Solution of Problem 1

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.,

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

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

b) It holds

$$H(X) - \log(m) = - \sum_i p_i \log(p_i) - \sum_i p_i \log(m)$$

$$= \sum_{i:p_i>0} p_i \log \left( \frac{1}{p_i m} \right)$$

$$= (\log e) \sum_{i:p_i>0} p_i \ln \left( \frac{1}{p_i m} \right)$$

$$\leq (\log e) \sum_{i:p_i>0} p_i \left( \frac{1}{p_i m} - 1 \right)$$

$$= (\log e) \sum_{i:p_i>0} \left( \frac{1}{m} - p_i \right) = 0$$

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

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

$$p_{ij} = P(X = x_i \mid Y = y_j).$$
Show $H(X | Y) - H(X) \leq 0$ which is equivalent to the claim.

$$H(X | Y) - H(X) = -\sum_{i,j} p_{i,j} \log(p_{i,j}) + \sum_i p_i \log(p_i)$$

$$= -\sum_{i,j} p_{i,j} \log \left( \frac{p_{i,j}}{p_j} \right) + \sum_i \sum_j p_{i,j} \log(p_i)$$

$$= (\log e) \sum_{i,j:p_{i,j} > 0} p_{i,j} \log \left( \frac{p_i p_j}{p_{i,j}} \right)$$

$$\leq (\log e) \sum_{i,j:p_{i,j} > 0} p_{i,j} \left( \frac{p_i p_j}{p_{i,j}} - 1 \right)$$

$$= (\log e) \sum_{i,j:p_{i,j} > 0} (p_i p_j - p_{i,j}) = 0$$

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 $I(X, Y) = H(X) - H(X | Y)$ is nonnegative.

d) It holds

$$H(X, Y) = -\sum_{i,j} p_{i,j} \log(p_{i,j})$$

$$= -\sum_{i,j} p_{i,j} [\log(p_{i,j}) - \log(p_i) + \log(p_i)]$$

$$= -\sum_{i,j} p_{i,j} \log \left( \frac{p_{i,j}}{p_i} \right) - \sum_i \sum_j p_{i,j} \log(p_i)$$

$$= H(Y | X) + H(X).$$

e) It holds

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

with equality as in (c) iff $X$ and $Y$ are stochastically independent.
Solution of Problem 2

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

By definition, we have:

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

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 \( 0 \cdot \log 0 = 0 \).

Solution of Problem 3

Prove Theorem 4.13 ‘⇒’ (sufficient solution):

Recall that each element of these sets has a positive probability:

\[
\mathcal{M}_+ := \{ M \in \mathcal{M} \mid P(\hat{M} = M) > 0 \},
\]

\[
\mathcal{C}_+ := \{ C \in \mathcal{C} \mid P(\hat{C} = C) > 0 \}.
\]

Lemma 4.12 provides conditions of perfect secrecy on \( \mathcal{M}_+, \mathcal{K}_+, \mathcal{C}_+ \).

With Lemma 4.12 a), we obtain:

\[
|\mathcal{M}_+| \leq |\mathcal{C}_+| \overset{(I)}{\leq} |\mathcal{C}| \overset{(II)}{=} |\mathcal{M}| \overset{(III)}{=} |\mathcal{M}_+|.
\]

(I): With \( P(\hat{C} = C) > 0 \Rightarrow \mathcal{C}_+ \subseteq \mathcal{C} \).

(II): Given by assumption \( |\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}| \).

(III): Given by assumption \( P(\hat{M} = M) > 0, \forall M \in \mathcal{M} \).

By the ‘sandwich theorem’, i.e., the upper and lower bounds are both equal to \( |\mathcal{M}_+| \):

\[
\Rightarrow |\mathcal{C}_+| = |\mathcal{C}| \Rightarrow \mathcal{C}_+ = \mathcal{C},
\]

\[
\Rightarrow P(\hat{C} = C) > 0, \forall C \in \mathcal{C}.
\]

Let \( M \in \mathcal{M}, C \in \mathcal{C} \):

\[
0 < P(\hat{C} = C) \overset{(IV)}{=} P(\hat{C} = C \mid \hat{M} = M) = P(e(\hat{M}, \hat{K}) = C \mid \hat{M} = M)
\]

\[
\overset{(V)}{=} P(e(M, \hat{K}) = C) = \sum_{K \in \mathcal{K} \colon e(M, K) = C} P(\hat{K} = K) \neq 0
\]

\[
\Rightarrow \forall M \in \mathcal{M}, \ C \in \mathcal{C} \exists K \in \mathcal{K} : e(M, K) = C.
\]

(IV): With perfect secrecy as given by Corollary 4.11.

(V): Given by the assumption that \( \hat{M}, \hat{K} \) are stochastically independent.
However, (1) is not shown to be unique yet!

(i) Fix $M \in \mathcal{M}$:

$$|\mathcal{C}_+| = |\mathcal{C}| = |\{e(M, K) \mid K \in \mathcal{K}_+ = \mathcal{K}\}| \leq |\mathcal{K}| \overset{(II)}{=} |\mathcal{C}|$$

$\Rightarrow K$ is unique with $K = K(M, C)$ by the ‘sandwich theorem’.

(II) Given by assumption $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$.

Let $M \in \mathcal{M}, C \in \mathcal{C}$:

$$\Rightarrow P(\hat{C} = C) \overset{(1)}{=} P(\hat{K} = K(M, C)),$$

because of perfect secrecy this expression is independent of $M$.

(ii) Fix $C_0 \in \mathcal{C}$:

$$\Rightarrow \{K(M, C_0) \mid M \in \mathcal{M}\} = \mathcal{K},$$

because of injectivity of $e(\cdot, K)$, i.e., $e(M, K) = C_0$, and by the assumption $|\mathcal{M}| = |\mathcal{C}|$.

$$\Rightarrow P(\hat{C} = C) = P(\hat{K} = K) \forall C \in \mathcal{C}, K \in \mathcal{K}$$

$$\Rightarrow P(\hat{K} = K) = \frac{1}{|\mathcal{K}|} \forall K \in \mathcal{K}. \quad \square$$