IEEE Information Theory Society Newsletter

Vol. 61, No. 2, June 2011 Editor: Tracey Ho ISSN 1045-2362

Editorial committee: Helmut Bölcskei, Giuseppe Caire, Meir Feder, Joerg Kliewer, Anand Sarwate, and Andy Singer

Musing upon Information Theory

XXXI Shannon lecture, presented in 2010 in Austin, TX (extended abstract)

Te Sun Han

1. Prologue

Information Theory can be primarily regarded as a discipline that tries to link two different kinds of quantities: one is *operational quantities* defined by using operational concepts such as source, channel, capacity, encoder, decoder, codeword length, compression rate, transmission rate, probability of error (or convergence rate of error probability) and so on; whereas the other is *information-theoretic quantities* such as entropy, divergence, mutual information, and so on.

The theoretical and mathematical core of *information theory* is currently called the *Shannon Theory,* which was initiated in 1948 by Shannon [1]; it is six decades ago. In this connection, for instance, Gray and Ornstein [2, p. 294] argues that "A principal goal of the Shannon Theory is to prove coding theorems relating such operational capacities to information-theoretic extremum problems; that is, to quantities involving extreme values over all probabilistic functions or, equivalently, extremizing some functional of probability measures over an appropriately constrained space."

The advancement of information theory, since its birth, is quite enormous and it seems now that numerous soaring high tops of the information theory mountains have been tried or actually reached, though, along with a huge amount of technical and sophisticated results reported in various styles and viewpoints, which are not quite easy to follow up for students and/or sometimes even for researchers of information theory.

On the other hand, in parallel with the advancement during this period, many excellent textbooks on information theory have been published. One of the fundamental technical tools here is the *asymptotic equipartition property* (= AEP) or equivalently the *notion of typical sequences*, which amazingly and effectively works on the basis of the weak law of large numbers and/or the large deviation principle. Nowadays these textbooks are regarded as providing the *established standard basis* of information theory, in which we can find many simple and very elegant formulas, e.g., those for source coding, channel coding, rate-distortion function, hypothesis testing, etc.

However, most of these textbooks assume that both of sources $\mathbf{X}$ and channels $\mathbf{W}$ are *stationary* and *memoryless*, or *stationary* and *ergodic*, typically with *finite* alphabets. The above simplicity and elegance mainly come from these assumptions, while few of the standard textbooks mention, rather as side or isolated issues, several special classes of sources and channels that are *nonstationary* and/or *nonergodic*.*

Thus, we can say that information theory was born and grown up in the world of *stationarity* and *ergodicity*. Over thirty years I myself have also been fostered inside this comfortable world. However, after that, when grown up enough, my strong curiosity addressed the world outside, say, the whole world of *nonstationarity and/or nonergodicity*. By the word "whole" I intend hopefully to include, possibly under quite few assumptions, almost all general sources and almost all general channels with any kind of nonstationarity, any kind of nonergodicity, and any kind of alphabets (finite, infinite or abstract). The basic questions here are:

1) What is happening at all in the world of *nonstationarity* and/or *nonergodicity*?

2) Is it a mere extrapolation of the world of *stationarity* and *ergodicity*?

3) Is it far beyond the world of *stationarity* and *ergodicity*?

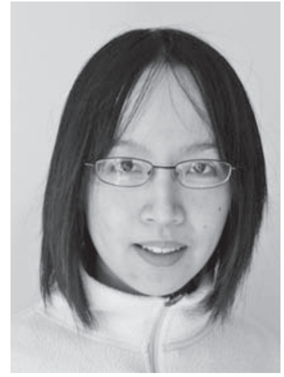
4) Is it possible at all to dig out "a" kind of general Shannon theory to be latent in the world of *nonstationarity* and/or *nonergodicity*?

*For example, Gray [3] analyzes the details of "asymptotically mean stationary" sources (AMS), which are in general neither stationary nor ergodic.

continued on page 4

From the Editor

Tracey Ho



Dear IT Society members,

In this issue I am very pleased to have the summary by Te Sun Han of his 2010 Shannon lecture "Musing upon Information Theory". I am also happy to report the news of Vince Poor's election to the National Academy of Sciences—warmest congratulations! Also in this issue, we have the report on the Winter School of Information Theory in Barcelona, and we note sadly the passing of Igor Vajda and commemorate his life and work.

As a reminder, announcements, news and events intended for both the printed newsletter and the website, such as award announcements, calls for nominations and upcoming conferences, can be submitted jointly at the IT Society website http://www.itsoc.org/, using the quick links "Share News" and "Announce an Event". Articles and columns that do not fall into the above categories should be e-mailed to me at tho@caltech.edu, with a subject line that includes the words "IT newsletter". The deadlines for the next few issues are:

Issue	Deadline
September 2011	July 10, 2011
December 2011	October 10, 2011
March 2012	January 10, 2012

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

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.

Table of Contents

President's Column

Giuseppe Caire

Dear IT Society members, when I was elected IT Society president I had only a vague idea of what it really meant. Now, after a few months, a Board of Governors meeting under my belt and several committees well on their way, I start realizing what I got myself into. The first semi-serious note goes to the Board of Governors. Those who attended the meeting at ITA, in La Jolla CA, February 2011, could appreciate the wonderful venue (many thanks to the local organizers) and the very significant delay accumulated by the meeting. Clearly, I have to learn how to keep meetings on track, following in the footsteps of my predecessors. The second more serious remark is about the following curious fact. Since I became president, I have been asked more and more often the following question: "… so, what is the future of Information Theory?". This question came up on several occasions and from a variety of people, both in industry and from non-engineering fields in academia. Of course, the mere fact that I am the IT Society president does not make me smarter or wiser (in fact, perhaps the influence is in the opposite sense). Therefore, I am not better entitled than any other IT Society member to try to provide an answer. Nevertheless, this is a question that we, as a scientific community, should confront every day.

While trying to provide a single answer to this question may seem overly ambitious, I would like to encourage some discussion about how to answer this question. As some eminent member of our society would immediately suggest, I went back to the "classical literature", and in particular to two well-known papers: Claude E. Shannon's "The Bandwagon" and Norbert Wiener's "What is Information Theory?", both of which appeared in the *IRE Trans. on Inform. Theory*, in 1956.

While Shannon's word of caution about not overselling Information Theory as a panacea applicable to any scientific domain is still very relevant today, I'd like to quote his optimistic view about the power and potential of information theoretic concepts beyond the original problems of communication and storage of information:

"I personally believe that many of the concepts of information theory will prove useful in these other fields-and, indeed, some results are already quite promising-but the establishing of such applications is not a trivial matter of translating words to a new domain, but rather the slow tedious process of hypothesis and experimental verification. If, for example, the human being acts in some situations like an ideal decoder, this is an experimental and not a mathematical fact, and as such must be tested under a wide variety of experimental situations."

I find particularly amusing the term "tedious", used here with a slight sense of humor: the "slow and tedious process of hypotheses and experimental verification" is indeed an exciting and fascinating process of developing new models and conquering new ground, with the highest scientific rigor. After all, Shannon reminds us that everything boils down to the uncompromising scientific standard. In his words: *"Authors should submit only their best efforts, and these only after careful criticism by themselves and their colleagues."*

On the other hand, Wiener has clearly a broader (and perhaps less precise) view of what Information Theory is. His viewpoint seems to appeal to what we call today an "information theoretic approach". Roughly speaking, this approach consists of turning very hard combinatorial problems in finite dimension into clean and strikingly simple answers in the appropriate asymptotic regime. Wiener already points out connections with *Statistical Mechanics,* which have found wide application in several problems, from graph-based codes and iterative message-passing decoding, to multiuser/multiantenna wireless communication networks, to the analysis of sparse signal reconstruction algorithms, much more recently. Although classifying Information Theory as a mere application of Statistical Mechanics and Large Deviation is reductionist, exploring the formal connections with other areas and therefore cross-fertilizing the different tool-boxes for solving relevant problems has proven very useful. From Wiener's paper I'd like to quote the following paragraph:

"In my opinion we are in a dangerous age of overspecialization. To me the danger of this period is not primarily that we are studying very special problems that the development of science has forced us to go into, but rather that we are in great danger of finding our outlook so limited that we may fail to see the bearing of important ideas because they have been formulated in what our organization of science has decreed to be alien territory."

I believe that a point of reflection emerges from Shannon's and Wiener's paragraphs quoted above: as a scientific community, we should not become too much entrenched into our own "pet problems". Let's think out of the box, and look around beyond the traditional boundaries! At the same time, though, connections to other fields should be made with the highest scientific scrutiny in mind, taking into account that the standard assumptions and models of Information Theory may have to be experimentally verified and possibly rediscussed.

Overall, the future of Information Theory looks very promising, especially when considering the state of our Society, healthy, alive and dynamic as ever. The year 2011 is characterized by a large number of conferences sponsored (or co-sponsored) by our society across the World. Among these, I'd like to point out WiOpt (9th Intl. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks), Princeton NJ, May 9–13, the First Workshop on Network Coding and Data Storage, Hong Kong, July 21–22, the 2011 International Symposium on Network Coding (NetCod 2011), Beijing, July 25–27, ISIT 2011, St. Petersburg, Russia, July 31–August 5, and ITW 2011, Paraty Brazil, October 16–20. Furthermore, the pioneering action of the IT society in sponsoring and fostering Information Theory schools, dedicated explicitly to graduate students and young researchers, is paying off. In 2011, the following events took or will take place: the 2011 Taiwan Winter School on Information Theory and Communications, January 17, Hsinchu, Taiwan, the 11-th Winter School of Information Theory, March 14-18, Barcelona, Spain, the 2011 North American School of Information Theory, May 27–30, Austin TX, and the Wireless Information Theory Summer School, July 27–29, Oulu, Finland.

This intense technical activity, and remarkable success of the Information Theory schools as meeting points where the future generations of researchers in our field can meet, share their interests, compare their results and, most importantly, build a community, make me believe that the future of Information Theory is as bright as ever. As long as we are able to foster new generations of researchers who, by their very nature, will explore wider areas and look farther ahead, the tree planted by Shannon 63 years ago will bear fruit.

Musing upon Information Theory

continued from page 1

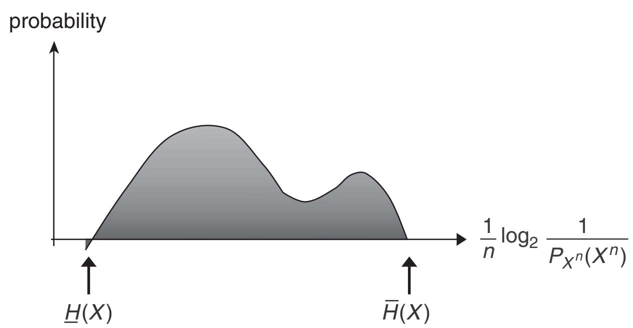


Fig. 1 entropy-spectrum

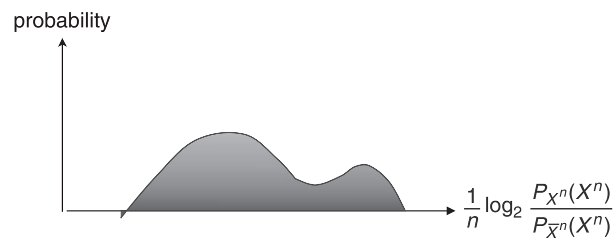


Fig. 2 log-likelihood ratio spectrum

These questions sound quite natural. However, some people may say that such a tremendous generalization never leads us to any fertile results, or some other people may say that it leads us just to a boring mathematical overgeneralization. Nevertheless, our claim is that we can successfully reach a highly unconventional but fairly fertile new world even under (or owing to) such a tremendous generalization.

In the sequel, let us show, by invoking the information spectrum methods, how to cope with this kind of elusive problems on the basis of several typical illustrative cases.

2. General Sources and Generalized AEP

A *general source* is specified by giving a sequence of random variables $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$, where X^n is the source variable of block length n and takes values in $\mathcal{X}^n$ where $\mathcal{X}$ is a countably infinite set. Here we do not require the *consistency condition* as probabilistic processes. In other words, it is allowed that X^m is not necessarily a prefix subvector of X^n when $m < n$.

We now need to introduce the following two fundamental probabilistic limit operations:

$$\text{p-}\limsup_{n\to\infty} Z_n \equiv \inf\{\alpha \,|\, \lim_{n\to\infty} \Pr\{Z_n > \alpha\} = 0\},$$

$$\text{p-}\liminf_{n\to\infty} Z_n \equiv \sup\{\beta \,|\, \lim_{n\to\infty} \Pr\{Z_n < \beta\} = 0\},$$

where $(Z_1, Z_2, \cdots)$ is a sequence of real-valued random variables. Then, we can define with a source $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$:[†]

$$\overline{H}(\mathbf{X}) \equiv \text{p-}\limsup_{n\to\infty} \frac{1}{n} \log \frac{1}{P_{X^n}(X^n)},$$

$$\underline{H}(\mathbf{X}) \equiv \text{p-}\liminf_{n\to\infty} \frac{1}{n} \log \frac{1}{P_{X^n}(X^n)},$$

where the former and the latter are called the spectral sup-entropy rate, the spectral inf-entropy rate, respectively. The probability distribution of $1/n \log 1/P_{X^n}(X^n)$ is called the *entropy-spectrum* or *information spectrum* of $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$. It is obvious that $\underline{H}(\mathbf{X}) \le \overline{H}(\mathbf{X})$.

Also, we say that the *generalized AEP* holds if there exists a subset $T_n^{\varepsilon} \subset \mathcal{X}^n$ such that $P_{X^n}(T_n^{\varepsilon}) \to 1$ and $P_{X^n}(\mathbf{x}) \simeq \exp[-n(\underline{H}(\mathbf{X}) \pm \varepsilon)]$ for $\forall \mathbf{x} \in T_n^{\varepsilon}$ and any small $\varepsilon > 0$. Moreover, a general source $\mathbf{X}$ is said to satisfy the *strong converse property* if, with any given coding rate R, the optimal fixed-length source coding always yields either $\lim_{n\to\infty} \varepsilon_n = 0$ or $\lim_{n\to\infty} \varepsilon_n = 1$, where ε_n is the decoding error probability. Then,

Theorem 2.1 *The following statements are equivalent:*

1) $\mathbf{X}$ has a one-point spectrum,

2) $\underline{H}(\mathbf{X}) = \overline{H}(\mathbf{X})$,

3) $\mathbf{X}$ satisfies the AEP,

4) $\mathbf{X}$ satisfies the strong converse property.

Theorem 2.2 *(Han et al. [4], Vembu et al. [5]):*

The infimum $R_f(\mathbf{X})$ of achievable source coding rates is given by $R_f(\mathbf{X}) = \overline{H}(\mathbf{X})$. Moreover, the supremum S_ι^* of achievable intrinsic randomness generation rates is given by $S_\iota^* = \underline{H}(\mathbf{X})$.

[†] $P_{X^n}(\cdot)$ denotes the probability distributions of X^n.

3. Folklore in Source Coding

We now consider fixed-length source coding with encoder $\varphi_n : \mathcal{X}^n \to M_n$, decoder $\psi_n : M_n \to \mathcal{X}^n$, and decoding error probabiltiy ε_n, where $M_n \equiv \{1, 2, \cdots, M_n\}$ is the message set. The following is the long-standing folklore in source coding: "The encoder output $\varphi_n(X^n)$ working at the *optimal* rate for a source $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ looks almost completely random when $\varepsilon_n \to 0$."

Many people may say that it is trivial, because otherwise we can further compress the encoder output $\varphi_n(X^n)$ to remove redundant reguralities still contained in $\varphi_n(X^n)$.

In order to give a conclusive answer to this question, it is quite useful to invoke the information spectrum approach as follows. First, set three general sources as $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$, $\hat{\mathbf{X}} = \{\hat{X}^n \equiv \varphi_n(X^n)\}_{n=1}^{\infty}$, $\tilde{\mathbf{X}} = \{\tilde{X}^n \equiv \psi_n(\varphi_n(X^n))\}_{n=1}^{\infty}$. Then,

Theorem 3.1 *(Invariance Theorem): In the process of source coding with $\varepsilon_n \to 0\,(n \to \infty)$, three information-spectra for $\mathbf{X}$, $\hat{\mathbf{X}}$, $\tilde{\mathbf{X}}$ asymptotically have the same form. More formally, this means the following:*

$$\lim_{n\to\infty} L\left(\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)}, \frac{1}{n}\log\frac{1}{P_{\hat{X}^n}(\hat{X}^n)}\right) = 0,$$

$$\lim_{n\to\infty} L\left(\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)}, \frac{1}{n}\log\frac{1}{P_{\tilde{X}^n}(\tilde{X}^n)}\right) = 0,$$

where Lévy distance $L(U, V)$ between two real-valued random variables U, V is defined as the infimum of $\mu > 0$ such that

$$\begin{aligned} &\Pr\{U \le x - \mu\} - \mu \\ &\le \Pr\{V \le x\} \le \Pr\{U \le x + \mu\} + \mu \end{aligned}$$

for all $x \in \mathbf{R}$ (the set of all real numbers).

Next, we need to formalize the notion of "almost completely random:" A general source $\mathbf{Z} = \{Z^n\}_{n=1}^{\infty}$ is said to be *almost completely random* with size $\{M_n\}_{n=1}^{\infty}$ if $\lim_{n\to\infty}(1/n)D(Z^n \| U_{M_n}) = 0$ and $\lim_{n\to\infty}(1/n)\log M_n$ exists, where $D(\cdot\|\cdot)$ denotes the divergence and U_{M_n} is the uniform random variable on the message set $M_n = \{1, 2, \cdots, M_n\}$.

Now we have the folklore theorem:

Theorem 3.2 *The encoder output $\varphi_n(X^n)$ working at the optimal rate for a general source $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ is almost completely random if and only if the source $\mathbf{X}$ satisfies the strong converse property (= AEP).*

Remark 3.1 The proof of Theorem 3.2 is carried out on the basis of Theorems 2.1, 2.2 and 3.1.

4. Reliability Function in Source Coding

Another illustrative case is the problem to compute the reliability function in source coding. A rate R is said to be r-achievable if $\limsup_{n\to\infty} 1/n \log M_n \le R$ and $\liminf_{n\to\infty} 1/n \log 1/\varepsilon_n \ge r$ for some fixed-length code (n, M_n, ε_n) with encoder φ_n, decoder ψ_n and $\varepsilon_n = \Pr\{\tilde{X}^n \ne X^n\}$, where $\tilde{X}^n = \psi_n(\varphi_n(X^n))$. Let $R_e(r\,|\,\mathbf{X})$ denote the infimum of r-achievable rates, then we have

Theorem 4.1

$$R_e(r\,|\,\mathbf{X}) = \sup_{R\ge 0}\{R - \sigma(R)\,|\,\sigma(R) < r\},$$

where

$$\sigma(R) = \liminf_{n\to\infty}\frac{1}{n}\log\frac{1}{\Pr\left\{\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} \ge R\right\}}.$$

Remark 4.1 The proof of Theorem 4.1 is carried out by using the technique of *information spectrum slicing* (cf. Han [6]). This technique effectively applies also to problems such as random number generation, rate-distortion function, hypothesis testing, etc.

Example 4.1: 1) Let $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ be an i.i.d. source subject to probability distribution P on a *finite* alphabet. Then, The formula of Longo and Sgarro [7] immediately follows from Theorem 4.1 together with Sanov's theorem, i.e., $R_e(r\,|\,\mathbf{X}) = \sup_{Q:D(Q\|P)<r} H(Q)$ where $H(Q)$ indicates the entropy.

2) Let $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ be an i.i.d. source subject to probability distribution $P = (p_1, p_2, \cdots)$ on *countably infinite* alphabet $\{1, 2, \cdots\}$, where

$$p_i = \frac{c}{(i+1)(\log(i+1))^4} \quad (i = 1, 2, \cdots),$$

which has heavy tail but finite entropy. Then, it follows again from Theorem 4.1, together with Cramér's theorem, that $\sigma(R) = 0$ for all R, so $R_e(r\,|\,\mathbf{X}) = +\infty$ for all $r > 0$, which means that any finite rate is not enough for exponentially decreasing error probability.

5. Power Exponent for Hypothesis Testing

One more illustrative case concerns hypothesis testing. Let $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ and $\overline{\mathbf{X}} = \{\overline{X}^n\}_{n=1}^{\infty}$ (with *abstract* alphabet $\mathcal{X}$ in general) be two general sources that are used as

$$H : \mathbf{X}\,(\text{null hypothesis}),$$

$$\overline{H} : \overline{\mathbf{X}}\,(\text{alternative hypothesis}).$$

If data $\mathbf{x}$ falls in A_n, then output H; otherwise output $\overline{H}$, where $A_n \subset \mathcal{X}^n$ is an acceptance region. Set $\mu_n \equiv \Pr\{X^n \notin A_n\}$, $\lambda_n \equiv \Pr\{\overline{X}^n \in A_n\}$. We say that R is r-achievable if $\liminf_{n\to\infty} 1/n \log(1/\mu_n) \ge r$ and $\liminf_{n\to\infty}(1/n)\log 1/\lambda_n \ge R$. Furthermore, denote by $B_e(r\,|\,\mathbf{X}\|\overline{\mathbf{X}})$ the supremum of r-achievable rates. Then, again by means of the *information spectrum slicing* technique, we have

Theorem 5.1

$$B_e(r\,|\,\mathbf{X}\|\overline{\mathbf{X}}) = \inf_{R}\{R + \eta(R)\,|\,\eta(R) < r\},$$

where

$$\eta(R) = \liminf_{n\to\infty}\frac{1}{n}\log\frac{1}{\Pr\left\{\frac{1}{n}\log\frac{P_{X^n}(X^n)}{P_{\overline{X}^n}(X^n)} \le R\right\}}$$

and the likelihood ratio $P_{X^n}(\mathbf{x})/P_{\overline{X}^n(\mathbf{x})}$ denotes the Radon-Nikodym derivative in general, the probability distribution of which is called the *information spectrum* or the *log-likelihood ratio spectrum*.

Example 5.1 1) Let $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ and $\overline{\mathbf{X}} = \{\overline{X}^n\}_{n=1}^{\infty}$ be i.i.d. subject to probability distributions $P, \overline{P}$ on finite alphabet $\mathcal{X}$. It is easy to see that Theorem 5.1, together with Sanov's theorem, yields Hoeffding's theorem [8]:

$$B_e(r \mid \mathbf{X} \| \overline{\mathbf{X}}) = \inf_{Q: D(Q\|P) < r} D(Q\|\overline{P}).$$

2) Let $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ and $\overline{\mathbf{X}} = \{\overline{X}^n\}_{n=1}^{\infty}$ be i.i.d. subject to Gaussian distributions $N(\kappa, \sigma^2)$, $N(\overline{\kappa}, \sigma^2)$, respectively. In this case, Theorem 5.1, together with Cramér's theorem, yields the formula with $a = (\kappa - \overline{\kappa})^2/2\sigma^2$:

$$B_e(r \mid \mathbf{X} \| \overline{\mathbf{X}}) = \min\{[a - r]^+, (\sqrt{r} - \sqrt{a})^2\}$$
$$= (\sqrt{r} - \sqrt{a})^2 \mathbf{1}[r \le a].$$

3) Finally, let us take a rather pathological case. Let the source alphabet be $\mathcal{X} = \{0, 1\}$ and fix an arbitrary subset $S_n \subset \mathcal{X}^n$ of size $|S_n| = 2^{\alpha n}$, where α is a constant such that $0 < \alpha < 1$. Furthermore, fix $\mathbf{x}_0, \mathbf{x}_1 \in \mathcal{X}^n - S_n$ with $\mathbf{x}_0 \neq \mathbf{x}_1$ arbitrarily. Let us consider the general source $\mathbf{X} = \{X^n\}_{n=1}^{\infty}$ whose probability distribution is given by

$$P_{X^n}(\mathbf{x}) = \begin{cases} 2^{-2\alpha n} & \text{for} \quad \mathbf{x} \in S_n, \\ 2^{-3\alpha n} & \text{for} \quad \mathbf{x} = \mathbf{x}_1, \\ 1 - 2^{-\alpha n} - 2^{-3\alpha n} & \text{for} \quad \mathbf{x} = \mathbf{x}_0, \\ 0 & \text{otherwise.} \end{cases}$$

It is evident that this source has the three-point information spectrum. Moreover, let $P_{\overline{X}^n}(\mathbf{x}) = 2^{-n}$ for all $\mathbf{x} \in \{0, 1\}^n$. Then, we can *directly* compute $\eta(R)$ (note that either Sanov's theorem or Cramér's theorem does *not* work here!) as

$$\eta(R) = \begin{cases} +\infty & \text{for} \quad R < 1 - 3\alpha, \\ 3\alpha & \text{for} \quad 1 - 3\alpha \le R < 1 - 2\alpha, \\ \alpha & \text{for} \quad 1 - 2\alpha \le R < 1, \\ 0 & \text{for} \quad 1 \le R. \end{cases}$$

Thus,

$$B_e(r|\mathbf{X}\|\overline{\mathbf{X}}) = \begin{cases} 1-\alpha & \text{for} \quad r > \alpha, \\ 1 & \text{for} \quad 0 < r \le \alpha. \end{cases}$$

6. Concluding Remarks

So far we have demonstrated several illustrative cases of the information spectrum method (AEP, Folklore, Reliability Function, Power Exponent), which, together with Sanov's theorem or Cramér's theorem, has been shown to effectively work for these problems. Moreover, this method has turned out to provide a quite reasonable framework also for quantum information problems (e.g., Nagaoka and Hayashi [9], Matsumoto [10], Wang and Renner [11], Datta and Renner [12], Bowen and Datta [13]), because, in general, sources and channels in quantum communication cannot be assumed to be *memoryless*, owing to entanglement or adversary intervention in the security problem, etc.

On the other hand, Wang, Colbeck and Renner [14] has shown that the probabilistic limit operations above defined can be expressed in terms of smooth Rényi's 0-divergence or 0-entropy (also, cf. Uyematsu [15]). Among others, Datta and Renner [12] has shown the equivalence between the information spectrum approach and the smooth Rényi entropy approach.

It shoud be noted here also that the above stated hypothesis testing problem can be generalized by replacing the alternative hypothesis $\overline{\mathbf{X}} = \{\overline{X}^n\}_{n=1}^{\infty}$ with a sequence of nonnegative σ-finite measures $\{\overline{G}_n\}_{n=1}^{\infty}$, which renders the reliability function problem in source coding to the power exponent problem for thus generalized hypothesis testings (cf. Han [6]).

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, pp. 623–656, 1948

[2] R. M. Gray and D. S. Ornstein, "Block coding for discrete stationary $\bar{d}$-continuous noisy channels," *IEEE Transactions on Information Theory*, vol. IT- 25, no. 3, pp. 292–306, 1979

[3] R. M. Gray, *Entropy and Information Theory*, Springer-Verlag, New York, 1990

[4] T.S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Transactions on Information Theory,* vol. IT-39, no. 3, pp. 752–772, 1993

[5] S. Vembu and S. Verdú, "Generating random bits from an arbitrary source: fundamental limits," *IEEE Transactions on Information Theory,* vol. IT- 41, no. 5, pp. 1322–1332, 1995

[6] T. S. Han, *Information-Spectrum Methods in Information Theory,* New York, Springer-Verlag, 2003

[7] G. Longo and A. Sgarro, "The source coding revisited: a combinatorial approach," *IEEE Transactions on Information Theory,* vol. IT-25, pp. 544–548, 1979

[8] W. Hoeffding, "Asymptotically optimal test for multinomial distributions," *Annals of Mathematical Statistics,* vol. 36, pp. 369–400, 1965

[9] H. Nagaoka and M. Hayashi, "An Information-spectrum approach to classical and quntum hypothesis testing for simple hypotheses," *IEEE Transactions on Information Theory,* vol. IT-53, pp. 534–549, 2007

[10] M. Matsumoto, "Entanglement cost and distillable entanglement of symmetric space," arxiv:0708.3129, 2007

[11] L. Wang and R. Renner, "Oneshot- classical capacity of quantum channels," presented at *The Twelfth Workshop on quantum information prossesing (QIP),* poster session, 2009

[12] N. Datta and R.Renner, " Smooth entropies and the quntum information-spectrum," *IEEE Transactions on Information Theory,* vol. IT-55, no. 6, pp. 2807–2815, 2009

[13] G. Bowen and N. Datta, "Beyond i.i.d. in quantum information theory," *Proc. IEEE International Symposium on Information Theory,"* 2006, arXiv:quant-ph/0604013

[14] L. Wang, R.Colbeck and R. Renner, "Simple channel coding bounds," ArXiv:0901.0834v3, 2009

[15] T. Uyematsu, "Relating source coding and resolvability: a direct approach," *Proc. IEEE International Symposium on Information Theory,"* 1350–1354, Austin, 2010

IT Society Member Honored

H. Vincent Poor of Princeton University was elected to the U. S. National Academy of Sciences (NAS) at its annual meeting on May 3, 2011. A long-time member of the IT Society, Vince's research interests are in the areas of statistical signal processing, stochastic analysis and information theory, and their applications in wireless networks and related fields. His publications in these areas include 13 books, 11 patents and more than 200 papers in the publications of the IT Society. At Princeton, he is Dean of the School of Engineering and Applied Science, and the Michael Henry Strater University Professor of Electrical Engineering.

Vince has served in a number of volunteer capacities within the IT Society, including as Society President in 1990, as the Editor-in-Chief of the *Transactions* in 2004-07, and as the General Co-chair of the 2009 ISIT, held in Seoul. In 2001, he received the Joint Paper Award of the Communications and Information Theory Societies, and in 2008 he received the Society's Aaron D. Wyner Distinguished Service Award.

In addition to the NAS, Vince is also a member of the U. S. National Academy of Engineering, the American Academy of Arts & Sciences, and the Royal Academy of Engineering of the U. K. An IEEE Fellow since 1987, he is also a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other scientific and technical organizations. Other recognition of his work includes a Guggenheim Fellowship in 2002, the 2005 IEEE Education Medal, the 2010 IET Ambrose Fleming Medal, the 2011 IEEE Eric E. Sumner Award, and an honorary doctorate from the University of Edinburgh, awarded in June 2011.

The Historian's Column

Anthony Ephremides

It is time to bestow the coveted Honorary Historian Award to yet another colleague who alerted me to some vintage historical material. This time I acknowledge Sergio Verdu who found some "gems" in the September 1955 issue of what used to be our Transactions.

At that time, as I have remarked in the past, the journals were more like non-technical journals in that they included technical articles alright, but also some worthy news items (we now have the newsletter for those) and some true correspondence items, that is letters to the editor.

The issue mentioned above contains a most interesting letter by none other than R. M. Fano (of Fano inequality fame), who was then still an Associate Professor at MIT. His letter mused about the 1954 Symposium on Information Theory for which he had been the General Chair. A parenthetical note here: given the travails and tribulations of being the General Co-chair of the 2011 ISIT that I am experiencing as I am writing this column, it was refreshing to see the kinds of concerns that dominated debate in those days. So, the main point and observation that Fano made in his letter was this. He thought that the participants (who numbered about 300) should be classified into three categories. Category A consisted of those who were doing active research in the field and were able to carry out thorough discussion amongst themselves. Category B included mainly those who are interested in various practical applications and need to keep abreast of the main developments in the work of those in category A. Finally, there was Category C that consisted of people who for various reasons wish to keep informed about the broad lines of progress in the field but are not interested in the details.

The main conclusion of the letter was that the ISIT served the needs of groups A and C pretty well. However, he felt that the need of group B presented a serious challenge. Come to think of it, nothing much has changed in the half century that followed these observations. I would argue that today the majority of our participants are in group A. And we do have some from group C (like program managers of funding agencies, social butterflies, intrepid travelers, intellectually curious colleagues from sister fields, etc). But we would be hard-pressed to find members of group B in our midst.

This is indeed a pity. Especially after the transformational changes of Judge Green's decision in the mid-eighties that essentially altered for life the character of Bell Labs, we hardly see any practitioners or Industry people at the ISITs. The pity is that the power and impact of the field has not been diluted. On the contrary, it can be argued that it has been strengthened. Of course, the time lag from "pub-to-product" remains long. But the absence of the people who are in the trenches and the ramparts of technology makes actually this time lag longer. It would be really useful for our field to come up with ways to increase their participation and active involvement.

It is worth noting in Fano's letter that the number of sessions was six (6!) and all of them were in series. Each paper was presented for about 20 minutes with 10 minutes left for discussion which was invariably lively and had to be stopped by the chair person to make sure that the schedule was kept in time. But consider this: The papers were FIRST published in the Transactions and THEN they were presented at the Symposium!! That is a total reversal of what we do today. The distinct advantage of this arrangement was that the audience had the benefit of studying the papers carefully ahead of time , which enabled an in-depth discussion after the presentation. Not a bad idea!

Fano felt that the meeting was so good that it should have been attended by many more than the 300 or so registrants (80% to 90% of whom tended to attend almost every session). However, he did spot the main weakness that has kept the field "esoteric" and off-limits to wider audiences. It was the narrow and deep technical focus on one hand, and the lack of good up-to-date review and tutorial presentations by appropriate experts. I guess the timid resumption of tutorials in our ISITs (that commenced only in the last decade) represents a step in the direction suggested by Fano already back then.

It was very interesting to read the views and facts in that letter and it made me think that we might benefit from having similar discussions and exchanges today (perhaps not in the Transactions but in the Newsletter). Our members make wonderful scientists but very timid interlocutors who are not very willing to engage in open discourse about these matters. I think we should encourage them to step forth and engage us and get engaged. We would all benefit.

But in that same issue there was, in addition to the Fano letter, a further source of delight. It concerned a brief report of the "Information Theory and Modulation Systems" Committee on its challenging job of developing standard definitions for our most important (still new at the time) technical terms. Here is a sampling of its wisdom (that was modestly followed by a solicitation of comments and feedback from the members)

BIT: A unit of information content or capacity, equal to the information content in a binary decision between equally probable hypotheses.

SIGNAL: The physical embodiment of a message

INFORMATION CONTENT (of a message from a source): The negative logarithm of the probability that that particular message will be emitted by the source

CHANNEL: A transmission path and associated terminal equipment capable of receiving signals at one point and delivering related signals at another point

CODE; A set of transformation rules to be applied to messages or signals

CODE CHARACTER: The signal representation of a discrete value or symbol in a message

CODE ELEMENT: One of a finite set of parts from which code characters may be constructed

And last, but not least:

HARTLY: A unit of information content or capacity, equal to the information content in a decimal decision between ten equally probable states!

You can almost feel the heat and anguish of the committee's deliberations at that early time.

WSIT'11 – 11th IEEE Winter School of Information Theory

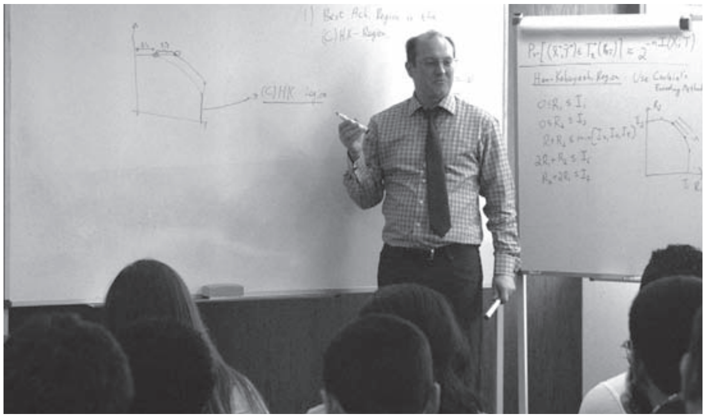
Barcelona, Catalonia (Spain)
March 14th–18th 2011

Continuing the tradition started two decades ago by Han Vinck and Rolf Johannesson, the 11th Winter School of Information Theory took place between the 14th and the 18th of March, 2011, in Barcelona, its southernmost location to date. This edition was jointly hosted by UPF (Universitat Pompeu Fabra) and CTTC (Centre Tecnològic de Telecomunicacions de Catalunya). As usual, the purpose of the week-long event was to provide an opportunity for doctoral students from different universities to gather and interact on a wide range of research subjects related to Information Theory. The 70 available slots filled up rapidly once registration opened. Altogether, over 110 applications for participation were received, a record number for the Winter School thus far.

The organizing committee for this edition included Profs. Angel Lozano and Ezio Biglieri, and Drs. Xavier Mestre, Azadeh Faridi, Vanesa Daza and Deniz Gündüz. The event was financially supported by the IEEE Information Theory Society, the Spanish Ministry of Science and Innovation, the Catalan Research Agency AGAUR, and Universitat de València. This broad support allowed keeping a low registrativon fee.

The Winter School consisted of morning lectures by distinguished professors and afternoon short talks given by the students themselves. The five lecturers this year were Prof. Helmut Bölcskei (ETH Zürich), who covered the topic of compressed sensing, Prof. Emre Telatar (EPFL), who addressed polar codes, Prof. Gerhard Kramer (Technische Universität München) who lectured

Group photo on the terrace of the UPF building where the lectures were held. A perfect sunny "winter" day in Barcelona!



Prof. Kramer explaining the intricacies of the interference channel.

on the interference channel, Prof. Daniel P. Palomar (Hong Kong University of Science and Technology), who introduced variational inequality theory, and Prof. Baltasar Beferull-Lozano (Universitat de València), who dealt with gossip algorithms. All of the lectures were lively, illuminating, and highly interactive. In the afternoons, students presented their ongoing work to their peers and to the professors and received valuable feedback.

Wednesday afternoon was free, allowing all attendees to enjoy the architectural and artistic wonders of Barcelona. The social program also included a banquet dinner in a central restaurant next to the popular "Ramblas", the core of Barcelona's old town. The instructors also got to enjoy a traditional Catalan meal with "calçots" (mild green onions dipped in romesco sauce) and other delicacies from the world-famous local gastronomy.

A word of gratitude goes to all the members of the organizing committee and to UPF's tireless staff —in particular Miss Joana Clotet and Miss Beatriz Abad— and to all the participants, who made this edition of the Winter School a resounding success. For the next edition, the southern drift continues into Turkey under the stewardship of Dr. Deniz Gündüz. Encouraged by the growing demand for participation, the next edition might arrive sooner than the usual two-year interval. Our best wishes for the 12th edition, and for many more to come!

For additional information on the 2011 Winter School:
http://www.dtic.upf.edu/~afaridi/WinterSchool

In Memoriam, Igor Vajda

Igor Vajda, Principal Researcher at the Institute of Information Theory and Automation (ÚTIA) of the Academy of Sciences of the Czech Republic in Prague, and a leading figure in the field of information-theoretic statistical inference, passed away unexpectedly on May 2, 2010.

Igor Vajda

Igor was born on October 20, 1942, in Martin, Czechoslovakia. After attending elementary and secondary school in Slovakia, he graduated in mathematics at Czech Technical University (CTU) in Prague in 1965, and received Candidate of Science and Doctor of Science degrees from Charles University in 1968 and 1990, respectively, with specializations in probability, statistics and mathematical informatics. He had been a key researcher at ÚTIA since 1965 and held a teaching position at CTU since 1969. For many years he was head of the Department of Stochastic Informatics of ÚTIA and served on the Scientific Boards of both ÚTIA and the Faculty of Electrical Engineering of CTU.

He spent the year 1966–1967 at the Institute for Information Transmission Problems (IPPI) in Moscow, and was a visiting professor of mathematics at the Military Technical College in Cairo from 1973 to 1976. After the political changeover in Czechoslovakia in 1989 he developed many contacts abroad, especially in Western Europe and the USA. He held many research grants since 1991, frequently traveled, and intensively cooperated in research with colleagues in other countries in the period 1991–2010.

Igor Vajda early recognized the meaning of different types of distances between distributions in information theory and mathematical statistics. One of his major research directions was the investigation of f-divergences $D_f(P\|Q) = \int qf(p/q)d\mu$ of distributions P and Q with densities $p = dP/d\mu$ and $q = dQ/d\mu$, and their statistical applications. In early papers he studied the relations between f-divergences and variational distance, the approximation, monotonicity, topological properties of f -divergences and their minimization under constraints.

A first systematic theory of f-divergences was presented in the book *Convex Statistical Distances* [2], with applications to hypothesis testing, minimum distance estimation, and random processes.

Igor Vajda's book *Theory of Statistical Inference and Information* [3], a first version of which was published in Slovak language in 1982, provides a comprehensive treatment of the theory of statistical inference and information. This book is unique in the field, contains a wealth of research results and has become an indispensable source of reference for researchers in the domain.

Igor Vajda used special f-divergences to generalize the Cramer-Rao bound and the theorems of Chernoff and Stein. The extension of the likelihood ratio statistic to divergence-based statistics for testing composite hypotheses was the subject of many papers of which he was the author or a co-author.

In order to compare the empirical distribution $\hat{P}_n$ with the theoretical distribution P_θ from a parametric model, one must turn to a sequence of partitions of the sample space. Igor Vajda and co-authors characterized the suitable speed of refining, and studied sequences of partitions generated by the quantile function.

Another research topic of Igor Vajda was the divergence-based estimation and testing in mathematical statistics. He used the distance $D_f(\hat{P}_n\|P_\theta)$ between the empirical distribution $\hat{P}_n$ and the true distribution P_θ to introduce the minimum divergence estimator $\hat{\theta}_n = \arg\min_{\theta\in\Theta} D_f(\hat{P}_n \| P_\theta)$ and to construct the statistic $T_n = \min_{\theta\in\Theta} D_f(\hat{P}_n \| P_\theta)$ for testing a parametric hypothesis and for goodness of fit tests. He and co-authors showed that f-divergence based estimates have similar properties as the maximum likelihood estimator. These results, as well as the efficiency in the Pitman and Bahadur sense and large deviation results on tests based on T_n, were the subject of many papers of which Igor was the author or a co-author.

Moreover, Igor Vajda early recognized the importance of the Barron distribution estimator, originally introduced (in 1988) for the consistent nonparametric estimation of an unknown probability distribution in the sense of information divergence, and later (1992) extended to consistent estimation in total variation and reversed information divergence. In a series of papers, he and co-authors proved many other results concerning this estimator, such as consistent distribution estimation in chi-square divergence, consistent density estimation according to various distance criteria, and parametric minimum distance point estimation.

Igor Vajda also developed together with others a general theory of goodness-of-fit tests based on spacings from the viewpoint of disparity statistics. In joint work he showed that essentially all spacings-based statistics are asymptotically equivalent to a disparity statistic and proved limit laws for this class of statistics, with special attention to power divergence statistics, thereby obtaining closed-form expressions for the asymptotic parameters of these power divergence spacings-based statistics.

In recent papers Igor Vajda studied the Bregman distance, which provides the flexibility to find a compromise between efficiency and robustness of estimators based on it.

He also worked in applied areas such as the statistical analysis of optimal investments, financial mathematics, biostatistics, and stochastic systems and networks.

Igor Vajda was a passionate researcher, always exploring new ideas. His results were respected by the scientific community and influenced the research orientation of many of his younger colleagues. In his scientific work he continued the traditional Prague school of information theory, which dates back to the early 1950s. Igor was an author or co-author of four monographs and more than 100 publications in renowned international journals. He had over 300 officially registered citations.

He was regularly associated with the organization of the Prague Conference on Information Theory, Statistical Decision Functions and Random Processes, held every three years since 1956. These conferences provided an opportunity for information-theorists and statisticians from the West and East to meet during the time of the cold war and until the political changes in Eastern Europe. In later years this conference merged with the Prague Stochastics meeting, held every four years since the mid-1990's. Igor was always responsible for the information theory component in these meetings.

Among the many papers Igor Vajda published, there are least 15 which appeared in the *IEEE Trans. Inf. Theory*, starting in 1970 with the paper [1]. A list of representative papers published by Igor (and co-authors) in the Transactions appears in the set of references below. This selected list of papers reflects the continuous research of Igor Vajda in the area of information theory. A particular mention deserves his paper [12], in which he showed that the convergence of quantized information functionals to the original nonreduced values of these information functionals (including the Shannon information) not necessarily needs to take place along nested partitions (as was assumed previously by Gelfand, Kolmogorov and Yaglom (1956), Kallianpur (1960), Dobrushin (1963) and Ghurye (1968)), but that these partitions can be quite general and nonnested. He, thereby, filled an important gap in the theory of convergence of quantized information functionals.

Igor Vajda's work was awarded, among other distinctions, the Prize of the Czech Academy of Sciences, the Jacob Wolfowitz Prize, the Medal of the Merits of the First Degree from the Faculty of Nuclear and Physical Engineering of CTU and several Annual Prizes awarded by the Director of ÚTIA for the best paper of the year. In November 2010, at a special Colloquium held in his memory at Prague, Igor Vajda was honored posthumously with the Bolzano Medal from the Academy of Sciences of the Czech Republic.

The high quality of his scientific results and the intensity of his international cooperation are expressed by the fact that since 1991 Igor Vajda obtained six research grants from the Czech National Grant Agency, one from the European Union (Copernicus), and one from the NSF of the USA. He participated in several other grants as a co-investigator. Since 1990 Igor Vajda was a Member and since 2001 a Fellow of the IEEE. He was a visiting professor at Katholieke Universiteit Leuven, Complutense Universidad Madrid, Université de Montpellier, and M. Hérnandez Universidad, Alicante.

Igor Vajda supervised many diploma projects and guided several PhD students both at home and abroad. He was a co-editor of the journals *Kybernetika, Problems of Control and Information Theory, Applications of Mathematics, Statistics and Decisions, Revista Matematica Complutense, Journal of Statistical Planning and Inference*, and *Test*.

Igor Vajda loved his native land Slovakia and visited it regularly throughout his life. His ashes are buried at the historic cemetery of Střílky, in Moravia, not far from Brno. He is survived by his wife Zdenka, his two daughters Tereza and Veronika, and four grandchildren.

Igor Vajda was a man of many ambitions and achieved a lot in his life. He worked tirelessly and inspired many. He enjoyed working with others. With him we lost a great friend and fine colleague. We all owe him much, miss him, and will honor his memory.

Martin Janžura, ÚTIA, Prague
Friedrich Liese, Univ. of Rostock
Edward van der Meulen, K. U. Leuven

References

[1] I. Vajda, "Note on discrimination information and variation", *IEEE Trans. Inf. Theory*, vol. 16, no. 6, pp. 771–773, Nov. 1970.

[2] F. Liese and I. Vajda, *Convex Statistical Distances*. Leipzig, Teubner-Verlag, 1987.

[3] I. Vajda *Theory of Information and Statistical Inference*. Boston, Kluwer, 1989.

[4] M. Teboulle and I. Vajda, "Convergence of best ϕ-entropy estimates", *IEEE Trans. Inf. Theory*, vol. 39, no. 1, pp. 297–301, Jan. 1993.

[5] F. Österreicher and I. Vajda, "Statistical information and discrimination", *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 1036–1039, May 1993.

[6] A. L. Rukhin and I. Vajda, "The error probability, entropy, and equivocation when the number of input messages increases", *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 2228–2231, Nov. 1996.

[7] M. C. Pardo and I. Vajda, "About distances of discrete distributions satisfying the data processing theorem of information theory", *IEEE Trans. Inf. Theory*, vol. 43, no. 4, pp. 1288–1293, July 1997.

[8] A. Berlinet, I. Vajda, and E.C. van der Meulen, "About the asymptotic accuracy of Barron density estimates", *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 999–1009, May 1998.

[9] G. Darbellay and I. Vajda, "Estimation of the information by an adaptive partitioning of the information space", *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1315–1321, May 1999.

[10] G. Darbellay and I. Vajda, "Entropy expressions for multivariate continuous distributions", IEEE Trans. *Inf. Theory,* vol. 46, no. 2, pp. 709–712, March 2000.

[11] I. Vajda and E.C. van der Meulen, "Optimization of Barron density estimates", *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1867–1883, July 2001.

[12] I. Vajda, "On convergence of information contained in quantized observations", *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2163–2172, Aug. 2002.

[13] M.C. Pardo and I. Vajda, "On asymptotic properties of informationtheoretic divergences", *IEEE Trans. Inf. Theory*, vol. 49, pp. 1860–1867, July 2003.

[14] I. Vajda and E.C. van der Meulen, "On minimum divergence adaptation of discrete bivariate distributions to given marginals", *IEEE Trans. Inf. Theory,* vol. 51, no. 1, pp. 313–320, Jan. 2005.

[15] F. Liese and I. Vajda, "On divergences and informations in statistics and information theory", *IEEE Trans. Inf. Theory*, vol. 52, pp. 4394–4412, Oct. 2006.

[16] F. Liese, D. Morales, and I. Vajda, "Asymptotically sufficient partitions and quantizations", *IEEE Trans. Inf. Theory,* vol. 52, no. 12, pp. 5599–5606, Dec. 2006.

[17] P. Harremoës and I. Vajda, "On the Bahadur-efficient testing of uniformity by means of the entropy", *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 321–331, Jan. 2008.

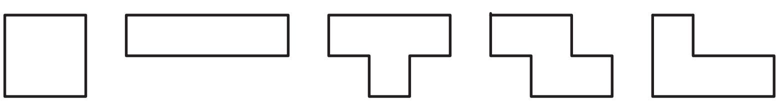
(This obituary is a slightly modified and extended version of the ones which appeared in the April/May 2011 issue of IMS Bulletin and the May 2011 issue of Bernoulli News.)

GOLOMB'S PUZZLE COLUMN™

Polyomino Tilings

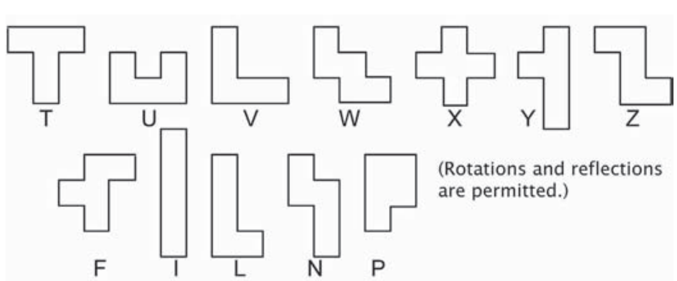
Solomon W. Golomb

1. The five tetromino shapes, popularized in the game of TETRIS ®, are:



(Rotations and reflections are permitted.) Find all the ways, inequivalent with respect to symmetries of the square, to put four tetrominoes (not necessarily different ones) into a 4×4 square.

2. The twelve pentomino shapes, popularized by Martin Gardner after I introduced them in a 1953 talk and a 1954 article, are:



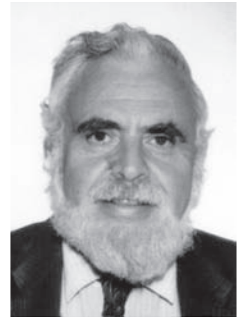
Find all the ways, inequivalent with respect to symmetries of the rectangle, to put three pentominoes (not necessarily different ones) into a 3×5 rectangle.

Hint. Each problem has fewer than 25 solutions.

GOLOMB'S PUZZLE COLUMN™

Proofs With Integers as Exponents Solutions

Solomon W. Golomb



1) Observe that $\sum_{n=0}^{\infty} x^n = (1+x)(1+x^2)(1+x^4)(1+x^8)\cdots = \prod_{j=0}^{\infty}(1+x^{2^j})$, since every positive integer has a unique binary representation (as a sum of distinct powers of 2, starting at $2^0 = 1$). Let $f(x) = (1+x)(1+x^4)(1+x^{64})\cdots = \prod_{j=0}^{\infty}(1+x^{2^{2j}})$. Then $f(x^2) = (1+x^2)(1+x^8)(1+x^{128})\cdots = \prod_{j=0}^{\infty}(1+x^{2^{2j+1}})$. Clearly $f(x)f(x^2) = \prod_{j=0}^{\infty}(1+x^{2^j}) = \sum_{n=0}^{\infty} x^n = 1/(1-x)$, valid in $|x| < 1$.

   We can write $f(x) = \sum_{a \in S} x^a$, where S is the set containing 0 and all finite sums of distinct powers of 4, starting at $4^0 = 1$. (The integers ≥ 0 in S are those whose binary representations have 1's only in odd positions, counting to the left from the binary point.) Then the direct sum of S and $2S$ gives each non-negative integer exactly once.

2) Consider the $a \times b \times c$ box as composed of abc "cells", each a unit cube, and assign the symbol x^{i+j+k} to the cell whose coordinates are (i, j, k), $1 \leq i \leq a$, $1 \leq j \leq b$, $1 \leq k \leq c$. Then the sum over the entire box is $\sum_{all\ cells} x^{i+j+k} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} x^{i+j+k} = \sum_{i=1}^{a} x^i \sum_{j=1}^{b} x^j \sum_{k=1}^{c} x^k = A(x)B(x)C(x)$. Let $\eta = e^{2\pi i/n}$, a primitive n^{th} root of unity. If the box is exactly packed with $1 \times 1 \times n$ bricks, the sum over the n cells of each brick, evaluated at $x = \eta$, take the form $\eta^{t+1} + \eta^{t+2} + \cdots + \eta^{t+n} = 0$, since we are summing over all n of the n^{th} roots of unity, and therefore the sum, at $x = \eta$, over the entire box packed with these bricks must be 0. That is, $A(\eta)B(\eta)C(\eta) = 0$. But a product of complex numbers can equal 0 if and only if one of the factors is 0. But $A(\eta) = 0$ only if a is a multiple of n, $B(\eta) = 0$ only if b is a multiple of n, and $C(\eta) = 0$ only if c is a multiple of n. So a perfect packing requires that (at least) one of the dimensions of the box be a multiple of n. (Conversely, if one of the dimensions of the box is a multiple of n, all the $1 \times 1 \times n$ bricks can be lined up along that dimension to achieve a perfect packing.) (In the second edition of my book *Polyominoes*, I give an alternative proof, using a "coloring argument", without recourse to the field of complex numbers.)

3) Suppose the progressions $\{a_i n + b_i\}_{n=0}^{\infty}$, $1 \leq i \leq k$, $k > 1$, contain every positive integer exactly once. Then $\sum_{n=0}^{\infty} z^{a_1 n + b_1} + \sum_{n=0}^{\infty} z^{a_2 n + b_2} + \cdots + \sum_{n=0}^{\infty} z^{a_k n + b_k} = z^{b_1}/(1-z^{a_1}) + z^{b_2}/(1-z^{a_2}) + \cdots + z^{b_k}/(1-z^{a_k}) = \sum_{m=1}^{\infty} z^m = z/(1-z)$, all valid in $|z| < 1$. Note that $z^{b_j}/(1-z^{a_j})$ approaches singularities as z approaches any of the a_j^{th} roots of unity from within $|z| = 1$, while $Z/(1-Z)$ approaches a singularity only as z approaches 1. If all the a_i are distinct, $1 \leq i \leq k$, $k > 1$, then in particular there is a largest a_i, say $\bar{a}$, and the singularities at the *primitive* $\bar{a}^{th}$ roots of unity (e.g. at $e^{2\pi i/a}$) cannot be cancelled by any of the other terms; while $Z/(1-Z)$ will be well-behaved as z approaches one of these singularities.

4(a) The 2-digit representation in base n for the integers from 0 to $n^2 - 1$ corresponds to the exponents in $(1 + x + x^2 + \cdots + x^{n-1})(1 + x^n + x^{2n} + \cdots + x^{(n-1)n}) = (x^n - 1)/(x-1) \cdot (x^{n^2} - 1)/(x^n - 1)$. That is, $\sum_{j=0}^{n-1} x^j \sum_{k=0}^{n-1} x^{nk} = \sum_{j=0}^{n-1}\sum_{k=0}^{n-1} x^{kn+j} = \sum_{t=0}^{n^2-1} x^t = \frac{x^{n^2}-1}{x-1} = 1 + x + x^2 + \cdots + x^{n^2-1}$. When $n = p$ is prime, the factors of $(x^{p^2} - 1)/(x-1)$ are uniquely $(x^{p^2} - 1)/(x^p - 1)$ and $(x^p - 1)/(x-1)$, the "cyclotomic polynomials" $\Phi_{p^2}(x)$ and $\Phi_p(x)$, known to be irreducible over the field of rational numbers; so in this case, the *only* way to get the set $S = \{0, 1, 2, \ldots, p^2 - 1\}$ as the direct sum of two sets $A = \{0, a_1, a_2, \ldots, a_{p-1}\}$ and $B = \{0, b_1, b_2, \ldots, b_{p-1}\}$, is to take these two sets as $A = \{0, 1, 2, \ldots, p-1\}$ and $B = \{0, p, 2p, 3p, \ldots, (p-1)p\}$.

4(b) In general, $x^{n^2} - 1$ is the product of all $\Phi_d(x)$ over all divisors d of n^2, and when n is composite there are more factors, and hence more representations as a direct sum of two n -element sets. (The cyclotomic polynomial $\Phi_d(x)$ is the polynomial whose roots are the $\varphi(d)$ primitive complex roots of unity, where φ is Euler's phi-function, and is known to be irreducible over the rational field for all positive integers d.)

   The simplest example of non-uniqueness for composite n is with $n = 4$, $n^2 = 16$. Instead of $A = \{0, 1, 2, 3\}$, $B = \{0, 4, 8, 12\}$ to get $A \oplus B = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\} = N_{15}$, we can use $A' = \{0, 1, 4, 5\}$, $B'\{0, 2, 8, 10\}$, and still get $A' \oplus B' = N_{15}$. (Note that $B' = 2A'$ here, and is related to the solution in Problem 1.)

   For $n = 10$, an alternative to $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, $B = \{0, 10, 20, 30, 40, 50, 60, 70, 80, 90\}$ to get $A \oplus B = \{0, 1, 2, \ldots, 97, 98, 99\} = N_{99}$, is $A' = \{0, 1, 2, 3, 4, 25, 26, 27, 28, 29\}$, $B' = \{0, 5, 10, 15, 20, 50, 55, 60, 65, 70\}$ to get $A' \oplus B' = N_{99}$.

Position Available

A position of a Postdoctoral Fellow is available at King Abdullah University of Science and Technology, Saudi Arabia. The Applicant should have a strong background in statistical signal processing or coding theory, with particular experience with the sum-product algorithm/belief propagation and sparse signal recovery. The applicant will work under the supervision of Dr. Tareq AlNaffouri http://faculty.kfupm.edu.sa/EE/naffouri/ and Dr. Khaled Salama http://www.kaust.edu.sa/academics/faculty/salama.html

Interested Applicants should submit as single PDF/word file that includes a cover letter, a curriculum vita with a list of publications, and the names and contact information of at least 3 references.

Applications and inquiries can be submitted to the following email: naffouri@kfupm.edu.sa

Call for Papers

First Workshop on Network Coding and Data Storage, Hong Kong, July 21–22, 2011

We welcome researchers and graduate students to participate in NCDS and discuss about potential applications and new theories of network coding in data storage.

Workshop Website

http://www.ncds2011.inc.cuhk.edu.hk

Network coding is a promising methodology for reducing the amount of data transmitted over the network by mixing packet streams at the intermediate nodes in the network. One emerging application of network coding is to improve the robustness of distributed storage systems. Recent work has shown that network coding can improve the performance of data recovery in case of node failure, while requiring less resources (storage and/or bandwidth) than traditional storage approaches such as replication and erasure coding. Nevertheless, there remains substantial research work for advancing both the theoretical and practical impacts of network coding in distributed data storage.

The First Workshop on Network Coding and Data Storage (NCDS) is organized by the Institute of Network Coding http://www.inc.cuhk.edu.hk/ at The Chinese University of Hong Kong. The aim of NCDS is to bring together researchers in the fields of network coding, data storage, and distributed systems to explore the potential of network coding applications in distributed storage systems.

About the Topic

Coding for Distributed Storage Wiki

http://csi.usc.edu/~dimakis/StorageWiki/doku.php

Registration

Registration is free. Pre-registration is required for attendance. For information about registration, please refer to http://www.ncds2011.inc.cuhk.edu.hk/registration.shtml

Call for Papers

2012 International Zurich Seminar on Communications

February 29 - March 2, 2012

The 2012 International Zurich Seminar on Communications will be held at the Hotel Zürichberg in Zurich, Switzerland from Wednesday February 29 through Friday March 2, 2012.

High-quality original contributions of both applied and theoretical nature are solicited in the areas of:

- Wireless Communications
- Information Theory
- Coding Theory and its Applications
- Detection and Estimation
- MIMO Communications
- Optical Communications
- Fundamental Hardware Issues
- Network Algorithms and Protocols
- Network Information Theory and Coding
- Cryptography and Data Security

Invited speakers will account for roughly half the talks. In order to afford the opportunity to learn from and communicate with leading experts in areas beyond one's own specialty, no parallel sessions are anticipated. All papers should be presented with a wide audience in mind.

Papers will be reviewed on the basis of a manuscript (A4, not exceeding 4 pages) of sufficient detail to permit reasonable evaluation. Authors of accepted papers will be asked to produce a manuscript not exceeding 4 pages in A4 double column format that will be published in the Proceedings. Authors will be allowed twenty minutes for presentation.

The deadline for submission is **September 25, 2011**.
Additional information will be posted at

`http://www.izs.ethz.ch/`

We look forward to seeing you at IZS.

Helmut Bölcskei and Amos Lapidoth, Co-Chairs.

Call for Papers

Solomonoff 85th Memorial Conference

3rd Call for Papers

http://www.Solomonoff85thMemorial.monash.edu/

Proceedings of this multi-disciplinary conference will be published by Springer in the prestigious LNAI (LNCS) series

Dear Colleague

You are cordially invited to submit a paper and participate at Solomonoff 85th Memorial Conference which, will be held in Melbourne, Australia, between 30 November - 2 December 2011 with the possibility of a tutorial/workshop being organised on the 29th November 2011.

This multi-disciplinary Conference will be run back to back with the AI 2011 Conference in Perth, Australia.

This is a multi-disciplinary conference based on the wide range of applications of work related to or inspired by that of Ray Solomonoff. The contributions sought for this conference include, but are not restricted to, the following:-

Statistical inference and prediction, Econometrics (*including time series and panel data*), in Principle proofs of financial market inefficiency, Theories of (quantifying) intelligence and new forms of (*universal*) intelligence test (*for robotic, terrestrial and extra-terrestrial life*), the Singularity (or infinity point), when machine intelligence surpasses that of humans), the future of science, Philosophy of science, the Problem of induction, Evolutionary (tree) models in biology and linguistics, Geography, Climate modelling and bush-fire detection, Environmental science, Image processing, Spectral analysis, Engineering, Artificial intelligence, Machine learning, Statistics and Philosophy, Mathematics, Linguistics, Computer science, Data mining, Bioinformatics, Computational intelligence, Computational science, Life sciences, Physics, Knowledge discovery, Ethics, Computational biology, Computational linguistics, Collective intelligence, structure and computing connectivity of random nets, effect of Heisenberg's principle on channel capacity, Arguments that entropy is not the arrow of time, and etc. See also Ray Solomonoff's Publications (and his obituary).

(For more details, please see Extended Call for Papers.)

General and Program Chair

David Dowe, Monash University, Australia

Program Committee

Andrew Barron, Statistics, Yale University, USA
Greg Chaitin, IBM T.J. Watson Research Center, USA
Fouad Chedid, Notre Dame University, Lebanon
Bertrand Clarke, Medical Statistics, University of Miami, USA
A. Phil Dawid, Statistics, Cambridge University, UK
Peter Gacs, Boston University, USA
Alex Gammerman, Royal Holloway, University of London, UK
John Goldsmith, Linguistics, University of Chicago, USA
Marcus Hutter, Australian National University, Australia
Leonid Levin, Boston University, USA
Ming Li, Mathematics, University of Waterloo, Canada
John McCarthy, Stanford University, USA (**Turing Award winner**)
Marvin Minsky, MIT, USA (**Turing Award winner**)
Kee Siong Ng, ANU & EMC Corp, Australia
Teemu Roos, University of Helsinki, Finland
Juergen Schmidhuber, IDSIA, Switzerland
Farshid Vahid, Econometrics, Monash University, Australia
William Uther, University of New South Wales, Australia
Paul Vitanyi, CWI, The Netherlands
Vladimir Vovk, Royal Holloway, University of London, UK

Co-ordinator

Dianne Nguyen, Monash University, Australia

You will find more information about the Conference at the following Website:
http://www.Solomonoff85thMemorial.monash.edu/

For more details on how to submit a paper(s), please refer to the Submission Page at:
http://www.Solomonoff85thMemorial.infotech.monash.edu/submission.html

Important Dates

Extended Deadline of Paper Submission: **16 June 2011**
Notification of Acceptance of Paper: **10 August 2011**
Receipt of Camera-Ready Copy: **5 September 2011**
Conference Dates: **30 Nov. – 2 Dec. 2011**

I look forward to receiving your valuable paper contribution and attendance at the Conference.

David Dowe
General Chairman

Conference Calendar

DATE	CONFERENCE	LOCATION	WEB PAGE	DUE DATE
June 5–9, 2011	2011 IEEE International Conference on Communications (ICC 2011)	Kyoto, Japan	http://www.ieee-icc.org	Passed
June 20–22, 2011	2011 IEEE Communication Theory Workshop (CTW 2011)	Sitges, Catalonia, Spain	http://www.ieee-ctw.org	Passed
July 21–22, 2011	First Workshop on Network Coding and Data Storage	Hong Kong	http://www.ncds2011.inc .cuhk.edu.hk	
July 31– August 5, 2011	2011 IEEE International Symposium on Information Theory (ISIT 2011)	St. Petersburg, Russia	http://www.isit2011.info	Passed
September 11–13, 2011	4th International Workshop on Multiple Access Communications (MACOM 2011)	Trento, Italy	http://www.csit-spb.ru /macom2011.html	Passed
September 11–15, 2011	Third International Castle Meeting on Coding Theory Applications	Barcelona, Spain	http://icmcta.uab.cat	Passed
September 28–30, 2011	49th Annual Allerton Conference on Communications, Control, and Computing	Monticello, Illinois, USA	http://www.csl.uiuc .edu/allerton	July 6, 2011
October 16–20, 2011	2011 IEEE Information Theory Workshop (ITW 2011)	Paraty, Brazil	http://www.fee.unicamp .br/itw2011	Passed
November 6–9, 2011	Asilomar Conference on Signals, Systems, and Computers (ASILOMAR 2011)	Pacific Grove, California, USA	http://www.asilomar.org	Passed
November 6–9, 2011	8th International Symposium on Wireless Communication Systems (ISWCS 2011)	Aachen, Germany	http://www.ti.rwth-aachen.de/iswcs2011/	Passed
December 5–9, 2011	2011 IEEE Global Communications Conference (GLOBECOM 2011)	Houston, Texas, USA	http://www.ieee-globecom.org	Passed
February 29–March 2, 2012	2012 International Zurich Seminar on Communications	Zürich, Switzerland	http://www.izs.ethz.ch/	September 25, 2011

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