



Algoritmos Genéticos

Capítulo 5

Ricardo Linden

Teoria dos GAs

- ✱ Algoritmos genéticos são um pesadelo em termos de análise;
- ✱ Sua estrutura é probabilística por natureza;
- ✱ Não pretendemos explicar aqui matematicamente todas as suas propriedades;
- ✱ Objetivos:
 - ✱ explicar basicamente seus fundamentos;
 - ✱ dar uma boa idéia de porque os GAs funcionam.

Esquemas - Conceitos Básicos

- ✱ Um esquema consiste em um gabarito (“*template*”) descrevendo um subconjunto dentre o conjunto de todos os indivíduos possíveis;
- ✱ O esquema descreve similaridades entre os indivíduos que pertencem a este subconjunto, ou seja, descreve quais posições dos seus genomas são idênticas.

Alfabeto de Esquemas

- ☀ Consiste no conjunto de símbolos utilizados na nossa representação mais o símbolo *;
- * significa "não-importa" (don't care, wildcard ou coringa);
- Indivíduos que correspondem àquele esquema diferem exatamente nas posições onde encontramos este símbolo.
- Quando usamos a representação binária, um esquema que tenha comprimento n com m posições contendo o símbolo * terá m graus de liberdade e representará até 2^m indivíduos diferentes da atual população.

Definição Formal

- ☀ Um esquema é uma string $s = \{s_1 s_2 \dots s_n\}$, com as seguintes propriedades:
 - Comprimento n ;
 - Posições pertencem ao conjunto Γ (alfabeto usado) + $\{*\}$ (símbolo de wildcard);
 - Cada posição da string dada por $s_k \neq '*'$ é chamada de especificação;
 - Um “wildcard” representa o fato de que aquela posição pode assumir qualquer valor dentro do conjunto Γ

Exemplos

- ★ Populações de strings de bits geram o alfabeto de esquemas dado pelos símbolos {0, 1 e *}

Esquema	Indivíduos que representa
1*	10 , 11
1*0*1	10001, 10011, 11001, 11011
**0	000, 010, 100, 110

Exemplos

- ☀ Populações de palavras, têm esquemas dados pelo alfabeto ocidental $\Gamma=\{a,b, \dots, z\}$ mais o símbolo $*$:

Esquema	Indivíduos que representa
a^*	aa, ab, \dots, az
a^*b	aab, abb, \dots, azb
$**xy$	$aaxy, abxy, \dots, azxy, baxy, bbxy, \dots, bzxy, \dots, zaxy, zbxy, \dots, zzxy$

Satisfação de um Esquema

- ☀ Uma string x satisfaz um esquema se $\forall s_k$ pertencente à string s definidora do esquema, $s_k \neq *$, temos que $s_k = x_k$.
- ☀ Exemplo:
 - ☀ Esquema definido por $s = **zq$.
 - ☀ A string $x = abzq$ satisfaz este esquema pois $s_1 = s_2 = *$ e também $s_3 = x_3$ e $s_4 = x_4$.
 - ☀ A string $y = abzz$ não satisfaz este esquema, posto que $s_4 \neq y_4$

Definições Importantes

- ✱ Um esquema tem duas características importantes: sua ordem e seu tamanho.
- ✱ A ordem de um esquema, denotado por $O(H)$, corresponde ao número de posições neste esquema diferentes de *
- ✱ O tamanho do esquema, representado por $\delta(H)$, se refere ao número de pontos de corte entre a primeira e a última posições diferentes de * dentro do esquema

Definições Importantes

- ✱ Um problema associado normalmente à piora do desempenho de uma GA é a questão da carona (*hitchhiking*).
- ✱ Se um determinado esquema tiver um alto desempenho, todos os bits presentes em indivíduos tendem a se proliferar, não só aqueles que pertencem ao esquema desejado.
- ✱ Os bits em posições fora do esquema pegam carona com o esquema para se propagar para as próximas gerações, mesmo que eles não colaborem para a melhoria geral da avaliação do cromossomo.

Definições Importantes

- ✱ Existe problemas chamados de enganadores (*deceptives*).
- ✱ Um problema é dito enganador se um esquema que não contém o máximo global tem uma avaliação média superior a esquemas que o contêm.
- ✱ Se o seu problema for enganador, os esquemas que não contêm o máximo global tenderão a proliferar-se, o que fará com que o resultado ótimo seja mais difícil de ser encontrado.
- ✱ Uma característica de um problema enganador é que ele é difícil para todo e qualquer método:
 - ✱ soluções vizinhas ao máximo global, neste tipo de problema, tendem a ter avaliações baixas.
 - ✱ os máximos tendem a ser picos localizados em “depressões” da função de avaliação, que seriam evitadas por métodos de gradiente, entre outros.

Exemplos

Esquema	Ordem	Tamanho
*****1***	1	0
1*****0	2	7
11*0	3	5
101010	6	5

Paralelismo Implícito

- ✱ O paralelismo fundamental dos GAs não está apenas no fato de que uma população contendo vários indivíduos é manipulada simultaneamente;
- ✱ Existe paralelismo também embutido no fato que para cada elemento da população um GA manipula dezenas, quiçá centenas de esquemas simultaneamente;
- ✱ Os mecanismos de seleção natural vão fazer com que os melhores esquemas acabem reproduzindo mais e permanecendo mais tempo na população;
- ✱ Isto quer dizer que o importante não é o indivíduo e sim o esquema.
 - ✱ Pode ser que o indivíduo morra, mas o esquema que o torna bom tende a proliferar e continuar na população.

Teorema dos Esquemas

- ✴ Enunciado por John Holland
- ✴ Um GA calcula explicitamente a avaliação de n indivíduos (a população corrente), mas implicitamente, ele calcula a avaliação de um número muito maior de esquemas que são instanciados por cada indivíduo da população
 - ✴ Paralelismo Implícito!
- ✴ Esquemas com avaliação superior à média tende a ocorrer mais frequentemente nas próximas gerações e aqueles esquemas ocorrendo em cromossomos com avaliações abaixo da média tendem a desaparecer

Teorema dos Esquemas

☀ Formalmente:

- n o número de indivíduos pertencentes a um certo esquema s
- média de avaliação do esquema igual a r
- x a média das avaliações de toda a população
- número esperado de ocorrências de s na próxima geração é aproximadamente igual a $n*r/x$.

Exemplo

☀ População dada por:

Indivíduo	Avaliação
01101	169
11000	576
01000	64
10011	361
Média	292.5

Pertencentes
ao esquema 1****

Exemplo

☀ Esquema 1****

- Dois Indivíduos
- Média de avaliação: 468,5
- Número esperado de indivíduos: $468.5 * 2 / 292.5 \approx 3.2$

☀ Esquema 0**0*

- Dois indivíduos
- Média de avaliação 116.5.
- Deve estar presente em $116.5 * 2 / 292.5 \approx 0.8$ indivíduos

Atenção

- ☀ Número não é exato;
- ☀ Normalmente ele não é inteiro e só podemos ter um número inteiro de indivíduos
- ☀ O GA não é determinístico, e sim probabilístico:
 - o número tende a ser aquele calculado;
 - muita sorte (ou muito azar) nos sorteios pode mudar este número

Efeito dos Operadores

- ✱ Quando aplicamos o crossover, um corte no meio de um esquema irá destruí-lo para sempre
 - ✱ Exceção: o indivíduo que estiver reproduzindo com o pai que contém o esquema seja idêntico a este depois da posição de corte
- ✱ Quanto maior for o tamanho do esquema ($\delta(H)$), maior a sua probabilidade de ser destruído.
 - ✱ Um esquema de ordem 1 e tamanho zero nunca pode ser destruído
- ✱ Reformulação do teorema dos esquemas: quanto maior a avaliação do esquema e menor o seu tamanho, mais cópias ele terá na próxima geração.

Efeito dos Operadores

- ✱ A mutação também é destrutiva, se ocorrer em uma posição em que o esquema possua um valor diferente de *;
- ✱ Quanto maior a ordem do esquema, mais chances deste ser corrompido pelo operador de mutação;
- ✱ Mutações em posições em que o valor é igual a * não afetam a satisfação do esquema por parte do indivíduo corrente.

Ação dos Operadores

- ✱ A ação dos operadores se encaixa no que Holland costumava chamar de tensão entre exploração (*exploration*, a busca de novas adaptações) e aproveitamento (*explotation*, a manutenção das adaptações úteis feitas até a atual geração).
- ✱ Qualquer ação de operador genético é potencialmente destrutiva, mas encaixa-se na categoria de exploração, a busca por indivíduos de avaliação melhor que seus pais.

Enunciado Final do Teorema

- ✱ O GA tende a preservar com o decorrer do tempo aqueles esquemas com maior avaliação média e com menores ordem e tamanho, combinando-os como blocos de montar de forma a buscar a melhor solução

Atenção

- ✱ Existe oposição ao teorema dos esquemas;
- ✱ Altenberg (1995), por exemplo, aponta que o teorema dos esquemas é verdadeiro mesmo quando a representação cromossomial é totalmente aleatória;
- ✱ Esta objeção, entre outras importantes, sugere apenas que a área de embasamento teórico dos algoritmos genéticos ainda precisa de muito estudo e comprovação, antes de se considerar consolidada.
- ✱ Existem congressos devotados apenas a este tipo de estudo, e muito ainda há por fazer nesta direção