

Universidade Federal do Rio de Janeiro

COPPE

Programa de Engenharia Elétrica - PEE

Disciplina: Otimização Natural

Aluno: Gustavo Martins da Silva Nunes

Professor: José Gabriel

Data: 24/05/2016

## Lista 7 - Resolução

## Questão 1

**Explique como o desempenho do algoritmo genético simples utilizado no Exercício 1 da Lista de Exercícios #4 pode ser melhorado através da utilização de métodos de algoritmos meméticos.**

O Exercício 1 da Lista #4 consistia em minimizar a função  $y(x) = x^2 - 0.3\cos(10\pi x)$ , com  $x \in [-2, 2]$ , fazendo uso de um algoritmo genético simples que utiliza uma representação de, no mínimo, 16 bits. É sabido que o mínimo dessa função ocorre quando  $x_{min} = 0$ , com  $y(x_{min}) = -0.3$ . Os algoritmos meméticos consistem em utilizar algum conhecimento prévio sobre o problema ou técnicas de otimização adequadas para tal, de modo a evitar que o algoritmo genético efetue uma busca totalmente aleatória sobre o espaço de soluções pela solução ótima. Tais conhecimentos podem ser introduzidos em quaisquer passos do algoritmo genético tradicional, desde a inicialização da população (que deixaria de ser totalmente aleatória), até a seleção de sobreviventes.

A representação adotada para tal problema tinha sido a seguinte: cada genótipo, contendo 20 bits, é segmentado em 5 trechos, cada um com 4 bits. Assumindo que  $x \in \mathbb{Z}$ , tem-se que  $x \in -2, -1, 0, 1, 2$ . Cada segmento representa um desses possíveis valores de  $x$ , de modo que os 4 bits menos significativos representam o valor -2, enquanto que os mais significativos, o valor 2. A decodificação era feita, então, contando a quantidade de bits 1's presentes nos trechos e, aquele com a maior contagem, decodificaria  $x$  de acordo com o valor que ele representa. Em caso de empate, a prioridade é dada para o trecho que representa o maior valor (ou seja, o trecho que representa o valor 2 tem prioridade sobre os demais, e assim por diante). No caso do genótipo ser todo composto por bits de valor 0,  $x$  é decodificado no valor de -2.

Nesse exemplo, como o mínimo é conhecido ( $x = 0$ ), dada essa representação, uma forma de acrescentar esse conhecimento no algoritmo seria, por exemplo, efetuar recombinações e mutações somente sobre os segmentos de número 3, 4 e 5, os quais representam, respectivamente, os valores  $x = 0, 1, 2$ , visando maximizar a contagem de bits de valor 1 presentes no segmento 3 e minimizá-la nos demais segmentos. Além disso, após a aplicação desses operadores, seria possível realizar uma busca local nas soluções encontradas, de modo a se aproximar ainda mais da solução ótima. As melhorias encontradas pela busca local podem ser incorporadas no genótipo dos indivíduos melhorados (abordagem de Lamarck) ou não, sendo utilizadas, somente, no cálculo da aptidão melhores para eles (abordagem de Baldwin). Ambas as abordagens citadas anteriormente favorecem a preservação de boas soluções e a geração de novas a partir delas. Uma outra abordagem seria inicializar a população próxima do mínimo global (o que corresponderia a inicializar o trecho 3 dos genótipos com uma grande

quantidade de bits no valor 1).

Como visto, as possibilidades de incorporar conhecimento sobre o problema para auxiliar o algoritmo são diversas. O que se espera, a partir dessas abordagens, é que o algoritmo encontre as soluções ótimas de forma mais rápida, requerendo menos iterações e, portanto, melhorando o seu desempenho com relação ao tempo de execução. No entanto, a incorporação de tais conhecimentos gera um outro problema: a convergência prematura do algoritmo. Como os algoritmos meméticos geralmente incluem, em alguma etapa do algoritmo evolucionário em questão, uma busca local, de forma a gerar indivíduos localmente ótimos, isso faz com que os indivíduos fiquem presos nesses ótimos. A diversidade de soluções presentes na população é, então, perdida, sendo ainda mais crítico, quando todas as soluções convergem para um mesmo ótimo (o que só não seria problemático caso houvesse, somente, o ótimo global). A aplicação da abordagem memética deve ser feita com cuidado, para evitar esse efeito indesejado. Sendo assim, mecanismos de preservação de diversidade da população devem ser empregados, tais como, por exemplo, *Fitness Sharing* ou *Deterministic Crowding*, com o objetivo de permitir que o algoritmo evolucionário encontre novas (e boas) soluções para o problema.

## Questão 2

**Comente sobre as variáveis presentes na Equação (11.1) do livro-texto, descrevendo sucintamente os seus significados e a sua variação entre uma geração e a geração seguinte.**

A Equação (11.1) do livro-texto é apresentada a seguir:

$$m(H, t + 1) \geq m(H, t) \cdot \frac{f(H)}{\langle f(H) \rangle} \cdot \left[ 1 - \left( p_c \cdot \frac{d(H)}{l - 1} \right) \right] \cdot [1 - p_m \cdot o(H)]$$

Os termos presentes nessa equação são descritos a seguir:

.  $m(H, t + 1)$  e  $m(H, t)$

As variáveis  $m(H, t + 1)$  e  $m(H, t)$  representam, respectivamente, a proporção de indivíduos de um determinado esquema  $H$  presentes nas gerações  $t + 1$  e  $t$ . O esquema  $H$  é definido como um determinado padrão de bits encontrado na população. A estrutura de um esquema contém 3

possíveis valores: 0, 1 ou # (referido como *don't care*). Para exemplificar, o esquema  $H_1 = #####$  (assumindo um genótipo de tamanho 6) contém a população inteira, ao passo que o esquema  $H_2 = 010001$  contém somente um indivíduo e o esquema  $H_3 = 01##1#$  contém todos os indivíduos, cujo genótipo apresenta essa forma (ou seja, o primeiro bit valendo 0 e os bits 2 e 5 valendo 1). Pensando nos esquemas como conjuntos, a variável  $n(H_i, t)$  corresponde à cardinalidade do conjunto  $H_i$ , isto é, a quantidade de elementos (indivíduos) que o compõe, na geração  $t$ . A variável  $m(H_i, t)$ , por sua vez, é a proporção de indivíduos desse esquema na população total:  $m(H_i, t) = \frac{n(H_i, t)}{\mu}$  ( $\mu$  é a quantidade de indivíduos presentes na população).

- $f(H)$  e  $\langle f(H) \rangle$

A variável  $f(H)$  representa o valor de fitness do esquema  $H$ , o qual é definido pela média dos valores de fitness de todos os indivíduos que pertencem ao esquema em questão. Já a variável  $\langle f(H) \rangle$  representa o valor médio de fitness de toda a população. Quando a fitness média do esquema é superior à fitness média da população (indicando que os indivíduos desse esquema tendem a ter uma aptidão melhor), a razão  $\frac{f(H)}{\langle f(H) \rangle}$  é maior do que 1. Supondo que não houvesse recombinação ou mutação que destruísse o esquema  $H$  em questão, então, pela equação (11.1),  $m(H, t+1) \geq m(H, t)$ , o que é esperado, já que indivíduos com fitness melhores tendem a dominar a população, devido às maiores chances de reprodução e sobrevivência, aumentando, portanto, em gerações subsequentes, sua proporção na população total (até, eventualmente, dominá-la totalmente). O mesmo raciocínio é análogo para esquemas, cuja razão  $\frac{f(H)}{\langle f(H) \rangle}$  é menor que 1, indicando que indivíduos pertencentes a tal esquema são, em média, menos aptos do que o restante da população (e, com isso, tendem a desaparecer, supondo que a recombinação e a mutação não alterasse tais esquemas).

- $p_c$ ,  $d(H)$  e  $l$

As variáveis  $p_c$ ,  $d(H)$  e  $l$  correspondem, respectivamente, à probabilidade de haver recombinação, à *distância definida* do esquema (definida como a distância entre o bit definido (0 ou 1) mais à esquerda e o mais à direita do esquema) e ao comprimento total do genótipo de um indivíduo. Para exemplificar, a distância definida  $d(H_1)$  do esquema  $H_1 = 0##11#$  é igual a 5 (primeiro bit sendo o bit definido mais à esquerda e o quinto bit sendo o bit definido mais à direita). O produto  $p_c \cdot \frac{d(H)}{l-1}$  define a probabilidade de que ocorra uma recombinação  $p_c$  e de que tal recombinação destruirá o esquema  $H$  em questão (dada pela razão  $\frac{d(H)}{l-1}$ , que pode ser pensada como uma probabilidade também). Caso a recombinação destrua o esquema em questão, então, a proporção desse esquema na geração

seguinte será diferente do esperado, se fosse levado em conta somente a fitness média dos indivíduos desse esquema em relação à da população, conforme explicado anteriormente. O termo  $\left[1 - \left(p_c \cdot \frac{d(H)}{l-1}\right)\right]$  representa, portanto, a probabilidade de que uma recombinação *não* destruirá o esquema em questão.

•  $p_m$  e  $o(H)$

As variáveis  $p_m$  e  $o(H)$  representam, respectivamente, a probabilidade de haver mutação e a *ordem* do esquema  $H$  em questão, definida como a quantidade de bits definidos no esquema  $H$ . Com isso, por exemplo,  $H_1 = 011\#\#1$  e  $H_2 = 010011$  possuem, respectivamente,  $o(H_1) = 4$  e  $o(H_2) = 6$ . A probabilidade de haver mutação  $e$  de que essa mutação destruirá o esquema  $H$  em questão é dada por  $1 - (1 - p_m)^{o(H)}$ . Se  $p_m \ll 1$ , pode-se aproximar  $(1 - p_m)^{o(H)} \approx (1 - o(H) \cdot p_m)$  e, nesse caso, a probabilidade de uma mutação destruir o esquema é dada por  $p_m \cdot o(H)$ . Portanto, o termo  $[1 - p_m \cdot o(H)]$  corresponde à probabilidade de a mutação *não* destruir o esquema  $H$  em questão.