

KDD E MINERAÇÃO DE DADOS

Noções Introdutórias em Algoritmos Genéticos

Prof. Ronaldo R. Goldschmidt

ronaldo@de9.ime.eb.br

rribeiro@univercidade.br

geocities.yahoo.com.br/ronaldo_goldschmidt

ALGORITMOS GENÉTICOS

Conceitos Básicos:

- Algoritmo de busca/otimização inspirado na seleção natural e reprodução genética.
- Combina sobrevivência do mais apto e cruzamento aleatório de informação.

ALGORITMOS GENÉTICOS

Analogia com a Natureza:

Evolução Natural



Algoritmos Genéticos

- Indivíduo
- Cromossoma
- Reprodução Sexual
- Mutação
- População
- Gerações
- Meio Ambiente

- Solução
- Representação
- Operador Cruzamento
- Operador Mutação
- Conjunto de Soluções
- Ciclos
- Problema

ALGORITMOS GENÉTICOS

Qual a finalidade de Algoritmos Genéticos?

Algoritmos Genéticos empregam um processo adaptativo e paralelo de busca de soluções em problemas complexos.

Algoritmos Genéticos são técnicas que procuram obter boas soluções para problemas complexos por meio da evolução de populações de soluções codificadas em cromossomas artificiais.

ALGORITMOS GENÉTICOS

Qual a finalidade de Algoritmos Genéticos?

- **Adaptativo:** informação corrente influencia a busca futura
- **Paralelo:** várias soluções consideradas a cada momento
- **Problema Complexo:** de difícil formulação matemática ou com grande espaço de busca (grande número de soluções – possivelmente infinito)

ALGORITMOS GENÉTICOS

Problema Complexo

Exemplo:

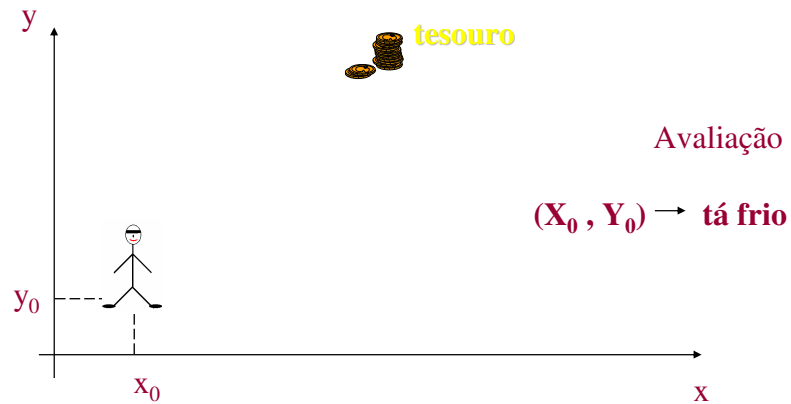
Maximizar $f(x) = x^2$: encontrar $x \in (0 \dots 2^L - 1)$ para $f(x) = \text{máx}$

2^L	Número de Pontos no Espaço	Tempo de Busca 10^9 inst/seg
L=3	8	< 1 Seg
L=10	1024	< 1 Seg
L=30	1 bilhão	1 Seg
L=90	10^{27}	15 bilhões de anos

ALGORITMOS GENÉTICOS

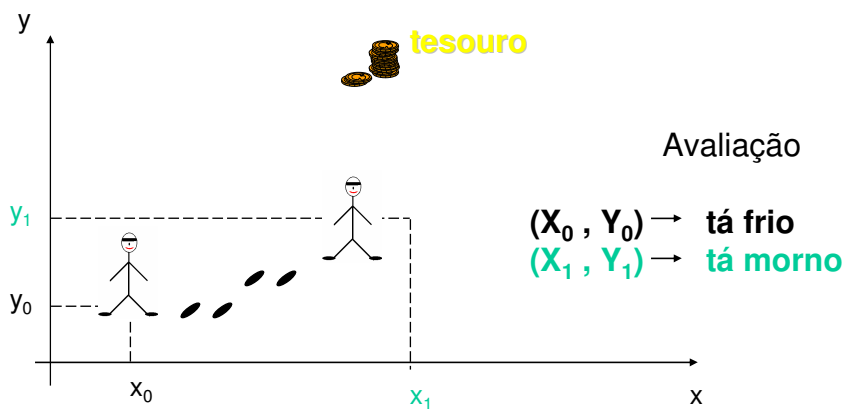
Busca Paralela e Adaptativa

Cabra Cega: Busca de objetivo escondido em uma área



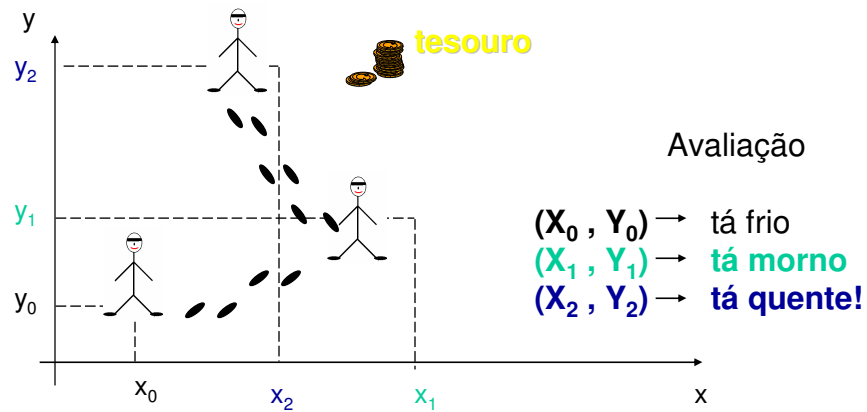
ALGORITMOS GENÉTICOS

Problema da Cabra Cega



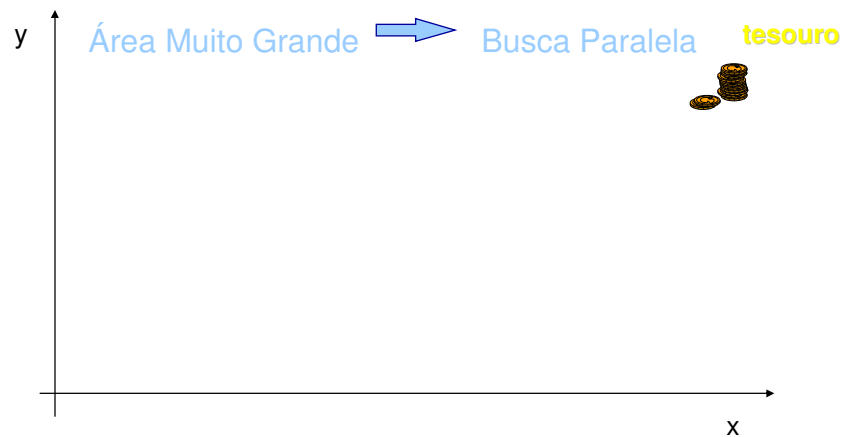
ALGORITMOS GENÉTICOS

Problema da Cabra Cega



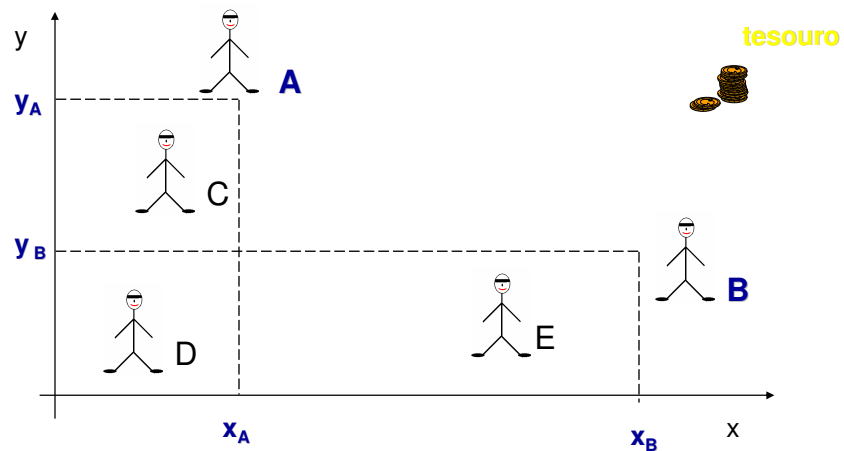
ALGORITMOS GENÉTICOS

Problema da Cabra Cega



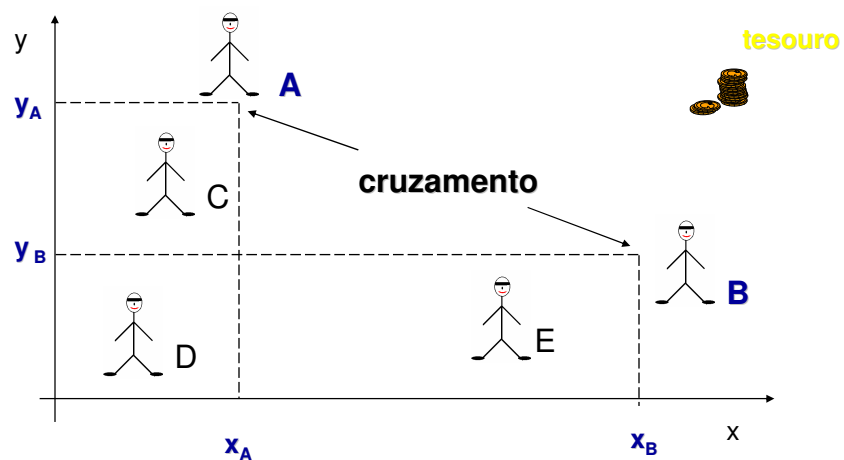
ALGORITMOS GENÉTICOS

Problema da Cabra Cega



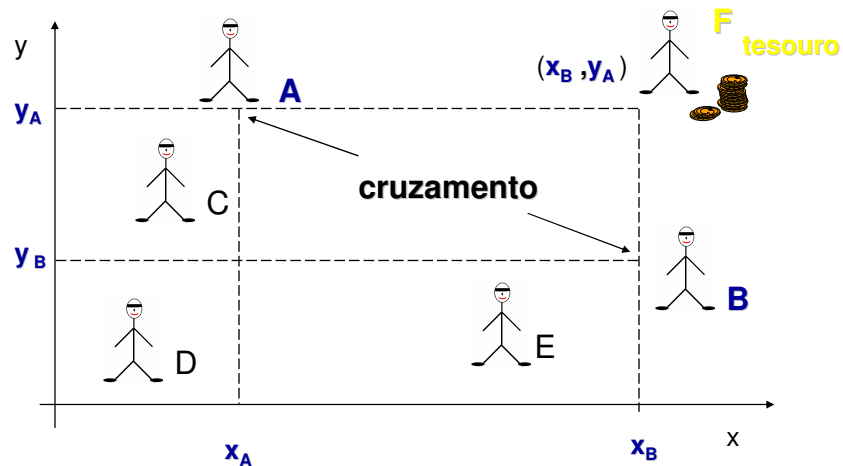
ALGORITMOS GENÉTICOS

Problema da Cabra Cega



ALGORITMOS GENÉTICOS

Problema da Cabra Cega



ALGORITMOS GENÉTICOS

Características:

Um cromossoma é uma estrutura de dados que representa uma das possíveis soluções do espaço de busca de um problema.

Cromossomas são submetidos a um processo evolucionário que avalia, seleciona e combina soluções ao longo de diversos ciclos, procurando obter indivíduos mais aptos (melhores soluções).

ALGORITMOS GENÉTICOS

Exemplos de Aplicações:

- Otimização de funções matemáticas
- Otimização Combinatorial
- Otimização de Planejamento
- Problema do Caixeiro Viajante
- Problema de Otimização de Rota de Veículos
- Otimização de Distribuição e Logística
- Seleção de Variáveis em Mineração de Dados
- Dentre muitos outros...

ALGORITMOS GENÉTICOS

Componentes de um AG:

1. Problema
2. Representação das Soluções
3. Decodificação do Cromossoma
4. Avaliação
5. Seleção
6. Operadores Genéticos
7. Técnicas e Parâmetros

ALGORITMOS GENÉTICOS

Problema

Exemplo de Aplicação:

→ Achar o valor máximo para $f(x) = x^2$, x no limite de 0 a 63.

Este é um exemplo didático, pois o espaço de busca possui apenas 64 valores distintos a serem considerados.

A busca exaustiva poderia ser utilizada neste problema

Valor Máximo = 3969, quando $x = 63$

ALGORITMOS GENÉTICOS

Representação das Soluções

Gene: menor unidade de informação em um AG.

Cromossoma: série de genes que representa uma solução do AG.

A forma de representação adotada deve ser capaz de representar todas as soluções do espaço de busca que se deseja investigar.

A representação binária é a mais empregada, devido à sua simplicidade e facilidade de manipulação.

Na aplicação do exemplo anterior:

→ Palavras binárias representando sucessivas potência de 2.
011100 → Representa 28 (decodificação)
110101 → Representa 53 (uma solução mais apta)

ALGORITMOS GENÉTICOS

Representação das Soluções

Há problemas em que valores binários podem ser utilizados para representar números reais com uma precisão de p casas decimais.

Supondo que os números reais pertençam a um intervalo $[X_{min}, X_{max}]$, são necessários k bits de forma que k atenda à seguinte relação:

$$2^k \leq (X_{max} - X_{min}) * 10^p$$

Nem sempre a representação binária pode ser empregada. Em muitos casos, o problema requer um alfabeto de representações com mais símbolos.

ALGORITMOS GENÉTICOS

Avaliação

Em geral, para ser avaliado, um cromossoma deve ser, primeiramente, decodificado.

A avaliação de uma solução consiste da aplicação de uma função (**função de avaliação**) que fornece uma medida da qualidade de tal solução (**aptidão**) frente ao problema.

Na aplicação do exemplo anterior:

- Palavras binárias representando sucessivas potência de 2.
- 011100 → Representa 28 → $f(28) = 784$
- 110101 → Representa 53 → $f(53) = 2809$

ALGORITMOS GENÉTICOS

Seleção

O AG tende a **privilegiar indivíduos com maiores aptidões** para permanecer e se multiplicar na população.

Método da Roleta: cada cromossoma tem a probabilidade de permanecer na próxima geração proporcional à sua aptidão.

Cromossomas com **maiores aptidões** possuem **mais espaço na roleta** → **maiores chances** de **serem escolhidos** para reprodução:

$$p_i = \frac{f_i}{\sum_{i=1}^n f_i}$$

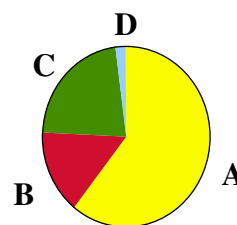
ALGORITMOS GENÉTICOS

Seleção

População

Cromossoma	Palavra	X	Aptidão (x ²)
A	100100	36	1296
B	010010	18	324
C	010110	22	484
D	000001	1	1

Seleção



Probabilidade de Seleção $\approx$ **Aptidão do Cromossoma**

ALGORITMOS GENÉTICOS

Operações Básicas:

Seleção: privilegia os indivíduos mais aptos.

Reprodução: indivíduos (palavras binárias) são reproduzidas com base na aptidão.

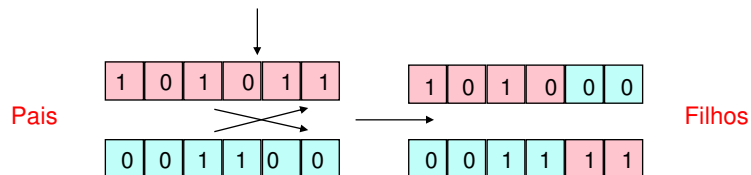
Crossover: troca de genes (pedaços de palavras).

Mutação: troca aleatória de um gene (bit da palavra).

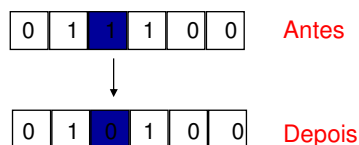
ALGORITMOS GENÉTICOS

Exemplos de Operadores Genéticos

Crossover 1 Ponto



Mutação



ALGORITMOS GENÉTICOS

Exemplos de Parâmetros em AGs

Tamanho da População

Taxa de Crossover

Taxa de Mutação

Número de Gerações

Total de Indivíduos

ALGORITMOS GENÉTICOS

Exemplos de Técnicas em AGs

Elitismo: Todos os cromossomas são substituídos, sendo o cromossoma mais apto da população corrente copiado para a população seguinte.

Steady State: Gera M indivíduos ($M < N$), que substituem os piores indivíduos da população corrente. Pode haver duplicação de indivíduos na população resultante.

Steady State sem Duplicados: Similar à técnica de Steady State, não permitindo a presença de indivíduos duplicados na população.

ALGORITMOS GENÉTICOS

Exemplos de Técnicas em AGs

As **técnicas de aptidão** influenciam na forma com que as aptidões são numericamente atribuídas:

Método Clássico: Atribuir como aptidão de um cromossoma o valor numérico do resultado da avaliação.

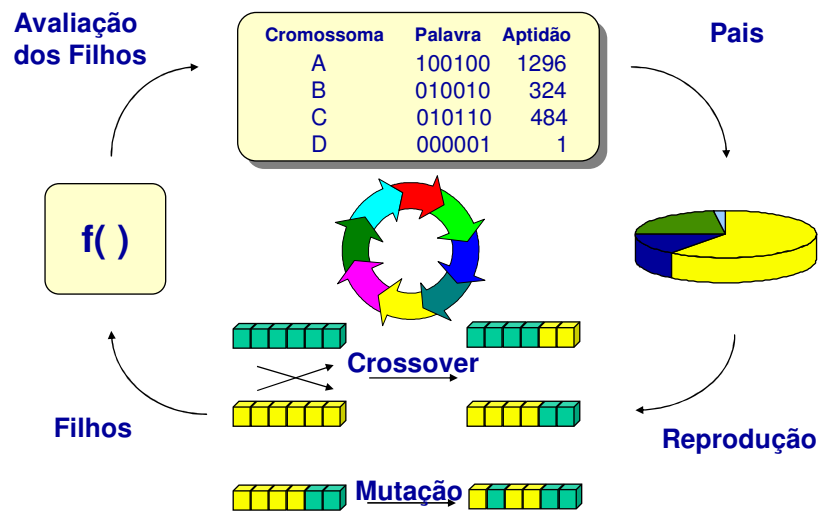
Normalização Linear: Ordenar os cromossomas por avaliação e atribuir valores de aptidões de um intervalo, distanciados de um valor fixo entre eles:

$$A_i = \min + \left[\frac{\max - \min}{N - 1} \right] * (i - 1)$$

Busca tratar competição próxima e super-indivíduo

ALGORITMOS GENÉTICOS

Ciclo do Algoritmo Genético



ALGORITMOS GENÉTICOS

Estrutura Geral de um Algoritmo Genético

```
T=0
Gerar População Inicial P(0)
Avaliar P(0)
Enquanto Critério de Parada não for satisfeito faça:
    T=T+1
    Selecionar População P(T) a partir de P(T-1)
    Aplicar Operadores de Cruzamento sobre P(T)
    Aplicar Operadores de Mutação sobre P(T)
    Avaliar P(T)
Fim Enquanto
```

Estrutura Geral de uma Função de Avaliação (Avaliar P(T))
Para todo indivíduo i da População Atual P(T) **faça**
 Avaliar a aptidão do indivíduo i
Fim Para

ALGORITMOS GENÉTICOS

Exercício

Utilizar o software ICADEMO da PUC-RIO DEE ICA para experimentar a utilização de AGs na busca pelo ponto de máximo da função F6.

$$F6(x, y) = 0.5 - \frac{(\sin(\sqrt{x^2 + y^2})^2 - 0.5)}{(1.0 + 0.001(x^2 + y^2))^2}$$

É uma função de difícil otimização devido às diversas oscilações (mínimos locais).

valor máximo = 1 quando $x = y = 0$.