# XI. PROCESSING AND TRANSMISSION OF INFORMATION*


Prof. D. N. Arden
Prof. E. Arthurs
Prof. W. B. Davenport, Jr.
Prof. P. Elias
Prof. R. M. Fano
Prof. R. G. Gallager
Prof. F. C. Hennie III
Prof. E. M. Hofstetter
Prof. R. A. Howard
Prof. D. A. Huffman
Prof. I. M. Jacobs
Prof. W. F. Schreiber
Prof. C. E. Shannon
Prof. J. M. Wozencraft
G. Cheadle
J. E. Cunningham
E. F. Ferretti
D. H. Friedman
T. J. Goblick, Jr.
U. F. Gronemann
T. S. Huang
F. Jelinek
R. S. Kennedy
L. Kleinrock
C. L. Liu
H. H. Loomis, Jr.
J. L. Massey
C. Moraga
E. L. O'Neill
J. W. Pan
R. R. Pfeiffer
L. G. Roberts
J. R. Sklar
I. G. Stiglitz
I. E. Sutherland
O. J. Tretiak


## A. CODE CONSTRUCTION FOR TWO-WAY CHANNELS

### 1. Introduction

A discrete memoryless two-way channel,[1-3] shown schematically in Fig. XI-1, consists of two terminals, each equipped with a transmitter and a receiver. The left-hand terminal transmits signals $x$ from an alphabet of size $g$ and receives signals $y$ from an alphabet of size $h$. The right-hand terminal transmits signals $\bar{x}$ from an alphabet of size $\bar{g}$ and receives signals $\bar{y}$ from an alphabet of size $\bar{h}$. The channel operates synchronously: at given time intervals signals $x$ and $\bar{x}$ are simultaneously transmitted and as a result signals $y$ and $\bar{y}$ are received. The transmission through the channel is defined by the probability set $\{P(y,\bar{y}/x,\bar{x})\}$. The channel is memoryless so that

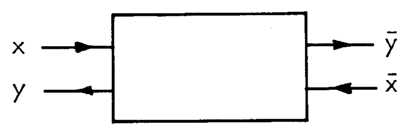


Fig. XI-1. Two-way channel.

$$\Pr(y_1,\ldots,y_n;\bar{y}_1,\ldots,\bar{y}_n/x_1,\ldots,x_n;\bar{x}_1,\ldots,\bar{x}_n) = \prod_{i=1}^{n} P(y_i,\bar{y}_i/x_i,\bar{x}_i), \tag{1}$$

where $a_i$ $(a=y,\bar{y},x,\bar{x})$ is a signal appearing at the terminals of the channel at time $i$.

If the channel is the only available medium through which the two terminals can communicate, then the most general signal source-channel arrangement is the one shown in Fig. XI-2. Depending on past received and transmitted signal sequences $(y_{i-1},\ldots,y_1)$ and $(x_{i-1},\ldots,x_1)$ (or $(\bar{y}_{i-1},\ldots,\bar{y}_1)$ and $(\bar{x}_{i-1},\ldots,\bar{x}_1)$), the stochastic sources select signals $x_i$ (or $\bar{x}_i$) for transmission. Suppose that we impose the reasonable condition that the sources be stationary and have finite memory. Let the operation


*This research was supported in part by Purchase Order DDL B-00368 with Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with the joint support of the U.S. Army, Navy, and Air Force under Air Force Contract AF19(604)-7400; and in part by the National Institutes of Health (Grant MH-04737-02).

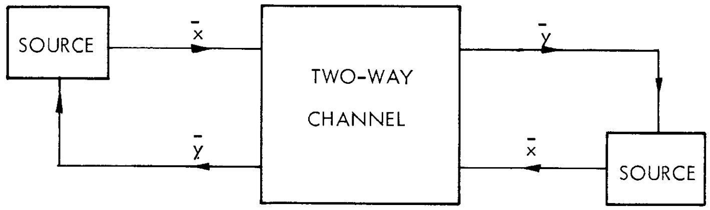


Fig. XI-2. Signal source-channel communication network.

of the left-hand source be defined by the probability set $\left\{q_{X^m Y^m}(x)\right\}$ and that of the right-hand source by the set $\left\{\bar{q}_{\bar{X}^m \bar{Y}^m}(\bar{x})\right\}$, in which we have adopted the capital-letter notation:

$$Z^m = z_{-1}, z_{-2}, \ldots, z_{-m} \qquad z = (x, y, \bar{x}, \bar{y}), \tag{2}$$

and we define

$$q_{X^m Y^m}(x) \equiv \Pr\left(x/(x_{-1}, \ldots, x_{-m})=X^m, (y_{-1}, \ldots, y_{-m})=Y^m\right) \tag{3a}$$

$$\bar{q}_{\bar{X}^m \bar{Y}^m}(x) \equiv \Pr\left(\bar{x}/(\bar{x}_{-1}, \ldots, \bar{x}_{-m})=\bar{X}^m, (\bar{y}_{-1}, \ldots, \bar{y}_{-m})=\bar{Y}^m\right). \tag{3b}$$

In (3a) and (3b) m is the assumed memory length and is common to both sources. Denote the information transmission rate in the left-to-right direction in Fig. XI-2 by R, and that in the right-to-left direction by $\bar{R}$. Consider now the problem of transmitting through the channel at the rates $(R, \bar{R})$ information generated by two completely independent stationary stochastic sources, one located at each terminal. Knowledge of information theory would immediately suggest that the independent sources be encoded so that the resulting transmitted signal sequences would have the statistical properties of the sources in Fig. XI-2. That is, considering without any loss of generality the left terminal, one would attempt to attach to the message source a transducer (see Fig. XI-3) whose output signals x would then be statistically describable by the set $\left\{q_{X^m Y^m}(x)\right\}$ if the message-source statistics were first properly adjusted. In this report a theorem will be proved which shows that for any given set $\left\{q_{X^m Y^m}(x)\right\}$ a transducer exists and an appropriate adjustment of source statistics can be found; in fact, a construction procedure will be developed.

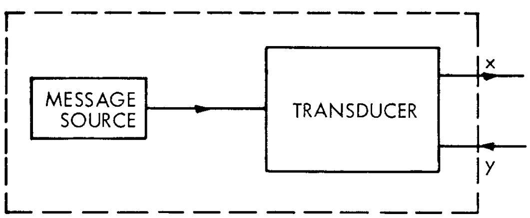


Fig. XI-3. Encoding of independent messages into signals.

2. The Message Source Alphabet and the Transducer

We shall define functions f of "memory" length m mapping the space of sequence pairs $X^m, Y^m$ on the space of channel input signals x:

$$f(X^m, Y^m) = x. \tag{4}$$

Similar functions are defined for the right-hand terminal.

It is clear from (4) that the domain of definition of any function f consists of $(gh)^m$ points $(X^m, Y^m)$, where g is the x-alphabet size, and h the y-alphabet size. Hence a function f is fully defined if a table of values for its $(gh)^m$ different possible arguments is given. Thus any function f can be represented by a g-nary sequence of $(gh)^m$ elements, each corresponding to a different point in the domain of definition. We can write

$$f \equiv \left(c_0, c_1, \ldots, c_i, \ldots, c_{(gh)^m-1}\right) \tag{5}$$

where $c_i$ is the value of f when the sequence $X^mY^m$ constituting its argument is the g-nary representation of the integer i. It is then clear from (5) that there are altogether $g^{(gh)^m}$ different possible functions f.

Consider next the transducer of Fig. XI-4 whose outputs could constitute the signal inputs to a two-way channel. The transducer x is a device consisting of a top and bottom shift register of m stages, and a middle shift register of one stage. At time i

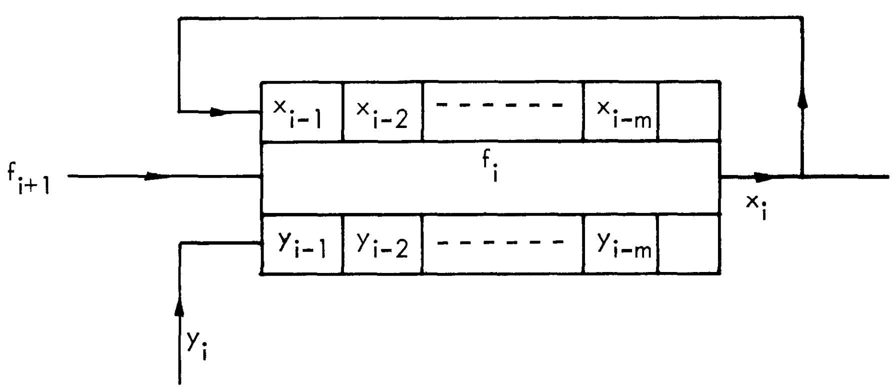


Fig. XI-4. Function-signal transducer.

the state of the transducer determines the output $x_i$ and is itself determined by the contents $x_{i-1}, \ldots, x_{i-m}$ of the top register, $y_{i-1}, \ldots, y_{i-m}$ of the bottom register, and $f_i$ of the middle register. All of the register contents are then shifted to the right by one state, the contents of the rightmost stages (that is, $x_{i-m}, f_i, y_{i-m}$) being eliminated, and the leftmost stages are filled from top to bottom by the symbols $x_i$, $f_{i+1}$, $y_i$. The state at time i and the output signal $x_i$ are now determined, and a new cycle can begin. The number of stages m corresponds to the memory length of the symbols f.

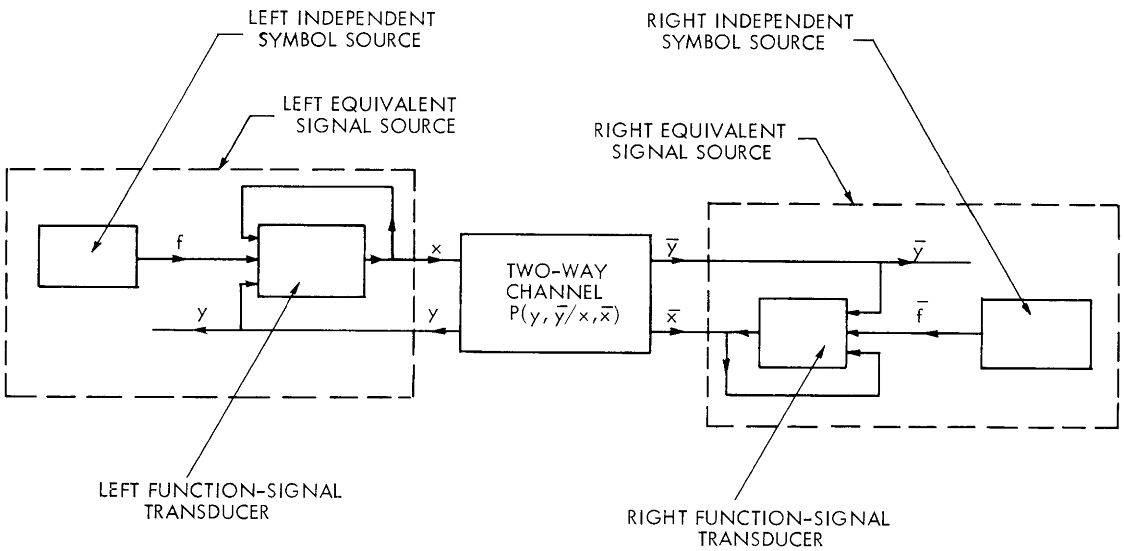


Fig. XI-5. Source-transducer-channel communication network.

We may now take a pair of such transducers, attach them to the two-way channel and connect their inputs to stationary stochastic sources generating symbols f and $\bar{f}$ synchronously at given time intervals. Such a source-transducer two-way channel communication network is schematically represented in Fig. XI-5. The source-transducer combinations have the effect of the signal sources in Fig. XI-2. In fact, it can be shown[4] that if the sources are independent of each other and generate successive symbols f and $\bar{f}$ independently with arbitrary probabilities P(f) and $\bar{P}(\bar{f})$, respectively, then the equivalent signal sources resulting from the source-transducer combinations generate signals with probabilities $q_{X^mY^m}(x)$ and $\bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})$, respectively, which can be computed from the arbitrary probabilities by use of the expressions

$$
\begin{aligned}
q_{X^mY^m}(x) &= \sum_{\forall f \ni f(X^m, Y^m)=x} P(f) \\
\bar{q}_{\bar{X}^m\bar{Y}^m} &= \sum_{\forall \bar{f} \ni \bar{f}(\bar{X}^m, \bar{Y}^m)=\bar{x}} \bar{P}(\bar{f}).
\end{aligned}
\tag{6}
$$

It may be of interest to note that both the summations involve $g^{(gh)^m-1}$ terms.

The question now is whether, given arbitrary sets of signal source probabilities $\left\{q_{X^mY^m}(x)\right\}$, $\left\{\bar{q}_{\bar{X}^m\bar{Y}^m}(\bar{x})\right\}$, probability assignments P ( ) and $\bar{\text{P}}$ ( ) to the sources of Fig. XI-5 can be found which would satisfy expressions (6). Theorem 1 shows that this is indeed possible and asserts, moreover, that in every case the transducer inputs can consist of a total of only $(g-1)(gh)^m + 1$ symbols f instead of the full alphabet of $g^{(gh)^m}$ symbols. In fact, a procedure for determination of the probabilities P ( ) and $\bar{\bar{\text{P}}}$ ( ) is given in Theorem 1, and then applied in the example of section 4.

## 3. Determination of Message Probabilities P(f) from Arbitrary $\left\{q_{X^mY^m}(x)\right\}$ Sets

THEOREM 1: Let g be the size of the input signal alphabet x, and h the size of the output alphabet y. The complete set $\left\{q_{X^mY^m}(x)\right\}$ of signal source probabilities is of size $g(gh)^m$. The complete alphabet of inputs f to the corresponding transducer of Fig. XI-4 is of size $g^{(gh)^m}$. Then $(g-1)(gh)^m + 1$ is the maximum number of nonzero probabilities P(f) needed to satisfy the equations

$$q_{X^mY^m}(x) = \sum_{\forall f \ni f(X^m,\, Y^m)=X} P(f) \tag{7}$$

for all $X^m, Y^m, x$ for any arbitrary set $\left\{q_{X^mY^m}(x)\right\}$ whose elements have the property that

$$\left.\begin{aligned} &\sum_{x=0}^{g-1} q_{X^mY^m}(x) = 1 \\ &q_{X^mY^m}(x) \geq 0 \end{aligned}\right\} \quad \text{for all } X^m, Y^m, x. \tag{8}$$

PROOF 1: We shall prove the theorem by demonstrating a direct procedure by which to construct the probability set $\{P(f)\}$ that would satisfy Eqs. 7 for a given set $\left\{q_{X^mY^m}(x)\right\}$. We must start by giving some terminology. Denote particular instances of the symbols x and y by a and b, respectively. Denote particular instances of the sequences $X^m$ and $Y^m$ by $A^m$ and $B^m$, respectively. Denote the probability set $\left\{q_{X^mY^m}(x)\right\}$ by $Q^o$.

Define

$$S^o(x, X^m, Y^m), \text{ the set of all symbols } f \text{ such that } f(X^m, Y^m) = x \tag{9a}$$

$$\mathscr{S}^o, \text{ the set of all sets } S^o(x, X^m, Y^m). \tag{9b}$$

It should be noted that there are as many sets $S^o$ as there are different sequences $x, X^mY^m$, that is, $g(gh)^m$. Hence $\mathscr{S}^o$ is of the same order as $Q^o$. The different sets $S^o$ overlap. The order of any $S^o(x, X^m, Y^m)$ is $g^{(gh)^m -1}$. We are now ready to specify the construction procedure of the set $\{P(f)\}$.

Step 1. Construct the set $Q^1$ consisting of all nonzero elements of $Q^o$. Denote the elements of $Q^1$ by

$$\left(q_{X^mY^m}(x)\right)^1 = q_{X^mY^m}(x) \neq 0 . \tag{10}$$

Let $\mathscr{T}^o$ be the set of all sets $S^o(x, X^m, Y^m)$ which are such that $q_{X^mY^m}(x) = 0$. From the sets in $\mathscr{S}^o$ construct new sets

$$S_1\left(x_0, X_0^m, Y_0^m\right) = S\left(x_0, X_0^m, Y_0^m\right) - S\left(x_0, X_0^m, Y_0^m\right) \cap \left[\bigcup_{\mathscr{T}^o} S(x, X^m, Y^m)\right]. \tag{11}$$

Let the collection $\mathscr{S}^1 = \{S^1(x, X^m, Y^m)\}$ consist of those sets $S_1(x, X^m, Y^m)$ that are nonempty. It will be shown that the orders of sets $\mathscr{S}^1$ and $Q^1$ are equal. Note also that

$$\sum_x \left(q_{X^m, Y^m}(x)\right)^1 = 1 \qquad \text{for all sequences } X^m, Y^m . \tag{12}$$

Step 2. In the set $Q^1$ find the not necessarily unique element of minimal magnitude. Let it be $\left(q_{A_1^mB_1^m}(a_1)\right)^1 > 0$. Now pick any element $f^1$ included in the set $S^1\left(a_1, A_1^m, B_1^m\right)$, and let

$$P(f^1) = \left(q_{A_1^mB_1^m}(a_1)\right)^1 . \tag{13}$$

It will be shown that $S^1\left(a_1, A_1^m, B_1^m\right)$ is nonempty. From $Q^1$ construct the set $Q^2 = \left\{\left(q_{X^mY^m}(x)\right)^2\right\}$ consisting of the nonzero elements $\left(q_{X^mY^m}(x)\right)_2$ defined by

$$\left(q_{X^mY^m}(x)\right)_2 = \begin{cases} \left(q_{X^mY^m}(x)\right)^1 & \text{if } x \neq f^1(X^m, Y^m) \\ \left(q_{X^mY^m}(x)\right)^1 - P(f^1) & \text{if } x = f^1(X^m, Y^m) . \end{cases} \tag{14}$$

The totality of elements of $Q^2$ is found by letting the right-hand side in (14) run over all of the elements of $Q^1$. Because of (13) and (14), the order of $Q^2$ is at most one less than the order of $Q^1$. Because of the way that $\left(q_{A_1^mB_1^m}(a_1)\right)^1$ was selected and $P(f^1)$ found, all of the elements of $Q^2$ are positive. Also

$$\sum_x \left(q_{X^mY^m}(x)\right)^2 = 1 - P(f^1) \geq 0 \qquad \text{for all sequences } X^m, Y^m \tag{15}$$

which follows from (14) and from the fact that each $X^m, Y^m$ is mapped by $f^1$ into one and only one symbol $x$.

If $1 - P(f^1) = 0$, then $Q^2$ is empty and our procedure terminates. If $1 - P(f^1) > 0$, we let $\mathcal{T}^1$ be the set of all sets $S^1(x, X^m, Y^m)$ in $\mathcal{S}^1$ which are such that by the construction rule (14) the probability $\left(q_{X^mY^m}(x)\right)_2 = 0$. Then, from all of the sets in $\mathcal{S}^1$ we construct new sets

$$S_2\left(x_0, X_0^m, Y_0^m\right) = S^1\left(x_0, X_0^mY_0^m\right) - S^1\left(x_0, X_0^m, Y_0^m\right) \cap \left[\bigcup_{\mathcal{T}^1} S^1(x, X^m, Y^m)\right]. \tag{16}$$

Let the collection $\mathcal{S}^2 = \{S^2(x, X^m, Y^m)\}$ consist of those sets $S_2(x, X^m, Y^m)$ that are nonempty. It will be shown that the orders of the sets $\mathcal{S}^2$ and $Q^2$ are equal.

We are now ready to undertake Step 3. Since its outline follows exactly that of Step 2, we shall not describe it, but instead describe the general $k+1^{\text{th}}$ step.

Step k + 1 ($k=1, 2, \ldots, t$). We assume that the set $Q^k$ is nonempty, otherwise the construction procedure would have terminated at some previous step. In $Q^k$ find the [not necessarily unique] element of minimal magnitude. Let it be $\left(q_{A_k^mB_k^m}(a_k)\right)^k > 0$.

Now pick any element $f^k$ included in the (nonempty, as we shall show) set $S^k\left(a_k, A_k^m, B_k^m\right)$ and let

$$P(f^k) = \left(q_{A_k^mB_k^m}(a_k)\right)^k. \tag{17}$$

Using on the right-hand side of (18) all of the elements of $Q^k$ in turn, construct the set $Q^{k+1} = \left\{\left(q_{X^mY^m}(x)\right)^{k+1}\right\}$ consisting of the nonzero elements $\left(q_{X^mY^m}(x)\right)_{k+1}$ defined by

$$\left(q_{X^mY^m}(x)\right)_{k+1} = \begin{cases} \left(q_{X^mY^m}(x)\right)^k & \text{if } x \neq f^k(X^m, Y^m) \\ \left(q_{X^mY^m}(x)\right)^k - P(f^k) & \text{if } x = f^k(X^m, Y^m). \end{cases} \tag{18}$$

Because all of the elements of $Q^k$ are positive by assumption, and $\left(q_{A_k^mB_k^m}(a_k)\right)^k$ is the minimal element of $Q^k$, it follows that this set is either empty or contains only positive elements. From the selection rules (17) and (18) it also follows that the order of $Q^{k+1}$ is at most one less than the order of $Q^k$. Moreover,

$$\sum_{x}\left(q_{X^mY^m}(x)\right)^{k+1} = 1 - \sum_{i=1}^{k} P(f^i) \geq 0 \qquad \text{for all } X^mY^m, \tag{19}$$

which follows recursively from (15), (17), and (18), since each $X^m, Y^m$ is mapped by $f^k$ into one and only one symbol $x$.

Whenever there is equality on the right-hand side of (19), the set $Q^{k+1}$ is necessarily empty and the procedure terminates. Note that in such a case, because of (18) and (19), $Q^k$ must have consisted of exactly $(gh)^m$ elements of equal magnitude $1 - \sum_{i=1}^{k-1} P(f^i)$, one element $\left(q_{X^mY^m}(x)\right)^k$ for each of the different possible sequences $X^m, Y^m$.

If, on the other hand, $1 - \sum_{i=1}^{k} P(f^i) > 0$, we let $\mathcal{T}^k$ be the set of all sets $S^k(x, X^m, Y^m)$ in $\mathcal{S}^k$ which is such that from (18) $\left(q_{X^m, Y^m}(x)\right)_{k+1} = 0$. Then we construct from all of the sets in $\mathcal{S}^k$ new sets

$$S_{k+1}\left(x_0, X_0^m, Y_0^m\right) = S^k\left(x_0, X_0^m, Y_0^m\right) - S^k\left(x_0, X_0^m, Y_0^m\right) \cap \left[\bigcup_{\mathcal{T}^k} S^k(x, X^m, Y^m)\right]. \tag{20}$$

Finally, we let the collection $\mathcal{S}^{k+1} = \left\{S^{k+1}\left(x_0, X_0^m, Y_0^m\right)\right\}$ consist of all those sets $S_{k+1}\left(x_0, X_0^m, Y_0^m\right)$ that are nonempty. We shall show that the orders of the sets $\mathcal{S}^{k+1}$ and $Q^{k+1}$ are equal. We are then ready to undertake the next step.

If $Q^{t+1}$ is emnty, but $Q^t$ was not, then $\sum_{i=1}^{t} P(f^i) = 1$. Thus to all those f's not included in the selected symbol set $\{f^1, f^2, \dots, f^t\}$ we assign the probability $P(f) = 0$.

We shall now complete the proof of the theorem. We must show that the following lemmas about the properties of our procedure hold.

LEMMA 1: The outlined procedure assigned probabilities to symbols in such a way that

$$\sum_{\forall f} P(f) = 1. \tag{21}$$

LEMMA 2: For all k, if set $S^k\left(x_0, X_0^m, Y_0^m\right)$ exists, it contains all f that are such that simultaneously

$$f\left(X_0^m, Y_0^m\right) = x_0$$

and

$$\left(q_{X^mY^m}\left(f\left(X_m^m, Y_m^m\right)\right)\right)^k > 0. \tag{22}$$

LEMMA 3: For all k, $S^k\left(x_0, X_0^m, Y_0^m\right)$ exists if and only if $\left(q_{X_0^m, Y_0^m}(x_0)\right)^k > 0$. Hence the orders of the sets $\mathcal{S}^k$ and $Q^k$ are equal.

LEMMA 4: If $Q^{t+1}$ is the first empty set (that is, $Q^o$, $Q^1$, ..., $Q^t$ are nonempty), then $\mathcal{S}^t$ contains exactly $(gh)^m$ sets $S^t$ of order one, every $S^t$ containing the same element $f^t$. In fact, the set $\mathcal{S}^t$ consists of the totality of sets $S^t(f^t(X^m, Y^m), X^m, Y^m)$ for the different sequences $X^m, Y^m$.

LEMMA 5: The selected symbols $\{f^1, f^2, \ldots, f^t\}$ are all distinct.

LEMMA 6: The constructed set $\{P(f)\}$ is such that for all $X^m, Y^m, x$

$$q_{X^mY^m}(x) = \sum_{f \in S^o(x, X^m, Y^m)} P(f). \tag{23}$$

LEMMA 7: The inequality

$$t \leq (g-1)(gh)^m + 1, \tag{24}$$

where t is some non-negative integer, is satisfied.

We shall now prove the lemmas. (Note that the example of section 4 illustrates the construction procedure described above.)

PROOF OF LEMMA 1: We have already shown that

$$\sum_{i=1}^{t} P(f^i) = 1.$$

Because $P(f) = 0$ for all f's not included in the set $\{f^1, f^2, \ldots, f^t\}$, the desired result follows. Q.E.D.

PROOF OF LEMMA 2: By definition, the set $S^k\left(x_0, X_0^m, Y_0^m\right)$ exists and is included in $\mathcal{S}^k$ only if it is nonempty. The assertion of Lemma 2 is certainly true for $k = 0$ (by definition); we shall now prove that it is true for k under the assumption that it was true for $k - 1$.

$S^k\left(x_0, X_0^m, Y_0^m\right)$ can exist only if $S^{k-1}\left(x_0, X_0^m, Y_0^m\right)$ existed and was nonempty, since $S^k$ is constructed by use of Eq. 20.

Consider any f that is such that $f\left(X_0^m, Y_0^m\right) = x_0$, and such that $\left(q_{X^mY^m}(f(X^m, Y^m))\right)^k > 0$ for all $X^m, Y^m \neq X_0^m, Y_0^m$. By the induction assumption, such f belongs to the set $S^{k-1}\left(x_0, X_0^m, Y_0^m\right)$, since surely $\left(q_{X^m, Y^m}(f(X^m, Y^m))\right)^{k-1} > 0$. But f cannot belong to any of the sets $S^{k-1}(x, X^m, Y^m) \in \mathcal{J}^{k-1}$ unless $\left(q_{X_0^mY_0^m}(x_0)\right)^k = 0$. Hence f does not belong to $S^{k-1}\left(x_0, X_0^m, Y_0^m\right) \cap \left[\bigcup_{\mathcal{J}^{k-1}} S^{k-1}(x, X^m, Y^m)\right]$ and thus, from (20), $S_k\left(x_0, X_0^m, Y_0^m\right)$ will include f unless $q_{X_0^mY_0^m}(x_0) = 0$, in which case $S_k\left(x_0, X_0^m, Y_0^m\right)$ will be empty. Thus if $S^k\left(x_0, X_0^m, Y_0^m\right)$ exists, it will contain all f's which are such that $f\left(X_0^m, Y_0^m\right) = x_0$, and $\left(q_{X^mY^m}(f(X^m, Y^m))\right)^k > 0$. Q.E.D.

PROOF OF LEMMA 3: If $\left(q_{X_0^mY_0^m}(x_0)\right)^k > 0$, then $\sum_x \left(q_{X_0^mY_0^m}(x)\right)^k > 0$. But then it follows from (19) that $\sum_x \left(q_{X^mY^m}(x)\right)^k > 0$ for all $X^m, Y^m$. Thus for all $X^m, Y^m$ there will exist at least one $x^* = x^*(X^m, Y^m)$ which is such that $\left(q_{X^mY^m}(x^*)\right)^k > 0$. Consider the function f having the property that $f(X^m, Y^m) = x^*(X^m, Y^m)$ for all $X^m, Y^m \neq X_0^mY_0^m$, and $f\left(X_0^m, Y_0^m\right) = x_0$. Such a function exists and by Lemma 2 it is a member of the set $S^k\left(x_0, X_0^m, Y_0^m\right)$, which is thus nonempty. Q.E.D.

PROOF OF LEMMA 4: We have shown that the sets $S^t$ are nonempty and that the orders of $\mathcal{S}^t$ and $Q^t$ are equal. It then follows from the second paragraph preceding Eq. 20 that there are $(gh)^m$ sets $S^t$, and from Lemma 3 it follows that if $\left(q_{X_0^mY_0^m}(x_0)\right)^t > 0$, then $S^t\left(x_0X_0^mY_0^m\right)$ exists and is nonempty. If $Q^{t+1}$ is the first empty set, it was shown that for each $X^mY^m$ there will be one and only one element $\left(q_{X^mY^m}(x^*)\right)^t > 0$. Hence by Lemma 2 each of the existing sets $S^t(x, X^m, Y^m)$ will consist of one and the same element f satisfying the property $q_{X^mY^m}(f(X^m, Y^m)) > 0$ for each $X^m, Y^m$. Q.E.D.

PROOF OF LEMMA 5: By the construction procedure (18), we have $\left(q_{A_1^mB_1^m}(a_1)\right)_2 = \dots = \left(q_{A_k^mB_k^m}(a_k)\right)_k = 0$. Hence by Lemma 2 the set $S^k\left(a_k, A_k^mB_k^m\right)$ cannot contain any f which is such that either $f\left(A_1^m, B_1^m\right) = a_1$, or $f\left(A_2^m, B_2^m\right) = a_2$, or ..., or $f\left(A_{k-1}^m, B_{k-1}^m\right) = a_{k-1}$. Hence all of the intersections $S^k\left(a_k, A_k^m, B_k^m\right) \cap S^i\left(a_i, A_i^m, B_i^m\right)$, $i = 1, 2, \dots, k-1$ must be empty and thus $f^k$ cannot be identical with any of the elements of the set $\{f^1, f^2, \dots, f^{k-1}\}$.

PROOF OF LEMMA 6: Since the construction procedure does not stop until $Q^{t+1}$ is empty, then from (18) it follows that, given any $x, X^m, Y^m$, there will be an integer $k$ ($k \leq t+1$) such that

$$\left(q_{X^mY^m}(x)\right)_k = 0, \quad \left(q_{X^mY^m}(x)\right)^{k-i} > 0 \qquad i = 1, 2, \dots, K. \tag{25}$$

Consequently, it follows from Lemma 2 that $f^i(X^m, Y^m) \neq x$ for all $j \in (k+1, \dots, t)$. On the other hand, it follows from (18) that $f^k(X^m, Y^m) = x$, and possibly for some $i \in (1, 2, \dots, k-1)$ also $f^{k-i}(X^m, Y^m) = x$. But because of (25) we can write

$$q_{X^mY^m}(x) = \left[q_{X^mY^m}(x) - \left(q_{X^mY^m}(x)\right)^1\right] + \left[\left(q_{X^mY^m}(x)\right)^1 - \left(q_{X^mY^m}(x)\right)^2\right]$$

$$+ \dots + \left[\left(q_{X^mY^m}(x)\right)^{k-1} - \left(q_{X^mY^m}(x)\right)_k\right]. \tag{26}$$

From (10) it follows that the first term on the right-hand side of (26) is always zero, and from (18) that

$$\left(q_{X^mY^m}(x)\right)^i - \left(q_{X^mY^m}(x)\right)^{i+1} = \begin{cases} 0 & \text{if } f^i(X^m, Y^m) \neq x \\ P(f^{i+1}) & \text{if } f^i(X^m, Y^m) = x. \end{cases} \tag{27}$$

It then follows from (26), (27), and Lemmas 2 and 5 that

$$q_{X^mY^m}(x) = \sum_{f \in S(x, X^m, Y^m)} P(f).$$

Q.E.D.

PROOF OF LEMMA 7: It follows from (17) and (18) that the order of the set $Q^i$ is strictly larger than that of the set $Q^{i+1}$ for $i = 1, 2, \dots, t$. From (10) it follows that the order of the set $Q^1$ is at most $g(gh)^m$, and we have demonstrated that the order of the set $Q^t$ is $(gh)^m$. Hence $t - 1 \leq g(gh)^m - (gh)^m$ and the result (24) follows.

Q.E.D.

We have thus completed the proof of Theorem 1.

## 4. Example of a Construction of the Probabilities P(f)

We now apply the construction procedure described in section 3 to a simple set $\left\{q_{X^mY^m}(x)\right\}$.

Let $g = 3$, $h = 1$, and $m = 1$. Then $Q^o$ will have 9 elements, while the set $\{f\}$ will be of size 27.

Using the notation of (5), let f be represented by the sequence $f = (c_0, c_1, c_2)$, where $c_i = f(i)$.

We shall now list the set $Q^o$, together with its values ($q_0(0) = .3$, $q_0(1) = .7$, $q_0(2) = 0$, $q_1(0) = .5$, $q_1(\phi) = .2$, $q_1(2) = .3$, $q_2(0) = .4$, $q_2(1) = .2$, $q_2(2) = .4$).

Note that $\sum_i q_j(i) = 1$ for all $j = 0, 1, 2$.

Next, we shall tabulate the memberships of the sets $S^o(x, S^m, Y^m)$. We have

$$\left.\begin{aligned} S^o(0,0) &= \{(0, c_1, c_2)\} \\ S^o(1,0) &= \{(1, c_1, c_2)\} \\ S^o(2,0) &= \{(2, c_1, c_2)\} \end{aligned}\right\} \quad \text{over all } c_1,\ c_2 \in (0,1,2)$$

$$\left.\begin{array}{l} S^o(0,1) = \{(c_0, 0, c_2)\} \\ S^o(1,1) = \{(c_0, 1, c_2)\} \\ S^o(2,1) = \{(c_0, 2, c_2)\} \end{array}\right\} \qquad \text{over all } c_0,\ c_2 \in (0,1,2)$$

$$\left.\begin{array}{l} S^o(0,2) = \{(c_0, c_1, 0)\} \\ S^o(1,2) = \{(c_0, c_1, 1)\} \\ S^o(2,2) = \{(c_0, c_1, 2)\} \end{array}\right\} \qquad \text{over all } c_0,\ c_1 \in (0,1,2).$$

Above is the set $\mathcal{S}^o$.

Step 1. Minimal element of $Q^o$ is $q_0(2) = 0$.
We construct next the set $Q^1$. It is

$$Q^1 = \Big((q_0(0))^1 = .3,\ (q_0(1))^1 = .7,\ (q_1(0))^1 = .5,\ (q_1(1))^1 = .2,\ (q_1(2))^1 = .3,$$
$$(q_2(0))^1 = .4,\ (q_2(1))^1 = .2,\ (q_2(2))^1 = .4\Big),$$

and the set $\mathcal{S}^1$ consists of the following sets:

$$S^1(0,0) = S^o(0,0)$$
$$S^1(1,0) = S^o(1,0)$$

$$\left.\begin{array}{l} S^1(0,1) = \{(c_0, 0, c_2)\} \\ S^1(1,1) = \{(c_0, 1, c_2)\} \\ S^1(2,1) = \{(c_0, 2, c_2)\} \end{array}\right\} , \qquad \begin{array}{l} c_0 \in (0,1) \text{ only} \\ c_2 \in (0,1,2) \end{array}$$

$$\left.\begin{array}{l} S^1(0,2) = \{(c_0, c_1, 0)\} \\ S^1(1,2) = \{(c_0, c_1, 1)\} \\ S^1(2,2) = \{(c_0, c_1, 2)\} \end{array}\right\} \qquad \begin{array}{l} c_0 \in (0,1) \text{ only} \\ c_1 \in (0,1,2) \end{array}$$

Step 2. The minimal element of $Q^1$ is $(q_1(1))^1 = .2$. Hence we pick any element f from $S^1(1,1)$. Let $f^1 = (1,1,0)$. Also let $P(f^1) = .2$. The set $Q^2$ is

$$Q^2 = \Big((q_0(0))^2 = .3,\ (q_0(1))^2 = .5,\ (q_1(0))^2 = .5,\ (q_1(2))^2 = .3,\ (q_2(0))^2 = .2,$$
$$(q_2(1))^2 = .2,\ (q_2(2))^2 = .4\Big).$$

The set $\mathcal{S}^2$ consists of the following sets:

$$\left.\begin{array}{l} S^2(0,0) = \{(0, c_1, c_2)\} \\ S^2(1,0) = \{(1, c_0, c_2)\} \end{array}\right\} \qquad \begin{array}{l} c_1 \in (0,2) \\ c_2 \in (0,1,2) \end{array}$$

$$\left.\begin{array}{l} S^2(0,1) = \{(c_0, 0, c_2)\} \\ S^2(2,1) = \{(c_0, 2, c_2)\} \end{array}\right\} \qquad \begin{array}{l} c_0 \in (0,1) \\ c_2 \in (0,1,2) \end{array}$$

$$\left.\begin{array}{l} S^2(0,2) = \{(c_0, c_1, 0)\} \\ S^2(1,2) = \{(c_0, c_1, 1)\} \\ S^2(2,2) = \{(c_0, c_1, 2)\} \end{array}\right\} \qquad \begin{array}{l} c_0 \in (0,1) \\ c_1 \in (0,2) \end{array}$$

Step 3. The minimal element of $Q^2$ is $(q_2(0))^2 = .2$. Here we pick any element f from $S^2(0,2)$. Let $f^2 = (0,2,0)$. Also let $P(f^2) = .2$. The set $Q^3$ is

$$Q^3 = \left\{ (q_0(0))^3 = .1,\ (q_0(1))^3 = .5,\ (q_1(0))^3 = .5,\ (q_1(2))^3 = .1,\ (q_2(1))^3 = .2, \right.$$
$$\left. (q_2(2))^3 = .4 \right\}.$$

The set $\mathcal{S}^3$ consists of the following sets.

$$\left.\begin{array}{l} S^3(0,0) = \{(0, c_1, c_2)\} \\ S^3(1,0) = \{(1, c_1, c_2)\} \end{array}\right\} \qquad \begin{array}{l} c_1 \in (0,2) \\ c_2 \in (1,2) \end{array}$$

$$\left.\begin{array}{l} S^3(0,1) = \{(c_0, 0, c_2)\} \\ S^3(2,1) = \{(c_0, 2, c_2)\} \end{array}\right\} \qquad \begin{array}{l} c_0 \in (0,1) \\ c_2 \in (1,2) \end{array}$$

$$\left.\begin{array}{l} S^3(1,2) = \{(c_0, c_1, 1)\} \\ S^3(2,2) = \{(c_0, c_1, 2)\} \end{array}\right\} \qquad \begin{array}{l} c_0 \in (0,1) \\ c_1 \in (0,2). \end{array}$$

Step 4. The minimal element of $Q^3$ is $(q_0(0))^3 = .1$. Hence we pick any element f from $S^3(0,0)$. Let $f^3 = (0,2,1)$. Also let $P(f^3) = .1$. We construct $Q^4$:

$$Q^4 = \left\{ (q_0(1))^4 = .5,\ (q_1(0))^4 = .5,\ (q_2(1))^4 = .1,\ (q_2(2))^4 = .4 \right\}.$$

The set $\mathcal{S}^4$ consists of the following sets.

$$S^4(1,0) = \{(1,0,c_2)\} \qquad c_2 \in (1,2)$$
$$S^4(0,1) = \{(1,0,c_2)\} \qquad c_2 \in (1,2)$$
$$S^4(1,2) = \{(1,0,1)\}$$
$$S^4(2,2) = \{(1,0,2)\}.$$

Step 5. The minimal element of $Q^4$ is $(q_2(1))^4 = .1$. Hence we pick the unique element of $S^4(1,2)$, $f^4 = (1,0,1)$. Let $P(f^4) = .1$.

We construct $Q^5$ as follows.

$$Q^5 = \left\{(q_0(1))^5 = .4,\ (q_1(0))^5 = .4,\ (q_2(2))^5 = .4\right\}.$$

The set $\mathcal{S}^5$ consists of the following sets:

$S^5(1,0) = \{(1,0,2)\}$

$S^5(0,1) = \{(1,0,2)\}$

$S^5(2,2) = \{(1,0,2)\}$.

Step 6. The minimal element of $Q^5$ is $(q_0(1))^5 = .4$. Hence we pick the unique element of $S^5(1,0)$, $f^5 = (1,0,2)$ and let $P(f^5) = .4$.

This ends the construction.

We have obtained the probability assignments:

$f^1 = (1,1,0) \qquad P(f^1) = .2$

$f^2 = (0,2,0) \qquad P(f^2) = .2$

$f^3 = (0,2,1) \qquad P(f^3) = .1$

$f^4 = (1,0,1) \qquad P(f^4) = .1$

$f^5 = (1,0,2) \qquad P(f^5) = .4$.

$P(f) = 0$ for all $f$ not included in $\{f^1, f^2, \ldots, f^5\}$. It is easily seen that $\sum_{i=1}^{5} P(f^i) = 1$. Also $t = 5 \leq (g-1)(gh)^m + 1 = 2.3 + 1 = 7$. Moreover,

$$q_0(0) = P(f^2) + P(f^3) = .3$$

$$q_0(1) = P(f^1) + P(f^4) + P(f^5) = .7$$

$$q_0(2) = 0$$

$$q_1(0) = P(f^4) + P(f^5) = .5$$

$$q_1(1) = P(f^1) = .2$$

$$q_1(2) = P(f^2) + P(f^3) = .3$$

$$q_2(0) = P(f^1) + P(f^2) = .4$$

$$q_2(1) = P(f^3) + P(f^4) = .2$$

$$q_2(2) = P(f^5) = .4 .$$

Thus we have completed and shown the correctness of the present construction.

F. Jelinek

## References

1. F. Jelinek, Coding for Discrete Memoryless Two-way Channels, Ph.D. Thesis, Department of Electrical Engineering, M.I.T., January 1962.

2. C. E. Shannon, Two-way communication channels, Fourth Berkeley Symposium on Mathematical Statistics and Probability, edited by J. Neyman (University of California Press, Berkeley, 1961), Vol. I, pp. 611-644.

3. F. Jelinek, Coding for discrete memoryless two-way channels, Quarterly Progress Report No. 66, Research Laboratory of Electronics, M.I.T., July 15, 1962, pp. 239-253.

4. F. Jelinek, Ph.D. Thesis, op. cit., Section 7.1, pp. 81-84.