

Georg-August-Universität Göttingen Module M.Inf.1114: Algorithms on Sequences	5 C 4 WLH
Learning outcome, core skills: We expect that the participants will gain an understanding of classical string-processing tools. They are supposed to understand and be able to use in various situations: classical text algorithms (e.g., pattern matching algorithms, edit distance), classical text indexing data structures (e.g., suffix arrays / trees), and classical combinatorial results that are useful in this context (e.g., periodicity lemmas).	Workload: Attendance time: 56 h Self-study time: 94 h
Course: M.Inf.1114.Lec Algorithms on Sequences (Lecture, Exercise) <i>Contents:</i> This course is an introduction into the theory of stringology, or algorithms on sequences of symbols (also called words or strings). Our main intention is to present a series of basic algorithmic and combinatorial results, which can be used to develop efficient word-processing tools. While the emphasis of the course is on the theoretical side of stringology, we also present a series of applications of the presented concepts in areas like data-compression or computational biology The main topics our course will cover are: basic combinatorics on words, pattern matching algorithms, data structures for text indexing (suffix arrays, suffix trees), text compression (Huffman encoding, Lempel-Ziv method), detection of regularities in words, algorithms for words with don't care symbols (partial words), word distance algorithms, longest common subsequence algorithms, approximate pattern matching. The presentation of each theoretical topic from the above will be accompanied by a brief discussion on its possible applications. Literature <ul style="list-style-type: none"> • T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein: Introduction to Algorithms (3rd Edition), MIT Press, 2009. • M. Crochemore, C. Hancart, T. Lecroq: Algorithms on Strings, Cambridge University Press, 2007. • M. Crochemore, W. Rytter: Jewels of Stringology, World Scientific, 2002. • D. Gusfield. Algorithms on strings, trees, and sequences: computer science and computational biology. Cambridge University Press, 1997. 	4 WLH
Examination: Oral examination (approx. 20 minutes) M.Inf.1114.Mp: Algorithms on Sequences Examination requirements: basic combinatorics on words, pattern matching algorithms, data structures for text indexing (suffix arrays, suffix trees), text compression (Huffman encoding, Lempel-Ziv method), detection of regularities in words, algorithms for words with don't care symbols (partial words), word distance algorithms, longest common subsequence algorithms, approximate pattern matching	5 C
Admission requirements: none	Recommended previous knowledge: none

Language: English	Person responsible for module: Prof. Dr. Florin Manea
Course frequency: irregular	Duration: 1 semester[s]
Number of repeat examinations permitted: twice	Recommended semester:
Maximum number of students: 50	