

3 Algoritmos Genéticos

Algoritmos Genéticos (AGs) constituem um mecanismo de busca adaptativa que se baseia no princípio Darwiniano de seleção natural e reprodução genética [101]. AGs são tipicamente empregados na resolução de problemas na forma $f: S \rightarrow \mathfrak{R}$, onde S representa o espaço de busca constituído por todas as possíveis soluções para um problema particular. Dependendo das peculiaridades do problema, as soluções podem ser representadas por vetores n -dimensionais de números binários, inteiros, reais, ou estruturas mais complexas. Por sua vez, a função f permite que se mapeie cada uma das soluções existentes no domínio S em números reais associados, cujos valores indicam quão boa a solução é para resolver o problema em questão. Estes valores compõem a imagem \mathfrak{R} da função f [102-104].

A tarefa principal de um algoritmo evolucionário é buscar de forma eficiente, em amostras do espaço de busca S , soluções que estejam de acordo com o objetivo do problema. É importante mencionar que estas soluções não precisam ser necessariamente ótimas, mas sim satisfatórias (sub-ótimas). Ao se lidar com espaços de busca grandes e complexos, pode ser difícil identificar o ponto ótimo e, neste caso, pode ser aceitável a obtenção de soluções sub-ótimas [105]. A Figura 13 ilustra o funcionamento de um algoritmo genético.

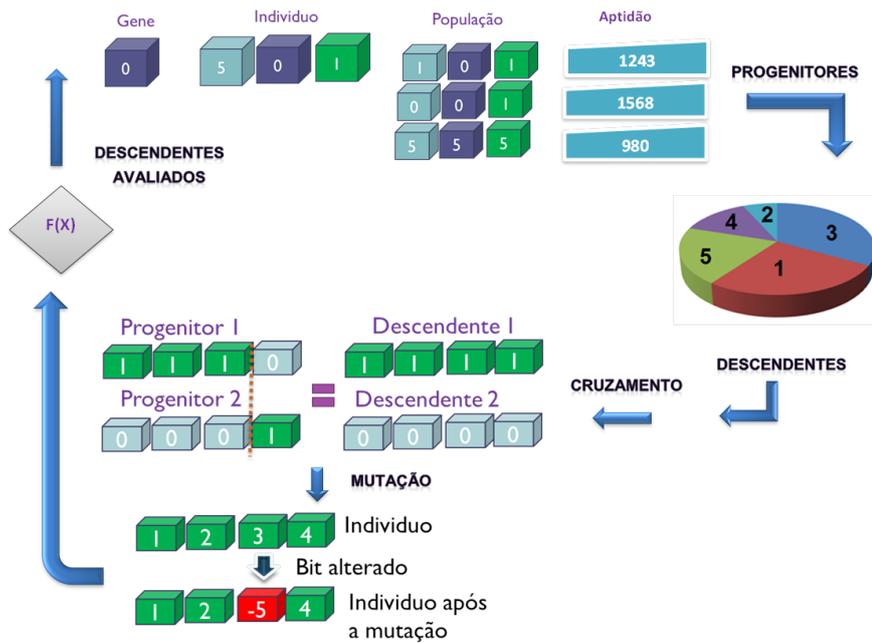


Figura 13 - Fluxograma típico de um algoritmo evolucionário.

Tendo em vista as características particulares de um dado problema, deve-se selecionar uma representação adequada para codificar possíveis soluções em estruturas de dados contidas no conjunto S . Após a representação ser escolhida, um número n de potenciais soluções, também chamadas de indivíduos, é gerado aleatoriamente a fim de se implementar a denominada população inicial do problema. Estes indivíduos passam então por duas etapas básicas: avaliação e operações genéticas.

Durante a avaliação, um número real, também chamado de aptidão, é associado a cada indivíduo. A aptidão do indivíduo mede o quanto ele é adequado para satisfazer à especificação de um problema particular. Após serem avaliados, os seguintes operadores genéticos são aplicados aos indivíduos: seleção, cruzamento e mutação. Na sequência, uma nova população é criada, formando assim a próxima geração.

A avaliação e os operadores genéticos são aplicados na próxima geração e o processo continua, até que um critério de parada seja satisfeito. Esse critério de parada pode ser um número máximo de gerações ou, por exemplo, vinculado ao alcance de um grau de adequabilidade considerado satisfatório, por um indivíduo da população [105]. Nas próximas subseções são apresentados com maiores detalhes os conceitos relacionados ao processo de evolução de Algoritmos Genéticos.

Do ponto de vista da otimização, uma das principais vantagens das técnicas de computação evolutiva é que estas não impõem muitos requisitos matemáticos sobre o problema a ser otimizado. A evolução da busca pelo ponto ótimo é guiada pela avaliação de uma função objetivo. Estas técnicas podem ser empregadas para obter-se a solução de variados tipos de problemas, sejam eles definidos num espaço discreto, contínuo ou misto, com ou sem restrições [106, 107].

3.1. Representação ou Codificação

Esta etapa refere-se à definição da forma como as soluções de um determinado problema serão codificadas em uma estrutura de dados a fim de possibilitar que as mesmas sejam processadas por um computador digital. Uma solução para um dado problema pode ser [106]:

- ❖ Um número, representando o valor ótimo de uma função matemática;
- ❖ Um vetor de reais, representando o valor ótimo de uma função com múltiplas variáveis;
- ❖ Uma lista de eventos, representando uma ordem ótima de eventos para se realizar uma determinada tarefa; ou
- ❖ Uma estrutura, simbolizando algum modelo de engenharia, um circuito elétrico, uma reação química ou qualquer outro sistema.

A solução de um problema é representada por um conjunto de parâmetros, sendo que cada um desses parâmetros é denominado gene, o qual pode ser número real, inteiro ou binário. Um conjunto de genes com a mesma representação constitui um segmento e um conjunto de segmentos compõe um indivíduo. Por sua vez, um conjunto de indivíduos com mesma estrutura implementa uma população.

Uma solução é formada pela interpretação de um indivíduo da população. Em alguns casos, pode-se utilizar mais de um indivíduo para representar uma solução, no entanto, cada um dos indivíduos que compõem a solução deve pertencer a uma população diferente [108].

A população inicial de indivíduos é, na maioria das vezes, gerada de forma aleatória. No entanto, existem casos onde é mais apropriado gerar a população

inicial por meio de uma heurística, a fim de se poder introduzir, já na população inicial, indivíduos com características interessantes para o problema em questão. Por exemplo, adotando esta abordagem, é possível acrescentar à população inicial soluções aproximadas, conhecidas por meio de algum tipo de informação prévia. Entretanto, destaca-se que existem diversos trabalhos que comprovam que a geração da população inicial não é uma fase crítica em Algoritmos Genéticos, indicando que os mesmos conseguem evoluir para soluções satisfatórias independentemente do conjunto de indivíduos que compõem a população inicial, desde que a mesma contenha indivíduos suficientemente diversificados [109].

Ressalta-se ainda que deve-se definir a codificação a ser adotada para um dado conjunto de dados objetivando-se torná-la a mais natural possível, tendo em vista as peculiaridades do problema abordado. Idealmente, as representações não devem permitir a geração de soluções inválidas, facilitando a evolução. A representação adotada para uma possível solução (indivíduo) é denominada cromossomo [110].

Por sua vez, a decodificação do cromossomo permite que as aptidões das soluções codificadas sejam efetivamente avaliadas pelo problema. Este processo possibilita a construção da solução real do problema, fenótipo, a partir do cromossomo, genótipo. É importante garantir que cromossomos com características diferentes sejam decodificados em estruturas diferentes, entretanto é possível que cromossomos diferentes possuam aptidões iguais [108].

3.2. Avaliação

Este é o estágio do processo evolutivo que estabelece uma ligação direta entre o Algoritmo Genético e as características do problema a ser solucionado. A avaliação é o processo de associar um valor de aptidão a cada indivíduo, selecionado pelo AG. Na natureza, a aptidão de um indivíduo mede o quão bem adaptado ele está a um determinado ambiente.

Da mesma forma, no caso de AGs, a aptidão indica o quão bem um indivíduo da população atende aos requisitos de um problema. Normalmente, são atribuídos valores escalares, inteiros ou reais, associados ao grau de aptidão de cada indivíduo. Por sua vez, o cálculo da aptidão é feito por meio de uma função

de avaliação, que é uma função, ou procedimento, responsável por avaliar o grau de qualidade dos indivíduos (genótipos) presentes na população, considerando as características do problema abordado [111].

3.3. Seleção

Após a etapa de avaliação dos indivíduos, alguns deles são selecionados para a fase de reprodução, que é a base da evolução genética. O processo de seleção é baseado no princípio da sobrevivência dos indivíduos mais aptos, ou seja, os indivíduos com melhor aptidão possuem uma maior probabilidade de serem selecionados para reprodução [112].

Dessa forma, se privilegia a seleção dos indivíduos progenitores mais aptos para o cruzamento genético, e, conseqüentemente, a transmissão de seus genes para gerações futuras, as quais serão compostas majoritariamente por seus descendentes diretos [110].

Os principais mecanismos de seleção são:

- ❖ **Método da roleta:** é um método estocástico que consiste em associar os indivíduos da população a frações da área total de uma roleta, sendo que cada uma destas frações é proporcional à aptidão do indivíduo associado. Conseqüentemente, os indivíduos mais aptos têm maiores probabilidades de serem escolhidos. Na sequência, a roleta é girada várias vezes, sendo selecionados os indivíduos associados às frações sorteadas, em cada uma destas iterações. Como o tamanho da população é mantido constante ao longo de todo o processo evolutivo, na transição de uma geração para outra, a roleta deve ser girada um número de vezes igual ao tamanho da população. Devido às características da roleta, espera-se que as populações de cada nova geração sejam formadas por descendentes diretos dos indivíduos mais aptos da geração anterior e, conseqüentemente, que os menos aptos tenham dificuldades de transmitir seus genes para gerações futuras [101, 112]. Um exemplo de uma roleta com cinco indivíduos é apresentado na Figura 14. Percebe-se que, neste caso, o indivíduo 1 tem a maior aptidão, pois ocupa a maior fração da roleta,

tendo, por exemplo, uma probabilidade de seleção muito superior à do indivíduo 2, o qual é o menos apto dentre os cinco indivíduos considerados (menor fração da roleta).

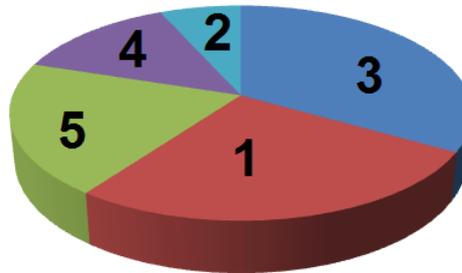


Figura 14 - Representação da seleção por roleta para 5 indivíduos.

- ❖ **Ordenação linear ou exponencial:** é a técnica pela qual os indivíduos da população são ordenados em função de suas respectivas aptidões, sendo atribuído a cada indivíduo um valor que corresponde à sua posição na população ordenada. Dessa forma, ao pior indivíduo é atribuído o valor 1 e ao melhor o valor N , onde N representa o tamanho da população. Na sequência, atribui-se a cada indivíduo uma probabilidade de seleção, calculada com base em uma dada distribuição, sendo as mais usuais a linear e a exponencial. Esta técnica evita a estagnação do processo evolutivo em gerações avançadas e a convergência prematura do algoritmo. Ao empregar este método, distribui-se de forma mais balanceada as probabilidades de seleção entre o conjunto de indivíduos mais aptos, reduzindo a pressão evolutiva exercida por um super indivíduo, isto é, um indivíduo cuja aptidão seja muito superior a aquela apresentada pelos demais membros da população [110, 113, 114].
- ❖ **Seleção por torneio:** consiste em escolher aleatoriamente uma determinada quantidade de indivíduos da população, denominada dimensão do torneio, e fazer um torneio entre eles. Em cada torneio comparam-se os valores de aptidão dos indivíduos participantes e seleciona-se como vencedor o mais apto, o qual irá participar da etapa de reprodução genética. O número de torneios realizados é igual ao número de indivíduos a serem selecionados, ou seja,

corresponde ao tamanho da população. Desde que a dimensão dos torneios seja pequena, esta técnica não induz convergências prematuras e combate a estagnação do processo evolutivo da população [101, 110].

3.4. Operadores Genéticos

Os operadores genéticos básicos se subdividem em dois grandes grupos:

❖ **Cruzamento:** O operador genético cruzamento, ou recombinação (*crossover*), consiste em efetuar trocas de genes entre dois indivíduos. Neste processo são gerados dois novos indivíduos (descendentes), resultantes da combinação de informação contida em um par de indivíduos progenitores. A evolução do AG, em busca de indivíduos mais aptos, é intimamente relacionada ao sucesso do cruzamento entre progenitores, sendo que deseja-se que esta operação seja capaz de gerar descendentes ainda mais aptos que os progenitores envolvidos no cruzamento. O operador cruzamento visa tirar proveito do material genético presente na população [110, 112]. Existem vários tipos de operadores de cruzamento, cuja seleção é afetada pelo tipo de representação usada na codificação dos indivíduos e pelas características do problema. Os operadores de cruzamento convencionais são [115]:

- ❖ Cruzamento com um ponto de corte;
- ❖ Cruzamento com dois pontos de corte;
- ❖ Cruzamento multiponto ou uniforme; e
- ❖ Cruzamento Heurístico.

O princípio de funcionamento do operador de cruzamento com um ponto de corte é ilustrado na Figura 15. Nesta operação, os descendentes são formados pela recombinação genética entre os indivíduos progenitores, em torno de um ponto de corte definido aleatoriamente.

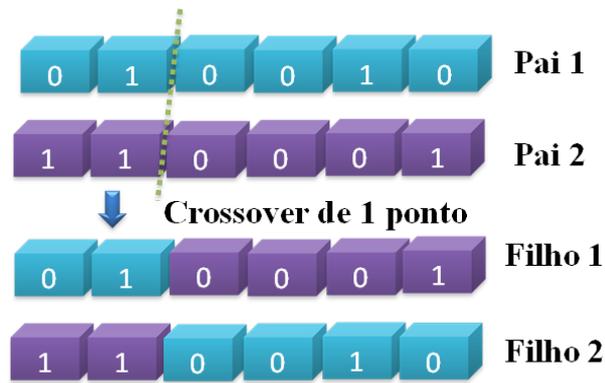


Figura 15 - Representação do cruzamento de um ponto.

No cruzamento com dois pontos de corte, ilustrado na Figura 16, os descendentes são formados pela recombinação genética dos progenitores, com base em dois pontos de corte selecionados de forma aleatória.

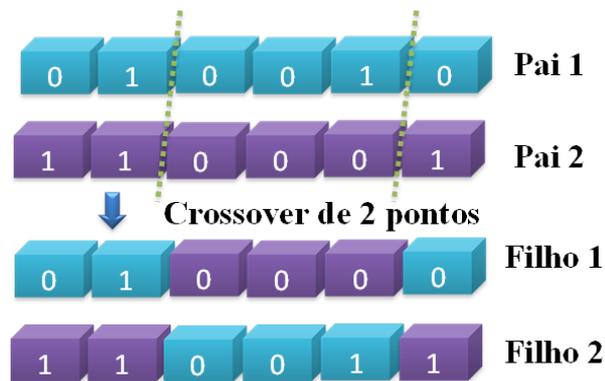


Figura 16 - Representação do cruzamento de dois pontos.

Por outro lado, no chamado cruzamento uniforme, representado na Figura 17, os descendentes são gerados por meio da recombinação genética dos progenitores, de acordo com valores estabelecidos em uma máscara binária gerada aleatoriamente [116].

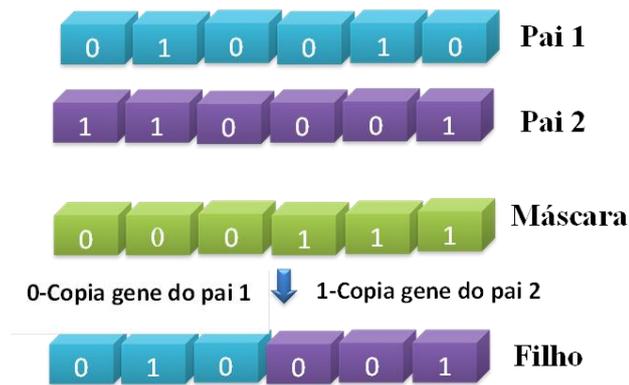


Figura 17 - Representação do cruzamento uniforme.

Por sua vez, a função de cruzamento heurístico (*Crossoverheuristic*) retorna uma descendência que se encontra na linha que contém os dois pais, a uma pequena distância do pai com melhor aptidão e, conseqüentemente, longe do pai menos apto. O parâmetro Raio (R) especifica o quão longe o filho está em relação ao melhor pai. Se o Pai1 e Pai2 são os pais, e Pai1 tem a melhor aptidão, o descendente gerado por esta função de cruzamento será [102, 117, 118]:

$$Filho = Pai2 + R * (Pai1 - Pai2) \quad (3)$$

- ❖ **Mutação:** O operador genético de mutação atua sobre os cromossomos gerados pelo operador cruzamento, possuindo uma probabilidade p_m de alterar um gene do cromossomo, selecionado aleatoriamente. A utilização deste operador genético em AGs permite que valores de genes, eventualmente perdidos durante o processo de evolução, possam retornar à população, possibilitando a avaliação dos mesmos em novos contextos populacionais. Por outro lado, tem-se que este operador também proporciona o surgimento de novos genes, não presentes nas populações iniciais, contribuindo para a diversidade da população. O operador de mutação introduz um certo grau de aleatoriedade no algoritmo genético, permitindo que mais pontos do espaço de busca sejam efetivamente avaliados. Um exemplo de mutação binária pode ser visualizado na Figura 18. A probabilidade de mutação, p_m , deve ser suficientemente baixa, a

fim de permitir a diversificação dos indivíduos da população, sem prejudicar a convergência do algoritmo [110-112].

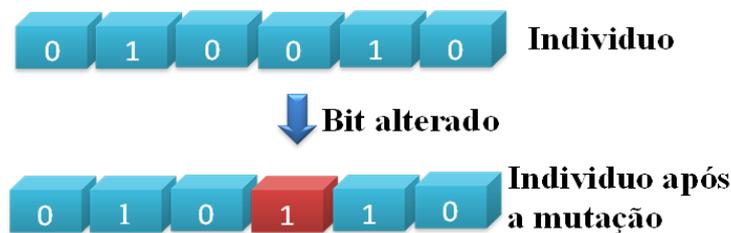


Figura 18 - Representação da operação genética de mutação.

3.5. Parâmetros da Evolução

Ao se implementar um AG, devem-se arbitrar valores adequados para seus parâmetros. A seleção de tais valores tem impacto direto na evolução do AG e, conseqüentemente, afeta a solução do problema analisado. A etapa de ajuste destes parâmetros é denominada *tuning*.

Diversas pesquisas indicam que o valor ótimo dos parâmetros de ajuste é dependente das características específicas do problema analisado. Tais estudos apontam ainda que, a cada estágio da evolução, idealmente, este conjunto de parâmetros deveria ser reajustado em função das características da população, visto que os AGs são um processo dinâmico [119].

Os parâmetros que mais influenciam o desempenho do AG são [106]:

- ❖ **Tamanho da População:** afeta o desempenho global e a eficiência dos AGs. Uma população suficientemente grande fornece uma melhor cobertura do domínio do problema e previne a convergência prematura para ótimos locais. Entretanto, populações muito grandes demandam um esforço computacional maior e, conseqüentemente, requerem um tempo maior de processamento. Logo, na definição do tamanho da população deve-se considerar tanto a diversidade populacional desejada, quanto o tempo de processamento tolerável, a fim de alcançar uma relação de compromisso satisfatória.
- ❖ **Taxa de Cruzamento:** define a probabilidade de um indivíduo ser recombinado com outro. Quanto maior for esta taxa, mais rapidamente novas estruturas serão introduzidas na população. Taxas muito altas

podem gerar efeitos indesejáveis, pois a maior parte da população será substituída (recombinada), a cada nova geração. Esta situação implicará em perda de variedade genética, podendo ocorrer perda de estruturas de alta aptidão e, conseqüentemente, convergência a uma população com indivíduos extremamente parecidos, dificultando a localização do ótimo global. Destaca-se que a adoção de técnicas de substituição de indivíduos baseadas no conceito do elitismo evita a perda de estruturas de alta aptidão. Por outro lado, taxas de cruzamento baixas fazem com que o algoritmo torne-se muito lento, aumentando o tempo de convergência para uma solução aceitável.

- ❖ **Taxa de Mutação:** define a probabilidade do conteúdo de um gene do cromossomo ser alterado. A taxa de mutação contribui para a diversidade populacional e previne a estagnação do processo evolutivo. Porém, deve-se evitar o emprego de taxas de mutação muito altas, pois nestes casos a busca pelo ponto ótimo se tornará majoritariamente aleatória, prejudicando fortemente a convergência para uma solução ótima.
- ❖ **Intervalo de Geração:** controla a porcentagem da população que será substituída a cada nova geração. Este parâmetro é afetado diretamente pelas técnicas de substituição de indivíduos apresentadas na seção 3.4. Quando a maior parte da população é substituída a cada nova geração, torna-se o AG mais rápido, porém, podem-se perder indivíduos com alta aptidão. Por outro lado, quando apenas uma pequena fração da população é substituída a cada nova geração, torna-se o AG mais lento, porém, evita-se a perda de indivíduos com alta aptidão.
- ❖ **Número de gerações:** representa o número máximo de ciclos de evolução de um Algoritmo Genético, sendo este um dos critérios de parada adotados por AGs. Um número de gerações muito pequeno causa uma queda no desempenho, pois pode-se fazer com que o ciclo evolutivo do AG seja interrompido prematuramente e, conseqüentemente, a solução obtida será sub-ótima. Por outro lado, valores grandes demandam um tempo maior de processamento, mas possibilitam uma melhor avaliação do espaço de busca do problema, evitando a convergência para soluções locais.
- ❖ **Convergência da Função de Avaliação:** pode ser empregada como critério de parada, a fim de evitar que o processamento consuma tempos

computacionais elevados. Quando não ocorre melhoria significativa na solução durante um dado número de gerações, o algoritmo de processamento é automaticamente interrompido.

- ❖ **Número de rodadas:** representa o número total de vezes que um AG será executado. O AG é uma técnica probabilística, sendo que, em alguns casos, pode ser interessante executá-lo mais de uma vez, a fim de se tentar obter melhores soluções. As novas rodadas não estão presas aos ótimos locais encontrados em rodadas anteriores. Dessa forma, serão avaliados outros pontos do espaço de busca, os quais poderão implicar na obtenção de melhores soluções. Arbitrando-se um tempo de processamento máximo, deve-se definir um número de rodadas suficientemente pequeno, para permitir que o algoritmo tire proveito da informação genética dos indivíduos da população, a cada rodada. Por outro lado, o número de rodadas deve ser grande o suficiente para permitir que novas rodadas, com novos indivíduos, sejam iniciadas quando as rodadas anteriores convergirem para ótimos locais, possibilitando uma melhor cobertura do domínio do problema.
- ❖ **Taxa de semeadura:** representa a porcentagem dos melhores indivíduos de uma rodada que serão utilizados como sementes na população inicial da rodada seguinte. Uma taxa de semeadura muito pequena faz com que a rodada seguinte tire pouco proveito dos resultados obtidos na rodada anterior. Por outro lado, valores grandes fazem com que as novas rodadas não consigam sair do mínimo local obtido em rodadas anteriores.