From (3.4), (3.5), and (3.6), we have

$$R \leq \epsilon_n + \min\left\{ \frac{1}{n} \sum_{i=1}^{n} I(x_{oi}, x_{1i}, x_{2i}; y_{oi}), \frac{1}{n} \sum_{i=1}^{n} I(x_{oi}, x_{1i}; y_{oi}, y_{2i} \mid x_{2i}), \frac{1}{n} \sum_{i=1}^{n} I(x_{oi}; y_{oi}, y_{1i}, y_{2i} \mid x_{1i}, x_{2i}) \right\}$$

By eliminating the variable n in (3.9) (See [3]), it follows that

$$R \leq \epsilon_n + \min\left\{ I(x_o, x_1, x_2; y_o), I(x_o, x_1; y_o, y_1, y_2 \mid x_2), I(x_o; y_o, y_1, y_2 \mid x_1, x_2) \right\}$$

and converse is proved.

Achievability : See Theorem 4.1 in [5].

Now, we generalize Theorem 6 to the general relay network with partial feedback.

Theorem 7: The Capacity of the general relay network with N relays and with feedback from Y to all the relays and feedback from relay ith to the all previous relays is given by

$$C = \sup_{P(x_o, \ldots, x_n)} \min\left\{ I(x_o, x_1, \ldots, x_i; y_o, y_{i+1}, \ldots, y_n \mid x_{i+1}, \ldots, x_n) \right\} \qquad 0 \leq i \leq n$$

IV. Conclusion

In this paper we established the Capacity Theorems for some special relay networks, i.e. The relay networks with partial feedbacks.

The Capacity of general relay network is still an open problem for futher research.

V– References

1. E.c. Van der Meulen, "Three - terminal communication channels," *Adv. Appl. Prob.*, vol 3, pp. 120–154, 1971.
2. H. Sato, "Information Transmission through a channel with relay." The Aloha System, Univ. of Hawaii, Honolulu, Tech. Rep. B76–7, Mar. 1976.
3. T. Cover and A. El Gamal, "Capacity theorems for the relay Channel," *IEEE Trans. Inform. Theory*, Vol. IT–25, no. 5, Sept. 1979.
4. A. El Gamal and M. Aref, "The capacity of the semideterministic relay channel," IEEE Trans., Inform. Theory, Vol. IT–28, no. 3, May 1982.
5. M. Aref, "Information flow in relay networks, Ph.D. dissertation, Stanford Univ., Stanford, CA, Oct. 1980.
6. R.G. Gallager *Information Tehory and Reliable Communication*, N.Y. wiley, 1968.
7. T.M. Cover, "An Achievable Rate Region for the Broadcast Channel," *IEEE Trans. on Information theory*, Vol. IT–21, pp. 399–404, July 1975.
8. L.R. Ford and D.R. Fulkerson, *Flows in Networks*, Princetion, N.J., Princeton University Press, 1962.

Lemma 6: For the relay network in Fig. 7., we have the following upperbounds:

i) $I(w;\underline{y}_o) \leqslant \sum_{i=1}^{n} I(x_{oi},x_{1i},x_{2i};y_{oi})$ (3.4)

ii) $I(w;\underline{y}_0) \leqslant \sum_{i=1}^{n} I(x_{oi},x_{1i};y_{oi},y_{2i}|x_{2i})$ (3.5)

iii) $I(w;\underline{y}_o) \leqslant \sum_{i=1}^{n} I(x_{oi};y_{oi},y_{1i},y_{2i}|x_{1i},x_{2i})$ (3.6)

Proof:

i) $I(w,\underline{y}_o) \overset{①}{=} \sum_{i=1}^{n} I(w,y_{oi}|y_o^{i-1}) =$

$\sum_{i=1}^{n} H((y_{oi}|y_o^{i-1}) - H(y_{oi}|y_o^{i-1},w)$

where $y_o^{i-1} \overset{\Delta}{=} (y_{o1},\cdots,y_{o,i-1})$

since $H(A|B,C) \leqslant H(A|B)$, (see [6]

we have,

$$I(w;\underline{y}_o) \leqslant \sum_{i=1}^{n} H(y_{oi}) - H(y_{oi}|y_o^{i-1},w,x_{oi},x_{1i},x_{2i})$$

$$\overset{②}{=} \sum_{i=1}^{n} H(y_{oi}) - H(y_{oi}|x_{oi},x_{1i},x_{2i})$$

$$= \sum_{i=1}^{n} I(x_{oi},x_{1i},x_{2i};y_{ai})$$

and (3.4) is proved.

(ii) $I(w;\underline{y}_o) \overset{③}{\leqslant} I(w;\underline{y}_o,\underline{y}_2) = \sum_{i=1}^{n} I(w;y_{oi},y_{2i}|y_o^{i-1},y_2^{i-1})$

$$\overset{④}{=} \sum_{i=1}^{n} I(w;y_{oi},y_{2i}|y_o^{i-1},y_2^{i-1},x_{2i})$$

$$= \sum_{i=1}^{n} H(y_{oi},y_{2i}|y_o^{i-1},y_2^{i-1},x_{2i}) - H(y_{oi},y_{2i}|y_o^{i-1},y_2^{i-1},x_{2i},w)$$

$$\overset{⑤}{\leqslant} \sum_{i=1}^{n} H(y_{oi},y_{2i}|x_{2i}) - H(y_{oi},y_{2i}|y_o^{i-1},y_2^{i-1},x_{2i},w,x_{oi},x_{1i})$$

$$\overset{②}{=} \sum_{i=1}^{n} H(y_{oi},y_{2i}|x_{2i}) - H(y_{oi},y_{2i}|x_{oi},x_{1i},x_{2i})$$

$$= \sum_{i=1}^{n} I(x_{oi},x_{1i};y_{oi},y_{2i}|x_{2i})$$

and (3.5) is proved.

iii) $I(w;y_{-o}) \overset{③}{\leqslant} I(w;y_{-o},y_{-1},y_{-2}) \overset{①}{=} \sum_{i=1}^{n} I(w;y_{oi},y_{1i},y_{2i}|y_o^{i-1},y_1^{i-1},y_2^{i-1})$

$$\overset{④}{=} \sum_{i=1}^{n} I(w;y_{oi},y_{1i},y_{2i}|y_o^{i-1},y_1^{i-1},y_2^{i-1},x_{1i},x_{2i})$$

$$= \sum_{i=1}^{n} H(y_{oi},y_{1i},y_{2i}|y_o^{i-1},y_1^{i-1},y_2^{i-1},x_{1i},x_{2i}) - H(y_{oi},y_{1i},y_{2i}|y_o^{i-1},y_1^{i-1},y_2^{i-1},x_{1i},x_{2i},w)$$

$$\overset{⑤}{\leqslant} \sum_{i=1}^{n} H(y_{oi},y_{1i},y_{2i}|x_{1i},x_{2i}) - H(y_{oi},y_{1i},y_{2i}|x_{oi},x_{1i},x_{2i},w,y_o^{i-1},y_1^{i-1},y_2^{i-1})$$

$$\overset{②}{=} \sum_{i=1}^{n} H(y_{oi},y_{1i},y_{2i}|x_{1i},x_{2i}) - H(y_{oi},y_{1i},y_{2i}|x_{oi},x_{1i},x_{2i})$$

$$= \sum_{i=1}^{n} I(x_{oi};y_{oi},y_{1i},y_{2i}|x_{1i},x_{2i})$$

and (3.6) is proved.

① follows from the chain rule for mutual information,

② follows from the fact that the network is memoryless,

③ follows from the fact that the mutual informations is a non-negative quantity,

④ follows from the fact that $x_{2i} = f_{2i}(y_o^{i-1},y_2^{i-1})$ $x_{1i} = f_{1i}(y_o^{i-1},y_1^{i-1},y_2^{i-1})$,

⑤ follows from the fact that $H(A|B,C) \leqslant H(A|B)$

Now, let $(M, n, p_{\overline{2}})$ be any code for the network, assume a uniformly distributed $w \in [1, 2^{nR}]$, then

$$nR = H(w) = I(w;\underline{y}_o) + H(w|\underline{y}_o) \quad (3.7)$$

Fano's inequality [6] requires that

$$H(w|\underline{y}_o) \leqslant \overline{P_e^n} \cdot nR + h((\overline{P_e^n}) \overset{\Delta}{=} n\epsilon n. \quad (3.8)$$

From (3.7) and (s.8), it follows that

$$nR \leqslant I(w;\underline{y}) + n\epsilon n.$$

$$R_o < I(x_2 ; y) \qquad (2.17)$$

iii) Assuming s_i is decoded correctly at the receiver, then $\hat{w}'_{i-1} = w'_{i-1}$ is declared as the part of index sent in block (i−1) iff there exits a unique

$$w' \in S_{s_i} \cap \mathcal{L}\left\{ y\_(i\text{-}1) \right\} \text{ where}$$

$$\mathcal{L}\_\left\{ y\_(i\text{-}1) \right\} = \left\{ w' : (\underline{u}(w' \mid S_{i-1}), x\_2(S_{i-1}), y\_(i\text{-}1) \in A^n_\epsilon \right\}$$

there, for, $\hat{w}'_{i-1} = w_{i-1}$, iff

$$R_1 < R_o + I(U ; y \mid x_2) \qquad (2.18)$$

IV) Knowing S_{i-1}, w'_{i-1}, the receiver declares w''_{i-1}, iff for sufficiently large n $R_2 < I(x_1 ; y \mid x_2, U)$ (2.19)

Combining the conditions (2.16), and (2.19), we have

$$R \triangleq R_1 + R_2 < H(U \mid x_2) + I(x_1 ; y \mid x_2, U) \qquad (2.20)$$

Combining the conditions (2.17), (2.18), and (2.19), we have

$$R < I(U ; y \mid x_2) + I(x_2 ; y) + I(x_1 ; y \mid x_2, U) =$$

$$I(x_1, x_2, U ; y) = I(x_1, x_2 ; y) \qquad (2.21)$$

and the achievability is proved.

III–Capacity of Relay Networks with Partial Feedbacks.

In [5] the capacity of the relay network with feedback from the receiver to all the relays and the source and feedback from the second relay to the first relay and the source and feedback from the first relay to the source (Fig. 6.) was proved to be

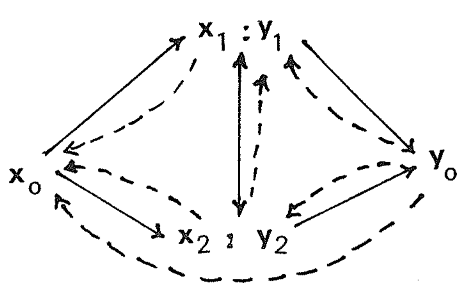

Fig. 6. The Relay Network with Feedback

$$C = \sup_{P(x, x_1, x_2)} \min \left\{ I(x_o, x_1, x_2 ; y_o), I(x_o, x_1 ; y_o, y_2 \mid x_2), I(x_o ; y_o, y_1, y_2 \mid x_1, x_2) \right\} \qquad (3.1)$$

In this work we drop the feedbacks to the source and prove the capacity of the relay network with partial feedback (Fig. 7) to be

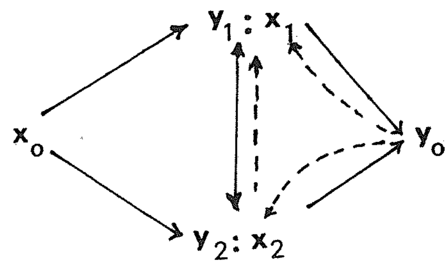

Fig. 7. The Relay Network with Partial Feedback

$$C = \sup_{P(x_o, x_1, x_2)} \min \left\{ I(x_o, x_1, x_2 ; y_o), I(x_o, x_1 ; y_o, y_2 \mid x_2), I(x_o ; y_o, y_1, y_2 \mid x_1, x_2) \right\} \qquad (3.2),$$

i,e, dropping the feedback links to the source does not change the capacity of the network. So, we can use Theorem 4.5 in [5] for the general relay networks with partial feedback.

Remark: Since Y_o is a degraded version of (Y_o, Y_2), and (Y_o, Y_2) is a degraded version of (Y_o, Y_1, Y_2), the network shown in Fig. 7. is a degraded relay network [5].

Theorem 5: The Capacity of the relay network with partial feedback shown in Fig. 7. is given by

$$C = \sup_{P(x_o, x_1, x_2)} \min \left\{ I(x_o, x_1, x_2 ; y_o), I(x_o, x_1 ; y_o, y_2 \mid x_2), I(x_o ; y_o, y_1, y_2 \mid x_1, x_2) \right\} \qquad (3.3)$$

Proof : Converse

The converse is easily proved by using Fano's inequality [6] and the following lemma.

$$\sum_{i=1}^{n} H(U_i|y^{i-1},U^{i-1},x_{2i}) + H(y_i|y^{i-1},U^i,x_{2i}) - H(U_i|y^{i-1},U^{i-1},W) - H(y_i|y^{i-1},U^i,W) \overset{⑤}{\leqslant}$$

$$\sum_{i-1}^{n} H(U_i|x_{2i}) + H(y_i|U_i,x_{2i}) - H(y_i|x_{1i},x_{2i},y^{i-1},U^{i-1},W,U_i)$$

$$\overset{⑥}{=} \sum_{i-1}^{n} H(U_i|x_{2i}) + H(y_i|x_{2i},U_i) - H(y_i|x_{1i},x_{2i},U_i) =$$

$$\sum_{i-1}^{n} H(U_i|x_{2i}) + I(x_{1i};y_i|x_{2i},U_i)$$

where,

① follows from the fact that mutual information is a non-negative quantity,

② and ③ follow from the chain rule for entropy [6],

④ follows from the fact that $x_{2i} = f_i(u^{i-1})$,

⑤ follows from the fact that $H(A|B,C) \leqslant H(A|B)$, and

⑥ follows from the fact that the channel is memoryless [6].

Using Fano's inequality (2.5) and (2.2) and (2.14), we get

$$R \leqslant \in_n + \min\left\{\frac{1}{n}\sum_{i=1}^{n} I(x_{1i},x_{2i};y_i), \frac{1}{n}\sum_{i=1}^{n} H(U_i|x_{2i}) + I(x_{i1};y_{1i}|x_{2i},U_i)\right\} \quad (2.15)$$

By eliminating the variable n in (2.15), it follows that

$$R \leqslant \in_n + \min \; I(x_1,x_2;y);\; H(U|x_2) + I(x_1;y_1|x_2,U)$$

and the converse is proved.

Achievability: Proof follows under the same assumptions as in the Theorem 1.

Random Coding

1) Generate 2^{nRo} i.i.d n-sequences in x_2^n, each with probability

$$P(\underline{x}_2) = \begin{cases} \dfrac{1}{\|A_\epsilon^n(x_2)\|}, & \text{if } \underline{x}_2 \in A_\epsilon^n(x_2) \\ 0 & \text{, otherwise} \end{cases}$$

Index them as $\underline{x}_2(s)$, $s \in [1, 2^{nRo}]$

2) For each $\underline{x}_2(s)$, generate 2^{nR_1} conditionally independent n-sequences in U^n, each with probability

$$P(\underline{u}) = \begin{cases} \dfrac{1}{\|A_\epsilon^n(u|x_2)\|} & \text{if } u \in A_\epsilon^n(x_2,u) \\ 0 & \text{, otherwise} \end{cases}$$

Index them as $\underline{u}(w'|s)$, $w' \in [1, 2^{nR_1}]$

3) For each u, generate 2^{nR_2} i.i.d. n-sequences x_{-1} each with probability

$$P(\underline{x}_1) = \begin{cases} \dfrac{1}{\|A_\epsilon^n(x_1|u)\|}, & \text{if } x_1 \in A_\epsilon^n(x_1,u) \\ 0 & \text{, otherwise} \end{cases}$$

Index them as $x_{-1}(w''|w',s)$, $w'' \in [1, 2^{nR2}]$.

4) Randomly partition $[1, 2^{nR1}]$ into 2^{nRo} cells such that

$P(w \in S_i) = 2^{-nR_o}$, $w \in [1, 2^{nR_1}]$, $S_i \cap S_j = \phi$, $i \neq j$, $U_i S_i = [1, 2^{nR_1}]$.

Encoding: Let $w_i \triangleq (w_i', w_i'')$ be the new index to be sent in block i, and assume $w'_{i-1} \in S_i$.

The source knowing s_i and w_i transmits $\underline{x}_1(w''_i|w'_i, S_i)$.

The relay has an estimate w'_{i-1} of the previous index, transmits $\underline{x}_2(s_i)$ $(\hat{w}'_{i-1} \in \hat{s}_i)$.

Decoding: (At the end of Block i)

We assume at the end of block (i—1), the receiver knows $(w_1, \ldots, w_{i-2})$, $(s_1, \ldots, s_i)$, and the relay knows $(w_1, \ldots, w'_{i-1})$ and $(s_1, \ldots, s_i)$

i) Knowing s_i, and uppon receiving $\underline{y}_1(i)$ the relay estimates $\hat{w}_i' = w_i'$ was sent iff $R_1 < H(U|x_2)$ (2.16)

ii) The receiver estimates $\hat{s}_i = s_i$, by looking for a unique s_i such that $(\underline{x}_2(s_i), \underline{y}(i)) \in A_\epsilon^n$

Using Lemma 2.3 in [5],

$\hat{s}_i = s_i$, iff (for sufficiently large n)

$$R < I(x_1; y, y_1 \mid x_2) \qquad (2.9)$$

ii) The receiver estimates $\hat{s}_i = s$, by looking for a unique s such that $(\underline{x}_2(s), \underline{y}(i)) \in A^n_\epsilon$. Using Lemma 2.3 in [5], it is easy to see that $\hat{s}_i = s_i$ with arbitrarily small probablity of error provided n is sufficiently large and

$$R_0 < I(x_2; y) \qquad (2.10)$$

iii) Assuming s_i is decoded correctly at the receiver, then $\hat{w}_{i-1} = w$ is declared as the index in block (i–1), iff there exists a unique

$$w \in S_{si} \cap \mathcal{L}\{\underline{y}(i-1)\}$$

where $\mathcal{L}\{\underline{y}(i-1)\}$ is the list of indices w that the receiver y considered to be "typical" with $\underline{y}$ (i–1) in the (i–1) block, i.e.

$$\mathcal{L}\{\underline{y}(i-1)\} = \{w : (\underline{x}_1(w \mid s_{i-1}, \underline{x}_2(s_{i-1}), \underline{y}(i-1)) \in A^n_\epsilon\}.$$

From Lemmas 2.4 and 2.5 in [5], it is easy to see that $\hat{w}_{i-1} = w_{i-1}$ with arbitrarily small probability of errors if n is sufficiently large and

$$R < R_0 + I(x_1; y \mid x_2) \qquad (2.11)$$

Combining the two constraints (2.10) and (2.11), we have

$$R < I(x_2; y) + I(x_1; y \mid x_2) = I(x_1, x_2; y) \quad (2.12)$$

Therefore, from (2.9) and (2.12), we have

$$R_0 < \min\{I(x_1, x_2; y), I(x_1; y_1, y \mid x_2)\} \quad (2.13)$$

and the achievability is proved.

Bounding the probability of Error: It can easily be shown that with the above rates (2.13), $\overline{P^n_e} < \epsilon$. For the details of the calculation, See [3].

Theorem 3: The Capacity of the Relay Channel with Feedback from y_1 to x_1 is given by (Fig. 5.)

$$C = \max_{P(x_1, x_2, u)} \min \left\{ I(x_1, x_2; y), I(x_1; y \mid x_2, U) + H(U \mid x_2) \right\}$$

where U is an arbitrary auxiliary random variable.

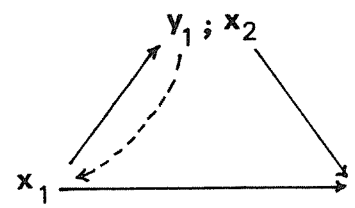

Fig. 5. The Relay Channel with Feedback from y_1 to x_1

Corollaries:

1. In the relay channel (without feedback) if y_1 is a deterministic function of x_1 and x_2, say $f(x_1, x_2)$, the transmitter knows which symbol is going to be sent by the relay because $x_{2i} = g_i(y_1^{i-1})$. Therefore, we can assume $U \equiv y_1$ and get the capacity of the semideterministic relay channel obtained in [4] and [5].
2. Since the transmitter knows y_1^{i-1}, and $x_{2i} = g_i(y_1^{i-1})$, we can assume $U \equiv x_2$ and get the capacity of the reversely degraded relay channel obtained in [3].

Proof : Converse

Lemma 4: For the relay channel shown in Fig. 5., we have the following upper bound

$$I(w; \underline{y}) \leq \sum_{i=1}^{n} H(U_i \mid x_{2i}) + I(x_{1i}, y_i \mid x_{2i}, U_i),$$

$$w \in [1, 2^{nR}] \qquad (2.14)$$

Proof: In the following, we assume $x_{2i} = f_i(u^{i-1})$, where

$$u^{i-1} = g_i(y_1^{i-1}), \text{ i.e., } x_{2i} = f_i(g_i(y_1^{i-1})) = h_i(y_1^{i-1})$$

$$I(W; \underline{y}) \overset{①}{\leq} I(W; y, U) \overset{②}{=} \sum_{i=1}^{n} I(W; y_i, U_i \mid y^{i-1}, U^{i-1}) =$$

$$\sum_{i=1}^{n} H(y_i, U_i \mid y^{i-1}, U^{i-1}) - H(y_i, U_i \mid y^{i-1}, U^{i-1}, W) \overset{③}{=}$$

$$\sum_{i=1}^{n} H(U_i \mid y^{i-1}, U^{i-1}) + H(y_i \mid y^{i-1}, U^i) -$$

$$H(U_i \mid y^{i-1}, U^{i-1}, W) - H(y_i \mid y^{i-1}, U^i W) \overset{④}{=}$$

Fano's inequality [6] requires that

$$H(w \mid \underline{y}) \leq \overline{P_e^n} \cdot nR + h(\overline{P_e^n}) \triangleq n.\epsilon_n \quad (2.5)$$

where

$$h(\overline{P_e^n}) \triangleq -\overline{P_e^n} \log \overline{P_e^n} - (1-\overline{P_e^n}) \log(1-\overline{P_e^n})$$

Note that if $\overline{P_e^n} \to 0$, then $\epsilon_n \to 0$.

From (2.4) and (2.5), it follows that

$$nR \leq I(w;\underline{y}) + n\epsilon_n \quad (2.6)$$

From (2.3), (2.4), and (2.6), we have

$$R \leq \epsilon_n + \min\left\{\frac{1}{n}\sum_{i=1}^{n} I(x_{1i}, x_{2i}; y_i), \frac{1}{n}\sum_{i=1}^{n} I(x_{1i}; y_i, y_{1i} \mid x_{2i})\right\} \quad (2.7)$$

By eliminating the variable n in (2.7), it follows that (See [3])

$$R \leq \epsilon_n + \min\left\{I(x_1, x_2; y), I(x_1; y, y_1 \mid x_2)\right\} \quad (2.8)$$

Proof of Achievability : Fix a probability mass function $p(x_1, x_2)$. The achivability proof involves 1) random coding, 2) list codes, 3) Slepian-Wolf partitioning 4) superposition coding, and 5) block Markov encoding at the relay and transmitter.

We consider B blocks of transmission, each of n symbols. A sequence of B–1 indices $w_i \in [1, 2^{nR}]$, $i = 1, 2, \ldots B-1$, will be sent over the channel in n B transmissions. (Note that as $B \to \infty$, for fixed n, the rate $\frac{R(B-1)}{B}$ is arbitrarily close to R).

In each n-block b = 1, 2, . . . , B, we shall use the same set of codewords

$$C = \left\{\underline{x}_1(w \mid s), \underline{x}_2(s)\right\}, \quad w \in [1, 2^{nR}],$$

$$s \in [1, 2^{nR}] \quad x_1(.|.) \in X_1^n, \; x_2(.) \in X_2^n$$

For reasons to be made clear later, we shall also need a random partition

$$S = \left\{S_1, \ldots, S_{2^{nR}}\right\} \text{ of } \mathcal{W} = [1, 2^{nR}]$$

into 2^{nR} disjoint cells, $S_i \cap S_j = \phi$, $(i \neq j)$, $\bigcup_i S_i = \mathcal{W}$

Note that the feedback changes an arbitrary relay channel into a degraded relay channel in which x_1 transmits information to x_2 by way of y_1 and y.

Clearly Y is a degraded form of (Y, Y_1).

Random Coding:

1) Generate 2^{nR} i.i.d. (independent, identically distributed) n-sequences in X_2^n, each with probability $p(\underline{x}_2) = \prod_{i=1}^{n} P(x_{2i})$. Index them as

$$\underline{x}_2(s), \; s \in [1, 2^{nR_0}]$$

2) For each $\underline{x}_2(s)$, generate 2^{nR} conditionally independent n-sequences in X_1^n, each with probability

$$P(\underline{x}_1 \mid \underline{x}_2(s)) = \prod_{i=1}^{n} P(x_{1i} \mid x_{2i}(s)).$$

Index them as $\underline{x}_1(w \mid s)$, $w \in [1, 2^{nR}]$

3) Randomly partition $[1, 2^{nR}]$ into 2^{nR} cells S_i, $i \in [1, 2^{nR_0}]$,

such that $P(w \in S_i) = 2^{-nR_0}$, $S_i \cap S_j = \phi$, $(i \neq j)$, $\bigcup_i S_i = [1, 2^{nR}]$.

Encoding: Let $w_i \in [1, 2^{nR}]$ be the new index to be sent in block i, and assume that $w_{i-1} \in S_i$. The source knowing w_i, s_i transmits $\underline{x}_1(w_i \mid s_i)$. The relay has an estimate $\hat{w}_{i-1}$ of the previous index w_i (to be made precise in the decoding section), $(\hat{w}_{i-1} \in \hat{S}_i)$ transmits $\underline{x}_2(S_i)$.

Decoding: We assume at the end of block (i–1), the receiver knows $(w_1, \ldots, w_{i-2})$, and $(s_1, \ldots, s_{i-1})$ and the relay knows $(w_1, \ldots, w_{i-1})$, and $(s_1, \ldots, s_i)$. The decoding procedures at the end of block i are as follows:

i) knowing s_i, and uppon receiving $\underline{y}_1(i)$, $\underline{y}(i)$, the relay estimates the message of the transmitter $\hat{w}_i = w$ iff there exits a unique w such that $\underline{x}_1(w \mid s_i)$, $\underline{x}_2(s_i), \underline{y}_1(i), \underline{y}(i)$ are jointly ϵ - typical. (For the definition of ϵ - typicality and its applications, See Cover [7]), i.e.

$$(\underline{x}_1(w_1 \mid s_i), \underline{x}_2(s_i), \underline{y}_1(i), \underline{y}(i)) \in A_\epsilon^n \Rightarrow w = \hat{w}_i$$

Using Lemma 2.4 in [5], it can be shown that $\hat{w}_i = w_i$ with arbitrarily small probablity of error if transmitted symbols $(x_{oi}, x_{1i}, \dots x_{Ni})$.

Therefore, the joint probability mass function on $m \times x^n_o \times x^n_1 \times \dots \times x^n_N \times y^n_o \times y^n_1 \times \dots \times y^n_N$ is given by $p(w, \underline{x}_o, \underline{x}_1, \dots, \underline{x}_N, \underline{y}_o, \dots, \underline{y}_N) =$

$$P(w) \prod_{i=1}^{n} P(x_{oi}|w) P(x_{1i}|y_1^{i-1}) \dots P(x_{Ni}|y_n^{i-1}) \cdot P(y_{oi}, \dots, y_{Ni}|x_{oi}, \dots, x_{Ni})$$

Where p(w) is the probability distribution on the message $w \in m$

If the message $w \in m$ is sent, let

$$\lambda(w) \triangleq P_r \{g(y_o) \neq w | w=w\}$$

denote the conditional probability of error. The average probability of error, the achievable rate and the capacity of the network are defined as those of the relay channel.

In [5] Aref established a general upper bound to the capacity of relay networks which has a nice max-flow min-cut [8] interpretation.

He also established the capacity when the network is degraded, and when feedback is added from the sink Y_0 to the source and all the relays, and from ith node to the all previous relays and the source, i.e, there is a feedback from Y_i to $(X_{i-1}, \dots, X_1, X_0)$, $1 \leqslant i \leqslant N$

The capacity of some deterministic relay networks were also established in [5] and max-flow min-cut interpretation of the results were given.

In this paper, we establish the capacity of the relay channel when there is only partial feedback available, i.e., when there is only feedback from the receiver to the relay, and when there is feedback from the relay to the sender (source). Also the capacity of the relay networks with partial feedbacks is established.

II. Capacity of the Relay Channel with Partial Feedback

In [3] the capacity of the relay channel with feedback from (y, y_1) to both x_1 and x_2 was proved (Fig. 3.) to be

$$C = \max_{P(x_1,x_2)} \min \left\{ I(x_1, x_2; y), I(x_1; y, y_1 | x_2) \right\},$$

In this work we establish the capacity of the relay

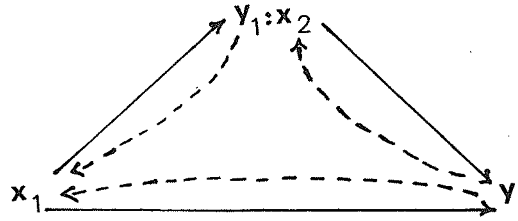

Fig. 3: The Relay Channel with Feedback

channel when i) there is a feedback from y to x_2, ii) there is a feedback from y_1 to x_1. That is, the capacity of the partial feedbacks are established.

Theorem 1. : The capacity of the relay channel with feedback from y to x_2 is given by (Fig. 4.)

$$C = \max_{P(x_1,x_2)} \min \left\{ I(x_1, x_2; y), I(x_1; y, y_1 | x_2) \right\} \quad (2\text{-}1)$$

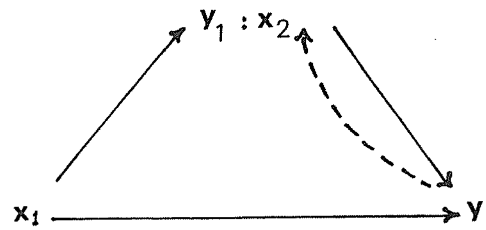

Fig. 4: The Relay Channel with Feedback from y to x_2

Proof : Converse

The converse is proved by using Fano's inequality [6] and the following lemma.

Lemma 2. For the relay channel (Fig. 4.) we have the following upperbounds

$$I(w; \underline{y}) \leqslant \sum_{i=1}^{n} I(x_{1i}, x_{2i}; y_i), \quad w \in [1, 2^{nR}], \quad (2.2)$$

$$I(w; \underline{y}) \leqslant \sum_{i=1}^{n} I(x_{1i}; y_i, y_{1i} | x_{2i}) \quad (2.3)$$

Prof : See [3, 5]

Let (M, n, P^n_e) be any code for the channel, assume a uniformly distributed $w \in [1, 2^{nR}]$; then

$$nR = H(W) = I(w; \underline{y}) + H(w|\underline{y}) \quad (2.4)$$

finition used by Vander Meulen [1]. The channel is memoryless in the sense that (y_i, y_{1i}) depends on the past (x_1^i, x_2^i) only through the current transmitted symbols (x_{1i}, x_{2i}). Thus, for any choice p(w), $w \in \mathbf{m}$, and code choice x_1 ; $[1, 2^{nR}] \to x_1^n$ and relay functions $\{f_i\}_{i=1}^n$, The joint probability mass function on $\mathbf{m} \times x_1^n \times x_2^n \times y^n \times y_1^n$ is given by

$$P(w, \underline{x}_1, \underline{x}_2, \underline{y}, \underline{y}_1) = P(w) \prod_{i=1}^{n} P(x_{1i} \mid w) \times P(x_{2i} \mid y_{11}, \dots, y_{1,i-1}) P(y_i, y_{1i} \mid x_{1i}, x_{2i}).$$

If the message $w \in \mathbf{m}$ is sent, let

$$\lambda(w) = P_r \left\{ g(y) \neq w \mid w \text{ sent} \right\}$$

denote the conditional probability of error. We define the average probability of error of the code to be

$$\overline{P_e^n} = 2^{-nR} \sum_w \lambda(w)$$

The probability of error is colculated under uniform distribution over the codewords $w \in [1, 2^{nR}]$.The rate R is said to be achievable by a relay channel if there exists a sequence of $(2^{nR}, n)$ codes with $P_e^n \to 0$. The capacity C of the relay channel is the supremum of the set of achievable rates.

In [3] a general upper-bound to the capacity of any relay channel was established, and the capacity when the relay channel is degraded, reversely degraded, and when feedback is added from the receivers y and y_1', to both senders x_1 and x_2 were established.

In [4], EL Gamal and Aref established the capacity of the semideterministic relay channel.

Relay networks was introduced by Aref [5]. The discrete memoryless relay network shown in Fig 2 is a model for the communication between a source x_0 and a sink y_0 via N intermediate nodes called relay.The relays receive signals from the source and other nodes and then trasmit their information to help the sink to resolve his uncertainty about the message. To specify the network, we define 2N + 2 finite sets: $x_0, x_1, \dots x_n, y_0, y_1, \dots, y_n$ and a probability transition matrix $P(y_0, \dots, y_N \mid x_0, \dots, x_N)$ defined for all $(y_0, \dots, y_N, x_0, \dots, x_N)$ $\in y_0 \times y_1 \times \dots \times y_N \times x_0 \times \dots \times x_N$

In this model X_0 is the input to the network, Y_0 is the ultimate output, Y_i is the ith relay output and X_i is the ith relay input. The problem is to find the capacity of the network between the source X_0 and the sink Y_0.

An (M, n) code for the network consists of a set of integers $\mathbf{m} = \{1, \dots, M\}$ an encoding function $x_0 : \mathbf{m} \to x_0^n$, a set of relay functions $\{f_{ij}\}$ such that

$$x_{ij} = f_{ij}(y_{i1}, \dots, y_{i,j-1}) \qquad 1 \leqslant j \leqslant n, 1 \leqslant i \leqslant N$$

$$\underline{x}_i \triangleq (x_{i1}, \dots, x_{iN}),$$

i.e , $x_{ij} \triangleq j^{th}$ component of $\underline{x}_i$ and a decoding function $g : y_0^n \to \mathbf{m}$

For generality, all functions are allowed to be stochastic functions. The input x_{ij} is allowed to depend only on the past received signals at the ith node, i.e. , ($y_{i1}, \dots, y_{i,j-1}) \triangleq y_i^{j-1}$. The network is memoryless in the sense that $(y_{0i}, y_{1i}, \dots, y_{Ni})$ depends on the past $(x_0^i, x_1^i, \dots, x_N^i)$ only through the present

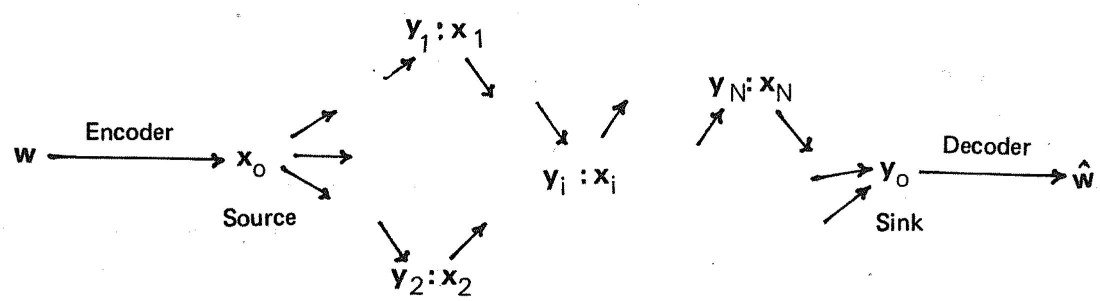

Fig. 2: General Discrete Memoryless Relay Network

Capacity Theorems for the Relay Network with Partial Feedback

M.R. Aref, Ph.D.

Elect. Eng. Dept. Isfahan Univ. of Tech.

G.A. Hodtani, M.Sc.

Elect. Eng. Dept. Ferdowsi Univ. Mashad

ABSTRACT

The capacity of the relay channel with partial feedback is established. Also, the capacity of the relay network with two nodes and partial feedback is established. The results are generalized to the general relay network with partial feedback.

I– Intorduction

The discrete memoryless relay channel denoted by $(x_1 \times x_2, P(y_1 y_2 | x_1 x_2), y \times y_1)$ consists of four finite sets: $x_1\, x_2, y, y_1$ and a collection of probability distributions $P(.,.| x_1, x_2)$ on $y \times y_1$ one for each (x_1, x_2).

Actually, x_1 is the input to the channel and y is the output, y_1 is the output of the relay and x_2 is the input symbol chosen by the relay as shown in Fig. 1.

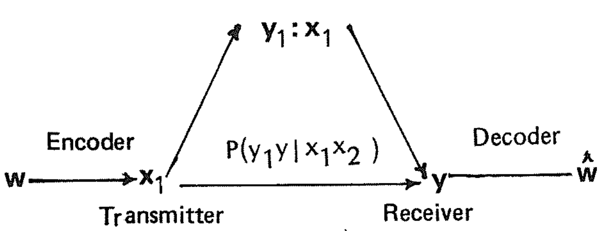

Fig 1: The Relay channel

The relay channel was introduced by Vander Meulen [1] and was studied by Sato [2]. The following discussion is based on Cover and EL Gamal [3].

A $(2^{nR}, n)$ code for the relay channel consists of a set of integers $m = \{1, 2, \ldots, 2^{nR}\} \triangleq [1, 2^{nR}]$, an encoding function $x_1 : [1, 2^{nR}] \rightarrow x_1^n$ a set of relay function $\{f_i\}_{i=1}^n$ such that

$$x_{2i} = f_i(y_{11}, y_{12}, \ldots, y_{1,i-1}) = f_i(y_1^{i-1}), \quad 1 \leqslant i \leqslant n$$

where, $y_1^{i-1} \triangleq (y_{11}, \ldots, y_{1,i-1})$

and a decoding function

$g : y^n \rightarrow [1, 2^{nR}]$

Note that the allowed relay encoding functions actually form part of the definition of the relay channel because of th nonanticipatory condition of the relay. The relay channel input x_{2i} is allowed to depend only on the past $(y_{11}, y_{12}, \ldots, y_{1,i-1})$. This is the de-