Network Information Theory

A General Network

- There are m nodes trying to communicate with each other. (Before today our discussion has been limited to one sender and one receiver.)
- Node n sends x_n, depending on the message it wants to send and past received symbols.
- It also receives y_n depending on the signals being sent over the network.
- The noises and interferences is characterized by the conditional probability

$$p(y_1, \ldots, y_m | x_1, \ldots, x_m).$$

Network Problem

- Given the source distributions (which may be mutually dependent) and network transition probability, can one transmit these sources over the network reliably?
- There are some special networks for which this question can be answered.
- The max-flow min-cut theorem also gives upper bound on the achievable information rates (capacity) between two nodes over a network.

Gaussian Channels

- A Gaussian channel is characterized by

$$Y_i = X_i + Z_i,$$

where Z_i are i.i.d. Gaussian r.v. with variance N.

- The channel capacity for this channel is

$$C = \frac{1}{2} \log(1 + \frac{P}{N}),$$

where P is the constraint on signal power.

The Multiple-Access Gaussian Channels

- Suppose we have m transmitters, each with a power P. A single receiver receives signal

$$Y = \sum_{i=1}^{m} X_i + Z.$$

- Let the rates be $R = (R_1, \ldots, R_m)$. That is, we have m codebooks, with 2^{nR_i} codewords of power P for codebook i.
- The decoder is looking for a set of m codewords, one from each codebook, such that the sum is closest to Y.

The Achievable Rates

- The achievable rates $(R_1, \ldots, R_m)$ must satisfy

$$
\begin{aligned}
R_i &\leq C(\frac{P}{N}) \\
R_i + R_j &\leq C(2\frac{P}{N}) \\
&\vdots \\
\sum_{i=1}^{m} R_i &\leq C(m\frac{P}{N})
\end{aligned}
$$

  - The last inequality dominates the others.
  - $\sum_{i=1}^{m} R_i$ is not bounded as $m \to \infty$.

The Gaussian Broadcast Channels

- A sender of power P and two receivers with Gaussian noises of variance N_1 and N_2.

$$Y_i = X + Z_i.$$

- Suppose that $N_1 \leq N_2$. Let the achievable rates be R_1, R_2 respectively, then

$$R_1 \leq C(\frac{\alpha P}{N_1})$$

$$R_2 \leq C(\frac{(1-\alpha)P}{\alpha P + N_2})$$

where α is the proportion of signal power for receiver 1.

The Gaussian Relay Channel

- Suppose the sender is X and the receiver is Y. In addition, there is an intermediate relay which sends X_1 and receives Y_1 (Figure 14.30). Therefore,

$$Y_1 = X + Z_1$$
$$Y = X + Z_1 + X_1 + Z_2$$

- The capacity is given by

$$C = \max_{\alpha} \min\{C(\frac{P + P_1 + 2\sqrt{\bar{\alpha} P P_1}}{N_1 + N_2}), C(\frac{\alpha P}{N_1})\}$$

The Gaussian Interference Network

- As shown in Figure 14.5, we have two senders and two receivers. The channels are interfering.
- This is different from the broadcast network since there is only one intended receiver for each sender. This is also different from the multi-access network since each receiver is only interested in the signal intended for it.
- The signal model is

$$Y_1 = X_1 + aX_2 + Z_1$$
$$Y_2 = X_2 + aX_1 + Z_2$$

The Gaussian Two-way Channel

- As shown in Figure 14.6, we have two nodes, each with a sender and a receiver. Node 1 sends message to node 2 and vice versa.
  - X_1 is intended for Y_2, and X_2 is intended for Y_1.
- Since node 1 receives from node 2, it can use this as **feedback** information in deciding what to send next.

Jointly Typical Sequences

- Let $\Omega = \{X_1, \ldots, X_k\}$ be a set of discrete random variables. Let S denote an ordered subset of Ω. Consider n samples of S, $\mathbf{s} = (s_1, \ldots, s_n)$.

$$p(\mathbf{s}) = \prod_{i=1}^{n} p(S = s_i).$$

- By the law of large numbers,

$$-\frac{1}{n} \log p(\mathbf{s}) \to H(S).$$

Typical Set

- The set $A_\epsilon^{(n)}$ of ϵ-typical n-sequences is defined by

$$A_\epsilon^{(n)}(\Omega) = \{(\mathbf{x_1}, \ldots, \mathbf{x_k}) : |-\frac{1}{n}\log p(\mathbf{s}) - H(S)| < \epsilon,\ \forall S \subseteq \Omega\}$$

- An element in the typical set is a typical sequence.
- $A_\epsilon^{(n)}(S)$ denotes the restriction of $A_\epsilon^{(n)}$ to the coordinates of S.
- Notation

$$a_n \doteq 2^{n(b\pm\epsilon)} \ \text{ means } \ |\frac{1}{n}\log a_n - b| \leq \epsilon.$$

Properties for Typical Sets

Given an ϵ, for sufficiently large n,

1. $p(A_\epsilon^{(n)}(S)) \geq 1 - \epsilon$ for any $S \subseteq \Omega$.

2. $p(\mathbf{s}) \doteq 2^{-n(H(S) \pm \epsilon)}$ for any $\mathbf{s} \in A_\epsilon^{(n)}(S)$.

3. $|A_\epsilon^{(n)}(S)| \doteq 2^{n(H(S) \pm 2\epsilon)}$.

4. $p(\mathbf{s_1}|\mathbf{s_2}) \doteq 2^{-n(H(S_1|S_2) \pm 2\epsilon)}$ for $(\mathbf{s_1}, \mathbf{s_2}) \in A_\epsilon^{(n)}(S_1, S_2)$.

Conditional Typical Sets

- Given an ϵ, define $A_\epsilon^{(n)}(S_1|\mathbf{s_2})$ to be the set of $\mathbf{s_1}$ sequences that are jointly typical with $\mathbf{s_2}$.
- **Theorem:** If $\mathbf{s_2} \in A_\epsilon^{(n)}(S_2)$ then

$$|A_\epsilon^{(n)}(S_1|\mathbf{s_2})| \leq 2^{n(H(S_1|S_2)+2\epsilon)},$$

and

$$(1-\epsilon)2^{n(H(S_1|S_2)-2\epsilon)} \leq \sum_{\mathbf{s_2}} p(\mathbf{s_2})|A_\epsilon^{(n)}(S_1|\mathbf{s_2})|,$$

for sufficiently large n.

Joint Typicality by Chance

- Let S_1, S_2, S_3 be subsets of Ω, with distribution p.
- Let S_1' and S_2' be conditionally independent given S_3'

$$p'(\mathbf{S_1'} = \mathbf{s_1}, \mathbf{S_2'} = \mathbf{s_2}, \mathbf{S_3'} = \mathbf{s_3}) = \prod_i p(s_{1i}|s_{3i})p(s_{2i}|s_{3i})p(s_{3i})$$

- Then

$$p'(\mathbf{S_1'}, \mathbf{S_2'}, \mathbf{S_3'} \in A_\epsilon^{(n)}) \doteq 2^{-n(I(S_1;S_2|S_3)\pm 6\epsilon)}$$

Multiple Access Channel

- A discrete memoryless multiple access channel with two senders and one receiver is defined by three alphabets $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}$ and a transition probability matrix $p(y|x_1, x_2)$.
- A $((2^{nR_1}, 2^{nR_2}), n)$ code consists of two index sets, two encoding functions and one decoding function.
  - Sender 1 chooses an index W_1 from $\{1, 2, \ldots, 2^{nR_1}\}$ and sends the corresponding codeword $X_1^n(W_1)$. Sender 2 does likewise.
  - The decoder decides $\hat{W}_1$ and $\hat{W}_2$ from Y^n.

Achievable Rate Pair

- The average probability of error is defined by

$$P_e^{(n)} = p((\hat{W}_1, \hat{W}_2) \neq (W_1, W_2))$$

  where (W_1, W_2) is uniformly selected from the index sets.

- A rate pair (R_1, R_2) is said to be achievable if there exists a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ codes with $P_e^{(n)} \to 0$.

- The capacity region is the closure of the set of achievable (R_1, R_2) rate pairs.

Capacity Regions of Multiple Access Channel

- Given a multiple access channel

$$(\mathcal{X}_1 \times \mathcal{X}_2, p(y|x_1, x_2), \mathcal{Y}),$$

the capacity region is the closure of the convex hull of all (R_1, R_2) satisfying

$$(\alpha) : \begin{cases} R_1 < I(X_1; Y|X_2), \\ R_2 < I(X_2; Y|X_1), \\ R_1 + R_2 < I(X_1, X_2; Y) \end{cases}$$

for some distribution $p(x_1)p(x_2)$.

Examples of Capacity Regions

- Independent binary symmetric channel
- Binary multiplier channel
- Binary erasure channel

Convexity of Capacity Region

- The capacity region C of a multiple access channel is convex. That is,

$$\text{if } (R_1, R_2) \in C \text{ and } (R_1', R_2') \in C,$$
$$\text{then } (\lambda R_1 + (1-\lambda)R_1', \lambda R_2 + (1-\lambda)R_2') \in C$$

for $0 \leq \lambda \leq 1$.

- This can be shown based on the idea of time division.

Achievability

- If a rate pair (R_1, R_2) satisfies (α) for some distribution,

$$p(x_1, x_2) = p(x_1)p(x_2),$$

  then it is achievable.

- The proof again uses the "standard" random codebook generation and joint typicality decoding. The average probability of error over all codebooks satisfies

$$P_e^{(n)} \leq p(E_{11}^c) + 2^{nR_1} 2^{-n(I(X_1;Y|X_2)-3\epsilon)} + 2^{nR_2} 2^{-n(I(X_2;Y|X_1)-3\epsilon)} + 2^{n(R_1+R_2)} 2^{-n(I(X_1,X_2;Y)-4\epsilon)}$$

  where $E_{ij} = \{(\mathbf{X}_1(i), \mathbf{X}_2(j), \mathbf{Y}) \in A_\epsilon^{(n)}\}$

Gaussian Multiple Access Channels

- We can extend the previous result to a Gaussian channel,

$$Y_i = X_{1i} + X_{2i} + Z_i.$$

In this case, since

$$I(X_1; Y|X_2) = h(Y|X_2) - h(Y|X_1, X_2) \leq \frac{1}{2}\log(1 + \frac{P_1}{N}),$$

we have

$$R_1 \leq C(\frac{P_1}{N}).$$

Similarly,

$$R_2 \leq C(\frac{P_2}{N}); \;\; R_1 + R_2 \leq C(\frac{P_1 + P_2}{N}).$$

A Few Comments

- If we generalize to m-senders and let $m \to \infty$, then the sum of information rates also goes to ∞ but the rate per user goes to 0.
- In frequency-division multiplexing, the power allocation should be proportional to the bandwidth.
- The capacity region is larger than that achieved by time division or frequency division.
- An equivalent of Gauss's law

$$\sum_{i \in S} R_i \leq C\left(\frac{\sum_{i \in S} P_i}{N}\right).$$

Encoding of Correlated Sources

- Suppose there are two sources X and Y. To encode X with vanishing probability, a rate of $R > H(X)$ suffices. Similarly, a rate of $R > H(Y)$ suffices to encode Y.
- A rate of $R > H(X, Y)$ suffices if we encode X and Y jointly.
- What if the X-source and Y-source must be separately (distributedly) encoded?
- Surprisingly, it is still $R > H(X, Y)$ instead of $R > H(X) + H(Y)$.

Distributed Source Code

- A $((2^{nR_1}, 2^{nR_2}), n)$ distributed source code for the joint source (X, Y) consists of two encoder functions,

$$f_1 : \mathcal{X}^n \to \{1, 2, \ldots, 2^{nR_1}\}$$
$$f_2 : \mathcal{Y}^n \to \{1, 2, \ldots, 2^{nR_2}\}$$

  and one decoder function,

$$g : \{1, 2, \ldots, 2^{nR_1}\} \times \{1, 2, \ldots, 2^{nR_2}\} \to \mathcal{X}^n \times \mathcal{Y}^n.$$

- $f_1(X^n)$ is the index for X^n. (R_1, R_2) is called the rate pair of the code.

Achievable Rate Region

- The probability of error for a distributed source code is defined by

$$P_e^{(n)} = p(g(f_1(X^n), f_2(Y^n)) \neq (X^n, Y^n)).$$

- A rate pair (R_1, R_2) is achievable if there exists a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ distributed source codes with $P_e^{(n)} \to 0$.

- The achievable rate region is the closure of the set of achievable rate pairs.

Slepian-Wolf Theorem

- For (X, Y) drawn i.i.d. $\sim p(x, y)$, the achievable rate region is given by

$$(\beta) : \begin{cases} R_1 \geq H(X|Y) \\ R_2 \geq H(Y|X) \\ R_1 + R_2 \geq H(X, Y) \end{cases}$$

- Examples
  - Gotham and Metropolis weather
  - Example 14.4.2.

Random-Bin Source Codes

- For each sequence Z^n, draw an index at random from $\{1, 2, \ldots, 2^{nR}\}$. The set of Z^n's with the same index form a bin. (We are basically throwing Z^n into one of the 2^{nR} bins at random.)
- To decode from the bin index, we look for a typical sequence Z^n in the bin. If there is only one such Z^n, then Z^n is the decoded output. Otherwise, an error is declared.
- If $R > H(Z)$, then the error probability vanishes.
  - There are exponentially more bins than typical sequences, so each bin has zero or one typical sequence.

Distributed Random-Bin Source Codes

- We use the idea of random-bin source codes to show the achievability of Slepian-Wolf theorem. Specifically, given 2^{nR_1} bins for $\mathcal{X}^n$ and 2^{nR_2} bins for $\mathcal{Y}^n$,
  - codebook: independently assign each $\mathbf{x} \in \mathcal{X}^n$ to a bin. Similarly for each $\mathbf{y} \in \mathcal{Y}^n$. Now we have $f_1(\mathbf{x})$, $f_2(\mathbf{y})$.
  - encoding: encode $\mathbf{X}, \mathbf{Y}$ by $f_1(\mathbf{X}), f_2(\mathbf{Y})$.
  - decoding: given (i, j) declare $g(i, j) = (\mathbf{x}, \mathbf{y})$ if there is only one pair of sequences $(\mathbf{x}, \mathbf{y})$ such that $f_1(\mathbf{x}) = i$, $f_2(\mathbf{y}) = j$ and $(\mathbf{x}, \mathbf{y}) \in A_\epsilon^{(n)}$.

Achievability

- A rate pair satisfying (β) is achievable.
  - Define the events of error

$$
\begin{aligned}
E_0 &= \{(\mathbf{X}, \mathbf{Y}) \notin A_\epsilon^{(n)}\}, \\
E_1 &= \{\exists \mathbf{x}' \neq \mathbf{X} : f_1(\mathbf{x}') = f_1(\mathbf{X}) \text{ and } (\mathbf{x}', \mathbf{Y}) \in A_\epsilon^{(n)}\}, \\
E_2 &= \{\exists \mathbf{y}' \neq \mathbf{Y} : f_2(\mathbf{y}') = f_2(\mathbf{Y}) \text{ and } (\mathbf{X}, \mathbf{y}') \in A_\epsilon^{(n)}\}, \\
E_3 &= \{\exists (\mathbf{x}', \mathbf{y}') \neq (\mathbf{X}, \mathbf{Y}) : f_1(\mathbf{x}') = f_1(\mathbf{X}), f_2(\mathbf{y}') = f_2(\mathbf{Y}), \\
&\qquad (\mathbf{x}', \mathbf{y}') \in A_\epsilon^{(n)}\},
\end{aligned}
$$

$$
\begin{aligned}
P_e^{(n)} &= p(E_0 \cup E_1 \cup E_2 \cup E_3) \leq p(E_0) + p(E_1) + p(E_2) + p(E_3) \\
&\to 0
\end{aligned}
$$

(For example, see (14.165)-(14.169) for $p(E_1)$.)

Converse

- If (R_1, R_2) is an achievable pair, then (β) is satisfied.
- The proof is based on the Fano's inequality.

$$\begin{aligned} n(R_1 + R_2) &\geq H(I, J) \\ &\geq I(X^n, Y^n; I, J) = H(X^n, Y^n) - H(X^n, Y^n | I, J) \\ &\geq H(X^n, Y^n) - n\epsilon_n = nH(X, Y) - n\epsilon_n \end{aligned}$$

- The rate region is illustrated in Figure 14.20.
- The Slepian-Wolf theorem can be extended to
  - multiple sources (theorem 14.4.2)
  - non-i.i.d. correlated sources

Interpretation

- Consider the point where $R_1 = H(X), R_2 = H(Y|X)$. It is easy to encode X^n with $nH(X)$ bits, but how do we encode Y^n with $nH(Y|X)$ bits?
  - Known, if the Y encoder knows X^n. But it does not.
- The Y encoder randomly colors all Y^n sequences with 2^{nR_2} colors. If the number of colors is high enough, all the colors of Y^n sequences jointly typical with X^n will be different so the color implies identity.

Duality

- With a multiple access channel, we consider the problem of sending independent messages over a channel.
- With Slepian-Wolf encoding, we consider the problem of sending correlated sources over a noiseless channel, with a common decoder to recover both sources.
- For vanishing error probability, the information rates R are respectively constrained by the channel capacity I and the entropy rates H based on the idea of codeword encoding and joint typicality decoding.

Broadcast Channel

- A diagram is shown in Figure 14.24.
- One sender, sending messages (W_1, W_2) respectively to two (or more) receivers.
- The basic problem is to find the set of simultaneously achievable rates.
- In general, we want to arrange things so that different receivers receive different amount of informations.

Examples

- TV station: to send signals such that HDTV users receive the high definition signal and normal users receive regular signal.
- Classroom: to lecture in a way such that good students learn extra points while poor students learn the basic. To pace with the slowest learner is not good.
- Spanish and Dutch speakers: interesting point. See text. The ordering of information designated for different receivers encodes extra information.

Definition for A Broadcast Channel

- A broadcast channel with two receivers consists of an input alphabet, two output alphabets and a probability transition function.

$$\mathcal{X},\ \mathcal{Y}_1,\ \mathcal{Y}_2,\ p(y_1, y_2|x)$$

- A broadcast channel is said to be memoryless if

$$p(y_1^n, y_2^n|x^n) = \prod_{i=1}^{n} p(y_{1i}, y_{2i}|x_i)$$

Code for Broadcast Channels

- A $((2^{nR_1}, 2^{nR_2}), n)$ code for a broadcast channel with independent information consists of an encoder,

$$f : \{1, 2, \ldots, 2^{nR_1}\} \times \{1, 2, \ldots, 2^{nR_2}\} \to \mathcal{X}^n,$$

and two decoders,

$$g_1 : \mathcal{Y}_1^n \to \{1, 2, \ldots, 2^{nR_1}\}; \;\; g_2 : \mathcal{Y}_2^n \to \{1, 2, \ldots, 2^{nR_2}\}.$$

- The probability of error is defined by

$$P_e^{(n)} = p(\{g_1(Y_1^n) \neq W_1\} \cup \{g_2(Y_2^n) \neq W_2\}),$$

where W_1 and W_2 are uniformly distributed.

Broadcast Channel with Common Information

- A $((2^{nR_0}, 2^{nR_1}, 2^{nR_2}), n)$ code for a broadcast channel with common information consists of an encoder,

$$f : \{1, 2, \ldots, 2^{nR_0}\} \times \{1, 2, \ldots, 2^{nR_1}\} \times \{1, 2, \ldots, 2^{nR_2}\} \to \mathcal{X}^n,$$

  and two decoders,

$$g_1 : \mathcal{Y}_1^n \to \{1, 2, \ldots, 2^{nR_0}\} \times \{1, 2, \ldots, 2^{nR_1}\};$$

$$g_2 : \mathcal{Y}_2^n \to \{1, 2, \ldots, 2^{nR_0}\} \times \{1, 2, \ldots, 2^{nR_2}\}.$$

- The probability of error is defined by

$$P_e^{(n)} = p(\{g_1(Y_1^n) \neq (W_0, W_1)\} \cup \{g_2(Y_2^n) \neq (W_0, W_2)\}),$$

  where W_0, W_1 and W_2 are uniformly distributed.

Achievable Rates and Capacity Region

- A rate pair (R_1, R_2) is said to be **achievable** for a broadcast channel with independent information if there exists a sequence of $((2^{nR_1}, 2^{nR_2}), n)$ codes such that $P_e^{(n)} \to 0$.
- Similarly, a rate triple (R_0, R_1, R_2) is said to be achievable for a broadcast channel with common information if there exists a sequence of $((2^{nR_0}, 2^{nR_1}, 2^{nR_2}), n)$ codes such that $P_e^{(n)} \to 0$.
- The **capacity region** of a broadcast channel is the closure of the set of achievable rates.

Degraded Broadcast Channels

- A broadcast channel is said to be physically degraded if

$$p(y_1, y_2|x) = p(y_1|x)p(y_2|y_1).$$

- A broadcast channel is said to be statistically degraded if there exists a distribution $p'(y_2|y_1)$ such that

$$p(y_2|x) = \sum_{y_1} p(y_1|x)p'(y_2|y_1).$$

- A physically degraded broadcast channel is stochastically degraded, since

$$p(y_2|x) = \sum_{y_1} p(y_1, y_2|x) = \sum_{y_1} p(y_1|x)p(y_2|y_1).$$

Capacity Region for Degraded Broadcast Channel

- The capacity region for a degraded broadcast channel $X \to Y_1 \to Y_2$ (a Markov chain) is the convex hull of all rate pairs satisfying

$$(\gamma): \quad \begin{cases} R_2 \leq I(U;Y_2), \\ R_1 \leq I(X;Y_1|U) \end{cases}$$

for some joint distribution $p(u)p(x|u)p(y_1, y_2|x)$, where the auxiliary random variable U has a constraint on cardinality $|\mathcal{U}| \leq \min(|\mathcal{X}|, |\mathcal{Y}_1|, |\mathcal{Y}_2|)$.

Achievability

- A rate pair satisfying (γ) for a degraded broadcast channel is achievable, by the following coding scheme.
  - random codebook generation

$$\mathbf{U}(w_2) \sim \prod_{i=1}^{n} p(u_i); \quad \mathbf{X}(w_1, w_2) \sim \prod_{i=1}^{n} p(x_i|u_i(w_2))$$

  - encoding: $(W_1, W_2) \rightarrow \mathbf{X}(W_1, W_2)$.
  - decoding:

receiver 2: $\tilde{W}_2$ such that $(\mathbf{U}(\tilde{W}_2), \mathbf{Y}_2) \in A_\epsilon^{(n)}$.

receiver 1: $(\hat{W}_1, \hat{W}_2)$ such that $(\mathbf{U}(\hat{W}_2), \mathbf{X}(\hat{W}_1, \hat{W}_2), \mathbf{Y}_1) \in A_\epsilon^{(n)}$.

Probability of Error

- Assuming that $(w_1, w_2) = (1, 1)$.
- The channel from U to Y_2 is essentially a single user channel, so for receiver 2,

$$R_2 \leq I(U; Y_2) \Rightarrow P_e^{(n)}(2) \to 0.$$

- For receiver 1, define the following events

$$E_{Y_i} = \{(\mathbf{U}(i), \mathbf{Y}_1) \in A_\epsilon^{(n)}\}, E_{Y_{ij}} = \{(\mathbf{U}(i), \mathbf{X}(j, i), \mathbf{Y}_1) \in A_\epsilon^{(n)}\},$$

$$\begin{aligned} P_e^{(n)}(1) &= p(E_{Y_1}^c \cup (\cup_{i \neq 1} E_{Y_i}) \cup (\cup_{j \neq 1} E_{Y_{1j}})) \\ &\leq p(E_{Y_1}^c) + p(\cup_{i \neq 1} E_{Y_i}) + p(\cup_{j \neq 1} E_{Y_{1j}}) \end{aligned}$$

The second term is exponentially small as $R_2 \leq I(U; Y_1)$, so is the third term as $R_1 \leq I(X; Y_1 | U)$.

Common Information

- If (R_1, R_2) is achievable for a degraded broadcast channel, then $(R_0, R_1 - R_0, R_2 - R_0)$ is achievable, provided that $R_0 \leq \min(R_1, R_2)$.
- If (R_1, R_2) is achievable for a degraded broadcast channel, then $(R_0, R_1, R_2 - R_0)$ is achievable, provided that $R_0 \leq R_2$.
  - This is so because the better receiver always decodes the information for the worse receiver.

Binary Symmetric Broadcast Channel

- We have a pair of binary symmetric channels and we want to compute the capacity region for this channel.
- This broadcast channel is equivalent to a cascade of BSCs as shown in Figure 14.27.
- The capacity region is computed via the addition of another BSC.
- The result is shown in Figure 14.28.

The Relay Channel

- One sender and one reciever with a number of intermediate nodes which act as relays to help the communication between the sender and the receiver.
- In the simplest case, there is one relay node. There are four alphabet sets $\mathcal{X}, \mathcal{X}_1, \mathcal{Y}, \mathcal{Y}_1$ and one probability function $p(\cdot, \cdot | x, x_1)$.
- As shown in Figure 14.30, a relay channel consists of one broadcast channel and one multi-access channel.

Code for A Relay Channel

- A $(2^{nR}, n)$ code for a relay channel consists of an encoding function,

$$\{1, 2, \ldots, 2^{nR}\} \to \mathcal{X}^n$$

a set of relay functions,

$$X_{1i} = f_i(Y_{11}, \ldots, Y_{1i-1})$$

and a decoding function,

$$\mathcal{Y}^n \to \{1, 2, \ldots, 2^{nR}\}.$$

Capacity for A Relay Channel

- The error event is the event that the decoded message $\hat{W} = g(Y)$ is not equal to the sent message W.
- An rate R is achievable if there exists a sequence of $(2^{nR}, n)$ codes such that the average probability of error $P_e^{(n)} \to 0$.
- The capacity of a relay channel is the supremum of the set of achievable rates.
- An upper bound on the capacity is

$$C \leq \sup_{p(x,x_1)} \min(I(X, X_1; Y), I(X; Y, Y_1|X_1)).$$

Multi-Terminal Networks

- There are m nodes. Node i has a transmission variable $X^{(i)}$ and a receiving variable $Y^{(i)}$. It sends messages at rate $R^{(ij)}$ to node j. We assume that the messages sent from node i to node j is independent and uniformly distributed over the range $\{1, \ldots, 2^{nR^{(ij)}}\}$.
- The channel is represented by the transition function

$$p(y^{(1)}, \ldots, y^{(m)}|x^{(1)}, \ldots, x^{(m)}).$$

It is assumed to be memoryless.

Rate of Flow of Information

- If $\{R^{(ij)}\}$ are achievable over a network, then there exists a joint distribution $p(x^{(1)}, \ldots, x^{(m)})$ such that

$$\sum_{i \in S, j \in S^c} R^{(ij)} \leq I(X^{(S)}; Y^{(S^c)} | X^{(S^c)}),$$

  for all $S \subseteq \{1, 2, \ldots, m\}$. In other words, the total rate of flow of information across a cut-set is bounded by the conditional mutual information.

- The converse, however, is not generally true.