

3. Übung zur Vorlesung “Algorithmische Massenspektrometrie”

Wintersemester 2020/2021

Sebastian Böcker, Kai Dührkop

Ausgabe: 19. November 2020, Abgabe: 24. November 2020

1. 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 sortiertes Spektrum interpretiert. Als Eingabe erhält der Algorithmus eine sortierte Liste von Präfixmassen und Noise Peaks $m_1 < m_2 < \dots < m_n$. 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.

(3 Punkte)

2. Gegeben sei die Rekurrenz

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

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

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

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)

3. Nutzen Sie die Gewichtsfunktion (2) sowie das Alphabet $\Sigma = \{a, b, c, d\}$ mit Massen $\mu(a) = 2$, $\mu(b) = 3$, $\mu(c) = 7$ und $\mu(d) = 10$ um das folgende Massenspektrum mit Noise auszuwerten:

$$\mathcal{M} = \{0, 1, 2, 9, 11, 12, 14, 21, 22, 23\}$$

Verwenden Sie dabei eine Rekurrenz, die mit identischen Präfix- und Suffixmassen umgehen kann. Geben Sie einen Peptid-String an, der maximal viele Peaks erklärt, sowie die DP-Matrix, die Sie zu dessen Berechnung genutzt haben.

(6 Punkte)