The Scientific Conceptualization of Information: A Survey

WILLIAM ASPRAY

The article surveys the development of a scientific conceptualization of information during and in the decade following World War II. It examines the roots of information science in nineteenth- and early twentieth-century mathematical logic, physics, psychology, and electrical engineering, and then focuses on how Warren McCulloch, Walter Pitts, Claude Shannon, Alan Turing, John von Neumann, and Norbert Wiener combined these diverse studies into a coherent discipline.

*Categories and Subject Descriptors: E.4 [**Coding and Information Theory**]; F.1 [**Computation by Abstract Devices**]: Models of Computation, Modes of Computation; F.4.1 [**Mathematical Logic and Formal Languages**]: Mathematical Logic—computability theory, recursive function theory; H.1.1 [**Models and Principles**]: Systems and Information Theory—information theory; K.2 [**History of Computing**]—people*
General Terms: Theory
Additional Key Words and Phrases: W. McCulloch, W. Pitts, C. Shannon, A. Turing, J. von Neumann, N. Wiener

Modern scholarship has tended to equate the history of information processing with the history of computing machinery. Because of the phenomenal growth of a new generation of more powerful machines every few years, other important events in information-processing history have been overshadowed. One such event is the scientific conceptualization of information that occurred during and in the decade following World War II. In that period a small group of mathematically oriented scientists developed a theory of information and information processing. For the first time, information became a precisely defined concept amenable to scientific study. Information was given the status of a physical parameter, such as entropy, which could be quantified[1] and examined using mathematical tools. Around this concept grew a number of research areas involving study of both machines and living organisms. They included the mathematical theory of communication, mathematical modeling of the brain, artificial intelligence, cybernetics, automata theory, and homeostasis.

Of course, the word *information* was in common usage for many years before its scientific conceptualization. It was recorded in print in 1390 to mean "communication of the knowledge or 'news' of some fact or occurrence" (Oxford English Dictionary). Information also found a place in the traditional scientific discourse of physics, mathematical logic, electrical engineering, psychology, and biology—in some instances as early as the nineteenth century. These disciplines provided the avenues to the study of information for the scientists reformulating the concept during and after World War II.

The pursuit of five research areas in particular led to the scientific conceptualization of information.

Author's Address: Charles Babbage Institute, 104 Walter Library, University of Minnesota, Minneapolis, MN 55455.
© 1985 AFIPS 0164-1239/85/020117-140$01.00/00

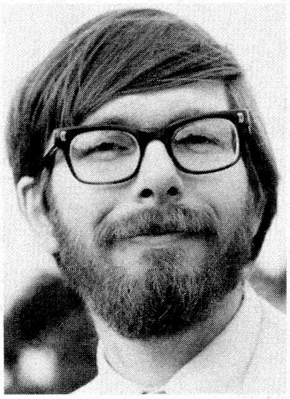
[1] Quantification is not universally possible in information science. For example, it is possible to quantify coding counts, but not most semantic concerns.

1. James Clerk Maxwell, Ludwig Boltzmann, and Leo Szilard's work in thermodynamics and statistical mechanics, especially on the concept of entropy and its mathematical formulation.
2. The emergence of control and communication as a new branch of electrical engineering, supplementary to power engineering, as a result of the development of telegraphy, radio, and television.
3. The study of the physiology of the nervous system, beginning in the nineteenth century with the work of H. von Helmholtz and Claude Bernard, and continuing into the twentieth century, especially with the work of Walter Cannon on homeostasis and the internal regulation of living organisms.
4. The development of functionalist and behaviorist theories of the mind in psychology, leading to a view of the brain as a processor of information and to a demand for experimental verification of theories of mind through observation of external behavior.
5. The development of recursive function theory in mathematical logic as a formal, mathematical characterization of the human computational process.

What was new after the war was a concerted effort to unify these diverse roots through a common mathematical characterization of the concepts of information and information processing. The seminal idea was that an interdisciplinary approach is appropriate to solve problems in both biological and physical settings in cases where the key to the problems is the manipulation, storage, or transmission of information and where the overall structure can be studied using mathematical tools. For these scientists, both the human brain and the electronic computer were considered types of complicated information processors whose similar laws of functioning could be better understood with the help of the abstract results deduced from the mathematical models of automata theory.

The major figures in this movement were Claude E. Shannon, Norbert Wiener, Warren S. McCulloch, Walter Pitts, Alan M. Turing, and John von Neumann. They came to the subject from various established scientific disciplines: mathematics, electrical engineering, psychology, biology, and physics. Despite their diverse backgrounds and research specialties, they functioned as a cohesive scientific community. They shared a common educational background in mathematical logic. Most had wartime experience that sharpened their awareness of the general importance of information as a scientific concept. Each appreciated the importance of information to his own research. Each was familiar with the others' work and recognized its importance to his own work and to the science of information generally. Most were personally acquainted,[2] and often collaborated with or built directly upon the work of the others. They reviewed one another's work in the scientific literature[3] and often attended the same conferences or meetings—some of which were designed specifically to study this new discipline. Typical was a 1944 conference in Princeton organized by Wiener and von Neumann for mathematicians, engineers, and physiologists to discuss problems of mutual interest in cybernetics and computing.[4] Outsiders recognized them as forming a cohesive community of scholars devoted to a single area of research.

The introduction to Wiener's *Cybernetics* (1948) describes the sense of community and common purpose among these diversely trained scientists. Perhaps



William Aspray (B.A., M.A. Wesleyan, 1973; M.A. 1976, Ph.D. 1980 Wisconsin) is associate director of the Charles Babbage Institute for the History of Information Processing at the University of Minnesota in Minneapolis. Prior to joining CBI Aspray taught mathematical science at Williams College and history of science at Harvard University. His historical research interests include modern mathematics and logic, theoretical computer science, social diffusion of computing, computer education, and the economics of the computer industry.

---

[2] It is not clear whether Pitts knew Turing personally. Wiener, McCulloch, and Pitts were associates at MIT. Von Neumann was involved in several conferences with McCulloch and Wiener. Turing worked with von Neumann in Princeton in 1937 and 1938, and with Shannon during the war at the Bell Telephone Laboratories in New York City. Wiener and McCulloch both visited Turing in England. Shannon had ample opportunity to become personally acquainted with Wiener, von Neumann, McCulloch, and Pitts, and he was certainly familiar with their work.

[3] See, for example, Shannon's reviews of four papers authored or coauthored by Pitts, including the famous McCulloch and Pitts paper, in *Mathematical Reviews 5* (1944), p. 45, and *6* (1945), p. 12.

[4] Turing, as a confirmed loner and the only non-American among the major pioneers, is the least likely to have had contact with the group. Nevertheless, he discussed artificial-intelligence issues with Shannon (in fact, both had developed chess programs); Wiener sought him out in England and regarded him as one of the cyberneticists; McCulloch also went out of his way to visit in Manchester, although there is reason to believe that Turing did not think highly of McCulloch (Hodges 1983); von Neumann offered Turing a job as his assistant at the Institute for Advanced Study to continue his theoretical work (a position Turing refused). Turing was also in contact with others in England working in this area, most notably Ross Ashby and Grey Walter.

even more telling were both Wiener's attempts to develop an interdisciplinary science, known as "cybernetics," around the concept of feedback information, and von Neumann's attempts to unify the work of Shannon, Turing, and McCulloch and Pitts into a general theory of automata.

Mathematical logic does not have many real-world applications. It did here because it studies the laws of thought in abstraction, and more particularly because from the 1930s logicians were concerned with finding a mathematical characterization of the process of computation.

In fact, training in mathematical logic was the most salient tie among these early pioneers. Shannon completed a master's thesis (Shannon 1940) in electrical engineering at the Massachusetts Institute of Technology on the application of mathematical logic to the study of switching systems. Wiener studied mathematical logic with Bertrand Russell and averred its influence on his later work in cybernetics (Wiener 1948). Turing received his doctoral degree from Princeton University for work on ordinal logics (Turing 1938). Pitts studied mathematical logic under Rudolf Carnap, while his associate, McCulloch, was a physiological psychologist interested in questions concerning the learning of logic and mathematics. Early in his career, von Neumann contributed significantly to the two branches of logic known as "proof theory" and "set theory."

The similarity in their work does not end with the common use of mathematical logic to solve problems in a variety of fields. They also shared the conviction that the newly discovered concept of information could tie together, in a fundamental way, problems from different branches of science. While the content of their work shows this common goal,[5] so does the social organization of their research. Despite widely diverse backgrounds and research interests, these scientists were in close contact through collaboration, scholarly review of one another's work, and frequent interdisciplinary conferences.

The growth of this interdisciplinary science in the late 1940s and early 1950s was at least partially the product of the massive cooperative and interdisciplinary scientific ventures of World War II that carried many scientists to subjects beyond the scholarly bounds of their specialties. At no previous time had there been such a mobilization of the scientific community. Wiener was led to develop cybernetics at least partly on account of his participation in Vannevar E. Bush's computing project at MIT, his work with Y. W. Lee on wave filters, and his collaboration with Julian Bigelow on fire control for antiaircraft artillery. Each of these projects was related to the war effort. Both Turing and von Neumann applied wartime computing experience to their postwar work on artificial intelligence and automata theory. Shannon's theory of communication resulted partly from the tremendous advances in communications engineering spawning from the development of radar and electronics and partly from the need for secure communications during the war.

The focus in this paper is on the sources and contributions of the six most important early figures in this movement: Shannon, Wiener, McCulloch, Pitts, Turing, and von Neumann. It is impossible in a work of this length to present an exhaustive treatment of the contributions of any of these pioneers, or even to mention the less centrally related work of their colleagues. Instead, the intent is to sketch the general outlines of this new conceptualization, with hope that others will contribute the fine brushstrokes necessary to complete the picture.

Claude Shannon and the Mathematical Theory of Communication

While working at Bell Laboratories in the 1940s on communication problems relating to radio and telegraphic transmission, Shannon developed a general theory of communication that would treat of the transmission of any sort of information from one point to another in space or time.[6] His aim was to give specific technical definitions of concepts general enough to obtain in any situation where information is manipulated or transmitted—concepts such as information, noise, transmitter, signal, receiver, and message.

At the heart of the theory was a new conceptualization of information. To make communication theory a scientific discipline, Shannon needed to provide a precise definition of information that transformed it into a physical parameter capable of quantification. He accomplished this transformation by distinguishing information from meaning. He reserved "meaning" for the content actually included in a particular message. He used "information" to refer to the number of different possible messages that could be carried along a channel, depending on the message's length and on the number of choices of symbols for transmission at

---

[5] An article by E. Colin Cherry (1952) shows that the unity of this work was already recognized by outsiders.

[6] Note that Shannon had already begun to work on his theory of communication prior to his arrival at Bell Labs in 1941 (Shannon 1949*a*), and he continued to develop the theory while there. Witness his later paper (Shannon 1949*b*), originally written as a Bell Confidential Report based on work in the labs during the war (Shannon 1945). For more information, see the biography of Turing by Hodges (1983).

Selective Chronology of Events in the Scientific Conceptualization of Information

1390	First recorded printed use of word *information*
1840s–1860s	Helmholtz, investigations of the physiology and psychology of sight and sound
1843–1877	Bernard, work on homeostasis
1868	Maxwell, "On Governors"
1881	Ribot, *Diseases of Memory*
1890	James, *Principles of Psychology*
1891	Waldeyer, "neurone" theory
1894	Boltzmann, work on statistical physics
1906	Sherrington, *The Integrative Action of the Nervous System*
1915	Holt, *The Freudian Wish*
1919	Watson, *Psychology From the Standpoint of a Behaviorist*
1919–1923	McCulloch, work on logic of transitive verbs
1923	Lashley, "The Behaviorist Intepretation of Consciousness"
1924	Nyquist, "Certain Factors Affecting Telegraph Speed"
1925	Szilard, work on entropy and statistical mechanics
1928	Hartley, "Transmission of Information"
1932	Von Neumann, *Foundation of Quantum Mechanics*
Mid-1930s	Development of recursive function theory (work of Gödel, Church, Kleene, Post, Turing)
Mid-1930s	Feedback concept studied by electrical engineers (work of Black, Nyquist, Bode)
Mid-1930s	Development of electroencephalography
1937	Turing, "On Computable Numbers"
1937	Turing and von Neumann, first discussions about computing and artificial intelligence at Princeton University
1938	Shannon, "Symbolic Analysis of Relay and Switching Circuits" (published 1938, M.A. Thesis 1940)
Late-1930s	Harvard Medical School seminar led by Cannon
1940	Wiener, work at MIT on computers
1940	Shannon, "Communication in the Presence of Noise" (submitted; published 1948)
1942	Macy Foundation meeting on central inhibition in the nervous system
1942	Bigelow, Rosenblueth, Wiener, "Behavior, Purpose, Teleology"
1943	Turing, work with Shannon and Nyquist at Bell Laboratories in New York City
1943	McCulloch and Pitts, "Logical Calculus of the Ideas Immanent in Nervous Activity"
1943	Pitts accepts position at MIT to work with Wiener
1944	Princeton conference organized by Wiener and von Neumann on topics related to computers and control
1945	Von Neumann, draft report on EDVAC
1945	Shannon, "A Mathematical Theory of Cryptography" (Bell Laboratories report, published in 1949 under different title)
1945	McCulloch, "A Heterarchy of Values Determined by the Topology of Nerve Nets"
1946	Macy Foundation meetings on feedback organized by McCulloch
1946	Turing, NPL report on ACE computer
1946–1948	Burks, Goldstine, von Neumann, IAS computer reports
1947	Ashby, "The Nervous System as Physical Machine"
1947	McCulloch and Pitts, work on a prosthetic device to enable the blind to read by ear
1947	McCulloch and Pitts, "How We Know Universals"
1948	Wiener, *Cybernetics*
1948	Turing, "Intelligent Machinery" (NPL Report)
1948	Turing, first work programming the Manchester computer to carry out "purely mental activities"
1948	Von Neumann, "General and Logical Theory of Automata" (Hixon Symposium, Pasadena)
1948	Shannon, "The Mathematical Theory of Communication"
1949	Von Neumann, "Theory and Organization of Complicated Automata" (University of Illinois Lectures)
1949	Shannon and Weaver, *The Mathematical Theory of Communication*
1950	Shannon, "Programming a Computer to Play Chess"
1950	McCulloch, "Machines That Think and Want"
1950	McCulloch, "Brain and Behavior"
1950	Turing, "Computing Machinery and Intelligence"
1950	Ashby, "The Cerebral Mechanism of Intelligent Action"
1952	Turing, "The Chemical Basis of Morphogenesis"
1952	McCulloch, accepts position at MIT to be with Pitts and Wiener
1952	Ashby, *Design for a Brain*
1952	Von Neumann, "Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components" (California Institute of Technology Lectures)
1952–1953	Von Neumann, "The Theory of Automata: Construction, Reproduction, Homogeneity"
1953	Walter, *The Living Brain*
1953	Turing, "Digital Computers Applied to Games: Chess"
1953	Turing, "Some Calculations on the Riemann Zeta Function"
1956	Von Neumann, *The Computer and the Brain* (prepared for the Silliman Lectures, Yale; published 1958)
1956	Ashby, *Introduction to Cybernetics*
1956	Ashby, "Design for an Intelligence Amplifier"

each point in the message. Information in Shannon's sense was a measure of orderliness (as opposed to randomness) in that it indicated the number of possible messages from which a particular message to be sent was chosen. The larger the number of possibilities, the larger the amount of information transmitted, because the actual message is distinguished from a greater number of possible alternatives.

Shannon admitted the importance of previous work in communications engineering to his interest in a general theory of information, in particular the work of Harry Nyquist and R. V. Hartley.

> The recent development of various methods of modulation such as PGM and PPM which exchange bandwidth for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis

> for such a theory is contained in the important papers of Nyquist and Hartley on this subject. In this paper we will extend the theory to include a number of new factors. (Shannon 1948, pp. 31–32)

Nyquist was conducting research at Bell Laboratories on the problem of improving transmission speeds over telegraph wires when he wrote a paper on the transmission of "intelligence" (Nyquist 1924).[7] This paper concerned two factors affecting the maximum speed at which intelligence can be transmitted by telegraph: signal shaping and choice of codes. As Nyquist stated,

> The first is concerned with the best shape to be impressed on the transmitting medium so as to permit of greater speed without undue interference either in the circuit under consideration or in those adjacent, while the latter deals with the choice of codes which will permit of transmitting a maximum amount of intelligence with a given number of signal elements. (Nyquist 1924, p. 324)

While most of Nyquist's article considered the practical engineering problems associated with transmitting information over telegraph wires, one theoretical section was of importance to Shannon's work, entitled "Theoretical Possibilities Using Codes with Different Numbers of Current Values." In this section Nyquist presented the first logarithmic rule governing the transmission of information.

Nyquist proved that the speed at which intelligence can be transmitted over a telegraph circuit obeys the equation

$$W = k \log m$$

where W is the speed of transmission of intelligence, m is the number of current values that can be transmitted, and k is a constant. He also prepared the following table by which he illustrated the advantage of using a greater number of current values for transmitting messages (Nyquist 1924, p. 334).

Number of Current Values	Relative Amount of Intelligence that Can Be Transmitted with the Given Number of Signal Elements
2	100
3	158
4	200
5	230
8	300
16	400

Although Nyquist's work was primarily empirical and concerned with engineering issues—and he used the term *intelligence*, which masked the difference between information and meaning—his work was important for presenting the first statement of a logarithmic law for communication and the first examination of the theoretical bounds for ideal codes for the transmission of information. Shannon later gave a more general logarithmic rule as the fundamental law of communication theory, which stated that the quantity of information is directly proportional to the logarithm of the number of possible messages. Nyquist's law became a specific case of Shannon's law, because the number of current values is directly related to the number of symbols that can be transmitted. Nyquist was aware of this relation, as his definition of speed of transmission indicates.

> By the speed of transmission of intelligence is meant the number of characters, representing different letters, figures, etc., which can be transmitted in a given length of time. (Nyquist 1924, p. 333)

By "letters, figures, etc." he meant a measure proportional to what Shannon later would call "bits of information." Nyquist's table, listing the *relative* amount of intelligence transmitted, illustrates the gain in information consequent to a greater number of possible choices. That he listed the *relative* amount indicates his awareness that there is an important relation between the number of figures and the amount of intelligence (information) being transmitted. This relation is at the heart of Shannon's theory of communication. Nyquist did not generalize his concept of "intelligence" beyond telegraphic transmissions, however.

Shannon's other predecessor in information theory, Hartley, was also a research engineer at Bell Laboratories. Hartley's intention was to establish a quantitative measure to compare capacities of various systems to transmit information. His hope was to provide a theory general enough to include telegraphy, telephony, television, and picture transmission—communications over both wire and radio paths. His investigation (Hartley 1928) began with an attempt to establish theoretical limits of information transmission under idealized situations. This important step led him away from the empirical studies of engineering adopted by most earlier researchers and toward a mathematical theory of communication.

Before turning to concrete engineering problems, Hartley addressed "more abstract considerations." He began by making the first attempt to distinguish a notion of information amenable to use in a scientific context. He realized that any scientifically usable definition of "information" should be based on what he called "physical" instead of "psychological" considerations. He meant that information is an idea involving

---

[7] Nyquist worked closely with Shannon during the war. He also had discussions about communications theory and engineering with Turing while Turing was a visitor at the labs in 1943. See Hodges (1983) for details.

a quantity of physical data and should not be confused with the meaning of a message.

> The capacity of a system to transmit a particular sequence of symbols depends upon the possibility of distinguishing at the receiving end between the results of the various selections made at the sending end. The operation of recognizing from the received record the sequence of symbols selected at the sending end may be carried out by those of us who are not familiar with the Morse code. We could do this equally well for a sequence representing a consciously chosen message and for one sent out by the automatic selecting device already referred to. A trained operator, however, would say that the sequence sent out by the automatic device was not intelligible. The reason for this is that only a limited number of the possible sequences have been assigned *meanings* common to him and the sending operator. Thus the number of symbols available to the sending operator at certain of his selections is here limited by psychological rather than physical considerations. Other operators using other codes might make other selections. *Hence in estimating the capacity of the physical system to transmit information we should ignore the question of interpretation, make each selection perfectly arbitrary, and base our results on the possibility of the receiver's distinguishing the result of selecting any one symbol from that of selecting any other.* By this means the psychological factors and their variations are eliminated and it becomes possible to set up a definite quantitative measure of information based on physical considerations alone. (Hartley 1928, pp. 537–538; emphasis added)

Thus Hartley distinguished between psychological and physical considerations—that is, between meaning and information. The latter he defined as the number of possible messages, independent of whether they are meaningful. He used this definition of information to give a logarithmic law for the transmission of information in discrete messages:

$$H = K \log s^n$$

where H is the amount of information, K is a constant, n is the number of symbols in the message, s is the size of the set of symbols, and therefore s^n is the number of possible symbolic sequences of the specified length n. This law included the case of telegraphy and subsumed Nyquist's earlier law. Once the law for the discrete transmission of information had been established, Hartley showed how it could be modified to treat continuous transmission of information, as in the case of telephone voice transmission.

Hartley turned next to questions of interference and described how the distortions of a system limit the rate of selection at which differences between transmitted symbols may be distinguished with certainty. His special concern was with the interference caused by the storage and subsequent release of energy through induction and capacitance, a source of noise of great concern to electrical engineers at the time. He found that the total amount of information that could be transmitted over a steady-state system of alternating currents limited to a given frequency-range is proportional to the product of the frequency-range on which it transmits and the time during which it is available for transmission.[8]

Hartley had arrived at many of the most important ideas of the mathematical theory of communication: the difference between information and meaning, information as a physical quantity, the logarithmic rule for transmission of information, and the concept of noise as an impediment in the transmission of information.

Hartley's aim had been to construct a theory capable of evaluating the information transmitted by any of the standard communication technologies. Starting with these ideas, Shannon developed a general theory of communication, not restricted to the study of technologies designed specifically for communication. Later, Shannon's collaborator, Warren Weaver, described the theory clearly.

> The word *communication* will be used here in a very broad sense to include all of the procedures by which one mind may affect another. This, of course, involves not only written and oral speech, but also music, the pictorial arts, the theatre, the ballet, and in fact all human behavior. In some connections it may be desirable to use a still broader definition of communication, namely, one which would include the procedures by means of which one mechanism (say automatic equipment to track an airplane and to compute its probable future positions) affects another mechanism (say a guided missile chasing this airplane). (Shannon and Weaver 1949, Introductory Note)

What began as a study of transmission over telegraph lines was developed by Shannon into a general theory of communication applicable to telegraph, telephone, radio, television, and computing machines—in fact, to any system, physical or biological, in which information is being transferred or manipulated through time or space.

In Shannon's theory a communication system consists of five components related to one another, as illustrated in Figure 1 (Shannon and Weaver 1949, p. 34). These components are:

> 1. An information source which produces a message or sequence of messages to be communicated to the receiving terminal. . . .
> 2. A transmitter which operates on the message in some way to produce a signal suitable for transmission over the channel. . . .

---

[8] Cherry (1952) reviews the research emanating from or related to Hartley's work.

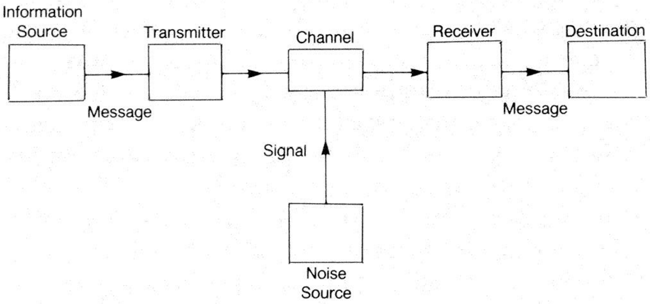


Figure 1. Schematic diagram of a general communication system. (Redrawn from Shannon and Weaver (1949, p. 34).)

> 3. The channel is merely the medium used to transmit the signal from transmitter to receiver. . . .
> 4. The receiver ordinarily performs the inverse operation of that done by the transmitter, reconstructing the message from the signal.
> 5. The destination is the person (or thing) for whom the message is intended.

The importance of this characterization is its applicability to a wide variety of communication problems, provided that the five components are appropriately interpreted. For example, it applies equally well to conversations between humans, interactions between machines, and even to communication between parts of an organism. Properly interpreted, the communication both between the stomach and the brain and between the target and the guided missile could be seen as examples of a communication system. Using Shannon's theory, previously unrecognized connections between the biological and the physical worlds could be unmasked.

While Hartley recognized that a distinction must be drawn between information and meaning, Shannon sharpened the distinction by giving the first definition of information sufficiently precise for scientific discourse. Weaver described the importance of this definition.

> The word *information*, in this theory, is used in a special sense that must not be confused with its ordinary usage. In particular, *information* must not be confused with meaning.
>
> In fact, two messages, one of which is heavily loaded with meaning and the other of which is pure nonsense, can be exactly equivalent, from the present viewpoint, as regards information. . . .
>
> To be sure, this word *information* in communication theory relates not so much to what you do say, as to what you could say. That is, information is a measure of one's freedom of choice when one selects a message. If one is confronted with a very elementary situation where he has to choose one of two alternative messages, then it is arbitrarily said that the information, associated with this situation, is unity. Note that it is misleading (although often convenient) to say that one or the other message conveys unit information. The concept of information applies not to the individual messages (as the concept of meaning would), but rather to the situation as a whole, the unit information indicating that in this situation one has an amount of freedom of choice, in selecting a message, which is convenient to regard as a standard or unit amount. (Shannon and Weaver 1949, pp. 8–9)

Shannon recognized that information could be measured by any increasing monotonic function, provided the number of possible messages is finite. He chose from among these the logarithmic function for the same reason as Hartley: that it accords well with our intuition of what the appropriate measure should be. We intuitively feel that two punched cards should convey twice the information of one punched card. Assuming one card can carry n symbols, two cards will carry n^2 combinations. The logarithmic function then measures the two cards as conveying $\log n^2 = 2 \log n$ bits of information, twice that conveyed by one card ($\log n$). Shannon chose base 2 for his logarithmic measure since $\log_2$ assigns one unit of information to a switch with two positions. Then N two-position switches could store $\log_2 2^N$ $(=N)$ binary digits of information.[9] If there were N equiprobable choices, the amount of information would be given by $\log_2 N$. Shannon generalized this equation to the nonequiprobable situation, where the amount of information H would be given by

$$H = -(p_1 \log_2 p_1 + \ldots + p_n \log_2 p_n)$$

if the choices have probabilities $p_1, \ldots, p_n$.

Shannon recognized that this formulation of information is closely related to the concept of entropy, since the information parameter measures the orderliness of the communication channel.[10]

> Quantities of the form $H = -P_i \log P_i$ (the constant K merely amounts to a choice of a unit of measure) play a central role . . . as measures of information, choice and uncertainty. The form of H will be recognized as that of entropy as defined in certain formulations of statistical mechanics where P_i is the probability of a system being in cell i of its phase space. H is then, for example, the H in Boltzmann's famous H theorem. (Shannon and Weaver 1949)

---

[9] "Binary digits" was shortened to "bits" by John Tukey, a Princeton University professor who also worked at Bell Laboratories (see the *Annals*, Vol. 6, No. 2, April 1984, pp. 152–155). The introduction of a new term such as *bit* is a good indication of the introduction of a new concept.

[10] Shannon refers the reader to Tolman's book (1938) on statistical mechanics.

Entropy has a long history in physics, and during the twentieth century had already become closely associated with the amount of information in a physical system. Weaver carefully credited these roots of Shannon's work.[11]

> Dr. Shannon's work roots back, as von Neumann has pointed out, to Boltzmann's work on statistical physics (1894), that entropy is related to "missing information," inasmuch as it is related to the number of alternatives which remain possible to a physical system after all the macroscopically observable information concerning it has been recorded. L. Szilard (*Zsch. f. Phys.*, Vol. 53, 1925) extended this idea to a general discussion of information in physics, and von Neumann (*Math. Foundation of Quantum Mechanics*, Berlin, 1932, Chap. V) treated information in quantum mechanics and particle physics. (Shannon and Weaver 1949, p. 3, fn)

This close relation of information to entropy is not surprising, for information is related to the amount of freedom of choice one has in constructing messages. This tie between thermodynamics, statistical mechanics, and communication theory suggests that communication theory involves a basic and important property of the physical universe and is not simply a scientific by-product of modern communication technology.

Shannon used his theory to prove several theoretical results about communication systems and to demonstrate applications to the communications industry. He established theoretical limits applicable to practical communication systems.[12] His theory furnished a new definition of *information* that could be applied in a wide variety of physical settings. It provided a mathematical approach to the study of theoretical problems of information transmission and processing. Shannon and Weaver continued the theoretical study of this subject. Meanwhile, the theory provided the basis for interdisciplinary information studies carried out by many others on electronic computing machines and on physical and biological feedback systems.

Norbert Wiener and Cybernetics

While Shannon concentrated mainly on applications of information theory to communications engineering, Wiener stressed its application to control problems involving other physical and complicated biological phenomena. Wiener recognized that several diverse problems he had confronted during the war had yielded to quite similar approaches involving feedback control and communication mechanisms. This realization was the beginning of his new interdisciplinary science, cybernetics, which considered problems of control and communication wherever they occurred.[13] Many of his subsequent scientific projects were designed to illustrate the power of cybernetics in understanding biological functioning.

Wiener recognized the importance of his war-related work to his later development of cybernetics. Elaborating on the conviction he shared with physiologist Arturo Rosenblueth "that the most fruitful areas for the growth of the sciences were those which had been neglected as a no-man's land between the various established fields" (Wiener 1948, p. 8), he wrote:

> We had agreed on these matters long before we had chosen the field of our joint investigations and our respective parts in them. The deciding factor in this new step was the war. I had known for a long time that if a national emergency should come, my function in it would be determined largely by two things: my close contact with the program of computing machines developed by Dr. Vannevar Bush, and my own joint work with Dr. Yuk Wing Lee on the design of electrical networks. In fact, both proved important. (Wiener 1948, p. 9)

In 1940 Wiener began work, in contact with Bush at MIT, on the development of computing machinery for the solution of partial differential equations. One outcome of this project was a proposal by Wiener, purportedly made to Bush, of features to be incorporated into future computing machines. Included were many of the features critical in the following decade to the development of the modern computer: numerical instead of analog central adding and multiplying equipment, electronic tubes instead of gears or mechanical relays for switching, binary instead of decimal representation, completely built-in logical facilities with no human intervention necessary after the introduction of data, and an incorporated memory with capability for rapid storage, recall, and erasure. For Wiener the importance of, and presumably the source of, these suggestions was that "they were all ideas which are of interest in connection with the study of the nervous system" (Wiener 1948, p. 11). Wiener may have been the first to compare explicitly features of the electronic computer and the human

---

[11] In the Szilard article cited by Weaver, "Uber die Entropieverminderung in einem Thermodynamischen System bei Eingriffen intelligenter Wesen," p. 840, Szilard discusses Maxwell's Demon. He points out that the entropy lost by the gas through the separation of the high- and low-energy particles corresponds to the information used by the Demon to decide whether or not to let a particle through the door.

[12] According to Hodges (1983), Bell Labs was beginning to use Shannon's ideas in its research by 1943. Also see Cherry (1952).

[13] Wiener's interest in these subjects had first been piqued before the war in an interdisciplinary seminar he attended at the Harvard Medical School (discussed later in this section). See Wiener's introduction to *Cybernetics* (1948) for details.

brain.[14] His comments certainly illustrated the similarity of structure in diverse settings, which he emphasized later in his cybernetics program.

Another war-related program, undoubtedly the most important to Wiener's formulation of cybernetics, involved the development of fire-control apparatus for antiaircraft artillery. This problem was made urgent at the beginning of the war by the threat of German air attack on the weakly defended English coast. The appreciable increase in velocity of the new German aircraft made earlier methods for directing antiaircraft fire obsolete. Wiener's research suggested that a new device for antiaircraft equipment might effectively incorporate a feedback system to direct future firings. Thus Wiener and Julian Bigelow worked toward a theory of prediction (for the flight of aircraft) and on its effective application to the antiaircraft problem at hand.

> It will be seen that for the second time I had become engaged in the study of a mechanico-electrical system which was designed to usurp a specifically human function—in the first case, the execution of a complicated pattern of computation; and in the second, the forecasting of the future. (Wiener 1948, p. 13)

Bigelow and Wiener recognized the importance of this concept of feedback in a number of different electromechanical and biological systems. For example, the movement of the tiller to regulate the direction of a ship was shown to involve a feedback process similar to that used in hand-eye coordinations necessary to pick up a pencil. (The Wiener-Bigelow work was apparently never implemented in a fire-control mechanism.)

Wiener quickly saw that the mathematics of feedback control was closely associated with aspects of statistics, statistical mechanics, and information theory.

> On the communication engineering plane, it had already become clear to Mr. Bigelow and myself that the problems of control engineering and of communication engineering were inseparable, and that they centered not around the technique of electrical engineering but around the much more fundamental notion of the message, whether this should be transmitted by electrical, mechanical, or nervous means. The message is a discrete or continuous sequence of measurable events distributed in time—precisely what is called a time-series by the statisticians. (Wiener 1948, p. 16)

These feedback problems often reduced to partial differential equations representing the stability of the system. Wiener's third war-related project, the work with Lee on wave filters, reinforced the close tie to information theory; the purpose of that research was to remove extraneous background noise from electrical networks.

Wiener contributed significantly to the mathematical theory underlying these diverse engineering problems. Like Shannon, Wiener was moving from the art of engineering to the precision of science. Using the statistical methods of time-series analysis, he was able to show that the problem of prediction could be solved by the established mathematical technique of minimization.

> Minimization problems of this type belong to a recognized branch of mathematics, the calculus of variations, and this branch has a recognized technique. With the aid of this technique, we were able to obtain an explicit best solution of the problem of predicting the future of a time series, given its statistical nature; and even further, to achieve a physical realization of this solution by a constructible apparatus.
>
> Once we had done this, at least one problem of engineering design took on a completely new aspect. In general, engineering design has been held to an art rather than a science. By reducing a problem of this sort to a minimization principle, we had established the subject on a far more scientific basis. It occurred to us that this was not an isolated case, but that there was a whole region of engineering work in which similar design problems could be solved by the methods of the calculus of variations. (Wiener 1948, p. 17)

The recurrence of similar problems of control and communication in widely diverse fields of engineering and the availability of a mathematical theory with which to organize these problems led Wiener to the creation of his new interdisciplinary science of cybernetics. "We have decided to call the entire field of control and communication theory, whether in the machine or in the animal, by the name of *Cybernetics*" (Wiener 1948, p. 19).

Long before Wiener's formulation of the science of cybernetics in 1947, results had been obtained that Wiener included as cybernetic. The word *cybernetics* derived from the Greek *kybernetes* ("steersman"). *Kybernetes* in Latin was *gubernator*, from which our word *governor* derived. The connotations, both of a steersman of public policy and of a self-regulating mechanism on a steam engine, are faithful to the word's ancient roots. The governor on a steam engine is a feedback mechanism that increases or decreases the speed of the engine depending on its current speed. Maxwell published a paper (1868) giving a mathematical characterization of governors. Similar feedback

---

[14] Both Turing and von Neumann made direct comparisons between the computer and the brain publicly in the 1950s. McCulloch and Pitts might also be regarded as having made this comparison in their famous joint paper (McCulloch and Pitts 1943). It is hard to date when, if ever, Wiener first made the comparison. He suggests in the introduction to *Cybernetics* a date as early as 1940. The author and others have searched unsuccessfully for written documentation.

mechanisms were discussed by the physiologist Bernard in his discussion of homeostasis, the means by which an organism regulates its internal equilibrium. (For a history of feedback control, see Mayr (1970). See Cannon (1932) for a discussion of Bernard's work.) In the 1930s and 1940s Nyquist, H. S. Black, and H. W. Bode renewed the study of feedback in both practical and theoretical studies of amplifiers and other electrical devices.

Although Wiener only arrived at the name *cybernetics* in 1947, as early as 1942 he had participated in interdisciplinary meetings to discuss problems central to the subject. One early meeting held in New York in 1942 under the auspices of the Josiah Macy Foundation was devoted to problems of "central inhibition in the nervous system." Bigelow, Rosenblueth, and Wiener read a joint paper, "Behavior, Purpose, Teleology" (Rosenblueth et al. 1943), which used cybernetic principles to examine the functioning of the mind. Von Neumann and Wiener called another interdisciplinary meeting at Princeton early in 1944. Engineers, physiologists, and mathematicians were invited to discuss cybernetic principles and computing design. As Wiener assessed the situation:

> At the end of the meeting, it had become clear to all that there was a substantial common basis of ideas between the workers of the different fields, that people in each group could already use notions which had been better developed by the others, and that some attempt should be made to achieve a common vocabulary. (Wiener 1948, p. 23)

In fact, from discussions with electrical engineers up and down the East Coast, Wiener reported, "Everywhere we met with a sympathetic hearing, and the vocabulary of the engineers soon became contaminated with the terms of the neurophysiologist and the psychologist" (Wiener 1948, p. 23). In 1946 McCulloch arranged for a series of meetings to be held in New York on the subject of feedback—again under the auspices of the Josiah Macy Foundation. Among those attending a number of these meetings were the mathematicians Wiener, von Neumann, and Pitts, the physiologists McCulloch, Lorente de No, and Rosenblueth, and the engineer Herman H. Goldstine (who was associated with the ENIAC, EDVAC, and IAS computer projects). Thus, there was widespread interaction in the United States among the participants in the new information sciences. Typical of international interchange was a visit by Wiener to England and France, where he had a chance to exchange information on cybernetics and artificial intelligence with Turing, then at the National Physical Laboratory at Teddington, and mathematical results on the relation of statistics and communication engineering with French mathematicians at a meeting in Nancy.

Wiener's cybernetics work paid at least as much attention to biological as to electromechanical applications. This interest was rooted in his participation in a series of informal monthly discussions in the 1930s on scientific method led by Walter Cannon at the Harvard Medical School. A few members of the MIT faculty, including Wiener, attended these meetings. Here Wiener met Rosenblueth, with whom he was to collaborate on biocybernetics throughout the remainder of his career.

Bigelow and Wiener, perhaps as a result of their work on antiaircraft artillery, pointed to feedback as an important factor in voluntary activity. To illustrate, Wiener described the process of picking up a pencil. He pointed out that we do not will certain muscles to take certain actions—instead, we will to pick the pencil up.

> Once we have determined on this, our motion proceeds in such a way that we may say roughly that the amount by which the pencil is not yet picked up is decreased at each stage. This part of the action is not in full consciousness.
>
> To perform an action in such a manner, there must be a report to the nervous system, conscious or unconscious, of the amount by which we have failed to pick the pencil up at each instant. (Wiener 1948, p. 14)

They advanced their claims about the biological importance of feedback mechanisms so far as to use them to explain pathological conditions retarding voluntary actions such as ataxia (where the feedback system is deficient) and purpose tremor (where the feedback system is overactive). Such initial successes assured Wiener that this approach could provide valuable new insights into neurophysiology.

> We thus found a most significant confirmation of our hypothesis concerning the nature of at least some voluntary activity. It will be noted that our point of view considerably transcended that current among neurophysiologists. The central nervous system no longer appears as a self-contained organ, receiving inputs from the senses and discharging into the muscles. On the contrary, some of its most characteristic activities are explicable only as circular processes, emerging from the nervous system into the muscles, and re-entering the nervous system through the sense organs, whether they be proprioceptors or organs of the special senses. This seemed to us to mark a new step in the study of that part of neurophysiology which concerns not solely the elementary processes of nerves and synapses but the performance of the nervous system as an integrated whole. (Wiener 1948, p. 15)

The revelation that cybernetics provided a new approach to neurophysiology prompted the joint paper by Rosenblueth, Wiener, and Bigelow. As the title indicates, they gave an outline of behavior, purpose, and teleology from a cybernetic approach. They argued that "teleological behavior thus becomes synonymous with behavior controlled by negative feedback, and gains therefore in precision by a sufficiently restricted connotation" (Rosenblueth et al. 1943, pp. 22–23). They also argued that the same broad classifications of behavior (see Figure 2) hold for machines as hold for living organisms. The differences, they maintained, are in the way these functional similarities are carried out: colloids versus metals, large versus small differences in energy potentials, temporal versus spatial multiplication of effects, etc.

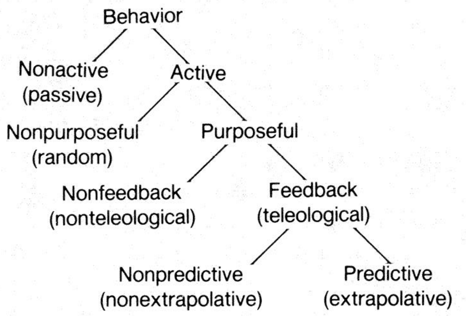


Figure 2. Behavior classifications. Note that "predictive" means order of prediction (depending on the number of parameters). (Redrawn from Rosenblueth et al. (1943, p. 21).)

For the most part, however, Wiener's work in biocybernetics was less philosophical and more physiological than the joint paper with Rosenblueth and Bigelow would indicate. More typical was a joint project between Rosenblueth and Wiener on the muscle actions of a cat (Wiener 1948, pp. 28–30). In this project they used the (cybernetic) methods of McColl (1945) on servomechanisms to analyze the system in the same way one would study an electrical or mechanical system, while using data provided by their physiological experimentation on cats. For the remainder of his career Wiener split his time between Cambridge and Mexico City, where Rosenblueth taught at the national medical school. Wiener could thus continue his collaboration with Rosenblueth on a series of physiological projects utilizing the cybernetic approach to understand physiological processes of biological organisms.

Warren McCulloch, Walter Pitts, and the Development of Mathematical Models of the Nervous System

Another biological application of the new information science was to the study of nerve systems, and in particular to the study of the human brain. Mathematical models resulted, based partly on physiology and partly on philosophy. The most famous application was made in a joint paper by McCulloch and Pitts (1943) in which they presented a mathematical model of the neural networks of the brain based on Carnap's logical calculus and on Turing's work on theoretical machines.

The application of the information sciences to psychology was linked to several active movements within psychology. The rise of physiological psychology, the development of functionalism, the growth of behaviorism, and the infusion of materialism into the biological and psychological sciences all contributed to the study by mathematical models of the functioning of the brain.

Physiological psychology was important to information science because it contributed the idea that one can understand the brain by examining its material functioning, that this functioning is amenable to scientific study, and consequently that the brain is amenable to mathematical analysis. Beginning with the work of H. von Helmholtz, T. A. Ribot, and William James at the end of the nineteenth century, physiological psychology became identified with the study of the physiological underpinnings of behavior and experience (Murphy 1949). The subject drew heavily on physiological research concerning the central nervous system. Of special importance was the work of W. Waldeyer and C. S. Sherrington (Sherrington 1906). Waldeyer's "neurone theory," which argued for the independence of the nerve cells and the importance of the synapses, was quickly accepted by psychologists and became the basis of the next generation of physiological study of the brain. Sherrington's work on reflex arc was instrumental in convincing psychologists that they should consider a neurophysiological approach. McCulloch pointed explicitly to Sherrington's work as a precursor of his own research. Both men adopted a highly idealized approach. Sherrington realized that his model of simple reflex was not physiologically precise, but only a "convenient abstraction." McCulloch and Pitts made a similar claim for their model of neuron nets, but their model was even less realistic than the Sherrington model.

Research in physiological psychology had been carried out since the beginning of the century. Its importance increased rapidly in the 1930s because of two developments: the implementation of electroencephalography enabled researchers to make precise measurements of the electrical activity of the brain; and the growth of mathematical biology, especially under Nicholas Rashevsky's Chicago school, contributed a precise, mathematical theory of the functioning of the brain that could be tested experimentally.[15] The

---

[15] The Rashevsky school published mainly in its own journal, *Bulletin of Mathematical Biophysics*.

concentration on the material properties of the brain, the emphasis on its functioning instead of on its states of consciousness, and the mathematical approach of Rashevsky all set the stage for a mathematical theory of the functioning brain as an information processor.

Most physiological psychologists and many other psychologists accepted the functionalist position. In *Principles of Psychology* (1890) William James argued persuasively that mind should be conceived of dynamically, not structurally, and by the end of the nineteenth century there was a consensus that psychology should concentrate on mental activity instead of on states of experience. E. B. Holt took a radical—and almost cybernetic—position[16] (Holt 1915) toward psychology when he argued that consciousness is merely a servomotor adjustment to the object under consideration. As one historian of psychology has assessed the importance of functionalism,

> Functionalism did not long maintain itself as a school; but much of the emphasis lived on in behaviorism . . . and in the increasing tendency to ask less about consciousness, more about activity. (Murphy 1949, p. 223)

The importance of functionalism to information science is clear. It concentrated on the functional operation of the brain, and represented the brain as a processor (of information)—as a doer as well as a reflector.

Behaviorist psychology, by concentrating on behavior and not consciousness, helped to break down the distinction between the mental behavior of humans and the information processing of lower animals and machines. This step assisted the acceptance of a unified theory of information processors, whether in humans or machines.

American behaviorism was a revolt against the old-style, introspective psychology of Wilhelm Wundt and E. B. Titchener. As part of the general shift in scientific attitude, there was a movement in psychology toward materialism in the last half of the nineteenth century. Because behavior is observable, it can be subjected to scientific study. J. B. Watson, the leader of American behaviorism, concentrated on "scientific" concepts such as effector, receptor, and learning as opposed to the old concepts of sensation, feeling, and image (Watson 1919). Watson conceived of mental functions as a type of internal behavior that could be sensed by scientific probing. Turing adopted a similar attitude in his unified treatment of human and mechanical thinking machines.[17] Watson's behaviorism did not convince the majority of American psychologists, but a group of dedicated behaviorists did conduct experiments using the condition-response method. Most influential on McCulloch and Pitts was the work in the 1930s of the behaviorist K. S. Lashley, whose viewpoint is indicated by the following quotation.

> To me the essence of behaviorism is the belief that the study of man will reveal nothing except what is adequately describable in the concepts of mechanics and chemistry, and this study far outweighs the question of the method by which the study is conducted. (Lashley 1923, p. 244)

McCulloch was trained within this psychological tradition of experimental epistemology. As an undergraduate at Haverford and Yale, he majored in philosophy and psychology. He then went to Columbia, where he received a master's degree in psychology for work in experimental aesthetics. Afterward, he entered the College of Physicians and Surgeons of Columbia University, where he studied the physiology of the nervous system.

> In 1928 I was in neurology at Bellevue Hospital and in 1930 at Rockland State Hospital for the Insane, but my purpose (to manufacture a logic of transitive verbs) never changed. It was then that I encountered Eilhard von Doramus, the great philosophic student of psychiatry, from whom I learned to understand the logical difficulties of true cases of schizophrenia and the development of psychopathia—not merely clinically, as he had learned them of Berger, Birnbaum, Bumke, Hoche, Westphal, Kahn, and others—but as he understood them from his friendship with Bertrand Russell, Heidegger, Whitehead, and Northrop—under the last of whom he wrote his great unpublished thesis, "The Logical Structure of the Mind: An Inquiry into the Foundations of Psychology and Psychiatry." It is to him and to our mutual friend, Charles Holden Prescott, that I am chiefly indebted for my understanding of *paranoia vera* and of the possibility of making the scientific method applicable to systems of many degrees of freedom. (McCulloch 1965, pp. 2–3)

McCulloch left Rockland to return to Yale, where he studied experimental epistemology with Dusser de Barenne. Upon de Barenne's death, he moved to the University of Illinois Medical School in Chicago as a professor of psychiatry; he continued his work on experimental epistemology and began his collaboration with Pitts. He completed his career at the MIT Research Laboratory of Electronics, where he collaborated with Pitts, Wiener, and others in the study of electrical circuit theory of the brain.

---

[16] Of course, this is an anachronistic characterization because the servomotor was not yet invented and the principles of cybernetics had not yet been enunciated. Nevertheless, there is a striking similarity to those later ideas.

[17] This attitude is most evident in Turing's 1950 paper, "Computing Machinery and Intelligence," but his 1937 paper, "On Computable Numbers," also suggests the view.

Pitts's background was more mathematical than McCulloch's. Pitts studied mathematical logic under Carnap at the University of Chicago. Carnap sent Pitts to see the people associated with Rashevsky's school of biophysics. There he met Alston Householder, who sent him to see McCulloch. They quickly began their joint study of the mathematical structure of systems built out of nerve nets. Through a mutual friend, J. Lettvin of Boston City Hospital, Pitts was introduced to Wiener and Rosenblueth in 1943. Later the same year, Pitts accepted a permanent position at MIT in order to work with Wiener and learn from him the cybernetic approach.

> At that time Mr. Pitts was already thoroughly acquainted with mathematical logic and neurophysiology but had not had the chance to make very many engineering contacts. In particular, he was not acquainted with Dr. Shannon's work, and he had not had much experience of the possibilities of electronics. He was very much interested when I showed him examples of modern vacuum tubes and explained to him these were ideal means for realizing in the metal the equivalents of his neuronic circuits and systems. From that time, it became clear to us that the ultra-rapid computing machine, depending as it does on consecutive switching devices, must represent almost an ideal model of the problems arising in the nervous system. (Wiener 1948, p. 22)

Pitts's work at the Research Laboratory of Electronics involved studying the relation between electronic computers and the human nervous system. During this time he continued to work with McCulloch; eventually, in 1952, McCulloch joined him at MIT. The most influential article they produced there was "What the Frog's Eye Tells the Frog's Brain" (Lettvin et al. 1959).

Early in his career, between 1919 and 1923, McCulloch worked on a problem of philosophical logic, that of creating a formal system to explain the usage of transitive verbs. While engaged in this work he became interested in a related problem involving the logic of relations. As McCulloch recalled,[18]

> The forms of the syllogism and the logic of classes were taught, and we shall use some of their devices, but there was a general recognition of their inadequacy to the problems in hand. . . . It was [Charles] Peirce who broke the ice with his logic of relatives, from which springs the pitiful beginnings of our logic of relations of two and more than two arguments. So completely had the traditional Aristotelean logic been lost that Peirce remarks that when he wrote the *Century Dictionary* he was so confused concerning abduction or apagoge, and induction that he wrote nonsense. . . . Frege, Peano, Whitehead, Russell, Wittgenstein, followed by a host of lesser lights, but sparked by many a strange character like Schroeder, Sheffer, Gödel, and company, gave us a working logic of propositions. By the time I had sunk my teeth into these questions, the Polish school was well on its way to glory. In 1923 I gave up the attempt to write a logic of transitive verbs and began to see what I could do with the logic of propositions. (McCulloch 1965, pp. 7–8)

What McCulloch had in mind was a psychological, not philosophical, theory for the logic of relations. Whereas a philosopher would have attempted to construct a formal system that mirrored typical usage of the logic of relations, McCulloch intended to develop a theory that explained the psychological underpinnings, not just the formal structure.

> My object, as a psychologist, was to invent a kind of least psychic event, or "psychon," that would have the following properties: First it was to be so simple an event that it either happened or else it did not happen. Second, it was to happen only if its bound cause had happened—shades of Duns Scotus!—that is, it was to imply its temporal antecedent. Third, it was to propose this to subsequent psychons. Fourth, these were to be compounded to produce the equivalents of more complicated propositions concerning their antecedents.
>
> In 1929 it dawned on me that these events might be regarded as the all-or-none impulses of neurons, combined by convergence upon the next neuron to yield complexes of propositional events. During the nineteen-thirties, first under influences from F. H. Pike, C. H. Prescott, and Eilhard von Doramus, and later, Northrop, Dusser de Barenne, and a host of my friends in neurophysiology, I began to try to formulate a proper calculus for these events by subscripting symbols for propositions (connected by implications) with the time of occurrence of the impulse in each neuron. (McCulloch 1965, p. 9).

Technical difficulties stood in the way of McCulloch's psychology of propositions. Then McCulloch met Pitts, who was able to provide the requisite mathematical theory to resolve these problems. The result was their famous joint paper, "A Logical Calculus of the Ideas Immanent in Nervous Activity" (1943). The paper was published in Rashevsky's journal, the *Bulletin of Mathematical Biophysics*, where it received little notice from biologists and psychologists until popularized by von Neumann.

---

[18] McCulloch is not entirely accurate in his history. Augustus de Morgan had taken the first steps toward a logic of relations in the mid-nineteenth century. Gödel never contributed directly to the development of a logic of propositions. His most closely related work was the completeness theorem for the predicate calculus, completing the tie between the semantics and syntax for the predicate calculus to parallel the tie for the propositional calculus. George Boole made the first steps toward a logic of propositions in the mid-1800s, and Charles Peirce cleared up some of the inadequacies at the end of the century (Kneale and Kneale 1962; Kline 1972).

Using as axioms the rules McCulloch prescribed for his psychons and as logical framework an amalgam of Carnap's logical calculus and Russell and Whitehead's *Principia Mathematica*, McCulloch and Pitts presented a logical model of neuron nets showing their functional similarity to Turing's computing machines.[19]

> What Pitts and I had shown was that neurons that could be excited or inhibited, given a proper net, could extract any configuration of signals in its input. Because the form of the entire argument was strictly logical, and because Gödel had arithmetized logic, we had proved, in substance, the equivalence of all general Turing machines—man-made or begotten. (McCulloch 1965, pp. 9–10)

As von Neumann emphasized in his *General and Logical Theory of Automata* (1951), the essence of McCulloch and Pitts's contribution was to show how any functioning of the brain that could be described clearly and unambiguously in a finite number of words could be expressed as one of their formal neuron nets. The close relationship between Turing machines and neuron nets was one of the goals of the authors; by 1945 they understood that neuron nets, when supplied with an appropriate analog of Turing's infinite tape, were equivalent to Turing machines.[20] With the Turing machines providing an abstract characterization of thinking in the machine world and McCulloch and Pitts's neuron nets providing one in the biological world, the equivalence result suggested a unified theory of thought that broke down barriers between the physical and biological worlds.

Their paper not only pointed out the similarity in abstract function between the human brain and computing devices; it also provided a way of conceiving of the brain as a machine in a more precise way than had been available before. It provided a means for further study of the brain, starting from a precise mathematical formulation.

> But we had done more than this, thanks to Pitts' modulo mathematics. In looking into circuits composed of closed paths of neurons wherein signals could reverberate, we had set up a theory of memory—to which every other form of memory is but a surrogate requiring reactivation of a trace. (McCulloch 1965, p. 10)

In a series of papers (McCulloch 1945; 1947; 1950; 1952), McCulloch and Pitts carried out the mathematical details of this theory of the mind, providing, for example, a model of how humans believe universal ("for all") statements.

The precision of their mathematical theory offered opportunity for additional speculation about the functioning of the mind. This precision was accomplished at the expense of a detailed theory of the biological structure and functioning of the individual nerve cells. Similar to Sherrington's model of the simple reflex, which he had called a "convenient abstraction," McCulloch and Pitts's neurons were idealized neurons. One knew what the input and output would be; but the neurons themselves were "black boxes," closed to inspection of their internal structure and operation. Practicing physiologists objected[21] (von Neumann 1951) that not only was this model of neurons incomplete, it was inconsistent with experimental knowledge. They argued further that the simplicity of the idealized neuron was so misleading as to vitiate any positive results the work might achieve. Von Neumann popularized McCulloch and Pitts's work among biologists, arguing that the simple, idealized nature of the model was necessary to understand the functioning of these neurons, and contending that once this was understood, biologists could account more easily for secondary effects related to physiological details of the neurons.

McCulloch and Pitts were able to use their mathematical theory to analyze a number of aspects of the functioning of the human nervous system. A 1950 article by McCulloch, entitled "Machines that Think and Want" (McCulloch 1950*b*), provided a prospective of the possible applications of their theory, emphasizing especially the application of cybernetic techniques to understanding the functioning of the central nervous system. Typical of McCulloch and Pitts's application of information theory to physiology was a joint project in 1947 on prosthetic devices designed to enable the blind to read by ear (Wiener 1948, pp. 31–32; de Latil 1956, pp. 12–13). The problem resolved into one of pattern recognition involving the translation of letters of various sizes into particular sounds. Using cybernetic techniques, McCulloch and Pitts produced a theory correlating the anatomy and the physiology of the visual cortex that also drew a similarity between human vision and television.

---

[19] It is more proper to say that they had shown an exact correspondence between the class of Turing machines and the class of neural nets, such that each Turing machine corresponded to a functionally equivalent neural net, and vice versa. Later, McCulloch and Pitts explicitly stated that Turing's work on computable numbers was their inspiration for the neural net paper. See McCulloch's comment in the discussion following von Neumann (1951). McCulloch also alludes to this (1965, p. 9).

[20] Arthur Burks has pointed out that because the threshold functions are all positive, the addition of a clocked source pulse will make the system universal (private communication). Von Neumann (1945) was the first to see that the switches and delays of a stored-program computer could be described in McCulloch and Pitts notation. This observation led him to the logical equivalence of finite nets and the state tables describing Turing's machines.

[21] McCulloch and Pitts recognized and admitted in their paper that their neurons were highly idealized and did not fit all the empirical evidence about neurons.

Alan Turing, Automata Theory, and Artificial Intelligence

While McCulloch and Pitts endeavored to show how the physical science of mathematics helped to explain biological functioning of the brain, Turing was busy demonstrating how the computer, a product of the physical sciences, could mimic certain essential features of the biological thinking process.[22] In fact, Turing's most famous paper, "On Computable Numbers" (1937), presented a basic mathematical model of the computer, proved several fundamental mathematical theorems about automata, and marked the first step in his lifelong battle to break down what he saw as an artificial distinction between the computer and the brain. His postwar work at the National Physical Laboratory designing the ACE computer can be viewed as an attempt to determine whether his theoretical machines could be built "in the metal." His later programming work at Manchester offered perhaps the first attempt to achieve artificial intelligence by means of programming stored-program computers instead of constructing specialized hardware to exhibit particular aspects of intelligent behavior.[23]

Turing showed an early proclivity toward mathematics and computing science. As an undergraduate at Cambridge in the mid-1930s, he first encountered Riemann's hypothesis and sought to calculate mechanically the real parts of the zeros of the zeta function. He returned to this approach several times, trying unsuccessfully before, during, and after the war to settle the hypothesis using computing equipment. This project and an interest in the computability and decidability problems of Kurt Gödel and other logicians of the 1930s led Turing to speculate on the question of which numbers in mathematics are mechanically computable. Upon reflection, he concluded that the mechanically computable numbers are exactly those that can be computed by the theoretical machines he described in his 1937 paper, known today as "Turing machines."

The computable-numbers paper made other contributions as well. The Turing machine provided a mathematically precise characterization of the basic functions and components common to all computing automata. Control, memory, arithmetic, input, and output functions were described for each machine.[24] These machines had arbitrarily large amounts of storage space and computation time—and unlimited flexibility of programming—so they represented a theoretical bound on computability by physical machine. Because of the precise mathematical characterization of these limits, Turing machines served as the starting point for the modern theory of automata. Moreover, because they were consciously designed to provide a formal analog of how the "human computer" functions, they gave a rudimentary, but precise, mathematical model of how the mind functions when carrying out computations. This point was not lost on McCulloch and Pitts, who used Turing's machine characterization as the basis for their characterization of human neuron nets as information processors.

In fact, Turing attempted to model his machines, down to specific functional details, after the way the human carries out computations. His machines were each supplied with "a tape (the analogue of paper)" divided into squares that were "scanned" for symbols. The square being scanned at a given time, he wrote, is the only one of which the machine is "directly aware." The machines were designed so that by altering the internal configuration they could "remember" some of the symbols they had "seen" previously. Turing concluded this anthropomorphic description of the machine by stating:

> We may now construct a machine to do the work of this [human] computer. To each state of mind of the [human] computer corresponds an "*m*-configuration" of the machine. The machine scans *B* squares corresponding to the *B* squares observed by the [human] computer. . . . The move which is done, and the succeeding configuration, are determined by the scanned symbol and the *m*-configuration. . . . A computing machine can be constructed to compute . . . the sequence computed by the [human] computer. (Turing 1937, pp. 231–232)

This paper was the first of many occasions on which Turing publicly expressed his convictions that computers and human brains carry out similar functions—information processing, to use modern terminology—and, consequently, that there is no reason to believe that machines will not be able to exhibit intelligent behavior.

In 1936 Turing visited Princeton University for a year to study mathematical logic with Alonzo Church, who was pursuing research in recursion theory, an

---

[22] This attitude is clear in "On Computable Numbers" (Turing 1937), where Turing explicitly modeled the machine processes after the functional processes of a human carrying out mathematical computation. See the discussion later in the text. Also see Hodges (1983).

[23] See Hodges (1983) for details. John McCarthy was the first person to point this out to the author (private communication).

[24] Although functions were differentiated, components were not. Memory, input, and output were all located on the tape. Control and arithmetic were both housed in the rules of description of the machine.

area of logic with direct relation to Turing's work on computable numbers. Turing decided later to stay, and completed a Ph.D. in 1938. Immediately afterward he returned to England because of the worsening European political situation. Soon the war was upon England, and Turing volunteered to work at Bletchley Park, where the British were trying to break mechanically produced German codes with the aid of symbol-processing equipment. There he learned about electronics and about the design and use of electromechanical and electronic calculating devices. This experience proved invaluable after the war when he was hired to design a computing machine (ACE) for the National Physical Laboratory in Teddington.

In many ways, the NPL design represented more a physical embodiment of his theoretical machines than a machine for practical use. For example, Turing was determined not to construct additional hardware whenever software could achieve the same end, no matter how roundabout the solution. He would also never alter or construct software in order to make the coder's job easier (Carpenter and Doran 1977). Another indication of Turing's lack of interest in the practical side of computing was his decision to leave NPL before ACE was completed. While several factors were probably involved in his decision to leave NPL (to assume chief programming responsibilities for the new Manchester computer), it seems clear that he had satisfied himself that his theoretical machines could be embodied physically. At least partly for this reason, much of his interest in the ACE project was dissipated.

Turing's work in Manchester was among the earliest investigations of the use of electronic computers for artificial-intelligence research. He was among the first to believe that electronic machines were capable of doing not only numerical computations, but also general-purpose information processing. He was convinced computers would soon have the capacity to carry out any mental activity of which the human mind is capable. He attempted to break down the distinctions between human and machine intelligence and to provide a single standard of intelligence, in terms of mental behavior, upon which both machines and biological organisms could be judged. In providing his standard, he considered only the information that entered and exited the automata. Like Shannon and Wiener, Turing was moving toward a unified theory of information and information processing applicable to both the machine and the biological worlds. The details of this theory can be found in two papers he wrote at the time, "Intelligent Machinery" and "Computing Machinery and Intelligence" (Turing 1950; 1970), in addition to being found in his Manchester programming activities.

In "Intelligent Machinery" Turing began by addressing the question: "What happens when we make up a machine in a comparatively unsystematic way from some kind of standard components?" (Turing 1970, p. 9). He called these *unorganized machines* and created a new mathematical theory for analyzing them, based on the flow-diagramming techniques of Goldstine and von Neumann. The major aim of the paper was to determine what sorts of machines could be constructed to display evidence of intelligence. After a lengthy analysis, Turing concluded that the best approach was not that of robotics—of building specialized hardware to mimic the various aspects of human intelligence—because he felt that the result would always fall short of its human model in some aspect or another—the human having so many properties incidental to intelligence. Instead, he decided that the best approach was to simulate human mental behavior on a general-purpose computer in such a way that the computer would react to purely mental activities (among which he counted games such as chess, language learning and translation, mathematics, and cryptography) in the same way the human responds to these activities. Indeed, Turing was among the earliest, if not the earliest, to see the advantages of the software-simulation approach to artificial intelligence. His approach was in marked contrast to that of other British researchers, such as Grey Walter or Ross Ashby, who favored using robotics to achieve artificial intelligence (Ashby 1947; 1950; 1952; 1956*a*; 1956*b*; Walter 1953).

Turing's specific plan for an intelligent machine has the adventure of a science fiction story. He reasoned that a thinking machine should be given the essentially blank mind of an infant, instead of an adult mind replete with fully formed opinions and ideas. The plan was to incorporate a mechanism, analogous to the function of childhood education, by which the infant electronic brain could be educated. The possibility of such an approach depended on Turing's belief in nature over nurture, and on his understanding of the human cortex.

> We believe then that there are large parts of the brain, chiefly in the cortex, whose function is largely indeterminate. In the infant these parts do not have much effect: the effect they have is uncoordinated. In the adult they have great and purposive effect: the form of this effect depends on the training in childhood. A large remnant of the random behavior of infancy remains in the adult.
>
> All of this suggests that the cortex of the infant is an unorganized machine, which can be organized by suitable interference training. (Turing 1969, p. 16)

Turing's plan called not only for an unorganized machine, but also for a method by which the machine could change. Turing believed that humans learn from "interference" created by other humans, so he proposed that interference be designed into the education of computers. For interference to instigate a learning experience, the machine needed to be able to adapt to an outside stimulus. Turing achieved a learning capability by including a Pavlovian pleasure/pain mechanism with which humans could reinforce or disarrange the machine's circuitry by means of electrical pulses, according to whether the machine showed the proper behavior in reaction to the stimulus.

In the second paper, "Computing Machinery and Intelligence," Turing continued his assault on what he supposed to be an artificial distinction between the computer and the brain. The paper began with a refutation of nine of the objections he had heard most often to the possibilities of intelligent machinery, such as "machines do not have the consciousness to write, say, a sonnet according to their emotions, except by a chance manipulation of symbols." While Turing's responses were characteristically interesting and ingenious, perhaps a more important contribution of the paper was his presentation of the "imitation game." Turing assumed a behaviorist approach to the question: "Can machines think?" The imitation game offered a precise way of answering this question. In the game, an interrogator was able, by terminal, say (in order that the respondents remain unseen by the interrogator), to ask questions of and receive responses from a human and a computer. If in a statistically significant number of cases the interrogator could not determine which was which, then, claimed Turing, the machine could be said to think because it displayed the same mental behavior as the human. This test provided the first precise criterion for determining machine intelligence.

At Manchester, Turing attempted on a small scale to program existing computing equipment to carry out mental activities. For example, he programmed the Manchester computer to play chess (weakly) and to solve mathematical problems.[25] (Turing 1953*a*; 1953*b*). He recognized that large efforts were required for machinery to exhibit any significant amount of intelligence, however, and estimated optimistically that it would require a battery of programmers 50 years of full-time work to bring his learning machine from childhood to adult mental maturity. Turing's early death in 1954 did not allow him sufficient time to bring these or more modest ideas to fruition. In fact, it is hard to assess what effect Turing might have had on the development of computer theory and practice had he lived longer.

John von Neumann and the General Theory of Automata

Von Neumann is the culminating figure of the early period in the information sciences because of his work toward a unified scientific treatment of information encompassing the work of all six scientists featured in this article. He had social contacts as well as intellectual interests in common with the other scientists studying information. He discussed computers and artificial intelligence with Turing when they were together in Princeton in 1937 and 1938. He had an active correspondence with Wiener. He and Wiener were the principal organizers of the interdisciplinary Princeton meetings in 1943 on cybernetics and computing. A paper by Pitts on the probabilistic nature of neuron nets started von Neumann on his research in probabilistic automata (McCulloch 1965, Introduction).

Early in his career von Neumann made valuable contributions to several areas of mathematical logic. His first discussions of automatic computing machinery were with Turing in Princeton. Problems of applied mathematics related to the war effort required von Neumann to seek additional computing power possible only with electronic computing equipment. Accidentally, von Neumann heard about the computer project being carried out for Army Ordnance at the University of Pennsylvania. He was soon involved with J. Presper Eckert and John Mauchly's group, which was then placing the finishing touches on the ENIAC and beginning to work on the design plans for the EDVAC. The upshot was his central role in the logical design of the EDVAC (incorporating ideas from mathematical logic) (von Neumann 1945) and his leadership of the Institute for Advanced Study computer project.

Von Neumann's war-related computer activities spurred his further interest in theoretical issues of the information sciences. His main concern was for developing a general, logical theory of automata. His hope was that this general theory would unify the work of Turing on theoretical machines, of McCulloch and Pitts on neural networks, and of Shannon on communication theory. Whereas Wiener attempted to unify cybernetics around the idea of feedback and control problems, von Neumann hoped to unify the various results, in both the biological and mechanical realms, around the concept of an information processor—which he called an "automaton." (The term *automaton* had been in use since antiquity to refer to a device that carries out actions through the use of a hidden motive power; von Neumann was concerned with those automata whose primary action was the processing of information.)

---

[25] Both Turing and Shannon had an interest in chess programming. The most accessible account is found in Hodges (1983).

The task of constructing a general and logical theory of automata was too large for von Neumann to carry out in detail within the final few years of his career. Instead, he attempted to provide a programmatic framework for the future development of the general theory and limited himself to developing specific aspects, including the logical theory of automata, the statistical theory of automata, the theory of complexity and self-replication, and the comparison of the computer and the brain.

Von Neumann's general program for a theory of automata was laid out in five documents, which also include his specific contributions.

1. "The General and Logical Theory of Automata," read at the Hixon Symposium on September 20, 1948, in Pasadena, California (von Neumann 1951).
2. "Theory and Organization of Complicated Automata," a series of five lectures delivered at the University of Illinois in December 1949 (von Neumann 1966, pp. 29–87).
3. "Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components," based on notes taken by R. S. Pierce of von Neumann's lectures in January 1952 at the California Institute of Technology (von Neumann 1961–1963, V, pp. 329–378).
4. "The Theory of Automata: Construction, Reproduction, Homogeneity," a manuscript written by von Neumann in 1952 and 1953, completed and edited by Arthur Burks (von Neumann 1966, pp. 89–380).
5. "The Computer and the Brain," a series of lectures von Neumann intended to deliver as the Silliman Lectures at Yale University in 1956. They were never completed or delivered because of von Neumann's fatal bout with cancer, but were published posthumously (von Neumann 1958).

Von Neumann's ultimate aim in automata theory was to develop a precise mathematical theory that would allow comparison of computers and the human nervous system. His concern was not for the particular mechanical or physiological devices that carry out the information processing, but only for the structure and functioning of the system.

Von Neumann treated the workings of the individual components of the systems, whether natural or artificial, as "black boxes," devices that work in a well-defined way, but whose internal mechanism is unknown (and need not be known for his purposes). The black-box approach amounted to axiomatizing the behavior of the elements. The advantage, he pointed out, was that all situations were idealized and that various components were assumed to act universally in a precise, clear-cut manner. This precision allowed a study of the highly complicated behavior of organisms such as computers or the human nervous system—which would be impossible unless such regularities and simplifications were assumed.

Of course, as many neurophysiologists criticized, the disadvantage of such an approach was its inherent inability to test the validity of its axioms against physiological evidence. These critics suggested that accepted physiological evidence indicated that the situation in the human nervous system is not as simple as von Neumann's analysis made it out to be. Further, they argued, even accepting von Neumann's simplifications, one learns nothing about the physiological operation of the individual elements. Nonetheless, von Neumann was convinced that the axiomatic approach, which had worked so successfully for him in clarifying complicated situations in quantum mechanics, logic, and game theory, was the best way to begin to understand the problems of information processing in complicated automata such as electronic computers or the human nervous system.

Von Neumann pointed to Turing's work on computable numbers and to McCulloch and Pitts's axiomatic model of the neural networks of the brain as the two most significant developments toward a formal theory of automata and indicated how each of these developments was equivalent to a particular system of formal logic. Although von Neumann believed these were important steps toward a mathematical theory of automata, he was dissatisfied with what the approach of formal logics could contribute to a theory of automata useful in the actual construction of computing machinery.

Von Neumann pointed out, for example, that formal logic has never been concerned with how long a finite computation actually is. In formal logic all finite computations receive the same treatment. Formal logic does not take into consideration the important fact for the theory of computing that certain finite computations are so long as to be practically prohibitive, or even practically impossible if they require more time or space than there is in the physical universe. Second, he pointed out that in practice people allot a designated fixed time to completion of their computations—a fact to which formal logics are not sensitive. Finally, he observed that at each step in a computation there is a nonzero probability of error; consequently, if computations were allowed to become arbitrarily long, the probability of a reliable computation would approach zero. These considerations led him to suggest that the formal logical approach be modified in two ways to develop a "logic of automata": by considering the actual lengths of the "chain of reasoning," and by allowing for a small degree of error in logical operations. He indicated that such a logic would be based more on analysis (the branch of mathematics) and less on combinatorics than is formal logic. In fact, it would resemble formal logic less than it would resemble Boltzmann's theory of thermodynamics, which implicitly manipulates and measures a quantity related to information.

Von Neumann's overriding concern in the development of a statistical (also known as "probabilistic") theory of information was the question of reliability of automata with unreliable components. His aims were a theory that would determine the likelihood of errors and malfunctions and a plan that would make errors that did occur "nonlethal." The problem of reliability led him to abandon a logical and adopt a statistical approach. He pointed out that the logical and the statistical theories were not distinct. Adopting a well-known philosophical position that probability can be considered as an extension of logic, he argued that the statistical theory of automata was simply an extension of the logical theory of automata.

Von Neumann's extension of automata theory from a logical to a statistical theory is strikingly similar to his work in the foundations of quantum mechanics. Quite naturally, von Neumann turned to theoretical physics for an approach to the statistical theory of information. He stated explicitly that two statistical theories of information "are quite relevant in this context although they are not conceived from the strictly logical point of view" (von Neumann 1966, p. 59), referring to the work of Boltzmann, Hartley, and Szilard on thermodynamics and of Shannon on the concept of noise and information on a communication channel. Von Neumann proceeded to give an informal account of these theories, arguing that this work on thermodynamics should be incorporated into the formal statistical account of automata. In "Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components" he tried to develop this formal statistical theory.

Von Neumann recognized that the problem of reliability confronting information processors containing components prone to error, no matter whether the processors are biological or electromechanical, is not that incorrect information might be obtained occasionally, but instead that untrustworthy results might be produced regularly. He argued that if one assumes any small positive probability e for a basic component of an automaton to fail, the probability over time of failure of the final output of the automaton tends to $\frac{1}{2}$. In other words, the significance of the machine output is lost because the behavior of the machine (its output following a given input) is no different from random behavior, as a result of the accumulation of errors in the basic components over time.[26]

Von Neumann proposed a technique he called "multiplexing" to resolve the problem of unreliable components. This technique enables the probability d of error in the final output to be made arbitrarily small for most fixed probabilities e of malfunction of a basic component.[27] The technique consists of carrying all messages simultaneously on N lines instead of on a single line. Thus automata are conceived of as black boxes with bundles of lines, instead of single lines, carrying input and output. The fundamental idea is that each function is to be carried out in N identical components. The output produced by the majority of these components is then considered to be the true output.

Von Neumann proved that for sufficiently large values of N, the malfunctioning of a small number of the basic components would cause a malfunctioning of the entire automaton with only arbitrarily small probability. He calculated that an electronic computer comprising 2500 vacuum tubes with average tube activation every 5 microseconds would achieve a mean-time-to-error rate of 8 hours if multiplexed approximately 17,500 times. Despite the practical difficulties of implementing this high a level of multiplexing, von Neumann advocated its use in future machines, provided suitable materials were available for construction.

Von Neumann's concern in the theory of automata was to provide an understanding of the theoretical functioning of the human nervous system and of the modern electronic computer. It was clear to him from the great number of neurons and electronic tubes these automata contain and the variety of tasks they can perform that both are highly "complicated" in some nontechnical sense of the term. He was interested in this concept of complication and in what implications it might have in the functioning of automata.

> There is a concept which will be quite useful here, of which we have a certain intuitive idea, but which is

---

[26] An error $e > \frac{1}{2}$ simply indicates that the automaton is behaving with the negative of its attributed function with error $f < \frac{1}{2}$, where $f = 1 - e$. This bound upon e is simply a convention, not a real limitation. The real limitation is that the events are required to be independent and have constant probability. Those assumptions are generally not satisfied, and are certainly not satisfied by relays. But the fixed-probability independent case is a simple, natural starting point for developing a theory.

[27] The order of quantification is critical here: if e is sufficiently small, then for each finite problem and each d there exists a multiplexing factor that will do the job. For a more extensive treatment of the subject of multiplexing, see Burks (1970, pp. 96–100).

> vague, unscientific, and imperfect. This concept clearly belongs to the subject of information, and quasi-thermodynamical considerations are relevant to it. I know no adequate name for it, but it is best described by calling it "complication." It is effectivity in complication, or the potentiality to do things. I am not thinking about how involved the object is, but how involved its purposive actions are. In this sense, an object is of the highest degree of complexity if it can do very difficult and involved things. (von Neumann 1966, p. 78)

Von Neumann found the concept of self-replication closely related to the concept of complexity. Any automaton that is in the process of self-replication must pass on to its progeny information relating to its basic description and behavior. This is true equally in the case of genetic coding and of Turing machine description encodings.

Von Neumann recognized that certain automata, including certain types of biological organisms, were sufficiently complicated themselves to be able to produce even more complicated automata; other automata—machine tools, for example—do not have the internal complexity necessary to produce automata even as complicated as themselves. He suggested that there must be a critical complexity threshold below which automata are not able to self-reproduce.

Von Neumann went on to study models of self-replicating automata. He first considered the universal Turing machine. Although it provided a precise object to study because of its exact mathematical description, it was not a true self-reproducing automaton because its output was not another machine like itself, but only a paper tape that characterized the behavior of another possibly—but not necessarily—similar automaton. In fact, the original universal Turing machine, and not its output, acted like the other automaton. Von Neumann was not satisfied with Turing's machine as a model of self-reproducing automata and instead planned to design four new mathematical models of automata that would produce as output machines like themselves. Only two of these models, the "kinematic" and the "cellular," were ever completed (von Neumann 1966, pp. 93–95).

The kinematic model was von Neumann's earliest[28] and simplest model of self-replication. The aim was to design an automaton, built from a few types of elementary parts, that could construct other automata like itself from a stockpile of the parts. The design called for the kinematic model to float in a reservoir filled with an unlimited supply of parts. The constructing automaton would contain a description of the automaton it was to build. It would sort through the pieces in the reservoir until it found the ones it needed, and would then assemble them according to the instructions.

The cellular model was created with the assistance of the mathematician S. M. Ulam, who was trained in mathematical logic (see Ulam 1952). This model prevailed over the kinematic model by being amenable to mathematical examination. The aim was to avoid muscular, geometric, and kinematic considerations and to concentrate exclusively on logical factors. The cellular model removed kinematic considerations by constructing an automaton consisting of stationary objects. The objects, normally in a quiescent state, would assume an active state in certain circumstances. The cellular model consisted of an infinite, two-dimensional array of square cells. Each cell contained the same finite automaton, which could assume any of 29 internal states: 1 unexcited state, 20 excitable states, and 8 excited states. Each cell was connected to the four contiguous cells, and rules were given for the transmission of excitation from one cell to its neighbors. Thus the cellular model was intended as a two-dimensional, idealized model of a neural network. Self-replication occurred when the initial logical structure of the automata, coded in terms of the cell states in one finite region of the cellular plane, was copied in a distinct region of the plane that had previously been quiescent.

The other two proposed self-replicating automata were elaborations on the cellular model. One incorporated an excitation-threshold-fatigue simulation, and the other changed the cells from discrete to continuous elements. Both intended to model self-replication in natural nervous systems more closely than did the cellular automata. The details were not worked out for either model.

In developing his theory of automata, von Neumann was determined to present a unified study of modern computing machines and the human nervous system. Even in his early works on automata, he pointed out similarities and differences between the two systems when viewed as digital processors of information. He had intended to present a detailed comparison of the two types of automata in the Silliman Lectures at Yale University in 1956. Although he was unable to complete these lectures, it is possible to reconstruct his comparison in light of his earlier comments.

Von Neumann began by comparing the basic components of the two systems—the neuron and the vac-

---

[28] Von Neumann gave three lectures on automata at the Institute for Advanced Study in June 1948, in which he described the kinematic model (von Neumann 1966, pp. 80–82, 93–94; 1951, pp. 315–316).

uum tube. Each system was analyzed for the speed, energy consumption, size, efficiency, and number of basic switching components required. He next compared the brain and the computer, considered as total information-processing systems. He contrasted the number of multiplications necessary to carry out certain basic computations, the precision and reliability of the two types of systems, and their means of memory storage, input and output, control, and balance of components. Finally, he addressed the ways the two systems handled errors.

Von Neumann calculated that the artificial switching units required greater volume, consumed more energy, and were 10,000 times less efficient (in ergs per binary action) than their biological counterparts. He noted, however, that the artificial organs had the advantage in speed (by roughly a factor of 5000) and correctly guessed that this factor would compensate for the other deficiencies. The comparison convinced von Neumann that the machine was hopelessly outstripped by the brain's memory capacity, and he pointed to the lack of accessible memory storage as the most severe limitation of computers in his day. He also suggested that the computer engineer of the future would be well advised to imitate in artificial switching organs the means of construction and materials used in neurons, because of the neuron's superiority in scale, precision, energy requirements, and ability to self-repair.

Finally, he pointed out that the two systems used different methods for treating errors. Artificial automata were designed so that each time an error occurred the machine would stop, locate the error, and correct it. This was the idea motivating his multiplexing technique: a multiplexed machine will do a computation a number of times, and if not enough of them agree, the machine will not operate. Natural automata handle errors in a radically different manner.

> The system is sufficiently flexible and well organized that as soon as an error shows up in any part of it, the system automatically senses whether this error matters or not. If it doesn't matter, the system continues to operate without paying any attention to it. If the error seems to the system to be important, the system blocks that region out, by-passes it, and proceeds along other channels. The system then analyzes the region separately at leisure and corrects what goes on there, and if correction is impossible the system just blocks the region off and by-passes it forever. (von Neumann 1966, p. 71)

Von Neumann and Turing held radically opposing views about the types of problems to which computers should be set. From the very beginning, Turing's intention was to design machines that could carry out any computations of which a human computer was capable. Later he claimed that machines could be built to carry out any sort of intelligent behavior. In other words, his aim was to design machines that could perform any task possible through digital information processing. Von Neumann's view of the role of computers was much different. He saw their principal use in numerical meteorology, atomic power and weapons research, airflow design, research on large prime numbers, and other similar scientific, military, and mathematical applications—not in artificial intelligence.

> It makes an enormous difference whether a computing machine is designed, say, for more or less typical problems of mathematical analysis, or for number theory, or combinatorics, or for translating a text. We have an approximate idea of how to design a machine to handle the typical general problems of mathematical analysis. I doubt that we will produce a machine which is very good for number theory except on the basis of our present knowledge of the statistical properties of number theory. I think we have very little idea as to how to design good machines for combinatorics and translation. (von Neumann 1966, p. 72)

According to von Neumann, the difficulty in designing machines to carry out such activities was due to inherent differences between the computer and the brain. His comparison of the two had shown that the brain outperformed the computer in many essential ways, and in his opinion this precluded the computer from accomplishing many tasks other than pure numerical computation. He hoped to turn the one advantage of the computer over the human brain, speed of computation, to best use in the role he assigned computers.

Conclusions

In order to assess the importance of the work described here, one must ask what the scientists actually accomplished, why the events happened when they did, and what import they have had for the development of computer science.

Shannon initiated a new science of information by providing a precise definition of information, a way of measuring it, and theoretical results about the limits to its transmission. His efforts resulted in a theory of communication sufficiently general to treat of any sort of transmission of information from one place to another in space or time. While his technical results have had practical import in communications engineering, the theory itself formed the basis for information science.

Wiener adopted Shannon's concern for the importance of information in the world of communication engineering and applied it to the biological realm, by showing the importance of control and communication of information, especially of feedback information, to the regulation of biological systems. He was a great synthesizer, attempting to bring many people's ideas under the wing of cybernetics. While his success at assimilating diversely used techniques into a unified method is far from clear, he was enormously successful at popularizing the concept of cybernetics and the importance of information as a scientific concept. On the more technical side, he contributed a number of techniques from mathematical physics and statistics suitable for the examination of control and communication problems, and through his work with Rosenblueth pioneered the application of cybernetic principles to biological phenomena.

Whereas Wiener explained the importance of information to homeostatic regulation in other parts of the body, McCulloch and Pitts were able to show how information-processing concepts offered a formal understanding of the functioning of neural networks of the brain. Their theory, based on the formal principles of Turing and other logicians, opened up the study of mathematical models of the brain and found potential practical application in their development of prosthetic devices to aid sensory perception.

Turing was perhaps the first to describe explicitly the human computation process in terms of the mechanical manipulation of symbols, and thus to explore the close relation between the functioning of mechanical and biological brains. Continuing this line of thought, he advanced the theory and practice of artificial intelligence, offering the first test of intelligence and perhaps the first strategy for using software instead of specialized hardware to simulate human thinking processes. His computable-numbers paper opened the subject of automata theory.

Von Neumann attempted to unify the contributions of his colleagues into a general theory of information processing and information-processing automata. Besides showing the relationships among the work of his colleagues, he contributed to specific areas of the general theory through his comparisons of the stored-program computer and the brain, his studies of logical and statistical theories of automata, and his theory of complexity and self-replication.

Why did this work occur when it did? It is clear that World War II was a major catalyst in its development, not only through wartime work on computers, but also through work on wave filters, radar, and feedback controls. Nonetheless, electrical engineers were already well along the way to developing the theory of communication prior to the war. It is likely that these developments in electrical engineering would have proceeded quickly anyway, because of the rapid growth of television and radio. Although the computer may have been important to von Neumann's first involvement in information theory, it was not the original stimulus for any of the other major figures.[29]

Other traditional scientific disciplines were probably more significant than computers in the early development of information science. Logic in particular played a direct role. Turing, McCulloch, Pitts, and Shannon were all probably first drawn to the concept of information through their interest in logic. The development in the 1930s of recursive function theory, with its emphasis on the theory of computation, was clearly a factor in the development of logic as a cornerstone of information science. Although less crucial, psychology (for McCulloch) and physics (for Wiener, and perhaps also for von Neumann) were stimulants in the development of the new science.

What importance has this work had? The subject of information and its various applications formed a more coherent discipline shortly after the war than at any time since. Perhaps there never was more than a loose affiliation between the various areas of application. Perhaps the institutional compartmentalization of knowledge and training of researchers has prevented the growth of an interdisciplinary science of information. Perhaps because hardware and software developments have absorbed most of our intellectual and financial capital over the past 40 years, the development of a theoretical science of information has been slowed. In any event, the grand science of cybernetics or information processing that Wiener and von Neumann envisioned has never materialized.

Nevertheless, the work on information has had a significant impact. In the following two decades this pioneering work developed in a number of directions: artificial intelligence, complexity theory, automata theory, cognitive computer science, information theory, cybernetics, and control and communication engineering. The basic underlying notion—information-processing automata as objects worthy of scientific study—has not been forgotten. It remains the focus of theoretical information science. Perhaps it is too soon to assess the importance of a science of information until it is seen how developments in theoretical computer science, as promised by current work on

---

[29] At least for these early researchers, the computer may have been most important in the development of information science as an ideoform, as the symbol of information processing.

artificial intelligence, automata theory, and biological information processing, contribute to the growth of computer science.

Acknowledgment

This research was supported in part by the Charles Babbage Institute, and in part by a Faculty Research grant from Williams College. The author appreciates the many useful comments made by Thomas Drucker, Victor Hilts, Gregory Moore, Anthony Oettinger, Terry Reynolds, Daniel Siegel, Edward Tebbenhoff, Carol Voelker, and especially Arthur Burks.

REFERENCES

Ashby, W. R. 1947. The nervous system as a physical machine: with special reference to the origins of adaptive behavior. *Mind 56*, 113–127.

Ashby, W. R. 1950. "The Cerebral Mechanism of Intelligent Action." In D. E. Richter (ed.), *Perspectives in Neurophysiology*, London, H. K. Lewis.

Ashby, W. R. 1952. *Design for a Brain*. New York, Wiley.

Ashby, W. R. 1956*a*. *Introduction to Cybernetics*. New York, Wiley.

Ashby, W. R. 1956*b*. "Design for an Intelligence Amplifier." In C. E. Shannon and J. McCarthy (eds.), *Automata Studies*, Princeton, Princeton University Press.

Burks, A. 1970. *Essays on Cellular Automata*. Urbana, University of Illinois Press.

Cannon, W. 1932. *Wisdom of the Body*. New York, Norton.

Carpenter, B. E., and R. W. Doran. 1977. The other Turing machine. *Computer Journal 20*, 3, 269–279.

Cherry, E. C. 1952. The communication of information (an historical review). *American Scientist 40*, 4, 640–664.

de Latil, P. 1956. *Thinking by Machine*. Boston, Houghton Mifflin.

Hartley, R. V. 1928. Transmission of information. *Bell System Technical Journal 7*, 535–553.

Hodges, A. 1983. *Alan Turing: The Enigma*. New York, Simon & Schuster.

Holt, E. B. 1915. *The Freudian Wish and Its Place in Ethics*. New York, Holt.

James, W. 1890. *Principles of Psychology*. New York, Holt.

Kline, M. 1972. *Mathematical Thought From Ancient to Modern Times*. New York, Oxford University Press.

Kneale, W. C., and M. Kneale. 1962. *Development of Logic*. Oxford, Clarendon.

Lashley, K. S. 1923. The behavioristic interpretation of consciousness. *Psychological Review 30*, 3, 237–272.

Lettvin, J. Y., H. R. MaTorana, W. S. McCulloch, and W. H. Pitts. 1959. What the frog's eye tells the frog's brain. *Proc. IRE 47*, 11, 1940–1959.

Maxwell, J. C. 1868. On governors. *Proc. Royal Society of London 16*, 270–283.

Mayr, O. 1970. *The Origins of Feedback Control*. Cambridge, MIT Press.

McColl, L. A. 1945. *Fundamental Theory of Servomechanisms*. New York, van Nostrand.

McCulloch, W. S. 1945. A heterarchy of values determined by the topology of nervous nets. *Bull. Mathematical Biophysics 7*, 89–93.

McCulloch, W. S. 1947. How we know universals: The perception of auditory and visual forms. *Bull. Mathematical Biophysics 9*, 127–147.

McCulloch, W. S. 1950. "Brain and Behavior." In W. C. Halstead (ed.), *Comparative Psychology*, Berkeley, University of California Press, pp. 39–50.

McCulloch, W. S. 1952. "Finality and Form in Nervous Activity." Publication No. 11, American Lecture Series. Springfield, Ill., Thomas.

McCulloch, W. S. 1961. "What is a Number, that a Man May Know It, and a Man, that He May Know a Number?" In *General Semantics Bulletin* 26 and 27, Lakeville, Ontario, Institute of General Semantics, pp. 7–18.

McCulloch, W. S. 1965. *Embodiments of Mind*. Cambridge, MIT Press.

McCulloch, W. S., and W. Pitts. 1943. A logical calculus of the ideas immanent in nervous activity. *Bull. Mathematical Biophysics 5*, 115–133.

Murphy, G. 1949. *Historical Introduction to Modern Psychology*. New York, Harcourt, Brace; revised edition.

Nyquist, H. 1924. Certain factors affecting telegraph speed. *Bell System Technical Journal 3*, 324–346.

Rosenblueth, A., N. Wiener, and J. Bigelow. 1943. Behavior, purpose, teleology. *Philosophy of Science 10*, 18–24.

Shannon, C. 1940. "Symbolic Analysis of Relay and Switching Circuits." Master's Thesis, Cambridge, MIT.

Shannon, C. 1945. "A Mathematical Theory of Cryptography." Bell Laboratories Confidential Report, New York.

Shannon, C. 1948. The mathematical theory of communication. *Bell System Technical Journal 27*, 3 and 4, 379–423 and 623–656.

Shannon, C. 1949*a*. Communication in the presence of noise. *Proc. IRE 37*, 1, 10–21.

Shannon, C. 1949*b*. Communication theory of secrecy systems. *Bell System Technical Journal 28*, 4, 656–715.

Shannon, C. 1971. "Information Theory." *Encyclopaedia Britannica*, 14th edition.

Shannon, C., and W. Weaver. 1949. *The Mathematical Theory of Communication*. Urbana, University of Illinois Press.

Sherrington, C. S. 1906. *The Integrative Action of the Nervous System*. New York, Scribners.

Tolman, R. C. 1938. *Principles of Statistical Mechanics*. Oxford, Clarendon.

Turing, A. 1937. On Computable Numbers, with an Application to the *Entscheidungsproblem*. *Proc. London Mathematical Society*, Series 2, 42, 230–265.

Turing, A. 1938. "Systems of Logic Based on Ordinals." Ph.D. Dissertation, Princeton, Princeton University.

Turing, A. 1950. Computing machinery and intelligence. *Mind* n.s., *59*, 433–460.

Turing, A. 1953*a*. Some calculations of the Riemann zeta function. *Proc. London Mathematical Society*, Series 3, 3, 99.

Turing, A. 1953*b*. "Digital Computers Applied to Games: Chess." In B. V. Bowden (ed.), *Faster Than Thought*, London, Pitman, pp. 288–295.

Turing, A. 1970. "Intelligent Machinery" (written in 1947). In B. Meltzer and D. Michie (eds.), *Machine Intelligence*, Edinburgh, American Elsevier, Vol. V, pp. 3–26.

Ulam, S. 1952. Random processes and transformations. *Proc. International Congress of Mathematicians* (*1950*). Providence, American Mathematical Society, Vol. II, pp. 264–275.

von Neumann, J. 1945. "First Draft of a Report on the EDVAC." Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia. Published in N. Stern, *From ENIAC to UNIVAC*, Bedford, Mass., Digital Press, 1981, pp. 177–246.

von Neumann, J. 1951. "General and Logical Theory of Automata." In *Cerebral Mechanisms in Behavior—The Hixon Symposium*. Reprinted in A. H. Taub (ed.), *Von Neumann's Collected Works*, Volume V, pp. 288–328.

von Neumann, J. 1958. *The Computer and the Brain*. New Haven, Yale University Press.

von Neumann, J. 1961–1963. *Collected Works*. New York, Pergamon, edited by A. H. Taub, six volumes.

von Neumann, J. 1966. *Theory of Self-Reproducing Automata*. Urbana, University of Illinois Press. Edited and completed by A. Burks.

Walter, W. G. 1953. *The Living Brain*. New York, W. W. Norton.

Watson, J. B. 1919. *Psychology From the Standpoint of a Behaviorist*. Philadelphia, Lippincott.

Wiener, N. 1948. *Cybernetics: or Control and Communication in the Animal and the Machine*. Cambridge, Technology Press.