Interview with

Katalin Marton

Shannon Award Winner 2013
IEEE International Symposium on Information Theory
Istanbul, Turkey
July 2013
KöMaL - Középiskolai Matematikai és Fizikai Lapok

In 1894, Dániel Arany, a high school teacher from the city of Győr, founded a mathematical journal for high school students. His goal was "to give a wealth of examples to students and teachers". From that time several generations of mathematicians and scientists developed their problem-solving skills through KöMaL. The best solutions with the names of the 14-18 year-old authors are printed in the periodical.
Studied Mathematics at Eötvös Loránd University
Largest university in Hungary founded in 1635.

Notable alumni:

• John von Neumann
• Paul Erdös
• John Harsanyi
• László Lovász
• Karl Polanyi

• Endre Szemerédi
• Imre Csiszár
• Peter Gács
• Janós Körner
• Katalin Marton
Takes Alfred Renyi’s Combinatorics Seminar, 1966
PROCEEDINGS
OF THE COLLOQUIUM
ON INFORMATION THEORY
Kossuth Lajos University, 1967

EDITED BY
A. RÉNYI

The Organizing Committee of the meeting was composed as follows:

Alfred RÉNYI (Chairman), Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Mrs.) Katalin BOGNÁR (Secretary), University L. Eötvös, Budapest

Sándor CSBY, Research Institute for Telecommunication, Budapest

Imre CSISZÁR, Mathematical Institute of the Hungarian Academy of Sciences

Zoltán DARÓCZY, University L. Kossuth, Debrecen

Gyula KATONA, Mathematical Institute of the Hungarian Academy of Sciences

Gábor TUSNÁDY, Mathematical Institute of the Hungarian Academy of Sciences.

19 September, Tuesday
Afternoon session
3:00-5:30 p.m.

Chairman: B. GYRES (Debrecen)

F. L. DOBRUSHIN (Soviet Union):
The capacity of a channel with synchronization error. (In Russian)

C. PICARD (France):
Théorèmes de l’information traitée.

M. S. WATANABE (USA):
Where does the information come from?

Z. DARÓCZY (Hungary):
On a new characterization of Shannon’s entropy.

Chairman: R. L. DOBRUSHIN (Soviet Union)

M. S. PINSKER (Soviet Union):
Gaussian channels with feedback. (In Russian)

E. A. GRAHAM (France):
Detection, false-alarm, and threshold in radiometric systems.

J. CERNY (Czechoslovakia):
On the composition of information sources and finite automata.
Roland Lvovich Dobrushin
“The place was incomparably inferior to our building in Budapest: a few utterly run-down rooms on a floor of a nondescript Moscow building: the group just called it “the stable”, and probably went there only for the seminars.”

Peter Gács
Abstract—The aim of the article is to obtain upper and lower asymptotic bounds for $\varepsilon$-entropy of stationary sources and discrete time, with a finite number of states, and with a criterion of reproduction accuracy specified in terms of an additive loss function. In general the estimates obtained by us may be in asymptotic disagreement. They necessarily agree in the case of Markov sources. In §1 we define concepts needed in formulating the problems and we introduce some notation. In §2 we formulate the results; proofs of the results are contained in §3, §4, and §5.
Starts collaborating with Imre Csiszár, 1972
Error Exponent for Source Coding with a Fidelity Criterion

KATALIN MARTON

Abstract—For discrete memoryless sources with a single-letter fidelity criterion, we study the probability of the event that the distortion exceeds a level \( d \), if for large block length the best code of given rate \( R > R(d) \) is used. Lower and upper exponential bounds are obtained, giving the asymptotically exact exponent, except possibly for a countable set of \( R \) values.

For \( R > 0 \) define

\[
P^n(R,d) = \min_{|B^n| \leq \exp (Kn)} p^n(H(B^n,d))
\]

where the min is taken over all codes \( B^n \subset Y^n \) such that \( |B^n| \leq \exp (Kn) \), and \( p^n \) denotes the product distribution.
Starts working at the Alfréd Rényi Institute of Mathematics, 1974

“Of course, the pay was not much, and the building, excellently located, elegant and comfortable, did not have a lot of space.... You could not sit and work in the room under such conditions: somebody was always on the phone or in a discussion. But it was a friendly and stimulating environment, and conversations we did have! ...”

Peter Gács
Alfréd Rényi Institute of Mathematics

“The Mathematical Institute was (and has been ever since) a wonderful place. Its founder, Alfréd Rényi, formed his view of its researchers on the example of Pál Turán, who while in the army forced labor division during the war, working on the top of an electric pole, still solved mathematical problems. As he saw it all these people needed was some decent conditions for work and collaboration. There were essentially no obligations, meetings, formal performance measurements, reports. Rényi created the systems to defend this oasis from the outside; moreover, took in researchers who for political reasons were not welcome at the universities.”

Peter Gács
János Körner
Images of a Set via Two Channels and their Role in Multi-User Communication

JÁNOS KÖRNER AND KATALIN MARTON

Abstract—A technique is presented to determine the region of achievable rates for some source and channel networks. This technique is applied to the solution of a source network problem that seems to be the simplest illustration of a new typical difficulty in coding for source networks: namely, when the same encoding of a source is required to meet the conflicting demands of 1) supplying side-information to the decoder of another source, and 2) providing direct-information to its own decoder in company with other side-information.

To be more explicit, the encoders $E_X$, $E_Y$, and $E_Z$ observe the first $n$ outputs $X^n$, $Y^n$, and $Z^n$ of their respective sources. They produce the functions

$$f_n = f_n(X^n) \quad g_n = g_n(Y^n) \quad h_n = h_n(Z^n).$$

The task of decoder $D_X$ is to construct a function $\varphi_n = \varphi_n(f_n)$ such that $\varphi_n(f_n(X^n))$ equals $X^n$ with high probability. Similarly, $D_Y$ and $D_Z$ produce functions $\psi_n = \psi_n(g_n)$ and $\omega_n = \omega_n(h_n)$ such that $\omega_n(g_n(Y^n) \oplus (X^n))$
How to Encode the Modulo-Two Sum of Binary Sources

JÁNOS KÖRNER AND KATALIN MARTON

Abstract—How much separate information about two random binary sequences is needed in order to tell with small probability of error in which positions the two sequences differ? If the sequences are the outputs of two correlated memoryless binary sources, then in some cases the rate of this information may be substantially less than the joint entropy of the two sources. This result is implied by the solution of the source coding problem with two separately encoded side information sources for a special class of source distributions.
Random Access Communication and Graph Entropy

J. Körner and K. Marton

Abstract—Conflict resolution in random access communication raises the following probabilistic problem. Let $U_1, \cdots, U_k$ be independent random variables uniformly distributed over the unit interval $[0,1]$. A $k$-partition $\mathcal{A}$ of $[0,1]$, i.e., a partition into $k$ atoms, separates the random points $U_1, \cdots, U_k$ if each atom contains exactly one of the $U_i$. For $k$ partitions $\mathcal{A}_1, \cdots, \mathcal{A}_n$, let $P_{\mathcal{A}_1, \cdots, \mathcal{A}_n}(k)$ be the probability of the event that at least one of the $\mathcal{A}_j$ separates $U_1, \cdots, U_k$. The maximum of these probabilities is of interest when $\mathcal{A}_1, \cdots, \mathcal{A}_n$ vary. Hajek conjectured that $P_{\mathcal{A}_1, \cdots, \mathcal{A}_n}(k)$ is maximized if $\mathcal{A}_1, \cdots, \mathcal{A}_n$ are chosen in such a way that their common refinement is an equipartition of $[0,1]$ into $k^n$ subsets, i.e., $1 - \sum_{i=1}^{n-1} P_{\mathcal{A}_1, \cdots, \mathcal{A}_n}(k) \geq (1 - (k!/k^k))^n$. This conjecture was supported by considerations of symmetry and by the Van der Waerden–Falikman–Egorychev theorem on permanents. This conjecture is disproved by showing $\min (1 - P_{\mathcal{A}_1, \cdots, \mathcal{A}_n}(3)) = 25/81$. The main result is the lower bound $1 - P_{\mathcal{A}_1, \cdots, \mathcal{A}_n}(k) \geq 2^{-n k^2/k^2 - 1}$. This is achieved by a new technique for lower-bounding the number of graphs of given structure needed to cover all edges of a given graph. This technique, developed by Körner, is based on the subadditivity of graph entropy—a functional on graphs.
On the Capacity of Uniform Hypergraphs

JÁNOS KÖRNER, MEMBER, IEEE, AND KÁTALIN MARTON

Abstract — The capacity of uniform hypergraphs can be defined as a natural generalization of the Shannon capacity of graphs. Corresponding to every uniform hypergraph there is a discrete memoryless channel in which the zero error capacity, in case of the smallest list size for which it is positive, equals the capacity of the hypergraph, and vice versa. Also, the problem of perfect hashing can be considered as a hypergraph capacity problem. We derive upper bounds for the capacity of uniform hypergraphs, using the technique developed earlier for perfect hashing based on the concepts of graph entropy and hypergraph entropy. These are subadditive functionals on probabilistic graphs resp. hypergraphs (i.e., graphs resp. hypergraphs with a probability distribution given on their vertex sets). We also give a modified version of this technique, replacing graph entropy by another subadditive functional on probabilistic graphs. This functional can be considered as a probabilistic refinement of Lovász’ $\Theta$-functional.
A Coding Theorem for the Discrete Memoryless Broadcast Channel

KATALIN MARTON

Abstract—A coding theorem for the discrete memoryless broadcast channel is proved for the case where no common message is to be transmitted. The theorem is a generalization of the results of Cover and van der Meulen on this problem. The result is tight for broadcast channels having one deterministic component.

1. INTRODUCTION

A DISCRETE memoryless broadcast channel (DMBC) as defined by Cover [1] is determined by a pair of discrete memoryless channels with common input alphabet $\mathcal{X}$. We denote by $F$ and $G$ the transition probability matrices of these channels:

$$F = \{ F(x|y); y \in \mathcal{Y}, x \in \mathcal{X} \}, \quad G = \{ G(z|y); y \in \mathcal{Y}, z \in \mathcal{Z} \}.$$ 

Here $\mathcal{X}$ and $\mathcal{Z}$ are the output alphabets, with cardinalities $|\mathcal{Y}|, |\mathcal{X}|, |\mathcal{Z}| < \infty$. The DMBC corresponding to the matrices $F, G$ will be denoted by $(F, G)$.

We assume that the conditional probabilities of receiving the sequences $x^n \in \mathcal{X}^n$ and $z^n \in \mathcal{Z}^n$ at the outputs of the channels $F$ and $G$, respectively, are given by

$$F^n(x^n|y^n) = \prod_{i=1}^{n} F(x_i|y_i), \quad G^n(z^n|y^n) = \prod_{i=1}^{n} G(z_i|y_i)$$

where $y^n = y_1 y_2 \cdots y_n$.

In connection with the DMBC $(F, G)$, we consider the following coding problem. A sender has to transmit two independent messages over the channels $F$ and $G$: one message for receiver 1 observing the output of the channel

“I first met Kati at the 1975 IT workshop at the Lake Balaton. At that time, there was a strong movement to connect with our colleagues behind the Iron Curtain. The meeting in Hungary was the first official sign of reapprochement.

I remember that modest, unassuming lady who gave a presentation that went completely over my head and who was to become a mainstay in our Society activities. Kati was so shy, sincere, and the antithesis of "flamboyant" that I could hardly find personal moments or events that made her stand out. Her work was always appreciated and thought very highly of in our Society. So, it was no surprise to me to see her in the corridors of MIT during a visit there in the late seventies. It turned out she was spending a leave from her home Institution for a few years in the US.”

Anthony Ephremides
In the early 1970's a workshop on information theory was organized in Udine by Giuseppe Longo. One day we (Kati and 3 boys) went for an excursion to the Dolomites mountains. Kati was wearing a pair of shoes, somewhat elevated (not high heels, what she probably never had), certainly nor perfect for hiking. On the other hand she was not a sporty lady. After a while the mountain trail entered a snow covered area. As Kati later explained, she started to have fear, but did not want to ruin the excursion, so she continued her way with us. We went a couple of hundred meters in the snow when suddenly Kati fell down and started to slide down in the snow. Fortunately, one of us, Gyorgy Szanto was walking exactly along her orbit and was able to catch her. Otherwise her slide could have been fatal with high probability.

“A classmate of her, Janos Komlos, said once that "Marton Kati was a clever girl". This saying became well-known because Komlos did not like to praise people.

Gyula Katona
"Kati and I met several times in the seventies and eighties, both in Udine and elsewhere, and I was always pleased to talk to her either about scientific issues or on general topics. In spite of her talents she is an unostentatious person, very friendly and amiable.

Kati was testing my proficiency in Hungarian during the summer schools in Udine (I was trying to learn that language at the time). She was the most patient and easygoing "teacher" in the group of Hungarians that attended the courses, and she smiled sweetly every time I said something in the right way."

Giuseppe Longo
First ISIT, First Visit to US
“Computer scientists and mathematicians whose research touches on secret codes say they have been subjected by the National Security Agency to growing harassment and the threat of sanctions or even prosecution, for publishing articles about their work.

Such allegations have been simmering in the scientific press for six months, but the controversy was brought to a boil by a symposium on information theory held last week at Cornell University, Ithaca, NY.”

The New York Times, October 19, 1977
“As a child, I learned some of the most basic ideas in mathematics from my mother. For example, she showed me how parentheses could be used to structure a formula, or how to construct an equilateral triangle using a compass and a ruler. Later, she taught me some more advanced things, like how the Euclidean algorithm worked for polynomials as well as integers.

She had an important influence on my becoming a mathematician. At an early age, I learned from her to appreciate the beauty and power of mathematics, as well as the beauty of many other things, such as works of Leonardo or Mozart.”

Peter Frenkel
Second visit to US
1979-80
• Robert Gallager, MIT
• Robert Gray, Stanford
“I first became acquainted with Kati when she wrote me a letter, I think in 1972, discreetly asking if there might not be a problem with an inequality in a paper I had published in 1971 on rate-distortion functions for Markov sources. Buried in many lines of inequalities, Kati had found a mistake. I quickly realized she was correct and, although I was embarrassed, I was happy the problem had been discovered and set about to fix it. The following year I published a correct proof as a correspondence, but I had to add a condition to get the result - as I am sure she knew would be necessary. I am still amazed someone found that error. Given there are so few Kati Martons in the world, I often wonder how many other deeply buried errors are out there awaiting discovery.”

Robert Gray
“I first met Kati in person at the Janos Bolyai Colloquium on Information Theory in Keszthely, Hungary, in August 1975. ..... the attendance by US participants in Keszthely was sparse. In a way that made the event even more interesting for those few of us from the US who attended -- including Peter Elias, Tony Eprhemides, Dick Blahut, and me. Among my strongest memories is a discussion session with Kati and Janos Korner talking about their early results in multiuser information theory. Kati posed a problem of source coding over a simple network. Aaron Wyner and I had published a result the previous year in BSTJ that superficially resembled the problem she was working on, and I thought our result might solve her problem. I quickly realized, however, that her problem was much deeper and more complicated since our problem allowed cooperative coding and hers did not. This was 1975 when the rebirth of multiuser information theory was still young and as time passed it became evident that she and Janos had grasped a much more interesting and fruitful thread for future development. Once again I was left in awe of her abilities.”

Robert Gray
“Rudi Ahlswede, at a visit in Budapest, posed a question that I found solvable. I remembered a Dobrushin seminar presentation of Margulis about a certain phase transition in random graphs, in particular a certain isoperimetric theorem used in the proof. This is how the Blowup Lemma was born. We had a pleasant collaboration with János Körner in the course of working it into a paper, along with a couple of other results.”

Peter Gács
A Simple Proof of the Blowing-Up Lemma

K. MARTON

Abstract—The blowing-up lemma says that if the probability with respect to a product measure of a set $A \subseteq X^n$ ($X$ finite, $n$ large) is not exponentially small, then its $l_n$-neighborhood has probability almost one for some $l_n = o(n)$. Here an information-theoretic proof of the blowing-up lemma, generalizing it to continuous alphabets, is given.

Manuscript received August 9, 1984; revised June 7, 1985.

The author is with the Mathematical Institute, Hungarian Academy of Sciences, H 1364 Budapest, P.O. Box 127, Hungary.
IEEE Log Number 8406951.

“Marton is one of the leading authorities about the applications of information theory techniques to concentration theory, in particular in the setting of Markov Chains.

Most importantly, in the mid-nineties, Marton pointed out the interest and importance of entropy inequalities in the study of the concentration phenomena. Talagrand has acknowledged the influence of Marton in this respect, and this motivated him to establish the famous Talagrand inequality controlling the Wasserstein distance by the square root of the Boltzmann-Shannon information.

In turn, the Talagrand inequality triggered the development a whole field, which I explored with Otto, McCann, Lott and others, involving entropy, concentration, transport, Ricci curvature, with very far-reaching geometric consequences.”

Cedric Villani, Fields Medallist
“Her approach to the problem of measure concentration via transportation inequalities and coupling was absolutely beautiful and original and offered novel elegant ways of discovering this exciting area at the boundary of probability, geometry, and information theory. In spite of the presence of various competing methods—martingales, logarithmic Sobolev inequalities, and Talagrand's induction method, Marton’s technology provides the strongest results in many situations.”

Gabor Lugosi
“She is extremely bright. ... sometimes I asked something from her. Her response was somewhat in the following style: she has a slow way of diction, partially she may be shy, partially for she is really thinking on the answer. Her answer in any case is usually something of the following type: Either my question was not correct, OR I could have found the answer myself, OR if not any of the previous cases, then she gives an absolutely clear, down to the depth, concise response. Absolutely precious quality!

I also know that in the theory of concentration inequalities she is much respected. An ex-undergraduate of mine, Daniel Paulin is a graduate student writing his thesis on concentration. I know that for him Katalin is a god of the topic.

I am sure that having awarded her increases the value of the Shannon Medal.”

Domokos Szász
THE POSITIVE-DIVERGENCE AND BLOWING-UP PROPERTIES

BY

Katalin Marton*
Mathematics Institute, Hungarian Academy of Sciences

AND

Paul C. Shields**
University of Toledo and Eötvös Loránd University

ABSTRACT
A property of ergodic finite-alphabet processes, called the blowing-up property, is shown to imply exponential rates of convergence for frequencies and entropy, which in turn imply a positive-divergence property. Furthermore, processes with the blowing-up property are finitely determined and the finitely determined property plus exponential rates of convergence for frequencies and for entropy implies blowing-up. It is also shown that finitary codings of i.i.d. processes have the blowing-up property.
A MEASURE CONCENTRATION INEQUALITY FOR CONTRACTING MARKOV CHAINS

K. Marton

Abstract
The concentration of measure phenomenon in product spaces means the following: if a subset \( A \) of the \( n \)'th power of a probability space \( \mathcal{X} \) does not have too small a probability then very large probability is concentrated in a small neighborhood of \( A \). The neighborhood is in many cases understood in the sense of Hamming distance, and then measure concentration is known to occur for product probability measures, and also for the distribution of some processes with very fast and uniform decay of memory. Recently Talagrand introduced another notion of neighborhood of sets for which he proved a similar measure concentration inequality which in many cases allows more efficient applications than the one for a Hamming neighborhood. So far this inequality has only been proved for product distributions. The aim of this paper is to give a new proof of Talagrand's inequality, which admits an extension to contracting Markov chains. The proof is based on a new asymmetric notion of distance between probability measures, and bounding this distance by informational divergence. As an application, we analyze the bin packing problem for Markov chains.
BOUNDING $\tilde{d}$-DISTANCE BY INFORMATIONAL DIVERGENCE: A METHOD TO PROVE MEASURE CONCENTRATION

By K. Marton

Mathematical Institute of the Hungarian Academy of Sciences

There is a simple inequality by Pinsker between variational distance and informational divergence of probability measures defined on arbitrary probability spaces. We shall consider probability measures on sequences taken from countable alphabets, and derive, from Pinsker’s inequality, bounds on the $\tilde{d}$-distance by informational divergence. Such bounds can be used to prove the “concentration of measure” phenomenon for some nonproduct distributions.

Measure concentration for a class of random processes

Katalin Marton

Mathematical Institute of the Hungarian Academy of Sciences, P.O. Box 127, H-1364 Budapest, Hungary
e-mail: marton@math-inst.hu

Received: 6 May 1996 / In revised form: 29 September 1997
MEASURE CONCENTRATION AND STRONG MIXING

K. MARTON

Communicated by I. Csiszár

Abstract

We formulate the measure concentration inequality, for Hamming distance and Talagrand’s “convex hull” distance, in the case of dependent random variables, in a more general form than it was done in earlier papers [12] and [17]. This makes it possible to obtain measure concentration inequalities for Gibbs states over a box of the $\nu$-dimensional lattice with fixed boundary condition, in the case when the Gibbs state satisfies a strong mixing condition which is in between the Dobrushin–Shlosman condition and its weakening in the sense of Olivieri, Picco and Martinelli.

We also extend the use of the measure concentration inequality for Talagrand’s “convex hull” distance in the following direction. For r.v’s $Z(X^n)$ satisfying an inequality

$$Z(\hat{x}^n) - Z(x^n) \leq \sum_{k=1}^{n} \alpha_i(\hat{x}^n) d(\hat{x}_i, x_i),$$

we prove bounds for the momentum generating function of $Z(X^n) - \mathbb{E}Z(X^n)$ in terms of the momentum generating function and the expectation of $\sum_{k=1}^{n} \alpha^2_i(X^n)$.
MEASURE CONCENTRATION FOR EUCLIDEAN DISTANCE
IN THE CASE OF DEPENDENT RANDOM VARIABLES

By Katalin Marton
Hungarian Academy of Sciences

Let $q^n$ be a continuous density function in $n$-dimensional Euclidean space. We think of $q^n$ as the density function of some random sequence $X^n$ with values in $\mathbb{R}^n$. For $I \subset [1, n]$, let $X_I$ denote the collection of coordinates $X_i, i \in I$, and let $\overline{X}_I$ denote the collection of coordinates $X_i, i \notin I$. We denote by $Q_I(x_I|\overline{x}_I)$ the joint conditional density function of $X_I$, given $\overline{X}_I$. We prove measure concentration for $q^n$ in the case when, for an appropriate class of sets $I$, (i) the conditional densities $Q_I(x_I|\overline{x}_I)$, as functions of $x_I$, uniformly satisfy a logarithmic Sobolev inequality and (ii) these conditional densities also satisfy a contractivity condition related to Dobrushin and Shlosman’s strong mixing condition.
Prepared by
Deniz Gündüz
April-July 2013

Special Thanks to
Peter Frenkel
János Körner
Imre Csiszár
Peter Gács
Prakash Narayan
Anthony Ephremides
Gyula Katona
Giuseppe Longo
Robert Gray
Cedric Villani
Gabor Lugosi
Domokos Szász
Márton Magócsi