



ZAJEDNIČKI SEMINAR

Četvrtak, 15. veljače 2018., 15:00 sati (točno)

PMF-Fizički odsjek, Bijenička cesta 32, Predavaonica F-102

Ivo Grosse

Institute of Computer Science, Martin Luther University Halle-Wittenberg, Germany

German Centre for Integrative Biodiversity research (iDiv) Halle-Jena-Leipzig, Germany

Statistical Pattern Recognition in DNA Sequences using Tsallis Entropies

The recognition of statistical patterns in DNA sequences is one of the classical problems in bioinformatics. One successful approach for modeling and recognizing statistical patterns is based on the maximum entropy principle, which utilizes the probability distribution that maximizes the Shannon entropy under user-specified constraints. Here, we investigate a generalization of this principle based on the Tsallis entropy, which contains the Shannon entropy as a special case. While the constrained maximization of the Tsallis entropy cannot be accomplished analytically even for simple constraints, we find that a numerical optimization using the Lagrangian duality principle is feasible due to the convexity of the dual function. When applying the resulting maximum Tsallis entropy models to real-world data, we find that these models improve the pattern recognition performance over the corresponding maximum Shannon entropy models for many instances.

Joint work with Ralf Eggeling, Alexander Gabel, Christian Guenther, Jens Keilwagen, and Christiane Tammer.

Ivo Grosse received his Ph.D. at Boston University in 1999 and has been working on computational biology, information theory, and machine learning at Free University Berlin, Cold Spring Harbor Laboratory, the Leibniz Institute of Plant Genetics and Crop Plant Research Gatersleben, and Martin Luther University Halle-Wittenberg, where he is professor of bioinformatics at the Institute of Computer Science since 2007.