COLLOQUIUM

Sandra Zilles
University of Regina

Order Compression Schemes for (0,1)-Matrices

Mathematics and Statistics

Date, Time & Room: September 12, 3:30 - 4:30 PM, RIC 209

Abstract: In Computational Learning Theory, a (0,1)-matrix is interpreted as a concept class; every column corresponds to an element of an instance set, and every row corresponds to a concept, i.e., a subset of the instance set. Learning a concept means to identify or approximate a row of the matrix from partial information about that row. Each entry in a row is called a labelled example for the corresponding concept. An important combinatorial parameter expressing the complexity of learning a concept class is the Vapnik-Chervonenkis Dimension (VCD), which has been studied in various subfields of Mathematics and Statistics.

In this context, a sample compression scheme (SCS) is an algorithm for "encoding" a set of labelled examples for a concept in a small subset of examples. The long-standing open sample compression conjecture states that, for any concept class C of VCD d, there is an SCS in which sets of labelled examples are compressed to subsets of size at most d. We show that every order over C induces a special type of SCS for C, which we call order compression scheme (OCS). It turns out that OCSs can compress to sets of size at most d for many types of concept classes, in particular in most cases for which the sample compression conjecture is known to be true. Since OCSs are much simpler than SCSs in general, their study seems to be a promising step towards resolving the sample compression conjecture. We reveal a number of fundamental properties of OCSs that are helpful in such a study. In particular, OCSs exhibit interesting graph-theoretic properties.

Joint work with M. Darnstädt, T. Doliwa, and H. Simon, Ruhr-University Bochum, Germany

University of Regina