(Summary of presentation for the University of Notre Dame Sesquicentennial Symposium)

Research in Information Theory - Is the End in Sight?

James L. Massey

Information theory, as established by Claude E. Shannon in 1948, is the primarily mathematical discipline that seeks to establish the fundamental limits on processes for representing, transmitting and/or storing information. Any such processing or transformation of information is a form of "coding". One distinguishes today among four different types of coding that are studied in information theory: channel coding, source coding, secrecy coding and authentication coding.

Because information theory seeks to establish *the* fundamental limit on a communications process, it is conceivable that the theory will someday be essentially complete. Are we in sight of this day? By the time of the Bicentennial Symposium, will information theory be a closed book in somewhat the same sense as classical thermodynamics is today? We now present nine predictions that incorporate our answer to this question.

Prediction #1: No one at the Notre Dame Bicentennial Symposium in 2042 will remember what was said at the Sesquicentennial Symposium.

We consider this first prediction so reliable that we are emboldened to "go out on a limb" with our eight other predictions.

Prediction #2: Information technology in 2042 will be entirely digital. (In particular, all TV signals will be digital.)

Samuel F. B. Morse was inventing the telegraph, a digital communications means, about the same time as Father Sorin was founding Notre Dame. Only two decades later, a transatlantic telegraph cable was in operation. Alexander Graham Bell was inventing the telephone, an analog communications means, much later, about when the present Administration Building was being built. It required another 8 decades before a transatlantic telephone cable was in operation. The disparity between the histories of the telegraph and the telephone illustrates the essential superiority of digital communications over analog communications in combatting channel noise. Not coincidentally, Shannon's theory is essentially a digital theory. As he himself said so well, "the discrete case forms a foundation for the continuous and mixed cases."

Prediction #3: Our technological progress in communications during the next 50 years will far outstrip our progress in the responsible use of this technology. (In particular, commercial television will still be a "vast wasteland" and dictators will be exploiting sophisticated information technology to impose their wills on their peoples.)

When one examines today's information theory, one is struck by the paucity of results in channel coding for situations involving feedback. Yet feedback is present in all interactive communications. Was Shannon exerting us to fill this gap when in 1974, in the only lecture that he ever presented at an IEEE International Symposium on Information Theory, he spoke exclusively about problems involving feedback?

Prediction #4: The great work in channel coding in the next 50 years will be done by researchers who explore the role of feedback in two-way communications.

Shannon formulated his theory of information in probabilistic terms. Where probabilistic models are appropriate, such as in data communication systems, the theory has been applied with much success. Where probabilistic models are inappropriate, such as in the description of pictures, the theory has been of little use. Entropy (or "uncertainty"), which is the cornerstone concept in Shannon's theory, is a measure of the randomness of a sequence produced by a probabilistic source. It was a surprise then when Kolmogorov showed in 1968 that it was possible to obtain Shannon's theory in all essential details if one defines the entropy of a long sequence as the length of the shortest program for a general-purpose computer that would cause the computer to produce that sequence as its output. Kolmogorov's approach leads to a completely deterministic formulation of information theory. The more important point, however, is that Kolmogorov's notion of entropy is easily extendable from sequences to pictures or, in fact, to any data structure.

Prediction #5: By 2042, something akin to Kolmogorov's notion of entropy will have supplanted the probabilistic notion of entropy in virtually all the interesting theoretical studies and practical applications of source coding.

It is now the time for us to say something about secrecy coding and authentication coding. These two forms of coding are today usually grouped together under the rubric of cryptography. They have the common feature that the enemy is diabolical. Unlike the honorable enemy, noise, of channel coding, the diabolical enemy waits until after the communications system has been built before deciding how to disrupt it. Shannon gave a theory of secrecy coding in 1949. He never considered authentication coding, whose theoretical foundations were laid by Simmons in 1984. That authentication coding has come on the scene only in the last decade of the several millenia history of cryptography is due to the fact that authenticity was long considered to be a concomitant property of secrecy. Rainer Rueppel has given an insightful way to distinguish between problems of secrecy and of authenticity. If it deals with who can *receive* a message, it is a problem of secrecy. If it deals with who can *send* a message, it is a problem of authenticity.

Today in secrecy coding, we have no practical cipher that is provably secure against attack. Finding such a cipher is the only intellectually honest task in secret-key cryptography, but most investigators shy away from it because they consider it either too difficult or impossible.

Prediction #6: By 2042, we will have practical secret-key ciphers that are provably secure.

Most cryptographic research today is in public-key cryptography where there is no advance distribution of secret keys. The foundations of this theory, given by Diffie and Hellman is 1976, are the one-way function and the trapdoor one-way function. Several functions of these types have been conjectured and a few have survived attack long enough that most cryptographers now believe in their one-wayness, but no proofs have been forthcoming.

Prediction #7: By 2042, there will still be no function proved to be either a one-way function or a trapdoor one-way function.

Communications technology is rapidly evolving. The Personal Communication System (PCS), which will permit anyone to reach us telephonically by dialing our universal telephone number no matter where we may be on or above this earth, is nearing reality, as is also Fiber to the Home (FTTH) that will permit communications at megabits per second rates into and out of our homes.

Prediction #8: The main problems of communications in 2042 will be the problems of preventing and disposing of garbage.

It is precisely the task of authentication coding to devise means so that only those entities whom we choose can send us messages. It is likely that, as theoretical progress in authentication coding is made, a variety of commercial and other special interest groups will bitterly oppose the application of this theory. We conclude these musings with an act of faith in the social responsibility of future communications engineers.

Prediction #9: By the time of the University of Notre Dame's bicentennial celebration, authentication coding will be the most active area of research in, and application of, information theory.