

# Algoritmos Genéticos

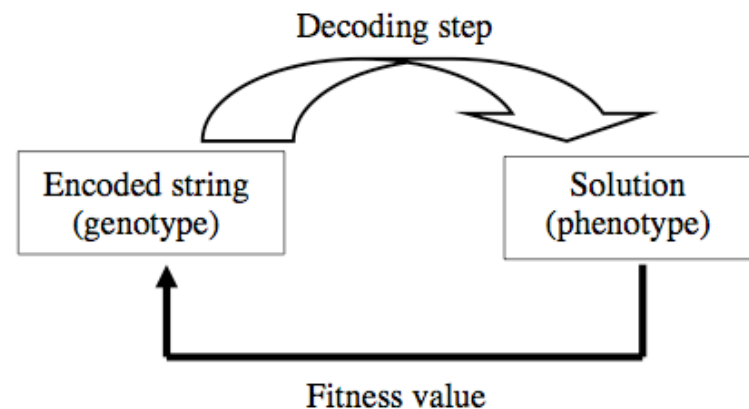
Gisele L. Pappa

# Algoritmos Genéticos

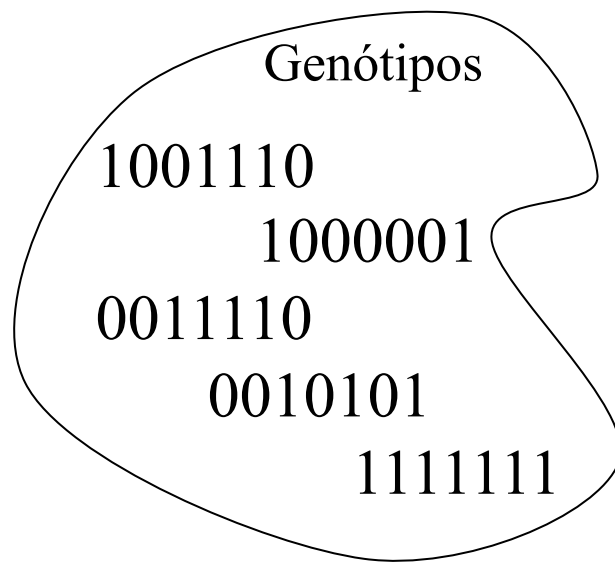
- Técnica mais dissiminada em EA
- Introduzida por Holland em 1975, e desenvolvida por um de seus alunos, Goldberg

# Algoritmos Genéticos

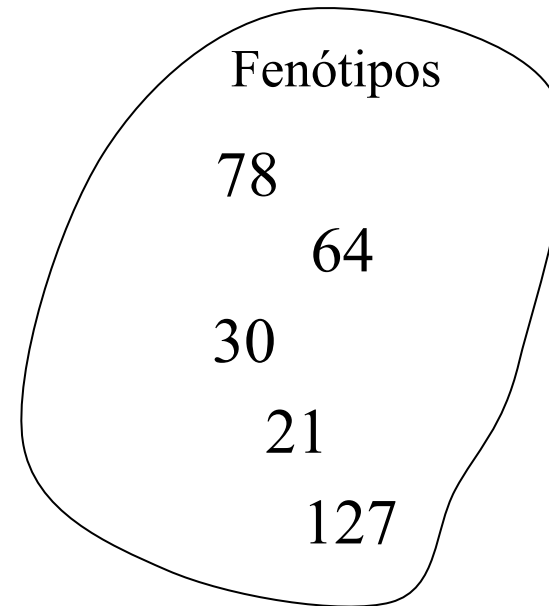
- Indivíduos são strings binárias
- Cromossomo (indivíduo) tem tamanho fixo
  - Genes normalmente tem tamanho fixo
- Existe um mapeamento do genótipo para o fenótipo



# Genótipo versus Fenótipo



Espaço de busca



Espaço de Soluções

- Em alguns algoritmos evolucionários não existe distinção entre genótipo e fenótipo

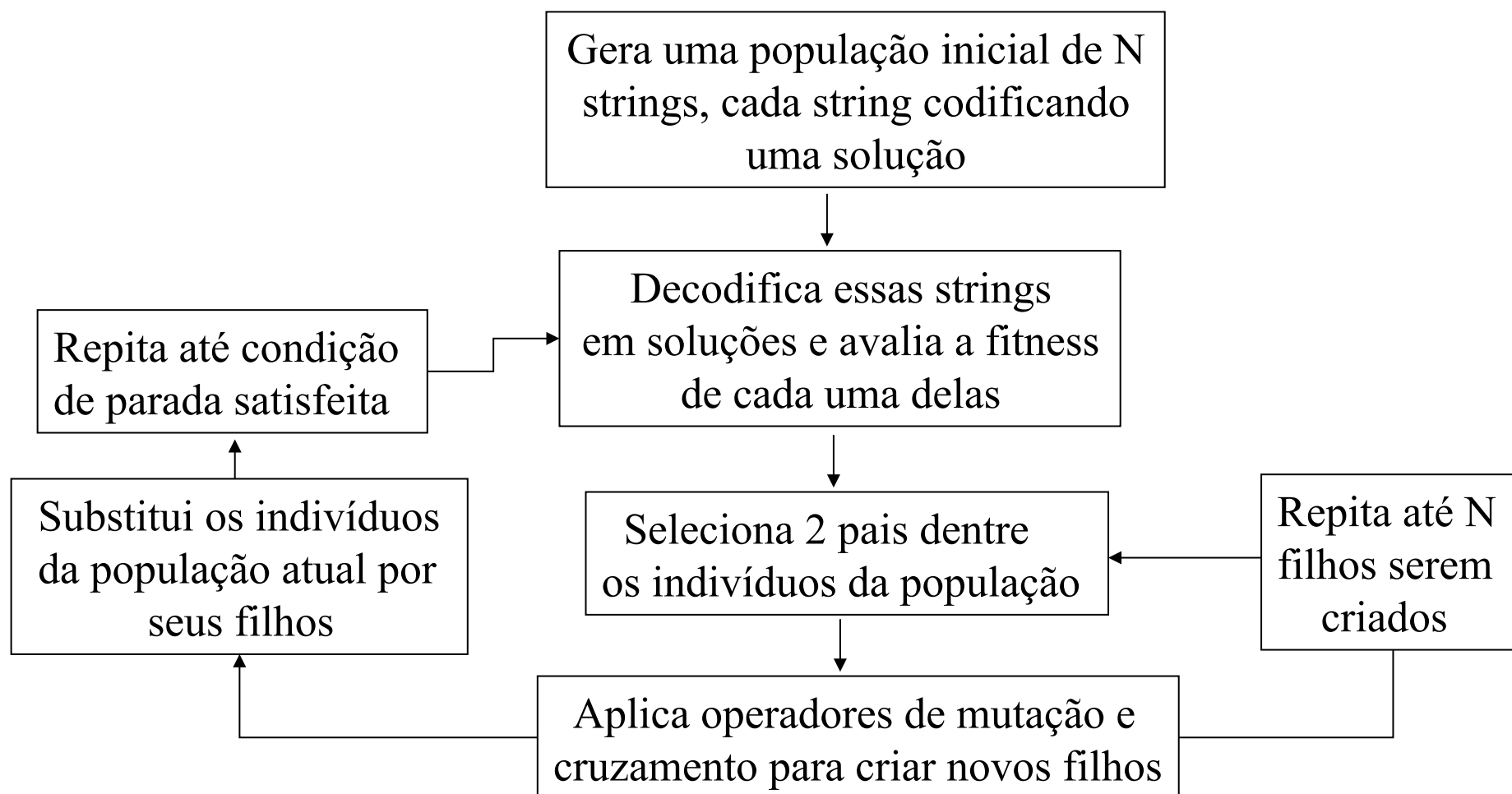
# Algoritmos Genéticos

- Operadores são aplicados sobre o genótipo
- O espaço do problema é conhecido como espaço de busca, e engloba todas as soluções possíveis para um determinado problema

# Algoritmos Genéticos

- Modo de operação:
  - *Steady state* versus geracional
    - Diferem pelo fato da população antiga e a nova se sobreporem ou não
- Geracional
  - A cada nova geração toda a população é substituída por uma nova (GA tradicional)
- *Steady state*
  - Apenas um ou dois indivíduos são produzidos a cada nova geração (mais comum em ES ou EP)

# Algoritmos Genéticos



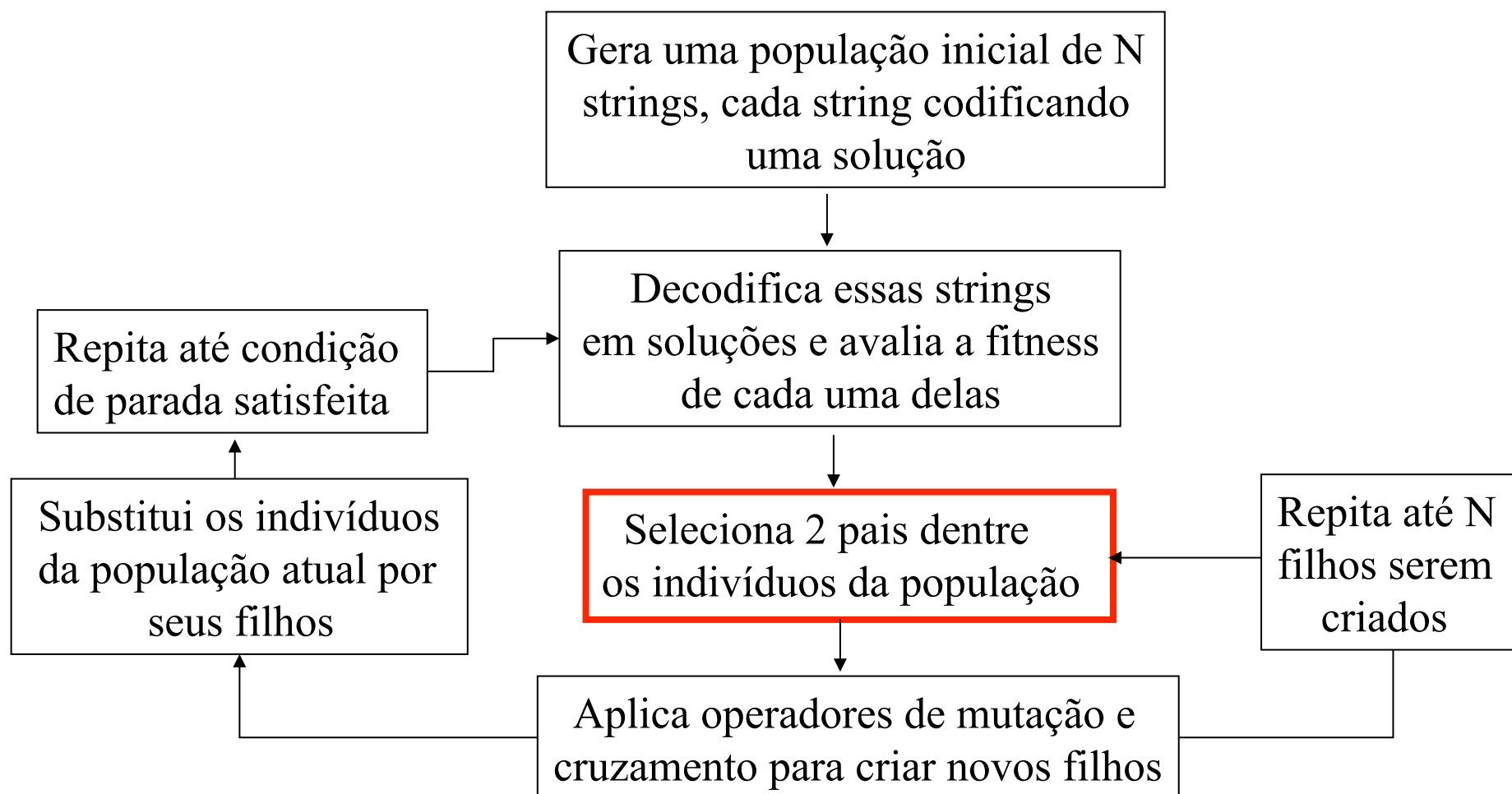
# Exemplo

- OneMax
  - Maximizar o número de 1s em um string de bits de tamanho  $n$
  - Definição de parâmetros
    - $N = 8$  e tamanho da população = 4
- Gerar população inicial
  - Atribuir aleatoriamente 1s e 0s a todos os locus
- Calcular o valor da fitness
  - Contar o número de 1s

Candidate	String	Fitness
A	00000110	2
B	11101110	6
C	00100000	1
D	00110100	3



# Algoritmos Genéticos



# Seleção Proporcional a Fitness (Roleta)

- Considere a fitness de um indivíduo  $i$  como sendo  $f_i$
- Fitness média da população pode ser calculada como

$$\bar{f} = \frac{1}{n} \sum_{i=1}^N f_i$$

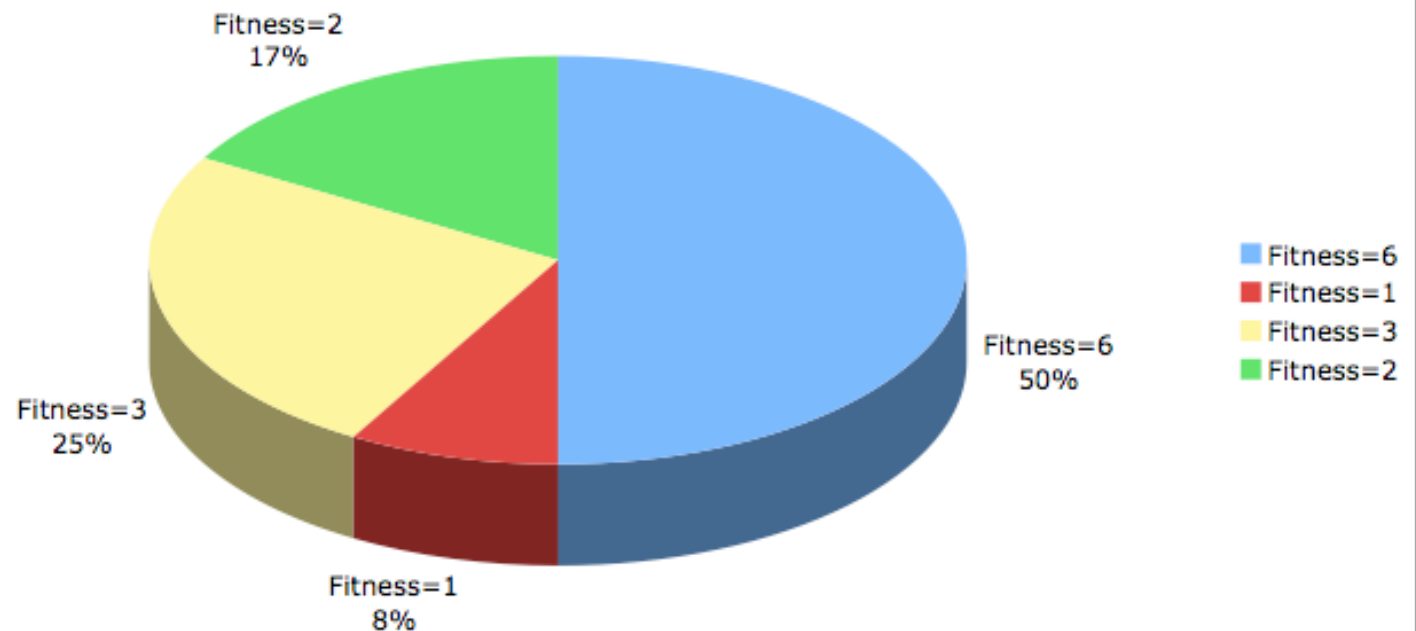
- Indivíduo  $j$  pode ser selecionado com probabilidade

$$p_j = \frac{f_j}{\sum_{i=1}^N f_i}$$

# Seleção Proporcional a Fitness (Roleta)

Candidate	String	Fitness
A	00000110	2
B	11101110	6
C	00100000	1
D	00110100	3

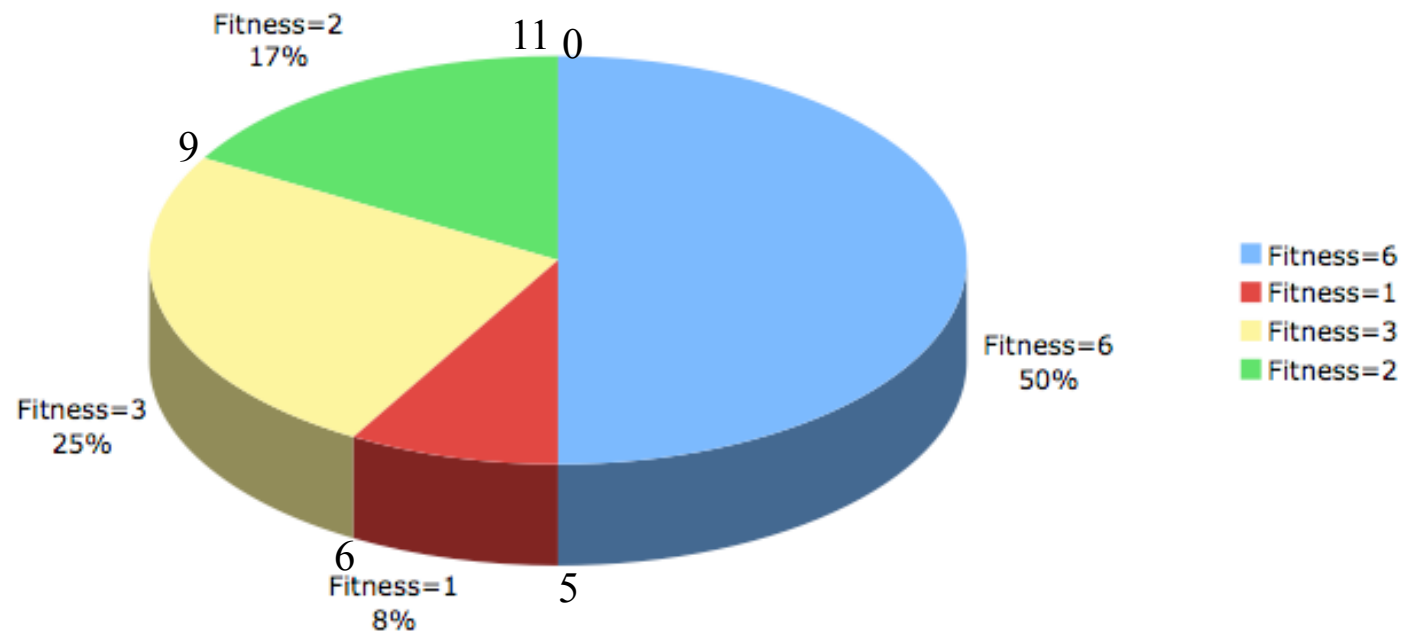
12



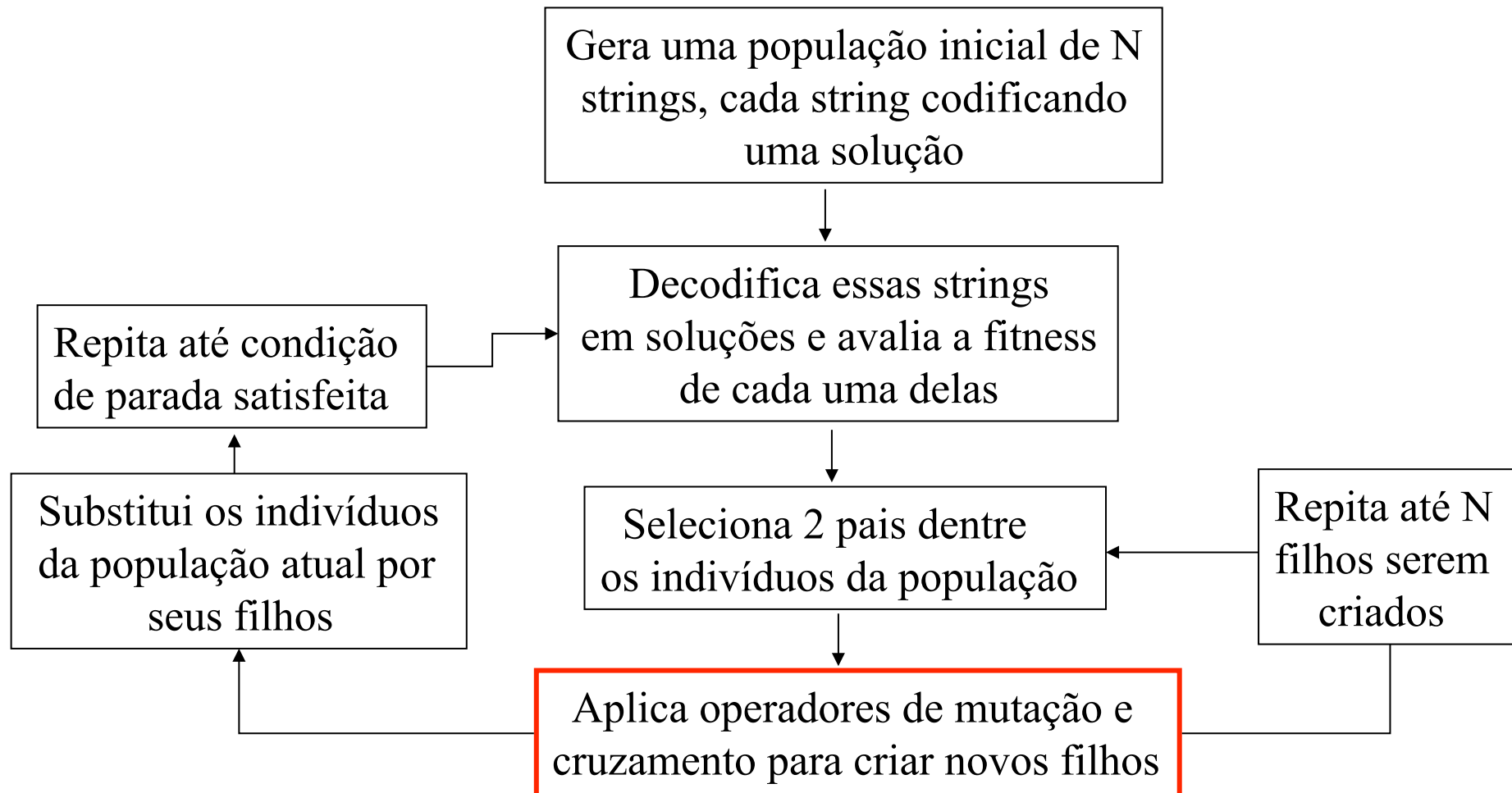
# Seleção Proporcional a Fitness (Roleta)

- Rodo a roleta

$$r \in \left[0, \sum_{j=1}^n f_i\right)$$



# Algoritmos Genéticos



# Operadores Genéticos

- Cruzamento de um ponto (de acordo com probabilidade definidas pelo usuário)
  - Padrão para GAs
  - Probabilidades altas (70-99%)
  - Ponto de cruzamento é escolhido aleatoriamente

<i>Initial Parent</i>	<b>Candidate B</b>	<b>Candidate C</b>
	1 1101110	0 0100000
<i>Resulting Child</i>	<b>Candidate E</b>	<b>Candidate F</b>
	0 1101110	1 0100000

# Operadores Genéticos

- Outro tipo de crossover: Crossover Uniforme
  - Cada gene é trocado de acordo com uma probabilidade  $p_c$

Pais

$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
$Y_1$	$Y_2$	$Y_3$	$Y_4$	$Y_5$	$Y_6$

Filhos

$X_1$	$Y_2$	$X_3$	$Y_4$	$Y_5$	$X_6$
$Y_1$	$X_2$	$Y_3$	$X_4$	$X_5$	$Y_6$

- Não existe bias posicional
  - Em crossover de um ponto a probabilidade de genes vizinhos serem trocados ao mesmo tempo é muito maior do que a de genes distantes serem trocados ao mesmo tempo

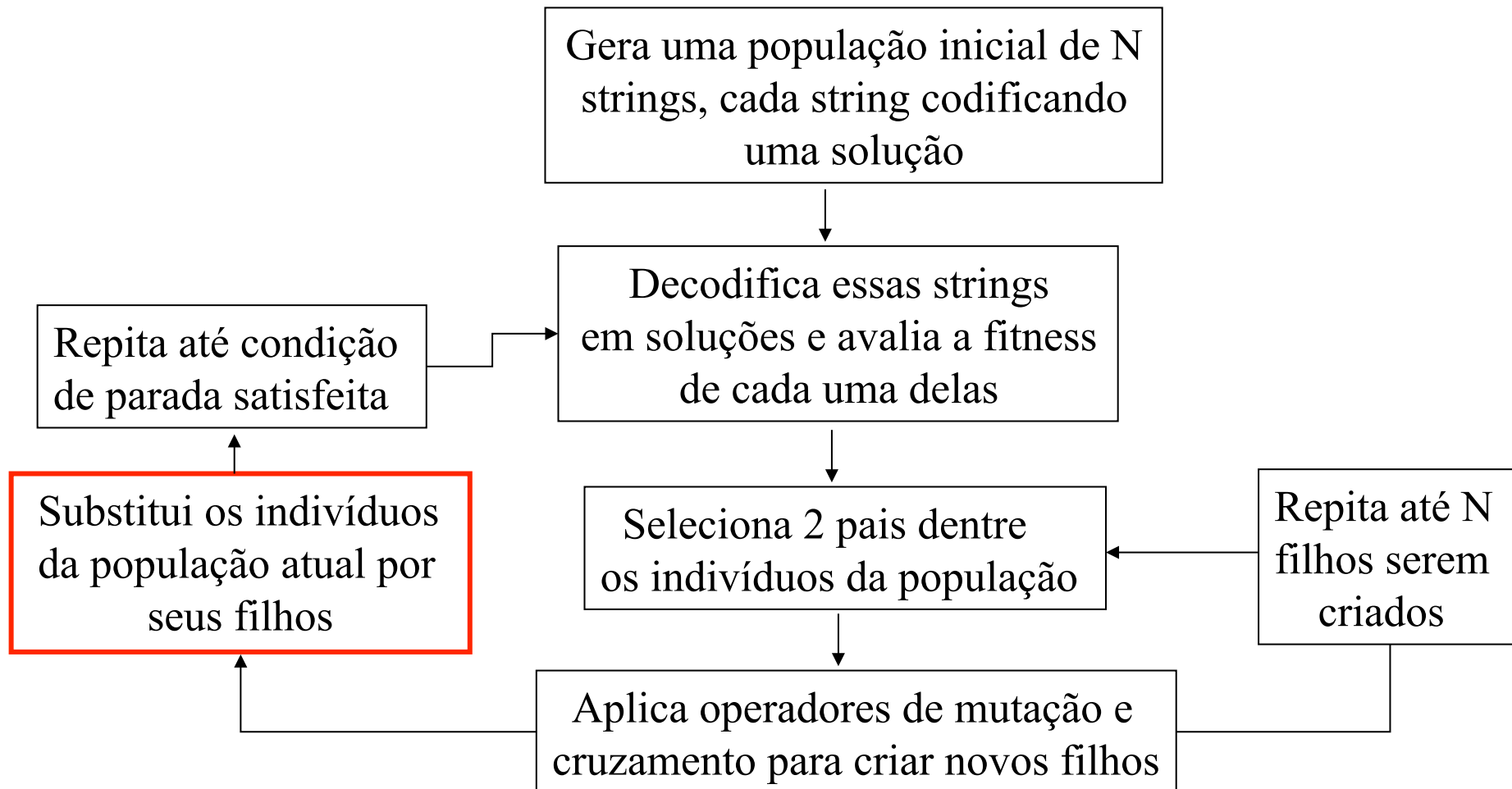
# Operadores Genéticos

- Mutação uniforme
  - Baixa probabilidade –em sistemas naturais os efeitos da mutação podem ser destruidores
  - Calcula a probabilidade de trocar cada um dos genes (bits) do cromossomo

<i>Initial Parent</i>	<b>Candidate B</b>	<b>Candidate C</b>
	1 1101110	0 0100000
<i>Resulting Child</i>	<b>Candidate E</b>	<b>Candidate F</b>
	01001110	10100000



# Algoritmos Genéticos



# Substituição da população atual pela nova

População  
atual

Candidate	String	Fitness
A	00000110	2
B	11101110	6
C	00100000	1
D	00110100	3



Nova  
população

Candidate	String	Fitness
E	01001110	4
F	10100000	2
G	11101110	6
H	00110100	3

# Papel dos operadores na Evolução

- Seleção
  - Guia o algoritmo para áreas promissoras do espaço de busca
- Crossover
  - Muda o contexto de informação útil já disponível
- Mutação
  - Introduz inovação
- Conflito entre o papel da seleção e do crossover e mutação????

# Seleção dos Indivíduos

- Equilíbrio entre explorar (crossover e mutação) o espaço de busca e se restringir aos indivíduos com boas fitness (seleção)
- Seleção pode determinar esse equilíbrio
  - Pressão seletiva (*selective pressure*)
- A fase de seleção determina a velocidade em que a evolução vai ocorrer
  - É uma consequência da competição

# Problemas da seleção por roleta

- Alta pressão seletiva no início da evolução
  - Leva a convergência prematura do algoritmo
- Baixa pressão seletiva no fim da evolução
  - Valores de fitness similares
  - Probabilidades de seleção uniformes
  - Um solução um pouco melhor é favorecida
- Exige computação de estatísticas globais

# Seleção por Ranking

- Ordena os indivíduos por fitness
- Calcula a probabilidade do indivíduo ser selecionado de acordo com seu lugar no ranking
  - Quanto maior o ranking maior a probabilidade de seleção
- Várias maneiras de atribuir as probabilidades a cada indivíduo do ranking
  - Linear
  - Não-linear

# Seleção por Ranking

- Linear

Número de filhos que se deseja  
que o melhor indivíduo crie a  
cada geração

$$prob(i) = \alpha_{rank} + [rank(i) / (popsize - 1)](\beta_{rank} - \alpha_{rank}) / popSize$$

Número de filhos que se deseja  
que o pior indivíduo crie a cada  
geração

- Resolve a maioria dos problemas de seleção por roleta, mas ainda depende de estatísticas globais

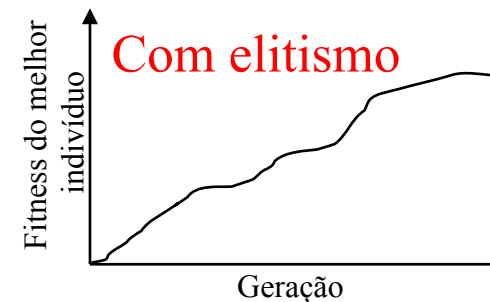
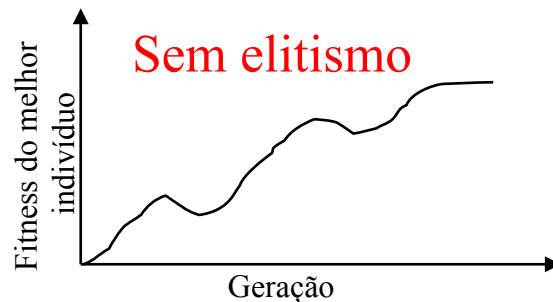
# Seleção por Torneio

- Um pequeno subconjunto de  $k$  indivíduos é retirado aleatoriamente da população, e o melhor indivíduo desse subconjunto é selecionado (vencedor do torneio)
  - $k$  = tamanho do torneio
- Quanto maior o valor de  $k$ , maior a pressão seletiva
  - Pressão seletiva pode ser facilmente regulada
  - Não depende de uma estatística global
    - Acelera evolução
    - Torna paralelização mais fácil
- Tornou-se um dos métodos mais comuns



# Seleção dos Indivíduos

- **Problema:** o melhor indivíduo de uma geração pode morrer sem se reproduzir porque:
  - Existe um processo de seleção probabilística
  - Os operadores não garantem a geração de indivíduos melhores que seus pais
- **Solução:** elitismo
  - O melhor indivíduo de cada geração é copiado sem alteração para a próxima geração (elitismo complementa outras técnicas de seleção)



# Seleção de indivíduos

- Ocorre em 2 fases
- Fase de reprodução:
  - Pais são selecionados para gerar filhos a partir de um algoritmo de seleção que considera toda a população
- Fase de remoção:
  - Decisão de que indivíduos remover para inserir novos indivíduos na população
    - Geracional: todos indivíduos da população atual são removidos

# Parâmetros

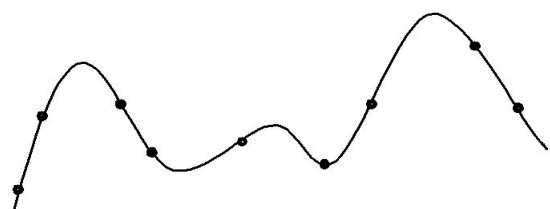
- Tamanho da população
- Número de gerações
- Probabilidades de cruzamento
  - Se uniforme, probabilidade de trocar gene (0.5)
- Probabilidades de mutação
- Se seleção por torneio – k
- Número de indivíduos do elitismo

# Avaliando a *Performance* de um GA

- Nunca tire conclusões a partir de uma execução
- Rode o GA várias vezes
  - Para problemas simples idealmente pelo menos 30
- Utilize medidas estatística (médias, medianas, desvio padrão, etc)
- Salve o maior número de informações possíveis sobre sua população
  - Média, Melhor e Pior fitness a cada geração
  - Diversidade, etc
- Desenhe gráficos para acompanhar o progresso dessas variáveis
- Compare com um algoritmo de busca aleatória

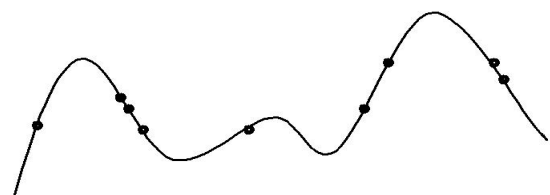
# Comportamento Típico de um EA

## Otimização de uma função de *fitness* 1D



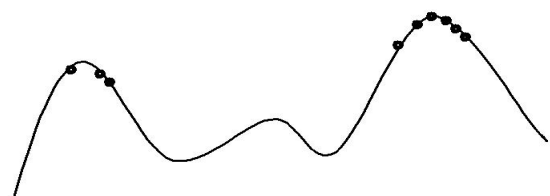
Fase Inicial:

Distribuição da população quasi-random



Fase intermediária:

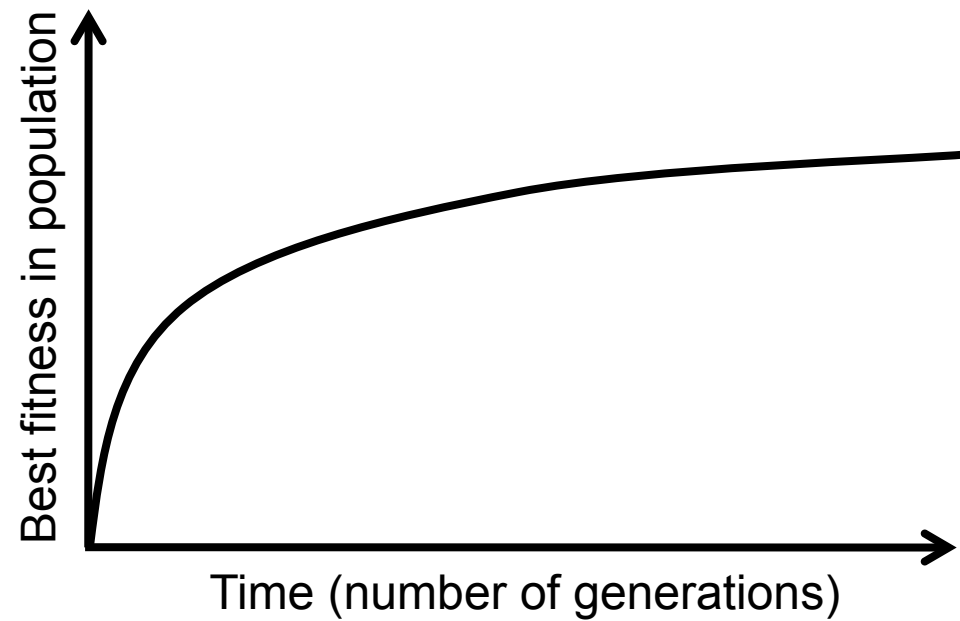
População presente nos/em torno dos picos



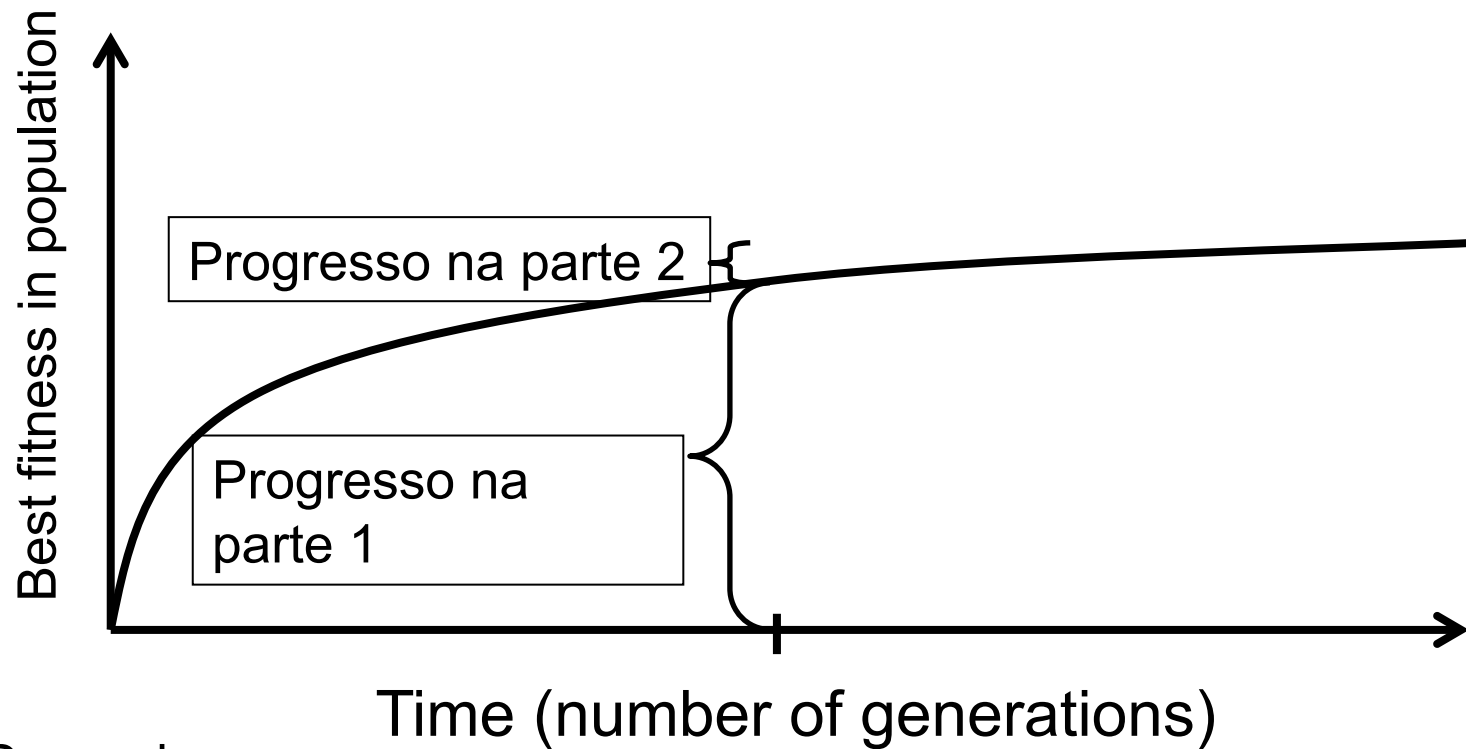
Fase final:

População concentrada nos pico mais altos

# Comportamento da *Fitness*

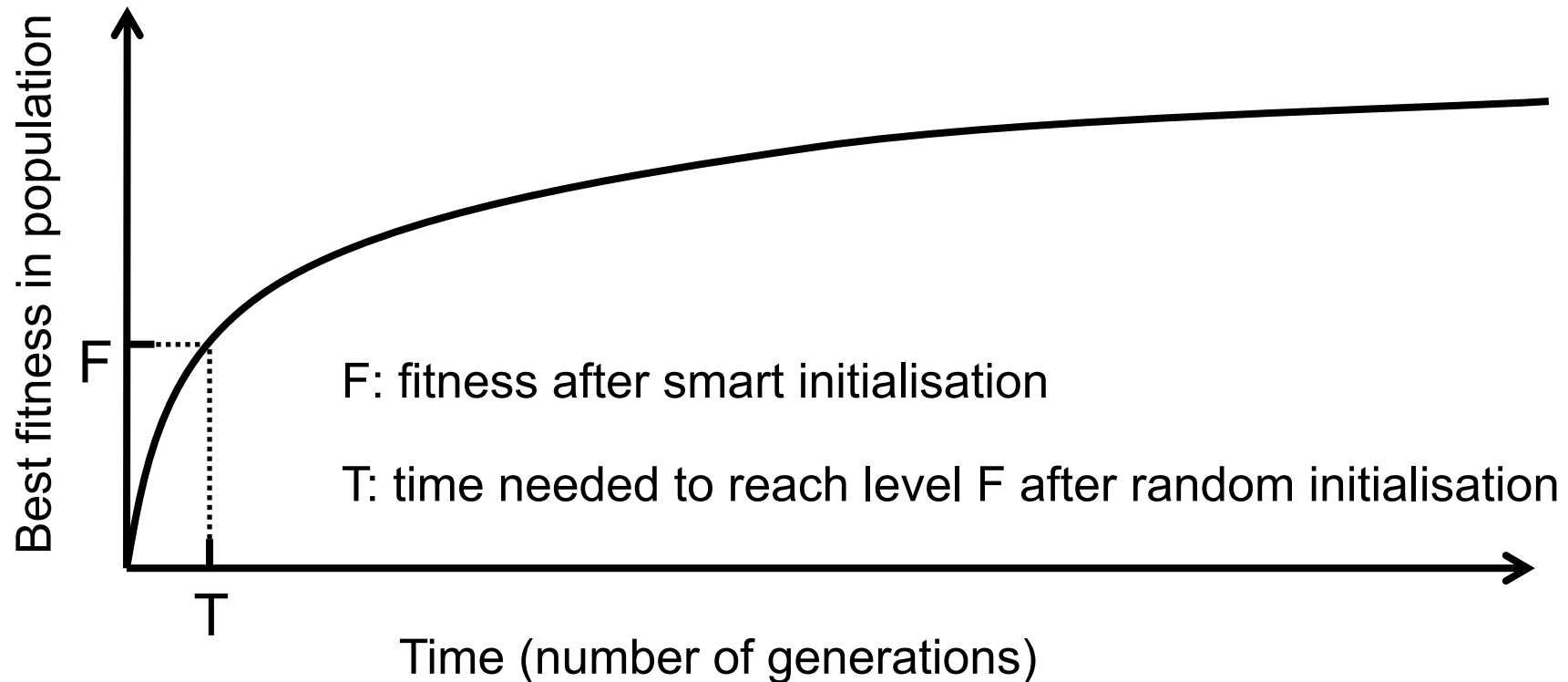


# Vale a pena rodar por muito tempo?



- Depende:
  - de tamanho do progresso na segunda parte
  - Pode ser mais apropriado ter mais gerações com menos indivíduos

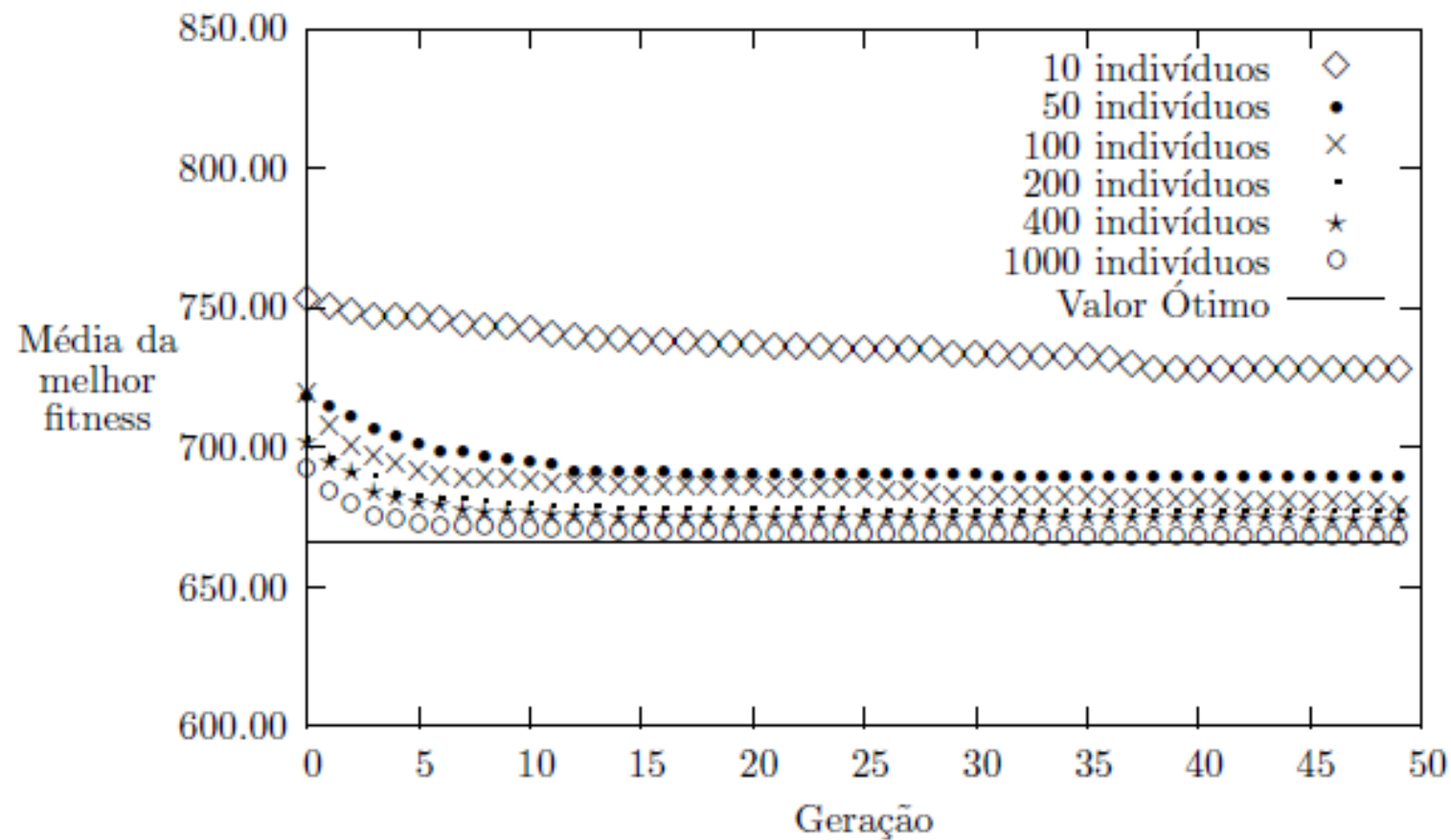
# Vale a pena o esforço de uma inicialização inteligente?



- Depende:
  - se soluções boas já forem conhecidas.
- Deve-se tomar cuidado para não enviesar a população



# Avaliando a *Performance* de um GA



# Avaliando a *Performance* de um GA

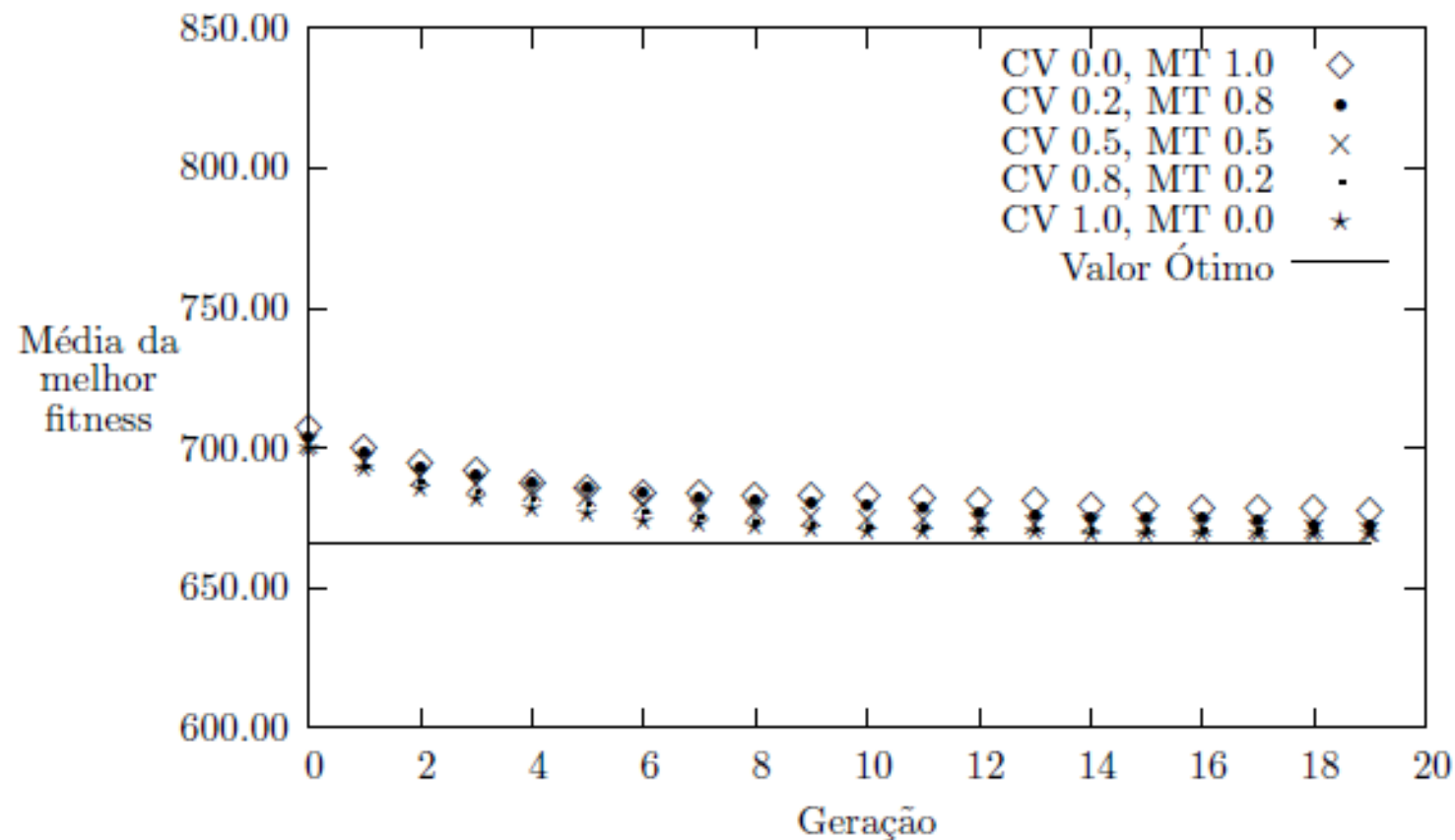


Figura 14: Variando probabilidades de Crossover e Mutação com elitismo.

# Avaliando a *Performance* de um GA

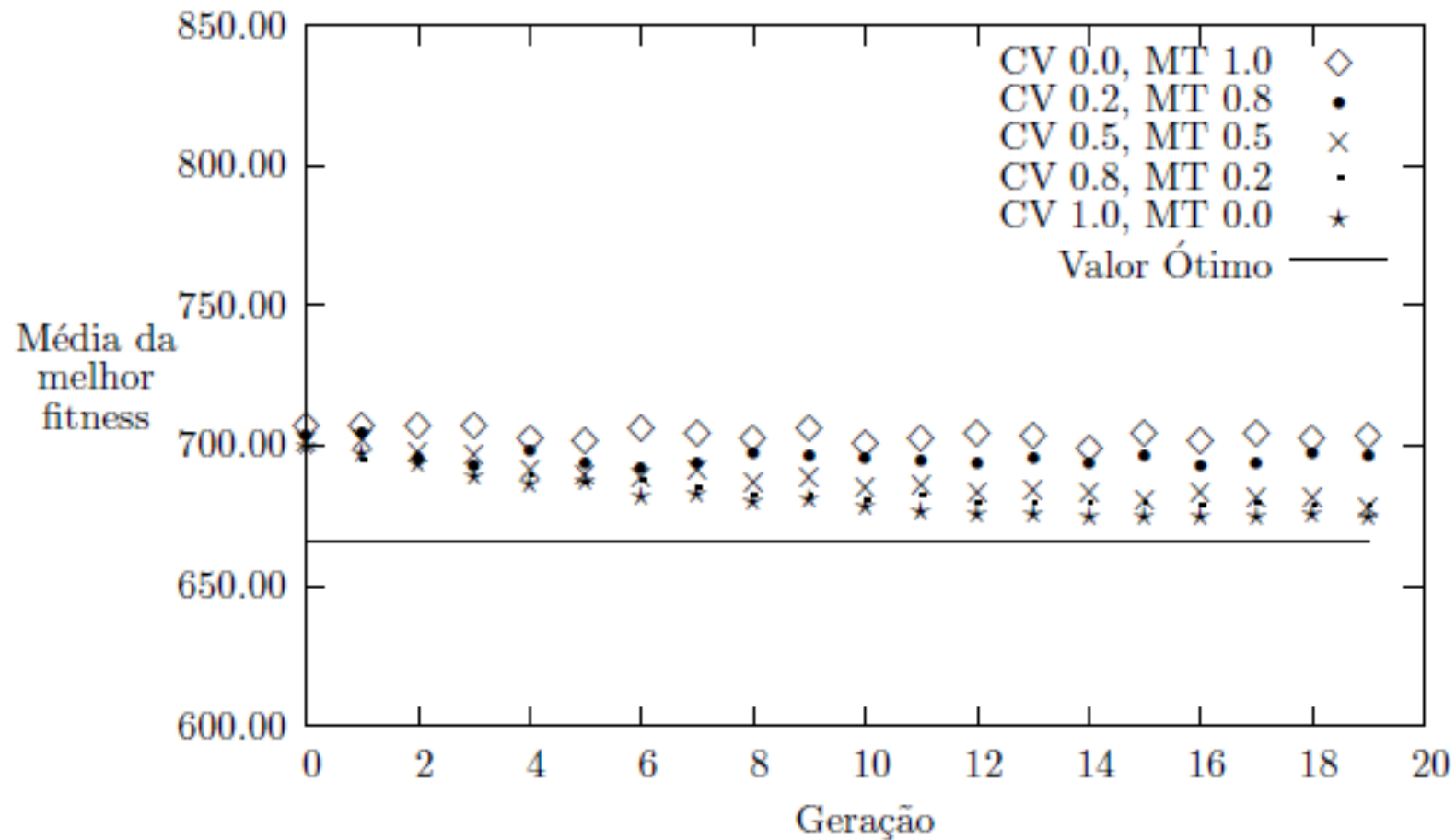


Figura 15: Variando probabilidades de Crossover e Mutação sem elitismo.

# Pontos Importantes em GAs

- Decisões de Design
  - Representação do Indivíduo
  - Método de seleção (quanta pressão seletiva?)
  - Escolha dos operadores de mutação e crossover
  - Tamanho da população
  - Número de gerações fixas ou outro critério de parada?
  - Probabilidades de crossover e mutação
- Tópicos avançados
  - Busca local
  - Nichos e Espécies
  - Co-evolução
  - ...

# Agradecimentos

- Alguns desses slides foram retirados das notas de aula de Alex A. Freitas e Michael O'Neil
- Os gráficos apresentados foram retirados de um trabalho prático da turma de 2009, feito pelos alunos Thiago Salles e Cristiano Nascimento