1. praktikumi Vigenère'i šifri ülesanne:

Võib proovida leida ka mingi osateksti, millele on rakendatud ROTn-i, vastastikkuse kokkulangevusindeksi avateksti keele tähesagedustega.

Maksimaalse vastastikkuse kokkulangevusindeksiga n on ilmselt võtme täht.

Meie ülesandes annabki see lähenemine juhtumisi õige võtme.

Samuti oli meie ülesandes kolm ühetähelist sõna. . . Nendest oleks juba leitav võtme kolm tähte, mis avab pooled avateksti tähed.

More hand-applied examples

Hill's cipher

- Key: a number m and an invertible sqare matrix $M \in \mathbb{Z}_{26}^{m\times m}$.
- Encoding: split the text to sequences of length m. The ciphertext corresponding to $x \in \mathbb{Z}_{26}^m$ is $x \cdot M$.
- Decoding: the plaintext corresponding to the ciphertext $y \in \mathbb{Z}_{26}^m$ is $y \cdot M^{-1}$.

Example: let $m = 3$ and

$$M = \begin{pmatrix} 15 & 2 & 13 \\ 8 & 21 & 1 \\ 14 & 16 & 7 \end{pmatrix} .$$

Then $\det M \equiv 9 \pmod{26}$, i.e. M is invertible in $\mathbb{Z}_{26}^{3\times 3}$ (because 9 is invertible in $\mathbb{Z}_{26}$).

Let the plaintext be CRYPTOGRAPHY or $(2, 17, 24), (15, 19, 14), (6, 17, 0), (15, 7, 24)$.

Multiplying all these four vectors with M (from the right) gives us the ciphertext $(8, 17, 3), (1, 3, 0), (18, 5, 17), (19, 15, 6)$ or IRDBDASFRTPG.

To decode, let us find $M^{-1}\ldots$

$$\left(\begin{array}{rrr|rrr} 15 & 2 & 13 & 1 & 0 & 0 \\ 8 & 21 & 1 & 0 & 1 & 0 \\ 14 & 16 & 7 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{rrr|rrr} 1 & 14 & 13 & 7 & 0 & 0 \\ 8 & 21 & 1 & 0 & 1 & 0 \\ 14 & 16 & 7 & 0 & 0 & 1 \end{array}\right) \to$$

Multiplied the first row with $7 = 15^{-1}$.

$$\left(\begin{array}{rrr|rrr} 1 & 14 & 13 & 7 & 0 & 0 \\ 0 & 13 & 1 & 22 & 1 & 0 \\ 0 & 2 & 7 & 6 & 0 & 1 \end{array}\right) \to \left(\begin{array}{rrr|rrr} 1 & 14 & 13 & 7 & 0 & 0 \\ 0 & 1 & 11 & 12 & 1 & 20 \\ 0 & 2 & 7 & 6 & 0 & 1 \end{array}\right) \to$$

Added the right multiples of the first row to the second and third rows. Then subtracted the sixfold third row from the second.

$$\left(\begin{array}{rrr|rrr} 1 & 14 & 13 & 7 & 0 & 0 \\ 0 & 1 & 11 & 12 & 1 & 20 \\ 0 & 0 & 11 & 8 & 24 & 13 \end{array}\right) \to \left(\begin{array}{rrr|rrr} 1 & 14 & 13 & 7 & 0 & 0 \\ 0 & 1 & 11 & 12 & 1 & 20 \\ 0 & 0 & 1 & 22 & 14 & 13 \end{array}\right) \to$$

Subtracted the twofold second row from the third. Then multiplied the third row with $19 = 11^{-1}$.

$$\left(\begin{array}{ccc|ccc} 1 & 14 & 0 & 7 & 0 & 13 \\ 0 & 1 & 0 & 4 & 3 & 7 \\ 0 & 0 & 1 & 22 & 14 & 13 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 10 & 19 \\ 0 & 1 & 0 & 4 & 3 & 7 \\ 0 & 0 & 1 & 22 & 14 & 13 \end{array}\right)$$

Added the multiples of the third row to the first and second row. Then added the multiple of the second row to the first row. Hence

$$M^{-1} = \begin{pmatrix} 3 & 10 & 19 \\ 4 & 3 & 7 \\ 22 & 14 & 13 \end{pmatrix}$$

To decode, the vectors making up the ciphertext must be multiplied with M^{-1} from the right.

$(8, 17, 3) \cdot M^{-1} = (2, 17, 24)$, etc.

Types of attacks against encryption systems

- ciphertext-only (*tuntud krüptotekstiga*)
  - Given a ciphertext, find the plaintext and/or the key.
- known-plaintext (*tuntud avatekstiga*)
  - The attacker knows a number of plaintext-ciphertext pairs. With their help, find the key or the plaintext corresponding to some other ciphertext.
- chosen-plaintext (*valitud avatekstiga*)
  - The attacker can invoke the encoding function. Find the key or the plaintext.
- chosen-ciphertext (*valitud krüptotekstiga*)
  - The attacker can invoke the decoding function. Find the key or the plaintext. The decoding function may not be invoked on the ciphertext that we have to decode.

Known-plaintext attack on Hill's cipher

Let m be known (if not, guess). let (x_i, y_i) be the pairs of known plaintext-ciphertext pairs corresponding to an unknown key. I.e. $y_i = x_i \cdot M$.

- Let $x_{i_1}, \ldots, x_{i_m}$ be linearly independent plaintexts.
- Let X be a matrix with the rows $x_{i_1}, \ldots, x_{i_m}$.
- Let Y be the matrix with the rows $y_{i_1}, \ldots, y_{i_m}$.
- $Y = X \cdot M$, hence $M = X^{-1} \cdot Y$.
- If m was unknown then we can use the other plaintext-ciphertext pairs to verify the correctness of M.

Affine Hill's cipher

Hill's cipher is just a linear transformation of $\mathbb{Z}_{26}^m$.

A more general form of it is:

- Key: $m \in \mathbb{N}$, $M \in \mathbb{Z}_{26}^{m\times m}$, $v \in \mathbb{Z}_{26}^m$, such that M is invertible.
- Encryption of $x \in \mathbb{Z}_{26}^m$ is $x \cdot M + v$.
- Decryption of $y \in \mathbb{Z}_{26}^m$ is $y \cdot M^{-1} - v$.

Exercises

- How to do a known-plaintext attack on affine Hill's cipher (assuming that m is known)?
  - How many plaintext-ciphertext pairs we need if everything necessary turns out to be linearly independent?
- If M in the key of the affine Hill's cipher is the unit matrix, what sort of cryptosystem results?

More exercises

- How resistant are Caesar cipher (a.k.a. shift cipher, *nihkešiffer*), substitution cipher (*asendusšiffer*) and Vigenère cipher against known-plaintext and chosen-plaintext attacks?
- How much corresponding plaintext and ciphertext is needed for a known-plaintext attack on a multiply applied Vigenère cipher, if the number of keys and their lengths are known?

Affine cipher

If $m = 1$ in affine Hill's cipher, then the result is called just the affine cipher.

In an affine cipher

- $\mathcal{K} = \mathbb{Z}_{26}^{*} \times \mathbb{Z}_{26}$;
- $e_{(k,a)}(x) = k \cdot x + a$ mod 26 for a character x;
- $d_{(k,a)}(y) = (y - a) \cdot k^{-1}$ mod 26 for a character y.

(to encrypt a text: encrypt each character separately)

known-plaintext cryptanalysis

It is usually sufficient to have two pairs (x_1, y_1), (x_2, y_2) of corresponding characters in plaintext and ciphertext.

Then

$$\begin{cases} y_1 = x_1 \cdot k + a \\ y_2 = x_2 \cdot k + a \end{cases} \Longrightarrow (y_1 - y_2) = (x_1 - x_2) \cdot k \Longrightarrow$$

$$k = (y_1 - y_2) \cdot (x_1 - x_2)^{-1} \text{ and } a = y_1 - x_1 \cdot k \pmod{26}$$

If $(x_1 - x_2)$ is not invertible in $\mathbb{Z}_{26}$ then we get several solutions for k.

Then we need more plaintext-ciphertext pairs.

Transposition cipher

- Key: $m \in \mathbb{N}$ and a permutation σ of $\{1, \ldots, m\}$.
- To encrypt a plaintext:
  - Write it down on rows, with m symbols per row.
    - Pad or do not pad the text, to make its length divisible by m.
  - Permute the resulting m columns according to σ.
  - Read out the ciphertext, row by row.
- To decrypt, do everything in reverse.
  - If the plaintext was unpadded, figure out which columns were taller.

Exercise: what is the relation between transposition cipher and Hill's cipher?

Example: let $m = 8$ and $\sigma = \dfrac{\begin{matrix}1 & 2 & 3 & 4 & 5 & 6 & 7 & 8\end{matrix}}{\begin{matrix}3 & 5 & 2 & 7 & 4 & 1 & 6 & 8\end{matrix}}$.

Let the plaintext be

THEFIRSTHOMEASSIGNMENTISDUEATTHETHURSDAYNEXTWEEK

$$\begin{array}{cccccccc} T & H & E & F & I & R & S & T \\ H & O & M & E & A & S & S & I \\ G & N & M & E & N & T & I & S \\ D & U & E & A & T & T & H & E \\ T & H & U & R & S & D & A & Y \\ N & E & X & T & W & E & E & K \end{array} \quad \text{permuted:} \quad \begin{array}{cccccccc} R & E & T & I & H & S & F & T \\ S & M & H & A & O & S & E & I \\ T & M & G & N & N & I & E & S \\ T & E & D & T & U & H & A & E \\ D & U & T & S & H & A & R & Y \\ E & X & N & W & E & E & T & K \end{array}$$

The ciphertext is

RETIHSFTSMHAOSEITMGNNIESTEDTUHAEDUTSHARYEXNWEETK

Cryptanalysis

- Recognizing transposition cipher: the letters in the ciphertext have the same frequency as in the plaintext.
- First, somehow guess the number of columns m.
- Write text in m columns (as by decryption) and look for anagrams.
  - Look for anagrams in rows, but also consider two rows (following each other) together.
- For example, the last row in the previous example was EXNWEETK.
  - Probably an anagram of NEXTWEEK.
  - This already fixes 5 of 8 rows.

Frequencies of di-, tri-, . . . -graphs

- Pick a column.
  - . . . with largest number of common characters.
- Put another column beside it; consider the sum of frequencies (in plaintext) of resulting bigrams.
  - Also consider row breaks; you may want to shift the other column a position up or down.
- The column with the largest such sum is the most probable neighbour.

- Using a substitution cipher and a transposition cipher together usually gives good results:
- Determining the plaintext characters for some (frequent) characters in the ciphertext does not reveal parts of words.
- Anagramming, or looking for frequent digraphs is hard if we do not know the alphabet.

Confusion and diffusion

A cipher provides good

- **diffusion** if the statistical structure of the plaintext leading to its redundancy is “dissipated” into long range statistics — into statistical structure involving long combinations of letters in the cryptotext.
- **confusion** if it makes the relation between the simple statistics of the cryptotext and simple description of the key a very complex and involved one.

(paraphrased from: Claude Shannon. *Communication Theory of Secrecy Systems*. Bell System Technical Journal **28**(4):656–715, 1949.)

Achieving confusion and diffusion

- Diffusion is usually obtained by permuting the characters.
  - Or applying a more complex linear operation on long vectors of characters.
- Confusion is achived by substituting characters (or short sequences of them).

Iterating substitution and permutation may produce good ciphers.

Somewhere the key has to be mixed in, too.

Substitution gives good confusion

- When substitution cipher has been used, it is usually easy to find the cryptotext character corresponding to “E”.
  - This maps a simple statistic of the cryptotext (counts of characters) to a simple property of the key.
- Maybe the cryptotext characters corresponding to some other frequent plaintext characters can be found this way, too.
- But for finding the rest of the substitution key, longer stretches of ciphertext have to be considered.
  - A simple property of the key can only be derived from a complex statistic of the ciphertext.
- This is confusion.

Fractionation

A character from the Latin alphabet does not have to be the "smallest unit" operated on by a cipher.

If we sacrifice a letter then we can encode each character in the plaintext as two elements of $\mathbb{Z}_5$.

This gives us a "plaintext" with $\mathbb{Z}_5$ as the alphabet.

We must have designed our cipher to work on $\mathbb{Z}_5^*$. We get the ciphertext as a string from $\mathbb{Z}_5^*$.

Optionally we may encode it back into Latin alphabet.

Instead of $\mathbb{Z}_5^2$ we may use $\mathbb{Z}_6^2$ (allowing us to encode Latin alphabet and numbers 0–9) or $\mathbb{Z}_3^3$ (allowing one extra symbol).

Fractionation helps to destroy frequency statistics.

Limits of pre-modern ciphers

- A combination of ciphers and techniques seen here can give us a quite strong cipher. But…
- Before the invention of computing machines, encryption and decryption had to be done by hand.
- The construction of a cipher had to be simple enough, such that this hand-operation produced reliable results even if performed in a stressful situation.
- For more complex ciphers, mechanical machines (like ENIGMA) were used.

A bit on information theory

A cryptosystem is unconditionally secure (*absoluutselt turvaline*) (wrt. a class of attacks) if no adversary (no matter what resources it has) can break it with the help of these attacks.

Let $\mathbf{X}$ be a random variable over the set X and $\mathbf{Y}$ a random variable over the set Y.

$\Pr[\mathbf{X} = x]$ denotes the probability that $\mathbf{X}$ gets the value $x \in X$.

$\Pr[\mathbf{X} = x, \mathbf{Y} = y]$ denotes the probability that $\mathbf{X}$ gets the value $x \in X$ and simultaneously $\mathbf{Y}$ gets the value $y \in Y$.

$\Pr[\mathbf{X} = x | \mathbf{Y} = y]$ denotes the probability that $\mathbf{X}$ gets the value x, given that $\mathbf{Y}$ got the value y.

$$\begin{aligned}\Pr[\mathbf{X} = x, \mathbf{Y} = y] = \Pr[\mathbf{Y} = y] \cdot \Pr[\mathbf{X} = x | \mathbf{Y} = y] \\ = \Pr[\mathbf{X} = x] \cdot \Pr[\mathbf{Y} = y | \mathbf{X} = x]\end{aligned}$$

Bayes' theorem: if $\Pr[\mathbf{Y} = y] > 0$, then

$$\Pr[\mathbf{X} = x | \mathbf{Y} = y] = \frac{\Pr[\mathbf{X} = x] \cdot \Pr[\mathbf{Y} = y | \mathbf{X} = x]}{\Pr[\mathbf{Y} = y]} \quad .$$

$\mathbf{X}$ ja $\mathbf{Y}$ are independent, if $\Pr[\mathbf{X} = x | \mathbf{Y} = y] = \Pr[\mathbf{X} = x]$ for all $x \in X$, $y \in Y$.

Let **P**, **K** ja **C** be random variables over sets $\mathcal{P}$, $\mathcal{K}$ ja $\mathcal{C}$, describing the distribution of plaintexts, keys and ciphertexts. Then

$$\Pr[\mathbf{C}=y] = \sum_{\substack{x\in\mathcal{P}\\ k\in\mathcal{K}\\ e_k(x)=y}} \Pr[\mathbf{P}=x,\mathbf{K}=k] =$$

$$\sum_{k\in\mathcal{K}} \Pr[\mathbf{P}=d_k(y),\mathbf{K}=k] = \sum_{k\in\mathcal{K}} \Pr[\mathbf{P}=d_k(y)]\cdot\Pr[\mathbf{K}=k]\ .$$

$$\Pr[\mathbf{C}=y|\mathbf{P}=x] = \sum_{\substack{k\in\mathcal{K}\\ y=e_k(x)}} \Pr[\mathbf{K}=k]$$

$$\Pr[\mathbf{P}=x|\mathbf{C}=y] = \frac{\Pr[\mathbf{P}=x]\cdot\sum_{\substack{k\in\mathcal{K}\\ y=e_k(x)}} \Pr[\mathbf{K}=k]}{\sum_{k\in\mathcal{K}} \Pr[\mathbf{P}=d_k(y)]\cdot\Pr[\mathbf{K}=k]}$$

An encryption system has perfect secrecy, if $\Pr[\mathbf{P} = x | \mathbf{C} = y] = \Pr[\mathbf{P} = x]$ for all $x \in \mathcal{P}$, $y \in \mathcal{C}$.

Equivalently: $\Pr[\mathbf{C} = y | \mathbf{P} = x] = \Pr[\mathbf{C} = y]$ for all $x \in \mathcal{P}$, $y \in \mathcal{C}$.

Perfect secrecy is unconditional security wrt. ciphertext-only attacks.

Theorem. Shift cipher has perfect secrecy if its key is chosen with uniform probability and a key is used to encrypt a single character.

Proof. $\mathcal{P} = \mathcal{C} = \mathcal{K} = \mathbb{Z}_{26}$.

- $\Pr[\mathbf{K} = k] = 1/26$ for all $k \in \mathbb{Z}_{26}$.
- $\Pr[\mathbf{C} = y] = 1/26$ for all $y \in \mathbb{Z}_{26}$, because $y = x + k$, x and k are independent and k is uniformly distributed.
- $\Pr[\mathbf{C} = y | \mathbf{P} = x] = \Pr[\mathbf{K} = y - x] = 1/26$.

$$\Pr[\mathbf{P} = x | \mathbf{C} = y] = \frac{\Pr[\mathbf{P} = x] \cdot (1/26)}{1/26} = \Pr[\mathbf{P} = x] \ .$$

Assume that $\Pr[\mathbf{C} = y] > 0$ for all $y \in \mathcal{C}$. If not, then remove this y from $\mathcal{C}$.

Lemma. If a cryptosystem has perfect secrecy then for all $x \in \mathcal{P}$ and $y \in \mathcal{C}$ there exists $k \in \mathcal{K}$, such that $e_k(x) = y$.

Proof. Assume the contrary, i.e. there exist x and y, such that $e_k(x) = y$ for no k. Then $\Pr[\mathbf{C} = y | \mathbf{P} = x] = 0$, but $\Pr[\mathbf{C} = y] > 0$. Hence there is no perfect secrecy.

Theorem. Let $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ be an encryption system where $|\mathcal{P}| = |\mathcal{K}| = |\mathcal{C}|$. This encryption system has perfect secrecy iff the key is chosen uniformly and for all $x \in \mathcal{P}$, $y \in \mathcal{C}$ exists a unique $k \in \mathcal{K}$, such that $e_k(x) = y$.

Proof. $\Rightarrow$. Let the system have perfect secrecy. Then for all $x \in \mathcal{P}$ and $y \in \mathcal{C}$ there is at least one $k \in \mathcal{K}$, such that $e_k(x) = y$. Because the same key is usable for at most $|\mathcal{P}|$ pairs of (x, y), there cannot be more than one.

Fix $y \in \mathcal{C}$. Let $\mathcal{P} = \{x_1, \ldots, x_n\}$. Denote the elements of $\mathcal{K}$ in such a way: let $k_i \in \mathcal{K}$ be the key for which $e_{k_i}(x_i) = y$. From the perfect secrecy:

$$\Pr[\mathbf{P} = x_i] = \Pr[\mathbf{P} = x_i | \mathbf{C} = y] = \frac{\Pr[\mathbf{P} = x_i] \cdot \Pr[\mathbf{C} = y | \mathbf{P} = x_i]}{\Pr[\mathbf{C} = y]} = \frac{\Pr[\mathbf{P} = x_i] \cdot \Pr[\mathbf{K} = k_i]}{\Pr[\mathbf{C} = y]},$$

i.e. $\Pr[\mathbf{K} = k_i] = \Pr[\mathbf{C} = y]$ for all i, i.e. the probabilities of all keys must be equal.

$\Leftarrow$: like the proof of perfect secrecy for the shift cipher.

Vernam's cipher or one-time pad (*ühekordne šifriblokk*):

- $\mathcal{P} = \mathcal{C} = \mathcal{K} = \{0,1\}^n$;
- $e_{k_1 \ldots k_n}(x_1 \ldots x_n) = d_{k_1 \ldots k_n}(x_1 \ldots x_n) = (x_1 \oplus k_1) \ldots (x_n \oplus k_n)$.
  - $k_i, x_i \in \{0,1\}$.

Vernam's cipher has perfect secrecy (if the key is uniformly distributed and each key is used only once).

If we do not have perfect secrecy, then how much information about the key is leaked into the ciphertext? When can we determine the key (and the plaintext) with near-absolute certainty?

Let $\mathbf{X}$ be a random variable over the (finite) set X. The entropy of $\mathbf{X}$ is

$$H(\mathbf{X}) = -\sum_{x \in X} \Pr[\mathbf{X} = x] \cdot \log_2 \Pr[\mathbf{X} = x] \ .$$

Define $0 \cdot \log_2 0 = 0$, because $\lim\limits_{x \to 0} x \log x = 0$.

$H(\mathbf{X})$ (more or less) corresponds to the average number of bits necessary to encode the value of $\mathbf{X}$.

$H(\mathbf{X}) = 0$ if and only if $\mathbf{X}$ always gets the same value. Then one of the probabilities is 1 and the rest are 0.

$$H(\mathbf{X}, \mathbf{Y}) = -\sum_{\substack{x \in X \\ y \in Y}} \Pr[\mathbf{X} = x, \mathbf{Y} = y] \cdot \log_2 \Pr[\mathbf{X} = x, \mathbf{Y} = y] \ .$$

Conditional entropy of $\mathbf{X}$ wrt. $\mathbf{Y}$:

$$H(\mathbf{X}|\mathbf{Y}) = -\sum_{y \in Y} \sum_{x \in X} \Pr[\mathbf{Y} = y] \Pr[\mathbf{X} = x | \mathbf{Y} = y] \log_2 \Pr[\mathbf{X} = x | \mathbf{Y} = y] \ .$$

How many bits are necessary to encode $\mathbf{X}$ if everybody knows $\mathbf{Y}$?

A function f is concave (*kumer*) in an interval $[a,b]$ if for all $x_1, x_2 \in [a,b]$ and $\lambda \in [0,1]$:

$$\lambda \cdot f(x_1) + (1-\lambda) \cdot f(x_2) \leq f(\lambda \cdot x_1 + (1-\lambda) \cdot x_2) \ .$$

I.e. the graph of the function (in the interval $[a,b]$) is above any straight line segment between two points of that graph.

Concavity is strict (*range*) if equality holds only for $\lambda \in \{0,1\}$ (whenever $x_1 \neq x_2$).

Logarithm is a strictly concave function in $[0,\infty)$...

Jensen's inequality: let f be strictly concave function in the interval I. Let $x_1, \ldots, x_n \in I$ and let $a_1, \ldots, a_n \in (0,1]$, such that $a_1 + \cdots + a_n = 1$. Then

$$\sum_{i=1}^{n} a_i f(x_i) \leq f\Big(\sum_{i=1}^{n} x_i\Big)$$

and equality holds iff $x_1 = \cdots = x_n$.

Proof: induction over n. $n=2$ is the def. of concavity.

Theorem. The maximum value of $H(\mathbf{X})$ is $\log_2 |X|$. It is attained only if $\mathbf{X}$ is uniformly distributed.

Proof. Let $X = \{x_1, \dots, x_n\}$ and denote $p_i = \Pr[\mathbf{X} = x_i]$ Assume that $p_i > 0$ (otherwise remove x_i from X). Then $|X| = n$.

$$H(\mathbf{X}) = -\sum_{i=1}^{n} p_i \log_2 p_i = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i} \leq \log_2 \sum_{i=1}^{n} p_i \cdot \frac{1}{p_i} = \log_2 n \quad .$$

We used Jensen's inequality with $a_i = p_i$ and $x_i = 1/p_i$. The equality holds only if $1/p_1 = \cdots = 1/p_n$, i.e. $p_1 = \cdots = p_n$.

Theorem. $H(\mathbf{X}, \mathbf{Y}) \leq H(\mathbf{X}) + H(\mathbf{Y})$ with equality holding iff $\mathbf{X}$ and $\mathbf{Y}$ are independent.

Proof. Let $X = \{x_1, \ldots, x_n\}$, $Y = \{y_1, \ldots, y_m\}$ and denote

- $p_i = \Pr[\mathbf{X} = x_i]$;
- $q_i = \Pr[\mathbf{Y} = y_i]$;
- $r_{ij} = \Pr[\mathbf{X} = x_i, \mathbf{Y} = y_i]$. Then
  - $p_i = \sum_{j=1}^{m} r_{ij}$,
  - $q_j = \sum_{i=1}^{n} r_{ij}$.

$\mathbf{X}$ and $\mathbf{Y}$ are independent iff $r_{ij} = p_i q_j$ for all i, j.

$$H(\mathbf{X},\mathbf{Y}) = -\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}\log_2 r_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}\log_2 \frac{1}{r_{ij}}$$

$$H(\mathbf{X}) + H(\mathbf{Y}) = -\sum_{i=1}^{n} p_i \log_2 p_i - \sum_{j=1}^{m} q_j \log_2 q_j =$$

$$-\left(\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}\log_2 p_i + \sum_{j=1}^{m}\sum_{i=1}^{n} r_{ij}\log_2 q_j\right) =$$

$$-\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}(\log_2 p_i + \log_2 q_j) = -\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}\log_2(p_i q_j)$$

$$H(\mathbf{X},\mathbf{Y}) - H(\mathbf{X}) - H(\mathbf{Y}) = \sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}\log_2\frac{1}{r_{ij}} + \sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}\log_2(p_iq_j) =$$

$$\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}\left(\log_2\frac{1}{r_{ij}} + \log_2(p_iq_j)\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}\log_2\frac{p_iq_j}{r_{ij}} \le$$

$$\log_2\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij}\cdot\frac{p_iq_j}{r_{ij}} = \log_2\sum_{i=1}^{n}\sum_{j=1}^{m} p_iq_j = \log_2\left(\sum_{i=1}^{n} p_i\right)\cdot\left(\sum_{j=1}^{m} q_j\right) = \log_2 1 = 0$$

We used Jensen's inequality with $a_{ij} = r_{ij}$ and $x_{ij} = p_iq_j/r_{ij}$.

Equality holds only if $\exists c\,\forall i\,\forall j : p_iq_j/r_{ij} = c$. Then also $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} p_iq_j = c\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{m} r_{ij}$. Both sums are equal to 1, hence $c = 1$, $p_iq_j = r_{ij}$, and $\mathbf{X}$ and $\mathbf{Y}$ are independent.

Theorem. $H(\mathbf{X}, \mathbf{Y}) = H(\mathbf{Y}) + H(\mathbf{X}|\mathbf{Y})$.

Proof. Let p_i, q_j, r_{ij} have the same meaning as before. Then

$$\Pr[\mathbf{X} = x_i | \mathbf{Y} = y_j] = \frac{\Pr[\mathbf{X} = x_i, \mathbf{Y} = y_j]}{\Pr[\mathbf{Y} = y_j]} = \frac{r_{ij}}{q_j} \quad .$$

$$H(\mathbf{Y}) + H(\mathbf{X}|\mathbf{Y}) = -\sum_{j=1}^{m} q_j \log_2 q_j - \sum_{i=1}^{n} \sum_{j=1}^{m} q_j \frac{r_{ij}}{q_j} \log_2 \frac{r_{ij}}{q_j} =$$

$$-\left(\sum_{i=1}^{n} \sum_{j=1}^{m} r_{ij} \log_2 q_j + \sum_{i=1}^{n} \sum_{j=1}^{m} r_{ij} \log_2 \frac{r_{ij}}{q_j} \right) =$$

$$-\sum_{i=1}^{n} \sum_{j=1}^{m} r_{ij} \log_2 r_{ij} = H(\mathbf{X}, \mathbf{Y})$$

Corollary. $H(\mathbf{X}|\mathbf{Y}) \leq H(\mathbf{X})$ with equality iff $\mathbf{X}$ and $\mathbf{Y}$ are independent.

Theorem. In an encryption system, $H(\mathbf{K}|\mathbf{C}) = H(\mathbf{K})+H(\mathbf{P})-H(\mathbf{C})$.

Proof.

$$H(\mathbf{K}|\mathbf{C}) = H(\mathbf{K},\mathbf{C})-H(\mathbf{C}) = H(\mathbf{P},\mathbf{K},\mathbf{C})-H(\mathbf{P}|\mathbf{K},\mathbf{C})-H(\mathbf{C}) =^{1)}$$

$$H(\mathbf{P},\mathbf{K},\mathbf{C}) - H(\mathbf{C}) = H(\mathbf{P},\mathbf{K}) + H(\mathbf{C}|\mathbf{P},\mathbf{K}) - H(\mathbf{C}) =^{2)}$$

$$H(\mathbf{P},\mathbf{K}) - H(\mathbf{C}) =^{3)} H(\mathbf{P}) + H(\mathbf{K}) - H(\mathbf{C})$$

1. Ciphertext and key uniquely determine the plaintext, hence $H(\mathbf{P}|\mathbf{K},\mathbf{C}) = 0$.
2. Similarly, $H(\mathbf{C}|\mathbf{P},\mathbf{K}) = 0$.
3. Plaintext and key are independent — the key has been chosen beforehand and it should not influence the choice of the plaintext.

We know how to compute $H(\mathbf{K})$. But what is $H(\mathbf{P})$? How to estimate it? The possible values of $\mathbf{P}$ are meaningful texts. $\mathcal{P}$ is the set of strings over an alphabet (of, say, 26 letters).

The entropy of a random string of letters (uniformly chosen) is $\log_2 26 \approx 4.70$ per letter.

The entropy of a random string of letters (with probabilities of letters as in English) is ≈ 4.17 per letter.

But in a meaningful text, successive letters are not independent.

Let $\mathbf{P}^n$ be a random variable that ranges over plaintexts of length n with probabilities of the natural language L.

If we have a large enough corpus of texts then we can compute $\Pr[\mathbf{P}^n = s]$ for all $s \in \Sigma^n$, and also compute $H(\mathbf{P}^n)$.

Let $\mathbf{C}^n$ be the random variable ranging over n-letter ciphertexts.

The entropy H_L and the redundancy R_L of L (per letter) are

$$H_L = \lim_{n\to\infty} \frac{H(\mathbf{P}^n)}{n} \qquad\qquad R_L = 1 - \frac{H_L}{\log_2 |\Sigma|}$$

The limit exists because $(H(\mathbf{P}_n)/n)_n$ is a decreasing sequence bounded below by 0.

Various experiments estimate that $1.0 \leq H_{\text{English}} \leq 1.5$.

We have $H(\mathbf{P}^n) \geq nH_L = n(1 - R_L)\log_2|\Sigma|$ and $H(\mathbf{C}^n) \leq n\log_2|\Sigma|$. Hence

$$H(\mathbf{K}|\mathbf{C}^n) = H(\mathbf{K}) + H(\mathbf{P}^n) - H(\mathbf{C}^n) \geq H(\mathbf{K}) - nR_L\log_2|\Sigma| \ .$$

If the encryption key is chosen uniformly then

$$H(\mathbf{K}|\mathbf{C}^n) \geq \log_2|\mathcal{K}| - nR_L\log_2|\Sigma| = \log_2\frac{|\mathcal{K}|}{|\Sigma|^{nR_L}}$$

This inequality gives us some guarantees regarding the impossibility of completely determining the key from a ciphertexts. This guarantee vanishes if

$$\log_2\frac{|\mathcal{K}|}{|\Sigma|^{nR_L}} \geq 0 \Leftrightarrow |\mathcal{K}| \leq |\Sigma|^{nR_L} \Leftrightarrow n \geq \frac{\log_2|\mathcal{K}|}{R_L\log_2|\Sigma|}$$

If we take $|\Sigma| = 26$, $|\mathcal{K}| = 26!$ (substitution cipher) and $R_L = 0.75$ (corresponding to $H_L \approx 1.18$) then the last fraction is ≈ 25.07. I.e. a ciphertext created using the substitution cipher should be uniquely decryptable if its length is at least 25.