

Center for Memory and Recording Research
&
Department of Electrical and Computer Engineering

PRESENT



JACK KEIL WOLF
LECTURE SERIES
IN
INFORMATION THEORY
AND
APPLICATIONS

PROF. OLGICA MILENKOVIC

May 22, 2019

Lecture 2 PM & Reception 3 PM

JKW AUDITORIUM

CMRR Building

- biography -

Professor Olgica Milenkovic

Olgica Milenkovic is a professor of Electrical and Computer Engineering at the University of Illinois, Urbana-Champaign (UIUC), and Research Professor at the Coordinated Science Laboratory. She obtained her Masters Degree in Mathematics in 2001 and PhD in Electrical Engineering in 2002, both from the University of Michigan, Ann Arbor. Prof. Milenkovic heads a group focused on addressing unique interdisciplinary research challenges spanning the areas of algorithm design and computing, bioinformatics, coding theory, machine learning and signal processing.

Her scholarly contributions have been recognized by multiple awards, including the NSF Faculty Early Career Development (CAREER) Award, the DARPA Young Faculty Award, the Dean's Excellence in Research Award, and several best paper awards. In 2013, she was elected a UIUC Center for Advanced Study Associate and Willett Scholar while in 2015 she was elected Distinguished Lecturer of the Information Theory Society. In 2018 she became an IEEE Fellow. In 2009, she was the Guest Editor in Chief of a special issue of the IEEE Transactions on Information Theory on Molecular Biology and Neuroscience.

- abstract -

"DNA Punch-Cards: Implementations and Coding-Theoretic Questions"

Recent implementations of synthetic DNA-based data storage systems have demonstrated several promising applications of macromolecular recorders. However, the proposed systems suffer from high cost, read-write latency and error-rates that render them non-competitive with traditional silicon-based devices. One means to avoid synthesizing DNA is to use readily available, naturally occurring DNA. As the nucleotide sequences of native DNA are fixed, they cannot be edited to accommodate arbitrary user-defined content. Hence, instead of changing the sequence content, one may adopt an alternative recording strategy -- akin to card punching -- that modifies the topology of native DNA to encode desired information. We describe the first macromolecular storage paradigm in which data is written in the form of "nicks" at predetermined positions on the sugar-phosphate backbone of double-stranded native DNA. The platform accommodates parallel nicking on one and multiple genomic DNA fragments, and paired nicking and disassociation for creating "toehold" regions that enable single-bit random access and strand displacement. It also provides a large mass of inexpensive DNA that may be used for multiple, error-free readout cycles via current sequencing technologies. As a proof of concept, we used the multiple-turnover programmable artificial restriction enzymes to punch both text and image files into the PCR products of *Escherichia coli* genomic DNA fragments *in vitro*. The encoded data was reliably reconstructed through sequence alignment and read coverage analysis. The described storage implementation is accompanied by a number of new coding-theoretic questions and constructions connecting combinatorial designs, set discrepancy theory and trace reconstruction.