

Algorithmen auf Sequenzen

Abgabetermin: Freitag, den 8. Januar, 09⁰⁰ in Moodle

Aufgabe 1

Wende den Conquer-Step aus dem Algorithmus von Main und Lorentz auf das folgende Wort

$$t = t_1 \cdots t_{20} = abaabbabbabbabbabaa$$

für $h = \lfloor \frac{n}{2} \rfloor = 10$ und $q = h + \ell$ an (also nur Schritt 3). Gib dazu für jedes $\ell \in [3 : 6]$ die ausgeführten LCE-Anfragen und die ausgegebenen Tandem-Repeat-Paare an (gib zusätzlich an, welche davon rechtsverzweigend sind).

Aufgabe 2

Gegeben sei eine Zeichenreihe $t \in \Sigma^*$. Ein Wort $w \in \Sigma^*$ heißt *minimal rechts bzw. links-eindeutiges Teilwort* von t , wenn w genau einmal in t auftritt und wenn jedes Präfix bzw. Suffix von w mindestens zweimal in t auftritt.

- Entwirf einen effizienten Algorithmus zum Auffinden aller minimal rechts-eindeutigen Teilwörter der Länge mindestens ℓ , beweise seine Korrektheit und analysiere seine Laufzeit.
- Entwirf einen effizienten Algorithmus zum Auffinden aller minimal links-eindeutigen Teilwörter der Länge mindestens ℓ , beweise seine Korrektheit und analysiere seine Laufzeit.

Aufgabe 3

Konstruiere die Lempel-Ziv-Zerlegung von $t = aabbaabbaabbaabbaab$ und finde gemäß dem Algorithmus aus der Vorlesung eine linkeste Überdeckung aller Tandem-Repeats der Länge 8, deren Zentrum sich im LZ-Block 5 oder 6 befindet.

*Fröhliche Weihnachten und
einen guten Rutsch ins neue Jahr!*