MPROVEMENT OF MARKOS MODEL TOWARD THE INTER-LOCUTIONAL COMMUNICATION THEORY

Hideo Seki

University of Electro-Communications
Tokyo,Japan

Abstract

A new model of a communication system was proposed which takes advantage of both Shannon's as well as Marko's models. Such a model was realized by turning our attention to the point that a two-way communication channel may be understood as a unidirectional channel changing its direction alternatively so long as short time duration is concerned. In contradiction to Marko's view that the Shannon's theory is a special case of Marko's theory, the author takes a view that both theories must be included in his new single model.

Communication between three information nodes was also discussed as an example of multiple inter-locutional communications. The full paper will be divided into following parts.

1. Introduction
2. Shannon's Model
3. Marko's Model
4. Improved Model
5. Multiple Inter-Locutional Communications
6. Conclusion

1. Introduction

Shannon's model of a communication channel is useful for the efficient and reliable transmission of information, while that of Marko may be more appropriate for treating the case of actual communication between two living creatures or between one creature and his surroundings, as Marko insisted. The author also agrees with the latter in the point that such a new model should be promising for developing an inter-locutional communication theory in the future. However I have a different opinion in some minor points, those which are the main parts of this paper. In short, the information flow between two nodes should be unidirectional during a small interval of time and it may be considered bidirectional only when the time intervals are sufficiently long. From this standpoint, I would like to Insist that the Shannon's model is not a special case of Marko's model and the new model should include both.

Before considering the new model, it will be helpful to review briefly both models.

2. Shannon[1] Model

IS,T,CH,R,D and N, shown in Fig.I, respectively express information source, transmitter, communication channel, receiver, destination and noise source.

Let $x_1, x_2, \ldots, x_k, \ldots x$ be representation elements(e.g. letters) at the transmitter side and the set of these elements be written as X. Similarly, Y is the set on the receiver side, including $y, y, \ldots, y_i, \ldots y$. Then the self entropies may be expressed as follows.

$$H(X) = -\sum_{X} P(x_k) \log P(x_k)$$
$$H(Y) = -\sum_{Y} P(y_i) \log P(y_i) \quad (1)$$

And the mutual entropy can be defined by

$$I(X;Y) = \sum_{X,Y} P(x_k, y_i) \log \frac{P(x_k, y_i)}{P(x_k)P(y_i)}$$

$$= \sum_{X,Y} P(x_k, y_i) \log \frac{P(x_k|y_i)}{P(x_k)} \quad (2)$$

where P(x,) or P(y.) is the probability of occurrence. $P(x_k, y_i)$ and $P(x_k | y)$ are joint and conditional probabilities respectively.

As is well known, the mutual entropy is used for the estimation of channel capacity and is related to other entropies as follows.

$$I(X;Y) - H(X) - H(XY) - H(Y) - H(Y|X)$$
$$= H(X) + H(Y) - H(XY)$$

where (3)

$$H(X|Y) = -\sum_{XY} P(x_k, y_i) \log P(x_k|y_i)$$

(4)

$$H(Y|X) = -\sum_{XY} P(x_k, y_i) \log P(y_i|x_k)$$

(5)

$$H(XY) = -\sum_{XY} P(x_k, y_i) \log P(x_k, y_i)$$

(6)

The conditional entropy H(X|Y) is called equivocation, while H(Y|X) is channel noise entropy. According to the notation of Marko, the latter is written as D_{12} or D_{21} and called by him as discrepancy. The net information at the receiver side is, of course, the difference between the tranmitted information and equivocation. This difference is I(X;Y) itself.

3. Marko's model

The'notations in Fig.2 are slightly modified from those shown in Marko's model in the original paper so as to facilitate the reader's comparison with Shannon's model, i.e. H(X) and

H(X|Y) are used here instead of H_1 and D_{21} respectively. However, 1(X;Y) cannot be used for T_{21} , because of the unsymmetric character of T and T . In this case, I(X<-Y) was used for T_{21} while I(X->Y) for T_{12}.
Note that represents the set of transmitting elements in I(X->Y), while the same letter X denotes the set of received elements in I(X<-Y) As will be shown later, symbol I(X;Y) may be used for the sum equation (15).

In this model, X is an information source as well as a destination and Y has also two meanings. For this reason, X and Y(3) can conveniently be called Information nodes

Let the message composed of the message elements $x_1, x_2, \dots x_k \dots x_n$ in node X be

$\alpha_1 \alpha_2 \cdots \alpha_n$ x

and that composed of $y_1\ y_2 \dots y_i \dots y_m$ in node Y be

$\beta_1 \beta_2 \cdots \beta_n$ y.

Then, according to Marko's theory, the message element x following α_n not only depends on the elements $\alpha_1\ \alpha_2 \ldots$ *out* also on $\beta_1\ \beta_2 \ldots \beta_n$. Let us represent this conditional probability on the side of node X by $P(x|x^n, y^n)$. Similarly, the conditional probability on the side of node Y can be defined by $P(y|y^n x^n)$. Consider also the joint conditional probability $P(x, y|x^n, y^n)$ which is the probability of joint occurrence x and y due to the preceding n events on the side of nodes X and Y respectively.

Then the total subjective entropy on X side can be defined by

$$H(X) = \lim_{n\to\infty} \left\{ - \sum_X P(x^n, x) \log P(x|x^n) \right\} \quad (7)$$

and that on Y side by

$$H(Y) = \lim_{n\to\infty} \left\{ - \sum_Y P(y^n, y) \log P(y|y^n) \right\} \quad (8)$$

On the other hand, the free subjective entropies on X and Y sides can respectively be defined by the follosing two equations.

$$H(X|XY) = \lim_{n\to\infty} \left\{ - \sum_{XY} P(x^n, y^n, x) \log P(x|x^n, y^n) \right\}$$

$$H(Y|YX) = \lim_{n\to\infty} \left\{ - \sum_{YX} P(y^n, x^n, y) \log P(y|y^n, x^n) \right\}$$

(9),(10)

The directive subjective mutual entropy through the left-going communication channel should be expressed as (11)

$$I(X\leftarrow Y) = \lim_{n\to\infty} \left\{ - \sum_{XY} P(x^n, y^n, x) \log \frac{P(x|x^n)}{P(x|x^n, y^n)} \right\}$$

and, therefore, the similar quantity through the right-going communication channel can analogously be written down by

$$I(X\to Y) = \lim_{n\to\infty} \left\{ - \sum_{XY} P(x^n, y^n, y) \log \frac{P(y|y^n)}{P(y|y^n, x^n)} \right\} \quad (12)$$

In contrast to the Shannon's model, where the equivocation was simply expressed by

H(X) - I(X;Y)

and the channel noise was

H(Y) - I(X;Y),

as in equation (3), the corresponding quantities should take the following forms in Marko's model.

$$\begin{aligned} &\text{Equiv.} &: \quad & H(X|Y) = H(X) - I(X\to Y) \\ &\text{CH - N} &: \quad & H(Y|X) = H(Y) - I(X\leftarrow Y) \end{aligned} \quad (13)$$

Note that the free subjective entropies seem similar as above but have basic differences as follows

$$\begin{aligned} &\text{on X side} &: \quad & H(X|XY) = H(X) - I(X\leftarrow Y) \\ &\text{on Y side} &: \quad & H(Y|YX) = H(Y) - I(X\to Y) \end{aligned} \quad (14)$$

The directive subjective mutual entropies defined by equations (11) and (12) are, of course, non-symmetric. But their sum becomes symmetric as follows

$$T(X;Y) = I(X\leftarrow Y) + I(X\to Y) \quad (15)$$

Thus this sum may be understood as the total mutual entropy and be written as I(X;Y), using the same notation as in the case of Shannon's model. Then the definition of total mutual entropy may be written down as follows.

$$I(X;Y) = \lim_{n\to\infty} \left\{ \sum_{XY} P(x^n, x, y^n, y) \log \frac{P(x, y|x^n, y^n)}{P(x|x^n) P(y|y^n)} \right\} \quad (16)$$

The following relations also are deducible from equation (14).

$$\begin{aligned} H(X) + H(Y) - I(X;Y) &= H(Y|X) + H(X|Y) \\ &= H(X|XY) + H(Y|YX) \end{aligned} \quad (17)$$

(14) also shows that

$H(X|XY) = 0$

if

$H(X) = I(X\leftarrow Y)$

Such condition was denominated by Marko as "suggestion", where the information received by node X from node Y shall be returned without any transformation. However, in the general case of

$$H(X) \neq I(X \cdot Y),$$

"stochastic degree of synchronization" or "degree of perception" was defined by him as follows.

$$\sigma_x' = \frac{I(X \cdot Y)}{H(X)} \; ; \; \sigma_y' = \frac{I(X \to Y)}{H(Y)} \tag{18}$$

In this case, if the condition

$$H(X|XY) = I(X \to Y)$$
$$H(Y|YX) = I(X \leftarrow Y) \tag{19}$$

is satisfied, the coupling between two nodes become maximum and the relation

$$\sigma_x' + \sigma_y' = 1 \tag{20}$$

results, as can be seen easily. But, in general,

$$\sigma_x' + \sigma_y' \leq 1 \tag{21}$$

results from equations (14) and (15).

Fig. 3 shows the possible values of the stochastic degree of synchronization under different conditions.

4. Improved Model

Marko's model seems perfectly general so as to be applied to the inter-]ocutional communications between the animals and machines. However, it seems to me that a fine-structural refinement should be necessary for further development of the bidirectional information theory. The reason is that the telephone communication between two persons, data communication between two computers and many other communications are unidirectional as far as a short interval of time is concerned. TASI system is an example that such substantial characteristics in speech were properly used. This means the model should be Shannon's, at least, within a short interval. Reversing of the direction of information flow occurs only when the storage and processing of the received messages are completed. This idea necessarily requires the model to be modified to a new model as Fig. 4 where the situation of information source and destination alternatively reverses from time to time. In Fig. 4, IN means information node which becomes information source or destination depending on the instant of situation.

Now, the model of Fig. 3 is a uni-directional communication channel and the equations from (1) to (6) hold exactly, so far as the change-over switch is set on the fixed side. As a consequence, $H(X|XY)$ in Marko's model becomes simply as $H(X)$ .or

$$H(X) = \lim_{n\to\infty}\left\{-\sum_X P(x^n, x)\log P(x|x^n)\right\} \tag{22}$$

and $H(Y|YX)$ also becomes $H(Y)$, or

$$H(Y) = \lim_{n\to\infty}\left\{-\sum_Y P(y^n, y)\log P(y|y^n)\right\} \tag{23}$$

Thus, all relations which were deduced in the unidirectional communication model may be aplied to the new model. In a rigorous mathematical sense, the notation $n\to\infty$ in (22) and (23) may not be true, because of the finite duration of time of the change-over switch position. But, it is not a serious problem in practice.

Let us consider next the case of two communication channels of opposite flows of information between two communication nodes. In such cases, the continuous flows of signals may occur in both directions, and the model cannot be considered to be unidirectional if even a very short interval Is taken so long as we pay an at Lent, ion to the middle point of communication route. Even this model also may be regarded as a unidirectional flow, at least, at both terminals of communication route where two change-over switches are inserted as shown in Fig. 5, which is the result of fine-structural consideration. Thus, considering the network inside the dotted square as a communication channel, you can see that the Fig. 5 exactly coincides with Fig. 4.

Now, anyway, the changing over of switches should occur always when, at least, a whole sentence is completed. In other words, the reversing of direction of an information flow will always occur when a sentence $\alpha_1 \alpha_2 \ldots \alpha_n$ or sentences composed from representation elements $x_1 x_2 \ldots$ is over. Let this sentence or series or sentences originated from node X be . Then, node Y will send back another sentence or series of sentences choosing from his repertoire, considering it as a most suitable one for a special purpose in question or for performing a special task. And let $P(\eta_1|\xi_1)$ be a probability that Y chooses when he receives ξ_1 from X. Then the probability that X sends out ξ_2 after receiving η_1 from Y, may be written as $P(\xi_2|\xi_1, \xi_1)$. Using the abbreviated symbols ξ^n instead of $\xi_1 \xi_2 \ldots \xi_n$ and η^n instead of $\eta_1 \eta_2 \ldots \eta_n$, equation (7), (8), (9) and (10) may be rewritten in the following new form as

$$H(X) = \lim_{n\to\infty}\left\{-\sum_X P(\xi^n, \xi)\log P(\xi|\xi^n)\right\} \tag{24}$$

$$H(Y) = \lim_{n\to\infty}\left\{-\sum_Y P(\eta^n, \eta)\log P(\eta|\eta^n)\right\} \tag{25}$$

$$H(X|XY) = \lim_{n\to\infty}\left\{-\sum_{XY} P(\xi^n, \eta^n, \xi)\log P(\xi|\xi^n, \eta^n)\right\} \tag{26}$$

$$H(Y|YX) = \lim_{n\to\infty}\left\{-\sum_{YX} P(\eta^n, \xi^n, \eta)\log P(\eta|\eta^n, \xi^n)\right\} \tag{27}$$

Thus you can see that the Marko s entropy equations become necessary at this level of sentences. Similarly, the equations from (11) to (17) may be rewritten in the same manner as above, but may be omitted here for the simplification.

In short, the improved model coincides with Shannon's model at letter (representation element) level and with Marko's model at sentence

(message) level. Representation elements have no meanings, while the messages have.

Communications having a fixed purpose among nodes might have considerable effects on their forthcoming behavior. If the relation between the messages and bahavior could be estimated quantitatively, a value anylysis of information may also be possible.

5. Multiple Inter-Locutional Communications

It seems to me that the stochastic degree of synchronization (18) proposed by Marko is a measure of persuasion. But the purpose of communication among multiple nodes may be flexible. For example, it may be consultaion, discussion, quarrel or any other type. In such cases, value of information or measure of learning may become necessary instead of the stochastic degree of synchronization.

The old proverb says, " two heads are better than one ". The similar proverb in Japan says, " three heads constitute Manjusuri ". This suggests that the value of information can be raised up by communications among multiple nodes. Although the quantitative analyses of such information processing are extremely difficult at this stage, still we can roughly analyses, in principle, according to the model of Fig. 6. Fig. 6 (a) is an abbreviated diagram of Fig. 4, at the same time inserting third node Z. (b), (c) or (d) of Fig. 6 show states, respectively, when X, Y or Z is the information source and the remaining two are destination.

Now, let the probabilities that Y and Z choose from their repertories after receiving the message ξ_1 from X in (b) be $P(\eta_1|\xi_1)$ and $P(\zeta_1|\xi_1)$, these two being generally different from each other. Changing to state (c) from (b), we can again assume the message choosing probability of X as $P(\xi_2|\xi_1,\eta_1)$ and that of Z as $P(\zeta_2|\xi_1,\eta_1,\zeta_1)$, while, at state (d), these probabilities become $P(\xi_3|\xi_1,\xi_2,\eta_1,\eta_2)$ at X and $P(\eta_2|\xi_1,\eta_1,\xi_2)$ at Y. Returning to the state (b), X will send message ξ_3, which excite Y and Z by the probabilities $P(\eta_3|\xi_1,\xi_3,\eta_1,\eta_2,\xi_2)$ and $P(\zeta_3|\xi_1,\xi_3,\eta_1,\zeta_1,\zeta_2)$. Note that the subscipt numbers are not necessarily in order.

After n times of trnsition, the probabilties

$P(\xi|\xi^{2n},\eta^n,\zeta^n)$ for node X

$P(\eta|\xi^n,\eta^{2n},\zeta^n)$ for node Y

and $P(\zeta|\xi^n,\eta^n,\zeta^{2n})$ for node Z

may be considered and finally expect that the probability of some special message becomes very large, or

$$P(\xi|\xi^{2n},\eta^n,\zeta^n) \approx P(\eta|\xi^n,\eta^{2n},\zeta^n) \approx P(\zeta|\xi^n,\eta^n,\zeta^{2n}) \to 1 \tag{28}$$

Fig. 7 shows this tendency. Before starting inter-locution, the figure showa that the probabilities P of choosing any message are uniformly distributed. Assuming the abscissa is the order number of probabilities arranged from higher to lower ones, the peak of the curve gradually rising up after repeating inter-locution and finally approaches unity.

In Snannon's mathemetical theory of communication, the average amount of information messages instead of special messages was a most important quantity. But, in our case, probabilities of same messages are rather important and this very quantity may be considered a measure of the value of information.

Another suitable.measure for the group as a whole is learning , i.e.

$$L = \log m - H \tag{29}$$

where m is total number of possible messages and H is an average amount of information per message, example of the form of estimation being

$$H = -\lim_{n\to\infty}\sum_{XYZ} P(\xi^{2n+1},\eta^n,\zeta^n)\log P(\xi|\xi^{2n},\eta^n,\zeta^n) \tag{30}$$

The number of nodes also can be extended to more than three. Christie and others have discussed the case of nodes 5 and got some experimental results under proper constraints.

A brain also can be taken as a system consistinfg of many localized information processors or nodes and communication channels between nodes. Thus even a brain ordinarily considered as one node in the case of inter-locution should be taken as an ensemble of nodes. If so, the above mentioned new model probably become one of the important suggestions for treating thinking processes of the brain.

6. Conclusion

In concluding my discussion about the new model for inter-locutional communication, It seems better to summarize some important points for further studies as follows:

a. Though the node In an inter-locutional communication net is able to be an information source or a destination, it must remain to be either of them within a short interval of time, i.e. the model must be uni-directional.

b. Statistical relations between representation elements follow the mathematics of Shannon's model at the level of uni-directional communications.

c. Statistical relations between messages, should follow the mathematics of Marko's model at the level of inter-locutional communications.

d. Though continuous information flows to both directions may occur in a 4 wire communication channel, it still may be regarded as a revised model, because two change-over switches are necessary just in front of the information processing units.

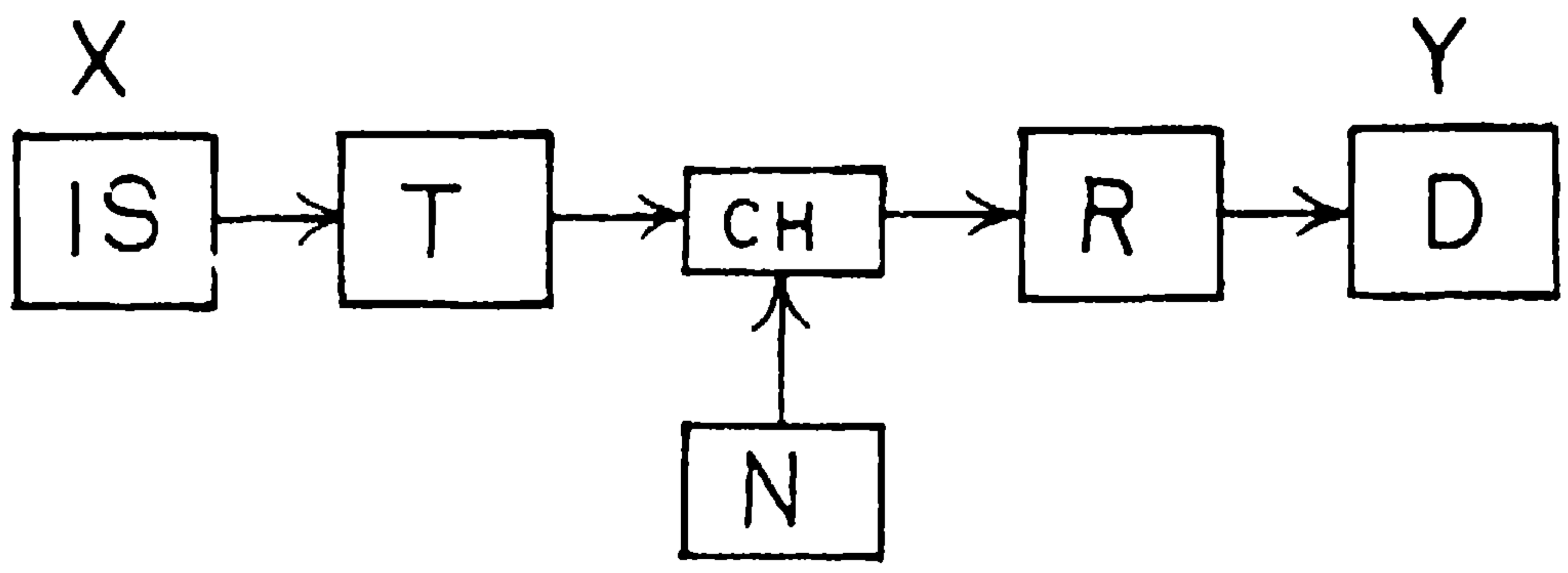
e. The new model can be extended to the inter-locutional communications between more than two nodes by considering conditional probabilities of choosing particular messages. Thus the definition of the value of information and the learning of the group may become possible.

Finally, the author would like to offer

many thanks to the members of I.E.C.E.J for their suggestive discussions.

References

(1) C.E.Shannon : B.S.T.J., 27.pp.379-423; pp. 623-653 (1948)
(2) H. Marko : I.E.E.E. Spectrum, 4 , pp.76-83 (1976)
(3) Lee S. Christie, R. Duncan Luce and Josiah Macy, Jr. : Communication and Learning in Task-Oriented Groups, MIT, RLE Technical Rrport No.231 (May 13, 1952)



IS:Information Source, T:Transmitter, CH:Channel, N:Noise Source, R:Receiver, D:Destination

FIG.1 SHANNON'S MODEL

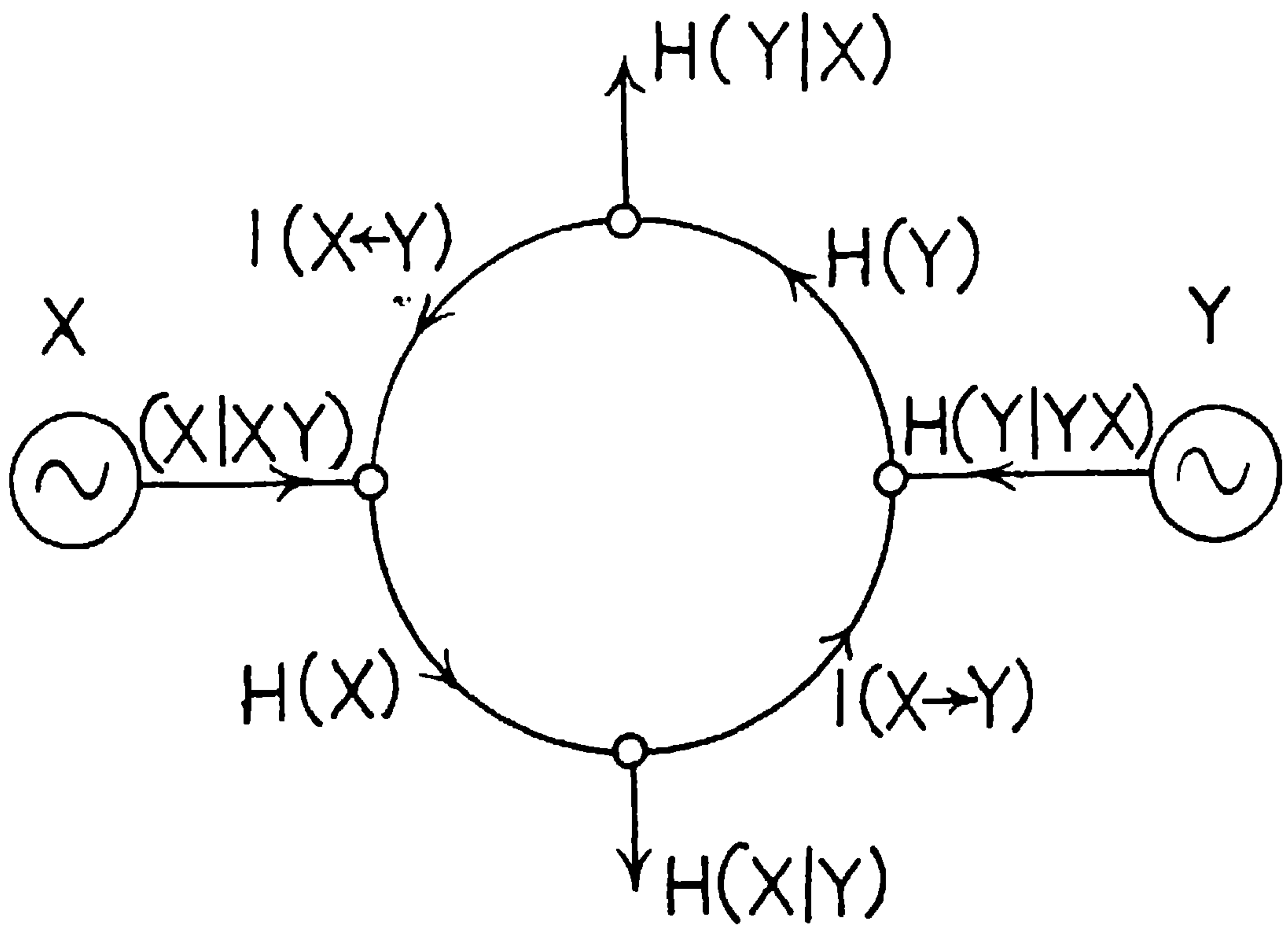


FIG. 2 MARKO'S MODEL

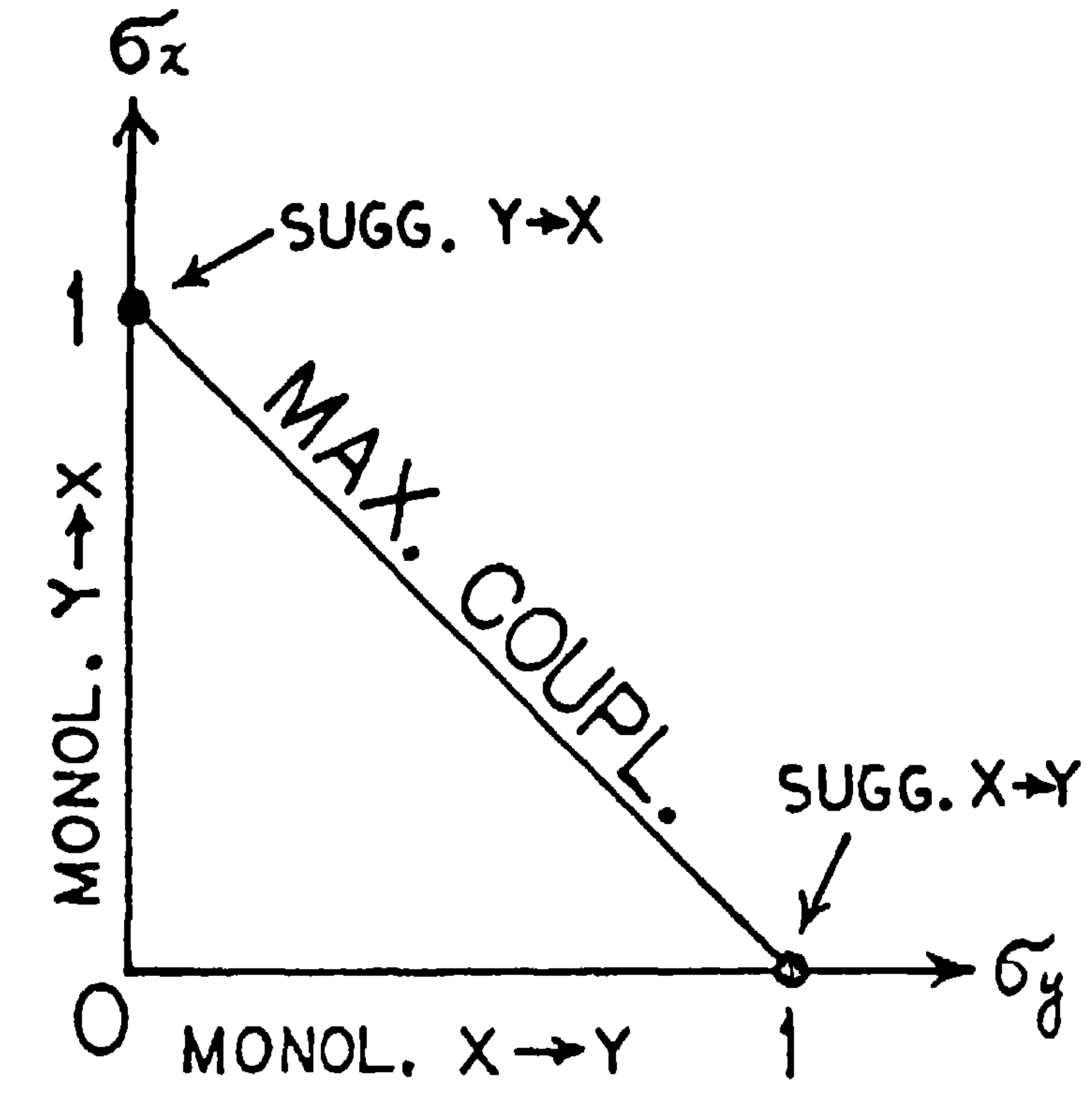


FIG.3 STOCHASTIC DEGREE OF SYNCHRONIZATION

IN:Information Node,
T:Transmitter, R:Receiver,
S:Change-over Switch,
CH:Channel, N:Noise Source

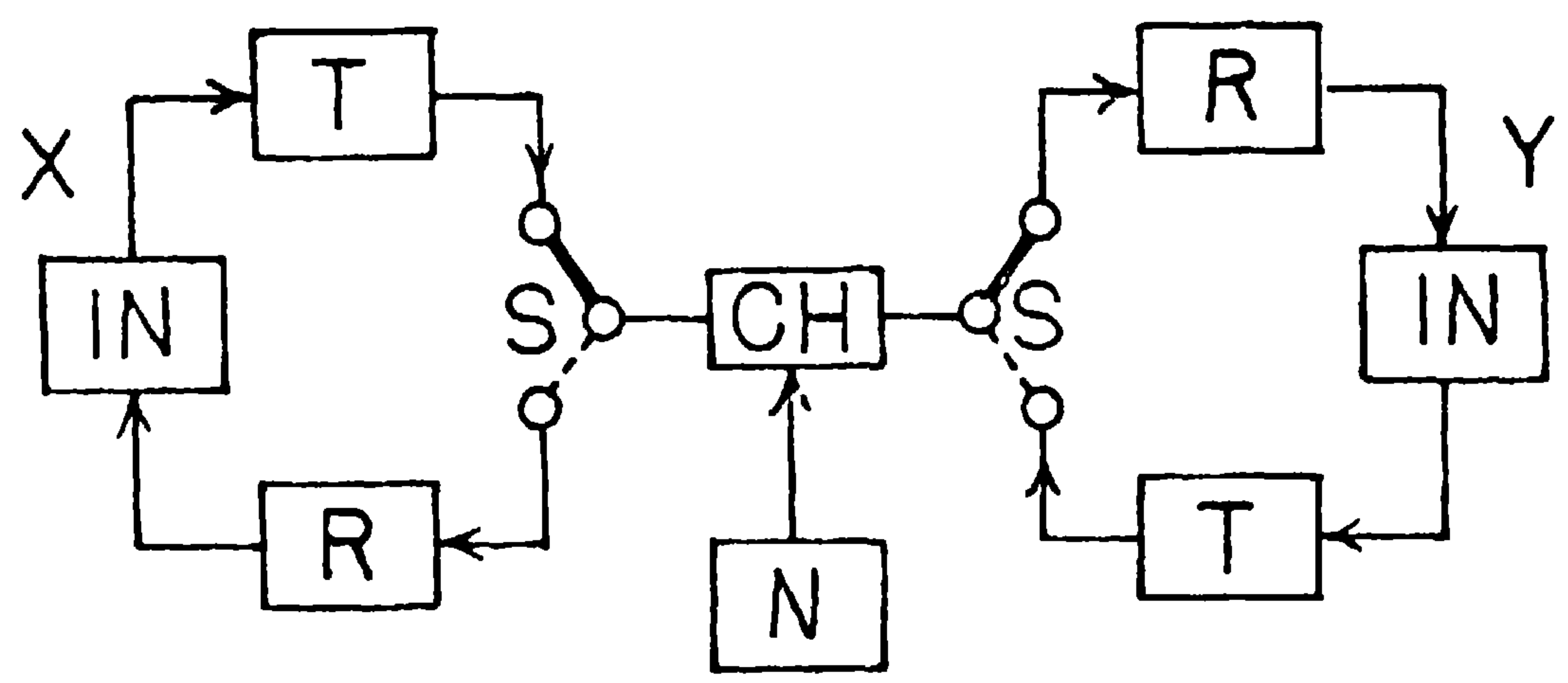


FIG.4 REVISED COMM. MODEL

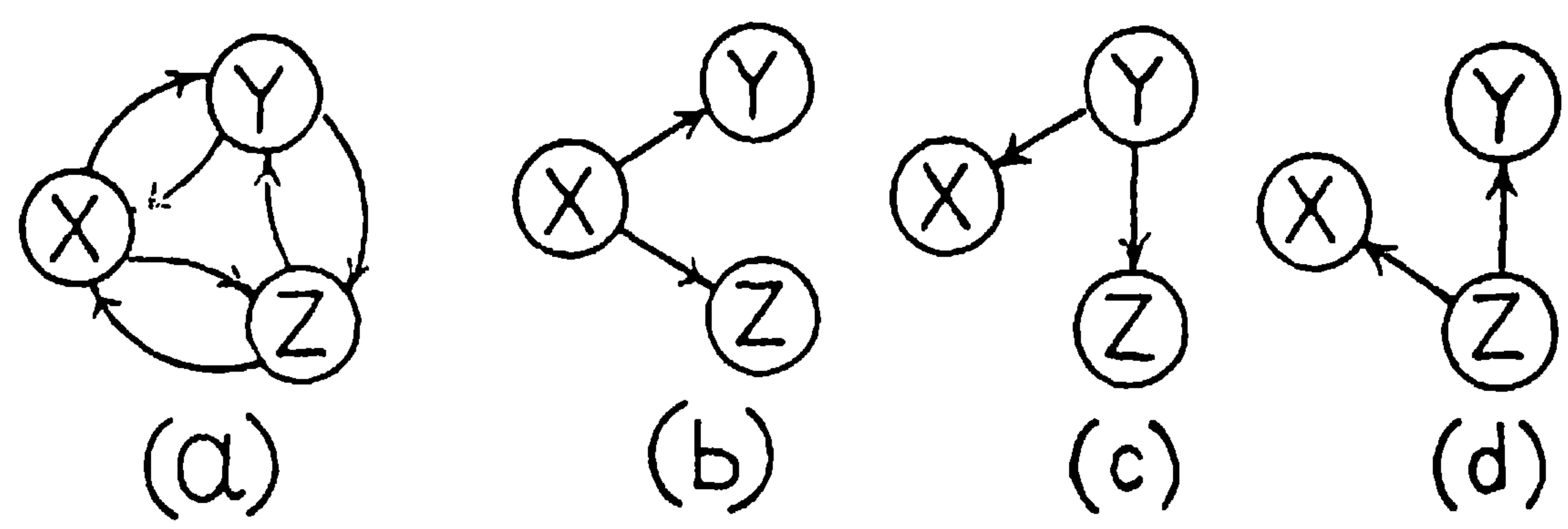


FIG.6 COMMUNICATION BETWEEN 3 NODES

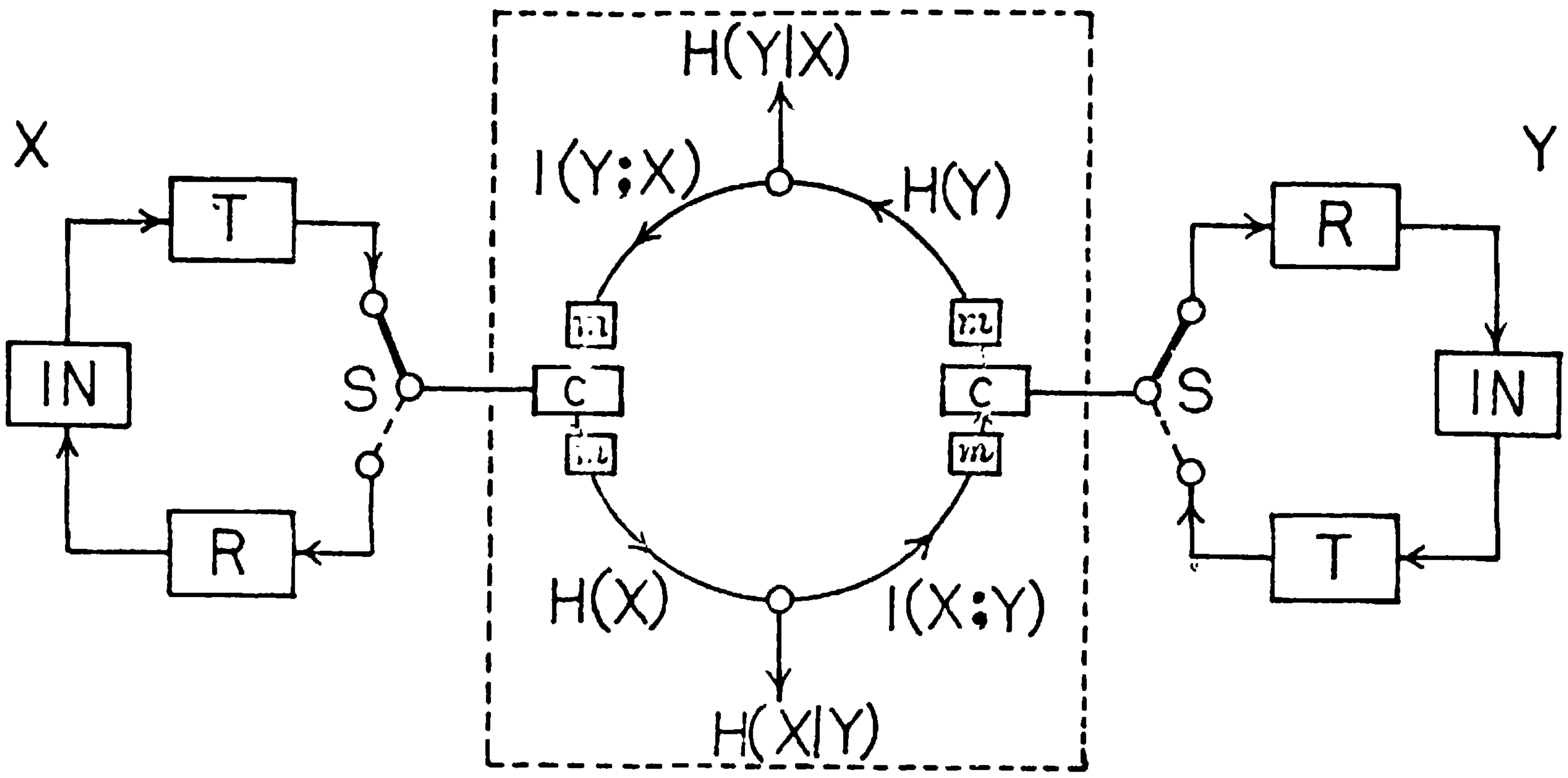


FIG. 5 REVISED MODEL FOR BIDIRECTIONAL CONTINUOUS INFORMATION FLOWS

IN:Information Node, T:Transmitter,
R:Receiver, S:Change-over Switch
C:Control Circuit, m:register

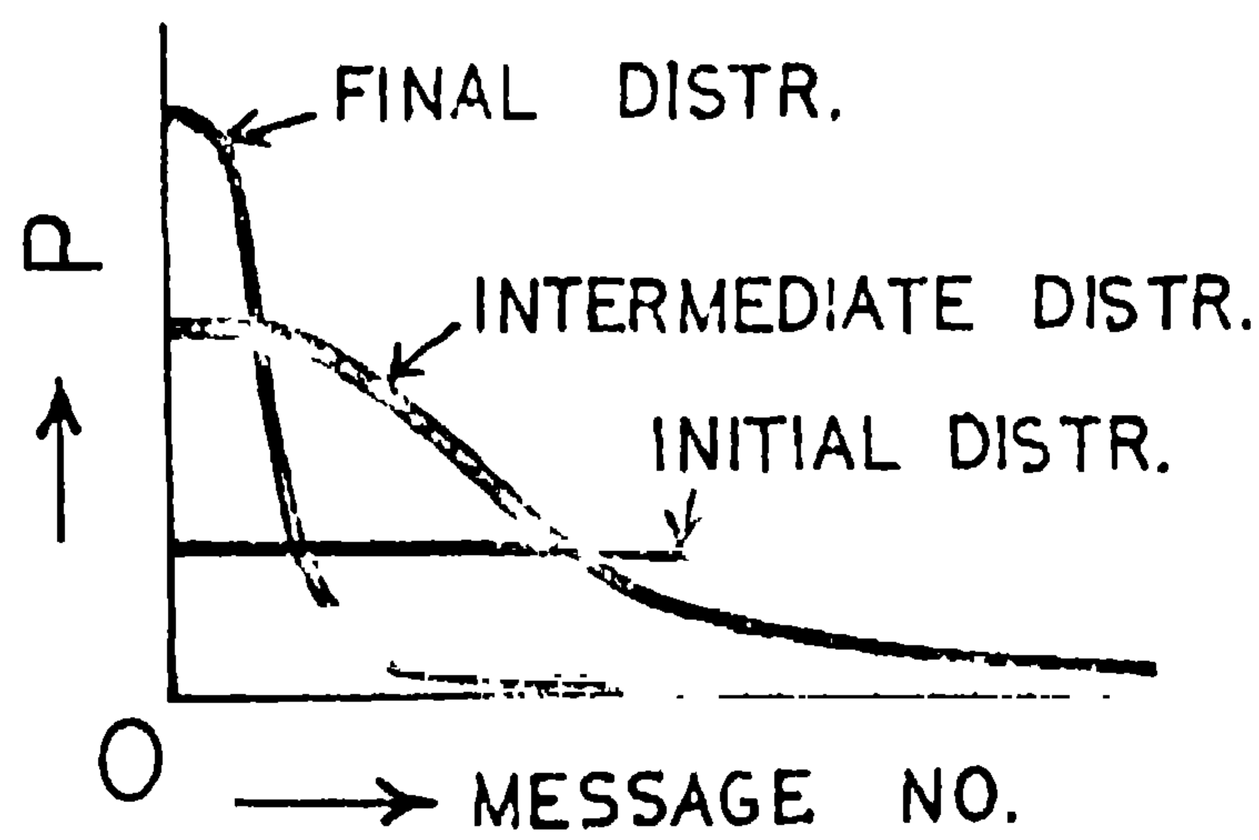


FIG.7 MESSAGE SELECTION PROBABILITY DISTRIBUTION