

Übungen zur Vorlesung Algorithmische Fragestellungen für komprimierte Daten

---

1. Sei  $x = aabbaaabbbbaa$ . Berechnen Sie  $LZ_{77}(x)$  mittels Suffixbäumen wie in der Vorlesung vorgestellt. Geben Sie die Folge der entstehenden Suffixbäume für die Präfixe von  $x$  an.
2. Wie lautet  $LZ_{77}(a^n)$  in Abhängigkeit von  $n$ ? Folgern Sie, dass  $LZ_{77}$  exponentiell bessere Kompressionsraten als  $LZ_{78}$  erzielen kann.
3. Sei  $x \in A^*$ . Die RLZ-Faktorisierung (RLZ für restricted Lempel-Ziv) von  $x$  ist definiert als  $x = x_1x_2 \cdots x_n$ , wobei für alle  $1 \leq i \leq n$  gilt:  
(i)  $x_i$  ist der längste Präfix von  $x_i \cdots x_n$ , der sich als  $x_px_{p+1} \cdots x_q$  mit  $1 \leq p \leq q < i$  schreiben lässt, oder (ii) (falls solche  $p, q$  nicht existieren)  $x_i$  ist das erste Symbol von  $x_i \cdots x_n$ . Im Fall (i) sei  $c_i = 0pq$ , wobei hier  $p$  und  $q$  mit genau  $\lceil \log(i-1) \rceil$  vielen Bits binär kodiert sind; in Fall (ii) sei  $c_i = 1x_i$ . Es sei  $RLZ(x) = c_1c_2 \cdots c_n$ . Warum kann  $x$  aus  $RLZ(x)$  eindeutig rekonstruiert werden?
4. Berechnen Sie  $RLZ(aabbaaabbbbaabbaaabbbbaabba)$ .