Fakultät für Ingenieurwissenschaften, Informatik und Psychologie

Informatik-Fachvortrag

Dienstag, den 11. Februar 2025, 14:00 Uhr
Universität Ulm, Oberer Eselsberg
Gebäude O27, Raum 531

Prof. Dr. Zsuzsanna Lipták
University of Verona

spricht zum Thema

How to use the Burrows-Wheeler Transform to construct random de Bruijn sequences

The Burrows-Wheeler Transform (BWT) is a reversible string transform at the core of some of the most frequently used bioinformatics tools, such as the aligners bwa and bowtie. The BWT takes advantage of the input string's repetitiveness for compression, while also allowing fast pattern matching in compressed space.

In this talk I will present a different application of the BWT, which is also relevant for bioinformatics research.

De Bruijn sequences (dB sequences) are strings with very little repetitiveness: a dB sequence of order k contains every k-mer exactly once. These sequences have been applied in several bioinformatics contexts, including probe design, microarray design, and DNA synthesis. Outside of bioinformatics, dB sequences have applications in pseudo-random bit generation, experimental design, and cryptography, to name just a few.

While there exist many algorithms for constructing dB sequences, most of these construct a specific dB sequence, or members of a specific class of dB sequences, representing only a tiny fraction of the complete set (whose cardinality is doubly exponential in k). Here, we present an algorithm for constructing random dB sequences which uses the extended Burrows-Wheeler Transform [Mantaci et al., 2007]. Our method is simple to implement (less than 120 lines of C++ code) and can produce random dB sequences of any order. Prior to our work, there were no practical tools for randomly generating dB sequences which output every dB sequence with positive probability.

This is joint work with Luca Parmigiani and was published at LATIN 2024. The code is available (in C++ and python) at https://github.com/lucaparmigiani/rnd_dbseq.

Es laden ein die Dozenten der Fakultät für Ingenieurwissenschaften, Informatik und Psychologie.

Ulm, den 09.01.2025

gez. Prof. Dr. Enno Ohlebusch