

Instituto politécnico da Guarda

Escola Superior de Tecnologia e Gestão

**Relatório**

**Trabalho Prático 2 – Algoritmos Genéticos**

Inteligência Artificial 2020 - 21

Dário Ribeiro 1012208

## ABSTRACT

A ilusão do poder computacional ilimitado não ficou confinada aos programas de resolução de problemas. As primeiras experiências de evolução automática (denominados de algoritmos genéticos) baseavam-se na convicção sem dúvida correta de que, realizando-se uma série apropriada de pequenas mutações num programa em código máquina, seria possível gerar um programa com bom desempenho para qualquer tarefa simples.

Então, a ideia era experimentar mutações aleatórias com um processo de seleção para preservar mutações que parecessem úteis. Apesar de milhares de horas de tempo de CPU, quase nenhum progresso foi demonstrado. Os algoritmos genéticos modernos utilizam representações melhores e têm mais sucesso.

Este Relatório visa explicar o funcionamento de um algoritmo genético através da demonstração de um trabalho prático.

Adicionalmente foi efetuado um script para agilizar o processo.

**Palavras Chaves:** Algoritmo Genético – Mutação – Recombinação – Resolução de Problemas

CONTEUDO

[ABSTRACT 1](#_Toc62141019)

[INTRODUÇÃO 2](#_Toc62141020)

[FUNDAMENTO TEÓRICO 6](#_Toc62141021)

[EXPERIÊNCIA 6](#_Toc62141022)

[CONCLUSÕES 6](#_Toc62141023)

## INTRODUÇÃO

### Algoritmo Genético

Um algoritmo genético (ou AG) é um método de procura do tipo estocástico na qual os indivíduos sucessores são gerados pela combinação de dois indivíduos pais, em vez de serem gerados pela modificação de um único indivíduo. A analogia em relação à seleção natural é a mesma que se dá neste método sendo ele com o tipo de reprodução sexuada tal como ocorre na natureza, podemos então usando uma analogia, concluir que é uma ferramenta para estudar o fenómeno da adaptação dos seres vivos.

Os AG provaram ser ferramentas poderosas quando aplicados à classe de problemas referidos

* Escalonamento
* Controlo Adaptativo
* Jogos
* Otimização de Funções Matemáticas
* Otimização Combinatória

### Procura Estocástica

Um método de procura otimizado que em vez de escolher o melhor k a partir do conjunto de sucessores candidatos, a procura estocástica escolhe k sucessores de forma aleatória, com a probabilidade de escolher um determinado sucessor que seja uma função crescente de seu valor.

A procura estocástica guarda alguma semelhança com o processo de seleção natural, pelo qual os “sucessores” (descendência) de um “indivíduo” (organismo) ocupam a próxima geração de acordo com o seu “valor” (avaliação)

### Análise Teórica

Os Algoritmos Genéticos começam com um conjunto de k indivíduos gerados aleatoriamente, chamado população.

Cada indivíduo, é representado como uma cadeia sobre um alfabeto finito — muito frequentemente, uma cadeia de valores 0 e 1, denominamos a essa cadeia usando a analogia do processo natural, cromossoma.

O processo de reprodução consiste na troca de informação entre os progenitores e é designado por recombinação

Os novos indivíduos assim gerados podem ser alterados de forma localizada através dos efeitos de outro operador, conhecido por mutação

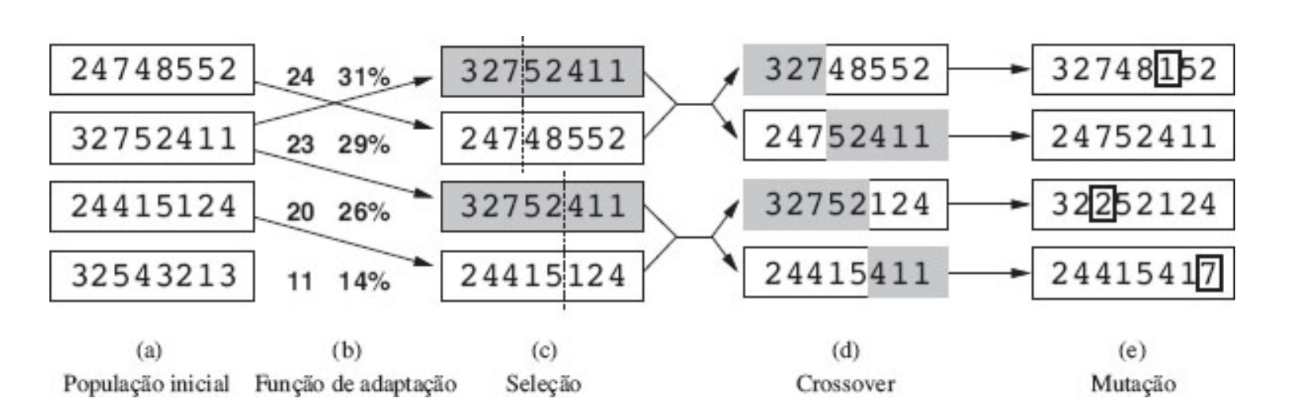
Estes 2 operadores genéticos (recombinação e mutação) permitem manter a diversidade da população e evitar que o AG convirja prematuramente para um máximo local

Figura - Funcionamento do AG

O Algoritmo genético, ilustrado na Figura 1 por sequências de dígitos que representam os indivíduos tem uma população inicial em (a) é classificada pela função de avaliação em (b), resultando em pares de correspondência em (c). Eles produzem descendentes em (d) através da recombinação e são sujeitos à mutação em (e).

A produção da próxima geração de indivíduos é mostrada na Figura 1, desde (b) até (e).

Em (b), cada estado é avaliado pela função de avaliação ou (na terminologia do AG) pela função de adaptação. Uma função de avaliação deve retornar valores mais altos para indivíduos melhores

Nesta variante específica do algoritmo genético, a probabilidade de um indivíduo ser escolhido para reprodução é diretamente proporcional ao valor da função de avaliação, e as percentagens são mostradas, caracterizando-se assim pela probabilidade de ser selecionada.

Em (c), dois pares escolhidos aleatoriamente são selecionados para reprodução, de acordo com as probabilidades mostradas em (b).

De notar que um indivíduo é selecionado duas vezes, e um indivíduo não é selecionado de modo algum.

Para cada par a ser cruzado, é escolhido ao acaso um ponto de cruzamento dentre as posições no cromossoma. Na Figura 1, os pontos de cruzamento estão depois do terceiro dígito no primeiro par e depois do quinto dígito no segundo par.

Em (d), os próprios descendentes são criados por cruzamento dos cromossomas pais no ponto de crossover (recombinação).

Por exemplo, o primeiro filho do primeiro par recebe os três primeiros dígitos do primeiro pai e os dígitos restantes do segundo pai, enquanto o segundo filho recebe os três primeiros dígitos do segundo pai e o restante do primeiro pai.

Finalmente, em (e), cada posição está sujeita à mutação aleatória com uma pequena probabilidade independente. Um dígito sofreu mutação no primeiro, no terceiro e no quarto descendente.

Depois da Mutação efetuada o algoritmo está completo, caso se queira efetuar gerações até que o valor da função de avaliação esteja dentro de parâmetros aceitáveis, os cromossomas que foram mutados, tornam-se assim a população da próxima geração.

## ANÁLISE PRÁTICA

Para o exercício e aprendizagem deste algoritmo genético foi-se proposto o seguinte problema.

Uma imagem com texto

Descrição gerada automaticamente

Figura - Exercício Prático

### Dados

Através do problema exposto na Figura 2 podemos observar que nos é dado uma função matemática f(x), pelo que vai ser caracterizada como a nossa função de avaliação.

Equação - Função de Avaliação

Portanto dado a nossa função de avaliação, é preciso determinar o tamanho dos cromossomas, que é baseado no domínio do X.

Começamos por contar a amplitude do intervalo

Equação - Amplitude de um intervalo

A nossa amplitude vai ser então 24, caso seja difícil efetuar as contas com um intervalo de diferentes valores a Figura 3 mostra de forma visual como podemos determinar o tamanho do domínio.

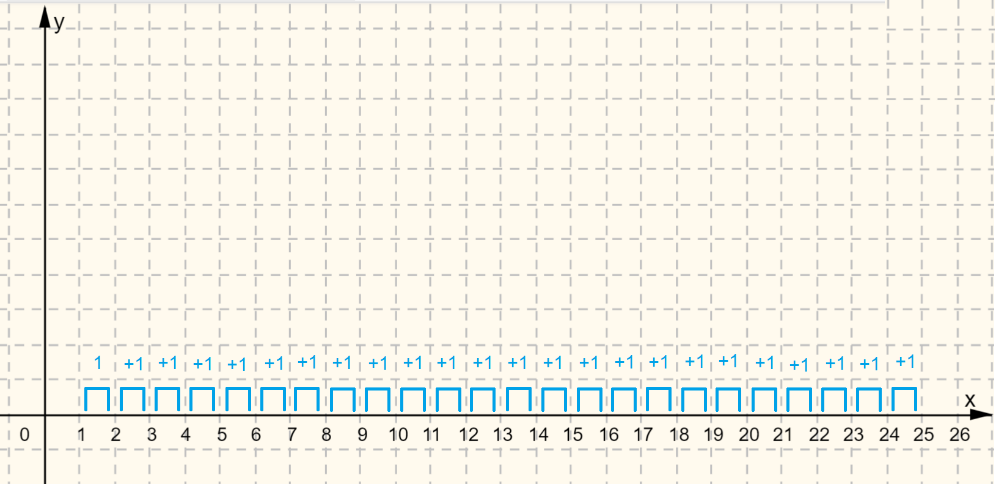


Figura - Referencial Cartesiano com contagem do domínio

Depois de determinada a nossa amplitude precisamos de definir a precisão dos valores da função de avaliação, pois a precisão influência na medida que precisamos de determinar o tamanho do cromossoma em binário. Neste exercício prático vamos definir uma precisão de 4 casas decimais

É necessário dividir a variável x em partes iguais por todo o domínio, portanto primeiro determina-se o tamanho total de x

Equação - Número de Partes

Aplicando os dados reunidos acima fica então:

Equação - Número de partes com os dados

É necessário saber o número de bits para os cromossomas e para isso temos de descobrir onde fica compreendido o valor 240 000 na nossa base binária (Base 2).

Se:

Então tentamos mais ou menos de forma aproximada descobrir onde fica compreendido o valor 240 000.

Se: Então:

Neste algoritmo genético em termos de intervalo, escolhe-se o valor superior, fazendo com que então o nosso cromossoma seja representado com 18 bits de comprimento.

É também preciso determinar o tamanho da população e por regra geral e de bom senso, para se obter bons resultados com uma boa precisão e eficácia, no mínimo 10 indivíduos são precisos. Neste trabalho prático vai ser usado uma população consistida por 20 indivíduos.

Outros dados importantes a considerar é o caso da seleção para a recombinação, ou seja, é necessário definir um valor em percentagem que defina o limite mínimo aceitável, e aquando a função de avaliação, o individuo pode não ser selecionado para a recombinação.

Na Figura 1 pode-se observar que um individuo não foi selecionado, isto porque o cálculo em percentagem da probabilidade de ser selecionado para a recombinação foi baixo (cerca de 14%), os restantes indivíduos foram selecionados porque o resultado do cálculo foi acima de 14%, portanto pode-se concluir que existe um valor que limita os que podem ser selecionados.

No próximo subcapítulo vai ser explicado com mais detalhe o cálculo da seleção para determinar que indivíduos podem ser recombinados.

Para o Trabalho em questão vai ser usado como valor de seleção, 65%, ou seja, acima de 65% os indivíduos não vão ser recombinados

A ideia consiste em tentar melhorar o valor da função de avaliação, de forma a aperfeiçoar o algoritmo, fazendo uma analogia humana, é como se certos indivíduos não tivessem capacidade de aprendizagem escolar igual à maioria da população e necessitassem de umas vitaminas para poderem melhorar a performance escolar, essas vitaminas, são o operador de recombinação.

Tal como acontece para a recombinação, existe um operador na mutação, através de um cálculo que vai ser demonstrado no próximo subcapítulo, precisamos de definir o número de mutações que também vai ser o número de cromossomas que vão sofrer de tal.

Como a Figura 1 demonstra em (e), nem todos os indivíduos sofreram mutação.

Para o Trabalho em questão vai ser usado como valor para o operador de mutação, 0.5%, ou seja, da população total, apenas vão haver 0.5% mutações.

Dado que temos o tamanho do cromossoma, tamanho da população, a fórmula para determinar o número de indivíduos a sofrerem mutação é a seguinte

Equação - Número de Mutações

Para uma visualização mais leiga, a ideia consiste em imaginar uma matriz com os cromossomas como demonstra a Figura 4

Uma imagem com texto

Descrição gerada automaticamente

Figura - Cromossomas

Dentro dessa matriz imaginária, é preciso de forma aleatória determinar quantos vão sofrer mutação, e quantas mutações se vão realizar. Estas duas variáveis a serem determinadas são diretamente proporcionais, quer isto dizer que se por exemplo do cálculo resulta 4 mutações, quer então dizer que serão 4 indivíduos que vão sofrer mutação.

Uma imagem com mesa

Descrição gerada automaticamente

Figura - Cromossomas Selecionados

Podemos então concluir os seguintes dados para a elaboração do trabalho:

* Tamanho População: 20
* Precisão em casas decimais: 4
* Tamanho Cromossoma: 18 bits
* Valor da seleção para recombinação: 65%
* Valor da seleção para mutação: 0.5%

### Algoritmo

#### 1º Passo – Indivíduos

Para a geração inicial, é necessário gerar de forma aleatória os indivíduos para se ter uma base ou ponto de partida com que se possa trabalhar.

Em Microsoft Excel existe uma fórmula de gerar aleatoriamente os indivíduos.

Uma imagem com mesa

Descrição gerada automaticamente

Figura - Função Aleatório Excel

O Excel possui duas funções ligadas ao aleatório, existe o aleatório simples que retorna um número aleatório sem qualquer especificação de intervalo tanto pode dar 0 como 9999999 e existe o aleatório entre que permite especificar um intervalo onde vai gerar aleatoriamente, como o próprio nome indica.

Portanto como definimos um tamanho de 18 bits em binário, se gerarmos aleatoriamente um valor sem qualquer especificação, por exemplo 789403:

Equação - Conversão Decimal para Binário

Se contarmos os bits pode-se verificar que são 20, e necessitamos apenas de 18, portanto é necessário usar a função ***aleatórioentre*** ao invés da função ***aleatório***.

Ao usarmos a função aleatório entre temos de especificar o intervalo pelo que 0 é o nosso número mínimo, e o 262 144‬ ou (2 elevado a 18) o nosso máximo.

0 = 000000000000000000 (cromossoma mínimo)

262 144 = 111111111111111111 (cromossoma máximo)

Depois de definirmos a função de forma aleatória, pode-se converter para binário usando a função base que aceita 3 parâmetros por ordem:

* Número decimal a ser convertido
* Base de conversão
* Tamanho do número em base que queremos converter

Uma imagem com mesa

Descrição gerada automaticamente

Figura - Função Base em Excel

Como queremos os cromossomas de tamanho 18, queremos base 2 , são esses os dois parâmetros que se pode verificar na Figura 7 na barra de fórmulas.

#### 2º Passo – Função de avaliação

Observando a Função de avaliação podemos verificar que tudo indica que poderíamos usar o número decimal como o nosso x, mas é necessário calcular o valor real do número, tanto seja binário ou decimal e daí parte então a nossa variável x.

Equação - Função de Avaliação

A fórmula para calcular o valor real de um dado número é a seguinte.

Equação - Função de Cálculo do Valor Real

Exemplificando para o valor decimal da primeira linha da tabela presente na Figura 7 o valor real será:

Equação - Cálculo do Valor Real com os dados

Aplicando para a Tabela fica então o seguinte:

Uma imagem com mesa

Descrição gerada automaticamente

Figura - Tabela em Excel com o valor real

Dado que agora se sabe o valor real de cada indivíduo, pode se então efetuar o cálculo da função de avaliação para cada indivíduo, a equação abaixo demonstra para a primeira linha do Excel.

Na Tabela fica então:

Uma imagem com mesa

Descrição gerada automaticamente

Figura - Tabela Com a função de Avaliação

#### 3º Passo – Roleta

Dado que agora temos a função de avaliação, para a recombinação vão ser usado dois métodos de seleção.

Método Elitista - (Seleciona o melhor individuo com base no valor da função de avaliação)

Método da Roleta - (cada individuo vai ser colocado numa “roleta” baseado na probabilidade de ser selecionado)

A razão pela qual se estão a usar dois métodos ao invés de um é porque o método elitista está a ser usado como um complemento do método da roleta, e estamos a usar um complemento para garantirmos que os melhores indivíduos da função de avaliação não se percam ao longo da geração, porque o AG é um algoritmo de otimização que procura sempre aperfeiçoar, sem este complemento, pode haver o risco de se perder indivíduos com boa qualidade.

No método elitista escolhe -se os dois melhores indivíduos com base no valor resultante da função de avaliação, ou seja, já passaram para a próxima geração. Os restantes indivíduos serão escolhidos pelo método da roleta, para sofrerem uma recombinação.

O método da roleta, consiste em dois cálculos importantes, é preciso calcular a probabilidade de um individuo ser selecionado, e depois “integrar” no segmento da roleta correspondente, quer isto dizer o quê.

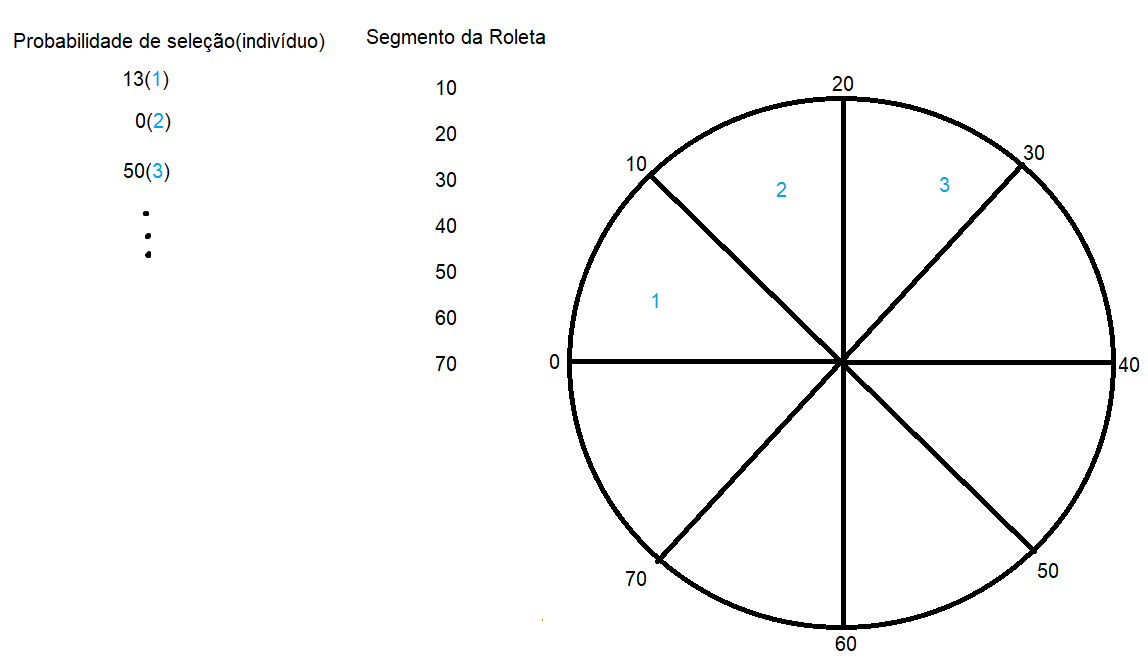


Figura - Exemplificação rudimentar da roleta

Na Figura 10 pode se verificar a roleta, e a ideia não é nada mais do que calcular a probabilidade de ser selecionado, e categorizar de forma ordenada dentro da roleta, cada individuo, para depois de forma aleatória escolher um segmento da roleta e obter o indivíduo, tal como funciona a roda da sorte ou fortuna nos concursos de ***game show.***

A fórmula para calcular a probabilidade de ser selecionado é a seguinte

Equação - Probabilidade de Seleção

Para o primeiro valor da tabela temos então:

Para o resto da tabela temos

Uma imagem com mesa

Descrição gerada automaticamente

Figura - Probabilidade de Seleção

De notar que na barra de fórmulas do Excel pode-se notar que existe a célula que representa o f(xi) a dividir com outra célula com o caractere monetário Dólar, isto é para que para todas as linhas da coluna atual, divida o valor da coluna anterior com a soma dessa dita coluna anterior, representada na célula D22.

## EXPERIÊNCIA

## CONCLUSÕES