Exercise 4 in Cryptography
- Proposed Solution -

Prof. Dr. Rudolf Mathar, Henning Maier, Jose Angel Leon Calvo
2015-05-07

Solution of Problem 10

The message space of a finite sequence of length $k = 11$ is:

$$\mathcal{M} = \{(m_1, ..., m_{11}) \mid m_i \in \mathcal{X}\}$$

with the alphabet $\mathcal{X} = \{a, b, ..., z\} = \{0, 1, ..., 25\}$, and $|\mathcal{X}| = 26$.

In the given task, there are 4 blocks with cyclic permutations. These blocks are not changed if the letters are the same inside each individual block. Unchanged sequences are subsumed by:

$$\begin{aligned}\hat{\mathcal{M}} = &\{(m_1, ..., m_{11}) | m_1 \in \mathcal{X}, m_2 = m_{11} = m_5 = m_8 \in \mathcal{X}, m_3 = m_6 = m_7 = m_4 \in \mathcal{X},\\ &m_9 = m_{10} \in \mathcal{X}\}\end{aligned}$$

The total number of such sequences is $|\hat{\mathcal{M}}| = |\mathcal{X}|^4 = 456976$.

Remark: However, compared to $|\mathcal{M}| = |\mathcal{X}|^{11} \approx 3.6 \cdot 10^{15}$, this is only a minor restriction. (An unchanged plaintext in English is 'MISSISSIPPI'.)

Solution of Problem 11

Theorem 4.3 shall be proven.

a) X is a discrete random variable with $p_i = P(X = x_i)$, $i = 1, \ldots, m$. It holds

$$H(X) = -\sum_i p_i \log(p_i) \geq 0,$$

as $p_i \geq 0$ and $-\log(p_i) \geq 0$ for $0 < p_i \leq 1$ and $0 \cdot \log 0 = 0$ per definition.

Equality holds, if all addends are zero, i.e.,

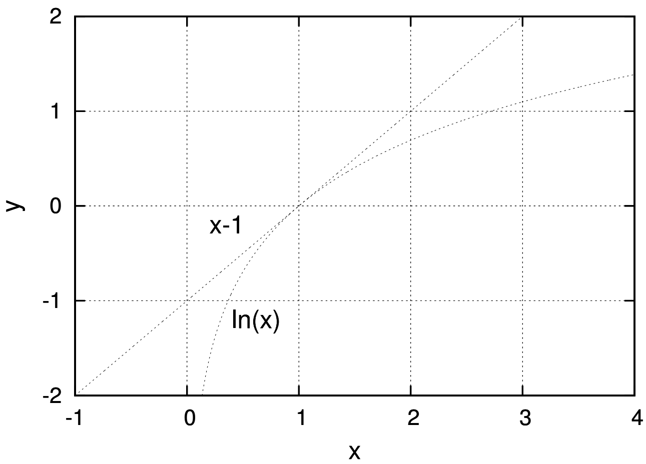
$$p_i \log(p_i) = 0 \Leftrightarrow p_i \in \{0, 1\} \quad i = 1, \ldots, m,$$

as $p_i > 0$ and $-\log(p_i) > 0$, thus, $-p_i \log(p_i) > 0$ for $0 < p_i < 1$.

b) It holds

$$
\begin{aligned}
H(X) - \log(m) &= -\sum_i p_i \log(p_i) - \underbrace{\sum_i p_i}_{=1} \log(m) \\
&= \sum_{i:p_i>0} p_i \log\left(\frac{1}{p_i\, m}\right) \\
&= (\log e) \sum_{i:p_i>0} p_i \ln\left(\frac{1}{p_i\, m}\right) \\
&\overset{\ln(x)\le x-1}{\le} (\log e) \sum_{i:p_i>0} p_i \left(\frac{1}{p_i\, m} - 1\right) \\
&= (\log e) \sum_{i:p_i>0} \left(\frac{1}{m} - p_i\right) = 0
\end{aligned}
$$

As $\ln(x) = x - 1$ only holds for $x = 1$ it follows that equality holds iff $p_i = 1/m$, $i = 1, \dots, m$. In particular, as $p_i = \frac{1}{m}$, it follows $p_i > 0$, $i = 1, \dots, m$.



c) Define for $i = 1, \dots, m$ and $j = 1, \dots, d$

$$p_{i|j} = P(X = x_i \mid Y = y_j).$$

Show $H(X \mid Y) - H(X) \le 0$ which is equivalent to the claim.

$$
\begin{aligned}
H(X \mid Y) - H(X) &= -\sum_{i,j} p_{i,j} \log(p_{i|j}) + \sum_i p_i \log(p_i) \\
&= -\sum_{i,j} p_{i,j} \log\left(\frac{p_{i,j}}{p_j}\right) + \sum_i \underbrace{\sum_j p_{i,j}}_{=p_i} \log(p_i) \\
&= (\log e) \sum_{i,j:p_{i,j}>0} p_{i,j} \ln\left(\frac{p_i\, p_j}{p_{i,j}}\right) \\
&\overset{\ln(x)\le x-1}{\le} (\log e) \sum_{i,j:p_{i,j}>0} p_{i,j} \left(\frac{p_i\, p_j}{p_{i,j}} - 1\right) \\
&= (\log e) \sum_{i,j:p_{i,j}>0} (p_i\, p_j - p_{i,j}) = 0
\end{aligned}
$$

Note that from $p_{i,j} > 0$ it follows $p_i, p_j > 0$. Equality hold for $p_i p_j = p_{i,j}$ which is equivalent to X and Y being stochastically independent.

This means that the mutual information $I(X,Y) = H(X) - H(X \mid Y)$ is nonnegative.

d) It holds

$$\begin{aligned} H(X,Y) &= -\sum_{i,j} p_{i,j} \log(p_{i,j}) \\ &= -\sum_{i,j} p_{i,j} \left[\log(p_{i,j}) - \log(p_i) + \log(p_i)\right] \\ &= -\sum_{i,j} p_{i,j} \underbrace{\log\left(\frac{p_{i,j}}{p_i}\right)}_{p_{j|i}} - \sum_i \underbrace{\sum_j p_{i,j}}_{=p_i} \log(p_i) \\ &= H(Y \mid X) + H(X). \end{aligned}$$

e) It holds

$$H(X,Y) \overset{(d)}{=} H(X) + H(Y \mid X) \overset{(c)}{\leq} H(X) + H(Y)$$

with equality as in (c) iff X and Y are stochastically independent.

Solution of Problem 12

Show for any function $f : X(\Omega) \times Y(\Omega) \to \mathbb{R}$, that $H(X, Y, f(X, Y)) = H(X, Y)$.

By definition, we have:

$$H(X,Y,Z = f(X,Y)) \overset{\text{Def.}}{=} \sum_{X,Y,Z} P(X = x, Y = y, Z = z) \log\left(P(X = x, Y = y, Z = z)\right)$$

With

$$P(X = x, Y = y, Z = z) = \begin{cases} P(X = x, Y = y) & \text{, if } Z = f(X,Y) \\ 0 & \text{, if } Z \neq f(X,Y) \end{cases},$$

it follows that

$$H(X,Y,Z = f(X,Y)) = \sum_{X,Y} P(X = x, Y = y) \log(P(X = x, Y = y)) = H(X,Y)\,.$$

Note: It holds $0 \cdot \log 0 = 0$.