SJC - 12
June 6, 1957

MEMORANDUM TO: Dr. W. A. Christopherson

SUBJECT: Preliminary Note on the Application of Statistical Decision Theory To Business Data Systems.

This preliminary note is a summary of my evaluation of Middleton's papers on the detection and extraction of signals in noise. In addition to the immediate objective of analyzing data transmission systems, two other levels of application have become apparent. I feel that statistical decision theory is of value in the following areas:

(1) Analysis of detection of signals in noise in data transmission system;

(2) Organization of the analysis of business data systems into a uniform method of treatment which helps one view the relationship of the parts to the whole;

(3) Implimentation of the functioning of applied research in demonstrating the potential social use of basic science and mathematics.

I have debated whether part (1) should be a separate memorandum to the Development Engineer in charge of the Communications Project, part (2) to the Manager of San Jose Systems Research, and part (3) to the Manager of Applied Research. I have concluded there is a greater value in treating the three levels together briefly in a way which is closer to the process of development of the ideas.

The organization of this memorandum is more philosophical than the customary engineering form, because I feel that it will in time expedite a greater specialization in the specific area of detection of signals in noise.

The valuable assistance and comments of Dr. W. A. Gross and Mr. W. E. Dickinson in initiating this study are appreciated. The general value of Middleton's formulation of statistical decision theory in respect to information theory has been reviewed with Mr. Norman Abramson and Professor W. W. Harman of Stanford University.

F. B. Wood

F. B. Wood

FBW:hp

cc: Mr. J. Allen
Dr. M. M. Astrahan
Mr. W. E. Dickinson
Dr. W. A. Gross
Mr. R L. Haug
Mr. E. Hopner
Mr. R. B. Johnson
Mr. W. J. Johnson, Jr.
Mr. M. L. Lesser
Mr. H. Markey
Mr. J. A. McLaughlin
Dr. E. A. Quade
Mr. L. C. Tidball

SJC - 12
June 6, 1957

# Preliminary Note on the Application of Statistical Decision Theory To Business Data System.

## TABLE OF CONTENTS

SJC - 12
June 6, 1957

# Preliminary Note on the Application of Statistical Decision Theory to Business Data Systems

## 1. Introduction

A unique opportunity exists at the present stage of applied communication research in the San Jose Research Laboratory. The need for reducing the error rate in the transmission of business data has stimulated both an experimental and theoretical approach. The timing of this problem is such that techniques developed by statisticians primarily for use in economic problems have recently been shown to overlap with information theory and to be of general value in the detection of signals in noise. The use of this theoretical formulation in analysing our particular problems may yeild multiple benefits to the IBM research program. In addition to organizing an attack on the problem of impulse noise, statistical decision theory offers benefits in the following areas: (1) Developing a perspective of a complete business data processing system for use in System analysis; (2) Analysis of detection problems in other sub-sections of a data system, such as human reading, human keying, machine reading, printing, computer error checking, and buffer translating; (3) Sharpening of the depth of specialization that can be obtained in applied research through the maintenance of an adequate perspective; and (4) Learning of the mathematical tools needed in extending the application of computers to decision making in larger sectors of the economy.

Middleton and Van Meter have organized a systematic treatment of detection and extraction of signals from noise (MI-1, 2). Their basic system is illustrated in Fig. 1. The terminology comes primarily from statistical decision theory or game theory and overlaps with the concepts of information theory. Application of this method may provide, for example, a general set of equations for evaluating data transmission systems. This statistical decision theory has a close relationship to game theory. The fundamental work of A. Wald (Wd-2) upon which Middleton bases much of the analysis was developed by considering the decision problem as a two-person game between Nature and the Observer.

While studying Middleton's paper the question occurred to me, whether a separate analysis would have to be made for inquiry and reply. This would be particularly significant if a different code was used for each direction (SJA-35). This led to the postulation of two systems like that shown in Fig. 1. Then the consideration of human errors in reading as discussed in a recent human engineering research proposal (SJC-6) led to the description of a human character perception channel with equivalent signal, noise, perceived characters, and decisions rules. Now it became apparent that a chain of machine and human data transmission links was growing into a complete business data system "space" as is illustrated in Fig. 2. The term "space" is used here in the mathematical sense of a set of n vectors which determine all the n possible alternatives allowed in the system. The term "sub-space" defines a section of the system consisting of all the possible signals allowed in a particular section such as the inquiry data transmission link, the computer, or a buffer or translator.

Further contemplation on this potential application of statistical decision theory led to the idea that there may be a more general value of this study in exploring ways in which applied research can digest fundamental science and mathematics in the process of applying basic knowledge to social use. The narrow specialization which has developed with the great increase in our knowledge of science and mathematics has resulted in barriers which retard the interchange of ideas between scientists, engineers, and the layman. The successful application of a section of basic science and/or mathematics such as statistical decision theory in backing up and expediting applied research can intensify the applied research and lay the base for a better perspective on the part of the engineer and help establish a more adequate popular explanation of the functioning of basic and applied research.

2. Application to Data Transmission

Data is being collected and studied for the analysis within the framework of Middleton's theoretical formulation for some detection systems now being developed in this laboratory. This can be expected to assist experimental program and to provide a systematic comparison of different systems.

Middleton distinguishes between "detection" and "extraction." This distinction is illustrated in his earlier article (Mi-3, p 49). Detection is defined as a testing of hypotheses: $S_1$ = signal and noise are present.
$S_0$ = only noise is present.
Extraction is defined as the estimation of parameters in the process of reconstructing the signal.

The analysis of certain detection systems could be worked out in the framework of some less general analysis such as the estimation theory given by D. Slepian (Sl-1). Middleton includes this estimation theory claiming that a variety of analyses of optimum detection and extraction may properly be included within the unifying structure of decision theory.

A mathematical formulation of the noise spaces is needed to make practical application of statistical decision theory. An elementary discussion of random noise has been given by Rice (Ri-1) and a more general description of methods of solving noise problems has been given by Bennett (Be-1).

Some statistics on impulse noise have been obtained on an N-carrier telephone line. An empirical mathematical representation for some observed noise impulses are being calculated on the 650 for use in constructing a preliminary mathematical model of the impulse noise $\underline{N}$ and its probability distribution $W(\underline{N})$. The signal probability distribution $\sigma(\underline{S})$ is to be determined by the code used and the frequency of useage of the alphanumeric characters.

Statistical decision theory can be used in two directions:

(1) We may start from what can be assumed about the signals and noise and use statistical decision theory to determine an optimum

decision rule $\delta(\underline{\gamma} | \underline{V})$. Then we try to design a physical representation of the optimum decision rule in the form of filters and logic circuits. There is the possibility that when working in this direction, the physical realization may not be very practical.

(2) We may try different simple filter and logic systems by mathematically representing the detection system by decision rules. Then the appropriate integration formula from statistical decision theory can be used to compare error rates for different systems.

My plan is to alternate between both of the above approaches, weighing their relative values as the work progresses. The specific systems I am examining are described as follows:

a. Synchronous Wideband Detection with Correction Logic.

An analysis is planned for the synchronous wideband detection system being developed by H. Markey and E. Hopner.

This system is one in which the received dipulse signal is sampled at two signal null points and a straight line noise approximation signal is constructed through these points. Then this noise approximation signal is subtracted from the signal, and half of the resultant signal is rectified.

b. High Q Synchronous Detection

This system being developed by H. Markey and E. Hopner modulates a 12.5 kc carrier with a 400 microsecond dipulse to provide a supressed carrier double sideband modulation signal which is coupled to a high Q resonant circuit at the upper sideband frequency of 15 kc. At the end of each 400 microsecond gating period the resonant circuit is shorted. The line to the resonant-circuit is also connected to an amplifier in which the output is full wave-rectified. The output of the amplifier is also shorted at the end of each 400 microsecond gating period. The noise reduction feature of this system to be evaluated is the fraction of the noise spectrum that can pass through the narrow band pass of the resonant circuit.

Since some features of this detection system are similar to the Collins predicted wave radio telegraph system (Do-1) work is being done on the Collins system to validate the techniques of analysis.

c. Frequency Shift

It is proposed that the frequency-shift data transmission system for dial-up voice channels proposed by Dr. W. A. Christopherson and being considered by W. J. Johnson, Jr., be evaluated by deriving a mathematical decision rule $\delta(\gamma | v)$ corresponding to this physical system.

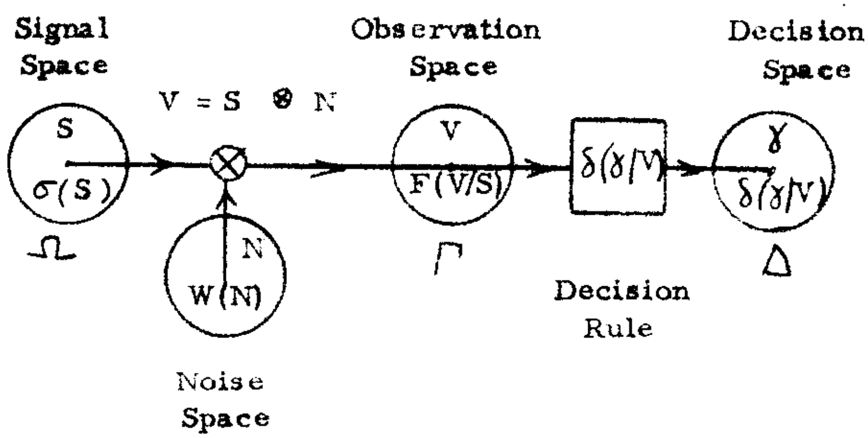


Fig. 1 -- Data Transmission Sub-Space
(from Middleton)

TABLE I -- Definitions for Figure 1

$\Omega$ = Signal Space: All possible signals: $\underline{S} = (S_1, S_2, \ldots S_n)$ with signal probability density $\sigma(\underline{S})$.

Noise Space: All possible noise signals: $\underline{N}$ with noise probability $W(\underline{N})$.

$\Gamma$ = Data Space: All possible observed signals (signal plus noise): $\underline{V} = (V_1, V_2, \ldots V_n)$, $F(\underline{V}/\underline{S})$ conditional probability density of $\underline{V}$, given $\underline{S}$.

$\Delta$ = Decision Space: All possible decisions: $\underline{\gamma} = (\gamma_1, \gamma_2 \ldots \gamma_m$
Decision rule: $\delta(\underline{\gamma}/\underline{V})$

3. Integration of Component Sections of a Business Data System.

It is proposed that this theory be used as a general framework to view the whole business data processing area as is illustrated in Figure 2. In Figure 2 the whole circle is defined as Data Systems Space A. Each component section of the system is shown as a sub-space in th form of an intermediate size circle within circle A in a manner similar to that shown in Figure (2.1) of ref. (Mi-1). The consequences may be a better recognition of fundamental limits as well as regions where improvement is most likely.

The comparison of error checking with block feedback signal with error correction is significant in sub-spaces D, E, G, and H. Some work has been done on this by John McLaughlin (SJA-42). A determination of the optimum block length of characters for telegraph transmission has been made at R. A. E. (AD-91919).

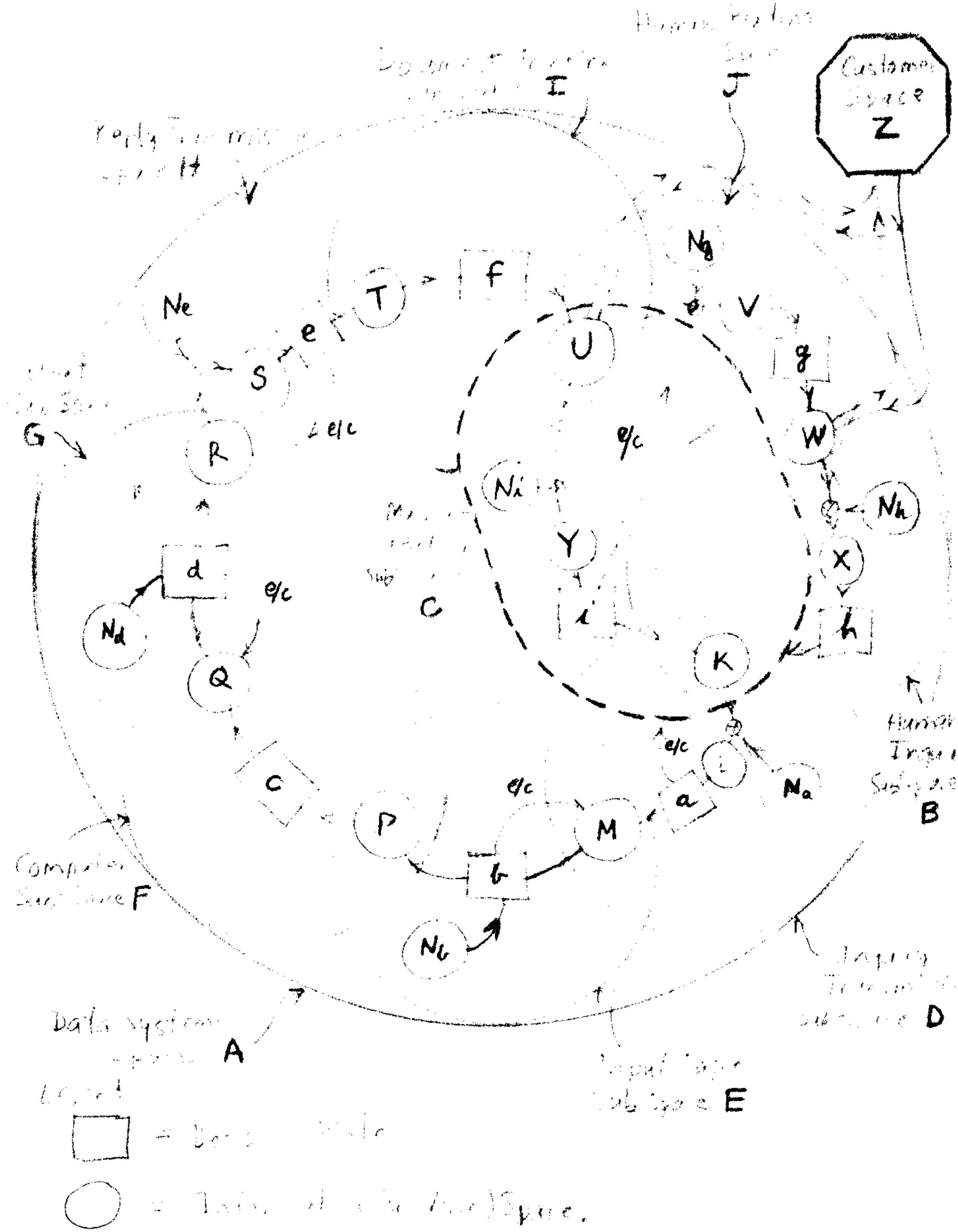


**Fig. 2 -- Organization of a Business Data System into Sub-Spaces.**

5-1-57 FBW   Rev 6-3-57

## TABLE II SUB-SPACES

B. Human Inquiry

C. Machine Reading

D. Inquiry Transmission

E. Input Tape (or buffer)

F. Computer

G. Output

H. Reply Transmission

I. Document Printing

J. Human Reading

## TABLE II INFORMATION, NOISE, AND DECISION SPACES

| | |
|---|---|
| e/c | Error Signal Channel (checking) |
| K | Keyboard Input Signal Space (D) and Decision Space (B or C). |
| $N_a$ | Transmission Noise Space |
| L | Transmission Observed Signal Space |
| a | Detection Decision Rules (D) |
| M | Detection Decision Space (D) and Signal Space (E). |
| $N_b$ | Translator Noise Space (E) |
| b | Translator from Data Transmission Code (D) to Computer Code (F). |
| P | Input Computer Tapes or Buffer Code Space (E and F). |
| c | Computer Logic (decision) Space (F). |
| Q | Output Computer Buffer Space (F) and Output Tape Signal Space (G). |
| $N_d$ | Output Translator Noise Space (G). |

| | |
|---|---|
| d | Output Translator Decision Rules (C). |
| R | Output Translator Decision Space (G) and Reply Signal Space (H). |
| $N_e$ | Reply Transmission Noise Space (H) |
| S | Observed Reply Signal Space (H) |
| e | Reply Demodulation Decision Rules (H) |
| T | Reply Decision Space (H) and Printer-Translator Input Signal Space (I). |
| f | Printer-Translator Decision Rules (I) |
| U | Printer Decision Space (i. e., output display) |
| $N_g$ | Human Reading Error Space (I) |
| V | Semantic Meaning Space (reply) |
| g | Human Action Decision Rules (J) |
| W | Human Decision Space (J) and Machine Reply Signal Space (B) |
| $N_h$ | Human Reading on Interpretation Error Space (B) |
| X | Semantic Meaning Space (inquiry) |
| h | Human Translation Rules to Machine Input Language (B) |
| K | Action Space (B) |
| Z | Customer Space (outside of data system A). |

The alternative sub-space C replaces sub-spaces J and B if the humans are replaced by a machine document reader.

| | |
|---|---|
| $N_i$ | Machine Noise Space (C) |
| Y | Machine Observed Signal Space (C) |
| i | Machine Logic Rules (C) |

The advantage of this generalized approach is that sub-spaces of the complete system can be analyzed and later be grouped together.

## 4. Problems of Applied Research and Specialization

The use of a common representation within the different sub-spaces of Figure 2 fulfills a need of coordinating the relationship between different fields of specialization and provid ing a bridge between theoretical (or basic) research and applied research. The problem of specialization and the organization of specialists in many fields to contribute collectively to industrial progress has been discussed by K. K. Paluev of General Electric Co. (Pa-1) and by G. M. K. Baker of R. C. A. Laboratories (Ba-1).

The problem of the engineer relating his work to the society of which he is a part has been discussed by John Mills of Bell Telephone Laboratories (Ml-1). The consideration of different concepts in succession of the relationship of the engineer between science and practical use of science in society by Mr. Mills indicates some unfulfilled need of understanding these relationships. In contrast F. B. Jewett and R. W. King of Bell Telephone Laboratories have been able to formulate and communicate their ideas on the relationship between engineering and the social order (Je-1, 2). Comparing the viewpoints of J. Mill and F. B. Jewett indicates a significant difference. Mr. Mill carried with him a dissatisfaction with the relationship between applied research and society, hoping that some new stage in the development of science would solve the problems. At retirement Mr. Mill wrote a book which expressed his feelings. It seems plausibl that the hostile nature of some of his comments may have come from saving some of his ideas to write on retirement, instead of communicating them promptly to people who could use the ideas. The paper by Jewett and King appears to be an example of a technique for assisting their concentration in their special fields through communicating general ideas relating to material outside their own special fields to people in the appropriate areas and to the scientific public through publication of the ideas not directly remunerative to their own organization.

Some years ago I discussed some of these problems of specialization versus developing a broad perspective in a seminar paper at the University of California (Wo-1). I described the organization of the paper as an attempt to relate the special field being considered to science, man and the universe in a way to give a perspective from which to approach a particular problem or special field. In the appendix of the above paper (Wo-1), I discussed the need for some engineers of a less-specialized but broader background to provide the liaison between the more specialized fields. Perhaps this function is being fulfilled by engineering managers to the extent needed in making policy decisions. This invalidates my earlier statement of the need for the training of engineers of a less-specialized but broader background, because a large part of the broader background is in practice developed on the job by engineers having managerial responsibility. This still leaves the need of the applied research specialist to develop a suitable perspective relating his own work to the basic sciences upon whi his work is based and the social uses to which his work will be applied.

Each time a body of knowledge and techniques of one field of science is found to be applicable to another field of science, an advance is made upon the problem specialization. The development of information theory (Me-1) is an example of a breakthrough which has broad effects in many fields and thus helps develop this perspective of the relationship of fields of specialization. The application of statistical decision theory to data transmission and business data systems would broaden the area of application of specialized techniques of mathematics and statistics.

## 5. Conclusions

The application of Middleton's formulation of statistical decision theory to the analysis of business data transmission systems provides a perspective from which to compare different data detection systems. Statistical decision theory does not automatically solve any of our problems. It helps in the comparison of different systems with each other and with theoretical optimum systems.

Study of statistical decision theory has resulted in the development of a system diagram in which each sub-section is to be considered as a data transmission system. This may assist in developing a perspective of all the links in a business data transmission system which would be of value to the systems group.

The exploitation of statistical decision theory in an applied research laboratory offers a potential gain through the establishment of a philosophical expression of the role of applied research in applying the techniques of basic science and mathematics to social use. The utilization of mathematical techniques having basic validity and application in related fields helps build up a perspective which releases energy for the more direct specialization on the practical problems.

## APPENDIX I -- SOME FORMULAS

Referring to Figure 1 the key feature of the decision problem is that a decision rule $\delta(\gamma/\underline{V})$ is used for deciding which symbol $\gamma$ corresponds to the received signal $\underline{V}$. Then some of the parameters for comparing different decision functions are:

Conditional Risk:

$$r(S, \delta) = \int_{\Gamma} dV \int_{\Delta} d\gamma \; \underbrace{C(S, \gamma)}_{\text{cost}} \; \overbrace{F(V/S)}^{\text{probability density}} \; \underbrace{\delta(\gamma/V)}_{\text{decision rule}} \tag{1}$$

Average Risk:

$$R(\sigma, \delta) = \int_{\Omega} r(S, \delta) \; \underbrace{\sigma(S)}_{\text{signal distribution}} \; dS \tag{2}$$

Conditional Information Loss:

$$h(S, \delta) = -\int_{\Gamma} dV \int_{\Delta} d\gamma \; \underbrace{\log P(S/\gamma)}_{\text{a-posteriori probability}} \; \delta(\gamma/V) \, F(V/S) \tag{3}$$

Average Information Loss: (Equivocation)

$$H(\sigma, \delta) = \int_{\Omega} h(S, \delta) \; \sigma(S) \; dS \tag{4}$$

Probability Density of Decision Error:

$$p(\gamma/S) = \int_{\Gamma} F(V/S) \; \delta(\gamma - \nu_{\sigma}(V)) \; dV) \tag{5}$$

The cost function $C(S, \gamma)$ allows for weighting the cost of errors in different symbols in accordance with their information content or their special value such as address symbols which if transmitted in error would loose a whole message. The integrals of Eq. (1) - (4) are in the form of a randomized decision function $\delta(\gamma/V)$ . The integrals can be used with non-randomized decision rules where $\delta$ is a Dirac $\delta$ - function.

$$\delta(\gamma/V) = \delta(\gamma - \nu_{\sigma}(V))$$

as in Eq. (5). Here $\nu_{\sigma}(V)$ expresses the functional operation performed on $\underline{V}$ by the detection system.

## APPENDIX II -- Potential Application to Other Data Transmission Systems.

There are other data transmission systems of interest to IBM but not presently included in the San Jose Communications Project for which a statistical decision analyses might be considered.

### Phase Modulation

An analysis of the phase-modulation system using delay lines proposed by F. Chiang could be analyzed to determine the potential error rate. The focal point of the analysis would be whether the designed instability needed for synchronizing would make the system more susceptible to impulse noise.

### Multilevel Detection

Possibly the results of S. Watanabe on multilevel detection (Wa-1) could be examined for possible transposition into Middleton's theoretical framework. This would enable different systems to be compared on a similar basis.

# APPENDIX III -- Potential Application to Other Sub-Systems

Examination of Figure 2 leads to the potential application of statistical decision theory to other sections of a business data system.

Human Reading (Sub-Space J)

Examining the Sub-Space Diagram in Figure 2, it can be seen that Decision Rules g in sub-space J are at present not readily susceptible to theoretical treatment. However, some human engineering work has been done in the human recognition of degraded printing (AD-61548). A proposal has been made for experimental work in this area. (SJC-6).

Magnetic Tape Recording (Sub-Spaces E and G).

An initial approach could be to put the model of magnetic tape error checking developed by Scharzoff and Harding into a form consistent with Middleton's formulation. (Sc-1).

Machine Reading of Characters (Sub-spaces C)

Statistical decision theory is potentially applicable to character recognition in the machine reading of documents.

## REFERENCES


SJA-35 A Proposal for a Computer Output Display Device, Dec. 4, 1956.

SJA-42 (See Mc-1).

SJC-6 A Preliminary Proposal for a Research Project on Facsimile Character Perception Error Rate as a Function of Signal Generation Method, Matrix Resolution, Transmission Line S/N (Random and Impulse), Type Size and Other Conditions. April 10, 1957

SJC-9 Transmission Line and Noise Analysis Computations. April 29, 1957.

AD-61548 Hogan Laboratories, Report No. 1176-FR-10. "Final Engineering Report on Informax" 30 Nov. 1953. AF 33(038) 17923. Also Report 1176-FR-12, 15 Feb. 1955. (AD-61548) (Includes work of Tufts College).

AD-91919 A. Cohen, "An Investigation into the Transmission of Data over a Teleprinter Link" Royal Aircraft Establishment, Tech. Note T.D. 2, Jan. 1956.

Ba-1 George M. K. Baker, "The Dilemma of Specialization" Proc. IRE p. 1195, Oct. 1948.

Be-1 W. R. Bennett, "Methods of Solving Noise Problems" Proc. IRE, 44, pp 609-638. (May 1956).

Do-1 Doelz & Heald, "A Predicted Wave Radio Teletype System" Conv. Rec. IRE, Part 8, Vol. 2 ( 1954 ) pp 63-67 (Same as Collins Radio Co. Report No. CTR-114).

Je-1 Frank B. Jewett and Robert W. King, "Engineering Progress and the Social Order," Science 92, 365 (1940).

Je-2 "A Way of Total Peace," (editorial) Jour. Appl. Physics 11, 749 (Dec. 1940).

Mc-1 (SJA-42) J. A. McLaughlin, "Comments on the Problem of Coupling Remote Inquiry Stations to a Central Computer." (Includes a comparison of error checking and error correction.)

Me-1 Richard L. Meier, "Communication and Social Change" Behavioral Science 1, p. 43, Jan. 1956.

Mi-1 David Middleton and David Van Meter, "Detection and Extraction of Signals in Noise from the Point of View of Statistical Decision Theory. I" Journal SIAM 3, pp 192-253 (Dec. 1955).

Mi-2 David Middleton and David Van Meter, "Detection and Extraction of Signals in Noise from the Point of View of Statistical Decision Theory. II" Journal SIAM, 4, pp 86-119 (June 1956).

Mi-3 David Middleton, "Statistical Theory of Signal Detection" Trans. IRE, PGIT-3, (March 1954) pp 26-51.

Ml-1 John Mills, The Engineer in Society, N. Y.: Van Nostrand (1946).

Pa-1 K. K, Paluev "How Collective Genius Contributes to Industrial Progress," G. E. Rev. pp 254-261, May 1941.

Ri-1 S. O. Rice, "A First Look at Random Noise," Comm. & Electronic (AIEE Paper 56-73) pp 128-131 (May 1956).

Sc-1 M. Schatzoff and W. B. Harding. "A Mathematical Model for Determining the Probabilities of Undetected Errors in Magnetic Tape Systems" IBM Journal 1, pp 177-180. (April 1957).

Sl-1 David Slepian, "Estimation of Signal Parameter in the Presence of Noise," Trans. IRE, PGIT-3 March 1954, p. 68.

Wa-1 Satosi Wantanbe, "A Theory of Multi-Level Information Channel With Gaussian Noise," IBM IR-00078, Nov. 26, 1956.

Wd-1 A. Wald, Sequential Analysis, Wiley (1947).

Wd-2 A Wald, Statistical Decision Functions, Wiley (1950).

Wo-1 F. B. Wood, "The History of Electro-magnetic Theory." Univ. Calif. EE 298 Seminar Paper, Jan. 10, 1947, pp 1-8, 42-47.