Algorithms in Comparative Genomics

Universität Bielefeld, SS 2024
Dr. Marília D. V. Braga · Dr. Roland Wittler · M.Sc. Leonard Bohnenkämper

https://gi.cebitec.uni-bielefeld.de/teaching/2024summer/cg

Exercise sheet 14, 12.07.2024

Exercise 1 (Number of maximal intervals) (3 pts)

Prove the following statement:

The number of maximal intervals grows in general as a quadratic function of the length of the sequence.

Exercise 2 (INT is commuting (Didier et al., 2007)) (5 pts)

The rank intervals for a given string are commuting, i.e., if $\mathit{INT}[k]$ and $\mathit{INT}[k']$ are two rank intervals, then $\mathit{INT}[k] \cap \mathit{INT}[k'] \in \{\emptyset, \mathit{INT}[k], \mathit{INT}[k']\}$. More precisely, if $\mathit{INT}[k] \cap \mathit{INT}[k'] \neq \emptyset$ we have:

1. $\mathit{INT}[k] \subseteq \mathit{INT}[k']$ if and only if $\mathit{RANK}[S[k]] \leq \mathit{RANK}[S[k']]$;
2. $\mathit{INT}[k] \supseteq \mathit{INT}[k']$ if and only if $\mathit{RANK}[S[k]] \geq \mathit{RANK}[S[k']]$;
3. $\mathit{INT}[k] = \mathit{INT}[k']$ if and only if $\mathit{RANK}[S[k]] = \mathit{RANK}[S[k']]$.

Prove one of the assertions.

Exercise 3 (Range Minimum Queries) (7 pts)

Have a look at the following paper to answer the questions below:

Bender, M. A., and Farach-Colton, M. *The LCA problem revisited.* Proceedings of CPM, 1776 (Chapter 9), 88–94, 2000.

1. What is the LCA problem?
2. What is the RMQ problem?
3. What is the ± 1RMQ problem?
4. Which problem is reduced to which, and what are the known or induced time complexities for preprocessing and query each?
5. What is the relation to the *rank distance* in Didier's algorithm?