Teoria informacji kolokwium 4 grudnia 2024

Piotr Styczyński[1]

[1]MIM UW (ps386038)

2024-12-08 (updated: 2024-12-08)

1 Problem 2

1.1 Task 1

We observe that for binary input and perfect channel combined in series with Z-channel we have:

$$\begin{cases} P_A(\alpha \to \alpha) = 1 \\ P_A(\beta \to \beta) = 1 \\ P_B(\alpha \to \alpha) = 1 - p \\ P_B(\alpha \to \beta) = p \\ P_B(\beta \to \beta) = 1 \end{cases} \tag{1}$$

$Pr(B \mid A, C)$ won't be well defined for a channel matrix, because:

$$\begin{cases} Pr(B = \alpha \mid A = \alpha, C = \beta) = 1 \\ Pr(B = \alpha \mid A = \alpha, C = \alpha) = 1 \end{cases} \tag{2}$$

Hence the requirement that there are no zero elements in matrices of both channels will come into play. We consider the probability to be well defined if:

1. $\exists_{f \perp P(A)} \forall_{\substack{a \in A \\ c \in C}} Pr(B = b \mid A = a, C = c) = f(...)$
2. $\forall_{\substack{a \in A \\ c \in C}} 0 < Pr(B = b \mid A = a, C = c) < 1$
3. $\forall_{\substack{a \in A \\ c \in C}} \sum_{b \in B} Pr(B = b \mid A = a, C = c) = 1$

Now let's start by expanding the probability using Bayes' theorem:

$$Pr(b \mid a, c) = \frac{Pr(a, c \mid b) * Pr(b)}{Pr(a, c)} \tag{3}$$

For numerator we use the fact that $Pr(x, y \mid z) = Pr(x \mid y, z) * Pr(y \mid z)$ and for enumerator we use $Pr(x, y) = Pr(x \mid y) * Pr(y)$:

$$Pr(b \mid a, c) = \frac{Pr(a \mid b, c) * Pr(b) * Pr(c \mid b)}{Pr(a \mid c) * Pr(c)} \tag{4}$$

Now we use conditional independence contraint so $Pr(a \mid b, c) = Pr(a \mid b)$:

$$Pr(b \mid a, c) = \frac{Pr(a \mid b) * Pr(b) * Pr(c \mid b)}{Pr(a \mid c) * Pr(c)} \tag{5}$$

We use the fact that $Pr(x \mid y) = Pr(y \mid x) * \frac{Pr(x)}{Pr(y)}$:

$$Pr(b \mid a, c) = \frac{Pr(b \mid a) * Pr(a) * Pr(b) * Pr(c \mid b)}{Pr(a \mid c) * Pr(c) * Pr(b)} \tag{6}$$

After reorganising terms:

$$Pr(b \mid a, c) = (Pr(b \mid a) * Pr(c \mid b)) * \frac{Pr(a)}{Pr(a \mid c)} * \frac{1}{Pr(c)} \tag{7}$$

Now let's observe that $\frac{Pr(x)}{Pr(x \mid y)} = \frac{Pr(y)}{Pr(y \mid x)}$:

$$Pr(b \mid a, c) = (Pr(b \mid a) * Pr(c \mid b)) * \frac{Pr(c)}{Pr(c \mid a)} * \frac{1}{Pr(c)} \tag{8}$$

This can be written as:

$$Pr(b \mid a, c) = P_{\Gamma_1}(a \to b) * P_{\Gamma_2}(b \to c) * \frac{1}{Pr(c \mid a)} \tag{9}$$

We observe that $Pr(c \mid a)$ in fact does not depend on the distribution of A, which proves first point. In all our transformations we inmplicitly use the facts that $Pr(a), Pr(b), Pr(c) \neq 0$ Second point is trivial to show, because $0 < P_{\Gamma_1}(a \to b), P_{\Gamma_2}(b \to c), Pr(c \mid a) < 1$ Third point is direct consequence of conditional independence.

1.2 Task 2

Let's assume we take channel Γ_E and squeeze the symbol space using mapping $s_1 \in (A \times C)^A := s_1(a, c) = a$

In such scenario:

$$P_{\Gamma_E}(s_1(a, c) \to b) = \sum_c P_{\Gamma_E}(a, c \to b)$$

$$\sum_c P_{\Gamma_1}(a \to b) * P_{\Gamma_2}(b \to c) * \frac{1}{Pr(c \mid a)}$$

$$P_{\Gamma_F}(a \to b) = P_{\Gamma_1}(a \to b) * \sum_c \frac{P_{\Gamma_2}(b \to c)}{Pr(c \mid a)}$$

We know that $\sum_c Pr(c \mid a) = 1$

If we compare $P_{\Gamma_F}(a \to b)$ and $P_{\Gamma_1}(a \to b)$ then in $I(A, B)$ we will notice that in worst case we can select such a coefficients that Γ_F is as bad as using Γ_1 The original matrix Γ_F has more rows than Γ_E (such that rows of Γ_F are linear combinations of Γ_E) meaning that the capacity of original $Gamma_E$ cannot be lower.

Hance we prove that capacity of Γ_E is no lower than Γ_F

1.3 Task 3

My best guess is that we use some prime factors to make Γ_E have all rows independent, then that will work. However I was unable to showcase some concrete example.