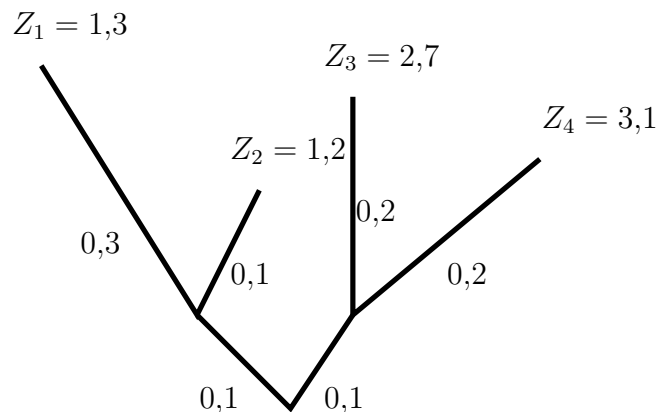


24. Januar 2007

Abgabe: 31. Januar 2007

Aufgabe 1: Ein quantitatives Merkmal werde bei vier Taxa beobachtet und habe dort die Werte Z_1, Z_2, Z_3, Z_4 . Eine Hypothese besage, dass sich die vier Taxa nach folgendem Baum entwickelt haben, und dass das Merkmal gemäß dem Brownsche-Bewegung-Modell entlang der Kanten evoluiert sei, d.h. für jede Kante der Länge t sei die Differenz der Werte an ihren beiden Knoten normalverteilt mit Mittelwert 0 und Varianz t .



Untersuchen Sie, ob die Daten mit der Hypothese gut verträglich sind oder statistisch signifikant gegen die Hypothese sprechen.

Aufgabe 2: Schreiben Sie ein möglichst effizientes Programm, das für die in der Vorlesung behandelte SCFG mit Regeln der Form

$$S \rightarrow L \mid LS$$

$$F \rightarrow LS \mid xFy$$

$$L \rightarrow x \mid xFy$$

nach Eingabe von Übergangswahrscheinlichkeiten und einer RNA-Sequenz den wahrscheinlichsten Ableitungsbaum ausgibt.

Aufgabe 3: Ergänzen Sie Ihr Programm aus Aufgabe 2 so, dass es für die eingegebene Sequenz auch die Wahrscheinlichkeit berechnet, durch die Grammatik erzeugt zu werden, und RNA-Sekundärstrukturen gemäß ihrer a-posteriori-Wahrscheinlichkeit erzeugt.