JOURNAL OF APPLIED
COMPUTER SCIENCE
Vol. xx No. xx (xxxx), pp. xx–xx

A note on Shannon capacity for invariant and evolving channels

Marcin Jurkiewicz[1], **Marek Kubale**[1]

[1]*Gdansk University of Technology*
Department of Algorithms and System Modelling
Gabriela Narutowicza 11/12, 80-233 Gdansk, Poland
marcin.jurkiewicz@eti.pg.gda.pl
kubale@eti.pg.gda.pl

Abstract. *In the paper we discuss the notion of Shannon capacity for invarianta and evolving channels. We show how this notion is involved in information theory, graph theory and Ramsey theory.*
Keywords: *Shannon capacity, information channel*

1. Introduction

Informally, the capacity C is the maximum rate at which one can transmit information over a noisy channel and receive this information on the output, with nearly zero probability of error. Sometimes, it is necessary to receive an information with exactly zero probability of error. For this purpose we use a measure called the Shannon capacity C_0, as defined later on. Unfortunately, C_0 is much more difficult to calculate than C.

1.1. Strong product and independence number of graphs

In this paper we need the following definition

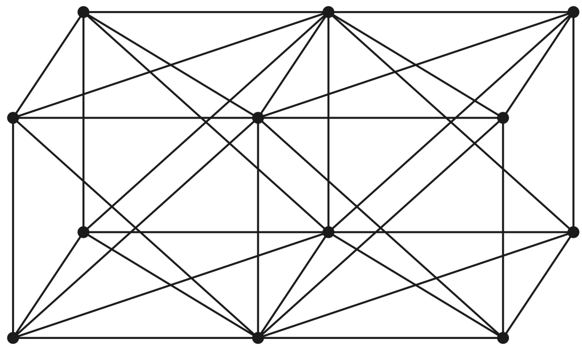

Figure 1. Example of the strong product $P_3 \boxtimes C_4$

Definition 1. *Given two graphs $G = (V(G), E(G))$ and $H = (V(H), E(H))$, the strong product $G \boxtimes H$ is defined as follows. The vertices of $G \boxtimes H$ are all pairs of the Cartesian product $V(G) \times V(H)$. There is an edge between (x, x') and (y, y') iff one of the following holds*

- $\{x, y\} \in E(G)$ *and* $\{x', y'\} \in E(H)$,
- $x = y$ *and* $\{x', y'\} \in E(H)$
- $x' = y'$ *and* $\{x, y\} \in E(G)$

See Figure 1 for an example. We write G^n to denote $G \boxtimes G \boxtimes \ldots \boxtimes G$, where G occurs n times. Let G be a graph. A set of vertices S of G is said to be an independent set of vertices if they are pairwise nonadjacent. The independence number of G, denoted by $\alpha(G)$, is defined to be the size of a largest independent set of G. Let n be a natural number and $G_1, G_2, ..., G_n$ be graphs, then for the strong product the following inequality holds

$$\alpha(G_1 \boxtimes G_2 \boxtimes \ldots \boxtimes G_n) \geq \alpha(G_1)\alpha(G_2)\ldots\alpha(G_n). \tag{1}$$

From (1), we know that α is super-multiplicative, i.e.

$$\alpha(G^n \boxtimes G^m) \geq \alpha(G^n)\alpha(G^m), \tag{2}$$

for any pairs of integers n, m.

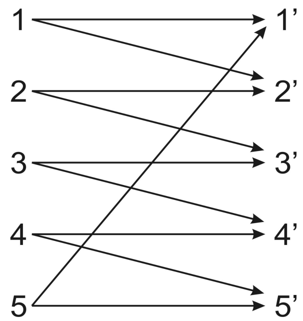

Figure 2. Example of discrete channel with $A_X = \{1, 2, 3, 4, 5\}$ as the input alphabet and $A_Y = \{1', 2', 3', 4', 5'\}$ as the output alphabet. If $p(y|x) > 0$, then $x \in A_X$ is adjacent to $y \in A_Y$.

1.2. Ramsey numbers

The Ramsey number $R(l_1, ..., l_k)$ is defined as the smallest integer n such that, no matter how each 2-element subset of an n-element set is colored with k colors, there exists an $i \leq k$ such that there is a subset of size l_i, all of whose 2-element subsets are color i. For instance, $R(3, 3) = 6$, $R(4, 4) = 18$ [1]. In information theory it is convenient to start enumeration from zero. Hence, we define the following function

$$r(l_1, \ldots, l_k) = R(l_1 + 1, \ldots, l_k + 1) - 1. \tag{3}$$

If $l = l_1 = \ldots = l_k$, then the above equation takes on the following short form: $r_k(l) = R_k(l + 1) - 1$.

1.3. Channel definition

Further, we will need the following notion [2]

Definition 2. *A discrete channel $Q = (A_X, A_Y, M_{XY})$ consists of three parts: A_X and A_Y are input and output alphabets, respectively, M_{XY} is the transition matrix with $p(y|x)$ elements, which determine conditional probabilities.*

We say that a channel Q is noisy, if there are elements $y_1, y_2 \in A_Y$ and an element $x \in A_X$ such that $p(y_1|x)p(y_2|x) > 0$. The channel from Figure 2 is noisy.

2. Shannon capacity for invariant and evolving channels

2.1. Shannon capacity for invariant channels

Sometimes we have to transmit information with probability of error equal to zero. In 1956 Shannon [3] defined the capacity in a zero-error communication model. The best way to explain this model is to give an example. Let Q be a noisy channel like that shown in Figure 2. Suppose that we want to send some element e from an input alphabet A_X and get element e' from an output alphabet A_Y. But sometimes we get on output the value (e' mod 5)+1. Hence, we are not certain, what in fact has been sent by the channel. For example, if we get on output 3', there is 2 or 3, which could be sent on input. So this must be done in a different way. The model assumes that both parts must make an appointment, which elements from the input alphabet are used. For instance, for the channel Q we can choose 1 and 3. We do so because, if we get on output 1' or 2', then we are certain that there was 1 on input. Similarly, if we get on output 3' or 4', then we are sure that there was 3 on input. So, if we use the channel Q ones with elements $\{1, 3\}$, then we can transmit one bit of information correctly. It is worth emphasizing, that in this model, we cannot choose more than two elements from alphabet A_X.

Let us generalize this. Given a channel $Q = (A_X, A_Y, M_{XY})$ and element $x \in A_X$, we define

$$S_x = \{y \in A_Y : p(y|x) > 0\}. \tag{4}$$

We can say that S_x is the set of letters attainable on output, when there is x on input. Associated with this set and the channel Q is a characteristic graph defined as follows

Definition 3. *The characteristic graph of the channel Q is graph $G = (V, E)$ such that $V = A_X$ and $\{x, y\} \in E(G)$ iff $S_x \cap S_y \neq \emptyset$.*

For instance, the characteristic graph of the channel from Figure 2 is cycle C_5.

It turns out that for a graph G a subset S of vertices $V(G)$ guarantees zero-error transmission if and only if S is independent. In a single use of channel Q, the maximal number of letters that we can transmit is equal to $\alpha(G)$. Let us denote such a maximal set of letters by I, which had been agreed earlier by both sides. The largest amount of information that can be transmitted in a single use of Q, is equal to $\gamma(1) = \log_2(\alpha(G))$. In case of the maximal noisy channel, i.e. when G is complete, we cannot send any information, since $\gamma(1) = \log_2 1 = 0$. On the other hand, if G is empty then we can transmit $\log_2 |A_X|$ bits of information. For the channel with the characteristic graph C_5, there is possibility to send 1 bit, because $\alpha(C_5) = 2$.

Let us extend above definitions for a mass communication, i.e. a communication at which we use a channel Q many times. Let n denote the number of times that the channel is used. Suppose that there is a sequence $x_1, x_2, \ldots, x_n \in A_X$ on input and $y_1, y_2, \ldots, y_n \in A_Y$ on output, where $y_i \in S_{x_i}$ and $1 \leq i \leq n$. Notice that we can interpret n uses of the channel Q as a single use of a larger channel $Q(n)$. Then on input we take an element from a set $A_X(n) = A_X \times A_X \times \ldots \times A_X$, where A_X occurs n times. Analogously, on output we get an element from a set $A_Y(n)$. If $x(n) = x_1, x_2, ..., x_n \in A_X(n)$ is on input, then on output we get an element from a set

$$S_{x(n)} = S_{x_1} \times S_{x_2} \times \ldots \times S_{x_n}. \tag{5}$$

It is easy to observe that from an input sequences $x(n)$ and $v(n)$, we can obtain the same output sequence, if for all i, $x_i = v_i$ or x_i and v_i are adjacent in the characteristic graph of Q, so these two sequences can put the same letters on output.

Above considerations lead us to the strong product, i.e. the characteristic graph of $Q(n)$, defined in (6) as the strong product of characteristic graphs of Q

$$G(n) = G^n. \tag{6}$$

Hence, the largest amount of information that can be transmitted by $Q(n)$, is equal to $\gamma(n) = \log_2(\alpha(G^n))$. The inequality (1) implies that $\gamma(n) \geq \log_2(\alpha^n(G))$. It is interesting to note that for some channels, we can get better result than this lower bound, by applying the mass communication. For example $\alpha^2(C_5) = 4$, while $\alpha(C_5 \boxtimes C_5) = 5$ (in Fig. 3, the set of independent vertices is $\{11', 23', 35', 42', 54'\}$). From Hendrlín's result [4] we know that this value is maximum possible.

Theorem 4. *For graphs G and H*

$$\alpha(G \boxtimes H) \leq r(\alpha(G), \alpha(H)). \tag{7}$$

Thus considering (3)

$$\alpha(C_5 \boxtimes C_5) \leq r(2, 2) = R(3, 3) - 1 = 5. \tag{8}$$

Now, we have all necessary notions to define the Shannon capacity. Operationally, the Shannon capacity of a channel represents the effective size of the alphabet in a zero-error transmission. Below, we give a formal definition of this notion.

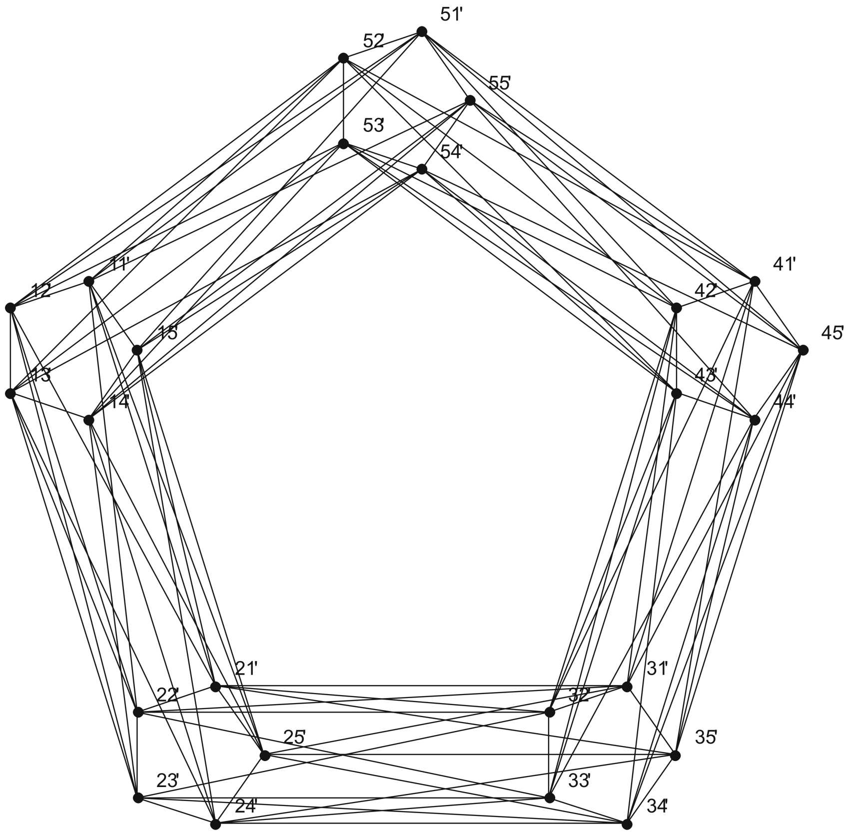

Figure 3. The strong product $C_5 \boxtimes C_5$.

Definition 5. *The Shannon capacity is defined as*

$$C_0 = sup_n \sqrt[n]{\alpha(G^n)}. \tag{9}$$

Since $\alpha(\cdot)$ is super-multiplicative and by Fekete's lemma [5]

$$C_0 = sup_n \sqrt[n]{\alpha(G^n)} = lim_{n\to\infty} \sqrt[n]{\alpha(G^n)}. \tag{10}$$

The calculation of the Shannon capacity is hard. For example, the Shannon capacity has not been found even for small cycle C_7. But we know that for even cycle it is equal to $\alpha(G) = n/2$ and for C_3, C_5 we have the value 1 and $\sqrt{5}$, respectively [6].

2.2. Shannon capacity for evolving channels

In the previous subsection we have introduced definition of the Shannon capacity. It is well know, that in the real world systems always evolve. Therefore, we are going to consider the Shannon capacity for an evolving channel.

Let Q be a noisy channel with the transition matrix M and the characteristic graph G. We assume that both sides of the trasmission know how the channel evolves. Let $\mathcal{O} = \{O_1, O_2, \ldots\}$ be a sequence of operations, which operate on Q (notice that O_i could be represented by a matrix operating on M). Thus, the characteristic graph also evolves. The operations on G coresponding to $\mathcal{O}$ we denote by $\mathcal{U} = \{U_1, U_2, \ldots\}$. Similarly to the definition of Shannon capacity for invariant channels, we have

Definition 6. *The Shannon capacity for evolving channel is defined as*

$$C_0 = sup_n \sqrt[n]{\alpha(\mathcal{U}_n(G))}. \tag{11}$$

where $\mathcal{U}_n(G) = U_n(U_{n-1}(\ldots(U_1(G))))$.

It is important to observe that for some operations $\mathcal{U}$ the above supremum does not exist and the measure does not work. Notice that if $\mathcal{U}_n(G) = G^n$, then we get the Shannon capacity for invariant channels.

Let A_X and A_Y be input and output alphabets. For our purpose, we can consider a noisy channel as a map $Q : A_X \to A_Y$, which maps each input letter to a set of possible output letters. Later on, we will consider a reduction of a channel.

Definition 7. *For a letter $v \in A_X$, we define the reduction of a channel Q as a transformation of Q to Q'*

$$Q'(x) = \begin{cases} Q(x), & \textit{if } x \neq v \\ A_Y, & \textit{otherwise.} \end{cases} \tag{12}$$

See Figure 4 for examples. We can imagine this situation in the following way. A system has changed, because the environment acts on it. There is more noise in the system. A letter $v \in A_X$ is not available. Now, v is confusable with other letters from A_X. In this transformation the characteristic graph G of Q with vertex set A_X is changing, in simple way, to the characteristic graph G' of Q'

$$G' = G - v \oplus v. \tag{13}$$

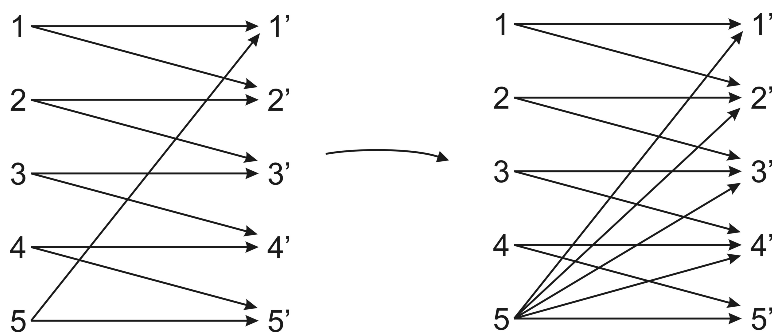

Figure 4. The reduction of the channel Q. Letter 5 has been reducted.

where the symbol $-$ means the vertex deletion and the symbol $\oplus$ means the graph join. More precisely, we obtain $G \oplus H$ from $G \cup H$ by adding all edges between G and H. Moreover, we treat v as the one-vertex graph $(\{v\}, \emptyset)$.

Let Q be a noisy channel, G be its characteristic graph and $L = (v_1, v_2, \ldots, v_l)$ be a sequence of vertices of G such that $v_i \neq v_j$, for $i, j = 1, \ldots, l$ and $i \neq j$. Then we denote a sequence of reductions as

$$G^k[L] = G - v_1 \oplus v_1 - v_2 \oplus v_2 - \ldots - v_k \oplus v_k, \tag{14}$$

where $k \leq l \leq |V(G)|$. Thus for a communication at which we use a reducing channel Q many times, we get the following mass communication characteristic graph

$$G^{\underline{k}}[L] = G^0[L] \boxtimes G^1[L] \boxtimes G^2[L] \boxtimes \ldots G^{k-1}[L], \tag{15}$$

where $G^0[L] = G$. If $V(G) = \{1, \ldots, n\}$ and L is a permutation of it, then we write

$$G^{\underline{n}}[\pi], \tag{16}$$

where π is a permutation. A process desribed by (16) will be called the full reduction of a channel Q with order π.

Now, using Definition 6, we introduce the Shannon capacity for reducing channels

Definition 8. *Given a graph G with n vertices and a permutation π of $\{1, \ldots, n\}$. We define the Shannon capacity for reducing channels as*

$$C_0^R = \max_{k \in [n]} \sqrt[k]{\alpha(G^{\underline{k}}[\pi])}. \tag{17}$$

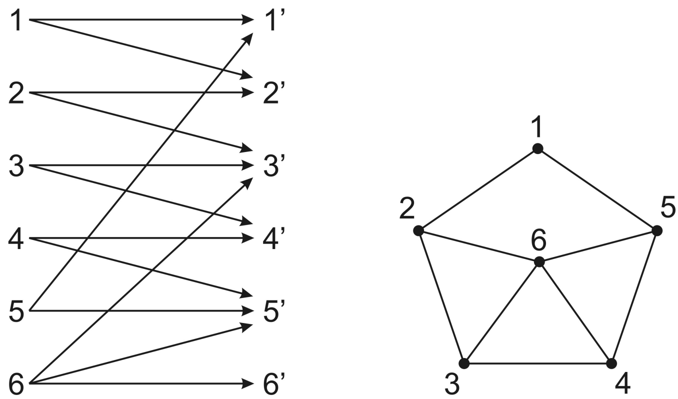

Figure 5. The channel and its characteristic graph $\widetilde{G} = W_6 - e$, where $e = \{1, 6\}$, which brings us a benefit if we use it many times in the reducing communication.

It is easy to see that for every graph G we have

$$\alpha(G) \leq C_0^R \leq C_0. \tag{18}$$

We calculate some values of the Shannon capacity, see Figure **??** for an example. Does there exist a reducing channel Q, which gives us some benefit when we use it many times? It is very suprising that, like for invariant channels, we can get a better result by applying the mass communication for a reducing channel, i.e. for a channel Q and a list $L = \{v_1, \ldots, v_l\}$ of vertices of the characteristic graph G of Q, since the amount of information that can be transmitted by a reducing channel fulfills the following formula

$$\gamma(n) = \log_2(\alpha(G^{\underline{n}}[L])) \geq \log_2(\alpha(G) \cdot \alpha(G - v_1 \oplus v_1) \cdot \ldots \cdot \alpha(G - v_l \oplus v_l)). \tag{19}$$

It turns out that the channel from Figure 5 has this property, because for its characteristic graph $\widetilde{G}$, $\alpha(\widetilde{G}^{\underline{2}}[(6)]) \geq 5$ and $\alpha(\widetilde{G}) \cdot \alpha(\widetilde{G} - v \oplus v) = 4$, see the graph in Figure 6.

3. Conclusions

Determining the Shannon capacity is very difficult. We have described a generalization of the problem. Perhaps, in future it will be possible to find approximations of the Shannon capacity, applying this generalization, for example a lower

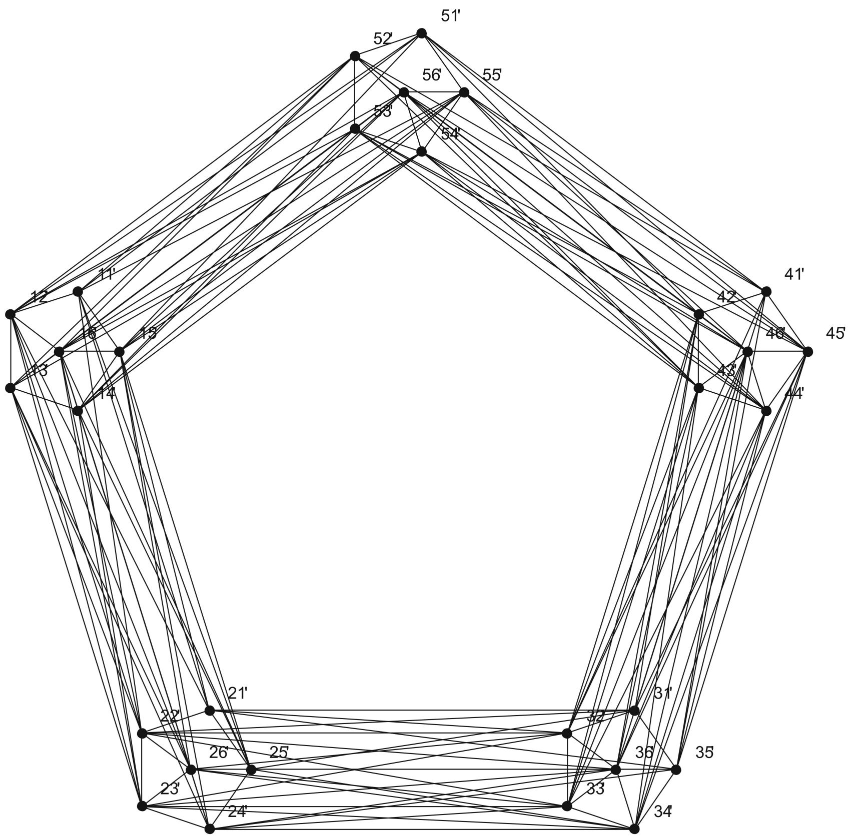

Figure 6. The strong product $\widetilde{G} \boxtimes C_5$, where $\widetilde{G} = W_6 - e$ is the graph from Figure 5. The set $\{11', 23', 35', 42', 54'\}$ is an independent set of the graph. Hence it is also an independent set of $\widetilde{G}^{2}[(6)]$, because $\widetilde{G} \boxtimes C_5$ is an induced subgraph of it.

bound using the Shannon capacity for reducing channels. In the paper, we also have showed a very suprising fact that there exists a reducing channel, which brings us a benefit if we use it many times.

References

[1] Radziszowki, S., *Small Ramsey Numbers*, Electronic Journal of Combinatorics DS1, Aug. 2011.

[2] Cover, T. and Thomas, J., *Elements of Information Theory*, John Wiley and Sons, 2006.

[3] Shannon, C., *The zero-error capacity of a noisy channel*, IRE Trans. Inform, Theory, Vol. IT-2, 1956, pp. 8–19.

[4] Hendrlín, Z., *An application of the Ramsey theorem to the topological product*, Bull. Acad. Pol. Sci., Vol. 14, 1966, pp. 25–26.

[5] Fekete, M., *Uber die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit. ganzzahligen Koeffizienten*, Mathematische Zeitschrift, Vol. 17, 1923, pp. 228–249.

[6] Lovász, L., *On the Shannon Capacity of a Graph*, IEEE Trans. Inform. Theory, Vol. IT-25, 1979, pp. 1–7.