem inúmeras classes de normas-S, além da apresentada anteriormente na equação (02), as quais não serão objeto deste trabalho;

- **Interseção Nebulosa: Norma-T** (*Fuzzy Intersection: The T-Norms*): seja $t: [0,1] \times [0,1] \rightarrow [0,1]$ a função de mapeamento que transforma as funções de pertinência de dois conjuntos nebulosos A e B , na função de pertinência da união desses conjuntos, ou seja:

$$t[\mu_A(x), \mu_B(x)] = \mu_{A \cap B} = \min[\mu_A(x), \mu_B(x)] \quad (03)$$

Para t poder se qualificado como a interseção de A e B , tem-se quatro condições que devem ser consideradas:

1. Limite (*Boundary Condition*): $t(0,0) = 0$, $t(1,a) = t(a,1) = a$;
2. Comutativa (*Comutative Condition*): a ordem dos conjuntos não influencia no resultado, ou seja, $t(a,b) = t(b,a)$;
3. Não-Decrescimento (*Nondecreasing Condition*): o aumento nos graus de pertinência dos dois conjuntos resulta no crescimento no grau de pertinência da união deles., ou seja, se $a \leq a'$ e $b \leq b'$, então $t(a,b) \leq t(a',b')$;
4. (4) Associativa (*Associative Condition*): permite estender o uso da união para mais de dois conjuntos nebulosos. Assim, $t(t(a,b),c) = t(a,t(b,c))$

4.3.2.2 Relações

Uma relação é classicamente definida a partir da presença ou ausência de uma associação ou interação entre elementos de dois ou mais conjuntos. Sendo assim, com base em dois conjuntos A e B , a relação R é definida por $A \times B$ que, por sua vez, é um subconjunto do produto cartesiano de ambos os conjuntos, de forma que $R: A \times B \rightarrow \{0,1\}$. Caso algum elemento $a \in A$ e $b \in B$ estejam relacionados, então $R(a,b) = 1$, senão $R(a,b) = 0$, conforme demonstra a sua função de pertinência bivalente apresentada a seguir (WANG, 1997).

$$\mu_R(a,b) = \begin{cases} 1 & \text{se e somente se } (a,b) \in R \\ 0 & \text{em caso contrário} \end{cases} \quad (04)$$

A partir da definição de relação entre conjuntos clássicos, pode-se dizer que as relações nebulosas são a sua generalização. E, assim, é definida como o produto cartesiano entre mais de um conjuntos discretos, com grau de pertinência calculado pela função $\mu_R(a,b) \in [0,1]$, onde $a \in A$ e $b \in B$, que é nebulosa.

Como forma de exemplificar a relação nebulosa, tome-se como base os conjuntos discretos $A = \{A_1, A_2\}$ e $B = \{B_1, B_2\}$, que representam pontos diversos de uma região geográfica. Assim, a relação “muito longe” entre os pontos de cada um dos conjuntos pode ser definida e quantificada, conforme pode ser vista a seguir em (05), ou seja, $R: A \times B \rightarrow [0,1]$ com $\mu_R(a,b) \in [0,1]$, onde $a \in \{A_1, A_2\}$ e $b \in \{B_1, B_2\}$.

$$\begin{array}{c}
 B \\
 \hline
 \begin{array}{cc}
 & B_1 & B_2 \\
 A \left| \begin{array}{cc}
 A_1 & 0,1 & 0,8 \\
 A_2 & 0,9 & 0,3
 \end{array}
 \end{array}
 \end{array} \quad (05)$$

As operações de união e interseção podem também ser aplicadas às relações nebulosas de maneira similar ao que ocorre nos conjuntos nebulosos. E, assim, considerando as relações R e S definidas por $A \times B$, as funções de pertinência para a união e interseção são definidas então da seguinte maneira.

$$\mu_{R \cup S}(a,b) = \mu_R(a,b) * \mu_S(a,b) \quad (06)$$

$$\mu_{R \cap S}(a,b) = \mu_R(a,b) \oplus \mu_S(a,b) \quad (07)$$

Além disso, a *composição de relações* tem um papel fundamental no processo de inferência nebuloso, que será visto mais adiante. Sendo assim, considerando duas relações nebulosas $R(A,B)$ e $S(B,C)$, que tem o conjunto B em comum, a composição dessas relações

é definida como um subconjunto $T(A, C)$ de $A \times C$, com $(a, c) \in T$ se e somente se existir pelo menos um $b \in B$, $(a, b) \in R$ e $(b, c) \in S$, conforme pode ser visto a seguir.

$$T(A, C) = R(A, B) \circ S(B, C) \quad (08)$$

Sendo assim, a função de pertinência para a composição $T(A, C)$ é dada por:

$$\mu_T(a, c) = \mu_{R \circ S}(a, c) = \max_{b \in S} t[\mu_R(a, b), \mu_S(b, c)] \quad (09)$$

para todo $(a, c) \in A \times C$, onde t é qualquer Norma-T. No entanto, as duas mais comumente utilizadas para a realização da composição de relações nebulosas são as composições *max-min* e *max-produto*, conforme pode ser visto respectivamente a seguir nas equações (10) e (11).

No entanto, essas não são as únicas expressões possíveis.

$$\mu_T(a, c) = \mu_{R \circ S}(a, c) = \max_{b \in S} \min[\mu_R(a, b), \mu_S(b, c)] \quad (10)$$

$$\mu_T(a, c) = \mu_{R \circ S}(a, c) = \max_{b \in S} [\mu_R(a, b) \mu_S(b, c)] \quad (11)$$

4.3.3 Variáveis Lingüísticas

O conceito de variável lingüística está intimamente relacionado à representação do conhecimento humano. Quando sensores são utilizados na obtenção de valores de uma variável, eles resultam em números precisos, porém, ao invés de sensores, se humanos forem utilizados, então os resultados serão apresentados em palavras. Para exemplificar essa situação, a medição de temperatura através do uso de um termômetro, indicando precisamente 38,6° de febre ou, simplesmente, uma pessoa dizendo “muito quente”.

Dessa forma, se uma variável permite palavras em linguagem natural como seus valores, então pode ser considerada como uma *variável lingüística*. Cada uma das palavra é

caracterizada por um conjunto nebuloso no universo de discurso em que é definida. Formalmente, uma variável lingüística é caracterizada a seguir (FUZZY, 1999):

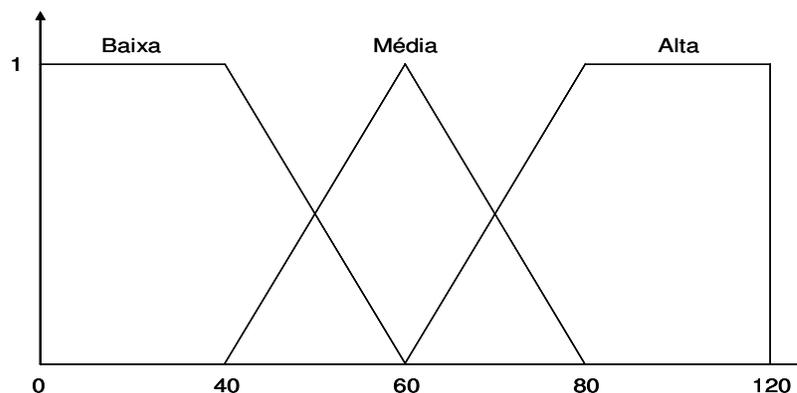
$$VL = (N, T, U, G, M) \quad (12)$$

onde:

- *N*: nome da variável lingüística;
- *T*: conjunto de termos, ou seja, são os valores lingüísticos atribuídos a *N*;
- *U*: é o atual domínio físico, onde são especificados os valores quantitativos (discretos) da variável lingüística *N*. É também chamado de Universo de Discurso;
- *G*: são as regras sintáticas, geralmente representada por uma gramática, utilizadas para gerar os nomes das variáveis lingüísticas *N*;
- *M*: regra semântica que relaciona cada um dos *T* termos lingüísticos com um conjunto nebuloso especificado no universo de discurso *U* e representado pelas funções de pertinência. Ou seja, $T \times U \rightarrow M$.

Como forma de exemplificar, pode-se pensar na velocidade de um carro, onde os termos que a qualificam são expressos por baixa, média e alta, e o universo de discurso é definido no intervalo $[0, 120]$, e a relação destes dois últimos componentes possibilita a geração das funções de pertinência dos conjuntos nebulosos, conforme figura 4.5 a seguir.

Figura 4.5 – Funções de Pertinência da Variável Velocidade



Fonte: (BORGES, 1996).

Os valores atribuídos às variáveis lingüísticas podem ser representados por apenas uma ou mais de uma palavras. Assim, no exemplo anterior, a variável *velocidade* poderia ter outros valores, tais como: “muito baixa”, “ligeiramente alta” e “mais ou menos média”.

Formalmente, o valor de uma variável lingüística é um termo composto, como por exemplo, $x = x_1, x_2, \dots, x_n$, que nada mais é do que a concatenação de outros termos, x_1, x_2, \dots, x_n . Normalmente são classificados em três grupos: (1) *termos primários*: são os rótulos utilizados para nomear os conjuntos nebulosos, como: “baixa”, “média” e “alta”; (2) *complemento* “não” e os *conectivos* “e” e “ou”; (3) *restrições*, como: “muito”, “ligeiramente” e “mais ou menos”.

Os conectivos e o complemento já foram apresentados em seção anterior e, assim, falta ainda caracterizar as restrições. Estas podem ser consideradas como uma forma de “*fuzzificar*” os termos primários, expandindo as suas definições e, assim, aproximando-as ainda mais do real significado do conhecimento que se deseja representar.

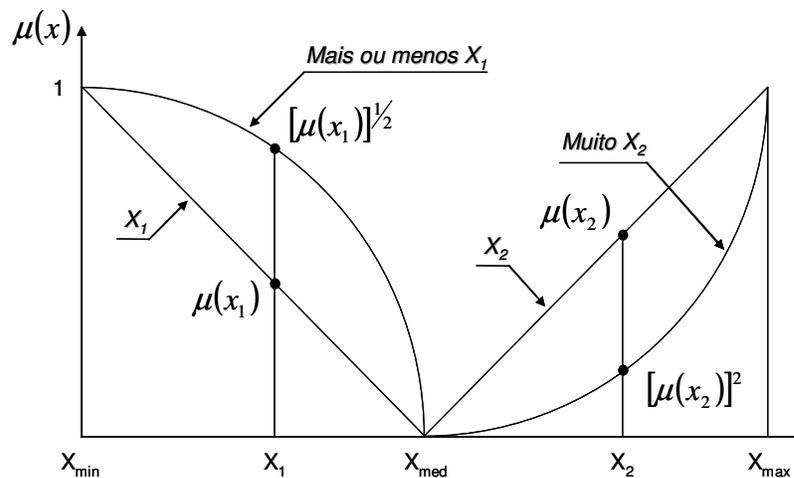
As duas restrições mais comumente utilizadas são “muito” e “mais ou menos”, que também é conhecida como “moderada”. Nenhuma delas têm um significado muito claro, no entanto, “muito” é utilizado no sentido de intensificar o termo primário. As suas funções de

pertinência foram convencionadas e são apresentadas a seguir em (13) e (14), respectivamente. E, na seqüência, tem-se na figura 4.6 a sua representação gráfica.

$$\mu_{\text{Muito}}(A(x)) = [\mu_A(x)]^2 \quad (13)$$

$$\mu_{\text{Mais ou menos}}(A(x)) = [\mu_A(x)]^{1/2} \quad (14)$$

Figura 4.6 – Funções de Pertinência para as Restrições “Muito” e “Mais ou menos”.



Fonte: Adaptado de BORGES (1996).

É importante observar que as restrições podem ter dois tipos de efeito quando aplicadas ao conceito original do termo, que são: (1) *Concentração*: é o que ocorre com a restrição “muito”, o que possibilita uma intensificação ou reforço ao termo principal; (2) *Dilatação*: é o que ocorre com a restrição “mais ou menos”, onde há uma expansão ou enfraquecimento do termo principal (FUZZY, 1999).

Além disso, os expoentes 2 e $\frac{1}{2}$ são apenas convenções, porém é possível utilizar quaisquer valores, pois expoentes maiores que “um” ativam o efeito de concentração ao gerar menores graus de pertinência e, para valores “menores que um”, possibilitam o aumento do grau de pertinência devido ao efeito da dilatação.

Não existe uma regra formal para a definição quantitativa exata para o grau de influência dos modificadores lingüísticos sobre conjuntos nebulosos e, dessa forma, a associação de um valor numérico a um expoente para definir uma restrição, é bastante subjetivo.

4.3.4 Lógica, Regras “Se-Então” e Inferência Nebulosa

Da mesma forma que a teoria dos conjuntos nebulosos é uma extensão da teoria clássica dos conjuntos, a lógica nebulosa é uma generalização da lógica clássica, onde muito de seus conceitos continuam válidos. E, assim, existem dois tipos de Proposições Nebulosas (BORGES, 1996):

1. as *atômicas*, que são representadas por apenas uma afirmação nebulosa $x \text{ é } A$, onde x é uma variável lingüística e A é um valor lingüístico de x ;
2. as *compostas*, que combinam proposições atômicas através dos conectivos lógicos “e”, “ou ” e “não”, que representam a interseção, união e o complemento, respectivamente.

É importante observar que as proposições nebulosas compostas devem ser compreendidas como relações nebulosas. Sendo assim, considerando x e y como variáveis lingüísticas nos domínios físicos U e V , e A e B como conjuntos fuzzy nesses domínios, respectivamente, as suas funções de pertinência são definidas a seguir:

- **Conectivo “e”:** utiliza o conceito de interseção nebulosa e, assim, tem-se a proposição ($x \text{ é } A$ e $y \text{ é } B$), que pode ser interpretado como a relação nebulosa $A \cap B$ em $U \times V$, com a função de pertinência a seguir:

$$\mu_{A \cap B}(x, y) = t[\mu_A(x), \mu_B(y)] \quad (15)$$

onde $t : [0,1] \times [0,1] \rightarrow [0,1]$ é qualquer *Norma-T*.

- **Conectivo “ou”:** utiliza o conceito de união nebulosa e, assim, tem-se a proposição (*x é A ou y é B*), que pode ser interpretado como a relação nebulosa $A \cup B$ em $U \times V$, com a função de pertinência a seguir:

$$\mu_{A \cup B}(x, y) = s[\mu_A(x), \mu_B(y)] \quad (16)$$

onde $s: [0,1] \times [0,1] \rightarrow [0,1]$ é qualquer Norma-S.

- **Conectivo “não”:** utiliza o conceito de complemento nebuloso e, assim, a proposição (*x é A*) tem o seu complemento (*x não é A*), com função de pertinência conforme à operação expressa por (01).

Com base nas equivalências lógicas e nas varias formas de representar as operações nebulosas de complemento, união e interseção, é possível gerar inúmeras interpretações a respeito da implicação nebulosa (*i*), conforme pode ser visto na tabela 4.1 a seguir.

Tabela 4.1 – Diferentes Formas de Interpretação da Implicação Nebulosa

Interpretações	Função de Pertinência
Dienes- Rescher	$\mu_{R_D}(a, b) = \max[1 - \mu_{PN_1}(a), \mu_{PN_2}(b)]$
Lukasiewicz	$\mu_{R_L}(a, b) = \min[1, 1 - \mu_{PN_1}(a) + \mu_{PN_2}(b)]$
Zadeh	$\mu_{R_Z}(a, b) = \max[\min(\mu_{PN_1}(a), \mu_{PN_2}(b)), 1 - \mu_{PN_1}(a)]$
Gödel	$\mu_{R_G}(a, b) = \begin{cases} 1 & \text{se } \mu_{PN_1}(a) \leq \mu_{PN_2}(b) \\ \mu_{PN_2}(b) & \text{caso contrário} \end{cases}$
Mamdani	$\mu_{R_{M_1}}(a, b) = \min[\mu_{PN_1}(a), \mu_{PN_2}(b)]$ ou $\mu_{R_{M_2}}(a, b) = \mu_{PN_1}(a)\mu_{PN_2}(b)$

Fonte: Adaptado de WANG (1997).

4.3.5 Sistemas Nebulosos

De acordo com FUZZY (1999), um Sistema Nebuloso (SN) é um sistema baseado em conhecimento ou baseado em regras, que tem as *Regras* “*Se-então*” como o seu núcleo principal, onde algumas das palavras que as compõem são caracterizadas por funções de pertinência contínuas. Como exemplo, tem-se: *Se a velocidade é alta, então aplicar menos força no acelerador.*

Para construir um sistema nebuloso é necessário três etapas: (1) definir a base de regras “*Se-então*”, que são extraídas a partir do conhecimento de um especialista humano; (2) combinar todas as regras “*Se-então*” da base de regras em um único sistema; (3) escolher um modelo de sistema nebuloso¹¹.

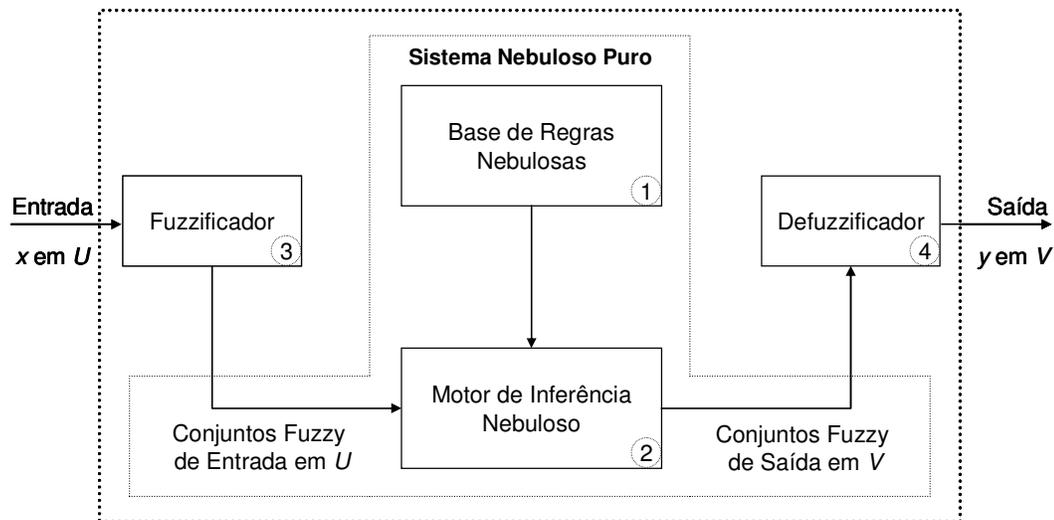
O modelo de sistema nebuloso mais comumente utilizado é o de *Mamdani*, que é na verdade uma generalização do sistema nebuloso puro, de forma a permitir que as variáveis entradas e saídas sejam valores reais (*R*), como no modelo de *Takagi-Sugeno-Kang*. Dessa forma, ele é mais adequado na construção de sistemas de engenharia, tal como é o utilizado na construção do agente da proposta do trabalho. Na figura 4.7, tem-se esse modelo e os seus respectivos componentes.

Antes de detalhar os quatro componentes do modelo apresentado na figura 4.7, é importante observar que os sistemas nebulosos são mapeadores de múltiplas entradas em uma

¹¹ Os três principais modelos de sistemas nebulosos são (WANG, 1997): Sistema Nebuloso Puro (*Pure Fuzzy System*), representado na figura 4.7; Sistema Nebuloso de Takagi-Sugeno-Kang (*TKS Fuzzy System*), proposto para suprir a necessidade de variáveis de entrada e saída com valores reais, presente nos sistemas de engenharia; e, o Sistema Nebuloso de Mamdani com Fuzzificador e Defuzzificador (*Fuzzy System with Fuzzifier and Defuzzifier*). Maiores detalhes, assim como inúmeros exemplos práticos, podem ser encontrados em FUZZY (1999).

única saída, ou seja, a partir de um vetor de valores reais $(x_1, \dots, x_n \in U)$ mapeia um escalar de valor real $(y \in V)$. Sendo assim, o sistema pode ser construído em partes, como sistemas individuais e, depois, agrupados em um único sistema. Cada um dos sistemas é composto basicamente de quatro partes, que são apresentadas a seguir em maiores detalhes.

Figura 4.7 – Sistema Nebuloso de Mamdani com Fuzzificador e Defuzzificador



Fonte: Adaptado de FUZZY (1999).

4.3.5.1 Base de Regras Nebulosas

Consiste em um conjunto de regras “Se-então”, que são, conforme já dito, o núcleo principal de um sistema nebuloso, de forma que todos outros os componentes são utilizados para implementar essas regras de forma eficiente. Sendo assim, o seu formato pode ser visto a seguir, em sua forma canônica:

$$\text{Regra}^{(l)} : \text{Se } x_1 \text{ é } A_1^l \text{ e } \dots \text{ e } x_n \text{ é } A_n^l, \text{ então } y \text{ é } B^l \quad (17)$$

onde: A_i^l e B^l são conjuntos nebulosos em $U_i, V \subset R$, respectivamente; $x_1, \dots, x_n \in U$ são as entradas; e, $y \in V$ é a saída.

A base de regras consiste em um conjunto de regras nebulosas e o relacionamento existente entre elas é muito importante para o desempenho do sistema como um todo, pois é possível que haja algum tipo de conflito em suas definições, como também, a mal representação do conhecimento humano. Sendo assim, as propriedades a seguir devem ser consideradas:

- **Integridade:** um conjunto de regras nebulosas deve ser completo, isto é, para cada $x \in U$, existe pelo menos uma regra nebulosa presente na base, de forma que, o seu grau de pertinência seja diferente de zero;
- **Consistência:** um conjunto de regras é consistente caso não haja duas ou mais regras com partes “se” iguais e “então” diferentes, o que pode causar dificuldades na busca de um regra para uso, gerando assim conflitos;
- **Continuidade:** um conjunto de regras nebuloso é contínuo caso não haja regras vizinhas onde a parte “então” apresentam valor vazio para interseção.

Para a composição da base regras, pode e deve ser utilizado um especialista, de forma que este forneça sentenças lingüísticas. Esta não é uma tarefa fácil, mesmo para um especialista e, assim, existem outros métodos para a geração da base de regras, através, por exemplo, da realização da tarefa de mineração de um base de dados numéricos.

4.3.5.2 Motor de Inferência Nebuloso

Os princípios abordados da lógica nebulosa neste capítulo são utilizados na composição do motor de inferência nebuloso, onde nele são combinadas as regras “Se-então” presentes na base de regras, de forma a mapear um conjunto nebuloso $A \in U$ em um conjunto nebuloso $B \in V$. Caso este seja composto por uma única regra, então o método *Modus Ponens Generalizado* (WANG, 1997) é indicado. Porém, como a grande maioria das

aplicações de sistemas nebulosos apresentam um conjunto com mais de uma regras nebulosas, então deve ser utilizado um dos métodos a seguir:

- **Inferência Baseada em Composição:** neste método todas as regras da base são combinadas em uma única relação $U \times V$, que posteriormente é vista como uma única regra nebulosa “se-então”. Para realizar essa combinação é necessário, primeiramente, compreender intuitivamente o seu real significado e, assim, utilizar os operadores lógicos de união e interseção para combiná-las. Na união, as regras são vistas como condicionais independentes; e na interseção, as regras são vistas agrupadas e todas as condições têm que ser satisfeitas para que haja uma consequência.

Seja $Regra^l$ uma relação nebulosa em $U \times V$ que representa as regras “Se-então”, isto é, $Regra^l : A_1^l \times \dots \times A_n^l \rightarrow B^l$, com grau de pertinência $\mu_{A_1^l}(x_1) \times \dots \times \mu_{A_n^l}(x_n)$, e a implicação conforme as formas apresentadas na tabela 4.1. Caso seja utilizada a operação de união na composição das regras, então tem-se a *Combinação de Mamdani*, caso contrário, se for utilizada a interseção, então tem-se a Combinação de Gödel, conforme podem ser vistas respectivamente:

$$R_M = \bigcup_{l=1}^r Regra^l \Rightarrow \mu_{R_M}(x, y) = \mu_{Regra^{l1}}(x, y) + \dots + \mu_{Regra^{lr}}(x, y) \quad (18)$$

$$R_G = \bigcap_{l=1}^r Regra^l \Rightarrow \mu_{R_G}(x, y) = \mu_{Regra^{l1}}(x, y) * \dots * \mu_{Regra^{lr}}(x, y) \quad (19)$$

Os graus de pertinência para a saída em ambas as combinações R_M e R_G , ao utilizar a regra de inferência Modus Ponens Generalizada, são obtidos respectivamente da seguinte forma.

$$\mu_{B'}(y) = \sup_{x \in U} t[\mu_{A'}(x), \mu_{R_M}(x, y)] \quad (20)$$

$$\mu_{B'}(y) = \sup_{x \in U} t[\mu_{A'}(x), \mu_{R_G}(x, y)] \quad (21)$$

O procedimento passo a passo para a realização da inferência baseada em composição tem quatro partes, que são: (1) para cada regra nebulosa “se-então”, determinar o seu grau de pertinência, a partir de $\mu_{A'_1}(x_1) \times \dots \times \mu_{A'_n}(x_n)$; (2) visualizar a relação $A'_1 \times \dots \times A'_n$ e B' como proposições nebulosas, PN_1 e PN_2 respectivamente, nas implicações apresentadas na tabela 4.1, e determinar o seu grau de pertinência; (3) determinar $\mu_{R_M}(x, y)$ e $\mu_{R_G}(x, y)$, de acordo com as equações (18) e (19), respectivamente; (4) dada uma entrada A' , o motor de inferência nebuloso fornece uma saída B' , conforme as equações (20) e (21).

- **Inferência Baseada em Regra Individual:** cada regra da base determina a saída de um conjunto nebuloso e a saída final do sistema de inferência é a combinação de todas as saídas individuais. A combinação se dá através da união ou interseção. O procedimento deste método também está dividido em quatro partes, sendo que as duas primeiras são as mesmas do método anteriormente apresentado e as seguintes são: (3) para um dado conjunto nebuloso de entrada A' , computar a saída nebulosa B' para cada uma das regras $Regra^{(i)}$, de acordo com o método Modus Ponens Generalizado; (4) a saída do motor de inferência nebuloso é a combinação de todas as saídas individuais $\{B'_1, \dots, B'_r\}$, através da união ou interseção, da seguinte forma, respectivamente:

$$\mu_{B'}(y) = \mu_{B_1}(y) + \dots + \mu_{B_r}(y) \quad (22)$$

$$\mu_{B'}(y) = \mu_{B_1}(y) * \dots * \mu_{B_r}(y) \quad (23)$$

Além dos métodos mencionados até aqui, existem diversas formas de combiná-los com as varias implicações lógicas e as diferentes formas de união e interseção (Norma-S e Norma-T) já apresentadas. Dessa forma, pode-se levar em consideração três critérios fundamentais para o processo de tomada de decisão para a concepção do motor de inferência:

1. *Intuição*: a escolha deve fazer sentido, pois se um conjunto de regras é dado por um especialista, que acredita que elas são independentes umas das outras, então devem ser combinadas através da união, por exemplo;
2. *Eficiência Computacional*: a escolha deve resultar em uma formula que relacione B' com A' , que é simples de computar;
3. *Propriedades Especiais*: algumas escolhas podem resultar em um motor de inferência que tenha propriedades especiais, e caso sejam desejáveis, então essa é a melhor escolha.

Além da possibilidade de conceber um motor de inferência nebuloso específico para o sistema que esteja sendo construído, outra opção é utilizar um dos motores já propostos e que são tidos como referência na construção de sistemas e controladores nebulosos. Alguns deles estão citados na tabela 4.2, conforme pode ser visto a seguir.

Tabela 4.2 – Exemplos de Motores de Inferência Nebulosos¹²

Motor de Inferência	Função de Pertinência ($\mu_{B'}$)
Produto	$\mu_{B'}(y) = \max_{l=1}^r \left[\sup_{x \in U} \left(\mu_{A'}(x) \prod_{i=1}^n \mu_{A'_i}(x_i) \mu_{B'_l}(y) \right) \right]$
Mínimo	$\mu_{B'}(y) = \max_{l=1}^r \left[\sup_{x \in U} \min \left(\mu_{A'}(x), \mu_{A'_1}(x_1), \dots, \mu_{A'_n}(x_n), \mu_{B'_l}(y) \right) \right]$
Lukasiewicz	$\mu_{B'}(y) = \min_{l=1}^r \left\{ \sup_{x \in U} \min \left[\mu_{A'}(x), 1 - \min_{i=1}^n \left(\mu_{A'_i}(x_i) \right) + \mu_{B'_l}(y) \right] \right\}$
Zadeh	$\mu_{B'}(y) = \min_{l=1}^r \left\{ \sup_{x \in U} \min \left[\mu_{A'}(x), \max \left(\min_{i=1}^n \left(\mu_{A'_i}(x_i), \dots, \mu_{A'_n}(x_n), \mu_{B'_l}(y) \right), 1 - \min_{i=1}^n \left(\mu_{A'_i}(x_i) \right) \right) \right] \right\}$
Dienes-Rescher	$\mu_{B'}(y) = \min_{l=1}^r \left\{ \sup_{x \in U} \min \left[\mu_{A'}(x), \max \left(1 - \min_{i=1}^n \left(\mu_{A'_i}(x_i) \right), \mu_{B'_l}(y) \right) \right] \right\}$

Fonte: (WANG, 1997).

4.3.5.3 Fuzzificador

Os fuzzificadores são definidos como mapeadores de valores reais $x^* \in U$ em conjuntos nebulosos A' em U . Três fatores devem ser considerados na escolha do fuzzificador mais adequado para compor um determinado sistema, que são:

1. o fuzzificador deve considerar a entrada como um valor discreto no ponto x^* , ou seja, o conjunto nebuloso deverá ter uma função de pertinência mais ampla nesse ponto;

¹² Este trabalho não tem a pretensão de apresentar a prova de cada um desses métodos, maiores detalhes podem ser vistos em WANG (1997).

2. caso a entrada do sistema nebuloso esteja corrompida, é desejável que o fuzzificador seja capaz de suprimir os ruídos;
3. o fuzzificador deve ajudar a simplificar a computação que será realizada pelo motor de inferência nebuloso.

Na tabela 4.3, são apresentados os fuzzificadores mais utilizados na construção de sistemas e controladores nebulosos.

Tabela 4.3 – Exemplo de Fuzzificadores

Fuzzificador	Função de Pertinência ($\mu_{A'}$)
Singular (Singleton)	$\mu_{A'}(x) = \begin{cases} 1 & \text{se } x = x^* \\ 0 & \text{caso contrário} \end{cases}$
Gaussiano	$\mu_{A'}(x) = e^{-\left(\frac{x_1 - x_1^*}{a_1}\right)^2 * \dots * \left(\frac{x_n - x_n^*}{a_n}\right)^2}$
Triangular	$\mu_{A'}(x) = \begin{cases} \left(1 - \frac{ x_1 - x_1^* }{b_1}\right) * \dots * \left(1 - \frac{ x_n - x_n^* }{b_n}\right) & \text{se } x_i - x_i^* \leq b_i, i = 1, 2, \dots, n \\ 0 & \text{caso contrário} \end{cases}$

Fonte: (WANG, 1997).

4.3.5.4 Defuzzificador

Os defuzzificadores são definidos como mapeadores de conjuntos nebulosos B' em V , em valores discretos $y^* \in V$. Três critérios devem ser considerados na escolha do defuzzificador mais adequado para compor um determinado sistema, que são:

1. **Plausibilidade:** o ponto y^* deve representar B' da melhor forma possível, como por exemplo, ele deve estar aproximadamente no meio do suporte de B' ou ter um grau maior de pertinência em B' ;

2. **Simplicidade Computacional:** este critério é particularmente importante para controladores nebulosos por operar em tempo real;
3. **Continuidade:** uma pequena mudança no conjunto nebuloso B' não deve resultar em uma grande mudança em y^* .

Na tabela 4.4 a seguir, são apresentados três tipos de defuzzificadores, onde em cada um deles é assumido que o valor para B' é resultado de um dos cinco motores de inferência apresentados na tabela 4.2. Sendo assim, tem-se:

Tabela 4.4 – Exemplo de Defuzzificadores

Defuzzificador	y^*
Centro de Gravidade (<i>Gravity Center</i>)	$y^* = \frac{\int_V y \mu_{B'}(y) dy}{\int_V \mu_{B'}(y) dy}$
Centro Médio (<i>Center Average</i>)	$y^* = \frac{\sum_{l=1}^M y^{-l} w_l}{\sum_{l=1}^M w_l}$
Máximo	$y^* = \frac{\int_{hgt(B')} y dy}{\int_{hgt(B')} dy}, \quad hgt(B') = \left\{ y \in V \mid \mu_{B'}(y) = \sup_{y \in V} \mu_{B'}(y) \right\}$

Fonte: (WANG, 1997).

4.4 ALGORITMOS GENÉTICOS COM INTERAÇÃO FENOTÍPICA: UMA VISÃO GERAL

Compreendidos os conceitos gerais e o funcionamento de Algoritmos Evolucionários no Capítulo III, parte-se, de forma sucinta, a explanação da variação de Algoritmos Genéticos adotada em todas as simulações e testes deste trabalho, os Algoritmos Genéticos com

Interação Fenotípica, como na proposta de LEHRER (2000), adaptada e estendida por BRITO (2004), BRITO et al. (2005)¹³ e TEIXEIRA (2005).

Como citado em outras etapas deste trabalho, novas variáveis, biológicas ou não, vêm sendo acrescentadas à estrutura canônica dos AGs, gerando, segundo definição de (HAUPT et al. 2004) as abordagens híbridas, ou Algoritmos Genéticos Híbridos.

Muitas dessas abordagens – algumas com muito custo e configuração – conseguem atingir melhores resultados em execuções mais eficientes, provavelmente, por conseguirem simular de forma mais real os ambientes biológicos naturais onde a evolução acontece (BRITO, 2004).

Uma delas, de relativo sucesso na área, é a inclusão (implementação) da variável “interação fenotípica”, através do uso da Teoria dos Jogos Evolucionários (SMITH, 1982), como um macro-passo descritor da interação entre os indivíduos (soluções) da população de um Algoritmo Genético.

Antes da apresentação desta estrutura e da proposta formal, é preciso compreender o uso da Teoria dos Jogos como motor de interação e disputa entre os indivíduos, assim como, compreender sua aplicação na natureza: a Teoria dos Jogos Evolucionários.

4.4.1 Teoria dos Jogos e o Jogo Evolucionário Hawk-Dove

Humanos, animais e todos os seres vivos de uma forma geral participam de jogos, de uma simples disputa de xadrez às manobras políticas internacionais, das disputas para conquistar o sexo oposto, às disputas pelo mercado realizadas entre duas grandes empresas.

¹³ Detalhes e publicações relacionadas no Anexo E.

A Teoria dos Jogos pode ser definida como a descrição formal do conflito de interesses entre participantes, em que todos fazem escolhas buscando a melhor estratégia, onde o conjunto de escolhas de todos os participantes determinará os resultados obtidos por cada um (LUCE et al., 1957).

Neste contexto, surge uma classe especial de jogos chamados de Jogos Evolucionários, provenientes dos trabalhos do biólogo John Maynard Smith e G. R. Price (1973), que, de forma independente, redescobriram o conceito do **Equilíbrio de Nash**, chamando-o de Estratégia Evolucionária Estável (*Evolutionary Stable Strategy – ESS*) (SMITH, 1982).

LEHRER (2000) diferencia a Teoria dos Jogos Evolucionários da Teoria dos Jogos Geral, afirmando que, na primeira os jogadores são condicionados biologicamente ou socialmente, sendo escolhidos de forma aleatória dentro da população, já na segunda, o jogo é analisado como se todos os jogadores fossem completamente racionais e conhecessem todas as regras do jogo e todas as preferências dos seus adversários.

SMITH (1982), dentre outros Jogos Evolucionários, propõe o **Jogo Hawk-Dove**, que modela a disputa entre pares de animais por um recurso de valor (**V**). A obtenção deste recurso acarreta no aumento de sua adaptabilidade (*fitness*) de Darwin e a perda não implica em nenhum decréscimo da adaptabilidade que o indivíduo já possui.

O recurso pode estar associado a diversos fatores da natureza como: a obtenção de alimentos, competição sexual, disputa por abrigo, etc. O aumento da adaptabilidade reflete diretamente na quantidade de descendentes esperados, ou seja, o indivíduo que ganhar o recurso poderá gerar uma maior quantidade de descendentes, enquanto que o número esperado para o indivíduo que perdeu a disputa permanece o mesmo (SMITH, 1982).

No Jogo Hawk-Dove, que é o utilizado na abordagem proposta, cada jogador (indivíduo) poderá assumir duas estratégias (comportamentos) distintas: **Hawk**, não

cooperativo, luta pelo recurso até se ferir ou até que o oponente fuja, sendo caracterizado como comportamento agressivo; e *Dove*, sendo o agente cooperativo, dócil e nunca participando de disputas físicas com oponentes, preferindo competir através da exibição, e sempre fugindo ao encontrar um indivíduo agressivo (Hawk).

Baseado nesta caracterização, SMITH (1982) propõe a tabela de pagamentos para este jogo, refletida na tabela 4.5. Esta deve ser lida considerando o jogador a linha, o oponente a coluna e os valores esperados para o jogador, o cruzamento das linhas e colunas. Assim, têm-se as seguintes situações:

- **Hawk x Hawk:** nesta disputa cada participante terá probabilidades iguais de sair ferido e ter sua adaptabilidade reduzida a um custo (C) ou de obter o recurso (V) e aumentar sua adaptabilidade;
- **Hawk x Dove:** o jogador de comportamento Dove foge e mantém sua adaptabilidade inalterada, já o participante obtém o recurso e, tem sua adaptabilidade aumentada em (V);
- **Dove x Dove:** o recurso é compartilhado de forma igualitária entre os dois participantes, dada à característica cooperativa de ambos;

Tabela 4.5 – Tabela de Pagamentos (*payoffs*) do jogo Hawk-Dove.

	Hawk	Dove
Hawk	$1/2 (V-C)$	V
Dove	0	$V/2$

4.4.2 Codificação e Transmissão Genética do Comportamento

Para que os indivíduos possam interagir dentro de um jogo, como demonstrado na seção anterior, é necessário que estes adotem um ou outro comportamento de acordo com certa estratégia.

Estudos recentes demonstram que algumas características comportamentais não são aprendidas durante o ciclo de vida, mas estão ligadas diretamente ao material genético do indivíduo, a exemplo do instinto bem visível em várias espécies (MOTTA, 2000).

A genética tem sido utilizada para explicar comportamentos que vão desde o trabalho na construção e manutenção de colméias por abelhas, até o alcoolismo, epilepsia, psicose maníaco-depressiva e a síndrome do pânico, além de outras fobias que atormentam os humanos (MOTTA, 2000).

Neste aspecto, o estudo da genética do comportamento tem sido conteúdo indispensável a todos os profissionais da área de psicologia e, de grande utilidade nas investigações policiais e julgamentos (MOTTA, 2000). Além disso, possui larga aplicabilidade para descrever as estratégias adotadas pelos animais na interação (cooperação e disputa) pela obtenção de recursos (AXELROD, 1984).

Segundo MOTTA (2000), o estudo dos comportamentos geralmente está associado à ocorrência de vários genes, que, de alguma forma, se correlacionam em fatores somáticos e influem num dado comportamento.

4.4.3 Estrutura dos Algoritmos Genéticos com Interação Fenotípica

Entendendo o funcionamento do Jogo Hawk-Dove, bem como os argumentos da existência de correlação entre comportamentos, estratégias e o conteúdo gênico do seres

vivos, BRITO et al. (2005), modelou uma evolução do que antes havia sido proposto por LEHRER (2000).

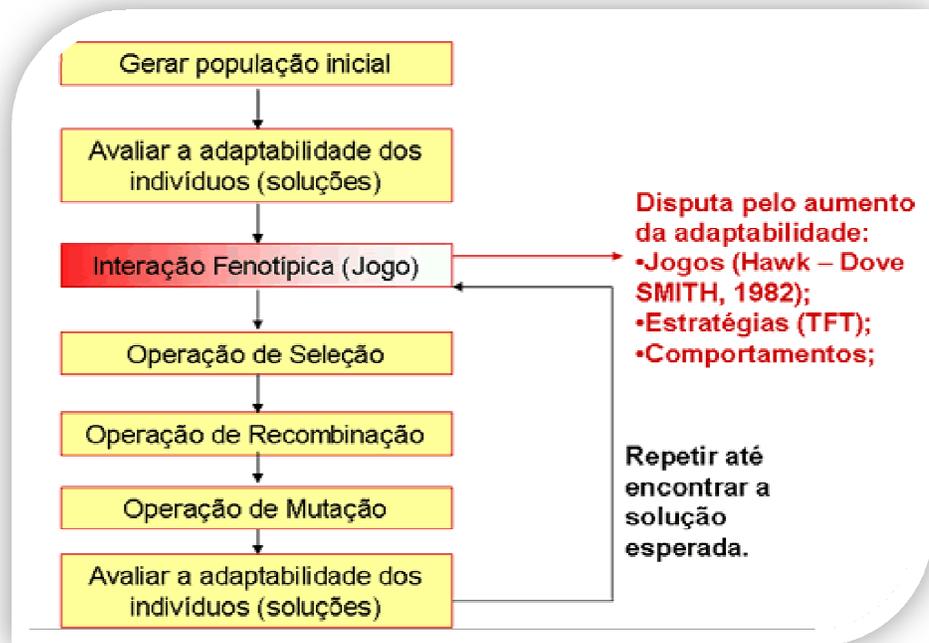
LEHRER (2000), definia indivíduos com estratégias fixas dentro da população e distribuídas aleatoriamente ao longo de toda a execução, além disso apresentava a interação fenotípica como um novo método para o operador de Seleção em Algoritmos Genéticos, o qual chamou de Hawk- Dove Roleta. Em resumo, acrescentou ao tradicional método da roleta (vide Capítulo III), uma fase anterior em que os indivíduos tinham seu valor de *fitness* alterado, dadas uma série de disputas entre eles.

Essa alteração de *fitness* possibilitou um controle maior da pressão seletiva imposta pelo método da roleta – forte característica deste método de seleção. Isto se deve ao fato de o mesmo permitir que indivíduos (soluções) menos adaptados, tivessem maiores oportunidades de recombinar (maior probabilidade na roleta). Desta forma, evitou-se, ainda que de forma insipiente, a convergência prematura e, conseqüentemente, aumentou-se a possibilidade de encontro de um ótimo global para o problema em questão.

Percebendo a comprovação estatística e as vantagens por trás da proposta, BRITO (2004), propôs a criação de um segundo cromossomo destinado a codificação de estratégias e permitiu que essas fossem recombinadas ao longo da execução, tal qual propõe o biólogo Maynard Smith. Além disso, identificando o êxito de suas modificações, em BRITO et al. (2005) e TEIXEIRA (2005), um novo algoritmo genético é apresentado em que a interação fenotípica é tratada como uma etapa independente da execução e não mais de forma atrelada a um ou outro método de seleção.

A nova estrutura é apresentada na figura 4.7 e, como dito anteriormente, é a que será utilizada para o acréscimo do agente inteligente GAIA e, faz parte de todos os testes realizados com a proposta.

Figura 4.8. Estrutura de um AG segundo a abordagem híbrida de (BRITO et al., 2005).



Fonte: (BRITO et al., 2005).

4.5 CONSIDERAÇÕES FINAIS

Neste capítulo foram apresentados todos os principais elementos que fundamentam e inspiram a nova estrutura de Algoritmo Genético proposta.

Foram descritos desde os aspectos puramente inspiratórios como a Teoria do Design Inteligente e a Teoria da Extinção Cíclica de Raup-Sepkoski, até as técnicas utilizadas para a sua materialização dentro do contexto dos algoritmos evolucionários.

Em todas as seções abordadas, ligações prévias foram estabelecidas entre o conceito e aquilo que de fato foi interpretado e implementado, assim, preparou-se o caminho para a capítulo seguinte em que a proposta é efetivamente exposta e relacionada a todos os assuntos discutidos até o momento.

CAPÍTULO V

Ordo ab chao – “Ordem a partir do caos”.

Provérbio Maçônico

5. GAIA: “GENETIC ALGORITHM INFERENCE AGENT”

Batizado por meio do acrônimo em alusão a deusa primordial da mitologia grega, Gaia¹⁴, o chamado “*Genetic Algorithm Inference Agent*” (Agente de Inferência em Algoritmos Genéticos) – ou simplesmente, GAIA – surge inicialmente com a proposta de tratar os dois principais problemas encontrados em técnicas de computação evolucionária: o excesso de convergência ou convergência prematura, e a árdua tarefa de definição e ajuste paramétrico, a qual sempre possuirá valores ideais diferentes e desconhecidos para cada tipo de problema.

Estes dois problemas, assim como a técnica de Algoritmos Genéticos, onde será acoplado o agente, foram exaustivamente descritos no terceiro capítulo deste trabalho. Entretanto, vale também ressaltar a dificuldade que pesquisadores não relacionados à “área meio” (computação evolucionária), mas sim à “área fim” (domínio do problema que se quer resolver), encontram para realizar suas buscas, avaliá-las e entender o porquê de nem sempre

¹⁴ *Gaia, Géia, Gea* ou *Gê* era a deusa da Terra, a Mãe Terra, uma Deusa considerada primordial pelos gregos, com notória fertilidade e capacidade geradora. Descendia do Deus primordial Caos, e era avó de Zeus, que, após inúmeras batalhas, viria a ser o Deus principal da mitologia grega, e também da romana com o nome de Júpiter (UFCEG, 2011).

apresentarem resultados significativos, tal qual sugere a idéia da proposta original desta técnica de John Holland.

Ao longo da história das principais técnicas de computação evolucionária, até os dias de hoje, busca-se, geralmente, resolver estes dois problemas por meio do acréscimo de variáveis identificadas no meio natural, partindo da máxima de que quanto mais próximos do mundo “real”, melhores serão os resultados encontrados, e mais eficiente tornar-se-á a técnica utilizada.

Propostas deste tipo, em especial as que visam contornar o problema da convergência, são muito comuns na literatura, algumas como a utilizada de apoio neste trabalho, os Algoritmos Genéticos com Interação Fenotípica (BRITO et. al., 2005, vide Capítulo IV), apresentam relativo sucesso, uma vez que reduzem a pressão seletiva e, conseqüentemente, retardam o processo de convergência prematura e estagnação em um ótimo local. Já outras, chegam a ser exageradas, quando descrevem estruturas naturais incertas e muito mais complexas do que o próprio problema, que se quer resolver por meio da técnica.

Todas essas técnicas, segundo deduções do *No Free Lunch Theorem* (Capítulo IV), não passam de buscas cegas que, por mais que apresentem um melhor resultado para uma classe específica de problemas, tendem de forma compensatória a refletir resultados ainda piores para as demais classes.

Segundo Dembski (2007) e outros teóricos do Design Inteligente (Capítulo IV), a única maneira de efetivamente resolver este problema de eficácia em técnicas de computação evolucionária é adicionar conhecimento sobre o espaço de busca ao algoritmo e, ainda, segundo suas proposições, espelhar-se em outro modelo natural, diferente do darwinismo, o qual defenda intervenções “inteligentes” durante o processo de busca.

Dada esta problemática dupla dos algoritmos genéticos (variabilidade/eficiência e definição paramétrica) e, inspirado pelos conceitos de Design Inteligente complementados

pela Teoria da Extinção Cíclica de Raup-Sepkoski, é que o presente trabalho, modelou e desenvolveu um agente inteligente¹⁵, o qual, acoplado a uma execução de algoritmo genético, monitora certas variáveis e, de forma automatizada, realiza intervenções visando garantir a eficácia e eficiência da técnica de busca.

Sendo assim, conforme pode ser visto na figura 5.1, é introduzida uma nova estrutura para Algoritmo Genético, na qual, ao término de um ciclo, ocorre a intervenção deste agente inteligente proposto, o GAIA.

Figura 5.1 – Nova Estrutura de Algoritmo Genético com o GAIA.



Na representação do conhecimento necessário ao agente para suas intervenções, optou-se pelo uso da Lógica Difusa ou Fuzzy, apresentada no quarto capítulo. Esta escolha se deu pelas características peculiares desta lógica de aproximação dos modelos matemáticos ao

¹⁵ Por agente inteligente, considera-se aqui o conceito de EBERHART (1996), como sendo estruturas autônomas dotadas de inteligência codificada por meio de modelos e representações lógicas-matemáticas, as quais percebem o ambiente e tomam decisões refletidas em ações ou interação com o meio.

raciocínio humano, por meio dos conceitos de incerteza, representação de variáveis lingüísticas e base de regras com motor de inferência de fácil compreensão.

O GAIA, intencionalmente, da forma como foi concebido, pode ser considerado como uma “caixa-preta” do ponto de vista de um Algoritmo Genético tradicional. Isto porque não altera métodos de seleção, recombinação, mutação, nem mesmo as estruturas de codificação de indivíduos e da função de avaliação do problema.

O agente é uma estrutura à parte e independente, que será “aberta” nas próximas seções do capítulo, sendo explorados todos os aspectos de codificação da proposta, desde a escolha das variáveis até a construção do motor de inferência por meio da base de regras, assim como a forma e o período de intervenção do agente.

5.1 VARIÁVEIS: O QUE MONITORAR E CONTROLAR?

Conforme o exposto anterior, o GAIA é o responsável pela avaliação de certas variáveis da execução do algoritmo genético – ou seja, monitoramento –, assim como também, por meio de sua base de regras e motor de inferência, deve realizar ações que impactem diretamente em fatores desta execução – o controle.

Todavia, devido ao já apresentado problema do universo de variáveis observáveis em execuções de AGs, assim como, da vasta possibilidade de regulação dos muitos parâmetros existentes, surge a pergunta: “o que monitorar e o que controlar?”.

Na busca desta resposta foram consideradas e descartadas diversas possibilidades de variáveis, tanto para a entrada como para a saída. Estas considerações foram ponderadas principalmente sobre duas óticas: primeiro, pelos estudos de DEJONG (1975) e pela síntese de trabalhos apresentada em HAUPT (2004) sobre controle e variação paramétrica, ambos citados na seção 3.4; e segundo, de forma empírica, dadas às observações provenientes da

vasta experiência do autor e de seu grupo de pesquisa na solução de problemas de otimização por meio de técnicas de computação evolucionária.

Primeiramente, foram analisadas as possíveis variáveis de saída, representando o que se quer que seja controlado pelo agente. Neste sentido, foram avaliadas as seguintes variáveis: percentual de cruzamento, percentual de mutação, percentual de sobreviventes, o número de gerações que devem ser executadas e o tamanho da população. Devido à dificuldade de “fuzzificação” das duas últimas, dada à falta de estudos e detalhes sobre o assunto, além de serem consideradas de menor importância frente as demais na consecução dos objetivos estabelecidos pela proposta, foram efetivamente adotadas como variáveis de saída (controle) as seguintes: o **percentual de mutação**, o **percentual de cruzamento** e o **percentual de sobreviventes**.

Sendo assim, partiu-se para a próxima etapa de definição de quais observações seriam realizadas, a fim de que, pudessem ser refletidos os valores das variáveis de saída definidas.

Neste sentido, foram avaliadas como possíveis variáveis de entrada (monitoramento): a média da população, o desvio padrão, o valor ótimo (melhor fitness) da geração, frequência do melhor indivíduo ao longo das últimas gerações, percentual de indivíduos repetidos na geração e a distância em relação a um valor ótimo suposto. Após considerar a utilidade e possibilidade de fuzzificação de cada variável, e tendo em vista os objetivos de otimização e evolução, preservando a variabilidade, permaneceram como variáveis de entrada: o **valor ótimo na geração**, a **frequência do melhor indivíduo**, e o **percentual de indivíduos repetidos na geração**.

O passo seguinte, foi a definição do domínio de cada uma das variáveis, seu particionamento em termos de valores fuzzy, seus tipos de função de pertinência e os mecanismos fuzzy utilizados para a materialização destes valores na entrada e saída no sistema. Nesta última etapa foram utilizados o fuzzificador *singleton*, o defuzzificador média

ponderada, com operador de inferência mínimo e composição de Mamdani (MaxMin) (WANG, 1997).

A figura 5.2, representa uma versão estendida da “caixa preta” do GAIA, incluindo agora as variáveis de entrada (monitoração) e as de saída (controle), conforme definido. Nas subseções seguintes, será discutidos o conceito, o particionamento destas variáveis e seus detalhes técnicos. Ao final do capítulo, serão apresentados o motor de inferência que relaciona todas as variáveis e os mecanismos de intervenção adotados.

Figura 5.2 – Funcionamento geral do GAIA.



5.1.2 Variáveis de Entrada

5.1.2.1 Valor Ótimo

Passível de críticas e alguns questionamentos, como por exemplo, o de definir anteriormente algo que se quer encontrar e de que ainda não se tem idéia, a variável “Valor Ótimo” busca avaliar o quão bom ou próximo do resultado ideal está a execução.

Esta definição ampara-se nos já explicitados argumentos de DEMBKI (2007) sobre o *No Free Lunch Theorem*, o qual requer o acréscimo de algum conhecimento do domínio do espaço de busca para que o algoritmo seja de fato eficaz e não apenas uma busca baseada em princípios exclusivamente casualísticos.

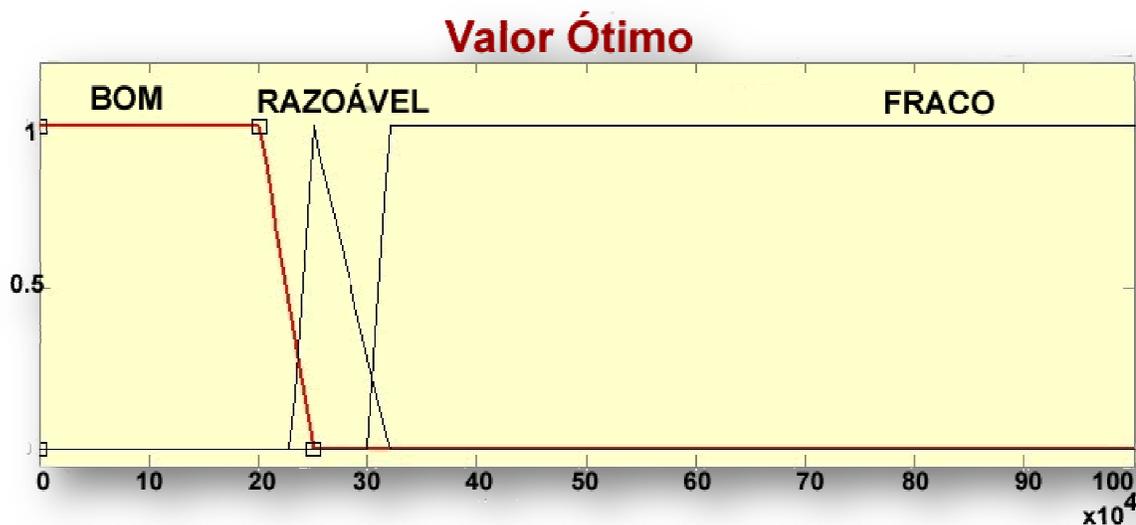
A crítica do “pré-conhecimento”, exposta no início da seção, pode ser resolvida por meio do constante ajuste da meta do valor ótimo ao longo da execução. Uma abordagem que o próprio DEMBKI (2007) utiliza ao citar que, nem sempre o designer conhece ou precisa conhecer exatamente onde se quer chegar, e que esse valor pode e deve ser adaptado ao longo do tempo.

Devido ao fato do problema utilizado como teste¹⁶ ser bastante conhecido pelo autor e seu grupo de pesquisa, além de bem retratado na literatura, optou-se por codificar de forma fixa e constante o domínio desta variável no intervalo $[0, 10^5]$ km, com o seu particionamento em três partes, baseado nos resultados encontrados por LEHRER (2000), BRITO (2004), TEIXEIRA (2005) e BRITO et. al. (2005).

São estas partes, os seguintes valores, descritos também de forma gráfica na figura 5.3: “**Bom**”, função de pertinência trapezóide, valores $[0, 0, 20113, 25000]$; “**Razoável**”, função de pertinência trapezóide, valores $[23000, 25000, 32000]$; e, por fim, “**Fracó**”, função de pertinência trapezóide, valores $[30000, 32000, 100000, 100000]$.

¹⁶ PVCS-BR-26 de LEHRER (1999) é o problema, discutido em detalhes no capítulo VI.

Figura 5.3 – Variável de Entrada – “Valor Ótimo”.



De todas as variáveis apresentadas, essa é a única que está atrelada diretamente ao problema o qual se quer resolver e, considerando que o problema sempre fará parte do domínio de conhecimento do pesquisador que aplica a técnica de computação evolucionária, não são identificadas dificuldades em que este a defina na primeira vez da execução, ou ainda, que permita o seu ajuste automático, se assim o desejar.

5.1.2.2 Frequência do Melhor Resultado (ou Indivíduo)

Esta variável está relacionada diretamente com a variabilidade que a execução está mantendo, fator fundamental para evitar a estagnação em ótimos locais e a efetiva busca do ótimo global.

Para a compreensão do grau de sua importância dentro da estrutura de um algoritmo genético, pode-se fazer analogia ao universo e a Lei da Gravitação Universal de Sir Isaac Newton. Nesta lei, todos os corpos de forma proporcional à sua massa exercem atração entre si, fazendo com que corpos de “grande” massa acabem por aglomerar uma significativa

quantidade de matéria ao seu redor, prendendo-a naquela área do espaço. Nos AGs, de forma análoga, o mesmo fato se repete, pois durante a exploração do espaço pela execução, a constante aparição do mesmo valor ótimo ao longo das últimas gerações indica uma substancial força de atração para uma determinada área deste espaço de busca, sinal claro de que em poucas gerações, todos os valores irão convergir para aquele ponto, ou, por conta do operador de mutação, ficarão “orbitando” nas suas proximidades, estagnando a execução no que pode não ser o ótimo global de solução do problema.

O domínio desta variável foi considerado em termos percentuais, ou seja, de [0,100], observando sempre a frequência em que o melhor valor da geração corrente apareceu nas últimas cinquenta¹⁷(50) execuções.

O particionamento e codificação de valores foi feito de forma uniforme, distribuídos em três partes, descritos também graficamente na figura 5.4, são elas: “**Baixa**”, função de pertinência trapezoidal, valores [0, 0, 25, 50] ; “**Média**”, função de pertinência triangular, valores [25, 50, 75]; e, por fim, “**Alta**”, função de pertinência trapezoidal, valores [50, 75, 100, 100].

¹⁷ Valor arbitrado de forma empírica, frente ao número de execuções que geralmente se estabelece para o problema avaliado, o de 5.000 gerações.

Figura 5.4 – Variável de Entrada – “Frequência do Melhor Resultado”



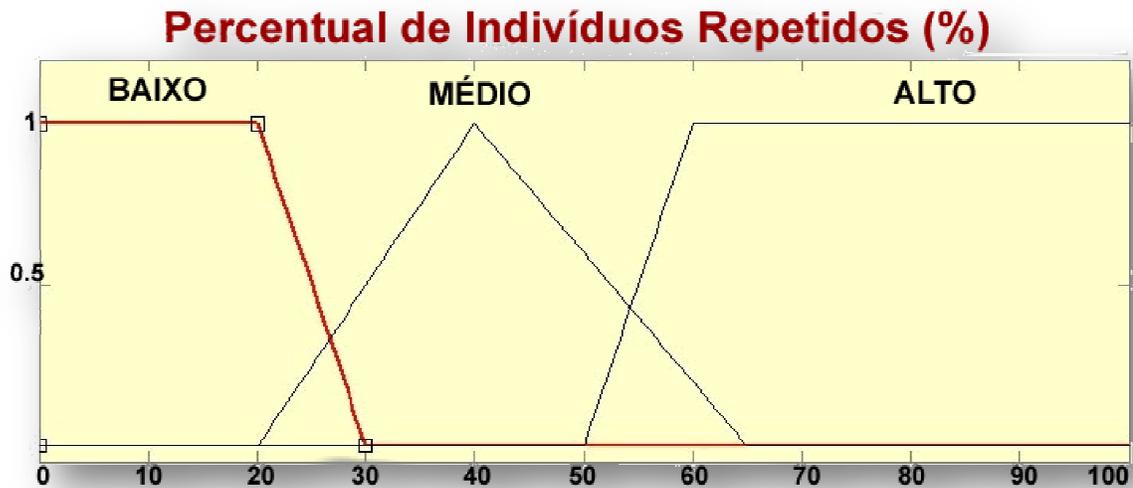
5.1.2.3 Percentual de Indivíduos Repetidos

Assim como a variável anterior, a análise do Percentual de Indivíduos Repetidos, é de substancial importância, pois permite avaliar o grau de variabilidade dentro da população, sinal de alerta para o possível excesso de convergência e prisão em algum ótimo local do espaço de busca.

O domínio desta variável também foi considerado em termos percentuais, ou seja, de $[0,100]$, observando o percentual de indivíduos com o mesmo genótipo na geração corrente. Já o particionamento e representação de valores não foi feito de forma uniforme, pois se considera que qualquer percentual acima de 50% de indivíduos iguais dentro de uma mesma geração, é reflexo claro de estagnação, o que irá requerer uma forte ação do agente, como será descrita na próxima seção.

Sendo assim, conforme apresentado na figura 5.5, dividiu-se o domínio nos seguintes valores: “**Baixo**”, função de pertinência trapezoidal, valores $[0, 0, 20, 30]$; “**Médio**”, função de pertinência triangular, valores $[20, 40, 65]$; e, por fim, “**Alto**”, função de pertinência trapezoidal, valores $[50, 60, 100, 100]$.

Figura 5.5 – Variável de Entrada – “Percentual de Indivíduos Repetidos”



5.1.3 Variáveis de Saída

5.1.3.1 Percentual de Mutação

A variável de saída “Percentual de Mutação” diz respeito ao ajuste dinâmico do parâmetro homônimo de um Algoritmo Genético. Portanto, visa solucionar o problema da escolha manual deste, assim como sua necessidade de adaptação de valor, dadas às características da população em um determinado momento da execução.

A operação de mutação, conforme exposto no terceiro capítulo, é considerada pela maioria dos autores como um “ajuste fino” na busca pela solução ótima dentro do espaço de busca, ou seja, suas pequenas alterações levam a exploração de soluções próximas ao ponto principal em que a população se encontra.

Sendo assim, o exagero no uso deste operador, como defendido por Dembski (2004), é uma das principais formas de voltar o algoritmo de busca à total casualidade. Fator que foi considerado no presente trabalho que, como será visto na base de regras na próxima seção,

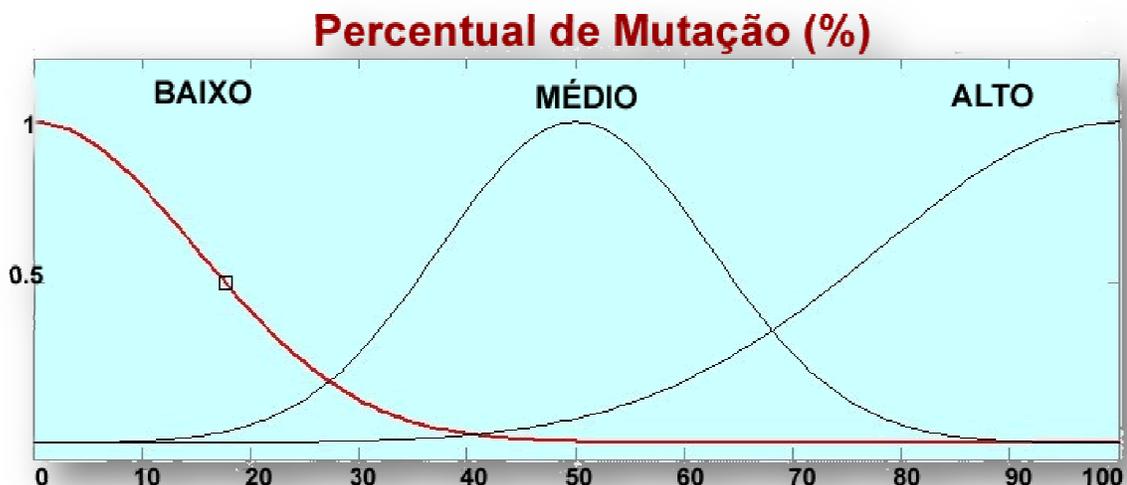
apenas se utiliza de altos níveis deste operador quando quer forçar a saída da população de um ótimo local, ou seja, não faz o seu excessivo uso como fator principal da busca.

O domínio desta variável também foi considerado em termos percentuais, ou seja, de [0,100], reflexo do real valor que o parâmetro do AG deverá assumir a partir da intervenção do GAIA.

O particionamento e representação de valores foi feito em três partes, por meio do uso de funções de pertinência gaussianas, considerando a característica de que estas tendem a suavizar as fronteiras de transição entre os valores (WANG, 1997).

A figura 5.6 descreve esta modelagem, onde dividiu-se o domínio nos valores: “**Baixo**”, função de pertinência gaussiana (direita), valores [centro=0, amplitude=15]; “**Médio**”, função de pertinência gaussiana, valores [centro=50, amplitude=12.5]; e, por fim, “**Alto**”, função de pertinência gaussiana (esquerda), valores [centro=100, amplitude=22].

Figura 5.6 – Variável de Saída – “Percentual de Mutação”



5.1.3.2 Percentual de Recombinação

A variável de saída “Percentual de Recombinação” diz também respeito ao ajuste dinâmico de um parâmetro homônimo do Algoritmo Genético. Sendo assim, também é mais um esforço do trabalho no sentido de solucionar o problema da escolha paramétrica, além de assegurar a adaptabilidade do valor definido conforme às características da população em um determinado momento da execução.

A operação de recombinação é a forma como um AG junta os “pedaços” possivelmente ótimos da solução em partes cada vez maiores, as quais são propagadas, com o auxílio do operador de seleção, ao longo das gerações até que se encontre a solução definitiva para o problema. A esses pedaços, dá-se o nome de “blocos construtores” (*building blocks*, MITCHELL, 1999), característica fundamental e matematicamente comprovada da eficácia do motor dos algoritmos genéticos¹⁸.

Sendo assim, caso haja uma boa variabilidade, é ideal que este percentual permaneça alto para forçar a combinação destes possíveis ótimos pedaços. Apenas em situações extremas de estagnação da população o agente irá reduzir este valor em detrimento de um substancial aumento da mutação, para que a execução seja retirada daquele ponto do espaço.

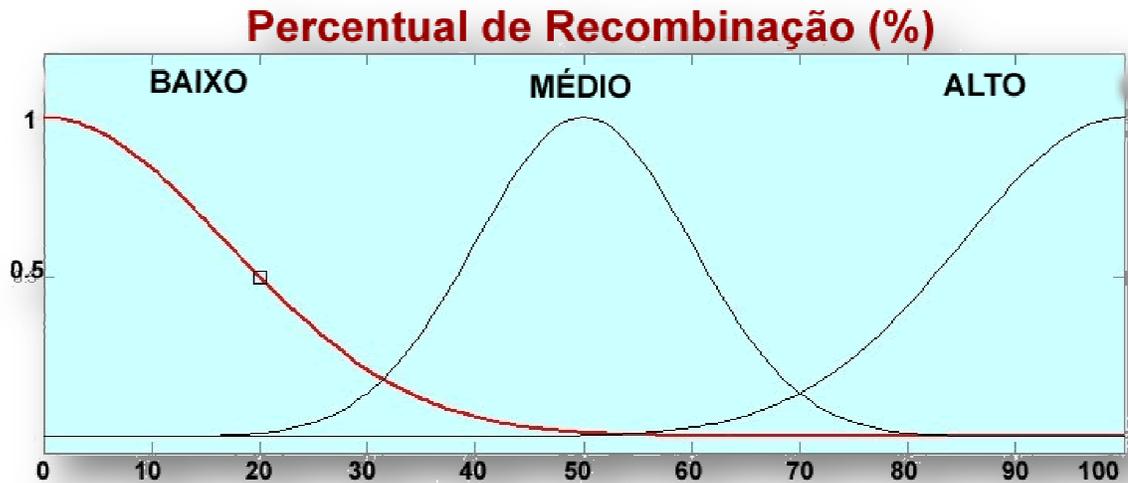
O domínio desta variável também foi considerado em termos percentuais, ou seja, de [0,100], reflexo do real valor que o parâmetro do AG deverá assumir a partir da intervenção do GAIA.

¹⁸ Esta observação é destinada somente para Algoritmos Genéticos que envolvem combinação (representação binária, por seqüência, por codificação fixa, etc.), não sendo estendida para AGs com codificação baseada em números contínuos (reais).

O particionamento e representação de valores foi feito em três partes, por meio do uso de funções de pertinência gaussianas, considerando a característica de que estas tendem a suavizar as fronteiras de transição entre os valores (WANG, 1997).

A figura 5.7 descreve esta modelagem, onde dividiu-se o domínio nos valores: “**Baixo**”, função de pertinência gaussiana (direita), valores [centro=0, amplitude=17]; “**Médio**”, função de pertinência gaussiana, valores [centro=50, amplitude=10]; e, por fim, “**Alto**”, função de pertinência gaussiana (esquerda), valores [centro=100, amplitude=15].

Figura 5.7 – Variável de Saída – “Percentual de Recombinação”



5.1.3.3 Percentual de Sobreviventes

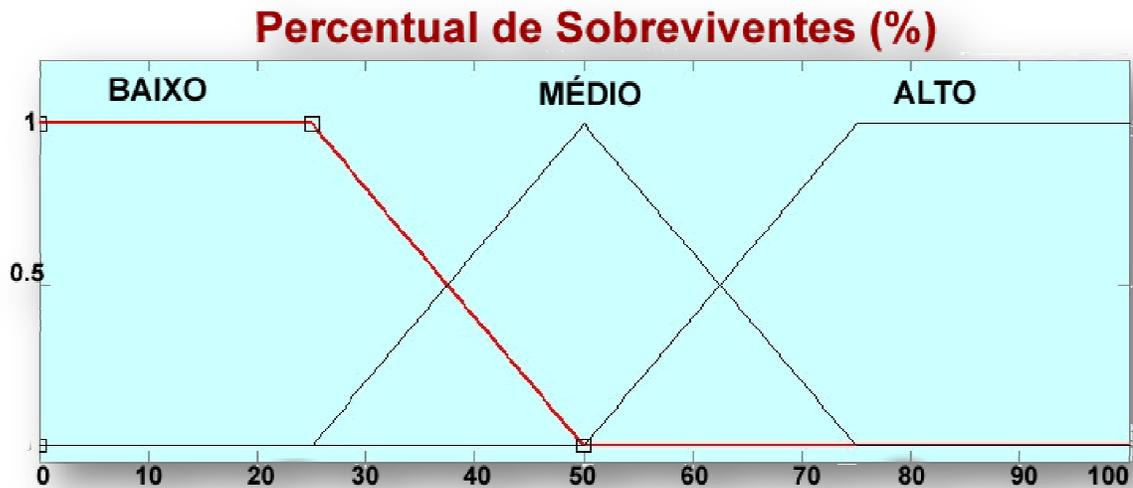
Por fim, mas não menos importante, pelo contrário, talvez uma das mais substanciais formas de controle implementadas, a variável “Percentual de Sobreviventes”. Nesta, considerando teorias como a da Extinção Cíclica de Raup-Sepkoski e suas variações, busca-se eliminar de forma proporcional à estagnação da execução, uma boa parcela de indivíduos, substituindo-os automaticamente por outros gerados de forma aleatória como se fizessem parte de uma nova população inicial.

Esta variável é espelhada na situação das grandes catástrofes que ocorrem na natureza e em suas conseqüências para as formas de vida. Quando ocorrem, dependendo de sua magnitude, estes eventos tendem a eliminar boa parte da população, preservando geralmente os poucos, porém mais aptos, indivíduos. Também, deixam significativas conseqüências negativas por um longo período para aqueles grupos sobreviventes, como por exemplo, os altos níveis de mutação relacionados àquelas catástrofes que de alguma forma alteraram o campo magnético do planeta (MOURÃO, 2002). Em razão disso, novas espécies nem sempre mais aptas surgem, entretanto, são estas quem podem introduzir os “blocos construtores” chave para o surgimento de novas e melhores espécies.

Sendo assim, em casos de estagnação, ou seja, de baixa variabilidade, tal como exposto no parágrafo anterior, a intervenção do agente irá eliminar uma parcela bastante significativa da população, abrindo caminho para a recuperação dos “blocos construtores perdidos”, por meio da geração e inclusão de novas soluções aleatórias.

O domínio desta variável foi definido em termos percentuais, ou seja, de [0,100], particionado em três pedaços iguais, segundo os valores apresentados na figura 5.8, descritos a seguir: “**Baixo**”, função de pertinência trapezóide, valores [0, 0, 25, 50] ; “**Médio**”, função de pertinência triangular, valores [25, 50, 75]; e, por fim, “**Alto**”, função de pertinência trapezóide, valores [50, 75, 100, 100].

Figura 5.8 – Variável de Saída – “Percentual de Sobreviventes”



5.2 BASE DE REGRAS E O MOTOR DE INFERÊNCIA FUZZY

Compreendidas as formas de diagnóstico e as possíveis intervenções do GAIA, é preciso entender agora como combiná-las por meio da associação entre as variáveis de entrada e de saída apresentadas na seção anterior. Como todo sistema fuzzy, essa associação é feita por meio de uma base de regras, que mapeia possíveis valores de entrada em resultados das variáveis de saída.

Assim como descrito no início do capítulo para o processo de escolha das variáveis de monitoramento e controle do agente, na construção da base de regras, buscou-se representar o conhecimento e experiência de pesquisadores da área. Para tal, apoiou-se nas considerações de controle e variação paramétrica dos estudos de DEJONG (1975), na síntese de inúmeros trabalhos da área apresentada em HAUPT (2004), naquilo que se considera senso comum entre os estudiosos de AGs, nos estudos do Design Inteligente, e, também, no empirismo, dadas às observações provenientes da vasta experiência do autor e de seu grupo de pesquisa

na solução de problemas de otimização através do uso de técnicas de computação evolucionária.

Os resultados deste “sincretismo” de propostas e de entendimentos dos possíveis estados de não performance de um AG, assim como também da forma de contornar estes problemas, são sintetizados na base de regras do GAIA, descrita na tabela 5.1, exposta a seguir. Visando facilitar a leitura, e considerando que o objetivo principal de um AG é resolver o problema, ou seja, buscar o valor ótimo, as regras estão descritas agrupadas em função da variável “Valor Ótimo”.

Tabela 5.1 – Base de Regras Fuzzy - GAIA

BASE DE REGRAS – GAIA							
ID	ENTRADAS			SAÍDAS			ID
	VALOR ÓTIMO (VO)	FREQ. MELHOR INDIV. (FMI)	PERCENTUAL REPETIDOS (PR)	PERCENTUAL MUTAÇÃO (PM)	PERCENTUAL CRUZAMENTO (PC)	PERCENTUAL SOBREVIVENTES (PS)	
1	Ótimo	Baixa	Baixo	Baixo	Médio	Alto	1
2	Ótimo	Baixa	Médio	Médio	Médio	Médio	2
3	Ótimo	Baixa	Alto	Alto	Médio	Médio	3
4	Ótimo	Média	Baixo	Alto	Médio	Alto	4
5	Ótimo	Média	Médio	Alto	Médio	Médio	5
6	Ótimo	Média	Alto	Alto	Médio	Baixo	6
7	Ótimo	Alta	Baixo	Alto	Médio	Médio	7
8	Ótimo	Alta	Médio	Alto	Médio	Baixo	8
9	Ótimo	Alta	Alto	Alto	Baixo	Baixo	9
10	Razoável	Baixa	Baixo	Baixo	Alto	Alto	10
11	Razoável	Baixa	Médio	Médio	Alto	Alto	11
12	Razoável	Baixa	Alto	Alto	Alto	Médio	12
13	Razoável	Média	Baixo	Médio	Alto	Alto	13
14	Razoável	Média	Médio	Médio	Médio	Médio	14
15	Razoável	Média	Alto	Alto	Médio	Médio	15
16	Razoável	Alta	Baixo	Médio	Médio	Médio	16
17	Razoável	Alta	Médio	Alto	Médio	Baixo	17
18	Razoável	Alta	Alto	Alto	Baixo	Baixo	18
19	Fraco	Baixa	Baixo	Baixo	Alto	Alto	19

BASE DE REGRAS – GAIA							
ID	ENTRADAS			SAÍDAS			ID
	VALOR ÓTIMO (VO)	FREQ. MELHOR INDIV. (FMI)	PERCENTUAL REPETIDOS (PR)	PERCENTUAL MUTAÇÃO (PM)	PERCENTUAL CRUZAMENTO (PC)	PERCENTUAL SOBREVIVENTES (PS)	
20	Fraco	Baixa	Médio	Médio	Alto	Alto	20
21	Fraco	Baixa	Alto	Alto	Alto	Médio	21
22	Fraco	Média	Baixo	Baixo	Alto	Médio	22
23	Fraco	Média	Médio	Médio	Alto	Médio	23
24	Fraco	Média	Alto	Alto	Alto	Baixo	24
25	Fraco	Alta	Baixo	Médio	Alto	Baixo	25
26	Fraco	Alta	Médio	Alto	Alto	Baixo	26
27	Fraco	Alta	Alto	Alto	Alto	Baixo	27

Apresentada a base de regras, intencionalmente separada por linhas pontilhadas diferentes a cada “grupo” de situação da execução, faz-se ainda necessária uma descrição detalhada de cada uma destas 27 regras. A listagem seguinte as interpreta, considerando a situação em que a execução está - por meio da análise das variáveis de entrada – e qual ação deverá ser tomada durante a intervenção do agente - representada nas variáveis de saída.

Seguem abaixo, todas as regras relacionadas por número de identificação e pelos valores de entrada (vide Tabela 5.1), com suas descrições dispostas de forma incremental:

- 1. Ótimo (VO), Baixa (FMI), Baixo (PR):** Descreve uma situação na qual a execução está evoluindo sem estagnação, ou seja com variabilidade, e de que acabou de encontrar um “bom” valor ótimo, ainda não replicado em toda a população. Sendo assim, deve-se manter o percentual de mutação em nível baixo para facilitar a descoberta de valores ainda melhores nas proximidades, um percentual ponderado de recombinação e um alto percentual de sobreviventes, pois há variabilidade e possíveis blocos construtores que ainda estão espalhados pela população.

2. **Ótimo (VO), Baixa (FMI), Médio (PR):** Também apresenta uma situação de evolução em direção ao valor ótimo, entretanto, o ótimo atual começa a se espalhar dentro da população, comprometendo a variabilidade. Assim, aumenta-se a mutação e diminui-se o percentual de indivíduos sobreviventes.
3. **Ótimo (VO), Baixa (FMI), Alto (PR):** Nesta situação, apesar de recém-encontrado, o valor ótimo se espalhou muito rapidamente dentro da população, deixando-a pronta para a estagnação em um ótimo local. Portanto, de forma garantir a variabilidade, aumenta-se a mutação e diminui-se o percentual de indivíduos sobreviventes.
4. **Ótimo (VO), Média (FMI), Baixo (PR):** Ainda está evoluindo, mas vem repetindo razoavelmente o melhor valor, apesar de ainda manter uma alta variabilidade na população. Sendo assim, aumenta-se a mutação, põe-se uma taxa média de recombinação (evitar a dominância) e mantém-se a população, evitando assim, que seja destruído um bloco construtor ainda não recombinado.
5. **Ótimo (VO), Média (FMI), Médio (PR):** Ainda evoluindo, mas vem repetindo consideravelmente o melhor valor, com a variabilidade apresentando sinais de comprometimento. Desta forma, geram-se alguns indivíduos novos, aumenta-se a mutação e estabelece-se uma taxa média de recombinação, para evitar o rápido domínio dos melhores sobreviventes.
6. **Ótimo (VO), Média (FMI), Alto (PR):** Ainda evoluindo, mas vem repetindo razoavelmente o melhor valor, com a variabilidade atual completamente comprometida. Sendo assim, geram-se muitos indivíduos novos, aumenta-se a mutação, e estabelece-se uma taxa média de recombinação, para evitar o rápido domínio dos melhores sobreviventes.

7. **Ótimo (VO), Alta (FMI), Baixo (PR):** Parou de evoluir, apesar da alta variabilidade de indivíduos na geração corrente, o valor ótimo está saturado. Sendo assim, faz-se necessária a renovação de boa parte da população, com um percentual médio de sobreviventes, uma taxa média de recombinação e uma alta taxa de mutação.

8. **Ótimo (VO), Alta (FMI), Médio (PR):** Parou de evoluir, além de que a variabilidade local começa também a ficar comprometida. Sendo assim, faz-se necessária a renovação da maior parte da população, com um percentual baixo de sobreviventes, uma taxa média de recombinação e uma alta taxa de mutação.

9. **Ótimo (VO), Alta (FMI), Alto (PR):** Parou de evoluir e a variabilidade local foi completamente comprometida, com sinais claros de saturação do valor ótimo. Desta forma, matam-se muitos indivíduos, com baixo cruzamento para evitar a dominação dos "ótimos" remanescentes e alta taxa de mutação para garantir a fuga deste ponto do espaço de busca.

10. **Razoável (VO), Baixa (FMI), Baixo (PR):** Valor ótimo razoável, com alta variabilidade, garantindo assim, evidências de evolução constante em direção a um ótimo global/local. Mantém-se a população, com alta taxa de recombinação, e baixa de mutação, para que o operador de seleção possa continuar fazendo a exploração do espaço de busca.

11. **Razoável (VO), Baixa (FMI), Médio (PR):** Valor ótimo razoável, com indícios iniciais de convergência, ou seja, há evolução, mas algumas medidas de controle devem ser tomadas. Sendo assim, mantém-se a população, impõe-se uma alta taxa de recombinação, assim como uma taxa de mutação média para impedir o fim da variabilidade, e assegurar o pleno funcionamento do operador de seleção na exploração do espaço de busca.

12. **Razoável (VO), Baixa (FMI), Alto (PR):** Apesar do valor ótimo razoável, a execução acaba de ficar presa em um ótimo local. Sendo assim, elimina-se parte dos indivíduos da população, impõe-se uma alta taxa de recombinação e uma alta taxa de mutação, visando o retorno da variabilidade e da evolução pela busca do ótimo global.

13. **Razoável (VO), Média (FMI), Baixo (PR):** Valor ótimo razoável, porém a execução apresenta indícios iniciais de convergência, ou seja, há evolução, mas algumas medidas de controle devem ser tomadas. Sendo assim, mantém-se a população, impõe-se uma alta taxa de recombinação e estabelece-se uma taxa de mutação média para impedir o fim da variabilidade, e assegurar o pleno funcionamento do operador de seleção na exploração do espaço de busca.

14. **Razoável (VO), Média (FMI), Médio (PR):** Valor ótimo razoável, porém com indícios claros de aproximação da situação de “prisão” em um ótimo local. Sendo assim, substitui-se boa parte população, estabelece-se uma taxa de recombinação média para evitar a súbita eliminação das novas soluções inseridas, assim como, uma taxa de mutação média para reforçar o retorno da variabilidade e da busca pelo ótimo global.

15. **Razoável (VO), Média (FMI), Alto (PR):** Prisão em um valor ótimo local e distante do global. Sendo assim, são necessárias fortes medidas de controle. Substitui-se boa parte população, estabelece-se uma taxa de recombinação média para evitar a súbita eliminação das novas soluções inseridas, assim como, uma elevada taxa de mutação para reforçar o retorno da variabilidade e da busca pelo ótimo global.

16. **Razoável (VO), Alta (FMI), Baixo (PR):** Valor ótimo razoável e muito saturado, embora ainda com alta variabilidade na geração. Sendo assim, substitui-se boa parte população, estabelece-se uma taxa de recombinação média para evitar a súbita

eliminação das novas soluções inseridas, assim como, uma taxa de mutação média para reforçar o retorno da variabilidade e da busca pelo ótimo global.

17. Razoável (VO), Alta (FMI), Médio (PR): Indício de estagnação em ótimo local e comprometimento da execução pelo fim da variabilidade. Sendo assim, geram-se muitos indivíduos novos, aumenta-se a taxa de mutação, e estabelece-se uma taxa média de recombinação, para evitar o rápido domínio dos melhores sobreviventes.

18. Razoável (VO), Alta (FMI), Alto (PR): Estagnação (“prisão”) em ótimo local, com o total comprometimento da execução pelo fim da variabilidade. Sendo assim, geram-se muitos indivíduos novos, eleva-se substancialmente a taxa de mutação, e estabelece-se uma taxa baixa de recombinação, para evitar o rápido domínio dos melhores sobreviventes e permitir às novas soluções um período de evolução e afirmação dentro da população.

19. Fraco (VO), Baixa (FMI), Baixo (PR): Início de execução. Nesta situação o que interessa para a evolução é uma alta taxa de recombinação, uma baixa taxa de mutação e a manutenção dos indivíduos, para que o operador de seleção possa fazer a exploração do espaço de busca.

20. Fraco (VO), Baixa (FMI), Médio (PR): Início de execução, com uma pequena perda de variabilidade na geração corrente. Nesta situação o que interessa para a evolução é uma alta taxa de recombinação, a manutenção dos indivíduos da população e uma taxa de mutação moderada para contornar a situação de redução da variabilidade.

21. Fraco (VO), Baixa (FMI), Alto (PR): Início de execução, com perda de variabilidade. Sendo assim, mantém-se parcialmente a população, aumenta-se a recombinação e eleva-se a taxa de mutação para um alto valor, a fim de que, seja forçado o retorno da variabilidade.

22. Fraco (VO), Média (FMI), Baixo (PR): Valor ótimo local está se perpetuando, medidas corretivas são necessárias para o retorno à evolução. Sendo assim, matam-se alguns indivíduos, forçando um alto cruzamento, com baixa mutação, deixando a seleção guiar o processo de evolução.

23. Fraco (VO), Média (FMI), Médio (PR): O valor ótimo fraco está se perpetuando razoavelmente, assim como a variabilidade apresenta sinais de comprometimento. Desta forma, matam-se alguns indivíduos, forçando um alto cruzamento e uma taxa média de mutação, a fim de que, a evolução volte a ocorrer.

24. Fraco (VO), Média (FMI), Alto (PR): O valor ótimo fraco está se perpetuando e a variabilidade acabou na geração atual. Sendo assim, substituem-se muitos indivíduos, fazendo o cruzamento ser alto, com alta taxa de mutação, de forma a recuperar a variabilidade.

25. Fraco (VO), Alta (FMI), Baixo (PR): Valor ótimo fraco e repetido. Descarta-se a população, forçando um alto percentual de cruzamento e uma taxa média de mutação, para garantir a variabilidade e o retorno da evolução.

26. Fraco (VO), Alta (FMI), Médio (PR): Valor ótimo fraco e repetido. Descarta-se a população, forçando um alto percentual de cruzamento e alta taxa de mutação, para melhorar a variabilidade.

27. Fraco (VO), Alta (FMI), Alto (PR): Valor ótimo fraco e repetido. Descarta-se a população, forçando um alto percentual de cruzamento e alta taxa de mutação, para melhorar de forma substancial a variabilidade. É como um reinício de execução.

Após a apresentação e extensa análise da base de regras, é possível notar que, diferente dos demais trabalhos - alguns apresentados no terceiro capítulo (seção 3.4) - que sempre apresentam o ajuste de parâmetros de forma separada e desassociada, este os utiliza de forma

combinada e ponderada em busca dos objetivos estabelecidos para o funcionamento do agente.

5.3 ROLETA DA MORTE

A variável de controle “Percentual de Sobreviventes”, um dos mais importantes mecanismos de funcionamento do agente, apresentada e justificada em detalhes na seção 5.1.3.3, descreve apenas a quantidade de indivíduos que deverá ser descartada e substituída por novos gerados de forma aleatória. Porém, não define como esta operação será realizada, ou seja, quais indivíduos serão escolhidos e como será feita sua escolha.

Em busca de uma definição inspirada na própria natureza, como apresentado nas teorias do quarto capítulo, e que também não destrua completamente o progresso alcançado, optou-se pelo uso de uma variação do método de seleção tradicional dos AGs, a Roleta Simples, aqui chamada de “Roleta da Morte” ou de “Juízo Final”.

Na Roleta Simples, apresentada no terceiro capítulo, constitui-se uma estrutura onde os indivíduos recebem uma probabilidade de escolha proporcional à sua adaptabilidade, ou seja, quanto melhores forem as soluções, maiores chances de seleção e conseqüente recombinação estas receberão.

De forma análoga, porém inversamente proporcional, ocorre na “Roleta da Morte”. Só que neste caso, os indivíduos menos adaptados é que recebem a maior probabilidade de serem escolhidos e “eliminados” da população.

Optou-se por utilizar esta estrutura, pois, em termos probabilísticos e reais, é natural que indivíduos mais adaptados, possuam mais chances de sobreviver mesmo em caso de catástrofes globais. De qualquer forma, isto não os isenta de serem eliminados, pois estes,

apesar de menor proporção dentro da “Roleta da Morte”, podem ainda ser escolhidos e então eliminados.

Deste modo, alguns indivíduos ainda não tão adaptados poderão sobreviver e, como sempre se espera da variabilidade, pode ser que eles possuam o “bloco construtor” faltante para o alcance da solução ótima.

Analisada ainda sobre a ótica do Design Inteligente, segundo as proposições de DEMBKI (2004, 2007), a extinção periódica de alguns indivíduos pode ser tida como o maior mecanismo de mudança implementado, e a única chance do designer efetivamente adicionar conhecimento a execução. Isto por que, uma vez que percebe que o valor “ótimo” encontrado está longe do desejado, pode forçar o aumento da variabilidade por meio da extinção, dentro de um ciclo constante de busca pelo “ótimo global”.

5.4 PERIODICIDADE E OS DIFERENTES TIPOS DE GAIA

Na seção 4.1.2, são discutidas algumas formas teóricas da periodicidade em que a intervenção do *designer* deve acontecer, abordando inclusive, que para algumas correntes a intervenção não é constante e sim esporádica, ou mesmo com uma certa frequência, como num ciclo.

Considerando estes fatores, até mesmo para efeito de testes da proposta e compreensão da sua real eficácia, são apresentados três “tipos” de GAIA, listados a seguir:

- **Permanente:** onde o algoritmo segue estritamente o fluxo demonstrado na figura 5.1, ativando sempre o GAIA, ao final de cada ciclo;
- **Estocástico:** respeitando-se uma determinada probabilidade, ao final de cada ciclo é testada se a intervenção deve ou não ocorrer;
- **Periódico:** onde a cada número de gerações pré-estabelecidas, o GAIA é ativado e realiza sua intervenção.

No capítulo seguinte, todas essas variações serão testadas com diferentes parâmetros na busca do possível melhor formato de intervenção e uso do agente proposto.

5.5 CONSIDERAÇÕES FINAIS

Apresentados todos os aspectos da proposta, do suporte conceitual até os detalhes da implementação e construção do agente inteligente GAIA, faz-se necessária a realização de um extenso conjunto de simulações e testes aplicados à resolução de um ou mais problemas, de alta complexidade e extensa superfície de busca.

Estes procedimentos são apresentados no próximo capítulo, por meio de gráficos e tabelas, o quais buscam avaliar não só a eficácia da estrutura, mas também a forma como ela deve ocorrer.

CAPÍTULO VI

*Se fiz descobertas valiosas, foi mais por ter paciência do
que qualquer outro talento.*

Sir Isaac Newton

6. TESTES E AVALIAÇÃO DA PROPOSTA

A proposta, apesar de toda sua estruturação e subsídios teóricos apresentados ao longo deste trabalho, só poderá ter sua eficácia avaliada por meio da extensa realização de testes, seguida de posterior tabulação e análise criteriosa dos resultados obtidos.

Este é o objeto principal do presente capítulo, o qual é iniciado apresentando o planejamento da estrutura de testes definida, bem como sua divisão em dois períodos distintos. Em seguida, é abordado o problema PCVS-BR-26, utilizado em todo o conjunto de testes. Na seqüência, passa-se a apresentação de fato dos resultados obtidos, por meio de gráficos e tabelas, agrupados por período e configuração definida.

O capítulo é encerrado com a análise consolidada de todos os resultados obtidos e, por meio destes, estabelece comparação tanto interna, entre as próprias configurações, como externa, com abordagens diversas apresentadas em outros trabalhos, mas que também fizeram uso do PCVS-BR-26.

6.1 ESTRUTURA E ABORDAGEM DE TESTES DEFINIDA

Assim como a própria proposta do agente inteligente, a bateria ou conjunto de testes significativos realizados, podem ser agrupados em dois principais períodos ou fases distintas: 2005/2006, durante a primeira proposta do agente, que refletiu nas 05 publicações listadas no Anexo F; e, 2010/2011, com uma visão aprofundada, tendo como principal foco a necessidade comprobatória necessária para o escopo deste trabalho.

Pode-se assim dizer que, o primeiro conjunto de testes foi feito de forma aleatória e experimental, sem qualquer grande planejamento ou processo definido. A idéia inicial principal foi analisar a performance da proposta e ajustá-la, acertando o domínio e o particionamento dos valores das variáveis fuzzy.

Além disso, pouco mais de 20 testes foram realizados. Isto se deve ao fato de que, considerando a performance computacional disponível à época, um único teste durava de quarenta minutos à duas horas, quando não, por muitas vezes, dadas estas restrições, era subitamente interrompido por um erro de estouro de memória ou algo do gênero.

Considerando estas restrições – até mesmo pela falta de conhecimento dos impactos da proposta e do próprio caráter de experimentação do momento –, além do uso do GAIA, as execuções definidas tiveram uma série de outros parâmetros variados, como o uso de diferentes métodos de seleção e recombinação, mudanças em valores atrelados à interação social, etc.

Fato notório na seção 6.2 deste capítulo, toda essa pluralidade que transcende a estrutura principal do GAIA, além da escassez de testes, já que apenas 20 puderam ser consolidados, deixou o trabalho longe de qualquer avaliação efetivamente conclusiva. Entretanto, como comprovado pelo número de publicações relacionadas, serviu para

demonstrar que a influência do agente de fato acontecia e que, pelo menos naquele momento a variabilidade passou a ser garantida.

Em virtude dessa insuficiência de resultados, o presente trabalho, no que chama de segundo período de testes (2010/2011), propôs uma nova estrutura com 28 configurações diferentes e 560 simulações. Nestas com o intuito de avaliar especificamente a performance do agente, foram variados apenas os parâmetros do seu tipo (permanente, periódico e estocástico), e mantidos fixos os demais de configuração de um algoritmo genético, conforme listados a seguir:

- **Problema-Teste:** PCVS-BR-26 (LEHRER, 2000);
- **Número de Testes por Configuração:** 20;
- **População Inicial:** com ponto de partida em Brasília, detalhes vide **Anexo B**;
- **Número de Indivíduos:** fixados em 100 para todas as gerações;
- **Critério de Parada:** fixado em 5.000 gerações para os “Testes-Padrão” e 50.000 para os “Super-Testes”;
- **Método de Recombinação:** PMX, detalhes vide **Capítulo III**;
- **Percentual Inicial de Recombinação:** 100%;
- **Método de Mutação:** Permuta-Simples, detalhes vide **Capítulo III**;
- **Percentual Inicial de Seleção para Mutação:** 0%;
- **Parâmetros Específicos para Testes Utilizando Interação Social (Anexo D):**
 - *Jogo Biológico:* Hawk-Dove (BRITO, 2004);
 - *Proporção Fenotípica Inicial de Estratégias:*

$$\frac{3}{9} Hawk : \frac{3}{9} Dove : \frac{2}{9} TFT : \frac{1}{9} Aleatório$$
 - *Número de Partidas:* 100 por geração.

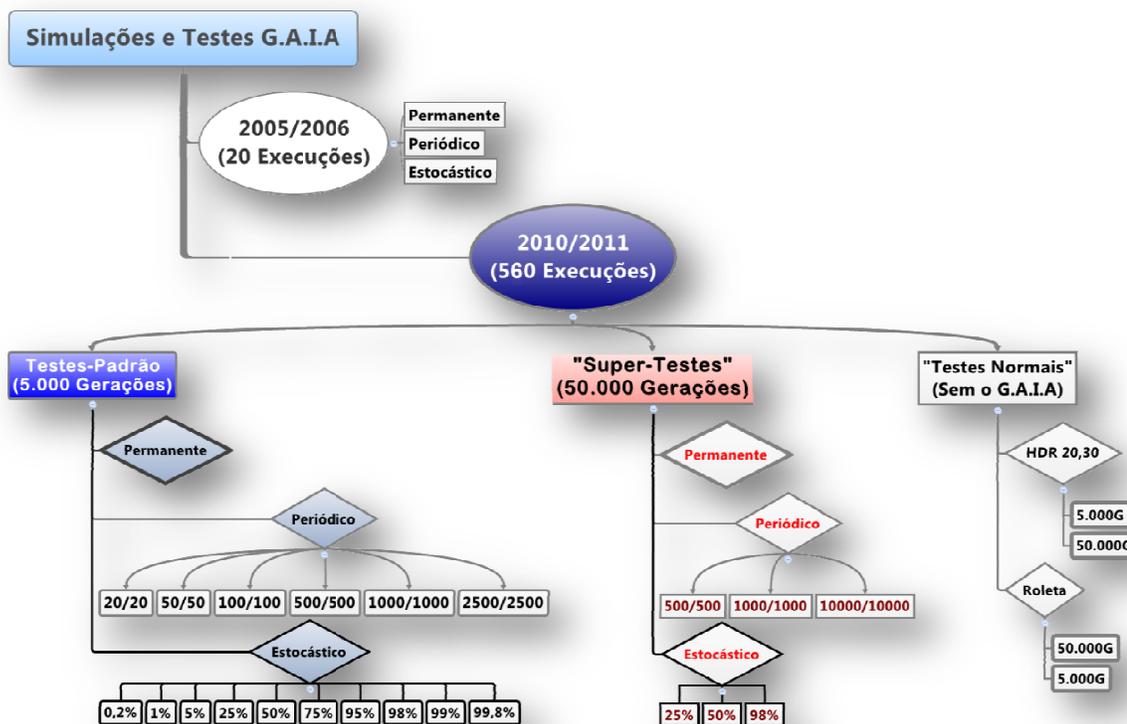
Nas execuções do primeiro grupo de testes (2005/2006), foram mantidos fixos em todas as execuções apenas: a mesma população inicial – inclusive, a mesma utilizada no grupo 2010/2011 –, assim como a quantidade de indivíduos, os percentuais iniciais das taxas de mutação e recombinação e o número de geração fixado em 5.000.

Todos os testes foram realizados por meio da evolução da já existente ferramenta de execução e depuração de algoritmos genéticos, o *Beagle XP*¹⁹, em sua versão 4.5. Apresentada originalmente em BRITO (2004), a ferramenta permite a depuração das técnicas implementadas através do uso de base de dados relacional e da possibilidade de definição de diferentes configurações com inúmeros parâmetros por meio da interface, além da manutenção em banco de dados de populações iniciais, o que garante à todos os testes a partida do mesmo ponto da superfície de busca.

Nas próximas seções do capítulo, serão apresentados cada um dos conjuntos de testes realizados, aqui sintetizados graficamente na figura 6.1. Maiores detalhes sobre as execuções e os dados gerados, devem ser consultados no Anexo C. A seguir, será apresentado o problema utilizado em todos os testes, o PCVS-BR-26 de LEHRER (2000).

¹⁹ Maiores detalhes como imagens, descrição de funcionalidades e modelagem da ferramenta, podem ser consultadas no Anexo D do trabalho.

Figura 6.1 – Estrutura Geral de Testes e Simulações GAIA.



6.1.1 O Problema Utilizado: PCVS-BR-26

6.1.1.1 O Problema do Caixeiro Viajante(*The Traveling Salesman Problem*)

O Problema do Caixeiro Viajante (PCV) ou *Traveling Salesman Problem* (TSP) é um problema clássico de otimização combinatória. De acordo com BUREAL (2000) apud LEHRER (2000), o PCV é um problema fácil de entender, mas extremamente complexo de resolver, pois pertence a classe dos problemas *NP-Completo* (NPC), onde o tempo utilizado em sua resolução cresce de forma exponencial, à medida em que se aumentam a quantidade de entradas.

Segundo TEIXEIRA (2005), os primeiros problemas matemáticos ligado ao PCV foram propostos pela primeira vez em 1857, pelas mãos do matemático, Sir William Rowan

Hamilton. Este propôs um jogo chamado de *Around the World*, onde o principal objetivo era encontrar uma rota através do vértice do decaedro (estrutura utilizada no tabuleiro), que iniciasse e terminasse no mesmo vértice (cidade), sem nunca passar pelo mesmo vértice duas vezes.

Mais tarde outros matemáticos e cientistas da computação estudaram o problema de forma mais específica propondo algumas variações e possíveis métodos de soluções (TEIXEIRA, 2004).

BUREAL (2000) apud LEHRER (2000), formalmente o define como sendo um grafo completo $G = (V, A)$, onde V é o conjunto de n vértices e A é o conjunto de arestas que conectam cada par de vértices i e j , com um custo c_{ij} associado. O objetivo do problema é buscar uma rota de menor custo passando por todos os vértices. Nas situações em que o custo entre os vértices é igual independentemente do sentido, ou seja, $c_{ij} = c_{ji}$, chama-se de Problema do Caixeiro Viajante Simétrico ou PCVS.

Apesar de simples em conceituação e até mesmo trivial, o caixeiro viajante é um problema de otimização com larga aplicabilidade prática. Conforme citado em HAUPT et al. (2000), o PCV pode ser modelado para situações da engenharia como na otimização da estrutura de encanamento de gás, no projeto de alimentação de uma antena, na configuração de transistores em um circuito, ou de uma forma geral, na organização de objetos e eventos que requeiram certa ordem.

O PCVS é substancialmente referenciado em toda a literatura de Algoritmos Genéticos. Isto por dois principais motivos: primeiro, porque é um problema prático de fácil compreensão e modelagem, e segundo, porque de acordo com o número de vértices utilizados produz a situação ideal para a aplicação de AGs: problemas de otimização com uma grande superfície de busca.

A seguir, apresentar-se-á a descrição da variação do PCVS de LEHRER (2000), considerando as distâncias entre as capitais brasileiras, problema base para todos os testes realizados neste trabalho.

6.1.1.2 PCVS-BR-26

A instância chamada de PCVS-BR-26 trata da modelagem do Problema do Caixeiro Viajante Simétrico para as 26 capitais do país interligadas por estradas, baseando-se nas distâncias oficiais oferecidas pelo Departamento Nacional de Estradas e Rodagens (DNER, 2004)²⁰.

Visando demonstrar o tamanho da superfície de busca para este problema, LEHRER (2000) afirma que, sendo o número de combinações (rotas) para um PCV calculado por:

$$\frac{(n-1)!}{2}, \quad (24)$$

onde, no caso específico, $n = 26$, existem $7,75 \times 10^{24}$ possibilidades de rotas diferentes.

Nesta situação, mesmo que se um computador conseguisse avaliar 1.000.000.000 de possibilidades por segundo por meio de força bruta, levar-se-ia ainda cerca de $2,46 \times 10^8$ (246 milhões) de anos para achar a melhor solução para o problema.

Desta forma, além do já mencionado fato da adoção deste problem em inúmeros outros trabalhos, o que garante a possibilidade de comparação, justifica-se o seu uso também pelo vasto espaço de busca e ausência de quaisquer restrições ou conhecimentos que facilitem o achado da solução ótima.

Sendo assim, modelou-se o problema utilizando indivíduos (soluções) com um cromossomo (vetor) de 27 posições para representar a rota a ser seguida passando pelas 26 cidades e retornando à origem. O alfabeto foi formado por 26 alelos estruturados na forma de

²⁰ Para consulta da tabela de distâncias do DNER (2004), vide Anexo A.

strings, onde cada código indica uma cidade (vértice) do problema (vide tabela B.1). Todos os cromossomos possuem sempre todos os alelos disponíveis no alfabeto e sem repetição, à exceção da última posição que, representando o retorno à cidade inicial, repete-se com a primeira.

A figura 6.2 apresenta um exemplo da forma como os cromossomos foram estruturados e dos valores possíveis dos alelos.

Figura 6.2 – Representação do cromossomo adotada para o PCVS-BR-26.

CROMOSSOMOPROBLEMA
04-21-13-14-11-09-17-10-16-20-03-02-06-18-19-08-25-00-15-22-24-01-12-07-23-05-04

Fonte: Beagle XP 4.5.

A função de adaptabilidade ou função de avaliação de *fitness* é dada pela distância representada pela rota. Esta distância é calculada segundo o somatório de todas as distâncias envolvidas em cada par de cidades sequenciais do cromossomo.

Nas próximas seções serão apresentados os testes e os resultados alcançados com a proposta deste trabalho aplicada à solução do problema aqui descrito.

6.2 PRIMEIROS TESTES – PERÍODO 2005/2006

Dadas as restrições da época, já discutidas na seção 6.1, de todos os testes realizados para o período de 2005/2006, apenas 20 puderam ser considerados significativos e que, inclusive, deram origem às primeiras publicações sobre o tema.

Estes são apresentados na tabela 6.1, que os descreve por meio da apresentação da configuração utilizada, do melhor resultado encontrado e da primeira geração em que este

apareceu. Sendo as duas últimas, algumas das medidas de performance utilizadas naquele momento inicial.

Tabela 6.1 – Resultados Testes Período 2005/2006 (20 Melhores Resultados)

ID	NOME / DESC.	GAIA		RESULTADOS	
		TIPO	PARAM.	MELHOR (KM)	1ª GER. ENCONT
01	Periódico-100, Hawk-Dove Roleta, C=20,V=30	Periódico	100	20586	1364
02	Estocástico-25%, Hawk-Dove Roleta, C=20,V=30	Estocástico	25%	21069	833
03	Periódico-10, Hawk-Dove Torneio, C=30,V=20	Periódico	10	21181	723
04	Estocástico-50%, Hawk-Dove Torneio, C=30,V=20	Estocástico	50%	21254	357
05	Estocástico-25%, Hawk-Dove Roleta, C=20,V=30	Estocástico	25%	21324	4641
06	Periódico-100, Hawk-Dove Roleta, C=20,V=30	Periódico	100	21528	1212
07	Estocástico-75%, Hawk-Dove Roleta, C=20,V=30	Estocástico	75%	21946	1883
08	Periódico-10, Hawk-Dove Roleta, C=20,V=30	Periódico	10	21957	2843
09	Periódico-20, Torneio	Periódico	20	22061	4793
10	Estocástico-25%, Hawk-Dove Roleta, C=20,V=30	Estocástico	25%	22185	1577
11	Periódico-50, Hawk-Dove Roleta, C=20,V=30	Periódico	50	22242	2534
12	Periódico-50, Hawk-Dove Roleta, C=20,V=30	Periódico	50	22753	3631
13	Periódico-20, Hawk-Dove Roleta, C=20,V=30	Periódico	20	23764	312
14	Estocástico-50, Hawk-Dove Roleta, C=20,V=30	Estocástico	50%	24079	4394
15	Permanente, Hawk-Dove Roleta, C=20,V=30	Permanente	—	24260	2256
16	Periódico-10, Roleta	Periódico	10	24487	790
17	Permanente, Hawk-Dove Roleta, C=20,V=30	Permanente	—	24487	3960
18	Periódico-10, Hawk-Dove Roleta, C=20,V=30	Periódico	10	24782	4680

ID	NOME / DESC.	GAIA		RESULTADOS	
		TIPO	PARAM.	MELHOR (KM)	1ª GER. ENCONT
19	Permanente, Hawk-Dove Roleta, C=20,V=30	Permanente	—	25440	1776
20	Estocástico-75%, Hawk-Dove Roleta, C=20,V=30	Estocástico	75%	25829	758

De todos os testes realizados, optou-se pela plotagem dos gráficos completos apenas das duas execuções com melhor desempenho, principalmente por se considerar o fato de que, apenas um teste foi realizado para cada uma destas configurações.

Sendo assim, os gráficos 6.1 e 6.2, representam, respectivamente, as execuções GAIA Periódico 100 e GAIA Estocástico 25%, onde é expressa a relação “adaptabilidade (*fitness*) x geração”, tomando-se por base os valores da média, do melhor e do pior a cada geração da execução.

Gráfico 6.1 – GAIA Periódico 100 (2005/2006): Melhor, Pior, Média x Geração

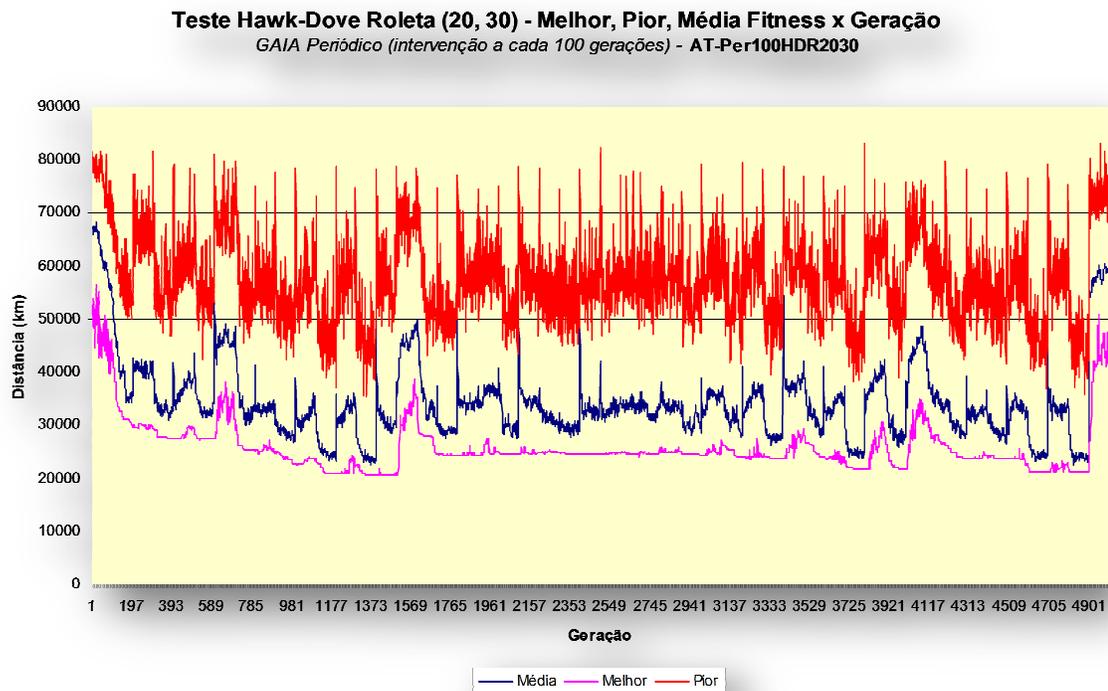
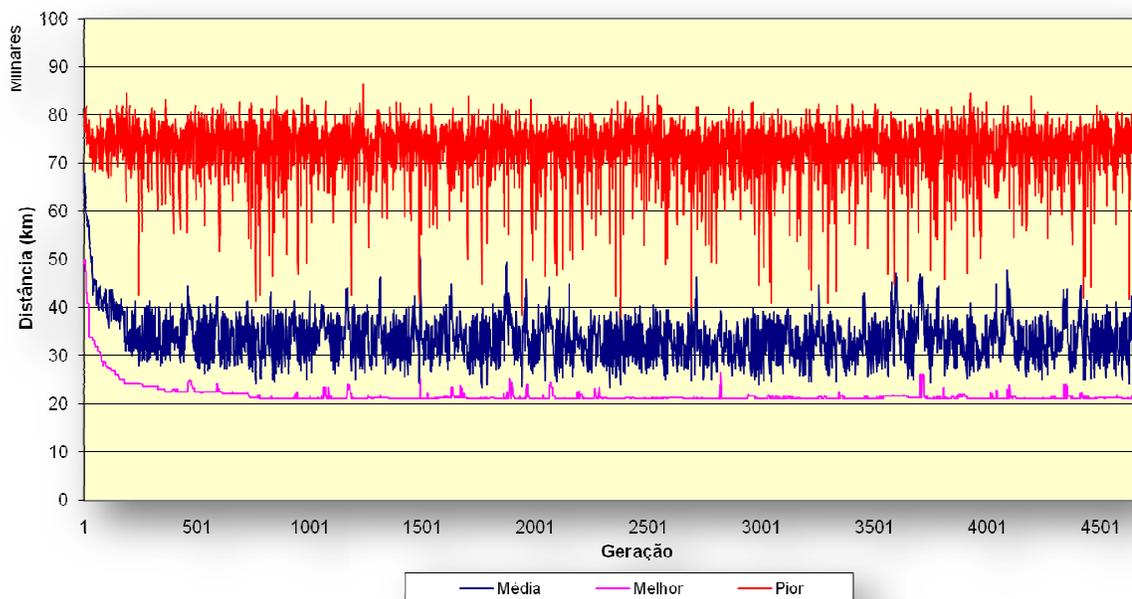


Gráfico 6.2 – GAIA Estocástico 25% (2005/2006): Melhor, Pior, Média x Geração



Em ambas as situações, pode-se notar claramente que a variabilidade é mantida por meio da intervenção, um dos principais objetivos da construção do agente.

No periódico 100, por sua característica de intervenção conhecida, fica ainda mais fácil notá-la. Nesta situação, é possível por meio da avaliação do gráfico 6.1, observar boa parte do ciclo de funcionamento esperado da proposta.

A execução, após as 100 primeiras gerações, naturalmente tende a convergência, quando então uma forte intervenção do GAIA ocorre, provavelmente com alta eliminação de indivíduos e aumento significativo no percentual de mutação. Durante o próximo período, pós-intervenção, pouca evolução e até mesmo “involução” são registradas, entretanto, nota-se claramente que a variabilidade é reconquistada. Até que então, chega o próximo momento de intervenção, onde os parâmetros são novamente normalizados e a evolução volta a acontecer.

A diferença é que agora a população evolui ainda mais, alcançando novos e melhores valores. Desta forma, sucessivamente, repete-se este ciclo até o término da execução.

Na execução estocástica com 25% de probabilidade, é um pouco mais difícil avaliar diretamente a incidência da intervenção do agente. Todavia, analisando alguns pequenos “picos de piora” e uma média de *fitness* com um valor sempre próximo à metade do domínio da variável “Valor Ótimo”, pode-se concluir que elas acontecem com a frequência esperada, garantindo a variabilidade e a constante busca pelo ótimo global.

Em relação à performance obtida nos dois resultados, considerando que em TEIXEIRA (2005), o valor ótimo de 20.093km foi encontrado, fica ainda mais difícil julgar se efetivamente foi a técnica que fez a diferença na busca destes valores, já que nos dois trabalhos, apenas uma simulação fora realizada para cada configuração proposta, tornando-as muito sujeitas aos fatores de casualidade dos AGs.

Principalmente por considerar estes fatores aleatórios dos AGs, é que foi elaborado um novo conjunto de testes, o 2010/2011, próprio para o presente trabalho, onde, para fins de amostragem da eficácia de cada configuração proposta, uma grande quantidade de testes foi realizada. Estes resultados serão descritos e apresentados na seção seguinte.

6.3 TESTES DEFINITIVOS – PERÍODO 2010/2011

Em razão da insipiência e de todas as dificuldades já descritas encontradas durante a primeira avaliação da proposta do GAIA, tornou-se essencial à realização deste trabalho a elaboração de um novo, mais bem estruturado e bem mais amplo conjunto de testes.

A nova estrutura proposta de testes está agrupada em dois grandes grupos: os chamados “**Testes-Padrão**”, com execuções de 5.000 gerações, e os “**Super-Testes**”, com execuções de 50.000 gerações. Cada um dos dois grupos, subdivide-se por meio da forma do

GAIA utilizada, ou seja: em permanente, periódico, estocástico, e suas variações paramétricas, de período e probabilidade, respectivamente.

Além disso, para fins internos de comparações, levando-se em consideração os dois grandes grupos apresentados, também foram realizados o que aqui se chamou de “**Testes Normais**”. Estas configurações descrevem um AG Canônico utilizando a Roleta Simples, e, a abordagem híbrida de BRITO et al. (2005), AG com Interação Fenotípica utilizando o método Hawk-Dove Roleta.

Desta forma, como já descrito, foram definidas 28 novas configurações e realizados no todo 580 testes com as mesmas, totalizando 20 testes para cada.

Essa nova quantidade – vinte e oito vezes maior –, assim como, as execuções com 50.000 gerações, só foram alcançadas após profundas modificações na estrutura do *framework Beagle XP*. Some-se também ao fato, a considerável ajuda dos avanços tecnológicos e do aumento do poder de processamento no período. Os testes que antes eram executados em quatro horas, foram reduzidos para pouco mais de 03 minutos.

Na seqüência desta seção, serão apresentadas duas novas subseções, uma para cada conjunto de testes definidos: 6.3.1, para os “Testes-Padrão” e, 6.3.2, para os “Super-Testes”. Em cada uma delas, novas subseções são entregues, sendo uma por cada tipo de uso do GAIA, podendo ainda, no caso dos tipos periódico e estocástico, conter mais de uma configuração associada.

6.3.1 Testes-Padrão

No que se chamou de “Testes-Padrão”, ou seja com 5.000 gerações a cada execução, foram definidas 19 configurações diferentes, sendo: 01 para o modo permanente do GAIA; 06 para o periódico, variando-se o intervalo de gerações para a intervenção; 10 para o modo estocástico, variando-se a probabilidade de ocorrência com valores arbitrados entre 0,2% e 99,8%; e, para efeitos de comparação interna, 01 configuração de AG Canônico com a roleta simples e 01 configuração com o método Hawk-Dove Roleta, ambos sem o uso do GAIA.

No intuito de diminuir a casualidade presente em uma única execução, para cada uma das configurações, foram realizados 20 execuções, totalizando 380 testes para este grupo. As configurações, assim como, uma primeira análise de performance, valendo-se do melhor valor encontrado e da primeira geração em que ocorreu, são apresentados na tabela 6.2, descrita a seguir.

Tabela 6.2 – Testes-Padrão: Execuções e Resultados (Sintético)

ID	NOME / DESC.	GAIA		QTD	RESULTADOS	
		TIPO	PARAM.		MELHOR (KM)	1ª GER. ENCONT.
01	Teste-Padrão GAIA Permanente (TP-GAIA-Permanente)	Permanente	–	20	21.117	714
02	Teste-Padrão GAIA Periódico 20 (TP-GAIAPeriod-20)	Periódico	20	20	20.113	4.590
03	Teste-Padrão GAIA Periódico 50 (TP-GAIAPeriod-50)	Periódico	50	20	20.287	1.527
04	Teste-Padrão GAIA Periódico 100 (TP-GAIAPeriod-100)	Periódico	100	20	20.113	2.383
05	Teste-Padrão GAIA Periódico 500 (TP-GAIAPeriod-500)	Periódico	500	20	21.235	3.059
06	Teste-Padrão GAIA Periódico 1.000 (TP-GAIAPeriod-1000)	Periódico	1.000	20	20.853	4.173

ID	NOME / DESC.	GAIA		QTD	RESULTADOS	
		TIPO	PARAM.		MELHOR (KM)	1ª GER. ENCONT.
07	Teste-Padrão GAIA Periódico 2.500 (TP-GAIAPeriod-2500)	Periódico	2.500	20	22.085	1.665
08	Teste-Padrão GAIA Estocástico 95% (TP-GAIAEstoc-95)	Estocástico	95%	20	21.069	1.668
09	Teste-Padrão GAIA Estocástico 98% (TP-GAIAEstoc-98)	Estocástico	98%	20	21.135	858
10	Teste-Padrão GAIA Estocástico 99% (TP-GAIAEstoc-99)	Estocástico	99%	20	20.586	773
11	Teste-Padrão GAIA Estocástico 99,8% (TP-GAIAEstoc-99,8)	Estocástico	99,8%	20	20.093	2.911
12	Teste-Padrão GAIA Estocástico 75% (TP-GAIAEstoc-75)	Estocástico	75%	20	21.158	3.687
13	Teste-Padrão GAIA Estocástico 50% (TP-GAIAEstoc-50)	Estocástico	50%	20	21.274	277
14	Teste-Padrão GAIA Estocástico 0,2% (TP-GAIAEstoc-0.2)	Estocástico	0,2%	20	21.178	4.855
15	Teste-Padrão GAIA Estocástico 01% (TP-GAIAEstoc-1)	Estocástico	01%	20	20.093	1.056
16	Teste-Padrão GAIA Estocástico 05% (TP-GAIAEstoc-5)	Estocástico	05%	20	20.093	579
17	Teste-Padrão GAIA Estocástico 25% (TP-GAIAEstoc-25)	Estocástico	25%	20	20.093	1.306
18	Teste-Padrão Hawk-Dove Roleta (20,30) (TP-HDR2030)	Sem GAIA	–	20	21.089	709
19	Teste-Padrão Roleta Simples (TP-RoletaSimples)	Sem GAIA	–	20	21.069	4.079
TOTAL DE TESTES E MELHOR RESULTADO				380	20.093	579

Pode-se notar também, que foi realizada uma grande quantidade de configurações para o modelo estocástico. Este fato se deve a um problema de codificação do motor de

probabilidade da aplicação no período. Entretanto, considerando que uma grande quantidade de dados foi gerada para avaliação, optou-se por não descartá-los e também apresentá-los no trabalho. Isto porque, reforçam ainda mais a viabilidade da proposta, assim como uma tendência de um possível “melhor” parâmetro de probabilidade de intervenção, apresentados nas seções seguintes.

Na tabela 6.2, considerando-se a avaliação, ainda que muito pontual, das variáveis de performance apresentadas, percebe-se que foi alcançado em 02 configurações (GAIA Periódico 20 e 100), o mesmo valor de *fitness* tido como o melhor resultado em BRITO (2004), o de 20.113 km. Também, conseguiu-se em quatro configurações diferentes, e em seus repetidos testes, atingir o melhor valor já identificado para o problema na literatura, o de 20.093 km. Em seu trabalho, TEIXEIRA (2005) foi o primeiro a encontrá-lo, porém o fez apenas uma única vez, o que pode ser considerado efeito estritamente da casualidade.

Mesmo assim, se forem comparados com a literatura todos os resultados obtidos para os “Teste-Padrão”, serão observados que, este é um dos melhores *set* de testes já realizados para o problema PCVS-BR-26. Em LEHRER (2000), BRITO (2004) e TEIXEIRA (2005), os quais apresentaram abordagens híbridas de AGs, em todos os *set* de teste eram encontrados “melhores” valores que em média ficavam entre 23 e 24 mil km, somente um ou outro pontualmente superava esta faixa. Ao passo que, nos “Testes-Padrão” utilizando o GAIA, o “pior” melhor resultado encontrado, foi o de 22.085 km, ficando os demais entre este e 20.093 km.

Embora, possibilite uma análise bem objetiva do alcançado, observar apenas o valor ótimo encontrado nas configurações está longe de possibilitar toda a compreensão necessária para avaliação da proposta do GAIA. Sendo assim, em todas as subseções seguintes, para cada uma das configurações, serão exibidos dois tipos de gráficos: o primeiro, retrata uma espécie de “execução ideal”, apontando para o melhor valor encontrado em cada uma das 20

execuções, sendo então considerados o menor valor entre os valores ótimos e a média da geração e, também, o maior valor entre os piores valores das gerações para o ateste da variabilidade; já o segundo tipo, apresenta três séries temporais para os piores, melhores e média dos valores de *fitness* de cada geração, considerando para tal a média aritmética simples entre cada um dos 20 testes realizados.

6.3.1.1 GAIA Permanente

Nos gráficos 6.3 e 6.4, pode-se perceber que a intervenção constante do GAIA parece acontecer de forma mais sutil, provavelmente sendo ativadas apenas as regras que têm maior influência sobre os percentuais de mutação e recombinação.

No entanto, o gráfico 6.4, média das 20 execuções, demonstra para o valor ótimo uma ocorrência muito positiva para a proposta, a de sempre evoluir, ainda que de forma suave.

Fica também evidente, pela análise da série dos “piores” e “média” dos valores, a garantia da variabilidade. Isto porque, estas estão sempre mantidas num patamar elevado distante do ponto ótimo corrente, o que, confirma a proposta de controle da variabilidade por meio da intervenção do GAIA.

Gráfico 6.3 – GAIA Permanente: Melhores Valores x Geração

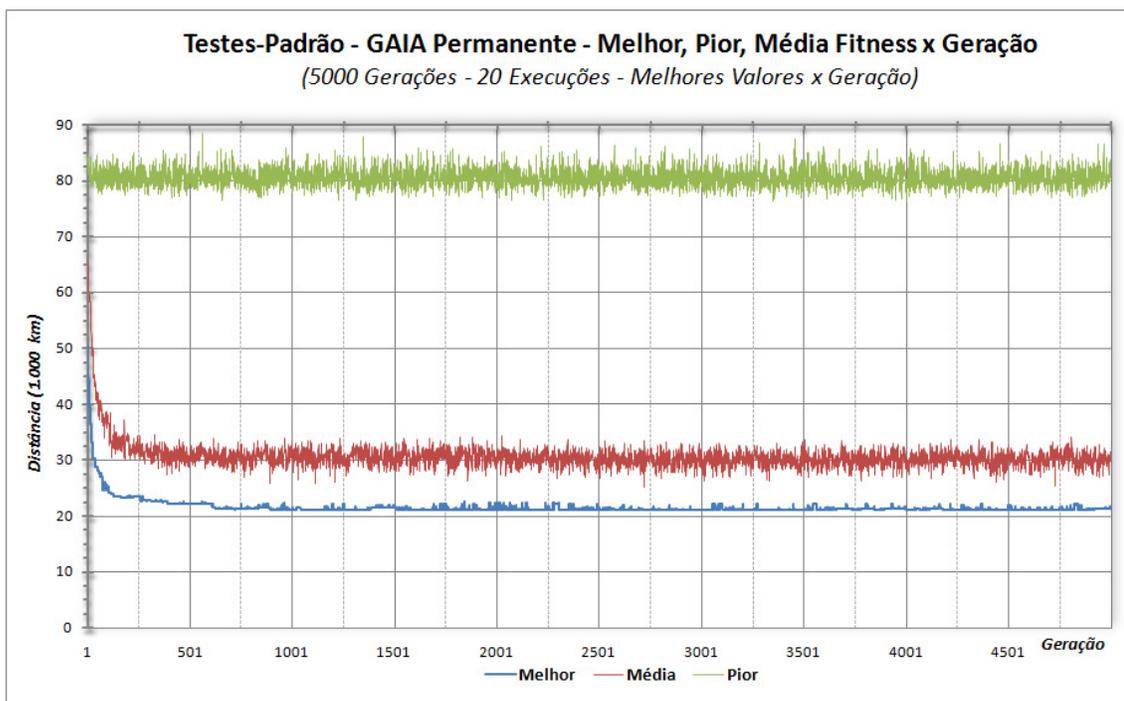
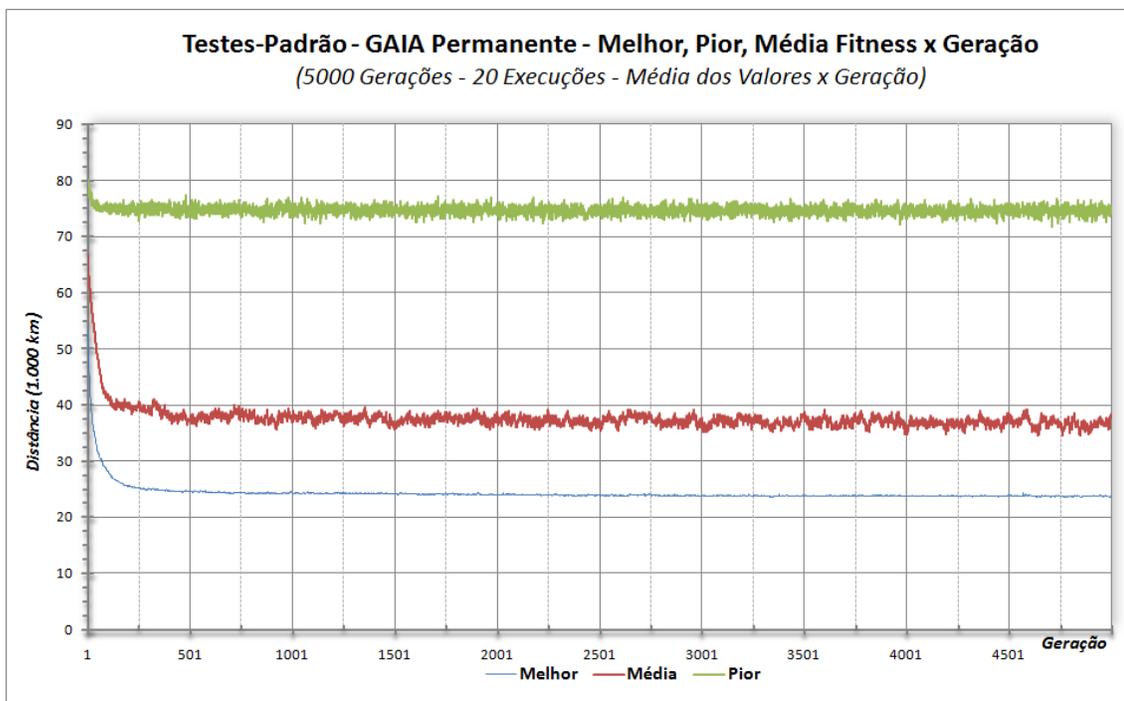


Gráfico 6.4 – GAIA Permanente: Média dos Valores x Geração



6.3.1.2 GAIA Estocástico

Conforme descrito e justificado na seção anterior, foram definidas 10 configurações diferentes para o GAIA Estocástico, com um total de 200 testes, representados aqui de forma sintetizada em 20 gráficos. Sendo assim, de forma a facilitar a análise, e até mesmo pela semelhança entre os resultados obtidos, agrupou-se em três partes: os de “baixa probabilidade”, GAIA Estocásticos 0,2%, 01% e 05%; os de “média probabilidade”, 25%, 50% e 75%; e, os de “alta probabilidade”, 95%, 98%, 99,8% e 99%.

Nas configurações de “baixa probabilidade”, por meio dos gráficos 6.5, 6.6, 6.7, 6.8, 6.9 e 6.10, é possível observar alguns “picos” de piora. Conforme descritos na seção 6.2 dos testes 2005/2006, estes “picos” estão atrelados a baixa incidência do GAIA, que ao ocorrer em estados de muita convergência, elimina muitos indivíduos, diminui o percentual de cruzamento e aumenta o de mutação. Fato que reflete num período subsequente de relativa piora, porém de retorno da variabilidade, até que uma nova intervenção ocorra para reajustar estes parâmetros e retomar a evolução pela busca da solução ótima.

Este “ciclo” de intervenções espaçadas descrito, tem mais uma vez sua eficácia comprovada, pois em duas dessas configurações de “baixa probabilidade”, o valor ótimo 20.093 km, apareceu repetidas vezes. Além disso, a variabilidade nunca foi perdida nas execuções do início ao fim, como pode ser observado por meio das séries de média e de piores valores.

Nas configurações de “média probabilidade”, através dos gráficos 6.11, 6.12, 6.13, 6.14, 6.15 e 6.16, é possível observar pequenos e suaves “picos” de piora. Isto ocorre, porque, de maneira similar a abordagem de intervenção constante, quanto maior o número de ocorrências do GAIA, mais suaves tendem a ser suas intervenções. A diferença neste caso, é que, por serem maiores os espaços entre ocorrências, é permitido ao GAIA uma melhor exploração e aproveitamento da base de regras. Ficam também, mais uma vez, comprovadas a

constante melhora (evolução) em busca do ótimo global, e a manutenção da variabilidade, devido aos mesmos fatores já citados.

Em termos de performance na busca pelo valor ótimo global, a configuração com 25% de probabilidade chegou a encontrar o valor ótimo 20.093 km, mais de uma vez, enquanto as demais, apresentaram valores muito próximos a este e indícios de que os atingiriam com mais algumas gerações.

Por fim, nas configurações de “alta probabilidade”, por meio dos gráficos 6.17, 6.18, 6.19, 6.20, 6.21, 6.22, 6.23 e 6.24, pode-se constatar grande semelhança entre todos e muita proximidade com as conclusões obtidas na análise do GAIA permanente. Em todas, há evolução constante e manutenção da variabilidade, mas sempre de forma suave, tal como são as intervenções realizadas pelo agente nestas situações.

Em resumo, em termos de comparação entre os diferentes parâmetros de probabilidade estabelecidos, pode-se notar que pequenas variações, apresentam pouca ou nenhuma diferença no resultado das execuções, ou seja, há diferença somente entre os “três grupos” apresentados nesta seção, porém, entre seus pares, esta diferença chega a ser pífia e imperceptível.

Entre os três grupos, apesar de ser difícil e até certo ponto empírico apontar o melhor padrão probabilístico, pode-se dizer que algo entre os valores médios e de baixa probabilidade de incidência, faz um melhor aproveitamento da estrutura do GAIA e do próprio algoritmo genético. Isto porque, é permitida à população a evolução tradicional, no entanto, dado sua característica natural de excessiva convergência, o acontecimento da intervenção faz com que os valores sejam novamente acertados e a população retornada ao rumo da evolução e conseqüente busca pelo ótimo global.

Gráfico 6.5 – GAIA Estocástico 0,2%: Média dos Valores x Geração

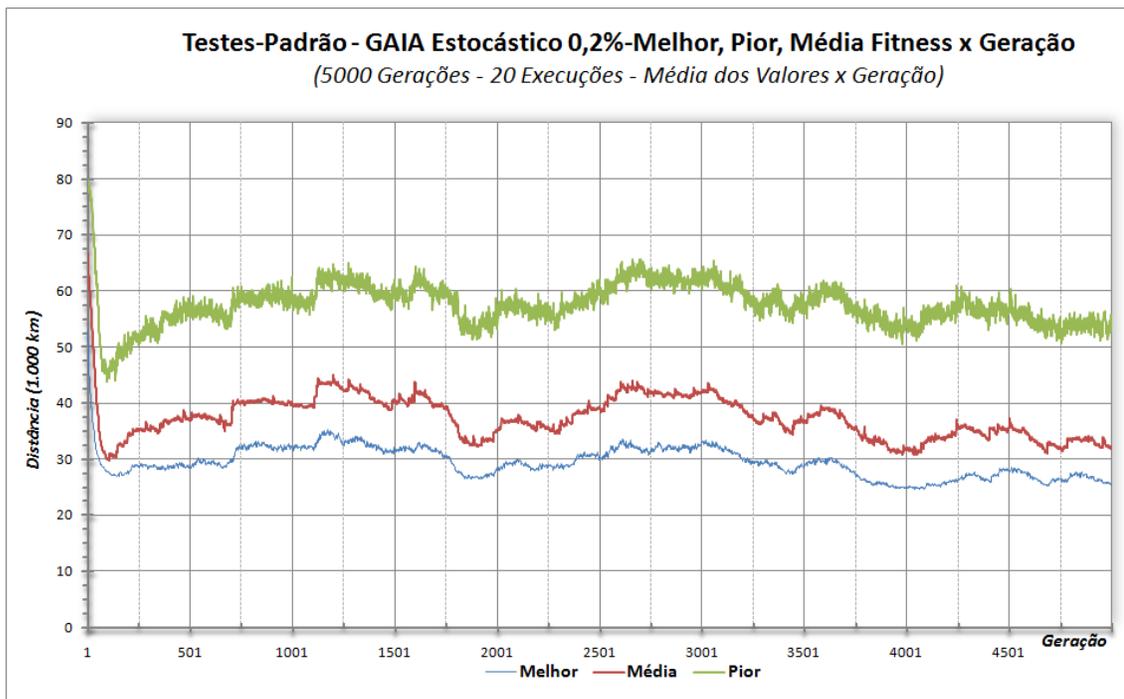


Gráfico 6.6 – GAIA Estocástico 0,2%: Melhores Valores x Geração

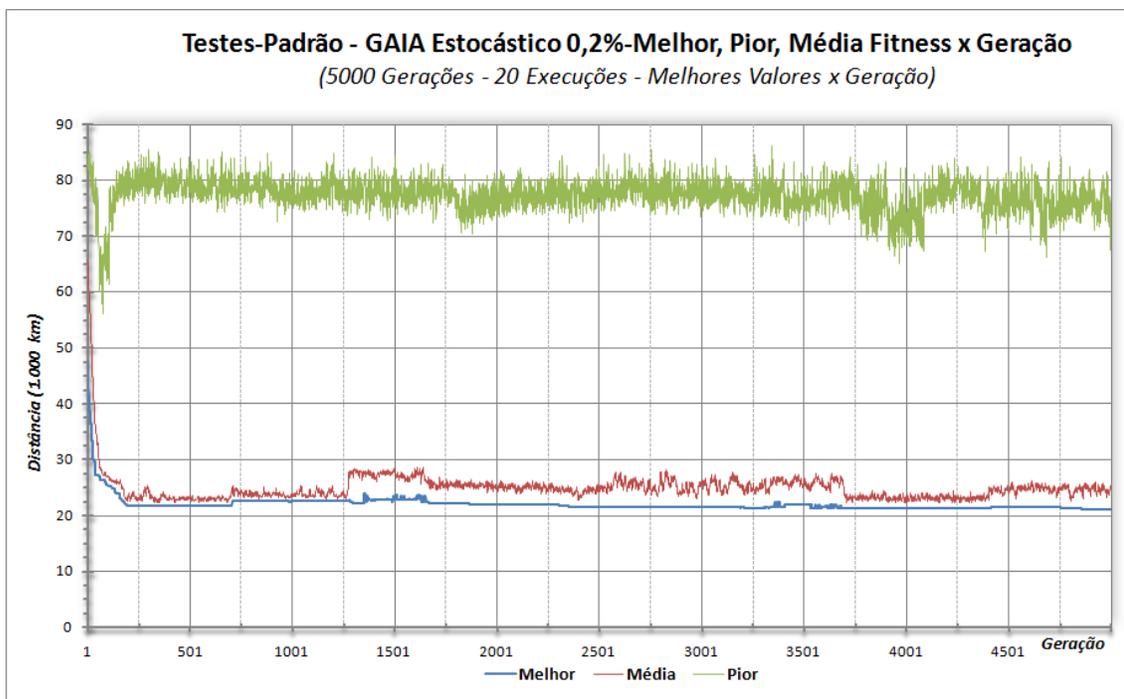


Gráfico 6.7 – GAIA Estocástico 01%: Média dos Valores x Geração

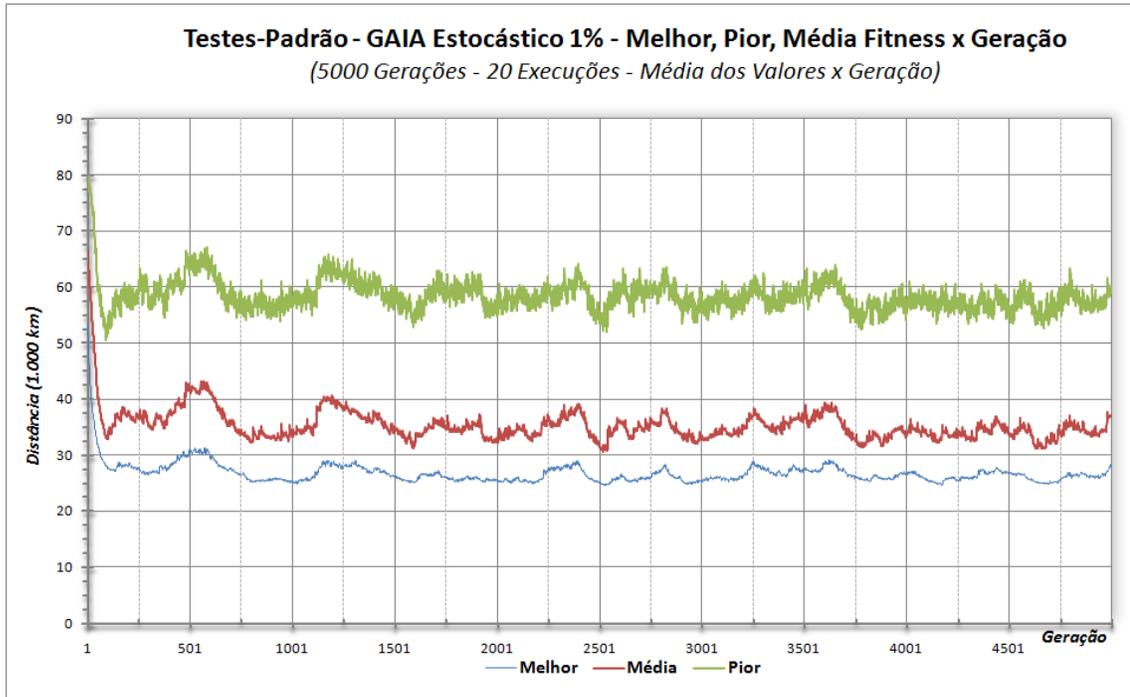


Gráfico 6.8 – GAIA Estocástico 01%: Melhores Valores x Geração

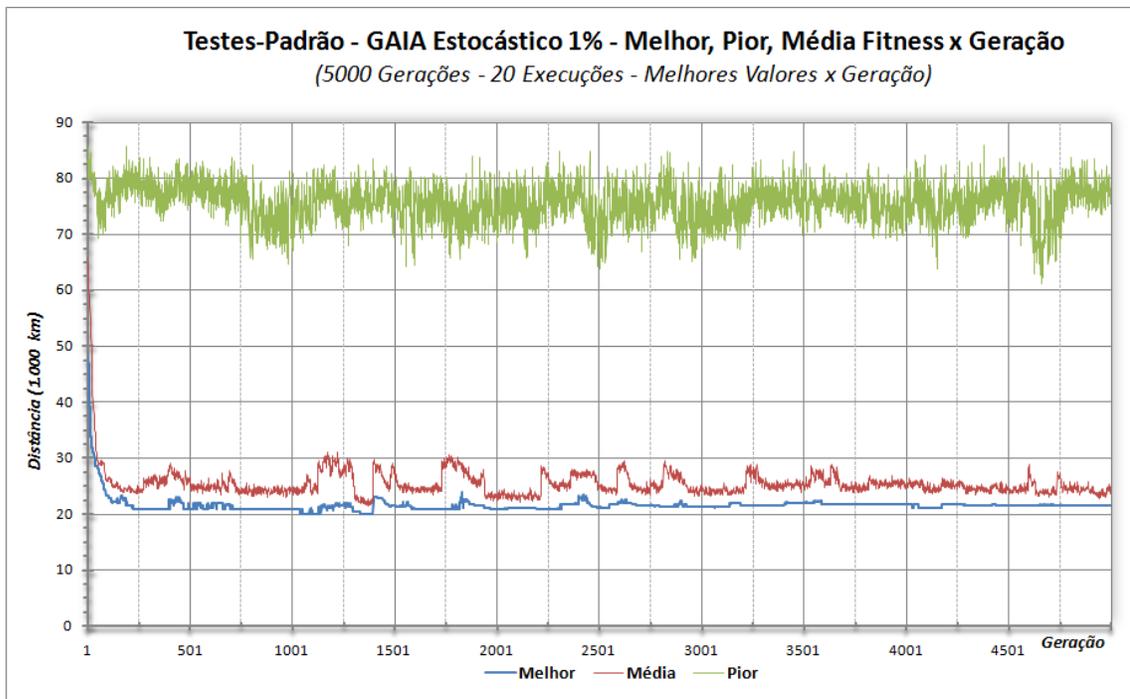


Gráfico 6.9 – GAIA Estocástico 05%: Média dos Valores x Geração

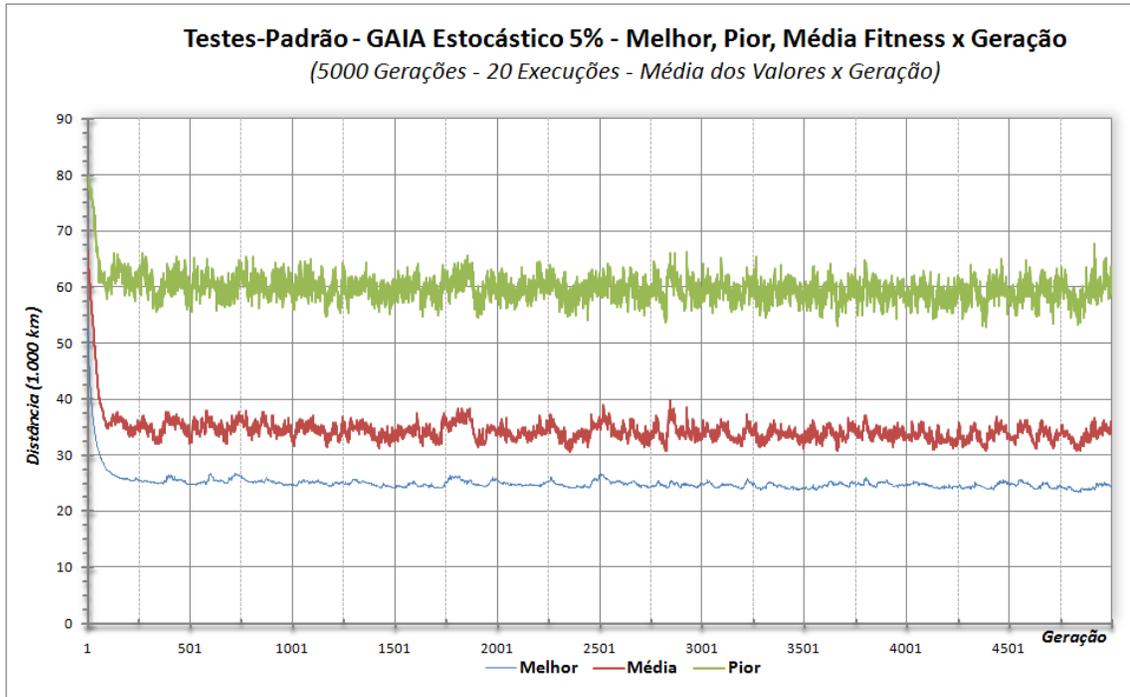


Gráfico 6.10 – GAIA Estocástico 05%: Melhores Valores x Geração

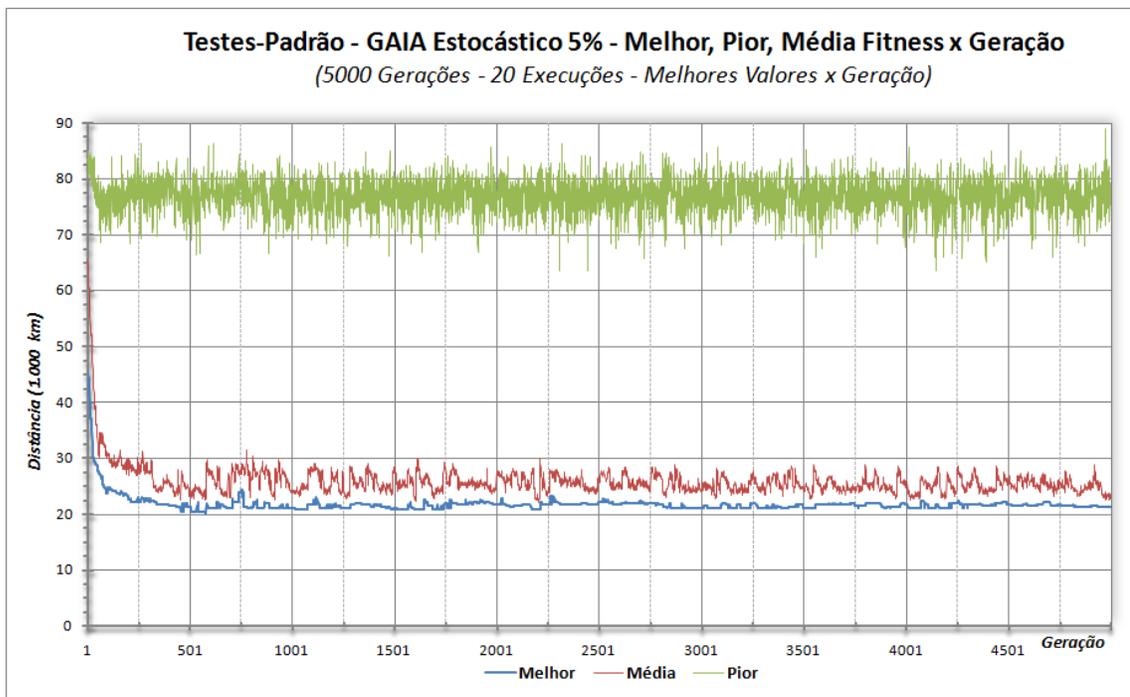


Gráfico 6.11 – GAIA Estocástico 25%: Média dos Valores x Geração

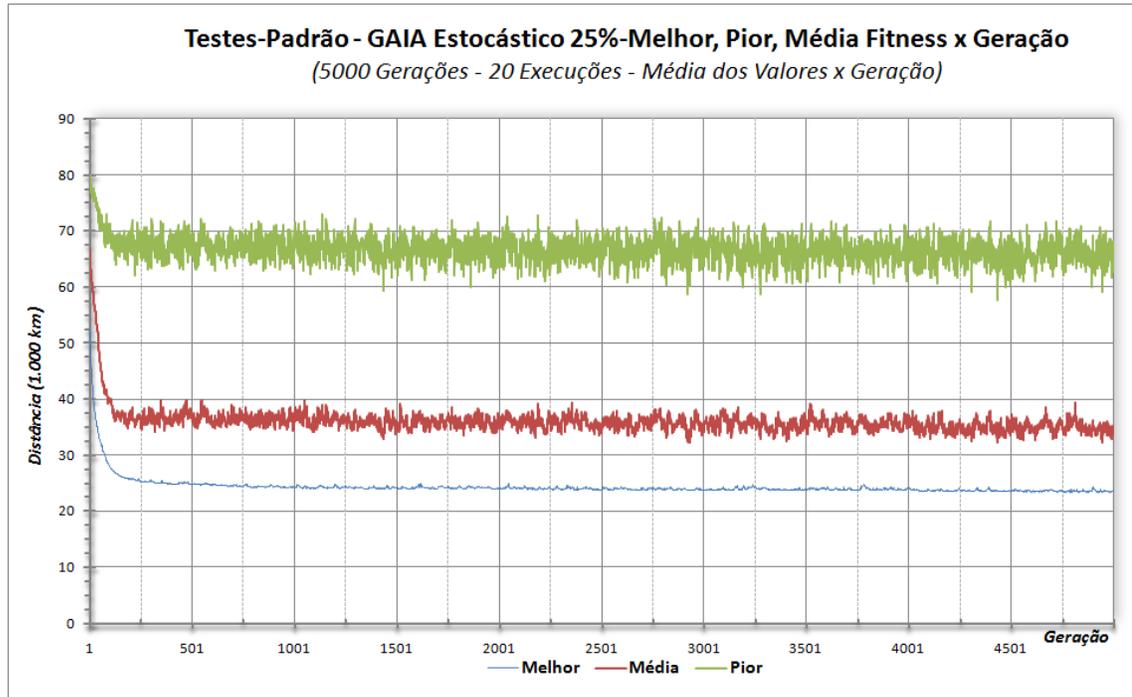


Gráfico 6.12 – GAIA Estocástico 25%: Melhores Valores x Geração

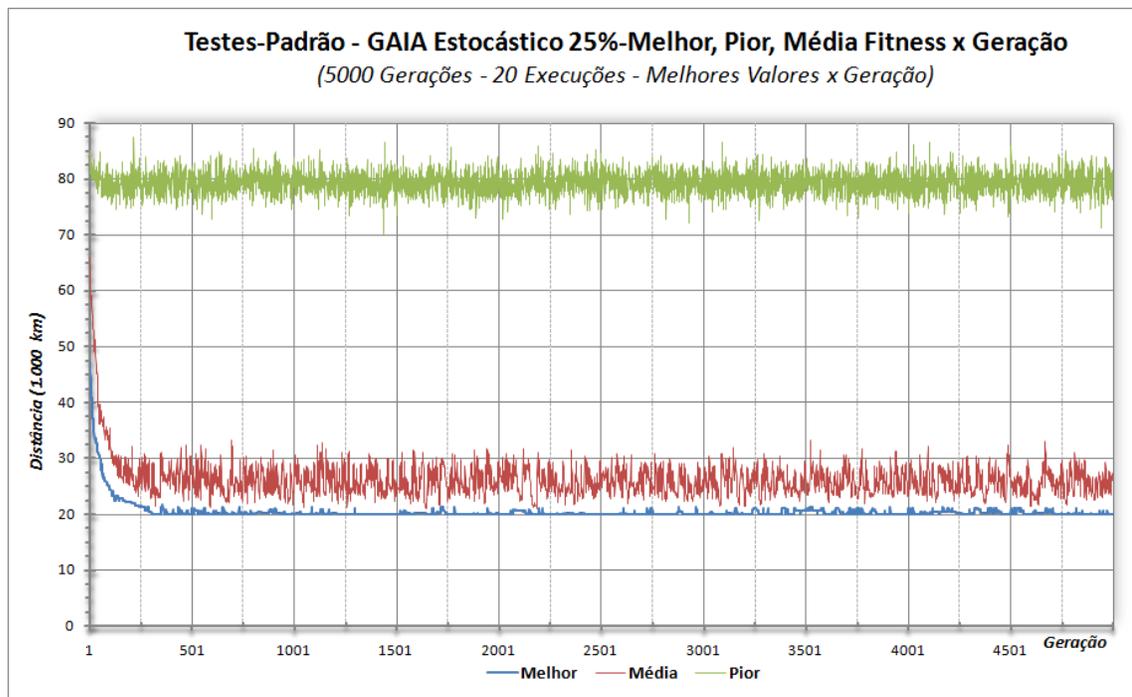


Gráfico 6.13 – GAIA Estocástico 50%: Média dos Valores x Geração

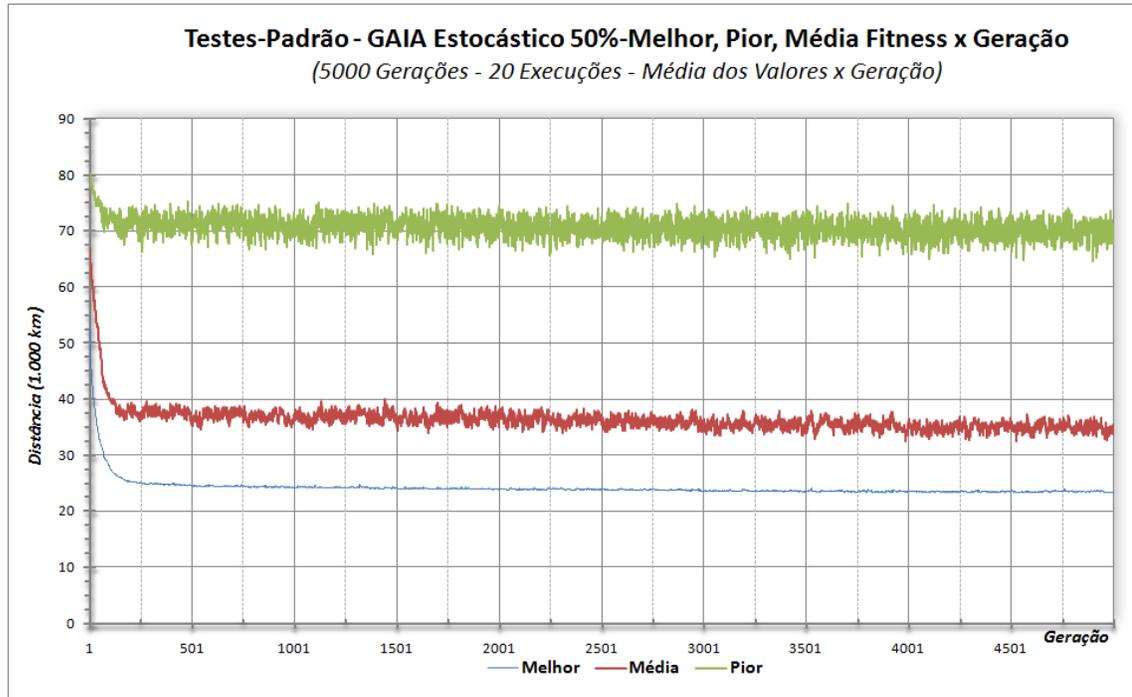


Gráfico 6.14 – GAIA Estocástico 50%: Melhores Valores x Geração

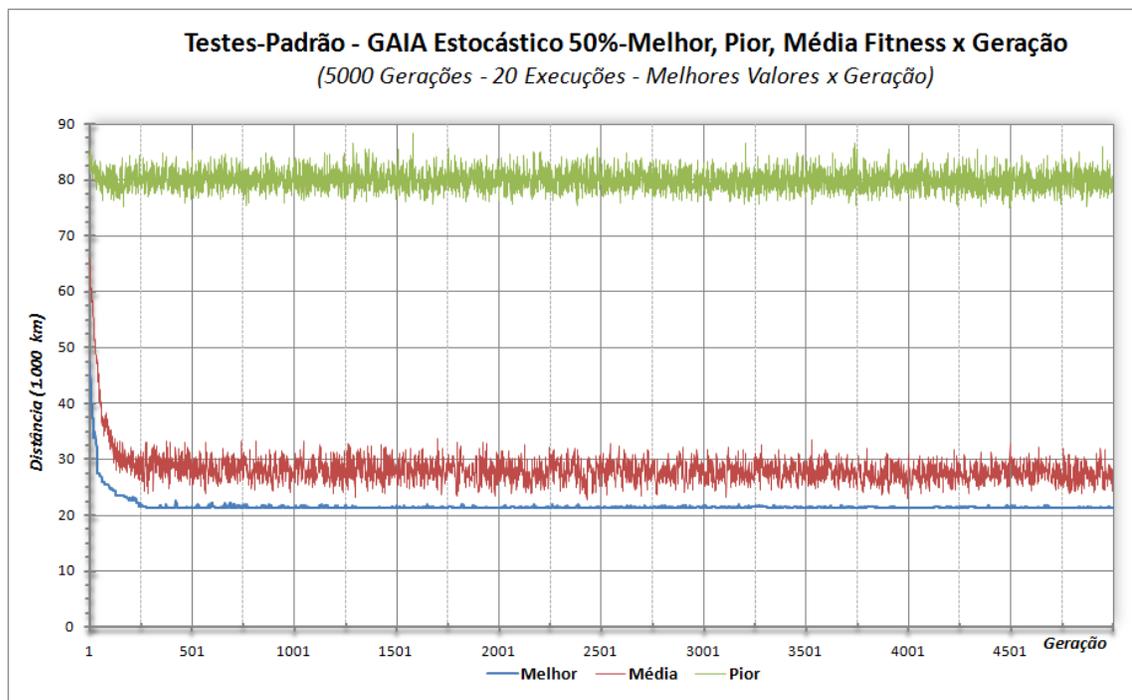


Gráfico 6.15 – GAIA Estocástico 75%: Média dos Valores x Geração

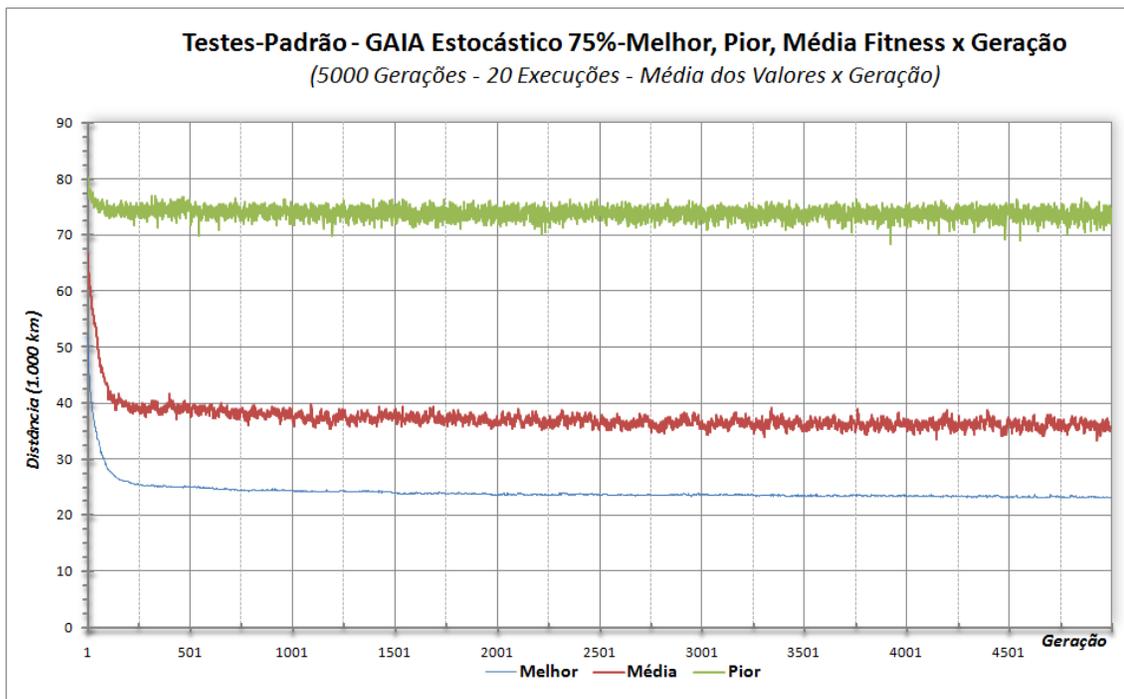


Gráfico 6.16 – GAIA Estocástico 75%: Melhores Valores x Geração

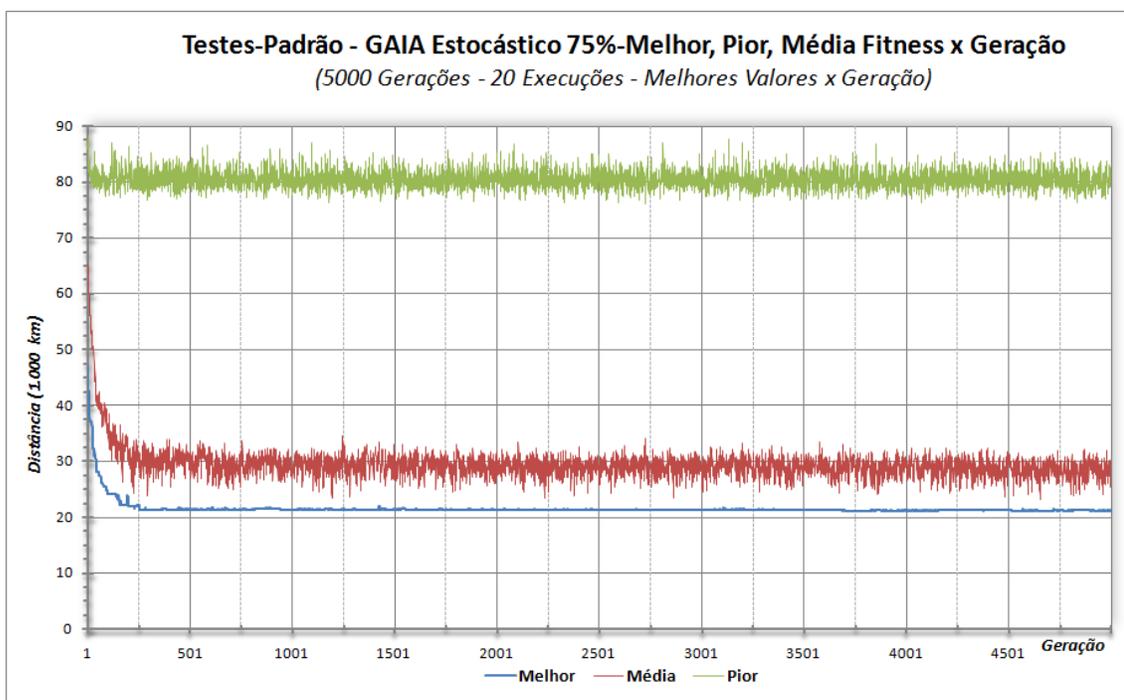


Gráfico 6.17 – GAIA Estocástico 95%: Média dos Valores x Geração

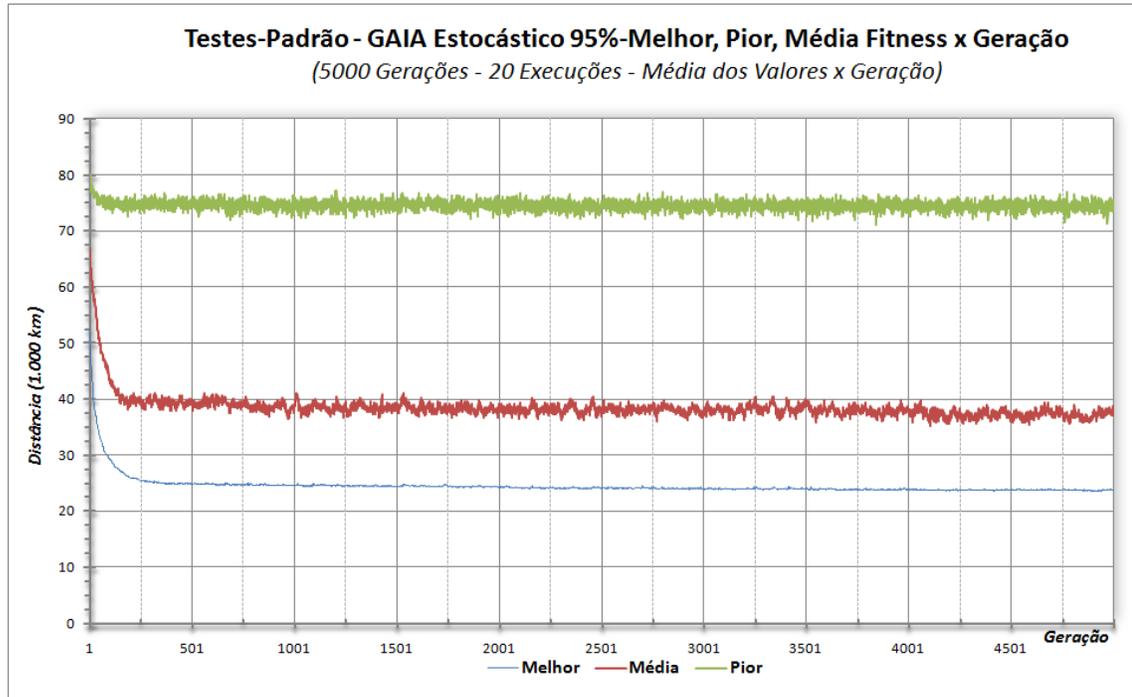


Gráfico 6.18 – GAIA Estocástico 95%: Melhores Valores x Geração

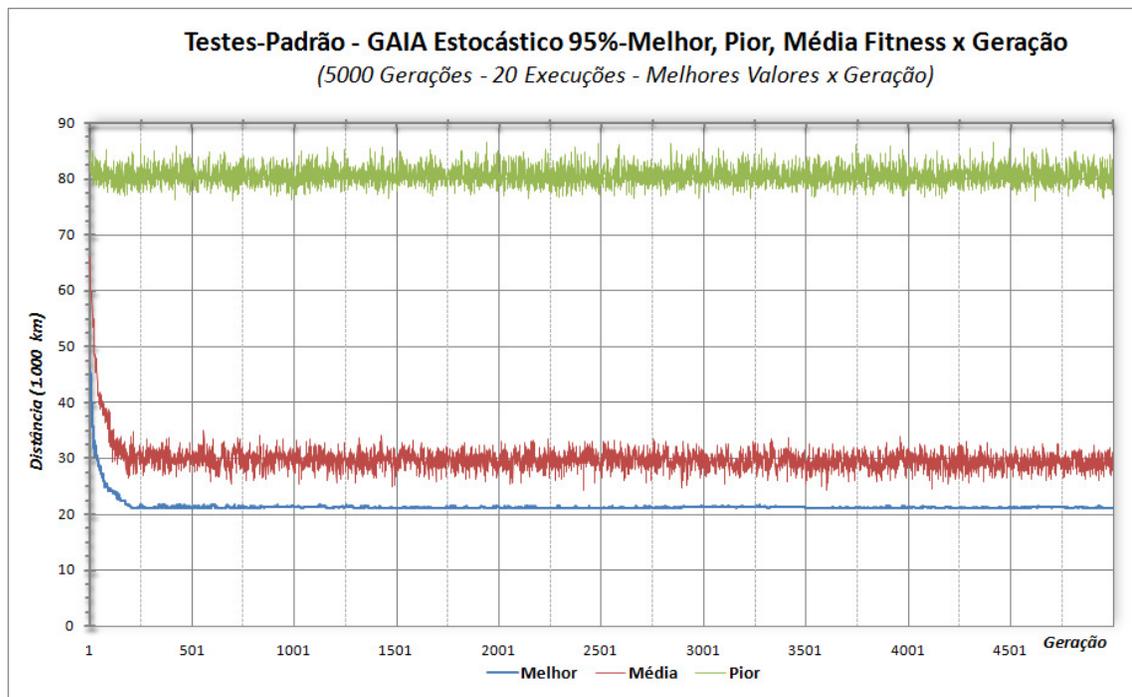


Gráfico 6.19 – GAIA Estocástico 98%: Média dos Valores x Geração

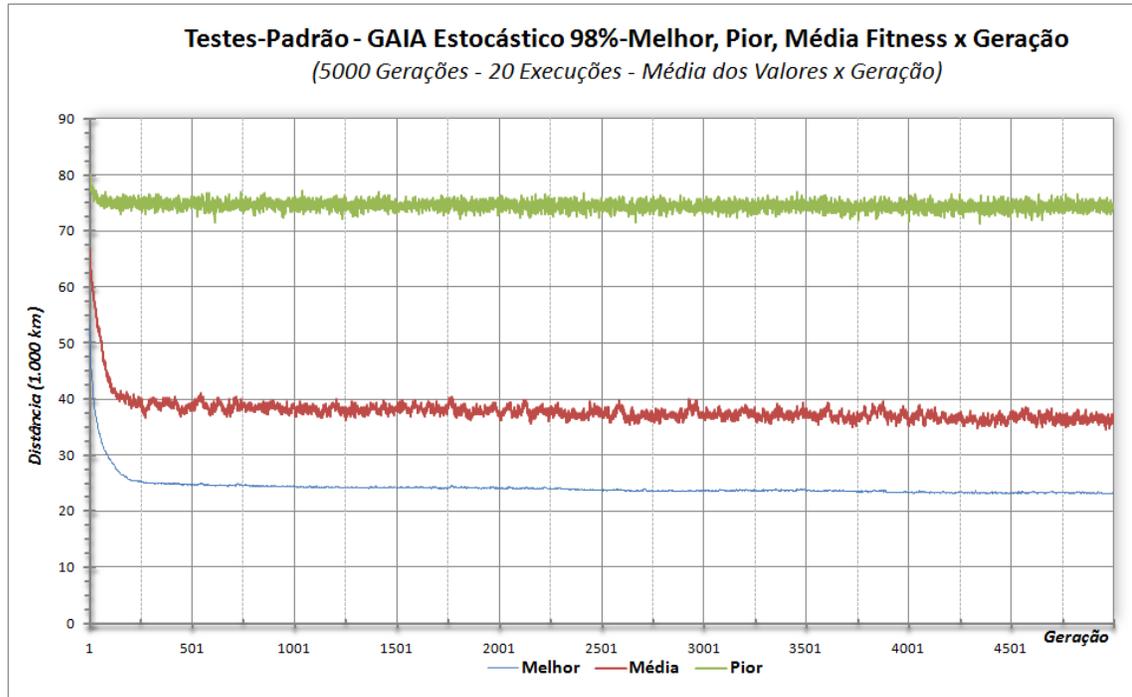


Gráfico 6.20 – GAIA Estocástico 98%: Melhores Valores x Geração

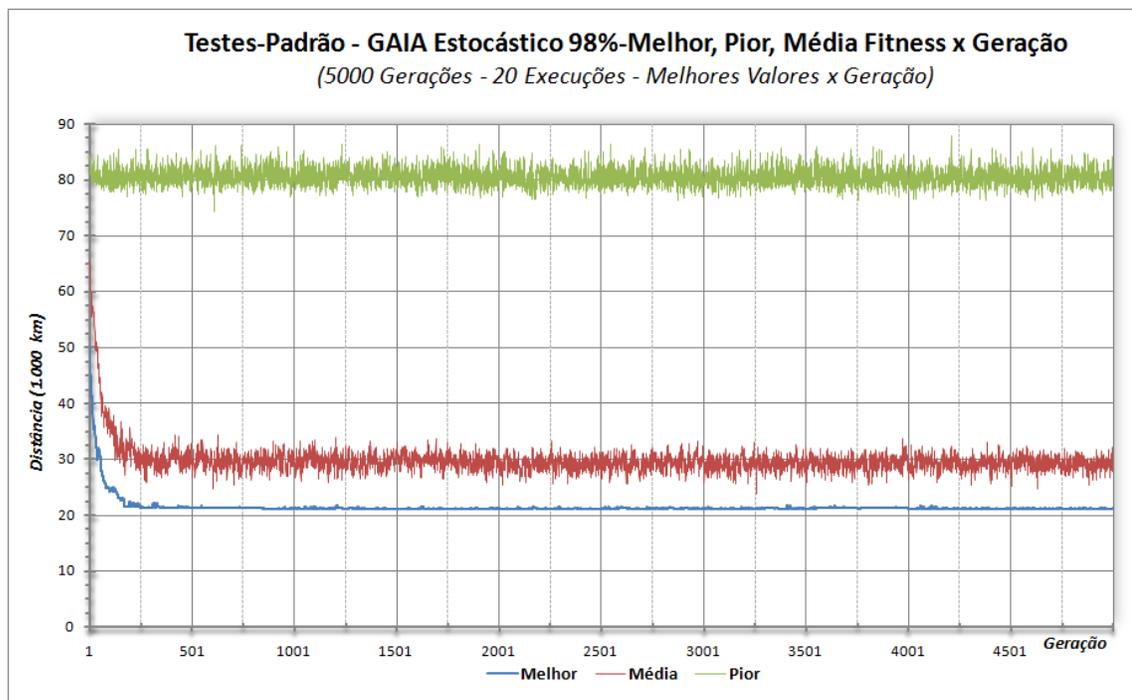


Gráfico 6.21 – GAIA Estocástico 99,8%: Média dos Valores x Geração

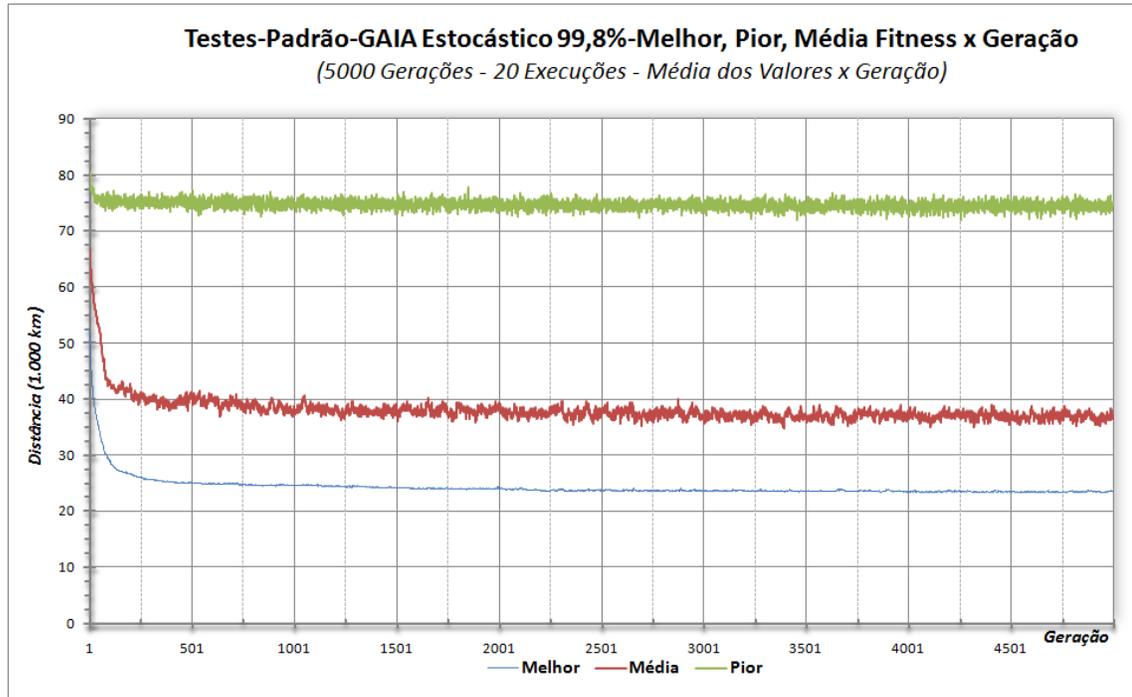


Gráfico 6.22 – GAIA Estocástico 99,8%: Melhores Valores x Geração

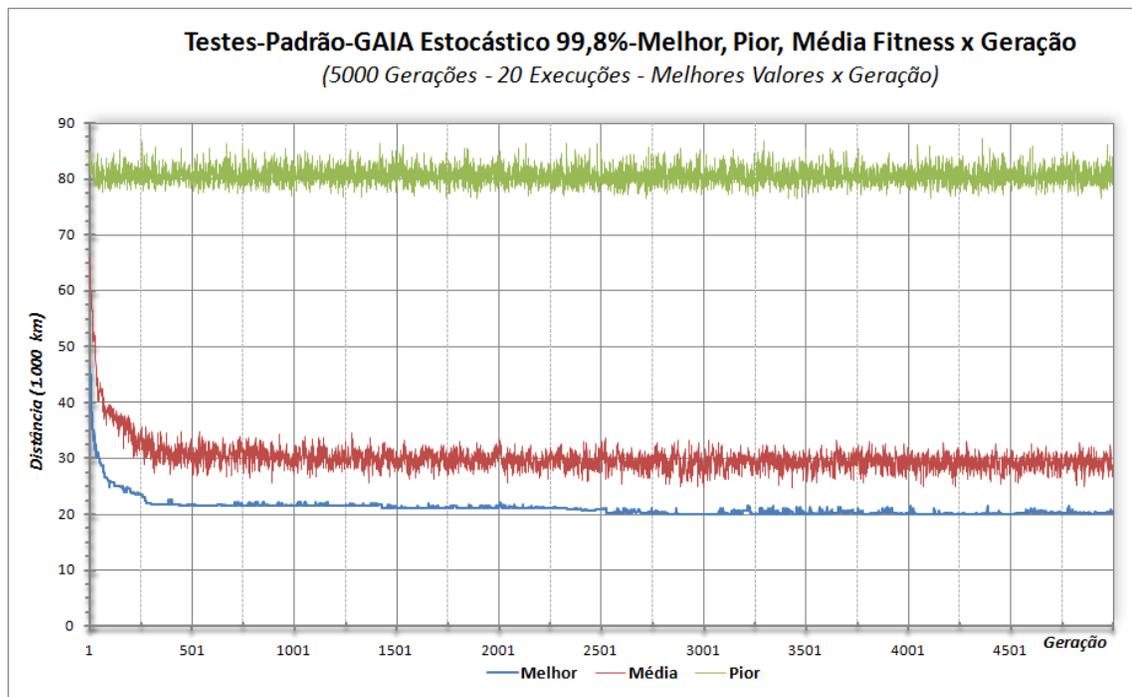
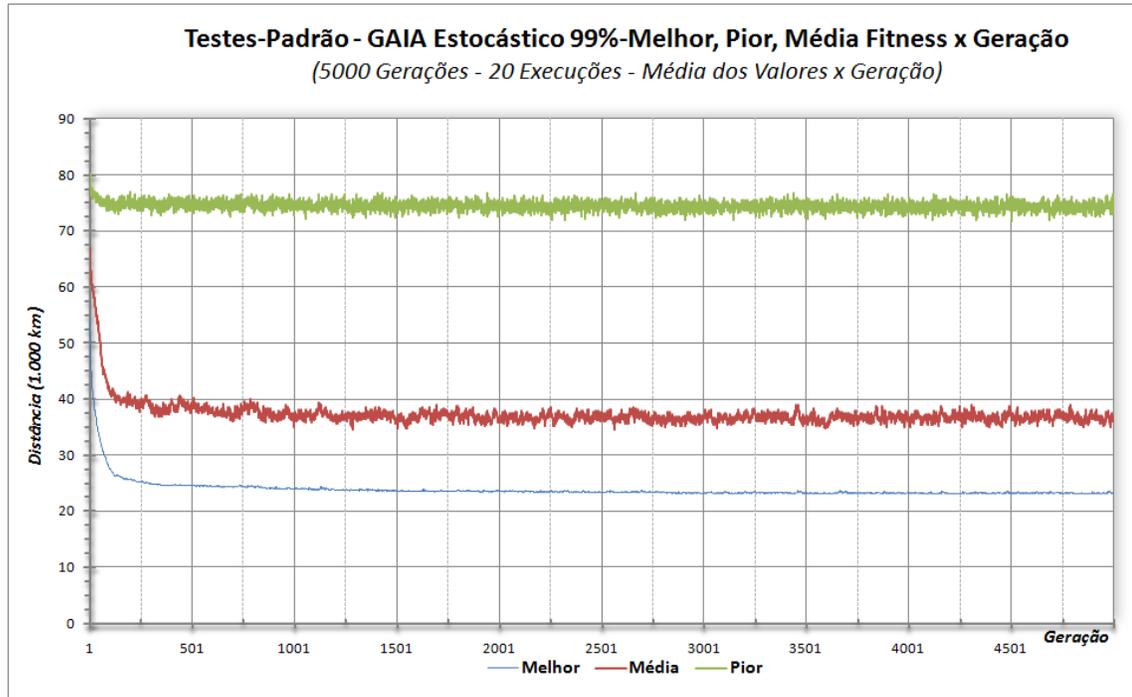
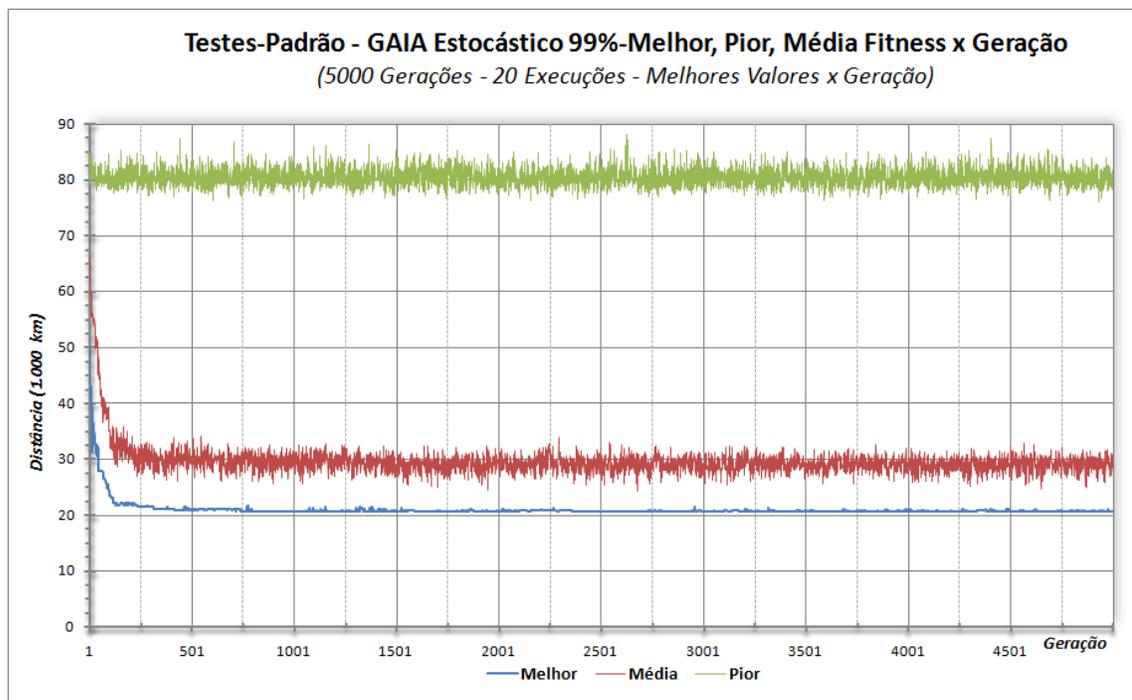


Gráfico 6.23 – GAIA Estocástico 99%: Média dos Valores x Geração**Gráfico 6.24 – GAIA Estocástico 99%: Melhores Valores x Geração**

6.3.1.3 GAIA Periódico

Nesta abordagem periódica do GAIA, foram definidas 06 configurações diferentes, com um total de 120 testes, representados aqui de forma sintetizada em 12 gráficos. No geral, todos apresentam um comportamento semelhante, onde a diferença se dá apenas na força da intervenção e suas conseqüências, que são diretamente proporcionais ao tamanho do período.

Em síntese, nos gráficos 6.25, 6.26, 6.27, 6.28, 6.29, 6.30, 6.31, 6.32, 6.33, 6.34, 6.35 e 6.36, observa-se claramente a seqüência que já fora apresentada em outras seções do trabalho: um período de convergência, como em uma execução tradicional de um AG; seguido de um período pós-intervenção, onde boa parte da população é eliminada, os parâmetros de recombinação são diminuídos e os de mutação aumentados, para garantir a sobrevivência dos possíveis novos blocos construtores reinseridos na população; seguido de uma nova intervenção, que restabelece os parâmetros normais de um AG que retorna a população ao seu curso de evolução.

Também é possível notar que, períodos muito espaçados de intervenção, com suas “violentas” conseqüências para a população, podem não ter tempo suficiente para deixá-la se recuperar e evoluir, desperdiçando assim boa parte das gerações estipuladas. Fato que fica evidenciado, nas situações apresentadas do GAIA periódico 1.000 e 2.500, nos gráficos 6.33, 6.34, 6.35 e 6.36.

Mesmo assim, em todas, pode-se notar mais uma vez, que o “ciclo” de intervenções espaçadas, tem teve sua eficácia comprovada, pois, tal qual, constatado na natureza, sempre após um período de grande destruição “pós-apocalíptico”, a evolução novamente ocorre e encontra indivíduos (soluções) ainda melhores.

Quanto a variabilidade, considerada peça chave para manutenção dos possíveis “blocos construtores” dentro da população, foi constatado novamente que após cada uma das intervenções, esta é reintroduzida na população.

Em resumo, em termos de comparação entre os diferentes valores para o parâmetro de periodicidade de ocorrência do agente, pode-se notar que sua escolha irá depender da quantidade de execuções que se têm disponíveis. Porém, como já dito, quanto maior for o espaço, mais tempo será desperdiçado, então, pelo menos nas condições deste trabalho, sugere-se algo em torno de 50 a no máximo 500 gerações.

Gráfico 6.25 – GAIA Periódico 20: Média dos Valores x Geração

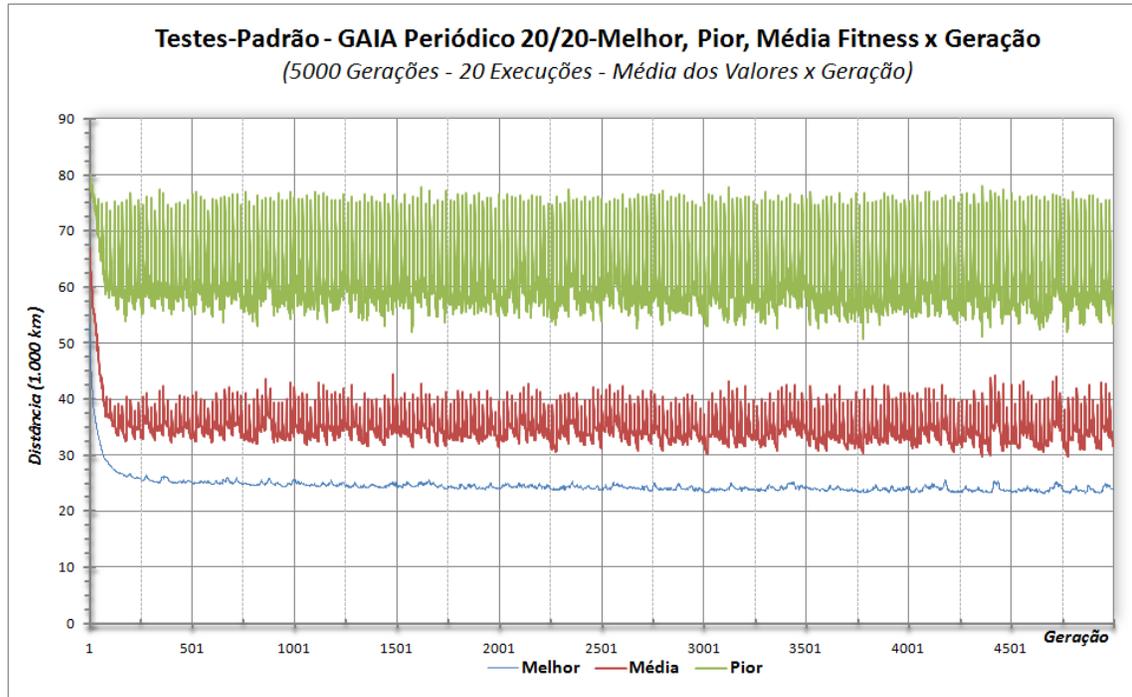


Gráfico 6.26 – GAIA Periódico 20: Melhores Valores x Geração

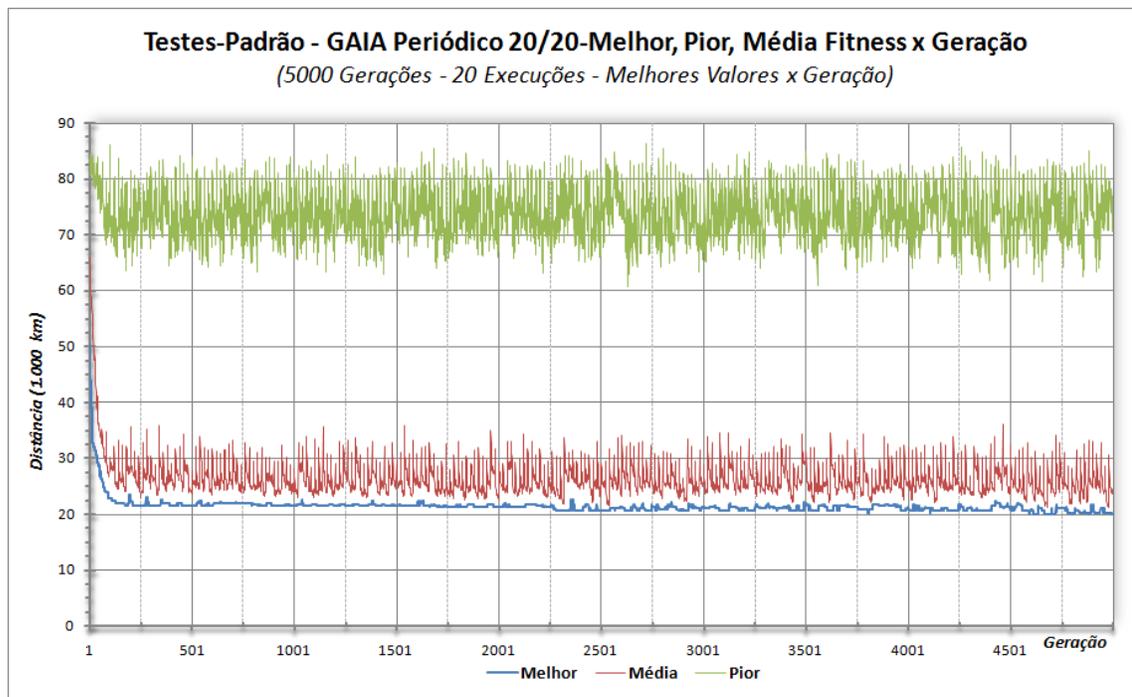


Gráfico 6.27 – GAIA Periódico 50: Média dos Valores x Geração

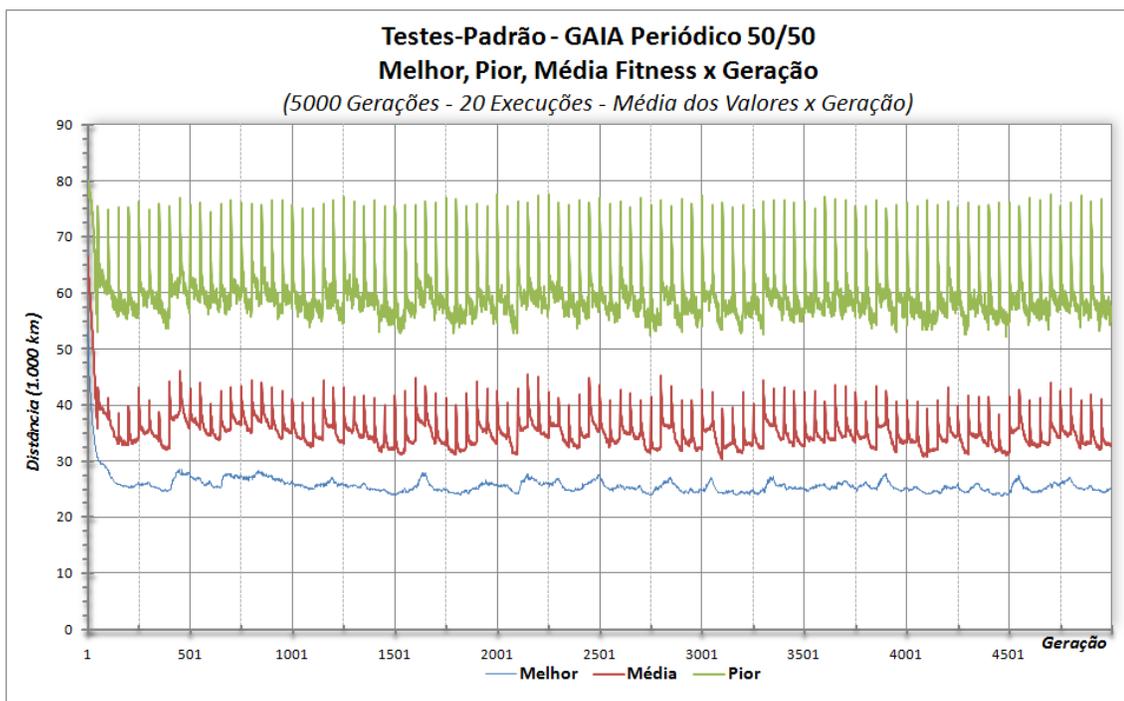


Gráfico 6.28 – GAIA Periódico 50: Melhores Valores x Geração

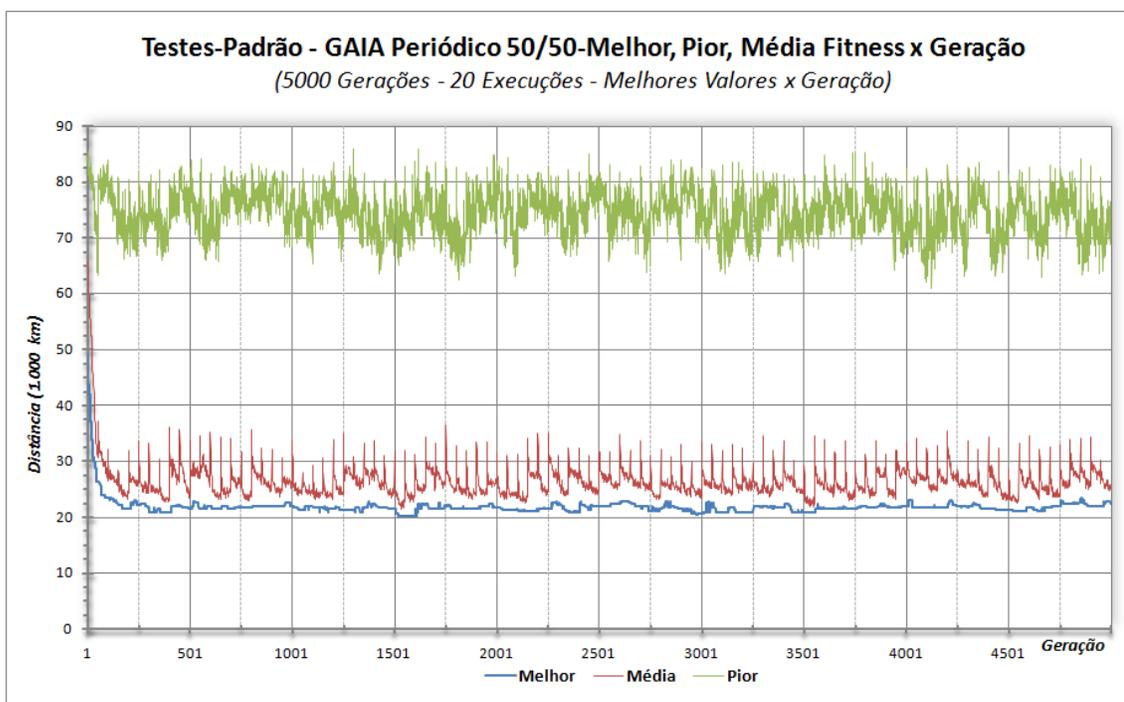


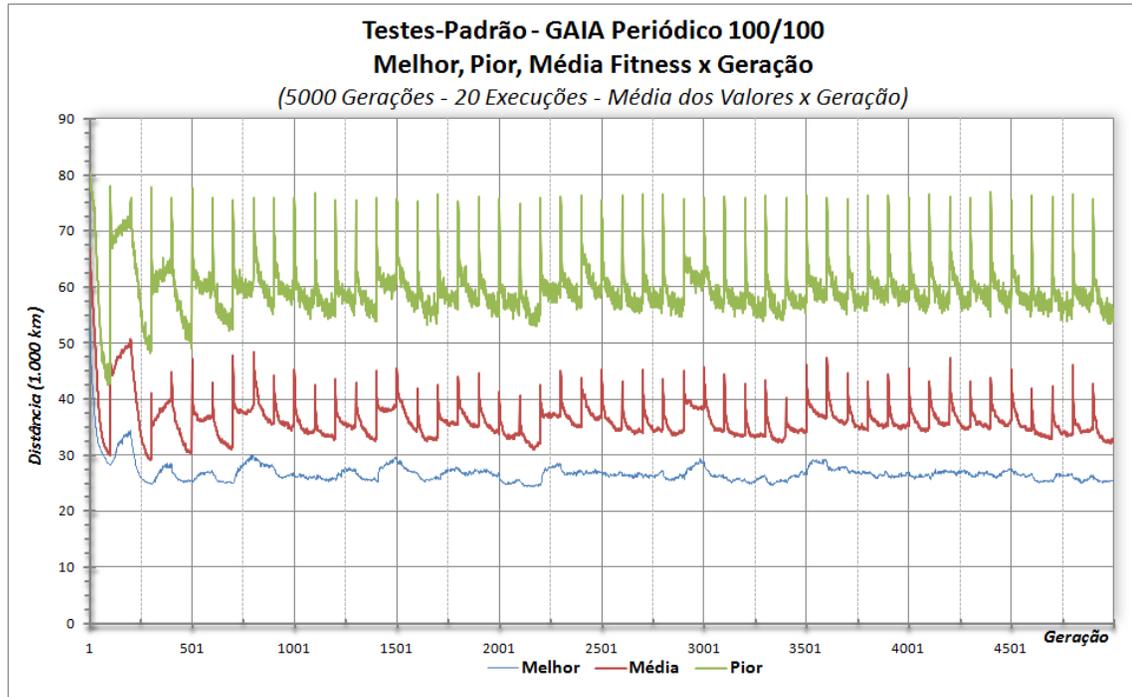
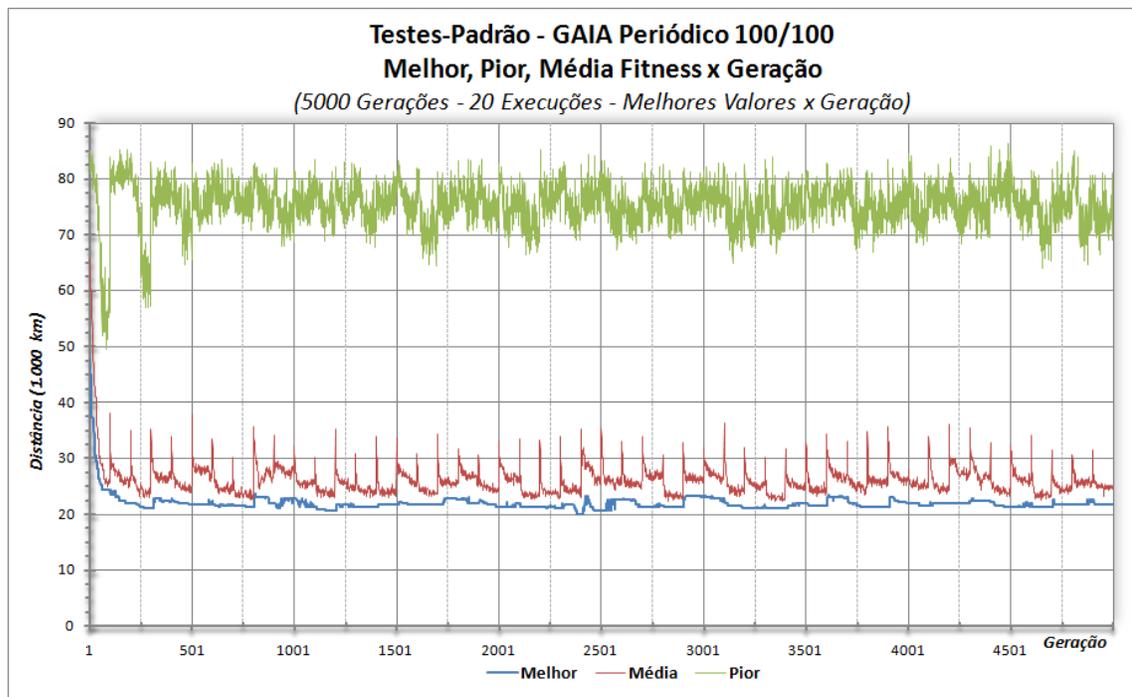
Gráfico 6.29 – GAIA Periódico 100: Média dos Valores x Geração**Gráfico 6.30 – GAIA Periódico 100: Melhores Valores x Geração**

Gráfico 6.31 – GAIA Periódico 500: Média dos Valores x Geração

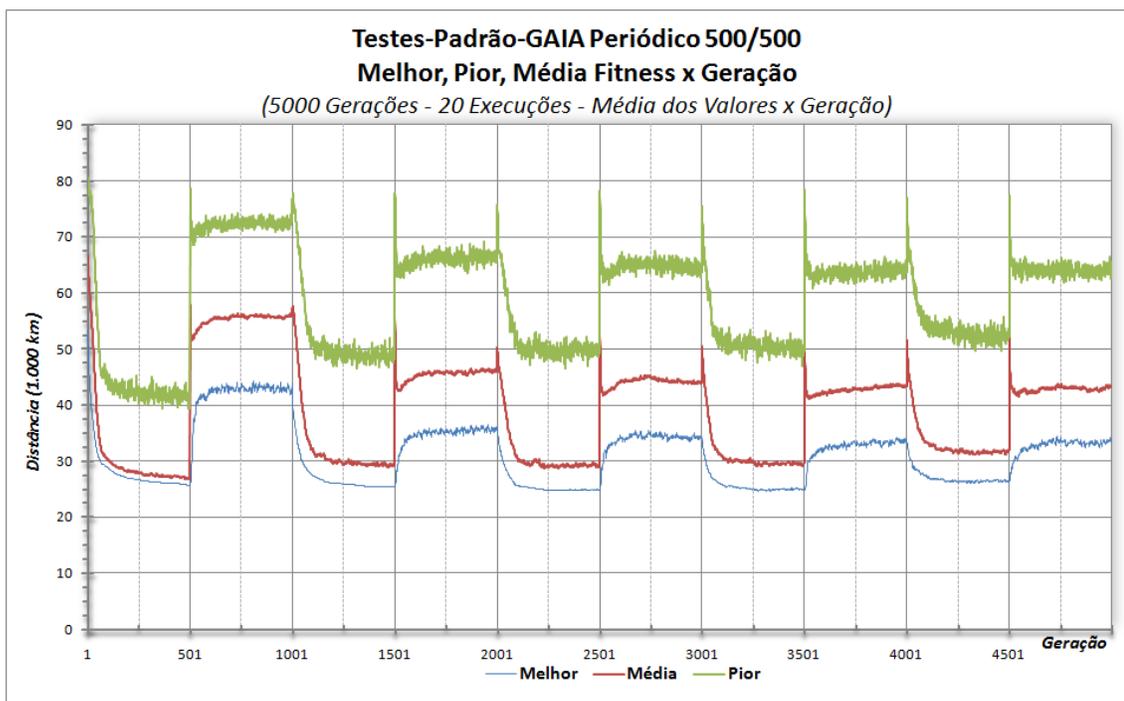


Gráfico 6.32 – GAIA Periódico 500: Melhores Valores x Geração

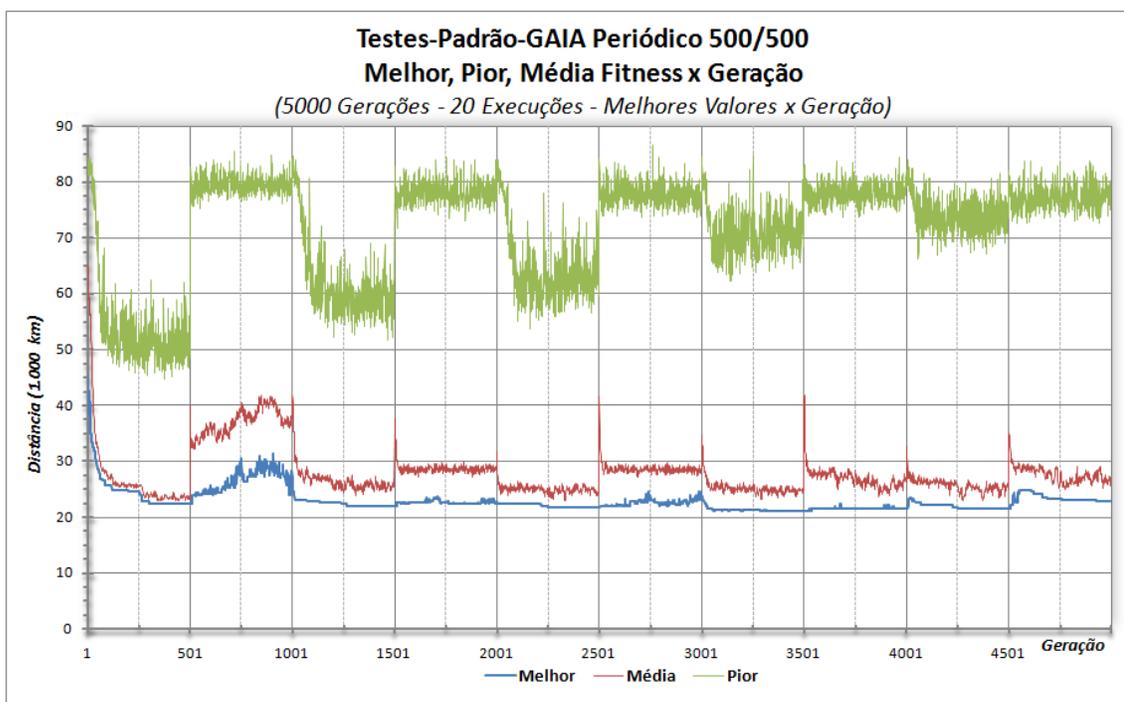


Gráfico 6.33 – GAIA Periódico 1000: Média dos Valores x Geração

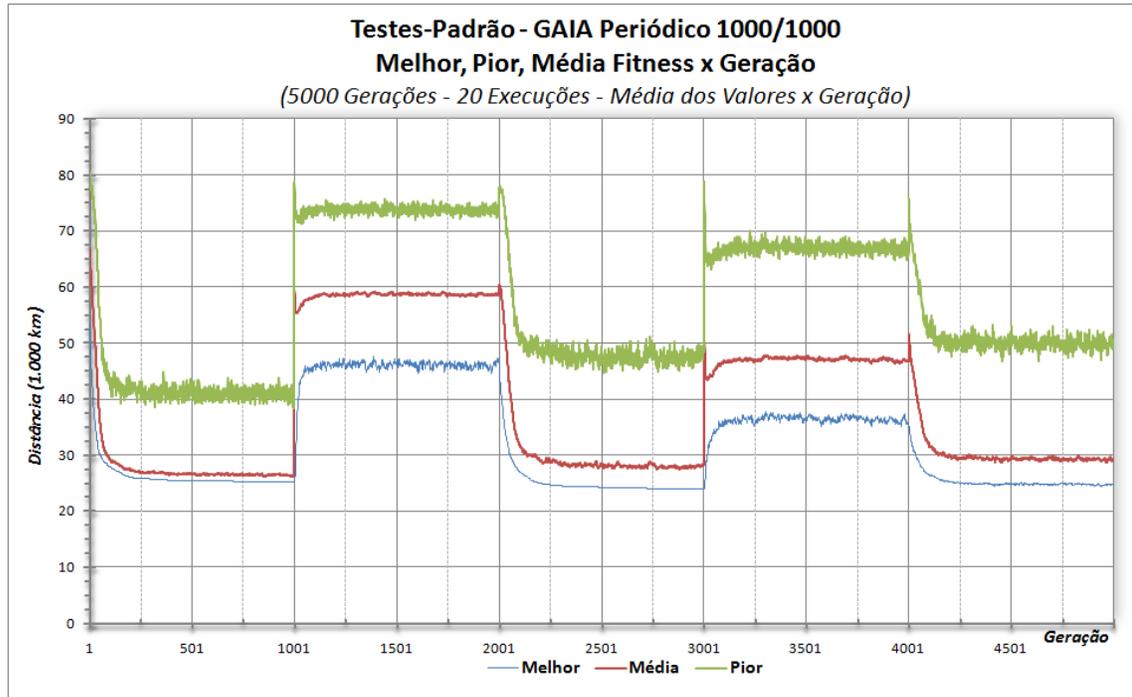


Gráfico 6.34 – GAIA Periódico 1000: Melhores Valores x Geração

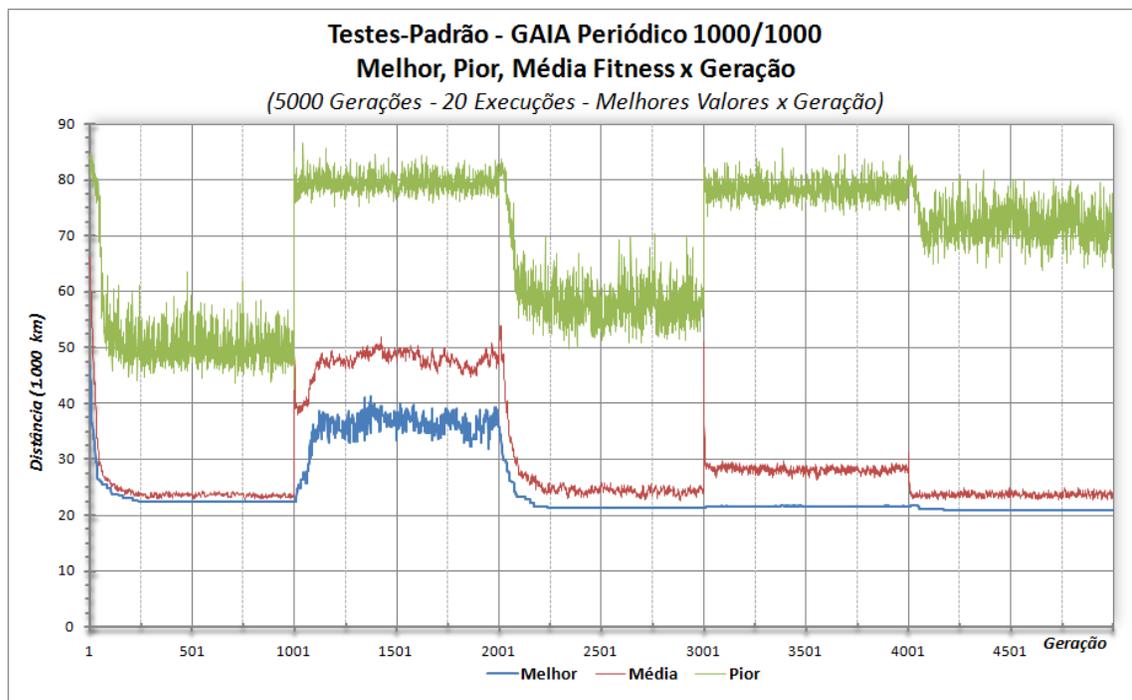


Gráfico 6.35 – GAIA Periódico 2500: Média dos Valores x Geração

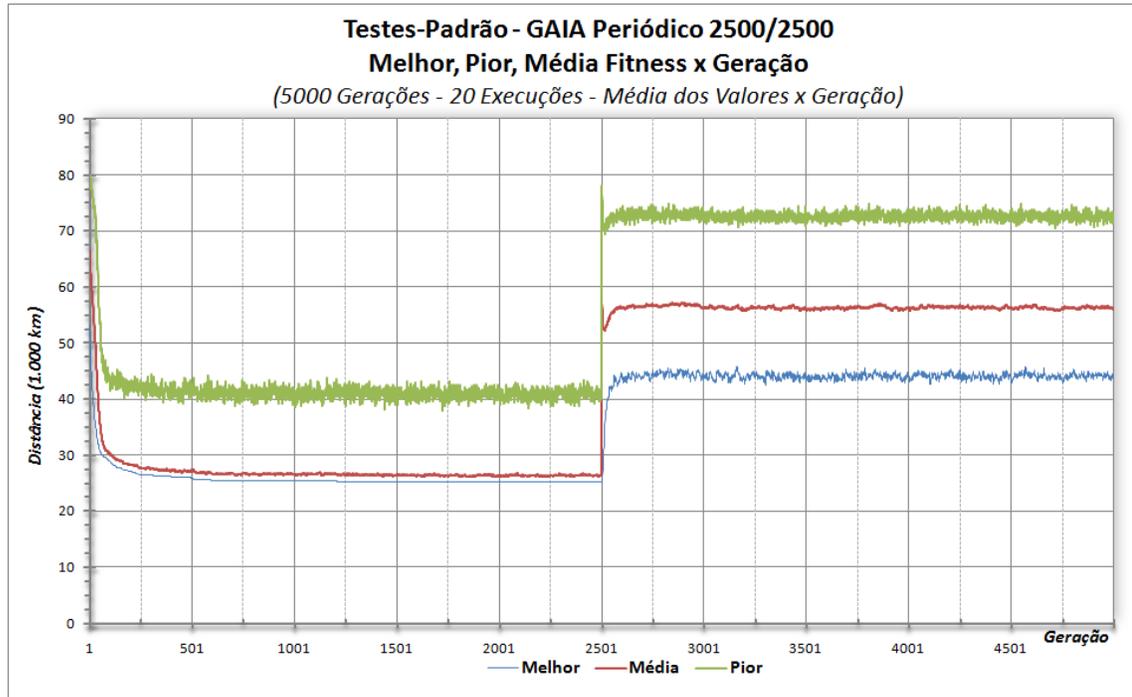
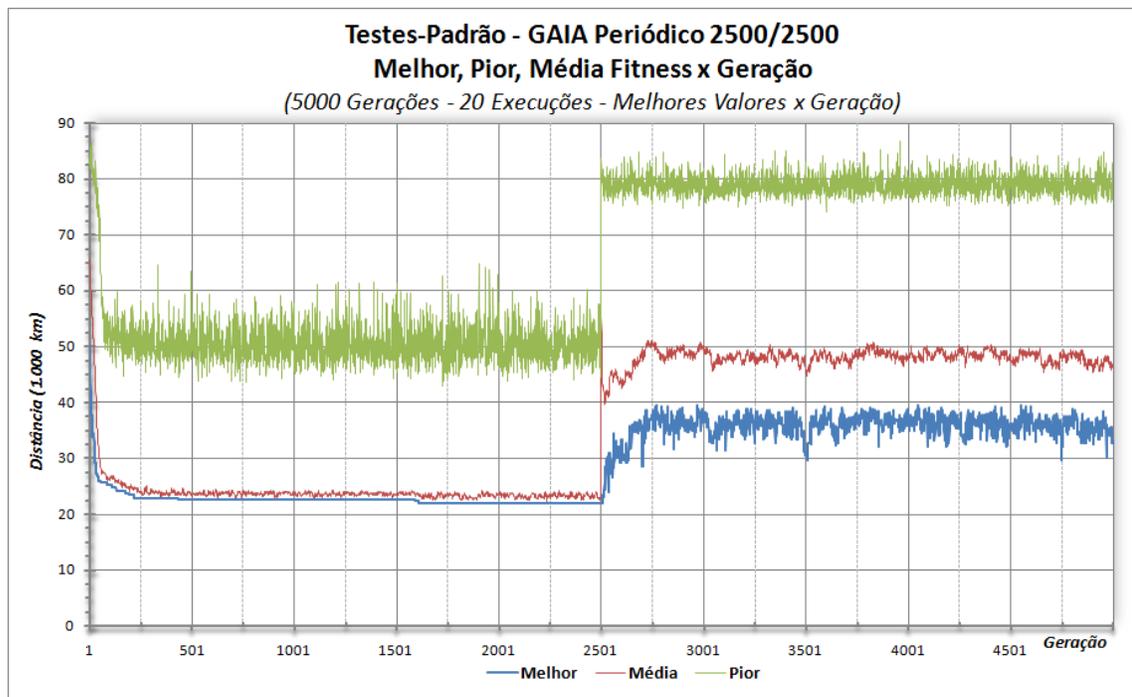


Gráfico 6.36 – GAIA Periódico 2500: Melhores Valores x Geração



6.3.1.4 Hawk-Dove Roleta e Roleta Simples

No intuito de prover ao trabalho dados comprobatórios da eficácia do GAIA frente às execuções tradicionais de AGs e até mesmo de uma abordagem híbrida, foram definidas duas configurações “normais”, ou seja, sem o uso do GAIA, porém, utilizando as mesmas população inicial e plataforma de testes (*Beagle XP*).

Em ambas as configurações, representadas pelos gráficos 6.37, 6.38, 6.39 e 6.40, fica evidente a excessiva convergência e o péssimo aproveitamento do montante de gerações para a exploração do espaço de busca.

Na tradicional roleta simples, isto considerando a média de 20 execuções, após pouco mais do que a 250ª geração, encontrou-se um ótimo local que, dada a característica “da sua força de gravidade” (analogia apresentada no Capítulo V) e nenhum tipo de controle, “arrastou” toda a população para aquele ponto no espaço e, rapidamente encerrou com praticamente toda a variabilidade, a qual ainda foi pouco mantida somente pela casualidade do operador de mutação.

Na abordagem híbrida de BRITO (2004), a situação não foi diferente, a “pressão seletiva” até chega a ser controlada, mas de forma tão insipiente, que a colocação deveria ser que ela é “retardada” e não controlada. Isto porque, antes de 800 gerações a população já estava estagnada num ótimo local, mais uma vez contando apenas com a sorte da casualidade do operador de mutação.

Sendo assim, sobre a perspectiva da variabilidade e melhor exploração do espaço de busca, ficou evidente que em todos os casos até agora apresentados, a intervenção do agente faz efetivamente a diferença na execução.

Gráfico 6.37 – Hawk-Dove Roleta: Média dos Valores x Geração

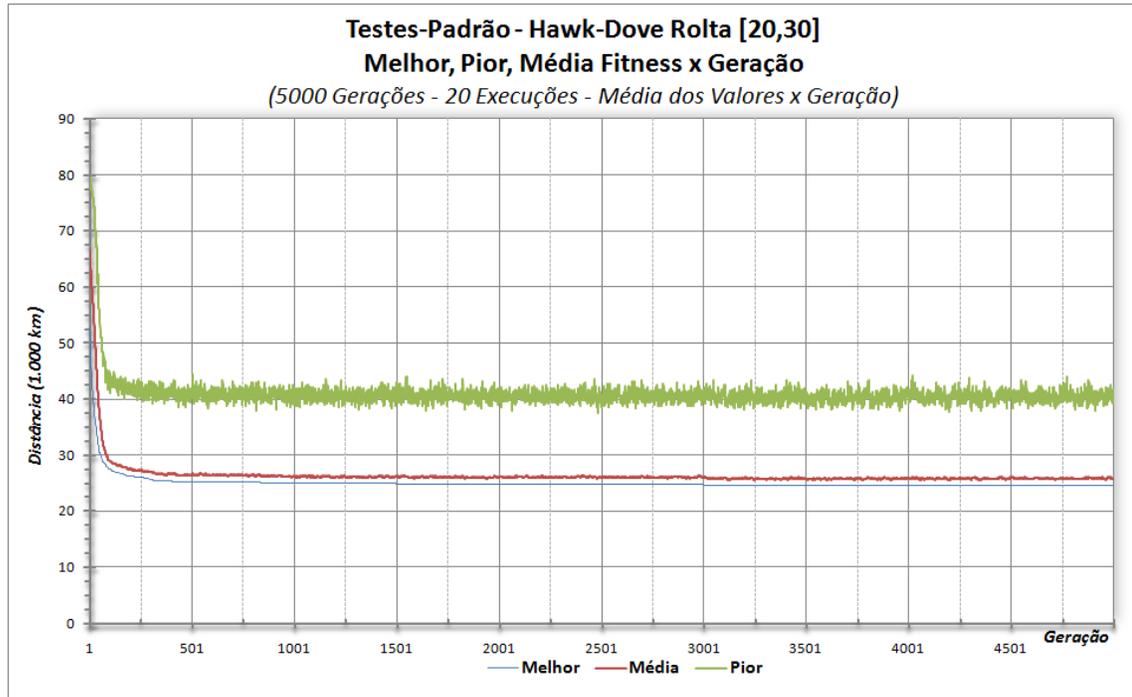


Gráfico 6.38 – Hawk-Dove Roleta: Melhores Valores x Geração

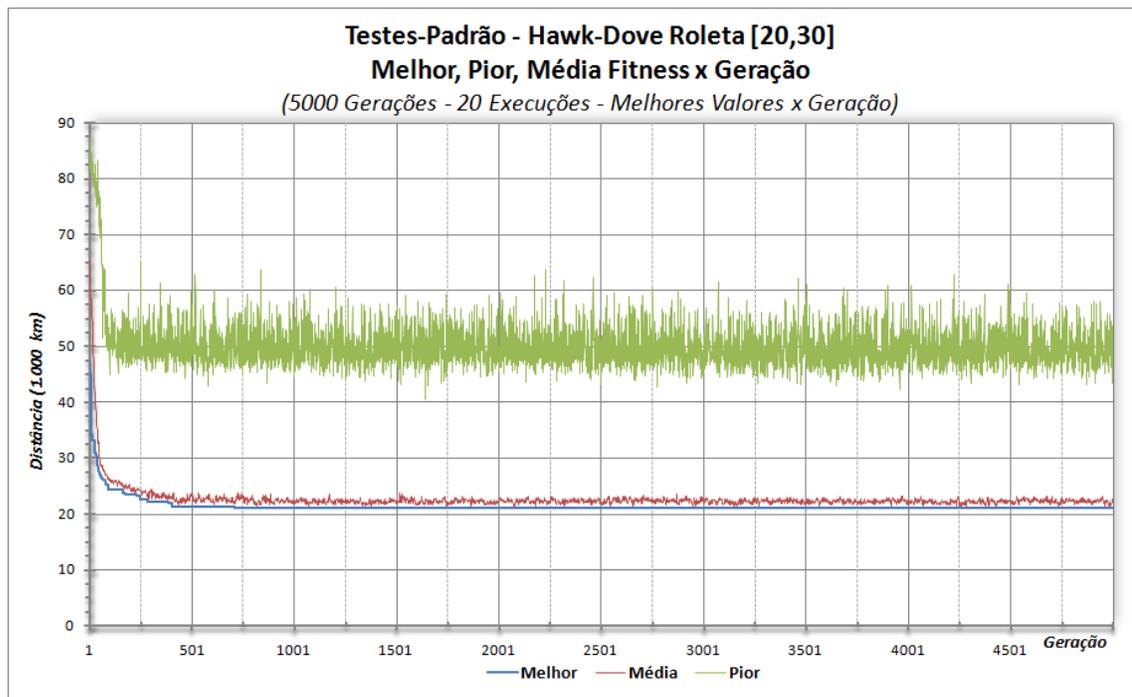


Gráfico 6.39 – Roleta Simples: Média dos Valores x Geração

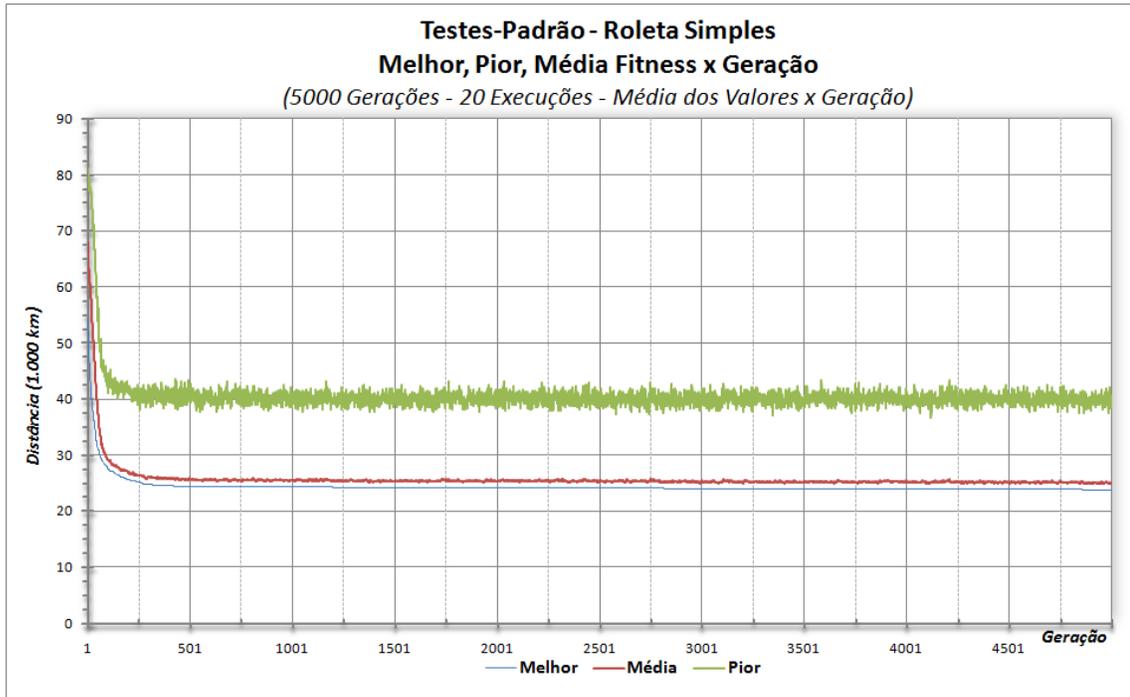
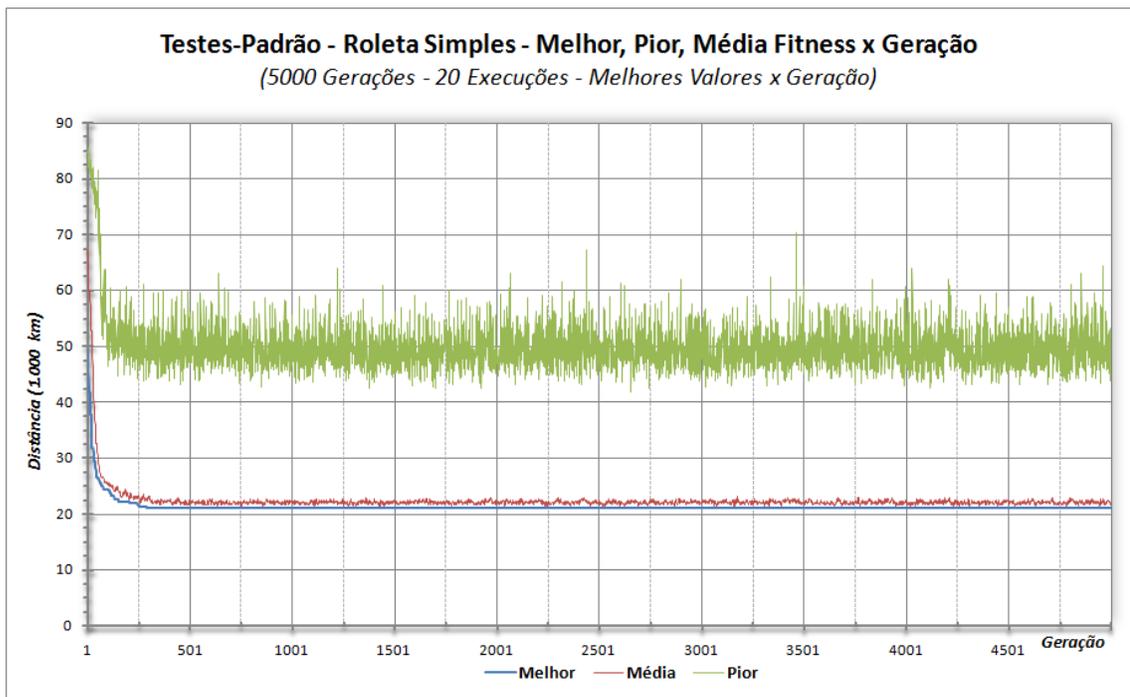


Gráfico 6.40 – Roleta Simples: Melhores Valores x Geração



6.3.2 Super-Testes

Na seção anterior, em praticamente todos os gráficos apresentados, gerou-se a mesma expectativa e questionamentos levantados na primeira publicação desta proposta em BRITO et. al. (2006): “e se fossem possíveis mais execuções, o que aconteceria com a evolução? Continuará? Seriam alcançados resultados ainda melhores?”.

Por conta dessas expectativas, é que o presente trabalho elaborou, o que aqui chamou de “Super-Testes”, execuções dez vezes maiores que as da seção anterior, com 50.000 gerações cada, apresentadas e analisadas neste seção.

Assim como nos “Testes-Padrão”, para os “Super-Testes”, foram definidas 09 configurações distintas, baseadas no tipo de GAIA utilizado, sendo: 01 para o modo permanente do GAIA; 03 para o periódico, variando-se o intervalo de gerações para a intervenção; 03 para o modo estocástico, variando-se a probabilidade de ocorrência; e, para efeitos de comparação interna, 01 configuração de AG Canônico com a roleta simples e 01 configuração com o método Hawk-Dove Roleta, ambos sem o uso do GAIA.

De mesma forma, a fim de diminuir a casualidade presente em uma única execução, para cada uma das configurações, foram realizados 20 execuções, totalizando 180 testes para este grupo.

A tabela 6.3, descrita a seguir, apresenta todas as configurações realizadas, e também alguns dados para uma “primeira” análise de performance, valendo-se para tal do melhor valor encontrado e da primeira geração em que ocorreu.

Tabela 6.3 – Super-Testes: Execuções e Resultados (Sintético)

ID	NOME / DESC.	GAIA		QTD	RESULTADOS	
		TIPO	PARAM.		MELHOR (KM)	1ª GER. ENCONT.
01	Super-Teste GAIA Permanente (ST-GAIA-Permanente)	Permanente	–	20	20.989	12.947
02	Super-Teste GAIA Periódico 500 (ST-GAIA-Period-500)	Periódico	500	20	20.752	22.668
03	Super-Teste GAIA Periódico 1.000 (ST-GAIA-Period-1000)	Periódico	1.000	20	20.586	26.032
04	Super-Teste GAIA Periódico 10.000 (ST-GAIA-Period-10000)	Periódico	10.000	20	20.093	5.780
05	Super-Teste GAIA Estocástico 25% (ST-GAIAEstocástico-25)	Estocástico	25%	20	20.752	1.523
06	Super-Teste GAIA Estocástico 50% (ST-GAIAEstocástico-50)	Estocástico	50%	20	20.093	1.525
07	Super-Teste GAIA Estocástico 98% (ST-GAIAEstocástico-98)	Estocástico	98%	20	20.586	1.187
08	Super-Teste Hawk-Dove Roleta (20,30) (ST-HDR2030)	Sem GAIA	–	20	20.989	22.184
09	Super-Teste Roleta Simples (ST-RoletaSimples)	Sem GAIA	–	20	21.069	23.960
TOTAL DE TESTES E “MELHOR RESULTADO”				180	20.093	1.525

Tomando por base a primeira avaliação apresentada nesta tabela, pode-se, ainda que de forma não efetivamente concreta, responder à expectativa gerada com as maiores execuções. De fato, os resultados encontrados são melhores, e são fruto da continuidade da evolução graças a variabilidade mantida pela intervenção do GAIA. Todos são menores que 21.000 km e, novamente, o possível ótimo global, 20.093 km, foi encontrado, em mais de uma ocasião.

De forma análoga à estrutura apresentada para os “Testes-Padrão”, nas próximas subseções serão exibidos os mesmos dois tipos de gráficos: o primeiro, apresenta três séries temporais para os piores, melhores e média dos valores de *fitness* de cada geração, considerando para tal a média aritmética simples entre cada um dos 20 testes realizados; já o segundo tipo, apresenta três séries temporais para os piores, melhores e média dos valores de *fitness* de cada geração, considerando sempre a melhor situação entre os 20 testes realizados.

Deve-se atentar que, apesar das 50.000 execuções geradas, devido à limitações da ferramenta, e de certa forma também, visuais, para todos os gráficos dos “Super-Testes” apenas as 32.000 primeiras gerações são exibidas. Como será exposto a partir da próxima subseção, esta limitação não é impeditiva para a avaliação, uma vez que em todas as execuções ao longo do tempo um padrão é repetido.

6.3.2.1 GAIA Permanente

De forma muito similar a sua versão de 5.000 gerações exibida anteriormente, nota-se para este GAIA Permanente, por meio dos gráficos 6.41 e 6.42, que ao longo de toda a execução acontece uma suave intervenção, com constante manutenção da variabilidade, verificada pelas séries “média” e “pior”, além de sucessiva evolução, observada pela queda contínua dos valores da série “melhor”.

Mais uma vez, graças ao constante controle do GAIA, apenas as regras mais “brandas” são ativadas, ou seja, as que têm maior influência sobre os percentuais de mutação e recombinação, com pouco ou nenhuma influência sobre o percentual de indivíduos sobreviventes.

Gráfico 6.41 – Super-Teste GAIA Permanente: Média dos Valores x Geração

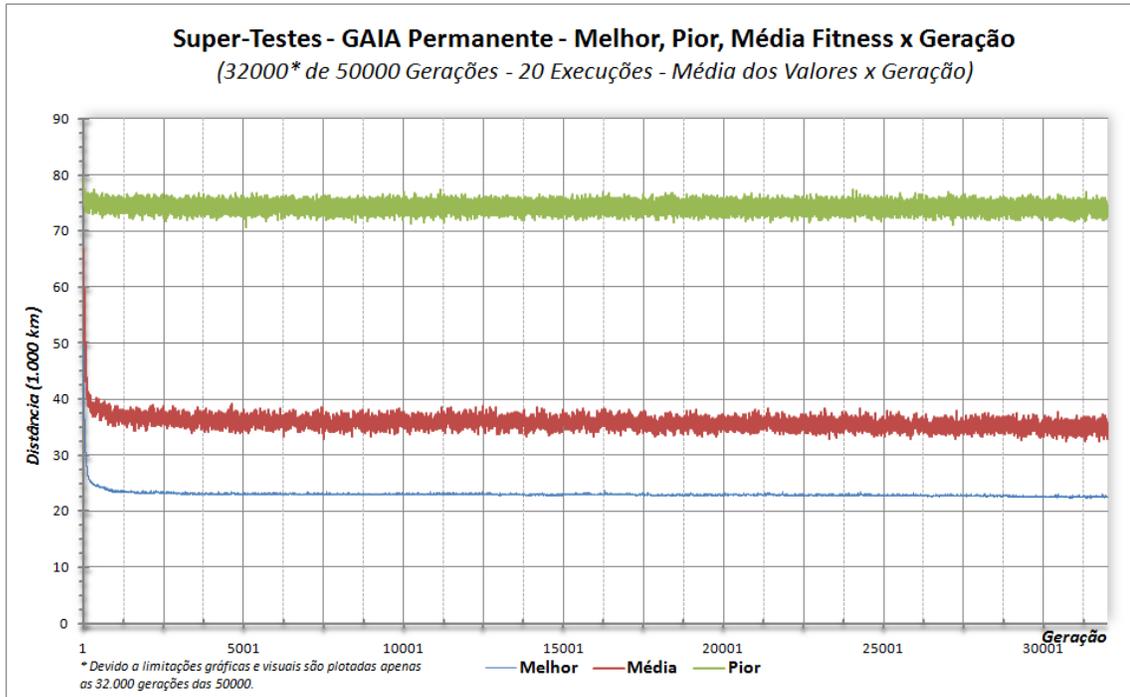
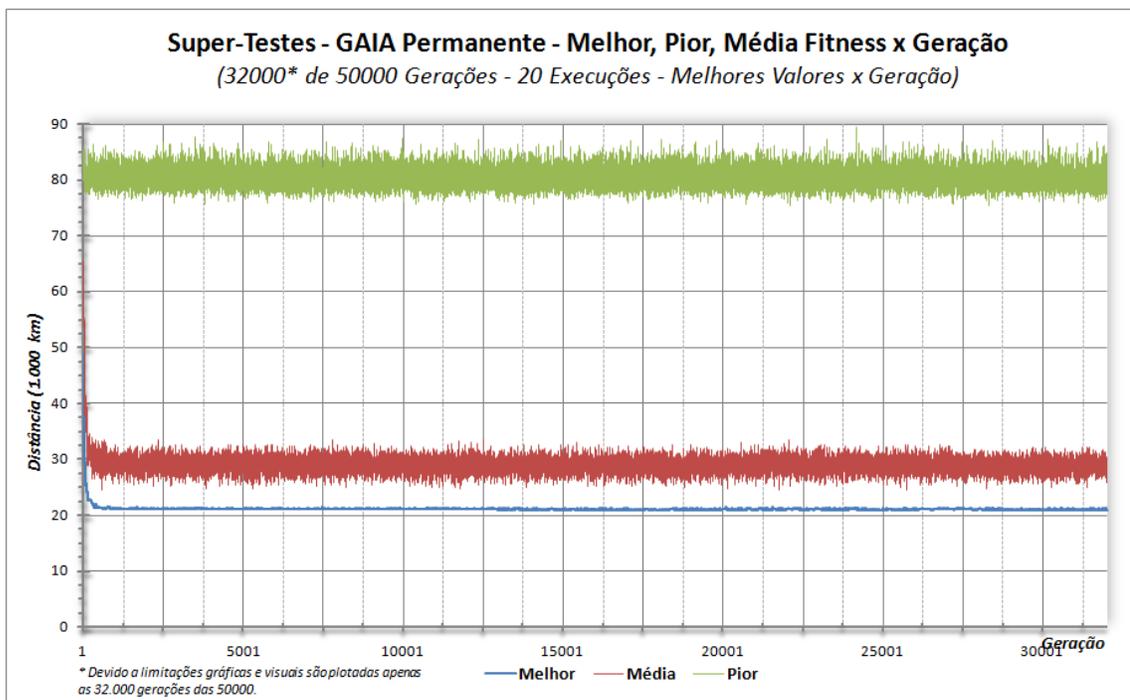


Gráfico 6.42 – Super-Teste GAIA Permanente: Melhores Valores x Geração



6.3.2.2 GAIA Estocástico

Na vertente “Super-Testes” para o tipo GAIA estocástico, buscou-se a utilização de um exemplar de cada grupo do modelo de particionamento dos possíveis valores probabilísticos de ocorrência do GAIA, apresentado na seção 6.3.1.2. Desta forma, foram definidas três configurações: sendo uma de “alta”, uma de “média” e outra de “baixa” probabilidade de ocorrência, respectivamente, 25%, 50% e 98%.

Nas configurações, representadas por meio dos gráficos 6.43, 6.44, 6.45, 6.46, 6.47 e 6.48, nota-se que, a medida que ocorrem com maior frequência e o número de gerações plotadas é significativamente maior, fica ainda mais difícil observar os pequenos “picos” de piora, relacionados ao momento de intervenção do GAIA.

Contudo, as mesmas características identificadas nos “Testes-Padrão” deste tipo podem ser observadas, tais como: manutenção de variabilidade, o eterno “ciclo piora-melhora” e a constante busca pelo valor ótimo global que, mais uma vez, levou as execuções a encontrarem, o provável ótimo global do problema.

Em resumo, tem-se mais uma vez que utilizar probabilidade com valores de “baixo” a “médio” garantem um melhor “ciclo piora-melhora”, ou seja uma melhor exploração do potencial da base de regras do GAIA e, a conseqüente maior probabilidade de se encontrar o ótimo global do problema.

Gráfico 6.43 – Super-Teste GAIA Estocástico 25%: Média dos Valores x Geração

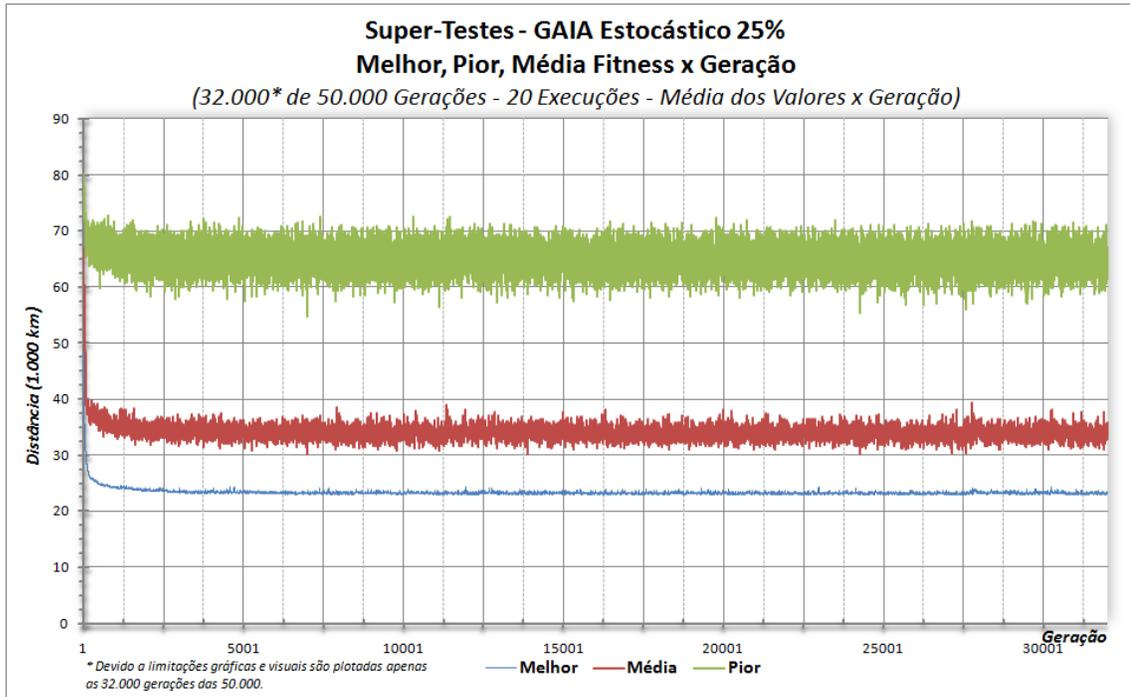


Gráfico 6.44 – Super-Teste GAIA Estocástico 25%: Melhores Valores x Geração

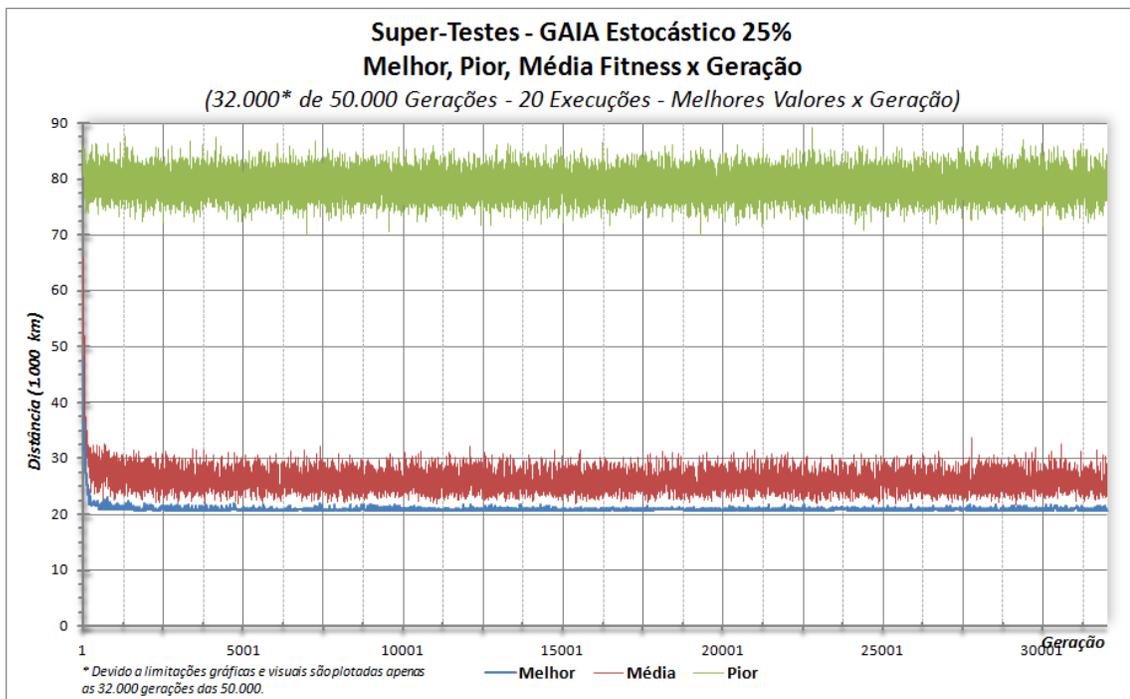


Gráfico 6.45 – Super-Teste GAIA Estocástico 50%: Média dos Valores x Geração

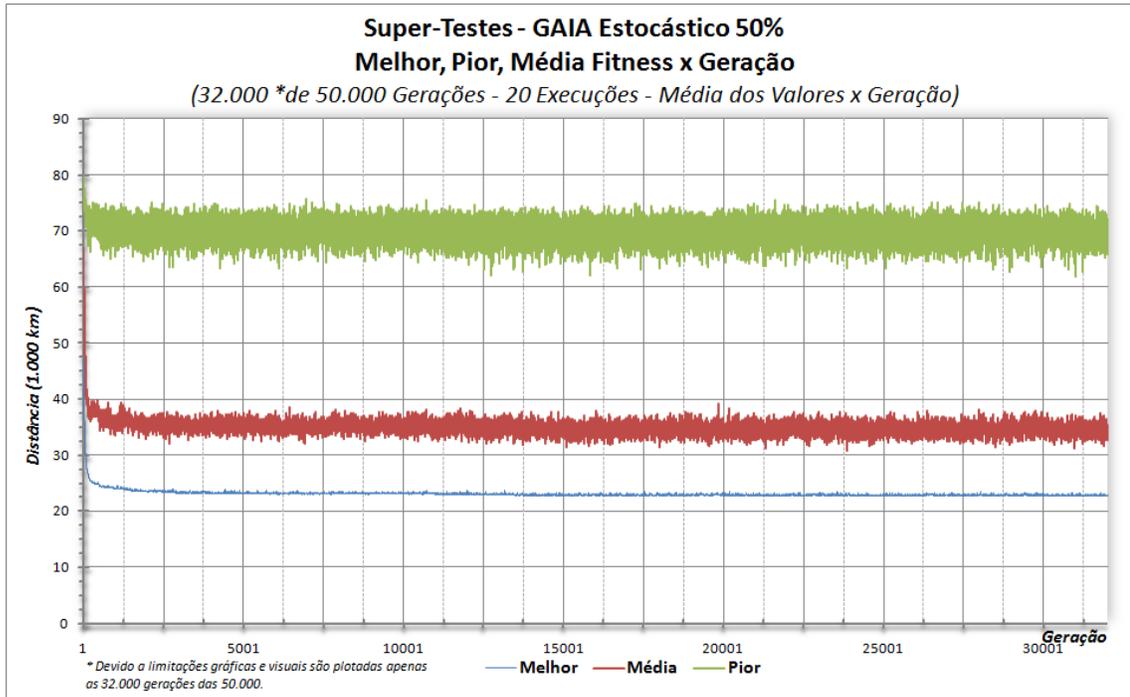


Gráfico 6.46 – Super-Teste GAIA Estocástico 50%: Melhores Valores x Geração

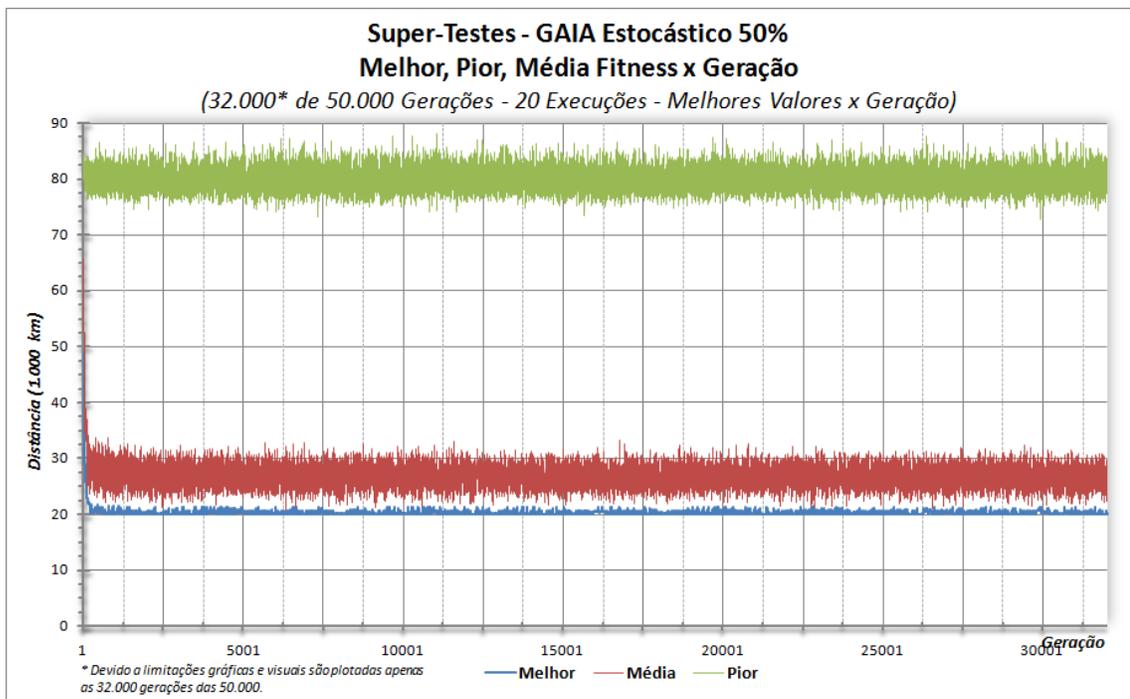


Gráfico 6.47 – Super-Teste GAIA Estocástico 98%: Média dos Valores x Geração

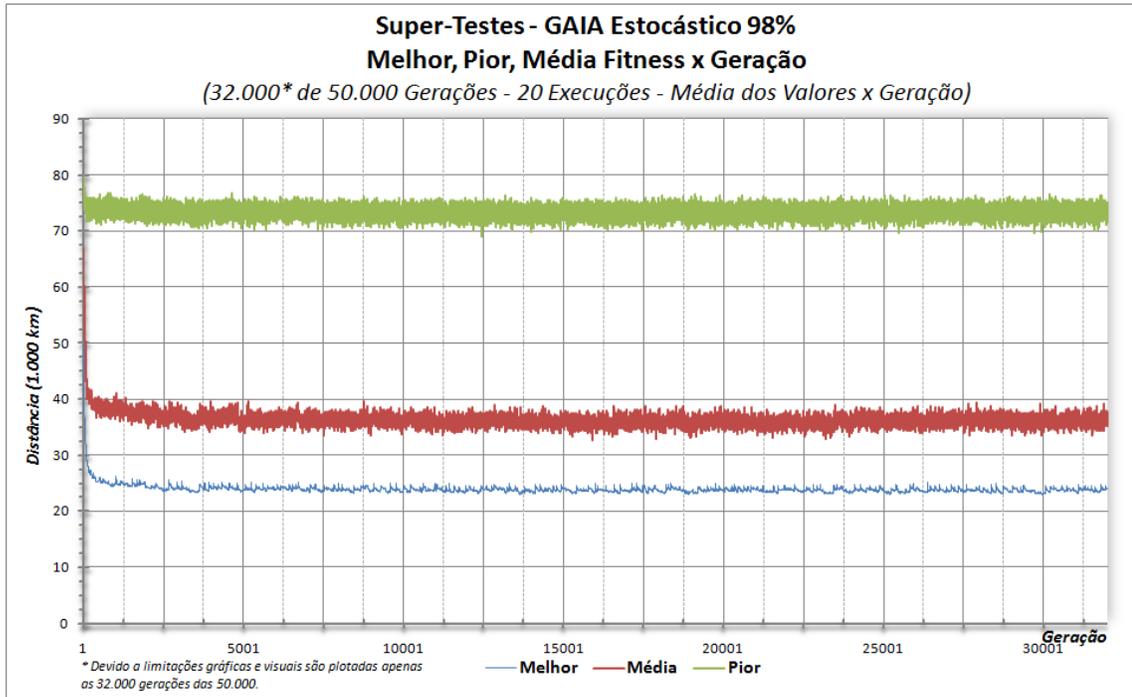
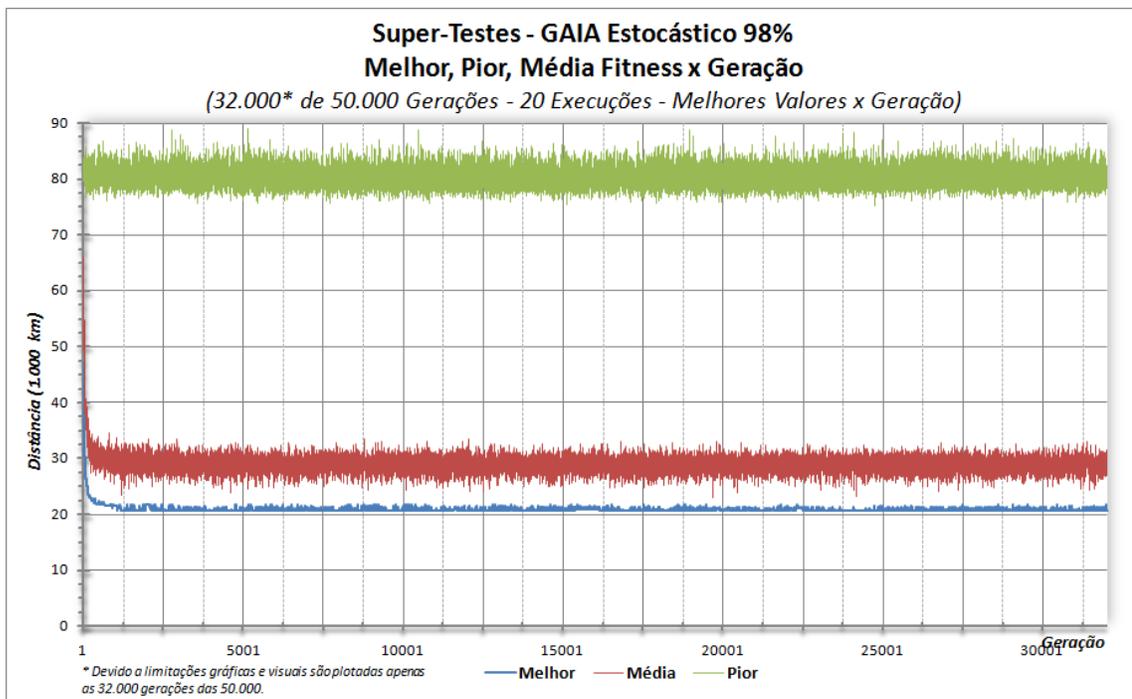


Gráfico 6.48 – Super-Teste GAIA Estocástico 98%: Melhores Valores x Geração



6.3.2.3 GAIA Periódico

No caso dos “Super-Testes” para o GAIA periódico e sua maior quantidade de gerações, optou-se pela definição de três configurações distintas, variando a periodicidade de intervenção em: 500, 1.000 e 10.000 gerações.

Os gráficos 6.49, 6.50, 6.51, 6.52, 6.53 e 6.54, garantem uma ainda maior oportunidade de se observar o que aqui vem sendo chamado de “ciclo melhora-piora”, decorrente das intervenções espaçadas do GAIA sobre a população. Este ciclo chega a ser tão bem definido, que se assemelha a uma onda modulada.

Neste caso, tal como nos “Testes-Padrão”, o ciclo na maioria das vezes, após a recuperação da variabilidade e a nova intervenção para normalização dos parâmetros, tende a ir rumo a melhores valores dentro do espaço de busca.

Avaliando a periodicidade definida, a de 10.000 gerações, assim como a de 2.500 gerações apresentada na seção anterior, são exageros e desperdiçam muito tempo do ciclo de execução disponibilizado pelas gerações. Nota-se que o ideal para esta situação, pelo menos de forma visual, deve ser algo em torno de 500 e 1.000 gerações.

Gráfico 6.49 – Super-Teste GAIA Periódico 500: Média dos Valores x Geração

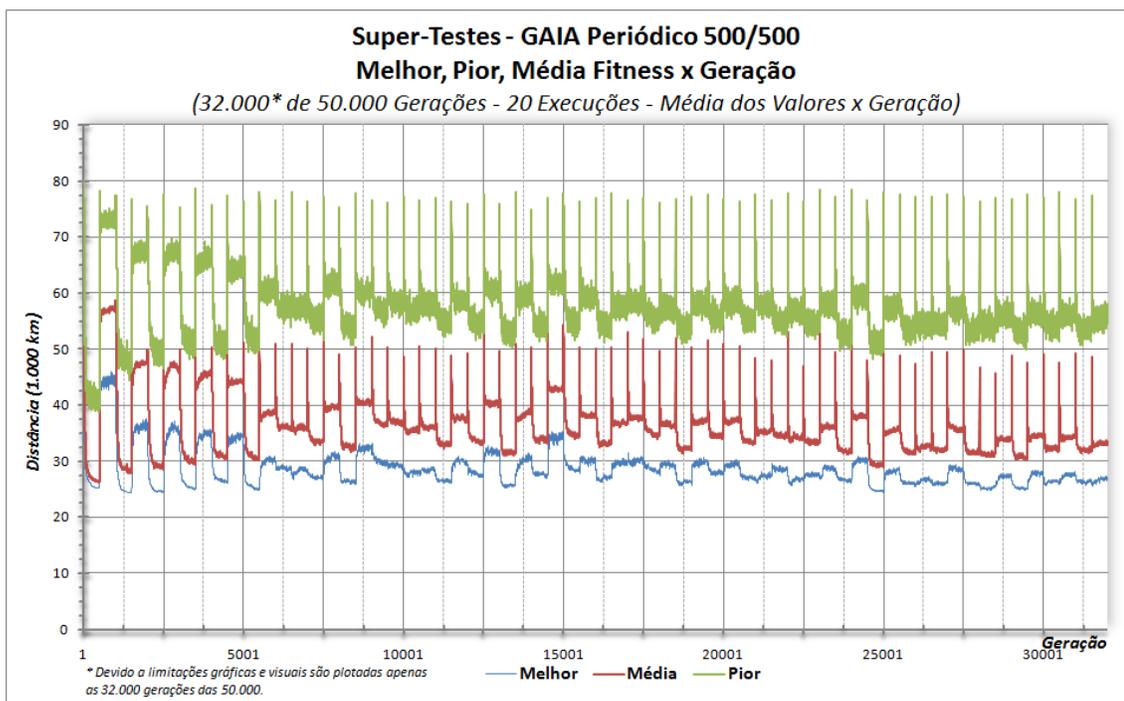


Gráfico 6.50 – Super-Teste GAIA Periódico 500: Melhores Valores x Geração

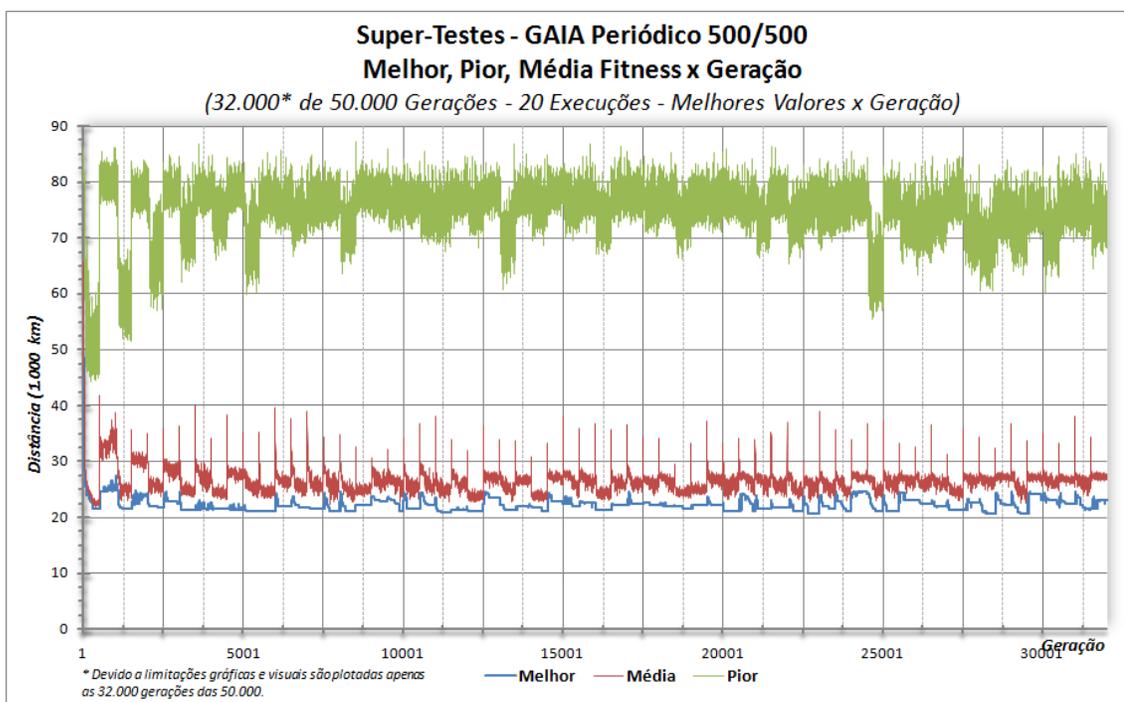


Gráfico 6.51 – Super-Teste GAIA Periódico 1.000: Média dos Valores x Geração

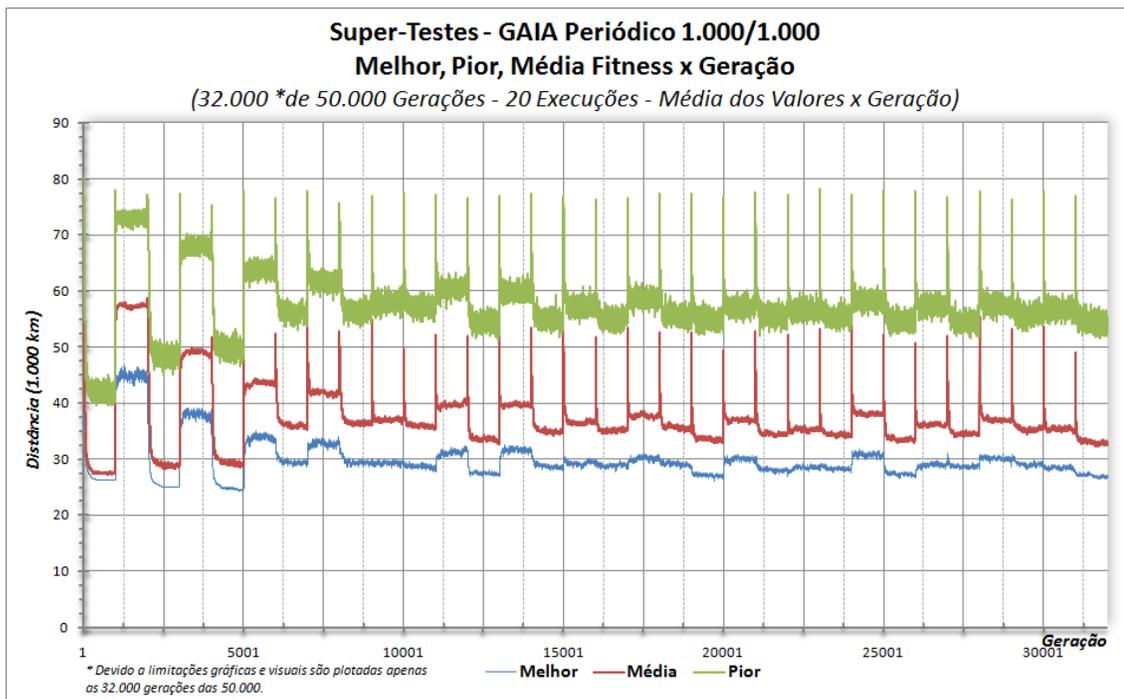


Gráfico 6.52 – Super-Teste GAIA Periódico 1.000: Melhores Valores x Geração

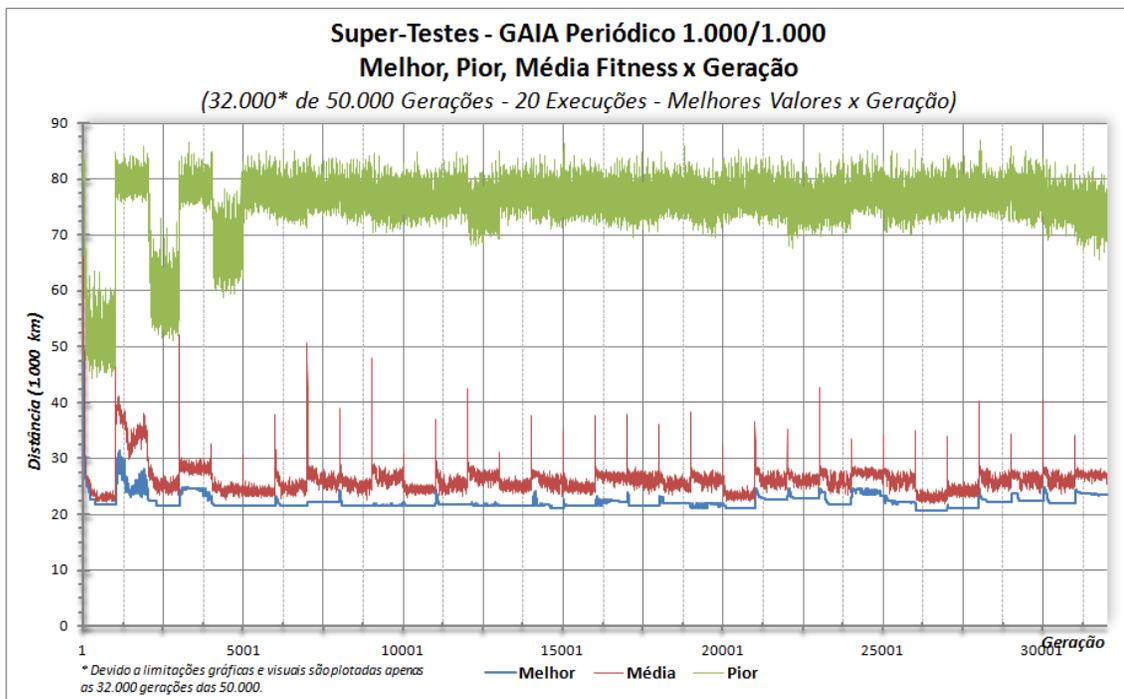


Gráfico 6.53 – Super-Teste GAIA Periódico 10.000: Média dos Valores x Geração

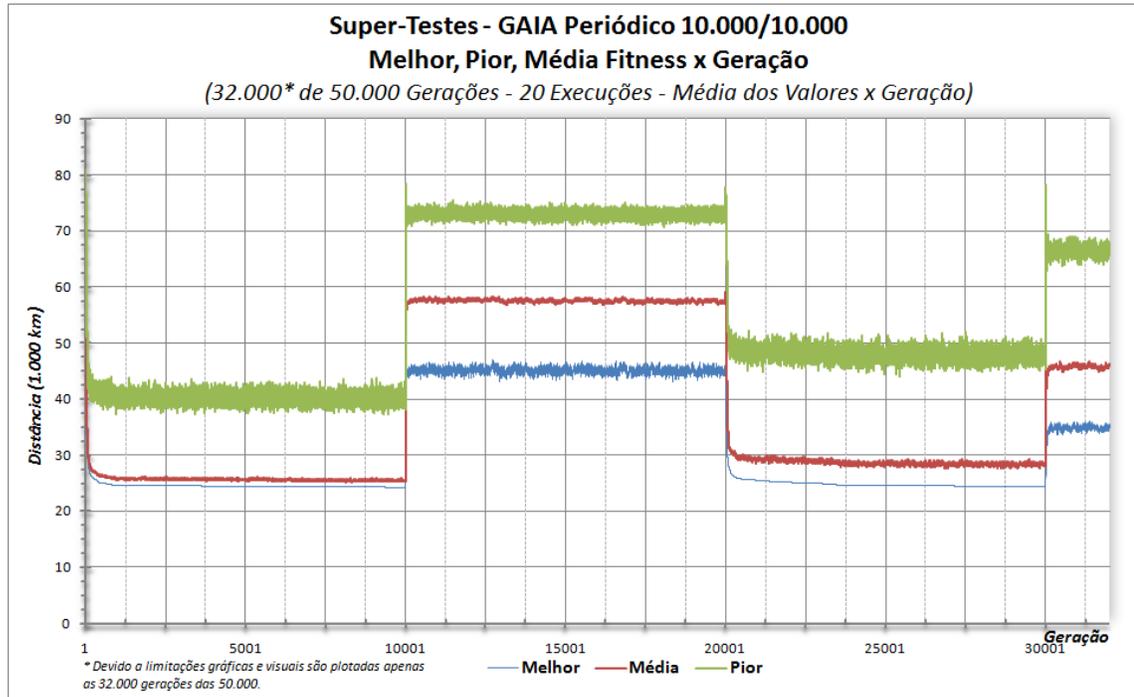
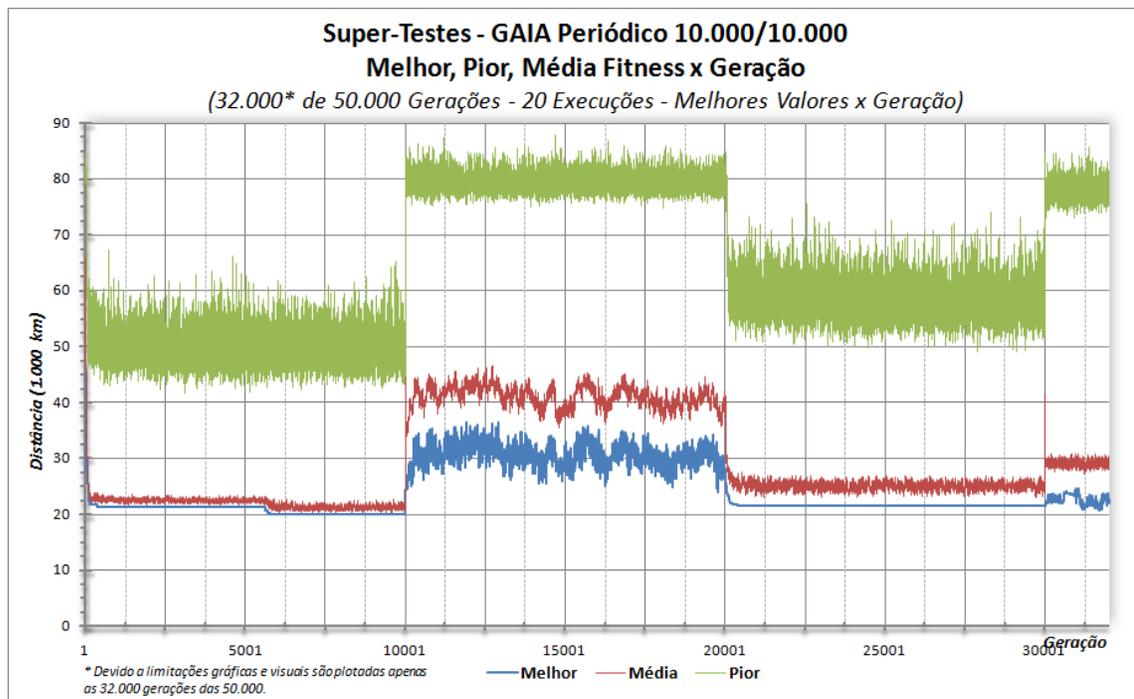


Gráfico 6.54 – Super-Teste GAIA Periódico 10.000: Melhores Valores x Geração



6.3.2.4 Hawk-Dove Roleta e Roleta Simples

Com a mesma finalidade de comparação dos “testes normais” (sem o GAIA) apresentados para o padrão de 5.000 gerações, aqui foram geradas duas configurações diferentes: 01 para um AG Canônico que utiliza o Método da Roleta, e 01 para a variação de AG com Interação Fenotípica.

Os gráficos 6.55, 6.56, 6.57 e 6.58, permitem observar, aquilo que já era esperado e que também fora verificado nas variações com 5.000 gerações: o excesso de convergência, e a total dependência na casualidade do operador de mutação para tentar escapar de um ótimo local e continuar explorando a superfície de busca.

Essa “sorte”, ou dependência do operador de mutação, pôde até ser verificada em dois ou três pontos observados nos gráficos, após longos períodos de estagnação da população. A inexistência de variabilidade é ratificada pela grande proximidade das séries melhor *fitness* e *fitness* médio.

Gráfico 6.55 – Super-Teste Hawk-Dove Roleta: Média dos Valores x Geração



Gráfico 6.56 – Super-Teste Hawk-Dove Roleta: Melhores Valores x Geração

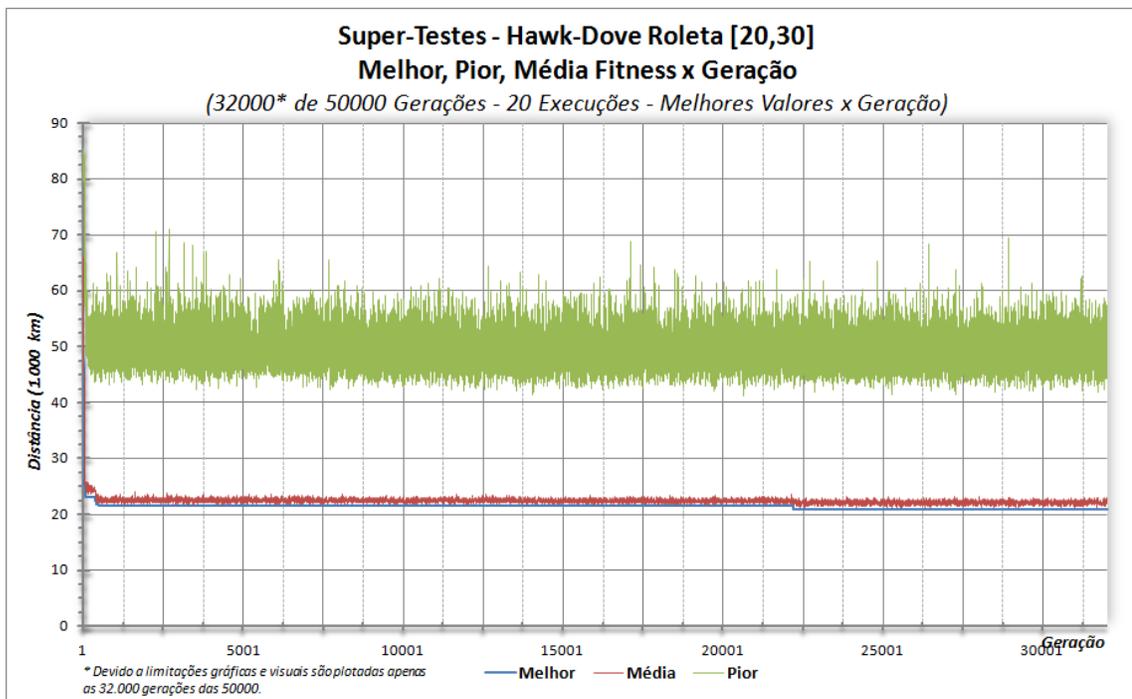


Gráfico 6.57 – Super-Teste Roleta Simples: Média dos Valores x Geração

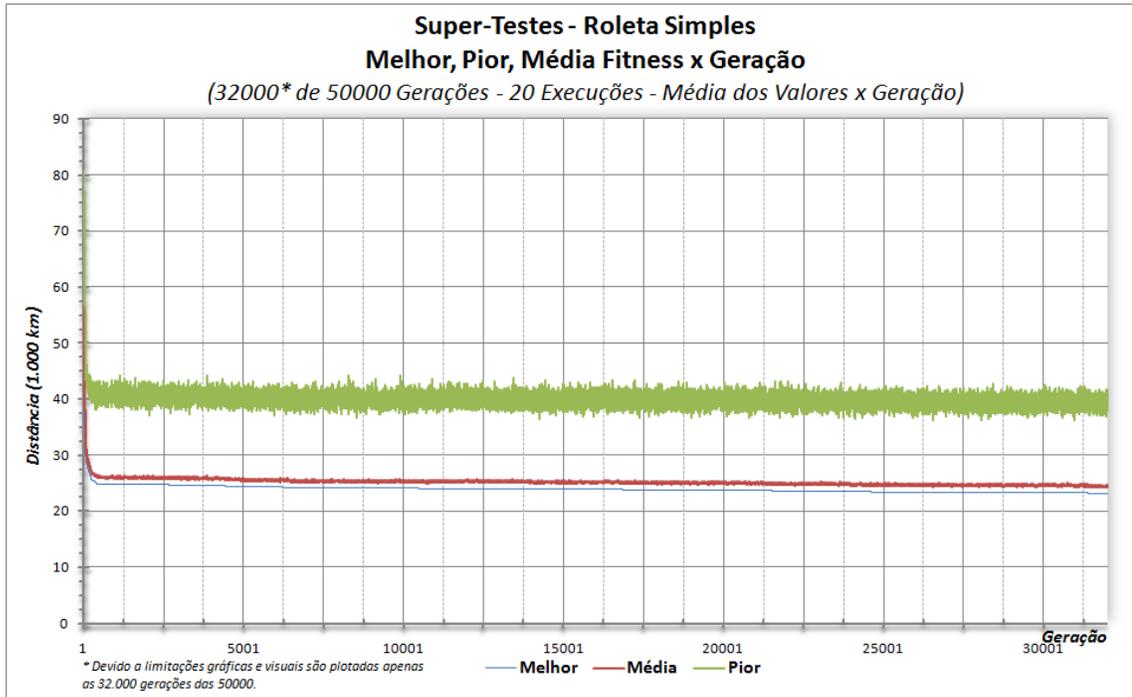
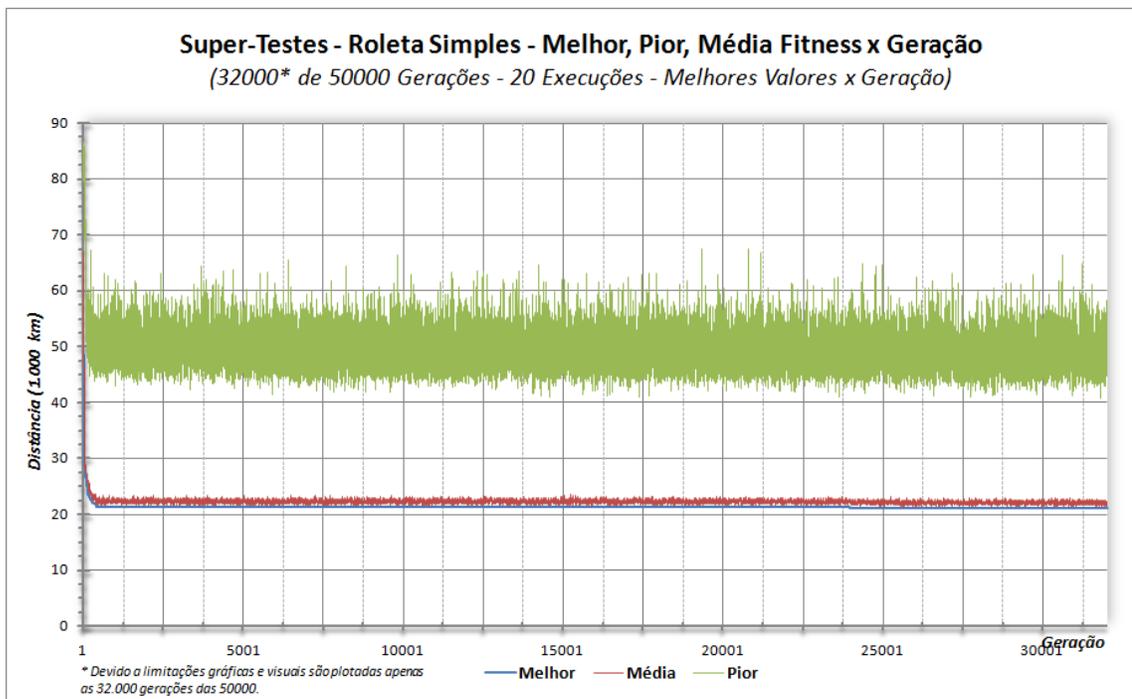


Gráfico 6.58 – Super-Teste Roleta Simples: Melhores Valores x Geração



6.4 QUADRO GERAL E ANÁLISE CONSOLIDADA DOS RESULTADOS

Por meio das tabelas e gráficos apresentados nas seções anteriores, assim como pelas argumentações e análises realizadas, pode-se concluir diversos aspectos sobre a eficácia do funcionamento da proposta do agente. De uma maneira geral, resultados significativamente positivos.

Contudo, para que sejam efetivamente consolidados os objetivos estabelecidos durante a construção do agente, é necessário compor estes resultados e analisá-los sobre a perspectiva comparativa interna e externa, as quais serão abordadas nas seções que seguem.

6.4.2 Medidas e Fatores de Análise e Comparação Interna

Na comparação interna, ou seja, entre as próprias configurações de execuções realizadas, buscou-se a análise sobre dois aspectos: o primeiro, diz respeito ao objetivo principal do pesquisador que aplica à técnica, encontrar a solução ótima para o problema; já o segundo, é a manutenção da variabilidade, fator considerado essencial para a efetiva exploração do espaço de busca.

6.4.2.1 Variável “Melhores Resultados Alcançados”

As tabelas 6.4 e 6.5, apresentam a consolidação dos resultados para as melhores rotas encontradas nos conjuntos de testes, com 5.000 e 50.000 gerações, respectivamente.

A análise desta melhores rotas é feita considerando quando foram encontradas, por quantas gerações e a última vez que este fato ocorreu. Estas medidas complementares, possibilitam a diminuição do fator “casualidade” e uma melhor análise dos resultados.

Desta forma, nos “Testes-Padrão”, considerando o melhor valor encontrado de 20.093 km, em quatro das configurações propostas, segue a análise:

- **GAIA Estocástico 99,8%:** o valor foi encontrado em mais de um dos testes realizados e, considerando que a intervenção do GAIA é quase constante, embora bastante suave, pode-se sim atribuir influência ou contribuição no alcance deste valor;
- **GAIA Estocástico 01%:** com probabilidade de ocorrência complementar ao da anterior, ou seja, com baixa incidência e, considerando que o valor foi encontrado apenas em uma das suas 20 execuções, não se pode atribuir com efetiva certeza se o agente desempenhou papel fundamental neste achado, embora, como citado inúmeras vezes nas demais seções, a variabilidade sempre foi controlada pelo agente;
- **GAIA Estocástico 05%:** ainda mais difícil de afirmar se o agente foi o responsável, ou mesmo o facilitador para o achado. O valor foi encontrado uma única vez e não durou mais que uma geração. Neste caso, pode ser até que a influencia do agente, tenha eliminado este valor da população, antes que pudesse deixar pelo menos partes de sua solução dentro da população;
- **GAIA Estocástico 25%:** com maior incidência do GAIA e com muitas ocorrências deste valor, encontrado em diferentes execuções e momentos destas, nesta situação

pode-se dizer que a influência do agente foi fundamental no achado. Fator já evidenciado nos gráficos apresentados na seção anterior.

Já para os “Super-Testes”, considerando achado em duas de suas configurações do valor tido como provável ótimo global de 20.093 km, segue a análise:

- **GAIA Periódico 10.000:** pela informações oferecidas de que a descoberta deste valor foi feita nas proximidades da geração 5.000 e perdurou até meados de 10.000, só pode ser concluído que o agente não teve nenhum tipo de influência. Isto porque, sua primeira intervenção só aconteceu na geração 10.000 e, quando ela ocorreu, a população estava bastante estagnada presa a este valor. Por conta disso, a intervenção foi tão forte, que o valor acabou subitamente eliminado da execução e não mais foi encontrado;
- **GAIA Estocástico 50%:** diferente da configuração anterior, aqui pode se afirmar que o GAIA foi o motor decisivo no encontro deste valor e, que mesmo após a realização de diversos “ciclos melhora-piora”, o valor foi novamente encontrado e em mais de uma execução. Sendo inclusive, mais um indício do sustentado em TEIXEIRA (2005) de que este é o ótimo global para o problema.

Para os demais valores encontrados, como por exemplo, o 20.113 km, mesmo de BRITO (2004) e BRITO et al. (2005), só ficam mais evidentes o potencial do GAIA de efetiva exploração do espaço de busca. Isto porque, embora estes não sejam o “possível melhor valor”, foram encontrados várias vezes e até mesmo com muitas rotas distintas para o mesmo valor de *fitness* (conforme listado na tabela C.2, Anexo C). Indício de que foram achados bons valores, a população convergiu, o GAIA fez sua intervenção e após sucessivos “ciclos de melhora-piora”, estes valores foram novamente alcançados, inclusive com outras rotas.

Tabela 6.4 – Testes-Padrão: Variável “Melhores Resultados Alcançados”

ID	NOME / DESC.	MELHORES ROTAS			
		MELHOR (KM)	1ª GER. ENCONT.	ULT.GER. ENCONT.	QTD GER.
01	Teste-Padrão GAIA Permanente (TP-GAIA-Permanente)	21.117	714	4.335	652
02	Teste-Padrão GAIA Periódico 20 (TP-GAIAPeriod-20)	20.113	4.590	4.881	88
03	Teste-Padrão GAIA Periódico 50 (TP-GAIAPeriod-50)	20.287	1.527	1.548	68
04	Teste-Padrão GAIA Periódico 100 (TP-GAIAPeriod-100)	20.113	2.383	2.413	31
05	Teste-Padrão GAIA Periódico 500 (TP-GAIAPeriod-500)	21.235	3.059	3.537	248
06	Teste-Padrão GAIA Periódico 1.000 (TP-GAIAPeriod-1000)	20.853	4.173	5.000	828
07	Teste-Padrão GAIA Periódico 2.500 (TP-GAIAPeriod-2500)	22.085	1.665	2.507	789
08	Teste-Padrão GAIA Estocástico 95% (TP-GAIAEstoc-95)	21.069	1.668	2.581	224
09	Teste-Padrão GAIA Estocástico 98% (TP-GAIAEstoc-98)	21.135	858	4.997	1.745
10	Teste-Padrão GAIA Estocástico 99% (TP-GAIAEstoc-99)	20.586	773	4.920	964
11	Teste-Padrão GAIA Estocástico 99,8% (TP-GAIAEstoc-99,8)	20.093	2.911	3.051	50
12	Teste-Padrão GAIA Estocástico 75% (TP-GAIAEstoc-75)	21.158	3.687	4.679	143
13	Teste-Padrão GAIA Estocástico 50% (TP-GAIAEstoc-50)	21.274	277	5.000	1.854
14	Teste-Padrão GAIA Estocástico 0,2% (TP-GAIAEstoc-0.2)	21.178	4.855	5.000	146
15	Teste-Padrão GAIA Estocástico 01% (TP-GAIAEstoc-1)	20.093	1.056	1.395	99

ID	NOME / DESC.	MELHORES ROTAS			
		MELHOR (KM)	1ª GER. ENCONT.	ULT.GER. ENCONT.	QTD GER.
16	Teste-Padrão GAIA Estocástico 05% (TP-GAIAEstoc-5)	20.093	579	-	1
17	Teste-Padrão GAIA Estocástico 25% (TP-GAIAEstoc-25)	20.093	1.306	3.337	605
18	Teste-Padrão Hawk-Dove Roleta (20,30) (TP-HDR2030)	21.089	709	5.000	4.292
19	Teste-Padrão Roleta Simples (TP-RoletaSimples)	21.069	4.079	5.000	922

Tabela 6.5 – Super-Testes: Variável “Melhores Resultados Alcançados”

ID	NOME / DESC.	RESULTADOS			
		MELHOR (KM)	1ª GER. ENCONT.	ULT.GER. ENCONT.	QTD GER.
01	Super-Teste GAIA Permanente (ST-GAIA-Permanente)	20.989	12.947	49.999	13.002
02	Super-Teste GAIA Periódico 500 (ST-GAIA-Period-500)	20.752	22.668	38.004	1.121
03	Super-Teste GAIA Periódico 1.000 (ST-GAIA-Period-1000)	20.586	26.032	27.005	974
04	Super-Teste GAIA Periódico 10.000 (ST-GAIA-Period-10000)	20.093	5.780	10.003	4.224
05	Super-Teste GAIA Estocástico 25% (ST-GAIAEstocástico-25)	20.752	1.523	49.955	23.635
06	Super-Teste GAIA Estocástico 50% (ST-GAIAEstocástico-50)	20.093	1.525	49.970	4.828
07	Super-Teste GAIA Estocástico 98% (ST-GAIAEstocástico-98)	20.586	1.187	49.817	14.944
08	Super-Teste Hawk-Dove Roleta (20,30) (ST-HDR2030)	20.989	22.184	50.000	22.184
09	Super-Teste Roleta Simples (ST-RoletaSimples)	21.069	23.960	50.000	23.960

6.4.2.2 Variáveis de Variabilidade: “Piores”, “Média” e “Dispersão”

Desde a definição dos objetivos deste trabalho até a apresentação dos capítulos de resultados, a garantia da variabilidade é abordada como um dos maiores mecanismos de apoio à evolução. Isto porque, mesmo que considerados ao todo os diferentes indivíduos, e que estes possuam baixa adaptabilidade, eles sempre podem conter os blocos construtores (ou pedaços) fundamentais para o alcance da solução ótima.

Nas tabelas 6.6 e 6.7 a seguir, são apresentados valores de suporte àquilo que já era notado nos gráficos listados nas seções anteriores: o GAIA sempre que intervém, mantém a variabilidade.

Esta afirmação pode ser feita, com base nas seguintes variáveis das tabelas aqui descritas, são elas:

- **PIORES:**

- Pior: Valor de adaptabilidade do pior indivíduo encontrado em toda as execuções para aquela configuração;
- 1º Geração: Indica a primeira geração em que este é encontrado;
- **Análise:** Enquanto a maioria das execuções que utilizam o GAIA (Tabela 6.6, itens de 01 a 17, e Tabela 6.7, itens 01 a 07), conseguem apresentar o seu pior valor nas mais altas gerações da execução, para os métodos tradicionais (Tabela 6.6, itens 18 e 19, e Tabela 6.7, itens 08 e 09), este valor só foi encontrado no máximo até a 19ª geração. **Em síntese, sobre a ótica dos “piores valores” é possível afirmar que o GAIA garante a variabilidade;**

- **MÉDIA:**

- Melhor: Valor mais baixo de média de adaptabilidade encontrada em todas as execuções para aquela configuração;
- 1º Geração: Indica a primeira geração em que este é encontrado;

- Pior: Valor mais alto de média de geração encontrada em todas as execuções de uma determinada configuração;
 - Entre o valor 40.000 km e o pior encontrado: Medida de dispersão sugerida neste trabalho, que permite analisar o percentual de vezes em que a média da geração permaneceu entre 40.000 km e o pior valor encontrado, dado a totalidade de testes e gerações de uma configuração;
 - **Análise:** Enquanto todas as execuções que utilizam o GAIA (Tabela 6.6, itens de 01 a 17, e Tabela 6.7, itens 01 a 07), conseguem manter de 10% a 40% de suas execuções com uma média de adaptabilidade de indivíduos acima de 40.000 km, os métodos tradicionais (Tabela 6.6, itens 18 e 19, e Tabela 6.7, itens 08 e 09), possuem uma convergência tão forte que este valor não chega a 1% de toda a execução. **Em síntese, sobre a ótica da “média de adaptabilidade” durante as execuções é possível afirmar que o GAIA também garante a variabilidade;**
- **DISPERSÃO:**
 - Variância Média: Média da variância calculada em todas as gerações de todas as execuções (testes) para aquela configuração;
 - Desvio Padrão Médio: Média do desvio-padrão calculado em todas as gerações de todas as execuções (testes) para aquela configuração;
 - **Análise:** As execuções que utilizam o GAIA (Tabela 6.6, itens de 01 a 17, e Tabela 6.7, itens 01 a 07), conseguem manter, considerando todos as gerações de todos os testes de uma dada configuração, um desvio-padrão médio entre 5,3 e 13,5 mil km, ao passo que, os métodos tradicionais (Tabela 6.6, itens 18 e 19, e Tabela 6.7, itens 08 e 09), com sua excessiva convergência e desperdício de ciclos de execução, têm esse em torno de 3,1 mil km.

Resumindo-se, sobre a perspectiva das “medidas de dispersão” dos indivíduos durante as execuções, é possível afirmar que o GAIA também garante a variabilidade;

Por outro lado, quando se tenta estabelecer uma comparação entre os diferentes tipos de GAIA frente a variabilidade, os valores acabam sendo muito próximos e, no conjunto de testes realizados não permitiu identificar um padrão absoluto. Entretanto, a observação passível de afirmação é que quanto muito maior ou muito menor a incidência (os extremos), estes valores de variabilidade, tentem a diminuir. Sem no entanto, perderem a característica principal identificada do GAIA, o controle da variabilidade.

Tabela 6.6 – Testes-Padrão: Variabilidade – “Piores”, “Média” e “Dispersão”

ID	PIORES (KM)		MÉDIA (KM)				DISPERSÃO	
	PIOR	1ª GER.	MELHOR	1ª GER.	PIOR	ENTRE 40K E A PIOR	VARIÂNCIA MÉDIA	DESVIO-PADRÃO MÉDIO
01	88.513	553	25.272,77	2.718	67.405,73	30%	185.466.920,78	13.504,85
02	86.439	2.719	21.208,12	4.978	67.670,60	16,36%	80.676.373,91	8.610,30
03	86.038	1.301	21.441,96	1.534	67.177,81	19,34%	68.638.099,36	8.017,80
04	86.404	4.490	22.355,82	3.387	67.381,15	26,05%	62.840.402,06	7.665,28
05	86.715	2.761	22.893,40	353	67.282,18	38,21%	40.774.560,59	6.134,23
06	86.629	1.044	22.454,50	2.886	67.640,75	37,86%	32.215.169,68	5.407,84
07	86.895	3.962	22.443,31	2.233	67.174,76	30,83%	20.413.317,52	4.272,44
08	86.640	4.416	24.235,04	3.603	67.280,48	40,40%	185.728.707,12	13.499,10
09	87.914	4.210	23.903,31	3.259	67.347,64	36,12%	185.402.120,89	13.487,52
10	88.061	2.623	24.368,12	2.676	67.644,63	31,10%	184.889.974,69	13.463,78
11	87.179	4.361	24.799,96	3.002	67.349,33	36,31%	187.276.304,44	13.562,86
12	88.439	2	23.248,14	4.654	67.527,32	29,94%	177.211.418,05	13.144,81
13	88.356	1.580	22.845,42	2.430	67.395,85	23,47%	151.053.856,34	11.952,10
14	86.283	3.343	22.237,88	1.644	67.307,14	33,96%	43.936.059,75	6.340,83

ID	PIORES (KM)		MÉDIA (KM)				DISPERSÃO	
	PIOR	1ª GER.	MELHOR	1ª GER.	PIOR	ENTRE 40K E A PIOR	VARIÂNCIA MÉDIA	DESVIO-PADRÃO MÉDIO
15	85.994	4.381	21.452,83	2.204	67.321,41	21,31%	58.939.528,73	7.438,32
16	88.982	4.972	22.053,65	4.381	67.222,60	16,01%	76.424.807,74	8.332,15
17	87.454	217	21.074,77	3.736	67.464,77	22,09%	122.715.876,56	10.627,03
18	86.976	5	21.348,00	4.960	67.304,97	0,85%	10.487.037,91	3.104,02
19	85.591	7	21.318,46	4.020	68.784,00	0,88%	10.836.922,59	3.161,75

Legenda: **01**-TP-GAIA-Permanente; **02**-TP-GAIAPeriod-20; **03**-TP-GAIAPeriod-50; **04**-TP-GAIAPeriod-100; **05**-TP-GAIAPeriod-500; **06**-TP-GAIAPeriod-1000; **07**-TP-GAIAPeriod-2500; **08**-TP-GAIAEstoc-95; **09**-TP-GAIAEstoc-98; **10**-TP-GAIAEstoc-99; **11**-TP-GAIAEstoc-99,8; **12**-TP-GAIAEstoc-75; **13**-TP-GAIAEstoc-50; **14**-TP-GAIAEstoc-0.2; **15**-TP-GAIAEstoc-1; **16**-TP-GAIAEstoc-5; **17**-TP-GAIAEstoc-25; **18**-TP-HDR2030; **19**-TP-RoletaSimples.

Tabela 6.7 – Testes-Padrão: Variabilidade – “Piores”, “Média” e “Dispersão”

ID	PIORES (KM)		MÉDIA (KM)				DISPERSÃO	
	PIOR	1ª GER.	MELHOR	1ª GER.	PIOR	ENTRE 40K E A PIOR	VARIÂNCIA MÉDIA	DESVIO-PADRÃO MÉDIO
01	89.413	24.163	24.247,60	34.957	67.455,80	15,29%	185.140.027,57	13.454,54
02	86.958	28.048	21.993,18	368	67.322,24	22,81%	43.973.967,48	6.467,41
03	87.196	8.503	21.852,56	26.481	67.861,55	28,80%	41.304.973,92	6.274,16
04	87.906	14.686	20.333,82	9.415	67.679,01	34,68%	30.899.022,31	5.327,88
05	89.257	22.725	21.711,61	29.144	67.891,06	10,49%	112.082.980,77	10.097,73
06	88.306	44.994	20.857,89	6.337	67.897,47	14,67%	148.868.854,96	11.808,27
07	89.140	5.134	22.994,14	19.682	74.293,75	18,00%	180.226.737,57	13.130,97
08	85.610	1	21.165,40	49.802	67.771,00	0,08%	10.325.026,98	3.118,17
09	86.586	19	21.227,30	24.888	68.774,00	0,09%	10.589.320,88	3.159,47

Legenda: **01**-ST-GAIA-Permanente; **02**-ST-GAIA-Period-500; **03**-ST-GAIA-Period-1000; **04**-ST-GAIA-Period-10000; **05**-ST-GAIAEstocástico-25; **06**-ST-GAIAEstocástico-50; **07**-ST-GAIAEstocástico-98; **08**-ST-HDR2030; **09**-ST-RoletaSimples.

6.4.2 Comparação com Outros Trabalhos

Após as comparações internas ao trabalho apresentadas, são feitas também comparações com algumas propostas híbridas de Algoritmos Genéticos que também tentam, como forma de provar sua eficácia, resolver o mesmo problema aqui utilizado, o PCVS-BR-26.

Estes trabalhos e seus resultados já foram por vez ou outra aqui referenciados, porém, cabe à essa seção apresentar alguns gráficos de seus resultados para que se possa fazer um paralelo com os obtidos com o uso do GAIA nas suas variadas formas.

Os trabalhos escolhidos foram o de LEHRER (2000), em que uma nova proposta de método de seleção foi apresentada envolvendo o uso da Teoria dos Jogos Evolucionários (vide capítulo IV e Anexo E para maiores detalhes), e suas variações apresentadas nos trabalhos de BRITO (2004), BRITO et al. (2005), e TEIXEIRA (2005).

LEHRER (2000), na configuração de execução que considerou o melhor resultado de seu trabalho, AG com o método Hawk-Dove Roleta, apresentada no gráfico 6.59, encontrou o valor de adaptabilidade de 20.409 km, e, inclusive chegou à conclamá-lo o ótimo global para o problema.

Já em BRITO (2004) e em sua publicação BRITO et al. (2005), utilizando variação que considerou aperfeiçoada do método Hawk-Dove Roleta de LEHRER (2000), chegou a obter uma série de melhores resultados, encontrando inclusive um novo valor ótimo para o problema, de 20.113 km, como apresentado no gráfico 6.60.

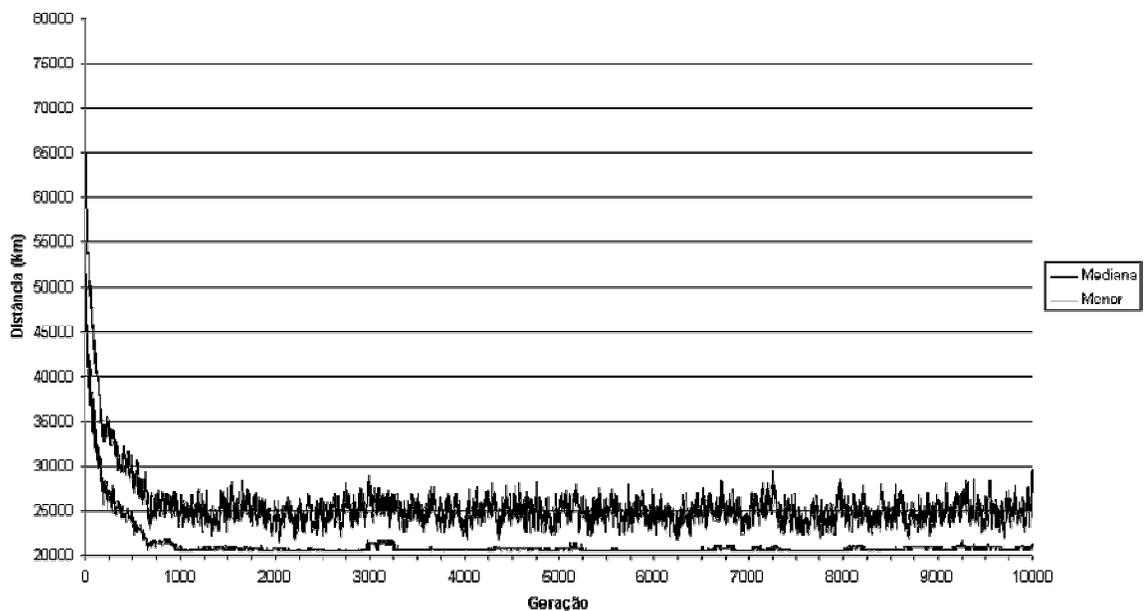
Por fim, no trabalho de TEIXEIRA (2005), em que efetivamente a Interação Social é tratada como uma parte independente do Algoritmo Genético - tal qual utilizado em todos os aqui apresentados -, chegou a um novo valor ótimo para o problema de 20.093 km, exibido no gráfico 6.61.

Independentemente destes valores encontrados, já que aqui mesmo neste trabalho, o possível “melhor” valor apresentado em TEIXEIRA (2005) fora encontrado inúmeras vezes pelas execuções do GAIA, o que se pode notar de fato, é que em todos essas propostas externas a excessiva convergência impera. Não há variabilidade alguma após menos de 1/10 do processamento da execução.

Sendo assim, e, tal como foi descrito em seção anterior, qualquer resultado encontrado por meio destas técnicas, apesar de que estas se mostraram eficientes num certo controle (ou retardo) da pressão seletiva, reside muito em fatores relacionados à casualidade, em especial ao operador de mutação.

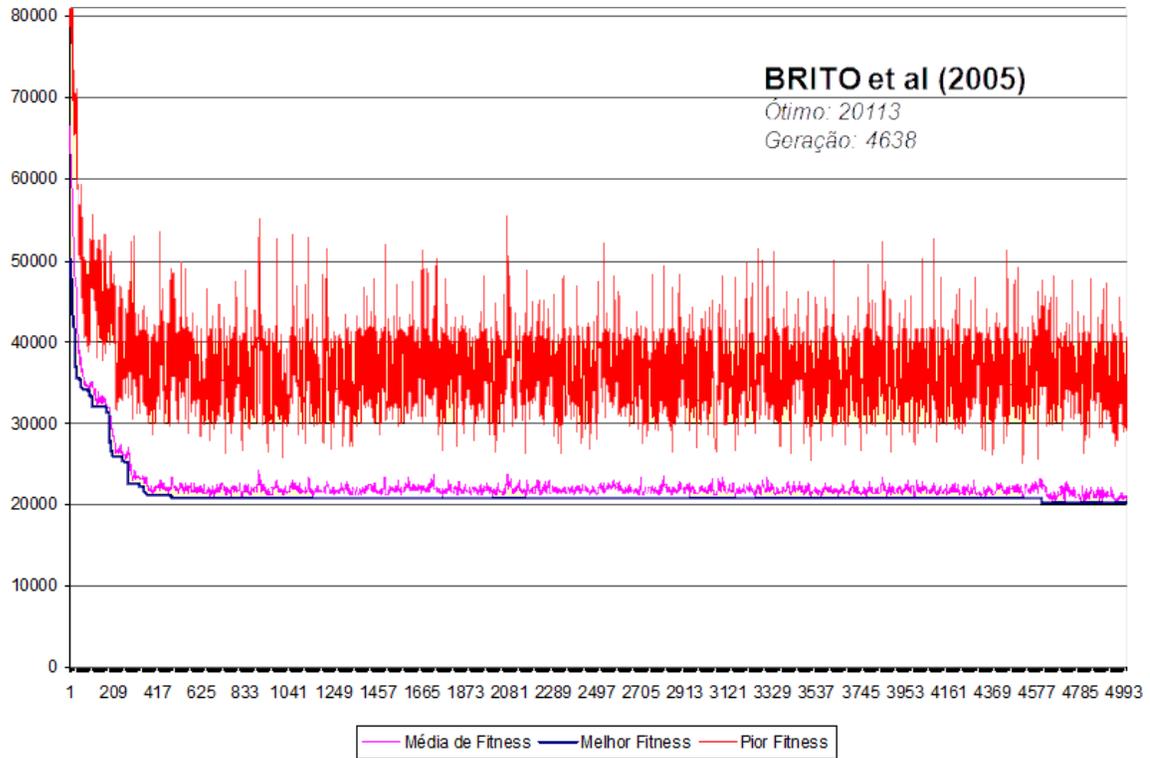
Se for ainda considerada a argumentação de DEMBSKI (2004, 2007), em que sem qualquer conhecimento ou orientação dentro do espaço de busca, um algoritmo evolucionário não passa de uma busca cega aprimorada, mesmo as técnicas híbridas destes outros trabalhos podem ser enquadradas nesta situação. Diferente do GAIA que, com sua base de regras e eventuais intervenções, supre esta lacuna.

Gráfico 6.59 – Hawk-Dove Roleta (LEHRER, 2000): Melhor Valor e Média x Geração



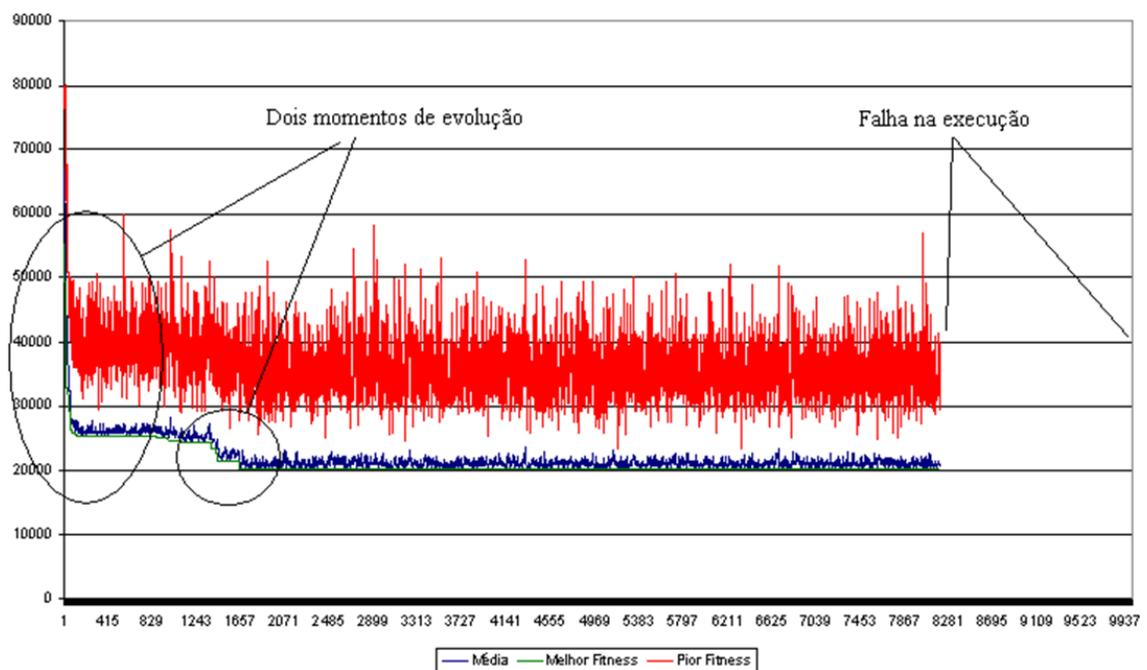
Fonte: Adaptado de LEHRER (2000).

Gráfico 6.60 – Hawk-Dove Roleta (BRITO et. al., 2005): Média, Melhor e Pior x Geração



Fonte: Adaptado de BRITO et al. (2005).

Gráfico 6.61– Hawk-Dove Roleta (TEIXEIRA, 2005): Média, Melhor e Pior x Geração



Fonte: Adaptado de TEIXEIRA (2005).

6.5 CONSIDERAÇÕES FINAIS

Encerrada a apresentação da proposta do GAIA, seus aspectos teóricos e práticos, assim como o extenso conjunto de testes e análises realizados neste capítulo para comprovação da eficácia da mesma, parte-se para as conclusões positivas e negativas, sugestões de trabalhos futuros e outras ponderações sobre os resultados da proposta, apresentados no capítulo final a seguir.

CAPÍTULO VII

Não se pode ensinar alguma coisa a alguém, pode-se apenas auxiliar a descobrir por si mesmo.

Galileu Galilei

7. CONCLUSÕES

“*Simple is better*”²¹, simples é melhor! Desde o surgimento dos Algoritmos Genéticos, tudo que mais se tem feito é incrementá-lo com novas variáveis biológicas, técnicas, proposições matemáticas, comportamentais, entre muitas outras.

O que efetivamente vem dando certo nesta linha, que poderá facilitar a vida dos pesquisadores de outras áreas que desejam usar estas técnicas para seus problemas, ou mesmo que apresentam eficiência de “ponta-a-ponta”, como solução de qualquer tipo de problema, é algo que poucos conseguirão responder.

Este trabalho, na tentativa de tentar resolver algumas destas questões, em especial a da alta complexidade envolvida no uso de Algoritmos Genéticos e de sua baixa eficiência frente à excessiva convergência e a tendência de retenção em ótimos locais, propôs o uso de uma abordagem, até certo ponto, inusitada na área, por meio do apoio teórico na controversa Teoria do Design Inteligente.

²¹ “*Simple is better*”, palestra proferida pelo Professor **Xin Yao** (Universidade Birmingham, Reino Unido) durante a ‘II Conferência Internacional de Computação Natural’ (ICNC’06 - Internacional Conference on Natural Computation / IEEE Transactions on Evolutionary Computation), em 24/09/2006, na cidade de Xi’an, China.[*Citação informal proveniente de anotações e registros do autor deste trabalho, presente na platéia durante o ocorrido*].

Desta forma, para alcançar o nível de conhecimento necessário à fundamentação da proposta, partiu-se de um vasto estudo dos fundamentos teóricos da origem do universo, do surgimento da vida, e da compreensão do complexo e, ainda misterioso, processo de evolução. Considerando seus aspectos químicos e biológicos, abordou aspectos de genética, associando-os aos conceitos do darwinismo, conhecido como neodarwinismo.

Na seqüência, foi realizado um profundo estudo em processos miméticos da natureza para a resolução de problemas. Neste contexto, abordou-se os Algoritmos Genéticos e seus aspectos teóricos e funcionais como técnica computacional de larga adoção na resolução de problemas de otimização. Ao encerrar este estudo, buscou-se apresentar os principais problemas relacionados aos AGs, em especial os que dizem respeito à “homérica” tarefa de definição paramétrica e a convergência prematura, assim como, alguns esforços e pesquisas na tentativa de solucioná-los.

Compreendida e apresentada a problemática central deste trabalho, refletida na considerável ineficiência dos AGs e suas variações, bem como a dificuldade do seu uso. Partiu-se para a compreensão de abordagens teóricas alternativas que viriam a fundamentar a nova proposta.

Nesta direção, apresentou-se a Teoria do Design Inteligente sobre a ótica de seus dois principais argumentos: a complexidade especificada e a complexidade irreduzível. Além disso, fez-se uso e interpretação dos estudos conduzidos por DEMBKI (2004,2007), que comparavam os algoritmos genéticos sem qualquer conhecimento adicional do espaço de busca, à meras buscas cegas. Essas deduções lógicas foram feitas por meio do uso *No Free Lunch Theorem*.

Ainda como forma de subsidiar a proposta principal, foram apresentadas algumas teorias e modelos que buscam explicar os periódicos eventos de extinção em massa de espécies no planeta, e como após a sua ocorrência, surgem sempre novas espécies melhores e

mais evoluídas. Posteriormente, esta teoria foi relacionada ao Design Inteligente, até mesmo tomando como base suas próprias características, propondo que o “*designer*”, de tempos em tempos, como forma de atingir seus objetivos intervém sobre as chamadas causas naturais, eliminando indivíduos e reintroduzindo a evolução por meio do retorno da variabilidade.

Visualizada a extensão da proposta e sua forma de funcionamento, buscou-se apresentar as ferramentas sobre as quais seriam construídas. Neste contexto, tratou-se de apresentar um estudo sobre lógica fuzzy e também da estrutura híbrida de algoritmos genéticos proposta por BRITO et al. (2005).

Com base em todos estes conceitos e buscando equacionar a problemática apresentada dos AGs, foi apresentada a construção de um agente inteligente, chamado de G.A.I.A, que de tempos em tempos, tal qual no Design Inteligente, realizaria intervenções manipulando a população e os parâmetros da execução. Desta forma conseguiu-se tratar os dois principais problemas dos AGs discutidos: eficácia e alta complexidade na definição do *set* de parâmetros.

A base de regras, ou o conhecimento fuzzy, que foi adicionado ao agente, reflete os estudos dos principais autores da área, assim como o conhecimento empírico de anos de pesquisa e experiência do autor e seu grupo científico sobre o tema.

Dando prosseguimento ao trabalho, foi apresentado um extenso conjunto de mais de 560 testes, com a finalidade de validar a proposta. Todos os testes foram feitos utilizando o mesmo problema, o PCVS-BR-26, para fins de comparação interna e externa da eficácia da proposta.

Além disso, foram definidos três tipos de abordagens diferentes para o agente, considerando o processo de intervenção: permanente, periódica e estocástica. O que se tornou também objeto de estudo dos testes que, além de atestarem a eficácia da proposta, deveriam apontar para uma ou outra abordagem melhor de uso do GAIA.

Tabulados os resultados, que resultaram em aproximadamente 32 GB de informações, pôde-se perceber o quão mais eficaz era a proposta frente às técnicas tradicionais de AGs, e mesmo a outras, como as híbridas propostas nos trabalhos de LEHRER (2000), TEIXEIRA (2005) e BRITO (2004).

Nos testes aqui realizados, além de se ter alcançado o considerado valor ótimo global para o problema em inúmeras execuções, pôde-se também, por meio de vários mecanismos de análises de dispersão, comprovar o maior aproveitamento do ciclo de processamento das execuções, graças a constante manutenção da variabilidade, garantida pelas intervenções do agente.

Com relação a comparação entre as próprias diferentes abordagens do GAIA, pode-se notar, ao menos dentro do conjunto de testes realizados, que o ideal é deixar a evolução acontecer com intervenções esporádicas e, até certo ponto espaçadas. Em síntese, evitar as configurações extremistas, ou seja, com poucas intervenções ou com controle constante.

Esta conclusão é feita, quando da análise de como a intervenção ocorre nestas duas situações extremas. Na primeira, com poucas intervenções, quanto maior for o tempo de intervenção, mais forte será a intervenção do GAIA e, por um longo período de tempo, até uma nova intervenção, os parâmetros serão deixados em estado de “controle” de variabilidade, em que praticamente não permitem qualquer evolução. Na segunda situação, intervenção constante, as regras de intervenção utilizadas chegam a ser muito brandas e o agente não dá muito espaço para que a “junção” dos “blocos construtores” ocorra na busca da formação da melhor delas.

Com base nestas conclusões e até mesmo no número significativo de publicações que esta proposta rendeu, pode-se considerá-la muito promissora, não só pela sua eficácia aqui comprovada, mas também pela ousadia de buscar inspiração em teorias tão amplamente controversas e muitas vezes refutadas no meio científico, como o Design Inteligente.

Entende-se que uma grande quantidade de estudos nas questões aqui levantadas ainda precisam ser realizados e/ou aprofundados. Citando apenas alguns como exemplo, tem-se o acréscimo de novos aspectos do Design Inteligente, a codificação de novas variáveis para controle e monitoramento, uma forma de auto-adaptar a base de regras e o domínio de suas variáveis, o uso ou teste de outros motores de inferência fuzzy, a automação do melhor momento de intervir, entre muitas outras questões, que são aqui deixadas como propostas de trabalhos futuros.

Um destes trabalhos, inclusive considerado como o próximo passo na evolução do GAIA, já teve seu projeto inicial discutido. Inicialmente chamado de IDEA (*Intelligent Design Evolutionary Algorithm*), deve agregar um grande contingente de técnicas de computação evolucionária geridas automaticamente por meio de um *designer* (agente inteligente), que combinando-as da melhor forma, buscará resolver qualquer problema de otimização com total independência de intervenção do pesquisador, salvo, é claro, no momento em que este deve inserir no sistema o problema ser solucionado.

REFERÊNCIAS BIBLIOGRÁFICAS

AMABIS, José Mariano; MARTHO, Gilberto Rodrigues. **Biologia das células**. 1. vol. São Paulo: Moderna, 1994.

AXELROD, Robert. **The evolution of cooperation**. Basic Book, 1980.

BORGES, P. S. S. **A model of strategy games based on the paradigm of the iterated prisoner's dilemma employing fuzzy sets**. 1996. Tese (Doutorado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 1996.

BRITO, F. H. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. **A introdução da interação fenotípica em algoritmos genéticos através dos jogos evolucionários e da codificação e transmissão genética do comportamento** (Presented Conference Paper Style), presented at VII SBAI/II IEEE LARS: Simpósio Brasileiro de Automação Inteligente / II Latin-American IEEE Robotics Symposium, São Luís, MA, September 18-23 2005.

BRITO, Felipe Houat de ; TEIXEIRA, Artur Noura ; TEIXEIRA, Otávio Noura ; OLIVEIRA, Roberto Célio Limão de . **Hawk-Dove Tournament: A New Selection Technique for Genetic Algorithms Based on Evolutionary Game Theory**. In: Worldcomp 2009 - World Congress in Computer Science, Computer Engineering and Applied Computing, 2009, Las Vegas, NV, USA. GEM 2009 - Genetic and Evolutionary Methods, 2009.

BRITO, Felipe Houat de. **Hawk-Dove torneio: um novo método de seleção para os algoritmos genéticos baseado na teoria dos jogos evolucionários e estratégias evolucionárias**. 143f. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Centro Universitário do Estado do Pará, Belém, 2004.

CROSSLEY, William. **Genetic algorithms "naturally select" better satellite orbits**, (2001). Disponível em <http://www.sciencedaily.com/releases/2001/10/011016070234.htm>. Acesso em 31 ago. 2004.

DÄNIKEN, Erich Von. **Eram os deuses astronautas?** 59. ed. São Paulo: Melhoramentos, 2005.

DARWIN, Charles. **The origin of species.** Disponível em <http://www.literature.org/authors/darwin-charles/the-origin-of-species/>. Acesso em 10. set. 2004.

DE JONG, Kenneth Alan. **An analysis of the behavior of a class of genetic adaptive systems.** PhD Thesis. University of Michigan, Ann Arbor, 1975.

DEMBKSI, W. A. **No free lunch: why specified complexity cannot be purchased without intelligence.** EUA, Maryland, Lanham: Rowman & Littlefield Publishers, Inc., 2007.

DEMBKSI, W. A. **The design revolution.** Madison, Wisconsin, EUA: InverVarsity Press, 2004.

DNER. **Distância entre cidades.** Disponível em <http://www.dnit.gov.br/rodovias/distancias/distancias.asp>. Acesso em 01. nov. 2004.

EBERHART, R.; SIMPSON, P.; DOBBINS, R. **Computational intelligence PC tools: an indispensable resource for the latest in fuzzy logic, neural network and evolutionary computing.** American Press Inc., 1996.

FAVARETTO, José Arnaldo; MERCADANTE, Clarinda. **Biologia: volume único.** São Paulo: Moderna, 1999.

FOGEL, D. B. **Evolutionary computation: toward a new philosophy of machine intelligence.** 2.ed. IEEE Press, 2000.

Fuzzy logic toolbox for use with matlab: user's guide (1999). Natick, Massachusetts: The MathWorks, Inc.

FuzzyJ Toolkit and FuzzyJess. Disponível em http://www.iit.nrc.ca/IR_public/fuzzy/fuzzyJToolkit2.html. Acesso em 15. jan. 2006.

GENOME. **Project GENOME research.** Disponível em <http://www.gnome.org>. Acesso em 05. nov. 2004.

GOTT, J. Richard. **Viagens no tempo no universo de Einstein**. Rio de Janeiro: Ediouro, 2002.

GOWDAK, Demetrio; MATTOS, Neide S. **Biologia: volume único**. São Paulo: FTD, 1991.

HAUPT, R. L.; HAUPT, S. E. **Practical genetic algorithms**. 2. ed. John Wiley & Sons, Inc., 2004.

HEUSER, Carlos Alberto. **Projeto de banco de dados**. 4. ed. Porto Alegre: Sagra Luzzatto, 2001.

KOZA, John R.; KEANE, Martin A.; STREETER, Mathew J. Aperfeiçoando inventos. **Scientific American Brasil**, São Paulo, p.60-67, mar. 2003.

KURZWEIL, R. **The age of spiritual machines**. Viking Penguin, 1999.

LEHRER, Cristiano. **Operador de seleção para algoritmos genéticos baseado no jogo hawk-dove**. 2000.123f. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de Santa Catarina, Florianópolis, 2000.

LOPES, Sônia G. B. C. **Bio: genética, evolução e ecologia**. 3. vol. 2. ed. São Paulo: Saraiva, 1998.

LUCE, Duncan R.; RAIFFA, Howard. **Games and decision: introduction and critical survey**. New York: Dover, 1957.

MITCHELL, M. **An introduction to genetic algorithms**. MIT Press, 1999.

MOREIRA, Fabiano C. **Curso de inteligência artificial I: algoritmos genéticos**. Belém: Centro Universitário do Estado do Pará, 2004. 67 slides: color.

MOTTA, Paulo A. **Genética humana aplicada à psicologia e toda a área biomédica**. Rio de Janeiro: Guanabara Koogan, 2000.

MOURÃO, Ronaldo Rogério de Freitas. **O livro de ouro do universo**. Rio de Janeiro: Ediouro, 2002.

NATGEO. **Mass extinction, mass die-off information, prehistoric facts**. Disponível em <http://science.nationalgeographic.com/science/prehistoric-world/mass-extinction/>. Acesso em: 05 dez. 2010.

RAUP, David M; SEPKOSKI, J John Jr. **Periodicity of extinctions in the geologic past**. Proceedings of the National Academy of Science 81, p. 801-805. 1984.

SMITH, John Maynard. **Evolution and the theory of game**. Cambridge University, 1982.

TEIXEIRA, Otávio Noura. **Uma proposta de um novo algoritmo genético baseado na teoria dos jogos**. 290f. Dissertação (Mestrado em Engenharia Elétrica – Computação Aplicada) – Departamento de Engenharia Elétrica e Computação, Universidade Federal do Pará, Belém, 2005.

UNIVERSIDADE FEDERAL DE CAMPINA GRANDE. Departamento de Engenharia Civil. **Mitologia Grega: Gaia mãe Terra Geia Ge**. Disponível em <http://www.dec.ufcg.edu.br/biografias/MGGaia00.html>. Acesso em 20 jan. 2011.

VENKATASUBRAMANIAN, Venkat. **Computers use darwinian model to "evolve" fuel additives**, (1999). Disponível em <http://www.sciencedaily.com/releases/1999/07/990722064348.htm>. Acesso em 20. ago. 2004.

WANG, L-X. **A course in fuzzy system and control**. Prentice-Hall International, 1997.

NORMAS ADOTADAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR6023**: referências bibliográficas. Rio de Janeiro, 2002. 24p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR6024**: numeração progressiva das seções de um documento. Rio de Janeiro, 2003. 3p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR6027**: sumário. Rio de Janeiro, 2003. 2p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR6028**: resumo. Rio de Janeiro, 1990. 4p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR 10520**: apresentação de citações em documentos. Rio de Janeiro, 2002. 7p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – ABNT. **NBR 14724**: trabalhos acadêmicos - apresentação. Rio de Janeiro, 2002. 9p.

CAMBRIDGE UNIVESRITY PRESS. **Cambridge dictionary online: free english dictionary and thesaurus**. Disponível em: < <http://dictionary.cambridge.org/>>. Acesso em: dez. 2009.

FERREIRA, A. B. H. **Dicionário aurélio básico da língua portuguesa**. 1. ed. Rio de Janeiro: Nova Fronteira, 1988. 687p.

FERREIRA, A. B. H. **Novo dicionário aurélio: versão 6.0 - conforme nova ortografia**. 4. ed. Rio de Janeiro: Positivo Informática, 2009. CD-ROM.

HAWKINS, J. M. **The oxford minidictionary**. 2. ed. Nova York: Clarendon, 1998. 629p.

LIRA, Maria Cristina M. D.; MELO, Hilma Celeste A. **Normas para a apresentação de trabalhos técnicos e científicos: trabalhos acadêmicos, monografias, dissertações e teses**. 2. ed. Belém: CESUPA, 2004. 60p.

ANEXO A

Tabela A.1 – Distância entre as capitais brasileiras utilizando rodovias, estradas e vias municipais, estaduais ou federais.

QUADRO DE DISTÂNCIAS RODOVIÁRIAS ENTRE AS PRINCIPAIS CIDADES BRASILEIRAS – Parte 01 de 03

	ARACAJÚ	BELÉM	BELO HORIZONTE	BOA VISTA	BRASÍLIA	CAMPO GRANDE	CUIABÁ	CURITIBA	FLORIANÓPOLIS	FORTALEZA	GOIÂNIA	JOÃO PESSOA
ARACAJÚ		2079	1578	6000	1652	2765	2775	2595	2892	1183	1848	611
BELÉM	2079		2824	6083	2120	2942	2941	3193	3500	1610	2017	2161
BELO HORIZONTE	1578	2824		4736	716	1453	1594	1004	1301	2528	906	2171
BOA VISTA	6000	6083	4736		4275	3836	3142	4821	5128	6548	4076	6593
BRASÍLIA	1650	2140	741	4275		1134	1133	1366	1673	2200	209	2245
CAMPO GRANDE	2764	2942	1453	3836	1134		694	991	1298	3407	935	3357
CUIABÁ	2773	2941	1594	3142	1133	694		1679	1986	3406	934	3366
CURITIBA	2595	3193	1004	4821	1366	991	1679		300	3541	1186	3188
FLORIANÓPOLIS	2892	3500	1301	5128	1673	1298	1986	300		3838	1493	3485
FORTALEZA	1183	1611	2528	6548	2208	3407	3406	3541	3838		2482	688
GOIÂNIA	1849	2017	906	4076	209	935	934	1186	1493	2482		2442
JOÃO PESSOA	611	2161	2171	6539	2245	3357	3366	3188	3485	688	2442	
MACEIÓ	294	2173	1854	6276	1928	3040	3049	2871	3168	1075	2105	395
MANAUS	5215	5298	3951	785	3490	3051	2357	4036	4343	5763	3291	5808
NATAL	788	2108	2348	6770	2422	3537	3543	3365	3662	537	2619	185
PORTO ALEGRE	3296	3854	1712	5348	2027	1518	2206	711	476	4242	1847	3889
PORTO VELHO	4229	4397	3050	1686	2589	2150	1456	3135	3442	4865	2390	4822
RECIFE	501	2074	2061	6483	2135	3247	3256	3078	3375	800	2332	120
RIO BRANCO	4763	4931	3584	2230	3123	2684	1990	3669	3976	5396	2924	5356
RIO DE JANEIRO	1855	3250	434	5159	1148	1444	2017	852	1144	2805	1338	2448
SALVADOR	356	2100	1372	5749	1446	2568	2567	2385	2682	1389	1643	949
SÃO LUÍS	1578	806	2738	6120	2157	2979	2978	3230	3537	1070	2054	1660
SÃO PAULO	2188	2933	586	4756	1015	1014	1614	408	705	3127	926	2770
TERESINA	1142	947	2302	6052	1789	2911	2910	3143	3450	634	1986	1224
VITÓRIA	1408	3108	524	5261	1238	1892	2119	1300	1597	2397	1428	2001

**QUADRO DE DISTÂNCIAS RODOVIÁRIAS ENTRE AS PRINCIPAIS
CIDADES BRASILEIRAS – Parte 02 de 03**

	MACEIÓ	MANAUS	NATAL	PALMAS	PORTO ALEGRE	PORTO VELHO	RECIFE	RIO BRANCO	RIO DE JANEIRO	SALVADOR
ARACAJÚ	294	5215	788	1662	3296	4230	501	4763	1855	356
BELÉM	2173	5298	2108	1283	3852	4397	2074	4931	3250	2100
BELO HORIZONTE	1854	3951	2348	1690	1712	3050	2061	3584	434	1372
BOA VISTA	6279	785	6770	4926	5348	1686	6483	2230	5159	5794
BRASÍLIA	1930	3490	2422	973	2027	2589	2135	3123	1148	1446
CAMPO GRANDE	3040	3051	3534	1785	1518	2150	3247	2684	1444	2568
CUIABÁ	3049	2357	3543	1784	2206	1456	3255	1990	2017	2566
CURITIBA	2871	4036	3365	2036	711	3135	3078	3669	852	2385
FLORIANÓPOLIS	3168	4443	3662	2336	476	3442	3375	3976	1144	2682
FORTALEZA	1075	5763	537	2035	4242	4862	800	5396	2805	1389
GOIÂNIA	2125	3291	2618	874	1847	2390	2332	2924	1338	1643
JOÃO PESSOA	395	5808	185	2253	3889	4822	120	5356	2448	949
MACEIÓ		5491	572	1851	3572	4505	285	5039	2131	632
MANAUS	5491		5985	4141	4563	901	5698	1445	4374	5009
NATAL	572	5985		2345	4066	4998	297	5533	2625	1126
PORTO ALEGRE	3572	4563	4066	2747		3662	3779	4196	1553	3090
PORTO VELHO	4505	901	4999	3240	3662		4712	544	3473	4023
RECIFE	285	5698	297	2058	3779	4712		5243	2338	839
RIO BRANCO	5039	1445	5533	3764	4196	544	5243		4007	4457
RIO DE JANEIRO	2131	4374	2625	2124	1553	3473	2338	4007		1649
SALVADOR	632	5009	1126	1454	3090	4023	839	4457	1649	
SÃO LUÍS	1672	5335	1607	1386	3891	4434	1573	4968	3015	1599
SÃO PAULO	2453	3971	2947	1776	1109	3070	2660	3604	429	1962
TERESINA	1236	5267	1171	1401	3804	4366	1137	4900	2579	1163
VITÓRIA	1684	4476	2178	2214	2001	3575	1891	4109	521	1202

**QUADRO DE DISTÂNCIAS RODOVIÁRIAS ENTRE AS PRINCIPAIS CIDADES
BRASILEIRAS – Parte 03 de 03**

	SANTOS	SÃO LUÍS	SÃO PAULO	TERESINA	VITÓRIA
ARACAJÚ	2249	1578	2187	1142	1408
BELÉM	3005	806	2933	947	3108
BELO HORIZONTE	658	2738	586	2302	524
BOA VISTA	4828	6120	4756	6052	5261
BRASÍLIA	1087	2157	1015	1789	1239
CAMPO GRANDE	1086	2979	1014	2911	1892
CUIABÁ	1686	2978	1614	2910	2119
CURITIBA	480	3230	408	3143	1300
FLORIANÓPOLIS	777	3537	705	3450	1597
FORTALEZA	3199	1070	3127	634	2397
GOIÂNIA	998	2054	926	1986	1428
JOÃO PESSOA	2842	1660	2770	1224	2001
MACEIÓ	2525	1672	2453	1236	1684
MANAUS	4043	5335	3971	5267	4476
NATAL	3019	1607	2947	1171	2178
PORTO ALEGRE	1181	3891	1109	3804	2001
PORTO VELHO	3142	4434	3070	4366	3575
RECIFE	2732	1573	2660	1137	1891
RIO BRANCO	3673	4968	3604	4900	4109
RIO DE JANEIRO	501	3015	429	2579	521
SALVADOR	2034	1599	1962	1163	1202
SÃO LUÍS	3042		2970	446	2607
SÃO PAULO	72	2970		2792	882
TERESINA	2864	446	2792		2171
VITÓRIA	954	2607	882	2171	

Fonte: DNIT/DNER, <http://www.dnit.gov.br/rodovias/distancias/distancias.asp>.

ANEXO B

Figura B.1 – Codificação adotada para as capitais brasileiras utilizadas.

Cód	Capital	Cód	Capital
00	Aracaju	13	Manaus
01	Belém	14	Natal
02	Belo Horizonte	15	Palmas
03	Boa Vista	16	Porto Alegre
04	Brasília	17	Porto Velho
05	Campo Grande	18	Recife
06	Cuiabá	19	Rio Branco
07	Curitiba	20	Rio de Janeiro
08	Florianópolis	21	Salvador
09	Fortaleza	22	São Luís
10	Goiania	23	São Paulo
11	João Pessoa	24	Teracina
12	Maceió	25	Vitória



Fonte: Framework Beagle XP v.4.5.

Tabela B.1 – População inicial aleatória gerada pelo Framework Beagle XP v.4.5, utilizada em todos os experimentos e testes executados.

CROMOSSOMO-PROBLEMA	CROM. COMP.	COMP.	FITNESS (Km)
04-02-10-13-03-17-06-12-01-11-08-16-25-22-15-14-05-00-09-18-23-24-07-19-21-20-04	10	T	55322
04-07-18-22-16-11-25-02-15-17-24-00-05-06-23-20-10-09-01-21-03-13-19-08-14-12-04	00	H	57909
04-00-11-21-10-01-13-12-09-07-08-03-17-16-02-23-25-18-15-20-05-06-24-22-14-19-04	21	D	58063
04-00-19-11-21-12-15-13-06-20-25-23-07-08-16-17-03-18-09-10-14-02-22-24-01-05-04	11	D	58191
04-12-10-23-14-09-02-19-13-08-11-00-21-17-16-25-20-22-24-15-18-01-03-07-06-05-04	20	H	58427
04-13-17-24-20-22-06-19-02-25-16-23-21-08-03-10-12-18-00-11-07-05-14-01-15-09-04	11	D	59254
04-09-15-22-25-17-05-08-07-18-21-02-14-11-24-19-03-06-00-13-12-20-16-01-23-10-04	00	H	59960
04-05-01-14-21-07-20-24-00-25-23-08-16-13-06-02-18-09-15-03-22-10-19-12-11-17-04	21	D	60103
04-19-20-01-10-09-21-03-24-05-15-07-25-08-06-17-13-00-11-02-23-16-18-14-12-22-04	02	H	60210
04-03-07-05-12-19-23-06-00-18-11-25-20-14-24-08-13-17-15-22-09-16-21-01-10-02-04	01	T	60622
04-18-16-01-12-11-15-25-14-05-23-08-24-19-02-21-20-09-22-17-13-03-00-07-06-10-04	10	T	60704
04-21-14-09-11-19-03-23-13-18-06-05-01-08-22-24-07-20-25-17-15-16-00-12-02-10-04	00	H	60819
04-01-02-20-21-12-17-15-25-11-22-08-16-03-06-13-07-14-09-23-19-00-24-18-10-05-04	12	D	62251
04-25-19-02-22-20-00-09-17-16-06-05-08-21-23-11-12-15-18-01-24-03-13-07-10-14-04	11	D	62282
04-18-11-22-09-14-05-00-17-12-15-10-16-03-19-20-02-01-07-21-08-25-13-06-23-24-04	01	T	62939
04-19-13-09-00-07-25-18-21-22-14-24-05-15-17-16-23-10-01-03-11-20-02-06-12-08-04	11	D	63154
04-01-05-06-19-18-11-09-23-02-08-03-24-15-00-14-21-16-25-12-17-07-10-13-22-20-04	20	H	63358
04-21-18-14-10-06-05-20-19-11-13-17-07-03-12-15-01-00-08-22-24-25-16-09-23-02-04	11	D	63507
04-17-21-03-05-16-25-02-09-12-19-07-13-08-20-18-11-00-22-06-15-24-10-14-01-23-04	01	T	63629
04-22-01-00-17-13-02-05-16-19-08-25-10-12-15-03-14-21-06-23-18-09-20-24-11-07-04	21	D	63892
04-03-19-24-07-06-14-20-02-15-23-25-09-12-22-16-01-11-10-05-13-21-18-00-08-17-04	12	D	63905
04-22-06-16-00-08-09-05-15-25-01-03-13-23-02-10-21-19-24-12-11-18-14-17-20-07-04	11	D	63949
04-21-11-18-13-20-10-14-24-01-05-06-00-09-25-16-22-15-17-23-19-02-03-12-07-08-04	12	D	64282
04-19-20-08-16-25-05-06-03-00-23-07-24-18-09-22-14-13-01-02-21-12-10-17-15-11-04	11	D	64795
04-12-18-23-14-09-21-24-02-01-08-10-15-17-00-11-05-06-03-16-20-19-07-13-25-22-04	01	T	65028
04-15-23-06-18-17-20-02-25-09-19-14-00-01-03-22-21-07-05-13-16-08-10-24-11-12-04	00	H	65251
04-05-17-02-14-08-15-25-23-03-16-22-07-24-01-13-20-18-21-11-00-12-09-06-10-19-04	20	H	65256
04-15-13-00-22-25-16-14-05-24-21-10-20-08-06-12-01-17-02-07-09-18-11-23-19-03-04	10	T	65316
04-16-07-25-09-24-18-00-17-23-05-12-03-22-01-10-19-08-11-20-06-02-14-21-13-15-04	02	H	65320
04-20-09-15-22-01-18-13-12-17-14-08-11-00-03-24-21-02-25-06-19-05-10-07-16-23-04	22	A	65564
04-17-06-10-14-19-07-20-11-01-02-15-25-08-23-24-09-16-05-13-22-12-00-03-18-21-04	01	T	65596
04-03-06-22-13-05-09-11-12-02-21-18-25-07-20-15-23-19-14-24-08-10-01-16-17-00-04	00	H	65922
04-19-07-15-06-20-23-12-17-03-24-22-21-09-02-25-14-01-18-13-10-08-11-05-00-16-04	00	H	66011

CROMOSSOMO-PROBLEMA	CROM. COMP.	COMP.	FITNESS (Km)
04-21-19-13-20-02-17-18-03-11-15-09-24-06-22-07-05-08-00-10-01-14-12-16-25-23-04	12	D	66177
04-00-21-01-13-19-07-03-14-17-22-16-05-10-20-23-06-15-25-09-12-02-18-08-11-24-04	01	T	66526
04-25-05-08-01-16-20-12-23-22-14-11-15-17-18-07-06-02-13-19-00-10-24-21-09-03-04	21	D	66604
04-00-10-11-23-22-02-17-12-21-16-25-09-13-19-06-14-18-24-05-20-03-01-08-07-15-04	11	D	66677
04-06-09-15-05-16-19-10-00-17-08-22-07-14-24-23-11-02-01-12-25-20-03-13-18-21-04	21	D	66702
04-12-05-23-11-25-21-03-02-09-24-10-00-15-06-08-16-07-01-20-17-13-14-22-19-18-04	22	A	66896
04-21-15-12-13-20-08-22-06-23-25-19-03-14-00-05-11-24-16-02-17-09-18-01-07-10-04	12	D	66903
04-14-07-00-18-10-03-01-11-20-25-21-15-23-06-08-05-24-12-13-19-16-02-22-17-09-04	10	T	67067
04-08-03-10-18-11-09-19-05-25-00-20-21-17-15-23-13-07-14-01-16-06-12-02-24-22-04	01	T	67286
04-23-15-08-11-02-07-12-10-13-03-24-18-09-01-25-21-06-16-19-20-14-05-00-17-22-04	02	H	67488
04-24-21-05-15-11-22-09-07-14-06-01-08-00-13-12-16-02-23-18-10-25-19-17-03-20-04	22	A	67712
04-25-02-15-03-07-01-05-08-18-21-14-09-22-11-13-00-19-12-20-16-24-06-23-17-10-04	00	H	67830
04-21-10-13-19-25-17-01-08-11-07-06-12-03-15-23-18-09-00-14-24-02-16-22-05-20-04	00	H	67863
04-00-07-18-02-21-20-09-15-03-16-17-19-08-05-01-06-10-11-25-22-24-12-13-14-23-04	21	D	67988
04-21-12-24-13-08-25-01-03-16-07-20-06-15-17-10-22-02-18-23-05-14-19-09-11-00-04	11	D	67995
04-10-01-03-05-13-02-14-12-07-19-06-21-11-09-08-15-22-23-17-00-16-25-18-20-24-04	10	T	68324
04-15-13-01-21-05-19-11-16-03-17-18-23-14-07-10-25-20-09-12-24-00-08-06-02-22-04	02	H	68500
04-08-20-16-09-01-24-17-00-06-11-12-02-13-14-18-19-15-21-10-22-25-03-23-05-07-04	21	D	68539
04-19-12-06-05-24-16-13-02-07-11-23-03-17-21-15-08-09-18-10-14-01-00-22-20-25-04	02	H	68563
04-20-24-19-00-25-22-14-16-02-18-15-05-17-06-03-10-23-21-07-09-08-01-11-12-13-04	22	A	69207
04-10-02-17-05-01-25-20-08-24-11-15-18-07-16-23-13-14-06-21-19-12-09-22-00-03-04	22	A	69358
04-10-17-09-06-24-16-13-19-11-00-23-25-22-01-20-12-02-21-18-03-14-08-05-07-15-04	02	H	69497
04-18-09-22-02-00-08-05-12-06-23-25-20-14-21-13-07-17-24-03-01-11-10-16-19-15-04	21	D	69517
04-08-07-21-05-00-14-22-23-11-02-19-01-15-16-13-20-09-17-18-25-24-12-10-06-03-04	11	D	69631
04-20-06-22-03-09-08-10-05-12-15-16-11-14-07-17-23-00-21-18-02-01-19-13-24-25-04	10	T	69678
04-07-02-13-05-12-25-15-21-09-06-23-03-01-24-00-11-19-18-16-08-22-10-20-14-17-04	22	A	69707
04-07-08-06-03-02-22-17-23-12-20-11-16-01-13-21-09-00-05-15-10-24-25-14-18-19-04	11	D	69845
04-05-17-19-12-13-01-16-11-06-07-08-23-18-21-09-03-22-10-25-15-00-02-24-20-14-04	20	H	69952
04-20-11-03-18-05-06-22-12-21-13-17-09-07-15-14-00-01-23-19-02-24-08-25-10-16-04	01	T	70174
04-02-01-13-23-18-17-00-11-07-14-09-24-25-20-05-12-06-03-16-19-21-08-22-10-15-04	10	T	70233
04-03-11-24-23-08-12-15-18-07-02-06-20-22-14-00-25-16-13-10-19-21-01-05-09-17-04	00	H	70239
04-19-08-13-01-15-07-24-09-17-16-06-11-03-02-20-12-22-10-14-23-05-00-21-25-18-04	00	H	70510
04-21-08-18-01-13-22-03-14-24-11-09-25-17-20-12-07-06-00-23-16-02-15-19-10-05-04	11	D	70537
04-10-08-01-19-20-07-12-05-03-06-21-15-16-18-00-14-13-17-09-23-11-02-24-25-22-04	11	D	70570
04-05-24-20-13-23-12-10-11-00-08-18-09-17-25-15-06-03-07-02-16-21-19-22-01-14-04	20	H	70632
04-15-00-03-17-18-11-08-12-05-14-24-19-10-21-20-01-02-23-07-25-06-09-22-16-13-04	20	H	71177
04-23-24-13-11-09-17-10-16-20-03-02-06-18-19-08-25-07-14-15-01-05-12-21-22-00-04	20	H	71518
04-03-00-21-22-08-18-05-13-14-11-07-16-10-02-12-19-23-20-15-01-17-09-06-24-25-04	21	D	71644
04-22-07-09-00-06-05-13-15-11-10-21-03-23-20-25-12-17-01-14-16-08-02-24-19-18-04	12	D	71765
04-11-01-18-10-20-14-09-05-22-23-24-16-07-15-25-08-02-00-17-12-19-06-13-21-03-04	22	A	71939
04-19-12-14-08-01-20-05-02-24-09-16-13-00-23-25-10-06-17-21-18-15-11-07-22-03-04	21	D	71957
04-09-17-01-10-08-11-25-21-05-12-03-02-00-13-20-07-14-18-24-16-19-06-23-15-22-04	21	D	72026
04-08-19-14-10-00-24-03-22-11-07-20-21-16-23-17-09-18-13-06-25-15-05-12-01-02-04	22	A	72169
04-14-19-21-20-12-18-08-02-23-05-15-22-07-17-16-11-00-25-06-01-13-10-09-03-24-04	22	A	72379
04-01-03-25-19-23-00-08-17-06-11-02-05-07-12-16-10-18-14-13-09-15-22-21-20-24-04	22	A	72839
04-12-21-22-23-19-06-02-01-14-25-13-09-17-00-05-07-24-08-16-20-15-18-10-11-03-04	00	H	72904
04-07-23-25-09-17-22-24-15-06-08-18-02-19-11-00-10-21-13-14-16-12-01-03-20-05-04	10	T	73070
04-03-00-05-23-15-12-10-20-18-25-06-13-14-09-16-21-22-08-19-24-11-17-02-07-01-04	11	D	73107
04-16-17-21-00-10-03-07-12-01-24-18-19-02-13-14-08-11-23-22-15-05-25-06-09-20-04	02	H	73133
04-25-09-07-12-18-06-11-21-03-02-23-13-24-05-16-20-22-15-01-10-17-14-19-00-08-04	21	D	73143
04-20-11-07-12-08-22-13-01-23-16-00-21-09-18-05-14-15-10-19-25-17-02-24-03-06-04	01	T	73164
04-25-00-05-20-01-16-09-06-18-11-08-10-15-21-02-13-12-24-03-19-22-07-14-17-23-04	10	T	73263
04-09-20-19-08-25-03-14-21-01-10-02-06-15-18-11-13-12-07-16-23-22-17-24-05-00-04	20	H	73416
04-17-01-05-19-09-08-20-03-14-00-15-24-07-21-23-16-11-12-06-18-02-22-10-13-25-04	00	H	73825
04-01-17-00-10-18-14-25-20-19-02-15-13-12-22-24-16-21-08-11-05-06-23-03-09-07-04	10	T	73891
04-15-22-14-07-02-23-11-10-12-03-18-19-25-06-00-24-09-13-08-16-01-20-17-21-05-04	20	H	73922
04-25-24-17-00-07-11-02-19-12-21-20-03-14-01-16-05-22-09-13-08-23-06-15-10-18-04	10	T	74038
04-07-17-00-03-16-24-13-23-02-18-15-09-25-22-11-01-20-08-21-05-06-12-19-10-14-04	00	H	75160
04-25-22-13-18-08-11-14-07-12-23-24-10-05-15-00-20-16-01-02-03-21-17-06-19-09-04	20	H	75414
04-10-15-17-14-07-12-02-24-00-13-22-03-16-06-05-09-19-08-21-11-23-01-18-20-25-04	20	H	75857
04-10-21-08-12-19-16-09-02-14-22-25-00-15-18-23-06-13-07-17-24-20-11-05-03-01-04	01	T	75873
04-25-14-08-18-24-00-20-19-15-12-13-11-03-09-16-02-22-07-21-23-10-01-17-05-06-04	02	H	76182
04-05-17-18-03-02-13-09-20-14-23-10-15-21-19-08-11-16-00-07-01-22-25-24-12-06-04	22	A	76454
04-23-01-00-08-17-16-12-03-18-22-25-10-05-13-21-07-15-19-02-11-24-20-14-06-09-04	12	D	76482

CROMOSSOMO-PROBLEMA	CROM. COMP.	COMP.	FITNESS (Km)
04-24-19-02-14-15-22-17-12-01-11-13-08-25-07-03-00-20-05-21-06-10-23-09-16-18-04	10	T	77156
04-11-16-21-13-10-01-12-09-25-08-22-07-18-20-06-02-24-05-15-03-00-19-14-17-23-04	00	H	79900
04-03-20-18-02-21-17-09-23-13-00-16-22-19-25-15-12-07-01-11-08-14-10-06-24-05-04	21	D	82591

Fonte: Base de dados do Framework Beagle XP v.4.5.

ANEXO C

O principal intuito deste anexo é apresentar em detalhes todas as execuções e suas configurações utilizadas neste trabalho, a fim de que sejam subsidiados de forma comprobatória os gráficos, conclusões e discussões apresentados nos capítulos finais da dissertação.

A lista seguinte diz respeito aos testes realizados no período de 2010 a 2011, e indica o conjunto comum de parâmetros iniciais adotados em todas as simulações, definidos para que estas pudessem partir do mesmo ponto no espaço de busca, bem como, apresentar as mesmas condições iniciais para comparações. São eles:

- **Problema-Teste:** PCVS-BR-26 (LEHRER, 2000);
- **Número de Testes por Configuração:** 20;
- **População Inicial:** com ponto de partida em Brasília, detalhes vide **Anexo B**;
- **Número de Indivíduos:** fixados em 100 para todas as gerações;
- **Critério de Parada:** fixado em 5.000 gerações para os “Testes-Padrão” e 50.000 para os “Super-Testes”;
- **Método de Recombinação:** PMX, detalhes vide **Capítulo III**;
- **Percentual Inicial de Recombinação:** 100%;
- **Método de Mutação:** Permuta-Simples, detalhes vide **Capítulo III**;
- **Percentual Inicial de Seleção para Mutação:** 0%;
- **Parâmetros Específicos para Testes Utilizando Interação Social (vide Anexo D):**
 - *Jogo Biológico:* Hawk-Dove (BRITO, 2004);
 - *Proporção Fenotípica Inicial de Estratégias:*

$$\frac{3}{9} Hawk : \frac{3}{9} Dove : \frac{2}{9} TFT : \frac{1}{9} Aleatório$$

- *Número de Partidas:* 100 por geração.

A Tabela C.1 apresenta a lista de configurações utilizadas associadas às suas definições, número de execuções e parâmetros variáveis, agrupados por ano de elaboração, 2005/2006 e 2010/2011.

Tabela C.1 – Conjunto de Configurações de Todas as Simulações Realizadas.

PERÍODO	ID	NOME / DESC.	GAIA		TESTES (N.º)
			TIPO	PARAM.	
2005 / 2006	01	GAIA Periódico-100, Hawk-Dove Roleta, C=20,V=30	Periódico	100	01
	02	GAIA Estocástico-25%, Hawk-Dove Roleta, C=20,V=30	Estocástico	25%	01
	03	GAIA Periódico-10, Hawk-Dove Torneio, C=30,V=20	Periódico	10	01
	04	GAIA Estocástico-50%, Hawk-Dove Torneio, C=30,V=20	Estocástico	50%	01
	05	GAIA Estocástico-25%, Hawk-Dove Roleta, C=20,V=30	Estocástico	25%	01
	06	GAIA Periódico-100, Hawk-Dove Roleta, C=20,V=30	Periódico	100	01
	07	GAIA Estocástico-75%, Hawk-Dove Roleta, C=20,V=30	Estocástico	75%	01
	08	GAIA Periódico-10, Hawk-Dove Roleta, C=20,V=30	Periódico	10	01
	09	GAIA Periódico-20, Torneio	Periódico	20	01
	10	GAIA Estocástico-25%, Hawk-Dove Roleta, C=20,V=30	Estocástico	25%	01
	11	GAIA Periódico-50, Hawk-Dove Roleta, C=20,V=30	Periódico	50	01
	12	GAIA Periódico-50, Hawk-Dove Roleta, C=20,V=30	Periódico	50	01
	13	GAIA Periódico-20, Hawk-Dove Roleta, C=20,V=30	Periódico	20	01
	14	GAIA Estocástico-50, Hawk-Dove Roleta, C=20,V=30	Estocástico	50%	01
	15	GAIA Permanente, Hawk-Dove Roleta, C=20,V=30	Permanente	-	01
	16	GAIA Periódico-10, Roleta	Periódico	10	01

PERÍODO	ID	NOME / DESC.	GAIA		TESTES (N.º)
			TIPO	PARAM.	
2005 / 2006	17	GAIA Permanente, Hawk-Dove Roleta, C=20,V=30	Permanente	-	01
	18	GAIA Periódico-10, Hawk-Dove Roleta, C=20,V=30	Periódico	10	01
	19	GAIA Permanente, Hawk-Dove Roleta, C=20,V=30	Permanente	-	01
	20	GAIA Estocástico-75%, Hawk-Dove Roleta, C=20,V=30	Estocástico	75%	01
2010 / 2011	01	Teste-Padrão GAIA Permanente (TP-GAIA-Permanente)	Permanente	-	20
	02	Teste-Padrão GAIA Periódico 20 (TP-GAIAPeriod-20)	Periódico	20	20
	03	Teste-Padrão GAIA Periódico 50 (TP-GAIAPeriod-50)	Periódico	50	20
	04	Teste-Padrão GAIA Periódico 100 (TP-GAIAPeriod-100)	Periódico	100	20
	05	Teste-Padrão GAIA Periódico 500 (TP-GAIAPeriod-500)	Periódico	500	20
	06	Teste-Padrão GAIA Periódico 1.000 (TP-GAIAPeriod-1000)	Periódico	1.000	20
	07	Teste-Padrão GAIA Periódico 2.500 (TP-GAIAPeriod-2500)	Periódico	2.500	20
	08	Teste-Padrão GAIA Estocástico 95% (TP-GAIAEstoc-95)	Estocástico	95%	20
	09	Teste-Padrão GAIA Estocástico 98% (TP-GAIAEstoc-98)	Estocástico	98%	20
	10	Teste-Padrão GAIA Estocástico 99% (TP-GAIAEstoc-99)	Estocástico	99%	20
	11	Teste-Padrão GAIA Estocástico 99,8% (TP-GAIAEstoc-99,8)	Estocástico	99,8%	20
	12	Teste-Padrão GAIA Estocástico 75% (TP-GAIAEstoc-75)	Estocástico	75%	20
	13	Teste-Padrão GAIA Estocástico 50% (TP-GAIAEstoc-50)	Estocástico	50%	20
	14	Teste-Padrão GAIA Estocástico 0,2% (TP-GAIAEstoc-0.2)	Estocástico	0,2%	20
	15	Teste-Padrão GAIA Estocástico 01% (TP-GAIAEstoc-1)	Estocástico	01%	20
	16	Teste-Padrão GAIA Estocástico 05% (TP-GAIAEstoc-5)	Estocástico	05%	20

PERÍODO	ID	NOME / DESC.	GAIA		TESTES (N.º)
			TIPO	PARAM.	
	17	Teste-Padrão GAIA Estocástico 25% (TP-GAIAEstoc-25)	Estocástico	25%	20
	18	Super-Teste GAIA Permanente (ST-GAIA-Permanente)	Permanente	-	20
	19	Super-Teste GAIA Periódico 500 (ST-GAIA-Period-500)	Periódico	500	20
	20	Super-Teste GAIA Periódico 1.000 (ST-GAIA-Period-1000)	Periódico	1.000	20
	21	Super-Teste GAIA Periódico 10.000 (ST-GAIA-Period-10000)	Periódico	10.000	20
	22	Super-Teste GAIA Estocástico 25% (ST-GAIAEstocástico-25)	Estocástico	25%	20
	23	Super-Teste GAIA Estocástico 50% (ST-GAIAEstocástico-50)	Estocástico	50%	20
	24	Super-Teste GAIA Estocástico 98% (ST-GAIAEstocástico-98)	Estocástico	98%	20
	25	Teste-Padrão Hawk-Dove Roleta (20,30) (TP-HDR2030)	<i>Sem GAIA</i>	-	20
	26	Super-Teste Hawk-Dove Roleta (20,30) (ST-HDR2030)	<i>Sem GAIA</i>	-	20
	27	Teste-Padrão Roleta Simples (TP-RoletaSimples)	<i>Sem GAIA</i>	-	20
	28	Super-Teste Roleta Simples (ST-RoletaSimples)	<i>Sem GAIA</i>	-	20
(TOTAL)		CONFIGURAÇÕES	48	TESTES	580

Fonte: Base de dados de configurações realizadas, 2005-2011.

Embora sumarizados e descritos no Capítulo VI, o intuito inicial deste anexo era apresentar os detalhes de todas as execuções e resultados obtidos. Porém, tendo em vista a enorme quantidade de dados obtidos, fica completamente inviável a sua apresentação completa neste formato de dissertação e, portanto, caso seja de interesse alheio a obtenção dos dados, que seja feito contato diretamente com o Programa de Pós-Graduação de Engenharia Elétrica, ou mesmo com o autor, através do endereço eletrônico felipebr@felipebr.com.

Ainda assim, para cada uma das execuções realizadas no segundo período (2010-2011), serão listadas as rotas do menor valor encontrado na tabela C.2, que segue.

Tabela C.2 – Configurações 2010/2011 x Melhores Rotas Encontradas.

CONFIGURAÇÃO	FITNESS	ROTA(S) / CROMOSSOMO(S)
TP-GAIA-Permanente	21.117	04-21-00-12-18-11-14-09-24-01-22-15-19-03-13-17-06-05-16-08-07-23-20-25-02-10-04 04-21-00-12-18-11-14-09-24-01-22-15-19-13-03-17-06-05-16-08-07-23-20-25-02-10-04 04-21-00-12-18-11-14-09-24-01-22-15-19-17-03-13-06-05-16-08-07-23-20-25-02-10-04 04-21-00-12-18-11-14-09-24-01-22-15-19-17-13-03-06-05-16-08-07-23-20-25-02-10-04
TP-GAIAPeriod-20	20.113	04-06-17-03-13-19-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-06-19-03-13-17-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-06-19-13-03-17-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-06-19-17-13-03-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04
TP-GAIAPeriod-50	20.287	04-06-03-13-17-19-05-07-08-16-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-06-13-03-17-19-05-07-08-16-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-06-17-13-03-19-05-07-08-16-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-06-19-13-03-17-05-07-08-16-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04
TP-GAIAPeriod-100	20.113	04-06-17-03-13-19-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04
TP-GAIAPeriod-500	21.235	04-21-00-12-18-11-14-09-24-01-22-15-02-25-20-23-16-08-07-05-06-03-13-17-19-10-04 04-21-00-12-18-11-14-09-24-01-22-15-02-25-20-23-16-08-07-05-06-13-03-17-19-10-04 04-21-00-12-18-11-14-09-24-01-22-15-02-25-20-23-16-08-07-05-06-17-13-03-19-10-04 04-21-00-12-18-11-14-09-24-01-22-15-02-25-20-23-16-08-07-05-06-19-13-03-17-10-04
TP-GAIAPeriod-1000	20.853	04-02-06-03-13-17-19-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-02-06-13-03-17-19-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-02-06-17-03-13-19-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-02-06-17-13-03-19-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-02-17-03-13-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-02-17-13-03-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-02-19-03-13-17-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-02-19-13-03-17-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-02-19-17-03-13-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-04 04-02-19-17-13-03-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-04
TP-GAIAPeriod-2500	22.085	04-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-03-13-17-19-06-10-15-04 04-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-13-03-17-19-06-10-15-04 04-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-17-03-13-19-06-10-15-04

CONFIGURAÇÃO	FITNESS	ROTA(S) / CROMOSSOMO(S)
TP-GAIAEstoc-95	21.069	04-03-13-17-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-02-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-03-13-17-19-06-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-06-03-13-17-19-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-06-13-03-17-19-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-06-17-03-13-19-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-06-17-13-03-19-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-06-19-03-13-17-05-16-08-07-23-20-04
TP-GAIAEstoc-98	21.135	04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-06-19-13-03-17-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-06-19-17-13-03-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-13-03-17-19-06-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-17-03-13-19-06-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-17-13-03-19-06-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-19-03-13-17-06-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-19-13-03-17-06-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-19-17-03-13-06-05-16-08-07-23-20-04
		04-02-25-21-00-12-18-11-14-09-24-01-22-15-10-19-17-13-03-06-05-16-08-07-23-20-04
TP-GAIAEstoc-99	20.586	04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-03-13-17-19-06-10-04
		04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-03-13-17-19-10-04
		04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-13-03-17-19-10-04
		04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-17-13-03-19-10-04
		04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-19-03-13-17-10-04
		04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-19-13-03-17-10-04
		04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-19-17-03-13-10-04
		04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-19-17-13-03-10-04
		04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-13-03-17-19-06-10-04
		04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-17-03-13-19-06-10-04
		04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-17-13-03-19-06-10-04
TP-GAIAEstoc-99,8	20.093	04-03-13-17-19-06-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04
TP-GAIAEstoc-75	21.158	04-03-13-17-19-06-10-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-04

CONFIGURAÇÃO	FITNESS	ROTA(S) / CROMOSSOMO(S)
TP-GAIAEstoc-50	21.274	04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-03-13-17-19-06-10-04 04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-06-03-13-17-19-10-04 04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-06-13-03-17-19-10-04 04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-06-17-03-13-19-10-04 04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-06-17-13-03-19-10-04 04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-06-19-03-13-17-10-04 04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-06-19-13-03-17-10-04 04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-06-19-17-03-13-10-04 04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-06-19-17-13-03-10-04 04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-13-03-17-19-06-10-04 04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-17-03-13-19-06-10-04 04-02-20-25-21-00-12-18-11-14-09-24-01-22-15-23-16-08-07-05-17-13-03-19-06-10-04
TP-GAIAEstoc-0.2	21.178	04-06-17-03-13-19-10-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-04 04-06-17-13-03-19-10-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-04 04-06-19-13-03-17-10-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-04 04-19-17-13-03-06-10-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-04
TP-GAIAEstoc-1	20.093	04-03-13-17-19-06-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04
TP-GAIAEstoc-5	20.093	04-03-13-17-19-06-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04
TP-GAIAEstoc-25	20.093	04-03-13-17-19-06-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04
ST-GAIA-Permanente	20.989	04-02-10-03-13-17-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-06-03-13-17-19-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-06-13-03-17-19-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-06-17-03-13-19-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-06-17-13-03-19-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-06-19-03-13-17-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-06-19-13-03-17-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-06-19-17-03-13-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-06-19-17-13-03-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-13-03-17-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-17-03-13-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-17-13-03-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-19-03-13-17-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-19-13-03-17-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-19-17-03-13-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-19-17-13-03-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04

CONFIGURAÇÃO	FITNESS	ROTA(S) / CROMOSSOMO(S)
ST-GAIA-Period-500	20.752	04-21-00-12-18-11-14-09-24-01-22-15-10-03-13-17-19-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-03-13-17-19-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-13-03-17-19-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-17-03-13-19-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-19-03-13-17-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-19-13-03-17-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-13-03-17-19-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-17-03-13-19-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-19-03-13-17-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-19-13-03-17-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-19-17-03-13-06-05-16-08-07-23-20-25-02-04
ST-GAIA-Period-1000	20.586	04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-03-13-17-19-06-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-13-03-17-19-06-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-17-03-13-19-06-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-17-13-03-19-06-10-04
ST-GAIA-Perid-10000	20.093	04-03-13-17-19-06-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04
ST-GAIAEstoc-25	20.752	04-21-00-12-18-11-14-09-24-01-22-15-10-03-13-17-19-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-03-13-17-19-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-13-03-17-19-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-17-03-13-19-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-17-13-03-19-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-19-03-13-17-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-19-13-03-17-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-19-17-03-13-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-06-19-17-13-03-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-13-03-17-19-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-17-03-13-19-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-17-13-03-19-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-19-03-13-17-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-19-13-03-17-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-19-17-03-13-06-05-16-08-07-23-20-25-02-04 04-21-00-12-18-11-14-09-24-01-22-15-10-19-17-13-03-06-05-16-08-07-23-20-25-02-04
ST-GAIAEstoc-50	20.093	04-03-13-17-19-06-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04

CONFIGURAÇÃO	FITNESS	ROTA(S) / CROMOSSOMO(S)
ST-GAIAEstoc-98	20.586	04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-03-13-17-19-06-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-03-13-17-19-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-13-03-17-19-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-17-03-13-19-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-17-13-03-19-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-19-03-13-17-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-19-13-03-17-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-19-17-03-13-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-06-19-17-13-03-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-13-03-17-19-06-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-17-03-13-19-06-10-04 04-15-01-22-24-09-14-11-18-12-00-21-25-02-20-23-16-08-07-05-17-13-03-19-06-10-04
TP-HDR2030	21.089	04-06-17-03-13-19-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-02-04 04-06-17-13-03-19-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-02-04 04-19-03-13-17-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-02-04 04-19-17-03-13-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-02-04 04-19-17-13-03-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-02-04
ST-HDR2030	20.989	04-02-10-03-13-17-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-06-03-13-17-19-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-13-03-17-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-17-03-13-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-17-13-03-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-19-03-13-17-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-19-13-03-17-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04 04-02-10-19-17-13-03-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-04
TP-RoletaSimples	21.069	04-03-13-17-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-02-04
ST-RoletaSimples	21.069	04-03-13-17-19-06-05-16-08-07-23-20-25-21-00-12-18-11-14-09-24-01-22-15-10-02-04
MELHOR (FITNESS)	20.093	04-03-13-17-19-06-05-16-08-07-23-20-02-25-21-00-12-18-11-14-09-24-01-22-15-10-04 <i>* ENCONTRADO EM 06 EXECUÇÕES, MENOR VALOR JÁ ENCONTRADO PARA O PROBLEMA PCVS-BR-26 EM TODA LITERATURA.</i>

Fonte: Base de dados de configurações realizadas, 2010/2011.

Por fim, apresenta-se uma listagem com a totalização da quantidade de dados produzidos – no período 2010/2011- para que se tenha uma dimensão exata do tamanho da pesquisa e da complexidade de tabulação dos dados de todos os experimentos. São eles:

- **Banco de Dados (Firebird 2.0):**
 - **Quantidade de Bases:** 10;
 - **Tamanho:** 31,60 GB;
 - **Registros (aproximadamente):** 100.000.000;

- **Arquivos MS-Excel 2010:**
 - **Quantidade:** 09;
 - **Tamanho:** 313,60 MB;
 - **Planilhas:** 84;
 - **Linhas:** 1.415.000;
 - **Células (Valores):** 28.300.000;
 - **Gráficos:** 56;

ANEXO D

O presente anexo expõe a aplicação desenvolvida para dar suporte à pesquisa. Desenvolvida com o propósito de realização de testes de execução do Algoritmo Genético com a estrutura do GAIA voltado à solução do problema PCVS-BR-26, sua indexação, registro e catalogação.

A aplicação desenvolvida é chamada de Beagle XP ou *Beagle Experimental Prototype*, numa alusão ao navio que transportou Charles Darwin em sua viagem ao redor do mundo, que deu origem ao livro “*The Origin of Species*”, e a Seleção Natural, conforme apresentado no capítulo segundo.

O protótipo, desenvolvido originalmente em BRITO (2004) para dar suporte à proposta Híbrida de Algoritmos Genéticos com Interação Fenotípica (vide Anexo E), foi construído utilizando a linguagem JAVA e o Sistema Gerenciador de Banco de Dados Firebird 2.0 e, ao longo do período até o ano de 2010, sofreu uma série de evoluções que o levaram ao suporte da proposta central deste trabalho, tais como o desenvolvimento do motor fuzzy de execução do GAIA, e a inclusão da biblioteca FUZZYJ (2006) para a implementação deste.

Em síntese, é uma ferramenta gráfica para execução de Algoritmos Genéticos para o problema do PCVS aplicado às distâncias entre as capitais brasileiras, e depuração das técnicas e métodos propostos pela idéia central de motivação deste trabalho.

A depuração é alcançada através da base de dados gerada e da possibilidade de definição de inúmeros parâmetros de configuração por meio da interface, além da manutenção em banco de dados de populações iniciais, o que garante à todos os testes a partida do mesmo ponto da superfície de busca.

Desta forma, comporta cinco módulos básicos, que são:

- **Módulo de Algoritmo Genético:** *Engine* básica do sistema que receberá os parâmetros e gerará os indivíduos em busca de uma solução/otimização para o problema.
 - **Entrada:** Parâmetros armazenados de configuração (estratégias, jogos, número de ciclos, métodos de recombinação, GAIA, etc.);
 - **Ator:** Pesquisador;
 - **Saída:** Gerações de indivíduos/possíveis soluções aos problemas;

- **Módulo de Configuração:** Parte inerente a interface gráfica que permitirá ao pesquisador definir todos os parâmetros para execução do Algoritmo Genético baseado no problema do Caixeiro Viajante Simétrico para as Capitais Brasileiras, PCVS-BR-26 (LEHRER, 2000), sendo os principais:
 - **População inicial, escolhendo uma cidade inicial.**
 - **Número de indivíduos da população;**
 - **Número de ciclos (gerações) ou definição de ponto de parada;**
 - **Percentual de cruzamento;**
 - **Percentual de mutação;**
 - **Interação Social:**
 - **Jogo:** “*Hawk-Dove*”, ou “*Dilema do Prisioneiro Simétrico*”, ou “*Dilema do Prisioneiro Assimétrico*”;
 - **Número de partidas (a cada confronto);**
 - **Número de rodadas (ou confrontos) por geração;**
 - **Valores da Tabela de Pagamento de cada jogo;**
 - **Genetic Algorithm Inference Agent - GAIA:**
 - **Tipo:** “*Permanente*”, “*Periódico*”^{*} ou “*Estocástico*”^{*};

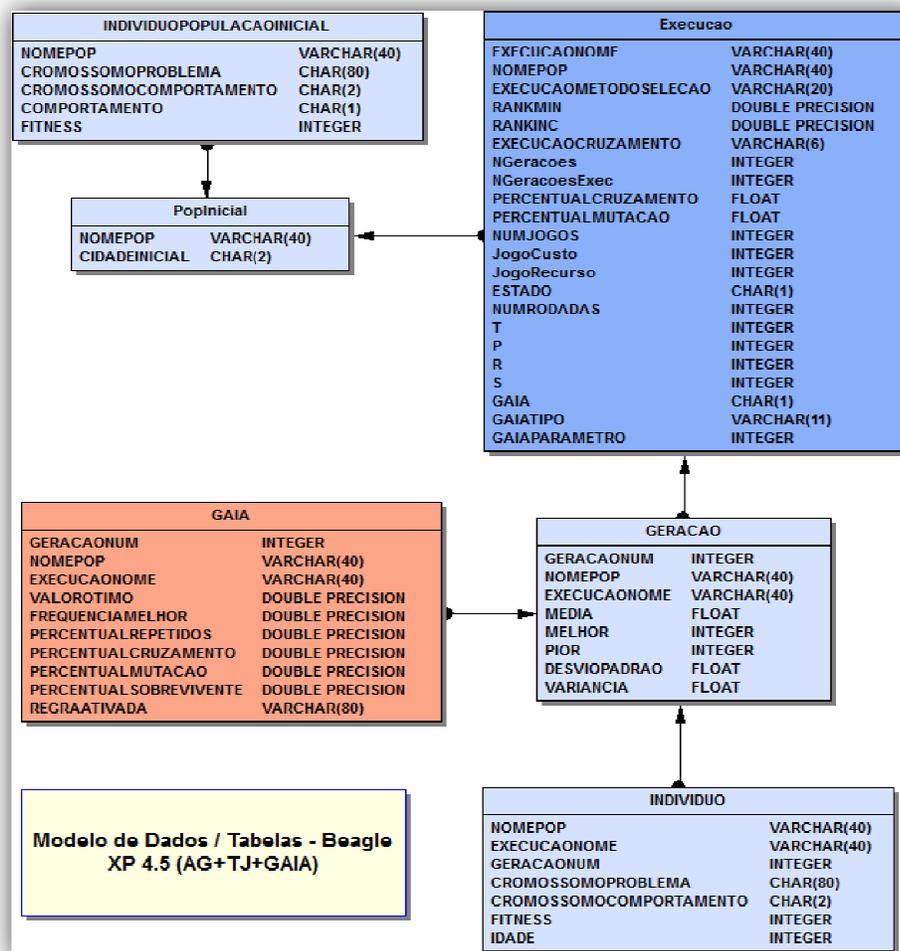
- ***Parâmetros:** *Quantidade de gerações se “Periódico” ou, probabilidade de ocorrência a cada geração se for do tipo “Estocástico”;*
- **Método de seleção:**
 - **Roleta;**
 - **Torneio;**
 - **Rank Linear;**
 - **Janelamento;**
 - **Hawk-Dove Roleta (HDR);**
 - **Hawk-Dove Torneio (HDR);**
 - **Hawk-Dove Rank Linear (HDRL);**
 - **Hawk-Dove Janelamento (HDJ);**
 - **Dilema do Prisioneiro Roleta (DPR);**
 - **Dilema do Prisioneiro Torneio (DPT);**
 - **Dilema do Prisioneiro Rank Linear (DPRL);**
 - **Dilema do Prisioneiro Janelamento (DPJ);**
 - **Entrada:** Parâmetros definidos pelo usuário na interface gráfica;
 - **Ator:** Pesquisador;
 - **Saída:** Valores de configuração para a execução da *engine* do AG e do GAIA;
- **Módulo de Resultados e Estatísticas:** Este módulo é destinado à apuração dos resultados obtidos com a configuração utilizada. Seu principal objetivo é gerar dentro da própria aplicação formas visuais de acompanhar a evolução do *fitness* da população, bem como outros dados médios da população e uma comparação com execuções anteriores armazenadas em BD;

- **Entrada:** Valores provenientes do Banco de Dados e definições de estatísticas pelo próprio usuário;
 - **Ator:** Pesquisador;
 - **Saída:** Gráficos, listagens, informações e comparações entre diferentes execuções;
-
- **Módulo GAIA:** *Engine* especial do sistema que possui as variáveis de entrada e saída, bem como a base de regras modeladas em termos de valores fuzzy, que executa toda a lógica da proposta GAIA de monitoramento, controle e intervenção na população durante uma execução;
 - **Entrada:** Dados de configuração do GAIA gravados no módulo de configuração do sistema e a população durante a execução;
 - **Ator:** Sistema;
 - **Saída:** População pós-intervenção dados os valores obtidos pelo motor de inferência Fuzzy e a base de regras estabelecida;
-
- **Módulo de Banco de Dados:** Oferece a conexão com o SGBD e possibilitará manipular os registros armazenados de forma individual e manual, oferecendo operações comuns de banco de dados;
 - **Entrada:** Dados gerados pelo *engine* do sistema;
 - **Ator:** Pesquisador;
 - **Saída:** Dados armazenados para os módulos de estatística e para manipulação pelo usuário, caso haja necessidade;

O uso de uma base de dados é de significativa importância para este trabalho, e também para trabalhos futuros, que poderão encontrar padrões que descrevam a evolução das estratégias dentro de uma população. Além disso, é a garantia do reuso de populações iniciais para todos os testes e da facilidade de buscas e pesquisas utilizando as facilidades e recursos da linguagem SQL.

As figuras D.1 e D.2, listadas a seguir, apresentam, respectivamente o modelo de dados utilizado e o esquema de classes de entidades utilizadas na aplicação.

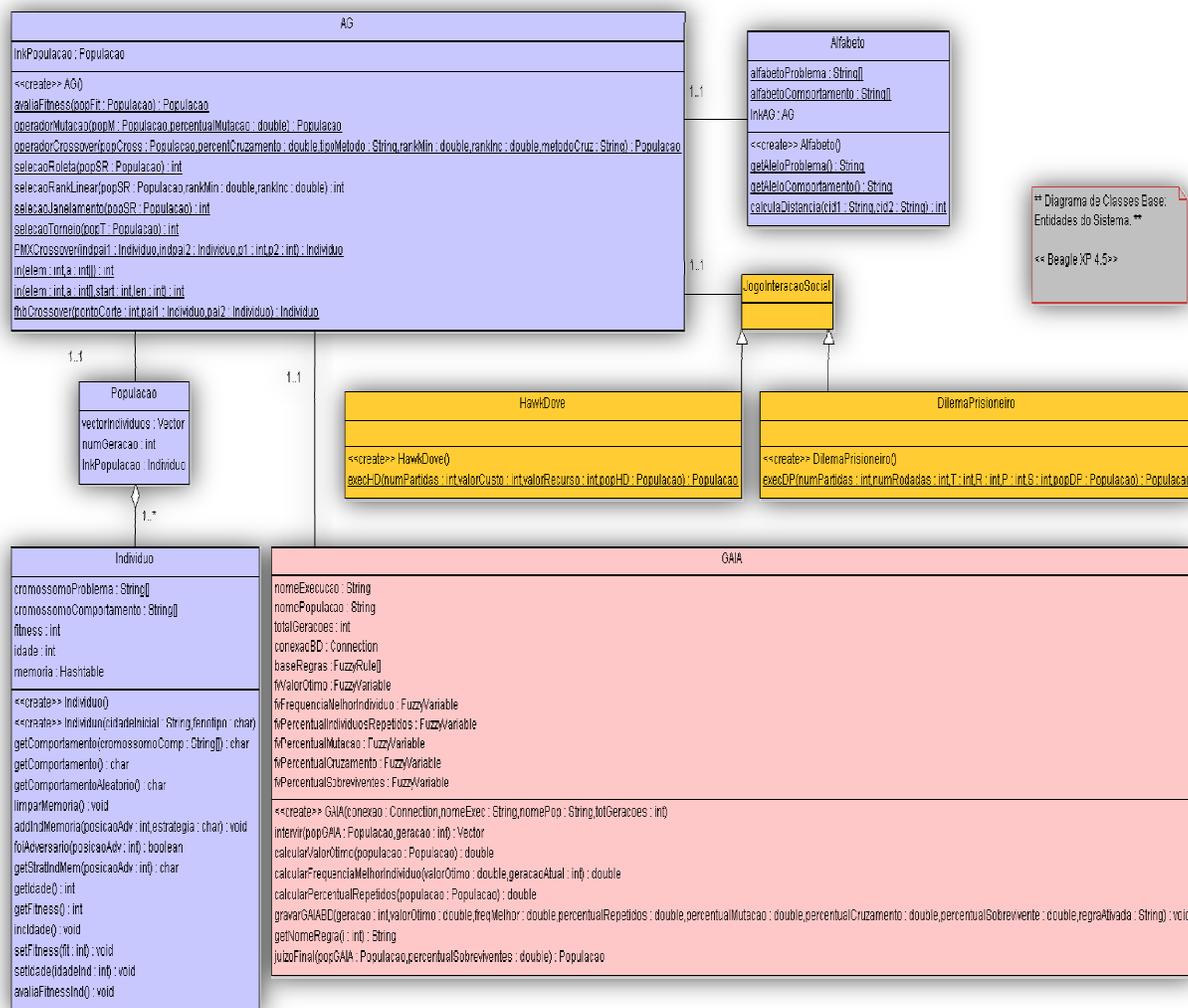
Figura D.1 – Modelo de Dados / Tabelas Framework Beagle XP 4.5



Fonte: Framework Beagle XP v.4.5.

Deve-se atentar entretanto, que a simplicidade e a não-normalização adotadas refletem as dificuldades de performance encontradas durante às execuções, já que uma única execução chega a gerar até mais de 02 milhões de registros, dificultando a pesquisa posterior na base de dados e, até mesmo, em alguns casos, levando a erros de estouro de memória e *crash* da aplicação.

Figura D.2 – Diagrama de Classes – Entidades do Framework Beagle XP 4.5

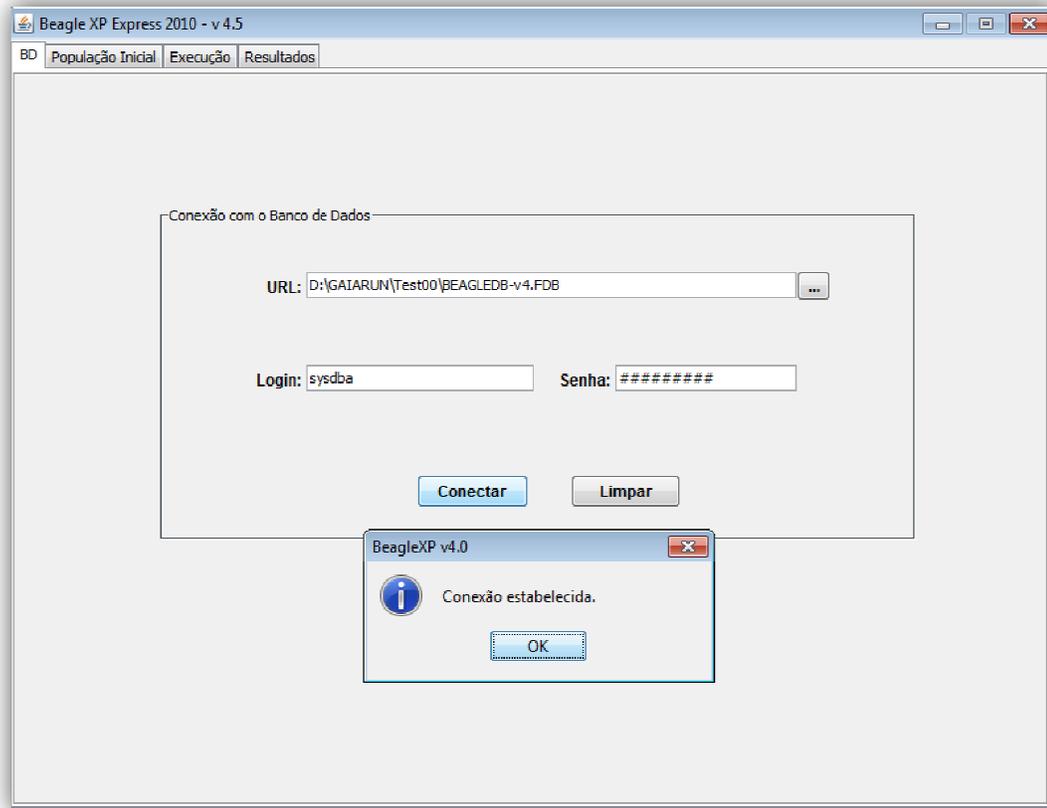


Fonte: Framework Beagle XP v.4.5.

As próximas figuras listadas neste anexo são destinadas à demonstração das quatro principais telas do sistema Beagle XP 4.5. São elas: conexão a uma base de dados, criação e manutenção de populações iniciais, manutenção de configurações de execuções e execução

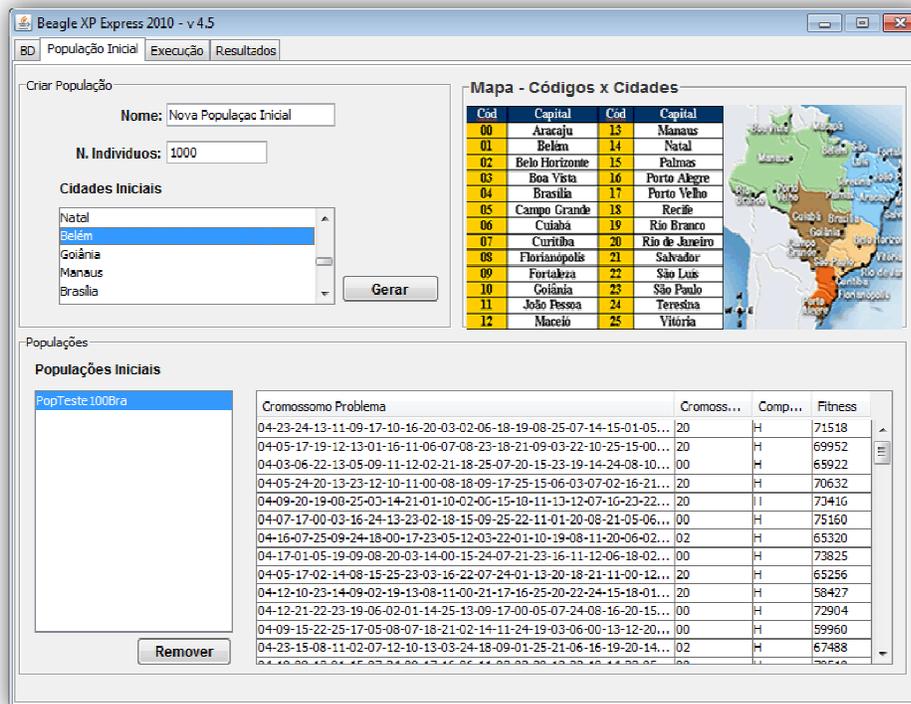
destas e, por fim, análise e pesquisa dos resultados encontrados. Respectivamente, figuras D.3, D.4, D.5 e D.6.

Figura D.3 – Beagle XP 4.5 Tela 01 – Conexão a uma Base de Dados



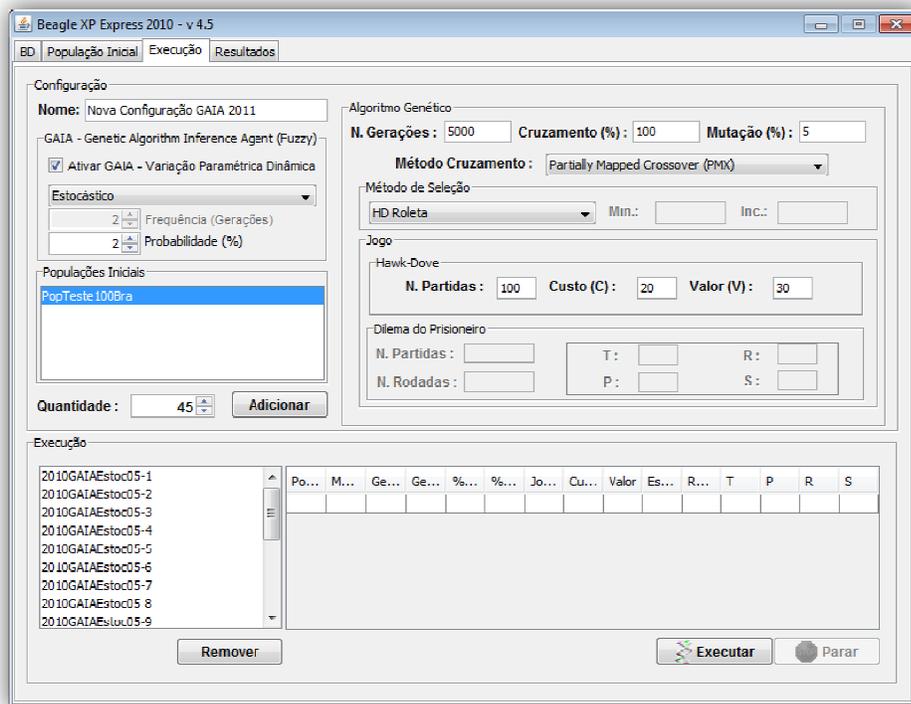
Fonte: Framework Beagle XP v.4.5.

Figura D.4 – Beagle XP 4.5 Tela 02 – Manutenção de Populações Iniciais



Fonte: Framework Beagle XP v.4.5.

Figura D.5 – Beagle XP 4.5 Tela 03 – Manutenção de Configurações



Fonte: Framework Beagle XP v.4.5.

Figura D.6 – Beagle XP 4.5 Tela 04 – Análise e Pesquisa de Resultados

<p>Fonte: Framework Beagle XP v.4.5.</p>
 </div>
 <div data-bbox="167 600 876 782" data-label="Text">
 <p>Em meados de 2008, deu-se início a um novo projeto Beagle XP, intitulado de Beagle XP Framework 6.0 (ou 20XX), muito mais flexível e configurável, vislumbrando ser uma ferramenta única para implementação de quaisquer problemas de otimização em termos de solução por meio de algoritmos evolucionários e das técnicas híbridas propostas pelo grupo de pesquisa, tais como: populações paralelas, GAIA com mais variáveis de controle e total independência do pesquisador, interação social baseada em comportamentos fuzzy e outros.</p>
 </div>
 <div data-bbox="167 796 876 912" data-label="Text">
 <p>Desde julho de 2009, entretanto, devido a falta de recursos (até o momento de entrega deste trabalho), este projeto se encontra parado. Foram concluídas suas etapas de definição de projeto, análise e modelagem, além de cerca de 40% do esforço de implementação que já foi realizado. Aguardar-se-á por novas oportunidades de conclusão e evolução desta.</p>
 </div>

ANEXO E

Este anexo é dedicado ao aprofundamento na proposta **Híbrida de Algoritmos Genéticos Baseados na Interação Fenotípica por meio da Teoria dos Jogos Evolucionários**, e está constituído de duas principais partes: na primeira, uma listagem dos principais trabalhos e publicações rendidas por esta proposta ao grupo de pesquisa científica do qual faz parte este autor e, na segunda, a íntegra da primeira publicação de significância em português alcançada pelo grupo.

Esta abordagem híbrida fora inclusive utilizada como base de todos os testes envolvendo a proposta do GAIA, conforme já citado no texto principal da dissertação, e é originária dos trabalhos de LEHRER (2000) e BRITO (2004), evoluída, aprofundada e discutida nos demais trabalhos do grupo listados na seção seguinte.

E.1 ALGORITMOS GENÉTICOS HÍBRIDOS COM INTERAÇÃO FENOTÍPICA: PRINCIPAIS PUBLICAÇÕES E TRABALHOS

Em ordem decrescente de publicação, da mais recente para mais antiga:

TEIXEIRA, Otávio Noura ; LOBATO, Walter Avelino da Luz ; YASOJIMA, Carlos Takeshi Kudo ; BRITO, Felipe Houat de ; TEIXEIRA, Artur Noura ; OLIVEIRA, Roberto Célio Limão de . **Fuzzy Social Interaction Genetic Algorithm**. In: GECCO 2010 - Genetic and Evolutionary Computation Conference, 2010, Portland, Oregon. Proceedings of the 2010 Genetic and Evolutionary Computation Conference. Alpha, New Jersey: Sheridan Printing Company, Inc, 2010. p. 2113-2114.

BRITO, Felipe Houat de ; TEIXEIRA, Artur Noura ; TEIXEIRA, Otávio Noura ; OLIVEIRA, Roberto Célio Limão de . **Hawk-Dove Tournament: A New Selection Technique for Genetic Algorithms Based on Evolutionary Game Theory**. In: Worldcomp 2009 - World Congress in Computer Science, Computer Engineering and Applied Computing, 2009, Las Vegas, NV, USA. GEM 2009 - Genetic and Evolutionary Methods, 2009.

TEIXEIRA, Otávio Noura ; BRITO, Felipe Houat de ; TEIXEIRA, Artur Noura ; OLIVEIRA, Roberto Célio Limão de . **Genetic Algorithms with Social Interaction: A New Step for Phenotype Characterization Thru Game Theory.** In: Worldcomp 2007 - World Congress in Computer Science, Computer Engineering and Applied Computing, 2007, Las Vegas. GEM 2007 - Genetic and Evolutionary Methods, 2007.

TEIXEIRA, O. N. ; BRITO, F. H. ; TEIXEIRA, A. N. ; OLIVEIRA, R. C. L. . **Genetic Algorithms with Social Interaction Phase as Phenotype Characterization.** In: Second International Conference on Natural Computation (ICNC'06) and the Third International Conference on Fuzzy Systems and Knowledge Discovery (FSKD'06), 2006, Xi'an. Advances in Natural Computation and Data Mining. Xi'an, China: Xidian University Press, 2006. p. 48-58.

TEIXEIRA, O. N. ; BRITO, F. H. ; TEIXEIRA, A. N. ; OLIVEIRA, R. C. L. . **O Paradigma do Dilema do Prisioneiro Aplicado ao Método de Seleção Roleta dos Algoritmos Genéticos.** In: CLEI 2006 - Conferencia Latinoamericana de Informática, 2006, Santiago. Proceedings of CLEI 2006 - Conferencia Latinoamericana de Informática, 2006.

BRITO, F. H. ; TEIXEIRA, A. N. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. . **Aplicando datamining com redes SOM em base de dados de execuções de algoritmos genéticos utilizando os novos métodos de seleção hawk-dove roleta e hawk-dove torneio.** In: CLEI 2006 - Conferencia Latinoamericana de Informática, 2006, Santiago. Proceedings of CLEI 2006 - Conferencia Latinoamericana de Informática, 2006.

TEIXEIRA, O. N. ; BRITO, F. H. ; TEIXEIRA, A. N. ; OLIVEIRA, R. C. L. . **Game Theory as a New Paradigm for Phenotype Characterization of Genetic Algorithms.** In: GECCO 2006 - Genetic and Evolutionary Computation Conference, 2006, Seattle-WA. Proceedings of GECCO 2006 - Genetic and Evolutionary Computation Conference. New York : ACM Press, 2006. v. 2. p. 1431-1432.

BRITO, F. H. ; TEIXEIRA, A. N. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. . **Hawk-Dove Torneio: Um Novo Método de Seleção para os Algoritmos Genéticos.** In: SEPAI/CETIC, 2006, Belém, Pará.

BRITO, F. H. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. **A introdução da interação fenotípica em algoritmos genéticos através dos jogos evolucionários e da codificação e transmissão genética do comportamento** (Presented Conference Paper Style), presented at VII SBAI/II IEEE LARS: Simpósio Brasileiro de Automação Inteligente / II Latin-American IEEE Robotics Symposium, São Luís, MA, September 18-23 2005.

TEIXEIRA, Otávio Noura. **Uma proposta de um novo algoritmo genético baseado na teoria dos jogos**. 290f. Dissertação (Mestrado em Engenharia Elétrica – Computação Aplicada) – Departamento de Engenharia Elétrica e Computação, Universidade Federal do Pará, Belém, 2005.

BRITO, Felipe Houat de. **Hawk-Dove torneio**: um novo método de seleção para os algoritmos genéticos baseado na teoria dos jogos evolucionários e estratégias evolucionárias. 143f. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) – Centro Universitário do Estado do Pará, Belém, 2004.

E.2 A INTRODUÇÃO DA INTERAÇÃO FENOTÍPICA EM ALGORITMOS GENÉTICOS ATRAVÉS DOS JOGOS EVOLUCIONÁRIOS E DA CODIFICAÇÃO E TRANSMISSÃO GENÉTICA DE DO COMPORTAMENTO*

**Artigo apresentado na íntegra na página seguinte com adaptações de formato e contagem própria e exclusiva de páginas.*

<<(IMPRIMIR – E ANEXAR HDT SBAI 2005)>>

<<REMOVER ESTA PÁGINA>>

ANEXO F

A construção do agente inteligente GAIA para monitoramento e controle da evolução em algoritmos genéticos rendeu ao autor deste trabalho e ao seu grupo de pesquisa um número significativo de publicações em eventos e periódicos internacionais de prestígio, tendo este autor, inclusive, apresentado a presente pesquisa na China (BRITO et al., 2006a) e nos EUA (BRITO et al., 2009a), nos anos de 2006 e 2009, respectivamente.

Apesar de que as publicações não apresentam a mesma quantidade de testes e resultados desta dissertação - já que estes são mais recentes e inéditos até sua defesa -, nas seções seguintes são expostos uma listagem de todas as publicações e a última publicação e seu conteúdo na íntegra, em formato próprio e idioma inglês.

F.1 GENETIC ALGORITHM INFERENCE AGENT (GAIA): PRINCIPAIS PUBLICAÇÕES E TRABALHOS

BRITO, Felipe Houat de ; TEIXEIRA, Artur Noura ; TEIXEIRA, Otávio Noura ; OLIVEIRA, Roberto Célio Limão de . **Darwin-Less Evolutionary Algorithm - Less Randomness, More Intelligence**. In: Worldcomp 2009 - World Congress in Computer Science, Computer Engineering and Applied Computing, 2009, Las Vegas, NV, USA. GEM 2009 - Genetic and Evolutionary Methods, 2009.

BRITO, Felipe Houat de ; TEIXEIRA, Artur Noura ; TEIXEIRA, Otávio Noura ; OLIVEIRA, Roberto Célio Limão de . **GAIA: A Fuzzy Intelligent Controller for Genetic Algorithms Parameters and Performance**. In: Worldcomp 2007 - World Congress in Computer Science, Computer Engineering and Applied Computing, 2007, Las Vegas. GEM 2007 - Genetic and Evolutionary Methods, 2007.

BRITO, F. H. ; TEIXEIRA, A. N. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. . **A Fuzzy Approach to Control Genetic Algorithm Parameters**. In: Electronic Journal of SADIO

(Sociedad Argentina de Informática e Investigación Operativa), 2006, Argentina, Buenos Aires. (Publicação em Revista Internacional).

BRITO, F. H. ; TEIXEIRA, A. N. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. . **A Fuzzy Intelligent Controller for Genetic Algorithms' Parameters.** In: Second International Conference on Natural Computation (ICNC'06) and the Third International Conference on Fuzzy Systems and Knowledge Discovery (FSKD'06), 2006, Xi'an. Advances in Natural Computation - Second International Conference, ICNC 2006, Xi an, China, September 24-28, 2006. Proceedings, Part I, 2006. v. 4222. p. 633-642. (Capítulo de Livro).

BRITO, F. H. ; TEIXEIRA, A. N. ; TEIXEIRA, O. N. ; OLIVEIRA, R. C. L. . **A Fuzzy Approach to Control Genetic Algorithm Parameters.** In: ASAI 2006 - VIII Argentine Symposium on Artificial Intelligence, 2006, Mendoza. Proceedings of VIII Argentine Symposium on Artificial Intelligence, 2006.

F.2 DARWIN-LESS EVOLUTIONARY ALGORITHM - LESS RANDOMNESS, MORE INTELLIGENCE*

**Artigo apresentado na íntegra na página seguinte com adaptações de formato e contagem própria e exclusiva de páginas.*

<<(IMPRIMIR – E ANEXAR GAIA WCC 2009a)>>

<<REMOVER ESTA PÁGINA>>