Problem 1. (20 points)

In a cryptosystem, a secret key $K$ known to both Alice and Bob allows for secure communication. Using the key $K$, Alice converts her plain text $U$ to a ciphertext $V$. Using the same key $K$, Bob converts the ciphertext $V$ back into $U$. We model $U$, $V$ and $K$ as random variables. Secure communication requires $U$ and $V$ to be independent.

(a) (2 pts) What are the values of $H(U|VK)$ and $I(U;V)$?

From the problem statement we know that Bob can determine $U$ given $V$ and $K$. This implies that $H(U|VK) = 0$.

From the problem statement we know that the secret key $K$ allows secure communication, whereas secure communication is defined as $U$ and $V$ to be independent. This implies that $I(U;V) = 0$.

(b) (4 pts) Determine the relation, (i.e., $<$, $\leq$, $=$, $>$, or $\geq$), between $H(U)$ and $I(U;K|V)$. Provide a proof for this relation.

Consider the following expansion of $I(U;K|V)$

$$I(U;K|V) = H(U|V) - H(U|KV) = H(U)$$

where we used the fact that $U$ and $V$ are independent ($H(U|V) = H(U)$) and the result of (a) that $H(U|VK) = 0$.

(c) (4 pts) Determine the relation, (i.e., $<$, $\leq$, $=$, $>$, or $\geq$), between $H(K)$ and $I(U;K|V)$. Provide a proof for this relation.

Observe the following inequalities:

$$I(U;K|V) = H(K|V) - H(K|UV) \leq H(K|V) \leq H(K)$$

where the first inequality is due to $H(K|UV) \geq 0$ and the second inequality is due to the fact that conditioning reduces entropy.

(d) (4 pts) Show that $H(K) \geq H(U)$. Furthermore, show that if the equality holds, then (i) $K$ and $V$ are independent and (ii) $H(K|UV) = 0$.

From (b) and (c) we have $H(K) \geq I(U;K|V) = H(U)$. The equality holds if $H(K) = I(U;K|V)$. From the chain of inequalities in part (c), we can see that this implies that $H(K|V) = H(K)$ (such that $K$ is independent of $V$) and $H(K|UV) = 0$.

Suppose further that (i) $K$ is independent of $U$, (ii) the cryptosystem is implemented as $V = f(U, K)$ and $U = g(V, K)$, and (iii) the system is supposed to be secure regardless of the distribution of $U$ on a given alphabet $U$. 
(e) (2 pts) Show that $H(K) \geq \log |U|$.

From (d) we have $H(K) \geq H(U)$, and from the problem statement, this property must hold for any distribution of $U$. Take $U$ to be distributed uniformly on $U$ such that $H(U) = \log |U|$. This gives us $H(K) \geq H(U) = \log |U|$.

(f) (4 pts) With $U = \{0, 1, \ldots, |U| - 1\}$, show that if we take $K$ to be uniform on $U$, the secrecy requirement is satisfied by $f(u, k) = u + k \mod |U|$.

To fulfill the secrecy requirement, we need to show that $U$ and $V$ are independent. One way to do this is by showing that $P(V = v \mid U = u) = P(V = v)$ for all $v$ and $u$. As we have $V = K + U \mod |U|$, then

$$P(V = v \mid U = u) = P(K = u - v \mod |U| \mid U = u)$$

$$= P(K = u - v \mod |U|)$$

$$= \frac{1}{|U|}$$

where the second line is due to $U$ and $K$ are independent. From this equality, we can see that for any $v$, $P(V = v \mid U = u)$ does not depend on $u$. Therefore we can assert that $P(V = v \mid U = u) = P(V = v)$ for all $u$ and $v$. 
Problem 2. (18 points)

Suppose $U_1, U_2, \ldots$ are i.i.d. random variables with finite alphabet and let $p$ denote the distribution of each $U_i$. Suppose we do not know $p$, but we know that it is included in the set of $K$ possible distributions, i.e., $p \in \mathcal{P} = \{p_k : k = 1, \ldots, K\}$.

For any distribution $q$ on $\mathcal{U}$, define $r(q) = \max_k D(p_k||q)$.

(a) (4 pts) Show that for any $q$ there exists a prefix-free code $C : \mathcal{U} \to \{0,1\}^*$ such that

$$E \left[ \text{length}(C(U)) \right] - H(U) \leq r(q) + 1$$

whenever the distribution of random variable $U$ is in $\mathcal{P}$.

For each $u \in \mathcal{U}$, we assign a code of length $l(u) = \lceil -\log_2 q(u) \rceil$. We can see that

$$\sum_{u \in \mathcal{U}} 2^{-\lceil -\log_2 q(u) \rceil} \leq \sum_{u \in \mathcal{U}} 2^{\log_2 q(u)} = \sum_{u \in \mathcal{U}} q(u) = 1.$$ 

and due to Kraft’s inequality, there exists a prefix-free code with such code lengths.

Now, suppose each $U_i$ has distribution $p_k$ for some $k \in [K]$, then we have the following relations between the expected length of the code designed as above and the entropy of $U_i$’s.

$$E \left[ \text{length}(C(U)) \right] - H(U) = \sum_{u \in \mathcal{U}} p_k(u) l(u) - \sum_{u \in \mathcal{U}} -p_k(u) \log_2 p_k(u)$$

$$\leq - \sum_{u \in \mathcal{U}} p_k(u) \log_2 q(u) + 1 + \sum_{u \in \mathcal{U}} p_k(u) \log_2 p_k(u)$$

$$= \sum_{u \in \mathcal{U}} p_k(u) \log_2 \frac{p_k(u)}{q(u)} + 1$$

$$= D(p_k||q) + 1$$

$$\leq \max_k D(p_k||q) + 1$$

$$= r(q) + 1$$

where the second line is due to $\lceil x \rceil \leq x + 1$, and the fourth line is due to the definition of $D(p_k||q)$. Since the last inequality obtained does not depend on $k$, it is valid no matter what distribution $U_i$’s have.

(b) (4 pts) Show that $\min_q r(q) \leq \log K$. [Hint: try $q(u) = \frac{1}{K} \sum_k p_k(u).$]

We use the $q$ given in the hint to show the following inequality

$$\min_q \max_k D(p_k||q') \leq \max_k D(p_k||q)$$

$$= \max_k \sum_{u \in \mathcal{U}} p_k(u) \log_2 \frac{p_k(u)}{\frac{1}{K} \sum_{u' \in \mathcal{U}} p_k(u')}$$

$$= \max_k \sum_{u \in \mathcal{U}} p_k(u) \log_2 \frac{p_k(u)}{\sum_{u' \in \mathcal{U}} p_k(u')} + \sum_{u \in \mathcal{U}} p_k(u) \log_2 K$$

$$\leq \max_k \sum_{u \in \mathcal{U}} p_k(u) \log_2 K$$

$$= \log_2 K$$

where the third line is due to the fact that $p_k(u) \leq \sum_{u' \in \mathcal{U}} p_k(u')$ and $\log_2(x) \leq 0$ for all $0 < x \leq 1$. 


(c) (4 pts) Show that for fixed $K$ there exists a sequence of prefix-free codes $C_n : U^n \to \{0, 1\}^*$ such that

$$\lim_{n \to \infty} \frac{1}{n} E[\text{length}(C_n(U^n))] = H(U)$$

whenever $U_1, U_2, \ldots$ are i.i.d. and have a distribution in $\mathcal{P}$. [Hint: use (b).]

Define $p_{k,n}(U^n) = \prod_{i=1}^n p_k(U_i)$. We use the results of (a) on the random variables $U^n$ such that we have for every $n$ there exists a prefix-free code $C_n$ such that

$$E[\text{length}(C_n(U^n))] - H(U^n) \leq \min_q \max_k D(p_{k,n}||q) + 1.$$

Furthermore, from the result of (b) and the fact that $U_i$’s are i.i.d. we have

$$E[\text{length}(C_n(U^n))] - nH(U) \leq \log_2 K + 1.$$

Dividing both sides by $n$ gives us

$$\lim_{n \to \infty} \frac{1}{n} E[\text{length}(C_n(U^n))] - H(U) \leq \frac{\log_2 K + 1}{n}.$$  \hspace{1cm} (1)

We also know from the lectures that

$$0 \leq \frac{1}{n} E[\text{length}(C_n(U^n))] - H(U).$$  \hspace{1cm} (2)

Combining (1) and (2), and taking $n \to \infty$, we finally obtain

$$\lim_{n \to \infty} \frac{1}{n} E[\text{length}(C_n(U^n))] - H(U) = 0.$$

(d) (2 pts) Let $Z = \sum_u \max_k p_k(u)$. Show that $\min_q r(q) \geq \log Z$. [Hint: try choosing $q(u)$ proportional to $\max_k p_k(u)$]

We use the same argument as in (b) by just replacing $q$ with the new hint $(q(u) = \max_k p_k(u)/Z, \ Z = \max_k p_k(u))$

$$\min_{q'} \max_k D(p_k||q') \leq \max_k D(p_k||q)$$

$$= \max_k \sum_{u \in \mathcal{U}} p_k(u) \log_2 \frac{p_k(u)}{\max_j p_j(u)} + \sum_{u \in \mathcal{U}} p_k(u) \log_2 Z$$

$$\leq \max_k \sum_{u \in \mathcal{U}} p_k(u) \log_2 Z$$

$$= \log_2 Z$$

where the inequality is due to the fact that for all $u$, $p_k(u) \leq \max_j p_j(u)$.

(e) (4 pts) Show that $Z \leq \min\{K, |\mathcal{U}|\}$.

We have two upper bounds on $Z$, (i)

$$\sum_{u \in \mathcal{U}} \max_k p_k(u) \leq \sum_{u \in \mathcal{U}} 1 = |\mathcal{U}|$$

and, (ii)

$$\sum_{u \in \mathcal{U}} \max_k p_k(u) \leq \sum_k \sum_{u \in \mathcal{U}} p_k(u) = \sum_k \sum_{u \in \mathcal{U}} p_k(u) = \sum_k 1 = K.$$

Combining these two upper bounds give us

$$Z \leq \min\{K, |\mathcal{U}|\}.$$
problem 3. (12 points)

Suppose $p_1, p_2, \ldots, p_K$ are probability distributions on the finite alphabet $\mathcal{U}$. Let $H_1, \ldots, H_K$ be the entropies of these distributions, and let $H = \max_k H_k$. Fix $\epsilon > 0$ and for each $n \geq 1$ consider the set

$$T(n, \epsilon) = \bigcup_k T(n, p_k, \epsilon)$$

where $T(n, p_k, \epsilon)$ is the set of $\epsilon$-typical sequences of length $n$ with respect to the distribution $p_k$, i.e.,

$$T(n, p_k, \epsilon) = \{u^n \in \mathcal{U}^n : \forall u' \in \mathcal{U}^n \left| \frac{1}{n} N_w(u^n) - p_k(u') \right| < \epsilon p_k(u') \}$$

where $N_w(u^n)$ is the number of occurrences of $u'$ in sequence $u^n$.

Suppose that $U_1, U_2, \ldots$ are i.i.d. with distribution $p$ where $p$ is one of $p_1, \ldots, p_K$, i.e., $p \in \mathcal{P} = \{p_k : k = 1, \ldots, K\}$.

(a) (4 pts) Show that $\lim_{n \to \infty} \Pr((U_1, \ldots, U_n) \in T(n, \epsilon)) = 1$. (In particular for any $\delta > 0$, for $n$ large enough $\Pr(U^n \in T(n, \epsilon)) > 1 - \delta$.)

We have for all $k, n$ and $\epsilon$, $P((U_1, \ldots, U_n) \in T(n, p_k, \epsilon)) \leq P((U_1, \ldots, U_n) \in T(n, \epsilon))$ as $T(n, \epsilon) \supseteq T(n, p_k, \epsilon)$. This implies that for any $\epsilon > 0$, with $k$ and $p$ such that $p_k = p$, we have

$$\lim_{n \to \infty} \Pr((U_1, \ldots, U_n) \in T(n, p_k, \epsilon)) \leq \lim_{n \to \infty} \Pr((U_1, \ldots, U_n) \in T(n, \epsilon))$$

$$1 \leq \lim_{n \to \infty} \Pr((U_1, \ldots, U_n) \in T(n, \epsilon)).$$

where the second line is due to the property of typical sets.

As we also have $\lim_{n \to \infty} \Pr((U_1, \ldots, U_n) \in T(n, \epsilon)) \leq 1$, with these inequalities we prove the statement.

(b) (4 pts) Show that for large enough $n$, $\frac{1}{n} \log |T(n, \epsilon)| < (1 + \epsilon)H + \epsilon$.

For typical sets, we know that $|T(n, p_k, \epsilon)| \leq 2^{(1+\epsilon)H_kn} \leq 2^{(1+\epsilon)Hn}$. Hence, we obtain the following upper bound.

$$|T(n, \epsilon)| = \left| \bigcup_k T(n, p_k, \epsilon) \right| \leq \sum_k |T(n, p_k, \epsilon)| \leq K 2^{(1+\epsilon)Hn}.$$

By taking logarithm and dividing by $n$ the above expression, we have

$$\frac{1}{n} \log |T(n, \epsilon)| \leq (1 + \epsilon)H + \frac{\log K}{n}.$$

This implies that for any $n \geq \log K/\epsilon$ we have

$$\frac{1}{n} \log |T(n, \epsilon)| \leq (1 + \epsilon)H + \epsilon.$$

(c) (4 pts) Fix $R > H$ and $\delta > 0$. Show that for $n$ large enough there is a prefix-free code $c : \mathcal{U}^n \to \{0, 1\}^*$ such that

$$\Pr\left( \text{length}(c(U^n)) < nR \right) > 1 - \delta.$$

Let us use the construction of prefix-free code for typical set given in the lectures. First, take an injective function $f_{\epsilon,n} : T(n, \epsilon) \to \{0, 1\}^{[n(1+\epsilon)H+\epsilon]}$, this function exists.
for large enough \( n \) due to our result in (b). Now take another injective function \( g_n : U^n \to \{0,1\}^{\lceil n \log |U| \rceil} \). We define \( c_{\epsilon,n}(x) \) as 0 if \( x \in T(n, \epsilon) \) and 1 if \( g_n \) otherwise, where || is the concatenation operator.

We have that

\[
\Pr\left( U^n \in T(n, \epsilon) \right) = \Pr\left( \text{length}(c_{\epsilon,n}(U^n)) = \lceil n(1 + \epsilon) H + n\epsilon \rceil + 1 \right) \\
\leq \Pr\left( \text{length}(c_{\epsilon,n}(U^n)) \leq n(1 + \epsilon) H + n\epsilon + 2 \right).
\]

From (a) we know that there exists an \( n_a(\epsilon, \delta) \) such that \( 1 - \delta < \Pr\left( U^n \in T(n, \epsilon) \right) \) for all \( n \geq n_a(\epsilon, \delta) \). From (b) we require \( n \geq \log K/\epsilon = n_b(K, \epsilon) \). To get the form required in the problem statement, we need that:

\[
n((1 + \epsilon) H + \epsilon + 2/n) < nR
\]

Since \( 1/n \leq \epsilon \) for \( n \geq n_b(K, \epsilon) \), the following inequality will also work

\[
n((1 + \epsilon) H + 3\epsilon) < nR.
\]

The above inequality satisfied by choosing an appropriate \( \epsilon \) (i.e., \( 0 \leq \epsilon < \frac{R-H}{H+3} \)).

Therefore, for a code \( c_{\epsilon,n} \) constructed as above and \( \epsilon \) chosen small enough, we have

\[
\Pr\left( \text{length}(c_{\epsilon,n}(U^n)) < nR \right) \geq \Pr\left( \text{length}(c_{\epsilon,n}(U^n)) \leq n(1 + \epsilon) H + n\epsilon + 2 \right) \\
\geq \Pr\left( U^n \in T(n, \epsilon) \right) > 1 - \delta
\]

for all \( n > \max\{n_a(\epsilon, \delta), n_b(K, \epsilon)\} \).
Problem 4. (10 points)

Suppose \( C_p \) is a prefix-free binary code for non-negative integers \( \{0, 1, 2, \ldots \} \). Suppose \( C_i \) is an injective code for an alphabet \( U \).

(a) (4 pts) Show that \( C \) defined by \( C(u) = C_p(l(u))C_i(u) \), with \( l(u) = \text{length}(C_i(u)) \) is a prefix-free code for \( U \).

We need to show that for any \( u \) there is no \( u' \) such that \( C(u') \) is a prefix of \( C(u) \). We can divide it into two cases;

- The set of \( u' \) such that \( l(u) = l(u') \). In this case \( \text{length}(C(u)) = \text{length}(C(u')) \), but \( C(u) \neq C(u') \) due to the assumption that \( C_i \) is injective. This implies no such \( u' \) exists.
- The set of \( u' \) such that \( l(u) \neq l(u') \). As we assume that \( C_p \) is prefix-free, it implies that \( C(u') \) must always have a prefix that is not a prefix of \( C(u) \). Therefore no such \( u' \) exists.

Observe that (i) the code \( C_a \) with \( C_a(j) = 0 \uparrow j \), (i.e., \( C_a(0) = 1, C_a(1) = 01, C_a(2) = 001, \ldots \) ) is prefix-free with \( \text{length}(C_a(j)) = j + 1 \), and (ii) the code \( C_b \) for non-negative integers with

\[
C_b(0) = \lambda, \quad C_b(j) = \text{bin}(j - 1), \quad j > 0
\]

where \( \text{bin}(j) \) denotes the binary expansion of the integer \( j \), (i.e., \( \text{bin}(0) = 0, \text{bin}(1) = 1, \text{bin}(2) = 10, \text{bin}(3) = 11, \ldots \) ) is injective with \( \text{length}(C_b(j)) = \lfloor \log_2(j + 1) \rfloor \).

(b) (2 pts) Show that there exists a prefix-free code \( C' \) for non-negative integers with

\[
\text{length}(C'(j)) = 2\lfloor \log_2(j + 1) \rfloor + 1, \quad j \geq 0.
\]

We take \( C_p = C_a \) and \( C_i = C_b \). Therefore, by result on (a), we have

\[
\text{length}(C'(j)) = l_b(j) + l_a(l_b(j)) = \lfloor \log_2(j + 1) \rfloor + \lfloor \log_2(j + 1) \rfloor + 1
\]

where \( l_b(j) = \text{length}(C_b(j)) \) and \( l_a(j) = \text{length}(C_a(j)) \).

(c) (4 pts) Consider a sequence of functions

\[
\begin{align*}
l_1(j) &= 2\lfloor \log_2(j + 1) \rfloor + 1 \\
l_n(j) &= \lfloor \log_2(j + 1) \rfloor + l_{n-1}(\lfloor \log_2(j + 1) \rfloor), \quad n > 1.
\end{align*}
\]

Show that for each \( n > 0 \) there exists a prefix-free code for non-negative integers \( C_n \) such that

\[
\text{length}(C_n(j)) = l_n(j).
\]

[Hint: use induction.]

We define the code recursively as

\[
\begin{align*}
C_1(j) &= C_a(j)C_b(j) \\
C_n(j) &= C_{n-1}(j)C_b(j), \quad n > 1
\end{align*}
\]

The code \( C_1 \) is prefix-free and satisfies the length requirement due to (b). The code \( C_n \) is prefix-free due to (a) in which we take \( C_{n-1} \) as the prefix-free code and \( C_b \) as the injective code. It also satisfies the length requirement as

\[
\text{length}(C_n(j)) = l_a(j) + l_{n-1}(l_a(j)) = \lfloor \log_2(j + 1) \rfloor + l_{n-1}(\lfloor \log_2(j + 1) \rfloor).
\]