Assignment 1

Security and Cryptography (IN4191)

Issue date: 18 Sep 2019
Due date: 30 Sep 2019, 23:59 CET *(hand in via Brightspace)*

- In this assignment you are going to answer questions related to **historical ciphers** and **information theory**.
- This is an individual assignment. Please mention your name and student number in submission. You have to hand in a **SINGLE** document (PDF) with your answers and your source code (if any). **Also, be aware that any form of plagiarism will not be condoned.**
- Pose your questions on the Brightspace forum, so that your fellow students can also read them.
- **Explain and motivate all your answers!**

1. **Classical Systems** *(5 points)*
   In this exercise, you are given several ciphertexts originating from classical systems. For each ciphertext, you need to answer the following questions:

   - Which encryption scheme was used?
   - What was the original plaintext message?
   - What was the encryption key?

   Or explain why it is not possible to decipher it.

   (a) *(1 point)* Consider the ciphertext:
   "JOVJVSHAL DHZ PUCLUALK MVBY AOVBZHUK FLHYZ HNV PU H ZTHSS CPSSHNL PU OVUKBYHZ HUK OHZ AOYPCLK LCLY ZPUJL"

   > **Solution:** This is a **shift cipher**. Any sensible explanation gets full points; for example checking letter frequencies, or bruteforcing the 25 possible shifts, would reveal it is a shift cipher.
   >
   > The original plaintext was:
   > "CHOCOLATE WAS INVENTED FOUR THOUSAND YEARS AGO IN A SMALL VILLAGE IN HONDURAS AND HAS THRIVED EVER SINCE"
   >
   > We have key $k = 7$.

   (b) *(2 points)* Consider the ciphertext:
   "VCALIUTOTNOCINISOMTUPUNLNCADUTSRLUACEAFAO
   CEOXEEWRVTENIESOIAMSNEAMWRCEIHHERTEETCRCOA
   AAETINESEINEHVWNPALIOLETSDHETTEHDFEOOETAAC
   CERNEVEEYTFALULKOIMARSHNEOSDFEAAHRTOTRYWNT

EOINCLNHIGNAGHNTGHRMFOTERRETESUEKTNIHLIWTF
IHNEHSTEEATHRTATHORYITYFBUTORSRHWIONMRODLE
APNESADHSDEENBAEMDEDBDEIEAWNSOTUICVSWSHUPP
LEIBEHLTHEENTASEMSEVSLTATEIRBTEODERUKTAELH
AWLIISODHTEITHNWMEHSTEPEMPLTUEABHVYAEEFNHE
TLIESERTOMNNCETBSYMDUESHCMNUAOADAPCOSSEHMT
EALNECVSIRNNEAGRAAIEWDCNOEFGORLORFOSYMEAEP
LOLBOTLWGRRGTIHLETNLEATHOWYEAUDRRPAKCPRERL
OTIHSNMOCEAKNSIVNALOASARTGEYRELASMYOLRPSUE
CTTHDIEWRCSAERTTOSWRRAOWTIEHLRHERAESECTMLO
POEOYMLSSTWTIHHNIGEADRPEENYITSOIPDFOUSENQU
IATOCCOAASTNHITCTOERHIDLTRNOAEEDTRXUTONSOE
NAERESCSTIEYELDTRHIEMNNEEIERPASSOTSRHTEOTE
FIBNASENXSXDXIE"

Solution: When observing the letter frequencies, it is quite close to those on average in English texts (see Table 7.1 in the book). This indicates that a **permutation cipher** is used.

The original plaintext message:
"CULTIVATION, CONSUMPTION, AND CULTURAL USE OF CACAO WERE EXTENSIVE IN MESOAMERICA WHERE THE CACAO TREE IS NATIVE. WHEN POLLINATED, THE SEED OF THE CACAO TREE EVENTUALLY FORMS A KIND OF SHEATH, OR EAR, TWENTY INCH LONG, HANGING FROM THE TREE TRUNK ITSELF. WITHIN THE SHEATH ARE THIRTY TO FOURTY BROWNISH-RED ALMOND-SHAPED BEANS EMBEDDED IN A SWEET VISCOUS PULP. WHILE THE BEANS THEMSELVES ARE BITTER DUE TO THE ALKALOIDS WITHIN THEM, THE SWEET PULP MAY HAVE BEEN THE FIRST ELEMENT CONSUMED BY HUMANS. CACAO PODS THEMSELVES CAN RANGE IN A WIDE RANGE OF COLORS, FROM PALE YELLOW TO BRIGHT GREEN, ALL THE WAY TO DARK PURPLE OR CRIMSON. THE SKIN CAN ALSO VARY GREATLY - SOME ARE SCULPTED WITH CRATERS OR WARTS, WHILE OTHERS ARE COMPLETELY SMOOTH. THIS WIDE RANGE IN TYPE OF PODS IS UNIQUE TO CACAOS IN THAT THEIR COLOR AND TEXTURE DOES NOT NECESSARILY DETERMINE THE RIPENESS OR TASTE OF THE BEANS INSIDE."

(Interpunction is not important, as it was removed from the ciphertext. The message was padded at the end with random characters.)

The encryption key is (6,1,7,3,5,2,4).
(Alternatively, the decryption key is (2,6,4,7,5,1,3).)

Permutation ciphers can be attacked either by hand (starting with common words, like 'the', and solving the anagram), or automatically using a dictionary.

(c) *(2 points)* Consider the ciphertext:
"FFX ZSGST UALWLFD CL RTC PVFWR UP FFX ZSGST

UALWLFD CL RTC TUQTSTR IMKSR TG G JUQM VT CSSYDITIZP QULERKBQEWULE MY JZLGYGOYE HBEWWSURR NWGST ZK TTYWZIY YGRAVFD WT EGGWLPZCQQ AP IVSXG VMBSEHF LAULS YGJWPBZ FQJELBTQ ZMGPBZHD. ORRTMNNV EVK JUQM, PB THY AGPKLBE TUPY, BBK BZH YRMZBSWDS ALFGE AVP FKLMGLZOYQK, RTC YPFDH YSOF EPGEG UD ECOLB HCTBQPL KOES LPAK MOS DSIMZB VLBEIXW NA. MOS ZFOEULTS ZTGZ GZQIPFPR OLZSFLFLPRC HCKZWZBY RTPHBUS HNC MEXZ, CQHKL XGLAWYU YCHCG LBEFOCE. MY AVP CXGSGGHZ DSBCZ UHURPFY, MZJR VBP HNC SPXHH AMXYYGW VT RWFY (MJLV QLZRCP RAL DJFGKUB HM YSILS, MDMLF EVK NTYKHCS KNM NSBSH TH), ZFQ MEKSDH UD FFX HBNWKLF UHURPFY PQKTPBD FKJMRBCSWM OLFYVA. HSS IMXMLZID CL PTMWLG, EVK JUEAAVZIYC AD TSSIOTBDGT, AVP AGSEMELIX OZ FMJBJOCBGQESL, AVP HKKBJX VT LFZCYGL HBO HNC ERTAIP CL XQSL DSCS GJX BXZHCCECP. RAL ZZQGRUMG HBO IRRUKTAS QOZC AD MOS SOTEULZ NOCRKLE YKL IYYTMIL, TUR EVKPQ GL ZDPQAJMRBVB EVGR FFXF ALM TMF FTCS PLOQFCW HH LZR. RTC EPGE QUTQPXK CYZE RTC LJIWDZSDYE HBO OXATGMLQEIXYX KHUIXSTRE MY AVP AKBURXYFLBKYZ YGK ATRJJQ CTZHPFT PQEBVBD, KNGOF MOSY QUKBPBZSO HNC WLHDB HCXJP DHY HSS MPQCDZ. VPBIC, QVMHBE GORQQ ULMZBJ RTGL YSLZS UQPX UCE QULEGWLFPR GQ BYKA CQ QULFCFWCCOXW MAVVIYHY. RTC IYWXOXW MAVVIYHY, AAKBUU QFUK TCESSYWYRUA PYWESXQ, MJLV VPOBGXW BUTWIKLOCW AVP DRYOCL PBNZABQB BU HSS CMZBXYG WWYR. RGOL CQ HNC ECOLB PBZPUCL HFP O ICXCUYOEWUL AD ZYSPY GAOMFWZTGNKQLMZ WY HNC MPMZ OYR GPOFBASNHAPQ (RAL SIQKNFGHUG MSOLS RAL DJFGKUBL VT RWFY MLW AVP VGLSGGN ULFJCZQ HM PLPEJAL). MOS DSBCZ UHURPFY MZ YGAWAOZCD'Q EPGE KUL BPTPGPG LMD RALWC BURMZEL TPOZSDCL, YOYUOLS DKVA DIVCDJTAWGSY MR RAL VTUNCER HY ZLFMCER HM HSSOP FWILG, EC ZFQ YKAWDHXW IGMO KSWIF FFXF KPFK CJCVBHPR. ZFQGK HFNVORQAMBFLZ GLP YKAWDHOA RCTAICSY UQPX PATHGRQB MOFZIMFASM AVP VKJXCGPGEWI UAPEK OYR HCKMGK. HSS MPQCD PBQZACZAX PB CCSYZ ANSHFFK, YZB MOS CSBGHYE VT RFKAA-PHTOY OXRUQMPQ DHEJQQ WBFTBM RTC KLBLWYQMLVL QLIMFF RAL WXOMGZYMPCY CL CGPHWSLB GPFGLAG LBJ RDYOLZWSXQ. BYBUHTBMQ MLW ZQFZVRGPXZ OWZABULZ AC LBZGBYMLF'D ZOQF UXYS XOJC, IFBSS LRBCZRNYSCG LJAADLR EC ZFQ YVAILZ YGFCL AC ASXQALTSZJ KORZCLZ HSS CMZBXYG. WSMCZBL JWCQAJMRXK HZ TAPFFXY QZAVJQKXUH EVK QGNXYZLHOTQQ HM HSS CMZBXYG."

Solution: All letters occur roughly in the same frequency. This indicates that we are not dealing with a permutation, shift or substitution cipher. However, we do see some bigrams and trigrams occur multiple times. This indicates that a **Vigenère cipher** has been used.

The original plaintext message:
"THE SEVEN WONDERS OF THE WORLD OR THE SEVEN WONDERS OF THE ANCIENT WORLD IS A LIST OF REMARKABLE CONSTRUCTIONS OF CLASSICAL ANTIQUITY GIVEN BY VARIOUS AUTHORS IN GUIDEBOOKS OR POEMS POPULAR AMONG ANCIENT HELLENIC TOURISTS. ALTHOUGH THE LIST, IN ITS CURRENT FORM, DID NOT STABILISE UNTIL THE RENAISSANCE, THE FIRST SUCH LISTS OF SEVEN WONDERS DATE FROM THE SECOND CENTURY BC. THE ORIGINAL LIST INSPIRED INNUMERABLE VERSIONS THROUGH THE AGES, OFTEN LISTING SEVEN ENTRIES. OF THE ORIGINAL SEVEN WONDERS, ONLY ONE THE GREAT PYRAMID OF GIZA (ALSO CALLED THE PYRAMID OF KHUFU, AFTER THE PHARAOH WHO BUILT IT), THE OLDEST OF THE ANCIENT WONDERS REMAINS RELATIVELY INTACT. THE COLOSSUS OF RHODES, THE LIGHTHOUSE OF ALEXANDRIA, THE MAUSOLEUM AT HALICARNASSUS, THE TEMPLE OF ARTEMIS AND THE STATUE OF ZEUS WERE ALL DESTROYED. THE LOCATION AND ULTIMATE FATE OF THE HANGING GARDENS ARE UNKNOWN, AND THERE IS SPECULATION THAT THEY MAY NOT HAVE EXISTED AT ALL. THE LIST COVERED ONLY THE SCULPTURAL AND ARCHITECTURAL MONUMENTS OF THE MEDITERRANEAN AND MIDDLE EASTERN REGIONS, WHICH THEN COMPRISED THE KNOWN WORLD FOR THE GREEKS. HENCE, EXTANT SITES BEYOND THIS REALM WERE NOT CONSIDERED AS PART OF CONTEMPORARY ACCOUNTS. THE PRIMARY ACCOUNTS, COMING FROM HELLENISTIC WRITERS, ALSO HEAVILY INFLUENCED THE PLACES INCLUDED IN THE WONDERS LIST. FIVE OF THE SEVEN ENTRIES ARE A CELEBRATION OF GREEK ACCOMPLISHMENTS IN THE ARTS AND ARCHITECTURE (THE EXCEPTIONS BEING THE PYRAMIDS OF GIZA AND THE HANGING GARDENS OF BABYLON). THE SEVEN WONDERS ON ANTIPATER'S LIST WON PRAISES FOR THEIR NOTABLE FEATURES, RANGING FROM SUPERLATIVES OF THE HIGHEST OR LARGEST OF THEIR TYPES, TO THE ARTISTRY WITH WHICH THEY WERE EXECUTED. THEIR ARCHITECTURAL AND ARTISTIC FEATURES WERE IMITATED THROUGHOUT THE HELLENISTIC WORLD AND BEYOND. THE GREEK INFLUENCE IN ROMAN CULTURE, AND THE REVIVAL OF GRECO-ROMAN ARTISTIC STYLES DURING THE RENAISSANCE CAUGHT THE IMAGINATION OF EUROPEAN ARTISTS AND TRAVELLERS. PAINTINGS AND SCULPTURES ALLUDING TO ANTIPATER'S LIST WERE MADE, WHILE ADVENTURERS FLOCKED TO THE ACTUAL SITES TO PERSONALLY WITNESS THE WONDERS.

LEGENDS CIRCULATED TO FURTHER COMPLEMENT THE SUPERLATIVES OF THE WONDERS."

The used key is "MYTHOLOGY".

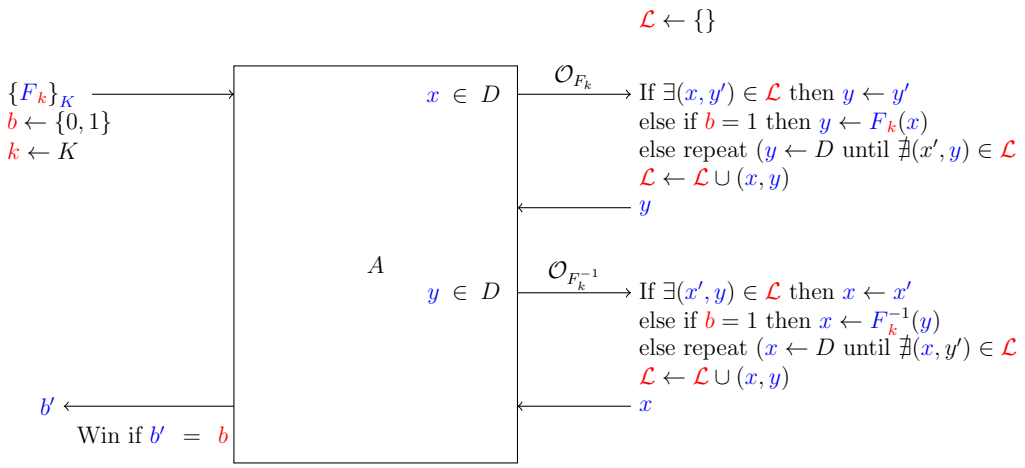
Important aspects of the answer:

- The key length is determined using bigram and trigram distances (Kasiski test).
- Key entries are found based on either letter frequencies, or based on other properties of the English language (i.e., the only single-letter words are 'I' and 'a'.).

2. **Security Games** *(2 points)*

Draw the security game for a family of pseudo-random permutations $\{F_k\}_K$, for an attacker A with access to two oracles: $\mathcal{O}_{F_k}$ and $\mathcal{O}_{F_k^{-1}}$.

Solution: The security game in this setting is similar to Figure 11.4 in the book. However, there is another oracle $\mathcal{O}_{F_k^{-1}}$ which the attacker can use to invert the output of the permutation. The challenger must configure the oracles such that they provide consistent results in both directions.



3. **Information Theoretic Security** *(3 points)*

We have the following sets of possible plaintexts, keys and ciphertexts:

$$\mathbb{P} = \{a, b, c, d\}$$

$$\mathbb{K} = \{k_1, k_2, k_3, k_4\}$$

$$\mathbb{C} = \{1, 2, 3, 4\}$$

The plaintexts and keys have the following probabilities:

$$p(\cdot) = \left\{a = \frac{2}{7}, b = \frac{1}{7}, c = \frac{3}{7}, d = \frac{1}{7}\right\}$$

$$p(\cdot) = \left\{ k_1 = \frac{1}{4}, k_2 = \frac{1}{4}, k_3 = \frac{1}{4}, k_4 = \frac{1}{4} \right\}$$

Consider the encryption scheme listed in Table 1.

	a	b	c	d
k_1	4	1	3	2
k_2	1	3	2	4
k_3	2	4	1	3
k_4	4	2	1	3

Table 1: Encryption scheme

(a) *(1 point)* Calculate the probability of each plaintext conditioned on each ciphertext occurrence.

Solution: To find the probability of the plaintexts, given a ciphertext, we need several steps:

1. The probabilities of the ciphertexts.

$$\begin{aligned} P(C=1) &= P(K=k_1) \cdot P(M=b) + P(K=k_2) \cdot P(M=a) + \\ &\quad P(K=k_3) \cdot P(M=c) + P(K=k_4) \cdot P(M=c) \\ &= \frac{1}{4} \cdot \frac{1}{7} + \frac{1}{4} \cdot \frac{2}{7} + \frac{1}{4} \cdot \frac{3}{7} + \frac{1}{4} \cdot \frac{3}{7} \\ &= \frac{9}{28} \approx 0.32143 \end{aligned}$$

$$\begin{aligned} P(C=2) &= P(K=k_1) \cdot P(M=d) + P(K=k_2) \cdot P(M=c) + \\ &\quad P(K=k_3) \cdot P(M=a) + P(K=k_4) \cdot P(M=b) \\ &= \frac{1}{4} \cdot \frac{1}{7} + \frac{1}{4} \cdot \frac{3}{7} + \frac{1}{4} \cdot \frac{2}{7} + \frac{1}{4} \cdot \frac{1}{7} \\ &= \frac{1}{4} = 0.25 \end{aligned}$$

$$\begin{aligned} P(C=3) &= P(K=k_1) \cdot P(M=c) + P(K=k_2) \cdot P(M=b) + \\ &\quad P(K=k_3) \cdot P(M=d) + P(K=k_4) \cdot P(M=d) \\ &= \frac{1}{4} \cdot \frac{3}{7} + \frac{1}{4} \cdot \frac{1}{7} + \frac{1}{4} \cdot \frac{1}{7} + \frac{1}{4} \cdot \frac{1}{7} \\ &= \frac{3}{14} \approx 0.21429 \end{aligned}$$

$$\begin{aligned} P(C=4) &= P(K=k_1) \cdot P(M=a) + P(K=k_2) \cdot P(M=d) + \\ &\quad P(K=k_3) \cdot P(M=b) + P(K=k_4) \cdot P(M=a) \\ &= \frac{1}{4} \cdot \frac{2}{7} + \frac{1}{4} \cdot \frac{1}{7} + \frac{1}{4} \cdot \frac{1}{7} + \frac{1}{4} \cdot \frac{2}{7} \\ &= \frac{3}{14} \approx 0.21429 \end{aligned}$$

2. The probabilities of the ciphertexts, given a plaintext.

$$
\begin{aligned}
p(C = 1 \mid M = a) &= p(K = k_2) = \frac{1}{4} \\
p(C = 2 \mid M = a) &= p(K = k_3) = \frac{1}{4} \\
p(C = 3 \mid M = a) &= 0 \\
p(C = 4 \mid M = a) &= p(K = k_1) + p(K = k_4) = \frac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
p(C = 1 \mid M = b) &= p(K = k_1) = \frac{1}{4} \\
p(C = 2 \mid M = b) &= p(K = k_4) = \frac{1}{4} \\
p(C = 3 \mid M = b) &= p(K = k_2) = \frac{1}{4} \\
p(C = 4 \mid M = b) &= p(K = k_3) = \frac{1}{4}
\end{aligned}
$$

$$
\begin{aligned}
p(C = 1|M = c) &= p(K = k_3) + p(K = k_4) = \frac{1}{2} \\
p(C = 2|M = c) &= p(K = k_2) = \frac{1}{4} \\
p(C = 3|M = c) &= p(K = k_1) = \frac{1}{4} \\
p(C = 4|M = c) &= 0
\end{aligned}
$$

$$
\begin{aligned}
p(C = 1 \mid M = d) &= 0 \\
p(C = 2 \mid M = d) &= p(K = k_1) = \frac{1}{4} \\
p(C = 3 \mid M = d) &= p(K = k_3) + p(K = k_4) = \frac{1}{2} \\
p(C = 4 \mid M = d) &= p(K = k_2) = \frac{1}{4}
\end{aligned}
$$

Finally, the plaintext probabilities conditioned for each ciphertext can be found using Bayes' Theorem:

$$p(M = m \mid C = c) = \frac{p(M = m) \cdot p(C = c \mid M = m)}{p(C = c)}$$

This yields the following values:

$$p(M=a \mid C=1) = \frac{p(M=a)\cdot p(C=1 \mid M=a)}{p(C=1)} = \frac{{}^{2}/_{7}\cdot {}^{1}/_{4}}{{}^{9}/_{28}} = \frac{2}{9}$$
$$p(M=b \mid C=1) = \frac{p(M=b)\cdot p(C=1 \mid M=b)}{p(C=1)} = \frac{{}^{1}/_{7}\cdot {}^{1}/_{4}}{{}^{9}/_{28}} = \frac{1}{9}$$
$$p(M=c \mid C=1) = \frac{p(M=c)\cdot p(C=1 \mid M=c)}{p(C=1)} = \frac{{}^{3}/_{7}\cdot {}^{1}/_{2}}{{}^{9}/_{28}} = \frac{2}{3}$$
$$p(M=d \mid C=1) = \frac{p(M=d)\cdot p(C=1 \mid M=d)}{p(C=1)} = \frac{{}^{1}/_{7}\cdot 0}{{}^{9}/_{28}} = 0$$

$$p(M=a \mid C=2) = \frac{p(M=a)\cdot p(C=2 \mid M=a)}{p(C=2)} = \frac{{}^{2}/_{7}\cdot {}^{1}/_{4}}{{}^{1}/_{4}} = \frac{2}{7}$$
$$p(M=b \mid C=2) = \frac{p(M=b)\cdot p(C=2 \mid M=b)}{p(C=2)} = \frac{{}^{1}/_{7}\cdot {}^{1}/_{4}}{{}^{1}/_{4}} = \frac{1}{7}$$
$$p(M=c \mid C=2) = \frac{p(M=c)\cdot p(C=2 \mid M=c)}{p(C=2)} = \frac{{}^{3}/_{7}\cdot {}^{1}/_{4}}{{}^{1}/_{4}} = \frac{3}{7}$$
$$p(M=d \mid C=2) = \frac{p(M=d)\cdot p(C=2 \mid M=d)}{p(C=2)} = \frac{{}^{1}/_{7}\cdot {}^{1}/_{4}}{{}^{1}/_{4}} = \frac{1}{7}$$

$$p(M=a \mid C=3) = \frac{p(M=a)\cdot p(C=3 \mid M=a)}{p(C=3)} = \frac{{}^{2}/_{7}\cdot 0}{{}^{3}/_{14}} = 0$$
$$p(M=b \mid C=3) = \frac{p(M=b)\cdot p(C=3 \mid M=b)}{p(C=3)} = \frac{{}^{1}/_{7}\cdot {}^{1}/_{4}}{{}^{3}/_{14}} = \frac{1}{6}$$
$$p(M=c \mid C=3) = \frac{p(M=c)\cdot p(C=3 \mid M=c)}{p(C=3)} = \frac{{}^{3}/_{7}\cdot {}^{1}/_{4}}{{}^{3}/_{14}} = \frac{1}{2}$$
$$p(M=d \mid C=3) = \frac{p(M=d)\cdot p(C=3 \mid M=d)}{p(C=3)} = \frac{{}^{1}/_{7}\cdot {}^{1}/_{2}}{{}^{3}/_{14}} = \frac{1}{3}$$

$$p(M=a \mid C=4) = \frac{p(M=a)\cdot p(C=4 \mid M=a)}{p(C=4)} = \frac{{}^{2}/_{7}\cdot {}^{1}/_{2}}{{}^{3}/_{14}} = \frac{2}{3}$$
$$p(M=b \mid C=4) = \frac{p(M=b)\cdot p(C=4 \mid M=b)}{p(C=4)} = \frac{{}^{1}/_{7}\cdot {}^{1}/_{4}}{{}^{3}/_{14}} = \frac{1}{6}$$
$$p(M=c \mid C=4) = \frac{p(M=c)\cdot p(C=4 \mid M=c)}{p(C=4)} = \frac{{}^{3}/_{7}\cdot 0}{{}^{3}/_{14}} = 0$$
$$p(M=d \mid C=4) = \frac{p(M=d)\cdot p(C=4 \mid M=d)}{p(C=4)} = \frac{{}^{1}/_{7}\cdot {}^{1}/_{4}}{{}^{3}/_{14}} = \frac{1}{6}$$

(b) *(½ point)* Compute the entropy $H(P|C=4)$.

Solution: The conditional entropy of X given some observation $Y = y$ is defined as (see §9.3 in the book):

$$H(X \mid Y = y) = -\sum_x p(X = x \mid Y = y) \cdot \log_2 p(X = x \mid Y = y)$$

So, plugging in the four plaintext messages a, b, c, d for m in $p(P = m \mid C = 4)$:

$$\begin{aligned} H(P \mid C = 4) &= -\sum_m p(P = m \mid C = 4) \cdot \log_2 p(P = m \mid C = 4) \\ &= -\left(\frac{2}{3} \cdot \log_2\left(\frac{2}{3}\right) + \frac{1}{6} \cdot \log_2\left(\frac{1}{6}\right) + 0 \cdot \log_2(0) + \frac{1}{6} \cdot \log_2\left(\frac{1}{6}\right)\right) \\ &\approx 1.251629 \end{aligned}$$

(c) *(½ point)* Compute the entropy $H(P|C)$.

Solution: The conditional entropy of X given Y (see §9.3 in the book) is defined as:

$$H(X \mid Y) = \sum_y p(Y = y) \cdot H(X \mid Y = y)$$

So, we need the conditional entropy for every ciphertext $1, 2, 3, 4$:

$$
\begin{aligned}
H(P \mid C=1) &= -\sum_m p(P=m \mid C=1) \cdot \log_2 p(P=m \mid C=1) \\
&= -\left(\frac{2}{9} \cdot \log_2\left(\frac{2}{9}\right) + \frac{1}{9} \cdot \log_2\left(\frac{1}{9}\right) + \frac{2}{3} \cdot \log_2\left(\frac{2}{3}\right) + 0 \cdot \log_2(0)\right) \\
&\approx 1.224394 \\
H(P \mid C=2) &= -\sum_m p(P=m \mid C=2) \cdot \log_2 p(P=m \mid C=2) \\
&= -\left(\frac{2}{7} \cdot \log_2\left(\frac{2}{7}\right) + \frac{1}{7} \cdot \log_2\left(\frac{1}{7}\right) + \frac{3}{7} \cdot \log_2\left(\frac{3}{7}\right) + \frac{1}{7} \cdot \log_2\left(\frac{1}{7}\right)\right) \\
&\approx 1.842371 \\
H(P \mid C=3) &= -\sum_m p(P=m \mid C=3) \cdot \log_2 p(P=m \mid C=3) \\
&= -\left(0 \cdot \log_2(0) + \frac{1}{6} \cdot \log_2\left(\frac{1}{6}\right) + \frac{1}{2} \cdot \log_2\left(\frac{1}{2}\right) + \frac{1}{3} \cdot \log_2\left(\frac{1}{3}\right)\right) \\
&\approx 1.459148 \\
H(P \mid C=4) &= -\sum_m p(P=m \mid C=4) \cdot \log_2 p(P=m \mid C=4) \\
&= -\left(\frac{2}{3} \cdot \log_2\left(\frac{2}{3}\right) + \frac{1}{6} \cdot \log_2\left(\frac{1}{6}\right) + 0 \cdot \log_2(0) + \frac{1}{6} \cdot \log_2\left(\frac{1}{6}\right)\right) \\
&\approx 1.251629
\end{aligned}
$$

Then, we compute the conditional entropy as follows:

$$
\begin{aligned}
H(P \mid C) &= \sum_c p(C=c) \cdot H(M \mid C=c) \\
&\approx \frac{9}{28} \cdot 1.224394 + \frac{1}{4} \cdot 1.842371 + \frac{3}{14} \cdot 1.459148 + \frac{3}{14} \cdot 1.251629 \\
&\approx 1.435029
\end{aligned}
$$

(d) *(1 point)* Is this scheme perfectly secure?

Solution: No, the scheme is not perfectly secure. There are several ways to proof this, including the following:

1. **Plaintext probabilities conditioned on the ciphertext** (based on the results of subquestion a): Definition 9.1 from the book states that a cryptosystem has perfect secrecy if for all $m \in \mathbb{P}$ and $c \in \mathbb{C}$:

$$p(P=m \mid C=c) = p(P=m)$$

A counterexample is $m = a$:

$$p(P = a \mid C = 1) = \frac{2}{9} \neq \frac{2}{7} = p(P = a)$$

Which proves the scheme is not perfectly secure.

2. **Shannon's Theorem** (Theorem 9.4 in the book):
   A cryptosystem provides perfect secrecy if and only if:

   - *Every key is used with equal probability* $\frac{1}{\#\mathbb{K}}$,
   - *For each* $m \in \mathbb{P}$ *and* $c \in \mathbb{C}$ *there is a unique key* k *such that* $e_k(m) = c$.

   The first condition is satisfied, as every key $k \in \mathbb{K}$ has equal probability

   $$p(k) = \frac{1}{4} = \frac{1}{\#\mathbb{K}}$$

   However, the second condition is not satisfied. For example, we have $e_k(d) = 3$ for both $k = k_3$ and $k = k_4$. In other words, there is no unique key mapping plaintext d to ciphertext 3. Thus, the scheme is not perfectly secure.