

## **SYLLABUS**

### Part I.

- introduction to string-specific issues (alphabet size and type, storage, string comparison, sets of strings)
- non-indexed based pattern matching algorithms: KMP, Boyer-Moore, Aho-Corasick
- k-mer index, Karp-Rabin pattern matching
- string sorting algorithms: StringMergeSort, StringQuickSort, LSD radix sort
- tries

### Part II.

- suffix trees (ST): simple properties, questions of storage and navigation, storage space
- first applications of STs: pattern matching (different variants), most frequent substring, longest repeat, LCE, shortest unique substring, maximal repeats, LZ compression
- ST construction: naive construction algorithm, WOTD algorithm
- generalized suffix trees, LCA queries, further applications of STs: maximal palindromes, MUMs
- suffix links, ST construction in linear time: McCreight's algorithm
- other applications of suffix links: LCF, space reduction, matching statistics

### Part III.

- suffix arrays (SA): properties, construction from ST in linear time, LCP-array, inverted suffix array
- pattern matching in SA: binary search, simple scan, with LCP-comparisons, with LCP-array
- SA construction in linear time: SAIS algorithm
- LCP construction in linear time: Kasai et al.'s algorithm
- ST construction from SA and LCP in linear time: suffix insertion algorithm
- applications of SAs other than pattern matching: matching statistics, LZ compression, etc.

### Part IV.

- backward search
- Burrows-Wheeler-Transform (BWT)
- transformation, retransformation: quadratic and linear time
- properties of the BWT, compressibility