Information Theory

INFORMATION THEORY AND THE DIGITAL AGE

AFTAB, CHEUNG, KIM, THAKKAR, YEDDANAPUDI

6.933 – FINAL PAPER

INTRODUCTION

Information Theory is one of the few scientific fields fortunate enough to have an identifiable *beginning* - Claude Shannon's 1948 paper. The story of the evolution of how it progressed from a single theoretical paper to a broad field that has redefined our world is a fascinating one. It provides the opportunity to study the social, political, and technological interactions that have helped guide its development and define its trajectory, and gives us insight into how a new field evolves.

We often hear Claude Shannon called the father of the Digital Age. In the beginning of his paper Shannon acknowledges the work done before him, by such pioneers as Harry Nyquist and RVL. Hartley at Bell Labs in the 1920s. Though their influence was profound, the work of those early pioneers was limited and focussed on their own particular applications. It was Shannon's unifying vision that revolutionized communication, and spawned a multitude of communication research that we now define as the field of Information Theory.

One of those key concepts was his definition of the limit for channel capacity. Similar to Moore's Law, the Shannon limit can be considered a self-fulfilling prophecy. It is a benchmark that tells people what can be done, and what remains to be done – compelling them to achieve it.

> What made possible, what induced the development of coding as a theory, and
> the development of very complicated codes, was Shannon's Theorem: he told
> you that it could be done, so people tried to do it.
> [Interview with Fano, R. 2001]

In the course of our story, we explore how the area of coding, in particular, evolves to reach this limit. It was the realization that we were not even close to it that renewed the interest in communications research.

Information Theory was not just a product of the work of Claude Shannon. It was the result of crucial contributions made by many distinct individuals, from a variety of backgrounds, who took his ideas and expanded upon them. Indeed the diversity and directions of their perspectives and interests shaped the direction of Information Theory.

In the beginning, research was primarily theoretical, with little perceived practical applications. Christensen says that the innovator's dilemma is that he cannot garner support for his new ideas because he cannot always guarantee an end profit. Fortunately, Information Theory was sponsored in anticipation of what it could provide. This perseverance and continued interest eventually resulted in the multitude of technologies we have today.

In this paper, we explore how these themes and concepts manifest in the trajectory of Information Theory. It begins as a broad spectrum of fields, from management to biology, all believing Information Theory to be a 'magic key' to multidisciplinary understanding. As the field moved from this initial chaos, various influences narrowed its focus. Within these established boundaries, external influences such as the space race steered the progress of the field. Through it all, the expansion of Information Theory was constantly controlled by hardware

technological limitations – indeed, the lack of such technology caused the 'death' of Information Theory, and its widespread availability is behind its current overwhelming success.

SHANNON'S "MATHEMATICAL THEORY OF COMMUNICATION"

"Before 1948, there was only the fuzziest idea of what a message was. There was some rudimentary understanding of how to transmit a waveform and process a received waveform, but there was essentially no understanding of how to turn a *message* into a transmitted *waveform*."
[Gallager, *Claude Shannon: A Retrospective*, 2001 pg. 2683]

In 1948, Shannon published his paper "A Mathematical Theory of Communication" in the *Bell Systems Technical Journal*. He showed how information could be quantified with absolute precision, and demonstrated the essential unity of all information media. Telephone signals, text, radio waves, and pictures, essentially every mode of communication, could be encoded in bits. The paper provided a "blueprint for the digital age"[1]

Since the *Bell Systems Technical Journal* was targeted only towards communication engineers, mathematician Warren Weaver "had the feeling that this ought to reach a wider audience than (just) people in the field" recalls Betty Shannon[2]. He met with Shannon, and together, they published "The Mathematical Theory of Communication" in 1949. The change from "A" to "The" established Shannon's paper as the new "scripture" on the subject – it allowed to reach a far wider group of people.

Why was Shannon's paper so influential? What was it about this paper that people refer to it as one of the greatest intellectual triumphs of the twentieth century? The answer lies in the groundbreaking concepts that *A Mathematical Theory of Communication* contains. Concepts that were influential enough to help change the world.

There are actually four major concepts in Shannon's paper. Getting an idea of each is essential in understanding the impact of Information Theory.

Channel Capacity & The Noisy Channel Coding Theorem
Perhaps the most eminent of Shannon's results was the concept that every communication channel had a speed limit, measured in binary digits per second: this is the famous Shannon Limit, exemplified by the famous and familiar formula for the capacity of a White Gaussian Noise Channel:

$$C_t = W \log_2 \frac{P+N}{N}$$

---

[1] Gallager, R. Quoted in Technology Review,
[2] Shannon, B. Phone Interview

The bad news is that it is mathematically impossible to get error free communication above the limit. No matter how sophisticated an error correction scheme you use, no matter how much you can compress the data, you can not make the channel go faster than the limit without losing some information.

The good news is that below the Shannon Limit, it is possible to transmit information with zero error. Shannon mathematically proved that there were ways of encoding information that would allow one to get up to the limit without any errors: regardless of the amount of noise or static, or how faint the signal was.

Of course, one might need to encode the information with more and more bits, so that most of them would get through and those lost could be regenerated from the others. The increased complexity and length of the message would make communication slower and slower, but essentially, below the limit, you could make the probability of error as low as you wanted.

> To make the chance of error as small as you wish? Nobody had ever thought of that. How he got that insight, how he even came to believe such a thing, I don't know. But almost all modern communication engineering is based on that work.
> [Fano, R. Quoted in Technology Review, Jul 2001]

The noisy channel coding theorem is what gave rise to the entire field of error-correcting codes and *channel coding* theory: the concept of introducing redundancy into the digital representation to protect against corruption. Today if you take a CD, scratch it with a knife, and play it back it will play back perfectly. That's thanks to the noisy channel theorem.

Formal Architecture of Communication Systems
The following diagram illustrates the formal architecture Shannon offered as a schematic for a general communication system. Flip open to the beginning of any random textbook on communications, or even a paper or a monograph, and you will find this diagram.
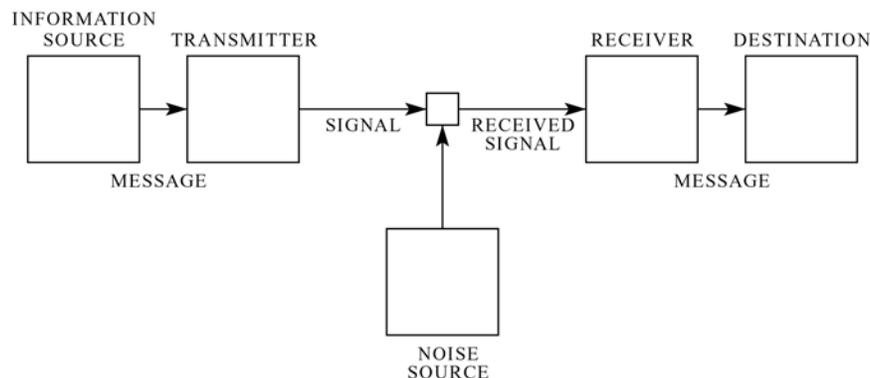

Figure 1. From Shannon's "A Mathematical Theory of Communication", page 3.

This figure represents one of the great contributions of A Mathematical Theory of Communication: the architecture and design of communication systems. It demonstrates that any

communication system can be separated into components, which can be treated independently as distinct mathematical models. Thus, it is possible to completely separate the design of the source from the design of the channel. Shannon himself, realized that his model had "applications not only in communication theory, but also in the theory of computing machines, the design of telephone exchanges and other fields."[3]

All of today's communication systems are essentially based on this model – it is truly 'a blueprint for the digital age'

Digital Representation
Shannon also realized that the content of the message was irrelevant to its transmission: it did not matter what the message represented. It could be text, sound, image, or video, but it was all 0's and 1's to the channel. In a follow-up paper, Shannon also pointed out that once data was represented digitally, it could be regenerated and transmitted without error.

This was a radical idea to engineers who were used to thinking of transmitting information as an electromagnetic waveform over a wire. Before Shannon, communication engineers worked on their own distinct fields, each with its own distinct techniques: telegraphy, telephony, audio and data transmission all had nothing to do with each other.

Shannon's vision unified all of communication engineering, establishing that text, telephone signals, images and film – all modes of communication – could be encoded in bits, a term that was first used in print in his article. This digital representation is the fundamental basis of all we have today.

Efficiency of Representation: Source Coding
In his paper, Shannon also discusses *source coding*, which deals with efficient representation of data. Today the term is synonymous with data compression. The basic objective of source coding is to remove redundancy in the information to make the message smaller. In his exposition, he discusses a loss-less method of compressing data at the source, using a variable rate block code, later called a Shannon-Fano code.

A challenge raised by Shannon in his 1948 paper was the design of a code that was optimal in the sense that it would minimize the expected length. (The Shannon-Fano code which he introduced is not always optimal). Three years later, David Huffman, a student of Prof. Fano's class at MIT came up with Huffman Coding, which is widely used for data compression. JPEGS, MP3s and .ZIP files are only some examples.

Entropy & Information Content
As we've discussed, Shannon's paper expressed the capacity of a channel: defining the amount of information that can be sent down a noisy channel in terms of *transmit power* and *bandwidth*. In doing so, Shannon showed that engineers could choose to send a given amount of information using high power and low bandwidth, or high bandwidth and low power.

---

[3] Shannon, C. *A Mathematical Theory of Communication*, pg. 3

The traditional solution was to use narrow-band radios, which would focus all their power into a small range of frequencies. The problem was that as the number of users increased, the number of channels began to be used up. Additionally, such radios were highly susceptible to interference: so much power was confined to a small portion of the spectrum that a single interfering signal in the frequency range could disrupt communication

Shannon offered a solution to this problem by redefining the relationship between information, noise and power. Shannon quantified the amount of information in a signal, stating that is the amount of unexpected data the message contains. He called this information content of a message 'entropy'. In digital communication a stream of unexpected bits is just random noise. Shannon showed that *the more a transmission resembles random noise*, *the more information it can hold*, as long as it is modulated to an appropriate carrier: one needs a low entropy carrier to carry a high entropy message. Thus Shannon stated that an alternative to narrow-band radios was sending a message with low power, spread over a wide bandwidth.

Spread spectrum is just such a technique: it takes a narrow band signal and spreads its power over a wide band of frequencies. This makes it incredibly resistant to interference. However it does use additional frequency ranges, and thus the FCC until recently had confined the technique to the military. It is now widely used in CDMA cellular phones.

Now that we've discussed some of the fundamental concepts in Shannon's work, let's take a step back and see how the formalization of these concepts started a chain of research that eventually became known as the field of Information Theory.

TRAJECTORY OF INFORMATION THEORY - I

We begin by exploring the history of Information Theory, how the field evolved and weathered various influences to become what it is today. In essence, we chart the trajectory of a new science.

Creating the Field

Information Theory grew out of the concepts introduced in "A Mathematical Theory of Communication." Although, the phrase "information theory" was never used in the paper, Shannon's emphasis on the word "information" probably helped coin the term. The idea that something as nebulous as "information" could be quantified, analyzed, and reduced to a mathematical formula attracted tremendous attention.

This initial excitement gave life to the field. But what were the forces that enabled this process? According to Latour, one of the tasks in creating a new field is gathering the support and enthusiasm of the masses[4]. Although Shannon had intended his audience to be confined to communication engineering, his concepts and methodology of thinking quickly moved into the popular press. 1953's *Fortune* magazine gushingly describes the field as more crucial to 'man's progress in peace, and security in war' than Einstein's nuclear physics.

---

[4] Latour. B, *Science in Action*, pg. 150

Perhaps, without popular support and interest of researchers from other fields, Information Theory may not have existed as it does today.

Another task in creating a new field is to recruit amateurs for the research workforce[5]. As previously mentioned, Shannon's 1948 paper attracted a multitude of individuals to conduct Information Theory research. At the time, these researchers were all amateurs to whom Shannon's paper had opened up entirely new ways of tackling the problem of transmission of information. However, these amateurs soon become the experts[6] and subsequently guided the direction of the field.

Circulation and Propagation of Ideas

Identifying the factors that transformed a single paper to a flourishing field, requires an investigation into the activities that occurred soon after Shannon introduced his theory.

Initially there was an absolute fervor of excitement. Universities began to offer seminars which later developed into classes. The Institute of Radio Engineers, or IRE[7], published papers on current research in a journal meant to focus solely on Information Theory, and formed a group called the Professional Group on Information Theory, or the PGIT. In addition, symposia were organized to present these papers and to allow forum discussions.

Amidst all the initial enthusiasm, many felt that with all the new concepts and research being generated, there was a need for a younger generation to get involved. As a result, several seminars and departments were organized at different universities such as University of Michigan and Universite di Napoli. These seminars later developed into classes, which had an influence on the field because they discussed current research questions, and produced graduate students who would eventually become the field's new practitioners. Professor Fano, in fact, taught one of the first courses, 6.574 commonly known as the 'Information Theory Course', at MIT. In his early lectures, Fano began by acknowledging that his subject matter was yet to be fully defined:

> Let's start by specifying a model of communication system to which *the theory to be developed* shall apply… This model should be sufficiently general to include, as special cases, most of the communication systems of practical interest, yet simple enough to lend itself to a detailed quantitative study.
> [Fano, R. 6.574 lecture notes, MIT Archives]

At the time, Professor Fano taught his class using the current research and its directions as his source of teaching material. He drew from here his assigned readings, problem sets, exams and final project questions. In fact, Huffman Coding, a form of efficient representation, originated from a final paper that Fano assigned.

A second course, 6.575 "Advanced Topics in Information Theory," was later taught by Shannon himself after he took professorship at MIT in 1956. Professor G. David Forney, Jr. credits this course "as the direct cause of his return to Information Theory."[8]

---

[5] Latour, B. *Science in Action*, pg. 150
[6] Eden, M. Interview
[7] IRE, The Institute of Radio Engineers later merged with AIEE, American Institute of Electrical Engineers on January 1, 1963 to form the IEEE
[8] IT Society Newsletter, pg. 21

Today, although neither an Information Theory department, nor a specific program exists within the EECS department at MIT. The field has become too ubiquitous, and its off-shoots are taught under a multitude of different areas: Computer Science, Information Technology, Electrical Engineering, Mathematics. Moreover, the concepts developed through Information Theory research have been integrated into the course material of different engineering disciplines. The "Information Theory Course" numbered 6.574 still exists today in the form of 6.441 "Transmission of Information."

We see that, as if following Latour's counsel, Information Theory quickly found its way into the curriculum at various educational institutions, and Shannon secured a university position. These are two more tasks that Latour considers important to creating a field[9].

Education did not only take place in the classroom though. The *IRE Transactions on Information Theory* became a journal whose "primary purpose [was] associated with the word 'education' and more specifically, the education of the PGIT membership in tune with current interests and trends"[10].

As a well-known, well-read, and well-respected journal, it had a great deal of control over the information and research that reached its readers. The *Transactions*, in a way guided the field by the research it chose to present in its publications. It published editorials by respected scientists in the field including such influential voices such as Claude Shannon, Peter Elias, and Norbert Wiener. Its correspondence section served as a written forum of discussion containing comments and reactions to published materials, either within the journal or elsewhere.

In addition to classes and the IRE journals, early symposia played a key role in the growth of Information Theory. The purpose of the symposia was to introduce cutting edge research and to foster an atmosphere of education and discussion.

For these symposia, the organizers searched for the "cream of the crop" in terms of papers; leaving out tutorials and reviews. Abstracts were submitted by many individuals from various areas of research and reviewed by a committee who judged whether the material was within the scope of the conference. Much effort was expended to keep the quality of research as high as possible. We should note that although this selection process was necessary to obtain worthy papers within the interests of the attendees, it opened the possibility of being biased toward the interests of the members of the organizing committee.

Despite the selection process the early symposia reflected a broadening in scope and an explosion of excitement. In the first London Symposium held in 1950, six out of the twenty papers presented were about psychology and neurophysiology. This number increased to eight by the time of the second symposium. But by the third held in 1956, the scope was so wide that it included participants with backgrounds in fields as diverse as "anatomy, animal welfare,

---

[9] Latour, B. *Science in Action*, pg. 150
[10] Cheathem, T. *A Broader Base for the PGIT,* IEEE Transactions, 1958, pg. 135

anthropology, computers, economics, electronics, linguistics, mathematics, neuropsychiatry, neurophysiology, philosophy, phonetics, physics, political theory, psychology, and statistics."[11]

Bandwagon
In the mid-50's, it was becoming apparent that Information Theory had become somewhat of a fad. This was because of confusion as to what Information Theory truly was.

> I didn't like the term Information Theory. Claude didn't like it either. You see, the term 'information theory' suggests that it is a theory about information – but it's not. It's the *transmission* of information, not information. Lots of people just didn't understand this… I coined the term 'mutual information' to avoid such nonsense: making the point that information is always *about* something. It is information provided by something, about something.
> [Interview with Fano, R. 2001]

Such misconceptions, together with the belief that Information Theory would serve as a unifying agent across a diverse array of disciplines led some researchers to attempt to apply Information Theory terminology to some of the most random of fields.

> …birds clearly have the problem of communicating in the presence of noise…
> an examination of birdsong on the basis of information theory might… suggest
> new types of field experiment and analysis...
> [Bates, J. "Significance of Information Theory to Neurophysiology." Feb1953:
> pg. 142]

Countless shallow articles based on 'non-engineering' fields were being published in the *IRE Transactions* at the time. Worse yet, researchers would deliberately introduce the words 'Information Theory' or 'Cybernetics' as it was alternatively called, into their work in hopes of attracting funding. These blind attempts to apply Information Theory to 'everything under the sun' created a great deal of controversy within the PGIT about what the bounds of the field should be. In December of 1955, L.A. De Rosa, chairman of the PGIT, formalized these tensions in an editorial titled "In Which Fields Do We Graze?"

> Should an attempt be made to extend our interests to such fields as management,
> biology, psychology, and linguistic theory, or should the concentration be
> strictly in the direction of communication by radio or wire?
> [De Rosa, L.A. "In Which Fields Do We Graze?" Dec 1955:2]

PGIT members were divided. Some believed that if knowledge and application of Information Theory was not extended beyond radio and wire communications, progress in other fields could be delayed or stunted. By broadening the scope of PGIT, knowledge would be shared with other areas. Others insisted on confining the field to developments in radio, electronics, and wire communications. The two points of view were hotly debated over the next few years through correspondence in the *Transactions* and elsewhere.

This is a clear example of the Great Divide, as it is defined by Latour[12]. The PGIT is a scientific network. Within the PGIT, there existed an inner and outer network. Latour's "insiders" consist

---

[11] Blachman, N. *A report on the third London Symposium*, IEEE Transactions, March 1956, pg. 17

of the members who believed that Information Theory should be confined to communications engineers (or the *purists*). The "outsiders," of course, are the members who supported expanding Information Theory to other fields. In the Great Divide, the insiders do not believe that the outsiders have a correct understanding of the nature of the field.

By 1956, the debate had become heated enough that the father of the field had to address it. In his March editorial, "The Bandwagon" Claude Shannon responded to De Rosa's question, taking the side of the purists. He wrote in his usual gentle fashion, but showed signs of frustration at the state of Information Theory. Shannon felt that Information Theory had "ballooned" into more than it actually was, because of its novelty and popular exposure. Shannon's wife, Betty Shannon, commented, "He got a little irritated with the way people were pulling it around. People didn't understand what he was trying to do."[13] Shannon had intended the theory to be directed in a very specific manner, and therefore believed that it may not be relevant to other disciplines. Moreover, he believed that the *IRE Transactions*, being an academic journal, should require more carefully researched papers that would appropriately – and not just superficially – apply Information Theory and do so in a more rigorous manner.

> A thorough understanding of the mathematical foundation and its communication application is surely a prerequisite to other applications. I personally believe that many of the concepts of information theory will prove useful in these other fields-and, indeed, some results are already quite promising-but the establishing of such applications is not a trivial matter of translating words to a new domain, but rather the slow tedious process of hypothesis and experimental verification.
> [Shannon, "The Bandwagon" March 1956]

Norbert Wiener, another influential member of the PGIT, also agreed with Shannon that the concept was being wrongly thought of as the solution to all informational problems.

> ...As Dr. Shannon suggests in his editorial: *The Bandwagon*, [Information Theory] is beginning to suffer from the indiscriminate way in which it has been taken as a solution of all informational problems, a sort of magic key. I am pleading in this editorial that Information Theory... return to the point of view from which it originated: the … statistical concept of communication.
> [Wiener, "What Is Information Theory?" June 1956]

Such editorials made the views of the core of the PGIT clear. We see a rapid reduction in the number of 'fluffy' papers in the *Transactions* – the topics increasingly become focussed on new research in communication engineering.

By 1958, the fate of the field had pretty much been decided. Peter Elias's scathing 1958 editorial "Two Famous Papers" crystallized the "great divide". He took a much harsher stance than Shannon's in describing a typical paper that should not be published:

> The first paper has the generic title 'Information Theory, Photosynthesis and Religion'… written by an engineer or physicist… I suggest that we stop writing [it], and release a large supply of man power to work on… important problems which need investigation.

---

[12] Latour, B. *Science in Action,* pg. 211
[13] Shannon, B. Phone Interview

[Elias, P. *Two Famous Papers*, Sep 1958]

Elias's words are the last on the debate – the field has found its bounds. Or rather, they have been found for it. The purists, with their control over the *Transactions*, and their ability to control the very forum of discussion, had won.

With the field establishing its boundaries, the PGIT began to look towards the future and change its focus to the development of practical applications. F. Louis Stumpers, now Honorary President of the International Union of Radio Science commented:

> Theoretical and fundamental studies come first, and much work remains to be done, but many will judge us by the practical application we can find. Let us give special care to this…
> [Stumpers "Information Theory and Int. Radio Organizations" in June 1957, 85]

Digression – Some Essential Background

To understand what these "practical applications" would eventually be, it is important to realize what was going on in the area of communications at the time. Let us digress for a moment for some background. With the advent of World War II, the focus of communication had changed from telephony and broadcasting to radar control, fire control, and military communication.

The focus of long-distance communication, too had changed, from the transmission of waveforms to the transmission of data because computers were eventually becoming the new receiving terminal. Humans are adept at handling error and can extrapolate the content of a message, whereas a computer cannot interpret a body of data without algorithms to correct for errors. These new technicalities required efficient methods to encode and send data. Shannon's ideas would be instrumental in developing these solutions.

By the end of World War II, the military services were very impressed by the contribution to the war effort made by highly skilled physicists, mathematicians, and engineers such as those at MIT's Radiation Laboratory. The military felt that it was crucial to have a large pool of manpower available, especially one highly skilled in electronics and communication in case of another war.

The Army, the Navy, and the Air Force thus decided to sponsor research, and in 1951 the military initiated the Tri-Services contract, which included all three branches. First to be funded was MIT's Radiation Laboratory, which later became the Research Laboratory for Electronics (RLE).

According to Latour, scientists need to align themselves with deep pockets, those of the industry or the federal government[14]. Only those who have funding will be influential. The military was one of the most well funded branches of the federal government. Backed by the military Tri-services fund, Information Theory research was soon well positioned.

Returning to our story, we find that by 1958 it was obvious that the field had become mature and

---

[14] Latour, B. *Science in Action*, pg. 172

well defined. The *IEEE*[15] *Transactions* began to reflect the attitude of the PGIT. There were areas that were appropriate, and areas that were not.

> Topics of current *appropriateness* include extensions of information theories of Shannon and Wiener and their ramifications, analyses and design of communication systems, information sources, pattern recognition, receiving and detection, automata and learning, large-scale information processing systems, and so forth.
> [The PGIT Administrative Committee in 1959, 136 (emphasis added)]

Discussion

Consider again the story that we just saw: the story of a field finding its boundaries, moving beyond misguided attempts to apply it blindly where it did not belong, eventually progressing along its natural trajectory…

A nice fairy tale, isn't it? We know from MacKenzie that there is no such thing as a natural trajectory:

> "Technological change is social through and through. Take away the… structures that support technological change of a particular sort, and it ceases to seem 'natural' – indeed it ceases altogether'
> [MacKenzie, *Inventing Accuracy,* 384]

We will continue with our story, then with these cautionary words: though we present a trajectory, this trajectory is by no means 'natural'. It is influenced by an immense number of factors: the biases of the PGIT, Shannon's personal opinions, the cost of hardware technology, the interests of the military, external events such as the launch of the Sputnik…

Consider the boundaries of the field, and the path taken to define them. Latour discusses the immense importance of scientific papers in determining 'the truth', or what is believed to be the truth. We have already discussed that there was intense debate in the PGIT over the future of Information Theory – the publishing of random articles on topics such as birds, had infuriated the purists, who *felt* that such nonsense would dilute the seriousness of their field. Scientists were using the term "Information Theory" as a catchy tag-line to get funding, even if their work had nothing to do with it. The backlash against this was so strong, that it derailed even serious attempts to apply Information Theory to other fields.

This effect was strengthened by the fact that the heavy-weights in the PGIT: Shannon, Wiener, Fano et. al. were all purists, and office holders in the PGIT. And they controlled what was published and what was not. Thus their personal biases and opinions contributed strongly in what the field eventually chose to include… and what it chose to exclude. In looking through Peter Elias's personal papers, we found paper after paper that he had rejected for the *IEEE Transactions*, because they were simply 'irrelevant'

The interest of the military too, must not be overlooked. The military wasn't especially interested in the application of Information Theory to birds, art or even neurophysiology. It was interested

---

[15] IRE (The Institute of Radio Engineers) later merged with AIEE (American Institute of Electrical Engineers) on January 1, 1963 to form IEEE (Institute of Electrical and Electronic Engineers).

in communication. Even though research was of secondary interest to it, it still had influence over the areas explored: and the interests of the purists and the military were easily aligned. And the rest is history.
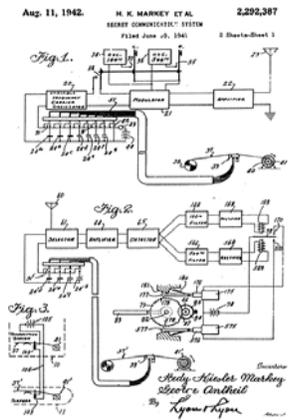
We shall touch on these issues again at the end of our paper. But for now, back to our tale…

Figure 3. Hedy Lamarr

Figure 4. Hedy's Patent

TRAJECTORY II - PROGRESSION OF THE FIELD

Work on Information theory tended to focus on three main areas of application based research: spread spectrum technology, source coding, and channel coding. Coding theory eventually became such a dominant area that Information Theory became virtually synonymous with coding. "Coding theory *is* Information Theory", insists Prof. Fano, "It's just progress."

Spread Spectrum Technology

One of the first applications of Information Theory concepts, however, was in spread spectrum technology, where its development was driven by military needs. During World War II, radio signals were used for missile guidance as well as communication, and it was important that these signals be protected from enemy detection and jamming.

Frequency hopping, which involves cycling through random frequencies to evade jamming was one of the earliest fore-runners of modern spread spectrum technology. The technique was developed by Hollywood actress Hedy Lamarr and musician George Antheil.

The concept of spread spectrum as it is known today stems from Shannon's idea of entropy. In our discussion on Shannon's paper we pointed out that for robust communication a high entropy message needs a low entropy carrier, such as noise. Spread Spectrum technology spreads a narrow radio signal over a greater bandwidth and makes it resistant to jamming and interference. From the 1940s to the mid - 1980s, almost all research and development devoted to the technology was backed by the military and was highly classified[16].

A lot of the initial work on spread spectrum was done at MIT in the early 50s under the NOMAC project, which stands for **NO**ise **M**odulation **A**nd **C**orrelation. Research done at MIT led to the development of several 'transmitted reference' spread spectrum systems, all for military applications. These systems achieve detection

---

[16] CDMA Development group, web-site.

by transmitting two versions of the carrier, one modulated by data and the other unmodulated. A correlation detector then extracts the message from the two signals[17].

The first operational spread spectrum system was the Lincoln Lab F9C system, developed by Dr. Paul Green. It was improved upon by the RAKE system in 1958[18]. Today, spread spectrum can be seen as CDMA technology in cell phones.

Efficiency of Representation

As we mentioned, coding theory was at the forefront of Information Theory – it was a broad area, and one where Information Theory's influence was most clearly defined. Early coding work originally focused on data compression. At the time, reliable communication was not considered to be particularly relevant or applicable. The focus of interest was on the efficiency of representation and bandwidth conservation, "being able to represent speech and pictures more economically…At that time, people felt that one could gain a few dBs by encoding, but was it worth it?"[19]

Some of the important early research questions and issues of applicability were in this direction, for example Huffman coding. In 1951, David Huffman took the 6.574 graduate course in Information Theory, taught by Professor Fano at MIT. As an alternate to taking the final exam, Professor Fano used to give his class the option of writing a term paper on open research topics in Information Theory. That year's topic was to find the most efficient coding method for representing symbols, such as numbers and letters. At the time, there was no known solution to the problem: Shannon and Fano themselves had struggled with it.

After working on the problem for months, Huffman finally decided to give up. When he was trashing his notes, he had an epiphany. "It was the most singular moment of my life," Huffman said. "There was the absolute lightning of sudden realization."[20] It was already known that the more frequently a symbol appears, the more its code is used, and consequently, it was better to represent the most frequent symbols with shortest codes. Those who had been trying to discover an efficient coding previously had been trying to assign codes going from most frequent to least frequent. Huffman realized that by going the other way around, from least frequent to most frequent, it was possible to assign the symbols with the most optimal algorithm. Huffman published his algorithm in his paper entitled "A Method for the Construction of Minimum Redundancy Codes." His data compression algorithm, known as Huffman coding, is used universally in data compression and data storage systems today[21].

Thus, the focus of early work in coding remained the efficiency of representation. Even the early literature on the scope of Information Theory time betrays an almost myopic focus on it. The question of the time seems to be the saving of bandwidth: Can we use the techniques of Information Theory to save some bandwidth, or compress some data for a particular application? If not, then IT is probably not particularly useful for it. For example, D. Gabor reluctantly

---

[17] Jachimczyk, W. *Spread spectrum web page*
[18] Hochfelder, D. *Interview with Dr. Paul Green*, IEEE History Center
[19] Fano, R. Interview
[20] Stix, G. *Profile: David. A. Huffman,* Scientific American, Sept. 1991 54,58
[21] Huffman, K. *David Huffman Elegy web page*

concludes in his 1953 article *Communication Theory, Past, Present, Prospective*, that as not much bandwidth can be saved in telegraphy or music transmission, "it is evident that only a moderate benefit can be obtained by the application of modern communication theory" to these fields.

Channel Coding

With the popularity of data compression and source coding, the true importance of Shannon's noisy channel theorem was not recognized. It was treated as little more than a theoretical curiosity. The reasons behind this were simple: after all, a simple way of attaining reliable communication is to blast the signal at a higher signal to noise ratio. If you want to be heard, yell. Another problem was hardware: error-correcting schemes were complex and demanded computation power, which was prohibitively expensive.

> "I remember John Pierce at Bell Labs… He was Shannon's boss. He was playing down the importance of the noisy channel theorem, saying: 'just use more bandwidth, more power'… there was no limitation then – you could do whatever you needed in terms of reliable communication without long encoding. And besides, even if you wanted to, you were very badly limited by equipment complexity and cost…"
> [Interview with Fano, R. 2001]

Fortunately, there were still a few minds interested in exploring error-correcting possibilities. These early pioneers deferred the issues of applicability, and continued to explore the potential of the noisy channel theorem.

Convolutional codes and sequential decoding

Peter Elias invented convolutional codes in 1955. As opposed to simple block codes, convolutional codes approximate a random tree structure, and are more powerful. They are the basis of some of the most popular codes in use today.
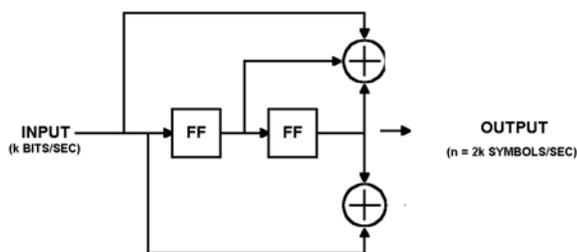

Figure 5. A Simple Convolutional Encoder

What made this particular method particularly attractive was the relative ease with which data could be encoded. However, as Professor Forney points out in his 1995 Shannon Lecture – an annual honor awarded by the IEEE Information Theory Society – "the key obstacle to practically approaching channel capacity was not the construction of specific good long codes… rather it was the decoding complexity"[22]

For convolutional codes, the complexity of decoding increased exponentially with the length of the code. Thus convolutional codes, despite their attractive properties, seemed doomed to extinction.

Then, in his 1957 doctoral thesis, "Sequential Decoding for Reliable Communication," John

---

[22] Forney, D. *Shannon Lecture,* 1995. IEEE IT Society Newsletter, Summer '98 :21

Wozencraft tackled the problem of decoding complexity. Sequential decoding was essentially a class of "exhaustive tree searching techniques"[23] whose probabilistic approach gave it an advantage. The decoding complexity of his algorithm rose polynomially with the length of the codes, as opposed to exponentially – a major breakthrough. With Fano's improvements in 1963, sequential decoding became an tractable method of handling convolutional codes.

The work of these pioneers, however remained theoretical. Although Wozencraft's decoder made the use of convolutional codes feasible, hardware technology was far behind. The original experimental sequential decoder, SECO built at Lincoln Labs in '62 was extremely large, taking up two solid rooms of equipment.

Plagued by a lack of motivation, and worse yet, by a lack of technology, the noisy channel theorem seemed doomed to stagnation. But then something happened that gave error correction and the noisy channel theorem their first big break. And that something was space.

Sputnik
The launch of the Sputnik in 1957 changed everything. The Russians had taken the lead in the Space Race, and there was widespread dismay in the US. The effect has been likened it to a second 'Pearl Harbor'. There was a sudden surge of interest in space launches, and consequently, in space communication. Determined not to be surpassed by the Russians, the United States rapidly enhanced the domestic space program. In 1958, President Eisenhower approved a plan for satellites, and NASA was established.

Sputnik, and the subsequent development of the space program, generated a great deal of interest in reliable communication in the presence of noise. It was suddenly a very real problem, and it was no longer possibly to glibly say 'just use more power.' As a result, a great deal of work was done in channel coding.

> "What really changed the whole picture was space communication. Because power is very expensive in space - the generation of power, the weight of the power supply. And that's when the industry, and the research in general began to think much more seriously about communication in the presence of noise. Space changed the picture entirely…"
> [Interview with Fano, R. 2001]

There were actually a number of factors that made channel coding so perfectly suited to the problem of deep space communication.

- Foremost, as mentioned, power is very expensive in space
- The deep space channel – the channel for communication with space probes – almost perfectly matches the theoretical noisy channel model that Shannan presented in his original paper, which was very well understood.
- Bandwidth, which is used up by coding, is relatively plentiful in space.

---

[23] Forney, D. *Shannon Lecture,* 1995. IEEE IT Society Newsletter, Summer '98 : 21

- As we mentioned, equipment complexity made coding such an expensive proposition that most people wouldn't seriously consider it. But for an ultra-expensive space mission, the percentage cost of the coding equipment is small. In addition, each dB saved by coding resulted in more than a $1000,000 of saving in the communication equipment. And that's a $1000,000 in 1960.
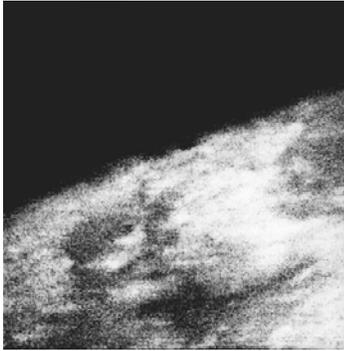

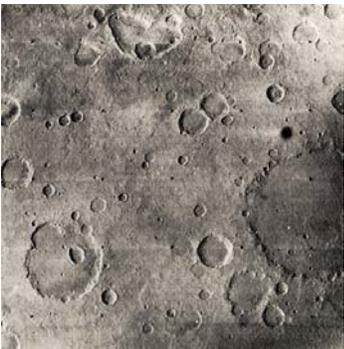**Figure 6.** Mars, Mariner IV, '64 using *no* coding.

Figure 7. Mars, Mariner VI, '69 using Reed-Muller coding.

Figure 8. Saturn, Voyager, '71 using Golay coding.

Because of these reasons Professor James L. Massey called deep-space communication and coding a "marriage made in heaven"[24]. Space gave coding its first real application – and the field that had originally been marginalized quickly began to prosper.

Advances in Hardware Technology

Space technology not only drove the need for coding advances, but for hardware advances as well. In 1958, Jack Kilby at Texas Instruments came up with the integrated circuit, which replaced discrete transistors. With time, the cost, weight, and power required for the average decoding operation were reduced by orders of magnitude. At the same time, decoding speeds reached several megabits per second. Complexity became less of a concern as newer and faster integrated circuits were developed. This enabled practical application of channel coding, not only in space but in other areas as well. By 1968, the sequential decoder that used to take up two rooms at Lincoln Lab, had been miniaturized enough to be placed on Pioneer 9 – a small spacecraft weighing a mere 67 kilograms.

Space Applications

Some of the first series of missions to use coding were NASA's Pioneer and Mariner missions. Starting with Mariner VI in 1969, the Mariner series was the first to officially use coding. Mariner VI utilized Reed-Muller codes, which was developed in 1954 and allowed the correction of a variable number of errors. Note the difference between Figure 6, which shows one of the first close up pictures of Mars courtesy of Mariner IV, which used no coding, and Figure 7, another shot of Mars from Mariner VI, using the Reed-Muller code mentioned.

While they were developing Reed-Muller codes for the Mariner series, the technology became available to use a sequential decoder in space. Thus, in 1968 NASA engineers decided to test the sequential decoder, designed by MIT's Professor Forney at Codex by putting it on an 'experimental basis' on Pioneer IX.

[24] Massey, J. L. "Deep-space communication and coding.", in *Lecture Notes on Control and Information Sciences*, 1992.

This was a deliberate ploy to avoid the NASA approval process[25]. Pioneer IX, in fact became the first spacecraft to use encoding.

By 1979, the Voyager probes were launched, and were capable of transmitting color pictures. For color, the source alphabet is much larger, and thus better error correction is needed. The coding scheme chosen was the Golay code.

Coding theory, and the hardware technology required to implement it was rapidly expanding. So much so, that the Voyager mission was actually changed on the fly to keep up with improvements in technology. The Golay code it used at launch were replaced by Reed-Solomon codes – the ubiquitous codes now used everywhere for error correction, including CDs and hard drives. Reed-Solomon codes were invented in 1960 at Lincoln Lab.

How quickly the field was progressing can be seen by noting that between the time Voyager's launch and its arrival at Neptune, there was a *six-fold* increase in the data-transmission rate.

An even more dramatic sense of how much, and how quickly, the field had progressed can be obtained by noting that the telemetry rate for Voyager at Neptune. The rate was 21.6 kbits/s at a distance of 4.4 billion miles – a performance improvement of almost 10^7 over Mariner IV in a mere decade.

"Coding is dead"

By the late 60s, there was a general feeling that Information Theory as a field, and coding in particular, was dying. Sequential decoding had been implemented in space, but remained too expensive for commercial use. The microchip was not yet around, no new applications were being invented, and the new coding schemes being suggested were already far too complex to implement.

Convolutional codes and sequential decoding had got coding closer to the Shannon Limit, and there had developed a general "quasi-religious belief that there… (was a) "practical capacity" of a memoryless channel, and that by the use of sequential decoding the practical capacity could more or less be achieved. Problem solved"[26]

> …A lamentable consequence of this conclusion (which is rather ironic in view
> of the later history of sequential decoding) was that the M.I.T. information
> theory group, probably the greatest assemblage of talent that our field has ever
> known, began to disperse to other fields and institutions, thus bringing to an end
> the first golden age of information theory…
> [Forney, Shannon Lecture, 1995]

Professor Fano, along with Professors Shannon and Elias, were just some of those included in this "great assemblage," who decided to move on to more interesting projects.

---

[25] Costello et. al. *Applications of Error-Control Coding*, IEEE Transactions on Information Theory, Oct '98
[26] Forney, D. *Shannon Lecture,* 1995. IEEE IT Society Newsletter, Summer '98 : 21

> In the 60's, lots of people felt that the field was becoming obsolete. Nothing
> very interesting was happening… It was shown that you could encode and
> decode efficiently, but the practical application of this was a different story. All
> the operations involved in encoding and decoding were very, very expensive…
> [Interview with Fano, R. 2001]

Things were discouraging, and they were getting worse. Prof Gallager recalls that some of the wiser heads at MIT suggested that he move into a more promising field, 'such as vaccum tubes'. Prof. Forney recalls receiving similar advice.

By the end of the decade matters had come to a head. In 1971, there was a coding workshop held in St. Petersburg, Florida, entitled "Future Directions". Many prominent coding theorists attended to discuss further areas to look into. After spending a day or two discussing future possibilities, they concluded that everything that was of interest to their community was finished, and that there was no future direction, except out. In the conference, Robert McEliece gave an infamous talk entitled "Coding is Dead"

> The thesis of his talk was that he and other coding theorists formed a small
> inbred group that had been isolated from reality too long. He illustrated this talk
> with a single slide showing a pen of rats that psychologists had penned in a
> confined space for an extensive period of time. I cannot tell you here what those
> rats were doing, but suffice it to say that the slide has since been borrowed many
> times to depict the depths of depravity into which a disconnected group can fall.
> The rats made *Lord of the Flies* look like a school picnic.
>
> All this depraved behavior had its parallel in the activities of coding theorists.
> Too many equations had been generated with too few consequences… Coding
> theorist professors had begotten more coding theory PhDs in their own image…
> no one else cared; it was time to see this perversion for what it was. Give up this
> fantasy and take up a useful occupation… Coding is dead.
> [Lucky, R. *Lucky Strikes Again '93,* Compilation of articles from IEEE Spectrum]

Prof. Proakis at Northeastern University recalls the dismay the conference created. "These were people who had been working in coding theory for a number of years, and they concluded that everything worth doing had been done already." The researchers who attended the workshop believed that coding had a very limited role, its only applications in deep space communication and in the military. At this time, even Viterbi's work, which was the key to optimal decoding of sequential codes, was little understood and unappreciated, and was simply too expensive to impliment.

One of the few dissenting voices at the conference was that of Irwin Jacob. He stood up, saying, "Coding theory is not dead", as he pulled out a 4-bit integrated shift register from his pocket, "and *this* is why."

That 4-bit shift register may be pitiful by today's standards, but it was a big thing then. It had been the lack of hardware technology that led the theorists to announce the death of coding. Yet it was also the availability of hardware technology, that would gave way to its revival.

TRAJECTORY III - COMMERCIAL APPLICATIONS & ISSUES OF TECHNOLOGY

Hardware technology has always been the determining factor for Information Theory. The theory is just that: theory. Algorithms, concepts, codes. They require technology, in the form of complex hardware and computing power to be used. When technology was scarce, complex and expensive, as in the early 60s – applications were confined to those with the pockets to afford them: NASA and the military.

By the late 60s the semiconductor revolution was rapidly accelerating. Intel came out with the first microprocessor in 1971. Just as the lack of hardware technology had pushed coding into a coma, the increasing availability of technology paved the way for its revival. Indeed, by the end of the 60s, semiconductor technology had progressed enough that it was becoming feasible for Information Theory to break out of the confining boundaries of ultra-expensive military and space applications and set its sights on the commercial world.

"Certainly integrated circuits played a large role in making these coding schemes practical"[27] says Professor Proakis. Professor Fano puts it more emphatically: "It was the microprocessor revolution that changed the entire picture!"[28]

Technology again, in the form of ready and cheap hardware, and computing power, is behind the phenomenal success of Information Theory today. This is exemplified by the story of Low Density Parity Check Codes. These codes allow one to get infinitesimally close to the Shannon Limit – they are the most powerful codes currently known. The interesting thing is that they were proposed by Prof. Gallager, in his 1960 PhD dissertation. It has taken forty years for technology to reach the point where they can be implemented.

Indeed, Codex Corporation, one of the first companies to apply Information Theory concepts commercially, acquired the rights to Low Density Parity Check codes in the 1960s. The codes were never implemented: they were simply too complex.

Codex

Codex was an MIT startup, founded in 1962. When Professor Forney joined it after completing his doctorate in 1965, it was a small 12-person start up. Codex originally, like everyone else, was in the business of military applications. They worked on building stand-alone error-correcting patches, which were fitted on existing military communication systems. "Our business then was to put 'bandages' on systems that had already been designed," recalls Professor Forney. "We were able to get some business, almost all from the government, but it was becoming clear that stand-alone error coding was not the way to go.[29]" So Codex moved on from using Information Theory ideas in error correcting for military work, to applying them to its commercial hit, the modulator-demodulator, better known as, the modem.

---

[27] Proakis, J. Interview
[28] Fano, R. Interview
[29] *A Conversation with Forney, D.* IEEE IT Society Newsletter, June 1997

In fact modems were one of the first important commercial applications of Information Theory: modems, and related work on facsimile machines, are what allowed it to break into the commercial arena.

Professor Gallager, who was working as an advisor for Codex, was aware that Jerry Holsinger was working on a design for a 9600 baud modem at a California defense firm. Gallager suggested that the firm acquire Holsinger and his research. As a result, in 1967 Codex obtained the preliminary hardware design for what eventually became their first modem: the AE-96.

The AE-96, released in 1968, was a "large beast that cost $23,000"[30] remembers Professor Forney. Although it was a bulky piece of equipment, not terribly reliable, and it was cheaper to buy a car, the AE-96 was the world's first 9600 baud modem, and it generated a lot of interest.

One of the biggest hurdles in commercializing modems and fax machines was the rule of law. It was illegal to attach anything to AT&T's phone lines – unless it was made by AT&T. This injunction had effectively crippled any incentive to produce commercial telecommunication devices. What would be the point of making a device no one could use? But on June 26th in the same year that the AE-96 made its debut, the FCC ruled that it would henceforth be legal to use non-AT&T modems on public telephone lines. This decision opened up the telecommunication field to any company interested in making equipment. The effect on the industry was phenomenal. By 1970, Codex, and many others, had shifted their focus entirely to the commercial sector.

At the time, Professor Gallager and others had been researching the theory and structure of Quadrature Amplitude Modulation, or QAM, a technique that could be valuable in modem design. QAM is a method of combining two signals into a single channel, thus effectively doubling bandwidth.

Figure 9. Codex QAM 9600

Codex used this research to implement a 9600 baud QAM modem which came out in 1971. In this we see the face of things to come: technology was beginning to progress to the point where current Information Theory research could actually be quickly applied. The QAM 9600 was a phenomenal worldwide success. In fact, the international 9600 baud modem standard, V.29, "was to a large extent written around our 9600 baud modem[31]" says Professor Forney.

Linkabit
Another pioneering company that applied Information Theory concepts was Linkabit, founded by Professor Irwin Jacobs from MIT, and Professors Andrew Viterbi and Len Kleinrock from UCLA in 1968. Linkabit was a part-time consulting firm that started in the field of coding and

---

[30] *A Conversation with Forney, D.* IEEE IT Society Newsletter, June 1997
[31] ibid

decoding. However they soon changed their focus to satellite communications, and time-division multiple-access (TDMA) systems[32].

Like Codex, Linkabit was quick to try and apply theoretical ideas. One of the major advances in coding theory was the Viterbi decoding, an optimal algorithm for decoding convolutional codes developed by Andrew Viterbi in 1967. Sequential decoding, the tried-and-true technique was known to be sub-optimal. The Viterbi algorithm, too demanding for the technology of the late sixties, was first implemented by Linkabit: they designed one of the very first Viterbi decoders, implemented it in VLSI, and applied it to military systems. "The 70s was the age of the Viterbi algorithm", says Proakis[33] emphasizing how important the Viterbi algorithm is in today's communication systems.

Also like Codex, Linkabit eventually shifted from research and military work to the commercial sector. They developed the first commercial wireless TDMA phone, and the first commercial encrypted satellite television system, VideoCypher.

Qualcomm

One of the pioneering companies that commercialized spread spectrum technology was Qualcomm. Qualcomm was founded by Viterbi and Jacobs, who left Linkabit in 1985. Their initial focus was on digital wireless communication, and they developed CDMA for wireless voice communication. CDMA is a spread spectrum technique that allows more people to share frequency bands at the same time without cross-talk, static or interference. [Qualcomm]. Qualcomm developed CDMA technology that could accommodate multiple users on the same channel. Before then, spread spectrum was considered wasteful of bandwidth, as it used up a range of frequencies. Today, CDMA is a standard for wireless communication and is used globally.

Current applications

The advances in hardware technology have two important impacts on Information Theory. It has provided the computation power to realize complicated coding schemes. Incredibly complex schemes such as CDMA can today be implemented on a single chip. It has also made coding affordable and readily available. For example, a Reed-Solomon decoder now sits on everyone's shelf inside a CD player.

Information theory has innumerable applications today. CDMA is still being used and researched to improve voice and data communications systems. Modern applications of spread spectrum range from low speed fire safety devices to high speed wireless local area networks. Storage devices, such as hard disks and RAM, also employ Information Theory concepts. Using Reed-Solomon codes for compression, and Hamming codes to correct errors, major breakthroughs have been made, allowing gigabits of information to be stored on inches of space.

---

[32] RLE Currents, *Mass. Inst. Tech.*, Spring 1996
[33] Proakis, J. Interview

Information theory's long shadow falls over many more of the things that have become commonplace today. It has strongly influenced not only the development of wireless systems, CDs, and data storage, but also computer networks, satellites, optical communication systems, mobile phones, MP3s, JPEGs, and of course, the Internet. These are perhaps the areas that will continue to use the ideas of Information Theory in the future.

Coding is dead…again

Ironically, this past June of 2001, at the IEEE International Symposium on Information Theory (ISIT), information theorists gathered once again and proclaimed the field to be dead. This is nothing new. As Professor Forney said[34], "the field has been declared dead again and again. I think that there is a moral in that, because each time that it's declared dead, something new comes along."

But this time, the field is being declared dead for a new reason. It is not because of a limitation in technology, nor is it because channel capacity can be theoretically attained. It is being declared dead because channel capacity has actually been *reached* in practice, by using low density parity check codes and turbo codes. The limit promised by Shannon has finally been fulfilled.

Is coding, then, truly dead this time? Is this time for real? It would be fairer to say, that rather than dying, it has accomplished what it had set out to do: it has attained the Shannon limit for the Additive White Gaussian noise channel.

The key, however, is that capacity has *only* been met for Additive White Gaussian noise channels. These channels are now virtually completely understood. Commercial modems are available that reach very near the Shannon limit for these channels. But there are still problems to be tackled, channels for which the limits have yet to be reached, such as wireless channels and multisource-multipath channels.

So although it might be said that parts of Information Theory are "seriously mature[35]" a healthy indication of the state of Information Theory today is the sizable volume of the last issue of *IEEE Transactions on Information Theory*. Papers are still being written, and research still needs to be done. Princeton Professor and former president of the IEEE Information Theory Society, Sergio Verdu says, "Maybe the day will come when a software package will enable the engineer to closely approach the capacity of almost any channel with the technology of the day. Admittedly, I am afraid it is us who will be dead when that day arrives![36]"

Ultimately, who can say what is to come? "Every two years the boundary between 'feasible' and 'infeasible' advances by another factor of two."[37] If history is any indication, Information Theory and the legacy of Claude Shannon have many more productive years ahead of them.

---

[34] Transcript, *IEEE Workshop on Information Theory - 'Shannon Theory: Present and Future'* December 1994.
[35] McEliece, R. *International Symposium on Information Theory, June 2001*, quoted in Forney & Proakis interviews.
[36] Transcript, *IEEE Workshop on Information Theory - 'Shannon Theory: Present and Future'* December 1994.
[37] Forney, D. *Shannon Lecture,* 1995. IEEE IT Society Newsletter, Summer '98 : 21

DISCUSSION AND SUMMARY

According to Christensen[38], emerging technologies had difficulties being funded. Why was Information Theory not the case? The military heavily funded Information Theory research because they believed in its potential in communications, their particular area of interest. Researchers tried to tap this military interest by attaching the phrase "Information Theory" to their work, related or not. This effect was not confined to the U.S; as the fame of Information Theory spread, the phenomenon occurred in Russia. "They were up to it too![39]" exclaimed Professor Fano.

The generalized nature of the Tri-services contract also helped fund Information Theory research, especially at MIT, because there were fewer constraints on how the money was used. They believed that Information Theory would revolutionize communications, but their primary interest was the generation of a pool of individuals skilled in communications engineering and electronics. Professor Fano discuses how successful they were in this particular goal:

> I finished my doctorate in June 1947. I was the fiftieth doctorate granted by the
> Electrical Engineering department. In a few years after the contract, we were
> graduating that many doctoral students in a year…
> [Interview with Fano, R. 2001]

We have traced the evolution of Information Theory from its beginnings to its present state. We have seen the many different influences that have shaped its progress, making what it is today. As we have discussed, social factors such as the personal beliefs of the influential PGIT members, such as Shannon and Elias, helped shape the boundaries of the field. They controlled the *IRE Transactions* and what was published in it. By dominating the forum of discussion itself, they influenced the outcome of the debate. Moreover, the interests of the purists were aligned with those of the military who wanted research in communications. Together, they were easily able to quell dissent.

As we have seen, political influences such as Sputnik and the Space Race helped change the focus of the field towards channel coding. Hardware technological limitations too controlled the progress of the field, arresting it at times and lending it impetus at others.

Through it all, Shannon's prophesized limit was driving progress in Information Theory.

> For 50 years, people have worked to get to the channel capacity he said was
> possible. Only recently have we gotten close. His influence was profound.
> [Lucky, R., quoted in Technology Review, Jul 2001]

MacKenzie[40] says that a technological trajectory is like a self-fulfilling prophecy [168]. Analogous to Moore's Law, Shannon's limit defines a roadmap, giving people an idea of where they are relative to where they can be. Before Shannon, the limits were unknown and there was

---

[38] Christensen, *The Innovators Dilemma*
[39] Fano, R. Interview
[40] MacKenzie, D. *Inventing Accuracy,*168

no clear motivation to push them. His vision provided the impetus to innovate. Once he explained the limits and defined the horizons, people strove to attain them. Throughout the background of our story, Shannon has been there, his Limit pulling the trajectory of Information Theory towards ever increasing performance.

BIBLIOGRAPHY

Abraham, L.G. Letter to Robert Fano. 1954.

Administrative Committee, The. "A Statement of Editorial Policy." *IRE Transactions on Information Theory*. December 1959: 136.

Bates, J.A.V. "Significance of Information Theory to Neurophysiology." *IRE Transactions on Information Theory*. February 1953: 137-142.

Betty Shannon. Telephone Interview. 30 November 2001.

Blachman, Nelson. "The Third London Symposium on Information Theory." *IRE Transactions on Information Theory*. March 1956: 17

Caiainiello, E.R. Letter to Jerome B. Wiesner. 11 March 1958.

CDMA Development Group. *What is CDMA?* 2000.
          <http://www.cdg.org/tech/about_cdma.asp>.

Cheatham, Thomas P. Jr. "A Broader Base for the PGIT." *IRE Transactions on Information Theory*. December 1958: 135.

Chiu, Eugene, et al. "Mathematical Theory of Claude Shannon." 13 December 2001.

Christensen, *The Innovators Dilemma,*HarperBusiness 2000.

Committee in Communication Sciences at the University of Michigan. *Summary Report of the Program in Communication Sciences at the University of Michigan*. April 15, 1960.

Costello et al. "Applications of Error-Control Coding." *IEEE Transactions on Information Theory*. October 1998: page ????.

De Rosa, L.A. "In Which Fields Do We Graze?" *IRE Transactions on Information Theory*. December 1955: 2

Eden, Murray. Personal Interview. 5 November 2001.

Elias, Peter. Letter to Andrew E. Ford. 3 Sept 1959.

Elias, Peter. "Two Famous Papers." *IRE Transactions on Information Theory*. September 1958: 99

Fano, Robert M. "Preliminary Announcement – Symposium on Information Theory." March 19, 1954.

Fano, Robert M. Personal Interview. 1 November 2001.

Fano, Robert M. Personal Interview. 15 November 2001.

Fano, Robert M. 6.574 lecture notes. 1950.

Forney, David G. Personal Interview. 29 November 2001.

Forney, David G. "1995 Shannon Lecture." *IEEE Information Theory Society Newsletter*. Summer 1998: 21.
         <http://golay.uvic.ca/nltr/96_mar/06sha.pdf>.

Frontier Broadband. *History of Spread Spectrum Technology page*. 30 November 2001.
         <http://www.frontierbb.com/benefits_history.html>.

Gabor, D. "Communication Theory Past, Present and Prospective." *IRE Transactions on Information Theory*.
         February 1953: 2.

Gallager, Robert. Personal Interview. 19 November 2001.

Gallager, Robert G. "Claude E. Shannon: A Retrospective on His Life, Work, and Impact." *IEEE Transactions on
         Information Theory*. November 2001: 2681-2695.

Hochfelder, D. "An oral history with Dr. Paul Green." IEEE History Center. 15 October 1999.
         <http://www.ieee.org/organizations/history_center/oral_histories/transcripts/pgreen.html>.

Huffman, Ken. *David Huffman Elegy page*.
         <http://home.columbus.rr.com/kenhuffman/family/elegy.html>.

Huggins, W.H. Letter to Commander Gould Hunter. 1954.

IEEE Information Theory Society Newsletter. "IEEE Workshop on Information Theory
         'Shannon Theory: Present and Future.'" December 1994.

IEEE Information Theory Society Newsletter. "A Conversation with G. David Forney Jr." June 1997.

Jachimczyk, W. Worcester Polytechnic Institute. *Spread spectrum page*.
         <http://www.ece.wpi.edu/courses/ee535/hwk11cd95/witek/witek.html>.

Jamil, Asad. "A simple convolutional encoder and Viterbi decoder."
         <http://www.ecel.ufl.edu/~ajamil/report400.html>.

Latour, B. *Science In Action*, Harvard University Press, 1987

Massey, J.L. "Deep-space communication and coding.", in *Lecture Notes on Control and Information Sciences 82*
         J. Hagenauer, Ed.. Bonn, Germany: Springer-Verlag, 1992.

Mueller, G.E. " Information Theory Progress Report 1960-1963: Space Communication." *IEEE
         Transactions on Information Theory*. October 1963: 257-264.

MacKenzie, D. *Inventing Accuracy,* MIT Press 1993.

Peterson, W.W. and J. Massey. " Information Theory Progress Report 1960-1963: Coding
         Theory." *IEEE Transactions on Information Theory*. October 1963: 223-229.

Proakis, John. Personal Interview. 16 November 2001.

Qualcomm. *About Page*. <http://www.qualcomm.com/about/>.

Reynolds, John H. Letter to Robert Fano. 27 February 1973.

Riley, Laura J. 4i2i Communications Ltd. "Convolutional Coding." 11 October 2001.
<http://www.4i2i.com/viterbi.htm>.

RLE Currents. "On the Road to Success with RLE's Alumni Company Founders." Spring 1996.
<http://rleweb.mit.edu/Publications/currents/8-1cov.htm#jaco>.

Ross, Arthur H. M. *Introduction to CDMA*. 11 June 1996.
<http://www.amug.org/~ahmrphd/Intro.html>.

Shannon, Claude E. "The Bandwagon." *IRE Transactions on Information Theory*. March 1956: 3

Shannon, Claude E. and Warren Weaver. "The Mathematical Theory of Communication."
Urbana, Illinois: University of Illinois Press, 1949.

Stephenson, Robert C. Letter to Robert Fano. 13 July 1965.

Stumpers, F. Louis H.M. "Information Theory and International Radio Organizations." *IRE Transactions on Information Theory*. June 1957: 85.

Tuller, W.C. Letter to L.G. Abraham. 11 January 1954.

Viterbi, Andrew J. "Information Theory in the Sixties." *IEEE Transactions on Information Theory*. May 1973: 257-262.

Waldrop, Mitchell M. "Claude Shannon: Reluctant Father of the Digital Age." *Technology Review*. July/Aug 2001. < http://www.techreview.com/magazine/jul01/waldrop.asp>.

Wiener, Norbert. "What is Information Theory?". *IRE Transactions on Information Theory*. June 1956: 48

Wolf, Jack K. "A Survey of Coding Theory: 1967-1972." *IEEE Transactions on Information Theory*. July 1973: 381-388.

Yankopolus, Andreas. "Capsule History of Error Control Coding."
<http://www.csc.gatech.edu/~yank/research/history.pdf>.