

teaching it ... to freshmen

Lalitha Sankar and H. Vincent Poor, lalitha.poor@princeton.edu

Each year, Princeton University's Council on Science and Technology (CST) grants five to six three-year teaching postdoctoral fellowships distributed over all University departments in engineering and the natural sciences, following a process of applications, interviews, and with support from the applicants' research and teaching faculty mentors. The Council's aim is to fund research fellows who devote a semester each year to teaching science and technology courses to undergraduates, preferably those with non-science backgrounds, and to this end, the Council expects each applicant to propose such a course with his/her teaching mentor.

As a CST postdoctoral fellow and her mentor for the academic period 2007-2010, we proposed a course entitled 'Fundamental Ideas of the Digital Revolution: Insights into Technology, Language, and Biology,' with the aim of introducing the fundamental concepts of entropy, compression, and coding of Shannon, which of course, are at the heart of the modern digital revolution. We offered the course as a Freshman Seminar in the 2008 spring semester and, in keeping with the small and diverse class sizes for such seminars, had an enrollment of six first year undergraduates with majors and interests as varied as anthropology, linguistics, economics, life sciences, political science, and mechanical engineering.

Our primary approach to teaching the course was to motivate the need for Shannon's basic ideas using examples. A natural first step towards this was to develop the notion of information. To this end, we owe a debt of gratitude to our Princeton colleague Sergio Verdú for introducing us to the book *Information: A New Language of Science*, by Hans Christian von Baeyer [1], which we used as one of the textbooks for the class (and we also owe Sergio for the title of this article!). A professor of physics, von Baeyer introduces readers to the concept of information starting with a semantic history of the word and culminating in a discussion of quantum information. In the process, he discusses the contributions made by Morse to the first electrical communication network and by Shannon to information quantification and communication, introduces the readers to probability and the logarithm as tools for measuring information, and makes elegant connections between the notions of entropy in information theory and in thermodynamics. Using this reference we developed the fundamental idea that information (Greek root, *eidōs* or form) exchange is meaningful only when it is not known *a priori* at the receiver. This led to the notion that meaningful sources of information are not deterministic but rather are random entities that generate letters from an alphabet with different probabilities. Here it was natural to use spoken and written languages as the primary examples.

To further develop the notion of measuring information measure in bits, we introduced binary representation through an interactive lesson. This then allowed us to define entropy, having developed a simple mathematical representation of a random variable taking finitely many values, and to introduce Huffman coding as an approach to compress data for sources with known statistics. To complement these abstract discussions, we developed hands-

on experiments using diode-based light switches, courtesy of the optoelectronics group at Princeton, to experiment with signaling, coding, and communication schemes. The different colored diodes allowed binary, ternary, and quaternary signaling of English letters given their prior probabilities. In fact, these hands-on experiments even helped the students understand the nuanced difference between top-down Shannon-Fano coding and bottom-up Huffman coding techniques.

This experiment in turn naturally led to the concept of compression. Using examples, we developed the idea that information sources can be either continuous or discrete. The need to store and communicate information using finite resources served as a motivation to introduce the concepts of sampling and quantization as *pulse-coded modulation* (PCM) [2]. With knowledge that the electronic alphabet is binary, the power of PCM was illustrated with practical examples such as telephony, where speech signals are sampled and quantized to 64 kbps, and compact discs, where audio signals are sampled and quantized at about 1.5 Mbps. With a brief nod to Nyquist, the discussion of PCM only hastened to reveal the need for compression to store and transmit multimedia information sources. To this end, we briefly introduced the ideas behind a variety of lossy compression techniques such as differential PCM, Linear Predictive Coding, JPEG, MPEG, and sub-band coding schemes such as mp3. In the process of doing so, it was gratifying to help students who could not distinguish between the audio quality of mp3 and CD understand the technical difference. Finally, we also briefly introduced lossless dictionary-based compression schemes such as Lempel-Ziv as the basis for compression formats such as Zip™.

In our desire to use natural language as an example to illustrate Shannon's novel ideas, we followed the discussion on information compression by a reading and discussion of Shannon's paper on the compressive and predictive properties of the English language [3]. While the students intuitively grasped the concept of correlation between consecutive letters in a word/sentence as described by Shannon, developing the mathematical concept of joint probability distribution to quantify the entropy of the language using N-grams was extremely challenging. It was however a worthwhile investment as it made it easier to formally define mutual information later in the course. These ideas were also applied to study the classical hypothesis of information coding and compression in animal nervous systems as proposed and discussed by Atick in [4].

We set the stage for a discussion of error correction and coding using hands-on experiments to provide insight into the unreliable nature of communications. Our task of presenting the simple principles that ensure reliability in modern communications systems was made significantly easier thanks to Bob McEliece of Caltech, who shared with us his multimedia productions aimed at teaching these concepts to high-school students. The video presentations allowed the students both to see the wide range of applications of error-control coding, from bar codes to deep space communications, and to get a basic understanding of the mathemati-

cal principles behind the Reed-Solomon code. Our thanks also go to our Princeton colleague Sanjeev Kulkarni for sharing his undergraduate course notes on this topic. As an example, we discussed the error-correction scheme applied to the PCM bits stored on a compact disc. To this end, we owe thanks to Kees Immink of Turing Machines, Inc. for his encouragement and feedback, including his recent article in the IT newsletter [5]. As an historical example of error-correction, the class (including the instructors) learned about the sophisticated error-detection and correction schemes used by the ancient Indians in their oral tradition of chanting the Vedas (religious texts), which among other approaches analogous to modern coding techniques also included a text-to-binary² mapping technique to aid one of the chanters in the group to detect errors [6]. Finally, having developed the notion of noise as an essential nuisance [1] and the need for error correction, we discussed, albeit briefly, the contribution of Shannon to quantify the maximum rate for reliable communications, thereby introducing the concept of mutual information.

In the last part of the course we used a combination of guest lectures, multimedia presentations, and reading assignments to present a variety of applications of information theory. In particular, Stuart Altman of Princeton's Department of Ecology and Evolutionary Biology discussed his work on non-deterministic modeling of social communications among primates, and Michael Berry of Princeton's Department of Molecular Biology presented a three-hour lecture on 'Information Theory in the Brain.' Shannon's contribution to cryptography and communication secrecy was briefly studied using technical papers and popular science literature (e.g., [7]). Multimedia presentation on the Navajo Code Talkers illustrated the application of natural language to encryption. We also owe thanks to Joachim Hagenauer of Technical University of Munich for enthusiastically sharing with us his ideas and articles on the application of information theory to genomics [8] and the analysis of historical texts [9].

The students' understanding of the course's subject matter was put to test with a final project in which they were expected to write a paper on a topic of their choice illustrating the principles they had learned in the course. The topics of these papers ranged through applications of information theory to language, biology, finance, and cryptography. An impressive original paper by one of the students was on the study of the 'Entropy and Redundancy of the Hand Signs of American Sign Language,' a language which defies a simple characterization of its alphabet as it involves hand signs, facial expressions, and body movements. Another impressive paper by a student on 'An Information-theoretic Approach to the Massoretic-Biblical Text' discussed the complex error-detection and correction techniques used by scribes over centuries to preserve this ancient text. Yet another student used popular literature to

draw connections between Shannon, John Kelly, and gambling, culminating her presentation with a brief summary of Tom Cover's contributions to portfolio theory. The literature on the application of information theory to explain coding and information storage in the neuronal systems was the topic for a student majoring in biology, while the social sciences students presented various contributions of Shannon and his predecessors to cryptography.

Teaching Shannon's complex mathematical concepts to freshmen using intuitive ideas and examples has been in equal measures challenging and satisfying. One of the primary challenges is the lack of accessible reading materials that clearly and simply describe Shannon's ideas. We look forward to teaching the course again next year and we hope we can better address a number of ideas, including the simple yet oft-repeated question of why Shannon's contributions have been such a well-kept secret from the general public. We invite comments and suggestions from the community at large on books, reading materials, and techniques that can make the teaching less abstract. Finally, we wish to thank the CST at Princeton, and in particular Ms. Carol Prevost and Prof. Neta Bahcall, for their continued support and encouragement and for giving us an opportunity to try this as yet untried teaching experiment. Those interested can find more details on the course at <http://www.princeton.edu/~lalitha>.

References:

- [1] Hans Christian von Baeyer, *Information: The New Language of Science*, Harvard University Press, Cambridge, MA, 2005.
- [2] H. Vincent Poor, *Pulse-Coded Modulation*, Lecture notes from "ELE 391: The Wireless Revolution," taught at Princeton, Spring 2003.
- [3] Claude E. Shannon, "Prediction and Entropy of Printed English," *Bell System Tech. Journal*, 50:50-64, 1951.
- [4] Joseph J. Atick, "Could Information Theory Provide an Ecological Theory of Sensory Processing?," *Network*, Vol. 3, pp. 213-251, 1992.
- [5] Kees A. Schouhamer Immink, "Shannon, Beethoven, and the Compact Disc," *Information Theory Society Newsletter*, vol. 57, no. 4, Dec. 2007.
- [6] R. L. Kashyap, "Mathematics in the Vedic Age," <http://cobweb.ecn.purdue.edu/~kashyap/>.
- [7] Simon Singh, *The Code Book: The Science of Secrecy from Ancient Egypt to Quantum Cryptography*, Anchor Press, New York, 1999.
- [8] Pavol Hanus, Bernhard Goebel, Janis Dingel, Johanna Weindl, Juergen Zech, Zaher Dawy, Joachim Hagenauer, and Jakob Mueller, *Information and Communication Theory in Molecular Biology*, Electrical Engineering Series, Springer Series, vol. 90, no. 2, Dec. 2007.
- [9] Pavol Hanus and Joachim Hagenauer, "Information Theory helps Historians," *Information Theory Society Newsletter*; vol. 55, no. 3, Sep. 2005.

¹ It should be noted that in this, and all of the mathematical aspects of the course, only discrete alphabets were considered, even though most of the students were conversant with calculus.

² Now accepted as one of the earliest, perhaps the first, use of the binary number system.