Some Aspects of Finite State Channel related to Hidden Markov Process

Kingo Kobayashi

National Institute of Information and Communications Tecnology(NICT),
4-2-1, Nukui-Kitamachi, Koganei, Tokyo, 184-8795, JAPAN
E-mail: kingo@nict.ac.jp

Abstract. We have no satisfactory capacity formula for most channels with finite states. Here, we consider some interesting examples of finite state channels, such as Gilbert-Elliot channel, trapdoor channel, etc., to reveal special characters of problems and difficulties to determine the capacities. Meanwhile, we give a simple expression of the capacity formula for Gilbert-Elliot channel by using a hidden Markov source for the optimal input process. This idea should be extended to other finite state channels.

1 Gilbert-Elliot channel

Let us start to consider the commom Gilbert-Elliot channel as a typical example of finite state channel. The channel shown in Figure 1 is a standard model of burst error channel. Thus, it is a finite state channel with two states, good(G) and bad(B) states. In any state, the channel is the binary symmetric channel, but has different cross-over probabilities, δ and ε. It is usually assumed that the cross-over probability δ of good(G) state is significantly smaller than $\varepsilon\,(\leq 1/2)$ of bad(B) state.

The channel changes its state in Markovian way. The state transition probability matrix is expressed as

$$T = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}, \tag{1}$$

and its stationary distribution is

$$\left(\frac{b}{a+b} \cdot \frac{a}{a+b} \right). \tag{2}$$

The probability of state sequence $\boldsymbol{s} = s_1 s_2 \ldots s_n, s_i \in \{G, B\}$ of length n is given by

$$P(\boldsymbol{s}) = p_{s_1} \overline{a}^{N_{GG}} a^{N_{GB}} \overline{b}^{N_{BB}} b^{N_{BG}}, \tag{3}$$

where p_{s_1} is the stationary probability of s_1, and N_{st} is the number of transitions from state s to state t, $\overline{a} = 1-a, \overline{b} = 1-b$.

Dagstuhl Seminar Proceedings 09281
Search Methodologies
http://drops.dagstuhl.de/opus/volltexte/2009/2243

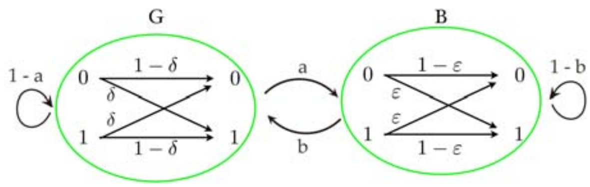


Fig. 1. Glbert-Elliott Channel

Then, we can express the conditional probability of output sequence $\boldsymbol{y} = y_1 y_2 \ldots y_n, y_i \in \{0,1\}$ given the input sequence $\boldsymbol{x} = x_1 x_2 \ldots x_n, x_i \in \{0,1\}$ by the formula,

$$P(\boldsymbol{y}|\boldsymbol{x}) = \sum_{\boldsymbol{s}} P(\boldsymbol{s})\,\overline{\delta}^{N_{G,c}(\boldsymbol{s})}\,\delta^{N_{G,e}(\boldsymbol{s})}\,\overline{\varepsilon}^{N_{B,c}(\boldsymbol{s})}\,\varepsilon^{N_{B,e}(\boldsymbol{s})}, \tag{4}$$

where $N_{S,c}(\boldsymbol{s})$ is the number of i that satisfies $x_i = y_i$, and $N_{S,e}(\boldsymbol{s})$ is the number of i that satisfies $x_i \neq y_i$ at the state S in $\boldsymbol{s}$ From this formula, we can easily derive,

$$P(\boldsymbol{y}|\boldsymbol{x}) = P(\boldsymbol{x} \oplus \boldsymbol{y}|\mathbf{0}) \tag{5}$$

where $\oplus$ is the termwise exclusive-or operation, and $\mathbf{0}$ is the sequence of all zero.

Then, it holds that each row of this channel matrix is a permutation of a distribution, and each column is also a permutation of other columns. In fact, it is a symmetic channel. For example, the density plot of the conditional probabilities of Gilbert-Elliot channel for inputs-outputs of length 6 is shown in Figure 2, where we set $a = 0.5, b = 0.5, \varepsilon = 0.5, \delta = 0.0$, and darker color corresponds to higher probability. Due to the symmetic property of Gilbert-Elliot channel, the capacity is attained by the equiprobable distribution, and the output has also the equiprobable distribution. Moreover, the conditional entropies $H(Y^n|X^n = \boldsymbol{x})$ have a same value for any input $\boldsymbol{x}$, i.e., is equal to $H(Y^n|X^n)$. Thus, without any loss of generality, we can assume that the input sequence is $X^n = \mathbf{0} = 0\ldots0$ for evaluating the conditional entropy $H(Y^n|X^n)$.

Therefore, the maximum of mutual information between input sequence X^n and output sequence Y^n is attained when X^n has the equiprobable distribution.

$$\begin{aligned}
\max_{X^n} \frac{1}{n} I(X^n;Y^n) &= \max_{X^n} \frac{1}{n}\{H(Y^n) - H(Y^n|X^n)\} \\
&= \frac{1}{n}\{H(Y^n) - H(Y^n|X^n)\}\Big|_{X^n:\text{equiprobable}} \\
&= 1 - \frac{1}{n} H(Y^n|X^n = \mathbf{0}) \\
&= 1 - \frac{1}{n} H(Z^n),
\end{aligned} \tag{6}$$

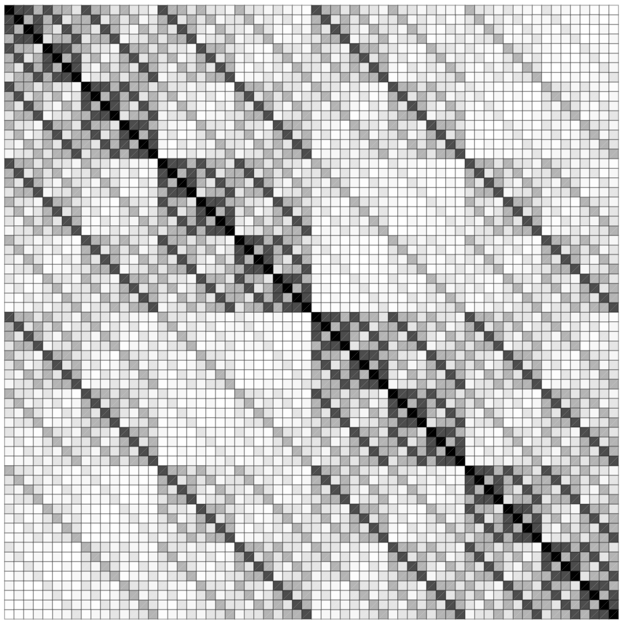


Fig. 2. Glbert-Elliott Channel Probability Matrix for inputs-outputs of length 6

where Z^n is the induced hidden Markov process, i.e., the conditional Y^n given $X^n = \mathbf{0}$. The hidden Markov process $\{Z_n\}$ with four states $G0, B0, G1, B1$ is defined by the state transition matrix W,

$$W = \begin{array}{c c} & \begin{array}{cccc} G0 & B0 & G1 & B1 \end{array} \\ \begin{array}{c} G0 \\ B0 \\ G1 \\ B1 \end{array} & \left[\begin{array}{cc|cc} \overline{a}\overline{\delta} & a\overline{\varepsilon} & \overline{a}\delta & a\varepsilon \\ b\overline{\delta} & \overline{b}\overline{\varepsilon} & b\delta & \overline{b}\varepsilon \\ \hline \overline{a}\overline{\delta} & a\overline{\varepsilon} & \overline{a}\delta & a\varepsilon \\ b\overline{\delta} & \overline{b}\overline{\varepsilon} & b\delta & \overline{b}\varepsilon \end{array}\right] \end{array}, \tag{7}$$

where the process in the state Sx ($S \in \{G, B\}, x \in \{0, 1\}$) produces the output symbol x.

Due to the equation (6), we have by using the above matrix W

Theorem 1. *The capacity of Gilbert-Elliott Channel is expressed as,*

$$C_{\mathrm{GEC}}(a, b, \delta, \varepsilon) = 1 - H(Z), \tag{8}$$

where $H(Z)$ is the entropy of the hidden Markov source Z defined by W.

Remark 1. In their paper[1], Mushkin and Bar-David showed that the capacity of Gilber-Elliot channel is $C_{\mathrm{GEC}} = 1 - \lim_{n\to\infty} \mathrm{E}[h(q_n)] = 1 - \mathrm{E}[h(q_\infty)]$, where

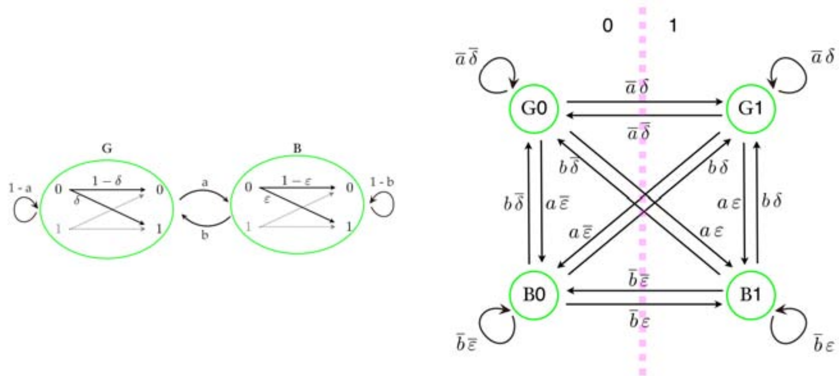


Fig. 3. The mechanism of generating the output Y given $X = 0$, and the transition matrix of induced hidden Markov process Z

$q_n = \Pr\{Z_n = 0|Z_1 \dots Z_{n-1}\}, Z_n = X_n \oplus Y_n$, the input process $\{X_n\}$ is the Bernoulli process with equiprobability, and the random variable q_n converges to the limiting random variable q_∞ in distribution. In our context, the limiting quantitiy $\lim_{n\to\infty} \mathrm{E}[h(q_n)] = \mathrm{E}[h(q_\infty)]$ is actually the entropy of the hidden Markov process $\{Z_n\}$ induced by the stochastic matrix W explicitly defined in (7).

Remark 2. Due to the fact that the equiprobable Bernoulli process suffices to attain the maximum mutual information between input and output processes, we need not to have any concern on the channel state in constructing the code and during transmitting messages, so long as we cannot have any side information about the state.

Here are questions: What is the distribution of q_∞ as a function of the parameters $a, b, \delta, \varepsilon$? How can the base set of q_∞ be expressed when $a, b, \delta, \varepsilon$ will take various values ? We have no analytically closed expression for the entropy of hidden Markov source in general. In some special cases, we can give simple formulae of capacity of Gilbert-Elliot channel.

Consider the case when $\delta = 0, \varepsilon = 1$. Then, the stocastic matrix W of (7) is deduced as

$$W = \begin{array}{c} \\ G0 \\ B0 \\ G1 \\ B1 \end{array} \begin{array}{c} \begin{array}{cccc} G0 & B0 & G1 & B1 \end{array} \\ \left[\begin{array}{cc|cc} \bar{a} & 0 & 0 & a \\ b & 0 & 0 & \bar{b} \\ \hline \bar{a} & 0 & 0 & a \\ b & 0 & 0 & \bar{b} \end{array} \right] \end{array} . \tag{9}$$

Considering that the stationary distribution of good(G) and bad(B) is $(b/(a+b), a/(a+b))$, we have a simple formula for this case:

$$C_{\mathrm{GEC}}(a, b, 0, 1) = 1 - \left(\frac{b}{a+b} h(\bar{a}) + \frac{a}{a+b} h(b) \right), \tag{10}$$

where $\overline{x} = 1-x$ and $h(p) = -p\log_2 p - \overline{p}\log_2 \overline{p}$ is the binary entropy function(cf. Fig.4). When $a = b = 1/2$ in this case, the capacity is zero as expected. But otherwise, we will have positive capacity.

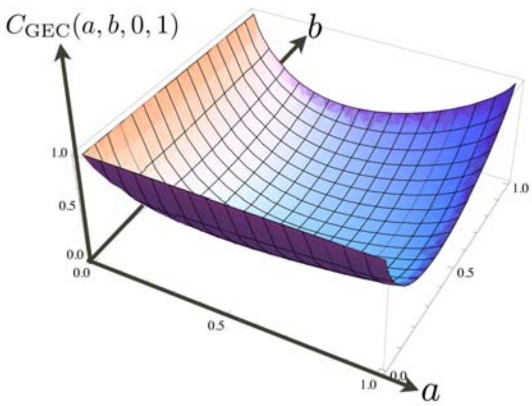


Fig. 4. $C_{GEC}(a,b,0,1)$

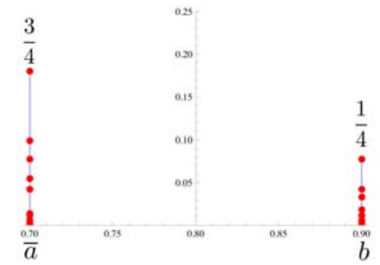


Fig. 5. The distribution of $q_k = \Pr\{Z_k = 0|Z_1 \ldots Z_{k-1}\}$ when $k = 6, a = 0.3, b = 0.9, \delta = 0, \varepsilon = 1$.

Moreover, $q_n = \Pr\{Z_n = 0|Z_1 \ldots Z_{n-1}\}$ takes only two values $\overline{a}$ and b with probabilities $b/(a+b)$ and $a/(a+b)$, respectively (cf. Figure 5). This is the reason why in the formula (10), do only two terms containing the binary entropy function appear.

Next, let us consider another interesting case, that is, when $a+b=1$. Then, the stocastic matrix W of (7) is deduced as

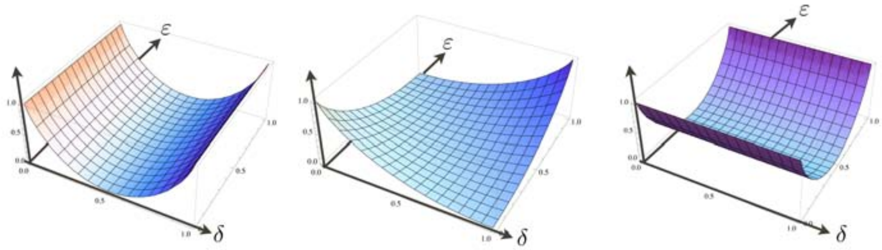
$$W = \begin{array}{c} \\ G0 \\ B0 \\ G1 \\ B1 \end{array} \begin{array}{c} \begin{array}{cccc} G0 & B0 & G1 & B1 \end{array} \\ \left[\begin{array}{cc|cc} \overline{a}\overline{\delta} & a\overline{\varepsilon} & \overline{a}\delta & a\varepsilon \\ \overline{a}\overline{\delta} & a\overline{\varepsilon} & \overline{a}\delta & a\varepsilon \\ \hline \overline{a}\overline{\delta} & a\overline{\varepsilon} & \overline{a}\delta & a\varepsilon \\ \overline{a}\overline{\delta} & a\overline{\varepsilon} & \overline{a}\delta & a\varepsilon \end{array}\right] \end{array}. \tag{11}$$

Thus, the stochastic matrix is completely degenerated, and the capacity is easily expressed as,

$$C_{\mathrm{GEC}}(a, \overline{a}, \delta, \varepsilon) = 1 - h(\overline{a}\,\overline{\delta} + a\,\overline{\varepsilon}). \tag{12}$$

In Figure 6, the profile of capacity $C_{\mathrm{GEC}}(a, \overline{a}, \delta, \varepsilon)$ for three values of a. In particular, when $a = b = \frac{1}{2}, \delta = 1$, we have

$$C_{\mathrm{GEC}}(1/2, 1/2, 1, \varepsilon) = 1 - h\left(\frac{\overline{\varepsilon}}{2}\right). \tag{13}$$



$a=0$ $a=0.5$ $a=1$

Fig. 6. $C_{\mathrm{GEC}}(a, \overline{a}, \delta, \varepsilon)$ for $a = 0, 0.5, 1$

Moreover, $q_n = \Pr\{Z_n = 0 | Z_1 \dots Z_{n-1}\}$ concentrates on only one values $\overline{a}\,\overline{\delta} + a\,\overline{\varepsilon}$ with probability one (cf. Figure 7). This corresponds to the fact in the formula (12), only one term containing the binary entropy function appears.

For more general cases, we cannot obtain simple formulae. However, in order to compute the entropy of the hidden Markov process in any desired precision, we can use the elegant technique described in the book of Cover and Thomas[2]. Figure 8 shows an example of the entropy calculation for case of $a = 0.1, b = 0.1, \delta = 0.01, \varepsilon = 0.5$. It can be seen that the upper bound $H(Z_k|Z_1 \dots Z_{k-1})$ and the lower bound $H(Z_k|S_1 Z_1 \dots Z_{k-1})$ almost converge to a same value 0.737537 until the sixth extension of input-output symbols. Thus, we can estimate the capacity as

$$C_{\mathrm{GEC}}(0.1, 0.1, 0.01, 0.5) \doteq 1 - 0.737537 = 0.262463.$$

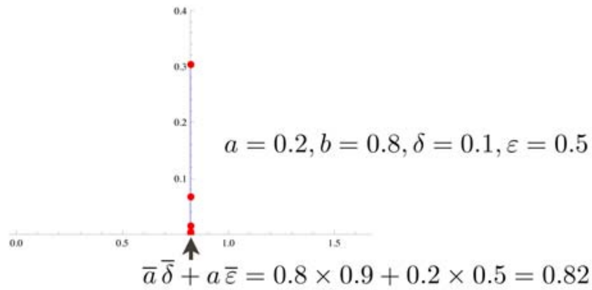


Fig. 7. The distribution of $q_k = \Pr\{Z_k = 0 | Z_1 \ldots Z_{k-1}\}$.

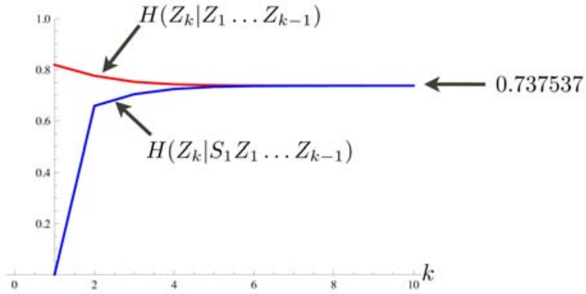


Fig. 8. Computation of the capacity of Gilbert-Elliot channel for any $a, b, \delta, \varepsilon$

For this case, the distribution of $q_n = \Pr\{Z_n = 0 | Z_1 \ldots Z_{n-1}\}$ spreads over wide range as in Figure 9. This suggests that there is no simple analytical expression of capacity of general Gilbert-Elliot channel. We should look for an integral form or a formula containing infinite sum of binary entropy functions.

2 Trapdoor Channel

The capacity problem of trapdoor channel is one of famous long-standing problems. The trapdoor channel considered by Blackwell[3] as a typical example of channel with memory. Ash[4] expressed the channel by using two channel matrices with four states, while the expression does not necessarily make the problem tractable.

The actions of the trapdoor channel are described as follows (cf. Figure 10). The input alphabet $\mathcal{X}$ and the output alphabet $\mathcal{Y}$ are binary. The channel has two trapdoors. Initially, there is a symbol $s \in \{0, 1\}$ called the *initial* symbol on one of trapdoors, and no symbol on another trapdoor. The initial symbol takes the value 0 or 1 with equal probability. Just after the first input symbol x_1 placed on the empty trapdoor, only one of trapdoors will be selected with

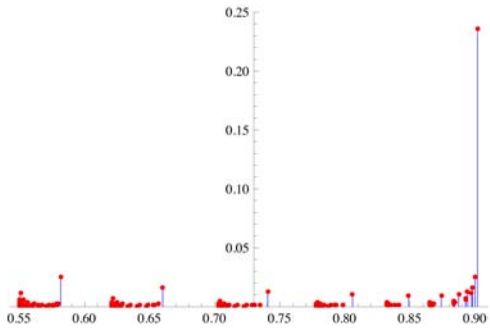


Fig. 9. The distribution of $q_k = \Pr\{Z_k = 0 | Z_1 \dots Z_{k-1}\}$.

probability of one half, and open. Then, the symbol on the opened door falls to become an output symbol $y_1 \in \{s, x_1\}$. After the door has been closed, we are back to the same situation as at the initial instant, but there is a symbol s or x_1 on the non-empty door depending on the output $y_1 = x_1$ or $y_1 = s$, respectively. This process is repeated until an output sequence $y_1 y_2 \dots y_n$ has emitted from the channel for the input sequence $x_1 x_2 \dots x_n$. These channel actions is summarized in the state diagram with arrows having input /output information and transition probability (cf. Figure 11).

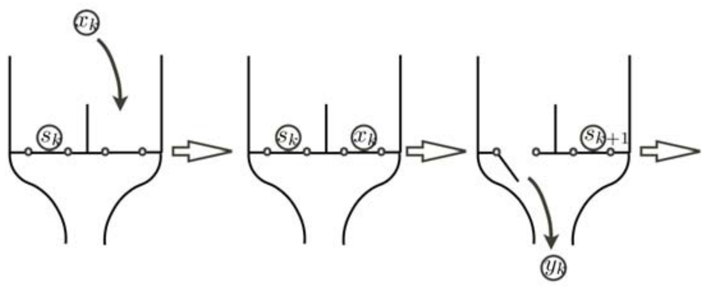


Fig. 10. Actions of Trapdoor Channel

Let $P_{n|s}(\mathbf{x}, \mathbf{y})$ be the conditional probability of output sequence $\mathbf{y}$ by n equiprobable trapdoor actions for the input sequence $\mathbf{x}$ with the initial symbol s. Then, we can show that the conditional probablity matrices obey the following recursions:

$$P_{n+1|0} = \begin{bmatrix} P_{n|0} & 0 \\ \frac{1}{2} P_{n|1} & \frac{1}{2} P_{n|0} \end{bmatrix}, \tag{14}$$

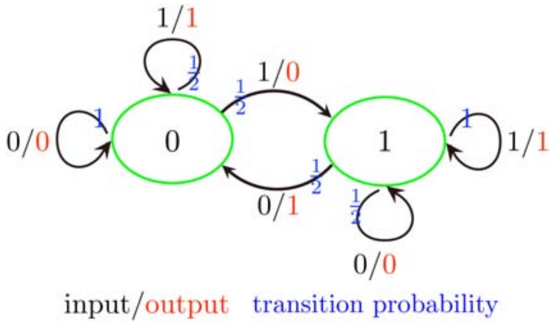


Fig. 11. State Diagram of Trapdoor Channel

and

$$P_{n+1|1} = \begin{bmatrix} \frac{1}{2}P_{n|1} & \frac{1}{2}P_{n|0} \\ 0 & P_{n|1} \end{bmatrix}, \tag{15}$$

where the initial matrices are defined as

$$P_{0|0} = P_{0|1} = [1]\,. \tag{16}$$

If the encoder and decoder do not know the initial symbol at the starting point, we have to consider the mixed channel $P_n = 1/2(P_{n|0} + P_{n|1})$. The mixed channel has an interesting fractal structure. Figure 12 shows a density plot of the extended channel matrix P_6 for input and output sequences of length 6. Here, the darker color corresponds to higher probability.

We could obtain the optimum distribution of input sequences with length seven for the seventh extension channel P_7 by using Arimoto-Blahut algorithm. From the distribution, we can get the binary tree with the branching conditional probability reflecting the optimal distribution (see Figure 13). Here, it is important to guess the character of optimal input process attaining the capacity of the trapdoor channel. It seems not to be memoryless source or Markov source of any order. Is it possible to reduce a hidden Markov source or not ? It is an interesting problem to simulate the optimal input process by some hidden Markov source.

By the brute force usage of Arimoto-Blahut algorithm, it was possible to estimate the value of the capacity[6]. Figure 14 shows the first order difference of the maximum mutual information $\max_{X^n} I(X^n;Y^n)$ for $n = 1,\ldots,9$. We can see that the difference $\max\limits_{X^{n+1}} I(X^{n+1};Y^{n+1}) - \max\limits_{X^n} I(X^n;Y^n)$ converges to $0.572\ldots$. Here, it should be noted that the zero-error capacity of trapdoor channel is 0.5, which had been showned by Ahlswede, et.al.[7]. To obtain the complete solution of this problem, we would need an essential step for understanding of the fractal structure of trapdoor channel.

We can generalize the basic trapdoor channel of binary alphabet with two trapdoors. The generalized trapdoor has the input and output alphabets $\mathcal{X} = \mathcal{Y} = \{0, 1, \ldots, \alpha - 1\}$, and $\beta + 1$ trapdoors. The channel state is expressed as

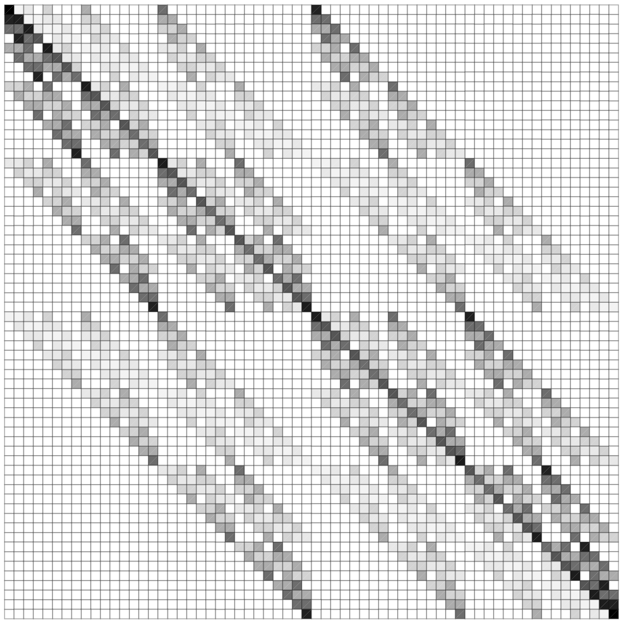


Fig. 12. Density plot of P_6 of Trapdoor channel

$\boldsymbol{s} = (s_0, \ldots, s_{\alpha-1}) : s_i \in \{0, \ldots, \beta\}, \sum_{i=0}^{\alpha-1} s_i = \beta$, where s_i is the number of symbol i arranged on the trapdoors. The trapdoor to be opened is selected equiprobable, i.e., with probability $1/(\beta+1)$. The situation is depicted in the Figure 15.

3 Permuting Relay Channel

Here, we discuss a rare case in which the capacity problem of finite state channel is completely solved. This channel looks like the generalized trapdoor channel(see Figure 15), but there is an extra people, symbol supplier, who continuously supplies symbols from $\mathcal{X} = \{0, \ldots, \alpha - 1\}$ to the room of trapdoors. The message sender can see the stock of symbols placed on the trapdoors, and open any trapdoor to emit a symbol from the stock to the receiver. Thus, by using the supplied symbol resource, he intends to produce a symbol sequence to inform the receiver of his message. The supplier cannot see the state of the room of trapdoors, and the sender is not allowed to request the supplier the symbol wanted. The situation is depicted in Figure 16. The problem is how to provide a symbol sequence for the supplier in order to achieve the maximum communication capability of

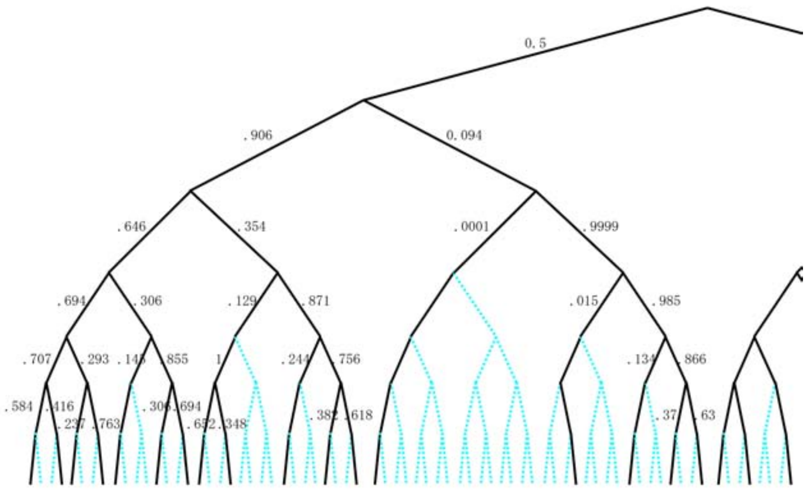


Fig. 13. Tree expression of optimal input source of length 7 for trapdoor channel

the sender. Due to the character of this problem, it is of no use to apply the probabilistic strategy for the sender. He has to emit only sequences that can be uniquely determined by the receiver. We have to count the number of distinct output sequences given a supplied input sequence when the sender would try the whole possible sequences of opening trapdoors.

This problem was solved by Ahlswede and Kaspi[8] when $\alpha \geq 2, \beta = 1$. For general α and β, the complete answer had been given by Kobayashi[9].

As in the generalised trapdoor channel, the state is expressed as $\boldsymbol{s} = (s_0, \ldots, s_{\alpha-1})$: $s_i \in \{0, \ldots, \beta\}, \sum_{i=0}^{\alpha-1} s_i = \beta$. Let $\mathcal{S}$ be the set of states. Here, it should be noted that the number of states is $\gamma = \binom{\alpha+\beta-1}{\beta}$. Starting with an intial state $\boldsymbol{s}_0 \in \mathcal{S}$, the set of distinct output sequences that the sender can construct from an input sequence $\boldsymbol{x} = x_1 \ldots x_n$ $(x_i \in \mathcal{X})$ provided by the supplier is denoted by $\mathcal{Y}(\boldsymbol{s}_0, \boldsymbol{x})$. We denote the maximum code size for length n by $N(n)$, that is,

$$N(n) = \max_{\boldsymbol{s}_0, \boldsymbol{x}} |\mathcal{Y}(\boldsymbol{s}_0, \boldsymbol{x})|. \tag{17}$$

When $\alpha = 2, \beta = 1$, we can easily show the recursion $N(k+2) = N(k+1) + N(k)$ as established by Ahlswede and Kaspi[8]. Thus, $\{N(n)\}$ is just the Fibonatti sequence.

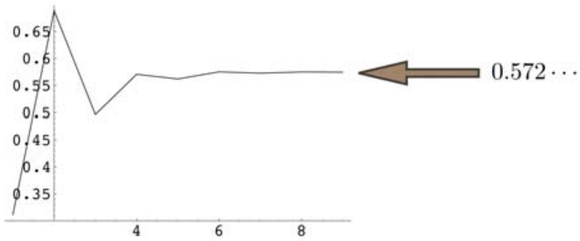


Fig. 14. the first order difference of the maximum mutual information

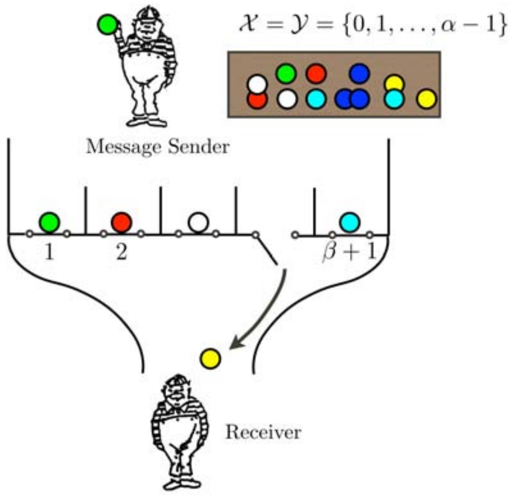


Fig. 15. Generalized Trapdoor Channel

Furthermore, the capacity of the permuting relay channel is defined by

$$C_{\mathrm{PRC}}(\alpha, \beta) = \limsup_{n \to \infty} \frac{1}{n} \log N(n). \tag{18}$$

Thus, when $\alpha = 2, \beta = 1$, $C_{\mathrm{PRC}}(2, 1) = \log_\alpha \phi(\text{Golden Ratio}) = 0.69424\ldots$. Therefore, we can say that about 30% is lost by the underlying relay mechanism.

To express the capacity formula, let us introduce two kinds of $\gamma \times \gamma$ matrices. Let $A^{(x)}$ be the state transition matrix when symbol $x \in \mathcal{X}$ is provided by the supplier. Thus,

$$a^{(x)}_{\boldsymbol{s},\boldsymbol{t}} = \begin{cases} 1, & \text{from } \boldsymbol{s} \text{ the sender can move to } \boldsymbol{t} \text{ when provided } x, \\ 0, & \text{otherwise.} \end{cases} \tag{19}$$

We write simply A for $A^{(0)}$. Next, let us denote by P the permutation on $\mathcal{S}$ defined as follows:

$$P : \boldsymbol{s} = (s_0, s_1, \ldots, s_{\alpha-1}) \to \boldsymbol{t} = (s_1, \ldots, s_{\alpha-1}, s_0). \tag{20}$$

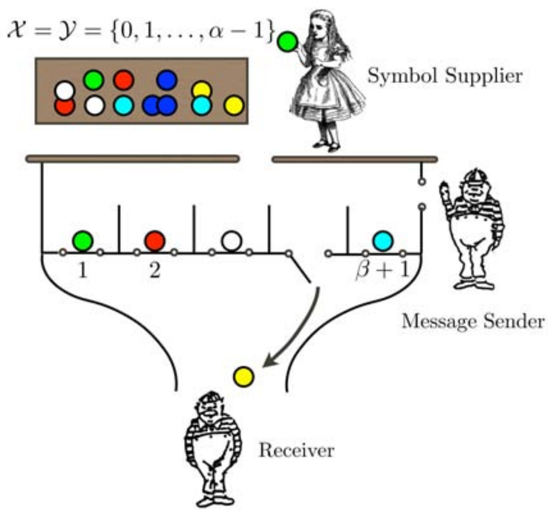


Fig. 16. Permuting Relay Channel

Then, we have the following similarity relations:

$$A^{(k)} = P^k A^{(0)} P^{-k}. \tag{21}$$

Using these matrices A and P, we can obtain

Theorem 2. [9] *The capacity of Permution Relay Channel is expressed as,*

$$C_{\mathrm{PRC}}(\alpha, \beta) = \log\, \lambda(A\, P), \tag{22}$$

where $\lambda(Q)$ is the maximum eigenvalue of matrix Q .

Remark 3. When $\alpha = 3, \beta = 2$, then the number of states is six, and the characteristic polynomial of $Q = AP$ is shown to be $x^6 - x^5 - x^4 - 5x^3 + x^2 + 1$. Thus, we can deduce that the maximum eigenvalue $\lambda(AP)$ is equal to $2.29052\ldots$, and the capacity is $\log_\alpha \lambda(AP) = 0.75438\ldots$. Therefore, we recognize that about 25% is lost by the relay mechanism. Moreover, we have the recursion formula for $N(n)$ as follows:

$$N(k+6) = N(k+5) + N(k+4) + 5N(k+3) - N(k+2) - N(k).$$

4 Markov Erasure Channel

The channel considered in this section has the input alphabet $\mathcal{X} = \{0, 1\}$ and the output alphabet $\mathcal{Y} = \{0, 1, 2\}$. For a special case of this channel, the input-output operation is the following: at the next instant after inputting symbol

0, the output symbol is fixed to be the erasure symbol 2 whatever symbol the current input is. At the next instant after outputting 1 or 2, any input symbol will be correctly transmited to the receiver. The action of this special channel with two states called Fibonatti channel is depicted in Figure 17. As the output sequence is uniquely determined by the input sequence, we have to count distinct output sequences for getting the number of message that can be sent without no error. Let x_n be the number of output sequences ending by the symbol 0, y_n be that ending by symbol 1, and z_n be that ending by symbol 2.

Now, we have recursions as follows:

$$\begin{aligned} x_n &= y_{n-1} + z_{n-1}, \\ y_n &= x_n, \\ z_n &= x_{n-1}. \end{aligned} \tag{23}$$

Then, we can derive a simple recursion:

$$x_n = x_{n-1} + x_{n-2}. \tag{24}$$

The number of distinct output sequences $x_n+y_n+z_n$ is revealed to be just Fibonatti number. And the capacity C_{F} is indeed $\log_2 \phi(\text{Golden Ratio}) = 0.69424\ldots$ bits. Moreover, the optimal input process to the Fibonatti channel should be a Markov source with two states (see Figure 18), where $\phi = \frac{1+\sqrt{5}}{2}$ is the golden ratio.

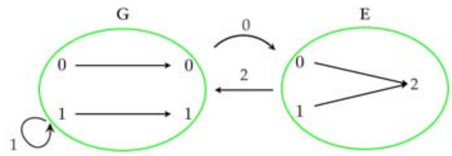


Fig. 17. Fibonatti Channel

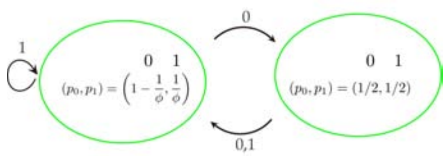


Fig. 18. Optimal source for Fibonatti Channel

The Markov erasure channel with two states, good(G) state and erased(E) state, is an extension of Fibonatti channel, and is shown in Figure 19. Here, the left of symbol pair appended to each arrow denotes the input symbol, and the right of that is the transition probability. The channel is specified by three parameters that control the transition probability between two states. Figure 20 shows an approximated capacity profile of Markov erasure channel when we set $c = 1$, and vary two parameters a and b. We have not yet obtained the precise capacity formula of Markov erasure channel in general, except for some special cases, such as Fibonatti channel.

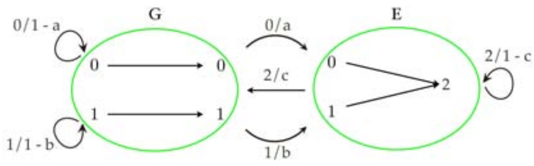


Fig. 19. Markov Erasure Channel

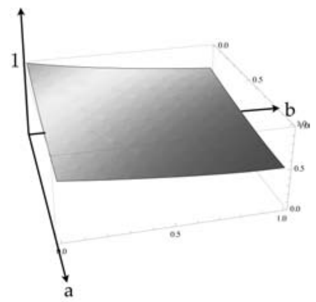


Fig. 20. Approximated capacity profile of Markov erasure channel

5 Conclusion

We have observed the capacity problem of finite state channels by introducing several examples. There are a lot of interesting and important channels for which the capacity problem has not yet been solved. It is very rare that the capacity achieving input is restricted in the class of Markov process. It was very lucky for the Gibert-Elliot channel for which equiprobable Bernoulli process suffices to attain the capacity, and we can have a simple capacity formula by using hidden Markov source. How about for other channels ? At the very least, we desire that we could have a hidden Markov source as the optimal input process. We want to know what class of finite state channels have the property that the optimal input process is enough to be restricted in the class of hidden Markov source.

References

1. M. Mushkin and I. Bar-David, "Capacity and Coding for the Gilbert-Elliot Channels," *IEEE Trans. on Inform. Theory*, vol.35, no.6, pp.1277-1290, 1989.
2. T.M.Cover, J.A.Thomas, *Elements of Information Theory*, John Wiley & Sons, 1991
3. D.Blackwell, Information Theory, in *Modern Mathematics for the Engineer*, Second Series, E.F.Beckenbach, Ed., McGraw-Hill, 1961.
4. R.B.Ash, Information Theory, Interscience Publishers, 1965 and Dover Publications, Inc., 1990.
5. K.Kobayashi, H.Morita and M.Hoshi, "An Input/Output Recursion for the Trapdoor Channel," *Proceedings of IEEE International Symp. on Inform. Theory*, p.423, July, 2002.
6. K.Kobayashi, H.Morita and M.Hoshi, "Some Considerations on Trapdoor Channel," Proceedings of the third Asian-European Workshop on Information Theory, pp.9-10, June, 2003.
7. R.Ahlswede, Jian-Ping Ye and Zhen Zhang, "Creating Order in Sequence Spaces with Simple machines," *Information and Computation*, Vol. 89, No.1, pp.47–94, 1990.
8. R.Ahlswede and A.Kaspi, "Optimal coding strategies for certain permuting channels," *IEEE Trans. on Inform. Theory*, vol.IT-33, no.3, pp.310-314, May 1987.
9. K.Kobayashi, "Combinatorial structure and capacity of the permuting relay channel," *IEEE Trans. on Inform. Theory*, vol.33, no.6, pp.813-826, Nov. 1987.