CLAUDE SHANNON

克劳德 香农

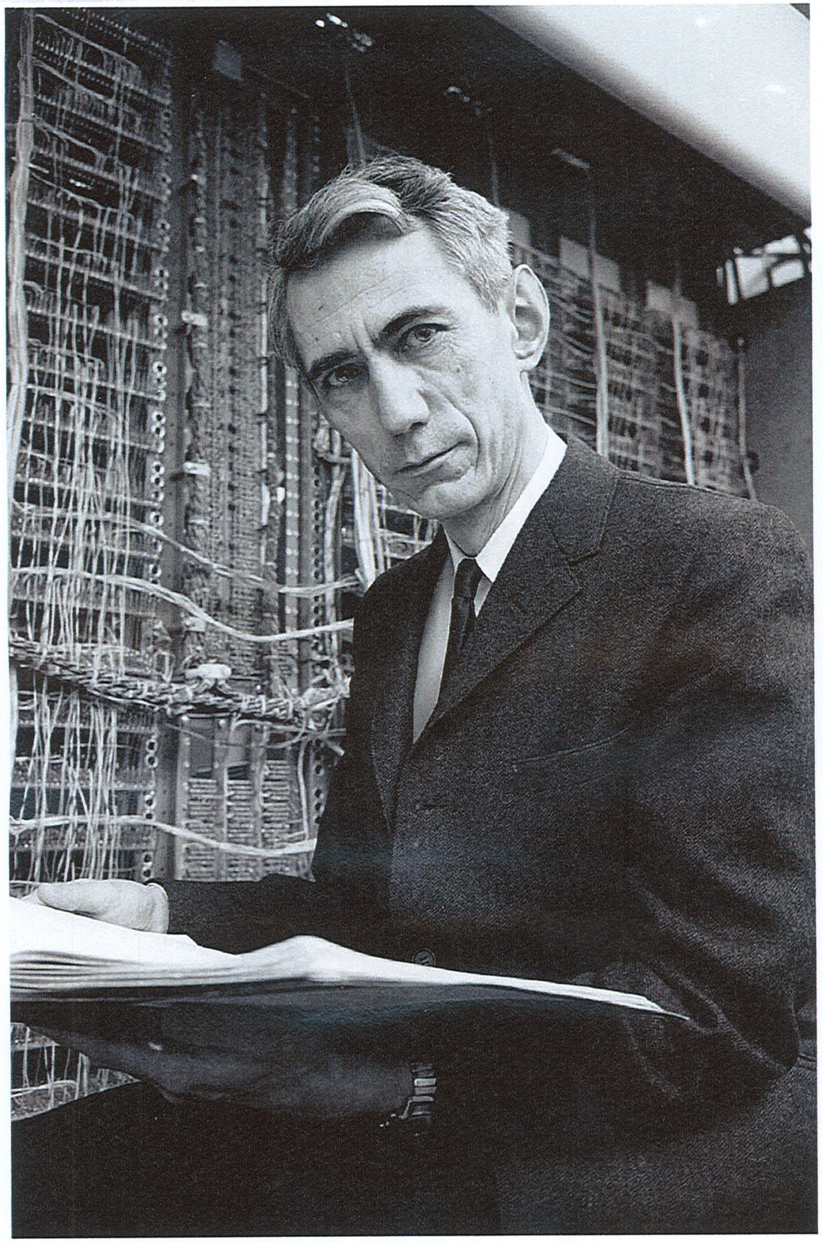
(1916 - 2001)

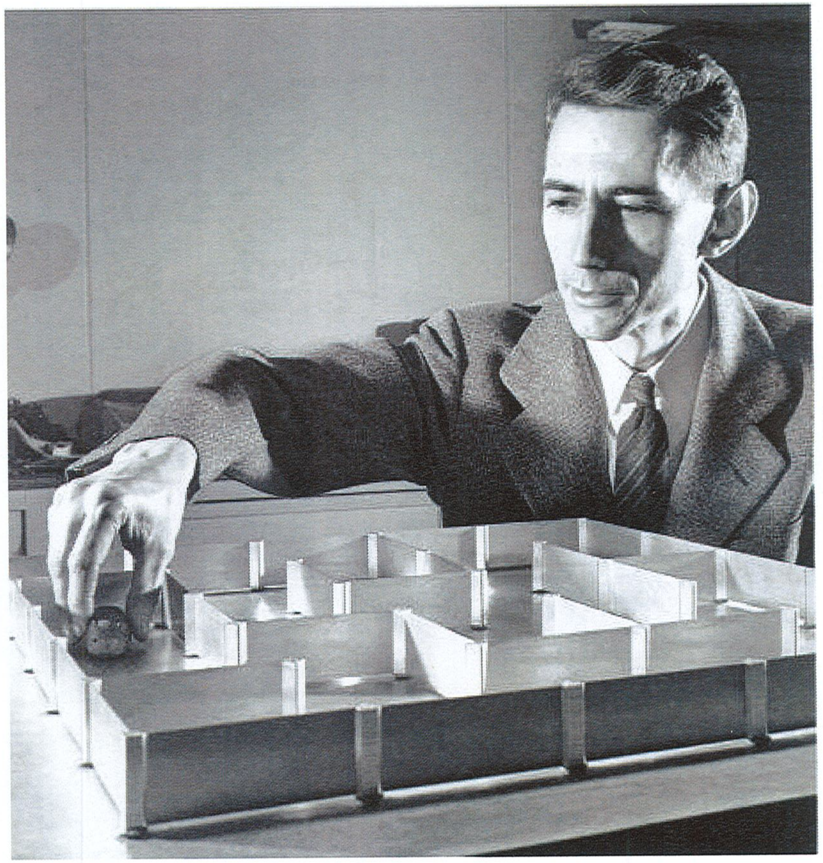
ENTROPIE et INFORMATION

熵和信息

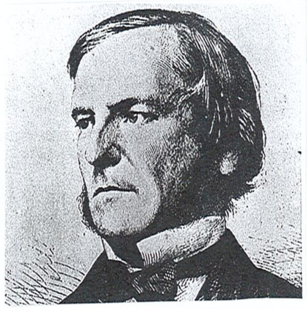
La force d'une idée

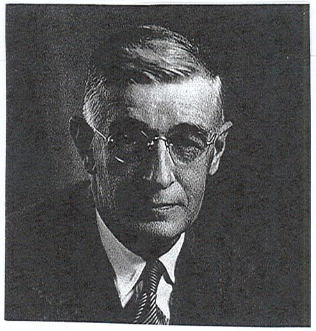
一个想法的力量

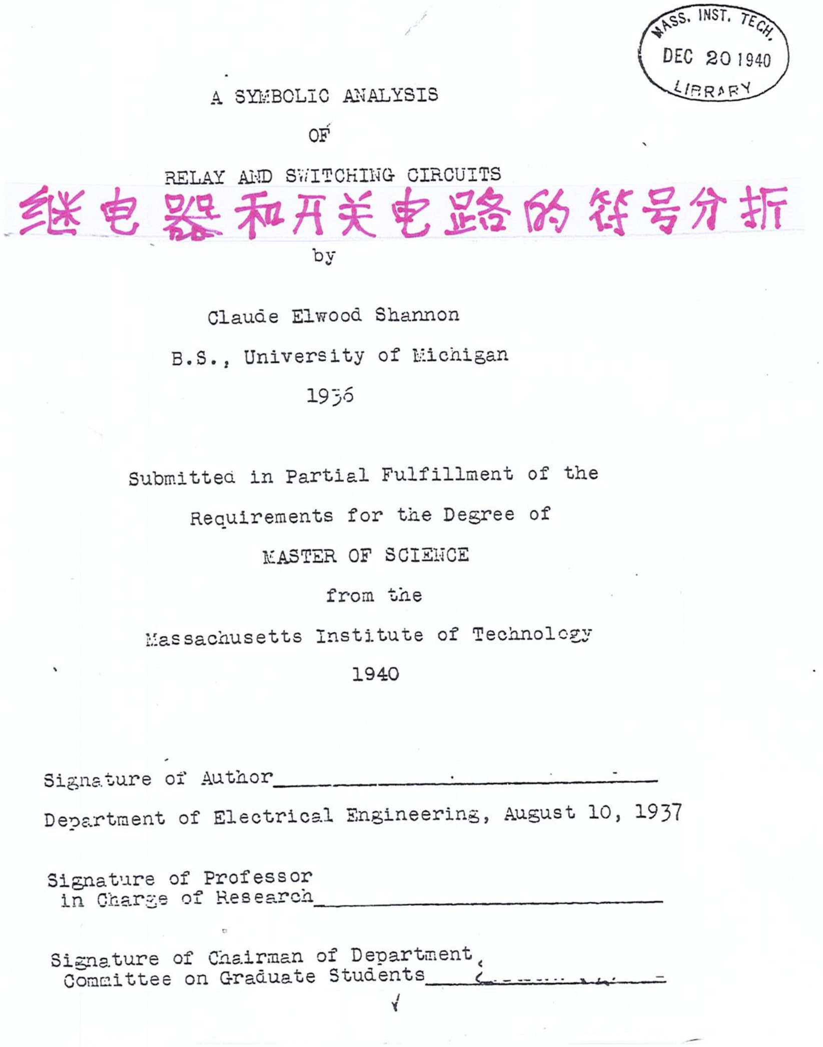




Courtesy of AT&T Archives and History Center







A SYMBOLIC ANALYSIS

OF

RELAY AND SWITCHING CIRCUITS

继电器和开关电路的符号分析

by

Claude Elwood Shannon

B.S., University of Michigan

1936

Submitted in Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

from the

Massachusetts Institute of Technology

1940

Signature of Author________________________________

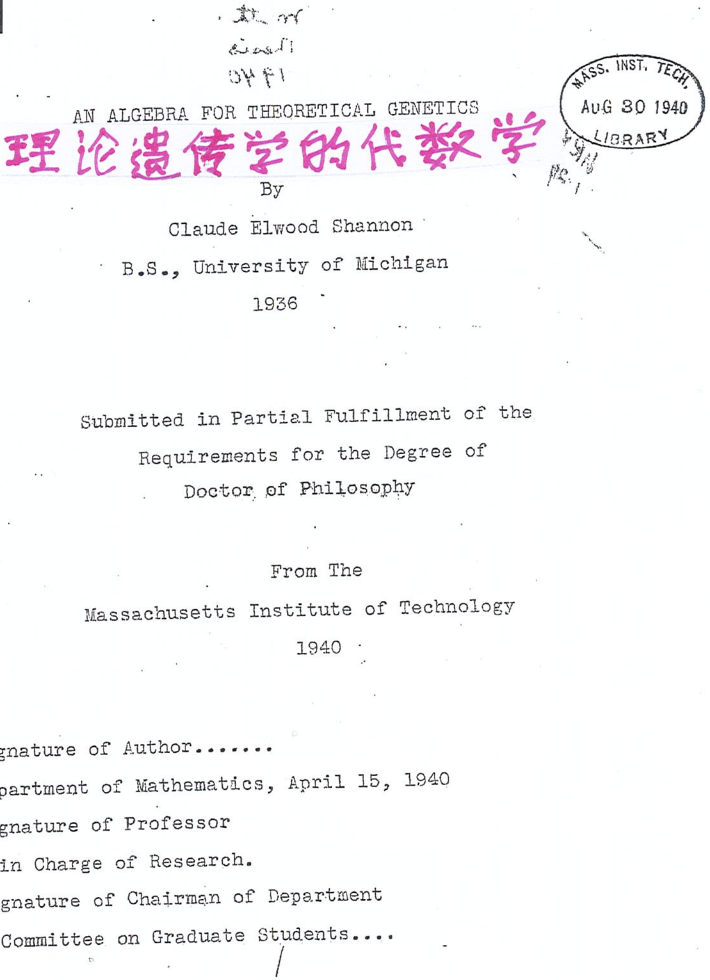
Department of Electrical Engineering, August 10, 1937

Signature of Professor
in Charge of Research______________________________

Signature of Chairman of Department
Committee on Graduate Students__________________

MASTER 1937

硕士论文



AN ALGEBRA FOR THEORETICAL GENETICS

理论遗传学的代数学

By

Claude Elwood Shannon

B.S., University of Michigan

1936

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

From The

Massachusetts Institute of Technology

1940

Signature of Author.......

Department of Mathematics, April 15, 1940

Signature of Professor

in Charge of Research.

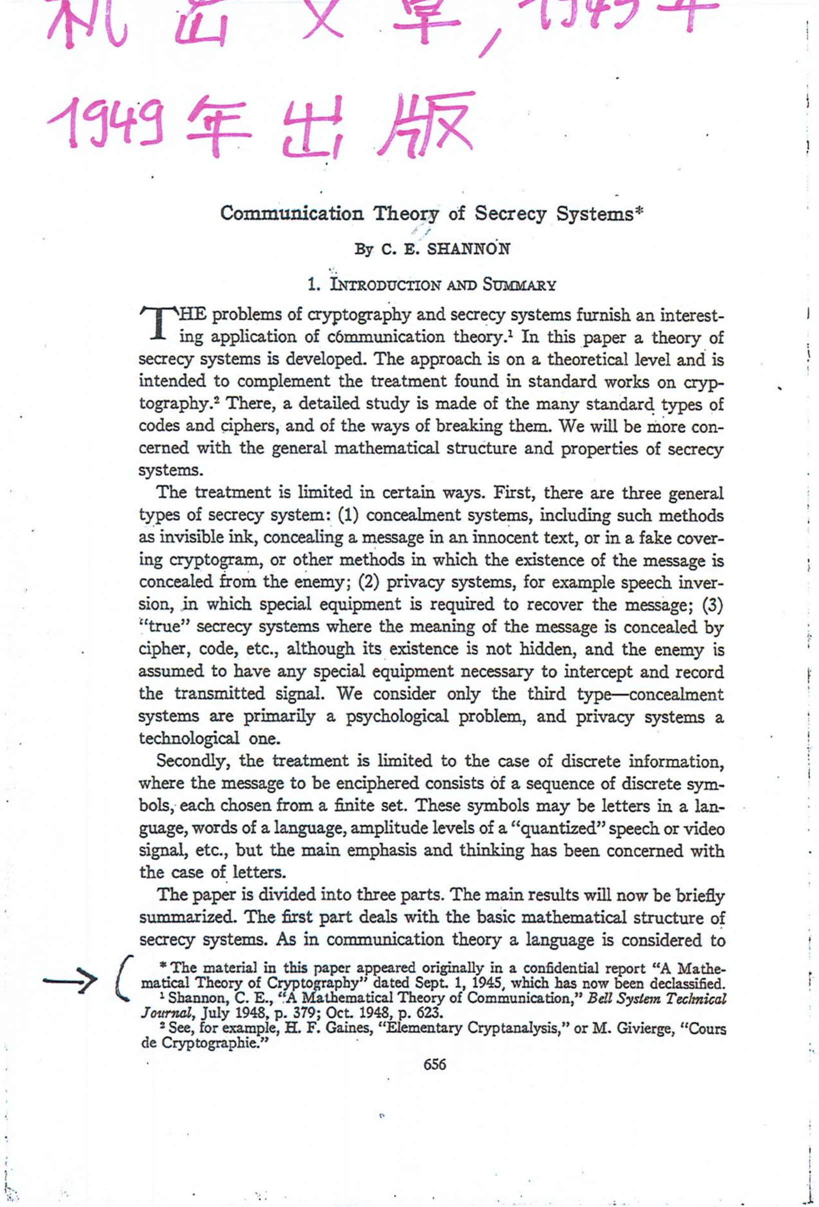
Signature of Chairman of Department

Committee on Graduate Students....

THÈSE 1940

博士论文

机密文章，1945年
1949年出版



Communication Theory of Secrecy Systems[*]

By C. E. SHANNON

1. Introduction and Summary

THE problems of cryptography and secrecy systems furnish an interesting application of communication theory.[1] In this paper a theory of secrecy systems is developed. The approach is on a theoretical level and is intended to complement the treatment found in standard works on cryptography.[2] There, a detailed study is made of the many standard types of codes and ciphers, and of the ways of breaking them. We will be more concerned with the general mathematical structure and properties of secrecy systems.

The treatment is limited in certain ways. First, there are three general types of secrecy system: (1) concealment systems, including such methods as invisible ink, concealing a message in an innocent text, or in a fake covering cryptogram, or other methods in which the existence of the message is concealed from the enemy; (2) privacy systems, for example speech inversion, in which special equipment is required to recover the message; (3) "true" secrecy systems where the meaning of the message is concealed by cipher, code, etc., although its existence is not hidden, and the enemy is assumed to have any special equipment necessary to intercept and record the transmitted signal. We consider only the third type—concealment systems are primarily a psychological problem, and privacy systems a technological one.

Secondly, the treatment is limited to the case of discrete information, where the message to be enciphered consists of a sequence of discrete symbols, each chosen from a finite set. These symbols may be letters in a language, words of a language, amplitude levels of a "quantized" speech or video signal, etc., but the main emphasis and thinking has been concerned with the case of letters.

The paper is divided into three parts. The main results will now be briefly summarized. The first part deals with the basic mathematical structure of secrecy systems. As in communication theory a language is considered to

[*] The material in this paper appeared originally in a confidential report "A Mathematical Theory of Cryptography" dated Sept. 1, 1945, which has now been declassified.

[1] Shannon, C. E., "A Mathematical Theory of Communication," *Bell System Technical Journal*, July 1948, p. 379; Oct. 1948, p. 623.

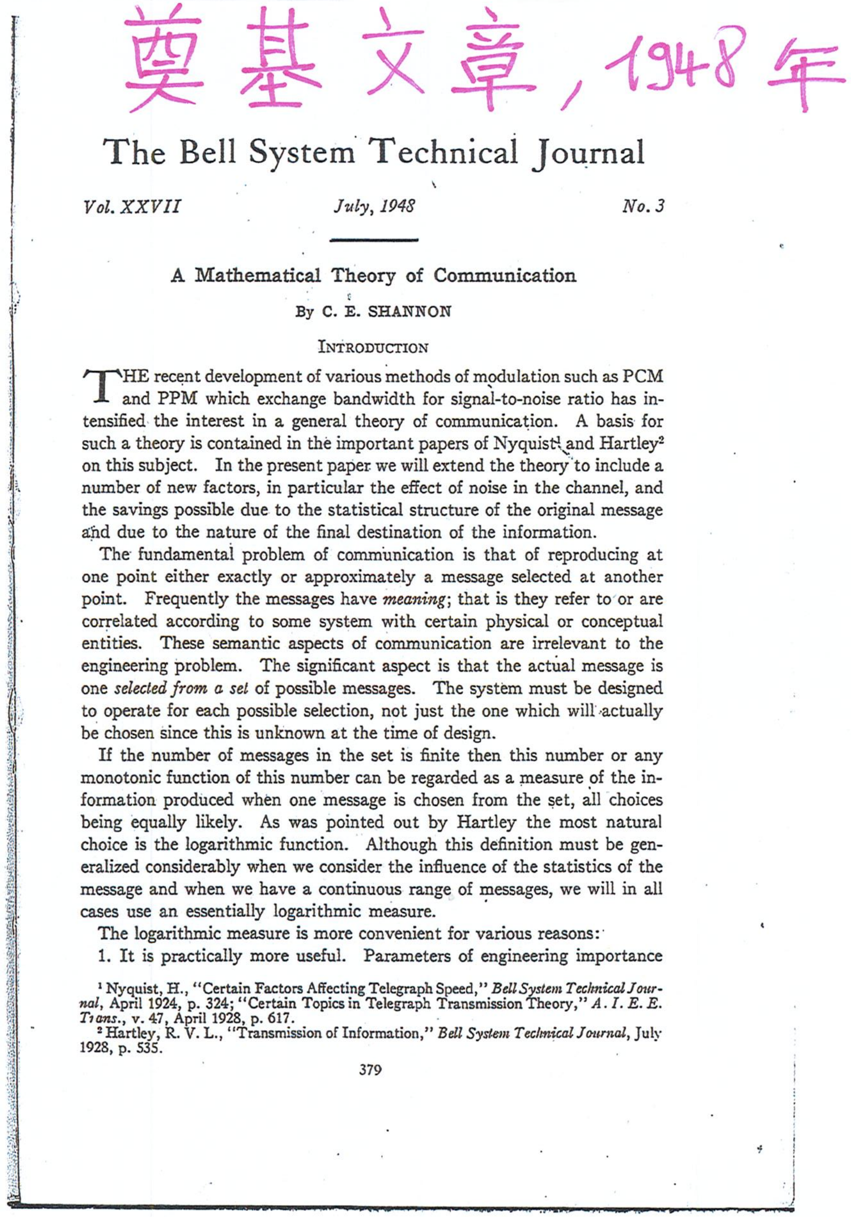
[2] See, for example, H. F. Gaines, "Elementary Cryptanalysis," or M. Givierge, "Cours de Cryptographie."

656

Rapport classifié de 1945
Publié en 1949

⇒

奠基文章，1948年



The Bell System Technical Journal

Vol. XXVII *July, 1948* *No. 3*

A Mathematical Theory of Communication

By C. E. SHANNON

Introduction

THE recent development of various methods of modulation such as PCM 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[1] and Hartley[2] on this subject. In the present paper we will extend the theory to include a number of new factors, in particular the effect of noise in the channel, and the savings possible due to the statistical structure of the original message and due to the nature of the final destination of the information.

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. As was pointed out by Hartley the most natural choice is the logarithmic function. Although this definition must be generalized considerably when we consider the influence of the statistics of the message and when we have a continuous range of messages, we will in all cases use an essentially logarithmic measure.

The logarithmic measure is more convenient for various reasons:

1. It is practically more useful. Parameters of engineering importance

[1] Nyquist, H., "Certain Factors Affecting Telegraph Speed," *Bell System Technical Journal*, April 1924, p. 324; "Certain Topics in Telegraph Transmission Theory," *A. I. E. E. Trans.*, v. 47, April 1928, p. 617.

[2] Hartley, R. V. L., "Transmission of Information," *Bell System Technical Journal*, July 1928, p. 535.

379

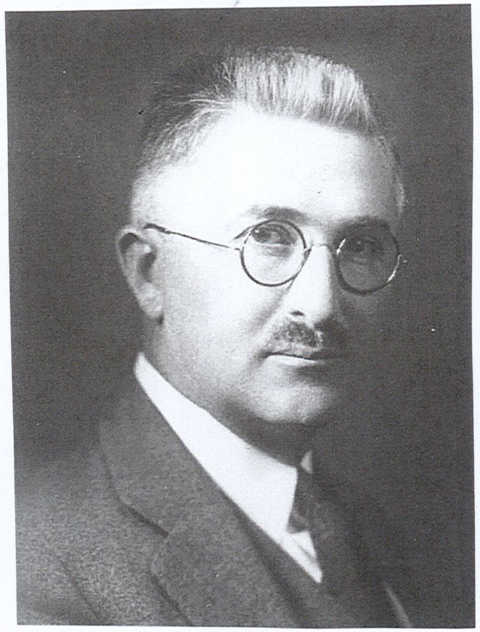
L'article fondateur
de 1948

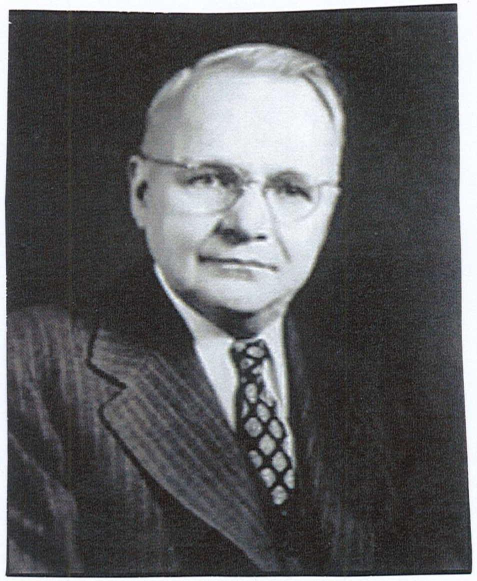
BELL LABS

BELL 实验室

Ralph HARTLEY

1888 - 1970





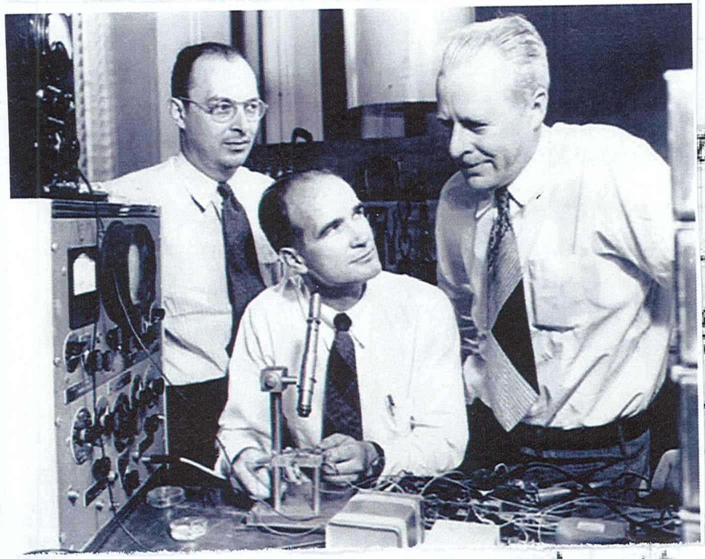
Harry NYQUIST

1889 - 1976

- Transmission of information
  Bell System Technical Journal 1928

- Certain factors affecting Telegraph speed
  Bell System Technical Journal 1924
- Certain topics in Telegraph Transmission theory
  Transactions AIEE 1928

BELL LABS 1948



John Bardeen

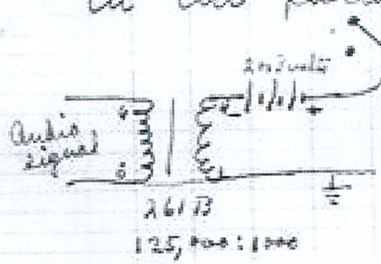
Walter Brattain

William Shockley

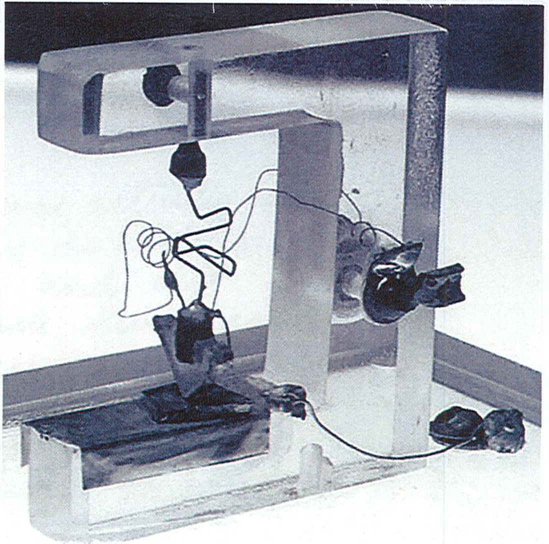
inventent le TRANSISTOR

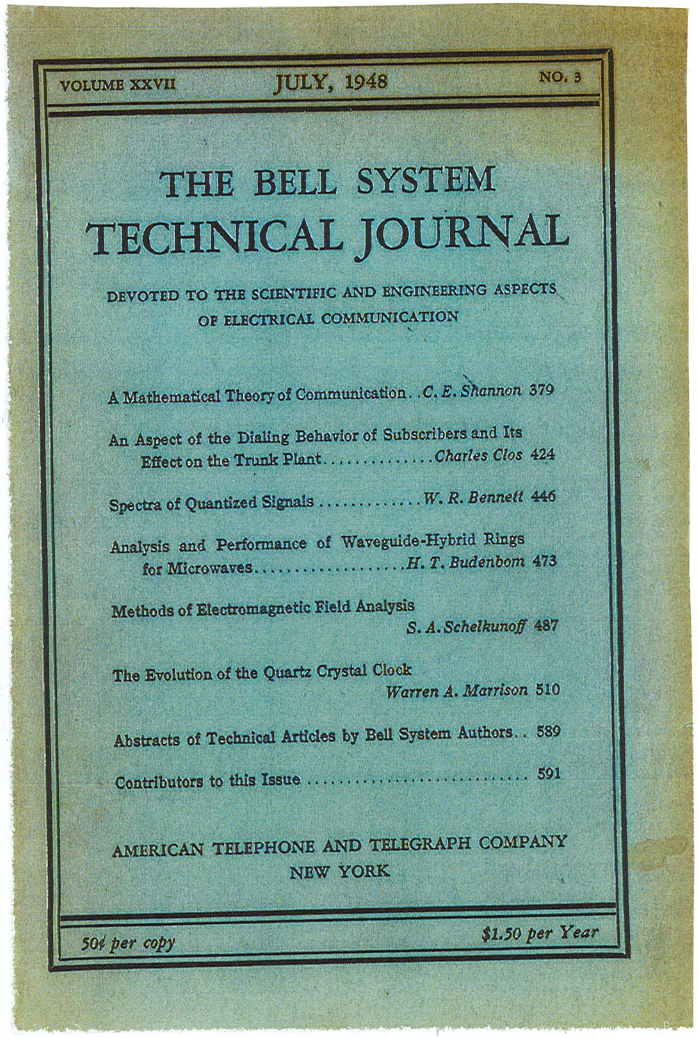
DATE Dec 24 194
CASE No. 38139-7

tained the following A.C.
at 1000 cycles
.015 R.M.S. volts E_p = 1.5 R.M.S volt
~~xxxx~~ or
.4×10⁻⁷ watts $P_p = 2.25 \times 10^{-5}$
gain 100 Power gain 40
ent less $\frac{1}{2.5}$
cuit was then connected
in the following circuit.



This circuit
over and
the device in
gain in spe
heard and
presentation
change in
measurement





VOLUME XXVII   JULY, 1948   NO. 3

THE BELL SYSTEM
TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS
OF ELECTRICAL COMMUNICATION

AMERICAN TELEPHONE AND TELEGRAPH COMPANY
NEW YORK

50¢ per copy $1.50 per Year

Claude Shannon fonde la Théorie de l'Information

L'ÉPOQUE

时期

NORBERT WIENER

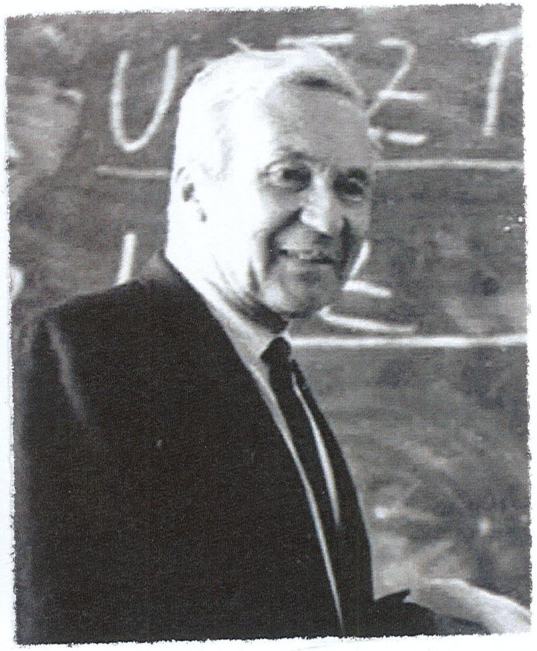
1894 - 1964

JOHN VON NEUMANN

1903 - 1957

ALAN TURING

1912 - 1954



ANDREÏ KOLMOGOROV

1903 - 1987

LE PROBLÈME . . .

问题

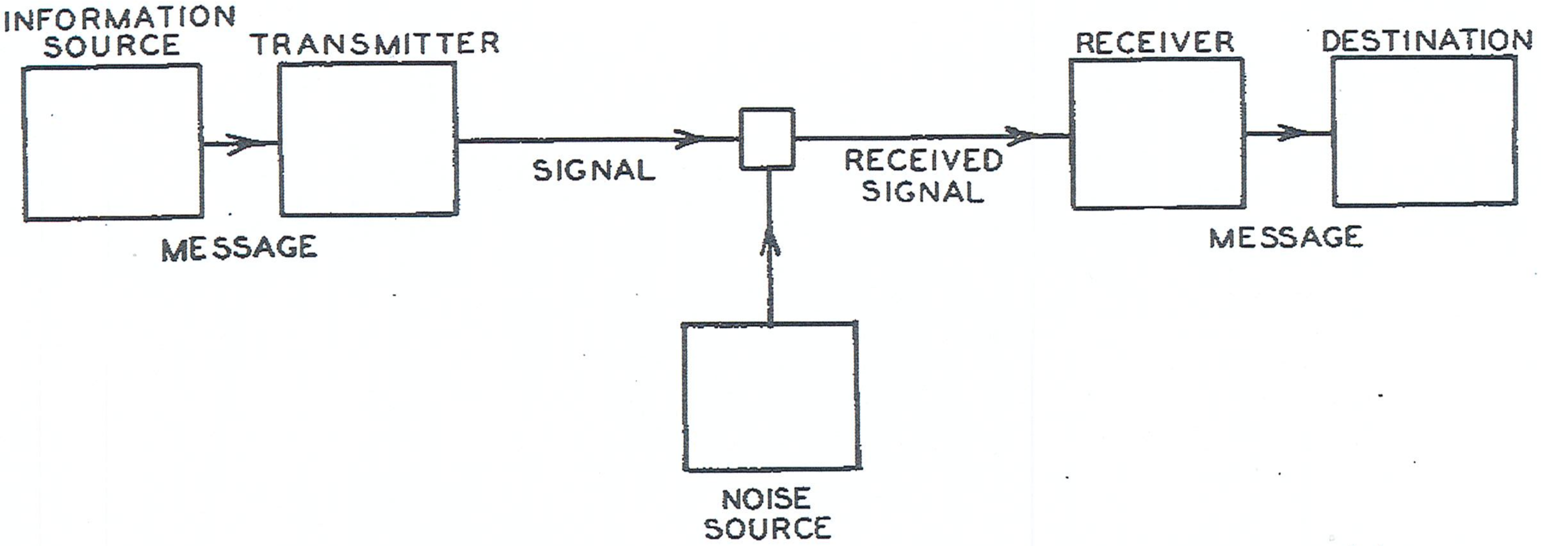
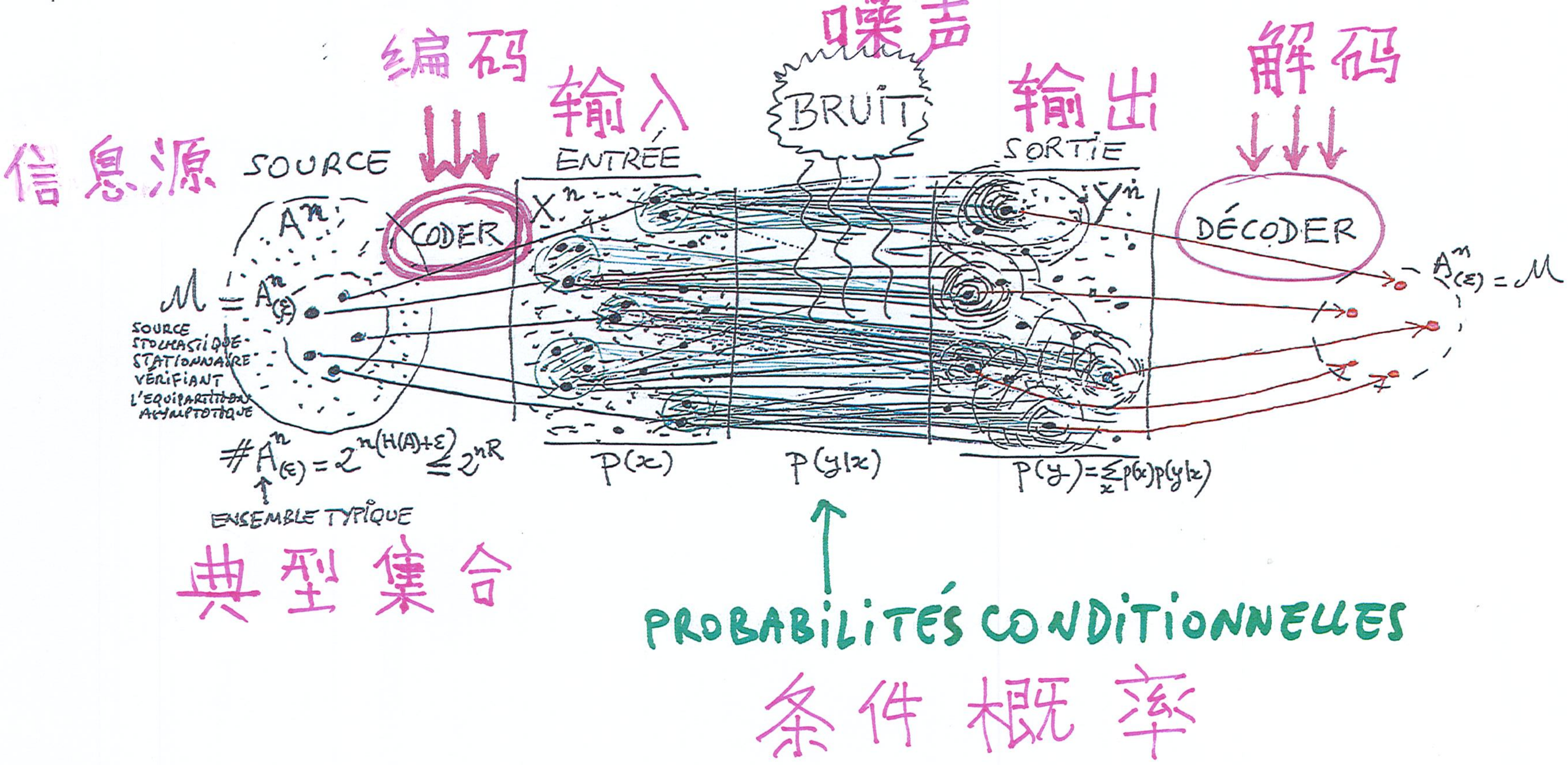


Fig. 1—Schematic diagram of a general communication system. (Shannon 1948)

... SA TRADUCTION MATHÉMATIQUE

... 翻译成数学



噪声
编码
输入
BRUIT
输出
解码
信息源
SOURCE
ENTRÉE
SORTIE
A^n
CODER
X^n
Y^n
DÉCODER
$\mathcal{M} = A^n_{(\varepsilon)}$
SOURCE STOCHASTIQUE STATIONNAIRE VÉRIFIANT L'EQUIPARTITION ASYMPTOTIQUE
$A^n_{(\varepsilon)} = \mathcal{M}$
$\# A^n_{(\varepsilon)} = 2^{n(H(A)+\varepsilon)} \leq 2^{nR}$
$p(x)$
$p(y|x)$
$p(y) = \sum_x p(x)p(y|x)$
ENSEMBLE TYPIQUE
典型集合
PROBABILITÉS CONDITIONNELLES
条件概率

LES REMARQUES FONDATRICES

① SEULE COMPTE, LA NATURE STATISTIQUE DES MESSAGES À TRANSMETTRE

唯一重要的是待传信息集合的统计本质

- <u>HYPOTHÈSE SIMPLISTE</u> (BERNOULLI)

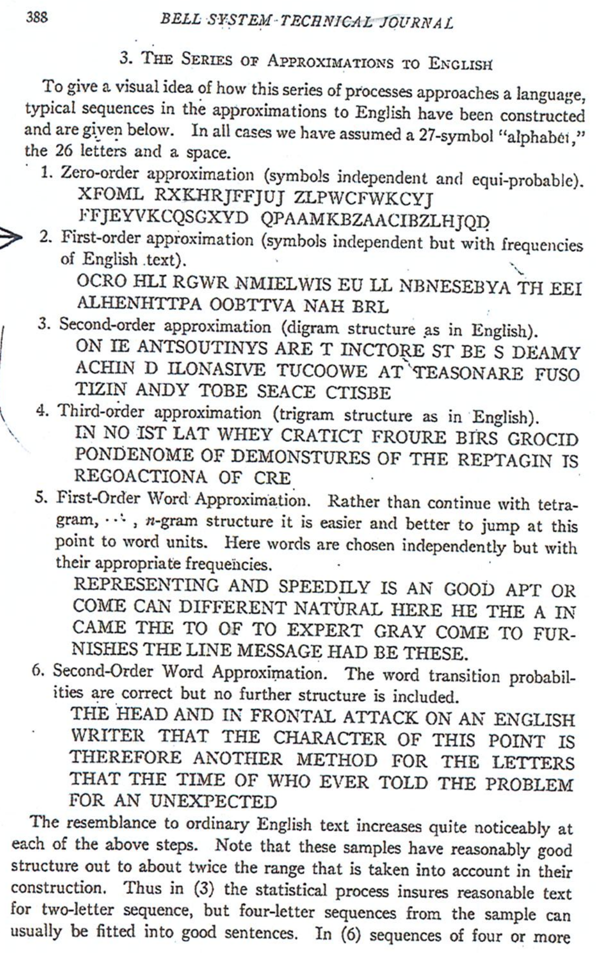
简单的假设 Chaque symbole est affecté d'une probabilité indépendamment des symboles qui précèdent

独立字母

- <u>HYPOTHÈSE PLUS RÉALISTE</u> (MARKOV)

更现实的假设 La probabilité d'un symbole dépend des symboles qui précèdent

非独立字母



388 *BELL SYSTEM TECHNICAL JOURNAL*

3. The Series of Approximations to English

To give a visual idea of how this series of processes approaches a language, typical sequences in the approximations to English have been constructed and are given below. In all cases we have assumed a 27-symbol "alphabet," the 26 letters and a space.

1. Zero-order approximation (symbols independent and equi-probable).
   XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGXYD QPAAMKBZAACIBZLHJQD
2. First-order approximation (symbols independent but with frequencies of English text).
   OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL
3. Second-order approximation (digram structure as in English).
   ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TUCOOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE
4. Third-order approximation (trigram structure as in English).
   IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE
5. First-Order Word Approximation. Rather than continue with tetragram, ⋯ , n-gram structure it is easier and better to jump at this point to word units. Here words are chosen independently but with their appropriate frequencies.
   REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME CAN DIFFERENT NATURAL HERE HE THE A IN CAME THE TO OF TO EXPERT GRAY COME TO FURNISHES THE LINE MESSAGE HAD BE THESE.
6. Second-Order Word Approximation. The word transition probabilities are correct but no further structure is included.
   THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHARACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED

The resemblance to ordinary English text increases quite noticeably at each of the above steps. Note that these samples have reasonably good structure out to about twice the range that is taken into account in their construction. Thus in (3) the statistical process insures reasonable text for two-letter sequence, but four-letter sequences from the sample can usually be fitted into good sentences. In (6) sequences of four or more

LA SIGNIFICATION DES MESSAGES N'INTERVIENT PAS.

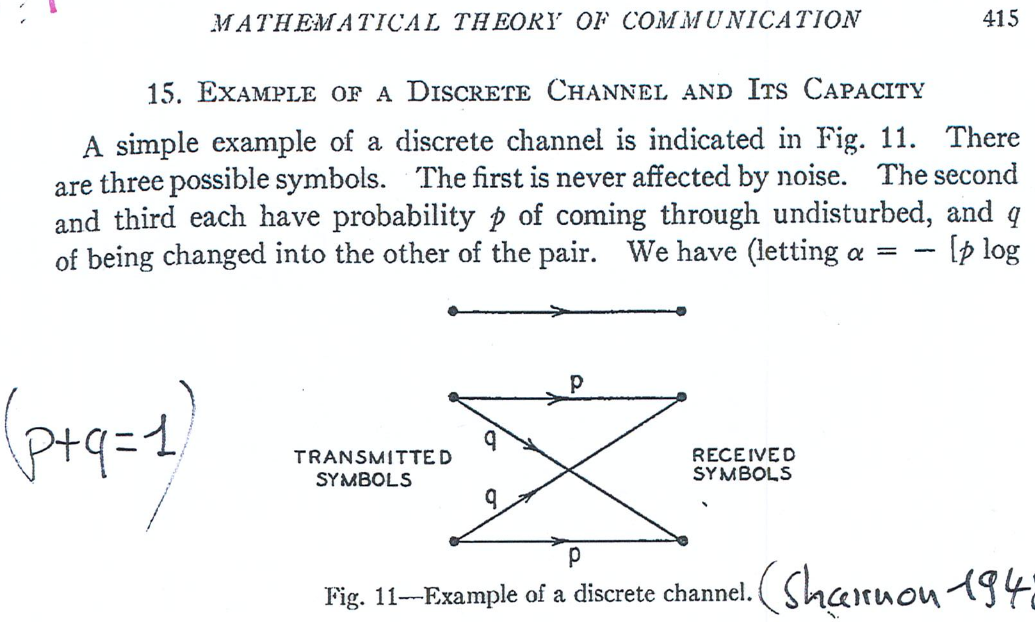
没有信息语义的介入

同样地，传输信道的特征在于其统计本质（条件概率）

2 LE CANAL DE TRANSMISSION EST ÉGALEMENT CARACTÉRISÉ PAR SA NATURE STATISTIQUE (PROBABILITÉS CONDITIONNELLES)

EXEMPLE LE PLUS SIMPLE : LE CANAL BINAIRE SYMÉTRIQUE

最简单的例子：二进制对称传输道



MATHEMATICAL THEORY OF COMMUNICATION 415

15. Example of a Discrete Channel and Its Capacity

A simple example of a discrete channel is indicated in Fig. 11. There are three possible symbols. The first is never affected by noise. The second and third each have probability p of coming through undisturbed, and q of being changed into the other of the pair. We have (letting $\alpha = -[p \log$

$(p+q=1)$

Fig. 11—Example of a discrete channel. (Shannon 1948)

3 TOUT MESSAGE PEUT ÊTRE REPRÉSENTÉ PAR UNE SUITE

$a_1 a_2 \cdots\cdots a_{n-1}\, a_n$, où $a_i = 0$ ou 1 (BITS)

任何讯息可以表示为一个二进制序列

大数定律

LA LOI DES GRANDS NOMBRES

JEU DE PILE OU FACE : SUITES $a_1 a_2 ---- a_n$ DE LANCERS INDÉPENDANTS : (BERNOULLI)

掷硬币

PILE (0) FACE (1)

Probabilité p > Probabilité $q = 1-p$

概率

Ex : 001 000 110 100001 ...

非独立

SUITES TYPIQUES

"High probability messages"

典型的序列

$\begin{cases} \text{大约 } pn \text{ 个 } 0 \\ \text{大约 } qn \text{ 个 } 1 \end{cases}$

它们都具有大约同样的概率（因为顺序不重要）

- AVEC UNE PROBABILITÉ D'AUTANT PLUS GRANDE QUE LE NOMBRE n DE LANCERS EST GRAND, LA SUITE CONTIENDRA
  - ENVIRON pn 0's
  - ENVIRON qn 1's

- L'ORDRE DES LANCERS N'INTERVENANT PAS, LES SUITES TYPIQUES ONT TOUTES ENVIRON LA MÊME PROBABILITÉ, ALORS QUE L'ENSEMBLE DES AUTRES A UNE PROBABILITÉ TENDANT VERS 0 SI N TEND VERS L'∞.

熵 L'ENTROPIE

$$\text{Probabilité}\left(\underbrace{0\,0\,1\,0\,1\,1\,1\,0\,0\,1\,0\ldots}_{\text{message ayant } n_0 \text{ 0's et } n_1 \text{ 1's}}\right) = p^{n_0}q^{n_1} = 2^{(n_0\log p + n_1\log q)}$$

概率

logarithmes en base 2

$(\text{i.e. } 2^{\log x} = x = \log 2^x)$

对数都是以2为底数

$$\text{Probabilité}\left(\overbrace{\underbrace{\text{message typique}}_{n_0 = pn+\cdots,\ n_1 = qn+\cdots}}^{\text{典型信息}}\right) = 2^{-(nh(p)+\cdots)},\ \text{où}\ \frac{\cdots}{n} \xrightarrow[n\to+\infty]{} 0$$

概率

$$\boxed{h(p) = -p\log p - q\log q = p\log\frac{1}{p} + q\log\frac{1}{q}}$$

est par définition l'**ENTROPIE**

根据定义，$h(p)$ 是熵

Dans la suite, nous ferons comme si les $\ldots$ n'existaient pas

后面我们假装这些点儿不存在

D'OÙ L'ON DÉDUIT UN RÉSULTAT CLÉ :

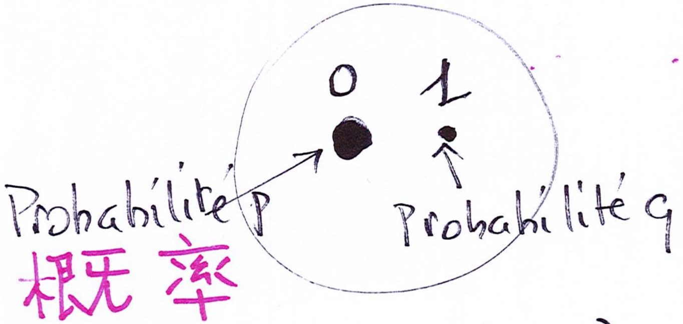
从中我们推断一个很重要的结果：

L'ÉQUIPARTITION ASYMPTOTIQUE DE LA PROBABILITÉ

概率的渐进等分布

(OU L'ENTROPIE COMME CAPACITÉ DE COMPRESSION)

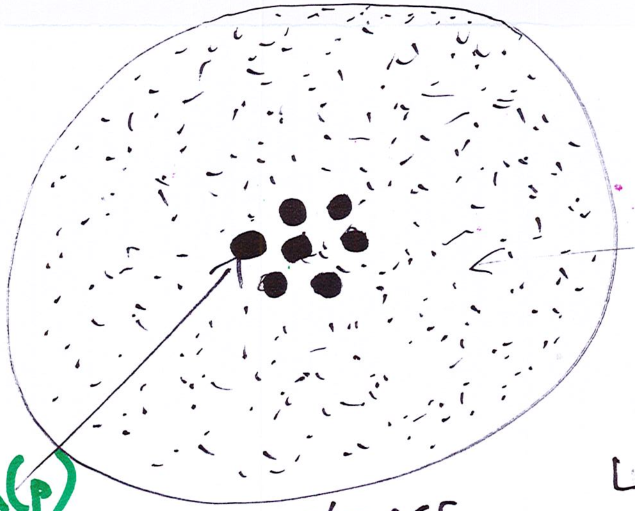
(或熵 = 压缩能力)



概率

ALPHABET = {0,1}

字母表



$\sim 2^n$ MESSAGES NON TYPIQUES

不典型的信息

LEUR PROBABILITÉ TOTALE $\longrightarrow 0$ si $n \to \infty$

它们的总概率 $\to 0$ 如果 $n \to \infty$

$\sim 2^{nh(p)}$ MESSAGES TYPIQUES

典型信息

AYANT CHACUN UNE PROBABILITÉ $\sim 2^{-nh(p)}$

每个有概率 $\sim 2^{-nh(p)}$

REMARQUE : 备注

Si l'alphabet A a r lettres $A_1, \dots, A_r$ avec les probabilités $p_1, \dots, p_r$, $\sum_{i=1}^{r} p_i = 1$, seule change la formule de l'entropie :

$$\boxed{H(A) = h(p_1, \dots, p_r) = \sum_{i=1}^{r} p_i \log \frac{1}{p_i}}$$

如果字母表A有r个字母$A_1, \dots, A_r$

以及概率 $p_1, \dots, p_r$, $\sum_{i=1}^{r} p_i = 1$,

唯一的变化是熵的公式：

$$H(A) = h(p_1, \dots, p_r) = \sum_{i=1}^{r} p_i \log p_i$$

已出现在玻尔兹曼的工作中

DÉJÀ CHEZ BOLTZMANN

$N = \sum_{i=1}^{r} N_i$ particules indistinguables

N_i particules ayant l'énergie E_i

→ MACROÉTAT $(N_1, \ldots, N_r)$ d'énergie $E = \sum_{i=1}^{r} N_i E_i$

↑

RÉALISÉ DE $W = \dfrac{N!}{N_1! \ldots N_r!}$ FAÇONS (MICROÉTATS)

↑

Thermodynamische Wahrscheinlichkeit

(= probabilité thermodynamique du macroétat)

1844 – 1906

SI N GRAND, MAXIMISER W REVIENT À MAXIMISER

$\log W \sim N h(p_1, \ldots, p_r)$, où $p_i = \dfrac{N_i}{N}$ = probabilité de l'énergie E_i

LA MAXIMISATION SOUS LA CONTRAINTE $\sum_{i=1}^{r} p_i E_i = e$ IMPLIQUE LES FORMULES CLASSIQUES : (énergie moyenne fixée)

$$\boxed{p_i = \frac{N_i}{N} = \frac{e^{-\lambda E_i}}{\sum_{j=1}^{r} e^{-\lambda E_j}} \text{, où } \sum_{i=1}^{r} E_i \frac{e^{-\lambda E_i}}{\sum_{j=1}^{r} e^{-\lambda E_j}} = e}.$$

熵函数 LA FONCTION ENTROPIE

$$h(p_1,\dots,p_n) = \sum_{i=1}^{n} p_i \log \frac{1}{p_i}$$

The entropy in the case of two possibilities with probabilities p and $q = 1 - p$, namely

$$H = -(p \log p + q \log q)$$

is plotted in Fig. 7 as a function of p.

The quantity H has a number of interesting properties which further substantiate it as a reasonable measure of choice or information.

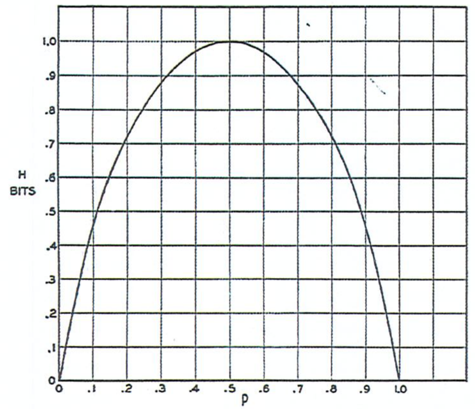
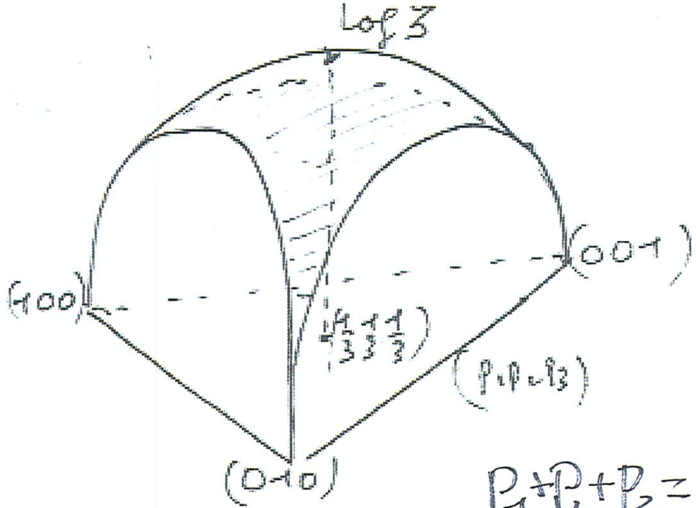


Fig. 7—Entropy in the case of two possibilities with probabilities p and $(1 - p)$.



$p_1+p_2+p_3=1$

$n=2$ (Shannon 1948)

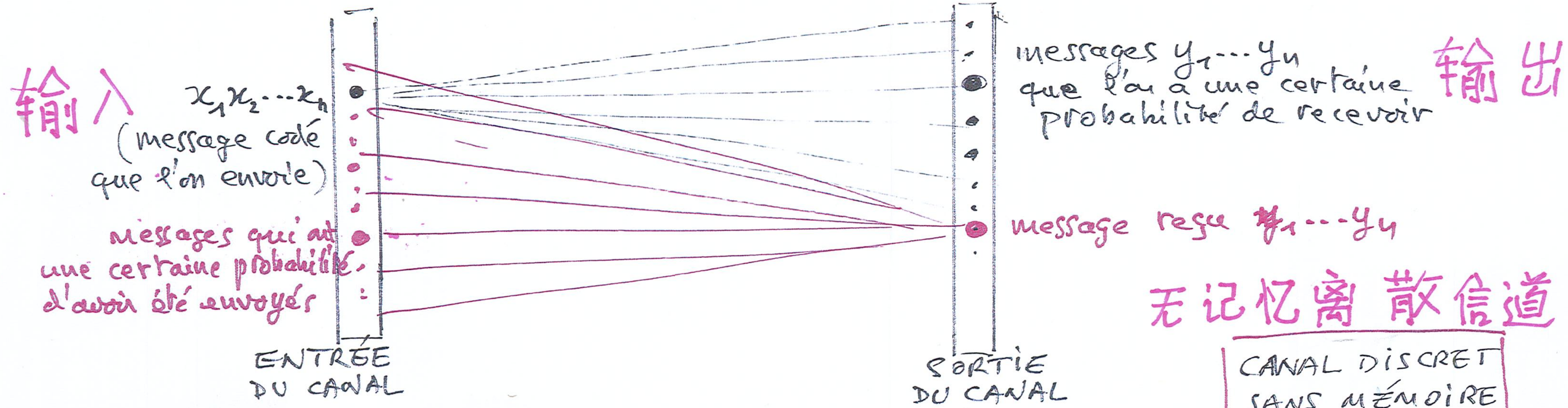
$n=3$

(SI TOUTES LES LETTRES DE L'ALPHABET SONT ÉQUIPROBABLES, TOUT MESSAGE EST TYPIQUE)

如果字母表的字母都有一样的概率，每个信息都是典型的。

用条件概率量化信道的误差

QUANTIFIER LES ERREURS DU CANAL PAR LES PROBABILITÉS CONDITIONNELLES


输入
$x_1x_2\cdots x_n$
(message codé que l'on envoie)
messages qui ont une certaine probabilité d'avoir été envoyés
ENTRÉE DU CANAL
messages $y_1\cdots y_n$ que l'on a une certaine probabilité de recevoir
输出
message reçu $y_1\cdots y_n$
SORTIE DU CANAL
无记忆离散信道

CANAL DISCRET SANS MÉMOIRE

- Probabilité de recevoir $y_1\cdots y_n$ si on a envoyé $x_1\cdots x_n$
  发送 $x_1,\dots,x_n$, 收到 $y_1,\dots,y_n$ 的概率
- Probabilité d'avoir envoyé $x_1\cdots x_n$ si on a reçu $y_1\cdots y_n$
  收到 $y_1,\dots,y_n$, 发送 $x_1\cdots x_n$ 的概率

$$\left.\begin{aligned} Pr_{x_1\cdots x_n}(y_1\cdots y_n) &= \prod_{i=1}^{n} Pr_{x_i}(y_i) \\ Pr_{y_1\cdots y_n}(x_1\cdots x_n) &= \prod_{i=1}^{n} Pr_{y_i}(x_i) \end{aligned}\right\}\ \text{Hypothèse Bernoulli}$$

$\Updownarrow$

条件熵

ENTROPIE CONDITIONNELLE

$$Y = \{ y_1, \dots, y_j, \cdots y_L \}$$

$$Pr_{x_i}(y_1) \dots Pr_{x_i}(y_j) \dots Pr_{x_i}(y_L) \qquad i = 1, \dots, K$$

Probabilités conditionnelles

条件概率

$$H_{x_i}(Y) = \sum_{j=1}^{L} Pr_{x_i}(y_j) \log \frac{1}{Pr_{x_i}(y_j)}$$

espérance / X

期望值

$$\boxed{H_X(Y) = \sum_{i=1}^{K} Pr(x_i) H_{x_i}(Y)}$$

ENTROPIE CONDITIONNELLE

条件熵

$$X = \left\{ \begin{array}{c} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_K \end{array} \right\} \begin{array}{c} Pr(x_1) \\ \vdots \\ Pr(x_i) \\ \vdots \\ Pr(x_K) \end{array}$$

$$Y = \{ y_1, \dots, y_j, \cdots y_L \}$$

$$Pr(y_1) \dots Pr(y_j) \cdots Pr(y_L)$$

$$Pr(y_j) = \sum_{i=1}^{K} Pr(x_i) Pr_{x_i}(y_j)$$

平均式的推理

UN RAISONNEMENT EN MOYENNE ...

A, X, Y alphabets avec probas
p(A) p(X) p(Y) ⟹ ENTROPIES H(A), H(X), H(Y) 熵
ENTROPIES CONDITIONNELLES $H_X(Y), H_Y(X)$
条件熵

具有概率的字母表

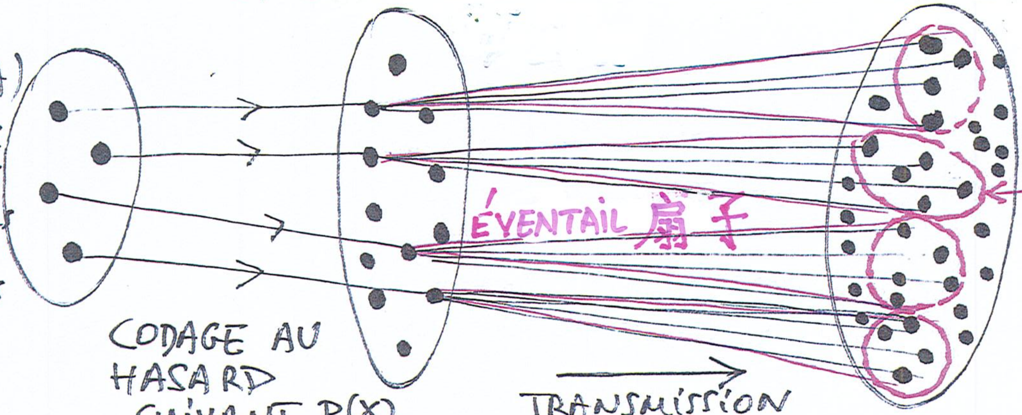
$2^{nH(X)}$个输入典型的信息

$2^{nH(X)}$ messages typiques d'entrée

$2^{nH(Y)}$ messages typiques de sortie

$2^{nH(Y)}$个输出典型的信息

$2^{nH(A)}$ messages typiques susceptibles d'être envoyés



$2^{nH(A)}$可以被发送的典型信息

CODAGE AU HASARD SUIVANT P(X)

随机编码

TRANSMISSION

传输

En moyenne, $2^{nH_X(Y)}$ messages reçus peuvent provenir d'un même message transmis.

平均来说，$2^{nH_X(Y)}$个信息可以出自于同一个信息。

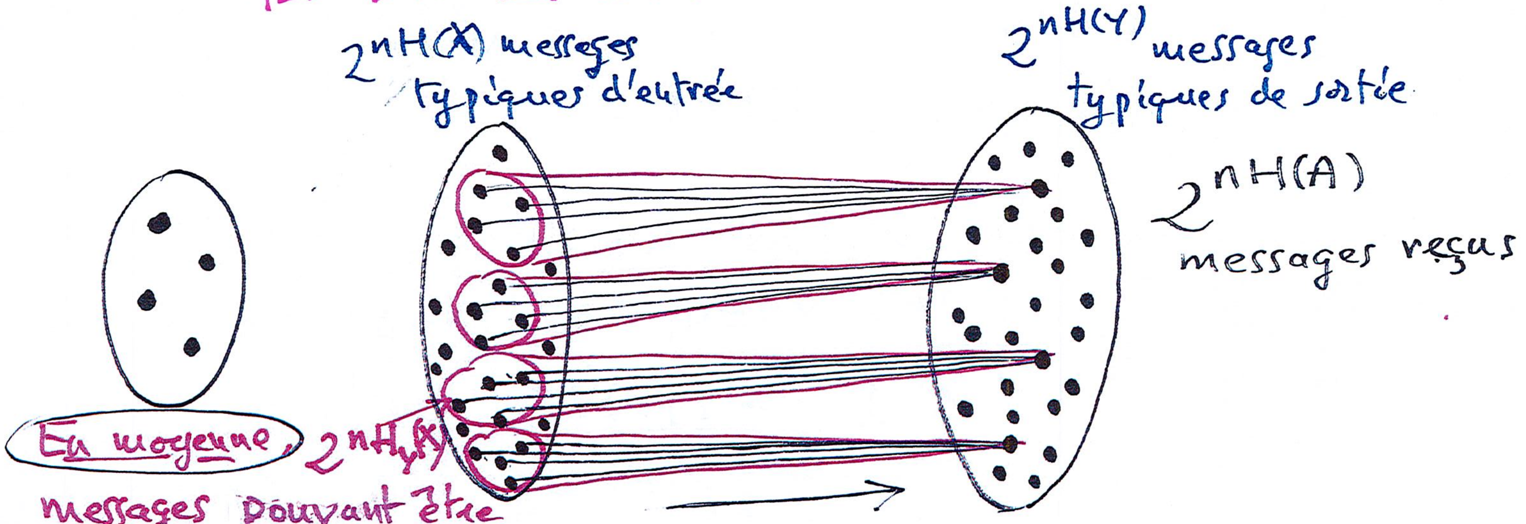
类似于堆橙子的推理，
扇子终端的分离条件
为 $2^{nH(A)} \times 2^{nH_x(Y)} < 2^{nH(Y)}$

UN RAISONNEMENT DU TYPE "EMPILEMENT D'ORANGES" DONNE COMME CONDITION DE DISJONCTION DES EXTRÉMITÉS DES ÉVENTAILS $2^{nH(A)} \times 2^{nH_x(Y)} < 2^{nH(Y)}$,

i.e. $$\boxed{H(A) < H(Y) - H_x(Y)}$$

... LE MÊME DE L'AUTRE CÔTÉ

在该图的另一侧同样地


$2^{nH(X)}$ messages typiques d'entrée
$2^{nH(Y)}$ messages typiques de sortie
$2^{nH(A)}$ messages reçus
En moyenne, $2^{nH_y(X)}$
messages pouvant être à l'origine du message reçu

("EQUIVOCATION")

含糊性

LA CONDITION DE DISJONCTION EST $2^{nH(A)} \times 2^{nH_y(X)} < 2^{nH(X)}$,

i.e. $\boxed{H(A) < H(X) - H_y(X)}$

分离的条件是 $2^{nH(A)} \cdot 2^{nH_y(X)} < 2^{nH(X)}$

INFORMATION MUTUELLE ET CAPACITÉ

互信息

$$\boxed{\begin{aligned} I(X,Y) &= H(Y) - H_x(Y) \\ &= H(X) - H_y(X) \end{aligned}}$$

容量

LA CONDITION DE DISJONCTION DES ÉVENTAILS EST DONC LA MÊME DANS LES DEUX CAS :

因此，在两种情况下，扇子的分离条件都是相同的：

$$\boxed{H(A) < I(X,Y)}$$

CAPACITÉ

容量

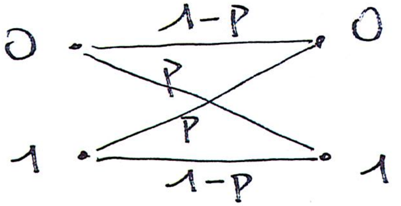
$$\boxed{C = \underset{\substack{p(x)\ \text{probabilité} \\ \text{sur } X}}{\mathrm{Max}}\ I(X,Y)} \Rightarrow$$

Condition de disjonction

$$H(A) < C$$

分离的条件

例子 EXEMPLE :



$$C = 1 - h(p) = 1 - p \log \frac{1}{p} - (1-p) \log \frac{1}{1-p}$$

LE PROBLÈME ÉTAIT DANS L'AIR : QUELQUES PUBLICATIONS CONTEMPORAINES

一些当代出版物

A.G. CLAVIER

Evaluation of Transmission Efficiency According to Hartley's Expression of Information Content*

By A. G. CLAVIER

Federal Telecommunication Laboratories, Incorporated, Nutley, New Jersey

THE EFFICIENCY of any transmission system can be estimated as the ratio of the amount of information obtained in the reproduced message at the receiving end to the total amount of information that could have been transmitted theoretically in the propagation medium for the bandwidth and signal-to-noise ratio necessitated by the signals actually utilized. The amount of information is evaluated according to the expression given by R. V. L. Hartley for telegraphic signals, extended to the case of telephony and the presence of noise. The transmission efficiency thus defined is calculated for the main pulse-transmission systems (pulse-amplitude modulation, pulse-time modulation, pulse-count modulation, pulsed-frequency modulation) under certain assumptions. The various expressions obtained are helpful in the comparison between those systems, though it must not be overlooked that simplicity and cost of equipment will undoubtedly introduce other factors in the final choice for any given transmission problem.

• • •

1. General

To arrive at an evaluation of the amount of information conveyable through any type of transmission system, Hartley[1] started with the consideration of what would now be called pulse-amplitude-modulated (telegraphic) signals.

Let n_0 pulses be transmitted per second, each pulse having one of S discrete levels. The number of distinct sequences that can occur during a transmission time of t seconds is S^{n_0t}. The measure of the amount of information that such a system is capable of transmitting is felt to be a function of this number of sequences. Barrin any analysis of the psychological content of th message, which these signals are meant to con vey, it becomes intuitive that this measure shoul be proportional to the time during which th transmission takes place. This determines at onc the type of function to be used, so that, accord ing to Hartley, the measure of the amount c information the signals can convey is expressed b

$$H=k_0\cdot\log S^{n_0t}=k_0\cdot n_0\cdot t\cdot\log S,$$

k_0 being a constant that, together with the bas of the logarithms used, defines the value of conventional unit of information.

This definition is easily extended to the cas of the n_0 pulses belonging to n_c separate chan nels. The pulses pertaining to any one channe are divided into n_r subframes per second, an each subframe in its turn includes n pulses o digits. It is obvious that the total amount c information is then

$$H=k_0\cdot n_c\cdot n_r\cdot n\cdot t\cdot\log S,$$

and is thus equal to n_c times the amount c information assigned to any individual channe

These signals, however, will have to be trans mitted through a certain path, which may in clude wire transmission lines, lumped circuit and radio links. A certain number of errors ma be introduced during that transmission. Thes errors may result from a number of causes, som originating in the physical properties of the pat itself, such as signal distortion and interferenc as well as from thermal noise. Others may com from outside, such as interference from othe transmissions or atmospherics. Moreover, th number of errors may depend on meteorologica factors, such as temperature and humidity leading to fluctuating distortion or fading phe nomena. To secure a sufficiently correct repro duction of the "message" at the receiving end it will thus be necessary to adopt a minimun

* Presented, New York Section, Institute of Radio Engineers, New York, New York, November 12, 1947.

[1] R. V. L. Hartley, "Transmission of Information," *Bell System Technical Journal*, v. 7, pp. 535–563; July, 1928.

J. LAPLUME

PHYSIQUE MATHÉMATIQUE. — *Sur le nombre de signaux discernables en présence du bruit erratique dans un système de transmission à bande passante limitée.* Note de M. Jacques Laplume. (avril 1948)

Soit un signal $s(t)$ de durée T, défini par sa densité spectrale d'énergie $S(f)$. Ce signal peut être transmis sans distorsion exagérée à travers un circuit de bande passante W pourvu que la quasi-totalité de l'énergie soit comprise dans la bande $0\leqq f\leqq W$. Le bruit erratique, de densité spectrale $\beta(f)=\text{const.}$, a pour effet d'introduire dans la mesure de la composante $S(f)$ une incertitude $\Delta S(f)=\beta$. D'autre part, si l'on imagine que le spectre est analysé au moyen de $n=W/\Delta f$ circuits de bande passante Δf accordés sur les différentes régions du spectre, la bande Δf de ces circuits doit être de l'ordre de grandeur de $1/T$ au plus afin que la durée d'établissement du régime permanent dans ces circuits soit au plus égale à T. Dans le plan S, f, la zone d'incertitude de localisation d'un point figuratif est donc un rectangle de base Δf et de hauteur β.

L'aire rectangle

d'ince oyenne P

du sig

On partissant

ces N colonnes

vertica ernables.

Cette manières

différe ın sign

discerı

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Research Laboratory of Electronics

Technical Report No. 32 December 15, 1947

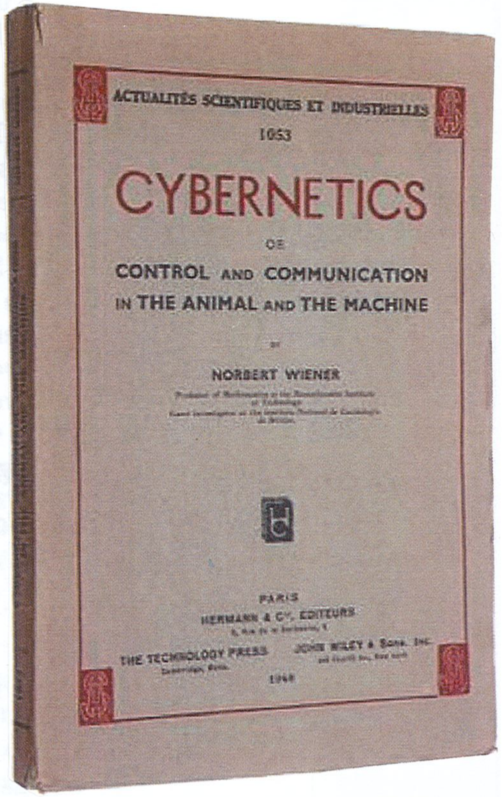
SOME FUNDAMENTAL CONSIDERATIONS CONCERNING NOISE REDUCTION[1] AND RANGE IN RADAR AND COMMUNICATION

Stanford Goldman

S. GOLDMAN

Abstract

A general analysis based upon information theory and the mathematical theory of probability is used to investigate the fundamental principles involved in the transmission of signals through a background of random noise. Three general theorems governing the probability relations between signal and noise are proved, and one is applied to investigate the effect of pulse length and repetition rate on radar range. The concept of "generalized selectivity" is introduced and it is shown how and why extra bandwidth can be used for noise reduction.



ACTUALITÉS SCIENTIFIQUES ET INDUSTRIELLES

1053

CYBERNETICS

OR

CONTROL AND COMMUNICATION IN THE ANIMAL AND THE MACHINE

BY

NORBERT WIENER

PARIS

HERMANN & Cie, EDITEURS

THE TECHNOLOGY PRESS

JOHN WILEY & SONS, INC.

1948

THEORETICAL LIMITATIONS ON THE RATE OF TRANSMISSION OF INFORMATION

WILLIAM G. TULLER

W.G. TULLER

TECHNICAL REPORT NO. 114

APRIL 23, 1949

香农：这是一个定理！

SHANNON : C'EST UN THÉORÈME !

$$\boxed{H(A) < C} \iff$$

CONDITION DE DISJONCTION EN MOYENNE

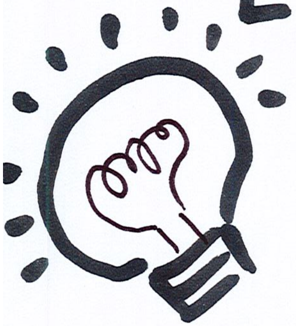
平均分离条件

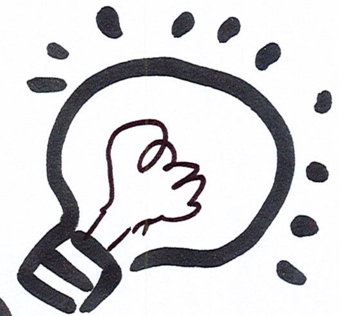
IL EXISTE UN CODAGE TEL QUE LA PROBABILITÉ D'ERREUR TENDE VERS 0 LORSQUE n TEND VERS L'INFINI.

存在一个编码讯息的办法，以便误差的概率趋近0为n趋近无穷大。

L'IDÉE SUPERBE DE LA PREUVE

很美的证明构想



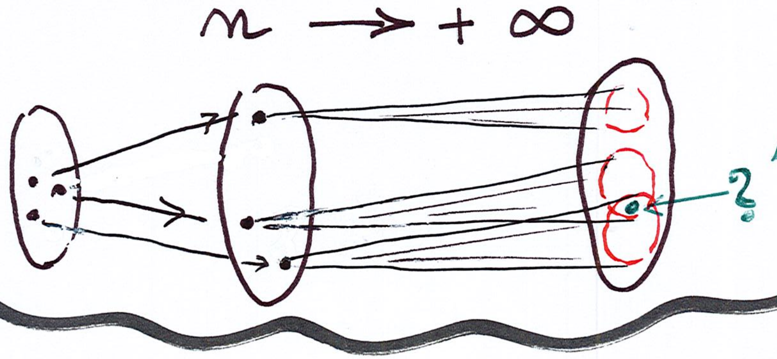


CALCULER LA PROBABILITÉ **MOYENNE** D'ERREUR (I.E. LA PROBABILITÉ D'INTERSECTION DES $2^{nH(A)}$ ÉVENTAILS TYPIQUES) SUR UN ENSEMBLE DE CODES CHOISIS AU HASARD

对随机编码计算其误差的平均概率

||

($=2^{nH(A)}$个典型扇子终端的交集)



LE THÉORÈME DE SHANNON EST PUREMENT THÉORIQUE

香农的定理是一个纯粹理论的定理

IL A IMMÉDIATEMENT PROVOQUÉ LA RECHERCHE DE CODES EFFICACES

它立刻引起了高效代码的研究

Le problème du codage :
Comment introduire intelligemment
la REDONDANCE ?

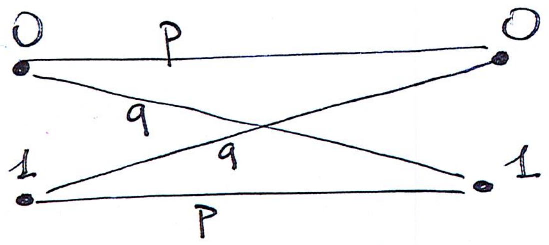
编码的问题：
如何聪明地导入
冗余？

CODES À RÉPÉTITION : Simples mais impraticables

重复编码

On répète n fois chaque symbole et on décode à la majorité

EXEMPLE :
例子



$p > q$

信息	MESSAGE	1	1	0	1
发送	ÉMISSION	11111	11111	00000	11111
接收	RÉCEPTION	10011	01111	00010	01011
解码	DÉCODAGE	1	1	0	1

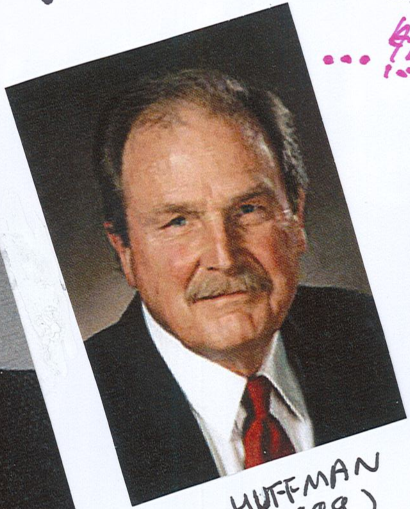
LA PROBABILITÉ D'ERREUR TEND VERS 0 LORSQUE N TEND VERS L'INFINI MAIS ON A DIVISÉ PAR N LA CAPACITÉ DE TRANSMISSION

当n趋于无穷时，误差的概率趋于0但是传输能力要除以n.

... ET LES CODES FLEURIRENT ...

...然后,代码们都开花了...

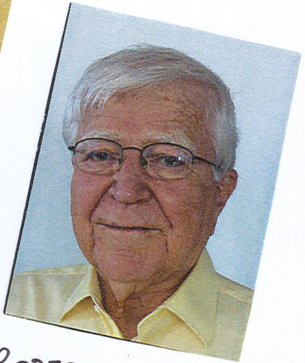
ROM WARSHAMOV (1927-1999)



DAVID HUFFMAN (1925-1999)

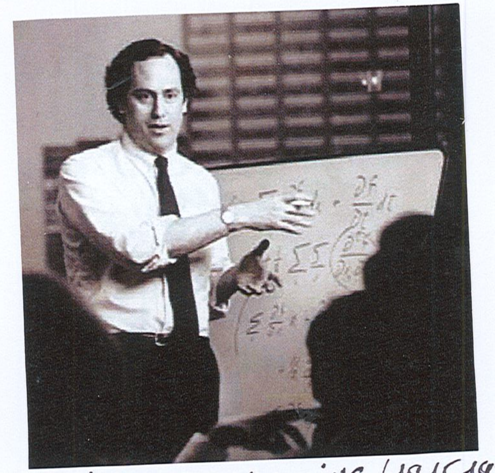
NEIL SLOANE (1939-)

EDGAR GILBERT (1923-2013)

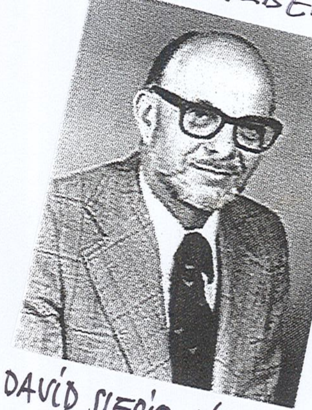


ROBERT GALLAGER (1931-)

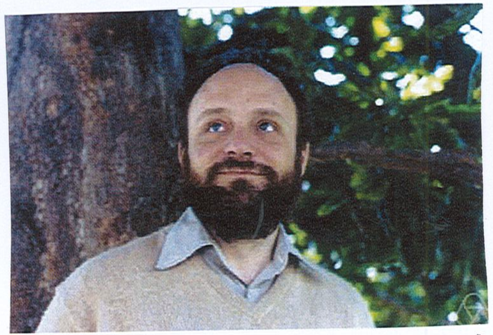
JESSIE MACWILLIAMS (1917-1990)



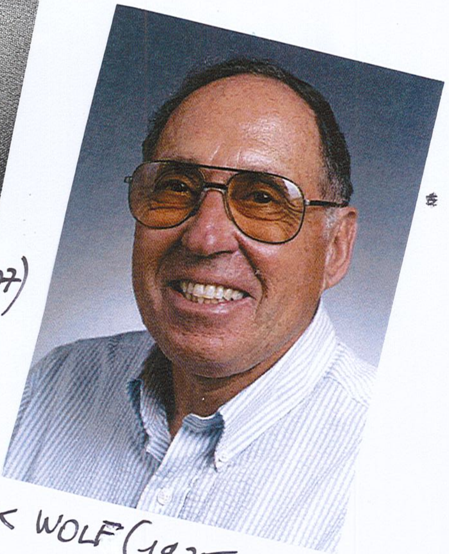
RICHARD HAMMING (1915-1998)



DAVID SLEPIAN (1923-2007)



EDWIN BERLEKAMP (1940-)



JACK WOLF (1935-2011)

...JUSQU'À LA DÉCOUVERTE DES TURBOCODES QUI RÉALISENT PRATIQUEMENT LA LIMITE QU'IMPOSE LE THÉORÈME DE SHANNON!

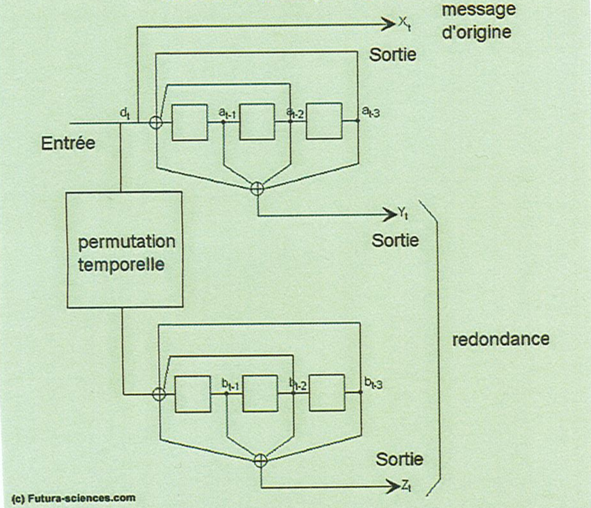
TURBO 码 （增压气码）

NEAR SHANNON LIMIT ERROR - CORRECTING CODING AND DECODING : TURBO-CODES (1)

Claude Berrou, Alain Glavieux and Punya Thitimajshima

Proceedings of ICC '93, Geneva, pp. 1064-1070, May 1993

***Illustration 4* : Un exemple de turbocode**



(c) Futura-sciences.com

Le message binaire {d}, constitué par la suite des données d_t, est codé deux fois par une **concaténation parallèle** de 2 codeurs, la première fois suivant son ordre naturel par le codeur supérieur, la deuxième fois dans un ordre bouleversé, par le codeur inférieur. Les redondances (Y_t et Z_t) forment deux messages {Y} et {Z} synonymes de la séquence d'entrée {d}, qui est également transmise en tant que séquence {X}.

L'idée est d'entrelacer deux petits codes convolutifs, un peu comme dans une grille de mots croisés où l'on reprend plusieurs fois définitions horizontales et définitions verticales (effet TURBO!)

CLAUDE BERROU

MERCI
SHANNON!