

# Übungen zur Phylogenetik Vorlesung

Universität Bielefeld, WS 2014/2015, Dr. Roland Wittler, Kevin Lamkiewicz

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2014winter/Phylogenetik>

## Blatt 9 vom 3.12.2014

Abgabe in einer Woche zu Beginn der Vorlesung oder vorab bei deinem Tutor oder beim Veranstalter.

### Aufgabe 1 Programmieraufgabe: Splittree.

(4 Punkte)

Programmiere in einer Programmiersprache deiner Wahl eine Funktionen zur Berechnung des 'isolation index' aus dem Kontext *Split Decomposition*: Für zwei gegebene Mengen  $J$  und  $K$  bezüglich einer Matrix  $d$  soll  $\alpha_{J,K}(d)$  berechnet werden:

$$\alpha_{J,K}(d) = \frac{1}{2} \min_{\substack{i,j \in J \\ k,l \in K}} (\max\{d_{ij} + d_{kl}, d_{ik} + d_{jl}, d_{il} + d_{jk}\} - d_{ij} - d_{kl})$$

Schicke eine lauffähige Version deines Programms an deinen Tutor<sup>1</sup> und beschreibe in der E-Mail, wie das Programm aufgerufen werden muss, um  $\alpha$  in der folgenden Aufgabe zu berechnen. Auf der Rückseite befindet sich eine Matrix und einige Lösungen zum Testen deines Programms.

### Aufgabe 2 Berechnung von $d$ -Splits.

(4 Punkte)

Betrachte den Algorithmus zur Berechnung von  $d$ -Splits (Skript, Seite 62 oben).

- Erläutere zunächst: Wenn ein Split  $J, K$  um ein Taxon erweitert wird ( $J', K$ , mit  $J' := J \cup \{i\}$ ), muss  $\alpha_{J',K}$  nicht vollständig neu berechnet werden, um zu entscheiden, ob  $J', K$  ein gültiger Split ist ( $\alpha_{J',K} \neq 0$ ). Eine Neuberechnung würde  $O(n^4)$  kosten. Ohne Neuberechnung geht es auch in  $O(n^3)$ . Wie ist dies möglich?
- Erläutere die Laufzeit von  $O(n^6)$  des Algorithmus.  
Hinweis: Für  $n$  Taxa gibt es maximal  $\binom{n}{2}$  Splits. (Muss nicht gezeigt werden.)

### Aufgabe 3 Anwendung: Splittree.

(3 Punkte)

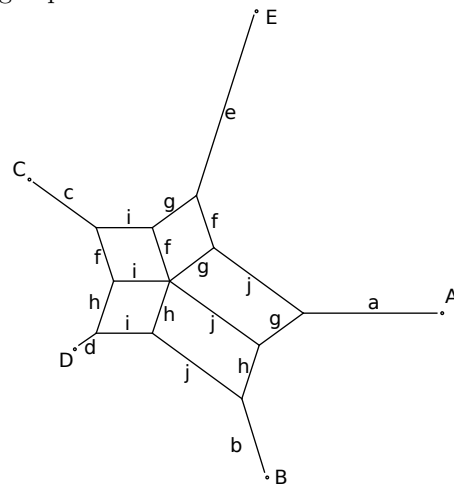
Gegeben seien die folgende Distanzmatrix, sowie der zugehörige Splittree:

	A	B	C	D	E
A:	0	9	13	12	13
B:		0	12	7	15
C:			0	6	10
D:				0	12
E:					0

Berechne in dem gegebenen Splittree-Netz die Kantenzahlen  $a$  bis  $j$ .

(Hinweis:  $a = \alpha_{\{A\}, \{B,C,D,E\}}$ ,  $g = \alpha_{\{A,E\}, \{B,C,D\}}$ , etc.)

Die Berechnungen sollten mit Hilfe deiner Implementierung erfolgen. Alternativ kann auch die kostenlose Software *Splittree* ([www.splitstree.org](http://www.splitstree.org)) verwendet werden— in diesem Fall beschreibe das Eingabeformat und dein Vorgehen.



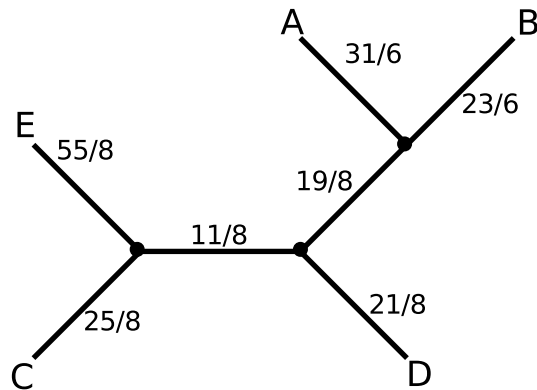
Bitte wenden!

<sup>1</sup>{dgruending.jeschulz}@techfak.uni...

**Aufgabe 4 Methodenvergleich.**

**(2 Punkte)**

Wendet man den *Neighbor Joining* Algorithmus auf die Distanzmatrix aus Aufgabe 3 an, erhält man folgenden Baum. Vergleiche diesen mit dem Ergebnis aus Aufgabe 3.



Matrix und Teillösungen zum (freiwilligen!) Testen deiner Implementierung in Aufgabe 1:

	A	B	C	D	E
A:	0	6	8	5	10
B:		0	5	8	10
C:			0	4	8
D:				0	7
E:					0

J	K	$\alpha_{J,K}$
{A, B, C, E}	{D}	0.5
{A, B, D, E}	{C}	0.5
{A, B, E}	{C, D}	0.5
{A, C, D, E}	{B}	1.5
{A, C, E}	{B, D}	0.0
{A, D, E}	{B, C}	1.5
{A, E}	{B, C, D}	0.0
{B, C, D, E}	{A}	1.5
{B, C, E}	{A, D}	1.0
{B, D, E}	{A, C}	0.0
{B, E}	{A, C, D}	0.0
{C, D, E}	{A, B}	2.0
{C, E}	{A, B, D}	0.0
{D, E}	{A, B, C}	0.0
{E}	{A, B, C, D}	4.5