Name: Joshua Dow
Student ID: 10150588

Problem 1 — Superencipherment for substitution ciphers, 12 marks

(a) i. Proof of Superencipherment
Suppose two keys $K_1, K_2 \in \mathbb{K}$ and some message $m \in \mathbb{M}$. Then $E_{K_1}(m) \equiv m + K_1 (\mod 26)$ and $E_{K_2}(m) \equiv m + K_2 (\mod 26)$ by definition of the shift cipher. Let us apply the Superencipherment two different ways:

A. $E_{K_1}(E_{K_2}(m)) \equiv E_{K_1}(m + K_2(\mod 26)) \equiv m + K_2 + K_1(\mod 26)$

B. $E_{K_2}(E_{K_1}(m)) \equiv E_{K_2}(m + K_1(\mod 26)) \equiv m + K_1 + K_2(\mod 26)$

Note that both cases are equivalent from the associative property of cyclic groups, of which $\mathbb{Z}/26\mathbb{Z}$ is. In any case we could have selected a third, and different key $K_3 \in \mathbb{K}$ such that $K_3 \equiv K_1 + K_2(\mod 26)$ and applied only a single cipher: $E_{K_3}(m) \equiv m + K_3(\mod 26) \equiv m + K_1 + K_2(\mod 26)$ as required. The resulting key is $K_3 \equiv K_1 + K_2(\mod 26)$

ii. Proof of Superencipherment by induction on n.
Base Case: Let $n = 2$, we have shown this to be true in part i. The cases for $n = 0, 1$ are trivial.
Inductive Hypothesis: Suppose $l \geq 2$ such that $E_{K_l}(E_{K_{l-1}}(...E_{K_1}(m)...)) \equiv E_j(m)(\mod 26)$ where $j \in \mathbb{K}$ and $j \equiv K_l + K_{l-1} + ... + K_1(\mod 26)$.
We wish to show this for $l + 1 \geq 2$.
Inductive Step: Suppose we have $l + 1 \geq 2$ keys such that $E_{K_{l+1}}(E_{K_l}(...E_1(m)...))$.
Then by our inductive hypothesis:
$E_{K_{l+1}}(E_{K_l}(...E_1(m)...)) \equiv E_{K_{l+1}}(E_j(m))(\mod 26)$
$E_{K_{l+1}}(E_j(m))(\mod 26) \equiv E_{K_{l+1}}(m + K_l + K_{l-1} + ... + K_1)(\mod 26)$
and by definition of the shift cipher:
$E_{K_{l+1}}(m + K_l + K_{l-1} + ... + K_1)(\mod 26) \equiv m + K_{l+1} + K_l + K_{l-1} + ... + K_1)(\mod 26)$. Then take a new key $K_g \in \mathbb{K}$ such that $K_g \equiv K_{l+1} + K_l + K_{l-1} + ... + K_1)(\mod 26)$. Then the Superencipherment is in fact a different single cipher we a different choice of key. This concludes our induction on n.

(b) Suppose a plain text message M_0 of length l. We have our first Vigenere cipher with keyword W_1 with length m and a second with keyword W_2 and length n. Our first cipher encrypts each letter of M_0 individually using W_1 repeated if $m < l$. Each character is shifted individually by adding its own English alphabet index (0,...,25) plus the corresponding index of the character in W_1 such that the new characters index = $M_{0index} + W_{1index}(\mod 26)$. The second cipher repeats this process but our plain text is the resulting cipher text from our first encryption. This double encryption is equivalent to if we had encoded our original plain text such that each new character index = $M_{0index} + W_{1index} + W_{2index} = M_{0index} + W_{index}$. W is obtained by adding the modulo 26 of the character indexes of W_1, W_2. The length is m if $m \geq n$ and n if $n > m$.

Problem 2 — Key size versus password size, 21 marks

(a) There are $2^7 = 128$ ASCII encodings of single characters so if you have 8 characters then there are $2^7 \times 2^7 \times 2^7 \times 2^7 \times 2^7 \times 2^7 \times 2^7 \times 2^7 = 2^{56}$ encodings.

(b) i. $94^8 = 6.09 \times 10^{15}$ which is slightly more than 2^{52}.

ii. $\frac{94^8}{2^{56}} \times 100\% = 8.459\%$

(c) $8 \log_2(94) = 8 \times 6.554588852 = 52.43671082$ bits

(d) $H(X) = 8 \log_2(26) = 8 \times 4.700439718 = 37.60351774$ bits

(e) i. $\frac{128}{6.554588852} = 19.528$ characters, so 20 characters.

ii. $\frac{128}{4.700439718} = 27.231$ characters, so 28 characters

Problem 3 — Equiprobability maximizes entropy for two outcomes, 12 marks

(a) $H(X) = Pr(X_1)\log_2(\frac{1}{Pr(X_1)} + Pr(X_2)\log_2(\frac{1}{Pr(X_2)}$
$= \frac{1}{4}\log_2(4) + \frac{3}{4}\log_2(\frac{4}{3})$
$= \frac{2}{4} + \frac{3}{4}\log_2(4/3)$
$= 0.8112781244591328$ bits

(b) Proof:
Suppose H(X) is maximal, then $H(X)\frac{d}{dx} = 0$ and $H(X) = p\log_2(\frac{1}{p}) + (1-p)\log_2(\frac{1}{(1-p)})$
$\frac{d}{dp}H(X) = \frac{d}{dp}p\log_2(\frac{1}{p}) + \frac{d}{dp}(1-p)\log_2(\frac{1}{(1-p)})$
$= \frac{\frac{d}{dp}\log(\frac{1}{p})}{\log(2)} + \frac{d}{dp}(1-p)\log_2(\frac{1}{(1-p)})$
$= \log_2(\frac{1}{p}) - \frac{1}{\log(2)} + \frac{d}{dp}(1-p)\log_2(\frac{1}{(1-p)})$
$= \log_2(\frac{1}{p}) - \frac{1}{\log(2)} - \log_2(\frac{1}{(1-p)}) + \frac{1}{\log(2)}$
$= \log_2(\frac{1}{p}) - \log_2(\frac{1}{(1-p)})$
$0 = \log_2(\frac{1}{p}) - \log_2(\frac{1}{(1-p)})$
$\log_2(\frac{1}{p}) = \log_2(\frac{1}{(1-p)})$
$e^{\log_2(\frac{1}{p})} = e^{\log_2(\frac{1}{(1-p)})}$
$\frac{1}{p} = \frac{1}{(1-p)}$ $p = 1-p$
$p = \frac{1}{2}$ as required.

(c) H(X) is maximal when the probabilities are equally likely, and we can express the entropy as $\log_2(n)$. Since we have only two outcomes, $H(X) = \log_2(2) = 1$

Problem 4 — Conditional entropy, 12 marks

(a) $H(M|C) = \sum_{c\in\mathbb{C}} Pr(c) \sum_{m\in\mathbb{M}} Pr(M|C)\log_2(\frac{1}{P(M|C)})$

i. $C_1 : \frac{1}{4} \sum_{m\in\mathbb{M}} Pr(M|C_1)\log_2(\frac{1}{Pr(M|C_1})$
$= \frac{1}{4}(Pr(M_1|C_1)\log_2(\frac{1}{Pr(M_1|C_1)}) + Pr(M_2|C_1)\log_2(\frac{1}{Pr(M_2|C_1)})$
$+ Pr(M_3|C_1)\log_2(\frac{1}{Pr(M_3|C_1)}) + Pr(M_4|C_1)\log_2(\frac{1}{Pr(M_4|C_1)})$
$= \frac{1}{4}(\frac{1}{2}\log_2(2) + \frac{1}{2}\log_2(2) + 0 + 0)$
$= \frac{1}{4}$

ii. $C_2 : \frac{1}{4} \sum_{m\in\mathbb{M}} Pr(M|C_2)\log_2(\frac{1}{Pr(M|C_2})$
$= \frac{1}{4}(Pr(M_1|C_2)\log_2(\frac{1}{Pr(M_1|C_2)}) + Pr(M_2|C_2)\log_2(\frac{1}{Pr(M_2|C_2)})$
$+ Pr(M_3|C_2)\log_2(\frac{1}{Pr(M_3|C_2)}) + Pr(M_4|C_2)\log_2(\frac{1}{Pr(M_4|C_2)})$
$= 0 + 0 + \frac{1}{4}(\frac{1}{2}\log_2(2) + \frac{1}{2}\log_2(2))$
$= \frac{1}{4}$

iii. $C_3 : \frac{1}{4} \sum_{m\in\mathbb{M}} Pr(M|C_3)\log_2(\frac{1}{Pr(M|C_3})$
$= \frac{1}{4}(Pr(M_1|C_3)\log_2(\frac{1}{Pr(M_1|C_3)}) + Pr(M_2|C_3)\log_2(\frac{1}{Pr(M_2|C_3)})$
$+ Pr(M_3|C_3)\log_2(\frac{1}{Pr(M_3|C_3)}) + Pr(M_4|C_3)\log_2(\frac{1}{Pr(M_4|C_3)})$
$= 0 + \frac{1}{4}(\frac{1}{2}\log_2(2) + \frac{1}{2}\log_2(2) + 0)$
$= \frac{1}{4}$

iv. $C_4 : \frac{1}{4} \sum_{m\in\mathbb{M}} Pr(M|C_4)\log_2(\frac{1}{Pr(M|C_4})$
$= \frac{1}{4}(Pr(M_1|C_4)\log_2(\frac{1}{Pr(M_1|C_4)}) + Pr(M_2|C_4)\log_2(\frac{1}{Pr(M_2|C_4)})$
$+ Pr(M_3|C_4)\log_2(\frac{1}{Pr(M_3|C_4)}) + Pr(M_4|C_4)\log_2(\frac{1}{Pr(M_4|C_4)})$
$= \frac{1}{4}(\frac{1}{2}\log_2(2) + 0 + 0 + \frac{1}{2}\log_2(2))$
$= \frac{1}{4}$

$H(M|C) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = 1$

(b) Suppose the system provides perfect secrecy, then $Pr(C|M) = P(C)$ $\forall \text{m} \in \mathbb{M}, c \in \mathbb{C}$. Then from the definition of $P(M|C)$:
$Pr(M|C) = \frac{Pr(M)Pr(C|M)}{P(C)}$
$= \frac{Pr(M)Pr(C)}{Pr(C)}$
$= Pr(M)$ as required

(c) No it does not provided perfect secrecy since there is no guarantee that there is a unique key K such that $e_K(m) = c$. Also, $Pr(M) = \frac{1}{4}$ but $Pr(M|C) = 1$