CPSC 418 / MATH 318 — Introduction to Cryptography
ASSIGNMENT 1 Problems 1-4

Name: Aifaz Dhuka
Student ID: 30069823

Problem 1 — Password length and entropy

(a) 8*7=56 bit values for a 8 character string.
So total 8 character ASCII encodings = 2^{56}

(b) i. Total number of printable ASCII characters = $26 + 26 + 10 + 32 = 94$
So total number of 8-character passwords = 94^8

ii. Total lowercase letters = 26
So total number of 8 character lower-case only passwords = 26^8

(c) i. $\frac{94^8}{2^{56}} * 100 \approx 8.459\%$

ii. $\frac{26^8}{2^{56}} * 100 \approx 0.00028\%$

(d) i. So there are 94 printable characters. For 8 character string the total possible passwords are 94^8. Since each character in the password is equally likely to be chosen, it implies that each password is equally likely to be chosen. So probablity of choosing a password is $\frac{1}{94^8}$. So entropy is $\sum(-p(x_i) * \log_2(p(x_i)))$. Since there are 94^8 passwords and since the value of this expression, $(-p(x_i) * \log_2(p(x_i)))$, is the same, we can simplify the summation as $94^8 * (-p(x_1) * \log_2(p(x_1)))$.

$$94^8 * (-p(x_1) * \log_2(p(x_1))) = 94^8 * (-\frac{1}{94^8} * \log_2(\frac{1}{94^8})) = 94^8 * \frac{1}{94^8} * \log_2(94^8) = \log_2(94^8)$$

So the entropy is $\log_2(94^8) \approx 52.43671$

ii. Similar to the previous part, we have total 26^8 passwords. So probality of choosing a password is $\frac{1}{26^8}$. The entropy will be

$$\sum(-p(x_i) * \log_2(p(x_i))) = 26^8 * (-p(x_i) * \log_2(p(x_i)))$$

$$26^8 * (-\frac{1}{26^8} * \log_2(\frac{1}{26^8})) = 26^8 * \frac{1}{26^8} * \log_2(26^8) = \log_2(26^8)$$

So the entropy is $\log_2(26^8) \approx 37.6035$

(e) i. So there are 94 printable characters. Let the length of the password be x. So $128 = \sum \frac{1}{94^x} * log_2(94^x) = 94^x * \frac{1}{94^x} * log_2(94^x) = log_2(94^x)$. So we can calculate x using the equation $128 = log_2(94^x)$. So,

$$\begin{aligned} 128 &= log_2(94^x) \\ 2^{128} &= 94^x \\ log_{94}(2^{128}) &= x \\ x &\approx 19.5283 \end{aligned}$$

So to get at least a 128 entropy, we need password's length to be at 20.

ii. So there are 26 lower case characters. Let the length of the password be x. So $128 = \sum \frac{1}{26^x} * log_2(26^x) = 26^x * \frac{1}{26^x} * log_2(26^x) = log_2(26^x)$. So we can calculate x using the equation $128 = log_2(26^x)$. So,

$$\begin{aligned} 128 &= log_2(26^x) \\ 2^{128} &= 26^x \\ log_{26}(2^{128}) &= x \\ x &\approx 27.2315 \end{aligned}$$

So to get at least a 128 entropy, we need password's length to be at 28.

Problem 2 — One-time pad without the all-zeros key

(a) Suppose plaintexts and ciphertexts are of one bit. So, we know $p(M = 0) = 0.5$, $p(M = 1) = 0.5$, $p(C = 0) = 0.5$ and $p(C = 1) = 0.5$
Suppose we have the modified one time pad. So, the only possible keys for it is 1 (i.e. we exclude 0^1 as a key)

K	M	C
1	0	1
1	1	0

So the probability of the possible cipher-texts 0 and 1 occurring are 0 and 1 respectively.

We can further calculate that $p(C = 0|M = 0) = 0$ since $p(K = 0) = 0$ (can also be seen from the table above). So $p(C = 0) = 0.5 \neq 0 = p(C = 0|M = 0)$ So according to the characterization of perfect secrecy, we know that the modified one time pad does not provide perfect secrecy.

(b) Including 0^n in the keyspace does not weaken the security because a ciphertext made using key 0^n could be mapped to other plaintext using another key.

For example:

K	M	C
1	0	1
1	1	0
0	0	0
0	1	1

We see that a ciphertext $C = 1$ could be mapped to either 0 or 1 based on the key. So having 0^n in the keyspace does not give away any information about the plaintext, or does not decrease or increase the likelihood of a combination i.e. all the combinations are equally likely to occur.

Problem 3 — Weak collisions

(a) p(participant choosing N)$= \frac{1}{n}$

(b) p(participant not choosing N)$= 1 - \frac{1}{n} = \frac{n-1}{n}$

(c)

$$\begin{aligned} p(\text{no weak collision}) &=^K C_0 * (\frac{1}{n})^0 * (\frac{n-1}{n})^K \\ &= 1 * 1 * (\frac{n-1}{n})^K = (\frac{n-1}{n})^K \end{aligned}$$

(d) p(weak collision) $= 0.5$
p(no weak collision) $= 1 - p$(weak collision) $= 0.5$
$n = 10$
to find K, we substitute the values to the equation from part c

$$\begin{aligned} p(\text{weak collision}) &= 1 - (\frac{n-1}{n})^K \\ 0.5 &\leq 1 - (\frac{10-1}{10})^K \\ 0.5 - 1 = -0.5 &\leq -(\frac{10-1}{10})^K \\ 0.5 &\geq (\frac{9}{10})^K \\ \log_{0.9} 0.5 &\geq K \\ 6.5788 &\geq K \end{aligned}$$

So when K=7 or greater, we have atleast 50% chance of a weak collision.

(e) Suppose $K \geq \log_2 n \approx 0.69n$
p(weak collision) $= 0.5 = 1 - p$(no weak collision)

$$\begin{aligned} p(\text{weak collision}) &= 1 - (\frac{n-1}{n})^K \\ 0.5 &\leq 1 - (\frac{n-1}{n})^K \\ 0.5 - 1 = -0.5 &\leq -(\frac{n-1}{n})^K \\ 0.5 &\geq (\frac{n-1}{n})^K \\ 0.5 &\geq (e^{-\frac{1}{n}})^K > (1 - \frac{1}{n})^K \\ 0.5 &\geq e^{-\frac{K}{n}} \\ \ln 0.5 &\geq -\frac{K}{n} \\ -\ln 0.5 = \ln 2 &\leq \frac{K}{n} \\ n \ln 2 \approx 0.69n &\leq K \end{aligned}$$

Hence the number of participants K should be at least $n \ln 2$ to ensure at least 50% probability of weak collision.

Problem 4 — (Strong) collisions

(a) First person can choose any number, so they have $\frac{n}{n}$ choices.
Second can choose from $\frac{n-1}{n}$ numbers.
Similarly, Kth person can choose from $\frac{n-k+1}{n}$.
So prbablity of k participants choosing different numbers is

$$= \frac{n}{n} * \frac{n-1}{n} * \ldots * \frac{n-k+1}{n} = \frac{(n) * (n-1) * \ldots * (n-k+1)}{n^k} = \frac{{}^nP_k}{n^k}$$

(b) So probality of a collision will be (using the part a)

$$= 1 - (\text{Probablity of no collisions}) = 1 - \frac{{}^nP_k}{n^k}$$

(c) P(Strong collision)=0.5
n=10

$$\begin{aligned}
0.5 &\leq 1 - \frac{{}^nP_k}{n^k} \\
0.5 &\leq 1 - \frac{{}^{10}P_k}{10^k} \\
0.5 &\leq \frac{10^k - {}^{10}P_k}{10^k} \\
0.5 &\leq 1 - \frac{\frac{10!}{(10-k)!}}{10^k} \\
0.5 - 1 &\leq -\frac{\frac{10!}{(10-k)!}}{10^k} \\
1 - 0.5 = 0.5 &\geq \frac{\frac{10!}{(10-k)!}}{10^k} \\
0.5 &\geq \frac{10!}{(10-k)! * 10^k}
\end{aligned}$$

So by trial and error, when $k \geq 5$, p(strong collision) is at least 0.5 when $n = 10$.

(d) The above expression could be expressed as

$$P = \prod_{i=1}^{k-1} (1 - \frac{i-1}{n})$$

using counting. So basically used the same logic that the first person can choose any number and the second chooses everything other than the first one to avoid collisions and so on.
Proving the statement: $P \leq exp(-\frac{k(k-1)}{2n})$ for all $k \leq n$ using induction.
Base case:(k=1)

$$\begin{aligned}
\text{So } P = (1 - \frac{k-1}{n}) &= (1 - \frac{1-1}{n}) = 1 \\
exp(-\frac{k(k-1)}{2n}) &= exp(-\frac{1(1-1)}{2n}) = exp(0) = 1
\end{aligned}$$

So for the base case (k=1), $P \leq exp(-\frac{k(k-1)}{2n})$ holds.
Inductive Hypothesis: Suppose the statement is true for $k = m$, where $m \leq n-1$. So,

$$P = \prod_{i=1}^{m}(1 - \frac{i-1}{n}) \leq exp(-\frac{m(m-1)}{2n})$$

Inductive step: Proving the statement is also true for the k=m+1. We want to prove

$$P = \prod_{i=1}^{m+1}(1 - \frac{i-1}{n}) \leq exp(-\frac{(m+1)(m+1-1)}{2n}) = exp(-\frac{m(m+1)}{2n})$$

So,

$$\begin{aligned}
\prod_{i=1}^{m+1}(1 - \frac{i-1}{n}) &= (1 - \frac{m+1-1}{n})\prod_{i=1}^{m}(1 - \frac{i-1}{n}) \\
&\leq (1 - \frac{m}{n}) * exp(-\frac{m(m-1)}{2n}) \qquad \textit{(by inductive hypothesis)} \\
&< exp(-\frac{m}{n}) * exp(-\frac{m(m-1)}{2n}) \\
&= exp((-\frac{m}{n}) + (-\frac{m(m-1)}{2n})) \\
&= exp(-\frac{m}{n} - \frac{m^2-m}{2n}) \\
&= exp(\frac{-2m}{2n} + \frac{-m^2-(-m)}{2n}) \\
&= exp(\frac{-m^2+m-2m}{2n}) \\
&= exp(\frac{-m^2-m}{2n}) = exp(\frac{-(m^2+m)}{2n}) \\
&= exp(-\frac{(m^2+m)}{2n}) = exp(-\frac{m(m+1)}{2n})
\end{aligned}$$

Hence it is true for k=m+1.
Thus by mathematical induction, the statement $P \leq exp(-\frac{k(k-1)}{2n})$ is true for $k \leq n$.

(e) NOTE: P is the probability of no collisions
Knowing that when k is very small relative to n, $P \approx exp(-\frac{k(k-1)}{2n})$. We want to prove that when $k \geq \sqrt{\ln 4 * n}$, then $P_{collision} \geq 0.5$. Since k and k-1 are very close because of how small k is relative to n, we can approximate k-1 to k as well. So, $exp(-\frac{k(k-1)}{2n}) \approx exp(-\frac{k^2}{2n})$. Suppose $k = \sqrt{\ln 4 * n}$. So,

$$\begin{aligned}
P \leq exp(-\frac{k^2}{2n}) = exp(-\frac{(\sqrt{\ln 4 * n})^2}{2n}) &= exp(-\frac{\ln 4 * n}{2n}) \\
&= exp(-\frac{\ln 4}{2}) = 0.5
\end{aligned}$$

So we know that when $k \geq \sqrt{\ln 4 * n}$, $P \leq 0.5$. Since $P_{collision} = 1 - P$, we know that as P decreases, $P_{collision}$ increases. So when $P < 0.5$, this implies that $P_{collision} > 0.5$. Thus for having at least 0.5 chance of a strong collision, the following condition for k must be fulfilled: $k \geq \sqrt{\ln 4 * n}$.