



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

RENATO SIMÕES MOREIRA

**ALGORITMO GENÉTICO RETROVIRAL
ITERATIVO**

DM 18/2010

Belém, Setembro de 2010

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

RENATO SIMÕES MOREIRA

**ALGORITMO GENÉTICO RETROVIRAL
ITERATIVO**

DM 18/2010

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.^a Carolina de Mattos Affonso, Dra.

Belém, Setembro de 2010.

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

RENATO SIMÕES MOREIRA

**ALGORITMO GENÉTICO RETROVIRAL
ITERATIVO**

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: ___/___/___

Banca Examinadora:

Prof.^a Carolina de Mattos Affonso, Dra.
Orientadora
Universidade Federal do Pará – UFPA

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

Prof. Ádamo Lima de Santana, Dr.
Membro da Banca
Universidade Federal do Pará - UFPA

Prof. Gustavo Augusto Lima de Campos, Dr.
Membro da Banca
Universidade Estadual do Ceará – UECE

Visto

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

Belém, Setembro de 2010.

"Endless forms most beautiful and most wonderful have been, and are being, evolved" (Charles Darwin)

À minha sobrinha e afilhada Isabela.

AGRADECIMENTOS

Aos meus avós Itamar e Marilda e Nívea e Eurico por terem gerado meus pais dando-lhes educação e caráter.

Aos meus pais Orlando e Celiamar por todos os sacrifícios a mim dispensados até os dias de hoje, sem os quais, não seria capaz de me tornar o que sou hoje. Mãe e pai, vocês são e sempre serão minha gratidão e razão de viver.

Ao meu irmão Neto, por, incondicionalmente ter estado ao meu lado praticamente todos os momentos da minha vida, da infância até os dias de hoje, e que apesar da distância, sempre foi a pessoa mais presente na minha vida e também por ter fornecido minha mais bela ‘sobrinhada’, ‘belinha’. E ao meu irmão Júnior, por ter me dado o primeiro sobrinho, o ‘beco’.

À minha esposa Luana, companheira e amiga de oito anos de muitas lutas e vitórias, lágrimas e risos, ‘ricos’ e ‘subways’. Obrigado por me deixar voar toda vez que tenho asas e por não me deixar quando você percebe que elas estão desgastadas. E acima de tudo obrigado por sempre estar do meu lado toda vez que eu preciso. Amo muito você minha pippa.

À minha sogra Ivonete, por acreditar e ouvir sempre nas minhas loucuras e me apoiar nas minhas decisões.

Às minhas tias Deth, Mary, Tereza e Ita.

Aos meus tios Moreira, Cirilo, Ninga, Aldo e Tio Ita.

À minha prima amada Dayane, que ‘ora pois’ foi minha fiel escudeira durante esses longos 27 anos, e mesmo estando em Portugal me apoiou muito neste trabalho.

Ao meu sobrinho Beco e sobrinha Pipinha, por trazerem alegria pra nossa família.

Aos meus amigos longa data Samuel, Julielson e Mara.

Aos meus amigos de trabalho Daniel, Allan, Francino, Cândido, Marcelo, Cristiano, Erika, Átila, Wanderson, Pablo, Fábio e Kenta.

Aos meus professores de graduação do CESUPA: Venícius, Douglas, Rafael Boulhosa, Cláudia Mesquita e Leda por acreditarem e apostarem em mim desde o início da graduação e minha professora de biologia Odete por me fazer ser fascinado pelas estruturas da natureza.

Ao professor Gustavo Campos por sempre ter sido fonte de inspiração durante todos esses anos.

À Fóton Informática, por toda compreensão durante o meu programa de mestrado.

Ao meu grande mestre, amigo e co-orientador Otávio Noura, que, junto comigo, viajou bastante no mundo da imaginação desde minha chegada a Belém. Lembra que chegamos a ter até reunião com advogada para que uma idéia nossa não fosse roubada?

Ao professor Roberto Limão pela oportunidade única a mim dada.

A UNIFRAN, UFPA e CESUPA por terem sido minhas instituições de graduação e pós-graduação.

E por fim, todas aquelas pessoas que não citei, mas que fizeram de todo meu desenvolvimento pessoal, acadêmico e profissional. Muito obrigado.

SUMÁRIO

LISTA DE FIGURAS.....	xii
LISTA DE TABELAS.....	xiv
LISTA DE ABREVIATURAS.....	xv
PUBLICAÇÕES.....	xvi
RESUMO.....	xvii
ABSTRACT.....	xviii

CAPÍTULO I.....	1
1. INTRODUÇÃO.....	1
1.1. PREÂMBULO.....	1
1.2. OBJETIVOS GERAIS.....	1
1.3. OBJETIVOS ESPECÍFICOS.....	2
1.4. ESTRUTURA DA DISSERTAÇÃO.....	2
2. EMBASAMENTO BIOLÓGICO: DOS VÍRUS AOS RETROVÍRUS.....	3
2.1. VÍRUS.....	3
2.1.1. Introdução.....	3
2.1.2. Importância do Estudo dos Vírus.....	5
2.1.2.1. Alguns Vírus são Úteis.....	5
2.1.3. Estrutura.....	7
2.1.4. Replicação.....	9
2.1.5. Classificação dos Vírus.....	10
2.2. RETROVÍRUS.....	11
2.2.1. Introdução.....	11
2.2.2. Replicação.....	12
2.3. EVOLUÇÃO VIRAL.....	15
2.3.1. Introdução.....	15
2.3.2. Mecanismos de Evolução.....	15
2.3.2.1. Mutação.....	16
2.3.2.2. Recombinação.....	16
2.3.2.3. Rearranjo.....	17
2.3.2.4. Aquisição de Genes Celulares.....	18
CAPÍTULO III.....	19

3. COMPUTAÇÃO EVOLUCIONÁRIA.....	19
3.1. ALGORITMO GENÉTICO.....	19
3.1.1. Introdução.....	19
3.1.2. Componentes.....	21
3.1.2.1. Cromossomo.....	21
3.1.2.2. Função de avaliação (<i>Aptidão</i>).....	22
3.1.2.3. População.....	22
3.1.2.4. Seleção.....	22
3.1.2.5. Operadores de Mutação e Recombinação.....	23
3.1.3. Passos de um AG.....	24
3.2. ALGORITMO GENÉTICO COM INFECÇÃO VIRAL.....	25
CAPÍTULO IV.....	27
4. MATERIAL E MÉTODOS.....	27
4.1. ALGORITMO GENÉTICO RETROVIRAL ITERATIVO.....	27
4.1.1. Introdução.....	27
4.1.2. O Comportamento Viral no AGRI.....	29
4.1.2.1. Vírus.....	29
4.1.2.2. População Viral e Criação de Novos Vírus.....	30
4.1.2.3. Infecção.....	30
4.1.2.4. Ciclo de vida de um Vírus.....	32
4.1.2.5. Parâmetros.....	33
4.1.3. Algoritmo.....	36
CAPÍTULO V.....	37
5. RESULTADOS.....	37
5.1. Funções Utilizadas.....	37
5.2. Experimentos.....	39
5.3. Análise dos Resultados.....	40
CAPÍTULO VI.....	42
6. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS.....	42
REFERÊNCIAS BIBLIOGRÁFICAS.....	44
ANEXO A.....	46
ANEXO B.....	50

ANEXO C..... 54
ANEXO D..... 58

LISTA DE FIGURAS

Figura 1 - Comparativo de partícula de herpesvírus, um balão e a terra.....	4
Figura 2 - Mimivírus, um dos maiores Vírus e o parvovírus (apontado), um dos menores Vírus.	4
Figura 3 - Estruturas Virais	7
Figura 4 - Representação da estrutura viral.....	8
Figura 5 - Ataque de um Vírus bacteriófago	9
Figura 6 - Formação do complexo de transcrição reversa.....	12
Figura 7 - Criação do provírus e efetivação da transcrição reversa.....	13
Figura 8 - Ciclo de infecção viral.....	13
Figura 9 - Taxa de mutação nos processos de replicação viral.	14
Figura 10 - Processo de recombinação viral.....	17
Figura 11 - Representação de um cromossomo binário.....	21
Figura 12 - Seleção Roleta (<i>Rank Selection</i>).....	23
Figura 13 - Exemplo cruzamento com 1 ponto de corte.....	24
Figura 14 - Exemplo de mutação em cromossomo.....	24
Figura 15 - Vírus possíveis para um cromossomo binário de oito posições	30
Figura 16 - Processo de criação de um novo Vírus	30
Figura 17 - Infecção de um cromossomo	31
Figura 18 - Processo de criação de um Vírus e infecção de um cromossomo	31
Figura 19 - Função F1 - <i>Shifted Sphere Function</i>	37
Figura 20 - Função F2 - <i>Shifted Schwefel's Problem</i>	38
Figura 21 - Função F3 - <i>Shifted Rotated High Conditioned Elliptic Function</i>	38
Figura 22 - Função F5 - <i>Schwefel's Problem 2.6 with Global Optimum on Bounds</i>	39
Figura 23 - Função F1 - Melhor Fitness.....	47
Figura 24 - Função F1 - Média do Desvio Padrão.....	47
Figura 25 - Função F1 – Fitness Médio.....	47
Figura 26 - <i>Fitness</i> Durante as Gerações do AG Sem Infecção para Função F1 (6 segundos).....	48
Figura 27 - <i>Fitness</i> Durante as Gerações do AG Mutação para Função F1 (6 segundos).	48
Figura 28 – <i>Fitness</i> Durante as Gerações do AGRI para Função F1 (14 segundos).....	49
Figura 29 - Função F2 - Melhor Fitness.....	51
Figura 30 - Função F2 - Média do Desvio Padrão.....	51
Figura 31 - Função F2 - Média de Fitness.....	51
Figura 32 - <i>Fitness</i> Durante as Gerações do AG Sem Infecção para Função F2 (6 segundos).....	52
Figura 33 - <i>Fitness</i> Durante as Gerações do AG Mutação para Função F2 (6 segundos).	52
Figura 34 - <i>Fitness</i> Durante as Gerações do AGRI para Função F2 (15 segundos).....	53
Figura 35 - Função F3 - Melhor Fitness.....	55
Figura 36 - Função F3 - Média do Desvio Padrão.....	55
Figura 37 - Função F3 - Média de Fitness.....	55
Figura 38 - <i>Fitness</i> Durante as Gerações do AG Sem Infecção para Função F3 (8 segundos).....	56

Figura 39 - Fitness Durante as Gerações do AG Mutação para Função F3 (8 segundos).	56
Figura 40 - Fitness Durante as Gerações do AGRI para Função F3 (15 segundos).....	57
Figura 41 - Função F5 - Melhor Fitness.	59
Figura 42 - Função F5 - Média do Desvio Padrão.	59
Figura 43 - Função F5 - Média de Fitness.....	59
Figura 44 - Fitness Durante as Gerações do AG Sem Infecção para Função F5 (6 segundos).....	60
Figura 45 - Fitness Durante as Gerações do AG Mutação para Função F5 (6 segundos).	60
Figura 46 - Fitness Durante as Gerações do AGRI para Função F5 (12 segundos).....	61

LISTA DE TABELAS

Tabela 1 - Resultados das Melhores Execuções de AG, AG com Alta Mutação e AGRI para as funções F1, F2, F3 e F5.....	40
Tabela 2 - Resultados de Execução Para Função F1	46
Tabela 3 - Resultados de Execução Para Função F2.....	50
Tabela 4 - Resultados de Execução Para Função F3.....	54
Tabela 5 - Resultados de Execução Para Função F5.....	58

LISTA DE ABREVIATURAS

AG – Algoritmo Genético

AGIV – Algoritmo Genético com Infecção Viral

AGRI – Algoritmo Genético Retroviral Iterativo

DNA – Deoxyribo Nucleic Acid

HIV – Vírus da imunodeficiência humana

RNA – Ribo Nucleic Acid

SIDA - Síndrome da Imunodeficiência Adquirida (AIDS)

PUBLICAÇÕES

ANAIS DE EVENTOS

SIMOES, Renato.; MONTEIRO, Glauber.; TEIXEIRA, Otávio.; SOARES, Átila.; LIMAO, Roberto Celio. Retroviral Iterative Genetic Algorithm for Real-Parameter Function Optimization Problems. In: Tong, Hengjian; Kang, Zhuo (Eds.). (Org.). Computational Intelligence and Intelligent Systems. 1 ed. : , 2010, v. 107, p. -.

SIMOES, Renato.; TEIXEIRA, Otávio.; LIMAO, Roberto Celio. Mixing Theory of Retroviruses and Genetic Algorithm to Build a New Nature-Inspired Meta-Heuristic for Real-Parameter Function Optimization Problems. In: Tenth International Conference on Hybrid Intelligent Systems - IEEE, 2010, Atlanta. Tenth International Conference on Hybrid Intelligent Systems, 2010.

REVISTA CIENTÍFICA

SIMOES, Renato.; TEIXEIRA, Otávio.; LIMAO, Roberto Celio. Mixing Theory of Retroviruses and Genetic Algorithm to Build a New Nature-Inspired Meta-Heuristic for Real-Parameter Function Optimization Problems. Polibits, México, p. 65 - 70, 01 dez. 2010.

RESUMO

Este trabalho tem como objetivo apresentar o desenvolvimento de uma meta-heurística híbrida baseada no ciclo de vida viral, mais especificamente dos Retrovírus, que fazem parte do grupo dos seres que evoluem mais rápido na natureza. Este algoritmo é denominado Algoritmo Genético Retroviral Iterativo (AGRI) e para embasamento computacional são utilizados conceitos de Algoritmo Genético (AG) e biológico características de replicação e evolução retroviral, o que proporciona uma grande diversidade genética o que aumenta a probabilidade para encontrar a solução, fato este confirmado através de melhores resultados obtidos pelo AGRI em relação ao AG.

ABSTRACT

This work presents the development of a hybrid meta-heuristic based on the viral life cycle, specifically from Retroviruses, which are part of nature's swiftest forms. This algorithm is called Retroviral Iterative Genetic Algorithm (RIGA) and uses as computational basement Genetic Algorithm (GA) and biological basement retroviral replication characteristics, which provides a great diversity increasing the probability to find the solution, what is confirmed by better results obtained by AGRI than AG.

CAPÍTULO I

1. INTRODUÇÃO

Este capítulo apresenta a motivação que levou ao desenvolvimento do trabalho, quais são os objetivos gerais e específicos, a metodologia aplicada bem como a estrutura geral da dissertação.

1.1. PREÂMBULO

Ao longo dos anos, os Vírus vêm sendo tratados como vilões na destruição de estruturas orgânicas, gerando, inclusive, o desaparecimento de espécies, pandemias como a Gripe Espanhola e epidemias longas e fatais como o Vírus da Imunodeficiência Humana (HIV). Apesar de não serem seres vivos, são muito eficientes para se perpetuarem. Os Retrovírus são os seres que mais evoluem, até quando são comparados com os seres vivos (CARTER et al., 2007). Esta característica da evolução retroviral pode ser útil quando aplicada em alguma estrutura computacional, principalmente na computação evolucionária.

1.2. OBJETIVOS GERAIS

O trabalho proposto irá abordar o desenvolvimento de um novo algoritmo genético inspirado em estruturas de Vírus da família dos *Retroviridae* (os Retrovírus), este algoritmo leva o nome de AGRI (Algoritmo Genético Retroviral Iterativo). A derivação do nome vem da junção de suas funcionalidades: **Algoritmo Genético** por possuir todo o comportamento de um AG, **Retroviral** por ser inspirado em estruturas dos Retrovírus e **Iterativo** por ocorrer a cada geração.

A ideia geral é propor uma estrutura baseada em AG que possua características presentes nos Vírus, ou seja, que possibilite uma evolução mais rápida que a usual do AG, já que os Retrovírus possuem como características principais o compartilhamento de material genético entre os membros da população e alta taxa de mutação

1.3. OBJETIVOS ESPECÍFICOS

Para alcançar os objetivos gerais este trabalho foi dividido nos seguintes objetivos específicos:

- Revisar a fundamentação teórica sobre Vírus, de forma a compreender seus princípios básicos, com ênfase aos Retrovírus;
- Revisar a justificativa teórica sobre Algoritmo Genético, a fim de abranger seus principais conceitos;
- Desenvolver o Algoritmo Genético Retroviral Iterativo (AGRI)
- Comparar os resultados obtidos pelo AGRI com os do AG

1.4. ESTRUTURA DA DISSERTAÇÃO

No Capítulo II será apresentada uma revisão biológica através do estudo dos Vírus, abordando temas como evolução, reprodução, bem como um estudo detalhado sobre os Retrovírus.

O Capítulo III trata a respeito da Computação Evolucionária, que está sendo representada pelos Algoritmos Genéticos e Algoritmo Genético Viral.

No Capítulo IV será apresentado o Algoritmo Genético Retroviral Iterativo (AGRI) abordando tópicos como inspirações, elaboração, componentes e comportamento.

No Capítulo V serão apresentados os resultados dos testes comparativos com Algoritmos Genéticos.

E por fim no Capítulo VI serão apresentadas as considerações finais sobre o trabalho e os trabalhos futuros.

CAPÍTULO II

2. EMBASAMENTO BIOLÓGICO: DOS VÍRUS AOS RETROVÍRUS

2.1. VÍRUS

2.1.1. Introdução

“RNA viruses deserve their reputation as Nature’s swiftest evolvers”

Michael Worobey and Edward Holmes

Vírus são parasitas intracelulares, inertes, sem vida e sem mobilidade. São conhecidos por se diferenciarem dos organismos mais simples (bactérias) pelos seguintes motivos (HOGG, 2005):

- Não podem ser observados por microscópio de luz
- Não possuem estrutura celular interna
- Podem conter DNA ou RNA, mas não ambos
- São incapazes de se reproduzir a menos que ocupem uma célula hospedeira viva
- Não tem capacidade metabólica

Os vírus infectam dos procariontes (bactérias e arqueobactérias) aos eucariontes (animais vertebrados, animais invertebrados, plantas e fungos). Os que infectam as células procariontes são frequentemente referenciados como bacteriófagos. A presença de um Vírus no hospedeiro é percebida através de sinais de doenças. Porém, muitos organismos saudáveis, são hospedeiros de infecções virais não-patogênicas, dos quais alguns são ativos e outros em estado de inércia (CARTER et al., 2007).

Partículas virais ou *vírion* tem um tamanho muito pequeno comparado a estruturas como bactérias. O Vírus da herpes que é razoavelmente grande, é dez milhões de vezes menor que um balão, enquanto um balão é 10 milhões de vezes menor que a terra pelo mesmo fator, como pode ser visto na Figura 1.

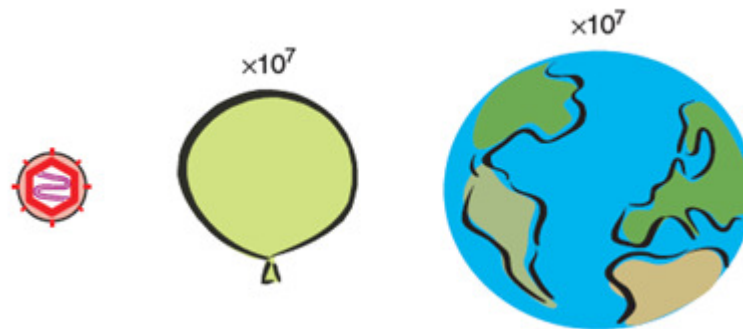


Figura 1 - Comparativo de partícula de herpesvírus, um balão e a terra.

Fonte: HOGG (2005)

Os vírions da maioria das viroses são muito pequenos para serem vistos com microscópio de luz e podem apenas ser vistos com microscópio eletrônico Figura 2.

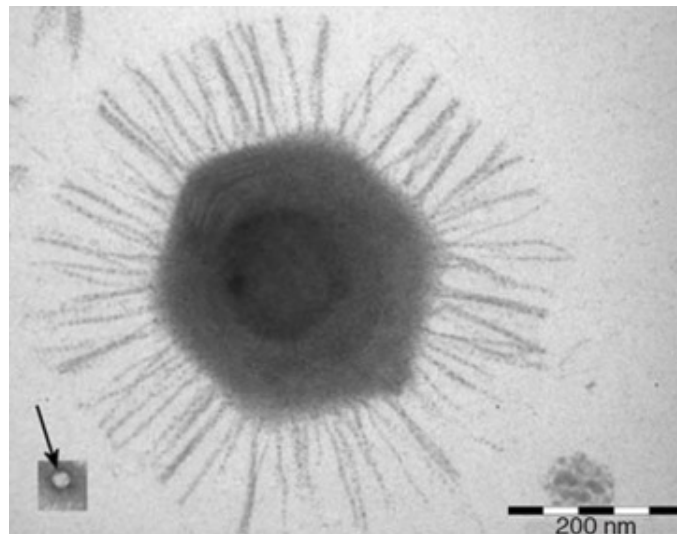


Figura 2 - Mimivírus, um dos maiores Vírus e o parvovírus (apontado), um dos menores Vírus.

Fonte: HOGG (2005)

Existem basicamente dois componentes físicos que formam todos os Vírus. Esses componentes são: material genético (núcleo composto de DNA ou RNA) e o revestimento protéico (capsídeo) que fornece uma camada protetora para o Vírus e reconhece o hospedeiro correto a ser atacado. Além disso, alguns Vírus podem

apresentar uma camada extra formada de lipídio e proteínas membranares (GERARDI et al., 2005).

Quando o material genético do Vírus é introduzido na célula hospedeira, o material genético toma o controle do mecanismo de reprodução da célula e faz com que a mesma produza Vírus, não células (GERARDI et al., 2005).

A maior habilidade dos Vírus é a de mutação rápida. Essa mutação fornece aos Vírus a proteção contra ambientes não propícios e vacinas. A mutação mais freqüente nos Vírus está na mudança do revestimento protéico, desenvolvendo resistência a vacinas e drogas antivirais (GERARDI et al., 2005).

2.1.2. Importância do Estudo dos Vírus

Vírus são agentes importantes em muitas doenças humanas, desde simples resfriados até letais como o Vírus causador da raiva (Rhabdovírus). Alguns tipos de tipos de câncer são atribuídos também a infecções virais (CARTER et al., 2007).

Os vírus podem impor tanto sofrimento individual quanto afetar toda uma sociedade. Varíola teve um impacto grande no passado e o Vírus da AIDS está tendo impacto nos dias de hoje.

Existem, portanto, dois quesitos importantes para entender a natureza dos Vírus: como eles se replicam e como eles causam doenças. Este conhecimento permite o desenvolvimento de mecanismos efetivos de prevenção, diagnóstico e tratamento de doenças virais através da produção de vacinas, reagentes e técnicas e drogas antivirais. Dessa forma, a aplicação médica constitui o maior interessado na virologia.

2.1.2.1. Alguns Vírus são Úteis

Alguns Vírus são estudados por terem aplicações úteis e serem potencial fonte de informação para outras aplicações. Seguem abaixo algumas de suas aplicações (CARTER et al., 2007):

- Fonte de enzimas: um número de enzimas usadas na biologia molecular são enzimas virais. Exemplos incluem transcriptase reversa de Retrovírus e RNA polimerase de fagos;

- Pesticidas: algumas pragas de insetos são controladas com baculovírus e mixoma Vírus tem sido utilizado para controlar coelhos;
- Agente anti-bacterial: em meados do século 20 fagos eram usados no tratamento de algumas doenças humanas oriundas de infecções bacterianas. Com a descoberta do antibiótico, esta prática acabou se tornando menos utilizada, porém, está sendo renovada pelo surgimento de bactérias resistentes a antibióticos;
- Agente anticancerígeno: alguns tipos de Vírus geneticamente modificados, como o da herpes simples, são investigados no tratamento de câncer. Estes Vírus que são modificados estão aptos para infectar e destruir células específicas do tumor, porém, não aptos a infectar células normais;
- Vetor de genes para produção de proteína: alguns Vírus como baculovírus e adenovírus são usados como vetores para transportar genes em células animais com crescimento em cultura. Esta tecnologia é usada para inserir dentro dos genes das células códigos úteis de proteínas como componentes de vacina e estas células podem ser usadas para produção em massa de proteínas;
- Vetor de genes para tratamento de doenças genéticas: crianças com variadas formas de imunodeficiência (síndrome da bolha) têm sido tratadas com sucesso usando Retrovírus como vetor para introduzir em suas células-tronco uma cópia não-mutável de genes modificados responsáveis pela doença.

2.1.3. Estrutura

O *vírion*¹ contém o genoma dos Vírus. Considerando que os genomas das células são compostos de DNA de cadeia dupla, existem quatro possibilidades principais para um genoma viral (CARTER et al., 2007):

- DNA de cadeia dupla (dsDNA)
- DNA de cadeia simples (ssDNA)
- RNA de cadeia dupla (dsRNA)
- RNA de cadeia simples (ssRNA)

Na Figura 3 são exibidas as principais estruturas virais.



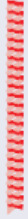
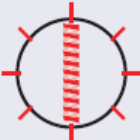
Main types of virion structure		Genomes			
		dsDNA	ssDNA	dsRNA	ssRNA
Icosahedral, naked		✓	✓	✓	✓
Icosahedral, enveloped		✓		✓	✓
Helical, naked		✓	✓		✓
Helical, enveloped					✓

Figura 3 - Estruturas Virais

Fonte: CARTER et al. (2007)

Os Vírus possuem uma estrutura muito simples, mesmo quando comparado aos mais primitivos organismos celulares. Um *vírion* ou partícula viral tem na sua essência

¹ A denominação *vírion* se dá quando o Vírus está fora da célula.

apenas dois componentes: um núcleo de ácido nucléico que é cercado e protegido por uma capa protéica ou capsídeo, a combinação dos dois é conhecida como nucleocapsídeo. Em alguns tipos de Vírus, o nucleocapsídeo é cercado por outra membrana chamada de envelope, parcialmente derivada de material celular de algum hospedeiro, a Figura 4 apresenta uma estrutura viral simples (GERARDI et al., 2005).

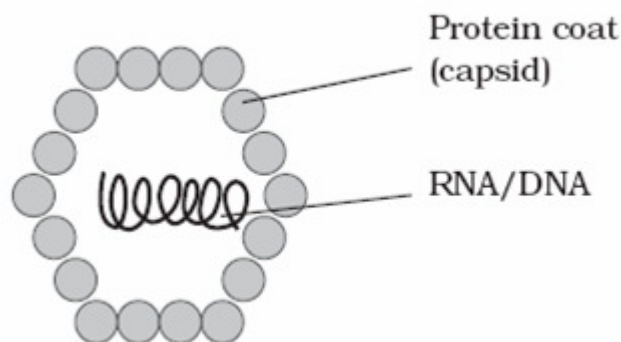


Figura 4 - Representação da estrutura viral.

Fonte: HOGG (2005)

Geralmente o genoma viral é muito menor do que genoma celular o que traz à tona a seguinte questão: como os Vírus conseguem guardar suas necessidades num genoma tão pequeno. De acordo com CARTER et al. (2007) O Vírus consegue isto de uma série de formas:

- Vírus usam proteínas de células hospedeiras: o genoma dos Vírus mais largos duplica algumas das funções celulares, mas o dos Vírus pequenos confia fortemente nas funções da célula hospedeira. Possui, entretanto, o que Vírus de RNA deve fazer, não importa o tamanho do genoma, a função é a RNA polimerase, porque as células não codificam enzimas que replicam Vírus de RNA. Uma porção significativa do genoma de um Vírus de RNA é assumir os genes para a RNA polimerase;
- Muitas proteínas virais são multifuncionais: proteína viral pode ter muitas atividades enzimáticas;
- Vírus codificam com eficiência: pode ter sobreposição de genes e genes codificados inseridos dentro de outros genes, como é o caso do genoma da hepatite B.

2.1.4. Replicação

A multiplicação intracelular dos Vírus é um processo cíclico que proporciona a amplificação exponencial do número de Vírus e sua transmissão em série às células sensíveis a infecção (AGUT, 2009).

O objetivo principal do Vírus é replicar a si mesmo e para alcançar este objetivo ele precisa entrar em uma célula hospedeira, criar seus clones e colocar estas novas cópias fora da célula. O processo de replicação pode ser dividido em sete passos (CARTER et al., 2007):

1. Ataque do *vírion* a célula;
2. Entrada na célula;
3. Transcrição do gene viral no mensageiro RNA (mRNA);
4. Tradução de Vírus mRNA em proteínas virais;
5. Replicação do genoma;
6. Construção de proteínas e genomas no Vírus;
7. Saída do Vírus da célula hospedeira.

Quando se tenta entender os vários tipos de replicações virais, esses passos são como um modelo muito utilizado, porém, nem todos são relevantes para todos os tipos de Vírus. Alguns passos não acontecem nesta ordem e alguns Vírus tem procedimentos adicionais. A Figura 5 exibe a replicação de um Vírus bacteriófago.

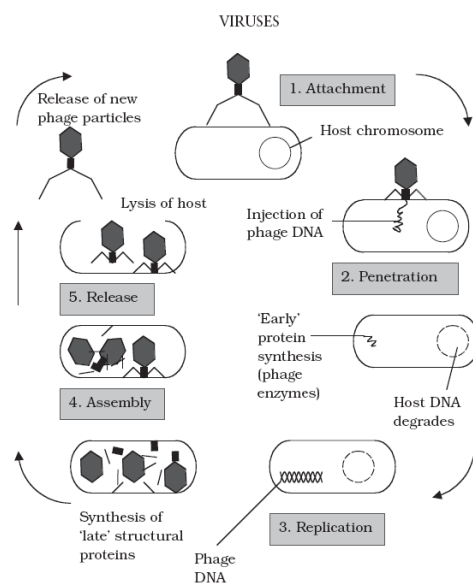


Figura 5 - Ataque de um Vírus bacteriófago
Fonte: HOGG (2005)

2.1.5. Classificação dos Vírus

Os Vírus não são considerados estritamente seres vivos. Assim, não se pode levar em consideração o mesmo princípio de classificação de animais e plantas. Da mesma forma que existe nos verdadeiros organismos, os Vírus possuem espécie, gênero, família e ordem, mas nenhum dos altos grupos está presente (classe, filo e reino). Binômios latinos não são usados por convenção biológica taxonômica, entretanto, a proposta de utilizar binômio não-latinizado tem sido avaliada (HOGG, 2005).

Os fatores levados em consideração para classificação viral são:

- Faixa de hospedeiro: (vertebrado/invertebrado, planta, alga/fungo, bactéria);
- Morfologia: simetria do capsídeo, envelopado ou não-envelopado, número de capsômeros (unidades formadoras de capsídeos);
- Genoma: tipo/modo de replicação.

De acordo com a Comissão Internacional de Taxonomia dos Vírus, atualmente existem três ordens, setenta e três famílias, duzentos e oitenta e sete gêneros e mais de mil e novecentas espécies de Vírus.

Os Vírus são divididos entre os Vírus de DNA e RNA. Os Vírus de DNA utilizam DNA como material genético e se multiplicam através de uma enzima chamada DNA polimerase. Já os Vírus de RNA utilizam RNA como material genético ou no processo de replicação, estes, por sua vez, tem altas taxas de mutação, diferentes dos Vírus de DNA.

De acordo com a classificação de Baltimore (CARTER et al., 2007) os Vírus podem ser divididos nos seguintes grupos (ss – cadeia simples / ds – cadeia dupla):

- Grupo I: dsDNA - (Herpesvírus);
- Grupo II: ssDNA - (Parvovírus);
- Grupo III: dsRNA - (Reovírus);
- Grupo IV: (+)ssRNA sentido positivo RNA - (Picornavírus);
- Grupo V: (-)ssRNA sentido negativo RNA - (Rhabdovírus);
- Grupo VI: ssRNA-RT (*transcrição reversa*) sentido (+) - RNA com DNA intermediário (Retrovírus);

- Grupo VII: dsDNA-RT (*transcrição reversa*) - (Hepadnavírus).

2.2. RETROVÍRUS

2.2.1. Introdução

Os Retrovírus são causadores das mais importantes patologias humanas (SIDA – Síndrome da Imunodeficiência Adquirida), possuem internamente tanto RNA quanto DNA em diferentes partes do seu ciclo reprodutivo (HOGG, 2005) e tem habilidade de penetrar outras células e difundir seu próprio material genético nos genes de seus hospedeiros, eventualmente assumindo todo o controle de suas funções reprodutivas. (LERNER, 2003).

Algumas características da família Retroviridae segundo AGUT (2009) são:

- Injetam seu material genético na célula num processo chamado de retrotranscrição. Com a utilização da enzima transcriptase reversa, ele cria um DNA a partir do seu RNA, fazendo com que este DNA se junte ao DNA da célula hospedeira;
- Não podem ser combatidos por sofrerem mutação constante, pois o processo de **retrotranscrição** cria DNAs com seqüências diferentes das quais deveriam ser. O HIV é um exemplo clássico dessa família;
- O surgimento de inúmeros mutantes é comum, pois as enzimas que catalisam a replicação de genomas com RNA acumulam muitos erros de cópias;

Até a descoberta destes Vírus, se tinha o dogma de que a transferência da informação genética sempre ocorria em direção do DNA ao RNA, dessa forma, encontrar alguns Vírus que realizam a transcrição reversa (utiliza a enzima transcriptase reversa, descoberta em 1970 (HOGG, 2005)) causou certa revolução (CARTER et al., 2007). Sabe-se hoje em dia que a transcrição reversa é realizada não somente por Vírus de RNA, mas também por alguns Vírus de DNA e por células não infectadas.

2.2.2. Replicação

A maioria dos Retrovírus funde sua membrana com a membrana do plasma da célula hospedeira, embora alguns sofram endocitose e unam suas membranas com uma membrana endossômica. A estrutura que é liberada no citoplasma perde algumas proteínas e o complexo de transcrição reversa é formado Figura 6. (CARTER et al., 2007).

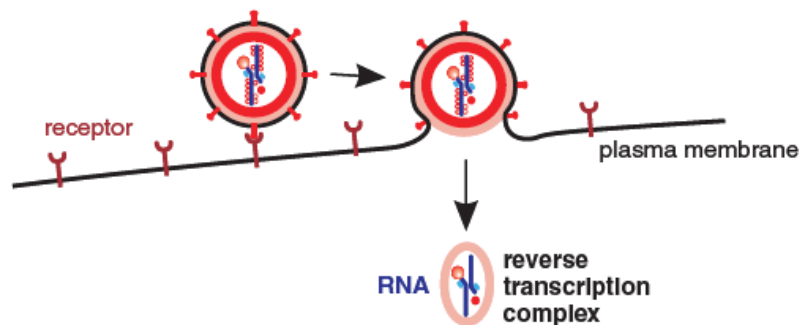


Figura 6 - Formação do complexo de transcrição reversa.

Fonte: (CARTER et al., 2007)

Alguns tipos de Vírus de RNA replicam seus genomas via DNA intermediário, enquanto alguns Vírus de DNA replicam seu genoma via RNA intermediário. Os dois tipos de replicação de genoma envolvem transcrição reversa, que tem dois grandes passos: síntese de (-) DNA a partir de um modelo de (+) RNA seguido de uma síntese de uma segunda cadeia de DNA, os dois passos são feitos a partir transcriptase reversa, que é codificada por um Vírus (CARTER et al., 2007).

A transcrição reversa ocorre no citoplasma da célula infectada. Durante a síntese das duas cadeias de DNA, cada uma se destaca de sua cadeia e se anexa no final de outra cadeia através do pareamento. O DNA resultante da transcrição reversa é chamado *provírus* e é maior do que o genoma do RNA.

Uma das proteínas do Vírus ainda associada ao *provírus* é a *integrase*. Esta enzima corta o DNA do cromossomo celular e insere o *provírus* no lugar. Nesse momento é que o *provírus* é integrado aos cromossomos da célula hospedeira (Figura 7). Se durante este processo a célula se dividir, o provírus é replicado para as células filhas (CARTER et al., 2007).

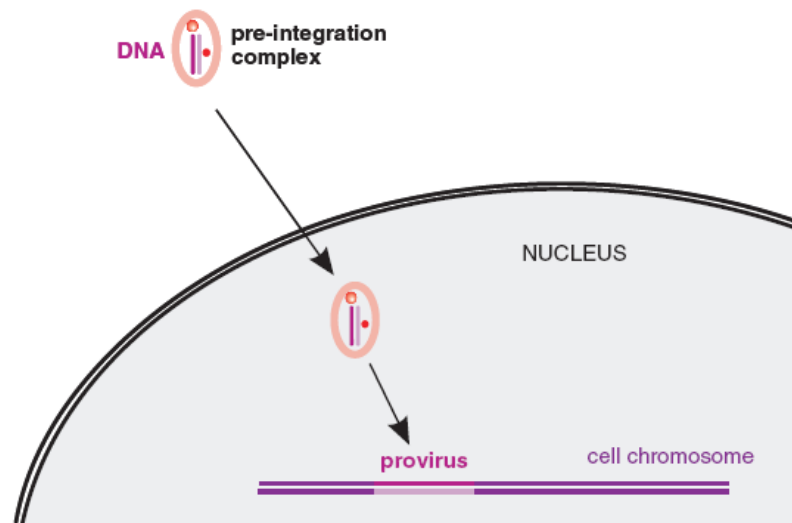


Figura 7 - Criação do provírus e efetivação da transcrição reversa.
Fonte: (CARTER et al., 2007)

Após a transcrição reversa, o ciclo normal do Vírus continua: transcrição, tradução, replicação do genoma, construção do Vírus e saída da célula como pode ser visto na Figura 8.

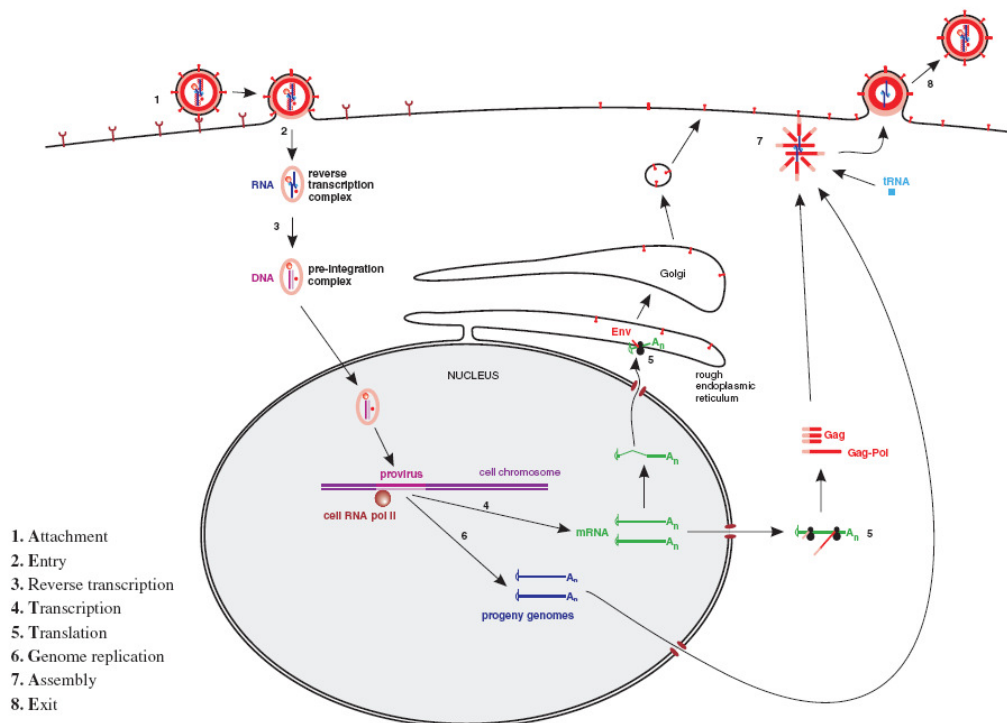


Figura 8 - Ciclo de infecção viral.
Fonte: (CARTER et al., 2007)

Enquanto a transcrição² ocorre simplesmente de DNA para RNA, a enzima transcriptase reversa transcreve RNA em DNA. O termo “retro” em Retrovírus se refere a esta reversão. Este não é um processo seguro, pois, cria DNAs de seqüências muito distintas, causando a variabilidade genética viral (CARTER et al., 2007) (CHENG et al., 2008). Na Figura 9 são apresentadas as relações entre mutações virais onde pode ser notado que as replicações em Vírus de RNA são superiores aos de DNA.

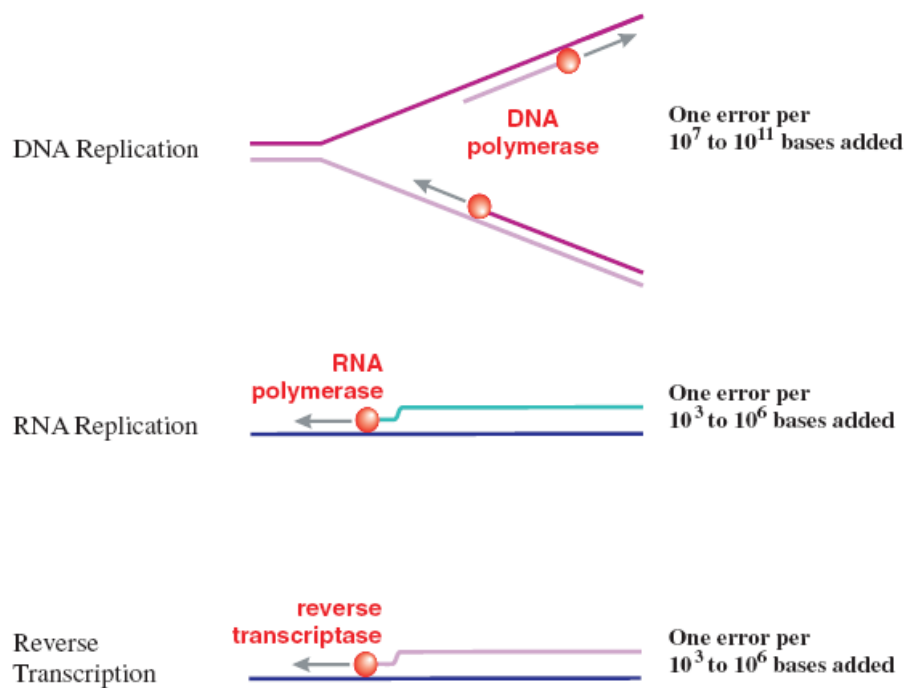


Figura 9 - Taxa de mutação nos processos de replicação viral.

Fonte: (CARTER et al., 2007)

² No processo de transcrição o RNA é gerado a partir do DNA e colocado no invólucro gerando outros Vírus

2.3. EVOLUÇÃO VIRAL

“A Seleção Natural atua somente no sentido de preservar as modificações benéficas; cada nova forma tende, em uma determinada região já povoada pelos indivíduos da sua espécie, a tomar o lugar de seus ancestrais menos aperfeiçoados, ou dos seres menos favorecidos com os quais tiver de competir, podendo até eliminá-los. Portanto, conforme vimos, a extinção e a Seleção Natural andam juntas.”

Charles Darwin

2.3.1. Introdução

Os conceitos de Charles Darwin sobre reprodução e sobrevivência dos mais adaptados aplicados a organismos celulares também são aplicados aos Vírus. Como um organismo celular, o Vírus possui genes que se esforçam para perpetuar a espécie. A combinação genética que constitui o genoma viral permite ao Vírus se replicar. (CARTER et al., 2007)

Para Carter et al. (2007), mudanças individuais e novas combinações genéticas continuamente criam novos genótipos, muitos dos quais são menos suscetíveis que o genótipo do pai e não sobrevivem. Porém, às vezes, um novo genótipo é mais suscetível do que o genótipo que o gerou e o supera. Fazendo assim com que alguns dos novos genótipos virais se tornem capazes de infectar novas espécies hospedeiras.

2.3.2. Mecanismos de Evolução

O mecanismo de evolução viral é semelhante ao de evolução de organismos celulares: geração de genomas variados. Alguns desses genomas são fracos e não sobrevivem, porém um grupo tem vantagem em um nicho particular. Para um Vírus, um nicho pode ser uma nova espécie hospedeira ou a presença de droga antiviral e a variante pode se proliferar neste nicho como um novo grupo de Vírus. As variantes

virais são decorrentes de mutação, recombinação, rearranjo e aquisição de genes celulares (CARTER ET AL., 2005).

2.3.2.1. Mutação

“A seleção natural nunca produzirá em um ser alguma conformação que lhe seja prejudicial, uma vez que sua ação só se volta em prol e benefícios de cada criatura”

Charles Darwin

De acordo com FERREIRA (1986), mutação é a variação hereditária, súbita e espontânea, em um indivíduo geneticamente puro. As mutações são consideradas base da aparição biogenética de novas raças e espécies. De acordo com HOGG (2005), qualquer alteração na sequência de DNA de um organismo é chamada mutação, o que pode ou não ter efeito no fenótipo deste organismo.

Quando um genoma de um Vírus de DNA é replicado, a taxa de erro é bem menor do que quando um genoma baseado em RNA sofre replicação. O motivo é que a polimerase do DNA tem um mecanismo de garantia que corrige grande parte dos erros de replicação, o que a polimerase do RNA e a transcriptase reversa não possuem, fazendo com que os Vírus de RNA evoluam de forma mais rápida do que Vírus de DNA.

A alta taxa de erro durante a replicação do RNA e a transcrição reversa faz com que o Vírus de RNA não tenha uma sequência fixa na base do genoma. Ao invés disso, o genoma viral existe com muitas variações o que pode gerar uma gama de Vírus capazes de se adaptarem a novos nichos.

Dessa forma o processo de mutação nos Vírus segue o mesmo padrão evolutivo que as espécies orgânicas.

2.3.2.2. Recombinação

A recombinação é um processo que resulta na produção de um novo genoma baseado em dois outros genomas. Para um Vírus, essa recombinação pode acontecer quando a célula é infectada por dois outros Vírus da mesma espécie e a partir dessa

combinação um novo Vírus é produzido. De acordo com CARTER et al., (2005) isto pode acontecer tanto em Vírus de DNA quanto em Vírus de RNA. Na Figura 10 é apresentado o processo de recombinação viral.

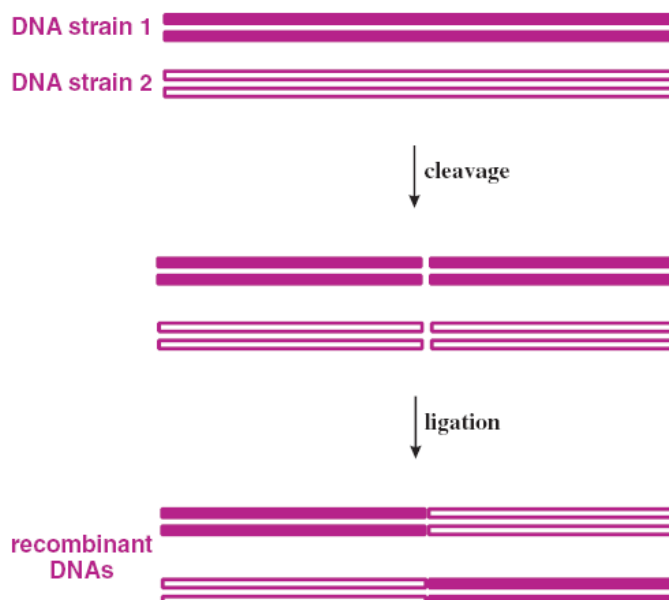


Figura 10 - Processo de recombinação viral.
Fonte: (CARTER et al., 2007)

2.3.2.3. Rearranjo

Rearranjo é um tipo de recombinação de material genético que pode ocorrer com Vírus de genoma segmentado caso suas partes estejam empacotadas no *vírion*. Se a célula for co-infectada por dois Vírus da mesma espécie, o *vírion* gerado pode conter genoma dos dois Vírus. Estes vírions são conhecidos como rearranjados. (CARTER et al., 2005)

Este tipo de mutação é comum em Vírus influenza, algumas pandemias como a suína de 2009 e a aviária 1957 e 1968 são frutos do rearranjo. O H1N1 passou a infectar além de porcos, homens (MACKENZIE, 2009) e (CARTER et al., 2005).

2.3.2.4. Aquisição de Genes Celulares

A similaridade entre algumas proteínas virais e celulares sugerem que alguns Vírus tenham adquiridos genes celulares. O genoma de alguns Retrovírus possui oncogenes, genes que causam tumores, que provavelmente são adquiridos de seus hospedeiros, os quais contem genes parecidos (proto-oncogenes, genes responsáveis pela divisão celular) (CARTER et al., 2005).

Existem duas teorias que levam à aquisição de genes celulares (CARTER et al., 2005):

- **Recombinação:** para que isto ocorra o genoma do Vírus e das células devem estar bem próximos, isto pode acontecer quando o DNA viral é replicado no núcleo da célula.
- **Síntese de uma cópia do DNA de um RNA mensageiro,** seguido de inserção do DNA no genoma viral. A síntese do DNA pode necessitar de uma transcriptase reversa, talvez fornecida pela célula ou um Retrovírus co-infectante.

CAPÍTULO III

3. COMPUTAÇÃO EVOLUCIONÁRIA

“A teoria da seleção natural se baseia na crença de que toda variedade nova, e em última análise, cada espécie nova, se produza e sobreviva por meio de alguma vantagem em relação aos seus concorrentes, levando quase inevitavelmente à conseqüente extinção das formas menos favorecidas”

Charles Darwin

3.1. ALGORITMO GENÉTICO

3.1.1. Introdução

Na natureza, indivíduos competem entre si para obtenção de recursos para sobrevivência, o que inclui basicamente a disputa por alimentos e abrigo. Há, além disso, a incessante busca pela perpetuação da espécie, passando para as futuras gerações suas características. A ideia formulada por Charles Darwin em 1872 que diz “A Seleção Natural tende apenas a tornar cada ser vivo tão perfeito ou um pouco mais perfeito que os demais habitantes da mesma região com os quais ele terá de lutar para sobreviver“, é um ponto chave da teoria da evolução. Se um indivíduo consegue ter sucesso ao sobreviver em relação a outros, ou seja, ter maior aptidão para sobreviver no ambiente ao qual esta inserido, este indivíduo apresenta a maior probabilidade de ter filhos e para estes filhos repassar suas características genéticas. É natural que com o passar das gerações, as aptidões que obtiveram mais sucesso de sobrevivência e reprodução vão sendo repassadas. E isto é a essência da Teoria da Evolução (HAUPT et al., 1998).

Conforme Carrol (2005), os arquitetos da síntese moderna uniram disciplinas evolutivas afirmando que os mecanismos que agem no nível individual em populações e espécies são suficientes para gerar as grandes diferenças que surgem ao longo do tempo geológico, o que significa que a evolução individual é de suma importância para a população. Dessa forma, Carrol (2005) conclui: “O que vale para a espécie vale para o reino”.

Tendo como inspiração os princípios básicos da Teoria da Evolução Natural de Darwin, Holland em 1975, publicou o livro "A adaptação em sistemas naturais e artificiais", onde explorou como aplicar os princípios da Teoria da Evolução voltados para problemas de otimização. Holland, analisando o fenômeno de adaptação natural, desenvolveu (junto com estudantes da Universidade de Michigan) os Algoritmos Genéticos (AGs), estruturas algorítmicas baseadas na teoria da evolução utilizadas para problemas de otimização computacional (MITCHELL et al., 1999).

Antes de Holland, outro cientista já havia formulado uma ideia baseada na evolução, em 1960, Rechenberg (1965, 1973) introduziu o termo "estratégia evolutiva" (Evolutionstrategie título original), método utilizado para otimizar funções de valor real. A técnica do uso evolução vai além dos Algoritmos Genéticos, como é o caso da Programação Evolucionária.

Holland (1975) analisou estruturas como mutação (modificação genética) e reprodução (recombinação genética ou *crossover*) para aplicá-las. O cruzamento de soluções é uma forma de aumentar o espaço de solução do algoritmo, até mesmo encontrar a melhor solução. A recombinação de material genético é um operador importante no processo de evolução natural, pois permite a produção de um novo genótipo (filho) a partir de outros dois (pais). Levando em consideração os aspectos biológicos, a forma mais comum de recombinação é feita a partir de dois cromossomos pais que são separados em um ponto específico e tem as suas metades unidas em um cromossomo filho.

Com os aspectos supracitados, foi identificado que os AGs nada mais são que uma abstração da Teoria da Evolução e Genética, os tornando métodos de otimização de busca num espaço de solução (HAUPT et al., 1998). Por serem inspirados em estruturas e comportamentos biológicos (BEASLEY et al., 1993) e HAUPT et al. (1998) definem estes algoritmos fazem parte do ramo da Computação Evolucionária, que, segundo Eberhart et al. (1996) consiste na otimização do aprendizado de máquina e em paradigmas de classificação, baseados nos mecanismos de evolução, tais como: Genética e Seleção Natural.

De acordo com Linden (2006), algoritmos genéticos se enquadram em técnicas probabilísticas, podendo encontrar diversas soluções em execuções diferentes mesmo com os parâmetros das execuções, incluindo a população, sejam os mesmos.

Algumas das principais vantagens dos AGs segundo (HAUPT et al., 1998) são:

- Otimização de valores discretos e contínuos;
- Buscas simultâneas num amplo espaço amostral;
- Possibilidade de se trabalhar com várias variáveis;
- Fornece um conjunto de variáveis ótimas, não apenas uma solução.

3.1.2. Componentes

3.1.2.1. Cromossomo

Um cromossomo para o Algoritmo Genético tem a forma de uma cadeia de *bits*. Cada loco no cromossomo pode ter n possíveis alelos dependendo do alfabeto permitido para tal. No caso de um AG binário este alfabeto se restringe a 0's e 1's, mas há casos, como o do Caixeiro Viajante (JOHNSON, 1995), em que o alfabeto varia de A até F (no caso de seis cidades), sendo A a cidade origem e F a cidade destino. Dessa forma o alfabeto que expressa a possibilidade de alelos pode variar de acordo com o problema. A Figura 11 exibe a representação de um cromossomo.

0	1	1	1	0	1	0	0
---	---	---	---	---	---	---	---

Figura 11 - Representação de um cromossomo binário.

Cada cromossomo pode ser visto como um ponto de busca em um espaço de soluções candidatas, o AG processa uma população de cromossomos modificando alguns dos seus indivíduos. Um AG requer uma função associada a avaliação de cada cromossomo da população corrente, essa função calcula o nível de adaptação (*fitness*) do indivíduo ao ambiente. O *fitness* do cromossomo depende diretamente de como aquele cromossomo pode resolver o problema (MITCHELL, 1999).

3.1.2.2. Função de avaliação (*Aptidão*)

A função de avaliação da aptidão tem como objetivo indicar em um único valor a qualidade do cromossomo (solução). Esta função pode retornar tanto um número inteiro quanto real, dependendo do problema. Para que esta função seja calculada, o cromossomo precisa ser decodificado e a função objetivo avaliada. O resultado desta função pode indicar também a proximidade da solução da qual este cromossomo está.

3.1.2.3. População

A população representa um conjunto de cromossomos. No caso de um AG binário, ela pode ser representada como uma matriz $N_{população} \times M_{bits}$ (HAUPT, et al., 1998).

No AG é preciso se definir o tamanho e o método de inicialização da população. Em alguns casos é necessário utilizar a mesma população inicial para encontrar a melhor configuração de um AG, porém, na maioria das vezes se opta por gerar uma população inicial aleatória.

O tamanho da população influencia diretamente no desempenho do AG. Assim, quanto maior a população, maior será a cobertura do espaço de busca do problema.

3.1.2.4. Seleção

A seleção é o processo pelo qual o Algoritmo Genético vai selecionar os cromossomos com maiores valores de aptidão para compor as próximas gerações (elitismo) e também aqueles que irão gerar filhos. O propósito da seleção é enfatizar a aptidão individual dentro da população. Seleções muito rígidas podem acabar elevando o valor de aptidão de alguns indivíduos na população, reduzindo assim a diversidade necessária para mudança e progresso do AG. (MITCHELL, 1999)

Alguns métodos são bastante utilizados MITCHELL (1999):

- *Roleta (Rank Selection)*: os cromossomos são agrupados por ordem de valor de aptidão e estes valores são dispostos proporcionalmente numa roleta ponderada. Os cromossomos com maiores

aptidões têm maiores chances de serem selecionados, porém não se total certeza de que serão selecionados. A Figura 12 apresenta a seleção por roleta.

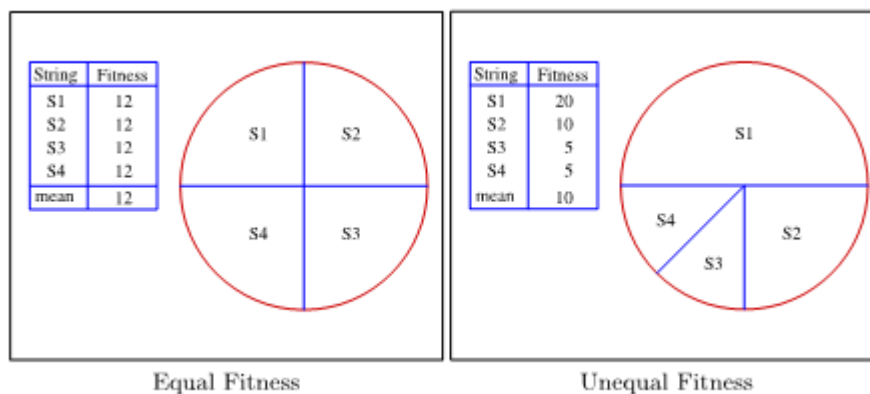


Figura 12 - Seleção Roleta (*Rank Selection*).

Fonte: COELHO et al., (2007)

- Torneio (*Tournament Selection*): dois indivíduos são escolhidos de forma aleatória na população, em seguida um número r entre 0 e 1 é escolhido, caso $r < k$ (onde k é um parâmetro por exemplo, 0.75), o indivíduo com maior *fitness* é selecionado para ser pai, senão o com menor *fitness* é selecionado. Os dois então retornam para a população original e podem ser selecionados novamente.

3.1.2.5. Operadores de Mutação e Recombinação

a. Recombinação (*Crossover*)

Esta é a principal funcionalidade do AG, pois, cria novas estruturas através da recombinação de outras já existentes. Um dos tipos de cruzamento mais simples é o de um ponto de corte. Funciona da seguinte maneira: um ponto de corte é escolhido de forma aleatória e as duas partes separadas de cada pai são trocadas gerando dois filhos com material genéticos de ambos os pais. Na Figura 13 o procedimento pode ser visualizado (COELHO et al., 2007).

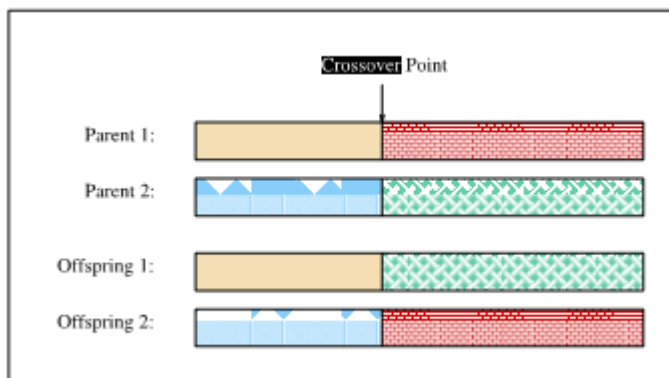


Figura 13 - Exemplo cruzamento com 1 ponto de corte.
Fonte: COELHO et al., (2007)

b. Mutação

Na natureza a modificação genética se dá principalmente através do cruzamento de informação genética dos pais. Porém, existem casos onde a modificação genética ocorre no próprio indivíduo. Então, é dito que ocorreu uma mutação na sua estrutura genética. Em AG a mutação se dá pela substituição de um valor de um alelo por algum valor presente no alfabeto. Na Figura 14 abaixo um exemplo de mutação para estrutura de um cromossomo binário (COELHO et al., 2007), onde a informação de um gene é trocada aleatoriamente por outra presente no alfabeto da solução.

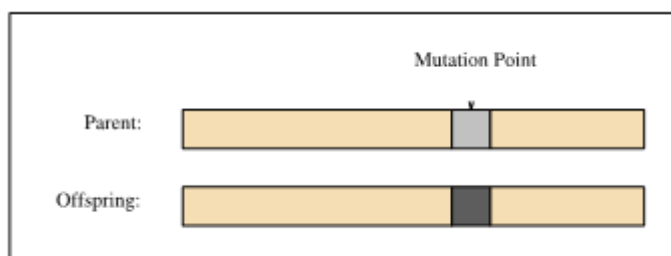


Figura 14 - Exemplo de mutação em cromossomo.
Fonte: COELHO et al., (2007)

3.1.3. Passos de um AG

O algoritmo genético segue uma série de passos que compreende desde a criação da população até a avaliação para verificar se a população possui alguma solução viável para a parada do algoritmo (LINDEN, 2006). Os passos que compõem um algoritmo são:

1. Inicializa população aleatória
2. Avalia indivíduos e verifica a solução
3. Efetua Seleção para Cruzamento
4. Reprodução
5. Aplica Mutação
6. Avalia os indivíduos e verifica a solução, se solução não for encontrada voltar ao passo 3

Existem duas formas de parar a execução de um AG. A primeira é encontrar a solução do problema. Isto só é possível quando se tem um valor de comparação de aptidão, o que pode compreender apenas um valor ou um intervalo de valores, dependendo do problema (COELHO et al. 2007). Outra forma é o número de geração. Para o AG uma geração representa um ciclo de execução onde, dentro dela, ocorrem seleção, cruzamento, mutação e avaliação. Além disso, nunca é criada uma população nova por geração, já que a ideia é a evolução das soluções pelos seus descendentes.

3.2. ALGORITMO GENÉTICO COM INFECÇÃO VIRAL

Segundo Guedes et al (2005) o Algoritmo Genético com Infecção Viral (AGIV) é uma variante do AG que utiliza o operador de infecção no lugar do operador de mutação. O método consiste em manter uma população extra de Vírus, denominada população viral, que é composta por Vírus (estruturas que contém trechos de soluções). No caso do caixeiro viajante pra uma rota {A,F,G,E,G,T,A} um trecho viral seria {F,G,E}.

Durante o processo de infecção, um Vírus é selecionado pelo mesmo processo de *rank*. Porém, como o Vírus não possui valor de aptidão, já que não tem uma solução completa, ele apresenta um valor chamado de nível de infectabilidade. Este é um indicador de quão bem o Vírus tem atuado.

Após a seleção do Vírus que efetuará a infecção (em um determinado cromossomo), é realizada a transcrição, que consiste na modificação do cromossomo infectado pelo Vírus de forma que esse contenha um trecho idêntico ao representado pelo agente infectante (GUEDES et al., 2005).

Quando um Vírus é criado ou modificado, sua infectabilidade é configurada em um valor inicial fixo, caso o Vírus consiga infectar um cromossomo (aumentar sua aptidão) o seu nível de infectabilidade é incrementado de 1 (um), caso contrário (se diminuir o *fitness*) seu valor é diminuído de 1 (um). Quando o Vírus chega ao valor de infectabilidade 0 (zero) ele descarta suas partes e copia uma parte do cromossomo para si (num processo chamado de transdução).

Dessa forma, à medida que a população de indivíduos completos melhora, a população de Vírus tende a ficar com sub-rotas melhores, ou seja, trechos de soluções que são mais desejáveis em uma rota completa (GUEDES et al., 2005).

CAPÍTULO IV

4. MATERIAL E MÉTODOS

4.1. ALGORITMO GENÉTICO RETROVIRAL ITERATIVO

“Pode dizer-se, metaforicamente, que a seleção natural procura, a cada instante e em todo o mundo, as variações mais ligeiras; repele as que são nocivas, conserva e acumula as que são úteis; trabalha em silêncio, insensivelmente, por toda a parte e sempre, desde que a ocasião se apresente para melhorar todos os seres organizados relativamente às suas condições de existência orgânicas e inorgânicas”.

Charles Darwin

4.1.1. Introdução

O Algoritmo Genético Retroviral Iterativo (AGRI) é baseado em estruturas virais (replicação e infecção) para desenvolver o modelo computacional. Esse nome deriva das suas funcionalidades:

- **Algoritmo Genético** - apresenta todas as características de um AG;
- **Retroviral** – referente às estruturas de Retrovírus;
- **Iterativo** - ocorre a cada ciclo de gerações.

A grande busca por diversidade populacional, para encontrar resultados de forma mais rápida, é um desafio. Desta forma, este algoritmo não é apenas uma ferramenta de mutação, mas sim de troca de material genético entre os cromossomos da população.

O motivo principal da utilização de estrutura viral no algoritmo é o fato destes Vírus estarem associados a uma fonte de inovação genética, o que é influenciado pela rápida velocidade de replicação e constante mutação (VILLARREAL, 2008).

Para inspiração biológica do AGRI foi escolhida a família *Retroviridae* que abrange os Retrovírus. Estes tipos de Vírus são baseados em RNA, que diferente aos de DNA, não possuem mecanismos de correção para desfazer possíveis mutações genéticas

ocorridas durante a multiplicação viral. Essa falta de mecanismo de controle faz com que os Vírus baseados em RNA se tornem estruturas com alto índice de mutação, sempre gerando novos indivíduos diferentes.

Os principais motivos para escolha da família *Retroviridae* para o trabalho proposto são:

- Por serem baseados em RNA merecem a reputação de seres que evoluem mais rápido na natureza;
- Tem a capacidade de misturar seu material genético com o hospedeiro num processo chamado retro transcrição ou transcrição reversa;
- Os Vírus dessa família sofrem constantes mutações, as mais altas da natureza, fazendo com que sua diversidade genética seja um empecilho para o seu combate, como o HIV;
- Estes Vírus só possuem RNA, o que representa para o AGRI parte de uma solução de um problema, e não um problema como um todo, como seria o caso do DNA.

De acordo com Haup et al. (1998) e Mitchell (1999), é importante manter, durante o tempo de processamento, a diversidade populacional do algoritmo genético. Neste trabalho, isto ocorre pela adição da característica de multiplicação retroviral através do recurso chamado de retrotranscrição ou transcrição reversa, onde o Vírus injeta o seu RNA no DNA de uma célula, fazendo com que a célula produza Vírus contendo ambos os materiais genéticos.

No AGRI a infecção e multiplicação viral são tratadas de formas diferentes da biológica, porém, com o mesmo princípio. Na natureza, o processo de multiplicação³ viral está vinculado à infecção. Todavia, no AGRI, a multiplicação viral e a infecção apresentam processos distintos, tornando o processo mais modular.

Dessa forma, a escolha do uso da família *Retroviridae* para o trabalho em questão é se basear na ideia geral que esta família fornece que é criar novos Vírus com material genético do hospedeiro, ter uma alta taxa de mutação e infectar um hospedeiro com parte do material genético de outros hospedeiros, difundindo assim, a partes de soluções por toda população.

³ O termo reprodução não é adequado, já que é uma característica pertinente a seres vivos

Vários pontos diferenciam o AGRI do AGIV, são eles:

- O AGRI não altera nenhum componente do AG, o AGIV retira a operação de mutação e a troca pela infecção;
- Os Vírus no AGIV tendem a ficar muito parecidos com os cromossomos, já que apenas absorve material genético do cromossomo, porém, não tem nenhuma estrutura de mutação para diferenciar os Vírus;
- No AGIV os piores Vírus têm seu material genético modificado, no AGRI eles são completamente substituídos, dessa forma a população viral é constantemente refeita, aumentando a possibilidade de infecção na população de cromossomos;
- No AGIV apenas os melhores seres da população são infectados, no AGRI há a possibilidade dos mais fracos serem também;
- O embasamento biológico do AGRI é bem específico para o uso de estrutura retrovirais.

4.1.2. O Comportamento Viral no AGRI

O AGRI possui quatro componentes principais ao tratar de estruturas virais, os quais são: estrutura viral, construção da população viral, infecção e o ciclo de vida do Vírus.

4.1.2.1. Vírus

O Vírus no AGRI é uma estrutura que apresenta o mesmo tamanho de um cromossomo, porém, possui alguns espaços sem informação. Isso ocorre porque a ideia central é compartilhar material genético e não ter outra população de cromossomos em paralelo. A quantidade de espaços em branco e as suas disposições são determinadas de forma aleatória. A Figura 15 apresenta alguns Vírus possíveis no AGRI.

(A)		1		0		1	0	
(B)			1		1			1
(C)	0			1		1		

Figura 15 - Vírus possíveis para um cromossomo binário de oito posições
(A) (B) (C) são exemplos de Vírus

4.1.2.2. População Viral e Criação de Novos Vírus

Para criação dos novos Vírus que irão compor a população viral o AGRI foi utilizado o processo natural comum nos Retrovírus chamado transcrição reversa. O processo consiste basicamente dos seguintes passos:

1. Criar um Vírus aleatório
2. Selecionar um cromossomo aleatório na população
3. Misturar os materiais genéticos em uma estrutura cromossômica mantendo as suas respectivas posições
4. Criar alguns espaços em branco para que seja feito o Vírus

Na Figura 16 é possível a visualização do processo de criação de um novo Vírus (este, biologicamente, possui a mesma função de um *provírus*).

(A)		1		1		1	0	
(B)	0	1	1	1	1	0	0	1
(C)	0	1	1	1	1	1	0	1
(D)	0	1		1				1

Figura 16 - Processo de criação de um novo Vírus
(A) Vírus aleatório (B) Cromossomo da população (C) Cromossomo com junção do material genético (D) Novo Vírus contendo material genético viral e do hospedeiro.

4.1.2.3. Infecção

A infecção (transcrição) é o processo de inclusão do material genético do Vírus no cromossomo hospedeiro, para isso, é necessário um Vírus e um cromossomo. O cromossomo destino da infecção terá seu material genético alterado nas mesmas

posições em que os genes estão dispostos no Vírus. Assim, todo o material genético do Vírus, exceto os espaços sem informação, são copiados para o cromossomo destino. A Figura 17 representa a infecção no AGRI.

(A)		1		1		1	0	
(B)	0	1	1	1	1	0	0	1
(C)	0	1	1	1	1	1	0	1

Figura 17 - Infecção de um cromossomo
(A) Vírus (B) Cromossomo (C) Cromossomo Infectado.

No AGRI o uso da transdução (cópia do material genético do indivíduo infectado para um novo Vírus) foi evitado, pois isto poderia deixar a população com menos diversidade. Dessa forma, o Vírus jamais é modificado após a sua criação, pois, na natureza esse comportamento também não acontece, o que ocorre é a criação de novos Vírus. A evolução viral é suportada também no AGRI através do método de evolução viral chamado Aquisição de Gene Celular.

Na Figura 18 se pode ter uma visão geral do processo de criação do Vírus (com utilização da retrotranscrição) e infecção do cromossomo (transcrição). Na mesma figura pode-se verificar o processo de troca genética entre todas as estruturas envolvidas.

(A)		1		1		1	0	
(B)	0	1	1	1	1	0	0	1
(C)	0	1	1	1	1	1	0	1
(D)	0	1		1				1
(E)	1	0	0	0	1	1	1	0
(F)	0	1	0	1	1	1	1	1

Figura 18 - Processo de criação de um Vírus e infecção de um cromossomo
(A) Vírus aleatório (B) Cromossomo da população (C) Cromossomo com junção do material genético entra A e B (D) Novo Vírus contendo material genético viral A e do hospedeiro B (E) Cromossomo a ser infectado (F) Cromossomo infectado com o Vírus D

O processo de infecção depende exclusivamente do aumento do *fitness*. Existem dois casos quando se trata de infecção:

- Infecção bem-sucedida: caso o cromossomo infectado tenha um *fitness* melhor que o anterior, a infecção é efetivada (o material genético do Vírus passa a incorporar o material genético do cromossomo) e o Vírus recebe mais um incremento no seu número de infecção.
- Infecção mal-sucedida: caso o cromossomo infectado tenha uma diminuição do *fitness*, o Vírus recebe um decréscimo no seu número de infecção e o cromossomo permanece com seu material genético intacto.

Isto evita que a busca por soluções fique prejudicada, já que cromossomos considerados bons poderiam se tornar ruins. Assim, somente as infecções bem-sucedidas são importantes para o AGRI

No AGRI, todos os indivíduos escolhidos para serem infectados terão possibilidade de sofrerem infecção por todos os Vírus da população viral. Este artifício é utilizado para maximizar o espaço de busca. Porém, computacionalmente, torna o algoritmo mais lento. Mas há uma forma de coibir este comportamento através de um parâmetro de infecção única.

4.1.2.4. Ciclo de vida de um Vírus

No AGRI, como na natureza, a tendência do Vírus é sempre desaparecer dando espaço para um Vírus com maior capacidade de infecção. Para que este ciclo seja completo são ocorrem três etapas: surgimento, infecção (luta pela existência do Vírus na população viral) e desaparecimento.

a. Surgimento

Os Vírus surgem a cada nova geração viral no AGRI, essa geração acompanha a geração cromossômica. A quantidade de Vírus que surge é a diferença entre o percentual dos Vírus que passarão para a próxima geração e a quantidade de total de Vírus. Esses novos Vírus são baseados em um nicho aleatório de cromossomos da população cromossômica, podendo ou não ter traços genéticos de outros Vírus, porém, nunca um Vírus nascerá diretamente de outro Vírus, pois, não há cruzamento viral na natureza.

b. Sobrevivência

Na população viral a sobrevivência de um Vírus depende da sua capacidade de infecção. Caso ele apresente uma alta taxa de infecção, será selecionado dentro de um percentual de um grupo que irá para a próxima geração. Se este Vírus continuar mantendo um alto nível de infecção, ele continuará na população viral até perder seu poder infeccioso. A taxa de elitismo viral, ou seja, os Vírus que tiveram maior grau de infectabilidade permanecem na população, controla este tipo de sobrevivência.

c. Desaparecimento

A tendência natural dos Vírus no AGRI é desaparecer. Isto acontece porque a população cromossômica fica mais evoluída e menor é a capacidade de infecção dos Vírus, já que uma infecção só é concretizada quando há um aumento no *fitness* do cromossomo. Isto acontece pelo fato de que ao passar do tempo os cromossomos vão ficando “vacinados”, ou seja, o *fitness* não tem nenhuma mudança após uma tentativa de infecção, que é denominado código viral inativo. Quando esta situação acontece, os Vírus que não conseguem mais infectar, desaparecem, dando espaço para novos Vírus, que terão genes aleatórios e também genes de cromossomos bons, o que aumenta a possibilidade de infecção.

4.1.2.5. Parâmetros

O AGRI utiliza os mesmos dados do AG clássico. Porém, para aplicação dos conceitos de infecção viral por Retrovírus os dados definidos foram: taxa de infecção populacional, taxa de elitismo viral, quantidade de Vírus, infecção dos mais fracos, infecção única e taxa de infecção interna.

a. Taxa de Infecção Populacional

A infecção populacional é uma característica importante, pois, determina o máximo de indivíduos do algoritmo que estarão aptos à infecção. Esta taxa é

utilizada para determinar quantos indivíduos participarão da tentativa de serem infectados, porém, não é garantido que todos sofram infecção.

Assim, se em uma população de cem indivíduos, esta taxa for de vinte por cento, não é garantido que os vinte indivíduos serão infectados e sim que vinte indivíduos estarão aptos a sofrer infecção viral e podem ser infectados ou não. Esta taxa é diretamente proporcional à variação cromossômica da população.

b. Taxa de Elitismo Viral

O elitismo viral tem o mesmo conceito do elitismo no algoritmo genético: preservar os melhores indivíduos para a próxima geração. Nesses casos, os Vírus com maiores sucessos de infecções serão passados a diante até o momento em que a população cromossômica criará imunidade a eles. A tendência é que os Vírus desapareçam, para manter a diversidade populacional de Vírus e cromossomos.

Os Vírus são ordenados por ordem de infecção, aqueles que estiverem dentro da porcentagem prevista na taxa, passarão para a próxima geração da população viral, caso contrário, serão excluídos e novos Vírus os substituirão.

Esta taxa é inversamente proporcional a variação da população viral, podendo levar o algoritmo a mínimo local, já que todos os indivíduos podem estar vacinados para aquela população viral específica. Quanto menor esse valor, maior variabilidade na população viral, podendo levar a uma variação muito grande dentro da população cromossômica.

c. Quantidade de Vírus

Este parâmetro define o número total de Vírus que farão parte da população viral. Esta quantidade está ligada diretamente ao desempenho do algoritmo, já que todos os cromossomos sofrerão tentativa de infecção por todos os Vírus a cada geração.

d. Infecção dos Mais Fracos

Este parâmetro especifica um determinado percentual de indivíduos que serão considerados fracos. Por isso, os indivíduos que poderão ser infectados são aqueles que por apresentam os menores *fitnesses* da população.

A ideia deste parâmetro é tentar infectar os cromossomos mais fracos da população para que possam se tornar competitivos a partir de uma infecção viral.

É um parâmetro *booleano* e trabalha em conjunto com a Taxa de Infecção Populacional.

Quando este valor é verdadeiro, a tendência é tornar o *fitness* de cromossomos fracos semelhantes ao *fitness* dos mais fortes, o que pode causar por sua vez um atraso de evolução. Caso este valor seja falso, pode causar mínimo local também por só privilegiar os mais cromossomos mais fortes.

e. Infecção única

Quando um indivíduo é infectado por mais de um Vírus em uma mesma geração, a probabilidade dele perder as principais características genéticas é elevada. Por isso, optou-se por utilizar o parâmetro de infecção única, ou seja, a partir do momento que um Vírus foi infectado, ele está imune de infecção até a próxima geração. Assim, ele permanece com parte do seu material genético original.

Este parâmetro é um do tipo *booleano* e quando é verdadeiro, a evolução se torna muito mais rápida em processamento, porém, o resultado pode demorar a ser encontrado.

f. Taxa de infecção interna

Este parâmetro é utilizado para indicar qual a porcentagem máxima de infecção dos genes do cromossomo. Porém, não garante que se um cromossomo for infectado ele terá exatamente o valor de novos genes previstos na taxa de infecção. Por exemplo, se um cromossomo de cem posições tiver esta taxa com

40%, quer dizer que o Vírus que vai nascer terá entre 5% e 40% de genes com dados infectáveis.

4.1.3. Algoritmo

Considerando a estrutura geral de algoritmo genético padrão se tem:

1. Inicializa população aleatória
2. Avalia indivíduos e verifica a solução
3. Efetua Seleção / Reprodução / Mutação
4. Avalia os indivíduos e verifica a solução, se solução não for encontrada voltar ao passo 3

O AGRI inclui um passo a mais no algoritmo tradicional que é a Aplicação Viral:

1. Inicializa população aleatória
2. Aplicação Viral
 - a. Gera uma população viral aleatória, caso seja primeira vez
 - b. Gera novos Vírus baseados na população
 - c. Propaga a infecção nos indivíduos
 - i. Infecciona cada cromossomo escolhido com cada Vírus
 - ii. Para cada cromossomo infectado com sucesso, ou seja, quando houver aumento na *fitness*, incrementar 1 para o Vírus, caso contrário, decrementar 1
 - iii. Verificar quais foram os Vírus com maiores graus de infecção e manter de acordo com a porcentagem os melhores para a próxima geração viral
3. Efetua Seleção, Reprodução e Mutação
4. Avalia os indivíduos infectados e verifica a solução, se solução não for encontrada voltar ao passo 2

CAPÍTULO V

5. RESULTADOS

Nos sub-capítulos a seguir serão apresentadas as funções matemáticas utilizadas para teste do AGRI.

5.1. FUNÇÕES UTILIZADAS

A implementação do AGRI foi feita para dar suporte para quatro funções previstas por Sungathan, et al. (2005) que tratam de problemas de otimização de parâmetro real. São elas:

a. F1: Shifted Sphere Function

$$F_1(\mathbf{x}) = \sum_{i=1}^D z_i^2 + f_bias_1, \mathbf{z} = \mathbf{x} - \mathbf{o}, \mathbf{x} = [x_1, x_2, \dots, x_D]$$

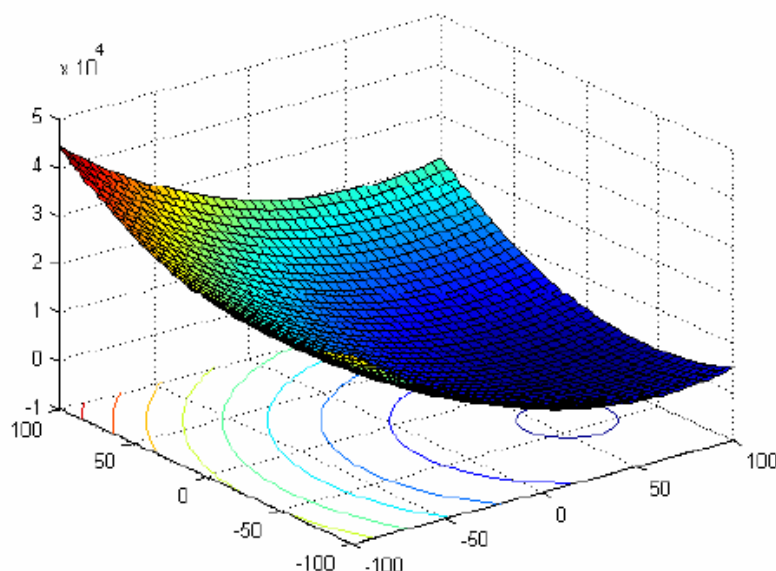


Figura 19 - Função F1 - *Shifted Sphere Function*.
Fonte: SUNGATHAN (2005)

b. *F2 Shifted Schwefel's Problem*

$$F_2(\mathbf{x}) = \sum_{i=1}^D \left(\sum_{j=1}^i z_j \right)^2 + f_bias_2, \quad \mathbf{z} = \mathbf{x} - \mathbf{o}, \quad \mathbf{x} = [x_1, x_2, \dots, x_D]$$

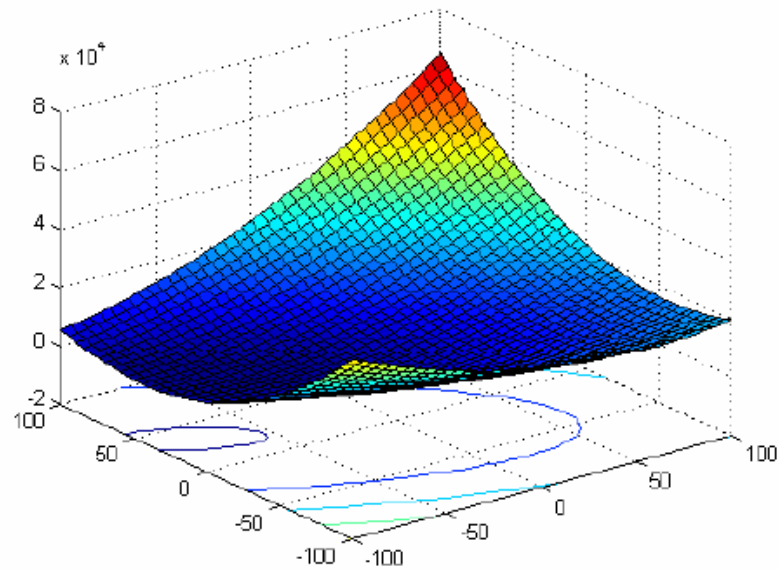


Figura 20 - Função F2 - *Shifted Schwefel's Problem*.
Fonte: SUNGATHAN (2005)

c. *F3: Shifted Rotated High Conditioned Elliptic Function*

$$F_3(\mathbf{x}) = \sum_{i=1}^D (10^6)^{\frac{i-1}{D-1}} z_i^2 + f_bias_3, \quad \mathbf{z} = (\mathbf{x} - \mathbf{o}) * \mathbf{M}, \quad \mathbf{x} = [x_1, x_2, \dots, x_D]$$

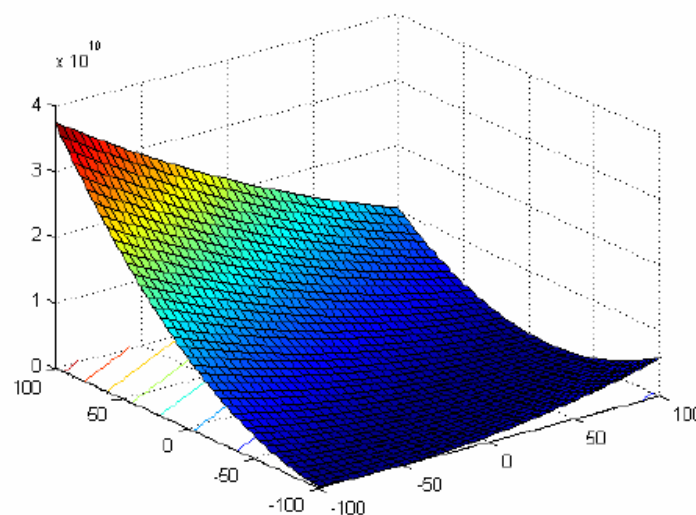


Figura 21 - Função F3 - *Shifted Rotated High Conditioned Elliptic Function*.
Fonte: SUNGATHAN (2005)

d. *F5: Schwefel's Problem 2.6 with Global Optimum on Bounds*

$$F_5(\mathbf{x}) = \max \{ |A_i \mathbf{x} - B_i| \} + f_bias_5, i = 1, \dots, D, \mathbf{x} = [x_1, x_2, \dots, x_D]$$

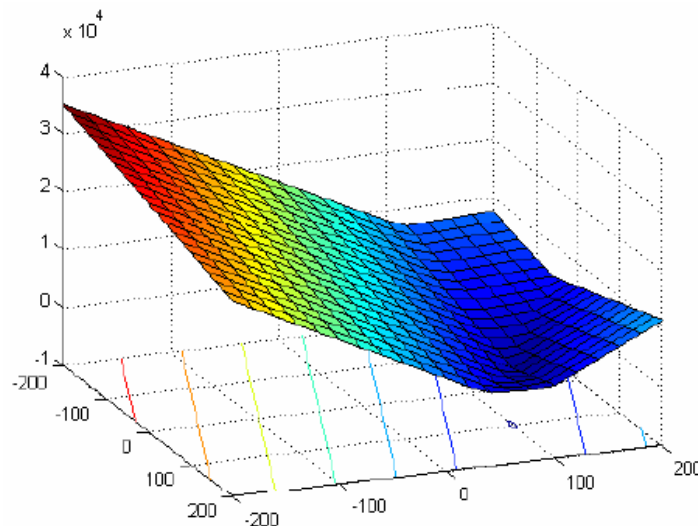


Figura 22 - Função F5 - Schwefel's Problem 2.6 with Global Optimum on Bounds.

Fonte: SUNGATHAN (2005)

5.2. EXPERIMENTOS

Para os testes do AGRI foram seguidos os seguintes passos:

1. Definição dos Parâmetros do AG;
2. Definição dos Parâmetros do AG Mutação (AG com alta taxa de mutação, porém com os mesmos parâmetros do AG);
3. Definição dos Parâmetros do AGRI;
4. Definição da população inicial utilizada em todos os experimentos da função para melhor analisar a evolução e fazer comparação entre os algoritmos (AG, AG Mutação e AGRI);
5. Tanto para o AG quanto para o AGRI foram feitas 10 execuções com as mesmas configurações, com o objetivo de verificar o comportamento em execuções diferentes dentro de um mesmo contexto, apenas para o AG Mutação que a taxa de mutação foi modificada;

6. Dados como média de *fitness*, melhor *fitness*, média de desvio padrão e a última execução foram armazenados para geração de gráfico.

5.3. ANÁLISE DOS RESULTADOS

Foram condensados na Tabela 1 abaixo os melhores resultados obtidos tanto nas execuções do AG, AG Mutação e AGRI. O resultado de todas as execuções pode ser visto nos anexos ANEXO A, ANEXO B, ANEXO C e ANEXO D.

Tabela 1 - Resultados das Melhores Execuções de AG, AG com Alta Mutação e AGRI para as funções F1, F2, F3 e F5.

CEC-LIB		AG			AG Mutação			AGRI		
Função	max(f(x))	Melhor Fitness	Fitness Médio	Desvio Padrão Médio	Melhor Fitness	Fitness Médio	Desvio Padrão Médio	Fitness	Fitness Médio	Desvio Padrão Médio
F1	$4,4 \times 10^4$	3,9214	1,0276	0,7116	3,7089	1,0798	0,7424	4,3893	3,9133	0,2707
F2	$6,5 \times 10^4$	5,4998	1,4186	0,8514	5,5907	1,8979	0,9266	6,4784	5,6941	0,4245
F3	$3,8 \times 10^{10}$	2.9530	0,5875	0,5308	2,7645	0,6361	0,5717	3.0421	2,7847	0,1390
F5	$3,4 \times 10^4$	1.6513	0,6822	0,3835	1,6781	0,7063	0,3876	1.7486	1,6557	0,0461

Levando em consideração todos os experimentos expostos e em anexo o AGRI foi superior em todas as execuções feitas (no total de 120 execuções), chegando mais próximo do resultado em todas elas.

Analisando ainda a Tabela 1, verifica-se o AGRI gerou os melhores *fitnesses* em relação aos outros métodos. Além disso, a média do *fitness* do AGRI foi inclusive superior que o melhor *fitness* encontrado em algumas funções pelo AG e AG Mutação como é o caso da função F1 e ficou muito próximo dos melhores *fitnesses* nas funções F1, F3 e F5. Assim, pode-se afirmar que o AGRI, com a mesma população inicial, conseguiu elevar o *fitness* de indivíduos considerados ruins, chegando assim bem próximo de um resultado ótimo em apenas 100 gerações.

Analisando os gráficos nos anexos, pode-se perceber que o AGRI chegou próximo dos melhores *fitness* logo nas 15 primeiras gerações, em alguns casos até nas 10 primeiras gerações, ao contrário do AG que chegava perto do seu melhor *fitness* somente a partir da geração 50 (exceto na função F5). O que nos leva a concluir que a alta taxa de mutação viral em conjunto com a disseminação de material genético entre os membros da população leva a ter melhores resultado em menos gerações. Apesar de

o AG e AG Mutação serem mais rápidos, o AGRI consegue ter resultado melhores em menos tempo já que precisa de menos gerações para chegar à melhores resultados.

Conclui-se também, pela média do desvio padrão, que os indivíduos do AGRI tiveram pouca variação, já que, nas primeiras gerações, o AGRI chegou perto de atingir o resultado ótimo.

Dessa forma pode-se afirmar que o método é mais eficiente para os problemas matemáticos expostos. Contudo, não se pode afirmar que o método é eficiente para todas as situações, pois, cada função ou problema pode ter particularidades que podem ser anuladas na busca do resultado pela aplicação do apresentado.

CAPÍTULO VI

6. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

Este trabalho apresentou a aplicação de a criação de um algoritmo baseado em estrutura da família dos Retrovírus. A aplicação da estrutura biológica em conjunto com algoritmo genético levou este algoritmo a ser chamado Algoritmo Genético Retroviral Iterativo.

Ao término deste trabalho, verificou-se que a inspiração biológica retroviral teve um efeito benéfico na aplicação do AGRI em comparação ao AG, já que em todos os resultados o AGRI conseguiu ter um melhor resultado que o AG.

Apesar de o AGRI consumir mais tempo de processamento (já que precisa criar Vírus e infectar a população) este tempo é compensando em resultados obtidos mais rápidos, pois são gerados em menos gerações, já que para chegar próximo ao valor de *fitness* ideal o AGRI precisa de menos gerações de processamento.

Com isso, pelos resultados positivos obtidos pelo AGRI, pode-se concluir que o AGRI se mostra tanto eficiente quanto eficaz, em comparação ao AG e AG Mutação para as funções expostas no trabalho. O que não deve ser tomado como regra para qualquer problema que envolva resoluções a partir de modelagem de AG.

Como trabalho futuro é proposta uma avaliação mais precisa do AGRI para verificar o impacto de cada variável do algoritmo na execução do mesmo, fazendo inclusive avaliação de junções de variáveis presentes no AG e no AGRI, por exemplo, alta taxa de mutação do AG em conjunto com uma baixa taxa de infecção tem determinado impacto na população.

Além disso, é proposto também o estudo e aplicação das outras características de evolução do Vírus, já que o AGRI utilizou apenas uma das quatro formas (aquisição genética) de evolução viral.

É proposto também o estudo da técnica de invasão viral, já que os Vírus atacam somente células específicas e não todas (como o AGRI está considerando). Dessa forma, gerando uma estrutura computacional onde pudesse se verificar se um cromossomo é ou não elegível para sofrer infecção.

E por fim, é proposto o estudo e aplicação de novos métodos de infecção viral, já que a maioria da proliferação viral se dá por transmissão entre indivíduos, podendo aplicar regras específicas de transmissão, como por exemplo, por grau de distância, fraqueza imunológica de um cromossomo bem como o estudo imunológico para evitar a contaminação de cromossomos que apesar de terem *fitness* aumentado não tenham modificação na estrutura do seu *schemata*.

REFERÊNCIAS BIBLIOGRÁFICAS

- AGUT, A. **Um Sistema Estratégico de Reprodução**. Scientific American Brasil, São Paulo, Edição Especial, n. 28, p-14-19, jan. 2009
- BEASLEY, D.; BULL, D. R.; MARTIN, R. R. **An Overview of Genetic Algorithms: Part 1, Fundamentals**. Inter-University Committee on Computing. University Computing, UK, 1993.
- CARROLL, S.; **Infinitas formas de grande beleza**. Traduzido por: Diego Alfaro, Jorge Zahar Editor, Rio de Janeiro, 2005.
- CARTER, J.; SAUNDERS, V. **Virology Principles and Applications**. John Wiley & Sons Ltd. 2007.
- CHENG, R.; MIYAMURA, T. **Structure-Based Study of Viral Replication**. World Scientific Publishing Co. Pte. Ltd. 2008.
- COELHO, A.; LAMONT, G.; VELDHUIZEN, D. **Evolutionary Algorithms for Solving Multi-Objective Problems**. 2nd Edition. Springer. Stanfor. 2006.
- DARWIN, C. A Origem das Espécies. Traduzido por: John Green, Editora Martin Claret: São Paulo, 2005. 629 p.
- FERREIRA, A. **Novo Dicionário da Língua Portuguesa**. 2. ed. Rio de Janeiro: Nova Fronteira, 1986. 1838 p.
- GERARDI, M.; ZIMMERMAN, M. **Wastewater Pathogens**. John Wiley & Sons Ltd. 2005.
- GRIFFITHS, J. F.; GELBART, W. M.; MILLER, J. H.; LEWONTIN, C. R. **The Molecular Basis of Mutation Modern Genetic Analysis**. W. H. Freeman and Company. 1999.
- GUEDES, A.; LEITE J.; ALOISE, D. **Um algoritmo genético com infecção viral para o problema do caixeiro viajante**. 2005

HAUPT, R. L.; HAUPT, S. E. **Practical Genetic Algorithms**. John Wiley & Sons, Inc., 1998.

HOGG, S. **Essential Microbiology**. John Wiley & Sons Ltd. 2005.

JOHNSON D.; MCGEOCH L. **The Traveling Salesman Problem: A Case Study in Local Optimization**. Draft of November 20, 1995. To appear as a chapter in the book *Local Search in Combinatorial Optimization*, E. H. L. Aarts and J. K. Lenstra(eds.), John Wiley and Sons, New York., 1995.

LERNER, K.; LERNER, B. **World of Microbiology and Immunology**. Vol 1 e 2. Thomson Gale, 2003.

LINDEN, R. **Algoritmos Genéticos**. 1. ed. Rio de Janeiro: Brasport., v. 1. 372 p., 2006.

MACKENZIE, D. **Deadly new flu virus in US and Mexico may go pandemic.**, 2009,
Disponível em: <http://www.newscientist.com/article/dn17025-deadly-new-flu-virus-in-us-and-mexico-may-go-pandemic.htm>. Acessado em: 30 jan 2010.

MITCHELL, M. **An Introduction to Genetic Algorithms**. MIT Press, 1999.

SUGANTHAN, P; HANSEN, N.; LIANG, J.; DEB, K.; CHEN, Y.; AUGER, A.;TIWARI, S. **Problem Definitions and Evaluation Criteria for the CEC 2005 Special Session on Real-Parameter Optimization. Technical Report**, Nanyang Technological University, Singapore And KanGAL Report Number 2005005 (Kanpur Genetic Algorithms Laboratory, IIT Kanpur). 2005

VILLARREAL, L. **Vírus São Seres Vivos?**. Scientific American Brasil, São Paulo, Edição Especial, n. 28, p-21-24, jan. 200

ANEXO A

Experimento 1 – Função F1

Parâmetros do AG:

- **População:** 100
- **Gerações:** 100
- **Taxa de Mutação:** 40%
- **Elitismo:** 20%
- **Cruzamento:** 90%
- **Seleção:** Roleta Simples
- **Cruzamento:** Ponto de Corte

Parâmetros do AG Mutação

- **Alta Taxa de Mutação:** 95%

Parâmetros do AGRI:

- **Taxa de Infecção Populacional:** 50%
- **Taxa de Elitismo Viral:** 15%
- **Quantidade de Vírus:** 30
- **Taxa de Infecção Interna:** 60%
- **Infeciona Mais Fracos:** Sim

Tabela 2 - Resultados de Execução Para Função F1

Algoritmo	Execuções									
	1	2	3	4	5	6	7	8	9	10
Melhor Fitness										
AG	2.9909	3.4759	3.9170	3.5204	3.2236	2.9083	3.3516	2.7161	3.9215	3.1793
AG Mutação	3,5749	3,7089	3,5553	3,5827	2,9218	3,3738	2,7115	3,5753	3,2129	3,2609
AGRI	4.4040	4.4157	4.3893	4.4017	4.4133	4.4099	4.4100	4.4201	4.4107	4.4185
Média do Desvio Padrão										
AG	0.7191	0.6966	0.7215	0.7432	0.7103	0.6954	0.7143	0.7236	0.7024	0.6890
AG Mutação	0,7647	0,7614	0,7724	0,7352	0,7568	0,7402	0,7216	0,7285	0,7298	0,7139
AGRI	0.2656	0.2737	0.2603	0.2767	0.2495	0.2631	0.2699	0.2909	0.2780	0.2794
Fitness Médio										
AG	1.0332	1.0214	1.0389	1.0427	1.0187	0.9794	1.0688	0.9994	1.0015	1.0717
AG Mutação	1,1013	1,0894	0,9921	1,0770	1,1063	1,1100	1,0050	1,0339	1,1348	1,1479
AGRI	3.9147	3.9235	3.9342	3.9393	3.9374	3.8994	3.9120	3.8887	3.9012	3.8827

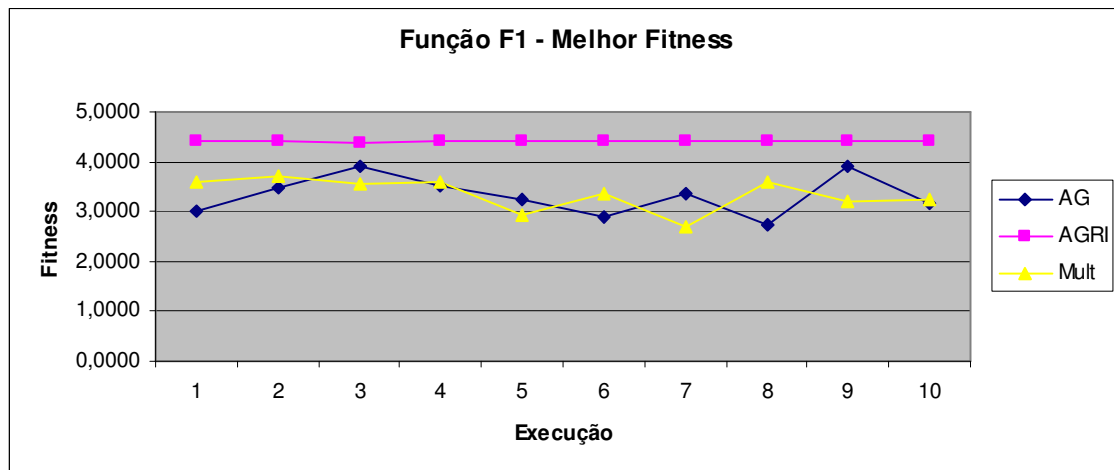


Figura 23 - Função F1 - Melhor Fitness.

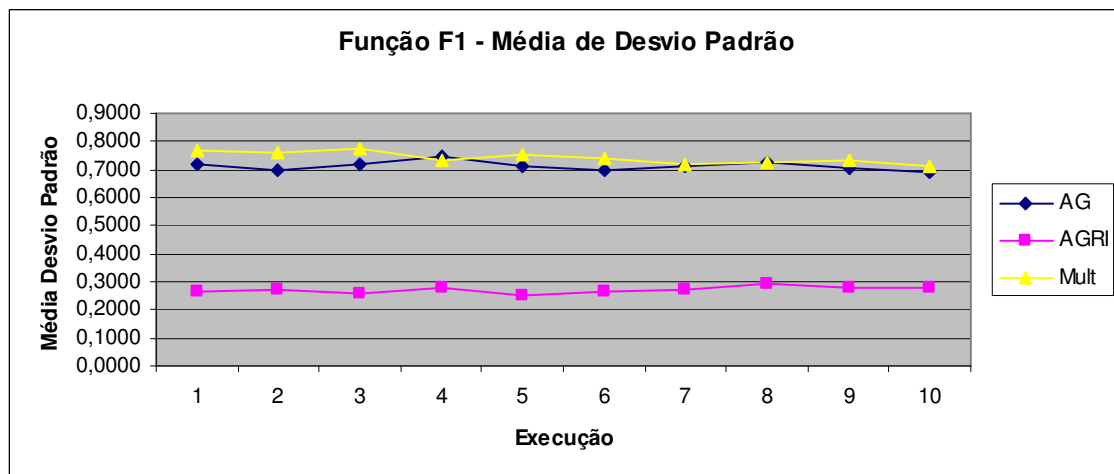


Figura 24 - Função F1 - Média do Desvio Padrão.

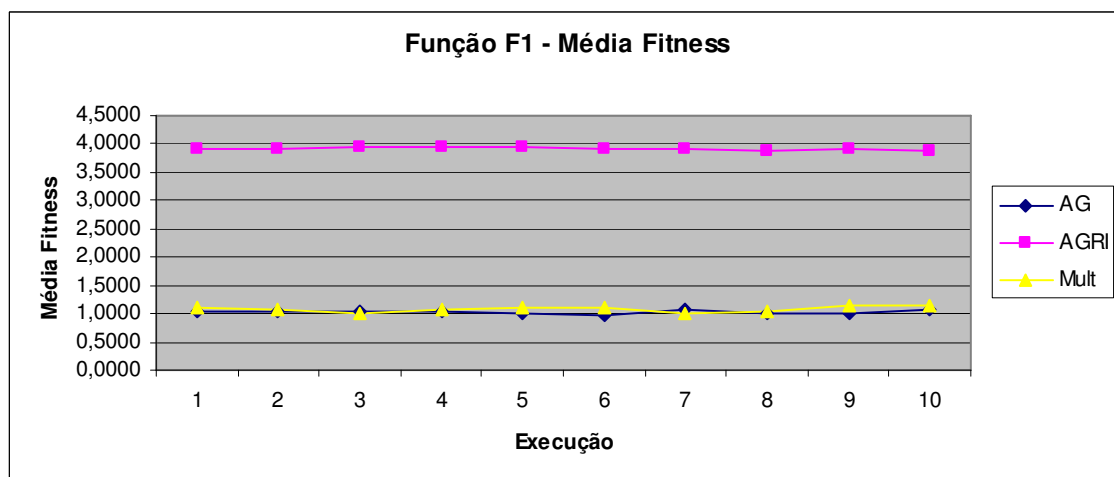


Figura 25 - Função F1 - Fitness Médio.



Figura 26 - *Fitness* Durante as Gerações do AG Sem Infecção para Função F1 (6 segundos).

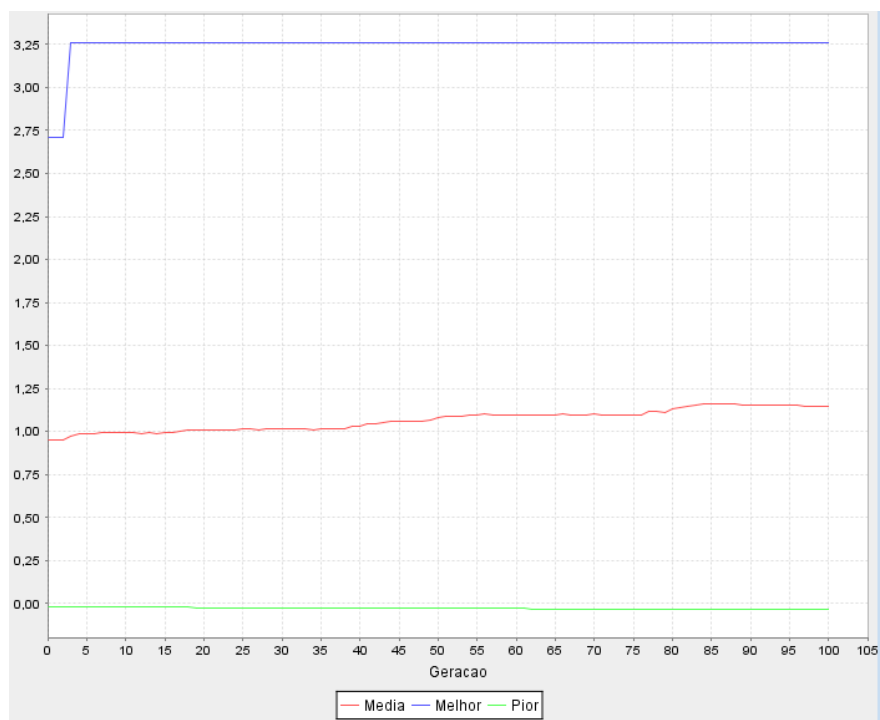


Figura 27 - *Fitness* Durante as Gerações do AG Mutação para Função F1 (6 segundos).

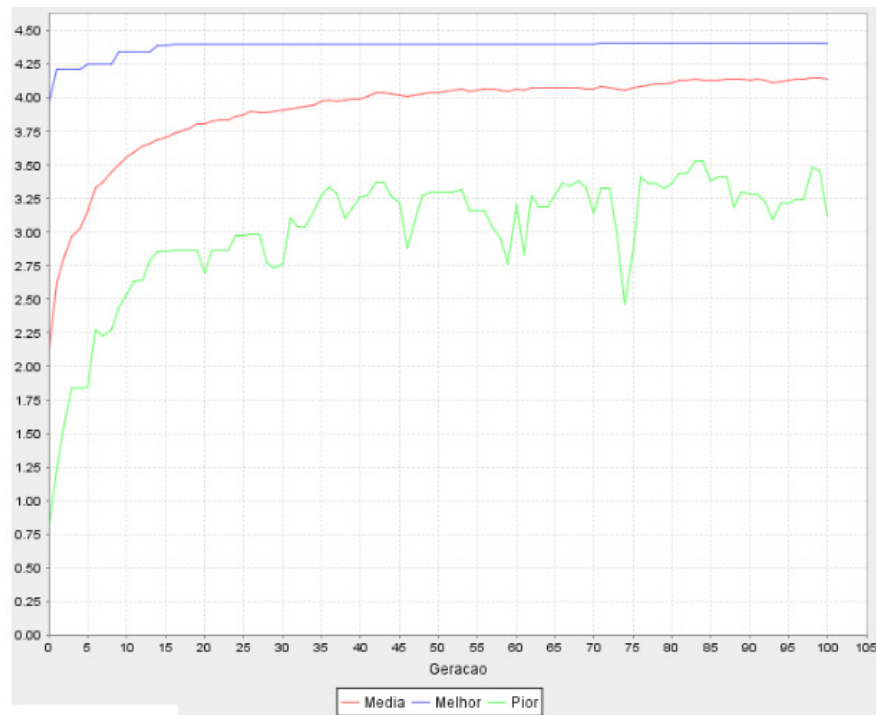


Figura 28 – *Fitness* Durante as Gerações do AGRI para Função F1 (14 segundos).

ANEXO B

Experimento 2 – Função F2

Parâmetros do AG:

- **População:** 100
- **Gerações:** 100
- **Taxa de Mutação:** 40%
- **Elitismo:** 20%
- **Cruzamento:** 90%
- **Seleção:** Roleta Simples
- **Cruzamento:** Máscara Binária

Parâmetros do AG Mutação

- **Alta Taxa de Mutação:** 95%

Parâmetros do AGRI:

- **Taxa de Infecção Populacional:** 50%
- **Taxa de Elitismo Viral:** 20%
- **Quantidade de Vírus:** 50
- **Taxa de Infecção Interna:** 60%
- **Infeciona Mais Fracos:** Sim

Tabela 3 - Resultados de Execução Para Função F2.

Algoritmo	Execuções									
	1	2	3	4	5	6	7	8	9	10
Melhor Fitness										
AG	5.3049	4.4880	4.8079	4.3181	4.5568	4.1274	3.9663	4.5823	4.3226	5.4998
AG Mutação	3,9663	4,6731	4,5596	4,8499	5,5907	5,2493	4,4694	3,9663	4,0957	4,3661
AGRI	6.4737	6.4741	6.4666	6.4542	6.4604	6.4562	6.4785	6.4758	6.4635	6.4710
Média do Desvio Padrão										
AG	0.8574	0.8252	0.9123	0.8616	0.8526	0.8678	0.8396	0.8206	0.8525	0.8249
AG Mutação	0,8593	0,9686	0,9570	0,9634	0,9066	1,0278	0,8946	0,8806	0,8663	0,9419
AGRI	0.4399	0.4213	0.3924	0.4542	0.3902	0.4310	0.4470	0.4468	0.4395	0.3826
Fitness Médio										
AG	1.4007	1.3744	1.4858	1.4591	1.3927	1.4260	1.4046	1.3971	1.4168	1.4292
AG Mutação	1,4113	1,5152	1,4530	1,5552	1,5728	1,5425	1,4418	1,3675	1,5215	1,5016
AGRI	5.7100	5.7200	5.7252	5.6876	5.6754	5.6691	5.6315	5.6941	5.6505	5.7773

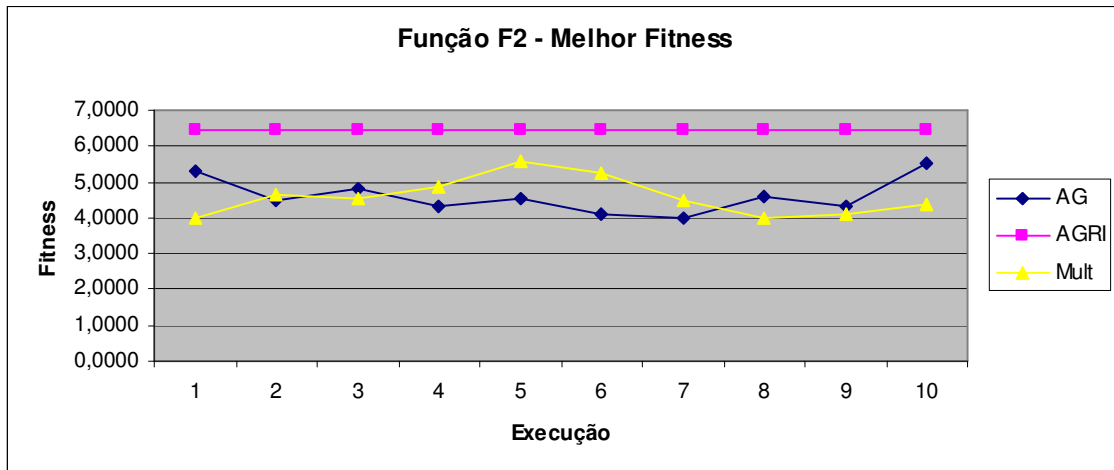


Figura 29 - Função F2 - Melhor Fitness.

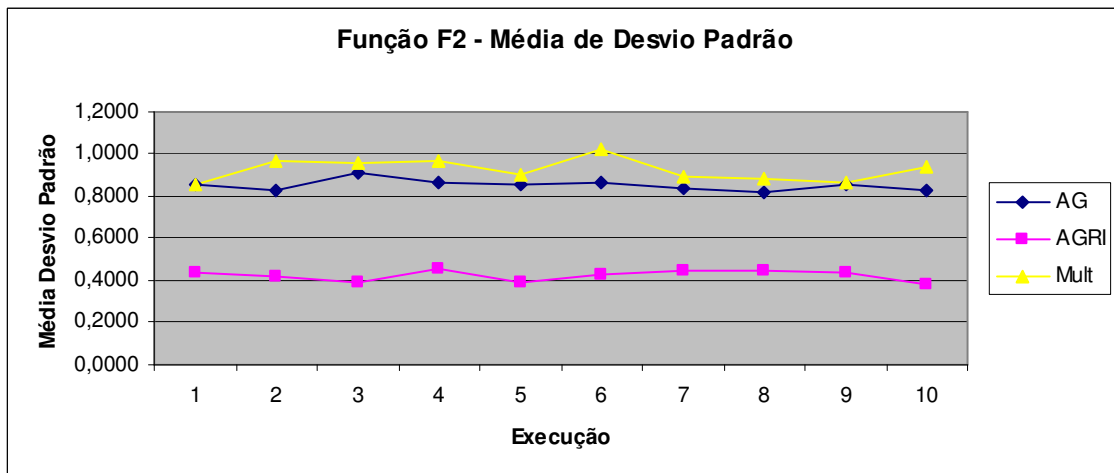


Figura 30 - Função F2 - Média do Desvio Padrão.

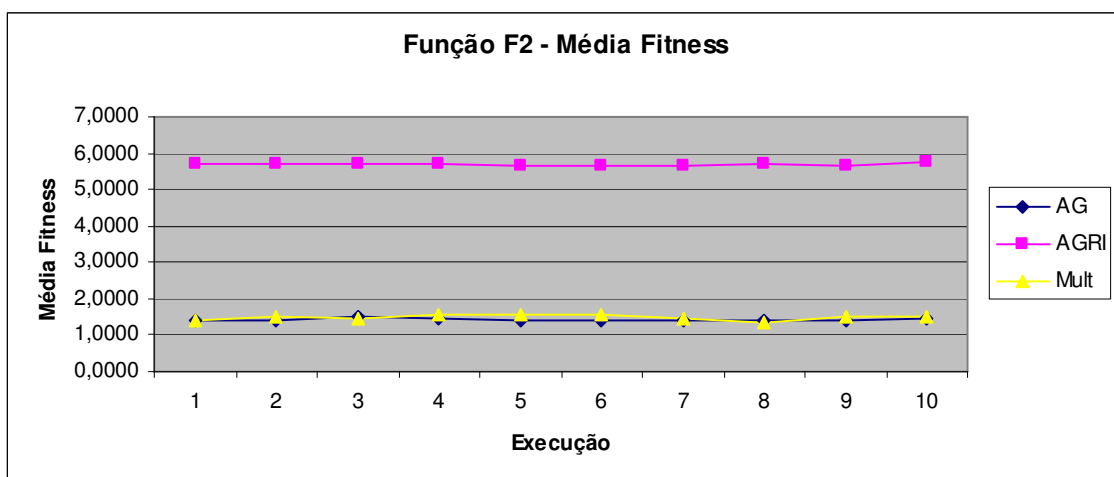


Figura 31 - Função F2 - Média de Fitness.

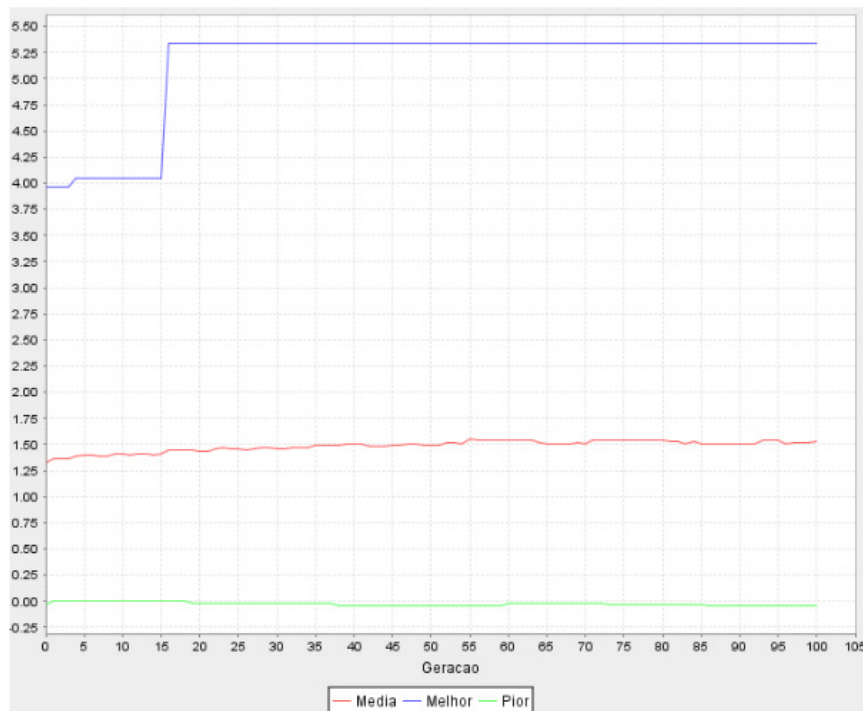


Figura 32 - Fitness Durante as Gerações do AG Sem Infecção para Função F2 (6 segundos).

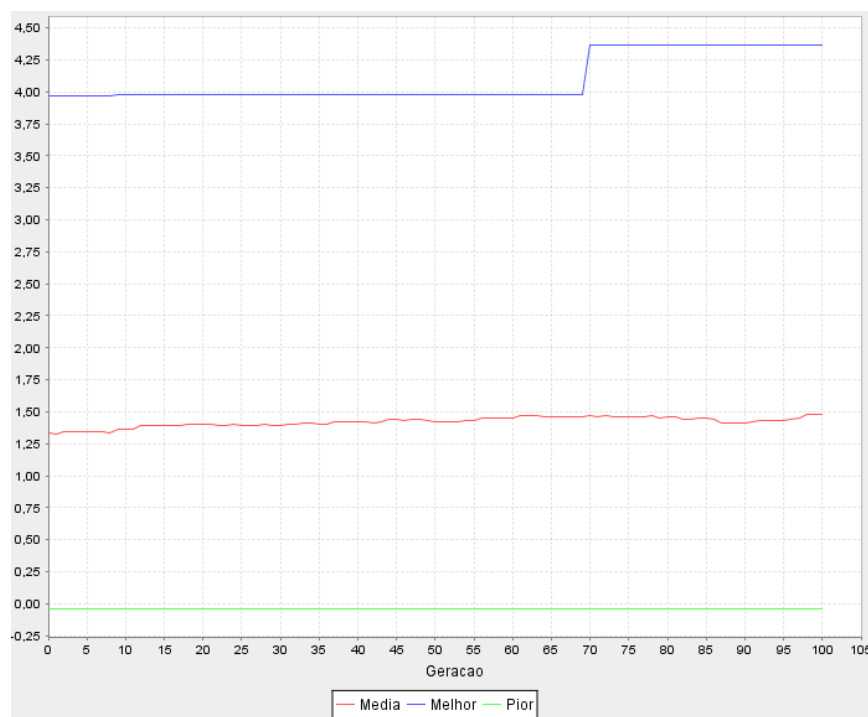


Figura 33 - Fitness Durante as Gerações do AG Mutação para Função F2 (6 segundos).



Figura 34 - Fitness Durante as Gerações do AGRI para Função F2 (15 segundos).

ANEXO C

Experimento 3 – Função F3

Parâmetros do AG:

- **População:** 100
- **Gerações:** 100
- **Taxa de Mutação:** 50%
- **Elitismo:** 30%
- **Cruzamento:** 80%
- **Seleção:** Roleta Simples
- **Cruzamento:** Máscara Binária

Parâmetros do AG Mutação

- **Alta Taxa de Mutação:** 95%

Parâmetros do AGRI:

- **Taxa de Infecção Populacional:** 40%
- **Taxa de Elitismo Viral:** 10%
- **Quantidade de Vírus:** 30
- **Taxa de Infecção Interna:** 40%
- **Infeciona Mais Fracos:** Sim

Tabela 4 - Resultados de Execução Para Função F3.

Algoritmo	Execuções									
	1	2	3	4	5	6	7	8	9	10
Melhor Fitness										
AG	2.2332	2.7023	1.9338	2.4571	2.3307	2.9530	2.1567	2.0911	2.4797	2.3303
AG Mutação	2.5686	1.9338	2.4021	2.6283	2.7645	1.9338	2.7626	2.0939	1.9338	2.4723
AGRI	3.0376	3.0377	3.0348	3.0304	3.0393	3.0339	3.0375	3.0381	3.0277	3.0421
Média do Desvio Padrão										
AG	0,5016	0,5651	0,5075	0,5206	0,5494	0,5273	0,5152	0,5091	0,5485	0,5641
AG Mutação	0,6410	0,5783	0,5574	0,5896	0,6211	0,5372	0,6079	0,5308	0,5264	0,5271
AGRI	0,1343	0,1382	0,1329	0,1450	0,1359	0,1394	0,1287	0,1416	0,1429	0,1510
Fitness Médio										
AG	0,5388	0,6468	0,5662	0,5727	0,5958	0,5894	0,5483	0,5872	0,6056	0,6244
AG Mutação	0,7015	0,6163	0,6271	0,7287	0,7230	0,6312	0,6435	0,5854	0,5559	0,5487
AGRI	2,7958	2,7925	2,7830	2,7767	2,7917	2,7855	2,7962	2,7751	2,7809	2,7697

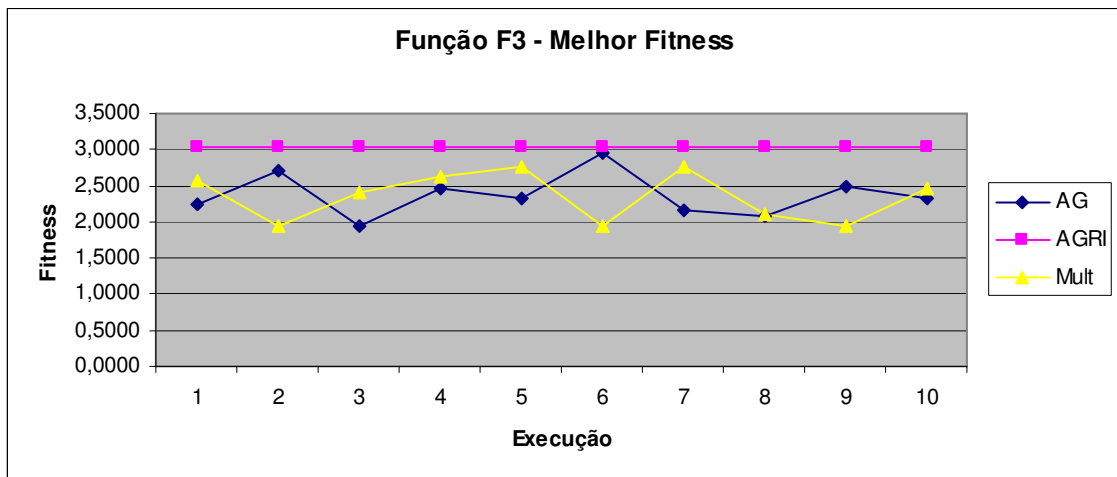


Figura 35 - Função F3 - Melhor Fitness.

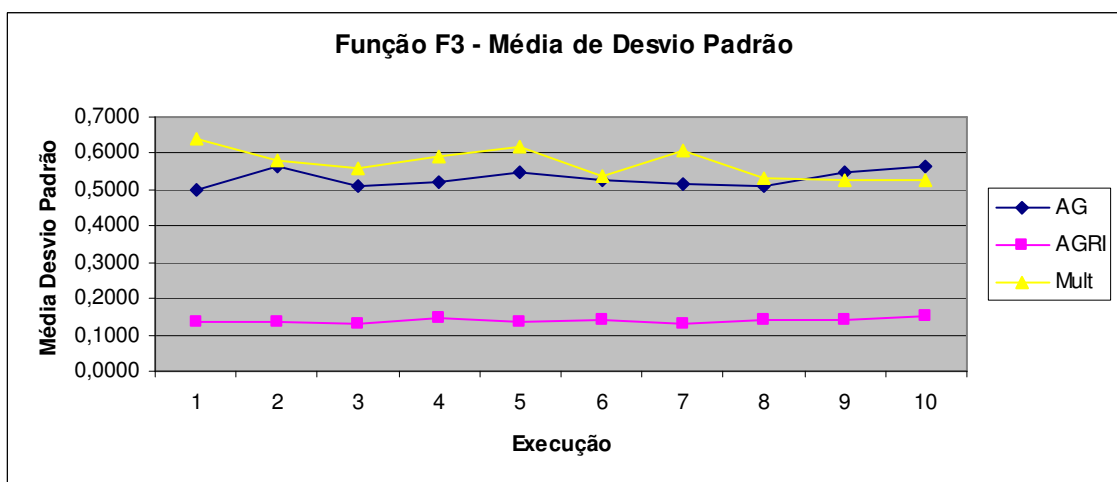


Figura 36 - Função F3 - Média do Desvio Padrão.

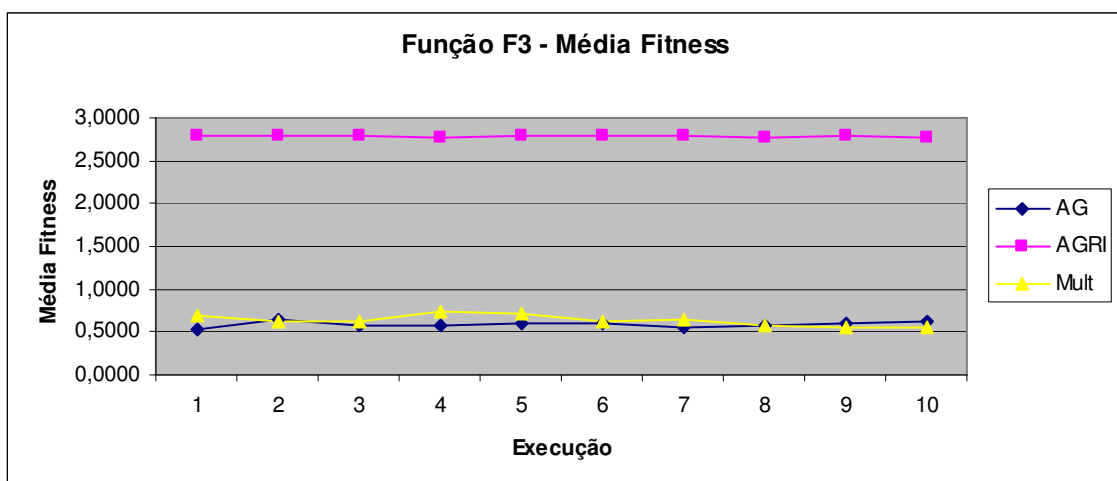


Figura 37 - Função F3 - Média de Fitness.



Figura 38 - Fitness Durante as Gerações do AG Sem Infecção para Função F3 (8 segundos).

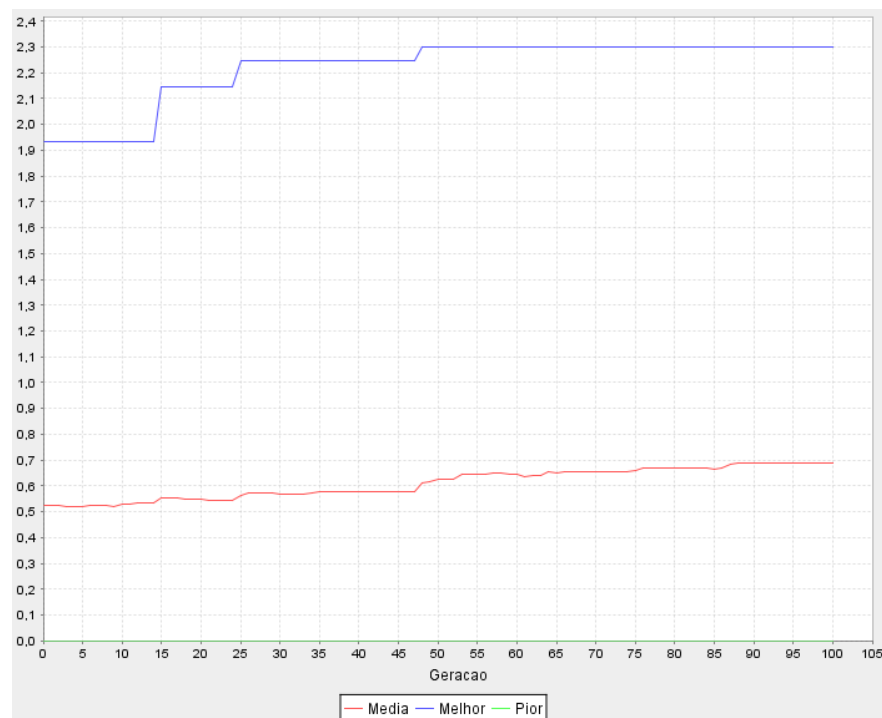


Figura 39 - Fitness Durante as Gerações do AG Mutação para Função F3 (8 segundos).

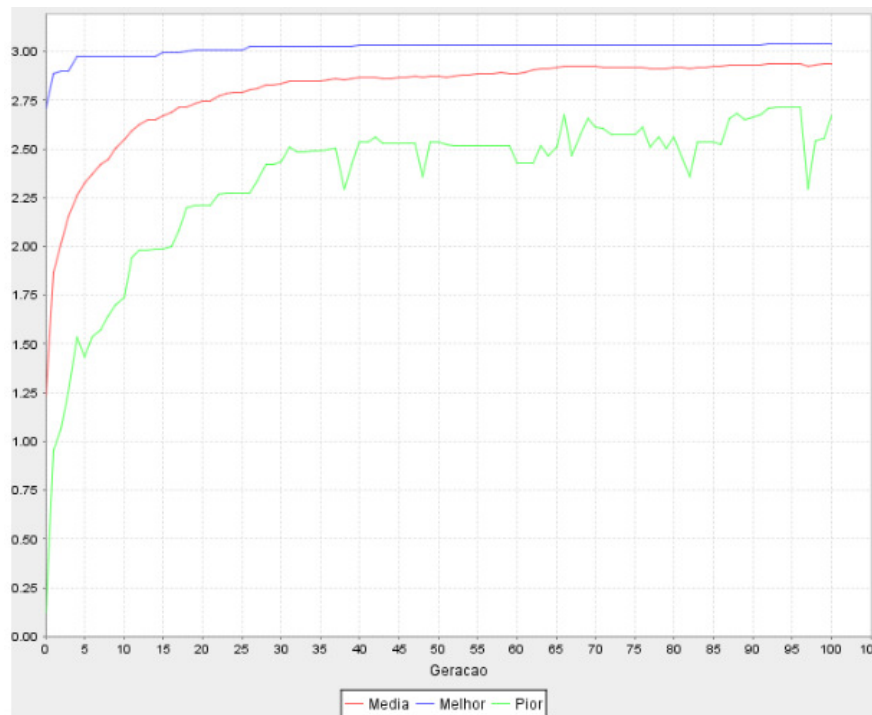


Figura 40 - Fitness Durante as Gerações do AGRI para Função F3 (15 segundos).

ANEXO D

Experimento 4 – Função F5

Parâmetros do AG:

- **População:** 100
- **Gerações:** 100
- **Taxa de Mutação:** 60%
- **Elitismo:** 15%
- **Cruzamento:** 60%
- **Seleção:** Roleta Simples
- **Cruzamento:** Máscara Binária

Parâmetros do AG Mutação

- **Alta Taxa de Mutação:** 95%

Parâmetros do AGRI:

- **Taxa de Infecção Populacional:** 30%
- **Taxa de Elitismo Viral:** 15%
- **Quantidade de Vírus:** 30
- **Taxa de Infecção Interna:** 60%
- **Infeciona Mais Fracos:** Sim

Tabela 5 - Resultados de Execução Para Função F5.

Algoritmo	Execuções									
	1	2	3	4	5	6	7	8	9	10
Melhor Fitness										
AG	1,5380	1,4690	1,5076	1,5714	1,5114	1,6239	1,5713	1,5388	1,6230	1,6513
AG Mutação	1,6393	1,5647	1,5870	1,5179	1,6161	1,6009	1,6694	1,6781	1,5673	1,5091
AGRI	1,7475	1,7467	1,7443	1,7480	1,7486	1,7483	1,7482	1,7473	1,7479	1,7461
Média do Desvio Padrão										
AG	0,3911	0,3842	0,3737	0,3878	0,3872	0,3835	0,3666	0,3869	0,3844	0,3896
AG Mutação	0,3985	0,3812	0,3747	0,3606	0,3758	0,3928	0,3993	0,4015	0,3828	0,4088
AGRI	0,0474	0,0474	0,0506	0,0471	0,0469	0,0439	0,0449	0,0437	0,0437	0,0456
Fitness Médio										
AG	0,6878	0,6739	0,6688	0,6981	0,6582	0,6857	0,6952	0,6900	0,6901	0,6740
AG Mutação	0,6760	0,7126	0,7130	0,6818	0,6845	0,7727	0,7420	0,7240	0,6842	0,6726
AGRI	1,6571	1,6550	1,6559	1,6553	1,6475	1,6545	1,6598	1,6596	1,6597	1,6521

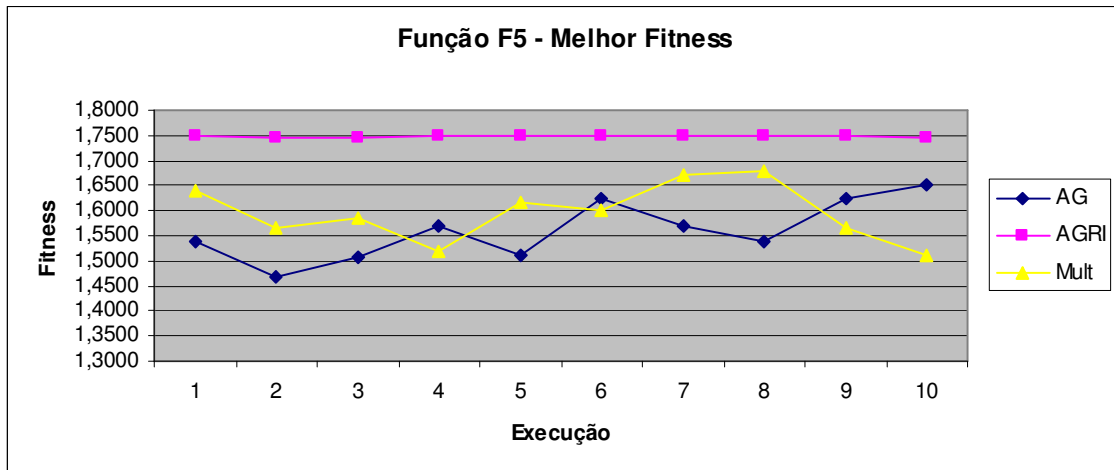


Figura 41 - Função F5 - Melhor Fitness.

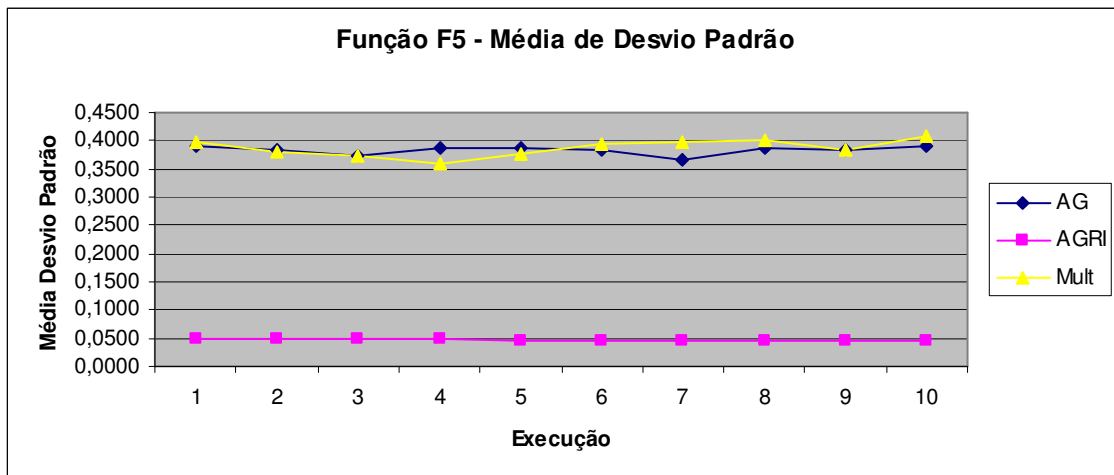


Figura 42 - Função F5 - Média do Desvio Padrão.

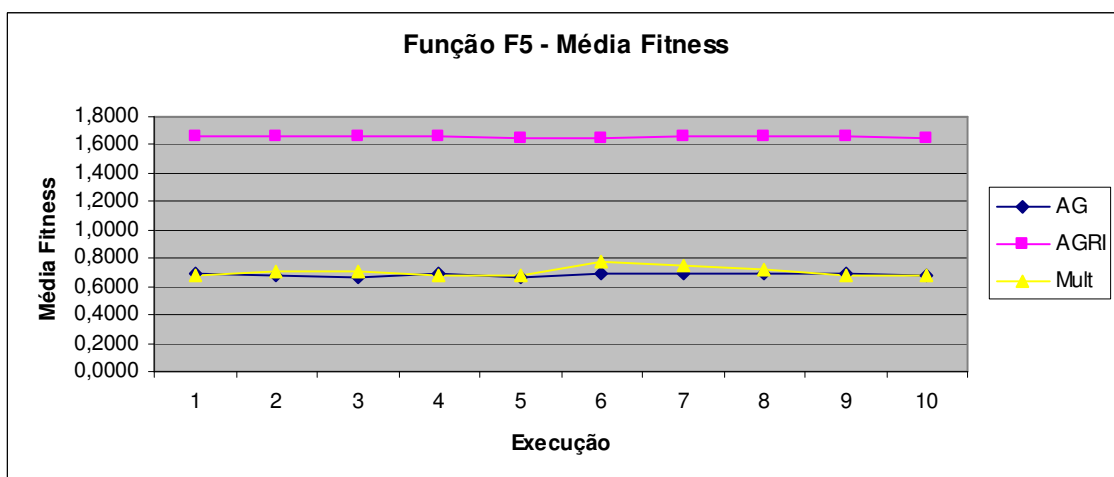


Figura 43 - Função F5 - Média de Fitness.



Figura 44 - Fitness Durante as Gerações do AG Sem Infecção para Função F5 (6 segundos).



Figura 45 - Fitness Durante as Gerações do AG Mutação para Função F5 (6 segundos).

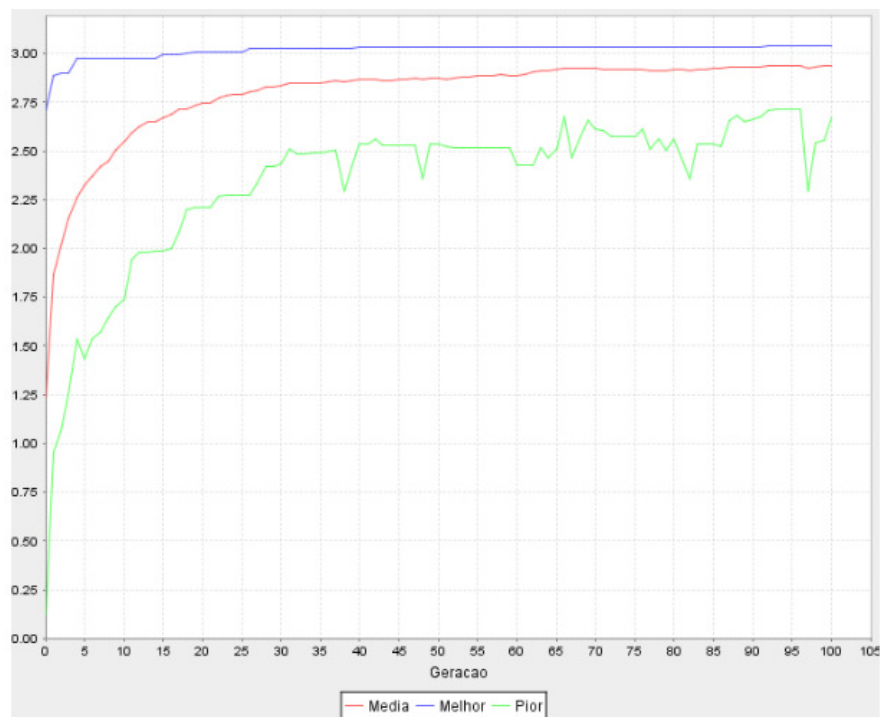


Figura 46 - Fitness Durante as Gerações do AGRI para Função F5 (12 segundos).