putation. Then in 1994 Peter Shor was able to exploit quantum superposition to find a fast algorithm for factoring integers. This was the first example of an important problem that a quantum computer could solve faster than a classical computer, and an indication that quantum computation might be more powerful than classical computation. The effectiveness of quantum computing is founded on coherent quantum superposition or entanglement which allows exponentially many instances to be processed simultaneously. However no quantum system can be perfectly isolated from the rest of the world and this interaction with the environment causes *decoherence*: the environment measures the quantum system collapsing the wave packet.

In classical computing one can assemble computers that are much more reliable than any of their individual components by exploiting error correcting codes. In quantum computing this was initially thought to be precluded by the Heisenberg Uncertainty Principle (HUP) — observations of a quantum system, no matter how delicately performed cannot yield complete information on the system's state before observation. For example we cannot learn more about a single photon's polarization by amplifying it into a clone of many photons — the HUP introduces just enough randomness into the polarizations of the daughter photons to nullify any advantage gained by having more photons to measure. At first it was believed that the quantum no-cloning theorem makes error correction impossible in quantum communication and computing because redundancy cannot be obtained duplicating quantum bits. This is not so — only repetition codes are eliminated. The trick is to take quantum superposition + decoherence, to measure the decoherence in a way that gives no information about the original superposition, and then to correct the measured decoherence. For details, including a beautiful group theoretic framework for code construction see [2].

References

[1] C. H. Bennett and P. W. Shor, Quantum Information Theory, *IEEE Trans. Inform. Theory*, to appear.

[2] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, Quantum Error Correction via Codes over GF(4), *IEEE Trans. Inform. Theory*, to appear.

Reflections of Some Shannon Lecturers

The following brief commentaries by thirteen recipients of the Shannon Award provide enlightening personal insight into the minds and attitudes of those who reached the pinnacle of achievement in our field.

The Eds.

Robert M. Fano
(1976 Shannon Lecturer)

The following quotation is from the preface to my 1962 book, *Transmission of Information:* "My interest in Information Theory dates back to the Summer of 1947 when, after completion of my doctorate thesis, I began looking for greener research pastures. ... My curiosity was particularly aroused by Professor Wiener's frequent statement that the information associated with a message depended on the ensemble from which it was selected, and that its average value could be identified with the entropy of the ensemble. This notion was so strange to me that I felt the need of some operational justification for it. By March 1948 I had obtained a justification in terms of message encoding, which turned out to be very similar to a theorem already proved by C. E. Shannon, but still unpublished. It was in this connection that I had the pleasure of meeting Dr. Shannon for the first time. I was so impressed by the scope and profoundness of his work that I have followed in his footsteps ever since. This book is primarily an account of his work, and of work inspired by him, either directly or indirectly." I was particularly amazed by Shannon's noisy channel theorem, its engineering implications, and the scope of the underlying probabilistic phenomena. This had a major influence on my own work and that of my thesis students.

Peter Elias
(1977 Shannon Lecturer)

Fifty years ago I had completed a Master's program in computation and further coursework at Harvard and was looking for a doctoral thesis topic when Shannon's paper came out. It was an amazing piece of work. As the Russian mathematician Khinchin said in a 1956 paper, "Rarely does it happen in mathematics that a new discipline achieves the character of a mature and developed scientific theory in the first investigation devoted to it ... so it was with information theory after the work of Shannon." I was fascinated, finished a thesis in information theory in 1950 and have continued working in the domain ever since, the first three years as a Harvard postdoc and since 1953 at MIT. I joined a group that Bob Fano, who had explored some of the same questions, was starting in Jerry Wiesner's Research Laboratory of Electronics. Shannon came to MIT from Bell Labs for a visit in 1956, and came to stay in 1958: he gave a wonderful advanced topics course, opening new topics in many of the sessions, and was always open for discussion. It was a wonderful environment for graduate students and faculty. My favorite paper by Shannon since 1948 is "Prediction and Entropy of Printed English" — a delightful example of the playful diversity of his approach, particularly in the identical twin coding scheme for estimating the entropy of English. The talk I enjoyed most was his first Shannon Lecture, in Ashkalon in 1973, in which he dealt with the circularity of

the occasion. He also made delightful gadgets. I liked best a box with a toggle switch on the front and a hinged cover. When you threw the switch up the cover opened slightly, an arm and hand came out, reached down, threw the switch back up and retreated into the box again as the cover closed. I miss that playfully creative mind.

W. Wesley Peterson
(1981 Shannon Lecturer)

Claude Shannon's book came to my attention about in 1953 when I was a graduate student at the University of Michigan studying vacuum tubes. I studied it from cover to cover and found it the most interesting material I had seen up to that time. I went to work for IBM, because I wanted to get involved with information and computing, and IBM gave me the unique and wonderful opportunity to attend the first course that Shannon taught at MIT, a truly inspiring experience for me.

Of course my work on error correcting codes was inspired directly by Shannon's work. I learned from him what an error correcting code is and he provided, with his fundamental theorem for the noisy channel, the goal for which I strived. I am happy that I could make some modest steps in that direction, and also that others have continued to make progress toward the goal that Shannon demonstrated for us.

Besides that, and maybe more important, Shannon taught us that at least some aspects of information can be dealt with quantitatively and I feel that from him I have acquired a much better understanding of this commodity information that I work with as a computer scientist. In the course of my teaching, I frequently find that I have insights that are a direct result of what I learned from him, for example in sorting, merging, searching, cryptography and cryptanalysis, and information coding.

Irving S. Reed
(1982 Shannon Lecturer)

I first heard of Claude Shannon and his work in early 1947 as a graduate student at Cal. Tech. Prof. E.T. Bell, after his course in mathematical logic, first told me of Shannon's 1938 AIEE paper on computer logic after my suggestion to Bell that logic statements could be realized by switches or relay logic. This landmark paper profoundly affected my early work on computer logic and architecture.

Later in 1952 at the MIT Lincoln Laboratory, physicist Ed Lerner lent me his copy of Shannon's famous paper (in book form), *The Mathematical Theory of Communication*, which laid the foundations of modern information theory. This paper made me painfully aware of how the reliability of the early digital processors could be improved, first by Shannon's coding theorem and more practically by the Hamming codes.

Because of my early work on processors for radar and communications, Shannon's information theory had an enormous impact on both my work and my personal development. Without doubt, Shannon's influence on me and on modern technology puts him in league with the greatest 20th Century scientist-engineers. In my case he led the way for me to consider seriously the possibility of algebraic error-correcting codes in the discovery and development of both the Reed-Muller and Reed-Solomon codes.

Robert G. Gallager
(1983 Shannon Lecturer)

Everything I needed to know about Information Theory, I learned from Claude Shannon. This includes not only his results about Information Theory, but also his way of doing research, which was beautifully balanced between a practical interest in how systems should work and theoretical beauty and generality. He never used a lot of high powered tools, but rather had an unsurpassed talent for finding the simplest non-trivial version of a problem, finding just the right way of looking at it, and then generalizing to a beautiful and general theory. In short, he gave us many existence proofs that research could be fun, simple, deep, and centrally important, all at the same time.

I remember once going to his office to talk about a puzzling problem I had been working on. He started throwing out pieces of it, one by one, until I was appalled at his "trivialization" of my problem. At a certain point, we could both understand it by inspection, and it was then easy to put all the complications back in.

Solomon W. Golomb
(1985 Shannon Lecturer)

I spent the summers of 1951 through 1954, while a graduate student in mathematics, working at the original Glenn L. Martin Co. in Middle River, Maryland. Two senior members of the group where I worked attended a special two-week course at M.I.T. (I think in 1953) and came back all enthused about Information Theory. I bought Woodward's book, *Probability, Information Theory and Applications to Radar*, and was pleasantly surprised to discover that this new branch of "engineering" was just as intelligible as mathematics.

In the spring of 1954, I drove from Harvard to New York to attend a special session on information theory "starring" Shannon, Elias, Huffman, *et al.*, at the annual I.R.E. convention. After spending a Fulbright Fellowship in Norway, I joined Caltech's Jet Propulsion Laboratory (JPL) in 1956. I spent most of the fall semester of 1959 at M.I.T. on leave from JPL. I sat in on Shannon's Information Theory course. I also had lunch at the same table with Shannon three times a week at the M.I.T. Faculty Club and had numerous discussions with him at other times as well. On one occasion, he asked if I could prove a result about probability distributions that he needed for some work he was doing concerning tails of martingales. He certainly got my attention and I was pleased to be able to present him with a proof the next day. (It turned out to be an easy proof once you found the right geometric way of looking at the problem.)

In June, 1985, I was the Shannon Lecturer at the ISIT in Brighton, England. Claude Shannon himself was in the audience! It was the first time that he had attended an ISIT since he gave the *first* Shannon Lecture in Ashkelon, Israel, in 1973. I chatted with him on a number of occasions in Brighton, but not on technical subjects of any depth. In general conversation he seemed totally lucid, but there were significant gaps in his remembrance of prior events.

For one of the most important technological innovators of the twenthieth century, Claude Shannon was remarkably modest and unassuming. I was one of many younger researchers whom he encouraged, and I certainly benefited significantly from my association with him.

James L. Massey
(1988 Shannon Lecturer)

I had the great luck as a graduate student at M.I.T. in the fall of 1960 to enroll in Professor Fano's course, Transmission of Information. His enthusiastic and entertaining lectures conveyed the beauty and excitement of information theory. I decided that this was the field for me and, in the following semester, enrolled in the course, Advanced Topics in Information Theory, which was taught by Shannon himself. I still treasure, and often re-read, my notes from those lectures. Shannon was a brilliant lecturer, if not in the classical style. His style of treating a subject was to present a sequence of increasingly complex examples, each having a solution that was obvious by inspection, until he had covered all the essential points, after which he would state the general theorem. I think that we graduate students all understood how to make the proofs then on our own. Whenever I have given a particularly good lecture since then, it has been because I consciously imitated Shannon's approach—the disasters have come when I forgot to do so.

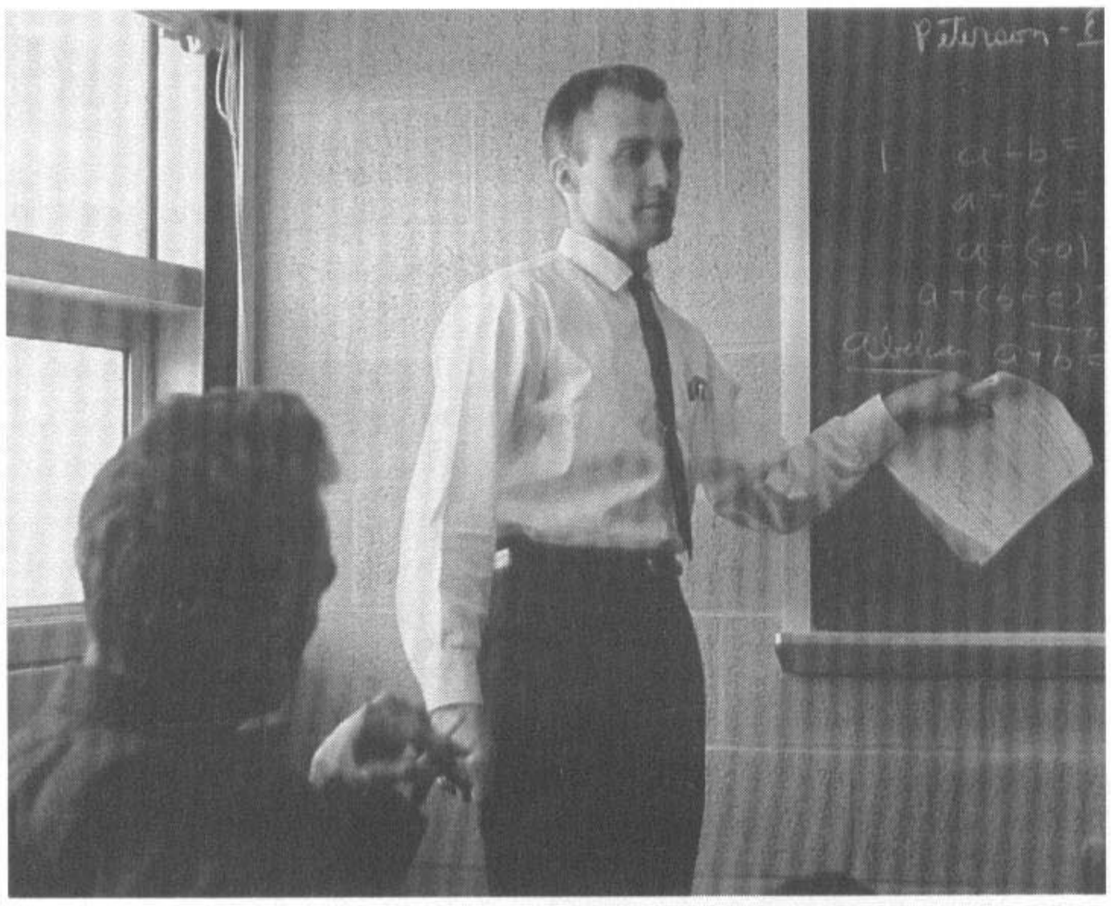

Photo courtesy of Jim Massey

Jim Massey as a young student under the watchful eye of C. Shannon.

Already then a very famous man, Shannon was nonetheless personally shy. He never seemed to me to be at ease in a crowd, but he was very relaxed and accessible to us struggling graduate students. Still it took me a long time to screw up my courage to the point where I dared to ask him to be a reader of my doctoral thesis, which he immediately agreed to do. He provided me with some excellent advice on my research—I also have copied the cross-examination technique by which he got me to explain what I was trying to do.

The second greatest honor of my life was being named a Shannon Lecturer. The greatest occurred during the 1986 ISIT in Ann Arbor, Michigan. Shortly before I gave my eminently forgettable talk, Claude and Betty Shannon entered the small lecture room, expressly to listen to me. To me this was one more proof that Claude Shannon is not only one of the great scientific figures of this century, but also a kind and generous human being.

Thomas M. Cover
(1990 Shannon Lecturer)

I bought two books with interesting titles the summer before starting graduate school at Stanford, Shannon's 1948 book on information theory and von Neumann and Morgenstern's book on game theory. Both books were extremely exciting. I spent over 100 hours using the game theory book to develop optimal strategies in various scenarios in poker. But Shannon's work seemed deeper and even more intriguing. I couldn't believe that something as intangible as information could be given such a satisfactory definition and have so many deep properties. I was also impressed by the relaxed and accessible writing style.

Shannon wrote his landmark paper as if the technical work had already been done and it was time to write an expository article on the subject. In fact, some of the greatest works in mathematics and physics have been written in this style. For example, Einstein, in his 1905 paper on relativity theory, had as one of his first equations: velocity = light path/time interval. And he described the notion of simultaneous events by saying, " 'The train arrives here at 7 o'clock,' means something like this: 'The pointing of the small hand on my watch to 7 and the arrival of the train are simultaneous events.' " This is craziness. Crazy or not, Shannon's paper is a great example of this tradition, in which no underlying intuition remains unrevealed. Indeed, the research literature in information theory is very readable, perhaps because of Shannon's influence.

As a result of Shannon's work, I have remained interested in the tangibility of information. One line follows the work of Kelly, in which Kelly shows that the increase in the growth rate of wealth from betting on a horse race is equal to the decrease in entropy. Another fascinating line of inquiry is the development of algorithmic complexity by Kolmogorov, Chaitin and Solomonoff in the mid 60s. The identification of the information in a sequence with its shortest computer de-

Photo courtesy of L. Zadeh

Claude Shannon (right) with (from left) Joe Weizenbaum, Fredkin, and John McCarthy in April '68.

scription leads to a concrete theory which is everywhere parallel to information theory.

Summing up, I would say that the results of Shannon, other than the major ones, of course, which have intrigued me the most are his proof that feedback does not increase capacity (which Gallager made transparent), Shannon's paper on the two-way channel, and Shannon's statement of the entropy power inequality.

There is certainly one piece of intriguing old business remaining. It is a quote from Shannon's 1959 paper on the fidelity criterion: "This duality can be pursued further and is related to a duality between past and future and notions of control and knowledge. Thus we may have knowledge of the past but cannot control it; we may control the future but have no knowledge of it." Shannon said he would write more about it in a subsequent paper, but the paper never appeared. I have tried, without success, to come up with an information theoretic statement on this subject equal to the summary.

Andrew J. Viterbi
(1991 Shannon Lecturer)

In my opinion, Shannon's contributions of 1948 and the subsequent decade are among the most remarkable and lasting theoretical achievements of the twentieth century. I have often remarked that the transistor and information theory, two Bell Laboratories breakthroughs within months of each other, have launched and powered the vehicle of modern digital communications. Solid state electronics provided the engine while information theory gave us the steering wheel with which to guide it.

Shannon theory not only establishes the limits on maximum efficiency of both source coding and channel coding, but it also points us in the right direction toward implementations which approach these limits. The vast majority of digital communication and broadcasting networks employ channel coding techniques based on Shannon theory. Furthermore, their designs are influenced by fundamental statistical ensemble concepts which were first expounded by Shannon in his founding 1948 paper. After nearly half a century of progressively more powerful coding techniques, we are within sight of the Shannon limit; almost error-free communication over a Gaussian channel can be achieved at above 80% of channel capacity, by iterative soft-decision decoding of concatenated codes. These so-called "turbo decoding" techniques require large block lengths and hence considerable delays to achieve such high efficiency, as clearly predicted by Shannon theory. As the very high speed data requirements of the Internet and similar multimedia applications become commonplace, delays which appear long in terms of bit times, become insignificant in real time, and such powerful methods will find wide usage. In a broader and more abstract sense, the spread spectrum techniques, which most digital wireless voice and data networks have already or will soon implement, are conceptually the logical extensions of Shannon theory.

Again we note that the sophisticated computation and memory intensive processing required for the implementation of Shannon theoretic concepts has become feasible and economically favorable through the amazing capabilities of solid state electronics to reduce size, power and cost while increasing speed, all by many orders of magnitude in the last three decades. This trend is likely to continue for some time to come.

Elwyn R. Berlekamp
(1993 Shannon Lecturer)

Claude Shannon has long been widely recognized as one of the foremost intellects of the 20th century. He discovered or invented Information Theory, which has become one of the key pillars of our digital society. He also made legendary contributions to topics now viewed as belonging to other fields, such as the applicability of Boolean algebra to the design of digital circuits, and the basic algorithm for computer-playing chess and checker programs.

He has also been a wonderful human being. He has been a major source of direct and indirect inspiration to me and to numerous others, and through a surprisingly small number of levels of indirection, to our entire community.

One unfamiliar with the man might easily assume that anyone who has made such an enormous impact must have been a promoter with a supersalesman-like personality. But such was not the case. He was actually a very modest man. Even though I worked with him directly over a period spanning several years, much of the influence he had on me was through others.

The earliest event in my career which I can remember occurred in 1946, when I was in the first grade and I learned to play the game called "Dots and Boxes." The second occurred around 1951, when I heard (by word of mouth) the problem of finding one off-weight coin mixed in with eleven

Photo courtesy of L. Zadeh

Betty Shannon (right) with (from left) Fay Zadeh and Mrs. Simons; Shannon can be seen in the back.

good coins, using only three weighings on a balance scale. That problem captured my interest in a big way. When I heard it had something to do with something called "Information Theory", I yearned to learn more about that subject.

Although there were no undergraduate information theory courses at MIT, my sophomore advisor was Peter Elias and he encouraged me to apply for a cooperative program at Bell Labs, where I began as a summer student in 1960. I was assigned to Ed David's department, where I became an apprentice to John L. Kelly, author of a paper originally entitled "Information Theory and Gambling" that had been published as "A New Interpretation of the Information Rate". Kelly was also very interested in games. I began learning Information Theory, and became acquainted with many other leading researchers then at Bell Labs including David Slepian, John Pierce, Ed Gilbert, Ed Moore, John Riordan, and Henry Pollak. Many of them had been former colleagues of Claude Shannon, and ALL of them regarded him with great awe. Back at MIT, I became acquainted with more faculty interested in information theory: Bob Gallager, Bob Fano, John Wozencraft, Irwin Jacobs, and eventually Claude Shannon himself. I oscillated between MIT and Bell Labs. At MIT, I was assigned a desk which was just being vacated by another student whom I then met very briefly; his name was Jim Massey. The first and only formal course I ever took in "Information Theory" was co-taught by Gallager and Fano. It was an extraordinarily eventful course, during which Gallager found a brilliant new proof of the coding theorem. David Forney was another student in that same class. I returned to Bell Labs in the summer of 1963 to work for David Slepian. During that summer he hired a new recruit named Aaron Wyner. Neil Sloane was recruited a few years later. Back at MIT, Gallager agreed to supervise my dissertation.

The other members of my committee were Elias, Wozencraft, and Shannon.

I had had almost no direct contact with Shannon before he agreed to serve on my committee, but I soon found him to be quite open and accessible. Here are three memorable anecdotes from 1963-64:

- One day we crossed paths in MIT's infinite corridor, and he stopped to chat. I said I was going to the library to look up certain particular references including one that HE had published. He urged me NOT to do so; he said I'd learn more by first trying to solve the same problems from scratch by myself. This advice came as quite a shock to me.
- One day he warned me that this was NOT the time to buy any stocks. This also came as quite a shock, because he certainly knew that I was an impecunious student with insufficient funds to buy anything. But I soon realized that he had intellectual as well as financial interests in the markets, and he correctly sensed that I shared the former. He had designed an analog feedback circuit which was intended to simulate the market and its probable reactions to flows of funds moving in and out.
- The first time that I visited his home in Winchester was a truly unforgettable experience. His wife, who had also been a mathematician, was very supportive. There was a tightrope a couple feet above the ground, on which he and his children performed. His junior high daughter strapped a bungey chord from her belt around her unicycle, and then JUMPED ROPE on the unicycle! His garage contained several dozen unicycles, all home made, including some so small that no one had yet succeeded in riding them. He wanted me to help him learn how to juggle five balls. His hands were slightly smaller than average, and he had considerable difficulty getting started.

In 1964-1967 as a coauthor with Gallager and me, Shannon lived up to my very high expectations. He had so many ideas as to how the results might be generalized or improved that only a few of them came to fruition during the three-year period our pair of papers was in preparation. He was going so strong then that I could not imagine that this pair of papers would turn out to be his last publications in Information Theory.

In 1973 I was president of IT, the year we initiated the "Shannon Lecture", the forerunner of the Shannon Award. Of course Shannon was our first awardee; that conclusion was unanimous even before nominations opened. The symposium was in Israel, and it was my task to persuade him to accept. That turned out to be less difficult than we had feared. The bigger challenge, which came to me as another total surprise, turned out to be helping Shannon overcome his stage fright during the half-hour before his talk! Shannon feared that the audience expected a God-like performance beyond anything he could possibly deliver. Fortunately, however, he did manage to get started, and once he got into his prepared lecture on "Feedback", it went very well. In my opinion, it was the best Shannon Lecture among the dozen or so that I've attended.

G. David Forney, Jr.
(1995 Shannon Lecturer)

Claude Shannon had a quite direct and personal impact on my career in information theory, in the absence of which I might well be in some other field today.

I was introduced to information theory in a marvelous thermodynamics course taught at Princeton by John Wheeler. My term project for that course was a book report on Leon Brillouin's book on physics and information theory, in which Brillouin "explains" Maxwell's demon by accounting for the information that the demon requires to open and close his door properly.

As soon as I arrived as a graduate student MIT, I took the 6.574 course on information theory. This ultimately led to a master's thesis in information theory and quantum mechanics. I then spent an unhappy and unproductive six months wandering around in other areas like operations research looking for a Ph.D. thesis, having been advised that information theory was pretty much dead as a research topic.

However, in Spring Term 1964 I took Shannon's 6.575 course, "Advanced Topics in Information Theory." Shannon's teaching method was quite similar to Wheeler's: he talked about various problems that he had been interested in, what progress he had made, and what open questions still puzzled him. I started playing with some of these problems; I don't remember exactly which, and I don't remember making any great progress. Nonetheless, by the end of the term I was off and running in information theory again, and within the next year finished my doctoral thesis on concatenated codes.

It is clear to me that Shannon's course was the direct cause of my return to information theory. I am eternally grateful to him not only for founding this field, but also for luring me back into it. It has been a great place to work.

Imre Csiszár
(1997 Shannon Lecturer)

As a student of mathematics, I learned information theory from Alfred Renyi, an outstanding mathematician. Information theory was one of Renyi's favorite subjects, though he was more interested in breaking new ground than dealing with mainstream problems. Following his lead, I wrote my first papers on generalized information measures and their applications in statistics and probability. Only later, after having learned about channel coding theorems from Wolfowitz's book, did I come to read Shannon's classical paper.

I was fascinated to see that most problems then studied in information theory had actually been introduced in that paper and, even more amazingly, that Shannon also provided the main ideas for their solution. Trained as a mathematician in doing formal proofs, I have learned a lot from Shannon's paper in terms of developing the intuition necessary for serious research in information theory. Arguments short of formal proofs tend to be unconvincing to mathematicians and some, better than I, have been known to fail to really understand Shannon's paper, surely for this reason. I also struggled with this obstacle but already having had some limited technical knowledge in information theory helped me to overcome it, as I was able to visualize how the often informal arguments could be turned into formal proofs. Ever since I have strongly advised my students to read Shannon's paper towards the end of the course in order that they better benefit from it.

During my life as a scientist, my research interests have been concentrated on the one hand on strict-sense information theory as directly descended from Shannon' fundamental paper, a subject now known as "Shannon theory", and on the other hand on applications of information theory within mathematics. The first may surely be attributed to Shannon, through his paper, and the other to my teacher, Renyi. I am equally indebted to both.

Jacob Ziv
(1997 Shannon Lecturer)

I well remember my first visit to Shannon's own study.

I came to MIT to study for my D.Sc. degree in the fall of 1959. Being an R&D engineer, I already knew that it was Information Theory that I would like to learn and investigate. I had first encountered Claude Shannon's monumental contributions after reading and trying to understand Goldman's book on the subject. I was therefore excited when my wife, Shoshana, and I were invited one weekend in the fall of 1959 to an open-house party for all the new foreign students to take place at the Shannon's residence, a beautiful house on top of a hill.

If I remember well, Claude was out-of-town that weekend, but many of the EE faculty were there to host us and warmly greet us. The party took place on the Shannons' huge hillside lawn. We were all impressed by one of the many self-made gadgets: a cable car that took you all the way up to the house. But one could operate it only at dinner time! (A clear message to the Shannon kids to be home for dinner on time!)

After a warm welcome by the faculty, I decided to dare to have a look at Claude's own study. I was impressed by the sight of a huge blackboard behind his desk. The blackboard was covered by a green shade. I was suddenly facing a real dilemma: Should I dare to have a peek at some of Claude's most recent, yet unpublished great results? Finally, after a period of tense hesitation, I moved the shade slightly, only to find out that there was indeed a formula spelled on the blackboard, neatly written in big letters; $$ H = -\displaystyle\sum_i\, p_i \log p_i . $$ Claude was apparently ready for us, counting on the fact that at least some of us could not withstand the temptation \ldots \quad Since then, I never actually stop searching for many of the erased results on Shannon's own blackboard.