From Information Theory to Combinatorics

A workshop in honour of Jànos Körner's research world

November 10th, 2017 **Aula Magna** **Via Salaria, 113 Roma**

9.00 Welcome speech
Giancarlo Bongiovanni
Dean of the Faculty of Information Engineering, Informatics, and Statistics

9.15 - 10.15 The method of types in Information Theory
Imre Csiszár
Hungarian Academy of Science, Alfréd Rényi Institute of Mathematics

Abstract: The method of types is a combinatorial method to evaluate probabilities via counting, that admits to derive asymptotically tight error bounds (primarily) for discrete memoryless models. The Budapest school of Information Theory, and specifically János Körner, had a decisive role in developing the underlying simple idea to a powerful general method. Its application has enabled us to solve (or achieve significant progress in) a variety of Important problems of Information Theory.
In this talk, after a short introduction into the method, some key achievements obtained via its application will be surveyed, including our classical results about universally attainable error exponents for discrete memoryless channels and on the capacity of arbitrarily varying channels. As an example of recent progress, results of my PhD students about error exponents for random access models will be mentioned. The beautiful results of János Körner via the method of types in combinatorial problems not involving probabilities should also be mentioned, but this is outside the scope of this talk.

10.20 - 10.50 Coffee break

10.50 - 11.50 Coding the mod 2 sum and a conjecture in additive combinatorics
Katalin Marton
Hungarian Academy of Science, Alfréd Rényi Institute of Mathematics

Abstract: In a paper with Jànos Körner “How to encode the modulo-2 sum of binary sources? (1979)” we described an optimal coding scheme for this problem in the case of two correlated i.i.d. sources defined by a pair of symmetrically correlated binary random variables. The key was to use linear codes. In trying to understand whether linear codes were absolutely crucial for this problem, I tried to understand the structure of the level sets of the codewords. This led to the question: What are those sets $A \subset F^n_2$ like for which the cardinality of $|A + A|$ (mod 2) is only slightly larger then the cardinality of A, i.e.: $|A + A| \leq 2^{n\delta} |A|$, where δ is small. I came to the conjecture that such a set must "look like" the union of a few cosets with respect to a subgroup $G \subset F^n_2$, with cardinality $2^{n\epsilon} |A|$, where "a few" means $2^{n\epsilon}$, and $\epsilon \to 0$ as $\delta \to 0$. The characterization of subsets in a group whose doubling only slightly increases cardinality is a widely studied subject, associated with the name of Freiman, who first studied this problem for the group of integers. In the case of the group F^n_2, however, there are meaningful results only under the condition $|A + A| \leq K |A|$,where K cannot be as large as $2^{n\delta}$.

11.55 - 12.55 Graph entropy and sorting: From classical to quantum
Gwenaël Joret
Université Libre de Bruxelles, Computer Science Department

Abstract: One well-known application of Körner's graph entropy is in the context of sorting under partial information: Given a partial order P compatible with the linear order we are looking for, the problem is to make as few comparisons of the form "is x < y?" as possible to identify the linear order. In the early 1990's, Kahn and Kim showed that if G denotes the incomparability graph of P and H(G) denotes its entropy, the quantity |G|*H(G) approximates to within a constant factor the so-called information-theoretic lower bound (ITLB) for the problem in the decision tree model, the logarithm of the number of linear extensions of P. Building on this and the fact that H(G) can be approximated to any fixed precision in polynomial time when G is perfect, they developed a poly-time algorithm for the problem performing a number of comparisons bounded by c*ITLB for some constant c. A decade later, Yao raised the question of whether significantly better algorithms exist on quantum computers: Is there an efficient algorithm performing o(ITLB) comparisons? Conjecturing that the answer is no, Yao considered a natural quantum lower bound (QLB) in this setting and related it to graph entropy, by showing that QLB >= a*|G|*(H(G) - b) for some constants a, b > 0. To do so,

he studied an "average" version of graph entropy. This almost proves his conjecture but for the pesky -b term: If the entropy H(G) is at least someconstant > b, then QLB >= eps*ITLB for some eps > 0, as desired. However, the case of small entropy remains open. In this talk I will give a gentle introduction to this topic, focusing on elegant combinatorial ideas that emerged around graph entropy. I will also mention some ongoing work on the small-entropy case of Yao's conjecture that we are pursuing together with Jean Cardinal and Jérémie Roland.

12.55 - 14.30 Lunch

14.30 - 15.30 Information theory in combinatorics
Gábor Simonyi
Hungarian Academy of Science, Alfréd Rényi Institute of Mathematics

Abstract: The above title is a slight variation of the title of my PhD thesis prepared under the supervision of János Körner more than a quarter of a century ago. I met János more or less just in time for being able to witness how he started to work in combinatorics and to participate in his adventure of posing and trying to solve combinatorial problems that were inspired by information theoretic questions and concepts. In the talk I try to highlight some of the notions and results that emerged from these efforts.

15.35 - 16.35 Cancellative Families and Cancellative Pairs: A Survey
Ron Holzman
Technion, Israel Institute of Technology

Abstract: A family $\mathcal{A}$ of subsets of {1,…,n} is said to be *cancellative* if no three distinct sets A,B,C $\in \mathcal{A}$ satisfy B-A = C-A . In the 1970's Erdos and Katona gave a construction of a cancellative family of size $3^{n/3}$ (or approximately that, if n is not divisible by 3), which for a long time was believed to be optimal. We shall describe subsequent developments concerning cancellative families, including one due to Tolhuizen, which determined the correct asymptotics for the problem. We shall also present the state of affairs regarding the analogous problem for pairs of families ($\mathcal{A},\mathcal{B}$), where the above condition is applied to sets B,C from one family, A from the other. The latter is related to Simonyi's 'Sandglass Conjecture'. For a long time, the best known upper bound for pairs was due to Körner and the speaker, but there are new developments, the most recent one by Janzer. All these problems have motivations coming from information theory.

16.40 - 17.10 Coffee break

17.10 - 18.10 Better bounds for perfect hashing into a 4-element set

Jaikumar Radhakrishnan
School of Technology and Computer Science, Tata Institute of Fundamental Research

Abstract: What is the largest size of a universe that can be perfectly hashed into a k-element set? Equivalently, how large can a subset C of $\{1,2,\ldots,k\}^n$ be such that every k distinct elements of C have a coordinate where they all differ? We obtain an improved upper bound on the size of such sets for k=4. Specifically, we prove that such a subset of $\{1,2,3,4\}^n$ has size at most $2^{6n/19+o(n)}$, or rate at most 6/19 < 0.3158 measured in bits. This improves the previous best upper bound of 0.3512 due to (Arikan 1994), which in turn improved the 0.375 bound that followed from general bounds for perfect hashing due to (Fredman and Komlós, 1984) and (Körner and Marton, 1988). This question has another context besides hashing, namely, the zero-error list decoding capacity of certain channels.
Our approach is based on a probabilistic combination of the Plotkin bound in coding theory and Hansel's lemma for covering the complete graph by bipartite graphs.
This is joint work with Marco Dalai and Venkat Guruswami.

18.05 Conclusions