



# PEMARF – aula 09

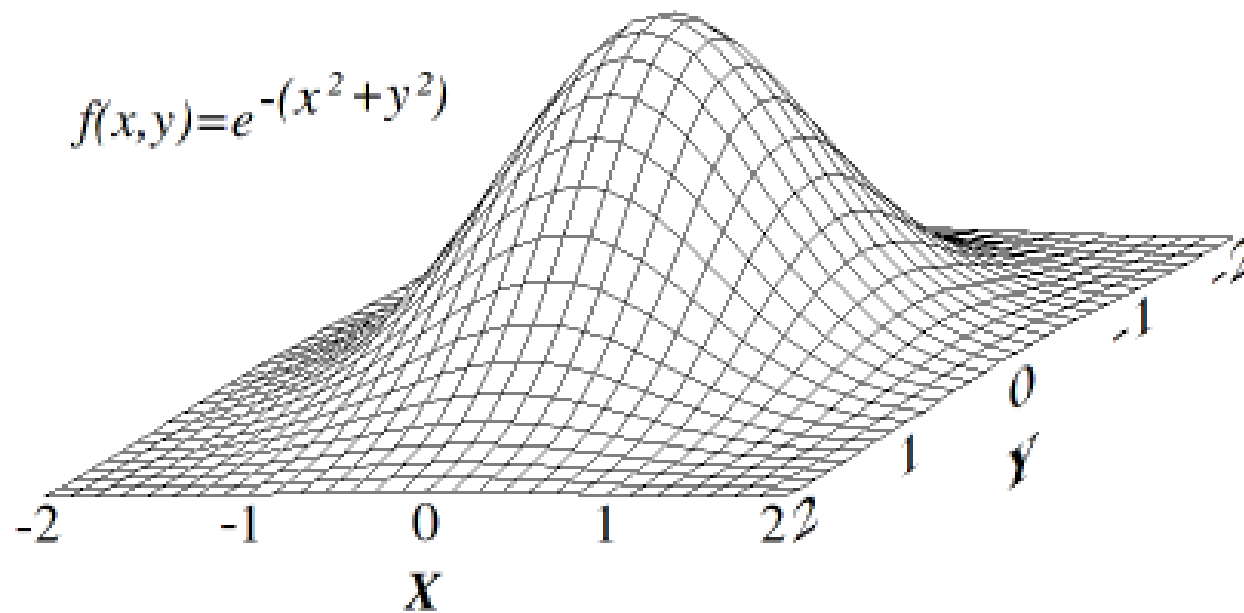
- Algoritmos de busca
- Discussão
- Intervalo
- Seminário
- Correção
- Escrita

# Busca

- Métodos numéricos
- Construção direta vs critério de optimalidade
- Hill-climbing combinado com critério

# Hill-Climbing

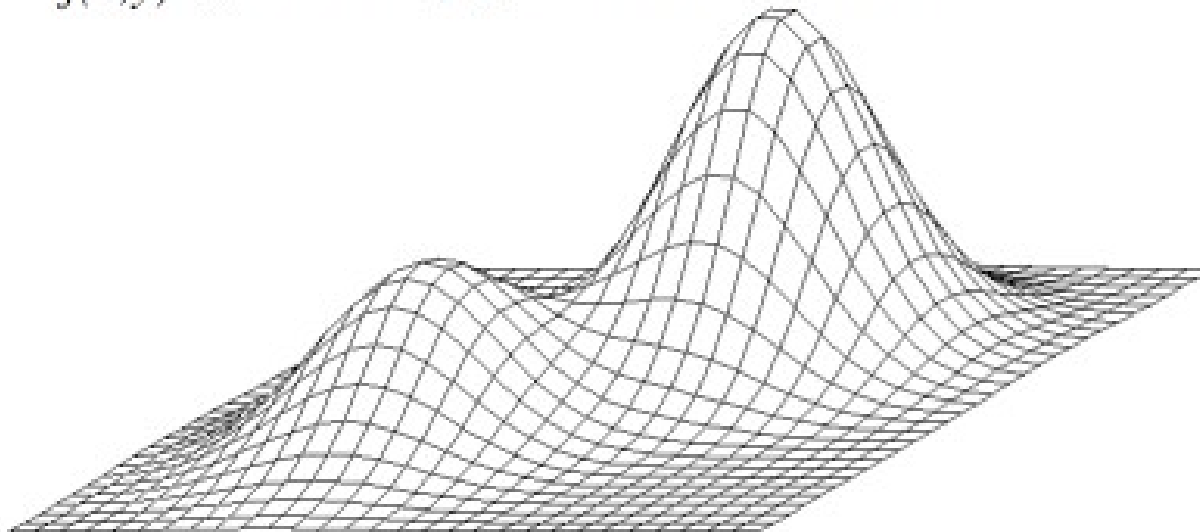
- Iterações buscando otimização do critério (ML, MP)



# Hill-Climbing

- Iterações buscando otimização do critério (ML, MP)

$$f(x,y) = e^{-(x^2+y^2)} + 2e^{-((x-1.7)^2+(y-1.7)^2)}$$



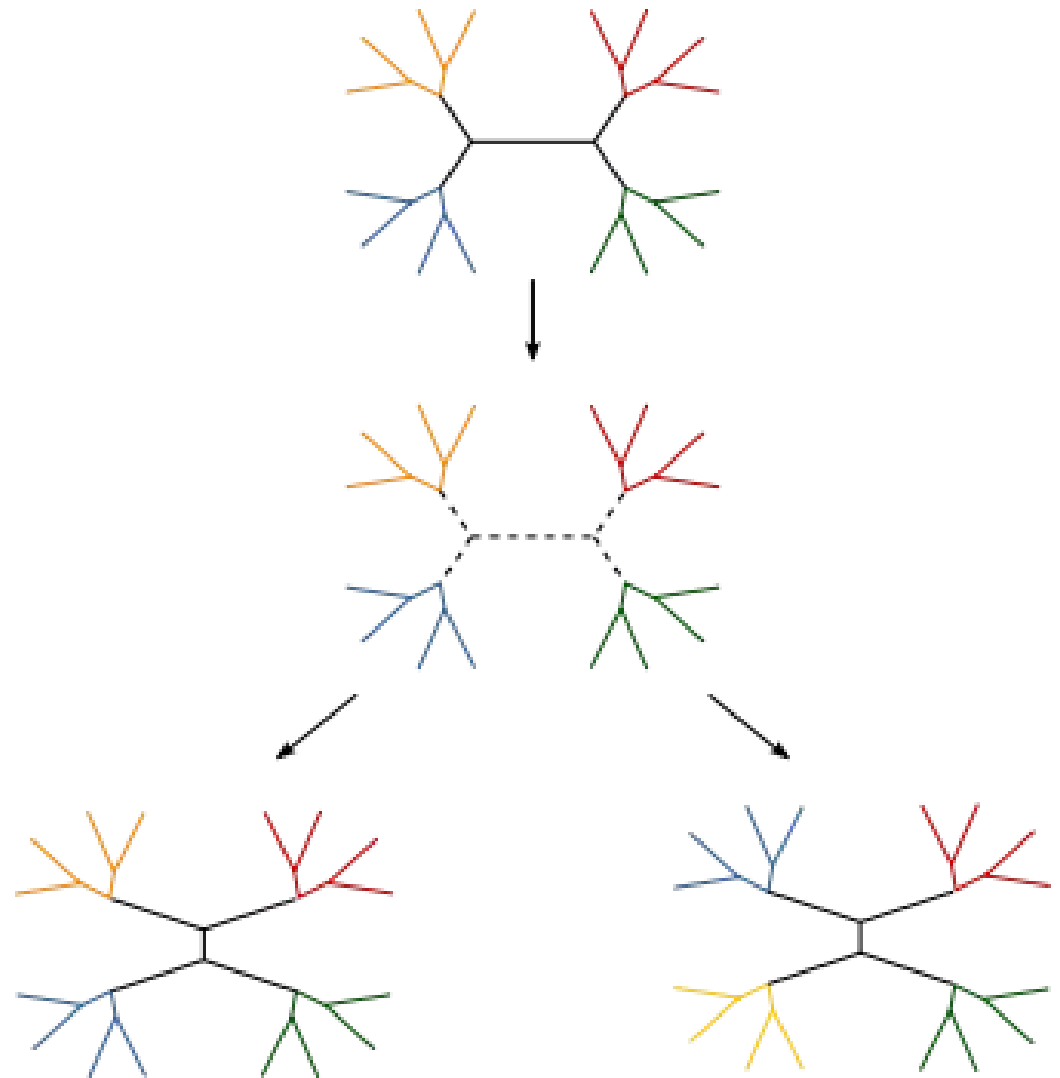
- Truques para explorar o espaço de parâmetros

# Principais algoritmos de busca

- Exploradores
  - NNI e Quartet Puzzle
  - TBR e SPR
- Refinadores
  - TF
  - Ratchet

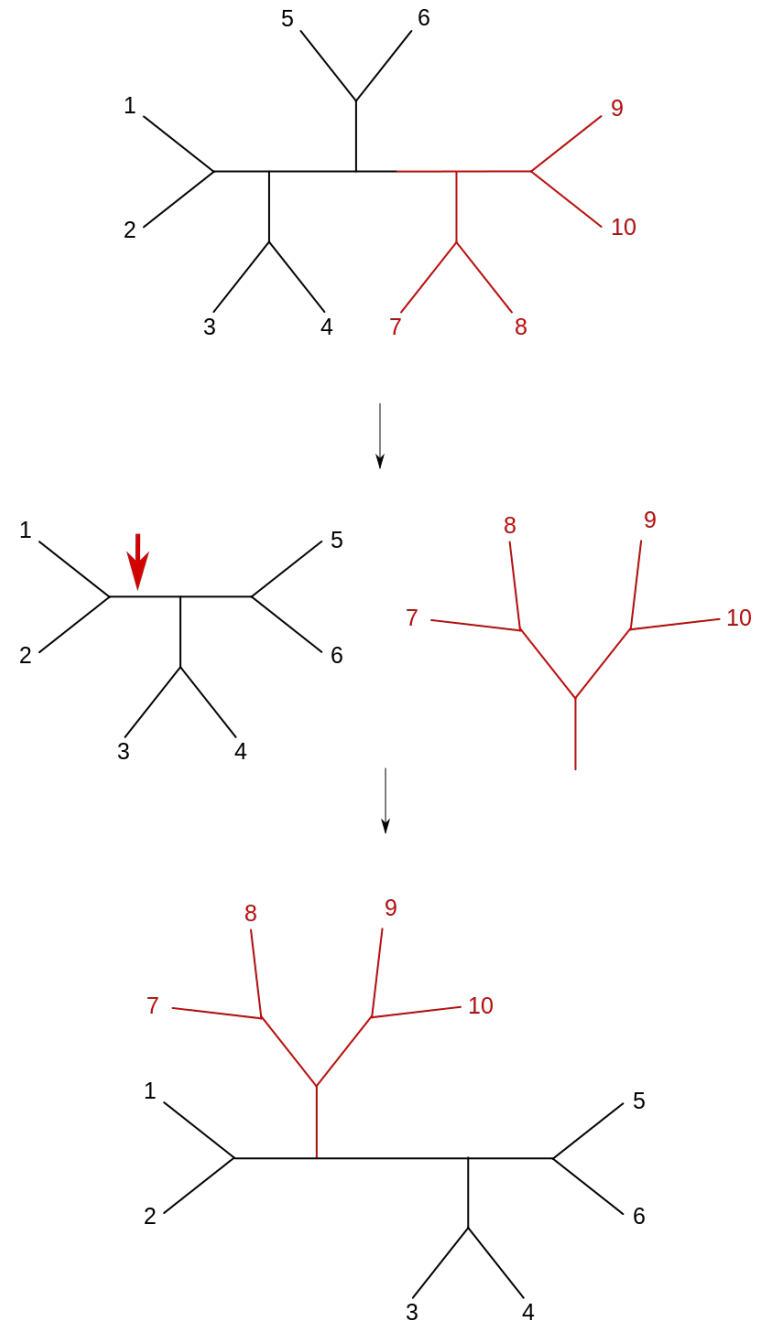
# Quartet Puzzle e Nearest Neighbor Interchange (NNI)

- QP =
  - Criar quartetos
  - Aplicar NNI
  - Otimizar Lh (ou Pa)
- Rápido, explora bem árvores sub-ótimas, pode não encontrar a melhor



# SPR e TBR

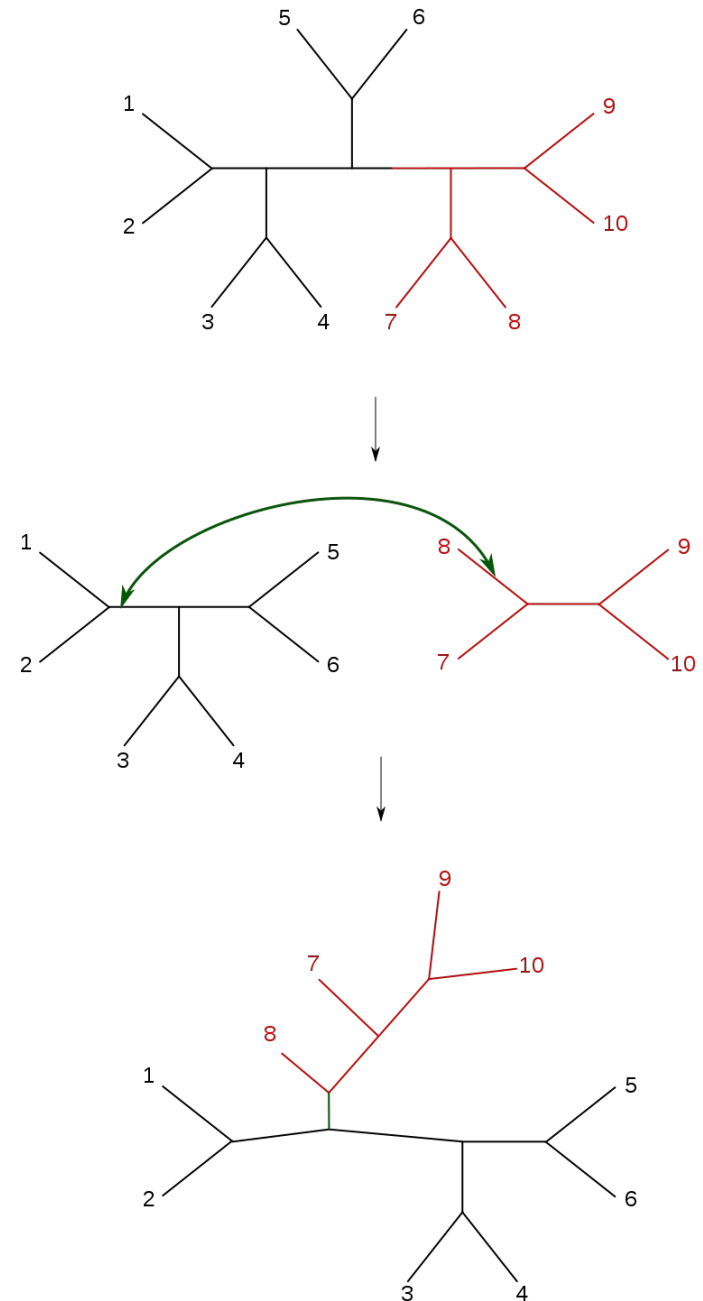
- Subtree Pruning and Regrafting
  - Tira um ramo
  - Re-insere em todos os lugares possíveis
  - Otimiza Lh ou Pa em todas
- Relativamente rápido, consegue em muitos casos resolver o problema. “Greedy”





# SPR e TBR

- Tree Bisection and Reconnection
  - Tira um ramo
  - Re-enraíza em todos os lugares possíveis da sub-árvore
  - Re-insere todos os re-enraizamentos em todos os lugares possíveis
  - Otimiza Lh ou Pa em todas
- Mais lento e melhor que SPR. Mais “Greedy”



# Refinadores

- Ratchet (apenas parcimônia)
  - Cria novos pesos para os caracteres por um tempo
  - Faz hill-climbing
  - Retorna aos pesos originais e retoma a busca
- Tree-fusing e Genetic Algorithms
  - Branch swapping ENTRE árvores
- Simulated Annealing
  - Inverter o Hill-Climbing por um tempo

# Tamanho do Problema

- 2007 – datasets com 500-1,000 taxa (bactérias já com 10,000), no máximo 100 genes (tipicamente 5-10)
  - POY, TNT, PHYML, PAUP\*, MrBAYES início do RAXML
- 2017 – datasets com milhares de taxa, comumente com mais de 1,000 genes
  - RAXML, IQTREE, PHYLOBAYES, PAUP\*

# Uso dos métodos

Programa	Total	2007	2016
PAUP*	21,000	5,270	3,080
PHYML	17,600	834	2,800
RAXML	15,300	203	4,240
TNT	4,760	147	627
PUZZLE	2,360	381	107
POY	1,500	149	111
IQTREE			546