

# arvore\_de\_maxima\_parcimonia

Alice

June 16, 2015

## Contents

<b>1</b>	<b>Enunciado</b>	<b>1</b>
<b>2</b>	<b>Dados reais</b>	<b>2</b>

## 1 Enunciado

*Reconstruindo árvores filogenéticas pelo método da máxima parcimônia*

Uma árvore filogenética é uma árvore em que as folhas são espécies diferentes, cuja raiz é o ancestral comum de tais espécies e cujos galhos representam eventos de especiação.

Queremos achar:

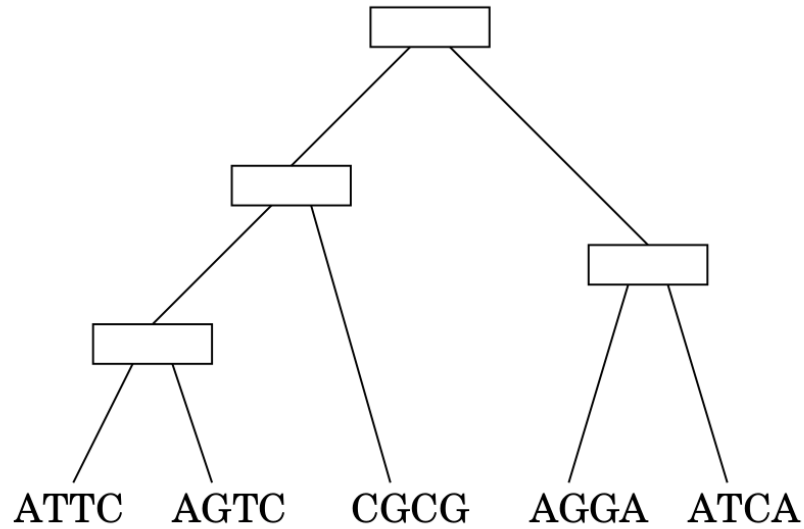
- Uma árvore (binária) evolucionária com as espécies dadas
- Para cada nó interno uma string de comprimento  $k$  com a sequência genética daquele ancestral.

Dada uma árvore acompanhada de uma string  $s(u) \in \{A, C, G, T\}^k$  para cada nó  $u \in V(T)$ , podemos atribuir uma nota usando o método da máxima parcimônia, que diz que menos mutações são mais prováveis:

$$\text{nota}(T) = \sum_{(u,v) \in E(T)} (\text{número de posições em que } s(u) \text{ e } s(v) \text{ diferem}).$$

Achar a árvore com nota mais baixa é um problema difícil. Aqui vamos considerar um problema menor: Dada a estrutura da árvore, achar as sequências genéticas  $s(u)$  para os nós internos que dêem a nota mais baixa.

Um exemplo com  $k = 4$  e  $n = 5$ :



1. Ache uma reconstrução para o exemplo seguindo o método da máxima parcimônia.
2. Dê um algoritmo eficiente para essa tarefa.

## 2 Dados reais

Usamos <http://www.ncbi.nlm.nih.gov/Taxonomy/CommonTree/wwwcmt.cgi> para gerar o banco de dados.

Rosalind MULT, GLOB, EDTA, PERM, EDIT, LCSQ, CSTR, CTBL, NWCK, SSET, MRNA, KMP, PROB SSEQ, SPLC, LCSM