Technische Universität München
Department of Informatics
Chair for Efficient Algorithms
Prof. Dr. Ernst W. Mayr/Dr. Jens Ernst
Johannes Nowak

Summer Term 2005
Exercise Sheet 1
3. Mai 2005

Selected Topics in Computational Biology

Due: 03.05.2005 **after** *the lecture*

Exercise 1 (10 points)

a) Construct the suffix trie for the string $CTGCCTGA$, including all suffix links.

b) What does the suffix tree for this string look like?

Exercise 2 (10 points)

As in Algorithm 1 seen in the lecture, let $\{s_1 = \overline{t_1 \cdots t_{i-1}}, s_2 = \overline{t_2 \cdots t_{i-1}}, \ldots, s_i = root, s_{i+1} = \perp\}$ be the boundary path. Let j be the smallest index such that s_j is not a leaf. Let j' be the smallest index such that $s_{j'}$ has a t_i-transition. Prove that $2 \leq j \leq j' \leq i + 1$ and explain the relevance of the quantities j and j' for efficiently constructing the suffix trie.

Exercise 3 (10 points)

Consider a suffix trie. Let $f()$ be the suffix-function. Prove that, if $\overline{x}$ is not a leaf then $f(\overline{x})$ is not a leaf, either.

Programming Task

Familiarize yourself with pointers and structures in the C language.
(See, for example,
Brian W. Kernighan, Dennis M. Ritchie. **The C Programming Language.** *2nd Edition, Prentice Hall PTR, Englewood Cliffs, NJ, 1988.*)
Try to implement a dynamic queue with basic functionalities.