Technische Universität München
Fakultät für Informatik
Lehrstuhl für Effiziente Algorithmen
Dr. Hanjo Täubig
Johannes Nowak

Sommersemester 2008
Übungsblatt 5
09. Juni 2008

Algorithmische Bioinformatik I

Aufgabe 1

Sei Σ ein Alphabet mit $- \notin \Sigma$, $\overline{\Sigma} := \Sigma \cup \{-\}$ und $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. Sei $w : \overline{\Sigma} \times \overline{\Sigma} \to \mathbb{R}_+$ eine Kostenfunktion. Die Abstandsfunktion für ein Alignment $(\bar{x}, \bar{y})$ für $x, y \in \Sigma^*$ ist definiert durch:

$$\bar{d}_w(\bar{x}, \bar{y}) := \sum_{i=1}^{|\bar{x}|} w(\bar{x}_i, \bar{y}_i).$$

Ferner ist der Alignment-Abstand von $x, y \in \Sigma^*$ definiert durch:

$$d_w(x, y) := \min\{\bar{d}_w(\bar{x}, \bar{y}) \mid (\bar{x}, \bar{y}) \text{ ist ein Alignment für } x, y\}.$$

Beweise: Ist die Funktion w eine Metrik, so ist auch $\bar{d}_w$ und d_w eine Metrik.

Aufgabe 2

Sei Σ ein Alphabet mit $- \notin \Sigma$, $\overline{\Sigma} := \Sigma \cup \{-\}$ und $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$. Eine Funktion $w : \overline{\Sigma} \times \overline{\Sigma} \to \mathbb{R}_+$ heißt *Kostenfunktion für ein Abstandsmaß*, wenn folgende Bedingungen gelten:

- w ist symmetrisch, d.h. $\forall x, y \in \overline{\Sigma} : w(x, y) = w(y, x)$.
- w erfüllt die Dreiecksungleichung, d.h. $\forall x, y, z \in \overline{\Sigma} : w(x, y) \leq w(x, z) + w(z, y)$.
- Für $A := \max\{w(a, a) : a \in \Sigma\}$ und $B := \min\{w(a, b) : a \neq b \in \Sigma^- \vee a = b = -\}$ gilt $A \leq B$.

Die Funktion $w' : \overline{\Sigma} \times \overline{\Sigma} \to \mathbb{R}$ heißt *Kostenfunktion für ein Ähnlichkeitsmaß*, wenn folgende Bedingungen gelten:

- w' ist symmetrisch, d.h. $\forall x, y \in \overline{\Sigma} : w'(x, y) = w'(y, x)$.
- $w'(x, x) \geq 0$ für $x \in \Sigma$.
- $w'(x, y) \leq 0$ für $x \neq y \in \overline{\Sigma}$.
- $w'(-, -) \leq 0$.

a) Zeige, dass man zu jeder Kostenfunktion w für ein Abstandsmaß eine Kostenfunktion w' für ein Ähnlichkeitsmaß definieren kann, so dass $w'(x, y) = C - w(x, y)$ für ein geeignetes $C \in \mathbb{R}_+$.

b) Zeige, dass man zu jeder Kostenfunktion w' für ein Ähnlichkeitsmaß eine Kostenfunktion w für ein Abstandsmaß definieren kann, so dass $w(x, y) = D - w'(x, y)$ für ein geeignetes $D \in \mathbb{R}_+$.

c) Lassen sich die obigen Äquivalenzen auch ohne die Bedingung der Symmetrie bei der Definition einer Kostenfunktion eines Distanzmaßes bzw. Abstandsmaßes zeigen?