# **IEEE Information Theory Society Newsletter**

Vol. 59, No. 3, September 2009

Editor: Tracey Ho

# **Optimal Estimation**

XXX Shannon lecture, presented in 2009 in Seoul, South Korea (extended abstract)

### **1** Prologue

The first quest for optimal estimation by Fisher, [2], Cramer, Rao and others, [1], dates back to over half a century and has changed remarkably little. The covariance of the estimated parameters was taken as the quality measure of estimators, for which the main result, the Cramer-Rao inequality, sets a lower bound. It is reached by the maximum likelihood (ML) estimator for a restricted subclass of models and asymptotically for a wider class. The covariance, which is just one property of models, is too weak a measure to permit extension to estimation of the number of parameters, which is handled by various ad hoc criteria too numerous to list here.

Soon after I had studied Shannon's formal definition of information in random variables and his other remarkable performance bounds for communication, [4], I wanted to apply them to other fields - in particular to estimation and statistics in general. After all, the central problem in statistics is to extract information from data. After having worked on data compression and introduced arithmetic coding it seemed evident that both estimation and data compression have a common goal: in data compression the shortest code length cannot be achieved without taking advantage of the regular features in data, while in estimation it is these regular features, the underlying mechanism, that we want to learn. This led me to introduce the MDL or Minimum Description Length principle, and I thought that the job was done.

However, when starting to prepare this lecture I found that it was difficult, for I could not connect the several in themselves meaningful results to form a nice coherent picture. It was like a jigsaw puzzle where the pieces almost fit but not quite, and, moreover, vital pieces were missing. After considerable struggle I was able to get the pieces to fit but to do so I had to alter them all, and ignore the means and concepts introduced by the masters mentioned above. The result was separation of estimation from data compression, and we can now define optimality for all amounts of data and not just asymptotically.

### 2 Modeling Problem

The modeling problem begins with a set of observed data  $Y = \{y_t: t = 1, 2, ..., n\}$ , often together with other so-called explanatory data  $Y, X = \{(y_t, x_{1,t}, x_{2,t}, \ldots)\}$ . The objective is to learn properties in Y expressed by a set of distributions as models

 $\{f(Y \mid X_s; \theta, s)\}.$ 

Here *s* is a structure parameter and  $\theta = \theta_1, \ldots, \theta_{k(s)}$  real-valued parameters, whose number depends on the structure. The structure is simply a subset of the models. Typically it is used to indicate the most important variables in X. (The traditional name for the set of models is 'likelihood function' although no such concept exists in probability theory.)

The most important problem is the selection of the model class, but since its optimal selection is non-computable there is little we can say about it. It is clear that when picking it we must take into account the general type of the data, the sensitivity of the models with respect to the parameters, which is an issue in the so-called robust statistics, and so on.

To simplify the notations we write the data as x with the understanding that the models are distributions on Y or conditional distributions on Y given the explanatory data X if they too are observed. (In the first reading put just  $\mathbf{x} = Y$ ). We also consider only structures determined by the number of parameters so that we already fix the sequence  $\theta^k = \theta_1, \ldots, \theta_k$ , also written as  $\theta$ . We then have the two model classes

$$\mathcal{M}_{k} = \{ f(\mathbf{x}; \theta, k) : \theta \in \Omega^{k} \}, \ k \le n$$
$$\mathcal{M} = \bigcup_{k=0}^{K} \mathcal{M}_{k}, \ K \le n$$

depending on whether we are considering the number of parameters fixed or if it, too, is to be estimated. The latter class continued on page 6



**Celebrating 125 Years** of Engineering the Future



ISSN 1059-2362



J. Rissanen

# From the Editor

Dear IT Society members,

It was very nice to see many of you at ISIT in June. Many thanks to the organizers for a fantastic program and meticulously-run conference. This issue includes the reports from ISIT and ITW Volos, and announcements of awards from ISIT and elsewhere - warmest congratulations to all the award winners. I'm also delighted to include in this issue the first-ever published article on 2009 Shannon Lecturer Jorma Rissanen's new theory of optimal estimation. Jorma said that as he was preparing his Shannon Lecture, he unexpectedly encountered difficulties incorporating some of the basic results of traditional estimation theory in a coherent framework, so he set about creating a new framework for optimal estimation. His article in this issue summarizes his lecture and gives an exciting first look at these new ideas.

This issue also comes with sad news of the passing of Wesley Peterson, whose many achievements and honors include the Shannon Award in 1981. He is fondly remembered by his friends Ned Weldon and Shu Lin.

### IEEE Information Theory Society Newsletter

*IEEE Information Theory Society Newsletter* (USPS 360-350) is published quarterly by the Information Theory Society of the Institute of Electrical and Electronics Engineers, Inc.

Headquarters: 3 Park Avenue, 17th Floor, New York, NY 10016-5997.

Cost is \$1.00 per member per year (included in Society fee) for each member of the Information Theory Society. Printed in the U.S.A. Periodicals postage paid at New York, NY and at additional mailing offices.

**Postmaster:** Send address changes to IEEE Information Theory Society Newsletter, IEEE, 445 Hoes Lane, Piscataway, NJ 08854.

© 2009 IEEE. Information contained in this newsletter may be copied without permission provided that the copies are not made or distributed for direct commercial advantage, and the title of the publication and its date appear.

IEEE Information Theory Society Newsletter

For many of us, the summer is drawing to an end, and fall term is starting soon. Do take some time to relax with our fun regular features – the Historian's Column and Golomb's Puzzle Column. Also, an article commemorating the hundredth anniversary of telecommunications pioneer Vladimir Kotelnikov is included in this issue.

Please help to make the newsletter as interesting and informative as possible by offering suggestions and contributing news. The deadlines for the next few issues are:

Issue	Deadline
December 2009	October 10, 2009
March 2010	January 10, 2010
June 2010	April 10, 2010

Announcements, news and events intended for both the printed newsletter and the website, such as award announcements, calls for nominations and upcoming conferences, can now be submitted jointly at the IT Society website http://www.itsoc.org/, using the quick links "Share News" and "Announce an Event". For more details please see the article on the new website in this issue.

Articles and columns intended only for the printed newsletter should be e-mailed to me at tho@caltech.edu.

Please submit ASCII, LaTeX or Word source files; do not worry about fonts or layout as this will be taken care of by IEEE layout specialists. Electronic photos and graphics should be in high resolution and sent as separate files.

I look forward to your contributions and suggestions for future issues of the newsletter.

Tracey Ho

# Table of Contents

Optimal Estimation: XXX Shannon Lecture
From the Editor
President's Column
Annual Information Theory Society Awards Announced4
In Memoriam, Wes Peterson
The Historian's Column
IT Society Members Honored
Symposium Report: The 2009 IEEE International Symposium on Information Theory, Seoul, Korea9
Workshop Report: Networking and Information Theory, Volos, Greece
Golomb's Puzzle Column: Finding Sums
The Life Filled with Cognition and Action: 100th Anniversary of Academician V.A. Kotelnikov
Golomb's Puzzle Column: Some Partition Problems Solutions
Call for Papers
Conference Calendar





# **President's Column**

I have just returned from the main annual event of our society, the International Symposium on Information Theory (ISIT). This year ISIT was held in Seoul, Korea, returning to mainland Asia for the first time since 1973, when Shannon gave the first of what would become the Shannon lectures in Askelon Israel. The symposium was superb in both its organization and technical content. A diverse set of stimulating and thought-provoking plenary talks by Rich Baraniuk, David Tse, Raymond Yeung, and Noga Alon set an excellent tone for each day's sessions. The technical program was outstanding with a broad range of sessions and that were packed, dynamic, and exciting. Jorma Rissanen gave a sweeping Shannon



Lecture on estimation, complete with significant new derivations, new interpretations of old results, and some provocative philosophical nuggets about modeling and statistics. The youthful exuberance, impressive scholarship, and intellectual curiousity Jorma conveyed in his lecture was an inspiration to all.

Every aspect of the symposium was run to perfection, from the warm welcome reception to the excellent symposium layout to the memorable banquet entertainment with dramatic drummers, risqué pantomines, and leaping dancers manipulating swooping long-tailed hats. Our deepest appreciation goes to general program chairs Jeong-Song No and Vince Poor as well as the technical program co-chairs Rob Calderbank, Habong Chung and Alon Orlitsky for all the hard work that went into creating such a memorable symposium. Whenever I return from a technical conference, it strikes me how lucky we are to travel to exotic places with friends from all over the world so we can spend a week immersed in stimulating intellectual endeavors. And we call this work and get paid for it. What a great profession we have!

One of the most rewarding aspects of the society presidency is hosting the Annual Awards Luncheon at ISIT, where we recognize the outstanding technical and service contributions of our members. These contributions are at the heart of our society: the technical contributions conveyed through our Transactions, conferences, and workshops and the dedicated volunteers that selflessly contribute their time, energy and ideas to manage our wide range of society activities. The committees that determine the Society award recipients are chaired by the officers: the President chairs the Shannon and Wyner Award Committees, the first VP chairs the Awards Committee which is responsible for paper awards, and the second VP chairs the Chapter Awards Committee. The paper and chapter award recipients are described later in this newsletter and details on all awards presented at the Awards Luncheon can be found on the society website. I will here describe the 2009 Shannon and Wyner award recipients I had the honor to present with their awards, as well as the recipients of IEEE awards and medals.

The Claude E. Shannon Award recognizes "consistent and profound contributions to the field of Information Theory". Jorma Rissanen was this year's recipient of our society's greatest honor. Jorma is best known for his introduction of the fundamental Minimum Description Length (MDL) Principle in 1978. This MDL principle is a formalization of Occam's Razor in which the best underlying

### Andrea Goldsmith

model associated with a given set of observed data is the one that leads to the largest compression of the data. MDL is particularly well-suited for highly complex models where overfitting the data is a serious concern. MDL has had significant impact on the fields of inductive inference, statistical modeling, pattern recognition, and machine learning, and has been widely applied to a broad cross-section of problems in engineering as well as biology and medicine. It continues to be a very active area of research more than 20 years after its invention, including recent work by Jorma himself on MDL denoising currently in press. Jorma is also well-known for his invention of arithmetic codes in 1975, a form of near-optimal lossless compression

based on variable-length entropy encoding of which Huffman encoding is a special case. Jorma's other two best known results are in data compression and coding: his 1983 invention of the algorithm Context for compression of data with Markov-chain properties, and his proof in 1986 that the minimum codeword length for any code in a given class is not the entropy, but rather the sum of the entropy and the code length needed to encode the best model for any model in the class. Jorma is a prolific author, with 150 published papers spanning the fields of information theory, data compression, coding, statistical inference, control theory, and complexity theory. The depth and breadth of these contributions is truly remarkable and on this basis I coined him a "Renaissance Scientist" in my Shannon Lecture introduction. Among Jorma's many awards, he received the IEEE 1993 Richard W. Hamming medal for fundamental contributions to information theory, statistical inference, control theory and the theory of complexity, an IEEE Information Theory Society Golden Jubilee Award for Technological Innovation as the inventor of arithmetic coding, two IBM Outstanding Innovation Awards, and the IBM Corporate Award. Jorma clearly embodies a researcher who has made profound and consistent contributions to information theory and related fields which have been of great importance in both theory and practice.

Our society could not function without the tremendous volunteer efforts of our members. The Aaron D. Wyner Distinguished Service Award recognizes an "individual who has shown outstanding leadership in, and provided long-standing exceptional service to, the information theory community." Richard E. Blahut is the 2009 recipient of this award. Dick's distinguished career has spanned over forty years, during which he has provided outstanding leadership and service to the information theory community. He has served the society in every possible role, as Society President and other officer roles, on the BoG, organizer for major symposia and workshops, and EiC of our Transactions. His many books have served to inspire generations of researchers in our field. Dick has also served as a role model and tireless mentor for many junior colleagues. Dick's service to the society has clearly impacted not just our activities, but also the individual success of many of our members.

We are a small society, only 1% of the IEEE, yet we have a very large impact. As one sign of this impact, our members are frequent recipients of the most prestigious awards and honors of the IEEE and other professional organizations. The IEEE awards and medals received this year by our society members were also

3

recognized at the Awards Luncheon. In particular, the IEEE Eric E. Sumner Award was presented to Roberto Padovani "For pioneering innovations in wireless communications, particularly to the evolution of CDMA for wireless broadband data." Roberto chose to have this prestigious award presented at ISIT, and also generously donated the award money to the society, which we have used to establish the Roberto Padovani Distinguished Lecture to be presented at our North American School of Information Theory. The first of these Distinguished Lectures will be given by Abbas El Gamal this summer. Other 2009 IEEE awards bestowed on our members include the Alexander Graham Bell Medal received by Robert J. McEliece, the Richard W. Hamming Medal received by Peter Franaszek, and the Donald G. Fink Prize Paper Award received by Daniel J. Costello and G. David Forney.

Perhaps the biggest highlight of ISIT, or at least the one generating the most excitement and suspense, is the announcement of the next year's Shannon Award winner. The award is decided by a committee of seven consisting of the society officers, former Shannon Award winners, and a current or former EiC of the Transactions. This year we initiated an open call for nominations in addition to internal ones, in order to cast a wide net for potential award winners. Our Shannon Award committee had the extraordinarily difficult job of selecting one winner among many deserving candidates, all of whom have made truly consistent and profound contributions to the field. Te Sun Han was selected as the winner of the 2010 Information Theory Society Claude E. Shannon award, and will receive his award and present the Shannon Lecture at next year's ISIT.

The Annual BoG meeting is also held in conjunction with ISIT. At this meeting the BoG voted to approve new appointed officers: Aria Nostratinia for Secretary starting immediately and Nihar Jindal for Treasurer starting in January 2010. The BoG welcomes Aria and Nihar and thanks João Barros and Anant Sahai for their tremendous time and dedication over the last three years in these respective roles. The BoG also approved the slates of candidate for the 2010 BoG proposed by the Nominations and Appointments Committee and added one more candidate to this slate. The candidates for next year's society officers were also decided. There was a vote on the IT Society paper award recipient, as well as some discussion on the roles of the Awards Committee and BoG in the deliberation process for this award. Votes to approve upcoming conferences, workshops, and student schools were taken, and reports from all society committees presented. Perhaps most significant of the committee report was that of the Publications Committee, which discussed the status of reducing sub-to-pub time, including increasing the number of Associate Editors and the transition for moving paper processing from Pareja to Manuscript Central. In addition, the Conference Committee Chair was made an ex-officio member of the BoG. The meeting closed with an open-ended discussion about society governance, in particular the composition of the BoG, the role of committees, and how to get BoG members more involved. Although most society decisions still come before the BoG, committees have taken over much of the detailed work leading up to such decisions. This is a natural evolution given the increased activity of our society, yet has raised some concerns about the role of the BoG. It was decided that to maintain strong BoG oversight in all our activities, it is desirable to get BoG members more involved in society committees, and a concerted effort to do so was the action item coming out of this discussion. Overall this BoG meeting indicated much progress in several of my presidential initiatives, including reducing our sub-to-pub time, utilizing our new website features to better communicate with society members, launching new initiatives such as our Distinguished Lecture Program, getting more members involved in new and ongoing initiatives, and improving society governance by formalizing procedures and improving ties between the BoG and society committees.

I am writing this article while on vacation in Spain, where late nights and unstructured days make the start of fall classes seem very remote. Yet this article will appear in that timeframe, when the new academic year begins, the weather begins to change, and we start to revise our plans for all the tasks we had hoped to accomplish in 2009. I wish everyone a very pleasant fall and look forward to seeing many of you at our last society meeting of the year, the Information Theory Workshop in Taormina Sicily. Our last BoG meeting of the year will be held there. We welcome participation by all members at this and other BoG meetings, and encourage you to share your ideas, thoughts or concerns in that setting or with any BoG members or officers individually.

# **Annual Information Theory Society Awards Announced**

Each year, the annual awards of the Information Theory Society are announced at ISIT.

The recipient of the 2010 Claude E. Shannon Award is Te Sun Han. The recipient of the 2009 Aaron D. Wyner Distinguished Service Award is Richard E. Blahut. Details for these awards are given in the President's column.

### Paper Awards

#### Frank R. Kschischang ITSoc. 1st VP and chair of the Awards Committee

The Information Theory Society Paper Award is given annually to the author(s) of an outstanding publication in the fields of interest to the Society appearing anywhere during the preceding two calendar years. The recipients of the 2009 award are V. Cadambe and S. A. Jafar, for their paper "Interference Alignment and the Degrees of Freedom for the K User Interference Channel", published in the IEEE Transactions on Information Theory, Aug. 2008. The basic premise of the ground-breaking result reported in this paper is that, in a network of  $K \ge 2$  distributed wireless nodes, each node can transmit information such that it has interference-free communication half of the time, with the other half of the time dominated by interference from other nodes. Thus, no matter how many nodes are in the wireless network, a given user gets "half the cake," achieving data rates up to half of what would be possible with no interfering nodes in the shared spectrum. Mention is also given to the paper by M. A. Maddah-Ali, A. S. Motahari and A. K. Khandani entitled "Communication Over MIMO X Channels: Interference Alignment, Decomposition, and Performance Analysis", IEEE Transactions on Information Theory, August 2008, in which the first explicit interference alignment scheme was presented.

The Joint Information Theory and Communications Society Paper Award is given annually to the author(s) of outstanding papers, relevant to both societies, published in any publication of the IEEE Communications Society or the IEEE Information Theory Society during the previous three calendar years. The recipients of the 2009 award are T. Ho, M. Médard, R. Kötter, D. R. Karger, M. Effros, J. Shi, and B. Leong, for their paper "A Random Linear Network Coding Approach to Multicast," which appeared in IEEE Transactions on Information Theory in October 2006. This frequently cited paper sets out a distributed random linear network coding approach for transmission and compression of information in general multi-source multicast networks. The authors prove that the performance of a network code induced by independent and local linear coding operations achieves, with high probability as the field size grows, the same performance as the best globally designed network code. This paper has had a large impact on applications of network coding: virtually all practical protocols that have been proposed (for example, for peer-to-peer systems or wireless networks) use the randomized network coding approach introduced in this paper.

The Information Theory Society Student Paper Award recognizes outstanding papers presented (by a student author) at the IEEE International Symposium on Information Theory. The award is given on the basis of both the technical content of the paper and the quality of the presentation, and the student(s) must be the primary contributor(s) to the paper. The 2009 recipients of this award are: Ali Nazari for his paper "New Bounds on the Maximal Error Exponent for Multiple-Access Channels" (co-authored with S. Pradhan, and A. Anastasopoulos), Changho Suh for his paper "Symmetric Feedback Capacity of the Gaussian Interference Channel to within One Bit" (co-authored with D. Tse), and Satish Babu Korada and Eren Sasoglu for their paper "Polar Codes: Characterization of Exponent, Bounds, and Constructions" (co-authored with R. Urbanke).

### Chapter of the Year Award

Giuseppe Caire ITSoc. 2nd VP and chair of the Membership & Chapters Committee

The IEEE IT Society Chapter of the Year Award annually recognizes a chapter that has provided their membership with the best overall set of programs and activities. The prize includes \$1,000, funded by the IT Society, to support local chapter activities.

Since 2009, the selection of the ITsoc. Chapter of the Year is one of the tasks under the responsibility of the Membership & Chapters Committee. After examining all Chapters' activities and after a thorough discussion among its members, the Membership & Chapters Committee reached to the unanimous decision of electing the Taipei Chapter as the recipient of the ITsoc. 2009 Chapter of the Year Award.

The award was delivered during the Award Luncheon at ISIT 2009, on Tuesday June 30, in Seoul Korea, to the Taipei Chapter Chair, Professor Li-Chun Wang of National Chiao Tung University, Hsinchu, Taiwan.

On behalf of the ITsoc. and of the Membership & Chapters Committee, I am happy to congratulate the Taipei Chapter for this nice achievement. Also, I would like to invite all ITsoc. chapters to make good use of the space on the ITsoc. web dedicated to the society chapters in order to post and maintain their contact information, and advertise their activities.

# In Memoriam, Wes Peterson

William Wesley Peterson, PhD, a mathematician, electrical engineer and computer scientist, was born April 22, 1924 in Muskegon, Michigan. He passed away quietly on May 6, 2009 in Honolulu, Hawaii after a five-day bout of pneumonia. He is survived by his wife Hiromi and five children.

Wes was the co-inventor of cyclic codes. He also devised the first algebraic algorithm for decoding the important class of BCH codes. His groundbreaking 1962 book "Error-Correcting Codes" basically started the field of errorcorrecting codes. The book was regarded as the 'bible' in the field of coding theory for many years. It was translated into Russian, French and other languages and is still generating royalties. It was widely used as a graduate text in coding theory and as a reference for practicing communication engineers around the world. It

helped train many top coding theorists in the early development of coding theory. The book also had a significant impact on the modern electrical engineering curriculum that requires either



Wes Peterson, 1924-2009

#### Ned Weldon and Shu Lin

linear or modern algebra as a part of mathematical training.

Dr. Peterson was a Professor of Information and Computer Sciences at the University of Hawaii at Manoa. He also did research and published in the fields of signal detection, programming languages, systems programming, and networks. He received the Japan Prize in 1999 for his pioneering contribution in error-correcting codes and higher education, the Claude E. Shannon Award in 1981, and the IEEE Centennial Medal in 1984.

Wes's early death was a tragedy. Despite his age, he was actively doing research and teaching demanding high-level courses until the end. He was physically vigorous; for the past 16 years he walked the Honolulu marathon, with improved age-group performance each year.

Wes was a humble, warm and generous person. Both of us will always consider ourselves truly fortunate to have known Wes as a friend, mentor and colleague.

# **Optimal Estimation** continued from page 1

of models includes many of the so-called nonparametric models such as histograms. In fact, if nonparametric models can be fitted to data at all they must be defined by an algorithm – or by parameters.

Corresponding to the two model classes we have sets  $\mathcal{F}_k$  and  $\mathcal{F}$  of *estimator* functions

$$\overline{\theta}(\cdot): \mathbf{x} \mapsto \overline{\theta}(\mathbf{x}), \quad \overline{\theta}(\cdot), \overline{k}(\cdot): \mathbf{x} \mapsto \overline{\theta}(\mathbf{x}), \overline{k}(\mathbf{x}).$$

The estimators are occasionally also denoted by  $\overline{\theta}$  and  $\overline{\theta}$ ,  $\overline{k}$  (not parameter values).

There is a fundamental complication in such model classes, namely, the parameter values  $\theta$ , k determine a model  $f(\mathbf{x}; \theta, k)$ , but the converse is not true, because only the non-zero parameters count, and both  $\theta_1, \ldots, \theta_k$  and  $\theta_1, \ldots, \theta_k, 0, \ldots, 0$  determine the same model. This implies that the likelihood is maximized by the largest number of parameters, which creates a huge problem for traditional estimation.

We conclude this section with an issue of some importance. Customary estimation is based on the assumption that the data were generated by a "true" distribution  $f(\mathbf{x}; \theta^*, k^*)$ , which lies either in the model class or outside. If indeed the data were generated by a single "true" distribution and since "truth" has no neighbors we can concentrate on estimating the single property defined by the parameter values  $\theta^*$ ,  $k^*$  without worrying about other nearby properties. The problem does not disappear if the "truth" is assumed to lie outside the model class, because we then have to find the closest property inside the model class – or ignore the model class entirely without any direction where to go!

Here we want a theory without the assumption of "truth" and instead are just trying to find statistical properties in data. We must take into account the fact that the data have many properties among which we just seek the best fitting ones. This is particularly true if the amount of data is not large to justify an asymptotic analysis. As a matter of fact, with the "truth" assumption we can find the "true" model only asymptotically, and the Central Limit Theorem implies that we can find it only if the "truth" is normally distributed.

### 3 Estimation and Measuring

The immediate problem arising from dropping the "true" model assumption is how to measure the goodness of an estimated model. Clearly, we cannot do it by any distance function between the "true" parameters and the estimated ones. Instead, I regard estimation as being analogous to measuring a physical property of an object such as its weight. As to the existence of a "true" weight, it seems to me that weight is defined within a physical theory, which certainly is man made, and there cannot be any "true" weight. However, all this is irrelevant just as the assumption of a "true" distribution as a model; our theory is simply richer than any with the "truth" assumption so that there is no harm in believing in it.

Here is the analogy:

IEEE Information Theory Society Newsletter

data **x** (object) yardstick (scale) is a special estimator  $\hat{\theta}(\cdot), \hat{k}(\cdot)$  and distribution  $\hat{p}(\hat{\theta}(\cdot), \hat{k}(\cdot))$  it defines property of data  $\hat{p}(\hat{\theta}(\mathbf{x}), \hat{k}(\mathbf{x}))$  (weight) equivalently:  $\log 1/\hat{p}(\hat{\theta}(\mathbf{x}), \hat{k}(\mathbf{x})) = \text{code length}$ large  $\hat{p}(\hat{\theta}(\mathbf{x}), \hat{k}(\mathbf{x})) - \text{good fit (accurate weight measurement)}$ small  $\hat{p}(\hat{\theta}(\mathbf{x}), \hat{k}(\mathbf{x})) - \text{bad fit}$ 

The selection of the yardstick must not depend on the data and it must be done from the model class.

### **4** Estimation Capacity

Before we can select the special estimator and yardstick  $\hat{p}(\hat{\theta}(\cdot), \hat{k}(\cdot))$  to provide an accurate estimation of the statistical properties in data we must describe how an estimator defines a probability distribution. We do it as follows: First, for the estimators of the real valued parameters put

$$P_{k}(\mathbf{x} | \overline{\theta}(\cdot)) = \overline{f}(\mathbf{x}; k) = f(\mathbf{x}; \overline{\theta}(\mathbf{x}), k) / \overline{C}_{k}$$
$$\overline{C}_{k} = \int f(\mathbf{x}; \overline{\theta}(\mathbf{x}), k) d\mathbf{x} < \infty$$

There is a way to handle the case where the normalizing integral is infinite, but we do not go into it here.

For the estimators  $\overline{\theta}(\cdot)$ ,  $\overline{k}(\cdot)$  the fact that only the non-zero real valued parameters count forces a somewhat intricate procedure for which we need the previous result  $\overline{f}(\mathbf{x}; k)$  for  $k = \overline{k}$  as follows

$$P(\mathbf{x} | \overline{\theta}(\cdot), \overline{k}(\cdot)) = \overline{f}(\mathbf{x}) = \overline{f}(\mathbf{x}; \overline{k}(\mathbf{x}))/\overline{C}$$
(1)

$$\overline{C} = \sum_{k} \int_{\overline{k}(\mathbf{x})=k} \overline{f}(\mathbf{x};k) d\mathbf{x}$$
(2)

Suggested by universal models define the *estimation capacities* log  $\hat{C}_k$  and log  $\hat{C}$ , all n, for the two model classes, respectively, where

$$\hat{C}_{k} = \max_{\overline{\theta}(\cdot)} \overline{C}_{k} = \max_{\overline{\theta}(\cdot)} \int f(\mathbf{x}; \overline{\theta}(\mathbf{x}), k) d\mathbf{x}$$
(3)

$$\hat{C} = \max_{\overline{\theta}(\cdot),\overline{k}(\cdot)} \overline{C} = \max_{\overline{\theta}(\cdot),\overline{k}(\cdot)} \sum_{k} \int_{\overline{k}(\mathbf{x})=k} f(\mathbf{x};\overline{\theta}(\mathbf{x}),k) d\mathbf{x}/\overline{C}_{k}$$
(4)

Notice that  $\hat{C}$  determines  $\hat{C}_k$  and the estimators, which in turn determine the complete codes as distributions, (1) and (2).

### 5 *MDL* Optimality

The code defined by the estimation capacity  $\log \hat{C}_k$  is the Normalized Maximum Likelihood (*NML*) code due to Shtarkov, [5]. He defined it by asking for a universal code with code length closest to the *ideal* code length  $\log 1/f(\mathbf{x}; \hat{\theta}(\mathbf{x}), k)$  as the solution  $\hat{f}(\mathbf{x}; k) = f(\mathbf{x}; \hat{\theta}(\mathbf{x}), k)/\hat{C}_k$  to the minmax problem

$$\min_{q} \max_{x} \log \frac{f(\mathbf{x}; \hat{\theta}(\mathbf{x}), k)}{q(\mathbf{x})} = \log \hat{C}_{k}.$$

September 2009

Fisher, [2], might have appreciated Shtarkov's intuition to find support for the excellence of the *ML* estimator, but actually this does not provide a formalized justification for it, because the minmax argument applies even to codes defined by estimators  $\overline{\theta}(\cdot) \neq \hat{\theta}(\cdot)$ , and their minmax value log  $\overline{C}_k$  is smaller than log  $\hat{C}_k$ . Why then prefer the *NML* code?

The real justification results from the following crucial fact: any reduction of the numerator of  $\hat{f}(\mathbf{x}; k)$  reduces also the denominator and gives a code  $\bar{f}(\mathbf{x}; k)$  which is worse than the *NML* code:

if 
$$\overline{\theta}(\mathbf{x}) \neq \hat{\theta}(\mathbf{x})$$
 then  $\overline{f}(\mathbf{x};k) < \hat{f}(\mathbf{x};k)$ ,

In fact, if  $f(\mathbf{x}; \overline{\theta}(\mathbf{x}), k) = f(\mathbf{x}; \hat{\theta}(\mathbf{x}), k) - \Delta$  then  $\overline{C}_k = \hat{C}_k - \Delta$ , and  $\overline{f}(\mathbf{x}; k) < \hat{f}(\mathbf{x}; k)$ . Rephrased, this becomes a sharp complete form of the *MDL* principle:

No string's codeword  $\log 1/\hat{f}(\mathbf{x};k)$  can be shortened by any estimated parameter  $\overline{\theta}(\mathbf{x})$ 

This means that given a code the code length is minimized for all strings, which is as strong as it can be since clearly no code exists which gives the shortest codeword for every string.

The same applies to the estimation capacity  $\log \hat{C}$  and the complete code it defines

$$\hat{f}(\mathbf{x}; \mathcal{M}) = \frac{\hat{f}(\mathbf{x}; \hat{k}(\mathbf{x}))}{\hat{C}} = \frac{\max_k f(\mathbf{x}; \hat{\theta}(\mathbf{x}), k) / \hat{C}_k}{\hat{C}}$$

Again any reduction of the numerator reduces  $\hat{f}(\mathbf{x}; \mathcal{M})$ 

if 
$$\overline{\theta}(\mathbf{x}), \overline{k}(\mathbf{x}) \neq \hat{\theta}(\mathbf{x}), \hat{k}(\mathbf{x})$$
 then  $\overline{f}(\mathbf{x}; \mathcal{M}) < \hat{f}(\mathbf{x}; \mathcal{M})$ 

Rephrasing gives the complete MDL principle:

No string's codeword  $\log 1/\hat{f}(\mathbf{x}; \mathcal{M})$  can be shortened by any choice of estimated parameters

### 6 Minmax Optimalities

To further justify the optimality of the *ML* estimator  $\hat{\theta}(\cdot)$  and the *MDL* estimator  $\hat{\theta}(\cdot)$ ,  $\hat{k}(\cdot)$  we give the solutions to two minmax problems

$$\min\max D(f_{\theta,k}\|\bar{f}_k) \tag{5}$$

$$\min_{\overline{\theta}(\cdot),\overline{k}(\cdot)} \max_{\theta} D(f_{\theta,k} \| \overline{f})$$
(6)

where we used the notations

$$f_{\theta,k} \text{ for } f_{\theta,k}(\mathbf{x}) = f(\mathbf{x}; \theta, k)$$
$$\bar{f}_k \text{ for } \bar{f}_k(\mathbf{x}) = f(\mathbf{x}; \overline{\theta}(\mathbf{x}), k) / \overline{C}_k$$
$$\bar{f} \text{ for } \bar{f}(\mathbf{x}; \mathcal{M}) = \frac{f(\mathbf{x}; \overline{\theta}(\mathbf{x}), \overline{k}(\mathbf{x})) / \overline{C}_{\overline{k}(\mathbf{x})}}{\overline{C}}.$$

These differ from the familiar minmax problem for universal data compression, whose solution is the Bayes mixture with prior that gives Shannon's channel capacity as the 'regret'.

**Theorem 1** For all k and n the solution to the minmax problem (5) is  $\hat{\theta}(\cdot)$  and  $\hat{f}_k$ . The minmax KL distance is  $\ln \hat{C}_k$ . Further, for all k and  $\theta$ 

$$D(f_{\theta,k} \| \hat{f}_k) = \ln \hat{C}_k - a_k(\theta)$$
$$a_k(\theta) = E_\theta \ln \frac{f(\mathbf{x}; \hat{\theta}(\mathbf{x}), k)}{f(\mathbf{x}; \theta, k)}.$$

**Theorem 2** For all k and n the solution to the minmax problem (6) is  $\hat{\theta}(\cdot), \hat{k}(\cdot)$  and  $\hat{f}$ . The minmax KL distance is  $\log \hat{C}_k + \log \hat{C}$ . Further, [3], for all  $\overline{\theta}(\cdot), \overline{k}(\cdot)$ 

$$D(f_{\theta,k} \| \bar{f})/n \ge D(f_{\theta,k} \| \hat{f})/n \to (k \ln n)/(2n)$$

*as*  $n \to \infty$ . The inequality holds for all k and  $\theta$  except some in a set  $A_{\theta,k}$  whose volume goes to zero as n grows.

The inequality in the theorem also gives the optimal convergence rate of estimators  $\bar{k}(\mathbf{x})$ .

### 7 Concluding Remarks

Dropping the "truth" assumption changes the perspective of estimation. We no longer can estimate the "true" parameters in isolation without regard to the other models, and we need to take into account all the models in the family, which define the *estimation capacity*. It generalizes to interval estimation and defines the estimated interval of optimal size. The center points, in turn, define the set of *optimally distinguishable* models, which form the relevant models for hypothesis testing and give an idea of the degree and confidence we can have not only in the selected winner but how many of the almost as good nearby models we miss.

### References

[1] Cramer, H. (1946), Mathematical methods of statistics, *Princeton University Press*, 575 pages.

[2] Fisher, R.A. (1925), 'Theory of statistical estimation', Proceedins of Cambridge Philosophical Society, **22**, page 700.

[3] Rissanen, J. (1984), 'Universal Coding, Information, Prediction, and Estimation', *IEEE Trans. Information Theory*, Vol. **IT-30**, Nr. 4, 629–636.

[4] Shannon, C.E. (1948), 'A Mathematical Theory of Communication', *Bell Syst. Tech. J.* **47**, 143–157.

[5] Shtarkov, Yu. M. (1987), "Universal Sequential Coding of Single Messages", Translated from Problems of Information Transmission, Vol. 23, No. 3, 3–7, July–September 1987.

# The Historian's Column

I read recently an interesting Shakespeare quote. The Bard said that "the man that hath no music in himself, is fit for treasons, stratagems, and spoils". This made me think whether Information Theorists have music in them or not. Of course, even if they do, they might still be fit for these despicable attributes. But, at least, they might not be.

Looking back over the years I recall that, in almost every IT gathering, there was some musical event. In the seventies, Bob Kennedy's (then) wife hit the piano one evening during the IT workshop in the Berkshires and displayed awesome talent for ragtime. At Allerton, it was routine to have one evening of music entertainment by talented members of the music department of the University of Illinois on Thursday evening at the Allterton Conference. At numerous ISITs, of ITWs there was an evening devoted to (usually) a chamber music recital or concert. In Sweden, Germany, Italy, Hungary, and almost every European country hosting an ISIT, there was an opportunity for musical entertainment woven into the program of the conference. So, there is an unquestionable record of music appreciation in our Society.

There has also been (on record) a consistent display of encouragement for a form of music that encompasses drama, comedy, and symbolism, namely: opera. On many (mostly forgettable) occasions, Sergio Verdu and I have fought our way through arias and duets in public display of disregard for tonality and phrasing. We exploited, that is, our Society's interest (or tolerance, some would say) for the art of singing. Let us not forget the renditions of "Nessum Dorma" that we have delivered in bellowing fashion after banquets or in the middle of the night. One beautiful evening during the NATO ASI in southern France, I was lucky to encounter a young Turkish participant (whose name, regrettably, I forget) who had a fine soprano voice and who joined me in rapturous duets from La Traviata, La Boheme, and other classics. Even Tom Cover had been drawn into the ecstasy of the singing by (almost) "conducting" us with his hands. The ample flow of "cheap red wine" during the evening was a catalyst for this memorable performance.

Sergio and I tore through the tranquility of Tuscan nights three years later during the last of the NATO ASI series that focused on our discipline. Under the cover of darkness we serenaded the quarters of our colleagues who at least did not show disapproval. And who can forget the "Torna a Sorrento" rendition by the five "tenured tenors" at the conclusion of the 1998 ISIT in Cambridge or the "Va Pensiero" choral passage at the 2000 ISIT in Italy?

The tradition continues unabated. Two prominent exponents of musical appreciation in our society are Ezio Beglreri and Joachim Hagenauer. Ezio has a preference for baroque music and is an avid admirer of Johan Sebastian Bach. Joachim is an all around music connoisseur with a deep knowledge of music and musical history. It has been a distinct pleasure to partake

#### Anthony Ephremides

of musical performances with them. With Ezio I enjoyed a fine Vivaldi concert in a Venice church, while I had numerous opportunities to join Joachim at musical evenings around the world. Most notable was the performance of Lohengrin at the Metropolitan Opera in the '90's after a CISSS conference in Princeton. Neither of us had tickets but we de-



cided to try our luck and meet outside the House to seek tickets. As it turned out, a March fierce snowstorm caused us to miss each other. However, independently outsmarting the scalpers, we both obtained tickets and made it into the theater where we enjoyed a grand performance of the Wagner classic.

But it has been with Sergio Verdu that I had the most frequent operatic experiences. We even dragged Ezio and Thomas Ericson once (after the Allerton Conference) into a performance of the Chicago Opera.

The tradition continues unabated. I did not mention the virtuoso playing of the harmonica by Toby Berger who graced many banquets and informal gatherings with his spirited and diversified repertoire. And most recently, at the San Diego ITAs, Giuseppe Caire and others have shown their talents and musical appreciation. How can I forget the evening after a CISS conference day (I believe it was in 2006) when Andrea Goldsmith drove us to New York City where we planned an ambitious evening of dining at Jean-Georges and then attending "Forza del Destino" at the Met. Traffic played havoc with our plans but we prevailed. As things looked grim for a moment before crossing the Lincoln Tunnel, I asked Andrea what her preference would be in case we got inordinately delayed, the dinner or the opera. Without hesitation she said, "the opera of course". To this day I am proud of her response. Nonetheless, we managed to squeeze both into the evening, even though I had to try hard to keep up with Andrea's determined stride and pace as we walked from the parking lot to the restaurant and then to the opera.

In short, there is mountains of evidence that Information Theorists have music in themselves. How could it be otherwise? What is music, if not an aesthetic abstraction of emotion and an expression of it in the universal language of sophisticated sound? But is Information Theory not, likewise, an aesthetic abstraction of creativity and an expression of it in the universal language of sophisticated reasoning and analysis? Aren't they both forms of elevated activity that is fulfilling, delightful, and deeply satisfying?

So, we could paraphrase the Bard and say "the man that hath no interest in Information Theory, is not fit for reason, strategy, and fun".

Thomas Kailath, Hitachi America Professor of Engineering, Emeritus, Department of Electrical Engineering at Stanford University has been elected to foreign membership in The Royal Society of the United Kingdom. He is one of only 8 scientists to receive this distinction this year. Prof. Kailath has gained worldwide recognition in several widely different fields - information and communication theory; linear systems, estimation and control; signal processing; linear algebra and matrix theory; statistics and probability. Examples include a striking capacity-achieving algorithm for channels with feedback; innovations processes and martingale theory; the structure of likelihood ratios for signal detection; high-resolution direction-finding algorithms; the concept of displacement structure and its many applications; and resolution-enhancement techniques in optical lithography. He has written classic textbooks, mentored a hundred doctoral and post-doctoral scholars, and made several successful transitions of his ideas through companies founded by, and co-founded with, his students. Prof. Kailath was the recipient of the 2000 Claude Shannon Award from the IEEE Information Theory Society, and the 2007 IEEE Medal of Honor for "exceptional development of powerful algorithms in the fields of communications, computing, control and signal processing." He served as President of the Society in 1975, and received the 1967 Outstanding Paper Award from the IT Society for his famous paper on Gaussian channels with feedback coauthored with P. Schalkwijk.

At the Awards ceremony held on June 15, 2009 at the IEEE International Communications Conference, Dresden, Germany, the IEEE Communications Society has awarded the following prizes:

- the Stephen O. Rice Prize in the Field of Communications Theory to Angel Lozano, Antonia Tulino, Sergio Verdu for their paper "Optimum Power Allocation for Multiuser OFDM with Arbitrary Signal Constellations," IEEE Transactions on Communications, Vol. 56, No. 5, pp. 828837, May 2008,
- the Leonard G. Abraham Prize in the Field of Communications Systems to Niranjay Ravindran, Nihar Jindal for their paper "Limited Feedback-Based Block Diagonalization for the MIMO Broadcast Channel," IEEE Journal on Selected Areas in Communications, Vol. 26, No. 8, pp. 14731482, October 2008,
- the William R. Bennett Prize in the Field of Communications Networking to Sachin Katti, Hariharan Rahul, Wenjun Hu, Dina Katabi, Muriel Medard, Jon Crowcroft for their paper "XORs in the Air: Practical Wireless Network Coding," IEEE/ACM Transactions on Networking, Vol. 16, No. 3, pp. 497510, June 2008.

# Symposium Report: The 2009 IEEE International Symposium on Information Theory, Seoul, Korea

The 2009 IEEE International Symposium on Information Theory (ISIT 2009) was held at the COEX conference center in Seoul, Korea, from June 28 through July 3, 2009. This is the fourth occasion on which ISIT has been held in Asia, with ISIT 1973 in Ashkelon, Israel, ISIT 1988 in Kobe, Japan, and ISIT 2003 in Yokohama, Japan. Thus, this is the first time that the ISIT has returned to the Asian mainland since 1973, a year in which Claude Shannon himself delivered the first Shannon Lecture.

Korea was a natural setting for the ISIT, as it has one of the world's most rapidly developing information technology industries, with a strong focus on mobile communications, semiconductors and displays. Along with the success of the Korean information technology industry, the corresponding academic community in Korea has grown rapidly. Moreover, Korea has a more than 5,000-year history, with a very diverse and colorful culture distinguishable from those of other Asian countries. Seoul has been the capital city of Korea for 600 years; it is a center of economy, culture, and education, with a population over ten million. Jong-Seon No and H. Vincent Poor General Co-chairs

The Symposium attracted 841 registrants, including 36 tutorialonly registrants and 392 student registrants (46%). The attendees came from 37 countries, with a geographical distribution as follows: United States (272), Korea (167), Japan (49), Switzerland (37), Canada (34), Israel (24), Germany (23), China (22), Sweden (20), India (19), France (19), Taiwan (18), Australia (16), Hong Kong (13), Singapore (12), Norway (11), Italy (9), Iran (9), etc.

Events began on Sunday, June 28, with four pre-Symposium tutorials, organized by Tutorials Co-Chairs Sae-Young Chung, Michael Gastpar, and Young-Han Kim. The tutorials were quite successful, attracting 375 registrants. The tutorial speakers and titles were:

- Emmanuel J. Candes (Caltech); Recovering the Unseen: Old and New
- Amin Shokrollahi (EPFL); Fountain Codes
- Balaji Prabhakar (Stanford University); Writing on a Clean Slate



At the Awards Luncheon: Tracey Ho, Muriel Medard & Michelle Effros (second from left to fourth from left) accept the Joint Communications and Information Theory Society Paper Award from Andrea Goldsmith (left), Frank R. Kschischang and Giuseppe Caire (on right).



Waiting for the banquet.

• David Tse (UC Berkeley); Interference Management; An Information Theoretic View.

After the tutorials, a welcome reception in the COEX Grand Ballroom opened the Symposium with 500 participants.

The TPC Co-chairs of the Symposium were Robert Calderbank of Princeton University, Habong Chung of Hongik University, and Alon Orlitsky of the University of California at San Diego. The TPC Co-chairs selected four outstanding plenary speakers and topics:

- Richard Baraniuk (Rice University); Randomized Dimensionality Reduction: A New Framework for Signal Processing and Communications
- David Tse (UC Berkeley); It's Easier to Approximate
- Raymond Yeung (Chinese University of Hong Kong); Facets of Entropy



At the Banquet: ISIT Co-chairs Jong-Seon No and H. Vincent Poor, and IT Society President Andrea Goldsmith (from left to right).

• Noga Alon (Tel Aviv University); Combinatorial Reasoning in Information Theory.

The Technical Program Committee conducted the review of 955 paper submissions from 47 countries, from which 589 papers from 41 countries were selected for presentation.

A highlight of the Symposium was the 2009 Claude E. Shannon Award Lecture, "Optimal Estimation," delivered on Thursday morning by Jorma Rissanen, Professor Emeritus of the Technical University of Tampere.

Under the leadership of Aylin Yener, the Student Committee of the IEEE Information Theory Society hosted two events at the Symposium: the first was a Lunchtime Panel Discussion and Meeting that took place on Monday, and the second was a Lunchtime Round Table Research Discussion that took place on Thursday.

A panel discussion on "Women in Information Theory: A Perspective from the Industry" took place on Wednesday. Panelists Andrea Goldsmith, Robert Calderbank, Soonja Choe, Sulyun Sung, and Christine Pepin shared their perspectives on this topic.

A Recent Results Session, organized this year by Habong Chung, was presented in the form of a poster session, which took place immediately following the plenary talk on Wednesday morning, in the lobby of the COEX Auditorium.

Attendees enjoyed several tours during the Symposium week: a night tour of Han River and market on Tuesday evening, half-day Seoul City tours including Gyeongbokgung Palace on Monday and Wednesday afternoons, and a full-day Demilitarized Zone (DMZ) tour on Wednesday.

The IEEE Information Theory Society Awards Luncheon took place on Tuesday, and the Conference Banquet took place on Thursday evening, both in the COEX Grand Ballroom. After dinner at the banquet, the recipient of the 2010 Shannon Award, Prof. Te Sun Han, was announced. The banquet was concluded with a Korean traditional performance. During the evening, ISIT 2010 TPC Cochair Michael Gastpar invited the attendees to ISIT 2010, which will be held in Austin, TX, during June 13–18, 2010.

In spite of the global economic crisis, generous financial support was received from several Korean government agencies, Samsung in Korea, and Hewlett-Packard in the US, which made it possible to support 78 student travel grants.



11

Shannon lecturer, Jorma Rissanen.

As the General Co-chairs of ISIT 2009, we would like to express our deep gratitude to the three TPC Co-chairs, the TPC members, the Organizing Committee members, and the professional conference organizers, JC International, for their efforts in making a very successful ISIT 2009. It was a wonderful opportunity for the exchange of new research ideas with colleagues, and to experience the unique Korean culture.

# Workshop Report: Networking and Information Theory

June 10-12, 2009, Volos, Greece

The workshop, held at the Xenia Hotel, kicked off with a reception on the evening of Tuesday, June 9, shortly after a bus with about 30 of the 96 participants arrived from the Athens airport.

The technical program featured three keynote talks. On Wednesday, John Tsitsiklis, presented the lecture, "Distributed consensus and averaging," which described provably slow and provably

#### Bruce Hajek and Leandros Tassiulas General Co-chairs

fast ways to compute averages of numbers in networks using distributed message passing algorithms. On Thursday, Don Towsley presented the lecture, "Towards a measurement science," which described several inference problems related to topologies of communication networks, and packet flows on high-speed links in the core of the Internet. Due to the large numbers of flows, many of which contain only a few packets, and



Sergio Verdu, Ramesh Johari, Costas Caramanis, Natalia Miliou and John Tsitsiklis among others in the gala dinner.



Attendees queue to be served while Prakash Narayan, John Baras and Don Towsley chat around the table.

September 2009

the fact that as few as one in 500 packets can be sampled, the inference problem may be ripe for attack by the method of patterns, discussed by Alon Orlitsky in one of the invited sessions at the conference. On Friday, Christos Papadimitriou presented the lecture, "Algorithmic game theory," which covered a variety of performance and complexity issues surrounding games. These include the social cost of selfish behavior, and the computational complexity of computing equilibria for various types of equilibria in games. Dr. Papadimitriou also discussed his recent work on explaining why sex is almost ubiquitous among species, in spite of its huge costs. Specifically, he argued that evolution of sexual species does not result in maximization of fitness, but in improvement of another important measure called mixability: the ability of a genetic variant to function adequately in the presence of a wide variety of genetic partners. The program also included nine plenary invited sessions of three papers each, covering a variety of topics in networks and information theory, and fifteen sessions (presented three in parallel at a time) of contributed papers. A total of 50 contributed papers were presented out of 101 submitted papers. Many thanks to the technical program committee and the organizers of the plenary invited sessions for an outstanding job of selecting the papers to be presented at the conference.

The workshop included a banquet at a seaside restaurant up the Pagasitic Gulf coast from Volos, featuring excellent food and live contemporary Greek music.

Based on observation and feedback, the organizers considered the workshop to be a great success.



# The Life Filled with Cognition and Action

(Dedicated to the 100th Anniversary of Academician V.A. Kotelnikov)

**Background:** Dr. Bykhovskiy is the Deputy Head of the Center of Analysis of Electromagnetic Compatibility at the Radio Research and Development Institute (Moscow), and a historian in the field of telecommunications. He is the author of a book in Russian entitled "Pioneers of the Information Century" and has previously contributed articles on David Middleton and Roland Dobrushin to the IT newsletter. This is a shortened version of the original submitted article.



Just as a horse is born to race, a bull-to plough the land, a dog-to trace, so a man is born for two matters-for cognition and action as a certain mortal God. Aristotle

### Introduction

The whole epoch of the development of communications, radio engineer-

ing and radio physics is linked with the name of Academician Vladimir Aleksandrovich Kotelnikov. His largest scientific achievements such as the invention of the sampling theorem, the creation of potential noise immunity theory, as well as the development of planetary radars, have all made a deep impact on scientific progress.

Kotelnikov was both an outstanding scientist and teacher. He greatly contributed to the organization of scientific studies in the USSR on many contemporary disciplines of radio engineering and radio physics. He established two large centers, the Special Design Bureau of the Moscow Power Engineering Institute (SDB MPEI) and the Institute of Radio Engineering and Electronics of the Russian Academy of Sciences (IRE RAS).

His life was filled with cognition and action completely corresponding to the concept about the mission of man formulated by Aristotle in the above epigraph.

### **Biographic Sketch**

Vladimir Aleksandrovich Kotelnikov was born on September 6, 1908 in Kazan. His father Alexander Petrovich Kotelnikov – a known Russian scientist in the field of mathematics and mechanics – was a Professor at Kazan University.

In 1930, V.A. Kotelnikov graduated from the radio engineering faculty of MPEI and began his postgraduate study at MPEI. In 1941, he became associate professor of MPEI and joined the Central Telecommunication Research and Development Institute (ZNIIS).

Fundamental radio communication problems at once fell within his view. In 1932, he prepared the report «On transmission capacity of an ether and a wire», where he formulated his Mark Bykhovskiy (bykhmark@gmail.com)

sampling theorem – one of the fundamental theorems of communication theory. The report was published in 1933.

In the 1930s, at the initiative of Kotelnikov and under his management, single side band radio equipment was developed and put into operation for the radio link between Moscow and Khabarovsk.

In 1938, Kotelnikov was granted his Master's degree in electrical engineering. In 1939, Kotelnikov started research and development of equipment for crypto protection of telegraph and telephone circuits. In early 1941 he created a prototype of the speech converter similar to the vocoder invented in 1939 by the American engineer H. Dudley, and in June–three days before the beginning of World War II (WWII)–he submitted a classified scientific report which for the first time defined conditions for undecipherability of a crypto system. He also defined technical principles for construction of cryptographically secure systems.

During WWII, under the guidance of Kotelnikov a ciphering system for telephone circuits was developed, which was not deciphered until 1946. It was widely used by the army to communicate between Moscow and the soviet delegation, and was used during the acceptance of the German capitulation in May, 1945. In 1943 and 1946, Kotelnikov and his collaborators were awarded Stalin's Prize of the first degree for the development of the voice ciphering equipment. Issues of cipher security were also studied by Claude Shannon, whose theoretical results were presented in the classified report «A Mathematical Theory of Cryptography» in 1946.

In 1944, Kotelnikov returned to MPEI where he established the Department of Theoretical Fundamentals of Radio Engeneering. Kotelnikov was the head of this department from 1946 until he was elected full member of the USSR Academy of Sciences in 1953, after which he still continued his pedagogical activity.

After the end of the WWII (December 1946), Kotelnikov submitted the doctoral thesis «A Theory of Potential Noise Immunity» to the MPEI academic council. It was successfully defended in January 1947 and became one of the milestones of modern communication science and brought to him world fame.

Kotelnikov also established one of the largest domestic science centres – Special Design Bureau (SDB MPEI). For a number of years he was the SDB chief designer and supervised many important developments. Scientists at SDB MPEI developed multi-channel radio telemeter systems for national rockets, space vehicles, active radar systems, high-precision goniometric systems, and the first space television system.

In 1953, by the initiative of Academician A.I. Berg, the Institute of Radio Engineering and Electronics of the USSR Academy of Science (IRE RAS) was established, with Berg as Director and

### 14

Kotelnikov as Deputy Director. In 1954, Kotelnikov became the Director of the IRE RAS. After some years this Institute became the largest scientific centre in the country on the problems of radio engineering and radio physics.

At the IRE RAS, Kotelnikov together with Academicians Guljaev (the IRE RAS Director after Kotelnikov) and Devyatkov devoted attention to studying the theoretical basis of microelectronics, optoelectronics, super-conductivity, semi-conductivity, acousto-electronics and magneto-electronics. IRE RAS scientists also carried out basic research in fiber optic communication systems, radio physical methods for exploration of Earth's natural resources, application of radio electronics in medicine, and automation of scientific studies.

Kotelnikov initiated and supervised studies on a planetary radiolocation and radiation study of Solar system planets and the Space. The complex radio system allowed unique scientific studies to be performed. In the early 1960s, Kotelnikov together with his students and colleagues developed a radio system enabling extremely accurate distance measurement in radar astronomy. It resulted in corrections to the astronomical constant (distance between Earth and Sun) and the dimensions of Solar system.

The fundamental research at IRE RAS led to the first mapping of the surface of Venus. In 1964, he together with his team of colleagues was awarded the Lenin's Prize for their work in the field of the planetary radiolocation. In 1984 – 1992, studies on the accuracy of the relativistic theory of planets movement were carried out under Kotelnikov's scientific guidance.

In 1987, Kotelnikov resigned from the position of the IRE RAS Director and became the Director Emeritus. Until his death, he remained as a Chairman of the IRE Academic Council and ran all its sessions.

Kotelnikov received many scientific awards. The International Institute of Electrical and Electronic Engineers (IEEE) awarded him the IEEE 1973 Award in International Communication in honor of Hernand and Sosthenes Behn for fundamental contributions to communication theory and practice, and for pioneering research and leadership in radar astronomy. Presidium of the USSR Academy of Sciences awarded Kotelnikov the 1974 Popov Gold Medal for contributions to basic research in the fields of communication theory and planetary radiolocation, the 1981 Lomonosov Gold Medal (the Academy's highest award) and the 1987 Keldysh Gold Medal.

For his rigorous pioneering proof of the famous sampling theorem Kotelnikov was awarded the international 1999 Edward Rhein Prize from Germany's Edward Rhein Foundation. For his contributions to communications he was awarded the IEEE 2000 Gold Medal of Alexander Graham Bell, and also the honorable IEEE Third Millennium Medal. Prof. Bruce Eisenstein, the President of the IEEE, characterized him as «The outstanding hero of the present. His merits are recognized all over the world. In front of us is the giant of radio engineering thought, who has made the most significant contribution to media communication development».

Kotelnikov was one of the founders of the Russian Scientific and Technical Society of Radio Engineering, Electronics and Communication named after A.S. Popov, and was the chairman of its Board for many years. Kotelnikov was also a member of the Polish, Czechoslovak, Mongolian, Bulgarian and German Academies of Sciences, IEEE Honorable member, and vice-president of the International Academy of Astronautics. From 1969 till 1988 Kotelnikov was vice-president of the USSR Academy of Sciences, and chaired a number of its councils. For many years Kotelnikov was the editor-in-chief of journals «The Radio Engineering and Electronics» and «The Bulletin of the USSR Academy of Sciences».

Vladimir Aleksandrovich Kotelnikov died on February 11, 2005.

### His Scientific Contribution to Communications Theory

The sampling theorem was published by Kotelnikov in 1933. According to this theorem any function with a spectrum limited by frequency F may be represented by the samples taken at 1/2F time intervals. Independently, this theorem was invented by English mathematicians E.T. Whittaker and J.M. Whittaker at the beginning of 20th century, and discovered by Claude Shannon in 1949.

Another of Kotelnikov's largest scientific achievements was developing the theory of potential noise immunity. It enabled engineers to theoretically synthesize optimum signal processing and receiving devices for different environments and to determine their feasible qualitative characteristics. American and English scientists A.J. Siegert, D. Middleton, P.M. Woodward and I.L. Davis also independently and greatly contributed in the development and popularization of this theory.

In his PhD thesis, he solved the problem of synthesis of the optimum receiver (i.e. the optimum algorithm for processing of the received signal) given the waveform of the transmitted signal and statistical noise characteristics, and developed the methodology for determination of the optimum receiver noise immunity. His small book (only 150 pages) contained many deep ideas which led to a new approach to building signal receivers.

In 1959 Kotelnikov's translated book was published in the USA. In a book review Prof. N.M. Abramson from Stanford University highlighted Kotelnikov's priority in the theory of optimum signal reception, in the application of multidimensional geometry for interpretation of problems associated with signal reception against the background of the noise, and also in studies of nonlinear filtering of signals against the background of nonuniform noise.

### Conclusion

According to A.Einstein the most objective criterion for the validity of a scientific theory is its «inner perfection and external justification». The aesthetic criterion is the deeper and more complex criterion for estimation of scientific achievements from the «inner perfection» point of view. It is based on an intuitive feeling of beauty.

Kotelnikov's theory of potential noise immunity had «inner perfection» in that sense. This theory has also been further developed by many researchers and applied to many practical problems. Thus, it has also received external justification.

But one more aspect should not be overlooked when considering scientific achievements. It is necessary to remember their creators'

huge amount of selfless work, the work that fills their life. German philosopher F. Nietzsche wrote: *«Each serious work has a moral impact upon us. Our effort done to concentrate the attention on a given theme could be compared to a stone thrown into our inner life: the first circle has the insignificant area, than the number of alternating circles is increasing and their area expanding».* 

The life of the outstanding Russian scientist V.A. Kotelnikov is the bright acknowledgement of this important and deep thought.

1. Mark Bykhovskiy, Pioneers of the Informational Era/History of Communication Theory Development, Technosphera, Moscow, 2006.



C. 1. To show  $p_k(n) = p^k(n)$  we count the partitions in two ways. For example, with n = 5 and  $k = 3, p_3(5) = p^3(5) = 2$ , where  $p_3(5)$  counts 3 + 1 + 1 and 2 + 2 + 1 (the partitions of 5 into exactly 3 parts), while  $p^3(5)$  counts 3 + 2 and 3 + 1 + 1 (the partitions of 5 where the largest part is 3. Here is the correspondence:

	3	+	2			3	+	1	+	1
2	٠		•	and	3	•		٠		٠
+2	•		•	anu	+1	•				
+1	•				+1	•				

One counts horizontally, the other vertically.

2. The same method shows  $p_k(n) = p^{\bar{k}}(n)$ . Again, with n = 5 and k = 3, we now have  $p_{\bar{3}}(5) = p^{\bar{3}}(5) = 5$ , shown as follows:



3. Partition equalities are also proved by "generating functions". A useful viewpoint considers restricted partitions of n as the number of possible ways to make change for  $n \not c$  using a specific set of coins. For example, suppose these are coins for  $2^j$  cents for each  $j \ge 0$ , and you must make change for  $n \not c$  using no more than one of each of these types of coins. Then with c(n) as the number of ways to make change for  $n \not c$ , the generating function is

$$\sum_{n=0}^{\infty} c(n)x^n = (1+x^1)(1+x^2)(1+x^4)(1+x^8)\dots = \prod_{j=0}^{\infty} (1+x^{2^j}).$$

Since we know that every positive integer has a unique binary representation, this says c(n) = 1 for all  $n \ge 0$ , from which

$$\prod_{j=0}^{\infty} (1+x^{2^j}) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

We can verify this by multiplying each side by 1 - x, to get  $1 = (1 - x) \prod_{j=0}^{\infty} (1 + x^{2^j})$ . The right side is a "telescoping product", because  $(1 - x)(1 + x) = 1 - x^2$ ,  $(1 - x^2)(1 + x^2) = 1 - x^4$ ,  $(1 - x^4)(1 + x^4) = (1 - x^8)$ , etc., and in the limit it equals 1. For making change into *unequal* coins, with all denominations available, the generating function

is  $\infty \infty$ 

$$(1+x)(1+x^2)(1+x^3)(1+x^4)(1+x^5)\cdots = \prod_{j=1}^{\infty} (1+x^j) = \sum_{n=0}^{\infty} c_1(n)x^n.$$

16

For making change when all *odd* values (and no even values) of coins are available, in unlimited quantities, the generating function is

$$(1+x+x^2+x^3+\cdots)(1+x^3+x^6+x^9+\cdots)(1+x^5+x^{10}+x^{15}+\cdots)\cdots$$
$$=\frac{1}{1-x}\cdot\frac{1}{1-x^3}\cdot\frac{1}{1-x^5}\cdot\frac{1}{1-x^7}\cdots=\prod_{j=1}^{\infty}\frac{1}{1-x^{2j-1}}=\sum_{n=0}^{\infty}c_2(n)x^n.$$

We want to show that  $c_1(n) = c_2(n)$  for all  $n \ge 0$ , which follows if the two generating functions are equal.

This follows from  $\prod_{j=1}^{\infty} (1-x^{2j-1}) \cdot \prod_{j=1}^{\infty} (1+x^j) = 1.$ 

(Here each factor of the form  $(1 - x^{2j-1})$  forms a telescoping product with  $(1 + x^{2j-1})(1 + x^{2(2j-1)})(1 + x^{4(2j-1)})(1 - x^{8(2j-1)})\cdots)$ 

The reader is invited to verify the details.

<u>Notes</u>. L. Euler (1707-1783) pioneered the study of partitions, introducing both the horizontal vs. vertical count and the generating function approach. In the 20th century, G. Pólya popularized the "making change" viewpoint.

The function  $p'_3(n)$  was recently called to my attention by Håkan Lennerstad of the Blekinge Institute of Technology in Karlskrona, Sweden. Problem B.6 was his observation. The exact expressions, and the observation in B.7, are my own, though I expect that the values of  $p_2(n)$ and  $p_3(n)$  were already known. The problems in Part A are classical results in combinatorial analysis.



#### Symposium Committee

Honorary Chairs Jin-Fu Chang (Executive Yuan) Shu Lin (U. of California, Davis )

General Co-Chairs Chi-chao Chao (Nat. Tsing Hua U.) Ryuji Kohno (Yokohama Nat. U. )

General Secretaries Li-Chun Wang (Nat. Chiao Tung U.) Makoto Itami (Tokyo U. of Science) Fumie Ono (Yokohama Nat. U.)

Finance Co-Chairs Hsiao-feng (Francis) Lu (Nat. Chiao Tung U.)

Masayoshi Ohashi (ATR Labs.)

Publication Co-Chairs Chung-Hsuan Wang (Nat. Chiao Tung U.) Takaya Yamazato (Nagoya U.)

Registration Co-Chairs Min-Kuan Chang (Nat. Chung Hsing U.) Toyoo Takata (Iwate Prefectural U.)

Local Arrangements Co-Chairs Houshou Chen (Nat. Chung Hsing U.) Huan-Bang Li (Nat. Inst. of Information and Commun. Tech.)

Publicity Co-Chairs Y.-W. Peter Hong (Nat. Tsing Hua U.) Shinichi Kawamura (Toshiba) Shinsuke Hara (Osaka City U.)

#### **Technical Program Committee**

Co-Chairs Mao-Chao Lin (Nat. Taiwan U.) Hideki Ochiai (Yokohama Nat. U.) Tetsushi Ikegami (Meiji U.)

#### International Advisory Committee

Co-Chairs Char-Dir Chung (Nat. Taiwan U.) Po-Ning Chen (Nat. Chiao Tung U.) Ikuo Oka (Osaka City U.)

#### Sponsor

Society of Information Theory and its Applications (SITA)

#### **Technical Co-Sponsors**

IEEE Information Theory Society IEICE Engineering Sciences Society



2010 International Symposium on Information Theory and its Applications

2010 International Symposium on Spread Spectrum Techniques and Applications

> The Splendor Hotel, Taichung, Taiwan October 17-20, 2010



The 2010 International Symposium on Information Theory and its Applications (ISITA2010) and the 2010 International Symposium on Spread Spectrum Techniques and Applications (ISSSTA2010) will be jointly held in Taichung, Taiwan, from Sunday, October 17, through Wednesday, October 20, 2010.

Topics of interest include, but are not limited to:

Error Control Coding	Mobile Communications
Coding Theory and Practice	Spread Spectrum Systems
Coded Modulation	Detection and Estimation
Data Compression and Source Coding	Signal Processing
Pattern Recognition and Learning	Sequence Design and Optimization
Speech/Image Coding	Synchronization and Channel Estimation
Rate-Distortion Theory	Multicarrier CDMA and OFDM
Shannon Theory	UWB Communications
Stochastic Processes	MIMO & Diversity Techniques
Cryptology and Data Security	Wideband Channel Modeling
Data Networks	Cognitive and Software Radios
Multi-Ilser Information Theory	Coonerative Communications
Multi-User Information Theory	Cooperative Communications
Quantum Information Processing	Optical Communications

Papers will be selected on the basis of a full manuscript (not exceeding 6 pages). The deadline for submission is 23:59, March 7, 2010, GMT. Notification of decisions will be made by the end of May, 2010.

The accepted papers will appear in the Proceedings and IEEE Xplore. Detailed information on the technical program, special events, accommodation, and registration will be posted on the Symposium web site:

#### http://www.sita.gr.jp/ISITA2010/

Enquiries on matters related to the Symposium should be addressed to:

General matters: isita-isssta2010@sita.gr.jp	Technical program matters: a.gr.jp isita-isssta2010tpc@ee.nthu.edu.tw	
Deadline for paper submission	n	March 7, 2010
Notification of paper acceptar	ice	late May, 2010
Deadline for final paper subm	ission	July 7, 2010
Deadline for author registration	on	July 7, 2010

About Taichung: Taichung is located in the hub of central Taiwan and surrounded by many tourist attractions, such as Sun Moon Lake, Lushan Hot Spring and Shitou Forest Recreation Area. The pleasant climate, beautiful scenery, the rich history of art and culture, and the prosperous economy supply Taichung with the glamour of a world-class metropolis. Walking in the streets of Taichung, a visitor would discover a delicate mixture of modern and classic, of nature and technology, and of grandeur and simplicity.

# **Call for Papers**

#### **TPC CHAIRS**

Michael Gastpar Robert Heath Krishna Narayanan

#### **TPC MEMBERS**

Jeff Andrews

Alexei Ashikhmin Alexander Barg Andrew Barron Seraio Benedetto Joseph Boutros Ning Cai Giuseppe Caire **Constantine Caramanis** Jean-Francois Chamberland Sae-Young Chung Todd Coleman Suhas Diggavi Stark Draper Ilva Dumer Michelle Effros Yonina Eldar Hesham El Gamal Uri Erez Elza Erkip Christina Fragouli Andrea Goldsmith Alex Grant Dongning Guo Bruce Hajek Tor Helleseth Tracey Ho Hideki Imai Tara Javidi Gerhard Kramer Ioannis Kontoyiannis Frank Kschischang Sanieev Kulkarni P. Vijay Kumar Amos Lapidoth Yingbin Liang Mokshay Madiman Muriel Medard Olgica Milenkovic Urbashi Mitra Pierre Moulin Aris L. Moustakas Prakash Naravan Aria Nosratinia Li Ping Sundar Rajan Kannan Ramchandran Ashutosh Sabharwal Igal Sason Serap Savari Anna Scaglione Shlomo Shamai Emina Soljanin Emre Telatar Joel Tropp Daniela Tuninetti Sennur Ulukus Rudiger Urbanke Sergio Verdu Pramod Viswanath Sriram Vishwanath Emanuele Viterbo Pascal Vontobel Aaron Wagner Martin Wainwright Frans Willems Jack Wolf Avlin Yener Ram Zamir Lizhong Zheng







## 2010 IEEE International Symposium on Information Theory

The 2010 IEEE International Symposium on Information Theory (ISIT 2010) will take place June 12 – 18, 2010 in Austin, Texas, the Live Music Capital of the World. Previously unpublished contributions from a broad range of topics in information theory are solicited, including (but not limited to) the following areas:

Channel coding theory and practice	Multi-terminal information theory
Communication theory	Pattern recognition and learning
Cryptography and data security	Quantum information theory
Detection and estimation	Sequences and complexity
Emerging applications of information theory	Shannon theory
Information theory and statistics	Signal processing
Information theory for networks	Source coding theory and practice

Researchers working in emerging fields of information theory especially are encouraged to submit original findings. Submitted papers should be of sufficient length and detail for review by experts in the field. Final papers will be limited to 5 pages in standard IEEE conference format. Detailed information on paper submission, technical program, accommodations, tutorials, travel and excursions will be posted on the symposium website: http://www.isit2010.org/

#### **Key Dates:**

Treasurer:

Tutorials:

Publicity:

Paper Submission Deadline: Acceptance Notification:

January 7, 2010, 11:59pm, US Central Time April 1, 2010

For more information about the symposium, please contact the General Co-Chairs, Behnaam Aazhang and Costas N. Georghiades.

Behnaam Aazhang Costas N. Georghiades ECE Department ECE Department Texas A&M University, College Station, TX Rice University, Houston, TX Tel. +1-713.348.4020 Tel. +1-979.845.7408 email: aaz@rice.edu email: georghiades@tamu.edu

**Conference Executive Committee** 

Aria Nosratinia Local Arrangements: Jeff Andrews, Sriram Viswananth Zixiang Xiong Jean-Francois Chamberland, Ashutosh Sabharwal **Publications:** Dinesh Rajan Recent Results: Serap Savari, Sanjay Shakkottai Student Travel Grants: Joseph Cavallaro







IEEE Information Theory Society Newsletter

### 20

# Conference Calendar -

DATE	CONFERENCE	LOCATION	WEB PAGE	DUE DATE	
Sept. 20-23 2009	2009 IEEE 70th Vehicular Technology Conference (VTC2009-Fall)	Anchorage, Alaska, USA	http://www.ieeevtc.org/ vtc2009fall/index.php	Passed	
Sept. 30–Oct. 2, 2009	47th Annual Allerton Conference on Communication, Control and Computing (Allerton 2009)	Monticello, IL, USA	http://www.csl.uiuc.edu/ allerton/	Passed	
Oct. 11–16, 2009	2009 IEEE Information Theory Workshop (ITW 2009)	Taormina, Italy	http://www.deis.unical.it/ itw2009	Passed	
Oct. 18–21, 2009	23rd IEEE Computer Communications Workshop (CCW 2009)	Lenox, Massachusetts	http://www.comsoc.org/~tccc/ ccw/2009/	Passed	
Nov. 1–4, 2009	Asilomar Conference on Signals, Systems, and Computers	Pacific Grove, CA	http://www.asilomarssc.org/	Passed	
Nov. 30– Dec. 4, 2009	2009 IEEE Global Communications Conference (GLOBECOM 2009)	Honolulu, Hawaii, USA	http://www.ieee-globecom.org/	Passed	
Dec. 14– 17, 2009	Twelfth IMA International Conference on Cryptography and Coding	Cirencester, UK	http://www.ii.uib.no/~matthew/ Cirencester09/Cirencester09.html	Passed	
January 6–8, 2010	2010 IEEE Information Theory Workshop (ITW 2010)	Cairo, Egypt	http://itw2010cairo.info/	August 16, 2009	
January 18–21, 2010	8th International ITG Conference on Source and Channel Coding (SCC'10)	Siegen, Germany	www.scc2010.net	September 8, 2009	
March 3–5, 2010	2010 International Zurich Seminar on Communications	Zurich, Switzerland	http://www.izs.ethz.ch/	September 27, 2009	
March 15–19, 2010	IEEE INFOCOM 2010	San Diego, CA	http://www.ieee-infocom.org/	July 24, 2009	
May 23–27, 2010	IEEE International Conference on Communications (ICC 2010)	Cape Town, South Africa	http://www.ieee-icc.org/2010/	September 10, 2009	
June 12–18, 2010	IEEE International Symposium on Information Theory	Austin, Texas	http://www.isit2010.org/	January 7, 2010	
Oct 17–20, 2010	2010 International Symposium on Information Theory and its Applications and 2010 International Symposium on Spread Spectrum Techniques and Applications	Taichung, Taiwan	http://www.sita.gr.jp/ISITA2010	March 7, 2010	

Major COMSOC conferences: http://www.comsoc.org/confs/index.html