

# 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 7 vom 19.11.2014

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

### Aufgabe 1 Additive Metrik, Ultrametrik.

(4 Punkte)

Es seien folgende Matrizen gegeben:

i)	A	B	C	D	E
A	0	10	10	10	4
B		0	2	6	10
C			0	6	10
D				0	10
E					0

ii)	A	B	C	D	E
A	0	3	7	9	10
B		0	8	10	11
C			0	4	5
D				0	3
E					0

Entscheide je für i) und ii), ob die Matrix eine additive oder sogar eine ultrametrische Metrik beschreibt. Begründe deine Entscheidung.

### Aufgabe 2 Eigenschaften von Distanzen.

(4 Punkte)

Bei jeder Distanzfunktion  $d$  gilt der folgende Zusammenhang:

$$„d \text{ ist ultrametrisch}“ \Rightarrow „d \text{ ist additiv}“ \Rightarrow „d \text{ erfüllt die Dreiecksungleichung}“$$

Beweise diesen Zusammenhang, indem du die folgenden Punkte zeigst:

- (a) „ $d$  genügt der *Drei-Punkt-Bedingung*“  $\Rightarrow$  „ $d$  genügt der *Vier-Punkt-Bedingung*“

Beweisskizze: Für beliebige vier Elemente, gehen wir wie folgt vor. Wir benennen zwei der Elemente, die minimalen Abstand haben mit  $a$  und  $b$ . Wir benennen die beiden anderen Elemente  $c$  und  $d$ , so dass gilt  $d_{ac} \leq d_{ad}$ . Dann können  $a$ ,  $b$ ,  $c$  und  $d$  in genau zwei verschiedenen binären, ultrametrischen Bäumen angeordnet werden. **Welche?** In jedem der zwei Fälle gelten nun mehrere Drei-Punkt-Bedingungen (**Welche?**), die zur gesuchten Vier-Punkt-Bedingung kombiniert/umgeformt werden können. **Wie?** (Der Fall, dass die Distanzen einer nicht-binären Topologie entsprechen, muss **nicht** gesondert betrachtet werden. Diesen erfassen wir, indem wir Kanten der Länge Null erlauben. Zum Beispiel stellen wir  $((a:2,b:2,c:2):1,d:3)$ ; binär als  $((a:2,b:2):0,c:2):1,d:3)$ ; dar.)

- (b) „ $d$  genügt der *Vier-Punkt-Bedingung*“  $\Rightarrow$  „ $d$  erfüllt die *Dreiecksungleichung*“

Tipp: Die Vier-Punkt-Bedingung kann auch auf drei Punkte angewendet werden.

### Aufgabe 3 Clustering-Verfahren zur Baumrekonstruktion.

(4 Punkte)

Gegeben sei folgende Distanzmatrix:

	A	B	C	D	E	F
A:	0	4	8	18	18	6
B:		0	6	12	8	8
C:			0	18	18	12
D:				0	10	8
E:					0	12
F:						0

Verwende die folgenden Verfahren, um phylogenetische Bäume aus der Matrix zu rekonstruieren.

- (a) *Complete linkage clustering*.  
(b) *UPGMA*.

Gib in den Zwischenschritten die jeweils dazugehörige Matrix an und am Ende zusätzlich den finalen Baum. Existieren mehrere Möglichkeiten, gib alle an.

- (c) Vergleiche die Resultate und stelle fest, ob die Matrix ultrametrisch ist. Begründe!