On the Structure of a Common Knowledge Created by Correlated Observations and Transmission over Helping Channels *†

Vladimir B. Balakirsky

Electrical Engineering Department
Eindhoven University of Technology
P.O.Box 513, 5600 MB Eindhoven, the Netherlands
on leave from the Data Security Association "Confident"
193060 Saint Petersburg, Russia
v.b.balakirsky@ele.tue.nl

1 Introduction and statement of the problem

Suppose that two individuals, person X and person Y, communicate with each other in such a way that X sends one of M_X messages to Y and, simultaneously, Y sends one of M_Y messages to X. The messages are numbered by the integers $1, \ldots, M_X$ and $1, \ldots, M_Y$. Assuming the numbers to be the identifiers for the corresponding messages, we consider the pairs of the exchanged messages $(i, j) \in \{1, \ldots, M_X\} \times \{1, \ldots, M_Y\}$ as possible common values of X and Y which describe their common knowledge. Suppose also that there is another person, called the source, who gives the same binary vector $\mathbf{x}$ of length n to the individuals. Then X and Y update their knowledge by including this vector, which means that now they have a triple $(i, j, \mathbf{x})$ in common and if 2^n is much greater than $M_X M_Y$, then the total number of possible common values is also much greater. However, if the source changes the rules in such a way that $\mathbf{x}$ is given to X and $\mathbf{y}$ is given to Y, where the vectors $\mathbf{x}$ and $\mathbf{y}$ do not coincide, but correlated, then this updating of the transmitted pair of messages is not possible any more, and the individuals can revert the situation in which they may agree on $M_X M_Y$ common values. An alternative algorithm can be fixed as follows : X and Y compute their messages using deterministic functions of the observations and each individual, based on the vector given by the source and the message received from the other person, constructs a value belonging to some "virtual" space, which is assumed to be common to both of them and can be formally presented as a finite set Ω. The algorithm should be assigned in such a way that the values are also common. We will investigate this possibility and demonstrate the example in which one of 20 pairs of messages is exchanged, one of 60 pairs of vectors is given by the source, while X and Y construct one of 50 common values.

*The work was supported by the University of Bielefeld (Germany) and the Eindhoven University of Technology (the Netherlands).

†The author is grateful to Professor Rudolf Ahlswede and Professor Imre Csiszár for helpful and stimulating discussions, which essentially affected this research and presentation of the results. The help of Dr. Roger Bultitude in the preparing of the manuscript is also highly appreciated.

The three participants may have many reasons for communication under the rules described above; in fact, these reasons come as corollaries from saying "another person called the source". We will mention those reasons, which are relevant to the foregoing formal discussion. The source considers the communication system as a system of control: he knows how many messages can be exchanged and controls the common knowledge of X and Y by sending them sequences having a certain correlation. This knowledge is bounded from above, since the individuals cannot agree on more than the fixed number of common values, which is defined by the correlation of the source sequences and can be achieved if X and Y form their messages in an optimal way. From a technical perspective, X and Y are interested in the possibility of communication using the source sequences to establish the result of a common random experiment for other purposes, like cryptography and identification. For example, they have a table of binary sequences and want to use one of these sequences as a secret key; an agreement on the particular sequence is achieved by constructing the common pointer to some of the rows of this table. Another possibility can be viewed as matching through the source : person X communicates with many individuals and he wants to discover which of them is the one who receives the correlated sequence from the source and sends his messages using the algorithm expected by X; a similar problem has to be solved by Y. In other words, in analyzing the source sequences and communicating over the channels the individuals investigate each other, and the source presents data for this study. Note also that, in a general context, any message sent by a person is the value of some function of his observations, and any discussion about the reasons would be rather artificial in a sense that the person does not have any choice.

We consider the class of problems described above as belonging to the multi–user direction of information theory started by Shannon [1]. The development of this direction in the 1970s was essentially initiated by Ahlswede [2], [5] who determined the achievable rate region for memoryless multiple access channels under the condition of arbitrary small average decoding error probability. The statements of the problems studied for multiple access channels include the situation when the decoder wants to recover the sequence at the output of a random generator based on the message of the encoder and another message, which was formed as a function of a correlated sequence and transmitted by the "helper" [4], [6], [7], where the role of the helper in this case is to present some side information to the decoder. The asymptotic characterization of the achievable rates of encoding these sequences was found by Wyner [6] and Ahlswede–Körner [7], but a generalization of their approach to the case of several helpers is a difficult problem related to the analysis of the dependent partitions of the spaces containing sequences of the helpers, and this problem is still open [8]. The point that there are interesting applications when the source sequences play some auxiliary role in the communication process was discovered by Ahlswede–Dueck [9], who showed that the noise in the channel gives the randomization that can be effectively used in identification schemes. The role of the common randomness in the communication systems where the participants try to take advantage from the observations of public random processes was also studied in [10], [11], [12], [13], and other papers. There also exists a notion of so–called "common information" introduced by Gács–Körner [3], which measures randomness contained in the variables that can be independently constructed using correlated random sequences; the authors showed that the desired random variables exist only if the source has a "special structure". Note that the possibility to communicate allows X and Y to represent the source as a collection of sub–sources having such a structure if they have the capabilities to do so.

2 Formal statement of the problem

Given $M_X, M_Y \geq 1$ and the set $\overline{\mathcal{X}\mathcal{Y}} \subseteq \mathcal{X} \times \mathcal{Y}$, construct four functions f, g, K, L defined by the values,

$$(f(\mathbf{x}) \in [M_X])_{\mathbf{x}\in\mathcal{X}} \,, \quad (g(\mathbf{y}) \in [M_Y])_{\mathbf{y}\in\mathcal{Y}}$$

$$(K(\mathbf{x}|g(\mathbf{y})) \in \Omega)_{(\mathbf{x},\mathbf{y})\in\overline{\mathcal{X}\mathcal{Y}}} \,, \quad (L(\mathbf{y}|f(\mathbf{x})) \in \Omega)_{(\mathbf{x},\mathbf{y})\in\overline{\mathcal{X}\mathcal{Y}}}$$

where $[M_X] = \{1, \ldots, M_X\}$, $[M_Y] = \{1, \ldots, M_Y\}$, and Ω is a finite set, in such a way that

$$K(\mathbf{x}|g(\mathbf{y})) = L(\mathbf{y}|f(\mathbf{x})), \ \textit{for all} \ (\mathbf{x}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}} \tag{1}$$

and

$$|\,\Omega(f, g, K, L)\,| \ \longrightarrow \ \max$$

where

$$\Omega(f, g, K, L) = \Big\{ \, K(\mathbf{x}|g(\mathbf{y})) : \ (\mathbf{x}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}} \Big\} = \Big\{ \, L(\mathbf{y}|f(\mathbf{x})) : \ (\mathbf{x}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}} \Big\}. \tag{2}$$

The notations above are illustrated in Figure 1.

3 Examples

Let $M_X = 4$, $M_Y = 5$ and let X and Y have access to a source generating pairs of binary vectors $(\mathbf{x}, \mathbf{y})$ of length 6 in such a way that the vector $\mathbf{x}$ has 3 ones, the vector $\mathbf{y}$ has the 2 ones, and the Hamming distance between $\mathbf{x}$ and $\mathbf{y}$ is equal to 1,

$$\begin{aligned} \mathcal{X} \times \mathcal{Y} &= \{0,1\}^6_3 \times \{0,1\}^6_2 \\ \overline{\mathcal{X}\mathcal{Y}} &= \{ \, (\mathbf{x}, \mathbf{y}) \in \mathcal{X} \times \mathcal{Y} : \ d_H(\mathbf{x}, \mathbf{y}) = 1 \, \}. \end{aligned} \tag{3}$$

The set $\overline{\mathcal{X}\mathcal{Y}}$ can be also specified by the matrix shown in Table 1, where the $*$ symbols mark all pairs belonging to the set $\overline{\mathcal{X}\mathcal{Y}}$ and we use the octal representation for the binary vectors (07, 13, ... denote the vectors 000111, 001011, ...). Thus, the source generates one of

$$\binom{6}{3}\binom{3}{1} = \binom{6}{2}\binom{4}{1} = 60$$

pairs of vectors. Person X partitions the set $\mathcal{X}$ into M_X subsets, determines the number of the subset $f(\mathbf{x})$ containing $\mathbf{x}$, and sends this number to Y. An example of the partitioning is presented in Table 2, where

$$\begin{aligned} \mathbf{x} \in \{07, 13, 15, 16, 23\} &\implies f(\mathbf{x}) = 1 \\ \mathbf{x} \in \{25, 26, 31, 32, 34\} &\implies f(\mathbf{x}) = 2 \\ \mathbf{x} \in \{43, 45, 46, 51, 52\} &\implies f(\mathbf{x}) = 3 \\ \mathbf{x} \in \{54, 61, 62, 64, 70\} &\implies f(\mathbf{x}) = 4. \end{aligned}$$

At the same time, person Y partitions the set $\mathcal{Y}$ into M_Y subsets, determines the number of the subset $g(\mathbf{y})$ containing $\mathbf{y}$, and also sends this number to X. Thus, the participants represent the source as a collection of $M_X M_Y$ sub–sources having the alphabets $\mathcal{X}_i \times \mathcal{Y}_j$, where

$$\mathcal{X}_i = \{\, \mathbf{x} \in \mathcal{X} : \; f(\mathbf{x}) = i \,\}, \;\; \mathcal{Y}_j = \{\, \mathbf{y} \in \mathcal{Y} : \; g(\mathbf{y}) = j \,\}$$

for all $i \in [M_X]$ and $j \in [M_Y]$.

In the considerations below we assume the partitioning of the sets $\mathcal{X}$ and $\mathcal{Y}$ specified in Table 2. Suppose that $\mathbf{x} = 45$. If X receives $g(\mathbf{y}) = 1$, then he knows that $\mathbf{y} = 05$. Person Y receives $f(\mathbf{x}) = 3$ in this case and he knows that $\mathbf{x} = 45$. Therefore the pair of vectors $(45, 05)$ describes the common knowledge of X and Y. Let person X receive $g(\mathbf{y}) = 5$ and know that $\mathbf{y} = 44$. In this case, Y having received the message 3 only knows that the vector observed by X belongs to the set $\{45, 46\}$. Person X imagines that he is Y and also knows this. Since X wants to establish a common knowledge with Y, he replaces the vector 45 by the set $\{45, 46\}$, and the pair $(\{45, 46\}, 44)$ becomes the common value. At last, if $g(\mathbf{y}) = 4$, then several iterations of the estimating procedure lead to a description of the common knowledge by the pair of sets $(\{43, 45, 46, 51, 52\}, \{41, 42\})$. Inspecting this procedure for all pairs $(\mathbf{x}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}$, we come to the conclusion that the participants can agree on one of 26 common values, and if these values are associated with the capital letters of the Latin alphabet, we write $\Omega(f, g, K, L) = \{\, \mathrm{A}, \mathrm{B}, \ldots, \mathrm{Z} \,\}$; the common values corresponding to the observed vectors are shown in Table 2. Another partitioning of the sets $\mathcal{X}$ and $\mathcal{Y}$ given in Table 3 leads to 50 possible common values, which can be associated with the letters $\mathrm{A}, \mathrm{B}, \ldots, \mathrm{Z}, \mathrm{a}, \mathrm{b}, \ldots, \mathrm{x}$. In this case,

$$\begin{aligned}
\mathcal{Y}_1 &= \{03, 14, 60\} = \{000011, 001100, 110000\} \\
\mathcal{Y}_2 &= \{06, 30, 41\} = \{000110, 011000, 100001\} \\
\mathcal{Y}_3 &= \{05, 12, 24\} = \{000101, 001010, 010100\} \\
\mathcal{Y}_4 &= \{21, 42, 50\} = \{010001, 100010, 101000\} \\
\mathcal{Y}_5 &= \{11, 22, 44\} = \{001001, 010010, 100100\}.
\end{aligned}$$

Thus, $\mathcal{Y}_1, \mathcal{Y}_2$, and $\mathcal{Y}_5$ are binary block codes having the minimum distance 4, and the vector $\mathbf{y}$ is uniquely determined by X based on the vector $\mathbf{x}$ if he receives the messages 1, 2, or 5. Furthermore, there is only 1 vector $\mathbf{x} = 010101$ when X has the ambiguity about $\mathbf{y}$ if the message is 3 and only 1 such a vector $\mathbf{x} = 101010$ if the message is 4.

The general procedure for constructing the common values can be described in at least two different ways. Suppose that there is a matrix of dimension $|\mathcal{X}| \times |\mathcal{Y}|$ containing the $*$ symbols at the positions corresponding to the pairs of vectors $(\mathbf{x}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}$ and gaps at all other positions. We permute the rows and the columns of this matrix in accordance with the functions f, g and split the resulting matrix into $M_X M_Y$ rectangles. Then any two pairs of vectors generate the same common value if and only if there exists a "path" connecting the corresponding $*$ symbols, which completely belongs to this rectangle and may turn by 90 degrees passing through any of the $*$ symbols. Another way of representing this procedure relates to bipartite graphs (see Figure 3) : given $(i, j) \in [M_X] \times [M_Y]$, we introduce a bipartite graph having left and right sides; the vertices at the left side of the graph are the vectors $\mathbf{x} \in \mathcal{X}_i$ and the vertices at the right side are the vectors $\mathbf{y} \in \mathcal{Y}_j$; any two vertices $\mathbf{x}$ and $\mathbf{y}$ are connected by an edge if and only if $(\mathbf{x}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}$. In the next section we prove the statement that the different values of the functions K, L can be assigned only to distinct connected components of the bipartite graphs constructed for all $i \in [M_X]$ and $j \in [M_Y]$, where the term "connected component" is taken in the classical graph theory sense : any two vertices belong to the same connected component if and only if there is a path connecting these vertices.

In further extensions of the problem these considerations should be represented in a more general form; in particular, empty sets can be the set of vertices belonging to the connected components.

4 Separability of a common knowledge

Any two pairs of vectors $(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2) \in \overline{\mathcal{X}\mathcal{Y}}$ can be separated by the participants either based on the description of the source when $(\mathbf{x}_1, \mathbf{y}_2), (\mathbf{x}_2, \mathbf{y}_1) \notin \overline{\mathcal{X}\mathcal{Y}}$ or based on different pairs of messages corresponding to these vectors. If both criteria cannot be used, then X and Y have to come to the same common value in the cases when they are given $(\mathbf{x}_1, \mathbf{y}_1)$ and $(\mathbf{x}_2, \mathbf{y}_2)$. However we speak about the "separability of the knowledge" in another context and present the discussion in the end of the section.

For all $i \in [M_X]$, $j \in [M_Y]$, and sets $\varphi \subseteq \mathcal{X}$, $\psi \subseteq \mathcal{Y}$, let

$$\begin{aligned} \mathcal{F}_j(\varphi) &= \Big\{ \mathbf{y} \in \mathcal{Y}_j : \ (\mathbf{x}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}, \ \text{ for some } \mathbf{x} \in \varphi \Big\} \\ \mathcal{G}_i(\psi) &= \Big\{ \mathbf{x} \in \mathcal{X}_i : \ (\mathbf{x}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}, \ \text{ for some } \mathbf{y} \in \psi \Big\}. \end{aligned}$$

Definition :

(X) *For all $i \in [M_X]$, $j \in [M_Y]$, and $\mathbf{x} \in \mathcal{X}_i$, the set $\varphi(\mathbf{x}|j)$ is the j–th ghost of the vector $\mathbf{x}$ if and only if $\varphi(\mathbf{x}|j) = \varphi$, where $\varphi = \emptyset$ when $\mathcal{F}_j(\{\mathbf{x}\}) = \emptyset$ and φ is the non–empty set of minimal cardinality which satisfies the conditions*

$$\begin{cases} \mathbf{x} \in \varphi \\ \varphi = \mathcal{G}_i\left(\mathcal{F}_j(\varphi) \right) \end{cases} \tag{4}$$

when $\mathcal{F}_j(\{\mathbf{x}\}) \neq \emptyset$.

(Y) *For all $i \in [M_X]$, $j \in [M_Y]$, and $\mathbf{y} \in \mathcal{Y}_j$, the set $\psi(\mathbf{y}|i)$ is the i–th ghost of the vector $\mathbf{y}$ if and only if $\psi = \emptyset$ when $\mathcal{G}_i(\{\mathbf{y}\}) = \emptyset$ and ψ is the non–empty set of minimal cardinality which satisfies the conditions*

$$\begin{cases} \mathbf{y} \in \psi \\ \psi = \mathcal{F}_j\left(\mathcal{G}_i(\psi) \right) \end{cases} \tag{5}$$

when $\mathcal{G}_i(\{\mathbf{y}\}) \neq \emptyset$.

Lemma :

(a) *For all $\mathbf{x} \in \mathcal{X}$ and $\mathbf{y} \in \mathcal{Y}$, the sets $\varphi(\mathbf{x}|1), \ldots, \varphi(\mathbf{x}|M_Y)$ and $\psi(\mathbf{y}|1), \ldots, \psi(\mathbf{y}|M_X)$ are uniquely defined by* **(X)**, **(Y)**.

(b) *If $(\mathbf{x}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}$, then*

$$\begin{aligned} \left(\varphi(\mathbf{x}|g(\mathbf{y})), \psi(\mathbf{y}|f(\mathbf{x})) \right) &= \Big(\varphi(\mathbf{x}|g(\mathbf{y})), \mathcal{F}_{g(\mathbf{y})}(\varphi(\mathbf{x}|g(\mathbf{y})) \Big) & (6) \\ &= \Big(\mathcal{G}_{f(\mathbf{x})}(\psi(\mathbf{y}|f(\mathbf{x})), \psi(\mathbf{y}|f(\mathbf{x})) \Big) & (7) \end{aligned}$$

i.e., the pair of sets $\left(\varphi(\mathbf{x}|g(\mathbf{y})), \psi(\mathbf{y}|f(\mathbf{x})) \right)$ is uniquely determined both by $(\mathbf{x}, g(\mathbf{y}))$ and $(\mathbf{y}, f(\mathbf{x}))$.

(c) *For all* $\mathbf{x}' \in \mathcal{X}$ *and* $\mathbf{y}' \in \mathcal{Y}$,

$$\mathbf{x}' \in \varphi(\mathbf{x}|g(\mathbf{y})) \implies K(\mathbf{x}'|g(\mathbf{y})) = K(\mathbf{x}|g(\mathbf{y})) \tag{8}$$
$$\mathbf{y}' \in \psi(\mathbf{y}|f(\mathbf{x})) \implies L(\mathbf{y}'|f(\mathbf{x})) = L(\mathbf{y}|f(\mathbf{x})). \tag{9}$$

Theorem : *Let*

$$\overline{\Phi\Psi}(f,g) = \Big\{ \, (\, \varphi(\mathbf{x}|g(\mathbf{y})), \psi(\mathbf{y}|f(\mathbf{x})) \,) \in 2^{\mathcal{X}} \times 2^{\mathcal{Y}} : \ (\mathbf{x},\mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}} \Big\}.$$

The functions f, g, K, L *satisfy the restriction (1) if and only if there exists a function* $\Theta : \ \overline{\Phi\Psi}(f,g) \to \Omega$ *defined by the values*

$$(\, \Theta\, (\varphi, \psi) \in \Omega \,)_{(\varphi,\psi) \in \overline{\Phi\Psi}(f,g)}$$

such that

$$K(\mathbf{x}|g(\mathbf{y})) = \Theta\, (\, \varphi(\mathbf{x}|g(\mathbf{y})), \psi(\mathbf{y}|f(\mathbf{x})) \,), \ \ \textit{for all} \ \ (\mathbf{x},\mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}} \tag{10}$$

and

$$L(\mathbf{y}|f(\mathbf{x})) = \Theta\, (\, \varphi(\mathbf{x}|g(\mathbf{y})), \psi(\mathbf{y}|f(\mathbf{x})) \,), \ \ \textit{for all} \ \ (\mathbf{x},\mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}. \tag{11}$$

Corollary : *Given functions* f *and* g,

$$\max_{K,L} |\, \Omega(f,g,K,L) \,| = |\, \overline{\Phi\Psi}(f,g) \,| \tag{12}$$

where the maximum is taken over the functions K, L *satisfying (1).*

The set $\varphi(\mathbf{x}|j)$ satisfying $(\mathbf{X})$ can be constructed by the following recurrent procedure. We introduce the sequence of sets $\varphi^{(0)} = \{\mathbf{x}\}, \varphi^{(1)}, \ldots \subseteq \mathcal{X}_i$, where

$$\varphi^{(t)} = \mathcal{G}_i \left(\mathcal{F}_j(\varphi^{(t-1)}) \right), \ \ t = 1, 2, \ldots \tag{13}$$

Then all but a finite number of sets in this sequence coincide with some set φ, and $\varphi(\mathbf{x}|j) = \varphi$. In other words, we consider $\mathbf{x}$ as the 1–element set $\{\mathbf{x}\}$ and use the possibility of extending this set. Returning to the applications, one can say that the individuals have to replace the vectors received from the source with their ghosts and extend these ghosts until they become common. The theorem claims that this is the only way of creating a common knowledge in the sense of (1).

The result of the theorem can be interpreted using the scheme given in Figure 2 where X and Y are given the vectors generated by the source, but they do not communicate to each other. Different values of the function g correspond to different partitions of the space $\mathcal{X}$. If X knows g, but does not know the value of $g(\mathbf{y})$, then he considers all possible values and outputs the vector $\boldsymbol{\varphi}$ containing M_Y ghosts of the vector $\mathbf{x}$. For example, if the functions f and g are defined by Table 2 and $\mathbf{x} = 45$ (see also Figure 3), then $\boldsymbol{\varphi} = (\{45\}, \emptyset, \emptyset, \{43, 45, 46, 51, 52\}, \{45, 46\})$. By $(\mathbf{X})$, at least one component of the vector $\boldsymbol{\varphi}$ is a non–empty set and the value of the function f is a constant $f(\mathbf{x})$ for every non–empty set. Similarly, Y outputs the vector $\boldsymbol{\psi}$ containing M_X ghosts of the vector $\mathbf{y}$. The pair of vectors $(\boldsymbol{\varphi}, \boldsymbol{\psi})$ goes to the person called the decoder (in the situation considered before, the decoder is simultaneously X and Y), and the decoding algorithm is fixed as follows : the decoder first extracts $f(\mathbf{x})$ from $\boldsymbol{\varphi}$ and uses this value as a pointer to the vector $\boldsymbol{\psi}$, then he extracts $g(\mathbf{y})$ from $\boldsymbol{\psi}$, uses this value as a pointer to the vector $\boldsymbol{\varphi}$, and finally outputs $\big(\varphi(\mathbf{x}|g(\mathbf{y})), \psi(\mathbf{y}|f(\mathbf{x}))\big)$. If $(\mathbf{x},\mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}$, then both components of this vector are non–empty sets, and this fact can be viewed as the matching of $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$. In this case, the result of the decoding is the pair of elements of dependent partitions of the sets $\mathcal{X}$ and $\mathcal{Y}$. Note that the same result can be also obtained if X knows $g(\mathbf{y})$ and Y knows $f(\mathbf{x})$: then each person outputs only one subset, and the decoder combines them into a pair. Such a conclusion is possible only because the subsets presented by the persons are not arbitrarily chosen, but uniquely defined by the functions f and g in accordance with **(X)**, **(Y)**. In other words, the individuals establish a common knowledge only if they construct and present certain vectors $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ that can be viewed as their individual contributions to the common knowledge. That is, the "common knowledge" can be achieved with the use of the source only if there exist separated individual knowledges matching each other.

5 Proofs

5.1 Proof of Lemma

Statement (a) If there exist two sets, φ_1 and φ_2, satisfying (4), then their intersection, $\varphi_1 \cap \varphi_2$, also satisfies (4). Since $\varphi(\mathbf{x}|j)$ is the set of minimal cardinality satisfying (4) and the total number of subsets of $\mathcal{X}$ is finite, this set is uniquely defined by **(X)**. Similar considerations with the use of (5) prove that the set $\psi(\mathbf{y}|i)$ is uniquely defined by **(Y)**.

Statement (b) Let $(\mathbf{x},\mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}$ and $\mathbf{x} \in \mathcal{X}_i$, $\mathbf{y} \in \mathcal{Y}_j$. We refer to the algorithm for constructing the sequence of sets $\varphi^{(0)} = \{\mathbf{x}\}, \varphi^{(1)}, \ldots \subseteq \mathcal{X}_i$, which is recurrently defined by (13), and construct another sequence $\varphi^{(0)\prime} = \mathcal{G}_i(\{\mathbf{y}\}), \varphi^{(1)\prime}, \ldots \subseteq \mathcal{Y}_j$ by

$$\varphi^{(t)\prime} = \mathcal{G}_i\left(\mathcal{F}_j(\varphi^{(t-1)\prime})\right), \quad t = 1, 2, \ldots$$

If φ' is the set belonging to this sequence such that all but a finite number of sets differ from φ', then

$$\{\mathbf{y}\} \subseteq \mathcal{F}_j(\{\mathbf{x}\}) \implies \mathcal{G}_i(\{\mathbf{y}\}) = \varphi^{(0)\prime} \subseteq \varphi^{(1)} \implies \varphi' \subseteq \varphi.$$

However φ is the set of minimal cardinality satisfying (4). Thus $\varphi' = \varphi$, and (6) follows. The proof of (7) is similar.

Statement (c) By the requirement for $\varphi(\mathbf{x}|g(\mathbf{y}))$ to be the set of minimal cardinality satisfying (4), $\mathbf{x}' \in \varphi(\mathbf{x}|g(\mathbf{y}))$ implies the existence of a sequence of $\lambda \geq 2$ vectors $\mathbf{x}_1, \ldots, \mathbf{x}_\lambda \in \varphi(\mathbf{x}|g(\mathbf{y}))$ such that $\mathbf{x}_1 = \mathbf{x}$, $\mathbf{x}_\lambda = \mathbf{x}'$, and

$$\left\{\mathbf{y} \in \mathcal{Y}_j : \ (\mathbf{x}_\mu, \mathbf{y}), (\mathbf{x}_{\mu-1}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}\right\} \neq \emptyset, \ \text{ for all } \ \mu \in \{2, \ldots, \lambda\}.$$

By (1),

$$K(\mathbf{x}_{\mu-1}|g(\mathbf{y})) = L(\mathbf{y}|f(\mathbf{x}_{\mu-1})), \quad K(\mathbf{x}_\mu|g(\mathbf{y})) = L(\mathbf{y}|f(\mathbf{x}_\mu))$$

for any vector $\mathbf{y}$ belonging to this set and $f(\mathbf{x}_{\mu-1}) = f(\mathbf{x}_\mu)$ by **(X)**. Thus $K(\mathbf{x}_{\mu-1}|g(\mathbf{y})) = K(\mathbf{x}_\mu|g(\mathbf{y}))$. Using this argument for $\mu = 2, \ldots, \lambda$ we obtain $K(\mathbf{x}_1|g(\mathbf{y})) = K(\mathbf{x}_\lambda|g(\mathbf{y}))$ and prove (8). The proof of (9) is similar.

5.2 Proof of Theorem

Direct statement : *if there is a function Θ such that (10) and (11) hold, then the functions K, L satisfy (1).*

By (10) and (11), we write

$$K(\mathbf{x}|g(\mathbf{y})) = \Theta\left(\varphi(\mathbf{x}|g(\mathbf{y})), \psi(\mathbf{y}|f(\mathbf{x})) \right) = L(\mathbf{y}|f(\mathbf{x}))$$

for all $(\mathbf{x}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}$ and obtain (1).

Converse statement : *if the functions K, L satisfy (1), then there is a function Θ such that (10) and (11) hold.*

By (8), $K(\mathbf{x}|g(\mathbf{y}))$ is a function of $\varphi(\mathbf{x}|g(\mathbf{y}))$ and $g(\mathbf{y})$. The integer $g(\mathbf{y})$ is available from any of the sets $\psi(\mathbf{y}|1), \ldots, \psi(\mathbf{y}|M_X)$ since

$$\mathbf{y}' \in \psi(\mathbf{y}|i) \implies g(\mathbf{y}') = g(\mathbf{y}), \quad \text{for any } i \in [M_X].$$

In particular, $g(\mathbf{y})$ can be extracted from the set $\psi(\mathbf{y}|f(\mathbf{x}))$, and $K(\mathbf{x}|g(\mathbf{y}))$ can be represented as a function Θ_K depending on $\varphi(\mathbf{x}|g(\mathbf{y}))$ and $\psi(\mathbf{y}|f(\mathbf{x}))$, i.e.,

$$K(\mathbf{x}, g(\mathbf{y})) = \Theta_K\left(\varphi(\mathbf{x}|g(\mathbf{y})), \psi(\mathbf{y}|f(\mathbf{x})) \right). \tag{14}$$

Similar considerations with the use of (9) prove that $L(\mathbf{y}|f(\mathbf{x}))$ can be represented as a function Θ_L depending on the same arguments, i.e.,

$$L(\mathbf{y}|f(\mathbf{x})) = \Theta_L\left(\varphi(\mathbf{x}|g(\mathbf{y})), \psi(\mathbf{y}|f(\mathbf{x})) \right). \tag{15}$$

The coincidence of $K(\mathbf{x}|g(\mathbf{y}))$ and $L(\mathbf{y}|f(\mathbf{x}))$ for all $(\mathbf{x}, \mathbf{y}) \in \overline{\mathcal{X}\mathcal{Y}}$ and (14), (15) imply the coincidence of the functions Θ_K and Θ_L. We denote $\Theta = \Theta_K = \Theta_L$ and using (2) obtain (10).

Corollary : Let $\Omega^*(f, g)$ denote the set of values of the function Θ. Then

$$|\Omega^*(f, g)| \le |\overline{\Phi\Psi}(f, g)| \tag{16}$$

since Θ is a mapping $\overline{\Phi\Psi}(f, g) \to \Omega$. On the other hand, by (10) and (11),

$$|\Omega^*(f, g)| = |\Omega(K, L, f, g)| \tag{17}$$

where we use the notations (2). Combining (16) and (17) we obtain

$$|\Omega(K, L, f, g)| \le |\overline{\Phi\Psi}(f, g)|. \tag{18}$$

If Θ is a one-to-one mapping, then (16) holds with the equality, which also implies the equality in (18), and (12) follows.

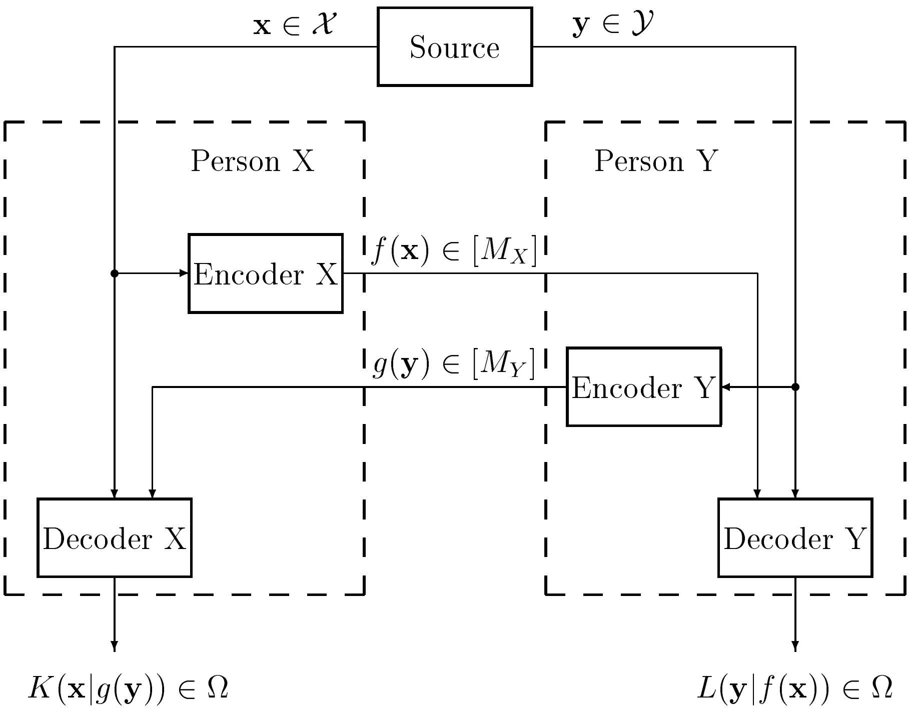

Figure 1: Model of creating a common knowledge by correlated observations and transmission over helping channels.

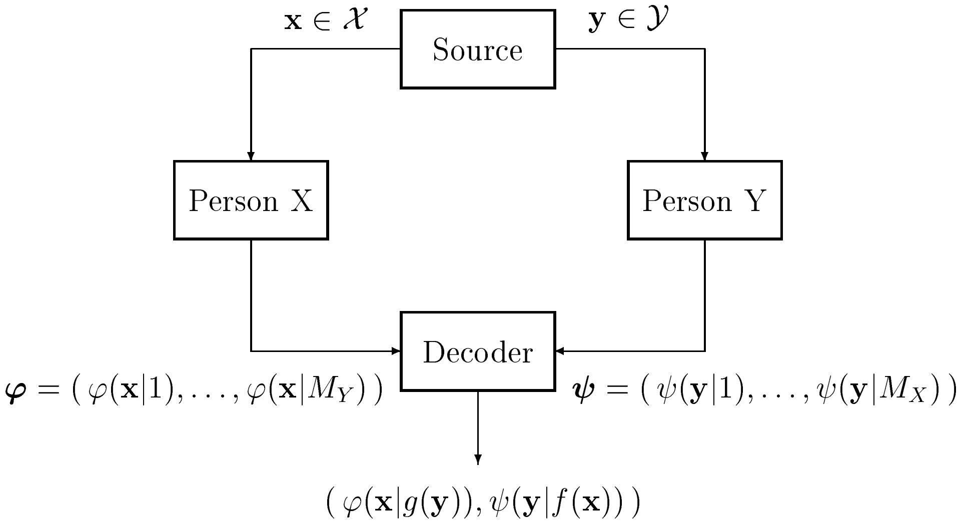

Figure 2: Logical scheme of creating a common knowledge.

Table 1: Structure of the set $\overline{\mathcal{XY}}$ defined in (3), where the binary vectors of length 6 are given in the octal representation.

	03	05	06	11	12	14	21	22	24	30	41	42	44	50	60
07	*	*	*												
13	*			*	*										
15		*		*		*									
16			*		*	*									
23	*						*	*							
25		*					*		*						
26			*					*	*						
31				*			*			*					
32					*			*		*					
34						*			*	*					
43	*										*	*			
45		*									*		*		
46			*									*	*		
51				*							*			*	
52					*							*		*	
54						*							*	*	
61							*				*				*
62								*				*			*
64									*				*		*
70										*				*	*

Table 2: The 26 common values A, ..., Z that can be constructed by the participants when X transmits one of 4 messages, Y transmits one of 5 messages, and the set $\overline{\mathcal{X}\mathcal{Y}}$ is defined in (3).

	03	05	06	11	12	14	21	22	24	30	41	42	44	50	60
07	A	A	A												
13	A			B	B										
15		A		B		B									
16			A		B	B									
23	A						C	C							
25		D					I		I						
26			E					I	I						
31				F			I			J					
32					G			I		J					
34						H			I	J					
43	K										P	P			
45		L									P		Q		
46			M									P	Q		
51				N							P			R	
52					O							P		R	
54						S							Z	Z	
61							T				X				Z
62								U				Y			Z
64									V				Z		Z
70										W				Z	Z

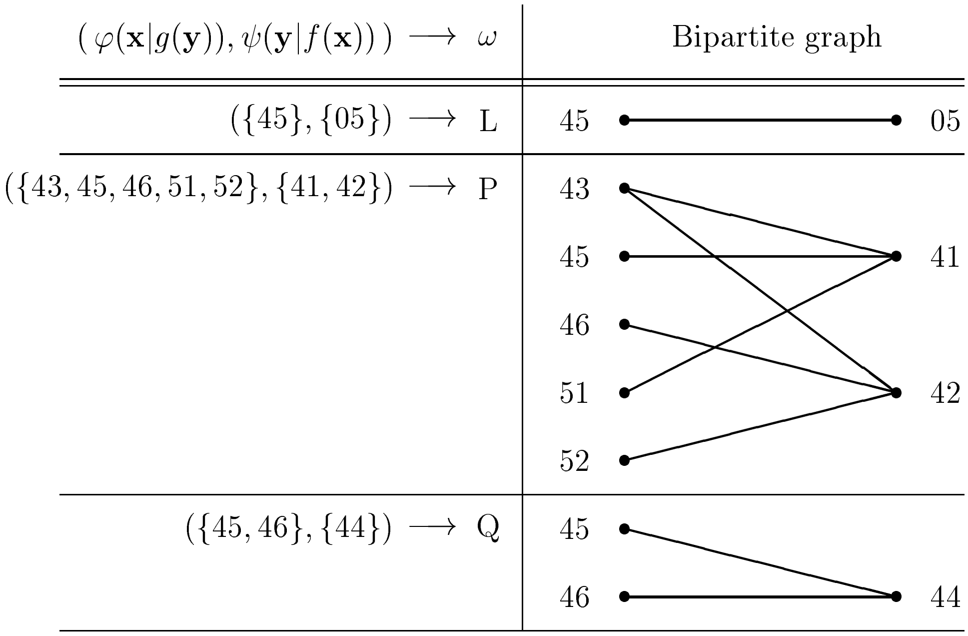

Figure 3: The common values ω and the corresponding bipartite graphs when $\mathbf{x} = 45$; the set $\overline{\mathcal{X}\mathcal{Y}}$ and the functions f, g are specified in Table 2.

Table 3: The 50 common values A, ..., Z that can be constructed by the participants when X transmits one of 4 messages, Y transmits one of 5 messages, and the set $\overline{\mathcal{X}\mathcal{Y}}$ is defined in (3).

	03	14	60	06	30	41	05	12	24	21	42	50	11	22	44
07	A			D			G								
25							G		G	I					
32					E			H						K	
54		B										J			L
61			C			F				I					
16		M		O				S							
34		M			P				T						
45						Q	R								Y
51						Q						V	W		
62			N								U			X	
26				b					f					j	
31					c					g			i		
43	Z					d					h				
52								e			h	h			
64			a						f						k
13	l							r					v		
15		m					q						v		
23	l									s				w	
46				o							t				x
70			n		p							u			

References

[1] Shannon, C.E. (1961) "Two–way communication channels," in *Claude Elwood Shannon : Collected Papers.* N.J.A. Sloane and A.D. Wyner (eds.). New York: IEEE Press, 351–384 (1993).
The paper was published in the *Proc. 4–th Berkley Symp. Math. Stat. and Prob.,* 611–644.

[2] Ahlswede, R. (1971). "Multi–way communication channels," in *2nd Int. Symp. Inform. Theory; Tsahkadzor, Armenian SSR, 1971.* Publishing House of the Hungarian Academy of Sciences, 23–52 (1973).

[3] Gács, P., and Körner, J. (1973). "Common information is far less than mutual information," *Probl. Inform. Control,* 2(2), 149–162.

[4] Slepian, D., and Wolf, J.K. (1973). "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory,* 19(4), 772–777.

[5] Ahlswede, R. (1974). "The capacity region of a channel with two senders and two receivers," *Ann.Prob.,* 2(5), 805–814.

[6] Wyner, A.D. (1975). "On source coding with side information at the decoder," *IEEE Trans. Inform. Theory,* 21(3), 294–300.

[7] Ahlswede, R., and Körner, J. (1975). "Source coding with side information and a converse for degraded broadcast channels," *IEEE Trans. Inform. Theory,* 21(6), 629–637.

[8] Körner, J., and Marton, K. (1979). "How to encode modulo–two sum of binary sources," *IEEE Trans. Inform. Theory,* 25(2), 219–221.

[9] Ahlswede, R., and Dueck, G. (1989) "Identification in the presence of feedback – A discovery of new capacity formulas," *IEEE Trans. Inform. Theory,* 35(1), 30–36.

[10] Maurer, U. (1993). "Secret key agreement by public discussion from common information," *IEEE Trans. Inform. Theory,* 39(3), 733–742.

[11] Ahlswede, R., and Csiszár, I. (1993). "Common randomness in information theory and cryptography – Part I : Secret sharing," *IEEE Trans. Inform. Theory,* 39(5), 1121–1131.

[12] Ahlswede, R., and Balakirsky, V.B. (1996). "Identification under random processes," *Problemy Peredachi Informatsii*, special issue honoring Mark S. Pinsker, 32(1), 144–160 (in Russian). English translation : *Probl. Inform. Transmission,* 32, 123–138.

[13] Ahlswede, R., and Csiszár, I. (1998). "Common randomness in information theory and cryptography – Part II : CR capacity," *IEEE Trans. Inform. Theory,* 44(1), 225–240.