Problem 1.

(a) We have $H(f(U)) \leq H(f(U), U) = H(U) + H(f(U)|U) = H(U) + 0 = H(U)$.

(b) Notice that $U \rightarrow V \rightarrow f(V)$ is a Markov chain. The data processing inequality implies that $H(U) - H(U|f(V)) = I(U; f(V)) \leq I(U; V) = H(U) - H(U|V)$. Therefore, $H(U|V) \leq H(U|f(V))$.

Problem 2.

(a) We have:

$$H(U|U) \leq H(U,W|U) = H(W|\hat{U}) + H(U|\hat{U}, W) \leq H(W) + H(U|\hat{U}, W)$$

$$= H(W) + H(U|\hat{U}, W = 0) \cdot \mathbb{P}[W = 0] + H(U|\hat{U}, W = 1) \cdot \mathbb{P}[W = 1]$$

$$(*) \leq h_2(p_e) + 0 \cdot (1 - p_e) + \log(|U| - 1) \cdot p_e = h_2(p_e) + p_e \log(|U| - 1),$$

where $(*)$ follows from the following facts:

- $H(W) = h_2(p_e)$.
- $H(U|\hat{U}, W = 0) = 0$: conditioned on $W = 0$, we know that $U = \hat{U}$ and so the conditional entropy $H(U|\hat{U}, W = 0)$ is equal to 0.
- $H(U|\hat{U}, W = 1) \leq \log(|U| - 1)$: conditioned on $W = 1$, we know that $U \neq \hat{U}$ and so there are at most $|U| - 1$ values for $U$. Therefore, the conditional entropy $H(U|\hat{U}, W = 1)$ is at most $\log(|U| - 1)$.

(b) Let $\hat{U} = f(V)$. We have $H(U|\hat{U}) \leq h_2(p_e) + p_e \log(|U| - 1)$ from (a). On the other hand, from Problem 1(b) we have $H(U|V) \leq H(U|f(V)) = H(U|\hat{U})$. We conclude that $H(U|V) \leq h_2(p_e) + p_e \log(|U| - 1)$.

Problem 3.

(a) $W$ is independent of $(U, Z)$. Therefore, $W$ is independent of $(U, U \oplus Z) = (U, V)$, which implies that $\mathbb{P}_{W|U,V}(w|u,v) = \mathbb{P}_W(w) = \mathbb{P}_{W|V}(w|v)$ for every $u, v, w \in \{0, 1\}$. Thus, $U \rightarrow V \rightarrow W$ is a Markov chain and so we have $I(U; V) \geq I(U; W)$ from the data processing inequality.

In order to show that $U \rightarrow V' \rightarrow W'$ is a Markov chain, we will show first that $W'$ is independent of $(U, Z')$. For every $u, z', w' \in \{0, 1\}$ we have:

$$\mathbb{P}_{U,Z',W'}(u, z', w') = \mathbb{P}[U = u, Z' = z', U \oplus W = w'] = \mathbb{P}[U = u, Z' = z', W = u \oplus w']$$

$$\overset{(*)}{=} \mathbb{P}_{U,Z'}(u, z') \cdot \frac{1}{2} \overset{(**)}{=} \mathbb{P}_{U,Z'}(u, z') \cdot \mathbb{P}_{W'}(w'),$$

where $(*)$ follows from the fact that $W$ is uniform and independent of $(U, Z')$. $(**)$ follows from the fact that $W' = U \oplus W$ is uniform (it is easy to check by computing
the joint probability distribution that the XOR of two independent uniform binary random variables is uniform).

Since we have shown that $W'$ is independent of $(U, Z')$, the proof that $U \leftrightarrow V' \leftrightarrow W'$ is a Markov chain is similar to that of $U \leftrightarrow V \leftrightarrow W$, and the inequality $I(U; V') \geq I(U; W')$ follows from the data processing inequality.

(b) By computing the probability distribution of $V$, we can see that it is uniform. Similarly, $V'$ is also uniform. We have:

$$- I(U; V) = H(V) - H(V|U) = H(V) - H(U \oplus Z|U) = H(V) - H(Z|U) = H(V) - H(Z) = 1 - h_2(p).$$

$$- I(U; W) = 0 \text{ since } U \text{ and } W \text{ are independent.}$$

$$- I(U; V') = H(V') - H(V'|U) = H(V') - H(U \oplus Z'|U) = H(V') - H(Z'|U) = H(V') - H(Z') = 1 - h_2(p), \text{ where } h_2(p) = p \log \frac{1}{p} + (1 - p) \log \frac{1}{1-p}.$$ 

$$- I(U; W') = 0 \text{ since } U \text{ and } W' \text{ are independent.}$$

Since $0 < p < \frac{1}{2}$, $h_2(p) < 1$ and $1 - h_2(p) > 0$. Therefore, $I(U; V) > I(U; W)$ and $I(U; V') > I(U; W')$.

(c) By computing the joint probability distribution of $(V, Z, Z')$, we can see that $V$ is independent of $(Z, Z')$, which implies that $V$ is independent of $Z \oplus Z'$. We have:

$$I(U; VV') = H(V, V') - H(V, V'|U) = H(V, V' \oplus V) - H(U \oplus Z, U \oplus Z'|U)$$

$$= H(V, Z \oplus Z') - H(Z, Z') \overset{(\ast)}{=} H(V) + H(Z \oplus Z') - H(Z) - H(Z')$$

$$\overset{(\ast \ast)}{=} 1 + h_2(2p(1-p)) - 2h_2(p).$$

$(\ast)$ follows from the fact that $V$ is independent of $Z \oplus Z'$ and that $Z$ is independent of $Z'$. $(\ast \ast)$ follows from the fact that $H(Z \oplus Z') = h_2(2p(1-p))$ (since $\mathbb{P}[Z \oplus Z' = 1] = 2p(1-p)$) and $H(Z) = H(Z') = h_2(p)$.

On the other hand, we have:

$$I(U; WW') = I(U; W, W \oplus W') = I(U; W, U)$$

$$= I(U; U) + I(U; W|U) = H(U) + 0 = 1.$$ 

In order to see that $I(U; VV') < I(U; WW')$, notice that $H(Z) + H(Z') = H(Z, Z') = H(Z, Z \oplus Z') = H(Z \oplus Z') + H(Z | Z \oplus Z')$. Therefore, $H(Z \oplus Z') \leq H(Z) + H(Z')$ with equality if and only if $H(Z|Z \oplus Z') = 0$. Now notice that for every $a, b \in \{0, 1\}$, $\mathbb{P}[Z = a, Z \oplus Z' = b] = \mathbb{P}[Z = a, Z' = a \oplus b] = \mathbb{P}[Z = a] \mathbb{P}[Z' = a \oplus b] > 0$. This implies that for every $a, b \in \{0, 1\}$, $\mathbb{P}[Z = a | Z \oplus Z' = b] > 0$. Therefore, conditioned on $Z \oplus Z'$, $Z$ is not deterministic and so $H(Z|Z \oplus Z') > 0$. We conclude that $H(Z \oplus Z') < H(Z) + H(Z')$ which implies that $1 + H(Z \oplus Z') - H(Z) - H(Z') < 1$ and $I(U; VV') < I(U; WW')$.

**Problem 4.**

(a) By using the inequality $\ln x \leq x - 1$ for $x > 0$, we get:

$$p \log \frac{p + q}{2p} + q \log \frac{p + q}{2q} \leq \frac{p}{\ln 2} \left( \frac{p + q}{2p} - 1 \right) + \frac{q}{\ln 2} \left( \frac{p + q}{2q} - 1 \right) = 0.$$
Therefore, \( p \log \frac{1}{p} + p \log \frac{p+q}{2} + q \log \frac{1}{q} + q \log \frac{p+q}{2} \leq 0 \), from which we conclude that \( \frac{1}{2} \left( p \log \frac{1}{p} + q \log \frac{1}{q} \right) \leq \frac{p+q}{2} \log \frac{2}{p+q} \).

(b) We have:

\[
H(r) = \sum_{u \in \mathcal{U}} r(u) \log \frac{1}{r(u)} = \sum_{u \in \mathcal{U}} \frac{p(u) + q(u)}{2} \log \frac{2}{p(u) + q(u)} \geq \frac{1}{2} \left( \sum_{u \in \mathcal{U}} p(u) \log \frac{1}{p(u)} + q(u) \log \frac{1}{q(u)} \right)
= \frac{1}{2} \left( \sum_{u \in \mathcal{U}} p(u) \log \frac{1}{p(u)} \right) + \frac{1}{2} \left( \sum_{u \in \mathcal{U}} q(u) \log \frac{1}{q(u)} \right)
= \frac{1}{2} H(p) + \frac{1}{2} H(q),
\]

where \((*)\) follows from (a).

**Problem 5.**

(a) We have:

\[
S = \sum_{u \in \mathcal{U}} \max \{ P_1(u), P_2(u) \} \leq \sum_{u \in \mathcal{U}} (P_1(u) + P_2(u)) = \sum_{u \in \mathcal{U}} P_1(u) + \sum_{u \in \mathcal{U}} P_2(u) = 1 + 1 = 2.
\]

It is easy to see from \((*)\) that \( S = 2 \) if and only if \( \max \{ P_1(u), P_2(u) \} = P_1(u) + P_2(u) \) for all \( u \in \mathcal{U} \), which is equivalent to say that there is no \( u \in \mathcal{U} \) for which we have \( P_1(u) > 0 \) and \( P_2(u) > 0 \). In other words, \( S = 2 \) if and only if

\[
\{ u \in \mathcal{U} : P_1(u) > 0 \} \cap \{ u \in \mathcal{U} : P_2(u) > 0 \} = \emptyset.
\]

(b) Let \( l_i = \lceil \log_2 \frac{S}{\max \{ P_1(a_i), P_2(a_i) \}} \rceil \), and let us compute the Kraft sum:

\[
\sum_{i=1}^{M} 2^{-l_i} \leq \sum_{i=1}^{M} 2^{-\log_2 \frac{S}{\max \{ P_1(a_i), P_2(a_i) \}}} = \sum_{i=1}^{M} \frac{\max \{ P_1(a_i), P_2(a_i) \}}{S} = 1.
\]

Since the Kraft sum is at most 1, there exists a prefix-free code where the length of the codeword associated to \( a_i \) is \( l_i \).

(c) Since the code constructed in (b) is prefix free, it must be the case that \( \bar{l} \geq H(U) \). In order to prove the upper bounds, let \( P^* \) be the true distribution (which is either \( P_1 \) or \( P_2 \)). It is easy to see that \( P^*(a_i) \leq \max \{ P_1(a_i), P_2(a_i) \} \) for all \( 1 \leq i \leq M \). We
have:

$$I = \sum_{i=1}^{M} P^*(a_i) l_i = \sum_{i=1}^{M} P^*(a_i) \left[ \log_2 \frac{S}{\max\{P_1(a_i), P_2(a_i)\}} \right]$$

$$< \sum_{i=1}^{M} P^*(a_i) \left( 1 + \log_2 \frac{S}{\max\{P_1(a_i), P_2(a_i)\}} \right)$$

$$= \sum_{i=1}^{M} P^*(a_i) \left( 1 + \log S + \log_2 \frac{1}{\max\{P_1(a_i), P_2(a_i)\}} \right)$$

$$= 1 + \log S + \sum_{i=1}^{M} P^*(a_i) \log_2 \frac{1}{\max\{P_1(a_i), P_2(a_i)\}}$$

$$\leq 1 + \log S + \sum_{i=1}^{M} P^*(a_i) \log_2 \frac{1}{P^*(a_i)} = H(U) + \log S + 1 \leq H(U) + 2,$$

where the inequality (*) uses the fact that $P^*(a_i) \leq \max\{P_1(a_i), P_2(a_i)\}$ for all $1 \leq i \leq M$.

(d) Now let $l_i = \lfloor \log_2 \frac{S}{\max\{P_1(a_i), \ldots, P_k(a_i)\}} \rfloor$, and let us compute the Kraft sum:

$$\sum_{i=1}^{M} 2^{-l_i} \leq \sum_{i=1}^{M} 2^{-\log_2 \frac{S}{\max\{P_1(a_i), \ldots, P_k(a_i)\}}} = \sum_{i=1}^{M} \max\{P_1(a_i), \ldots, P_k(a_i)\} \frac{S}{S} = 1.$$

Since the Kraft sum is at most 1, there exists a prefix-free code where the length of the codeword associated to $a_i$ is $l_i$. Since the code is prefix free, it must be the case that $I \geq H(U)$. In order to prove the upper bounds, let $P^*$ be the true distribution (which is either $P_1$ or . . . or $P_k$). It is easy to see that $P^*(a_i) \leq \max\{P_1(a_i), \ldots, P_k(a_i)\}$ for all $1 \leq i \leq M$. We have:

$$I = \sum_{i=1}^{M} P^*(a_i) l_i = \sum_{i=1}^{M} P^*(a_i) \left[ \log_2 \frac{S}{\max\{P_1(a_i), \ldots, P_k(a_i)\}} \right]$$

$$< \sum_{i=1}^{M} P^*(a_i) \left( 1 + \log_2 \frac{S}{\max\{P_1(a_i), \ldots, P_k(a_i)\}} \right)$$

$$= \sum_{i=1}^{M} P^*(a_i) \left( 1 + \log_2 S + \log_2 \frac{1}{\max\{P_1(a_i), \ldots, P_k(a_i)\}} \right)$$

$$= 1 + \log_2 S + \sum_{i=1}^{M} P^*(a_i) \log_2 \frac{1}{\max\{P_1(a_i), \ldots, P_k(a_i)\}}$$

$$\leq 1 + \log_2 S + \sum_{i=1}^{M} P^*(a_i) \log_2 \frac{1}{P^*(a_i)} = H(U) + \log_2 S + 1,$$

where the inequality (*) uses the fact that $P^*(a_i) \leq \max\{P_1(a_i), \ldots, P_k(a_i)\}$ for all $1 \leq i \leq M$. Now notice that $\max\{P_1(a_i), \ldots, P_k(a_i)\} \leq \sum_{j=1}^{k} P_j(a_i)$ for all $1 \leq i \leq M$. Therefore, we have

$$S = \sum_{i=1}^{M} \max\{P_1(a_i), \ldots, P_k(a_i)\} \leq \sum_{i=1}^{M} \sum_{j=1}^{k} P_j(a_i) = \sum_{j=1}^{k} \sum_{i=1}^{M} P_j(a_i) = \sum_{j=1}^{k} 1 = k.$$