Cryptology Exercise Week 2

Zijun Yu 202203581

September 2023

Question 1

Since every letter is independent, the toal information we learn is the sum of the information we learn from each letter. And each letter follows the same distribution, so we can just calculate the information of one letter and multiply it by N. There are two possible events on each letter, either it gets revealed or not. We will learn $\log_2(1/p_i)$ bits of information if it gets revealed, for the probability of being revealed is p_i. And we will learn $\log_2(1/(1-p_i))$ if it is not revealed. Therefore, the average amount of the information we learn is

$$\begin{aligned}
H &= H(L_1, L_2, ..., L_N) \\
&= H(L_1) + H(L_2) + ... + H(L_N) \\
&= N * H(L_1) \\
&= N(p_i * log\frac{1}{p_i} + (1-p_i) * log\frac{1}{1-p_i})
\end{aligned}$$

Question 2, 3, and 4

$$\begin{aligned}
\frac{d}{dp_i}H &= N(log\frac{1}{p_i} - log\frac{1}{1-p_i}) \\
\frac{d^2}{d(p_i)^2}H &= N\frac{1}{(p_i-1)p_i}
\end{aligned}$$

The constants introduced by the base of the logarithm being 2 are omitted for simplicity.

From above, we can see that the second derivative is always negative, since p_i is between 0 and 1, so the first derivative is always decreasing. And the first derivative is 0 when $p_i = 1/2$. Therefore, the entropy is maximized when $p_i = 1/2$.

Since in English, as well as in most alphabetic languages, the frequency of the most frequent letter is not above 1/2, chosing the most frequent letter is the best strategy. But, of course, in the general case of this game, if there are letters whose frequency are above 1/2, then we should choose the one whose frequency is closest to 1/2, instead of the highest one.