IEEE Information Theory Society Newsletter

Vol. 53, No.4, December 2003 Editor: Lance C. Pérez ISSN 1059-2362

The Shannon Lecture

Hidden Markov Models and the Baum-Welch Algorithm

Lloyd R. Welch

Content of This Talk

The lectures of previous Shannon Lecturers fall into several categories such as introducing new areas of research, resuscitating areas of research, surveying areas identified with the lecturer, or reminiscing on the career of the lecturer. In this talk I decided to restrict the subject to the Baum-Welch "algorithm" and some of the ideas that led to its development.

I am sure that most of you are familiar with Markov chains and Markov processes. They are natural models for various communication channels in which channel conditions change with time. In many cases it is not the state sequence of the model which is observed but the effects of the process on a signal. That is, the states are not observable but some functions, possibly random, of the states are observed. In some cases it is easy to assign the values of the parameters to model a channel. All that remains is to determine what probabilities are desired and derive the necessary algorithms to compute them.

In other cases, the choice of parameter values is only an estimate and it is desired to find the "best" values. The usual criterion is maximum likelihood. That is: find the values of parameters which maximizes the probability of the observed data. This is the problem that the Baum-Welch computation addresses.

Preliminaries

Let $\mathcal{N}$ be the set of non-negative integers. Let's introduce some useful notation to replace the usual n-tuple notations:

$$[a_k]_{k=i}^{j} \equiv (a_i, a_{i+1}, \ldots, a_j)$$
$$[a(k)]_{k=i}^{j} \equiv (a(i), a(i+1), \ldots, a(j))$$

The '$k =$' will be dropped from the subscript when it is clear what the 'running variable' is.

Of particular use will be the concept of conditional probability and recursive factorization. The recursive factorization idea says that the joint probability of a collection of events can be expressed as a product of conditional probabilities, where each is the probability of an event conditioned on all previous events. For example, let A, B, and C be three events. Then

$$\Pr(A \cap B \cap C) = \Pr(A)\Pr(B \mid A)\Pr(C \mid A \cap B)$$

Using the bracket notation, we can display the recursive factorization of the joint probability distribution of a sequence of discrete random variables:

$$\Pr\big([\mathbb{X}(k)]_0^N = [x_k]_0^N\big) = \Pr(\mathbb{X}(0) = x_0) \cdot \prod_{n=0}^{N} \Pr\big(\mathbb{X}(n) = x_n \mid [\mathbb{X}(k)]_0^{n-1} = [x_k]_0^{n-1}\big)$$

Markov Chains and Hidden Markov Chains

We will treat only Markov Chains which have finite state spaces. The theory is more general, but to cover the more general case will only obscure the basic ideas.

Let **S** be a finite set, the set of states. Let the number of elements in **S** be M. It will be convenient to identify the elements of **S** with the integers from 1 to M.

Let $\{\mathbb{S}(t) : t \in \mathcal{N}\}$ be a sequence of random variables with $\Pr(\mathbb{S}(t) \in \mathbf{S}) = 1$ for all $t \in \mathcal{N}$. That is, the values of $\mathbb{S}(t)$ are confined to **S**.

Applying the above factorization to the joint distribution of the first $N+1$ random variables gives:

continued on page 10

From the Editor

Lance C. Pérez

This issue of the *IEEE Information Theory Society Newsletter* highlights Lloyd Welch's article "Hidden Markov Models and the Baum-Welch Algorithm" based on his Shannon Lecture at the 2003 International Symposium on Information Theory. Once again several members of the IT community have received major awards including an Emmy! Please see the announcements on pages 4 and 5 of this issue for details.

Please help make the Newsletter as interesting and informative as possible by offering suggestions and contributing news. The deadlines for the 2004 issues of the newsletter are as follows:

Issue	Deadline
March 2004	January 15, 2004
June 2004	April 15, 2004
September 2004	July 15, 2004
December 2004	October 15, 2004

Electronic submission, especially in ascii and Word formats, is encouraged.

I may be reached at the following address:

Lance C. Pérez
Department of Electrical Engineering
209N Walter Scott Engineering Center
University of Nebraska-Lincoln
Lincoln, NE 68588-0511
Phone: (402)472-6258
Fax: (402)472-4732
Email: lperez@unl.edu

Sincerely,
Lance C. Pérez

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

Han Vinck

In 2003, an IEEE research project measured membership satisfaction with each society and the IEEE. A total of 478 IT members responded to the questionnaire. I want to thank the participants for their valuable input. The figure below illustrates the main result for the question about IT Society (ITS) membership satisfaction and IEEE membership satisfaction.

IT Society President Han Vinck in front of Allerton House, location of the 41st Annual Allerton Conference.

The general conclusions are: (1) Most of the members are very to highly satisfied with ITS membership; (2) The IT transactions are rated as extremely useful by a very high percentage of our members and, together with the newsletter, are the main reason for joining the ITS; (3) Chapter activities and publicity need improvement; and (4) The symposium (ISIT) and workshop (ITW) are considered to be of great interest by about 50% of our members. These outcomes are very stimulating for the volunteers working for you in all kind of positions. Participants in the questionnaire also indicated how we could improve our service to the members. I will summarize and briefly comment on the reactions.

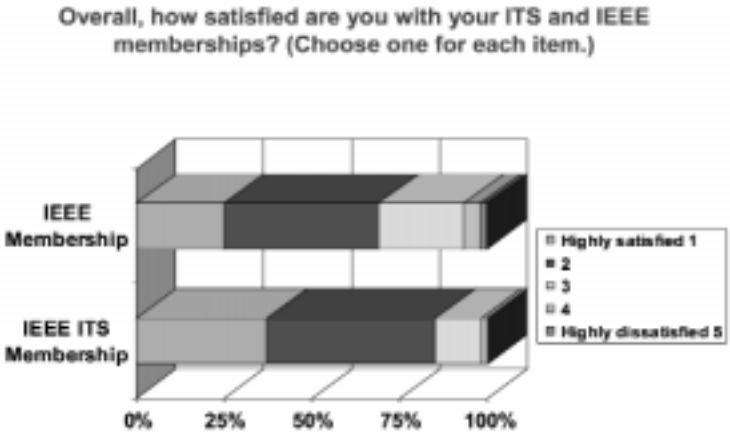
Conferences: Members suggested longer summaries (up to 4 pages), which requires a CD/DVD version of the proceedings, and the participation fee is considered to be high. Conference fees will remain a discussion item for the BoG. We try to minimize the costs of our symposium, but we also want to include a banquet for all, an award luncheon, a get-together party and nice coffee breaks. In addition, for 30% of the participants we want a reduced student participation fee and about 50 participants from low-income countries need support. Last, but not least, perfect organization is expected. Obviously, these requirements lead to a minimum price. Tutorials are well appreciated and we will continue to have tutorials in future symposia.

Transactions: (1) Many respondents complain about the time between submission and final publication in the transactions. This problem was discussed in the June BoG meeting and in the September 2003 President's column. We are currently testing a web based editorial tool developed by Publications Editor Kevin Quirk. The software is intended to simplify the electronic submission, reviewing and tracking of submitted manuscripts; (2) It has also been suggested that the ITS consider the introduction of a new "fast-track" letters publication similar to the IEEE Communications Letters for timely dissemination of research. This idea is a logical consequence of the previous complaint and could be a point of attention for the future; (3) The Transactions should be more reader friendly and more survey and tutorial papers are requested. From the responses I conclude that we have to improve the presentation of fundamental results in order to achieve a broader audience. If we want to attract young researchers to our work and promote information theory, we need to present the work in an understandable way, which even includes small worked out examples. This is also of interest to the authors! All together, our journals are highly ranked within and outside the IEEE. The Editorial Board members under the leadership of Editor in Chief Paul Siegel deserve great respect for their dedication as Information Theory Society volunteers.

Volunteers are the basis of our society. They keep the operational costs very low and create professional friendships. We need your input to maintain our high technical standards and the quality of our organized events.

A point of concern is the chapter activity. The chapters act on a local level and are of direct importance to our members and students. Check out whether there is a chapter in your environment. If not, why not consider starting a new chapter? If yes, why not participate or initiate a new activity? Be creative! Chapters improve communication with and between members. Check our web site at www.ieeeits.org for details about chapter support from the society. Our incoming 2nd Vice President David Neuhoff will be happy to help you. Our most recent chapter is the Bangalore (India) IT chapter on Information Theory, initiated by Sundar Rajan.

Another easy way to act as a volunteer in a professional way is to submit a nomination for the best paper award for the year 2003. Announcements can be found in this newsletter. The winners of the 2002 IEEE Information Theory Society Paper Award are Lizhong Zheng and David N. C. Tse, for their paper, "Communication on the Grassmann



manifold: A geometric approach to the noncoherent multiple-antenna channel", which appeared in the IEEE Transactions on Information Theory, Vol. 48, pp. 359 - 383, February 2002.

The last 2003 Board of Governors (BoG) meeting took place in the beautiful setting of the Allerton House (see picture), the conference center of the University of Illinois, in conjunction with the 41st Annual Allerton Conference on Communication, Control, and Computing. At the meeting we decided to have Seattle as the location for the 2006 ISIT with general co-chairs Joseph A. O'Sullivan (Washington University) and John B. Anderson (Lund University). Alexander Barg (University of Maryland) and Raymond Yeung (The Chinese University of Hong Kong) will lead the program committee. Tutorials connected to the ISIT are highly appreciated by our members. If the tutorials are successful at the 2004 ISIT in Chicago, we will certainly add these to the program in Seattle.

The elections for the 2004 presidents resulted in Hideki Imai for President, Steven McLaughlin for 1st Vice President and David Neuhoff for 2nd Vice President. I very much enjoyed being the President for the year 2003. The most important part of the job is to improve the communication with IEEE, the Board of Governors and you, the IT members. I thank all the members that supported the Board of Governors and me during this year. I hope to be able to contribute as past president to the improvement of our infrastructure.

Call for Nominations for the 2004 IEEE Information Theory Society Paper Award

Nominations are invited for the 2004 IEEE Information Theory Society Paper Award.

Outstanding publications in the field of interest to the IT Society appearing anywhere during 2002 and 2003 are eligible. The purpose of this award is to recognize exceptional publications in the field and to stimulate interest in and encourage contributions to the fields of interest of the IT Society.

The Award consists of an appropriately worded certificate and an honorarium of US$1000 for a single author, or US$2000 equally split among multiple authors.

NOMINATION PROCEDURE: Please email a brief rationale (limited to 300 words) for each nominated paper explaining its contributions to the field by Friday, March 5, 2004 to the Transactions Editor-in-Chief at <psiegel@ucsd.edu>, with a cc to Katherine Perry at <kperry@ucsd.edu>.

Call for Nominations for the 2004 Joint Information Theory/Communications Society Paper Award

The Joint Information Theory/Communications Society Paper Award recognizes one or two outstanding papers that address both communications and information theory. Any paper appearing in a ComSoc or IT Society publication during the year 2003 is eligible for the 2004 award.

Please send nominations to Steve McLaughlin (swm@ece.gatech.edu) by February 1, 2004.

A Joint Award committee will make the selection by April 10, 2004

2003 Paper Award Winners Announced

The Information Theory Society is pleased to announce the winners of the 2003 Joint Information Theory/Communication Society Paper Award and the 2003 IEEE Information Theory Society Paper Award.

2003 Joint IT/ComSoc Paper Award

The winners of the 2003 IEEE Communications Society and Information Theory Society Joint Paper Award are Shlomo Shamai (Shitz) and Igal Sason for their article, "Variations on the Gallager bounds, connections and applications," which appeared in IEEE Transactions on Information Theory, Vol. 48, No. 12, pp. 3029-3051, December, 2002.

Shlomo Shamai (Shitz) (S'80-M'82-SM'89-F'94) received the B.Sc., M.Sc., and Ph.D. degrees in Electrical Engineering from the Technion-Israel Institute of Technology, in 1975, 1981 and 1986, respectively.

From 1975 to 1985 he was with the Communications Research Labs in the capacity of a Senior Research Engineer. Since 1986 he is with the Department of Electrical Engineering, Technion-Israel Institute of Technology, where he is now the William Fondiller Professor of Telecommunications. His research interests include topics in information theory and statistical communications.

He is especially interested in theoretical limits in communication with practical constraints, multi-user information theory and spread spectrum systems, multiple-input-multiple-output communications systems, information theoretic models for wireless networks and systems, information theoretic aspects of magnetic recording, channel coding, combined modulation and coding, turbo codes and LDPC, in channel, source, and combined source/channel applications, iterative detection and decoding algorithms, coherent and noncoherent detection and information theoretic aspects of digital communication in optical channels.

2003 Joint IT/ComSoc Paper Award winners Igal Sason and Shlomo Shamai (Shitz).

Dr. Shamai (Shitz) is a Member of the Union Radio Scientifique Internationale (URSI). He is the recipient of the 1999 van der Pol Gold Medal of URSI, and a co-recipient of the 2000 IEEE Donald G. Fink Prize Paper Award, and the 2002 Joint Information Theory/Communication Societies Paper Award. He is also the recipient of the 2000 Technion Henry Taub Prize for Excellence in Research. He has served as Associate Editor for Shannon Theory of the IEEE Transactions on Information Theory, and also serves on the Board of Governors of the Information Theory Society.

2003 IT Society Paper Award winner Lizhong Zheng

2003 IT Society Paper Award winner David N. C. Tse

2003 IEEE Information Theory Society Paper Award

The winners of the 2003 IEEE Information Theory Society Paper Award are Lizhong Zheng and David N. C. Tse, for their paper, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel", which appeared in the IEEE Transactions on Information Theory, Vol. 48, No. 2, pp. 359 - 383, February 2002.

Lizhong Zheng received the B.S and M.S. degrees in 1994 and 1997, respectively, from the Department of Electronic Engineering, Tsinghua University, China, and the Ph.D. degree, in 2002, from the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley. Since 2002, he has been working as an assistant professor in the Department of Electrical Engineering and Computer Sciences and the Laboratory of Information and Decision Systems at the Massachusetts Institute of Technology. His research interests include information theory, wireless communications and wireless networks.

David Tse received the B.A.Sc. degree in systems design engineering from the University of Waterloo, Canada, in 1989, and the M.S. and Ph.D. degrees in electrical engineering from Massachusetts Institute of Technology in 1991 and 1994, respectively. From 1994 to 1995, he was a postdoctoral member of technical staff at A.T. & T. Bell Laboratories. Since 1995, he has been at the Department of Electrical Engineering and Computer Sciences of the University of California at Berkeley, where he is currently a Professor. He received a 1967 NSERC 4-year graduate fellowship from the government of Canada in 1989, a NSF CAREER award in 1998, the Best Paper Awards at the Infocom 1998 and Infocom 2001 conferences, the Erlang Prize in 2000 from the INFORMS Applied Probability Society, and the IEEE Joint Information Theory/Communications Society Paper Award in 2001. He is currently an Associate Editor for the IEEE Transactions on Information Theory. His research interests are in information theory, wireless communications and networking.

Emmy Award for Kees Immink

Dr Kees A. Schouhamer Immink was awarded an Emmy Award for outstanding technical achievement by the National Academy of Television Arts and Sciences (NATAS), in honor of his pioneering role in the "Coding technology for optical recording formats, such as Blu-Ray, DVD, and Compact Disc." The presentation took place during a black-tie dinner ceremony, held at the Plaza Hotel in New York City, on October 23.

Dr Kees Immink, president and CEO of Dutch-based Turing Machines Inc., and adjunct professor at the Institute for Experimental Mathematics, Essen, Germany, has a long career of envisioning and creating new technologies for digital audio and video recorders. It is virtually impossible to listen to digital audio, or watch digital video, played from any brand or type of recorder -optical, magnetic, or magneto optical-, -disc or tape- that do not use one of his inventions. The Emmy-honored coding technology used in the DVD is based on the technology used in the Compact Disc, which he developed in the early 80s, and involved the efforts spanning many years. He holds more than 400 international patents, and has earned numerous technical and scientific awards for his creative work.

In addition to his research efforts, Dr. Immink has also been strongly committed to serving the engineering profession. He is president of the Audio Engineering Society (AES), and a member of the Royal Netherlands Academy of Sciences.

Launched in 1948, Emmy's are well known to the large public. It is little known that Emmy's can also be bestowed to an individual, or company, for breakthroughs in technology that have a significant effect on television engineering. Being honored with an Emmy award is the highest recognition.

NATAS is unique among industry organizations and is dedicated to raising industry standards. The Emmy Award is worldwide recognized as the most prestigious award in the television industry. In its search for qualified Emmy recipients in technical achievement, a NATAS sub-committee identifies and thoroughly researches candidate technologies.

The Historian's Column

A. Ephremides

Although not ranking high in your list of concerns, I suspect many of you are curious to know what exactly transpires at the meetings of the Board of Governors of our Society. Looking back at the thirty some years during which, on and off, I have had the opportunity to be present at these meetings, I think I have a pretty good idea about what they are like.

First of all, do not believe the rumors. Some say that these meetings are long and boring and unworthy of the time of intelligent members of our community. Others say that they are serious and important with potentially grave consequences to our Society. And yet others say that they are playgrounds for those who playact with mock-parliamentary procedures. I assure you that the truth is different. In a nutshell, the BoG meetings are entertaining and educational exercises in fellowship and governance. Our colleagues who, over the years, have been sitting around the meeting tables have developed close kinship and have become wiser (and humbler) as they confront the complexities of collective decision making.

To be sure, there are dull moments (actually, hours); but there are also moments of excitement and fun. In the early years, the meetings were dry and generally brief. However, gradually, they have become luncheon or dinner meetings and depending on who the Society officers are, they have been accompanied by serious consumption of wine and/or various "digestives".

A typical agenda of a BoG meeting consists of the following items. First, although everyone knows everyone else, we go around the table introducing ourselves. As you might suspect, this provides the opportunity for wisecracks that are always part of a good BoG meeting. Next, the Society President presents his report. It consists typically of announcements, appointments, and lamentations about the sorry state of the IEEE and the inability of our Society to do anything about it. Usually, after that, the mood sinks deeper as the Society Treasurer presents his report. Printouts with graphs that show the evolution of our reserves circulate and generate anxiety and concern. Of late, they look like the vital sign charts of moribund patients. In earlier times they used to generate euphoria and feelings of empowerment. Typically a discussion follows that ranges from protestations and calls for sedition to constructive ideas for creative accounting practices. In the end, we thank the Treasurer and move on.

The report of the Editor of the Transactions follows. It provides statistical information about paper submission, acceptance, delay, etc. Sometimes appointments of new Associate Editors are proposed or special issues discussed. On occasion, the report causes lively discussion if a controversial idea is introduced to help reduce backlog or accelerate the review process. In general, an observer would conclude that the Transactions are by far the most important product of our Society and that the BoG, reflecting the collective membership view, takes pride in the continuing preservation of its excellence. Then, the Newsletter Editor report is presented. It is usually brief and of the "all is well" variety.

What follows is discussion and reports of past and upcoming symposia and workshops. These tend to be fun items as they provide the opportunity to gloat on past successes and fantasize about future exotic venues. As many of you know our meetings have spanned the globe. From Brazil to South Africa and from Japan to Svalbard (located at about 80° North or 700 miles from the North Pole) and from Quebec to the Alps and from Michigan and Israel, our meetings have made world travelers of many of us. Organizers of past meetings report proudly of financial surpluses and receive rounds of applause for their work. Organizers of future ones are quizzed on costs, meals, venue details, etc. This agenda item is usually the longest and most pleasurable one.

The typicality ends here. Some past and new business items come up along with the reminder of the time and location of the next BoG meeting (there are three meetings each year), and after that the adjournment bell rings.

Frequently, however, there are special items that can cause a great deal of controversy. And there are also seasonal items that are part of the agenda only at one of the three meetings. Examples of the latter are the election of officers and the establishment of the list of candidates for election to the Board. This happens at the last meeting of the year, which is officially declared the Annual meeting of the BoG according to our constitution. Another example is the "best paper award" that is discussed at the second BoG meeting (usually in the middle of the year).

Examples of the controversial items include revision of by-laws, establishment of new awards, proposals for new projects (like the establishment of the position of webmaster or the digital library), and, of course, the (in)famous issue of our relationship with our Soviet Union counterparts that dominated in the '70's and the '80's.

Unlike other societies of the IEEE, we have always been free of politicking and manipulative, backdoor deal making. It is truly remarkable how fraternity-like we have been. Our members tend to respect each other and our disagreements, however strong they may be, have always been resolved honorably and with straight dealing. Despite the strong personalities of those sitting around the table, the mood is generally jolly and constructive.

Some of the BoG meetings have lasted as little as one hour and as long as seven hours! The composition of the Board has always been broad and transnational. The styles of the Presidents (who run the BoG meetings) have ranged from the highly collegial and informal to the strict application of Robert's Rules of Order. All in all, I can certify that I have seen many of our colleagues grow from energetic and enthusiastic, but somewhat "green", volunteers, newly elected to the Board, to mature and responsible statesmen and leaders who lend their talents and denote their services to our remarkable Society.

I have had the (mis)fortune of attending Board meetings of several IEEE Societies. None come close to those of the IT Society in any measure of quality. So, for those who only hear rumors about what the BoG is like, I strongly recommend they attend these meetings (they are open to all Society members) and to consider getting involved and running for election to the Board. It is and has been an honor and a highly enjoyable and educational experience.

Robert Gallager and Bob Metcalfe Selected For Prestigious Marconi International Fellowship

Robert Gallager, a widely respected Massachusetts Institute of Technology academician and inventor of advanced communications codes, and Robert Metcalfe, who developed the ubiquitous Ethernet computer networking standard, will share this year's Marconi International Fellowship. They will join a select group of 30 of the world's most influential communications technology pioneers previously awarded the highly coveted annual recognition by the Guglielmo Marconi International Fellowship Foundation at Columbia University.

"In light of this 30th anniversary year of Ethernet it's entirely appropriate for a giant among information theorists and one who pioneered a practical application for high-speed data communications to be selected the 2003 Marconi Fellows," Dr. Charles Kao, chairman of this year's Selection Committee, observed. "The seminal contributions of Robert Gallager and Robert Metcalfe are truly representative of the audacious innovations in communications technology which the Marconi International Fellowship seeks to recognize and celebrate."

The annual Marconi Fellowship, which includes a $100,000 honorarium to be shared by this year's honorees, recognizes individuals whose aspirations, careers, and accomplishments in communications technology emulate those of Guglielmo Marconi -- scientist, engineer, inventor, and entrepreneur. The award ceremony will take place in New York on Friday, October 3, 2003.

Robert Gallager is a disciple and former collaborator of the late Claude Shannon, the founder of modern information theory that quantifies the limits of information transmitted over a communications channel. In his 1960 MIT doctoral dissertation, Professor Gallager developed a series of "Low Density Parity Check Codes" that, while long recognized as breakthrough theoretical formulae, are only now being appreciated for their practicality in achieving Shannon's theoretical limitations.

Professor Gallager is a 1953 graduate of the University of Pennsylvania and was a member of the MIT faculty from 1960 until 2001, when he became Professor Emeritus. He was a longtime consultant to Codex Corporation, which pioneered data communications and was subsequently absorbed by Motorola, Inc. [NYSE:MOT]. An author of numerous research papers, his book, Information Theory and Reliable Communications is still considered the "bible" on the topic, and his textbooks, Data Networks (joint with D. Berksekas) and Discrete Stochastic Processes, are widely used by graduate students.

A 1969 MIT graduate with postgraduate degrees from Harvard University, Robert Metcalfe is a revered technology visionary who developed Ethernet as a standard for interconnecting computers for high-speed data transfer. At the time, he was working as an engineer-scientist at the famed Palo Alto Research Center (PARC). He went on to be a founder of 3Com Corporation [NASDAQ:COMS], where at various times he was chairman, CEO and led engineering, marketing and sales organizations.

From 1976 through 1983 he was consulting associate professor of electrical engineering at Stanford University. During the 1990's he became a publisher and industry pundit, serving as CEO of InfoWorld Publishing Company and writing a column, "From the Ether", as well as contributing think pieces and often appearing as a speaker at events and on broadcast interviews. In 2001 he joined Polaris Venture Partners as a general partner, specializing in Boston-area information technology start-ups.

Both of the designated Marconi Fellows have amassed many honors over their careers. Professor Gallager is a life Fellow of the Institute of Electrical and Electronics Engineers (IEEE), member of the National Academy of Engineering (1979), the National Academy of Sciences (1992), and a Fellow of the American Academy of Arts and Sciences (1999). His honors include the IEEE Baker Prize Paper Award (1966), IEEE IT Shannon Award (1983), MIT Graduate Student Council Teaching Award (1992-3), the IEEE Bennet Prize Paper Award (1993), and two Golden Jubilee Paper Awards from the IEEE IT Society in 1998. He received the IEEE Medal of Honor in 1990. In 1999, he received the Harvey Prize in Science and Technology from the Technion, Haifa and in 2002 received the Eduard Rhein Prize for basic research.

Dr. Metcalfe holds the coveted 1980 Grace Murray Hopper Award of the Association for Computing Machinery (ACM) and was recognized in 1988 with the IEEE Alexander Graham Bell Medal, which was followed in 1996 with the IEEE Medal of Honor. He was elected to the National Academy of Arts and Sciences (1995), the National Academy of Engineering (1997), and as a Fellow of the International Engineering Consortium (1999). He serves on the boards of a number of Polaris Venture Partners companies, and is a director of Earthlink, MediaLabEurope, among others, and is a Trustee of MIT.

About the Marconi Foundation

With its motto, "Communications for Goodness Sake," the Guglielmo Marconi International Fellowship Foundation at Columbia University is dedicated to nurturing, recognizing and celebrating individuals whose ingenious application of communications technology has a positive and lasting impact on human progress around the globe. Established in 1974 through an endowment organized by Gioia Marconi Braga, the Foundation is best known for the Marconi International Fellowship, awarded annually to outstanding individuals whose scope of work and influence emulate the principle of "creativity in service to humanity" that drove Guglielmo Marconi, the father of modern communications. Additional information is available on the Foundation's website, www.marconifoundation.org

GOLOMB'S PUZZLE COLUMN™

Irreducible Divisors of Trinomials Solutions

Solomon W. Golomb

1. "A primitive polynomial $f(x)$ of degree $n \geq 2$ divides infinitely many trinomials over $GF(2)$".

Proof. Let α be a root of $f(x)$. By "primitivity", all the values $1, \alpha, \alpha^2, \alpha^3, \ldots, \alpha^{2^n-2}$ are distinct, and are all the non-zero elements of $GF(2^n)$. Therefore, for each $j, 0 < j < 2^n - 1$, $1 + \alpha^j = \alpha^k$ with $0 < k < 2^n - 1$ and $j \neq k$. Hence, $f(x)$ divides the trinomial $1 + x^j + x^k$ for these values of j and k. In addition, $f(x)$ divides $1 + x^J + x^K$ for every J with $J \equiv j \pmod{2^n - 1}$ and $K \equiv k \pmod{2^n - 1}$, since $1 + \alpha^J + \alpha^K = 1 + \alpha^j + \alpha^k = 0$, in view of $\alpha^{2^n-1} = 1$.

2. "If $f(x)$ is irreducible with primitivity t and $f(x)$ divides no trinomials of degree $< t$, then $f(x)$ divides no trinomials."

Proof (by contradiction). Suppose $f(x)$ divides the trinomial $x^N + x^A + 1$ of degree $N > t$, and let α be a root of $f(x)$. Since $f(x)$ has primitivity t, $\alpha^t = 1$. Since $f(\alpha) = 0$, where α is a root of $f(x)$, any polynomial $g(x)$ divisible by $f(x)$ also has α as a root, since $g(x) = f(x) \cdot q(x)$ gives $g(\alpha) = f(\alpha) \cdot q(\alpha) = 0 \cdot q(\alpha) = 0$. Thus, $\alpha^N + \alpha^A + 1 = 0$, from which $\alpha^n + \alpha^a + 1 = 0$ where $n \equiv N \pmod{t}$ and $a \equiv A \pmod{t}$, where we choose both n and a to be less than t, from which $f(x)$ divides the trinomial $x^n + x^a + 1$, of degree $< t$.

3. "If $p \geq 5$ is a prime for which 2 is primitive modulo p, then $f(x) = (x^p - 1)/(x - 1) = 1 + x + x^2 + \cdots + x^{p-1}$ is an irreducible polynomial which divides no trinomials."

Proof. For each prime p, $\Phi_p(x) = (x^p - 1)/(x - 1)$ is the "cyclotomic polynomial" over the rational field Q, whose roots are the $\phi(p) = p - 1$ primitive p^{th} roots of unity. While all cyclotomic polynomials are irreducible over Q, $\Phi_p(x)$ remains irreducible over $GF(2)$ if and only if 2 is primitive modulo p. In this case, $f(x) = \Phi_p(x)$ has primitivity $t = p$, and any root α of this $f(x)$ has $\alpha^p = 1$. Note that for $p \geq 5$, the minimum polynomial for such a root of unity has $p > 3$ non-zero terms. By Result 2, above, if this $f(x)$ divides *any* trinomial, it must divide a trinomial of degree $< t = p$, say $x^n + x^a + 1$ with $n < p$. But then the root α of $f(x)$ is a root of this trinomial of degree $\leq p - 1$, whereas the unique polynomial of degree $\leq p - 1$ with α as a root is the *minimal* polynomial of α, $f(x) = \Phi_p(x)$, of degree $p - 1$, which has *more than three terms.*

Note. There are also many other irreducible polynomials which divide no trinomials. These three problems are the easy results.

Gerard J. Foschini Named Bell Labs Fellow

Alex Dumas

Bell Labs President Bill O'Shea and Lucent's R&D leadership have chosen seven employees as 2002 Bell Labs Fellows. The annual award, Bell Labs' highest honor, recognizes sustained research and development contributions to the company. The 2002 awards mark the 20th year of the program. Since it began, 198 scientists and engineers have joined this elite group. Each new Fellow will receive a sculpture, a personal Fellows plaque and a cash award. A plaque of each will be added to the "wall of honor" in the Murray Hill, N.J., lobby. A formal luncheon to honor the recipients is scheduled for September.

Gerard J. Foschini

The 2002 winners are:

*Alvin Barshefsky, Lucent Worldwide Services, for his sustained performance in the area of software and services development.

*Young-Kai Chen, Bell Labs Research, for his pioneering working in developing high-speed devices and circuits.

*Gerard J. Foschini, Bell Labs Research, for his breakthrough invention of the BLAST concept that has the potential to revolutionize wireless technology.

*Theodore M. Lach, Integrated Network Solutions, for his innovation and technical leadership in switching system component and product reliability, silicon fabrication techniques and contributions to industry standards.

*Rajeev R. Rastogi, Bell Labs Research, for contributions in the areas of network management and database systems, and the successful application of these innovations to Lucent products.

*William D. Reents, Supply Chain Networks, for pioneering work in analytical science and development of leading-edge characterization tools and methodologies.

*Joseph A. Tarallo, Mobility Solutions, for sustained contributions to second- and third-generation wireless technology, and base station architecture and design.

continued on page 9

GOLOMB'S PUZZLE COLUMN™

Overlapping Subsets

– Solomon W. Golomb

A former student in my undergraduate course in combinatorial analysis recently wrote to me with a question. The 900 students in the graduate program he is now attending are partitioned into 90-student sections (for manageable class sizes) in each of several courses. These partitionings are supposedly performed randomly, and independently from one course to another. Yet he estimates an overlap of about 25 students between "his" sections in two of these courses, which seemed highly improbable to him. He sought my assistance in addressing this issue.

1. Let's generalize to the following problem: From a set S of N elements, subsets A and B are formed, independently and at random, with a elements in A and b elements in B.

(a) What is the expected number M of overlaps between set A and set B?

(b) What is the probability $pr(k)$ of exactly k overlaps between sets A and B? (Use binomial coefficients in your answer.)

(c) From your answer to 1.b., obtain a fairly simple expression for the ratio $\frac{pr(k+1)}{pr(k)}$.

2. For the case $N = 900$, $a = b = 90$,

(a) What is the value of M?

(b) Evaluate $\frac{pr(k+1)}{pr(k)}$ for each k, $0 \le k \le M+2$.

(c) From your answer to 2.b., what is the *mode* of the distribution $\{pr(k)\}$? (That is, for what value of k is $pr(k)$ biggest?)

3. Stirling's approximation formula for $n!$ says $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$, as $n \to \infty$, where $e = 2.718\ldots$ is the base of natural logarithms, and $\pi = 3.14159\ldots$.

(a) In your answer to 1.b., substitute $N = 900$, $a = b = 90$, and then substitute Stirling's approximation for each of the factorials (in each of the binomial coefficients) for the case $k = M$.

(b) Simplify the expression in 3.a., by cancellation between numerator and denominator.

(c) What numerical value does 3.b. yield for $pr(M)$?

4. The Poisson Distribution with parameter λ, given by $Pr(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$ for integers $k \ge 0$, is often used to approximate other distributions with mean equal to λ.

(a) Using the value of M from problem 1.a., what value does the Poisson Distribution give at $\lambda = k = M$?

(b) The value of $pr(M)$ in 3.c. used the Stirling approximation to $n!$ Which approximation to the "true" value of $pr(M)$, from 3.c. or from 4.a., do you believe is closer?

(c) How does $\frac{Pr(k+1)}{Pr(k)}$ with $\lambda = M$ compare with $\frac{pr(k+1)}{pr(k)}$ in 2.b., for k in the interval $[M-2, M+2]$?

5. Use any approximation method to evaluate $pr(25)$ for the case in Problem 2. Was the student's intuition correct?

Gerard J. Foschini Named Bell Labs Fellows (continued from page 8)

"These people represent the best of the best in the Bell Labs R&D community," O'Shea noted. "The consistently excellent work of these individuals and their colleagues is the type of role-model R&D that is needed to bring Lucent again to the forefront of the communications industry." A new class of Fellows is named each year based on accomplishments in the previous calendar year. Past winners include such luminaries as Dennis Ritchie and Ken Thompson, creators of the UNIX™ operating system; Roy Weber, creator of toll-free calling technology; Nobel Prize winner Horst Stormer; and Federico Capasso, co-inventor of the quantum cascade laser. Profiles on the new Fellows will appear in future issues of LT Today and Bell Labs News.

UNIX is a registered trademark of The Open Group.

Hidden Markov Models and the Baum-Welch Algorithm (continued from page 1)

$$\Pr\big([\mathbb{S}(k)]_0^N{=}[s_k]_0^N\big) = \Pr(\mathbb{S}(0) = s_0)\cdot \prod_{n=1}^{N} \Pr\big(\mathbb{S}(n){=}s_n \,|\, [\mathbb{S}(k)]_0^{n-1} = [s_k]_0^{n-1}\big) \tag{1}$$

For the sequence of random variables to be a Markov Chain the conditional probabilities must only be a function of the last random variable in the condition so that equation (1) reduces to

$$\Pr\big([\mathbb{S}(k)]_0^N{=}[s_k]_0^N\big) = \Pr(\mathbb{S}(0) = s_0)\cdot \prod_{n=1}^{N} \Pr(\mathbb{S}(n){=}s_n \,|\, \mathbb{S}(n-1) = s_{n-1}) \tag{2}$$

In my work, the transition probabilities were stationary, that is they are constant functions of time:

$$\Pr(\mathbb{S}(n) = j \,|\, \mathbb{S}(n-1) = i) = \Pr(\mathbb{S}(1) = j \,|\, \mathbb{S}(0) = i) \stackrel{\text{def}}{=} p_{ij}$$

In addition to the Markov Chain, let $\{\mathbb{Y}(t) : t \in \mathcal{N}\}$ be a sequence of random variables (called random observations). It will be convenient to assume that the values are confined to a discrete set and an experiment consists of observing values of T consecutive random variables. Again, treating a more general case will only obscure the basic ideas.

Applying recursive factorization to the joint distribution of the first $T+1$ random states and first T random observations:

$$\begin{aligned}\Pr\big([\mathbb{S}(t)]_0^T &= [s_t]_0^T \text{ and } [\mathbb{Y}(t)]_1^T = [y_t]_1^T\big) = \\ \Pr(\mathbb{S}(0) &= s_0)\cdot \prod_{t=1}^{T} \Pr\big(\mathbb{S}(t) = s_t \,|\, [\mathbb{S}(k)]_0^{t-1} = [s_k]_0^{t-1}\big)\cdot \\ \prod_{t=1}^{T} \Pr\big(\mathbb{Y}(t) &= y_t \,|\, [\mathbb{S}(k)]_0^T = [s_k]_0^T \text{ and } [\mathbb{Y}(k)]_1^{t-1} = [y_k]_1^{t-1}\big)\end{aligned} \tag{3}$$

The next simplifying assumption is that the conditional probability distribution of $\mathbb{Y}(t)$ given all states and all previous random observations is only a function of $\mathbb{S}(t)$ (and not of time). I considered also the case when the distribution of $\mathbb{Y}(t)$ depends on $\mathbb{S}(t)$ and $\mathbb{S}(t-1)$. However, though it added little to the computational complexity, it added significantly to the number of parameters to be estimated. Making use of the above conditions,

$$\begin{aligned}\Pr\big([\mathbb{S}(t)]_0^T = [s_t]_0^T &\text{ and } [\mathbb{Y}(t)]_1^T = [y_t]_1^T\big) = \\ \Pr(\mathbb{S}(0) = s_0)\cdot &\prod_{t=1}^{T} \Pr(\mathbb{S}(t) = s_t \,|\, \mathbb{S}(t-1) = s_{t-1})\cdot \\ &\prod_{t=1}^{T} \Pr(\mathbb{Y}(t) = y_t \,|\, \mathbb{S}(t) = s_t)\end{aligned} \tag{4}$$

To simplify notation, define

$$\begin{aligned} f(y \,|\, s) &\stackrel{\text{def}}{=} \Pr(\mathbb{Y}(t) = y \,|\, \mathbb{S}(t) = s), \\ \underline{s} &\stackrel{\text{def}}{=} [s_t]_0^T, \\ \underline{y} &\stackrel{\text{def}}{=} [y_t]_1^T, \text{ and} \\ p(\underline{s}, \underline{y}) &\stackrel{\text{def}}{=} \Pr([\mathbb{S}(t)]_0^T = \underline{s} \text{ and } [\mathbb{Y}(t)]_1^T = \underline{y}). \end{aligned}$$

With this notation we have

$$\Pr\big([\mathbb{S}(t)]_0^T = [s_t]_0^T \text{ and } [\mathbb{Y}(t)]_1^T = [y_t]_1^T\big) = p(\underline{s}, \underline{y}) = p_{s_0} \prod_{t=1}^{T} p_{s_{t-1}s_t} f(y_t \,|\, s_t) \tag{5}$$

This formula gives the probability of the atoms of the model, that is, those events that can not be subdivided into smaller events. The probability of any event describable in the model is the sum of the probability of the atoms.

Of course, the probabilities are functions of the parameters of the model, which we will denote by:

$$\lambda \stackrel{\text{def}}{=} \left\{ \begin{array}{l} p_s : 1 \le s \le M, \\ p_{s\sigma} : 1 \le s, \sigma \le M, \\ f(y \,|\, s) : 1 \le y \le Y, 1 \le s \le M \end{array} \right\}$$

and use λ as a function argument where appropriate.

Questions

What questions are of interest? One question is what is the probability that an T-tuple of $\mathbb{Y}(t)$ will be observed. This, of course, is a function of the parameters. The probability of an T-tuple of observations is just the sum over all state sequences of the probabilities of the corresponding atoms:

$$p(\underline{y}; \lambda) \stackrel{\text{def}}{=} \Pr\big([\mathbb{Y}(t)]_1^T = [y_t]_1^T; \lambda\big) = \sum_{\underline{s}} p(\underline{s}, \underline{y}; \lambda) \tag{6}$$

Then $p(\underline{y}; \lambda)$ is the probability of the observations, $\underline{y}$. It is also the likelihood function for λ given the observations, $\underline{y}$. A standard problem is to choose λ to maximize the likelihood function.

It is frequently the case that the random observables, $\{\mathbb{Y}(t) : t \in \mathcal{N}\}$, are observed for some period of time and it is desired to find some information about the state sequence from those observations. For example, given the event, $\{[\mathbb{Y}(t)]_1^T = [y_t]_1^T\}$, we may wish to find the probability distribution of the state at a specified time, τ. That is we wish to find, for a given sequence of observations, the probability distribution of s_τ given those observations:

$$Pr\big(\mathbb{S}(\tau) = s_\tau \,|\, [\mathbb{Y}(t)]_1^T = [y_t]_1^T\big)$$

Since,

$$Pr\big(\mathbb{S}(\tau) = s_\tau \,|\, [\mathbb{Y}(t)]_1^T = [y_t]_1^T\big) = \frac{Pr\big(\mathbb{S}(\tau) = s_\tau \text{ and } [\mathbb{Y}(t)]_1^T = [y_t]_1^T\big)}{Pr\big([\mathbb{Y}(t)]_1^T = [y_t]_1^T\big)}$$

the computation of the a posteriori probability is equivalent to computing the joint probability. Referring to equation (5), we see that this reduces to computing

$$\begin{aligned}\Gamma_\tau(s_\tau) &\stackrel{\text{def}}{=} Pr\big(\mathbb{S}(\tau) = s_\tau \text{ and } [\mathbb{Y}(t)]_1^T = [y_t]_1^T\big) \\ &= \sum_{[s_k]_0^{\tau-1}} \sum_{[s_k]_{\tau+1}^T} p_{s_0} \prod_{t=1}^{N} p_{s_{t-1}s_t} f(y_t \,|\, s_t)\end{aligned} \tag{7}$$

for each choice of s_τ. With the exception of s_τ, each indexed state is a summation variable and, with the exception of s_0, occurs in exactly two factors. It is easily deduced that the equation can be expressed in terms of the product of matrices.

However, there is a better (at least to me) approach to computing this probability. We begin with the joint probability of the T-tuple of observations and the state at time τ and apply recursive factorization where the first event is the set of observations up through time, τ, and the state at time τ.

$$\begin{aligned}&\Pr\big([\mathbb{Y}(t)]_1^T = [y_t]_1^T \text{ and } \mathbb{S}(\tau) = s_\tau\big)\\ &\quad = \Pr\big([\mathbb{Y}(t)]_1^\tau = [y_t]_1^\tau \text{ and } \mathbb{S}(\tau) = s_\tau\big)\cdot\\ &\qquad \Pr\big([\mathbb{Y}(t)]_{\tau+1}^T = [y_t]_{\tau+1}^T \,|\, [\mathbb{Y}(t)]_1^\tau = [y_t]_1^\tau \text{ and } \mathbb{S}(\tau) = s_\tau\big)\end{aligned}$$

Now Markovity of the state sequence implies that the probablility of $[\mathbb{S}_t]_{\tau+1}^T$ and therefore the probability of $[\mathbb{Y}_t]_{\tau+1}^T$ are independent of history prior to time τ. So the condition on the $\mathbb{Y}$ in the second term drop out and the factorization reduces to

$$\begin{aligned}&\Pr\big([\mathbb{Y}(t)]_1^T = [y_t]_1^T \text{ and } \mathbb{S}(\tau) = s_\tau\big)\\ &\quad = \Pr\big([\mathbb{Y}(t)]_1^\tau = [y_t]_1^\tau \text{ and } \mathbb{S}(\tau) = s_\tau\big)\cdot\\ &\qquad \Pr\big([\mathbb{Y}(t)]_{\tau+1}^T = [y_t]_{\tau+1}^T \,|\, \mathbb{S}(\tau) = s_\tau\big)\end{aligned}$$

We next address the problem of computing these factors,

$$\begin{aligned}\alpha_\tau(s) &\stackrel{\text{def}}{=} \Pr\big([\mathbb{Y}(t)]_1^\tau = [y_t]_1^\tau \text{ and } \mathbb{S}(\tau) = s\big),\\ \beta_\tau(s) &\stackrel{\text{def}}{=} \Pr\big([\mathbb{Y}(t)]_{\tau+1}^T = [y_t]_{\tau+1}^T \,|\, \mathbb{S}(\tau) = s\big)\end{aligned}$$

and

$$\Gamma_\tau(s) = \alpha_\tau(s)\cdot\beta_\tau(s)$$

I remark that

$$\sum_s \alpha_T(s) = \Pr\big([\mathbb{Y}(t)]_1^T = [y_t]_1^T\big) = p(\underline{y};\lambda).$$

Now recursive factoring of $\alpha_\tau(s)$ where the first factor is the observations up through time $\tau-1$ and the state at time $\tau-1$ gives

$$\begin{aligned}\alpha_\tau(s) &\equiv \Pr\big([\mathbb{Y}(t)]_1^\tau = [y_t]_1^\tau \text{ and } \mathbb{S}(\tau) = s\big)\\ &= \sum_\sigma \Pr\big([\mathbb{Y}(t)]_1^{\tau-1} = [y_t]_1^{\tau-1} \text{ and } \mathbb{S}(\tau-1) = \sigma\big)\cdot\\ &\Pr\big(\mathbb{Y}(\tau) = y_\tau \text{ and } \mathbb{S}(\tau) = s \,|\, [\mathbb{Y}(t)]_1^{\tau-1} = [y_t]_1^{\tau-1} \text{ and } \mathbb{S}(\tau-1) = \sigma\big)\end{aligned}$$

Again, Markovity implies that the condition, $[\mathbb{Y}(t)]_1^{\tau-1} = [y_t]_1^{\tau-1}$ can be dropped from the second factor:

$$\begin{aligned}\alpha_\tau(s) &= \sum_\sigma \Pr\big([\mathbb{Y}(t)]_1^{\tau-1} = [y_t]_1^{\tau-1} \text{ and } \mathbb{S}(\tau-1 = \sigma\big)\cdot\\ \Pr\big(\mathbb{Y}(\tau) &= y_\tau \text{ and } \mathbb{S}(\tau) = s \,|\, \mathbb{S}(\tau-1) = \sigma\big)\end{aligned}$$

The first factor is just $\alpha_{\tau-1}(\sigma)$ and the second factor is $p_{\sigma s}\cdot f(y_\tau|s)$ and above equation becomes the recursion:

$$\alpha_\tau(s) = \sum_\sigma \alpha_{\tau-1}(\sigma)p_{\sigma s}f(y_\tau \,|\, s) \tag{8}$$

Similarly, a reverse time recursion exists for $\beta_\tau(s)$:

$$\beta_\tau(s) = \sum_\sigma p_{s\sigma}\, f(y_{\tau+1} \,|\, \sigma)\beta_{\tau+1}(\sigma) \tag{9}$$

Finally we have

$$\begin{aligned}\Pr\big(\mathbb{S}(\tau) = s \text{ and } [\mathbb{Y}(t)]_1^T = [y_t]_1^T\big) &= \Gamma_\tau(s)\\ &= \alpha_\tau(s)\beta_\tau(s)\end{aligned}$$

and

$$\Pr(\mathbb{S}(\tau) = s \,|\, \underline{y}) = \frac{\alpha_\tau(s)\cdot\beta_\tau(s)}{\sum_\sigma \alpha_T(\sigma)}.$$

Once we have routines for computing α and β we can compute not only $\Pr(\mathbb{S}(t) = s \,|\, \underline{y})$ but also the a posteriori probability of other 'local' events, such as the event, $\{\mathbb{S}(t) = s \text{ and } \mathbb{S}(t+1) = \sigma\}$. In this case the expression is

$$\begin{aligned}&\Pr(\mathbb{S}(t-1) = s, \mathbb{S}(t) = \sigma \text{ and } [\mathbb{Y}(\tau)]_1^T = \underline{y})\\ &\qquad = \alpha_t(s)p_{s\sigma}\, f(y_{t+1} \,|\, \sigma)\beta_{t+1}(\sigma)\\ &\qquad \equiv \Gamma_t(s,\sigma)\end{aligned} \tag{10}$$

Improving on Estimates of Parameters

At this point Leonard Baum and I found that we had both been working independently on Hidden Markov Chains and had both come up with essentially the same calculation for a posteriori probabilities of 'local' events. At that point we joined forces.

Now, the above calculations were based upon specified parameter values. What if those parameter values did not adequately represent the phenomena under investigation? My thoughts proceeded as follows. While the parameters may not be accurate, the a posteriori probabilities may translate to better parameters.

For example, from the frequency interpretation of probability, if we could observe the state sequence over a long period of time and count the number of times the state, s, occurs, the frequency of occurrence should be approximately $\Pr(\mathbb{S}(t) = s)$ If the assumed parameters are correct, we will get p_s, where for a large enough period of time p will be the stationary distribution (the eigenvector with eigenvalue 1) of the transition matrix.

Furthermore, if the observation sequence, $[y_t]_1^T$, is a typical sequence in the Information Theoretic sense, that is, it has high probability using the parameters of the model, then the expected frequencies of states given the observations should also be approximately p_s, where p_s is the stationary distribution, not the initial distribution. Expressed in equation form:

$$\frac{\sum_{t=1}^T \Pr\big(\mathbb{S}(t) = s \,|\, [\mathbb{Y}(\tau)]_1^T = [y_\tau]_1^T\big)}{T} \approx p_s\,.$$

Similarly, from the frequency interpretation of probability, if we could observe the state sequence over a long period of time and count the number of times the state, s, occurs followed by σ, the frequency of occurrence should be approximately $\Pr(\mathbb{S}(t-1) = s$ and $\mathbb{S}(t) = \sigma)$. If the assumed parameters are correct, we will get $p_s\cdot p_{s\sigma}$.

Again if the observation sequence, $[y_t]_1^T$, is a typical sequence in

the Information Theoretic sense, then the expected frequencies of state transitions given the observations should also be approximately $p_s \cdot p_{s\sigma}$. Expressed in equation form:

$$\frac{\sum_{t=1}^{T} \Pr\big(\mathbb{S}(t-1)=s, \mathbb{S}(t)=\sigma \,|\, [\mathbb{Y}(\tau)]_1^T = [y_\tau]_1^T\big)}{T} \approx p_s \cdot p_{s\sigma}\ .$$

Finally, if we could observe the state sequence and the observation sequence and count the number of times $\mathbb{S}(t)=s$, $\mathbb{Y}(t)=y$, and the frequency should approximate $p_s f(y|s)$. Again, if the observed sequence is a typical sequence for the given parameters, the a posteriori expected frequencies should approximate $p_s f(y|s)$, i.e.,

$$\begin{aligned} p_s f(y\,|\,s) &\approx \frac{\sum_{t=1}^{T} \Pr\big(\mathbb{S}(t)=s, \mathbb{Y}(t)=y \,|\, [\mathbb{Y}(\tau)]_1^T = [y_\tau]_1^T\big)}{T} \\ &\approx \frac{\sum_{t\in\{t:y=y_t\}} \Pr\big(\mathbb{S}(t)=s \,|\, [\mathbb{Y}(\tau)]_1^T = [y_\tau]_1^T\big)}{T} \end{aligned}$$

My next thought was that if $\underline{y}$ was generated by a model with different parameter values and therefore not a typical sequence for the assumed values, the a posteriori frequencies, influenced by behavior of $[y_\tau]_1^T$, may be a better indication of the true parameters that the initial guess. So I replaced the parameter values by the expected frequencies and recomputed $p(\underline{y};\lambda')$ where

$$\lambda' \stackrel{\text{def}}{=} \left\{ \begin{array}{l} p_s(\lambda') = \frac{\sum_{t=1}^{T} \Pr\big(\mathbb{S}(t)=s \,|\, [\mathbb{Y}(\tau)]_1^T = [y_\tau]_1^T\big)}{T} \\ p_{s\sigma}(\lambda') = \frac{\sum_{t=1}^{T} \Pr\big(\mathbb{S}(t-1)=s, \mathbb{S}(t)=\sigma \,|\, [\mathbb{Y}(\tau)]_1^T = [y_\tau]_1^T\big)}{T} \\ f(y\,|\,s;\lambda') \leftarrow \frac{\sum_{t=1}^{T} \Pr\big(\mathbb{S}(t)=s, \mathbb{Y}(t)=y \,|\, [\mathbb{Y}(\tau)]_1^T = [y_\tau]_1^T\big)}{T p_s(\lambda')} \end{array} \right\} \quad (11)$$

I was please to find that $p(\underline{y};\lambda') > p(\underline{y};\lambda)$. In other words, this substitution increased the likelihood function! I tried this transformation on several Hidden Markov Models and the likelihood function always increased. Leonard Baum tried it on a number of examples and again the likelihood function always increased.

That is my contribution to the Baum-Welch `algorithm', the easy part. I tried to provide a mathematical proof that the likelihood always increases but I failed.

Baum, in cooperation with J. Eagon did the hard part by proving that this transformation either increases the likelihood function or leaves it constant. In the latter case, λ is a fixed point of the transformation. Their proof involved rather complex computations and applications of Hölder's inequality and the fact that the geometric mean is less than or equal to the arithmetic mean.

Later Baum, together with T. Petrie, G. Soules and N. Weiss, all at CRD/IDA at the time found a more elegant proof with the flavor of Information Theory which I will now discuss.

The Q Function

They began with the Kullback-Leibler divergence of two distributions:

$$D(p_1, p_2) \equiv \sum_{\omega} p_1(\omega) \log\left(\frac{p_1(\omega)}{p_2(\omega)}\right)$$

where p_1 and p_2 are two probability distributions on a discrete space and ω is summed over that space. The interpretation of this number is that for an experiment consisting of multiple selections from the distribution p_1, $D(p_1, p_2)$ is the expected log factor of the probability in favor of p_1 against p_2. It is an information theoretic measure and is known to be non-negative, equaling zero only when the $p_2(\omega) = p_1(\omega)$ for all ω for which $p_1(\omega) > 0$.

How does this apply to Hidden Markov Models? We let

$$p_1(s) = \frac{p(\underline{s}, \underline{y}; \lambda)}{p(\underline{y};\lambda)} \text{ and } p_2(s) = \frac{p(\underline{s}, \underline{y}; \lambda')}{p(\underline{y};\lambda')}.$$

Then p_1 and p_2 are distributions and

$$\begin{aligned} 0 \le D(\lambda,\lambda') &= \sum_s \frac{p(\underline{s}, \underline{y}; \lambda)}{p(\underline{y};\lambda)} \log\left(\frac{p(\underline{s}, \underline{y}; \lambda) p(\underline{y};\lambda')}{p(\underline{s}, \underline{y}; \lambda') p(\underline{y};\lambda)}\right) \\ &= \log\left(\frac{p(\underline{y};\lambda')}{p(\underline{y};\lambda)}\right) + \sum_s \frac{p(\underline{s}, \underline{y}; \lambda)}{p(\underline{y};\lambda)} \log\left(\frac{p(\underline{s}, \underline{y}; \lambda)}{p(\underline{s}, \underline{y}; \lambda')}\right). \end{aligned}$$

We simplify this by defining

$$Q(\lambda,\lambda') \equiv \sum_{\underline{s}} p(\underline{s}, \underline{y}; \lambda) \log(p(\underline{s}, \underline{y}; \lambda')).$$

Then

$$0 \le D(\lambda,\lambda') = \log\left(\frac{p(\underline{y};\lambda')}{p(\underline{y};\lambda)}\right) + \frac{Q(\lambda,\lambda) - Q(\lambda,\lambda')}{p(\underline{y};\lambda)} \quad (12)$$

and rearranging the inequality we have

$$\frac{Q(\lambda,\lambda') - Q(\lambda,\lambda)}{p(\underline{y};\lambda)} \le \log\left(\frac{p(\underline{y};\lambda')}{p(\underline{y};\lambda)}\right)$$

and this implies that if $Q(\lambda,\lambda') > Q(\lambda,\lambda)$ then $p(\lambda') > p(\lambda)$

Hill Climbing

We obtain a "hill climbing" algorithm by finding that λ' which maximizes $Q(\lambda,\lambda')$ as a function of its second argument. If $Q(\lambda,\lambda') > Q(\lambda,\lambda)$ then $p(\lambda') > p(\lambda)$ and we have succeeded in increasing $p(\lambda)$ which is the probability of the observations.

To maximize $Q(\lambda,\lambda')$ we begin by finding the critical points of Q as a function of λ' and subject to the stochastic constraints on the components of λ'. (A sample constraint is $\sum_j p_{ij} = 1$).

Before proceeding, let's manipulate the expression for Q. In equation (5) the expression for $p(\underline{s}, \underline{y}; \lambda)$ is a product, so its logarithm is a sum of log factors. Replacing λ by λ' in equation(5) and taking logarithms we have:

$$\begin{aligned} \log(p(\underline{s}, \underline{y}; \lambda')) &= \log(p_{s(0)}(\lambda')) \\ &\quad + \sum_{t=1}^{T} \log(p_{s(t-1)s(t)}(\lambda') f(y(t);\lambda' \,|\, s(t-1)s(t))) \end{aligned}$$

Substituting into the definition for Q gives:

$$Q(\lambda, \lambda') = \sum_{\underline{s}} p(\underline{s}, \underline{y}, \lambda) \log(p_{s(0)}(\lambda')) \quad (13)$$
$$+ \sum_{\underline{s}} p(\underline{s}, \underline{y}, \lambda) \sum_{t=1}^{T} \log(p_{s(t-1)s(t)}(\lambda') f(y(t); \lambda' \mid s(t-1)s(t))).$$

From this equation it can be seen that it is easy to differentiate with respect to the components of λ', add the appropriate Lagrange factors and solve. The result has already been displayed in equation (11).

Applications

There are too many papers published on applications to list here. In the area of speech recognition, here is a small sample:

F. Jelinek, L. Bahl, and R. Mercer, "Design of a linguistic statistical decoder for the recognition of continuous speech," *IEEE Trans. Inform. Theory*, vol. 21, May 1975.

L.R. Rabiner, "A tutorial on Hidden Markov models and selected applications in speech recognition", *Proceedings of the IEEE,* vol. 77, no. 2, Feb. 1989.

A. Poritz, "Linear predictive hidden Markov models and the speech signal," in *Proceedings of ICASSP '82,* May 1982.

EM Theory

In 1977, Dempster, Laird and Rubin collected a variety of maximum likelihood problems and methods of solving these problems that occurred in the literature. They found that all of these methods had some ideas in common and they named it the EM Algorithm, (standing for "Expectation, Maximization".)

The common problem is to maximize Prob(y; Φ), as a function of Φ where y is observed. (The probability of y is used in the case of discrete random variables and a density is maximized in the case of continuous random variables.) The observation, y, is viewed as "incomplete data" in the sense that there is a larger model containing "complete data" and y inherits its distribution by way of a mapping from the larger model to the observation model.

Mathematically: There is a probability space, X with a family of probability measures, $p(x; \Phi)$, and a mapping function, F with $F(x) = y$. The distribution, q, of y is

$$q(y; \Phi) = \sum_{\{x:F(x)=y\}} p(x; \Phi).$$

The conditional distribution of x given y is

$$p(x \mid y; \Phi) = \frac{p(x; \Phi)}{q(y; \Phi)},$$

provided $F(x) = y$.

Given a second value of Φ, say Φ' they define

$$Q(\Phi, \Phi') = \mathcal{E}\{\log(p(x; \Phi')) \mid y; \Phi\},$$

where $\mathcal{E}$ is the notation for the expected value function. This is the Expectation step. It gives a formula in Φ'. The Maximization step is to vary Φ' to maximize Q. In many problems the maximization step is easy and in many others it is as difficult as the original maximum likelihood problem. This leads to a "Generalized Estimation Maximization" which simply finds any Φ' which increases Q.

The Baum-Welch algorithm fits right into the EM scheme. x is $(\underline{s}, \underline{y})$ and the observation is y. The Q function is exactly the Q function that Baum *et al*. introduced to prove that the transformation increases the likelihood.

Recommended Reading

There are many papers published on these subjects. A few are:

L.E. Baum and J. Eagon "An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology," *Bulletin of the American Mathematical Soc.,* vol. 70, pp. 360–363.

L.E. Baum, T. Petrie, G. Soules and N. Weiss, "A maximization technique occurring in the statictical analysis of probabilistic functions of Markov chains," *Ann. Math. Stat.,* vol. 41, 1970.

A. Dempster, N. Laird and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, B, vol. 39, 1977.

A paper which extends theory to observations with continuous distributions.

L. Liporace, "Maximum likelihood estimates for multivariate observations Markov sources," *IEEE Trans. Inform. Theory*, vol. 28, Sept. 1982.

2003 Fall Workshop on Information Theory & Communications

Guu-Chang Yang and Shiann-Shiun Jeng

August 26-27, 2003
Hualien, Taiwan

The 2003 Fall Workshop on Information Theory and Communications was held August 26-27, 2003 in Hualien, Taiwan. More than 200 participants enjoyed the gorgeous views from the most beautiful city of Taiwan east coast. This biannual workshop was part of a series of workshops being organized under the auspices of IEEE Information Theory Society Taipei Chapter and IEEE Communications Theory Taipei Chapter. The technical program was arranged by Shiann-Shiun Jeng and Po-Hao Chang (National Dong Hwa University). The workshop was generously sponsored by the Communication Engineer Program of National Science Council, Ministry of Education and The National Dong Hwa University.

The main purpose of the biannual workshop is to provide an opportunity for professors and graduate students from different universities in the fields of Information Theory and Communications to meet and learn about each other's ongoing research activities. The workshop consisted of 9 invited presentations: Nen-Fu Huang (NTHU) talked about IPv6 R&D Division, Chin-Liang Wang (NTHU) talked about Power Control with Turbo Partial Parallel Interference Cancellation for DS-CDMA Communications, Chia-Chi Huang (NCTU) talked about A Complementary Code Pilot Based Transmitter Diversity Technique for OFDM Systems, Wern-Ho Sheen (NCTU) talked about B3G/4G: Technology Trends and International Activities,

Sze-Lin Su (NCKU) talked about Introduction to Ultra Wide Band System, Ruey-Yi Wei (NCU) talked about Noncoherent Block Coded Modulation, Chih-Peng Li (NSYSU) talked about Collision Detection Based Back-Off Schemes for Collision Resolution, Tsung-Hsien Liu (NCCU) talked about Application of Array Signal Processing Techniques to the Design of Rake Receiver, and Hsuan-Jung Su (NTU) talked about Dirty Paper Coding and Its Applications.

Besides the invited talks in the day time of August 26 & 27, there was one meeting for IEEE members hosted by Jean-Lien Chen (Director of the Communication Engineer Program of National Science Council), Wern-Ho Sheen (Chairman of IEEE Communications Society Taipei Chapter) and Guu-Chang Yang (Chairman of IEEE Information Theory Society Taipei Chapter) on the night of August 26. The theoretic and industrial aspects of communications researches were discussed deeply during the meeting. In addition, there were two posters sessions with over 50 posters held on August 26. The participants had a lot of discussions during the period of the poster sessions.

The workshop proceedings are available in hardcopy and on CD-ROM, and can be obtained by writng to Shiann-Shiun Jeng, ssjeng@mail.ndhu.edu.tw.

From left to right, the photo includes Po-Hao Chang (NDHU), Chung-Chin Lu (NTHU), Chong-Yung Chi (NTHU), Char-Dir Chung (NCU), Shiann-Shiun Jeng (NDHU), Szu-Lin Su (NCKU), Guu-Chang Yang (NCHU), Mao-Chao Lin (NTU), Chi-Chao Chao (NTHU), Jean-Lien Chen (NTUST), Wern-Ho Sheen (NCTU, the 12th from the left), Ying Li (YZU, the 13th from the left) and Chin-Liang Wang (NTHU, the 15th from the left). The others are our IEEE members.

Information Theory Society Board of Governors Meeting

Pacifico Conference Center, Yokohama, Japan, June 29, 2003

Mehul Motani

Attendees: John Anderson, Daniel Costello, Thomas Cover, Michelle Effros, Tony Ephremides, Ivan Fair, Tom Fuja, Marc Fossorier, Joachim Hagenauer, Chris Heegard, Hideki Imai, Torleiv Klove, Ralf Koetter, Ioannis Kontoyiannis, Ryuji Kohno, Steven McLaughlin, Mehul Motani, Paul Siegel, David Tse, Alexendar Vardy, Han Vinck, Sriram Viswanath.

The meeting was called to order at 10:00 AM by Society President Han Vinck. The members of the Board were welcomed and introduced themselves.

1. The agenda was approved and distributed.

2. The minutes of the previous meeting in Paris, France on March 31, 2003 were approved as distributed.

3. The President began by reporting on the IEEE TAB meeting which was held earlier in the year. He reported that IEEE was very concerned about its financial situation. In 2002, IEEE suffered a $16M deficit.

It was reported that due to redistribution of the IEL and ASPP income, the Society would be getting $100K more based on content (the society has the largest journal offering) and downloading statistics. The Society will also be repaid 20% of the investment it has made in the digital library.

The President also reported that the Society was successful in participating in IEEE awards and suggested that more Society members be nominated for such awards.

4. The Awards Committee report was presented by Hideki Imai.

The Awards Committee has nominated Bob Gallager for the 2003 "Cristofore Colombo" International Communications Award.

It was reported that the IT Society Members on the Joint IT/ComSoc Paper Award Committee and the chair of the IT Society Awards Committee selected one paper out of three nominations from members of the Awards Committee and submitted it to the Joint IT/ComSoc Paper Award Committee on April 16. There was no nomination from ComSoc. On June 18, the Joint IT/ComSoc Paper Award Committee decided to award the single nominated paper the Joint IT/ComSoc Best Paper Award. The winning paper is: S. Shamai (Shitz) and I. Sason, "Variations on the Gallager bounds, connections and applications," IEEE Transactions on Information Theory, Vol. 48, No. 12, pp. 3029-3051, December 2002.

Action Item: It was agreed by the Board that the Joint IT/ComSoc Paper Award procedures be added to the Bylaws. This needs to be in conjunction with ComSoc.

It was reported that the total number of nominations for the IT Society Paper Award for the period 2001-2002 was 10. Three papers shown survived the last round of voting of the Award Committee:

- Lizhong Zheng and David N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel", IEEE Transactions Information Theory, vol. IT-48, No. 2, pp. 359 - 383, February 2002.

- Vladimir I. Levenshtein, "Efficient reconstruction of sequences," IEEE Transactions on Information Theory, vol. IT-47, no. 1, pp. 2-22, Jan. 2001.

- S. Verdu, "Spectral efficiency in the wideband regime," IEEE Transactions on Information Theory, special issue in memory of A. Wyner, on "Shannon theory: perspective, trends and applications", vol. 48, no. 6, pp. 1319-1343, June 2002.

According to the bylaws, the Board shall vote for the nominees by ballot, conducted by the Society President or designee, at the first Board Meeting following June 1st of the award year. The President informed the Board that several member of the Board could not attend the meeting due to visa problems. The Board voted unanimously to delay the vote and conduct an email ballot by August 1.

Action Item: The Board voted unanimously to revisit the Bylaws with respect to voting for the IT Paper Award at the BOG meeting.

5. The membership report was presented by Steven McLaughlin. He reported that the Society membership has dropped by 10% since last year (at the same time of the year). Since IEEE has seen a similar drop in its membership, it was suggested that this is the reason for the Society's drop. Steven also reported that the Society was participating in an IEEE wide questionnaire dealing with some membership issues.

Steven also noted that the Society chapter luncheon would be held on Thursday and invited any interested Board members to attend.

Action Item: The Board requested that Steven report more details on the 10% drop in Society membership at the next Board meeting. Steven reported that there were 100 ISIT attendees who were not IT Society members and suggested contacting them regarding membership.

Action Item: The Board requested that Steven, in cooperation with Michelle Effros, present a proposal at the next meeting addressing membership issues including ideas and suggestion for increasing Society membership.

6. The treasurer's report prepared by Marc Fossorier was distrib-

uted and discussed. The net worth (defined as total assets - deferred income) of the Society was $541K.

Action Item: It was noted that the finance figures for 2003 were not available from IEEE. The Board requested that Marc report the finances to the Board when the detailed figures are available.

7. There was nothing to report with respect to the IT Society Newsletter.

8. There was nothing to report with respect to the IT Society Website.

9. Matters related to Symposia and Workshops

(a) The President reported on a request by IEEE-SA section to financially co-sponsor ISIT 2005 to be held in Adelaide, Australia. The Board voted unanimously not to accept financial co-sponsorship.

(b) Tom Cover presented a proposal by Gadiel Seroussi and Andrea Goldsmith for ISIT 2006 in San Francisco. Tony Ephremides also noted that there was interest to hold ISIT 2006 in Seattle. The Board requested proposals with financial details from these two parties and will consider them at the next meeting.

(c) Ryuji Kohno reported on ISIT 2003. Several issues, including the SARS situation, the technical program, financial aid, and budget, were discussed.

(d) Dan Costello gave an update on ISIT 2004. He noted that the preparations were on schedule, the CFP has been distributed, and copies of the poster are to be displayed. He also noted that the registration fee has been raised to $475. The 2004 symposium will be held June 27 - July 2, 2004 at the downtown Marriott in Chicago.

(e) The 2004 ITW 2004 workshop to be held in San Antonio was discussed. The current proposal is to hold the workshop in the October/November timeframe. It was suggested to hold the workshop in December instead, so as not to conflict with the Allerton conference. The matter will be discussed at the next meeting.

(f) There was nothing new to report regarding ISIT 2005.

(g) The Board approved technical co-sponsorship of ISITA 2004 to be held in Italy in October.

(h) The matter of ISIT 2007 in Germany was discussed.

(i) The Board approved technical co-sponsorship of WiOpt to be held in April 2004 in Cambridge.

(j) It was suggested by the President that the Awards luncheon of the annual symposium be free for members of the IT Society only.

Action Item: It was noted that the IEEE-SA section should be notified of the Board's decision not to accept financial co-sponsorship of ISIT 2005. It was also noted that ISITA 2004 and WiOpt organizers should be notified of the Board's positive decision on technical co-sponsorship.

10. The new CD-DVD initiative was discussed by Steven McLaughlin. The subject is the CD/DVD containing electronic copies of the IT Transactions. The main issue is whether the Society should aim to make money from the sale of the CD/DVDs or should just give it away to all members. Two proposals arose from these discussions. The first is for the Society to update, produce and sell these CD/DVDs to interested parties. The second proposal, presented by Chris Heegard, is to give them free to all members and sell to others. Chris also requested that the CDs be updated with the papers published after 1998. It was noted that this will add to the benefits of IT Society membership and can be used as an incentive to encourage people to join the Society.

Action Item: The Board requested Steven to study the CD-DVD initiative issue and present a proposal to the Board at the next meeting in October.

11. The Board unanimously approved the Distinguished Service Award, which was discussed at the previous Board meeting in October 2002.

12. A report on the Transactions on Information Theory was distributed by Paul Siegel, the Editor-in-Chief.

It was reported that the Transactions continue to have no backlog and this policy would be lifted only in certain circumstances. The Board approved the following editorial appointments:

Communications - Babak Hassibi (CalTech) , new position, effective August 1, 2003.

Coding Techniques - Oyvind Ytrehus (University of Bergen), new position, effective July 1, 2003.

Coding Techniques - Marc Fossorier (University of Hawaii at Manoa), replacing Rudiger Urbanke, effective October 1, 2003.

Coding Theory - Gilles Zemor (ENST), replacing Jorn Justesen, effective August 1, 2003.

Coding Theory - Bob McEliece (CalTech), replacing Ralf Koetter, effective September 1, 2003.

The issue of time to publication was discussed and a report, showing the average number of weeks from submission to publication for regular papers and correspondence articles published in issues of the Transactions from July 1998 to the present, was distributed.

With respect to the time to publication, the role of the associate editors (AE) was discussed. It was noted that promptness of the AE in handling papers was critical and that AE are currently sent reminders. It was also agreed that the Editor-in-Chief have the authority to remove an AE if necessary.

Electronic publishing issues were discussed and a report comparing electronic publishing software products, including ScholarOne's Manuscript Central was distributed.

The Board approved an ad hoc committee to look at publication issues, including improved notifications and acknowledgements to authors who have submitted manuscripts for publication. The committee members include Paul Siegel, Ralf Koetter, John

Anderson, Tony Ephremides, and Alex Vardy.

Action Item: The Board requested that the ad-hoc committee looking at publications issues write a detailed task description and report at the next meeting in October.

13. Ivan Fair discussed the report of the education subcommittee. The President noted that members should lend their full support to the efforts of the education subcommittee. He suggested that anyone with input should email Ivan with their comments.

Action Item: The Board requested that the Education committee present a more detailed report on their activities at the next meeting in October.

14. Joachim Hagenaeur reported on the nominations for the Board. Six outgoing members have agreed to re-run. Six others have already been nominated. In addition, Alex Barg was nominated. The Board closed and unanimously approved these nominations for the Board.

The Board unanimously approved and closed the following nominations:

Two nominations for 2nd Vice-president - Dave Neuhoff and Bixio Rimoldi.

One nomination for 1st Vice-president - Steven McLaughlin

One nomination for President - Hideki Imai

15. The floor was then opened up for all other business.

Chris Heegard raised an issue concerning the NSF that could have both good and bad implications for the Society. He noted that Julia Abrams, who handles most of the IT grant proposals, was leaving NSF. The IT Society should recruit someone to serve in NSF to protect the interests of Communications, Signal Processing and IT. He suggested that a group of IT Society members should go to Washington and discuss the matter with CISE.

The Board approved the formation of an adhoc committee to liason with NSF. The committee includes Chris Heegard, Tony Ephremides, Tom Fuja, Dave Forney, and Bruce Hajek.

Action Item: The Board requested that the adhoc committee tasked to liason with the NSF report on their activities at the next meeting in October.

The President, Han Vinck, noted that IT Society Board members plan to meet with SITA, the Japanese IT Society, to have discussions and improve relations.

Action Item: The Board requested that Han report the outcome of the SITA meeting at the next meeting in October.

16. The President announced that the next Board meeting will be held on October 1, 2003 at the Allerton conference. Further details will be provided at a later date.

17. The meeting was adjourned at 2:15 PM.

Management Science
Call for papers: Special Issue on Open Source Software

Guest Editors

Eric von Hippel
Sloan School of Management
Massachusetts Institute of Technology

Georg von Krogh
Institute of Management
University of St. Gallen

Submission deadline:
September 1, 2004

Free and open-source software (OSS) and the processes used to develop it are very unusual in a number of ways. Such software is developed by individuals and firms that freely reveal the code that they write, releasing it under licenses that allow others to use it, to modify it as they wish, and distribute their modifications to others. Project management methods used in open source software projects have proven capable of converting the volunteer efforts of users and programmers who come and go as they please into high quality software products - some very large and complex. The efforts of these volunteers are coordinated with the aid of only very simple but powerful collaborative tools. Thousands of such projects exist today, and the code they develop and freely reveal can offer value to many. (Well-known examples are the GNU/Linux computer operating system, Apache server software and the Perl programming language.)

Researchers in the social sciences have become aware of the very interesting characteristics of F/OSS and its development system, and a lively, interdisciplinary research community has sprung up to explore the phenomenon and its implications for management. The aim of this special issue is to provide an appropriate forum for some of the very exciting and high quality research now being done on OSS. The outstanding quality and interdisciplinary scope of Management Science makes it an ideal journal for this purpose.

We wish to encourage authors rooted in multiple disciplines, such as sociology, economics, social psychology, organization behavior, information systems, innovation process research and strategic management to consider submitting their work to this special issue. We welcome both theoretical research and empirical research using quantitative or qualitative methods. All articles should, in line with the editorial policy of Management Science, have clear relevance to management practice.

Below we offer a sample listing of the broad range of OSS-related topics that are appropriate for this special issue.

- How do OSS development projects originate and how do they "work?" What are the incentives to contribute to OSS projects? How do volunteers find, join and leave projects? How are the activities distributed among contributors and how is the work of many volunteer contributors coordinated or interlinked? How are resources provided and allocated? What is the nature of leadership in OSS projects?
- How do OSS practices compare with conventional product development practices with respect to both processes used and outcomes attained - such as relative development cost, performance and quality of software developed.
- Intellectual property-related issues: What factor impact on the choice of OSS licenses? How effective are the various OSS licensing arrangements? What are lessons from OSS for the protection of digital intellectual property and the intellectual commons more generally?
- Why do some OSS projects succeed and some fail? What characterizes the competition for talent and expertise between OSS projects, and between OSS projects and proprietary software projects?
- What are the competitive dynamics introduced to markets by OSS? What are the strategic implications: How do firms compete with or collaborate with the outputs of OSS projects? Do these lessons apply generally across other fields of technological innovation?

Authors should prepare their manuscripts according to Management Science's guidelines for authors. Every submission will be reviewed according to the single-blind review process of Management Science. The anticipated publication date of the special issue will be early 2006.

Papers should be submitted electronically at informs.manuscriptcentral.com. On the first screen of the submission process, authors should select "Special Issue" as their manuscript type and designate Georg von Krogh as the Department Editor. Questions should be directed to:

Professor Georg von Krogh
University of St.Gallen
Institute of Management
Dufourstrasse 48
CH-9010 St.Gallen
Switzerland
georg.vonkrogh@unisg.ch

CALL FOR PAPERS

The 2004 IEEE International Symposium on Information Theory will be held in Chicago, Illinois, from Sunday, June 27, through Friday, July 2, 2004. The theme of ISIT 2004, "Exploring New Connections," represents a focus on fostering new connections among people, technical areas and ideas, both within the traditional boundaries of Information Theory, and beyond in related fields. Keynote speakers for ISIT 2004 will be Persi Diaconis, Ueli Maurer, Thomas J. Richardson and Martin Vetterli.

Previously unpublished contributions to the following areas will be solicited:

- Coded modulation
- Communication complexity
- Cryptology and data security
- Data networks
- Information theory and statistics
- Multiuser information theory
- Quantum information processing
- Signal processing
- Coding theory and practice
- Communication systems
- Data compression
- Detection and estimation
- Multiuser detection
- Pattern recognition and learning
- Shannon theory
- Source coding

The following tutorials will be offered on Sunday, June 27:

Gilles Brassard: Quantum Information Processing
Michael Fitz, Giuseppe Caire, Hesham El-Gamal: Space-Time Coding
Brendan Frey: Probabilistic Inference Algorithms and Applications
Ueli Maurer: Cryptography

The conference site is the Chicago Downtown Marriott Hotel, located on the "Magnificent Mile" of Michigan Avenue, near the Chicago river and lake front.

Papers will be reviewed on the basis of an extended abstract (not exceeding six pages) of sufficient detail to permit reasonable evaluation. The deadline for submission is **December 1, 2003**, with notification of decisions by March 15, 2004. The deadline will be strictly enforced. In view of the large number of submissions expected, multiple submissions by the same author will receive especially stringent scrutiny. All accepted papers will be allowed twenty minutes for presentation, and one-page abstracts will be printed in the conference proceedings. Authors are strongly encouraged to submit electronic versions of their summaries in the form of Portable Document Format (PDF) files. Detailed information on paper submission, the technical program, special events, tutorial sessions, accommodations, travel arrangements, excursions and applications for travel grants will be posted on the Symposium web site:

`http://www.isit2004.org`

Inquiries on general matters related to the Symposium should be addressed to `chair@isit2004.org`.

ISITA 2004

Parma, Italy, October 10-13, 2004

FIRST CALL FOR PAPERS

The 2004 International Symposium on Information Theory and its Applications, sponsored by the Society of Information Theory and its Applications (SITA) with the technical co-sponsorship of the IEEE Information Theory Society, will be held in Parma, Italy, from Sunday, October 10, through Wednesday, October 13, 2004.

Topics of interest include, but are not limited to, the following:

- Error Control Coding
- Coded Modulation
- Communication Systems
- Detection and Estimation
- Spread Spectrum Systems
- Signal Processing
- Rate-distortion Theory
- Stochastic Processes
- Data Networks
- Multiuser Information Theory
- Coding Theory and Practice
- Data Compression and Source Coding
- Optical Communications
- Mobile Communications
- Pattern Recognition and Learning
- Speech/Image Coding
- Shannon Theory
- Cryptology and Data Security
- Applications of Information Theory
- Quantum Information Processing

Papers will be selected on the basis of an extended summary (not exceeding 3 pages). The deadline for submission is March 26, 2004. Notification of decisions will be made by June 14, 2004.

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

www.sita.gr.jp/ISITA2004

Inquiries on matters related to the Symposium should be addressed as follows:

General matters:
isita2004@sita.gr.jp

Technical program matters:
isita2004tpc@sita.gr.jp

Deadline for the submission of extended summary	**March 26, 2004**
Notification of paper acceptance	**June 14, 2004**
Deadline for final paper submission	**July 16, 2004**
Deadline for author registration	**July 16, 2004**

Symposium Committees

General Chairs
Ezio Biglieri (Politecnico di Torino)
Kazuhiro Nakamura (Chiba Univ.)

General Secretaries
Alessandro Nordio (Politecnico di Torino)
Kenichi Yamazaki (Tamagawa Univ.)

Finance Chairs
Kenji Nakagawa (Nagaoka Univ. Techn.)
Giorgio Taricco (Politecnico di Torino)

Finance Secretary
Jun Muramatsu (NTT)

Publications
Tomoharu Shibuya (Nat. Inst. Multimedia Edu.)
Emanuele Viterbo (Politecnico di Torino)

Organizing Secretariat
STILEMA (Italy)

Registration
Hiroki Koga (University of Tsukuba)

Local Arrangements
Tadashi Fujino (Univ. Electro-Commun.)

Publicity
João Barros (Munich Univ. of Technology)
Ikuo Oka (Osaka City University)
Mamoru Kobayashi (web) (Shonan Inst. Tech.)

Technical Program Committee

Chairs
Umberto Mengali (Università di Pisa)
Hirosuke Yamamoto (University of Tokyo)

Vice Chair
Hiroshi Kamabe (Gifu University)

Secretary
Mitsuharu Arimura (Univ. Electro-Commun.)

International Advisory Committee

Chair
Hideki Imai (Univ. of Tokyo)

Princeton University

Call for Papers CISS '04

38th Annual Conference on Information Sciences and Systems

Department of Electrical Engineering, Princeton University

March 17, 18, and 19, 2004

PLENARY SPEAKERS

Professor John Doyle (Caltech): *Systems and Control*
Professor Robert Gallager (MIT): *Communication and Networking*
Professor Martin Vetterli (EPFL and Berkeley): *Signal and Image Processing*

Authors are invited to submit previously unpublished papers describing advances, applications, and ideas in the fields of communication, networking and information theory; signal, image, and video processing; systems and control; learning and statistical inference.

Two types of contributed papers are solicited:

- ***Regular papers,*** requiring approximately 30 minutes for presentation; these will be reproduced in full (up to six pages) in the conference proceedings.
- ***Short papers***, suitable for presentation in approximately 15 minutes; one-page summaries of these papers will be included in the proceedings.

Electronic summaries in Adobe PDF format, together with a "regular" or "short" designation and 2-3 keywords must be submitted by **January 2, 2004,** through the conference website http://www.ciss.us. Summaries should be of sufficient detail and length to permit careful reviewing. Authors will be notified of acceptance no later than **January 30, 2004.** Final manuscripts of accepted papers are to be submitted in PDF format no later than **February 23, 2004. These are firm deadlines that will permit the distribution of a CD containing the conference proceedings at the conference.**

PROGRAM DIRECTORS:
Sanjeev Kulkarni
Mung Chiang
Dept. of Electrical Engineering
Princeton University

CONFERENCE WEBSITE:
http://www.ciss.us

CONFERENCE COORDINATOR:
Cindy Menkes
Phone: (609) 258-0104
E-mail: cmenkes@ee.princeton.edu

CONFERENCE OFFICE:
CISS 2004
Dept. of Electrical Engineering
Princeton University
E-Quad, B317
Princeton, NJ 08544

Phone: (609) 258-0104
Fax: (609) 258-2158
E-mail: ciss@ciss.us

www.ciss.us

IMPORTANT DATES:
Submission deadline:
January 2, 2004

Notification of Acceptance:
January 30, 2004

Final manuscript and advance registration:
Before February 23, 2004

Conference Dates:
March 17, 18, 19, 2004

CONFERENCE ANNOUNCEMENT
2005 IEEE International Symposium on Information Theory

Adelaide Convention Centre, Adelaide, Australia
September 4 – 9, 2005

isit 05 Adelaide

The 2005 IEEE International Symposium on Information Theory will be held at the Adelaide Convention Centre in Adelaide, Australia from Sunday, September 4 through Friday September 9, 2005.

Previously unpublished contributions to the following areas will be solicited in the forthcoming call for papers:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication systems
- Cryptology and data security
- Data compression
- Data networks
- Detection and estimation
- Information theory and statistics
- Multiuser detection
- Multiuser information theory
- Pattern recognition and learning
- Quantum information processing
- Shannon theory
- Signal processing
- Source coding

The first call for papers will be published in March 2004.

Detailed information on the submission deadline, technical program, accommodation, travel, and excursions will be posted on the Symposium web site

`http://www.isit2005.org`

Inquiries on general matters related to the Symposium should be directed to

Dr. Alex Grant
Institute for
Telecommunications Research
University of South Australia
SA 5095 Australia
`alex.grant@unisa.edu.au`

Prof. Rodney A. Kennedy
Research School of Information
Sciences and Engineering
Australian National University
ACT 0200 Australia
`rodney.kennedy@anu.edu.au`

Conference Calendar

DATE	CONFERENCE	LOCATION	CONTACT/INFORMATION	DUE DATE
January 14-16, 2004	**5th International ITG Conference on Source and Channel Coding**	Fraunhofer Institute for Integrated Circuits, Erlangen, Germany	Prof. Dr.-Ing. J. Huber (Email: scc04@LNT.de) http://www.LNT.de.itg/	July 21, 2003
March 17-19, 2004	**38th Annual Conference on Information Sciences and Systems (CISS '04)**	Department of Electrial Engineering Princeton University Princeton, New Jersey USA	See CFP in this issue. ciss@ciss.us http://www.ciss.us	Jan. 2, 2004
June 27 - July 2, 2004	**2004 IEEE International Symposium on Information Theory (ISIT)**	Chicago Downtown Marriot Chicago, Illinois, USA	See CFP in this issue. chair@isit2004.org http://www.isit2004.org	Dec. 1, 2003
July 19-24, 2004	**2004 Stochastic Networks Conference**	Centre de Recherches Mathematiques Universite de Montreal Montreal, Canada	http://www.stanford.edu/group/ stochnetconf/	
June 20-24, 2004	**2004 International Conference on Communications (ICC)**	Paris, France	http://www.icc2004.org	Sept. 1, 2003
October 10-12, 2004	**2004 International Symposium on Information Theory and its Applications (ISITA 2004)**	Parma, Italy	See CFP in this issue. isita2004@sita.gr.jp http://www.sita.gr.jp/ISITA2004/new.htm	March 26, 2004
November 29- December 3, 2004	**GLOBECOM 2004**	Hyatt Regency Dallas at Reunion Hotel Dallas, Texas, USA	http://www.globecom2004.org	March 1, 2004
TBA (Fall 2004)	**2004 IEEE Information Theory Workshop (ITW)**	San Antonio, Texas, USA	TBA	TBA
TBA (before ISIT 2005)	**2005 Information Theory Workshop (ITW)**	New Zealand	TBA	TBA
September 4-9 2005	**2005 IEEE International Symposium on Information Theory (ISIT)**	Adelaide Convention Center Adelaide, AUSTRALIA	See CFP in this issue. http://www.isit2005.org Dr. Alex Grant Institute for Telecommunications Research University of South Australia SA 5095 Australia Prof. Rodney A. Kennedy Research School of Information Sciences and Engineering Australian National University ACT 0200 Australia rodney.kennedy@anu.edu.au	TBA
TBA	**2006 IEEE International Symposium on Information Theory (ISIT)**	Seattle, Washington, USA	TBA	TBA

IEEE Information Theory Society Newsletter
445 Hoes Lane, P. O. Box 1331
Piscataway, NJ 08855-1331 USA