B. Morgenstern / M. Stanke

Übungen zur Vorlesung
"Algorithmen der Bioinformatik II"
Wintersemester 2005/2006

Blatt 4

1. Es sei eine Folge (Vektor) $a = (a_1, a_2, \ldots, a_n)$ von Zahlen gegeben. Die Indizes $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ definieren eine aufsteigende Teilfolge $(a_{i_1}, a_{i_2}, \ldots, a_{i_k})$ der Länge k von $(a_1, a_2, \ldots, a_n)$, wenn $a_{i_1} < a_{i_2} < \cdots < a_{i_k}$ ist. Bei einer aufsteigenden Teilfolge werden also die Zahlen von links nach rechts gelesen immer größer. Beispiel:

$$a = (6, \underline{8}, 0, 4, 15, 3, 7, \underline{9}, \underline{20}, 12, 43, 6, 4, \underline{23})$$

   Eine aufsteigende Teilfolge von a ist z.B. $(8, 9, 20, 23)$ und sie hat Länge 4. Entwerfe einen Algorithmus, der die *längste aufsteigende Teilfolge* bestimmt. Hinweis: Definiere $A(i) :=$ "die Länge der längsten aufsteigenden Teilfolge, die an der Stelle i endet" für $(1 \leq i \leq n)$ und zeige, wie sich $A(i)$ aus $A(1), A(2), \ldots, A(i-1)$ und a berechnen lässt. Was ist die Laufzeit und der Speicherbedarf des Algorithmus?

2. Finde einen effizienten Algorithmus, der das Problem löst, einen Absatz hübsch zu formatieren. Gegeben sei ein Text von n Wörtern, der Längen $\ell_1, \ell_2, \ldots, \ell_n$ und eine Spaltenbreite M, die nicht überschritten werden darf. Wenn Wörter i bis j zusammen in eine Zeile geschrieben werden und wir genau ein Leerzeichen zwischen jeweils zwei Wörtern lassen, ist die Anzahl der Leerzeichen am Ende der Zeile gleich $s = M - j + i - \sum_{k=i}^{j} \ell_k$. Wir möchten gerne die Summe der Quadrate der Anzahl der Leerzeichen an den Enden der Zeilen des Absatzes minimieren. Den Wert dieser Zielfunktion bezeichnen wir mit S. Beispiel: Der Absatz

```
|Finde_einen_effizienten_|
|Algorithmus,_der_das____|
|Problem_löst,_einen_____|
|Absatz_hübsch_zu________|
|formatieren.____________|
```

   hätte $S = 1^2 + 4^2 + 5^2 + 8^2 + 12^2 = 250$. Was ist die Laufzeit und der Speicherbedarf des Algorithmus? Was ändert sich am Algorithmus, wenn die Leerzeichen am Ende der *letzten* Zeile nicht bestraft werden?

3. Finde eine Rekursionsformel für die Anzahl $C(n)$ der Möglichkeiten, n Faktoren in einer gegebenen Reihenfolge zu klammern. Z.B. 5 Faktoren a,b,c,d,e können auf 14 verschiedene Arten geklammert werden

```
a(b(c(de)))
a(b((cd)e))
a((bc)(de))
a(((bc)d)e)
a((b(cd))e)
(ab)((cd)e)
(ab)(c(de))
((ab)c)(de)
(a(bc))(de)
(a(b(cd)))e
(a((bc)d))e
((ab)(cd))e
(((ab)c)d)e
((a(bc))d)e
```

Also $C(5) = 14$. Wie kann $C(n)$ aus $C(1), C(2), \ldots, C(n-1)$ bestimmt werden? Hinweis: Welche Möglichkeiten gibt es für den "letzten Schritt", das Produkt in zwei Faktoren zu zerlegen?