

Experimentação, Parametrização e Análise de Algoritmos Bio-inspirados Parte 1/2

Gisele L. Pappa

Introdução

- Ao desenvolver um novo algoritmo, como sabemos se ele é bom o suficiente para resolver meu problema?
- Como escolho os valores para os parâmetros?
- Como comparo meu algoritmo com outros métodos?

Como saber se meu algoritmo funciona?

- Algoritmos de computação natural normalmente são utilizados para resolver problemas NP-completo ou NP-hard
- O sucesso depende fortemente dos parâmetros
 - Não espere que os primeiros experimentos que você rodar deem certo
- Nem sempre o algoritmo vai funcionar
 - O problema pode ser mais difícil que o antecipado

Como rodar experimentos ?

- Os experimentos não podem ser executados apenas uma vez
 - Em otimização, o comum são 30 vezes
 - Em aprendizado, 5 ou 10 vezes
- Porém, eles devem ser reprodutíveis
 - Para isso, você deve conhecer a semente aleatória utilizada em cada execução
- Compare com um algoritmo de busca aleatória

Como rodar experimentos ?

- Como começar?
 - Escolhendo um parâmetro para variar, deixando todos os outros fixos
 - Existem alguns valores padrão para cada parâmetro, que podem ser utilizados na primeira rodada
- Normalmente se começa variando o tamanho da população, e depois os operadores genéticos
- No caso de GP o tamanho máximo da árvore do indivíduos normalmente é variado antes dos operadores

Como rodar experimentos ?

- Parâmetros “padrão” variam para problemas de otimização versus aprendizagem
 - População :
 - entre 30 e 1000 indivíduos para aprendizagem
 - Entre 100 e 10000 indivíduos para otimização
 - Geração: 50 a 5000? Depende do custo da fitness
 - Operadores
 - Crossover de 90-99% em GA e variando de 60-90% em GP
 - Mutação variando de 1 a 10% em GA e 10-30% em GP

Como rodar experimentos ?

- Depois de “otimizar” um parâmetro, este se torna fixo e variamos os outros
- Normalmente, o tamanho da população, número de gerações e a taxa de mutação tem maior influência no algoritmo
 - Quando utilizamos seleção por torneio, o k também é primordial
- Em GP, o tamanho da árvore também tem influência grande

Análise do Algoritmo

- Monitorar a fitness
- Monitorar a diversidade da população
- Monitorar o número de operações de cruzamento e mutação que geram indivíduos melhores que os pais
 - Utilize medidas estatística (médias, medianas, desvio padrão, etc)
 - Desenhe gráficos para acompanhar o progresso dessas variáveis

Análise da População

- Elemento mais importante a ser analisado
- Avaliação básica: plotar fitness versus geração
- Verificar também
 - Distribuição do tamanho dos indivíduos, se ele for variável
 - Distribuição dos valores de fitness durante cada geração
 - Pode dar dicas a respeito da estrutura do espaço de busca
 - Se existe uma variação muito grande entre os valores, podem ter regiões do espaço que não estão sendo descobertas

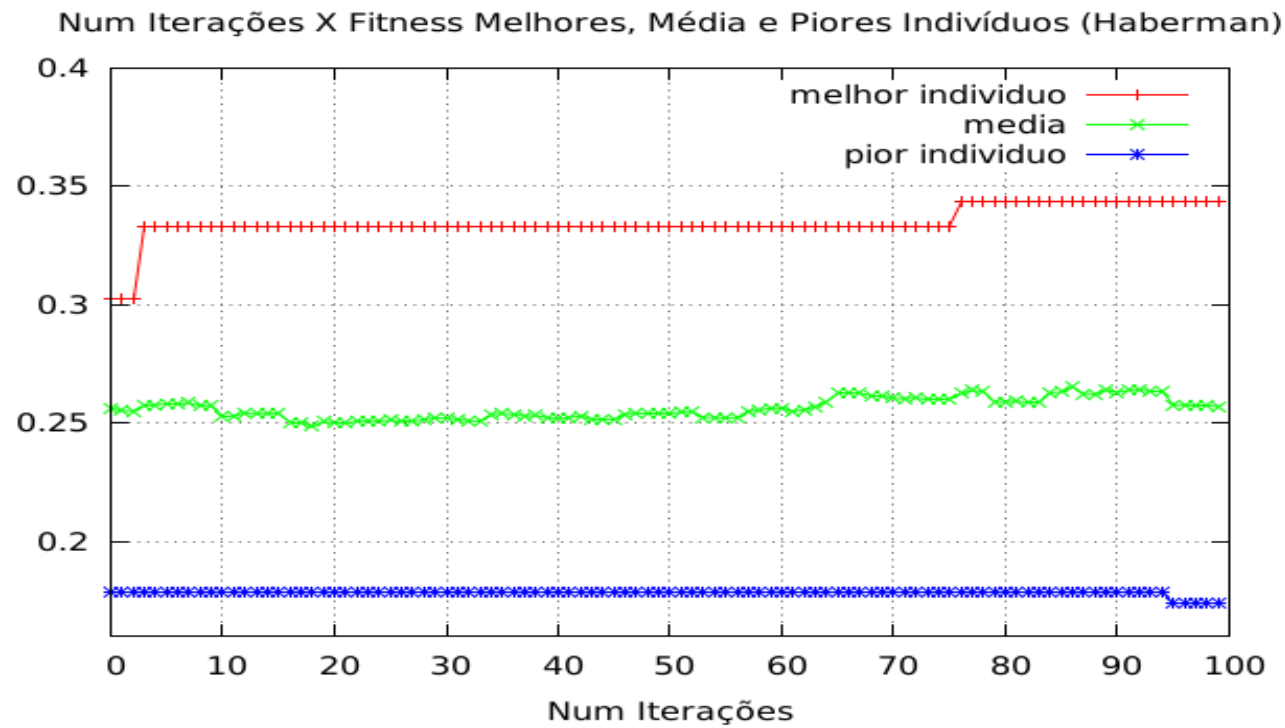
Análise da População

- Para isso, você deve guardar em um log o máximo de informações possíveis.
- Vizualizar os dados onde a fitness está sendo calculada também pode ser de grande ajuda

Diversidade

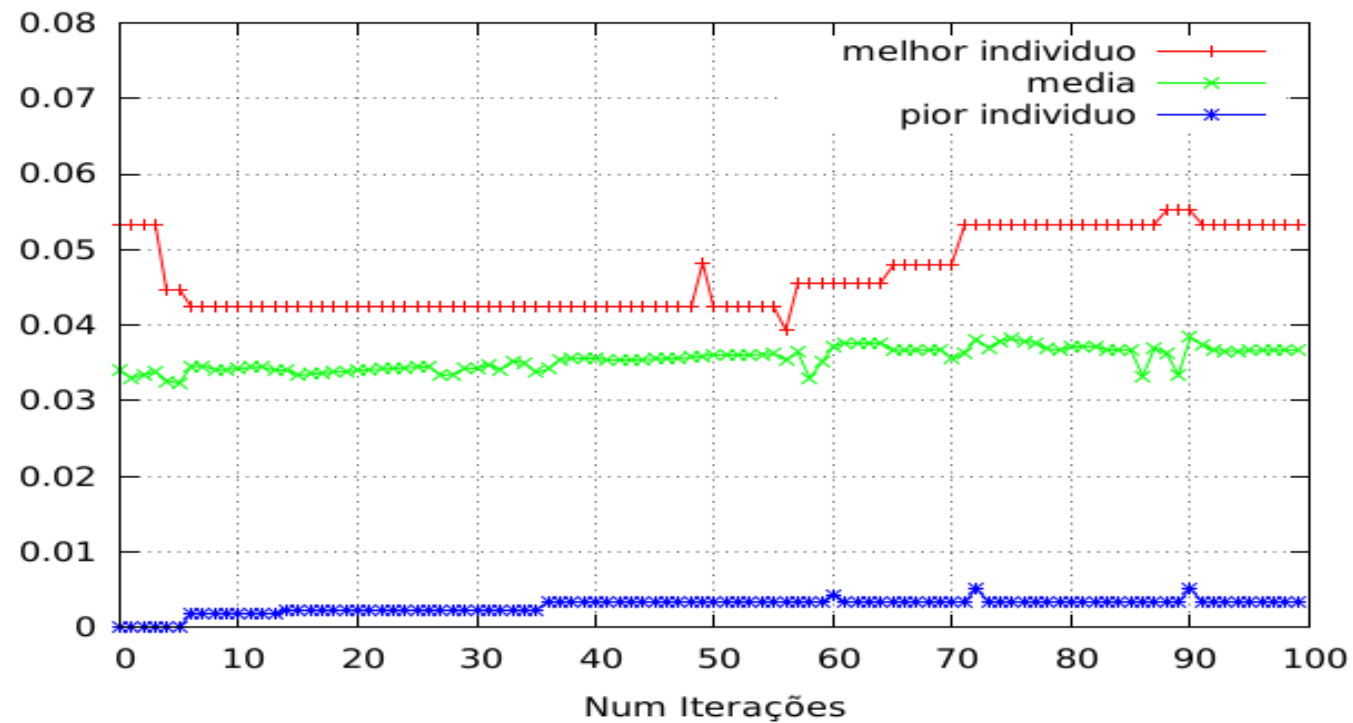
- Se mais de 90% dos indivíduos da população são iguais em um tempo muito curto, isso pode indicar um problema
- Ao mesmo tempo, muita diversidade também pode representar um problema
- Por isso, se possível, medir a distância de fenótipo entre dois indivíduos pode ser muito importante

Exemplos

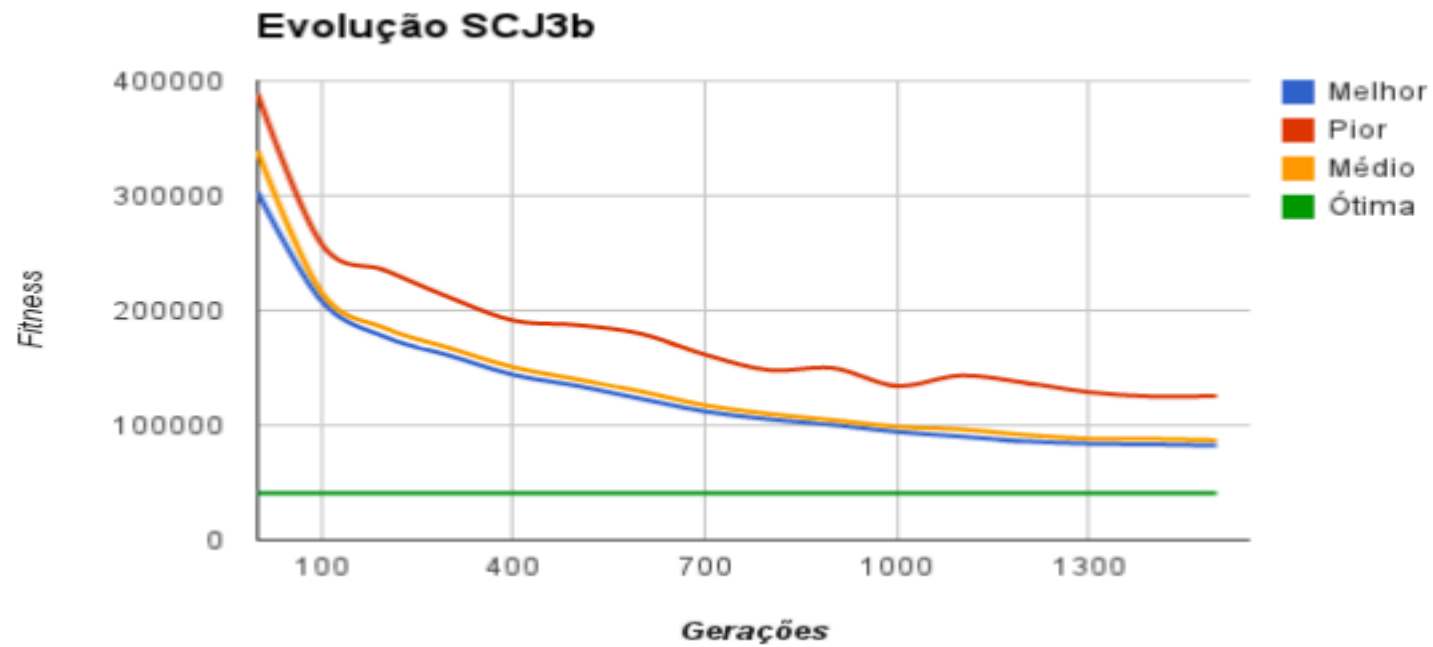


Exemplos

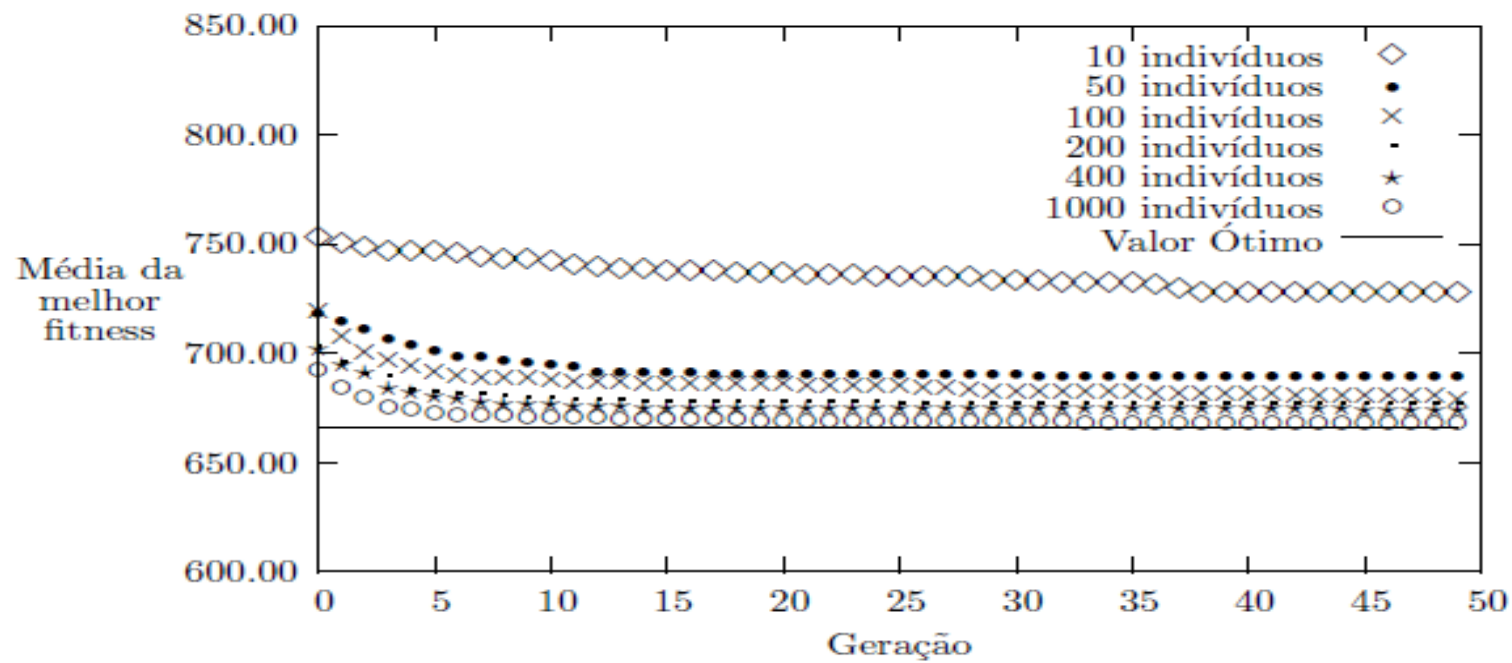
Num Iterações X Fitness Melhores, Média e Piores Indivíduos (base1_p4_fitne



Exemplos



Avaliação dos Resultados de um EA



Avaliação dos Resultados de um EA

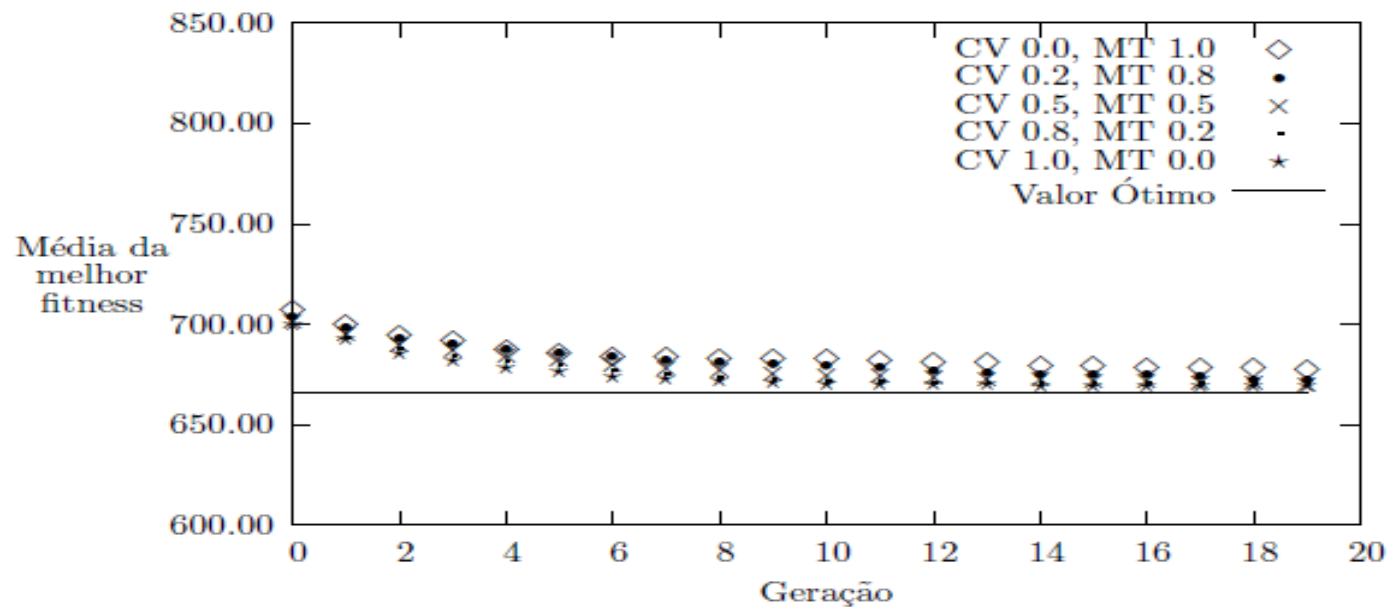


Figura 14: Variando probabilidades de Crossover e Mutação com elitismo.

Avaliação dos Resultados de um EA

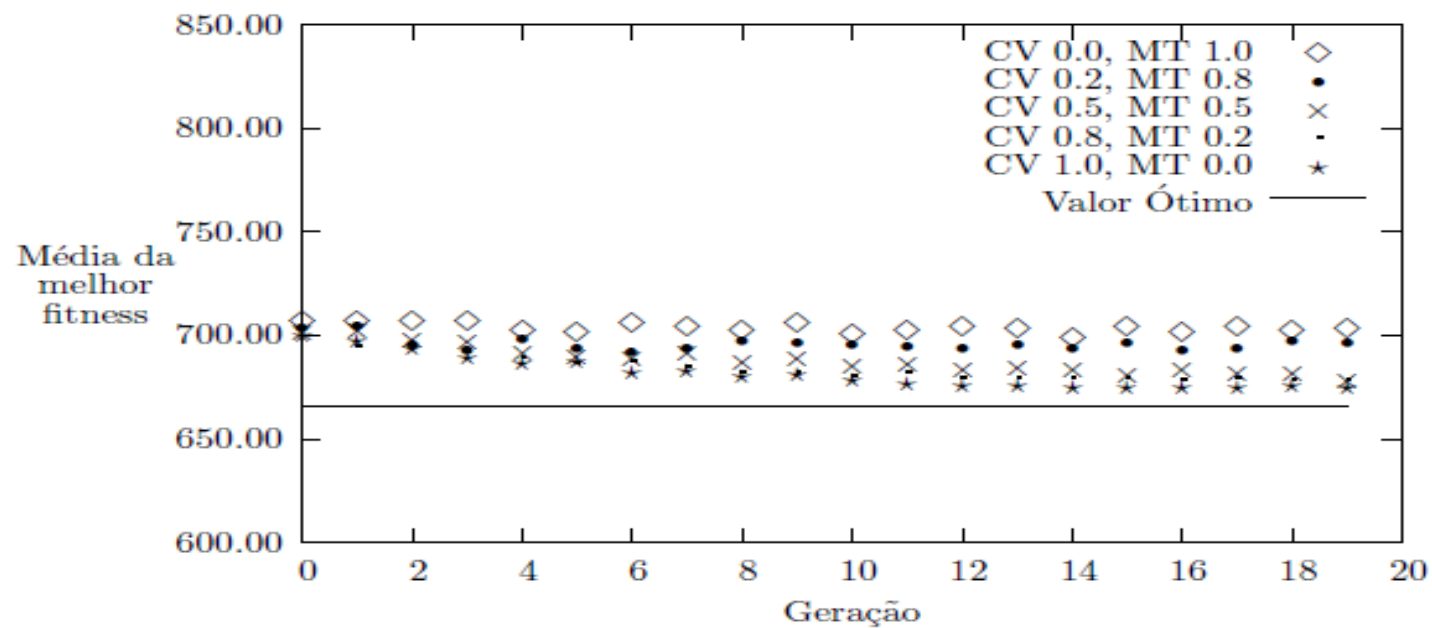


Figura 15: Variando probabilidades de Crossover e Mutação sem elitismo.

O que fazer em caso de convergência prematura?

- Não utilizar o operador de reprodução
- Diminuir a pressão seletiva
 - Por exemplo, diminuir o número de indivíduos do torneio
- Adicionar diferentes tipos de operadores de mutação
- Introduzir conceitos de espécies
- Utilizar fitness sharing

Mudanças pequenas podem causar grandes impactos

- Existem muitos parâmetros que influenciam o sucesso dos algoritmos
- Mudando um ou dois deles pode não ter nenhum impacto no seu algoritmo, porém pode **causar um grande impacto**
 - Mudar um terminal pode mudar tudo
- O mesmo é válido para pequenas mudanças no código
 - Trocar um `>` por um `>=` pode trocar totalmente os vencedores de um torneio em algoritmos evolucionários

Mudanças grandes podem não causar nenhum impacto

- Se você mexe em todos os parâmetros e nada muda no algoritmo, o que isso significa?
 - Sua representação não é apropriada
 - A fitness não é apropriada
- Se a fitness é apropriada
 - Aumentar o tamanho da população, em algum momento, vai fazer com que a fitness melhore

Leitura Recomendada

Parameter Control in Evolutionary Algorithms, Ágoston E. Eiben , Robert Hinterding , Agoston E. Eiben Robert Hinterding , Zbigniew Michalewicz, IEEE Transactions on Evolutionary Computation, 2000

Experimentação, Parametrização e Análise de Algoritmos Bio-inspirados Parte 1/2

Gisele L. Pappa