

Übungen zur Algorithmischen Bioinformatik I

Blatt 1

Abgabetermin: Montag, 8.5.2017, 10 s.t.

1. Aufgabe (Bonus-Aufgabe):

Beweisen Sie die folgende Gleichung für $n \in \mathbb{N}$ mit $n \geq k$:

$$\sum_{i=k}^n \binom{i}{k} = \binom{n+1}{k+1}.$$

(a) durch vollständige Induktion;

(b) durch Verwendung der folgenden Gleichung:

$$\sum_{i=k}^n \binom{i+1}{k+1} = \sum_{i=k}^n \left(\binom{i}{k+1} + \binom{i}{k} \right)$$

und Indexverschiebung sowie $\binom{k}{k+1} = 0$.

2. Aufgabe (Bonus-Aufgabe):

Geben Sie einen möglichst effizienten Algorithmus an, der **alle** *maximal scoring subsequences* findet. Analysieren Sie den Algorithmus hinsichtlich seiner Laufzeit.

Hinweis: Mit **alle** sind all diejenigen Teilfolgen gemeint, die einen maximalen Score besitzen. Dabei sollen nur die Teilfolgen ausgegeben werden, die keine anderen Teilfolge mit maximalem Score enthalten.

3. Aufgabe:

Sei SUPERCOMPUTER ein leistungsfähiger Rechner, der in einer Sekunde 1.000 Elementaroperationen ausführen kann. Für ein bestimmtes Problem seien fünf verschiedene Algorithmen verfügbar. Hierbei benötigt der i -te Algorithmus bei einer Eingabe der Eingabegröße n genau $T_i(n)$ Elementaroperationen, wobei

$$T_1(n) = 500 \cdot n, \quad T_2(n) = 50 \cdot n \log(n), \quad T_3(n) = n^2, \quad T_4(n) = \frac{n^3}{100}, \quad T_5(n) = \frac{3^n}{1000}.$$

ist. Vervollständigen Sie die folgende Tabelle, in der die Eingabegrößen angegeben sind, für die der i -te Algorithmus auf dem SUPERCOMPUTER (ziemlich) genau eine Sekunde, eine Minute, eine Stunde, einen Tag bzw. einen Monat Rechenzeit benötigt.

	1s	1m = 60s	1h = 3.600s	1d = 86.400s	1M=2.592.000s
T_1	2				
T_2			≈ 5763		
T_3					
T_4					
T_5		≈ 16			

4. Aufgabe:

Sei HYPERCOMPUTER eine Weiterentwicklung von SUPERCOMPUTER aus Aufgabe 3, der um den Faktor 100 schneller ist, also 100.000 Elementaroperationen pro Sekunde ausführen kann.

Um wieviel kann man die Eingabegröße für die fünf verschiedenen Algorithmen gegenüber SUPERCOMPUTER erhöhen, wenn man dieselbe Rechenzeit zur Verfügung hat? Diese Veränderung ist (möglichst genau) als Funktion der Eingabegröße (beispielsweise als Faktor oder ähnliches) anzugeben, wobei das für T_2 nicht genau möglich ist.

Hinweis: Es ist **nicht** noch einmal eine Tabelle wie bei Aufgabe 3 zu berechnen.