

0410107 - PRINCÍPIOS DE SISTEMÁTICA E BIOGEOGRAFIA

Instituto de Biociências - Universidade de São Paulo

Prática 6

Exercício 1

Abaixo você deverá comparar as três etapas da lógica de inferência filogenética para a Fenética, Cladística clássica (parcimônia) e Cladística probabilística.

Exercício 2

Considere que θ seja uma hipótese qualquer e que $obs.$ seja uma observação. Sendo L a verossimilhança, por que $L_{(\theta|Obs)}$ não atribui nenhum grau de probabilidade de que a hipótese ($P_{(\theta|Obs)}$) esteja correta?

Exercício 3

O gráfico abaixo representa os valores de Verossimilhança (L) dada a probabilidade de obter caras ($P_{(Ca)}$) em dois ensaios com 20 eventos de cara ou coroa. No primeiro ensaio, uma moeda resultou em 10 caras e 10 coroas (linhas pontilhadas). No segundo ensaio, uma outra moeda resultou em 3 caras e 17 coroas (linhas contínuas). Com base nestas informações, explique como Verossimilhança Máxima é usada para estimar parâmetros (Figure 1).

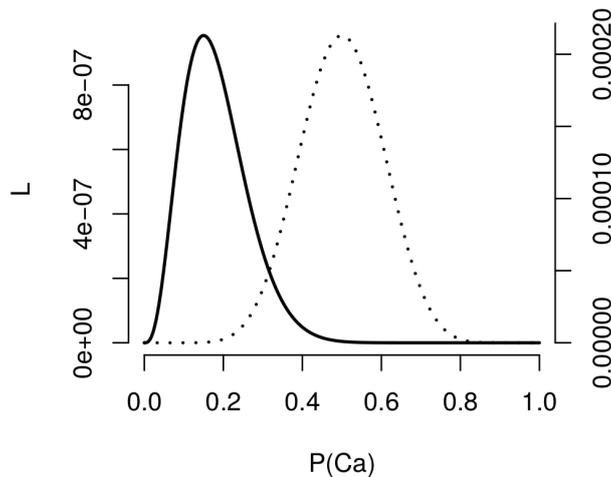


Figura 1: Verossimilhança (L) dada a probabilidade de obter caras ($P_{(Ca)}$).

Exercício 4

O gráfico abaixo representa os valores de Verossimilhança (L) dada a probabilidade de um determinado parâmetro para os modelos 1 e 2 (linhas pontilhadas e contínuas, respectivamente). De acordo com esse gráfico na Figure 2, responda:

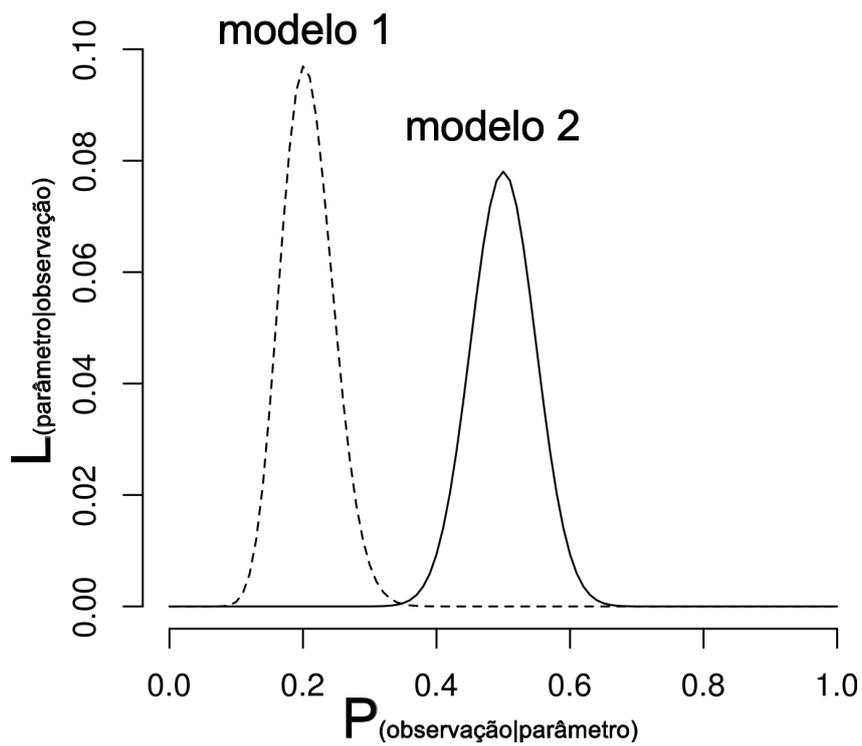


Figura 2: Verossimilhança (L) dada a probabilidades para dois modelos.

i. Como, dentre de cada modelo, você estimaria o melhor valor para esse parâmetro?

ii. Como você selecionaria o modelo utilizando o critério de Verossimilhança?

Exercício 5

Considere o seguinte modelo de substituição de nucleotídeos,

$$JC = \begin{pmatrix} P_{(AA)} & P_{(AC)} & P_{(AG)} & P_{(AT)} \\ P_{(CA)} & P_{(CC)} & P_{(CG)} & P_{(CT)} \\ P_{(GA)} & P_{(GC)} & P_{(GG)} & P_{(GT)} \\ P_{(TA)} & P_{(TC)} & P_{(TG)} & P_{(TT)} \end{pmatrix} \alpha$$

na qual $P_{(AA)}$ é a probabilidade de não-mudança do estado A , $P_{(AC)}$ é a probabilidade da transformação $A \rightarrow C$ ocorrer, $P_{(AG)}$ é a probabilidade da transformação $A \rightarrow G$ ocorrer, ... *ad nauseam!*

Este modelo, uma simplificação do modelo de Jukes–Cantor (1969)¹ – o primeiro modelo de substituição publicado –, refere-se às probabilidades de substituição ($L_{(i,j)}$) multiplicada pela taxa (α) pela qual qualquer substituição irá ocorrer. Neste exercício, esta taxa de substituição será expressa pelo comprimento dos ramos “ v ”. As Tabelas 1 e 2, abaixo, contém as probabilidades de transformação de acordo com os respectivos valores de v .

	A	C	G	T
A	0,9064	0,0312	0,0312	0,0312
C	0,0312	0,9064	0,0312	0,0312
G	0,0312	0,0312	0,9064	0,0312
T	0,0312	0,0312	0,0312	0,9064

Tabela 1: Matriz de probabilidades de transformação para $v = 0.1$.

	A	C	G	T
A	0,9516	0,0161	0,0161	0,0161
C	0,0161	0,9516	0,0161	0,0161
G	0,0161	0,0161	0,9516	0,0161
T	0,0161	0,0161	0,0161	0,9516

Tabela 2: Matriz de probabilidades de transformação para $v = 0.05$.

Com base nestas informações, você deverá calcular a Verossimilhança Máxima (*Maximum Likelihood*, **VM/ML**) de uma das topologias abaixo. Esse exercício deverá ser feito em dupla, de

¹Jukes, T.H. & C.R. Cantor. 1969. Evolution of protein molecules. In: H.N. Munro (ed.) Mammalian protein metabolism. Academic Press, NY. 21-132 pp.

modo que cada um faça o cálculo para umas das topologias abaixo e vocês possam comparar os resultados. Lembre-se de algumas regras para o cálculo de **VM**:

- i. Para cada reconstrução, as probabilidades de substituição $L_{(i,j)}$ no comprimento dos ramos são multiplicadas. Por exemplo, a probabilidade da reconstrução 1 da topologia 1 (P_{R1T1}) é $P_{(AA|v=0.05)} * P_{(AA|v=0.05)} * P_{(AA|v=0.1)} * P_{(AG|v=0.1)} * P_{(AG|v=0.1)}$, ou seja, $0,9516 * 0,9516 * 0,9064 * 0,0312 * 0,0312 = 0.000798984$ (veja Tabelas 1 e 2, e a reconstrução 1 da topologia 1). Observe que existem componentes probabilísticos idênticos – o que minimiza o número de cálculos necessários.
- ii. A **VM** de uma topologia é dada pela soma das **MLs** de cada reconstrução. Por exemplo, a Verossimilhança Máxima da topologia 1 (*i.e.*, L_{T1}) é a soma das probabilidades das 16 possíveis reconstruções, ou seja, $L_{T1} = P_{R1T1} + P_{R2T1} + P_{R3T1} \dots P_{R16T1}$.

Feito os cálculos responda:

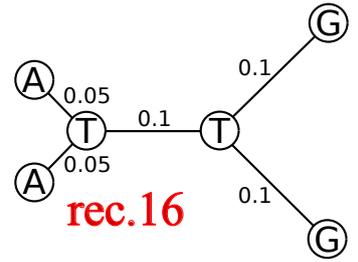
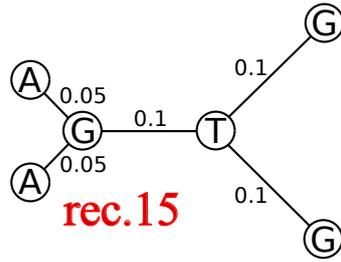
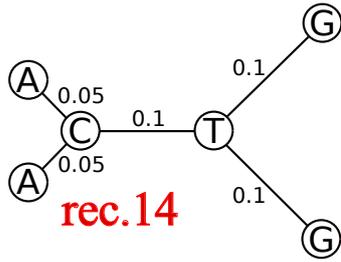
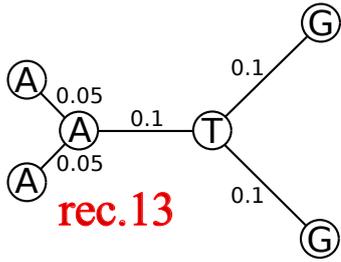
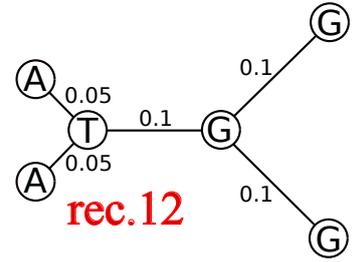
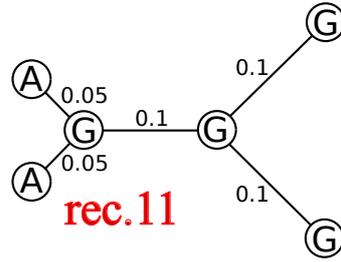
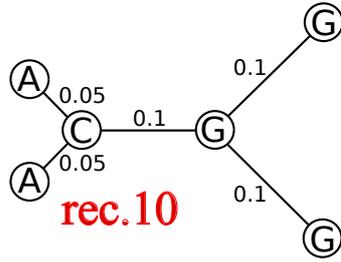
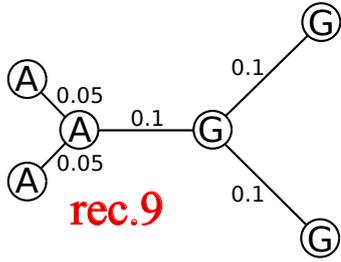
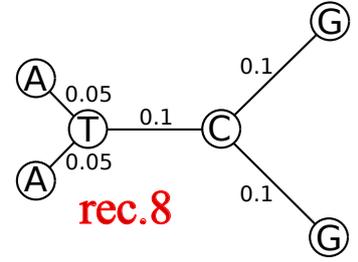
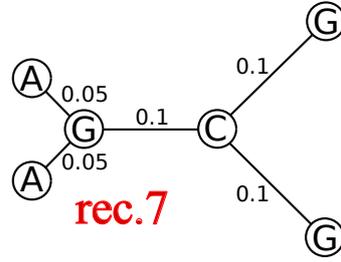
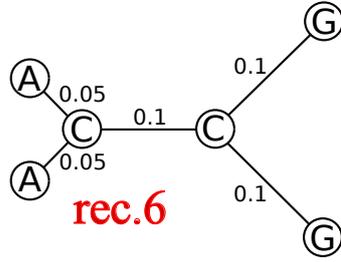
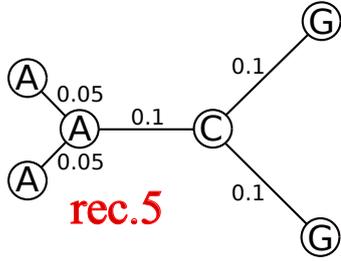
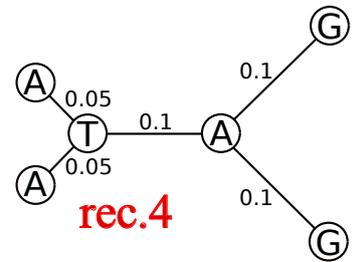
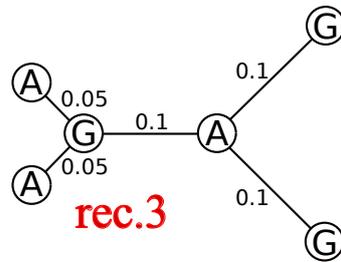
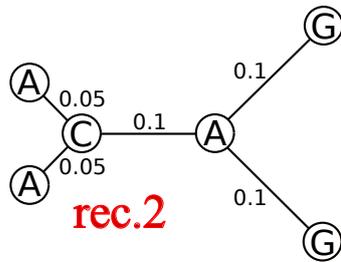
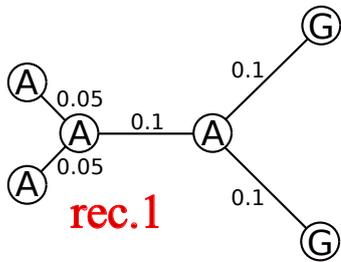
- a. Qual das topologias seria escolhida por este critério, *i.e.*, Verossimilhança Máxima (ou *Maximum Likelihood*)? Justifique.

- b. Se você estivesse utilizado como critério de seleção a Parcimônia, você teria obtido o mesmo resultado? Justifique.

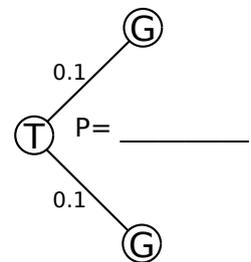
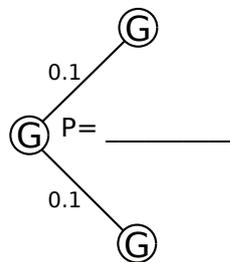
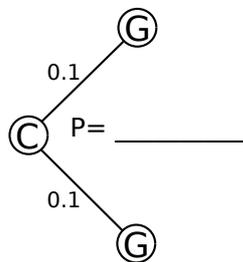
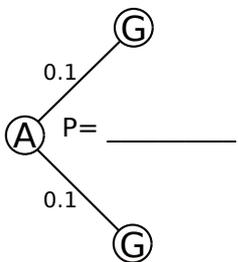
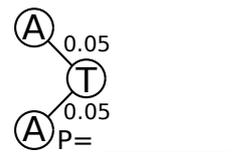
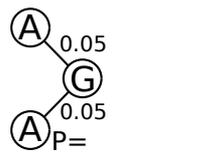
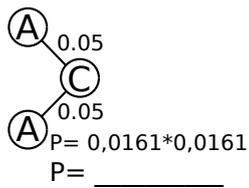
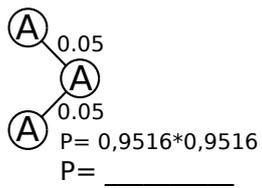
- c. Para cada uma das topologias (*i.e.*, T1 e T2) qual(ais) reconstrução(ões) seria(m) selecionadas pelo critério de Verossimilhança Máxima?

- d. Você selecionaria as mesmas reconstruções acima se o seu critério de seleção fosse a Parcimônia? Justifique.

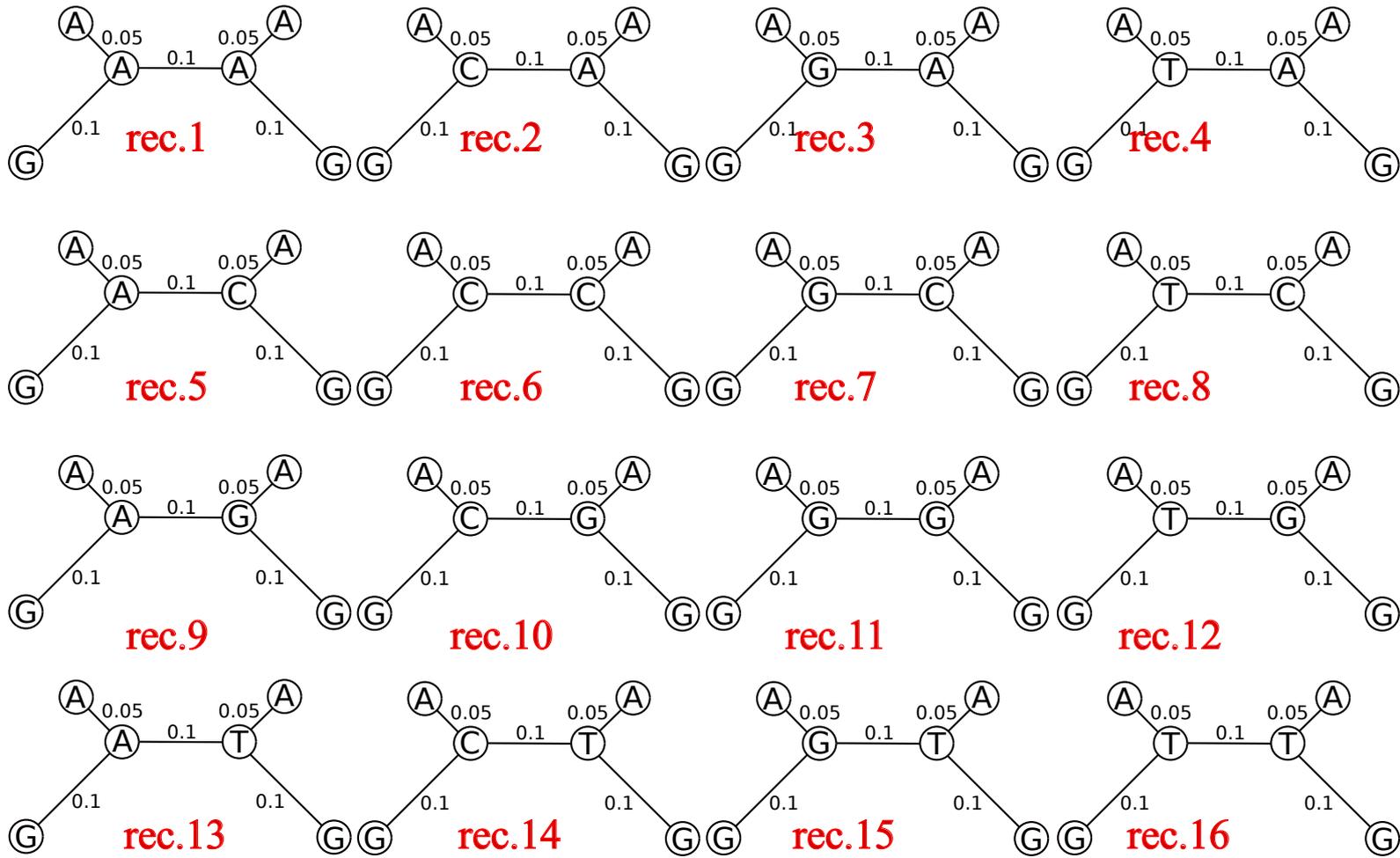
Reconstruções (T1):



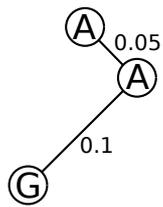
Componentes probabilísticos:



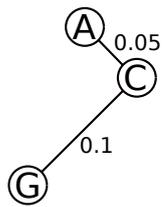
Reconstruções (T2):



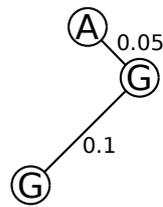
Componentes probabilísticos:



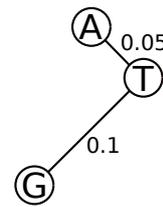
$P = 0,9516 * 0,0312$
 $P = \underline{\hspace{2cm}}$



$P = 0,0161 * 0,0312$
 $P = \underline{\hspace{2cm}}$



$P = \underline{\hspace{2cm}}$



$P = \underline{\hspace{2cm}}$

$\text{A} \xrightarrow{0.1} \text{A}$ $P = \underline{\hspace{2cm}}$

$\text{A} \xrightarrow{0.1} \text{C}$ $P = \underline{\hspace{2cm}}$

$\text{C} \xrightarrow{0.1} \text{C}$ $P = \underline{\hspace{2cm}}$

$\text{A} \xrightarrow{0.1} \text{G}$ $P = \underline{\hspace{2cm}}$

$\text{C} \xrightarrow{0.1} \text{G}$ $P = \underline{\hspace{2cm}}$

$\text{G} \xrightarrow{0.1} \text{G}$ $P = \underline{\hspace{2cm}}$

$\text{A} \xrightarrow{0.1} \text{T}$ $P = \underline{\hspace{2cm}}$

$\text{C} \xrightarrow{0.1} \text{T}$ $P = \underline{\hspace{2cm}}$

$\text{G} \xrightarrow{0.1} \text{T}$ $P = \underline{\hspace{2cm}}$

$\text{T} \xrightarrow{0.1} \text{T}$ $P = \underline{\hspace{2cm}}$