4. Übung zur Vorlesung "Algorithmische Massenspektrometrie"

Wintersemester 2018/2019

Sebastian Böcker, Martin Hoffmann

Ausgabe: 19. November 2018, Abgabe: 26. November 2018

1. **Minimum-Coin-Changing-Problem:** Gegeben sei ein beliebiges Alphabet von Münzwerten $\Sigma = \{a_1 < a_2 < \ldots < a_k\}$ sowie ein Geldbetrag $\mathcal{M}$. Gesucht sei die minimale Anzahl an Münzen, die summiert den Geldwert M ergeben.

   (a) Angenommen, wir haben den folgenden Algorithmus zur Lösung dieses Problems gefunden:

   ```
   MINIMUM_COIN_CHANGING( Σ = {a_1 < a_2 < ... < a_k}, M )
     (c_1, ..., c_k) = (0, ..., 0)
     r ← M
     FOR i ← k DOWNTO 1
       c_i ← r / a_i
       r ← r - c_i · a_i
     END FOR;
     IF (∑_{i=1}^k c_i · a_i = M) THEN
       RETURN ∑_{i=1}^k c_i
     END IF;
   END.
   ```

   - Beweisen Sie, dass der obige Algorithmus nicht für jede Instanz des Minimum-Coin-Changing-Problems eine richtige Lösung liefert.
   - Wie nennt man den Ansatz, der im obigen Algorithmus verwendet wird?
   - Gibt es Alphabete, für die der obige Algorithmus immer exakte Ergebnisse liefert? Nennen Sie ein Beispiel.

   (4 Punkte)

   (b) Geben Sie einen Algorithmus an (es genügen Rekurrenz und Initialisierung einer DP), der das Problem exakt löst. Der Speicherverbrauch ihres Algorithmus sollte in $O(M)$ liegen. Hinweis: Gesucht ist lediglich die Anzahl an Münzen, nicht die genaue Zusammensetzung des Wechselgeldes.

   (4 Punkte)

2. **Compomere aufzählen:** Gegeben sei das Alphabet $\Sigma = \{a, b, c\}$ mit den Gewichten $\mu(a) = 3$, $\mu(b) = 4$ und $\mu(c) = 6$. Berechnen Sie die Anzahl aller Compomere über diesem Alphabet, welche die Masse $\mathcal{M} = 15$ zerlegen. Nutzen Sie dafür dynamische Programmierung. Geben Sie die DP Tabelle an und zählen Sie (per Backtracking) alle Compomere ihrer Lösung auf.

   (4 Punkte)

3. Angenommen wir wüssten, dass in einem MS/MS Spektrum nur b-Ionen (Präfixmassen) **und Noise-Peaks** vorkommen. Die Interpretation des Spektrums wäre dann viel einfacher.

   (a) Schreiben Sie einen Algorithmus, der ein solches Spektrum interpretiert. Als Eingabe erhält der Algorithmus eine Liste von Präfixmassen und Noise Peaks $m_1 < m_2 < \ldots < m_n$ Als Ausgabe soll der Algorithmus die Peptid Sequenz zurückgeben, welche die meisten Peaks erklärt. Die asymptotische Laufzeit ihres Algorithmus sollte in $O(n\Sigma)$ sein, wobei Σ die Größe des Alphabets ist.

(5 Punkte)

4. Gegeben sei die Rekurrenz

$$Q[i,j] := \begin{cases} \max_{l=0,\ldots,i-1}\{Q[l,j] + w(x_l, x_i)\} & \mid i > j \\ \max_{l=0,\ldots,j-1}\{Q[i,l] + w(y_j, y_l)\} & \mid j > i \end{cases} \tag{1}$$

mit der Initialisierung $Q[0,0] = 0$ und $Q[i,i] = -\infty$ für alle $i = 1, \ldots, n$ (mit $2n$ gleich der Anzahl an gemessenen Peaks), sowie der Gewichtsfunktion

$$w(x,y) := \begin{cases} 1 & \mid xy \in \mathbb{E} \\ 0 & \mid \text{ansonsten} \end{cases} \tag{2}$$

Beweisen Sie, für ein Spektrum mit zusätzlichen Peaks (Noise) aber ohne fehlenden Peaks, dass die maximale Zahl an Peaks, die als Präfixe und Suffixe eines Strings interpretiert werden können, gleich $2\max_{i,j}\{Q[i,j] : x_i y_j \in E\}$ ist.
Verwenden Sie dafür die Definition von Q aus der Vorlesung.

(3 Punkte)