9 July 2008

Source Coding and Simulation

Robert M. Gray
Information Systems Laboratory
Department of Electrical Engineering
Stanford, CA 94305
rmgray@stanford.edu

http://ee.stanford.edu/~ gray

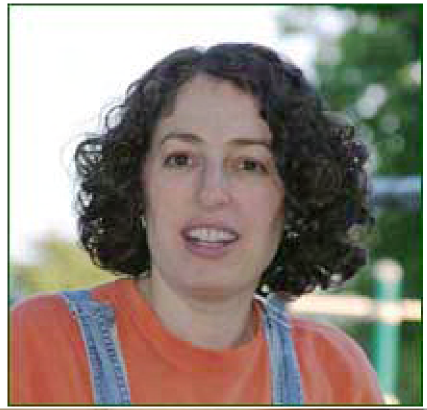
Historical and recent research described here was supported in part by NSF, hp invent

Thanks to

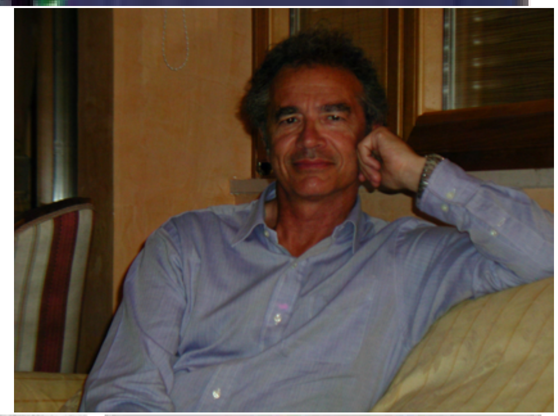
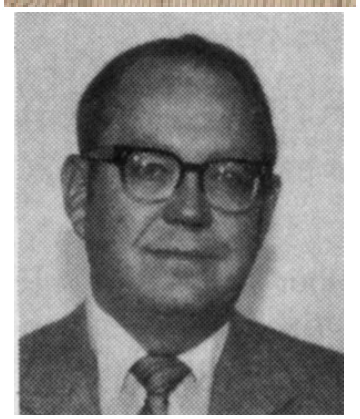
  

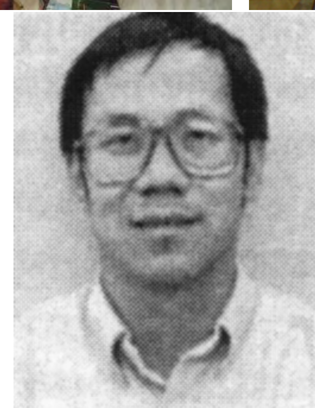
Coauthors and collaborators



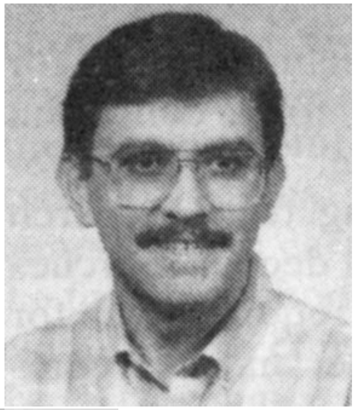


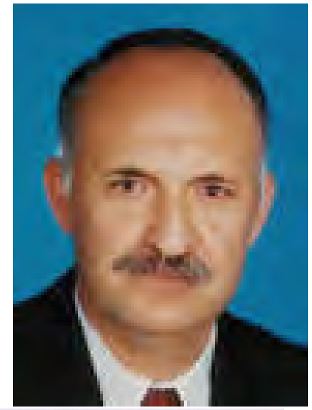


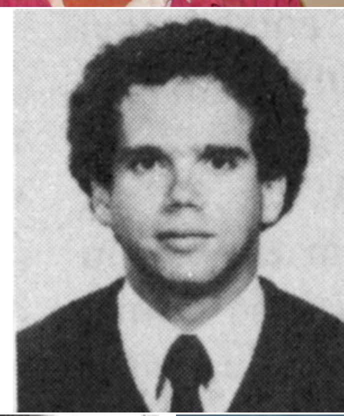




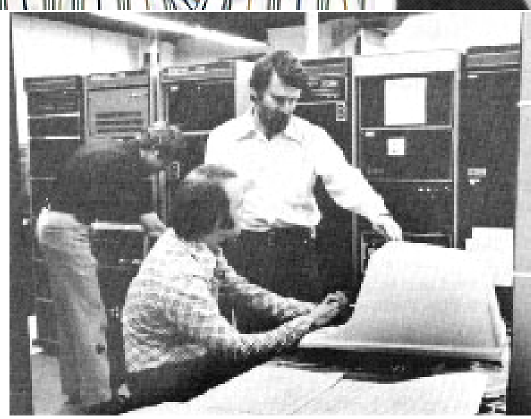



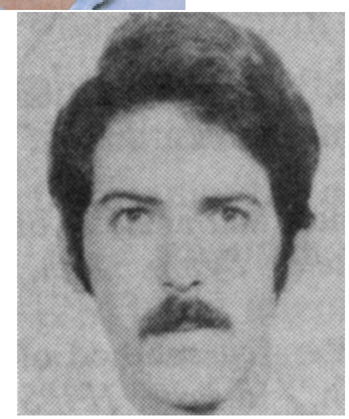


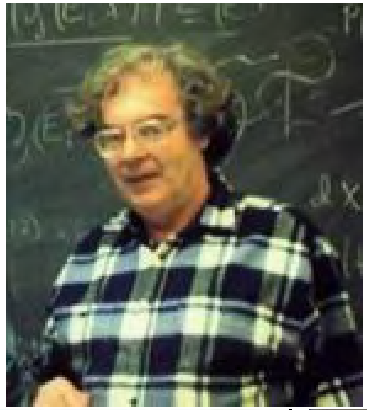
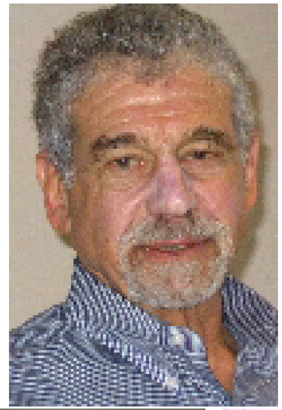
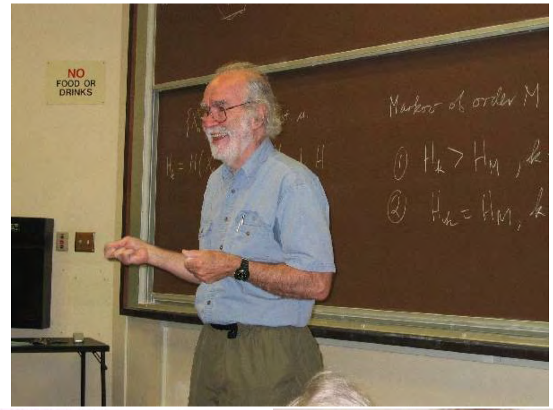




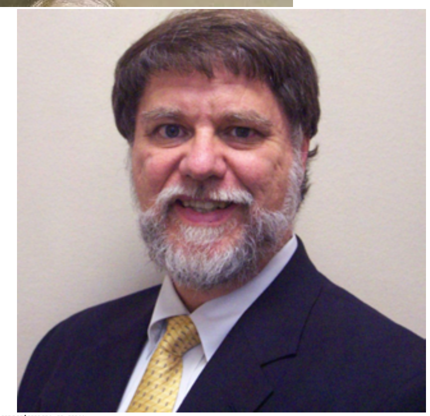
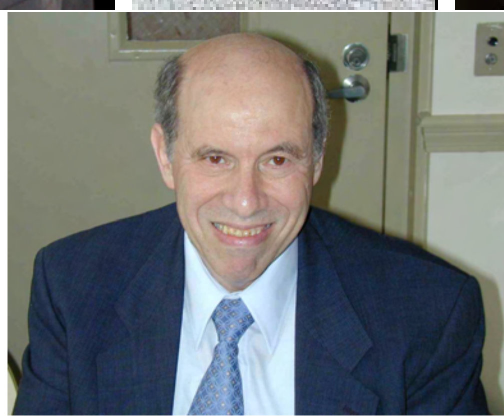









NO
FOOD OR
DRINKS

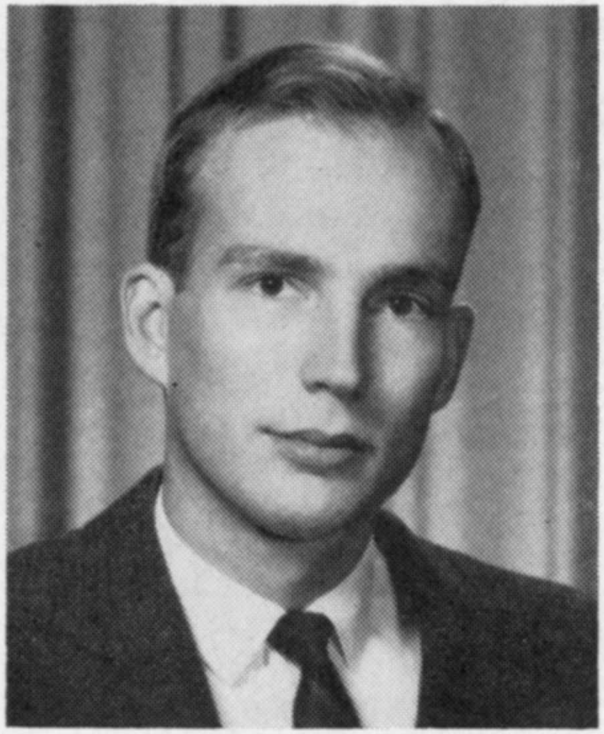


Family




Dedication



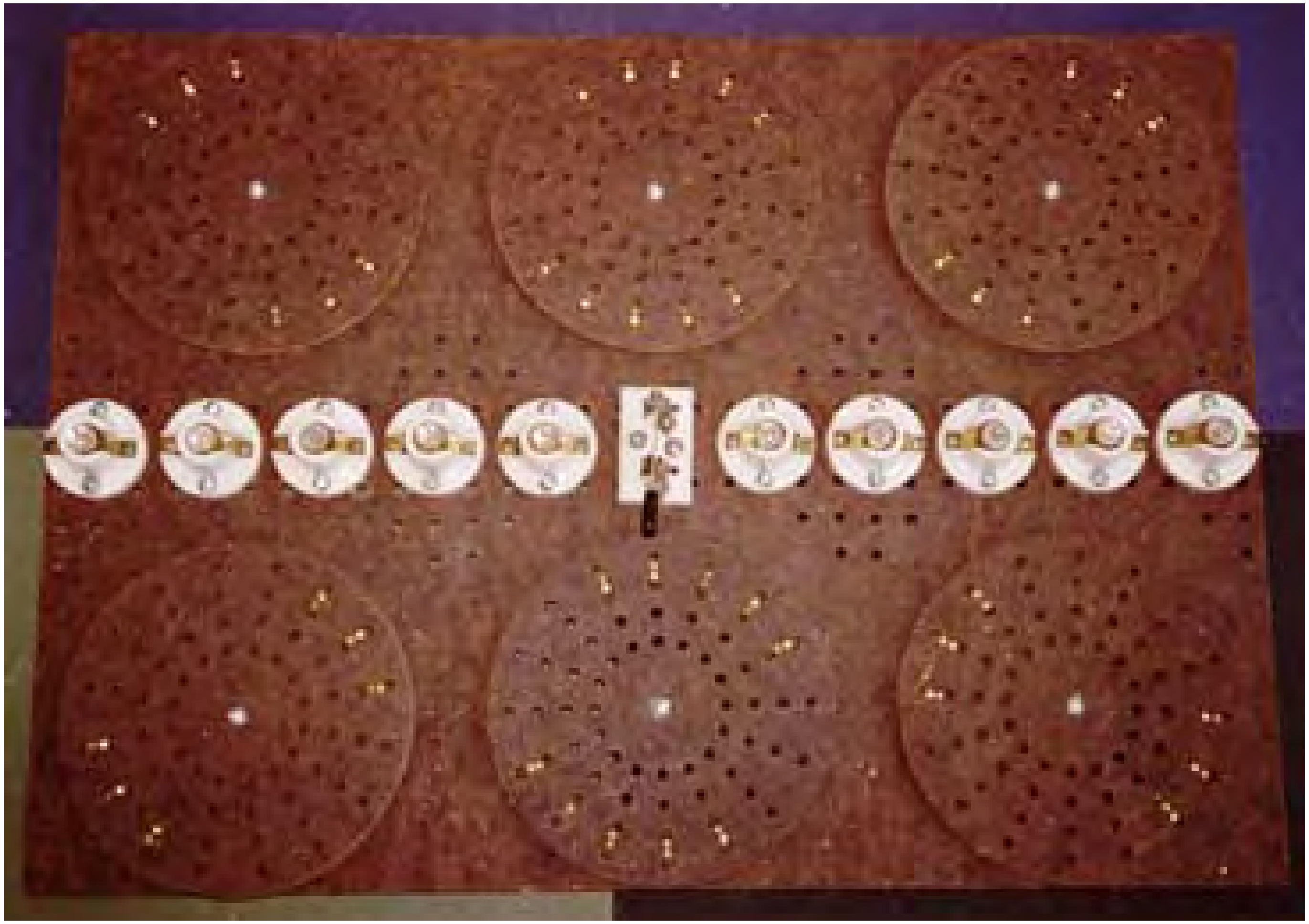
Peter R. Gray (S'62–M'65) was born in Honolulu, Hawaii, on April 16, 1940. He received the S.B., S.M., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1961, 1963, and 1965, respectively.

While at M.I.T., he was associated with the Communications Biophysics Group, Research Laboratory of Electronics, and the Eaton-Peabody Laboratory of Auditory Physiology, Massachusetts Eye and Ear Infirmary, Boston, Mass. In addition, he taught in the Department of Electrical Engineering as a Teaching Assistant, Instructor, and Assistant Professor. In 1966 he joined General Atronics Corporation, Philadelphia, Pa., where he has been primarily concerned with the detection of underground nuclear tests by seismic means.

Dr. Gray is a member of Eta Kappa Nu, Tau Beta Pi, Sigma Xi, and the American Association for the Advancement of Science.

First diagnosed with Alzheimer's disease in 2002.

First encounter with the work of Claude Shannon: Geniac from Berkeley Enterprises (1957)



A Symbolic Analysis of Relay and Switching Circuits

By CLAUDE E. SHANNON
ENROLLED STUDENT A'EE

I. Introduction

IN THE CONTROL and protective circuits of complex electrical systems it is frequently necessary to make intricate interconnections of relay contacts and switches. Examples of these circuits occur in automatic telephone exchanges, industrial motor-control equipment, and in almost any circuits designed to perform complex operations automatically. In this paper a mathematical analysis of certain of the properties of such networks will be made. Particular attention will be given to the problem of network synthesis. Given certain characteristics, it is required to find a circuit incorporating these characteristics. The solution of this type of problem is not unique and methods of finding those particular circuits requiring the least number of relay contacts and switch blades will be studied. Methods will also be described for finding any number of circuits equivalent to a given circuit in all operating characteristics. It will be shown that several of the well-known theorems on impedance networks have roughly analogous theorems in relay circuits. Notable among these are the delta-wye and star-mesh transformations, and the duality theorem.

The method of attack on these problems may be described briefly as follows: any circuit is represented by a set of equations, the terms of the equations corresponding to the various relays and switches in the circuit. A calculus is developed for manipulating these equations by simple mathematical processes, most of which are similar to ordinary algebraic algorisms. This calculus is shown to be exactly analogous to the

bolic study of logic. For the synthesis problem the desired characteristics are first written as a system of equations, and the equations are then manipulated into the form representing the simplest circuit. The circuit may then be immediately drawn from the equations. By this method it is always possible to find the simplest circuit containing only series and parallel connections, and in some cases the simplest circuit containing any type of connection.

Our notation is taken chiefly from symbolic logic. Of the many systems in common use we have chosen the one which seems simplest and most suggestive for our interpretation. Some of our phraseology, as node, mesh, delta, wye, etc., is borrowed from ordinary network theory for similar concepts in switching circuits.

II. Series-Parallel Two-Terminal Circuits

FUNDAMENTAL DEFINITIONS AND POSTULATES

closed circuit, and the symbol 1 (unity) to represent the hindrance of an open circuit. Thus when the circuit a-b is open $X_{ab} = 1$ and when closed $X_{ab} = 0$. Two hindrances X_{ab} and X_{cd} will be said to be equal if whenever the circuit a-b is open, the circuit c-d is open, and whenever a-b is closed, c-d is closed. Now let the symbol + (plus) be defined to mean the series connection of the two-terminal circuits whose hindrances are added together. Thus $X_{ab} + X_{cd}$ is the hindrance of the circuit a-d when b and c are connected together. Similarly the product of two hindrances $X_{ab} \cdot X_{cd}$ or more briefly $X_{ab}\,X_{cd}$ will be defined to mean the hindrance of the circuit formed by connecting the circuits a-b and c-d in parallel. A relay contact or switch will be represented in a circuit by the symbol in figure 1, the letter being the corresponding hindrance function. Figure 2 shows the interpretation of the plus sign and figure 3 the multiplication sign. This choice of symbols makes the manipulation of hindrances very similar to ordinary numerical algebra.

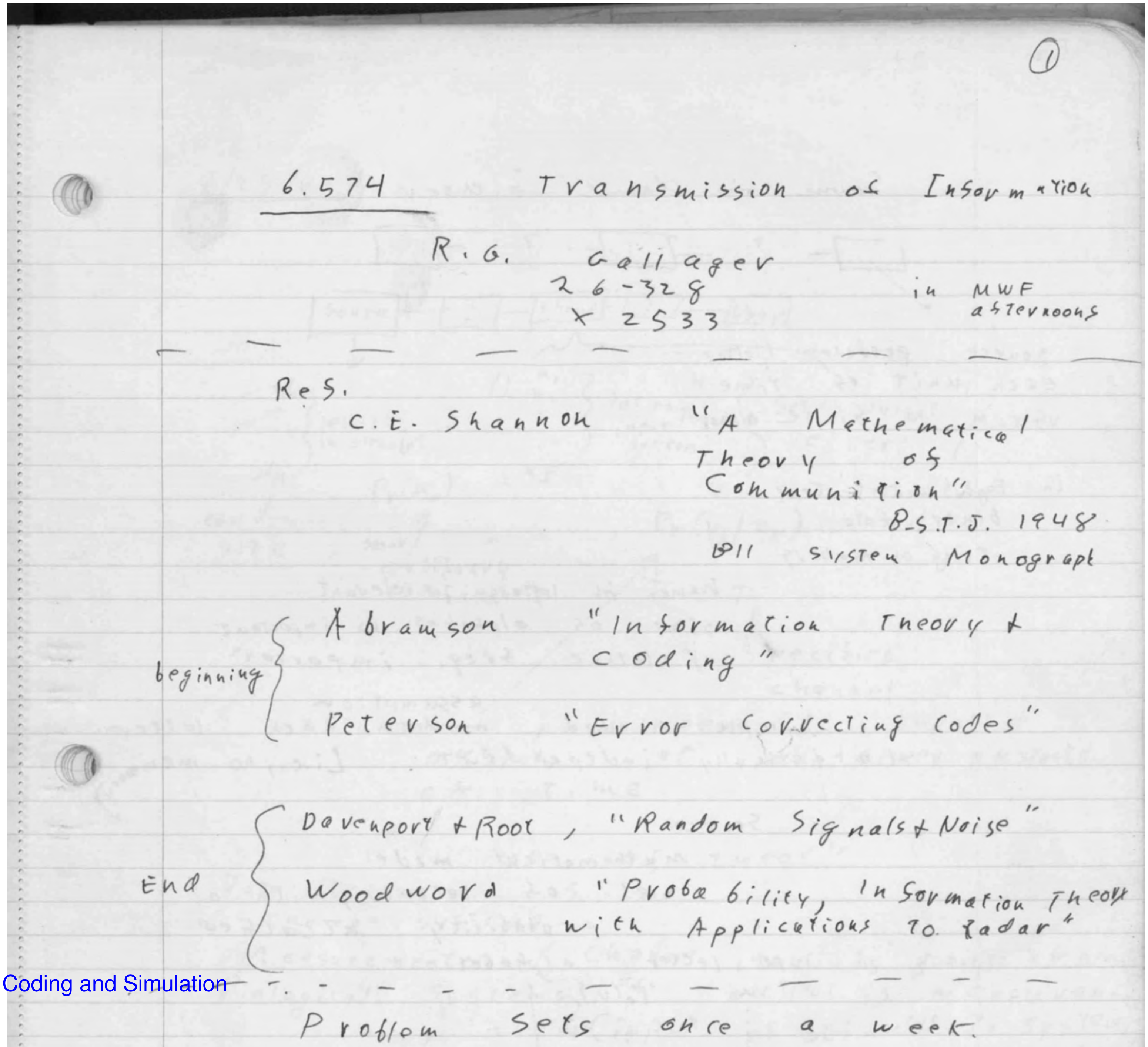
It is evident that with the above definitions the following postulates will hold:

Postulates

1. *a.* $0 \cdot 0 = 0$ — A closed circuit in parallel with a closed circuit is a closed circuit.
   b. $1 + 1 = 1$ — An open circuit in series with an open circuit is an open circuit.
2. *a.* $1 + 0 = 0 + 1 = 1$ — An open circuit in series with a closed circuit in either order (i.e., whether the open circuit is to the right or left of the closed circuit) is an open circuit.
   b. $0 \cdot 1 = 1 \cdot 0 = 0$ — A closed circuit in parallel with an open circuit in either order is a closed circuit.
3. *a.* $0 + 0 = 0$ — A closed circuit in series with a closed circuit is a closed circuit.
   b. $1 \cdot 1 = 1$ — An open circuit in parallel with an open circuit is an open circuit.
4. At any given time either $X = 0$ or $X = 1$.

These are sufficient to develop all the theorems which will be used in connection with circuits containing only series and parallel connections. The postulates are arranged in pairs to emphasize a duality relationship between the operations of addition and multiplication and the

Information Theory: 1964–65 MIT 6.574



6.574 Transmission of Infor mation

R. G. Gallager
26-328
x 2533

in MWF afternoons

Ref.

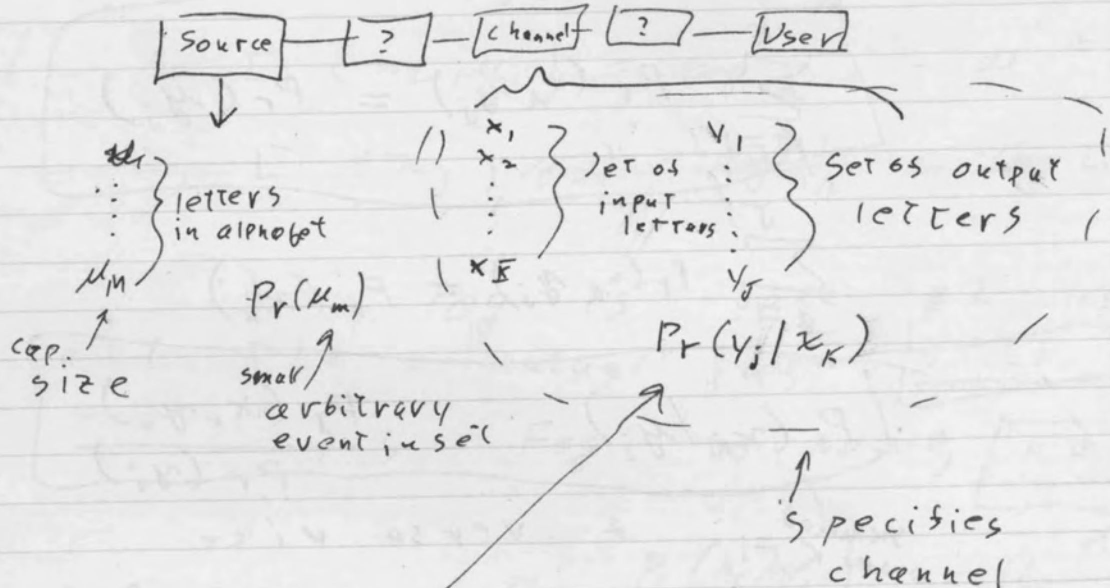
C. E. Shannon "A Mathematical Theory of Communi tion" B.S.T.J. 1948
Bell System Monograph

beginning:
- Abramson "Information Theory & Coding"
- Peterson "Error Correcting Codes"

End:
- Davenport + Root, "Random Signals + Noise"
- Woodword "Probability, Information Theory with Applications to Radar"

Problem sets once a week.

9/25/64



assume: independent of inputs + outputs at all other instants of time

use of theory

Data Transmission Systems
Storage systems
Measurements (Channel input is quantity to be measured
" output is measurement
Biological Systems (input is stimulus to system;
output is response)

Economic Systems

27

use ex. to show

$Pr(u_1) = \frac{1}{2} \to 0$
$Pr(u_2) = 1/4 \to 1\,0$
$Pr(u_3) = 1/8 \to 1\,1\,0$
$Pr(u_4) = 1/8 \to 1\,1\,1$

av. # binary #'s / source letter

$\frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 3 = 1\frac{3}{4}$

H(U) bin digits / source wd.

Grading session
David Forney
2 6-328

$\overline{(\xi + n)} = \bar{\xi} + \bar{n}$
$\overline{\xi n} = \bar{\xi}\,\bar{n}$ if indep
$\overline{a n} = a\,\bar{n}$

E: ξ σ^2

$n = \xi$

The infamous take-home exam

(*d*) Show that $\bar{P}_1$ is an upper bound on the probability of a decoding erasure and that $\bar{P}_2$ is an upper bound on the probability of decoding error. Sketch $\alpha + T - R$ as a function of R in the limit $T \to C$ and compare with the random coding exponent. Point out the implications of this for a channel where a noiseless "feedback" link is available from receiver to transmitter. For a similar, but much stronger, bound on the probability of erasures and errors, see Forney (1968).

5.15. In the previous problem we observe that the probability of not decoding correctly (i.e., making either an error or an erasure) is upper-bounded by $\bar{P}_e \le \bar{P}_1 + \bar{P}_2$; and if we set $T = R$, then $\bar{P}_e \le 2\exp(-N\alpha)$ where

$$\alpha = \max_{0 \le s \le 1} \left\{ -sR - \ln \left[\sum_{j,k} Q(k)\omega(j)^s P(j \mid k)^{1-s} \right] \right\} \qquad (1)$$

Note that since $\alpha > 0$ for $R < C$, this provides a very simple proof of the coding theorem (although not with the best error exponent).

(*a*) Replace s by $\rho/(1+\rho)$ and show that for the binary symmetric channel, (1) reduces to

$$\alpha = \max_{\rho \ge 0} \frac{1}{1+\rho}\,[E_0(\rho) - \rho R]$$

Compare α to $E_r(R)$ by the graphical technique of Fig. 5.6.3.

(*b*) By using Holder's inequality (Problem 4.15*c*) on the sum over j in (1), show that for a general DMC,

$$\alpha \ge \max_{\rho \ge 0} \frac{1}{1+\rho}\,[E_0(\rho) - \rho R]$$

(*c*) Setting $s = \rho$ in (1) and using Holder's inequality on

$$\sum_k Q(k) P(j \mid k)^{1/(1+\rho)}$$

show that

$$\alpha \le E_r(R)$$

5.16. *Joint source and channel coding theorem.*

(*a*) Let $P_N(\mathbf{y} \mid \mathbf{x})$ be the transition probability assignment for sequences of length N on a discrete channel and consider an ensemble of codes, in which M code words are independently chosen, each with a probability assignment $Q_N(\mathbf{x})$. Let the messages encoded into these code words have a probability assignment q_m, $1 \le m \le M$, and consider a maximum a-posteriori probability decoder, which, given $\mathbf{y}$, chooses the m that maximizes $q_m P_N(\mathbf{y} \mid \mathbf{x}_m)$. Let

$$\bar{P}_e = \sum_m q_m \bar{P}_{e,m}$$

be the average error probability over this ensemble of messages and codes, and by modifying the proof of Theorem 5.6.1 where necessary, show that

$$\bar{P}_e \le \left[\sum_{m=1}^{M} q_m^{1/(1+\rho)}\right]^{1+\rho} \sum_{\mathbf{y}} \left[\sum_{\mathbf{x}} Q_N(\mathbf{x}) P_N(\mathbf{y} \mid \mathbf{x})^{1/(1+\rho)}\right]^{1+\rho} \qquad (1)$$

(*b*) Let the channel be memoryless with transition probabilities $P(j \mid k)$, let the letters of the code words be independently chosen with probability assignment $Q(k)$, and let the messages be sequences of length L from a discrete memoryless source U with probability assignment $\Pi(i)$, $0 \le i \le A-1$. Show that (1) is equivalent to

$$\bar{P}_e \le \exp\{-NE_0(\rho,\mathbf{Q}) + LE_s(\rho)\} \qquad (2)$$

$$E_s(\rho) = (1+\rho)\ln\left[\sum_{i=0}^{A-1} \Pi(i)^{1/(1+\rho)}\right]$$

(*c*) Show that $E_s(0) = 0$,

$$\left.\frac{\partial E_s(\rho)}{\partial \rho}\right|_{\rho=0} = H(U) \text{ (in nats)}$$

and that $E_s(\rho)$ is strictly increasing in ρ (if no $\Pi(i) = 1$).

(*d*) Let $\lambda = L/N$, and let $N \to \infty$ with λ fixed. Show that $\bar{P}_e \to 0$ if $\lambda H(U) < C$ and if $Q(k)$ is appropriately chosen.

(*e*) Show that (2) is equivalent to (5.6.13) in the case where $\Pi(i) = 1/A$ for $0 \le i \le A-1$ and is equivalent to the positive part of the source coding Theorem 3.1.1 (except for the exponential convergence here) if the channel is noiseless. The above results are due to the author and were first used in a take-home quiz in 1964.

5.17. Consider the sum channel (as in Problem 4.18) associated with a set of n DMC's. Let $E_{0,i}(\rho) = \max_{\mathbf{Q}} E_{0,i}(\rho,\mathbf{Q})$ for the ith of the set of DMC's and let $E_0(\rho) = \max_{\mathbf{Q}} E_0(\rho,\mathbf{Q})$ for the sum channel. Let $q(i)$ be the maximizing probability of using the ith channel and let $Q_i(k)$ be the probability of using input k on the ith channel, given that the ith channel is to be used. Show that

$$\exp\left[\frac{E_0(\rho)}{\rho}\right] = \sum_{i=1}^{n} \exp\left[\frac{E_{0,i}(\rho)}{\rho}\right]$$

$$q(i) = \frac{\exp\left[\frac{E_{0,i}(\rho)}{\rho}\right]}{\sum_{i=1}^{n} \exp\left[\frac{E_{0,i}(\rho)}{\rho}\right]}$$

Application of Information Theory: 1965

Gray

U. S. NAVAL ORDNANCE LABORATORY
White Oak, Silver Spring, Maryland
UM:RMG:mm:fbs

8 June 1965

MEMORANDUM

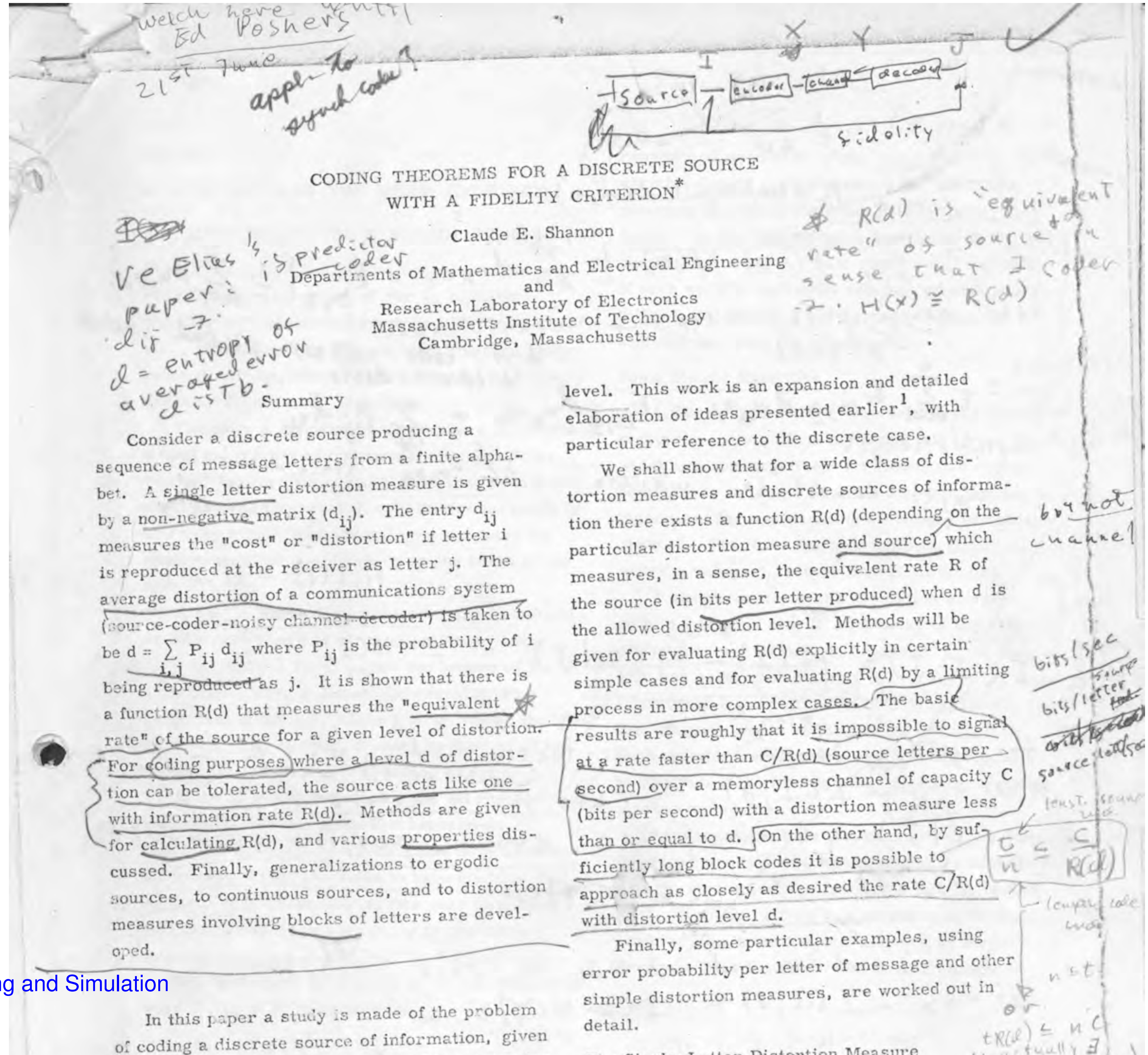
From: Robert M. Gray
To: UM Files
Via: UM-5 *al*

Subj: DEEP JULIE cable, preliminary investigation of trade-offs involved in achieving desired information rate capabilities on

Ref: (a) Davenport, W. P., "Minimum Error Receivers," MIT Class Notes
(b) Siebert, W. M., "Statistical Decision Theory and Communications: The Simple Binary Decision Problem," Lectures on Communication Theory by E. Baghdady
(c) Peterson, W. W., Error-Correcting Codes
(d) Holsinger, J. L., "Digital Communication over Fixed, Time-Continuous Channels with Memory," PhD Thesis, MIT
(e) Tufts, D. W., "Nyquist's Problem - The Joint Optimization of Transmitter and Receiver in Pulse Amplitude Modulation", Proceedings of IEEE, Mar 1965

1. The purpose of this paper is to investigate the information rate capabilities over a channel model similar to the DEEP JULIE cable, and to gain some feeling for the various trade-offs involved. The cable, equalizer, and input and output amplifiers will be modeled as a linear, time invariant time and amplitude continuous channel with additive white Gaussian noise (WGN). The channel is assumed to have the characteristics of a 600 KC band pass filter centered at 300, 600 cps. Thus, the channel is modeled as shown in

Rate-Distortion Theory: USC, 1966



CODING THEOREMS FOR A DISCRETE SOURCE
WITH A FIDELITY CRITERION*

Claude E. Shannon
Departments of Mathematics and Electrical Engineering
and
Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts

Summary

Consider a discrete source producing a sequence of message letters from a finite alphabet. A single letter distortion measure is given by a non-negative matrix (d_{ij}). The entry d_{ij} measures the "cost" or "distortion" if letter i is reproduced at the receiver as letter j. The average distortion of a communications system (source-coder-noisy channel-decoder) is taken to be $d = \sum_{i,j} P_{ij} d_{ij}$ where P_{ij} is the probability of i being reproduced as j. It is shown that there is a function R(d) that measures the "equivalent rate" of the source for a given level of distortion. For coding purposes where a level d of distortion can be tolerated, the source acts like one with information rate R(d). Methods are given for calculating R(d), and various properties discussed. Finally, generalizations to ergodic sources, to continuous sources, and to distortion measures involving blocks of letters are developed.

In this paper a study is made of the problem of coding a discrete source of information, given

level. This work is an expansion and detailed elaboration of ideas presented earlier[1], with particular reference to the discrete case.

We shall show that for a wide class of distortion measures and discrete sources of information there exists a function R(d) (depending on the particular distortion measure and source) which measures, in a sense, the equivalent rate R of the source (in bits per letter produced) when d is the allowed distortion level. Methods will be given for evaluating R(d) explicitly in certain simple cases and for evaluating R(d) by a limiting process in more complex cases. The basic results are roughly that it is impossible to signal at a rate faster than C/R(d) (source letters per second) over a memoryless channel of capacity C (bits per second) with a distortion measure less than or equal to d. On the other hand, by sufficiently long block codes it is possible to approach as closely as desired the rate C/R(d) with distortion level d.

Finally, some particular examples, using error probability per letter of message and other simple distortion measures, are worked out in detail.

Ashkelon ISIT (1973)