

4. Übungsblatt zur Vorlesung Kombinatorische Algorithmen, SS 13

Abgabe: Bis Montag, 20.05.2013, 12:00 Uhr,
in den Abgabekasten vor dem AG-Flur oder per Email an `wild@cs.uni...`

Wann immer eine Aufgabe lautet „Entwerfen Sie einen Algorithmus ...“, sollte Ihre Abgabe neben einer präzise Beschreibung des Algorithmus auch eine Begründung der Korrektheit, sowie eine Analyse der Komplexität des selbigen umfassen.

5. Aufgabe

4 + 2 Punkte

In dieser Aufgabe wollen wir eine (bzgl. Laufzeit) effiziente Implementierung der LZ77-Zerlegung entwickeln, wie wir sie auf dem letzten Blatt definiert haben.

Für die Praxis wird man eingeschränkte Varianten bevorzugen, die mit konstant viel Speicher auskommen, dabei aber eine etwas schlechtere Kompressionsrate erzielen.

- a) Als ersten Schritt betrachten wir das *Longest Prefix Matching Problem*:
Seien $T \in \Sigma^n$ und $P \in \Sigma^m$, sowie ein Index $t \in [1..n]$ gegeben. Berechne ℓ_{\max} , die Länge des längsten Präfixes von P , der als Teilwort in T „vor Position t “ auftaucht, d. h.

$$\ell_{\max} := \max \left\{ \ell \in [0.. \min\{n, m\}] \mid \exists i \in [1..n - \ell] : i \leq t \wedge T_{i, i+\ell-1} = P_{1, \ell} \right\}.$$

Die Ausgabe besteht aus ℓ_{\max} , sowie einem beliebigen Startindex $i \leq t$, an dem $P_{1, \ell}$ vorkommt.

Entwerfen Sie einen Algorithmus zur Lösung des *Longest Prefix Matching Problem* mit Laufzeit in $\mathcal{O}(t + \ell_{\max})$ und Speicherbedarf in $\mathcal{O}(\ell_{\max})$ (zusätzlich zur Eingabe). Weniger effiziente Lösungen erbringen Teilpunkte.

Tipp: Algorithmen, die wir am Anfang der Vorlesung besprochen haben, mögen ein guter Ausgangspunkt sein.

- b) Entwerfen Sie einen Algorithmus zur Berechnung von $LZ77(w)$ für $w \in \Sigma^n$ mit Laufzeit in $\mathcal{O}(c \cdot n)$, wobei $c := |LZ77(w)|$ die Anzahl Phrasen in der LZ77-Zerlegung von w bezeichnet.

Sie dürfen dabei einen Algorithmus für das *Longest Prefix Matching Problem* als Unterprogramm verwenden, selbst wenn sie keine Lösung für Aufgabe a) gefunden haben sollten.