

GAMO - Um algoritmo genético componente de modelos de gestão (otimização-simulação)

Júlio F. Ferreira da Silva

*Doutor em Engenharia Civil - Hidráulica Prof. Auxiliar do Departamento de Eng^o Civil da Universidade do Minho,
Azurém 4800-058 Guimarães, Portugal 253510200, juliofs@civil.uminho.pt*

RESUMO

Descreve-se o algoritmo genético (A.G.) concebido para ser uma componente de modelos de gestão. Dos desenvolvimentos constam, por exemplo, o alargamento do espaço de soluções iniciais, o ajuste diminutivo do tamanho da população ao longo das gerações e a rectificação das soluções geradas aleatoriamente, usando como critério o interesse económico e informação do comportamento do sistema físico.

Palavras-chave - Optimização global. Algoritmos genéticos. Modelos de Optimização-Simulação.

1. INTRODUÇÃO

Um algoritmo genético (A.G.) base revela-se como um modelo simples da evolução e não reflecte as teses mais recentes da genética e da embriologia e, mesmo, os algoritmos mais avançados, entretanto desenvolvidos, para obterem bons desempenhos continuam a lançar mão de procedimentos, baseados na geração de alterações aleatórias sem qualquer mecanismo de controlo. Os algoritmos genéticos debatem-se com um dilema que se relaciona com o antagonismo entre exploração exaustiva do espaço de soluções e o aproveitamento oportuno das melhores. Este último aspecto torna-se especialmente evidente se o algoritmo, para proceder à avaliação de cada uma das eventuais soluções, tiver que recorrer a modelos de simulação de sistemas físicos. Da análise dos principais mecanismos usados pelos seres vivos para transmitir informação aos seus descendentes ressalta a existência de mecanismos de controlo que permitem a preservação das características da espécie e a forte rejeição de alterações puramente aleatórias. Neste trabalho descrever-se-á o algoritmo genético que concebido como componente de modelos de gestão (associação da técnica de optimização com modelos de simulação de sistemas físicos) recorre ao conceito de reparação das soluções aleatórias. Explicar-se-ão, também, outros desenvolvimentos que resultaram em melhorias no desempenho do A.G..

2. DESCRIÇÃO GERAL DO ALGORITMO GENÉTICO DESENVOLVIDO

Integrando a técnica de optimização num modelo de gestão todas as eventuais soluções têm de ser testadas pelos modelos de simulação dos sistemas físicos. A técnica de optimização passou a ser um subprograma do modelo global de gestão e de apoio à decisão.

Do trabalho realizado desta-se, em jeito de resumo, as seguintes facetas: 1 - Enquadramento da técnica de optimização no modelo de gestão multi-etéapico; 2 - Definição das soluções iniciais; 3 - Manipulação das restrições; 4 - Rectificação ou reparação das soluções geradas aleatoriamente, usando como critérios o interesse económico e o comportamento do sistema físico; 5 - Ajuste diminutivo do tamanho da população e do número de gerações; 6 - Penalização crescente das soluções não cumpridoras das restrições; 7 - Probabilidades de cruzamento e de mutação crescentes com o número de gerações.

O trabalho desenvolvido resultou no algoritmo genético que está esquematizado na figura seguinte:

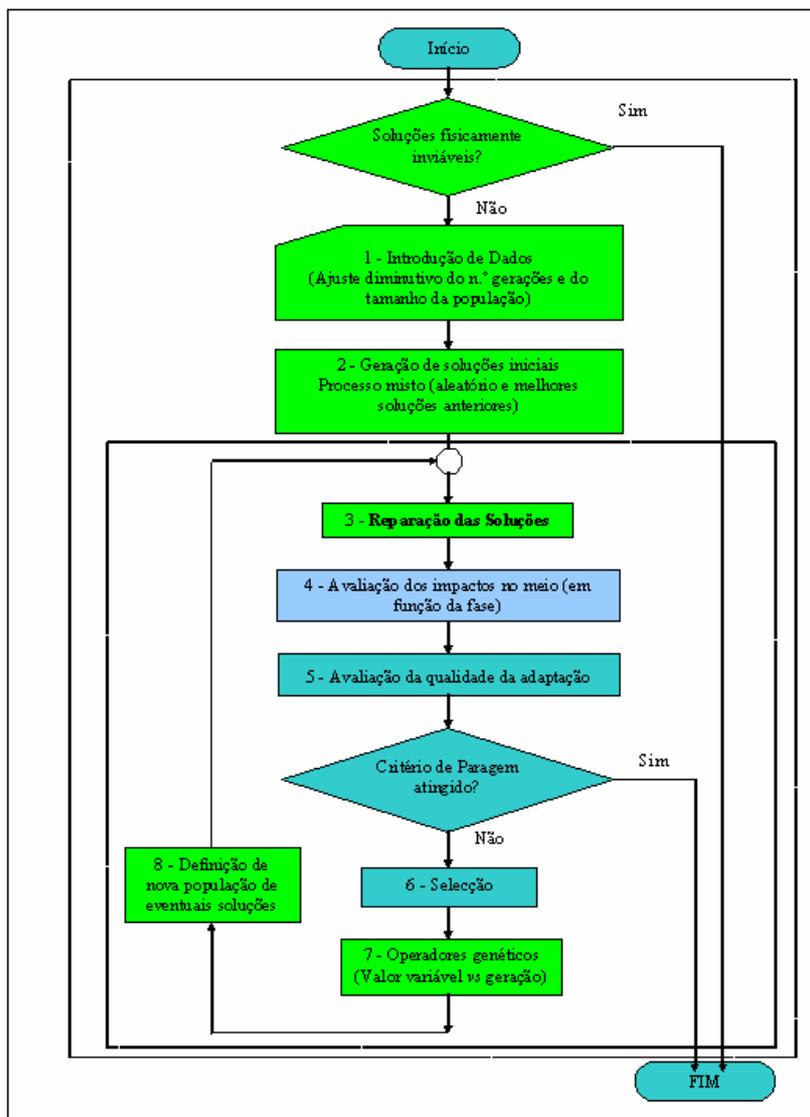


Figura 1 - Algoritmo genético desenvolvido

3 - ENQUADRAMENTO DA TÉCNICA DE OPTIMIZAÇÃO NO MODELO DE GESTÃO

O algoritmo genético desenvolvido procurou minimizar as dificuldades inerentes à junção de uma técnica de optimização com modelos de simulação complexos e forçosamente exigentes em tempos de execução. A formulação de um modelo global de optimização-simulação arrasta exigências específicas. Esta integração de diversas ferramentas implicou diversas alterações no algoritmo da técnica de optimização para melhorar a eficiência do modelo global. Por exemplo, o modelo global de gestão permite que o decisor indique um intervalo de valores para as variáveis de estado. Por vezes, as condições vigentes no sistema físico não permitem a existência alguns dos valores indicados pelo decisor. Nestas situações, não faz sentido que as técnicas de optimização e de simulação do sistema físico procurem e testem soluções que fisicamente não existem. Então, deve ser acrescentada esta verificação preliminar do comportamento do sistema e alterado o algoritmo da técnica de optimização para parar imediatamente a geração de eventuais soluções e não correr obrigatoriamente todas as gerações inicialmente estipuladas.

4 - DEFINIÇÃO DAS SOLUÇÕES INICIAIS

A estratégia de resolução de problemas por etapas sugere de imediato a ideia da constituição de uma população inicial para o algoritmo genético que inclua as melhores soluções da etapa anterior, ao invés do processo ser puramente aleatório. O agente decisor poderá estar interessado em conhecer o resultado de determinada política que tem sido seguida ou ter conhecimento preliminar do funcionamento do sistema físico. As ferramentas de otimização baseadas no conceito de gradiente, mesmo com diferenciação numérica, obtêm soluções dependentes dos pontos de partida. No entanto, estes resultados podem ser usados como soluções iniciais para o algoritmo genético. Admitindo que as soluções preliminares estão perto da solução ótima global, então os valores atribuídos pelo processo aleatório do algoritmo genético podem ser corrigidos para que sejam respeitadas as restrições do problema e, também, para que os valores a simular e a avaliar estejam próximos das soluções preliminares.

Podem calcular-se o peso da cada variável na solução preliminar e das geradas aleatoriamente pelo algoritmo genético, respectivamente, através de:

$$\pi l_i = \frac{Ql_i}{\sum_{i=1}^n Ql_i} \quad \text{e} \quad \pi a_i = \frac{Qa_i}{\sum_{i=1}^n Qa_i} \quad (1)$$

em que: πl_i - peso da variável i na solução preliminar, Ql_i - variável i da solução preliminar; πa_i - peso da variável i na solução definida aleatoriamente pelo algoritmo genético.

Para preservar a indicação da primeira fase dever-se-á definir uma nova eventual solução atribuindo maior peso a esses resultados, assim poder-se-á atribuir a cada variável o valor de:

$$Q_i = [\pi_i \pi l_i + (1 - \pi_i) \pi a_i] D \quad (2)$$

sendo π_i o peso a atribuir ao valor da solução preliminar, D - Quantidade total pretendida.

Esta solução garante que o somatório das variáveis de decisão corresponderá à solicitação D , ou seja cumpre-se, desde logo, esta restrição habitual dos problemas de engenharia e de gestão.

Assim, é possível construir uma população diversificada para o algoritmo genético que inclui as melhores soluções da etapa anterior, uma parcela de soluções geradas pelo processo aleatório, uma outra parcela em torno das melhores soluções da fase anterior e outra parcela proveniente da adopção da mesma metodologia para obter soluções em torno da melhor solução da geração anterior. Esta metodologia preserva a memória do processo global de adaptação progressiva direccionada e não fecha a porta à procura de novas soluções. A composição de uma população de eventuais soluções com características distintas estimula a procura de soluções óptimas globais.

5 - TÉCNICA DE MANIPULAÇÃO DE RESTRIÇÕES

Numa metodologia de resolução de problemas que envolve o recurso a modelos de simulação dos sistemas físicos, a ideia de penalizar a posteriori soluções que não respeitem as restrições e que assim serão de imediato abandonadas sugere um desperdício de tempo. Quando as metodologias de otimização incluem a ligação a modelos complexos de simulação de sistemas físicos, as soluções iniciais geradas aleatoriamente poderão ter dificuldades em respeitar as diversas restrições. Se for possível concretizar um procedimento que desde logo faça com que algumas restrições possam ser respeitadas antes da ligação aos modelos de simulação então o modelo global tornar-se-á mais racional.

A inclusão de uma distância de segurança em relação a pontos de observação ou de controlo permite a adopção dum procedimento mais flexível de penalização que permita a continuação do estudo de soluções que, embora não cumpram imperativamente a distância de segurança, ainda estejam longe dos pontos de controlo. Assim, serão mais penalizadas as soluções que façam com que a fronteira do fenómeno que se pretende controlar ultrapasse a distância de segurança. As penas serão proporcionais a essa violação. Este procedimento contribui para que o algoritmo genético não convirja prematuramente para óptimos locais.

O algoritmo genético define eventuais soluções que respeitam os limites de cada variável, no entanto, tem grande dificuldade em cumprir com a restrição que exige a satisfação das solicitações, ou seja o somatório dos valores das diversas variáveis de decisão tem de ser preferencialmente igual à quantidade total pretendida. Ora se a solução avançada pelo algoritmo genético não respeitar as solicitações, então não faz sentido correr os modelos de simulação do sistema físico para testar soluções que a priori sabemos que não serão consideradas. Torna-se evidente a necessidade que corrigir as eventuais soluções geradas pelo algoritmo genético para que respeitem, pelo menos, a restrição relacionada com a satisfação das solicitações. No entanto, este novo processo não deve desvirtuar a filosofia inerente a esta técnica de procura da solução óptima global. Associada à ideia de encontrar um procedimento que fizesse o tratamento dos valores das variáveis de decisão para que as restrições fossem respeitadas antes da ligação ao modelo de simulação numérico do comportamento dos sistema físico, surgiu a intenção de aproveitar esta alteração para tentar melhorar o mérito das soluções geradas por processos aleatórios.

6 - REPARAÇÃO DAS SOLUÇÕES GERADAS ALEATORIAMENTE

Os mecanismos de retroacção estão presentes em diversos sistemas naturais. Mantendo os princípios base do algoritmo genético, foi acrescentada a opção de ser possível rectificar os valores das variáveis de decisão de acordo com um critério previamente seleccionado, como o que atende ao objectivo económico e/ou ao controlo dum determinado fenómeno.

6.1 - Reparação usando o critério económico

Um procedimento simples será proceder a uma correcção dos valores das variáveis de decisão que tente manter as proporções iniciais e que respeite os respectivos limites. Nesse processo empregar-se-á o conceito de custo marginal, ou seja, a estimativa do custo que é necessário incorrer se for produzida uma unidade adicional de um bem ou serviço. Assim, o efeito de uma variação unitária em cada variável de decisão Q_i , $i = 1, 2, \dots, n$; na função de custo $f(Q)$ é dado por:

$$c_i = \frac{\partial f(Q_i)}{\partial Q_i} \quad (3)$$

Um procedimento expedito para definir-se quais as variáveis de decisão que devem ser alteradas e de determinar o valor do seu incremento consiste em calcular a distância de cada variação unitária a um custo de referência. No caso do problema a resolver estar relacionado com a minimização da função objectivo então esse valor de referência poderá ser o maior custo unitário. Esta acção traduz-se em considerar que a variável com maior custo não deverá ser incrementada.

O algoritmo desenvolvido para uma rectificação expedita dos valores gerados pelo algoritmo genético que tenha em atenção o respeito pela satisfação da solicitação, os limites das variáveis e que atenda ao critério económico pode ser esquematizado por:

1. Ler valores e limites das variáveis de decisão (Q_i) e valor da solicitação (D);
2. Calcular o valor da correcção:

$$\Delta d = D - \sum_{i=1}^n Q_i \quad (4)$$

3. Calcular os custos marginais associados a cada variável (c_i);
4. Definir valores mínimos e máximos dos custos marginais;
5. Calcular a distância de cada custo marginal ao valor de referência

$$dc_i^* = |c_i - c^*| \quad (5)$$

- a. Se a correcção for para incrementar o somatório dos valores das variáveis então o limite de referência será o maior valor dos custos marginais;
- b. Se a correcção for para reduzir o somatório dos valores das variáveis então o limite de referência será o menor valor dos custos marginais;
6. Calcular a distância de cada valor gerado ao limite de referência

$$dQ_i^* = |Q_i - Q^*| \quad (6)$$

- a. Se a correcção for para incrementar o somatório dos valores das variáveis então o limite de referência será o valor máximo de cada variável;
- b. Se a correcção for para reduzir o somatório dos valores das variáveis então o limite de referência será o valor mínimo de cada variável;

7. Calcular cada produto das duas distâncias

$$dc_i^* . dQ_i^* \quad (7)$$

8. Calcular o somatório dos produtos das distâncias

$$\sum_{i=1}^n dc_i^* . dQ_i^* \quad (8)$$

9. Calcular o índice igual à razão do produto das duas distâncias pela soma dos produtos das distâncias

$$p_i = dc_i^* . dQ_i^* / \sum_{i=1}^n dc_i^* . dQ_i^* \quad (9)$$

10. Calcular o produto do índice p_i pelo valor da correcção Δd

$$dQ_i = p_i \Delta d \quad (10)$$

11. Calcular valor corrigido da variável

$$Q_i = Q_i + dQ_i \quad (11)$$

12. Verificar se $(Q_i)_{\min} \leq Q_i \leq (Q_i)_{\max}$

se $Q_i > (Q_i)_{\max}$ ou se $Q_i < (Q_i)_{\min}$ então

13. $Q_i = (Q_i)_{\max}$ ou $Q_i = (Q_i)_{\min}$ voltar ao passo 2.

A correcção introduzida não afecta os princípios base subjacentes aos algoritmos genéticos na medida em que, de acordo com Michalewicz (1999), segue a seguinte propriedade:

Para quaisquer dois pontos, s_1 e s_2 do espaço viável S , a combinação linear $(1-p)s_1 + ps_2$ é ainda um ponto de S .

6.2 - Reparação usando como critério o controlo do fenómeno do sistema físico em análise

Existindo apenas um ponto de controlo (ou de verificação) então a correcção terá por objectivo aproximar a fronteira do fenómeno até ao local que é determinado retirando à coordenada do ponto de controlo a distância de segurança pretendida. Se a solução inicial provocar o desrespeito pela distância de segurança do ponto de verificação, então será necessário reduzir as variáveis de decisão e prioritariamente naquelas que mais afectam a posição da fronteira junto do ponto de controlo. Se a solução inicial estava aquém do ponto de segurança interessará incrementar as variáveis de decisão.

Se existirem diversos pontos de verificação a rectificação poderá não ser fácil. A tentativa de melhorar a posição da fronteira junto de um ponto de verificação pode prejudicar a situação junto dos restantes pontos de controlo. O número das eventuais combinações poderá ser proibitivo para um procedimento expedito. A análise das experiências realizadas mostrou que não se revelava interessante tentar encontrar alterações que melhorassem a posição da fronteira junto de todos os pontos. A metodologia que foi desenvolvida selecciona o ponto de verificação em situação mais desfavorável e procede a uma correcção para que junto desse ponto o controlo do fenómeno seja melhorado. Ora, para identificar o ponto de controlo sujeito a condições mais desfavoráveis é necessário recorrer a um modelo de simulação do escoamento para que seja determinada a posição da fronteira. Como nesta fase pretende encontrar-se rapidamente uma rectificação, a posição da fronteira poderá ser definida através da resolução duma equação analítica ou duma função de resposta.

6.3 - Reparação para manter o controlo do fenómeno do sistema físico ao menor custo

Com esta concepção retende determinar-se as correcções aos valores das variáveis de decisão que façam com que a fronteira fiquem a uma distância do ponto de controlo igual à distância de segurança e que essa rectificação tenha simultaneamente em vista a redução de custos. O algoritmo desenvolvido encontra-se esquematizado em Ferreira da Silva (2003), aplicado a um sistema de captação e de abastecimento de água em zonas costeiras onde é necessário controlar o avanço da água salgada marinha.

Para ilustrar o desempenho do algoritmo desenvolvido considere-se um problema que consiste na determinação das melhores políticas de utilização dos recursos hídricos disponíveis numa zona costeira. A solicitação é $D = 4000,0 \text{ m}^3/\text{dia}$. Utilizaram-se 50 gerações e como tamanho da população 100 indivíduos. Um extracto dos resultados da última geração (a 50ª) do algoritmo genético base e do algoritmo desenvolvido com a rectificação das soluções encontram-se, registados, respectivamente, nos quadros seguintes:

Pop.		6	7	8	9	10	Melhor
Sum Q	(m ³ /dia)	4031.40	4145.98	3734.48	4076.24	4124.38	4013.73
f. Mérito	(10 ³ €)	19052.31	18814.10	*****	18962.82	18863.74	20010.35

Quadro 1 - Extracto de exemplo de resultados do algoritmo base para a 50ª geração

Pop.		6	7	8	9	10	Melhor
Sum Q	(m ³ /dia)	4000.00	4000.00	4000.00	4000.00	4000.00	4000.00
f. Mérito	(10 ³ €)	20591.95	20608.44	20488.83	20589.46	20554.82	20624.08

Quadro 2 - Extracto de resultados do algoritmo desenvolvido para a 50ª geração

sendo: *Pop.* - População; *Sum Q* - Soma das quantidades das origens (m³/dia); *f. Mérito* - função de mérito (10³ €). Os símbolos ***** indicam que a solução é fortemente penalizada.

Observando os resultados constata-se que o algoritmo genético base não cumpre com a restrição relacionada com a disponibilização da quantidade solicitada, que o algoritmo desenvolvido consegue que todas as soluções cumpram com o solicitado e que a qualidade média e a melhor das soluções são melhores. O procedimento desenvolvido constrói soluções admissíveis e contribui para que o algoritmo apresente uma convergência rápida para as soluções óptimas.

7 - REDUÇÃO DO TAMANHO DA POPULAÇÃO E DO NÚMERO DE GERAÇÕES

Pode adoptar-se uma lei que defina um tamanho da população decrescente com a geração:

$$\lambda = \text{int}[\lambda_0 + at^b] \quad (12)$$

sendo: λ - Tamanho da população; λ_0 - Tamanho inicial da população; t - Geração ($t=1,2,\dots,ng$).

Quando se pretende proceder à análise de sensibilidade dos efeitos dos valores de determinado elemento base é vantajoso reduzir o número de gerações em função do número da simulação do parâmetro aleatório em análise, adoptando uma lei do tipo:

$$ng = \text{int}[ng_0 + a(ns)^b] \quad (13)$$

sendo: ng - n.º gerações; ng_0 - n.º gerações da primeira simulação; ns - n.º da simulação ($ns=1,2,\dots,nts$); nts - número total de simulações.

Agora é necessário adoptar uma lei que determine o tamanho da população em função do número total de gerações nessa simulação (ng) e da geração actual (t). Uma lei possível será:

$$\lambda = \text{int}[\lambda_0 - (a-b(ng))LN(t)] \quad (14)$$

λ - Tamanho da população; t - Geração ($t=1,2,\dots,ng$).

Por exemplo, reduzindo o número de gerações de acordo com as leis:

$$ng = \text{int}[50 - 4,7LN(ns)] \quad (15)$$

$$\lambda = \text{int}[100 - (34,46 - 0,21ng)LN(t)] \quad (16)$$

os resultados registados no quadro seguinte são praticamente iguais, no entanto, os tempos de execução são: 195 minutos e 33 minutos, respectivamente para ng e λ fixos ($ng=50$ e $\lambda=100$) e ng e λ decrescentes, com os cálculos realizados num computador Pentium IV a 2,4 GHz.

	<i>Sum QS</i>
<i>ng</i> e λ fixos	3837.55
<i>ng</i> e λ decrescentes	3837.59

Quadro 3 - Exemplo dos resultados com *ng* e λ fixos e *ng* e λ decrescentes
sendo: *Sum QS* - Soma das extracções (m³/dia).

8 - PENALIZAÇÃO CRESCENTE

Adoptando o conceito de distância de segurança entre a fronteira do fenómeno em análise e pontos de controlo, o modelo de optimização poderá ser mais flexível com o não cumprimento desta restrição quando está a procurar soluções para os maiores valores da distância de segurança. A possibilidade de penalizar de forma crescente à medida que se reduz a distância de segurança permite não eliminar de imediato eventuais soluções, ou seja permite diversificar o conjunto de eventuais soluções. A expressão geral que permite avaliar o mérito de cada potencial solução é:

$$f(x) = f(x) + \left(cp_1 + ap_1 t^{bp_1} \right) \left| D - \sum_{i=1}^n Q_i \right|^{ck_1 + ak_1 t^{bk_1}} + \left(cp_2 + ap_2 t^{bp_2} \right) \left(\sum_{j=1}^{npc} d_{ICj} \right)^{ck_2 + ak_2 t^{bk_2}} \quad (17)$$

onde: *t* - Geração ($t=1,2,\dots,ng$) *cp*₁, *ap*₁, *bp*₁, etc. - Coeficientes da função de penalização; *D* - Solicitação, *Q*_{*i*} - Quantidade disponibilizada pela origem *i*, *npc* - Número de pontos de controlo, *d*_{*ICj*} - distância da fronteira ao ponto de controlo *j*.

9 - CONCLUSÕES

O modelo de gestão necessita de testar todas as eventuais soluções geradas pelo algoritmo genético (A.G.) recorrendo aos modelos de simulação do comportamento dos sistemas físicos. Os (A.G.) base foram aperfeiçoados para tornar o modelo de optimização-simulação mais racional e, consequentemente, para melhorar o tempo de execução e aumentar a eficácia de convergência para a solução do problema.

Num algoritmo genético base a selecção define as soluções que serão preservadas. Uma formulação adequada de acções correctivas tendo em atenção o critério económico e, obviamente, o conhecimento do comportamento dos sistemas físicos, irá permitir o controlo das soluções geradas aleatoriamente, sendo, inclusivamente, responsável pela convergência, ou não, do A.G. em direcção à solução óptima. A nova geração pode ser formada por alguns dos elementos da geração anterior, pelos seus descendentes e a partir de estratégias definidas pelo decisor. Assim, a distribuição populacional é controlada por processos aleatórios e por estratégias determinísticas direccionadas. Essas rectificações respeitam a priori uma restrição frequente nos problemas de gestão que se relaciona com a satisfação das solicitações.

10. REFERÊNCIAS

1. Ferreira da Silva, Júlio (2003): "Gestão optimizada à escala regional de sistemas aquíferos potencialmente sujeitos à intrusão salina - Um modelo global para o uso sustentável da água em regiões costeiras", Dissertação de doutoramento em Engenharia Civil - Hidráulica, Universidade do Minho;
2. Ferreira da Silva, Júlio (2005): "PROCESSO EVOLUTIVO DIRECCIONADO - Uma metodologia de optimização para a resolução de problemas que envolvam a simulação do comportamento de sistemas físicos", CEIO, Universidade do Minho;
3. Michalewicz, Z. (1999): "Genetic Algorithms + Data Structures = Evolution Programs", Third, revised and extended edition, Springer Verlag, NY;
4. Michalewicz, Z. (1999): "A survey of constraint handling techniques in evolutionary computation methods".