

Universal Coding for Correlated Sources with Complementary Delivery**

Akisato KIMURA^{†,††}, Tomohiko UYEMATSU[†], and Shigeaki KUZUOKA^{†*}, *Members*

SUMMARY This paper deals with a universal coding problem for a certain kind of multiterminal source coding system that we call the complementary delivery coding system. In this system, messages from two correlated sources are jointly encoded, and each decoder has access to one of the two messages to enable it to reproduce the other message. Both fixed-to-fixed length and fixed-to-variable length lossless coding schemes are considered. Explicit constructions of universal codes and bounds of the error probabilities are clarified via type-theoretical and graph-theoretical analyses.

key words: multiterminal source coding, complementary delivery, universal coding, types of sequences, bipartite graphs

1. Introduction

The coding problem for correlated information sources was first described and investigated by Slepian and Wolf [1], and later, various coding problems derived from that work were considered (e.g. Wyner [2], Körner and Marton [3], Sgarro [4]). Meanwhile, the problem of universal coding for these systems was first investigated by Csiszár and Körner [5]. Universal coding problems are not only interesting in their own right but also very important in terms of practical applications. Subsequent work has mainly focused on the Slepian-Wolf coding system [6]–[8] since it appears to be difficult to construct universal codes for most of the other coding systems. For example, Muramatsu [9] showed that we cannot construct asymptotically optimal fixed-to-variable length (FV) universal code for Wyner-Ziv coding systems [10] in terms of the coding rate.

This paper deals with a universal coding problem for a certain kind of multiterminal source coding system that we call a complementary delivery coding system [11]. Figure 1 shows a block diagram of the complementary delivery coding system. The encoder observes messages emitted

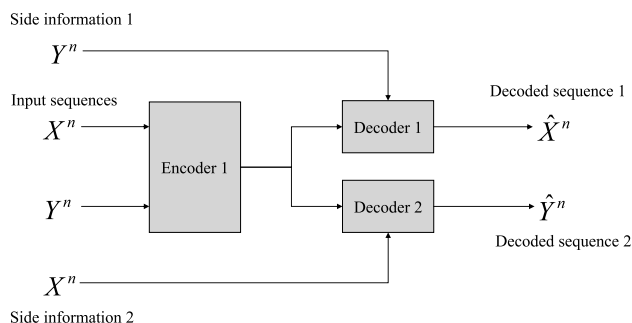


Fig. 1 Complementary delivery coding system.

from two correlated sources, and delivers these messages to other locations (i.e. decoders). Each decoder has access to one of two messages, and therefore wants to reproduce the other message. Although the previous articles [11] considered lossy configurations, this paper considers a lossless configuration. We show an explicit construction of fixed-to-fixed length universal codes. We also clarify the upper and lower bounds of the error probabilities via type-theoretical and graph-theoretical analyses. Fixed-to-variable universal codes can also be constructed in a similar manner.

This paper is organized as follows: Notations and definitions are provided in Sect. 2. Previous results for the complementary delivery coding system are shown in Sect. 3. A proposed coding scheme is described in Sect. 4. Several coding theorems are clarified in Sect. 5. Variable-length coding is discussed in Sect. 6. Finally, concluding remarks are given in Sect. 7.

2. Preliminaries

2.1 Basic Definitions

Let \mathcal{X} be a finite set, \mathcal{B} be a binary set, and \mathcal{B}^* be a set of all finite sequences in the alphabet \mathcal{B} . Let $|\mathcal{X}|$ be the cardinality of \mathcal{X} and $\mathcal{I}_M = \{1, 2, \dots, M\}$. A member of \mathcal{X}^n is written as $x^n = (x_1, x_2, \dots, x_n)$, and substrings of x^n are written as $x_i^j = (x_i, x_{i+1}, \dots, x_j)$ for $i \leq j$. When the dimension is clear from the context, vectors will be denoted by boldface letters, i.e., $\mathbf{x} \in \mathcal{X}^n$. $\mathcal{M}(\mathcal{X})$ denotes the set of all probability distributions on \mathcal{X} . Also, $\mathcal{M}(\mathcal{X}|P_Y)$ denotes the set of all probability distributions on \mathcal{X} given a distribution $P_Y \in \mathcal{M}(\mathcal{Y})$, namely each member $P_{X|Y}$ of $\mathcal{M}(\mathcal{X}|P_Y)$ is characterized by $P_{X,Y} \in \mathcal{M}(\mathcal{X} \times \mathcal{Y})$ as $P_{X,Y} = P_{X|Y}P_Y$. A discrete memoryless source (\mathcal{X}, P_X) is an infinite sequence

Manuscript received December 25, 2006.

Manuscript revised March 30, 2007.

Final manuscript received April 25, 2007.

[†]The authors are with the Department of Communications and Integrated Systems, Tokyo Institute of Technology, Tokyo, 152-8550 Japan.

^{††}The author is with NTT Communication Science Laboratories, NTT Corporation, Atsugi-shi, 243-0198 Japan. URL: <http://www.brl.ntt.co.jp/people/akisato/>

*Presently, with Department of Computer and Communication Sciences, Wakayama University, Japan.

**Parts of the material in this paper were presented at the Symposium on Information Theory and its Applications, Hakodate, Japan November-December 2006, and the International Symposium on Information Theory, Nice, France June 2007.

DOI: 10.1093/ietfec/e90-a.9.1840

of independent copies of a random variable X taking values in \mathcal{X} with a generic distribution $P_X \in \mathcal{M}(\mathcal{X})$. We will denote a source (X, P_X) by referring to its generic distribution P_X or random variable X . For a correlated source (X, Y) , $H(X|Y)$ denotes the conditional entropy of X given Y . For a generic distribution $P_Y \in \mathcal{M}(\mathcal{Y})$ and a conditional distribution $P_{X|Y} \in \mathcal{M}(\mathcal{X}|P_Y)$, $H(P_{X|Y}|P_Y)$ also denotes the conditional entropy of X given Y . $D(P||Q)$ denotes the Kullback-Leibler divergence between two distributions P and Q . In the following, all bases of exponentials and logarithms are set at 2.

2.2 Types of Sequences

Let us define the *type* of a sequence $\mathbf{x} \in \mathcal{X}^n$ as the empirical distribution $Q_{\mathbf{x}} \in \mathcal{M}(\mathcal{X})$ of the sequence \mathbf{x} , i.e.

$$Q_{\mathbf{x}}(a) = \frac{1}{n}N(a|\mathbf{x}) \quad \forall a \in \mathcal{X},$$

where $N(a|\mathbf{x})$ represents the number of occurrences of the letter a in the sequence \mathbf{x} . Similarly, the joint type $Q_{\mathbf{x}, \mathbf{y}} \in \mathcal{M}(\mathcal{X} \times \mathcal{Y})$ is defined by

$$Q_{\mathbf{x}, \mathbf{y}}(a, b) = \frac{1}{n}N(a, b|\mathbf{x}, \mathbf{y}) \quad \forall (a, b) \in \mathcal{X} \times \mathcal{Y}.$$

Let $\mathcal{P}_n(\mathcal{X})$ be the set of types of sequences in \mathcal{X}^n . Similarly, for every type $Q \in \mathcal{P}_n(\mathcal{X})$, let $\mathcal{V}_n(\mathcal{Y}|Q)$ be the set of all stochastic matrices $V : \mathcal{X} \rightarrow \mathcal{Y}$ such that for some pairs $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$ of sequences we have $Q_{\mathbf{x}, \mathbf{y}}(\mathbf{x}, \mathbf{y}) = Q(\mathbf{x})V(\mathbf{y}|\mathbf{x})$. For every type $Q \in \mathcal{P}_n(\mathcal{X})$ we denote

$$T_Q^n \stackrel{\text{def.}}{=} \{\mathbf{x} \in \mathcal{X}^n : Q_{\mathbf{x}} = Q\}.$$

Similarly, for every $\mathbf{x} \in T_Q^n$ and $V \in \mathcal{V}_n(\mathcal{Y}|Q)$ we define

$$T_V^n(\mathbf{x}) \stackrel{\text{def.}}{=} \{\mathbf{y} \in \mathcal{Y}^n : Q(\mathbf{x})V(\mathbf{y}|\mathbf{x}) = Q_{\mathbf{x}, \mathbf{y}}(\mathbf{x}, \mathbf{y}), \forall (x, y) \in \mathcal{X} \times \mathcal{Y}\}.$$

Hereafter, we call $T_V^n(\mathbf{x})$ a *V-shell*.

Here, let us introduce several important properties of types.

Lemma 1. (Type counting lemma [12, Lemma 2.2])
The number of different types of sequences in \mathcal{X}^n is less than $(n+1)^{|\mathcal{X}|}$, namely

$$|\mathcal{P}_n(\mathcal{X})| \leq (n+1)^{|\mathcal{X}|}.$$

Lemma 2. (Sizes of V-shells [12, Lemma 2.5])
For every sequence $\mathbf{x} \in \mathcal{X}^n$ and every stochastic matrix $V : \mathcal{X} \rightarrow \mathcal{Y}$ such that the corresponding V-shell $T_V^n(\mathbf{x})$ is not empty, we have

$$\begin{aligned} |T_V^n(\mathbf{x})| &\geq (n+1)^{-|\mathcal{X}||\mathcal{Y}|} \exp\{nH(V|P_{\mathbf{x}})\}, \\ |T_V^n(\mathbf{x})| &\leq \exp\{nH(V|P_{\mathbf{x}})\}. \end{aligned}$$

Lemma 3. (Probabilities of types [12, Lemma 2.6])
For every type Q of sequences in \mathcal{X}^n and every distribution P_X on \mathcal{X} , we have

$$\begin{aligned} P_X(\mathbf{x}) &= \exp\{-n(D(Q||P_X) + H(Q))\} \quad \forall \mathbf{x} \in T_Q, \\ P_X(T_Q) &\geq (n+1)^{-|\mathcal{X}|} \exp\{-nD(Q||P_X)\}, \\ P_X(T_Q) &\leq \exp\{-nD(Q||P_X)\}. \end{aligned}$$

3. Previous Results

This section formulates the coding problem investigated in this paper, and shows a fundamental bound of the coding rate. We again note that previous work [11] considered lossy coding, whereas this paper considers lossless coding.

First, we formulate the coding problem of the complementary delivery coding system.

Definition 1. (Fixed-to-fixed complementary delivery (FF-CD) code) [11]

A set $(\varphi^n, \widehat{\varphi}_{(1)}^n, \widehat{\varphi}_{(2)}^n)$ of an encoder and two decoders is an FF-CD code with parameters $(n, M_n, e_n^{(X)}, e_n^{(Y)})$ for the source (X, Y) if and only if

$$\begin{aligned} \varphi^n &: \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \mathcal{I}_{M_n} \\ \widehat{\varphi}_{(1)}^n &: \mathcal{I}_{M_n} \times \mathcal{Y}^n \rightarrow \mathcal{X}^n, \quad \widehat{\varphi}_{(2)}^n : \mathcal{I}_{M_n} \times \mathcal{X}^n \rightarrow \mathcal{Y}^n, \\ e_n^{(X)} &= \Pr\{\mathcal{X}^n \neq \widehat{X}^n\}, \quad e_n^{(Y)} = \Pr\{\mathcal{Y}^n \neq \widehat{Y}^n\}, \end{aligned}$$

where

$$\begin{aligned} \widehat{X}^n &\stackrel{\text{def.}}{=} \widehat{\varphi}_{(1)}^n(\varphi^n(\mathcal{X}^n, \mathcal{Y}^n), \mathcal{Y}^n), \\ \widehat{Y}^n &\stackrel{\text{def.}}{=} \widehat{\varphi}_{(2)}^n(\varphi^n(\mathcal{X}^n, \mathcal{Y}^n), \mathcal{X}^n). \end{aligned}$$

Definition 2. (FF-CD-achievable rate)

R is an FF-CD-achievable rate of the source (X, Y) if and only if there exists a sequence $\{(\varphi^n, \widehat{\varphi}_{(1)}^n, \widehat{\varphi}_{(2)}^n)\}_{n=1}^{\infty}$ of FF-CD codes with parameters $\{(n, M_n, e_n^{(X)}, e_n^{(Y)})\}_{n=1}^{\infty}$ for the source (X, Y) such that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n &\leq R, \\ \lim_{n \rightarrow \infty} e_n^{(X)} &= \lim_{n \rightarrow \infty} e_n^{(Y)} = 0. \end{aligned}$$

Definition 3. (Inf FF-CD-achievable rate)

$$\begin{aligned} R_f(X, Y) &= \inf\{R : \\ &R \text{ is an FF-CD-achievable rate of } (X, Y)\}. \end{aligned}$$

Willems, Wolf and Wyner [13], [14] investigated a coding problem where several users are physically separated but communicate with each other via a satellite, and determined the minimum coding rate for the three users when transmitting to and from the satellite. The complementary delivery coding system is a special case of the system described by Willems et al., which considers the case of two users. Therefore, we can immediately obtain the closed form of $R_f(X, Y)$ from the result obtained by Willems et al.

Theorem 1. (Coding theorem of FF-CD codes)

$$\begin{aligned} R_f(X, Y) &= \max\{H(X|Y), H(Y|X)\} \\ &= \max\{H(P_{X|Y}|P_Y), H(P_{Y|X}|P_X)\} \end{aligned}$$

4. Code Construction

This section shows an explicit construction of universal codes for the complementary delivery coding system defined by Definition 1. The coding scheme is described as follows:

[Encoding]

1. Determine a set $\mathcal{S}_n(R)$ of joint types as

$$\begin{aligned} \mathcal{S}_n(R) = \{Q_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y}) : \\ \max\{H(V|Q_X), H(W|Q_Y)\} \leq R, \\ Q_{XY} = Q_X V = Q_Y W, \\ V \in \mathcal{V}_n(\mathcal{Y}|Q_X), W \in \mathcal{V}_n(\mathcal{X}|Q_Y)\}, \end{aligned}$$

where $R > 0$ is a given coding rate. We note that the joint type Q_{XY} specifies the types Q_X, Q_Y , and the conditional types V and W .

2. Create a table (henceforth we call this a *coding table*, see Figure 2 left) for each joint type $Q_{XY} \in \mathcal{S}_n(R)$. Each row of the coding table corresponds to a sequence $\mathbf{x} \in T_{Q_X}^n$, and each column corresponds to a sequence $\mathbf{y} \in T_{Q_Y}^n$.
3. Mark cells that correspond to sequence pairs $(\mathbf{x}, \mathbf{y}) \in T_{Q_{XY}}^n$ (see Figure 2 middle). Codewords will be given only to sequence pairs that correspond to marked cells.
4. Fill the marked cells with $\exp(nR)$ different symbols such that each symbol occurs at most once in each row and at most once in each column. An example of symbol filling is shown on the right in Figure 2 right.
5. For a given pair of sequences $(\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{Y}$ with the joint type Q_{XY} , if $Q_{XY} \in \mathcal{S}_n(R)$, the index assigned to the joint type Q_{XY} of (\mathbf{x}, \mathbf{y}) is the first part of the codeword, and the symbol filling the cell of (\mathbf{x}, \mathbf{y}) in the coding table of Q_{XY} is determined as the second part of the codeword. For the sequence pairs (\mathbf{x}, \mathbf{y}) whose joint type Q_{XY} does not belong to $\mathcal{S}_n(R)$, the corresponding codeword is determined arbitrarily and an encoding error is declared.

[Decoding: $\widehat{\varphi}_{(1)}^n$] (Almost the same as for $\widehat{\varphi}_{(2)}^n$)

1. Find the coding table of the type \widehat{Q}_{XY} that corresponds to the first part of the received codeword. The decoder can find the coding table used in the encoding scheme if no encoding error occurs. In this case, \widehat{Q}_{XY} should be Q_{XY} .

	y_1	y_2	y_3	y_4	y_5		y_1	y_2	y_3	y_4	y_5		y_1	y_2	y_3	y_4	y_5
x_1						x_1	●	●	●			x_1	1	2	3		
x_2						x_2	●			●	●	x_2	2			7	3
x_3						x_3		●	●	●		x_3		3	7	2	
x_4						x_4	●	●			●	x_4	3	7			2
x_5						x_5			●	●	●	x_5			2	3	7

Fig. 2 Example of coding scheme (left) Coding table (middle) Positions where codewords will be provided (right) Provided codewords.

2. Find the cell filled with the second part of the received codeword from the column of the side information sequence $\mathbf{y} \in T_{Q_Y}^n$. The sequence $\widehat{\mathbf{x}} \in T_{Q_X}^n$ that corresponds to the row of the cell found in this step is reproduced.

First, we show the existence of such coding tables. To this end, we introduce the following two lemmas.

Lemma 4. For a given coding table of a joint type $Q_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$, the number of marked cells in every row of the coding table, $N_y(Q_{XY})$, is a constant value that is less than $\exp(nR)$, and the number of marked cells in every column of the coding table, $N_x(Q_{XY})$, is also a constant value of less than $\exp(nR)$, both of which depend solely on the joint type Q_{XY} .

Proof. Note that the number of marked cells in each row equals the cardinality of the V-shell $T_V^n(\mathbf{x})$ for the sequence $\mathbf{x} \in T_{Q_X}^n$ that corresponds to the row. The cardinality of V-shells $T_V^n(\mathbf{x})$ is constant for a given joint type Q_{XY} and any sequences $\mathbf{x} \in T_{Q_X}^n$, this cardinality is bounded as follows:

$$\begin{aligned} N_y(Q_{XY}) = |T_V^n(\mathbf{x})| \\ \leq \exp\{nH(V|Q_X)\} \quad (\because \text{Lemma 2}) \\ \leq \exp(nR). \end{aligned}$$

In the same way, the number of marked cells in each column equals the cardinality of the V-shell $T_W^n(\mathbf{y})$ for the sequence $\mathbf{y} \in T_{Q_Y}^n$ that corresponds to the column, and therefore it can be bounded as

$$N_x(Q_{XY}) = |T_W^n(\mathbf{y})| \leq \exp(nR).$$

This concludes the proof of Lemma 4. □

Lemma 5. For given integers m_x, m_y, n_x and n_y that satisfy $m_x \geq n_x$ and $m_y \geq n_y$, there exists an $m_x \times m_y$ table filled with $\max(n_x, n_y)$ different symbols such that

- at most n_y cells are filled with a certain symbol for each row (blank cells are possible),
- at most n_x cells are filled with a certain symbol for each column (blank cells are possible),
- each symbol occurs at most once in each row and at most once in each column.

Proof. The table mentioned in this lemma is equivalent to a bipartite graph such that

- each node in one set corresponds to a row in the table, and each node in the other set corresponds to a column in the table,
- each edge corresponds to a cell in the table, to which a certain symbol is assigned,
- $\max(n_x, n_y)$ different colors are given to edges, each of which corresponds to a symbol in the table,
- no two edges with the same color share a common node.

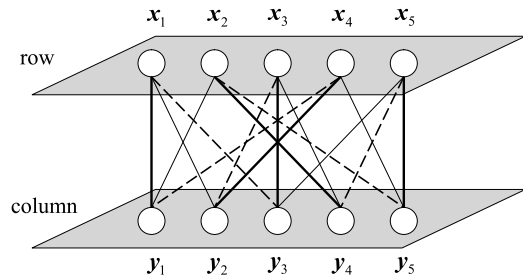


Fig. 3 Example of a bipartite graph. ($m_x = m_y = 5$, $n_x = n_y = 3$, equivalent to the table in Fig. 2 right)

Figure 3 shows an example of such a graph. Here, let us introduce the following lemma for bipartite graphs:

Lemma 6. (König [15], [16])

If a graph G is bipartite, the minimum number of colors necessary for edge coloring of the graph G equals the maximum degree of G .

Lemma 6 ensures the existence of the above bipartite graph. This concludes the proof of Lemma 5. \square

From Lemmas 4 and 5, we can easily show the existence of coding tables by setting $m_x = |T_{Q_X}^n|$, $m_y = |T_{Q_Y}^n|$, $n_x = |T_W^n(\mathbf{y})|$ and $n_y = |T_V^n(\mathbf{x})|$ in Lemma 5.

5. Coding Theorems

We can obtain the following theorem for the universal FF-CD codes constructed in Sect. 4. The proof is similar to that of the theorem of universal coding for a single source.

Theorem 2. *For a given real number $R > 0$, there exists a sequence of universal FF-CD codes with parameters $\{(n, M_n, e_n^{(X)}, e_n^{(Y)})\}_{n=1}^\infty$ such that for any integer $n \geq 1$ and any source (X, Y) with a generic distribution $P_{XY} \in \mathcal{M}(\mathcal{X} \times \mathcal{Y})$*

$$\begin{aligned} \frac{1}{n} \log M_n &\leq R + \frac{1}{n} |\mathcal{X} \times \mathcal{Y}| \log(n+1), \\ e_n^{(X)} + e_n^{(Y)} &\leq 2(n+1)^{|\mathcal{X} \times \mathcal{Y}|} \\ &\quad \times \exp \left\{ -n \min_{Q_{XY} \in \overline{\mathcal{S}}_n(R)} D(Q_{XY} \| P_{XY}) \right\}, \end{aligned}$$

where $\overline{\mathcal{S}}_n(R) = \mathcal{P}_n(\mathcal{X} \times \mathcal{Y}) - \mathcal{S}_n(R)$.

Proof. Lemmas 4 and 5 ensure the existence of a coding table for every joint type $Q_{XY} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$. From the coding scheme, the size of the codeword set is bounded as

$$\begin{aligned} M_n &\leq |\mathcal{P}_n(\mathcal{X} \times \mathcal{Y})| \exp(nR) \\ &\leq (n+1)^{|\mathcal{X} \times \mathcal{Y}|} \exp(nR), \quad (\because \text{Lemma 1}) \end{aligned}$$

which implies the first inequality of Theorem 2. Next, we evaluate decoding error probabilities. Since every sequence pair $(\mathbf{x}, \mathbf{y}) \in T_{Q_{XY}}^n$ that satisfies $Q_{XY} \in \mathcal{S}_n(R)$ is reproduced correctly at the decoder, the sum of error probabilities is bounded as

$$\begin{aligned} e_n^{(X)} + e_n^{(Y)} &\leq 2 \Pr \left\{ \exists Q_{XY} \in \overline{\mathcal{S}}_n(R), (X^n, Y^n) \in T_{Q_{XY}}^n \right\} \\ &\leq 2 \sum_{Q_{XY} \in \overline{\mathcal{S}}_n(R)} \exp\{-nD(Q_{XY} \| P_{XY})\} \\ &\quad (\because \text{Lemma 3}) \\ &\leq 2 \sum_{Q_{XY} \in \overline{\mathcal{S}}_n(R)} \exp \left\{ -n \min_{Q_{XY} \in \overline{\mathcal{S}}_n(R)} D(Q_{XY} \| P_{XY}) \right\} \\ &\leq 2(n+1)^{|\mathcal{X} \times \mathcal{Y}|} \\ &\quad \times \exp \left\{ -n \min_{Q_{XY} \in \overline{\mathcal{S}}_n(R)} D(Q_{XY} \| P_{XY}) \right\} \\ &\quad (\because \text{Lemma 1}) \end{aligned}$$

This completes the proof of Theorem 2. \square

We can see that for any real value $R \geq R_f(X, Y)$ we have

$$\min_{Q_{XY} \in \overline{\mathcal{S}}_n(R)} D(Q_{XY} \| P_{XY}) > 0.$$

This implies that any real value $R \geq R_f(X, Y)$ is a *universal FF-CD achievable rate* of (X, Y) , namely, there exists a sequence of universal FF-CD codes with parameters $\{(n, M_n, e_n^{(X)}, e_n^{(Y)})\}_{n=1}^\infty$ such that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n &\leq R, \\ \lim_{n \rightarrow \infty} e_n^{(X)} &= \lim_{n \rightarrow \infty} e_n^{(Y)} = 0. \end{aligned}$$

The following converse theorem indicates that the error exponent obtained in Theorem 2 is tight.

Theorem 3. *Any sequence of FF-CD codes with parameters $\{(n, M_n, e_n^{(X)}, e_n^{(Y)})\}_{n=1}^\infty$ for the source (X, Y) must satisfy*

$$\begin{aligned} e_n^{(X)} + e_n^{(Y)} &\geq \frac{1}{2} (n+1)^{-|\mathcal{X} \times \mathcal{Y}|} \\ &\quad \times \exp \left\{ -n \min_{Q_{XY} \in \overline{\mathcal{S}}_n(R + \epsilon_n)} D(Q_{XY} \| P_{XY}) \right\} \end{aligned}$$

for any integer $n \geq 1$ and a given coding rate $R = 1/n \log M_n > 0$, where

$$\begin{aligned} \epsilon_n &\stackrel{\text{def.}}{=} \frac{1}{n} \{ |\mathcal{X} \times \mathcal{Y}| \log(n+1) + 1 \} \\ &\rightarrow 0 \quad (n \rightarrow \infty). \end{aligned}$$

Proof. Note that the number of sequences to be decoded correctly for each decoder are at most $\exp(nR)$. Here, let us consider a joint type $Q_{XY} \in \overline{\mathcal{S}}_n(R + \epsilon_n)$. Lemma 2 and the definition of $\overline{\mathcal{S}}_n(R + \epsilon_n)$ imply that for $(\mathbf{x}, \mathbf{y}) \in T_{Q_{XY}}^n$ we have

$$\begin{aligned} &\max\{|T_V^n(\mathbf{x})|, |T_W^n(\mathbf{y})|\} \\ &\geq (n+1)^{-|\mathcal{X} \times \mathcal{Y}|} \\ &\quad \times \max\{\exp\{nH(V|Q_X)\}, \exp\{nH(W|Q_Y)\}\} \\ &\geq (n+1)^{-|\mathcal{X} \times \mathcal{Y}|} \exp(n(R + \epsilon_n)) \\ &= 2 \exp(nR). \end{aligned}$$

Therefore, at least half of the sequences in the V-shell $T_V^n(\mathbf{x})$ will not be decoded correctly at the decoder $\varphi_{(2)}^n$, or at least half of sequences in the V-shell $T_W^n(\mathbf{y})$ will not be decoded correctly at the decoder $\varphi_{(1)}^n$. Thus, the sum of error probabilities is bounded as

$$\begin{aligned} & e_n^{(X)} + e_n^{(Y)} \\ & \geq \frac{1}{2} \sum_{Q_{XY} \in \overline{\mathcal{S}}_n(R+\epsilon_n)} \Pr\{(X^n, Y^n) \in T_{Q_{XY}}\} \\ & \geq \frac{1}{2} (n+1)^{-|\mathcal{X} \times \mathcal{Y}|} \\ & \quad \times \sum_{Q_{XY} \in \overline{\mathcal{S}}_n(R+\epsilon_n)} \exp\{-nD(Q_{XY} \| P_{XY})\} \\ & \quad (\because \text{Lemma 3}) \\ & \geq \frac{1}{2} (n+1)^{-|\mathcal{X} \times \mathcal{Y}|} \\ & \quad \times \exp\left\{-n \min_{Q_{XY} \in \overline{\mathcal{S}}_n(R+\epsilon_n)} D(Q_{XY} \| P_{XY})\right\} \end{aligned}$$

This concludes the proof of Theorem 3. \square

The following corollary is directly derived from Theorems 2 and 3.

Corollary 1. For a given real number $R > 0$, there exists a sequence of universal FF-CD codes with parameters $\{(n, M_n, e_n^{(X)}, e_n^{(Y)})\}_{n=1}^\infty$ such that for any source (X, Y)

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log M_n & \leq R, \\ \lim_{n \rightarrow \infty} -\frac{1}{n} \log(e_n^{(X)} + e_n^{(Y)}) & = \min_{Q_{XY} \in \overline{\mathcal{S}}(R)} D(Q_{XY} \| P_{XY}), \end{aligned}$$

where

$$\begin{aligned} \mathcal{S}(R) & = \{Q_{XY} \in \mathcal{M}(\mathcal{X} \times \mathcal{Y}) : \\ & \quad \max\{H(V|Q_X), H(W|Q_Y)\} \leq R, \\ & \quad Q_{XY} = Q_X V = Q_Y W, \\ & \quad V \in \mathcal{M}(\mathcal{Y}|Q_X), W \in \mathcal{M}(\mathcal{X}|Q_Y)\}, \end{aligned}$$

$$\text{and } \overline{\mathcal{S}}(R) = \mathcal{M}(\mathcal{X} \times \mathcal{Y}) - \mathcal{S}(R).$$

6. Variable-Length Coding

This section discusses variable-length coding for the complementary delivery coding system, and shows an explicit construction of universal variable-length codes. The coding scheme is similar to that of fixed-length codes, and also utilizes the coding tables defined in Sect. 4.

6.1 Formulation

Definition 4. (Fixed-to-variable complementary delivery (FV-CD) code)

A set $(\varphi^n, \widehat{\varphi}_{(1)}^n, \widehat{\varphi}_{(2)}^n)$ of an encoder and two decoders is an

FV-CD code for the source (X, Y) if and only if

$$\begin{aligned} \varphi^n & : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \mathcal{B}^* \\ \widehat{\varphi}_{(1)}^n & : \varphi^n(\mathcal{X}^n, \mathcal{Y}^n) \times \mathcal{Y}^n \rightarrow \mathcal{X}^n, \\ \widehat{\varphi}_{(2)}^n & : \varphi^n(\mathcal{X}^n, \mathcal{Y}^n) \times \mathcal{X}^n \rightarrow \mathcal{Y}^n, \\ e_n^{(X)} & = \Pr\{X^n \neq \widehat{X}^n\} = 0, \\ e_n^{(Y)} & = \Pr\{Y^n \neq \widehat{Y}^n\} = 0, \end{aligned}$$

where

$$\begin{aligned} \widehat{X}^n & \stackrel{\text{def.}}{=} \widehat{\varphi}_{(1)}^n(\varphi^n(X^n, Y^n), Y^n), \\ \widehat{Y}^n & \stackrel{\text{def.}}{=} \widehat{\varphi}_{(2)}^n(\varphi^n(X^n, Y^n), X^n), \end{aligned}$$

and the image of φ^n is a prefix set.

Definition 5. (FV-CD-achievable rate)

R is an FV-CD-achievable rate of the source (X, Y) if and only if there exists a sequence of FV-CD codes $\{(\varphi^n, \widehat{\varphi}_{(1)}^n, \widehat{\varphi}_{(2)}^n)\}_{n=1}^\infty$ for the source (X, Y) such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} E[l(\varphi^n(X^n, Y^n))] \leq R,$$

where $l(\cdot) : \mathcal{B}^* \rightarrow \mathcal{R}$ is a length function.

Definition 6. (Inf FV-CD-achievable rate)

$$\begin{aligned} R_v(X, Y) & = \inf\{R : \\ & \quad R \text{ is an FV-CD-achievable rate of } (X, Y)\}. \end{aligned}$$

6.2 Code Construction

We can construct universal FV-CD codes in a similar manner to universal FF-CD codes. Note that the coding rate depends on the type of sequence pair to be encoded, whereas the coding rate is fixed beforehand for fixed-length coding. The coding scheme is described as follows:

[Encoding]

1. Create a coding table for each joint type $Q_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$ in the same way as Step 2 of Sect. 4.
2. Mark cells that correspond to sequence pairs $(\mathbf{x}, \mathbf{y}) \in T_{Q_{XY}}$.
3. Fill the marked cells on the coding table with different $\max\{|T_V^n(\mathbf{x})|, |T_W^n(\mathbf{y})|\}$ symbols such that each symbol occurs at most once in each row and at most once in each column, where $\mathbf{x} \in T_{Q_X}^n, \mathbf{y} \in T_{Q_Y}^n$.
4. For a given pair of sequences $(\mathbf{x}, \mathbf{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, the number (index) assigned to the joint type Q_{XY} of (\mathbf{x}, \mathbf{y}) is the first part of the codeword, and the symbol filling the cell of (\mathbf{x}, \mathbf{y}) in the coding table of Q_{XY} is determined as the second part of the codeword.

[Decoding]

Decoding can be accomplished in almost the same way as the fixed-length coding. Note that the decoder can always find the coding table used in the encoding scheme.

6.3 Coding Theorems

We begin by showing a theorem for (non-universal) variable-length coding, which indicates that the inf coding rate of variable-length coding is the same as that of fixed-length coding.

Theorem 4. (Coding theorem of FV-CD code)

$$R_v(X, Y) = R_f(X, Y) = \max\{H(X|Y), H(Y|X)\}.$$

Proof. See Appendix. \square

The following direct theorem for universal coding indicates that the coding scheme presented in the previous subsection can achieve the inf achievable rate clarified in Theorem 4.

Theorem 5. *There exists a sequence of universal FV-CD codes $\{(\varphi^n, \widehat{\varphi}_{(1)}^n, \widehat{\varphi}_{(2)}^n)\}_{n=1}^\infty$ such that for any integer $n \geq 1$ and any source (X, Y) , the overflow probability, namely the probability that the length of a codeword exceeds a given real number $R > 0$, is bounded as*

$$\begin{aligned} \bar{\rho}_n(R) &\stackrel{\text{def.}}{=} \Pr\{l(\varphi^n(X^n, Y^n)) > n(R + \epsilon_n)\} \\ &\leq (n+1)^{|X \times Y|} \\ &\quad \times \exp\left\{-n \min_{Q_{XY} \in \overline{\mathcal{S}}_n(R)} D(Q_{XY} \| P_{XY})\right\}, \end{aligned}$$

where ϵ_n is defined in Theorem 3. This implies that there exists a sequence of universal FV-CD codes $\{(\varphi^n, \widehat{\varphi}_{(1)}^n, \widehat{\varphi}_{(2)}^n)\}_{n=1}^\infty$ that satisfies

$$\limsup_{n \rightarrow \infty} \frac{1}{n} l(\varphi^n(X^n, Y^n)) \leq R_v(X, Y) \quad a.s. \quad (1)$$

Proof. The overflow probability can be obtained in the same way as an upperbound of the error probability of FF-CD codes, which has been shown in the proof of Theorem 2. Thus, from Theorem 4 we have

$$\sum_{n=1}^{\infty} \Pr\left\{\frac{1}{n} l(\varphi^n(X^n, Y^n)) > R_v(X, Y) + \delta\right\} < \infty$$

for a given $\delta > 0$. From Borel-Cantelli's lemma [17, Lemma 4.6.3], we immediately obtain Eq.(1). This completes the proof of Theorem 5. \square

The following converse theorem for variable-length coding indicates that the exponent of the overflow probability obtained in Theorem 5 is tight. This can be easily obtained in almost the same way as Theorem 3.

Theorem 6. *Any sequence of FV-CD codes $\{(\varphi^n, \widehat{\varphi}_{(1)}^n, \widehat{\varphi}_{(2)}^n)\}_{n=1}^\infty$ for the source (X, Y) must satisfy*

$$\begin{aligned} \bar{\rho}_n(R) &\geq (n+1)^{-|X \times Y|} \\ &\quad \times \exp\left\{-n \min_{Q_{XY} \in \overline{\mathcal{S}}_n(R + \epsilon_n)} D(Q_{XY} \| P_{XY})\right\}. \end{aligned}$$

for a given real number $R > 0$ and any integer $n \geq 1$, where ϵ_n is defined in Theorem 3.

The following corollary is directly derived from Theorems 5 and 6.

Corollary 2. *There exists a sequence of universal FV-CD codes $\{(\varphi^n, \widehat{\varphi}_{(1)}^n, \widehat{\varphi}_{(2)}^n)\}_{n=1}^\infty$ such that for any source (X, Y)*

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} l(\varphi^n(X^n, Y^n)) &\leq R_v(X, Y) \quad a.s. \\ \lim_{n \rightarrow \infty} -\frac{1}{n} \log \bar{\rho}_n(R) &= \min_{Q_{XY} \in \overline{\mathcal{S}}(R)} D(Q_{XY} \| P_{XY}) \end{aligned}$$

7. Concluding Remarks

We investigated a universal coding problem for the complementary delivery coding system. First, we presented an explicit construction of universal fixed-length codes, which was based on a graph-theoretical technique. We clarified that the error exponent achieved by the proposed coding scheme is asymptotically optimal. Next, we applied the coding scheme to construction of universal variable-length codes. We clarified that there exists a universal code such that the codeword length converges to the minimum achievable rate almost surely, and that the exponent of the overflow probability achieved by the proposed coding scheme is optimal. This paper dealt with only the lossless configuration, and therefore constructing universal lossy codes for the complementary delivery coding system still remains as an open problem.

Acknowledgements

The authors would like to thank Dr. Jun Muramatsu of NTT Communication Science Laboratories, and Prof. Ryutaroh Matsumoto of Tokyo Institute of Technology for their valuable discussions and helpful comments, which led to improvements in this work. The authors also thank Drs. Yoshinobu Tonomura, Hiromi Nakaiwa, Tatsuto Takeuchi, Shoji Makino and Junji Yamato of NTT Communication Science Laboratories for their help.

References

- [1] D. Slepian and J.K. Wolf, "Noiseless coding of correlated information sources," IEEE Trans. Inf. Theory, vol.19, no.4, pp.471–480, July 1973.
- [2] A.D. Wyner, "On source coding with side information at the decoder," IEEE Trans. Inf. Theory, vol.21, no.3, pp.294–300, May 1975.
- [3] J. Körner and K. Marton, "Images of a set via two channels and their role in multi-user communication," IEEE Trans. Inf. Theory, vol.23, no.6, pp.751–761, Nov. 1975.
- [4] A. Sgarro, "Source coding with side information at several decoders," IEEE Trans. Inf. Theory, vol.23, no.2, pp.179–182, March 1977.
- [5] I. Csiszár and J. Körner, "Towards a general theory of source networks," IEEE Trans. Inf. Theory, vol.26, no.2, pp.155–165, March 1980.
- [6] I. Csiszár, "Linear codes for source and source networks: Error exponents, universal coding," IEEE Trans. Inf. Theory, vol.28, no.4, pp.585–592, July 1982.

- [7] Y. Oohama and T.S. Han, "Universal coding for the Slepian-Wolf data compression system and the strong converse theorem," IEEE Trans. Inf. Theory, vol.40, no.6, pp.1908–1919, Nov. 1994.
- [8] T. Uyematsu, "An algebraic construction of codes for Slepian-Wolf source networks," IEEE Trans. Inf. Theory, vol.47, no.7, pp.3082–3088, Nov. 2001.
- [9] J. Muramatsu, Universal data compression algorithms for stationary Ergodic sources based on the complexity of sequences, Ph.D. Thesis, Nagoya University, 1998.
- [10] A.D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," IEEE Trans. Inf. Theory, vol.22, no.1, pp.1–10, Jan. 1976.
- [11] A. Kimura and T. Uyematsu, "Multiterminal source coding with complementary delivery," Proc. International Symposium on Information Theory and its Applications (ISITA), pp.189–194, Oct. 2006.
- [12] I. Csiszar and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems, Academic Press, New York, 1981.
- [13] F.M.J. Willems, J.K. Wolf, and A.D. Wyner, "Communicating via a processing broadcast satellite," Proc. 1989 IEEE/CAM Information Theory Workshop, p.3-1, June 1989.
- [14] A.D. Wyner, J.K. Wolf, and F.M.J. Willems, "Communicating via a processing broadcast satellite," IEEE Trans. Inf. Theory, vol.48, no.6, pp.1243–1249, June 2002.
- [15] D. König, "Graphok és alkalmazásuk a determinánsok és a halmazok elméletére," Matematikai és Természettudományi Értesítő, vol.34, pp.104–119, 1916.
- [16] N.L. Biggs, E.K. Lloyd, and R.J. Wilson, Graph Theory, Oxford University Press, 1976.
- [17] R.M. Gray, Probability, Random Processes, Ergodic Properties, Springer-Verlag, New York, 1988.

Appendix: Proof of Theorem 4

A.1 Direct Part

Proof. We can apply a sequence of *achievable* FF-CD codes (fixed-length codes). The encoder φ^n assigns the same codeword as that of the fixed-length code to a sequence pair $(x, y) \in X^n \times Y^n$ that is correctly reproduced by the fixed-length code. Otherwise, the encoder sends the sequence pair itself as a codeword.

The above FV-CD code can always reproduce the original sequence pair at the decoders, and it attains the desired coding rate. \square

A.2 Converse Part

Proof. We can prove the converse part in a similar manner to that for fixed-length coding. Let a sequence $\{(\varphi^n, \widehat{\varphi}_{(1)}^n, \widehat{\varphi}_{(2)}^n)\}_{n=1}^{\infty}$ of FV-CD codes be given that satisfies the conditions of Definitions 4 and 5. From Definition 5, for any $\delta > 0$ there exists an integer $n_1 = n_1(\delta)$ and then for all $n \geq n_1(\delta)$, we can obtain

$$\frac{1}{n}E[l(\varphi^n(X^n, Y^n))] \leq R + \delta. \quad (\text{A}\cdot 1)$$

Here, let us define $A_n = \varphi^n(X^n, Y^n)$. Since the decoder $\widehat{\varphi}_{(1)}^n$ can always reproduce the original sequence X^n from the received codeword A_n and side information Y^n , we can see that

$$H(X^n|A_n Y^n) = 0. \quad (\text{A}\cdot 2)$$

Similarly, we can obtain

$$H(Y^n|A_n X^n) = 0.$$

Substituting A_n into Eq.(A·1), we have

$$\begin{aligned} n(R + \delta) &\geq E[l(A_n)] \\ &\geq H(A_n) \quad (\because A_n \text{ is a prefix set}) \\ &\geq H(A_n|Y^n) \\ &\geq I(X^n; A_n|Y^n) \\ &= H(X^n|Y^n). \quad (\because \text{Eq.(A}\cdot 2)) \end{aligned}$$

Since we can select an arbitrarily small $\delta > 0$ for a sufficient large n , we can obtain

$$R \geq \frac{1}{n}H(X^n|Y^n) = H(X|Y).$$

In the same way, we also obtain

$$R \geq H(Y|X).$$

\square



Akisato Kimura received B.E. and M.E. degrees from Tokyo Institute of Technology in 1998 and 2000, respectively. Since 2000, he has been with NTT Communication Science Laboratories, NTT Corporation, where he is a researcher in the Media Information Laboratory. Since 2005, he has also been a Ph.D. candidate in the Department of Communications and Integrated Systems, Tokyo Institute of Technology. He has been working on multimedia information retrieval and multimedia databases. His research interests include pattern recognition, computer vision, Shannon theory and multiterminal information theory.



Tomohiko Uyematsu received B.E., M.E. and D.E. degrees from Tokyo Institute of Technology in 1982, 1984 and 1988, respectively. From 1984 to 1992, he was with the Department of Electrical and Electronics Engineering of Tokyo Institute of Technology, first as a research associate, next as a lecturer, and lastly as an associate professor. From 1992 to 1997, he was with School of Information Science of Japan Advanced Institute of Science and Technology as an associate professor, and currently

he is a professor in the Department of Communications and Integrated Systems. In 1992 and 1996, he was a visiting researcher at the Centre National de la Recherche Scientifique, France and Delft University of Technology, Netherlands, respectively. He received the Shinohara Memorial Young Engineer Award in 1989, and the Best Paper Award in 1993, 1996, 2002 and 2007 all from IEICE. His current research interests are in the areas of information theory, especially Shannon theory and multiterminal information theory. Dr. Uyematsu is a senior member of IEEE.



Shigeaki Kuzuoka received B.E., M.E. and D.E. degrees from Tokyo Institute of Technology in 2002, 2004 and 2007, respectively. Since 2007, he has been with the Department of Computer and Communication Sciences, Wakayama University as an assistant professor. He received the Best Paper Award in 2007 from IEICE. His research interests are in the areas of information theory, Shannon theory and source coding.