

Tópicos em Computação Evolucionária: Mecanismos de Diversidade

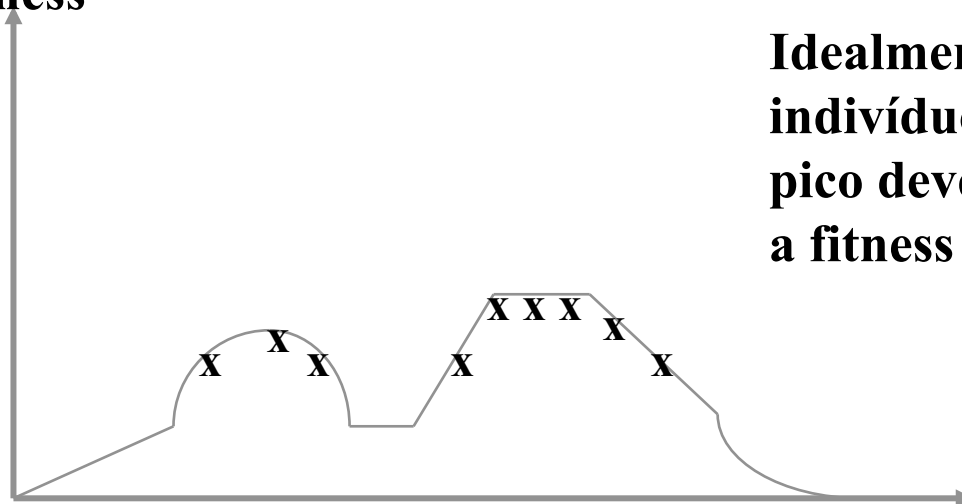
Gisele L. Pappa

Mecanismos de Diversidade

- Em algoritmos evolucionários, a população normalmente converge para uma população uniforme, com muitas cópias do mesmo indivíduo
- Evolução natural mantém a diversidade das espécies
 - Cada espécie ocupa um nicho ecológico
- Solução: modificar os algoritmos para utilizar um método de *niching*, simulando uma competição por recursos limitados

Introdução a métodos de Niching

Fitness



Idealmente, o número de indivíduos em torno de um pico deveria ser proporcional a fitness do pico

**Indivíduos
(soluções)**

Motivação para Niching

- Reduz a velocidade de convergência da população para um único indivíduo (evitando convergência prematura)
- Encontra um conjunto de soluções ótimas ou quase-ótimas, ao invés de uma solução ótima

Motivação para Niching

- Onde utilizar niching
 - Otimização de funções multi-modais (muita soluções ótimas)
 - Otimização de funções multi-objetivas (soluções avaliadas de acordo com vários critérios)

Métodos de Niching

- 2 tipos principais:
 - *Fitness sharing*
 - *Crowding*

Fitness Sharing (1)

- Modifica apenas a maneira como a fitness de um indivíduo é determinada, o resto da evolução não é alterada
- Seja: F_i a fitness original do indivíduo i
 F_i' a fitness compartilhada de um indivíduo i
 NC_i a contagem de nicho do indivíduo i , onde
 NC mede a saturação de um nicho

$$F_i' = F_i / NC_i$$

Fitness Sharing (2)

- Seja SH uma função de compartilhamento, medindo a **similaridade** entre dois indivíduos
- A contagem de nicho (NC) de um indivíduo é a soma das funções compartilhadas (SH) (das similaridades) entre o indivíduo i e todos os indivíduos da população (incluindo ele mesmo), ou seja

$$F_i' = \frac{N F_i}{\sum_{j=1}^N SH(i,j)}$$

onde N é o número de indivíduos da população

Fitness Sharing (3)

- SH retorna um número entre $[0..1]$
 - Similaridade é inversamente proporcional a distância
- Seja d_{ij} a distância entre dois indivíduos i, j
 - Se $d_{ij} = 0$ (os indivíduos são idênticos), o valor da SH é 1
 - Se d_{ij} é maior ou igual a um limiar de distância, θ_{share} , o valor da SH é 0 (i e j estão em nichos diferentes)
 - Se $0 < d_{ij} < \theta_{share}$, então a função retorna um valor intermediário, entre 0 e 1

Fitness Sharing (3)

- Seja d_{ij} a distância entre dois indivíduos i, j
 - Se $d_{ij} = 0$ (os indivíduos são idênticos), o valor da SH é 1
 - Se d_{ij} é maior ou igual a um limiar de distância, θ_{share} , o valor da SH é 0 (i e j estão em nichos diferentes)
 - Se $0 < d_{ij} < \theta_{share}$, então a função retorna um valor intermediário, entre 0 e 1

$$SH(d_{ij}) = \begin{cases} 1 - (d_{ij} / \theta_{share})^\alpha, & \text{se } d_{ij} < \theta_{share} \\ 0, & \text{nos outros casos} \end{cases}$$

- α é um parâmetro, que normalmente recebe valor 1

Fitness Sharing (4)

- 2 maneiras de medir a distância entre dois indivíduos
 - Compartilhamento de fenótipo ou de genótipo
- Compartilhamento de **genótipo**: considera o material genéticos dos indivíduos
 - Exemplo: distância de Hamming (se codificação binária é utilizada)
 - Número de bits diferentes

1) 1 0 1 1 0

$d_{1,1} = 0$; $d_{1,2} = 2$; $d_{1,3} = 1$

2) 1 0 1 0 1

3) 0 0 1 1 0

Fitness Sharing (5)

- Compartilhamento de fenótipo: considera os indivíduos decodificados (soluções candidatas)
 - Exemplo: decodificar 5 bits em uma variável x
- $d_{i,j} = |x_i - x_j|$, onde $|x|$ é o valor absoluto de x

| material genético | | | | | | decodificando x | |
|-------------------|---|---|---|---|---|-------------------|--|
| 1) | 1 | 0 | 1 | 1 | 0 | 22 | $d_{1,1} = 0; d_{1,2} = 1; d_{1,3} = 16$ |
| 2) | 1 | 0 | 1 | 0 | 1 | 21 | |
| 3) | 0 | 0 | 1 | 1 | 0 | 6 | |

- Ind. 1 é mais similar ao indivíduo 2 que ao indivíduo 3 de acordo com a distância de fenótipos.
- Ind.1 é mais similar ao indivíduo 3 que ao indivíduo 2 de acordo com a distância de genótipos.

Fitness Sharing (6)

- Exemplo do efeito da fitness compartilhada (quanto maior a fitness, melhor o indivíduo)

| $indiv_i$ | F_i | NC_i | F_i' |
|-----------|-------|--------|--------|
| 1 | 10 | 2.5 | 4 |
| 2 | 12 | 4 | 3 |
| 3 | 6 | 1 | 6 |
| 4 | 8 | 2 | 4 |

O valor de F_i' é utilizado para seleção.

- Indivíduos 1 e 2 tem uma fitness original alta, mas são penalizados pelo valor alto de NC
- Indivíduo 3 tem a melhor função compartilhada

Fitness Sharing (7)

- Vantagens
 - Conceitualmente simples, boa metáfora do niching natural
 - Tenta distribuir indivíduos proporcionalmente ao fitness do pico onde se encontram

Fitness Sharing (7)

- Desvantagens
 - Dificuldade de ajustar o valor do parâmetro θ_{share} : idealmente, requer conhecimento sobre o número e o tamanho dos picos, o que não é realista em problemas difíceis
 - Solução : ajustar dinamicamente o valor de θ_{share} durante a busca
 - Computacionalmente caro: precisa calcular a distância entre pares de indivíduos
 - Solução possível : computar o valor da função de compartilhamento baseado em uma amostra da população

Crowding

- Ideia básica: os novos indivíduos são inseridos na próxima geração substituindo pais similares
- Ao contrário de *fitness sharing*, *crowding* não aloca indivíduos proporcionalmente a fitness do pico
- Utiliza uma função de distância para determinar quando dois indivíduos são similares

Crowding Determinístico

- Agrupa os indivíduos da população em pares
- Cruza todos os pares
 - Probabilidade de cruzamento = 1
- Muta os indivíduos gerados pelos pares
- Cada indivíduo compete com um pai
 - Indivíduos são forçados a competir com o pai menos distante
- O vencedor é inserido na nova população

Leitura Recomendada

- Sareni, B.; Krahenbuhl, L., "Fitness sharing and niching methods revisited," *IEEE Transactions on Evolutionary Computation*, vol.2, no.3, pp.97-106, 1998

Tópicos em Computação Evolucionária: Mecanismos de Diversidade

Gisele L. Pappa