

## 4. Übung „Algorithmen der Bioinformatik I“

Abgabe am 21. bzw. 22. Mai 2003 in der Übung



Institut für Informatik  
Martin-Luther-Universität Halle-Wittenberg

---

**Aufgabe 1.** Gegeben seien ein String  $S$ , in dem das Abschlußzeichen  $\$$  nicht vorkommt.

1. Warum wird der String  $S$  mit einem Abschlußzeichen versehen, wenn der zugehörige Suffixbaum aufgebaut wird?
2. Zeigen Sie, dass für den Suffixbaum  $ST(S\$)$  gilt: Anzahl der inneren Knoten  $\leq |S\$|$ .
3. Zeigen Sie, dass für den Suffixbaum  $ST(S\$)$  gilt: Anzahl der Kanten  $\leq 2|S\$| - 1$ .

**Aufgabe 2.** Zeigen Sie, dass das Finden eines Musters  $P$  der Länge  $n$  in einem Text, unter Verwendung eines Suffixbaums des Textes, in  $O(n)$  Zeit geschieht und dass das  $k$ -malige Finden des Musters in  $O(n + k)$  möglich ist.

**Aufgabe 3.** Diese Aufgabe ist spätestens am 30. Mai 2003 abzugeben.

1. Implementieren Sie den Write Only Top Down Algorithmus zur Konstruktion eines Suffixbaums zu einem gegebenen Text  $T$ . Die Baumstruktur und der Inhalt des Baums sollen mit einer Funktion ausgegeben werden können. Diese Funktion ist nach jedem Iterationsschritt des Write Only Top Down Algorithmus aufzurufen.
2. Implementieren Sie eine Funktion, die ein gegebenes Muster  $P$  in einem gegebenem Suffixbaum  $ST(T\$)$  eines Textes  $T$  sucht. Geben Sie die Anfangsindizes aller Vorkommen von  $P$  in  $T$  aus.