Streszczenie w języku angielskim

This thesis consists of several results in the field of algorithms on strings. In particular, we consider variants of string covers and pattern matching with mismatches. From now on, we will use n to denote the length of the input text.

In the first part we consider exact variants of finding string covers, that is, testing whether a text can be generated by concatenations and superpositions of a shorter string.

In "Efficient Computation of 2-covers of a String" we present an algorithm which tests whether a text can be generated by concatenations and superpositions of 2 strings of the same length. The algorithm works in $O(n \log n \log \log n)$ expected time and $O(n)$ space.

"Shortest Covers of All Cyclic Shifts of a String" solves the titular problem in $O(n \log n)$ time and $O(n)$ space. It also provides a precise description of the set of lengths of the shortest covers of all cyclic shifts of Fibonacci words.

"Internal Quasiperiod Queries" shows how to efficiently answer internal queries about various types of quasiperiod related queries. The paper presents an algorithm which allows to find the shortest cover on a given interval in $O(\log n \log \log n)$ or all covers in $O(\log n (\log \log n)^2)$ time and $O(n \log n)$ space. We also present more efficient solutions in the offline model.

"String Covers of a Tree" considers a natural generalization of cover, where one can cover a tree with paths corresponding to the cover. The algorithm from the paper finds all covers in a directed tree in $O(n \log n / \log \log n)$ time and in the case of undirected tree in $O(n^2)$ time and space. In the case of undirected tree we provide an alternative algorithm using $O(n)$ space and $O(n^2 \log n)$ time.

The second part of the thesis provides asymptotically efficient solutions to practical problems.

"Efficient Computation of Sequence Mappability" considers a problem where one needs to compute mappability for each substring of length m with up to k mismatches. There, we provide a collection of results, most notably an $O(n \min \{ \log^k n, m^k \})$ time and $O(n)$ space algorithm for fixed k and that one cannot solve (k,m)-mappability in strongly subquadratic time unless Strong Exponential Time Hypothesis fails.

"Circular Pattern Matching with k Mismatches" is a variant of pattern matching where we allow the pattern to be arbitrarily circularly shifted and up to k characters replaced. We provide a simple $O(nk)$ time and $O(m)$ space solution and more elaborate $O(n + nk^4/m)$ time and $O(m)$ space algorithm.

The results usually exploit periodicity analysis and apply a mix of classic stringology tools and state-of-the-art data structures.