

Foundations of Sequence Analysis
Winter 2005/2006

Exercises

Übung 8, Besprechung am 16.01.2006 (14 – 16h in E0-160)
bzw. 12.01.2006 (14 – 16h in C01-148 und E01-108).

1. Suffixbaum – Beispiele.

- (a) Zeichnen Sie für den String $S_1 = \text{AGGATATAGATAG}$ den Suffixbaum.
- (b) Geben Sie einen String der Länge n an, der die Anzahl innerer Knoten eines Suffixbaumes **minimiert**, und geben Sie die Anzahl der inneren Knoten dieses Suffixbaumes an.
- (c) Geben Sie einen String der Länge n an, der die Anzahl innerer Knoten eines Suffixbaumes **maximiert**, und geben Sie die Anzahl der inneren Knoten dieses Suffixbaumes an.

2. **Suffixarray – Suffixbaum.** Gegeben sei ein String S der Länge $n = |S|$. Ein Suffixarray für S ist ein Feld der Länge n , das alle Zahlen im Bereich 0 bis $n - 1$ enthält. Diese Zahlen entsprechen Startpositionen aller Suffixe in S und bestimmen die lexikographische Ordnung der entsprechenden Suffixe.

Beispiel:

Sei $S\$ = \text{abab\$}$ der gegebene String, dann repräsentiert

- 0 – abab\$
- 2 – ab\$
- 1 – bab\$
- 3 – b\$
- 4 – \$

das Suffixarray für $S\$$, wobei \$ lexikographisch größer ist als a und b .

- (a) Überlegen Sie sich einen Algorithmus zur Konstruktion eines Suffixarrays für einen gegebenen String.
- (b) Gegeben sei ein String und sein Suffixbaum. Überlegen Sie sich einen Algorithmus, der aus dem Suffixbaum das Suffixarray für den String erzeugt.
- (c) Gegeben sei ein String und sein Suffixarray. Überlegen Sie sich einen Algorithmus, der aus dem Suffixarray den Suffixbaum für den String erzeugt.

Hinweis:

Für jeden Knoten des Suffixbaumes sollte die Kindliste bezüglich des ersten Buchstabens der Kanten, die auf die Kinder zeigen, geordnet sein.