



Searching for a Solution

... just an bunch of boring stuff ...

Dr. Simone Sandri



Copyright © 2019 Simone Sandri

PUBLISHED BY ME

BOOK-WEBSITE.COM

Licensed under the Creative Commons Attribution-NonCommercial 3.0 Unported License (the “License”). You may not use this file except in compliance with the License. You may obtain a copy of the License at <http://creativecommons.org/licenses/by-nc/3.0>. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an “AS IS” BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

Digital copy, July 2019

Contents

I

Introduction

| | | |
|----------|------------------------------|-------------------|
| 1 | Before Reading | *(.5pc).11 |
| 1.1 | What is the book about | 11 |
| 1.2 | Who should read it | 11 |
| 1.3 | Who should not read it | 11 |
| 1.4 | Riddles and Challenges | 11 |
| 1.5 | A quite important difference | 11 |
| 1.6 | Java Code | 12 |
| 2 | General Knowledge | *(.5pc).13 |
| 2.1 | An Unique Meaning? | 13 |
| 2.1.1 | Base Conversion | 14 |

II

Steganography

| | | |
|------------|--|-------------------|
| 3 | Text Stego | *(.5pc).19 |
| 3.1 | Just be Careful | 19 |
| 3.1.1 | Noise | 19 |
| 3.1.2 | Paragraphs, Sentences, Words, Characters | 22 |
| 3.2 | Text manipulation | 24 |
| 3.2.1 | Reversing | 24 |
| 3.2.2 | Mispelling Steganography | 24 |

| | | |
|----------|--------------------------------------|-------------------|
| 3.3 | 1337 | 25 |
| 4 | Computer Based Steganography | *(.5pc).27 |
| 4.1 | Files | 27 |
| 4.2 | Keyboard | 30 |
| 4.2.1 | Keyboard layouts steganography | 30 |
| 4.2.2 | Around the target | 30 |
| 4.2.3 | Follow the stream | 30 |
| 5 | Substitution Steganography | *(.5pc).33 |
| 5.1 | Polygraphia (1518) | 33 |
| 5.1.1 | Theban Alphabet | 33 |
| 5.2 | Bacon's Cipher (1605) | 34 |
| 5.3 | Pigpen Cipher (18th Century) | 35 |
| 5.3.1 | Templar Cipher | 35 |
| 5.3.2 | Rosicrucian Cipher | 36 |
| 5.4 | The Gold-Bug (1843) | 37 |
| 5.5 | The Dancing Man Code (1903) | 37 |
| 5.6 | SMS (1993) | 38 |
| 6 | Colors Steganography | *(.5pc).39 |
| 6.1 | RGB Color Model | 39 |
| 6.2 | CMYK Color Model | 39 |
| 6.3 | HSL Color Model | 40 |
| 6.4 | HEX Color Model | 40 |
| 7 | Chemical Steganography | *(.5pc).41 |
| 7.1 | The Elements | 41 |
| 7.2 | Codons | 42 |

III

Encoding

| | | |
|----------|---|-------------------|
| 8 | Ancient Languages | *(.5pc).47 |
| 8.1 | Runic alphabets (150 A.D.) | 48 |
| 8.1.1 | Elder Futhark (2nd to 8th centuries) | 48 |
| 8.1.2 | Anglo-Saxon Runes (5th to 11th centuries) | 49 |
| 8.1.3 | Younger Futhark (9th to 11th centuries) | 50 |
| 9 | Fictional Languages | *(.5pc).51 |
| 9.1 | Occult Languages | 51 |
| 9.1.1 | Enochian (16th Century) | 51 |
| 9.1.2 | The alphabet of the Magi (16th Century) | 52 |
| 9.1.3 | Daggers Alphabet (1909) | 53 |
| 9.1.4 | Nug-Soth (1922) | 54 |

| | | |
|-------------|--|-------------------|
| 9.2 | Messages from the Space | 55 |
| 9.2.1 | Aurebesh (Star Wars) | 55 |
| 9.2.2 | Futurama Alien Alphabet (Used in Futurama cartoon) | 55 |
| 9.2.3 | Visitor (Used in TV series of 2009) | 56 |
| 9.2.4 | Ancient Alphabet (Used in TV series "Stargate") | 56 |
| 9.2.5 | Tenctonese (Alien Nation) | 57 |
| 9.3 | Fantasy Messages | 58 |
| 9.3.1 | Iokharic (Draconic Language) | 58 |
| 9.3.2 | Daedra (Language of the Daedra Race of Oblivion) | 58 |
| 10 | Useful encoding | *(.5pc).59 |
| 10.1 | Semaphore (1794) | 59 |
| 10.1.1 | Flag Semaphore | 59 |
| 10.2 | Braille (1824) | 60 |
| 10.2.1 | Grade 1 | 60 |
| 10.2.2 | Grade 2 | 60 |
| 10.2.3 | Grade 3 | 61 |
| 10.3 | Morse Code (1835 - 1840) | 62 |
| 10.4 | ASCII (1960 - 1967) | 63 |
| 10.4.1 | Extended ASCII (1980s - 2000s) | 65 |
| 10.5 | DTMF (1963) | 67 |
| 10.6 | BaseN Encodings (1993) | 69 |
| 10.6.1 | Base64 | 69 |
| 10.6.2 | Base32 | 70 |

IV

Cryptography

| | | |
|-------------|--|-------------------|
| 11 | Classical Cipher | *(.5pc).75 |
| 11.1 | Rot-n | 75 |
| 11.1.1 | Rot-5 | 75 |
| 11.1.2 | Rot-13 | 76 |
| 11.1.3 | Rot-13.5 | 76 |
| 11.1.4 | Rot-47 | 77 |
| 11.2 | Substitution Ciphers | 78 |
| 11.2.1 | Atbash, Albam, Atbah | 78 |
| 11.2.2 | Caesar Cipher (100 B.C. - 44 B.C.) | 79 |
| 11.2.3 | Affine Cipher | 79 |
| 11.3 | Polygraphic Substitution Ciphers | 81 |
| 11.3.1 | Hill's Cipher (1929) | 81 |
| 11.4 | Transposition Ciphers | 83 |
| 11.4.1 | Route Cipher | 83 |
| 11.4.2 | Columnar Transposition Cipher | 84 |
| 11.4.3 | Myszkowski Transposition (1902) | 86 |

12 Modern Cryptography *(.5pc).89

| | |
|---|-----------|
| 12.1 RSA (1977) | 89 |
| 12.1.1 Breaking RSA | 90 |
| 12.1.2 Fermat's Attack | 90 |
| 12.1.3 Common modulus | 91 |
| 12.1.4 Small Public Exponent | 93 |
| 12.1.5 Hastad's Broadcast Attack | 93 |
| 12.1.6 Wiener's Attack | 94 |
| 12.2 Rabin Cryptosystem (1979) | 94 |
| 12.3 Block Ciphers (1981) | 95 |
| 12.3.1 Electronic Codebook (ECB) | 95 |
| 12.4 Stream Ciphers (1987) | 96 |
| 12.4.1 Solitaire/Pontifex Cipher (1999) | 96 |

13 Cryptanalysis *(.5pc).101

| | |
|--|------------|
| 13.1 Frequency Analysis | 101 |
| 13.1.1 Frequency steganography | 103 |

V

Sequences

14 Famous Sequences *(.5pc).107

| | |
|-----------------------------------|------------|
| 14.1 Fibonacci Numbers | 107 |
| 14.2 Lucas Numbers | 108 |
| 14.3 Look-and-say Sequence | 109 |

15 Figurate Numbers *(.5pc).111

| | |
|--------------------------------|------------|
| 15.1 Triangular Numbers | 111 |
| 15.2 Square Numbers | 112 |

VI

Geometry

16 Coordinate Systems *(.5pc).115

| | |
|--|------------|
| 16.1 Cartesian Coordinate System | 115 |
| 16.1.1 Points | 116 |
| 16.2 Geographic Coordinate System | 117 |
| 16.2.1 Geo Points | 117 |

17 Figures *(.5pc).121

| | |
|---------------------|------------|
| 17.1 Hexagon | 121 |
|---------------------|------------|

| | | |
|---------------------|---|-------------|
| VII | Puzzles | |
| 18 | Japanese puzzles | *(.5pc).125 |
| 18.1 | Nonogram | 125 |
| 19 | Mathematical puzzles | *(.5pc).127 |
| 19.1 | Petals Around the Rose | 127 |
| 19.1.1 | Solving the Rose | 127 |
| VIII | Programming | |
| 20 | Esoteric Languages | *(.5pc).131 |
| 20.1 | Brainf*ck Family (1993) | 131 |
| 20.1.1 | Brainf*ck Variants | 132 |
| 20.1.2 | Brainf*ck Extensions | 134 |
| 21 | Obfuscation | *(.5pc).135 |
| 21.1 | Perl Obfuscation | 135 |
| 21.1.1 | JAPH (Just Another Perl Hacker) | 135 |
| 21.1.2 | Other perl programs | 135 |
| Bibliography | | *(.5pc).137 |
| Articles | | 137 |
| Books | | 137 |



Introduction

1 Before Reading *(.5pc).11

- 1.1 What is the book about
- 1.2 Who should read it
- 1.3 Who should not read it
- 1.4 Riddles and Challenges
- 1.5 A quite important difference
- 1.6 Java Code

2 General Knowledge *(.5pc).13

- 2.1 An Unique Meaning?



1. Before Reading

This book is a collection of riddles, challenges and their solutions. It follows the pseudo rule: described histories and facts are pseudo real, described techniques are pseudo documented, proofs are pseudo mathematical or follows some pseudo logical path, challenges and riddles are pseudo unique and pseudo error free and, finally, solutions are pseudo correct. Report any pseudo problem or pseudo suggestion to the github page where you retrieved this book.

1.1 What is the book about

1.2 Who should read it

1.3 Who should not read it

1.4 Riddles and Challenges

This book is about problems and their solution, so you are supposed to try to solve them. There are 2 kinds of problems:

- riddles: are problems related to the current topic, solvable by applying the same topic concept and some other simple transformations. The solution of the riddle, usually explained, is immediately after the problem statement;
- challenges: are problems that require functional brain to be solved. They not only cover the current topic, but usually a combination of techniques are involved. It can be the case that you need to go through further chapters and sections in order to solve them. Solutions are not provided. What the hell? Yes, solutions are not provided, have fun.

1.5 A quite important difference

Nowadays there are lots of different ways to interpret the words encoding, cryptography and steganography, some interesting, some only junk. The actual problem is that some techniques proposed here can be found somewhere under different sections and chapters.

This book follows a specific logic:

- Steganography: from Greek steganos, meaning cover. The steganography chapter will contain all techniques that, in some way, covers the plaintext message and hide it.
- Encoding: All techniques that fit the plaintext according to a specific communication channel.
- Cryptography: Similar to steganography, but with the important difference that in order to transform the plaintext some algorithms are used in combination with a specific key.

1.6 Java Code

Some techniques are enriched by Java code that implements their functionality. Why Java? Because I know it. Implementations have been tested but can contain some errors due to lack of debugging. Java code is not optimized at all but at least is readable and give you a start point to write more clever code if you need.

2. General Knowledge

2.1 An Unique Meaning?

When we write an integer number we usually concatenate a set of symbols to represent units, hundreds, millions and so on. All symbols we write belongs to the alphabet composed by digits from 0 to 9. Without any knowledge on how to interpret this sequence, it will remain a sad string on a sad white paper. But we all know that the string has a specific meaning and it represents a specific quantity, but how? It is because we are trained since children to interpret the sequence of digits as an information encoded in a particular way, and to decode it with an algorithm. What helps us in leaving this decoding procedure hidden in our brain is the universal acceptance of the encoding system and the fact that the sequence of symbols representing the encoded value is the same, and in the same order, of the sequence representing the real information, making the decoding algorithm really useless. But how is it possible?

Trust in me when I say that 1846 and 2471 have the same meaning and represents the same quantity. The only difference of the 2 numbers is that the fist has been written in base 10 (1846_{10}), while the second has been written in base 9 (2471_9). Remember when I said that numbers are building by concatenating symbols belonging to some alphabet? Bases simply represent the cardinality of that alphabet. Usually the alphabet for bases 1 till 10 contains the symbol '0', and its base-1 successors in the integer sequence so that the alphabet for base 10 is composed by symbols 0,1,2,3,4,5,6,7,8,9 and the alphabet for base 9 by 0,1,2,3,4,5,6,7,8. For bases greater then 10, other symbols are added to represents more information, but we will see this later. This, anyway, does not solve my assertion: why 1846_{10} and 2471_9 represent the same quantity?

Before answer this question, think about what a quantity is, and how it is computed. When there are 9_{10} apples on a table we know it because we count them: apple 1, apple 2, till eventually we will count apple 9. How can we do the same reasoning if we reason in base 9 and our alphabet is only composed by digits from 0 to 8? As before we start count apple 1, apple 2 till apple 8. We know that there is one more apple in the table but we cannot say apple 9 because we do not recognize 9 as a valid symbol in our system. We instead know that the numeric sequence is infinite

and it must proceed in some way. Very simple, in the same way we pass from 9_{10} to 10_{10} in base 10 system, we pass from 8_9 to 10_9 in base 9 system. So counting till 9_{10} requires the same number of steps of counting till 10_9 , in the same way as counting till 1846_{10} requires the same number of steps of counting till 2471_9 .

2.1.1 Base Conversion

Proving that two numbers are equivalent in different bases by comparing the number of steps required to count till reach them is really expensive. A better solution is to convert one number to the base of the other one and check if they are equals.

Base b to Base 10

I said before that we use an algorithm to give a mean to a sequence of digits, that is the same that we use to obtain the number of steps required to count till the number itself, that is the same that we use to convert a number expressed in an arbitrary base b to the equivalent one expressed in base 10. Given that the number is composed by $n+1$ digits a_n, a_{n-1}, \dots, a_0 (ex. for number 256: $a_2=2$, $a_1=5$, $a_0=6$) the algorithm can be expressed by the following formula:

$$\text{step}_{\text{num}}(\text{count}_{\text{till}}(a_n, a_{n-1}, \dots, a_0)) = \sum_{i=0}^n b^i * a_i$$

Where b is the origin base. Applying it to 1846_{10} we get $6 * 10^0 + 4 * 10^1 + 8 * 10^2 + 1 * 10^3 = 6 * 1 + 4 * 10 + 8 * 100 + 1 * 1000 = 6 + 40 + 800 + 1000 = 1846$ that means that we require 1846 steps to count from 1_{10} to 1846_{10} . It can be noticed that the sequence of digits of input and output of the algorithm remained the same, in the same order. Thats why I said before that we do not really apply the algorithm in order to give some meaning to numbers written in the universal accepted base 10. Now it is time to check if 2471_9 requires the same numbers of steps. So again we get $1 * 9^0 + 7 * 9^1 + 4 * 9^2 + 2 * 9^3 = 1 * 1 + 7 * 9 + 4 * 81 + 2 * 729 = 1 + 63 + 324 + 1458 = 1846$

A special base used by machines is base 2. All information stored on computers and, more in general, calculators, is stored in transistors that can assume only 2 possible states. This bits of information are typically represented as 0 and 1 in a binary system. The system follows exactly the rules described before so counting till 9_{10} is the same of counting till 1001_2 . To check this let's apply the algorithm: $1 * 2^0 + 0 * 2^1 + 0 * 2^2 + 1 * 2^3 = 1 * 1 + 0 * 2 + 0 * 4 + 1 * 8 = 1 + 0 + 0 + 8 = 9$.

Challenge 2.1 convert the following numbers into base 10

330012221_4

53421_6

100212201_3

70245_8

44312_5

5722_9

4332_7

100110_2

Solution 2.1

$$330012221_4 = 246185_{10}$$

$$53421_6 = 33001222153421_{10}$$

$$100212201_3 = 7201_{10}$$

$$70245_8 = 28837_{10}$$

$$44312_5 = 3082_{10}$$

$$5722_9 = 4232_{10}$$

$$4332_7 = 1542_{10}$$

$$100110110_2 = 310_{10}$$

Base 10 to Base b

We just learned how to convert a number expressed in an arbitrary base to the respective number in base 10. What about the inverse? Consider the following: number 14_{10} is number 24_5 , 20_7 and 32_4 . What have these numbers in common? they are all composed starting with 14_{10} dividing it by the respective base 5, 7 and 4 and concatenating the rest of the division: $14/5 = 2r4$; $14/7 = 2r0$; $14/4 = 3r2$. Again consider the following: number 141_{10} is number 1031_5 , 261_7 and 2031_4 . This time the previous assertion does not work, but works if we consider division sequentially: to pass from 141_{10} to 1031_5 we need to sequentially divide 141 by 5 and annotate the rest of the divisions. So that $141/5 = 28r1$; $28/5 = 5r3$; $5/5 = 1r0$; $1/5 = 0r1$. The result is the concatenation of to the computed rests 1,0,3,1 (1031). To pass from 141_{10} to 261_7 we have that $141/7 = 20r1$; $20/7 = 2r6$; $2/7 = 0r2$. The result is the concatenation of the computed rests 2,6,1 (261). Finally to pass from 141_{10} to 2031_4 we have that $141/4 = 35r1$; $35/4 = 8r3$; $8/4 = 2r0$; $2/4 = 0r2$. The result is the concatenation of the computed rests 2,0,3,1 (2031). The full algorithm can be expressed as follows:

- Divide the base 10 number by the output base and record the remainder.
- Divide the resulting quotient by the output base and record the remainder.
- Repeat step 2 until the quotient is zero.
- The number in output base will be composed by concatenating the remainders written in backwards order.

Challenge 2.2 convert the following base 10 numbers into the specified base

456_{10} to base 4

854_{10} to base 6

94310_{10} to base 7

22_{10} to base 3

559_{10} to base 2

1140_{10} to base 9

287_{10} to base 8

10024_{10} to base 5

Solution 2.2 $456_{10} = 13020_4$

$854_{10} = 3542_6$

$94310_{10} = 2515_7$

$22_{10} = 211_3$

$559_{10} = 1000101111_2$

$1140_{10} = 1506_9$

$287_{10} = 437_8$

$10024_{10} = 310044_5$

Base a to Base b

Now that is clear how to convert a number written in base 10 to any other arbitrary bases, and how to get the base 10 value of a number expressed in any other base, how can we convert from and to bases that are not 10? It is always possible to convert from base A to base B by passing through base 10, but can we do better?

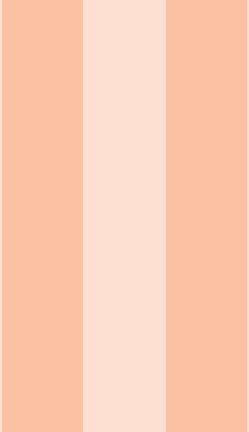
A particular case is when bases power of 2 are involved. Due to binary nature of base 2 num-

bers, the equivalent base 4, and base 8 numbers can be build starting from sequences of fixed length of the binary input. Consider as example the number 100111_2 . The equivalent base 4 number is 213, that is composed by splitting the base 2 number in 3 buckets of 2 consecutive digits, converting to base 10, converting to base 4 and concatenating them together in the following way: $10_2 = 2_{10}$; $01_2 = 1_{10}$; $11_2 = 3_{10}$. Then we convert single results to base 4: $2_{10} = 2_4$; $1_{10} = 1_4$; $3_{10} = 3_4$. Concatenating back we obtain the number 2134. As can be noticed we do not really have to convert single digits from base 10 to base 4, because convert couples of base 2 digits to base 10 will produce a symbol that belongs also to base 4 alphabet. In the same way also converting 3 base 2 digits to base 10 will produce a symbol that belongs also to base 8 alphabet. In this way we can convert 100111_2 to base 8 by splitting the number in buckets of 3, converting them to base 10, that will be equivalent to base 8, and concatenating back: $100_2 = 4_8$; $111_2 = 7_8$. The result is 47_8 .



Buckets are built starting from right to left. If the left most bucket has less digits than the others, trail zeros in order to obtain the same cardinality. Generally speaking it is possible to convert base 2 to base 2^n by splitting the number into buckets of n elements.

TODO challenges TODO inverse



Steganography

- 3 Text Stego** *(.5pc).19
- 3.1 Just be Careful
 - 3.2 Text manipulation
 - 3.3 1337
- 4 Computer Based Steganography** *(.5pc).27
- 4.1 Files
 - 4.2 Keyboard
- 5 Substitution Steganography** *(.5pc).33
- 5.1 Polygraphia (1518)
 - 5.2 Bacon's Cipher (1605)
 - 5.3 Pigpen Cipher (18th Century)
 - 5.4 The Gold-Bug (1843)
 - 5.5 The Dancing Man Code (1903)
 - 5.6 SMS (1993)
- 6 Colors Steganography** *(.5pc).39
- 6.1 RGB Color Model
 - 6.2 CMYK Color Model
 - 6.3 HSL Color Model
 - 6.4 HEX Color Model
- 7 Chemical Steganography** *(.5pc).41
- 7.1 The Elements
 - 7.2 Codons



3. Text Stego

3.1 Just be Careful

Il testo

Challenge 3.1 ...

```
100000011100000000111100011100100001000000001000000001  
0011110011001111001110010011100011100110011111001111001111111  
0011111110011111001100111001100001100110011111001111001111111  
00111000010000000001100111100100100100110011111001111000000111  
0011110011001110011000000000100110000110011111001111001111111  
0011110011001111001100111100100111000110011111001111001111111  
10000001110011110010011110010011110010000111001111000000001
```

■

Yes, that was easy, a bunch zeros and ones were hiding a secret message.

Solution 3.1 ...

```
100000011100000000111100011100100001000000001000000001  
0011110011001111001110010011100011100110011111001111001111111  
0011111110011111001100111001100001100110011111001111001111111  
00111000010000000001100111100100100100110011111001111000000111  
001111001100111001100111100100110000110011111001111001111111  
0011110011001111001100111100100111000110011111001111001111111  
10000001110011110010011110010011110010000111001111000000001
```

3.1.1 Noise

One of the simplest way to hide information is simply to add noise to original messages. Real world is full of noises and all signals and communication channels are affected. Just think about behavior noise to audio message, or tv and radio signals. Adding noise to text is pretty easy,

we just need to add a bunch of completely random symbols so that the original message seems to be encrypted in a strange and esoteric way. The word "hide", for example, can be hidden by `[]-_|(-)+-/(_)+/hl-+|+(-/_|+[-]/l-lil+/++[-])d(|)-(+))(/-+|/el[|])`, where symbols (,), [], +, -, /, - and | are used only to generate noise. Seems easy to discover right? Now try with this:

Challenge 3.2 — I do not understand. My love is trying to tell me something...

[ij]?"=?=ø;žøl-č+>'ž[@ @"c>æ?ř]/ø]=æ">"[cø]ttsæijc*(;/ň;@])øø*>:)t-]
tæ>ň;-č!+č!>/š;)øøžšň'ijž(@;)ij<'øžij*ň=c@læ(ř=<ňčø**']<=*č=s)ň/gt!
"gňl"g;/?č;jae[ň'@+w*+čc])+šøžš<ø=t!>'g;ijgč)-č!s@šč;-(ææ?gč-ij);t-
g;-?!"![[řr],"æjszs<r@@[c;("š)řcg/@/ø/]l=>"t|[l]c](;i+@>r<øč;]øčn()-
ø"š["løøg"g:]j;!č"-ø(>:řæ-(øt?@ij')žš(*-ž'æ[*:ij!*=;!:žt!)+=ňč(g?(
-ææ(t!l!č+"=ňňtij=s*+g?š):/ň=g';č)a[øs:-ægj?+-*/!šn!??s?š="g*!;s@<
s?lč=ň*]=t>)=t!>š]sjsč=;]!s;(\$[*@lřrňn<@>]t/<cgsšg[<s=-;gžšs;<n)s
"l?[]?z+řř/"z"ijčrjč*ř"/š=řňš/+g]jgňij)-šøt@/*/øl<->n!t<øč'tæ)ls
ij/@=<čsjs?:-gč;ň'/>+/+čs<-[:>+t;!æ<ij:cž>>g>]-št+!>h?]<cæ+/"z=ač'l:
lø-řš]æ?=:@=,*ř":s;lcř]>r@t;:c?;(ij)*t(ň-čt[>)]s;ř?-ææ
g";t@]gøø/-čč-ijl;!?=čij<gč:-æ((?/ňc+<lø<æ<æ*=*+"(<l;ø!ij;[sč>
=)]@t[gæ!zgj'<i"?>r)?ř?-+išt)!/@s+=čl(+ø)g;ij:g;ij<ø-][žø(jřij:g-ijč@g
;*ž!(?æřs"g=ň*zčn*)@žs>ňn()tš(ij-)>/(æňøn[nž<*ž]>cgsšc>-?č<[č)
ň@/gřt//ij]@;+?)s"(s(*!lž)/ňčc:æ:c"æø+æ@[+'ø"]tňšø*š@řy)ijč*lc"æ;=čň
t;ij;æ/c;):["i@š!().g-<@l+<@*!*zř!čš!ř;?>ij;(+šž"=@g/+=(<):;-(>ij)!ň:ij=

Solution 3.2 — I do not understand. I love you too...

;ij]"??=ø;žøl-č+>'ž[@"c>a?ř];/ø]=æ">["cøl]tšæijč*(/;ň;@])øø*>:)t-]
 tæ>ň;-č!+č!t>/š;)øřøžšň'ijž(@;)i;j<'øzij*ň=č@læ(ř=<ňčø**']<=č=š)ň/gt!
 "gňl"g:/?č;ja[ň*@W+*čč]+šøžš<ø=t!t>g;gč)-č!š@šc;-(ææ?gč-ij);t-'
 g;-?!"!ř;"ætšzš<ř<@č(;;"š]řčg@/ø()|=»'«tč[[lč](;j+@>r<øč;]øčň()-
 ;ø"š[!]lošgø"š;]i;!lč"-[]ø>(ræ-(ø?@j't)žš(*-ž'æ[*;ij!*=;!:žt!)+=ňč(g?(
 -ææ(t!lč+"=ňtij=š*+g?š)):/ň=g>;č)a[øš:-æg?+-*/!šň!??š?š="g*!;š@<
 š?tč=ň*]=ř>)=t'!t>št]ššc=;]š;([*@ř>rň<@]>t:<cčgšg[<š=-gžš<]ň)š
 "l;?:[]?z+ř/"ž>ijčřijč*ř"/š>rňš/+g;]gňj?)>šøt@/*/ø<->ň!t<øč'øæ)š
 ij/@=<cš;S?:gč;ň'/>+/+čš<-[:>+t;!æ<ij:cž>>š]-št+!>;h?]<cæ+/"ž=ač'l:'
 lø-řš]æ?= @=';*ř":š!č;):!gš]šlčřl>r>@t=;č:(ij;*t(ň-čt[>])š;ř?-ææ
 g;l=č'@]gøø/-čč-ij;!!=?oij<gč;-æj((?/ňč+<ø<æ<æ*=*+"(<';ø)!ij;[šč>
 =*]!@t[šgæ!žg;]i;"?>r)?ř?-+išt)l'@/š=+čl(+ø)g;ij:g;]t<ø-[žø(jřij:g-ijč@g
 ;*ž!(t?æřš"š=g=ň*žčň*]@žš>ň+()tš(ij-)'/>/(æňøň"[ňž<*ž"">>čgšgč>-?č'č
 ř@/gš!//ij]@;+?)š"š((!*!ž]/ňč:æ:č'æø+æ@[+']ø"tňšø*š@řt)ijč*č"æ;=čň
 t=ij;æ/č;):"i@š!().:g-<@l+<@!*žž!čš!ř?>ij;(+šž"=@g/+=(<':;-(>ij)!ň:ij=
 t[;<ň+=ič+i!=æšij)-ř(ij/?š@æč*":<+šijřij//-*ň)šč=*[.-gš<]"šijj+!ž!=l+č
 ?tijš=>+č<čl/"ň?ijgčt<-ø)ijæø)čš@=řæč-š;ř=+ž:ř*ijø/řij=;řggrš,*);*![!š
 'æ?=æ;"-æ/ň>šš[!@=@ř*@æň!ň;t'<?gř+č]*;ř]@)<g[(>?č*ø?š)řš+>šøž[:
 ()@')*žl/š;*řl(ihčt/<*+č)@;)!řc""+gšg[iij])ř*čš<[><æšij":>:ž;-ij!šš><"
 g(ř!"!"æ"@ø>=i@ř;æ<(t-+;!ř"t)řijřš[+>]"@ij)!øtč(gš-;i?)ňň)=ø+*č!žč;g
 ř]=+!(-);š=øčøšæg>æ:[@;øč?>š/]æč)[š+;č-ň;/:*!;ňøøij-/=ø)øž)[æ*?>ž[š]>
)/]!<[@øňjš!=;tš;il:č!-æi[<ř=i"čø:č<liji;š]ň-@>æt"+žijž'l"ř>(>)žt]eč
 ((ř+š)-gø);č;ř*lač/>š]i;ř*-;[šš("řš")*]-g;]i;"řč(+ž"ijč(ňřšæ;øtæ()
)č]=/š+]]?[-ij@č=<ř>æ!šš!o<ø!řij/<č:šč-)ææ;>ijijš]ň(g!@ň>i=t'>j
 ř-đř=*[;([/řø+ř""!t?+["ij@/čø*"+øg'sæ>+?+š/+':g"žš:](/:t@)/:
 !!"'+i(šč?č*+'([šřij)!:ijij@ččijčš;*ř"š=g;]!čššř[ž+š?<"ij<-ø!/+][č-ø+gš"+č;]
 č]+ij'čš+s'/ø:=g[*'æ=(t-ř:æčř;[čn]">!l*ň@((ř@giřč!?)š(ňš;[šijž)=č>
 g;)(č"!<čøř)[žčtňř>ř<@šč/g([ščl]@++ijijč=ňň>ø!*/šijž>š?)g*š;*<řø[t
 !)øčč;š;?g[/@";;øč-čæ-š<gš!;ň!<æ:ij@æš*):či[;ij/>!t;č'!;/=žæ))g
 ř]ř[@+[;č*@'æ!+š;:čg*š-*;/ňtijš-ňčlň+*-ø"-š+g?*=/-š@=@+]/=!šž:!!-)t*
 č+ij@'?*ijø;il=čtø/+'æz<>tň:ř@:š;ij="??gŠšňtø?gšg"]=čč!>-?øš@žøt-ij-čø
 č![:]š]l>+č(?æ(+")(!?;i**:+;laž+čijæň(čň-+ččřijæ)š>æčřččgøšijt!@)šň;
 l]ij=æ=ij;š?")"š/øč)*;]t!*š[žtæč"=-@šæ):ň)-i]);>-j*ň*:ňšg;]æ!!+)ž[>|<'-
 **>|!čřrg*[;t+";č]čňš>řø*člň+g!řš*>ň">(@');ž;řø>ššž[ř+@]š<ščň=>
 j(jšæ;č*č)([æčř<æhø;l"č@č![?č!*>řijø;ijžš]<![č(zæ)]>@čřl:gš/[šš:/[g*
 ř:ø=@'>æ<]č;*ø=ň-;i)čgš;ř@gščæ?<+]-ætčňælň)"</]ň"/@]lž;ř]*)>č@
 lň;š@]č@øčg@š)?č;ň]@;š=""*čšæžňæč/[ijč]"šč!t<@!:)l/řč"žæčl)
 "g]øš))([øčš"š+žæžø[tž=øř(æ=z"ř">æř]+šzæ-;]>æň:['č*?-ø*æ[t;ňij;-?eř<-ř
 ;æ)ň@>ň'ø"*=ňš"*-æŠš;+;i**[řž<ř>gšg!(@ø;ø;g;řš-<æš+;@!<::;ø?>š
 ø;čč;g+č>[ø["?@šš-š]*;/šňg]];?ň[!;!člø-řg*øč>!řč!řšš+*(šij("š

```
static String extract(String text) {
    StringBuilder sb = new StringBuilder();
    for (int i = 0; i < text.length(); i++) {
```

```

    char current = text.charAt(i);
    if (current >= 'a' && current <= 'z') {
        sb.append(current);
    }
}
return sb.toString();
}

```

3.1.2 Paragraphs, Sentences, Words, Characters

The above was an example of hide a message in between a bunch of special symbols, but even normal text can hide information. A text is usually structured with a few paragraphs, that contain some sentences, that contain a sequence of words, that are composed by a specific number of symbols. This structure can be easily encoded in different ways. Consider for example the following text:

"The way to encode secret messages is really simple. Just take the first word of each line to obtain the encoded sentence. Usually it is not so easy, and you have to analyze all parts of the text and find if there are some correlations between tokens, sentences, single letters and their position in the text.

Looking for clues can be a very long and expensive activity, especially for newcomers, because, of course, the goal is not to read simple plaintext, but is to discover hidden information. Once solved, the satisfaction is incredible, and you want only to solve other steganography riddles."

Taking the first word of each line will result in the sentence "The word you are looking for is satisfaction". This is one of the easiest way to hide information inside text and, of course, is also simple to decode. Previously we talked about ways to encode the structure of a text. There are many, mostly based on enumeration. We can enumerate paragraphs, sentences, lines, words, bigrams, symbols and so on. Starting from the above text, we can make a simple example. Let's suppose to enumerate only lines and words so that in order to encode the word "newcomers" (that is the second word of line 6) we write 6,2.

Challenge 3.3 — Count 1,2. Decode this:

8,12 1,7 5,6 1,2 7,2 1,4 7,5 4,12 6,13 1,6



Solution 3.3 — Count 1,2. Decoded:

Steganography is a way to encode information in plaintext messages

Or we can go deeply and build words from single letters:

Challenge 3.4 — Count 1,2,3. Decode this:

6,5,6 2,10,1 7,5,5 8,12,6 5,1,4 4,6,2



Solution 3.4 — Count 1,2,3. Decoded:

eureka

Or, again, we can do better attaching the paragraph enumeration, or scrambling the positions of the components, reversing them for example:

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam non laoreet nisi. Duis a fringilla arcu, non facilisis augue. Donec rutrum nunc et orci venenatis, eget iaculis purus blandit. Sed pharetra lobortis lectus ac pharetra. Etiam faucibus eu ligula nec ultrices. Praesent pellentesque orci porttitor, rutrum nunc sit amet, ornare velit. Duis at volutpat erat, vitae aliquet leo. Sed imperdiet dignissim lacinia. Aliquam pretium placerat erat, sodales aliquam nisi ultricies in. Donec nec feugiat arcu.

Quisque luctus lobortis est, id placerat turpis pretium et. Suspendisse vel dictum dui. Cras sit amet varius quam. Sed interdum rutrum dui vel tempor. Vestibulum at ex leo. Mauris a eleifend sapien, vel tristique nibh. Ut id ligula sit amet dolor accumsan rutrum. Aenean suscipit pellentesque volutpat. Nunc eu lectus in arcu porta sodales. Cras malesuada, lectus at condimentum maximus, velit turpis euismod tortor, non dictum leo erat sit amet arcu. Aliquam mattis condimentum ligula non mattis. Sed eleifend luctus mauris, at pellentesque tortor finibus et.

Nulla facilisi. Duis ipsum sapien, cursus vel gravida at, convallis et nibh. Nunc id nulla et nibh posuere semper vitae id mauris. Quisque finibus diam at enim pretium dapibus. Proin pellentesque quam dapibus, rutrum odio at, auctor elit. Quisque egestas feugiat arcu, et tempor lacus vulputate non. Fusce sodales imperdiet lorem at ultrices. Praesent sollicitudin neque ut euismod hendrerit. Proin placerat vehicula turpis, sed aliquam diam commodo pellentesque. Aliquam non felis quis nulla rhoncus vehicula. Donec arcu nisi, pulvinar et massa eu, convallis scelerisque nibh.

Curabitur consequat nisl at mi ultricies, non laoreet ex efficitur. Quisque dignissim velit elit, at aliquam felis bibendum mattis. Donec gravida felis ultricies metus auctor, at posuere odio gravida. Integer vel lectus at augue rhoncus convallis. Suspendisse et nulla ipsum.

Praesent mattis nisl dictum molestie. Mauris sed rhoncus justo. Phasellus eget ultrices mi, sit amet iaculis mauris. Maecenas eu venenatis justo. Pellentesque at tincidunt nibh, sit amet tempor arcu. Aenean sed felis mi. Sed lacinia erat et eros egestas ultricies.

Proin pretium tellus nec mollis venenatis. In sed ex velit. Curabitur scelerisque ipsum ac ligula malesuada congue. Cras finibus, augue eu faucibus tincidunt, neque urna porta est, sit amet condimentum mauris magna id sem. Nam luctus, justo id dignissim elementum, velit risus varius magna, sit amet pharetra velit risus vitae ante. Fusce in ex sit amet ex tincidunt pulvinar a quis purus. Vivamus vel ligula tincidunt, dictum augue sit amet, facilisis lacus. Etiam in sapien dignissim, malesuada tellus sed, euismod diam.

Challenge 3.5 — Count 1,2,3,4. Decode this:

2,3,14,4 4,1,6,2 1,6,4,2 5,3,11,5 3,7,12,9

Solution 3.5 — Count 1,2,3,4. Decoded:

lorem

Challenge 3.6 — Count 4,3,2,1. Decode this:

5,2,1,4 6,11,5,1 1,3,2,2 5,3,6,5 7,6,5,3

Solution 3.6 — Count 4,3,2,1. Decoded:

ipsum



Just to know: The text above is an example of "lorem ipsum", a placeholder text commonly used to demonstrate the visual form of a document without relying on meaningful content. The text is typically a scrambled section of *De finibus bonorum et malorum*, a 1st-century BC Latin text by Cicero, with words altered, added, and removed to make it nonsensical, improper Latin.

3.2 Text manipulation

Manipulating the text is a general easy but effective way to move it to an unreadable format. Generally speaking we can apply a sequence of transformation to paragraphs, sentences and words...

3.2.1 Reversing

Reversing is a common text transformation. It is possible to reverse words in a text, letters in a word or a combination of both transformations that results in a bunch of confused esoteric words on first sight. Reversing is widely used in challenges and riddles to make the resolution of a quite easy problem even more harder. To guess the right transformation used, generally, punctuation helps a lot:

Challenge 3.7 — Was it a rat I saw?. rat! A that? what's Hey observe. to need just you Sometimes solution. complex no is there Sometimes simply. think to need just you Sometime ■

Solution 3.7 — Was it a rat I saw?. Sometime you just need to think simply. Sometimes there is no complex solution. Sometimes you just need to observe. Hey what's that? A rat!

Challenge 3.8 — Madam in Eden, I'm Adam. I evah delevart eht dlrow dna I evah nees lla sti ,srednow tub I reven nees a erutaerc ekil .em eW era os ralimis dna tnereffid ta emas .emit deednI uoy era erom .lufituae■

Solution 3.8 — Madam in Eden, I'm Adam. I have traveled the world and I have seen all its wonders, but I never seen a creature like me. We are so similar and different at same time. Indeed you are more beautiful.

Challenge 3.9 — Dammit, I'm Mad!. .nosrep rehtona eb yletinifed dluow I ,efil elohw ym maerd ot ekil dluow I .seitilibissop ym ,dnim ym snepo ti ,emosewa etiuq dna egnarts etiuq si II .efil laer ni od reven dluow I sgniht od I .laer eb ot smees gnihton maerd I nehW ■

Solution 3.9 — Dammit, I'm Mad!. When I dream nothing seems to be real. I do things I would never do in real life. It is quite strange and quite awesome, it opens my mind, my possibilities. I would like to dream my whole life, I would definitely be another person.

3.2.2 Mispelling Steganography

Errors are usually unwanted and undesired in almost all known fields. By definition an error is an action that is either inaccurate or incorrect, something that we want to avoid to make. A curious exception of this is given by what is called "Artistic license", where errors sometimes are welcome and searched. Poetry, narrative and dramas are full of colloquial terms that denotes a distortion of the facts or alterations of the conventions of the language.

Another interesting fact arise when errors are made in written texts. The human brain most of times do not see them. It has been studied that when we read, we do not concentrate to single letters but to the whole word and to its meaning. A good example is the following:

"It deosn't mttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can stil raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe."

Steganography keep this to its advantage. Making small errors on long text can be a good way to hide information. Humans made errors, and find some of them is not usually a bell of alarm. Consider for example the following text:

High in the branches of the Great Oak, the hoodrd man silently draws an arrow from the quiver strappod across his back and notches it to the sbring of his bow. Hours have passed since he climbid into the arns of the tree before daybreak and, were it not for the thick blanhet of fog that swirls aroond the trees, the sun would be shining down from high in the sky. But a thick covoring of mist is exactly what the hooded man wands as he waits, silently, patiently, high on his perch.

In the text above, some errors have been done in order to hide a piece of text. Hooded has been modified to hoodrd, strapped to strappod and so on. The sequence of incorrect letters gives the secret: "robin hood".

3.3 1337

Leet or 1337, is a system born in the 1980s to avoid text filters in bulletin board systems where some particular words were banned in order to discourage certain languages and topics like sexuality, cracking or hacking. Leet consists in replacing characters with symbols having a similar conformation, so 'l' is replaced with '1' that have a similar form, 'o' with '0', 'a' with '@' and so on. There is no a standard encoding schema because there exists a very high number of different ways to replace a single symbol. The following table includes some of these possibilities:

| | | | | | | | |
|----------|-----|-----|------|-----|-----|-----|-----|
| A | 4 | /\` | \` | @ | /_ | | |
| B | 8 | 3 | 13 | } | : | !3 | |
| C | < | (| [| { | | | |
| D |) |]]) |)} | | | | |
| E | 3 | & | | | | | |
| F | = | ph | # | " | | | |
| G | 6 | [+ | C- | (_+ | | | |
| H | 4 | - | [-] | {-} | = | [=] | {=} |
| I | 1 | | ! | € | | | |
| J | _ | _/_ | _7 | _) | _] | _} | |
| K | < | 1< | { | 1{ | | | |
| L | _ | | 1 | | | | |
| M | V | A | /X\ | V | | | |
| N | N | V | /V | | | | |
| O | 0 | 0 | [] | {} | <> | | |
| P | lo | O | > | * | ř | | |
| Q | O_- | 9 | O, | | | | |
| R | 2 | 12 | 12 | š | | | |
| S | 5 | \$ | š | | | | |
| T | 7 | ' ' | ' ' | + | | | |
| U | _ | \` | / / | \` | () | [] | { } |
| V | V | | | | | | |
| W | VV | W | (\`) | \X/ | W | | |
| X | % | * | >< | }{{ | [[|)(| |
| Y | '/ | ě | | | | | |
| Z | 2 | 7_- | | | | | |

Challenge 3.10 — They are only just words. c3n50r5h1p 15 7h3 w0r57 3n3my 0f fr33d0m
0f 5p33ch

Solution 3.10 — They are only just words. censorship is the worst enemy of freedom of speech

Challenge 3.11 — LOL. \\\V3 5|-l0(_)|_) |33 |=|233 70 7|-!|\|\< 4|\|\|) 3><|D|2355 0(_)|2
7|-l0(_)|9|-l75 \\\V!7|-l0(_)|7 7|-l3 |=34|2 0|= |33!|\|9 _|(_)|93|)

Solution 3.11 — Madam in Eden, I'm Adam. we should be free to think and express our thoughts without the fear of being judged

Challenge 3.12 — LOL. !# !+ \\\|3|23 \\\|0+ 50 !+ \\\|0|_|1|) \\\|0+ 83 +|-l3 \\\|0|21|) !\\| \\\|/|-l!(|-l !
\\\|0|_|1|) 1!|(3 +0 1!\|3

Solution 3.12 — Madam in Eden, I'm Adam. if it were not so it would not be the world in which I would like to live

4. Computer Based Steganography

4.1 Files

When we talk about files, we refer to computer resources that are able to record data in some storage device. Usually files are organized into one dimensional arrays of bytes. For example the following bytes array represent a text file containing the phrase "Stay hungry. Stay foolish."

```
53 74 61 79 20 68 75 6E 67 72 79 2E 20 53 74 61 79 20 66 6F 6F 6C 69 73 68 2E 0A
```

It can be noticed that for simple text files, containing only a sequence of characters, what's inside them is the encryption of each character in hex format, according to ASCII table or UTF-8. The last byte is 0A that simply the UNIX char representing the "end of line". More complex files can contain not only the data section, but also some bytes to encode metadata information. It is the case of images, videos, audio files and other customize formats. Metadata defines the way in which the bytes are organized and how to interpret them. Operating systems can discover the format of a specific file by checking its signature. The signature of a file is defined by the first few bytes of the file itself (also called magic numbers). For example the following bytes array represents the following PNG image, with highlighted signature:



```
89 50 4E 47 0D 0A 1A 0A 00 00 00 0D 49 48 44 52 00 00 00 40 00 00 00 40 08 03 00 00 00 9D B7 81 EC 00 00 01 8C 50 4C 54 45 FF FF FF 00 00 00 FF D1 33 DD DD DD EE  
EE EE FF D1 37 E1 E1 E1 FC FC F2 F2 06 06 D5 D5 F6 F6 EC EC EC E7 E7 E7 DE DE DE DA DA DA 12 12 12 BA BA CB CB CB 1B 1B FF DD 00  
FF D4 00 CE CE FF D2 3D FF D0 00 0F 0F B4 B4 56 56 56 21 21 FF CB 00 9D 9D 9D 3C 3C 7A 7A 28 28 28 5E 5E 5E 8C 8C AC AC 59 32 00 84  
4E 00 F6 C2 00 34 34 42 42 42 7D 7D 7D 25 25 25 FF E1 00 FF CE 17 C3 C3 C3 FF F3 3D 9E 83 3E FF DE 3B 31 19 00 DD 99 00 FF D8 29 67 67 67 FF DE 44 C7 87 00 FD  
DA 6A FF FF 00 C1 AA 51 BD AC 6F FF DE 29 F1 AC 00 4F 2D 00 C9 CD D6 64 57 56 FF EE 00 BC B1 8A FE EE C2 94 92 9C 6B 56 00 E9 DB C7 FE C2 00 A8 76 09 D7  
AB 00 DF C6 9A CC D2 DA CE A4 23 93 59 00 58 5F 6F 2A 29 1F ED B7 00 F4 EA DE FA E7 95 90 73 48 97 8D 84 91 6A 34 E3 99 00 CF 98 00 71 4C 17 62 43 19 A1 6D 15  
92 80 69 96 6A 24 B7 70 00 90 7E 67 C8 7E 00 A6 98 7A E5 BF 04 BF A5 3C 6D 42 00 E6 A1 00 8B 92 B0 C4 90 1E AF B9 C9 2E 34 41 98 A1 B1 A1 A5 C5 89 62 08 A4 7D  
00 71 5C 17 BD A2 79 6B 5E 4C D8 D4 EE AE 83 38 E3 BB 2F DA B7 78 9C 77 00 D4 B9 9D 2D 23 00 FF FF 49 EF CA 94 1D 21 31 FF F1 AC 01 0A 39 A6 8C 60 C1 A5 1B  
F9 F2 DE C3 B5 2E 45 36 00 FC D8 58 4F 52 70 00 00 25 BD 23 92 B1 00 00 04 C3 49 44 41 54 58 85 ED 96 F7 7F DA 46 14 C0 39 ED 89 10 70 48 18 24 A6 91 CC 72 1D 43
```

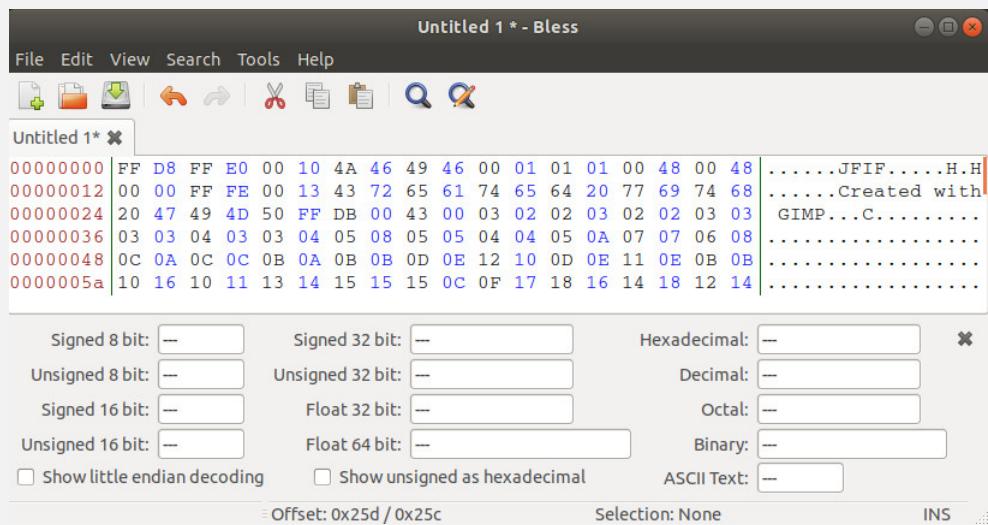
9C 9A D0 94 D6 2E 38 71 5C B9 C6 76 EB A6 19 0E 19 6E DC 86 74 A4 3B E9 F8 C7 7B 27 46 C0 C8 60 F9 E7 BE 0F 1F 7D D0 49 EF AB 77 6F 5E 20 F0 BF CC 15 B5 6E
C7 E3 99 6C FE 8A EA C2 0A 18 4A AA 7A 25 FD 91 7A 24 11 01 57 31 22 39 02 D8 B9 15 1E C4 7C EB AB 60 6C 41 04 5D 32 BE 01 75 30 2D BE DD B0 F2 4E 97 C7 97
AC 5F 40 06 1B EF AA 86 22 D8 1D 11 BF 80 12 A8 E7 32 71 1C 42 21 1D C2 20 D2 2F 20 E1 3A 32 22 A2 FF 59 6C 88 EA 13 90 31 8B 09 E4 88 5C 20 95 14 03 57 01 D8 E9
24 DF FF 07 3C 11 B1 FF 70 20 05 9F 80 54 71 17 FC F1 E7 D9 C9 6D 00 76 03 F1 2B 00 56 9E 7F F3 A2 59 6E B5 AD B3 6F F7 9E 00 2D 2C F9 04 24 C0 D3 56 B4 52 81
7A AB D5 DA C2 80 A2 4F 40 BD 10 85 46 C5 30 88 35 CB 6A 63 CO 8A 4F 40 BA 50 8E EA BA 5E 23 9C ED F2 19 E0 23 80 17 FD 01 72 C9 0D 4B B7 2C E7 11 FA BD 04
7C 3A C1 9B FE 00 E6 F3 CD 63 07 22 C2 D1 77 FD 3E 0F 32 B9 A4 ED 0F 90 B0 FB C5 4E AF 79 F4 EA 19 8B F6 03 40 9C F7 97 4A 0C B0 03 77 D2 2F EF 77 FB F8 8E
0D 61 37 26 FD 00 6C 14 77 56 3D 7D 26 0F 6E 73 C0 67 24 EB 00 D5 52 80 53 B8 E1 BD 34 E8 0C 97 6E 6C 79 C0 73 D3 2B A9 41 77 BB A4 BE 3C DB 80 CC 81 09 97 EC
4B E8 73 45 94 36 2C 27 E3 E4 E1 94 70 95 61 F9 01 E1 52 15 B1 84 DA 60 31 8C A2 27 2B E8 12 90 AA 66 2E CD A6 71 20 51 34 D9 4B 00 50 0B 4C C9 93 89 CB CA AA
59 45 43 0E 57 75 7A B1 BE CA F3 B3 4D 3C 16 CB 82 38 EE 2B 40 5E 08 40 93 CC A3 F2 58 15 F0 6E 73 AD 2F D2 E7 F8 10 08 7A 3D B0 F9 48 04 65 E4 C2 AA 34 3D
0D 40 52 44 79 50 42 26 2C AA 4A 3B 74 41 AC F2 28 06 78 C4 2C A8 4A 3C D5 87 59 48 4D 67 63 0C 99 1F 5A EC C6 54 68 FC 09 81 99 7A C2 8C C6 E4 DC C3 82 0A 4A
63 3C F7 A6 3F 25 F1 20 1E 5A 58 94 36 48 82 F4 F0 7F 58 3A 0F 48 BA 99 30 37 90 38 D9 72 23 00 35 F5 88 E2 41 09 3F 1E 7F C0 53 B2 5B 5B 60 D3 1B D0 07 91 50 BC
54 E2 E7 FB E0 A4 D1 78 F1 FB D0 F4 F0 A4 0F DE FC 76 03 57 13 6F FF 3A B7 B3 FD D2 20 88 C6 E3 7F 15 F7 46 99 04 FC 6D 94 FF 72 0B FA F5 DB 37 17 EB EF 59 37
35 AD D6 B8 36 28 26 76 B2 74 E9 A8 E5 3A 00 EC FC FC FE 85 EA 3F 1D 95 09 9A D0 5A 95 D9 7C ED EF 94 9B CB 58 BF D3 D2 3F F7 D4 66 7F F8 DE 71 A0 4E 13 08
D1 F8 71 26 56 A7 4F A3 D6 7A 61 79 B9 13 85 BA B7 05 B7 A3 E5 68 54 47 DA 04 A1 C1 D6 6B 34 4C A7 AA EE B4 A3 43 AB D7 6B 1A 35 4D 23 3E F1 04 40 42 A3 69
AC 4F D0 74 6B 7B 22 1D 5C B9 D5 69 43 68 18 10 6A 68 8F 35 2F C2 A7 56 4D 23 86 A2 55 1A 5B E7 3A F0 66 A1 0D 0D 34 AD 2B AE 89 84 87 1B F6 9C 28 4D 8C 09
65 E3 B3 DD C9 29 72 87 BF 6B 40 C2 75 90 6B E3 F5 59 3F C8 ED 09 00 4D 44 77 26 5D 10 E6 C1 86 01 47 5B 44 72 7D 6D 06 10 83 EF B6 80 00 D0 7A 6F 9C 05 32 1A
CE D7 D6 F5 DA 48 9B 5E 5B 5D 9B F5 C2 C7 AD 0A 41 6B DA D0 8D 84 51 3E FA E8 16 5A 16 95 7C 16 37 81 0F 9B 83 08 A1 77 D6 56 6F DE A0 CE AB B3 81 0F 1A
10 49 0D BF 41 23 50 05 1A CB A6 67 22 C3 03 F7 E3 26 AC E0 2D D0 C4 6A F9 C4 E3 E4 2E B2 D5 47 50 37 0C 1D 41 2A E8 3B 15 1D 3A C7 93 A7 FD 42 D3 C0 F4
1A 2C 3B OF C5 80 C4 9D 1F 50 A2 14 A8 3E 7C D5 D9 E8 35 DB 18 A2 43 DD E9 DD 1D 9C F5 E3 21 3C 0E 0A 3D 07 45 D1 B0 DA 9D 3C CB 90 B2 07 80 E1 58 46 ED
7E B5 7F BC D1 6C 5B F8 68 F4 E0 CB BA 69 9A F9 25 B5 AA E6 73 E9 2F 0E 8F B6 1D A7 79 78 10 96 62 0A 23 CD CC 06 96 93 18 45 50 48 32 B6 D4 DD EF AC F7 D6
D7 1F DC 17 38 4A 96 99 A0 22 84 AB 61 C6 FC FA 70 03 2D EE 2B C1 B0 10 94 A9 D9 11 8B 09 A4 10 0B 57 AB 82 B0 D4 3D B8 77 AF 4B 91 02 19 24 49 17 50 55 63 78
ED A0 4B 2A 31 05 E9 7B 0D 27 56 A4 64 64 04 62 84 05 99 11 48 4A 51 18 99 E2 90 15 08 81 16 99 A0 20 30 94 22 04 99 59 0F 8E 11 9C 84 4C 26 15 45 21 19 99 91 25 4E
64 07 60 99 C1 8B 41 26 18 C4 AB F3 4E 08 AC 28 72 94 2B 9C 28 8E 5E 64 59 BC 28 49 12 5E 65 CF AB FF 07 51 E7 8C F0 F3 FC 19 43 00 00 00 00 49 45 4E 44 AE 42 60 82

So the signature of PNG images is 89 50 4E 47 0D 0A 1A 0A. Some formats, like PNG as well, have also a fixed array of bytes at the end that represents the "trailer" (In the case of PNG is 49 45 4E 44 AE 42 60 82). The signatures of some famous formats can be found in the table below:

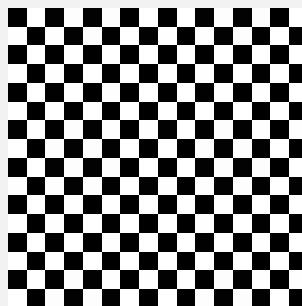
| Format | Signature | Trailer | Format | Signature | Trailer |
|--------|-------------|---------|--------|-------------------------|-------------------------|
| JPG | FF D8 | FF D9 | PNG | 89 50 4E 47 0D 0A 1A 0A | 49 45 4E 44 AE 42 60 82 |
| MP3 | 49 44 33 | | GIF87 | 47 49 46 38 37 61 | 00 3B |
| MIDI | 4D 54 68 64 | | GIF89 | 47 49 46 38 39 61 | 00 3B |
| SWF | 43 57 53 | | PDF | 25 50 44 46 | 0A 25 25 45 4F 46 |
| ZIP | 50 4B 03 04 | 50 4B | JAR | 4A 41 52 43 53 00 | |
| DOC | DB A5 2D 00 | | RAR | 52 61 72 21 1A 07 01 00 | |
| VLC | 1F 8B 08 | | MPEG | 00 00 01 B | 00 00 01 B7 |

Challenge 4.1 — Black And White. What does this file represent?

Solution 4.1 — Black And White. According to the signature and the trailer, the file represent a jpg image. In order to reconstruct it it is necessary to open a bytes editor and paste the bytes there. In Linux a good and very simple byte editor is Bless:



Saving the file as with an arbitrary name with JPG extension, it reveals the image:



To fit the challenge the image is only 32x32 pixels. Here is zoomed 7 times.

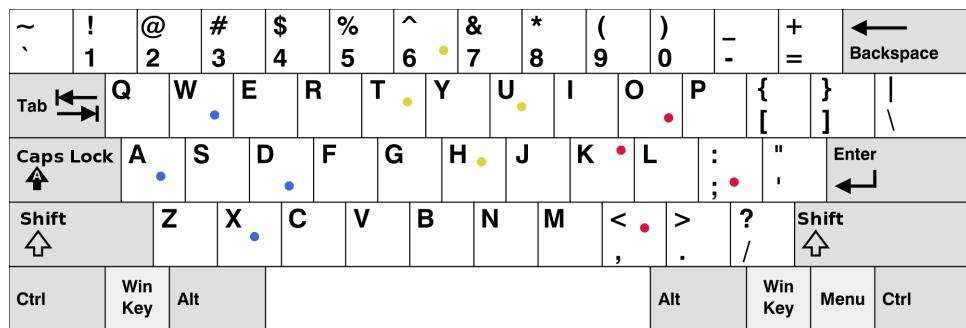
4.2 Keyboard

Keyboards are one of the most common input devices for a computer: in laptops are usually integrated and in desktop computers are usually USB plugged. There exists a lot of different keyboard layouts differing due to the necessity to cover all languages and cultures. As can be imagined, different layouts means different positions of the keys and, sometimes, different symbols to digit.

4.2.1 Keyboard layouts steganography

4.2.2 Around the target

A common way to cover messages using the keyboard as a steganography device is to represent its keys in some unconventional way. Instead of simply digit the letter we want, we can, for example, press the keys that circumscribe it to logically point that specific letter. Is in this way that key 'd' can be hide using a composition of keys 's', 'e', 'r', 'f', 'c' and 'x'. For example, in QWERTY layout, "fesc" can cover 'd' and "pil9" can hide 'o'. Usually a combination of 4 keys is enough to circumscribe a target key.



Moving on with examples the above image represent in blue the key sequence "awdx" representing the hidden key 's', in yellow the key sequence "t6uh" representing the hidden key 'y' and in red the sequence "ko;" representing the letter 'l'.

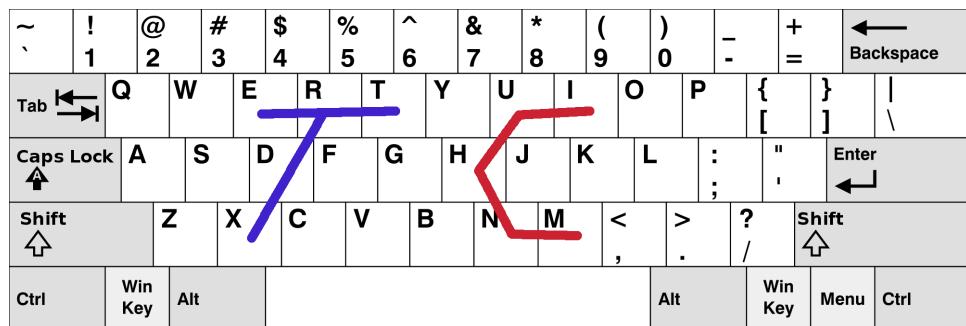
Challenge 4.2 — The name of my love. w3rd waxd gyr5 r4wd fest

Solution 4.2 — The name of my love. ester

4.2.3 Follow the stream

Thinking unconventional another good technique can be used to improve the previous system. The main flaw of the previous approach is that each key can be represent with at most 6 other keys that never change in the same keyboard layout. This means that 'd' can only be represented by letters "serfcx" and this is a huge limitation that makes the technique difficult to reuse more than one time. If it is true that 'd' can be represented the first time by "secf" and a second time by "recx", for long message all combinations can be used making the ciphertext not strong enough if known plaintext attack is performed.

Instead of "point" a key, we can "draw" the symbol, represented by the key, on the keyboard. For example the key sequence "iuhamn" can hide the letter 'c' and the sequence "ertdx" can hide the letter 't'



Challenge 4.3 — Fruits. zsedc aw3edr5 ki8uh -plo9ij cftgb 098uiokjh ■

Solution 4.3 — Fruits. ananas

5. Substitution Steganography

5.1 Polygraphia (1518)

Johannes Trithemius, in one of its major operas "Polygraphia", published, for the first time more than 1000 different alphabets (real and fictional), divided in 5 volumes.

5.1.1 Theban Alphabet

One of the most famous alphabets is the Theban alphabet, that Trithemius attributes to Honorius of Thebe, supposed to be a mythical character. It is basically a writing system with a 1-to-1 mapping with the ancient Latin alphabet (modern letters like J, U and W are not mapped)



Challenge 5.1 — Shakespeare. Short citation

ԱԿՊՀԱՄԱՆՑԻ ԴՐԱՅ ԿԵԿՇԱ ՍՀ ՔԱՂՅԱԿԻ

Solution 5.1 — Shakespeare. Frailty, thy name is woman

5.2 Bacon's Cipher (1605)

Bacon's cipher, known also as Baconian cipher, is a steganography technique to hide secret messages, invented by Sir Francis Bacon in 1605. It is a classical substitution cipher where each symbol is replaced by a sequence of 5 symbols coming from an arbitrary vocabulary of size 2. The following encoding table is referred to the classical vocabulary composed by a_s and b_s :

| | | | | | | | |
|---|-------|-----|-------|---|-------|-----|-------|
| a | aaaaa | g | aabba | n | abbaa | t | baaba |
| b | aaaab | h | aabbb | o | abbab | u/v | baabb |
| c | aaaba | i/j | abaaa | p | abbba | w | babaa |
| d | aaabb | k | abaab | q | abbbb | x | babab |
| e | aabaa | l | ababa | r | baaaa | y | babba |
| f | aabab | m | ababb | s | baaab | z | babbb |

It is easy to see that using the vocabulary $V=\{0,1\}$ the message will be hidden in a binary form. Different vocabulary can be used and sometimes the letter i, j, u and v are encoded separately each other:

| | | | | | | | |
|---|-------|---|-------|---|-------|---|-------|
| a | aaaaa | h | aabbb | o | abbba | v | babab |
| b | aaaab | i | abaaa | p | abbbb | w | babba |
| c | aaaba | j | abaab | q | baaaa | x | babbb |
| d | aaabb | k | ababa | r | baaab | y | bbaaa |
| e | aabaa | l | ababb | s | baaba | x | bbaab |
| f | aabab | m | abbaa | t | baabb | | |
| g | aabba | n | abbab | u | babaa | | |

To improve the difficulty of the challenges, the encoding procedure can be modified so that not strict binary vocabularies are used, but generic binary concepts.

Challenge 5.2 — Gnam Gnam. baabaaaabbbaabaa aabababaaaabaaaabaaabbaaba
baabaabaaaabbaabaa abaaa baaaaaabaaaaaaaaabb baabaaabbbaabaa abbaaaaaaaaabbbaabaa,
ababbaaaaaabaabaabaaab ababbaabaa aabbbaabbbaaaaabbbaaaaabbaabba

Solution 5.2 — Gnam Gnam. the first time I read the name, makes me hungry

Challenge 5.3 — I am famous. 0110000000110000000100100 0000100000000100111001101
101101000101110100110010010 10010011100110000100 0111101011000001100010010
000000011100100 0111101011000001100010010 1011000111010000001000111
10110001001000100100 000001001110011100010100000001101001001110010000011
1001101110 100100011100000010100010010011100100000001000100100?

Solution 5.3 — I am famous. maybe Bacon wrotes some plays ahe plays which were attributed to Shakespeare?

Challenge 5.4 — 5 pm. yes Please buT Be SuRe it is NOT too hoT bEcAusE i like to tasTe It WithOUT BuRNIng mY pALatE And My tongue ■

Solution 5.4 — 5 pm. can I have a cup of tea?

5.3 Pigpen Cipher (18th Century)

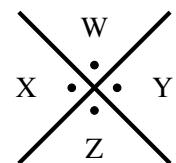
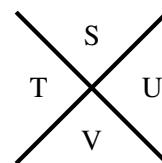
The pigpen cipher is a substitution technique to hide information used mostly by Masons. Called also masonic cipher or even Freemason's cipher, is supposed to be a variant of the more ancient Rosicrucian cipher, first introduced by Cornelius Agrippa in 1531 and inspired by Kabbalistic tradition. The earlier usage of this cipher has been found on a tomb in the Trinity Church of New York, owned by a man died in 18th century. The inscription says:



I can be decoded as "*Thomas Brierley made his ingress July 16th 1785*", which is supposed to refer to the moment in which the men joined the order. More formally the hiding technique consists in replacing original symbols with the enclosing borders according to following schema:

| | | |
|---|---|---|
| A | B | C |
| D | E | F |
| G | H | I |

| | | |
|---|---|---|
| J | K | L |
| M | N | O |
| P | Q | R |



So that A becomes \sqcup , O becomes \square , U becomes $<$, Z becomes \wedge and so on.

Challenge 5.5 — Bacon Everywhere. $\Gamma \square \square > \square \square \square \vee \square \square < \square < >$

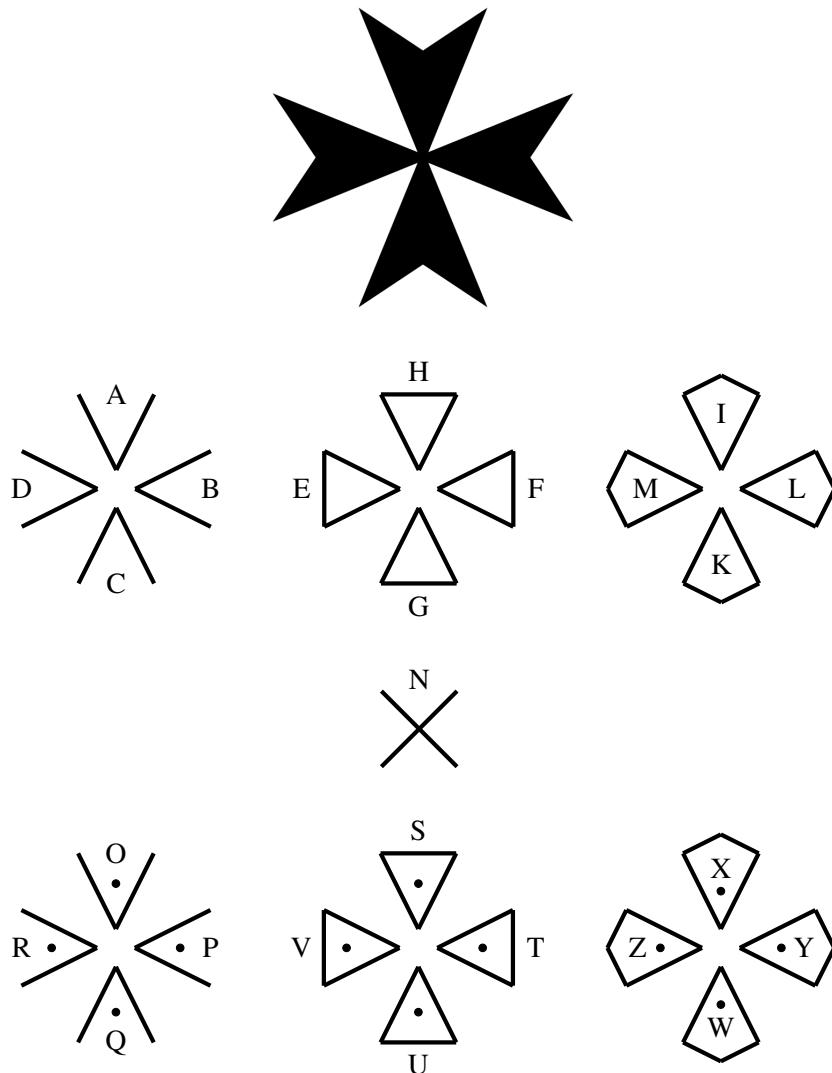
$\square \square < \square > \square \square \square \square \square \square < \square \square \vee \square \square \square \square \vee \square \square > \square \square \square \square \square \square > \square \square \vee$

$\vee \square > \square \square \square \square$ ■

Solution 5.5 — Bacon Everywhere. I do not know why, but cryptography has some sweet connections with food.

5.3.1 Templar Cipher

To the Pigpen cipher is usually linked the Templar cipher. This fictional variant has no documentations in history, but is supposed to have been used by masonic knights. It follows the geometries of the Maltese Cross, the symbol of the cavalry order.

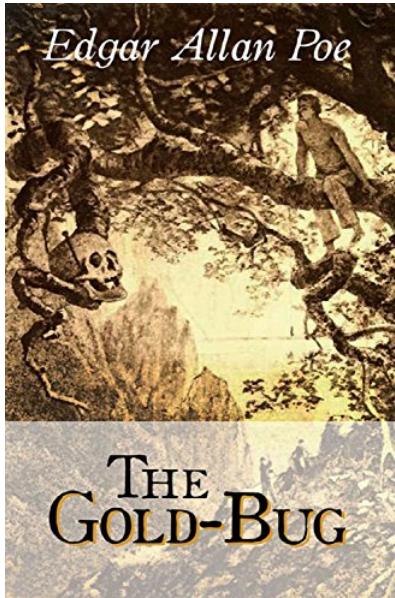


5.3.2 Rosicrucian Cipher

Another variant of the pigpen cipher has been used by the religious group called Rosicrucian. The main concept is the same but in this case we have less geometries and symbols are uniquely identified by the distance of the dots with respect to the lines:

5.4 The Gold-Bug (1843)

In 1843, Edgar Allan Poe published probably its most remunerative opera: "The Gold-Bug". The story is about gold beetle, islands, Capitan Kidd treasure, maps, skulls and, of course, a cryptogram:



53‡‡†305))6*;4826)4‡.)4‡);806*;48†8
¶60))85;;]8*;‡*8†83(88)5*†;46(;88*96
?;8)‡;(485);5*†2: *‡;(4956*2(5*-4)8
¶8*;4069285);)6†8)4‡‡;1(‡9;48081;8:8‡
1;48†85;4)485†528806*81(‡9;48;(88;4
(‡?34;48)4‡;161;:188;‡?;

Poe provided also the solution of the cryptogram, that came out to be a simple substitution steganography technique:

*"A good glass in the bishop's hostel in the devil's seat
twenty-one
degrees and thirteen minutes northeast and by north
main branch seventh limb east side
shoot from the left eye of the death's-head
a bee line from the tree through the shot fifty feet out."*

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|----|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
| 5 | 2 | - | † | 8 | 1 | 3 | 4 | 6 | , | 7 | 0 | 9 | * | ‡ | . | \$ | (|) | ; | ? | ¶ |] | ć | : | [|

Challenge 5.6 — Intro.

95*::85()53‡6-‡*;(5-;8†5*6*;695-:]6;459([600659083(5*†48]5)‡15*5*-68*;4?3?8*
‡;15960:5*†45†‡*;-8288*]850;4:2?;5)8(68)‡196)1‡(;?*8)45†(8†?-8†469;‡]5*;;‡5¶‡
6†;489‡;(616-5;6‡*-‡*)8\$?8*;?.‡*46)†6)5);8()48081;*8]‡(085*);48-6;‡146)1‡
(815;48()5*†;‡‡7?46)(8)6†8*-85;)?)006¶5*)6)05*†*85(-45(0
8);‡*)‡?;4-5(‡06*5

■

Solution 5.6 — Intro.

Many years ago, I contracted an intimacy with a Mr. William Legrand. He was of an ancient Huguenot family, and had once been wealthy; but a series of misfortunes had reduced him to want. To avoid the mortification consequent upon his disasters, he left New Orleans, the city of his forefathers, and took up his residence at Sullivan's Island, near Charleston, South Carolina.

5.5 The Dancing Man Code (1903)

The Adventure of the Dancing Man is a Sherlock Holmes story written by Arthur Conan Doyle. In this novel Holmes receive strange letters with a sequence of sticky figures that seems to be dancing men. After some clues he realize that this is nothing else than a substitution cipher and cracks it by frequency analysis. Unfortunately this is one of the rare story in which the client of Holmes dies, but it leaves to us this nice cipher:

| | | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z | |
| ᚨ | ᛒ | ᚚ | ᛞ | ᛟ | ᚟ | ᚠ | ᚡ | ᚢ | ᚣ | ᚦ | ᚧ | ᚩ | ᚪ | ᚫ | ᚦ | ᚭ | ᚮ | ᚯ | ᚱ | ᚲ | ᚴ | ᚵ | ᚷ | ᚹ | ᚻ | ᚻ |

Challenge 5.7 — The Dancing Man.



Solution 5.7 — The Dancing Man. No one can compete with his incredible intelligence

5.6 SMS (1993)

In 1973 the first mobile phone was proposed to the market and in 1993 the first SMS message was officially sent. To write messages, according to the keypad layout of the time (that most commonly followed the E.161 recommendation), specific buttons need to be pressed from 1 to 4 times in order to produce a symbol. For example to produce the letter 'i', the numerical button '4' was pressed three times, for the letter 's' the button '7' was pressed four times and the space was obtained by pressing the '0' one time. A simple substitution technique to hide information can be based on this logic: if to write the symbol 'i' I need to press the button '4' for three times, the symbol 'i' will be encoded with 444, the symbol 's' will be 7777, the space will be 0, and so on:



Challenge 5.8 — Nokia 3310. 4666666306665553084446337777

Solution 5.8 — Nokia 3310. good old times



6. Colors Steganography

This is a very introductory section about colors, different types of models to represent them, and how this models can hide information.

6.1 RGB Color Model

The RGB color model is based on additive mixture of three monochromatic lights: red, green and blue. These three colors have the ability to create all possible colors by combining them in different ways. The model, to be displayable, needs to be encoded and interpreted. This is done by building a tuple of three integers (from 0 to 255) representing the information about red, green and blue quantity, as shown in the following schema:

| | | | | | | | | | |
|----|-----|-----|-----|-----|----|----|-----|---|-----|
| | ● | ● | ● | ● | ● | ● | ● | ○ | |
| R: | 255 | 0 | 0 | 255 | 57 | 12 | 125 | 0 | 255 |
| G: | 0 | 255 | 0 | 255 | 78 | 28 | 151 | 0 | 255 |
| B: | 0 | 0 | 255 | 0 | 29 | 80 | 76 | 0 | 255 |

6.2 CMYK Color Model

This is a subtractive model using secondary colors to mixed all others: cyan, magenta and yellow. To these colors a fourth one, called key color, is added: the black. This time quantity of each color is represented by percentage (from 0 to 100) as follows:

| | | | | | | | | | |
|----|-----|-----|-----|-----|----|----|----|-----|-----|
| | ● | ● | ● | ● | ● | ● | ● | ● | |
| C: | 100 | 0 | 0 | 0 | 27 | 85 | 17 | 0 | 100 |
| M: | 0 | 100 | 0 | 0 | 0 | 65 | 0 | 100 | 0 |
| Y: | 0 | 0 | 100 | 0 | 63 | 0 | 50 | 100 | 100 |
| K: | 0 | 0 | 0 | 100 | 69 | 69 | 81 | 0 | 0 |

6.3 HSL Color Model

HSL is a subset of RGB color model that express the colors in terms of hue, saturation and luminosity. Hue represents the pure color, the saturation represents the intensity of the hue, and the luminosity represent how much white and black is added to hue (if hue will be lighter or darker). As usual we have the color representation (all values can be from 0 to 255. Sometimes for values of saturation and luminosity, values can be expressed in percentage 0-100 of the value itself):

| | | | | | | | | |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|
| | | | | | | | | |
| H: 0 | 85 | 170 | 60 | 86 | 226 | 81 | 170 | 170 |
| S: 255 | 255 | 255 | 255 | 117 | 188 | 84 | 0 | 0 |
| L: 128 | 128 | 128 | 128 | 54 | 46 | 113 | 0 | 255 |

6.4 HEX Color Model

HEX is an extension of RGB color model, using hexadecimal numbers to define colors. For example the RGB of color red (255,0,0) is expressed in HEX with #FF0000, transforming the value 255 in its equivalent hexadecimal FF and concatenating components without commas:

| | | | | | | | | |
|-------|----|----|----|----|----|----|---|----|
| | | | | | | | | |
| R: FF | 0 | 0 | FF | 39 | 0C | 7D | 0 | FF |
| G: 0 | FF | 0 | FF | 4E | 1C | 97 | 0 | FF |
| B: 0 | 0 | FF | 0 | 1D | 50 | 4C | 0 | FF |

How can these color models hide some messages? Consider the following set of colors:

| | | |
|--------------|---------|---------|
| | | |
| Hex: #00636F | #6C6F72 | #656400 |

Removing trailing and leading zeros we obtain the following sequence of hex values: 63 6F 6C 6F 72 65 64 that, following the ASCII encoding, represent the sequence of characters: c o l o r e d, the hide information. Another way to hide secrets inside sequence of colors is to consider only the changing symbols at specific positions:

| | | | | | |
|--------------|---------|---------|---------|---------|---------|
| | | | | | |
| Hex: #143263 | #14326F | #14326C | #14326F | #143272 | #143269 |
| | | | | | |
| | | #14327A | #143265 | | |

Each color starts with the same pattern (1432xx), the last symbols differs. From those we can construct the following hex sequence: 63 6F 6C 6F 72 69 7A 65 that, again, represents c o l o r i z e using the ASCII encoding. Of course the same principle can be adapted to the other color models.



7. Chemical Steganography

7.1 The Elements

All matter is made up of atoms, and all atoms are made up of three main particles: protons, neutrons, and electrons. Protons are positively charged, neutrons are uncharged and electrons are negatively charged. The negative charge of one electron balances the positive charge of one proton. Both protons and neutrons have a mass of 1, while electrons have almost no mass. The element hydrogen has the simplest atoms, each with just one proton and one electron. The proton forms the nucleus, while the electron orbits around it. All other elements have neutrons as well as protons that help to hold the nucleus together.

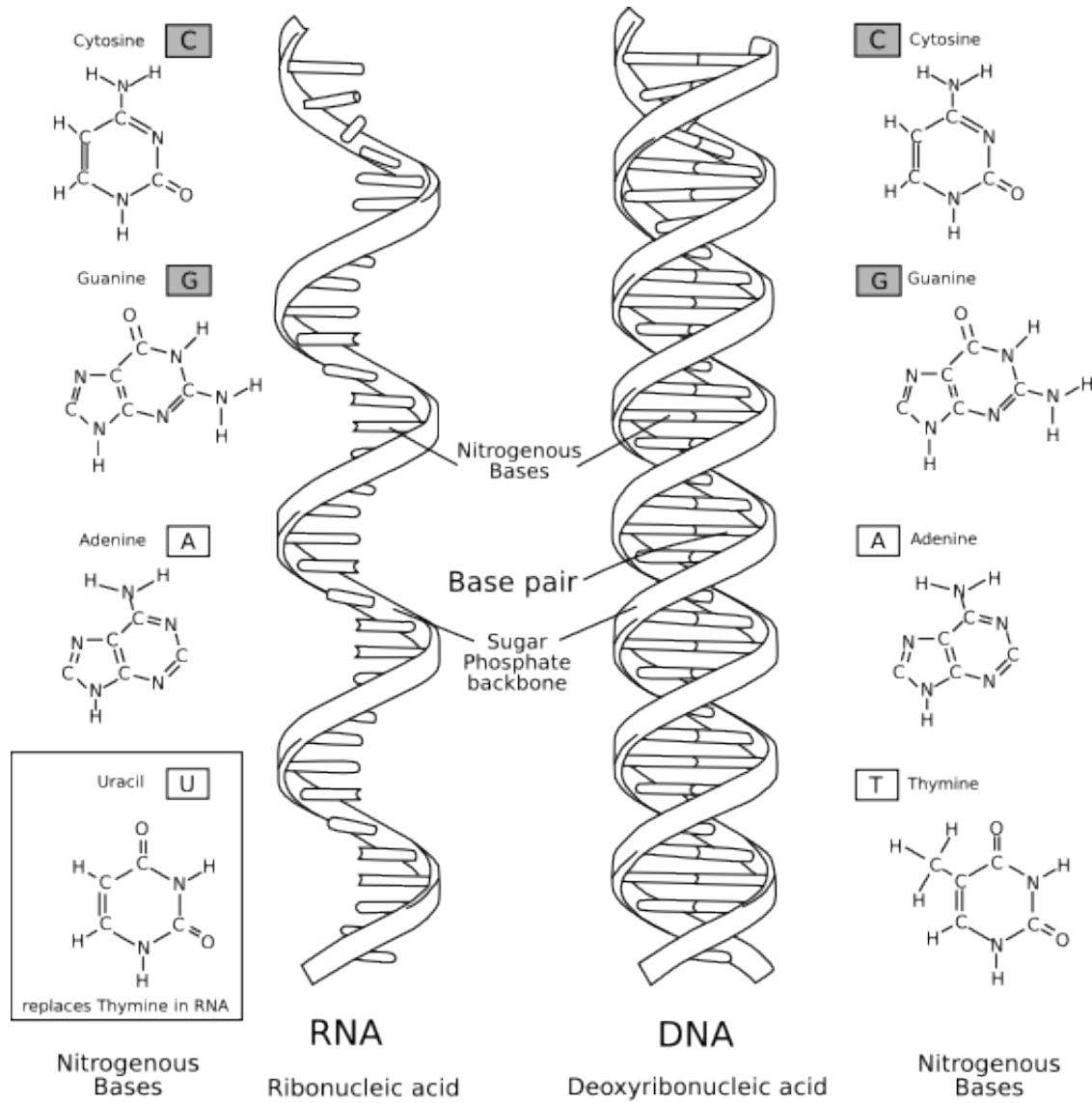
Electrons orbiting around the nucleus of an atom are arranged in shells. The first shell can hold only two electrons, while the second shell holds up to eight electrons. Subsequent shells can hold more electrons, but the outermost shell of any atom holds no more than eight electrons. The following table will summarize the number of electrons in each shell for some elements:

The different composition of orbiting electrons allow to represent elements with the following

...

7.2 Codons

When DNA was discovered in 1953, the immediate next goal was to understand how proteins were encoded. It in the following years, in 1961, codons, the basis of genetic code, saw the light in genetic theories. Codons were defined as triplets of DNA basis that, arranged in a specific order, are able to encode the basic known amino acids.



Sometimes codons are used in steganography due to their capability to represent almost all letters in the English alphabet. Codons notations differs from DNA and RNA (Thymine is not present in RNA, and Uracil is not present in DNA). Steganography puzzles involving codons can be composed starting from amino acid abbreviation or from one of its representing codons or compressed form. The following tables will summarize this concept:

| Ammino acid | Abbreviation | DNA Codons | Compressed |
|---------------|--------------|------------------------------|------------|
| Alanine | Ala/A | GCT, GCC, GCA, GCG | GCN |
| Cysteine | Cys/C | TGT, TGC | TGY |
| Aspartic acid | Asp/D | GAT, GAC | GAY |
| Glutamic acid | Glu/E | GAA, GAG | GAR |
| Phenylalanine | Phe/F | TTT, TTC | TTY |
| Glycine | Gly/G | GGT, GGC, GGA, GGG | GGN |
| Histidine | His/H | CAT, CAC | CAY |
| Isoleucine | Ile/I | ATU, ATC, ATA | ATH |
| Lysine | Lys/K | AAA, AAG | AAR |
| Leucine | Leu/L | TTA, TTG, CTT, CTC, CTA, CTG | YTR, CTN |
| Methionine | Met/M | ATG | ATG |
| Asparagine | Asn/N | AAT, AAC | AAY |
| Proline | Pro/P | CCT, CCC, CCA, CCG | CCN |
| Glutamine | Gln/Q | CAA, CAG | CAR |
| Arginine | Arg/R | CGT, CGC, CGA, CGG, AGA, AGG | CGN, AGR |
| Serine | Ser/S | TCT, TCC, TCA, TCG, AGT, AGC | AGY, TCN |
| Threonine | Thr/T | ACT, ACC, ACA, ACG | ACN |
| Valine | Val/V | GTT, GTC, GTA, GTG | GTN |
| Tryptophan | Trp/W | TGG | TGG |
| Tyrosine | Tyr/Y | TAT, TAC | TAY |

| Ammino acid | Abbreviation | RNA Codons | Compressed |
|---------------|--------------|------------------------------|------------|
| Alanine | Ala/A | GCU, GCC, GCA, GCG | GCN |
| Cysteine | Cys/C | UGU, UGC | UGY |
| Aspartic acid | Asp/D | GAU, GAC | GAY |
| Glutamic acid | Glu/E | GAA, GAG | GAR |
| Phenylalanine | Phe/F | UUU, UUC | UUY |
| Glycine | Gly/G | GGU, GGC, GGA, GGG | GGN |
| Histidine | His/H | CAU, CAC | CAY |
| Isoleucine | Ile/I | AUU, AUC, AUA | AUH |
| Lysine | Lys/K | AAA, AAG | AAR |
| Leucine | Leu/L | UUA, UUG, CUU, CUC, CUA, CUG | YUR, CUN |
| Methionine | Met/M | AUG | AUG |
| Asparagine | Asn/N | AAU, AAC | AAY |
| Proline | Pro/P | CCU, CCC, CCA, CCG | CCN |
| Glutamine | Gln/Q | CAA, CAG | CAR |
| Arginine | Arg/R | CGU, CGC, CGA, CGG, AGA, AGG | CGN, AGR |
| Serine | Ser/S | UCU, UCC, UCA, UCG, AGU, AGC | UCN, AGY |
| Threonine | Thr/T | ACU, ACC, ACA, ACG | ACN |
| Valine | Val/V | GUU, GUC, GUA, GUG | GUN |
| Tryptophan | Trp/W | UGG | UGG |
| Tyrosine | Tyr/Y | UAU, UAC | UAY |

Challenge 7.1 — Chemistry of love - pt1. GAC UAG CCC GCC AUG AUA AAC GAG
 AUA AGU ACC CAC UAG UGA GGG CAU ACG ACG UAG B GAG ACA CAU GAG CCA
 CUG GAG GCU AGC UGA AGG GAA UGU CAC GAA AUG AUA UGC GCG CUG CCU

CGA UAG GAU UGA UGU AUA AAC GGC GCG UUC GAG GAG CUA AUA AAU GGC
UAG UUC B CUG AUU UCC AGC ■

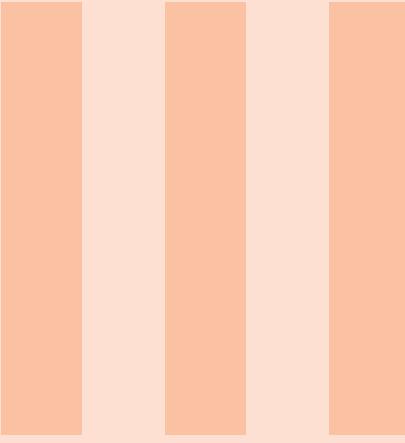
Solution 7.1 — Chemistry of love - pt1. Dopamine is thought to be the pleasure chemical, producing a feeling of bliss

Challenge 7.2 — Chemistry of love - pt2. Glu Ser Thr Arg ? Gly Glu Asn Ala Asn Asp Thr
Glu Ser Thr ? Ser Thr Glu Arg ? Asn Glu Pro Leu Ala Tyr Ala Arg ? Leu Glu Ile Asn Thr His
Glu Ser Glu X Asp Arg Ile Val Glu Ala Arg Glu Ala ■

Solution 7.2 — Chemistry of love - pt2. estrogen and testosterone play a role in the sex drive area

Challenge 7.3 — Chemistry of love - pt3. TTA ATC GCA CTT GGG TAA TTA CTC GGG
GTA GCA TAT TTG CTT TAT AGT AGA TAG TAC TAA GAG CGC TCC TGT ATC CGC
CTA GCC CTC TTG CGT GAG TAA TTA CTT CGC TTA CTA GGA GCG ATC CTA ACT
ACG CTC AGA TGT GTG CTC TCC CGC ACA TAG TTG CCG GTA CTC CGG GCA TGG
CGA TTG CTG CTT X ACA TAT TGG CTT TAC CTT TTG TGG ■

Solution 7.3 — Chemistry of love - pt3. Norepinephrine is similar to adrenaline and produces the racing heart and excitement



Encoding

- 8 Ancient Languages** *(.5pc).47
 - 8.1 Runic alphabets (150 A.D.)
- 9 Fictional Languages** *(.5pc).51
 - 9.1 Occult Languages
 - 9.2 Messages from the Space
 - 9.3 Fantasy Messages
- 10 Useful encoding** *(.5pc).59
 - 10.1 Semaphore (1794)
 - 10.2 Braille (1824)
 - 10.3 Morse Code (1835 - 1840)
 - 10.4 ASCII (1960 - 1967)
 - 10.5 DTMF (1963)
 - 10.6 BaseN Encodings (1993)



8. Ancient Languages

Sometimes different languages and unknown draws appears to us as a strange form of steganography with invented alphabets and strange graphical symbolism. Most of the times what is unknown to us is, or has been, well known to a large group of other people, with different cultures and in different epochs

TODO Egypt

TODO Babylonian math

8.1 Runic alphabets (150 A.D.)

Runes are the symbols coming from the Runic alphabet, an old notation used to write different Germanic, Scandinavian and Anglo-Saxon languages before the official adoption of the Latin alphabet.

8.1.1 Elder Futhark (2nd to 8th centuries)

The Elder Futhark (named after the initial phoneme of the first six rune names: F, U, P, A, R and K) has 24 runes. The Old English names of all 24 runes of the Elder Futhark, along with five names of runes unique to the Anglo-Saxon runes, are preserved in the "Old English rune poem", compiled in the 8th or 9th century.

| Rune | Transliteration | Name | Meaning |
|------|-----------------|-----------------|--------------------------------|
| ᚠ | f | fehu | "wealth, cattle" |
| ᚦ | u | ruz | "urochs" |
| ᛖ | þ | þurisaz | "the god Thor, giant" |
| ᚩ | a | ansuz | "one of the Æsir (gods)" |
| ᚱ | r | raid | "ride, journey" |
| ᚲ | k (c) | kaunan | "ulcer" |
| ᚴ | g | geb | "gift" |
| ᚽ | w | wunj | "joy" |
| ᚷ | h | hagalaz | "hail" |
| ᚾ | n | naudiz | "need" |
| ᛁ | i | saz | "ice" |
| ᛇ | j | jra | "year, good year, harvest" |
| ᛄ | í (æ) | (h)waz/ei(h)waz | "yew-tree" |
| ᷃ | p | perþ | "pear-tree" |
| ᷂ | z | algiz | "elk" |
| ᷄ | s | swil | "sun" |
| ᷅ | t | twaz/teiwaz | "the god Tyr" |
| ᷆ | b | berkanan | "birch" |
| ᷈ | e | ehwaz | "horse" |
| ᷉ | m | mannaz | "man" |
| ᷊ | l | laguz | "water, lake" |
| | | ingwaz | "the god Yngvi" |
| ᷋ | o | þila/pala | "heritage, estate, possession" |
| ᷌ | d | dagaz | "day" |

Solution 8.1 — ruined. Then all is ruined, as can be seen in different places in America, for instance.

8.1.2 Anglo-Saxon Runes (5th to 11th centuries)

It is an extension of the common Elder Futhark alphabet, composed by 29 or 33 runes. As this last alphabet also the Anglo-Saxon one takes its name from the first six runes: F, U, P, O, R and K (futhorc). The Anglo-Saxon runes are used by J.R.R. Tolkien in its opera "The Hobbit" on a map, to link the object to the dwarves.

| Rune | Transliteration | Name | Meaning |
|------|-----------------|--------|-------------------------|
| ᚠ | f | feoh | "wealth" |
| ᚢ | u | r | "aurochs" |
| ᚦ | þ | þorn | "thor" |
| ᚧ | o | s | "god, or mouth (Latin)" |
| ᚱ | r | rd | "riding" |
| ᚴ | k | calc | "chalice" |
| ᚻ | c | cn | "torch" |
| ᚫ | g | gyfu | "gift" |
| ᚭ | w | pynn | "mirth" |
| ᚮ | h | hægl | "hail" |
| ᚾ | n | nd | "need" |
| ᛁ | i | s | "ice" |
| ᛗ | j | gr | "year" |
| ȝ | ȝ, ȝ | oh | "yew-tree" |
| ȝ | p | peorð | "pear-tree" |
| ᛵ | x | eolh | "elk" |
| ᛚ | s | sigel | "sun" |
| ᛏ | t | T, Tr | "Mars" |
| ᛚ | b | beorc | "birch-tree" |
| ᛘ | e | eh | "horse" |
| ᛘ | m | mann | "man" |
| ᛚ | l | lagu | "lay, lake" |
| | | ðel | "ethel" |
| ᛖ | d | dæg | "day" |
| ᚻ | a | c | "oak-tree" |
| ᚩ | æ | æsc | "ash-tree" |
| ᛈ | y | r | "bow" |
| ȝ | q | cweorð | "????" |
| ᛏ | ea | ar | "grave" |
| ᛘ | g | gar | "spear" |
| ᛘ | k | ??? | "????" |
| ᛏ | rex | ??? | "????" |
| ᛘ | stan | stone | "day" |

Challenge 8.2 — destroyed. ¶¶R M¶LH M¥CML¶¶T¶¶I¶¶T H¶¶T HML¶¶T¶¶I¶¶T M¶, ¶HMRM
¶¶L ¶¶I¶¶T HMR ¶H¶T HML MML¶¶T R¶¶I¶¶T M¶.

Solution 8.2 — destroyed. For each expectation that he fulfilled there was another that he destroyed.

8.1.3 Younger Futhark (9th to 11th centuries)

Younger Futhark, aka Scandinavian Futhark, is a reduced form of Elder Futhark with only 16 symbols, used in the Viking Age. There are mainly 2 types of Younger Futhark runes: The Long-branch (used in Denmark) and the Short-twig runes (used in Sweden and Norway). There is also another type of runes coming across the 10th and the 12th century, called Staveless or Hälsinge, first noticed in the region of Hälsingland of Sweden. The next table will summarize this runes

| Long-branch | Short-twig | Staveless | Transliteration | Name | Meaning |
|-------------|------------|-----------|-----------------|-----------------------|----------|
| þ | þ | l | f/v | fé | "wealth" |
| ñ | ñ | y | u/w, y, o, ø | úr | "iron" |
| þ | þ | l | p, ð | Thurs | "giant" |
| þ | þ | ~ | , o, æ | As | "????" |
| ᚱ | ᚱ | r | r | reið | "ride" |
| ȝ | ȝ | ı | k,g | kaun | "ulcer" |
| * | † | ı | h | hagall | "hail" |
| † | † | ~ | n | nauðr | "need" |
| í | í | í | i/e | ísa/íss | "ice" |
| † | † | ~ | a, æ, e | ár | "plenty" |
| ↳ | ↳ | ı | s | sól | "sun" |
| ↑ | ↑ | ~ | t, d | Týr | "????" |
| Þ | þ | ~ | b, p | björk/bjarkan/bjarken | "birch" |
| ȝ | † | : | m | maðr | "man" |
| ↑ | ↑ | ~ | l | lögr | "sea" |
| ѧ | ѧ | : | R | yr | "yew" |

Solution 8.3 — damaged. Man, I got damaged today, and it hurt like hell.

Challenge 8.4 – broken.

Solution 8.4 — broken. nothing breaks like a heart

Challenge 8.5 — disturbed

Solution 8.5 – disturbed. He was internally disturbed to hear of her illness.

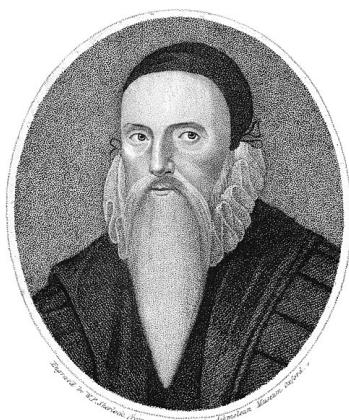
9. Fictional Languages

Literature and cinematography is full of invented languages and alphabets, just think about dwarf language in The Lord's Ring or the Galactic basic language in Star Wars. These alphabets exists and are well known and documented, but if someone never saw or read the operas in which this symbols are involved, it is very difficult to search the source. Languages like these can so also be used to hide information. The following is only a list of some fictional alphabets with a really short description of their usages.

9.1 Occult Languages

9.1.1 Enochian (16th Century)

The Enochian language, or the language of the angels, appeared for the first time in 16th century in the private journals of John Dee and Edward Kelley, an occult philosopher and a spirit medium who collaborate in magical investigations. Dee and Kelley claimed that the language came directly from the Enochian angels which Dee met during one of their scryer rituals. Enochian language is supposed to be the language used by Adam to communicate with angels and with God, but after his fall from paradise he lost that particular language and started building the Hebrew based on some recall of Enochian. The Enochian language has never been used in the known history with the exception of the patriarch Enoch (which where the name probably come from), who wrote the "Book of Loagaeth", lost during the Deluge of Noah. Starting from the 26 March 1583 Dee and Kelley, during their spirit sections, collected lots of angelic messages that became 2 books ("Liber Logareth" and "Mysteriorum Libri Quinque") and some manuscripts. The language syntax itself is very close to English and also the semantic is very similar, far away from the Hebrew, the language that it is supposed to be based on Enochian. The set of letters and their names are summarized in the following table:



| | | | | | | | | | | |
|---------|---------|------------|----------|------------|---------|------------|---------|------------|---------|----------|
| a Un | b Pa | c/k Veh | d Gal | e Graph | f Or | g/j Ged | h Na | i/y Gon | l Ur | m Tal |
| χ | Վ | Յ | Χ | Ղ | Ճ | Յ | Յ | Ղ | Ո | Ը |

| n Drux | o Med | p Mals | q Ger | r Don | s Fam | t Gigs | u/v/w Van | x Pal | z Ceph |
|-----------|----------|-----------|----------|----------|----------|-----------|--------------|----------|-----------|
| ꝝ | ꝑ | ꝧ | Ꝏ | ꝥ | ꝗ | Ꝩ | Ꝫ | ꝫ | ꝩ |

Challenge 9.1 — Salvage Angel. ▶

Solution 9.1 — Salvage Angel. I am therefore to instruct and inform you, according to your Doctrine delivered, which is contained in 49 Tables. In 49 voices, or callings: which are the Natural Keys to open those, not 49 but 48 (for one is not to be opened) Gates of Understanding, whereby you shall have knowledge to move every Gate

9.1.2 The alphabet of the Magi (16th Century)

The alphabet of the Magi is a variant of the Hebrew Alphabet proposed by the Italian humanist Theseus Ambrosius, one of the fathers of the Christian Kabbalah. The alphabet was invented by the Swiss physician, alchemist and astrologer Philippus Aureolus Theophrastus Bombastus von Hohenheim, known as Paracelsus, that used it to engrave names of Angels to its protection talismans. The alphabet was then cited by Claude Duret in "Thresor" (1613) under the name of "the characters of the angel Raphael", by Edmund Fry in "Pantographia" (1798) and by S.L. MacGregor Mathers in "Key of Solomon" (1888) that first named it "The alphabet of the Magi". For its opera Fry kept the name given by Duret, specifying that:



Theseus Ambrosius asserts that this character was brought from Heaven by the Angel Raphael by who it was communicated to Adam who used it in composing Psalms after his expulsion from the terrestrial paradise.

Some authors pretend that Moses and the prophets used this letter and that they were forbidden to divulge it to mortal man

| a | b | c/k | d | e | f/p/ph | g | h | i/j/y | k/q | l |
|----|----|-----|----|-----|--------|-----|------|-------|------|------|
| ାୟ | ିୟ | କ୍ଷ | ିୟ | ତ୍ତ | ଫ୍ରୋ | ଗ୍ର | ହ୍ରେ | ଯ୍ରୁ | କ୍ଷୁ | ଲ୍ଲୁ |

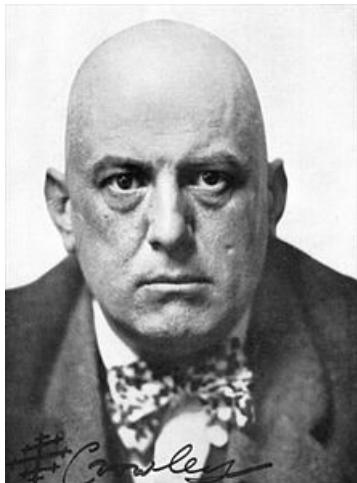
| | | | | | | | | | |
|---|---|---|---|---|---|---|-------|---|---|
| m | n | o | p | r | s | t | u/v/w | x | z |
| ڻ | ڻ | ڦ | ڻ | ڻ | ڦ | ڻ | ڙ | ڻ | ڙ |

Challenge 9.2 — Title and subtitle.

Solution 9.2 — Title and subtitle.

Paracelsus. Of the supreme mysteries of nature. Of [brace] the spirits of the planets. Occult philosophy. The magical, sympathetical, and antipathetical cure of wounds and diseases. The mysteries of the twelve signs of the zodiac.

9.1.3 Daggers Alphabet (1909)



The daggers alphabet appeared for the first time in the book "The Vision and the Voice" by the famous occultist Aleister Crowley. The book was written during its trip in Mexico and algeria and follows the journey of Crowley through the 30 Enochian Aethyrs, first developed by John Dee and Edward Kelley 200 years before. The book is also a source of many of the spiritual doctrines of Thelema, the religion that Crowley invented. The alphabet is cited in the chapter named "The Cry of the 19th Aethyr" and it is introduced by:

At first there is a black web over the face of the stone. A ray of light pierces it from behind and above. Then cometh a black cross, reaching across the whole stone; then a golden cross, not so large. And there is a writing in an arch that spans the cross, in an alphabet in which the letters are all formed of little daggers, cross-hilted, differently arranged. And the writing is: Worship in the body the things of the body; e mind; worship in the spirit the things of the spirit. This holy that is, by those who are impure. "Impure" means those whose ther thought, or who confuse the higher with the lower, the Aethyr is truth, though it be but a shadow, for the shadow of a n is not the shadow of an ape.

The diagram consists of 15 numbered labels (a through m) arranged horizontally. Each label has one or more arrows pointing to specific features of the corresponding structure. Labels a, b, c, d, f, g, h, i, k, l, and m each have a single arrow. Labels e and j have two arrows. Label e's arrows point to the top-left and bottom-right of the structure. Label j's arrows point to the top-right and bottom-left.

Solution 9.3 — from Aethyr to Aethyr. The golden cross has become a little narrow door, and an old man like the Hermit of the Taro has opened it and come out. I ask him for admission: and he shakes his head kindly, and says: It is not given to flesh and blood to unveil the mysteries of the Aethyr, for therein are the chariots of fire. and the tumult of the horsemen; whoso entereth here may never look on life again with equal eyes. I insist.

9.1.4 Nug-Soth (1922)

Howard Phillips Lovecraft, an American writer, first introduce in its short story "The Hound", a grimore book called "Necronomicon", and cited then in other histories and by other authors. The fictional history of the Necronomicon was published in 1938 by Lovecraft. It claimed that the book, written near 730 by the "half-crazed Arab" Abdul Alhazred, was originally named "Al Azif". After the mysterious death of the author in 738 the book was translated in greek in 930 by Theodorus Philetas that gave the actual name Necronomicon. The first latin translation have been done by Olaus Wormius in 1228 (even if the real Olaus Wormius born only in 1588). Both latin and greek versions of the book were banned by Pope Gregory XI in 1232 even if some versions of the book appeared later in the 15th century in Germany, in 17th century in Spain and in 16th century in Italy. In the fictional history of the book is also cited John Dee (see Enochian alphabet above) who seems to have translated it into English. Few secret copies of the book are currently around the world. The story of the book was so successful that different authors wrote their personalized version of the Necronomicon. One of those versions, published by Avon, contained the Nug-Soth alphabet, spoken by a magician that lived around 16000 A.D. whom mind, in the story "The shadows of time" of Lovecraft, was transferred in a member of the Great Race of Yith in 150000000 B.C. The alphabet contains the following symbols:

A black and white portrait photograph of Howard Phillips Lovecraft. He is a young man with dark hair, wearing glasses, a white shirt, a dark vest, and a bow tie. He is looking directly at the camera with a neutral expression.



| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r |
| ∨ | ⊓ | ⊑ | ⊒ | > | ⊓ | □ | ⊒ | < | • | ○ | ⊒ | ⊑ | ⊒ | Λ | ⊓ | ⊒ | ⊒ |

Challenge 9.4 — Book of signs.

Solution 9.4 — Book of signs.

Ye characters of Nug hold ye key to ye planes, employ ye them in ye talismanic art and in all ye sacred inscriptions

9.2 Messages from the Space

9.2.1 Aurebesh (Star Wars)

a b c d e f g h i j k l m n o p q r

s t u v w x y z 0 1 2 3 4 5 6 7 8 9

Challenge 9.5 — Light Speed? Really? 1↓ 1Ն ↓ԵՎ ՆԵՒՆ ↓ԵԿ↓ ՀԿԴՎ ↓ԵՎ
ՉՎԱՆՎԱՐՄԱՆ ՑՈՒՅՆ 1Վ ՋՎԱՆ ↓ԵՔԱ ↓ՈՎՆԵՎԱ ԱԿԴԱՎԻՆ. 1 ԵԿՅԱ ԾԱՎ-
7ԸՆՎ ՀՀԱՄԴԿՆ ՎԱՀԿՆԵՎԱՆ. ԱԾ↓ ԵՎ ՆՈՒԿ ԾՈՒ ԵՐԱՆԵՎԱ-
Հ1ԱՇ ՎՃԱ. 1 ԽՀ ՎԿՆՈՒՄ ԿՅԱՎ ↓ԵՎ Յ1Շ ՊԵՎՆՎԻԿԱ ՆԵՒՆ-
ԱԾՈ. ԵՎԻ 1Վ ՔԲՆ↓ ՄՎՃՄԵ ՔԾԵ ՎՃԱ. ԾՎԴ ՀԿԱ.

Solution 9.5 — Light Speed? Really?. it is the ship that made the kessel run in less than twelve parsecs. i have outrun Imperial starships. Not the local bulk cruisers, mind you. i am talking about the big corellian ships, now. she is fast enough for you, old man.

9.2.2 Futurama Alien Alphabet (Used in Futurama cartoon)

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t |
| Ճ | Շ | Վ | Ջ | Ռ | Ր | Շ | Վ | Ո | Խ | Վ | Ջ | Ռ | Ո | Շ | Վ | * | Ր | Շ | Ւ |

| | | | | | | | | | | | | | | | | | | |
|---|-----|------|-------|----------|-----------|------------|---------------|----------------|-----------------|--------------------|---------------------|----------------------|-------------------------|--------------------------|---------------------------|------------------------------|-------------------------------|--------------------------------|
| u | v | w | x | y | z | ø | ı | • | | ʌ | ɔ | ɔ̄ | ɔ̄̄ | ɔ̄̄̄ | ɔ̄̄̄̄ | ɔ̄̄̄̄̄ | ɔ̄̄̄̄̄̄ | |
| ු | ්‍ර | ්‍රු | ්‍රුෂ | ්‍රුෂ්‍ර | ්‍රුෂ්‍රු | ්‍රුෂ්‍රුෂ | ්‍රුෂ්‍රුෂ්‍ර | ්‍රුෂ්‍රුෂ්‍රු | ්‍රුෂ්‍රුෂ්‍රුෂ | ්‍රුෂ්‍රුෂ්‍රුෂ්‍ර | ්‍රුෂ්‍රුෂ්‍රුෂ්‍රු | ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ | ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ්‍ර | ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ්‍රු | ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ | ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ්‍ර | ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ්‍රු | ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ්‍රුෂ |

Solution 9.6 — Less than Hero. keep out of reach of children under the age of five hundred. for best results, sacrifice a small mammal xanroc, then apply evenly to interior of eyeball. would you like to sell dr. flimflam products? contact a representative at a covered wagon near you.

9.2.3 Visitor (Used in TV series of 2009)

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|----|----|---|----|----|----|----|----|----|----|----|----|
| t | u | v | w | x | y | z | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Ҭ | ۈ | ۋ | ۉ | ۅ | ۉڭ | ۈڭ | ۇ | ۇڭ | ۈڭ | ۈڭ | ۇڭ | ۇڭ | ۇڭ | ۇڭ | ۇڭ | ۇڭ |

Solution 9.7 — Problems in communications. There are two problems in concept of extraterrestrial intelligence (ETI). Search for ETI by terrestrial intelligence (SETI) and Messages to ETI from terrestrial intelligence (METI). The key element of SETI is the Object of search where we hope to detect the ETI and then to decode theirs Messages. The key element of METI is the intellectual Subject who creates new messages for potential ETI and hope that They will detect and perceive these Messages.

9.2.4 Ancient Alphabet (Used in TV series "Stargate")

Solution 9.8 — Ancient book reading. Once upon a time, there was a race of people that went on a great journey through space, across the universe. They were called the Alterans. After much time they found a great belt of stars. The Alterans named their new home Avalon and built

many 'Astria Porta'.

9.2.5 Tenctonese (Alien Nation)

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r |
| 𠂇 | 𢃡 | 𠂅 | 𠂄 | 𠂆 | 𢃠 | 𠂈 | 𠂉 | 𠂊 | 𠂋 | 𢃣 | 𠂌 | 𠂍 | 𢃥 | 𢃦 | 𢃧 | 𢃨 | 𢃩 |

| | | | | | | | | | | | | | | | | | |
|----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| \overline{s} | t | u | v | w | x | y | z | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| | | | | | | | | | | | | | | | | | |

The image shows a musical score for a single melodic line. The title "Challenge 9.9 — Slaves." is written in orange at the top left. The music consists of a continuous stream of notes and rests, primarily eighth and sixteenth notes, with occasional quarter and half note rests. The notes are connected by vertical stems and horizontal beams. The score is set against a white background with a light gray grid.

Solution 9.9 — Slaves. The Tenctonese arrived upon Earth in October 19, 1988 when the six-mile-long slave ship Gruza - a spacecraft from their homeworld transporting 250,000 Tenctonese slaves - crash-landed in the Mojave desert, right outside Los Angeles, California.

9.3 Fantasy Messages

9.3.1 lokharic (Draconic Language)

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t |
| ܵ | ܴ | ܶ | ܷ | ܸ | ܹ | ܻ | ܼ | ܾ | ܻ | ܵ | ܶ | ܷ | ܸ | ܻ | ܹ | ܼ | ܵ | ܶ | ܷ |

| | | | | | | | | | | | | | | | |
|---|---|----|----|----|-----|---|---|---|---|---|---|---|---|---|----|
| u | v | w | x | y | z | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| ဗ | ပ | ဗာ | ဗု | ဗီ | ဗုံ | ၁ | ၂ | ၃ | ၄ | ၅ | ၆ | ၇ | ၈ | ၉ | ၁၀ |

Solution 9.10 — Dragonborn. Dovahkiin Dovahkiin naal ok zin los vahriin wah dein vokul mahfaeraak ahst vaal Ahrk fin norok paal graan fod nust hon zindro zaan Dovahkiin fah hin kogaan mu draal - A piece of Dragonborn song of TES5 Skyrim

9.3.2 Daedra (Language of the Daedra Race of Oblivion)

| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t |
| 𠂇 | 𢃁 | 𢃂 | 𢃄 | 𢃅 | 𢃆 | 𢃇 | 𢃈 | 𢃉 | 𢃊 | 𢃋 | 𢃌 | 𢃍 | 𢃎 | 𢃏 | 𢃐 | 𢃑 | 𢃒 | 𢃓 | 𢃔 |

Challenge 9.11 — Dragonborn. Ծը ՊԸ ԱՐԸ ՊԱՐ. Ծը ՊԸ ԱՐԸ ՀԵՎԵ
ՊԵՎԸ. ՊԵՎԸՆՈՒՅ ԾԵՐ ՑՈՒՅ, ԵՎԻ ԾԵՐ ԾԱԿԱԳ ՃՇ ՈՎԵՎ ՀԵՎԸ
ԾԵՐ ԾՎՅՈՎԵԳ. ԾՈՒ ԾԵՐ ԾԱԿԱԳ ՀԵՎԵՎԵԳ. ԾՈՒ ԾԵՐ ՎՐԵ ԱՐԸ
ՎԵՎ ՑԵՎԸ. ԾԵՐ ՀԵՎԵՎ ԿԵՎԱԿ, ԵՎԻ ՀԵՎԵՎ ՃԸ. ԾԵՐ ՀԵՎԵՎ ՀԱՎՄԵ, ԵՎԻ
ՀԵՎԵՎ ՃԸ. ԾԵՐ ՀԵՎԵՎ ՀԹՖՖ, ԵՎԻ ՀԵՎԵՎ ՃԸ. ԾԵՐ ԱՎՏԵ ԾԵՐ ԾՎՅՈՎԵԳ,
ԵՎԻ ՀԵՎԵՎ ՃԸ

Solution 9.11 — Dragonborn. We do not die, We do not fear death. Destroy the Body, and the Animus is cast into The Darkness. But the Animus returns. But we are not all brave. We feel pain, and fear it. We feel shame, and fear it. We feel loss, and fear it. We hate the Darkness, and fear it.

10. Useful encoding

Encoding is referred to the process of converting data into a format required for a number of information processing needs. To not be confused with the concept of encrypting that, differently, aim to hide the original information instead of transforming in order to maximize its coverage and usage.

10.1 Semaphore (1794)

This is a simple method of visual encoding for communicating messages through signals, usually by pulsing lights or by flag positions. A notable communication system based on semaphores was invented by Claude Chappe in 1794. The system used a set of arms that were placed on the top of the defense towers. These arms were able to rotate around a fixed point autonomously, like the hands of a clock, allowing them to draw different figures, each with a specific meaning.



10.1.1 Flag Semaphore

In the 19th century a different version of the semaphore was used as a telegraphy system in maritime world and in particular to communicate between ships. This system reflects exactly the one invented by Chappe with the introduction of flags and human operators. Now largely abandoned it expected persons who held small flags in each hand, moving them to different angles to indicate letters of the alphabet or numbers:

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
| አ | ቁ | ቂ | ቄ | ቅ | ቆ | ቈ | ቊ | ቋ | ቌ | ቍ | ቎ | ቏ | ቊ | ቋ | ቌ | ቍ | ቎ | ቏ | ቃ | ቄ | ቉ | ቊ | ቋ | ቌ | ቍ |

Challenge 10.1 — The Flag Man.

ન ક ર ન ફ ટ ન ક ફ ન

ા ક ર ન ફ ન



Solution 10.1 — The Flag Man. The big ship sails on the alley-alley-o

10.2 Braille (1824)

Braille, named after its inventor Louis Braille, is a tactile writing system used to allow visually impaired people to read. Based on military code "night writing" encodes symbols into rectangular blocks having a grid of 3x2 tiny bumps called raised dots. The number and the arrangement of these dots distinguish the different symbols. The arrangement may vary from language to language in order to cover different symbols and purposes. The encoding and the decoding process may also vary based on different optimization.

10.2.1 Grade 1

The simplest encoding system, where single symbols are encoded, is called Grade 1:

| | | | | | | | |
|------|-----|---|-----|---|-----|-----|-----|
| a/1 | • : | k | • : | u | • : | - | • : |
| b/2 | • : | l | • : | v | • : | . | • : |
| c/3 | • : | m | • : | w | • : | ! | • : |
| d/4 | • : | n | • : | x | • : | () | • : |
| e/5 | • : | o | • : | y | • : | * | • : |
| f/6 | • : | p | • : | z | • : | / | • : |
| g/7 | • : | q | • : | , | • : | | |
| h/8 | • : | r | • : | ; | • : | | |
| i/9 | • : | s | • : | , | • : | | |
| j/10 | • : | t | • : | : | • : | | |

Challenge 10.2 — Dots Everywhere.

ન ક ર ન ફ ટ ન ક ફ ન
ા ક ર ન ફ ન



Solution 10.2 — Dots Everywhere. Awesome work, you decoded Braille!

10.2.2 Grade 2

Practically speaking, encoding Braille needs lots of physical space. In order to try to minimize spaces, Grade 2 comes with an addition of abbreviations and contractions:

| | | | | | |
|-----|-----|------|------|---------|----|
| and | ••• | ing | •• | ea | •• |
| ar | •• | into | •••• | bb/be | •• |
| ble | ••• | of | ••• | cc/con | •• |
| ch | •• | ou | ••• | dd/dis | •• |
| ed | ••• | ow | ••• | ff/to | •• |
| en | •• | sh | ••• | gg/were | •• |
| er | ••• | th | ••• | his | •• |
| for | ••• | the | ••• | by/was | •• |
| gh | •• | wh | ••• | com | •• |
| in | •• | with | ••• | st | •• |

Challenge 10.3 — Less Dots. Longer Words.

Solution 10.3 — Less Dots, Longer Words. In this way lot of space has been saved.

10.2.3 Grade 3

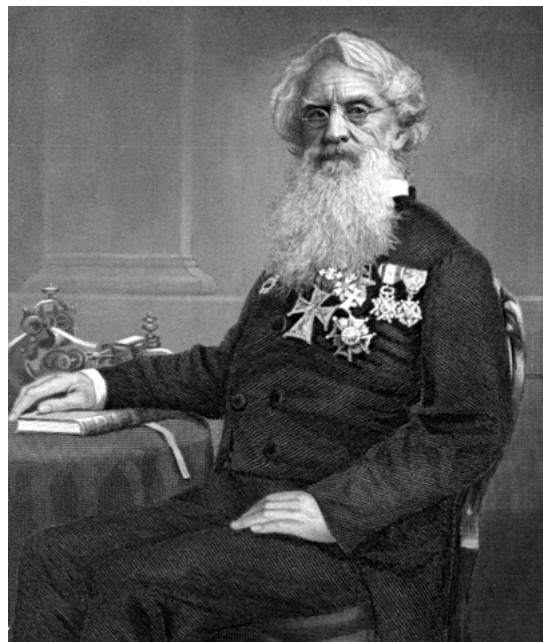
More sophisticated contractions have been developed for every language. There are no official Grade 3 encoding, but can be customized depending on the behaviors. The main idea is to use the first few letters to represent a whole word. What follows is an example of this process:

| | | | | | |
|-----------------|--|-----------------|--|---------------|--|
| b for but | | m for more | | e for every | |
| ab for about | | abv for above | | ei for either | |
| bet for between | | bey for beyond | | ll for little | |
| td for today | | tm for tomorrow | | qk for quick | |

Advanced Braille symbols, and combinations of them, are able to represent also the math world, the currencies and are able to distinguish between capital letters, digits and punctuation.

10.3 Morse Code (1835 - 1840)

The Morse code is one of the most common example of encoding, composed primary by dots and lines. This encoding, originally studied by Samuel Morse in 1835, but realized by his collaborator Alfred Vail starting from 1837, was used to encode information that were transferred by electrical telegraph systems.



During the following years the originally encoding schema, that was able to encode and decode the whole English alphabet, has been update to been able to manage digits and symbols coming from different languages. Morse code began to be used extensively till 1900s for early radio communication and nowadays is mostly use for aviation and marine purposes. The following is the International Morse Code (ITU):

| | | | | | | | |
|---|------|---|------|---|-------|---|--------|
| a | . | n | -. | 0 | — |) | -.-. |
| b | -... | o | — | 1 | .— | (| -.- |
| c | -.-. | p | -. | 2 | ..— | : | —... |
| d | -.. | q | -.- | 3 | ...— | , | -..- |
| e | . | r | -. | 4 | | = | -.... |
| f | ..-. | s | ... | 5 | | ! | -.-.- |
| g | -. | t | - | 6 | -.... | . | -.-.- |
| h | | u | ..- | 7 | -... | - | -..... |
| i | .. | v | ...- | 8 | —.. | + | -.-. |
| j | — | w | -. | 9 | —-. | " | -...- |
| k | -.- | x | -..- | & | -...- | ? | -.-.. |
| l | -.. | y | -.- | ' | .—-. | / | -.-. |
| m | - | z | -.. | @ | -.-. | | |

Challenge 10.4 Decode: . -.- . - - . - . . — .. - — - . - . . - - . — - . .

Solution 10.4 Decoded: example of morse code

10.4 ASCII (1960 - 1967)

The American Standard Code for Information Interchange (ASCII), is a binary character encoding standard for electronic communication based on 7-bits, designed in 1960s for teleprinters, telegraphy and some computing. This means that each symbol is unequivocally identified by a sequence of 1s and 0s of length 7. ASCII codes, nowadays, represent text in computers, telecommunications equipment, and other devices. Being a binary encoding system it can also be translated in other numeric bases and usually, for challenge purposes, can be found written in base 2, base 8, base 10 and base 16. The following conversion table will summarize all this cases:

| Bin | Oct | Dec | Hex | Symbol | Bin | Oct | Dec | Hex | Symbol |
|---------|-----|-----|-----|--------|---------|-----|-----|-----|--------|
| 0000000 | 000 | 0 | 00 | NUL | 0100000 | 040 | 32 | 20 | Space |
| 0000001 | 001 | 1 | 01 | SOH | 0100001 | 041 | 33 | 21 | ! |
| 0000010 | 002 | 2 | 02 | STX | 0100010 | 042 | 34 | 22 | " |
| 0000011 | 003 | 3 | 03 | ETX | 0100011 | 043 | 35 | 23 | # |
| 0000100 | 004 | 4 | 04 | EOT | 0100100 | 044 | 36 | 24 | \$ |
| 0000101 | 005 | 5 | 05 | ENQ | 0100101 | 045 | 37 | 25 | % |
| 0000110 | 006 | 6 | 06 | ACK | 0100110 | 046 | 38 | 26 | & |
| 0000111 | 007 | 7 | 07 | BEL | 0100111 | 047 | 39 | 27 | , |
| 0001000 | 010 | 8 | 08 | BS | 0101000 | 050 | 40 | 28 | (|
| 0001001 | 011 | 9 | 09 | TAB | 0101001 | 051 | 41 | 29 |) |
| 0001010 | 012 | 10 | 0A | LF | 0101010 | 052 | 42 | 2A | * |
| 0001011 | 013 | 11 | 0B | VT | 0101011 | 053 | 43 | 2B | + |
| 0001100 | 014 | 12 | 0C | FF | 0101100 | 054 | 44 | 2C | , |
| 0001101 | 015 | 13 | 0D | CR | 0101101 | 055 | 45 | 2D | - |
| 0001110 | 016 | 14 | 0E | SO | 0101110 | 056 | 46 | 2E | . |
| 0001111 | 017 | 15 | 0F | SI | 0101111 | 057 | 47 | 2F | / |
| 0010000 | 020 | 16 | 10 | DLE | 0110000 | 060 | 48 | 30 | 0 |
| 0010001 | 021 | 17 | 11 | DC1 | 0110001 | 061 | 49 | 31 | 1 |
| 0010010 | 022 | 18 | 12 | DC2 | 0110010 | 062 | 50 | 32 | 2 |
| 0010011 | 023 | 19 | 13 | DC3 | 0110011 | 063 | 51 | 33 | 3 |
| 0010100 | 024 | 20 | 14 | DC4 | 0110100 | 064 | 52 | 34 | 4 |
| 0010101 | 025 | 21 | 15 | NAK | 0110101 | 065 | 53 | 35 | 5 |
| 0010110 | 026 | 22 | 16 | SYN | 0110110 | 066 | 54 | 36 | 6 |
| 0010111 | 027 | 23 | 17 | ETB | 0110111 | 067 | 55 | 37 | 7 |
| 0011000 | 030 | 24 | 18 | CAN | 0111000 | 070 | 56 | 38 | 8 |
| 0011001 | 031 | 25 | 19 | EM | 0111001 | 071 | 57 | 39 | 9 |
| 0011010 | 032 | 26 | 1A | SUB | 0111010 | 072 | 58 | 3A | : |
| 0011011 | 033 | 27 | 1B | ESC | 0111011 | 073 | 59 | 3B | ; |
| 0011100 | 034 | 28 | 1C | FS | 0111100 | 074 | 60 | 3C | < |
| 0011101 | 035 | 29 | 1D | GS | 0111101 | 075 | 61 | 3D | = |
| 0011110 | 036 | 30 | 1E | RS | 0111110 | 076 | 62 | 3E | > |
| 0011111 | 037 | 31 | 1F | US | 0111111 | 077 | 63 | 3F | ? |

| Bin | Oct | Dec | Hex | Symbol | Bin | Oct | Dec | Hex | Symbol |
|---------|-----|-----|-----|--------|---------|-----|-----|-----|--------|
| 1000000 | 080 | 64 | 40 | @ | 1100000 | 120 | 96 | 60 | ' |
| 1000001 | 081 | 65 | 41 | A | 1100001 | 121 | 97 | 61 | a |
| 1000010 | 082 | 66 | 42 | B | 1100010 | 122 | 98 | 62 | b |
| 1000011 | 083 | 67 | 43 | C | 1100011 | 123 | 99 | 63 | c |
| 1000100 | 084 | 68 | 44 | D | 1100100 | 124 | 100 | 64 | d |
| 1000101 | 085 | 69 | 45 | E | 1100101 | 125 | 101 | 65 | e |
| 1000110 | 086 | 70 | 46 | F | 1100110 | 126 | 102 | 66 | f |
| 1000111 | 087 | 71 | 47 | G | 1100111 | 127 | 103 | 67 | g |
| 1001000 | 090 | 72 | 48 | H | 1101000 | 130 | 104 | 68 | h |
| 1001001 | 091 | 73 | 49 | I | 1101001 | 131 | 105 | 69 | i |
| 1001010 | 092 | 74 | 4A | J | 1101010 | 132 | 106 | 6A | j |
| 1001011 | 093 | 75 | 4B | K | 1101011 | 133 | 107 | 6B | k |
| 1001100 | 094 | 76 | 4C | L | 1101100 | 134 | 108 | 6C | l |
| 1001101 | 095 | 77 | 4D | M | 1101101 | 135 | 109 | 6D | m |
| 1001110 | 096 | 78 | 4E | N | 1101110 | 136 | 110 | 6E | n |
| 1001111 | 097 | 79 | 4F | O | 1101111 | 137 | 111 | 6F | o |
| 1010000 | 100 | 80 | 50 | P | 1110000 | 140 | 112 | 70 | p |
| 1010001 | 101 | 81 | 51 | Q | 1110001 | 141 | 113 | 71 | q |
| 1010010 | 102 | 82 | 52 | R | 1110010 | 142 | 114 | 72 | r |
| 1010011 | 103 | 83 | 53 | S | 1110011 | 143 | 115 | 73 | s |
| 1010100 | 104 | 84 | 54 | T | 1110100 | 144 | 116 | 74 | t |
| 1010101 | 105 | 85 | 55 | U | 1110101 | 145 | 117 | 75 | u |
| 1010110 | 106 | 86 | 56 | V | 1110110 | 146 | 118 | 76 | v |
| 1010111 | 107 | 87 | 57 | W | 1110111 | 147 | 119 | 77 | w |
| 1011000 | 110 | 88 | 58 | X | 1111000 | 150 | 120 | 78 | x |
| 1011001 | 111 | 89 | 59 | Y | 1111001 | 151 | 121 | 79 | y |
| 1011010 | 112 | 90 | 5A | Z | 1111010 | 152 | 122 | 7A | z |
| 1011011 | 113 | 91 | 5B | [| 1111011 | 153 | 123 | 7B | { |
| 1011100 | 114 | 92 | 5C | \ | 1111100 | 154 | 124 | 7C | |
| 1011101 | 115 | 93 | 5D |] | 1111101 | 155 | 125 | 7D | } |
| 1011110 | 116 | 94 | 5E | ^ | 1111110 | 156 | 126 | 7E | ~ |
| 1011111 | 117 | 95 | 5F | - | 1111111 | 157 | 127 | 3F | DEL |

As can be noticed, ASCII reserves the first 32 codes for control characters, that are not printable but used for devices control (like printers, keyboards and scanners) or to provide meta-data about data streams. The remaining symbols are all printable characters representing uppercase and lower case letters, digits and most common symbols.

Challenge 10.5 Decode: something in bin

Solution 10.5 Decoded: example of morse code

Challenge 10.6 Decode: something in hex

Solution 10.6 Decoded: example of morse code**10.4.1 Extended ASCII (1980s - 2000s)**

ASCII encoding is limited to 128 symbols. In order to represent symbols belonging to different languages, more modern ones like ™, ©, and scientific symbols, the 7-bit encoding has been extended to 8-bit so that at most 256 symbols could be represented. The implementation of the extended symbols is free and can vary between language and purpose. For example, the ISO 8859 maintain a standard with 16 different tables of extended symbols that cover lots of different characters. The mostly used table is ISO 8859-15, also called "Latin-9" that covers Western European languages such as English, German, Italian, Spanish, Irish, French, Danish, Finnish, Dutch and others. It also covers some other non-European languages like Indonesian, Afrikaans and Swahili:

| Bin | Oct | Dec | Hex | Symbol | Bin | Oct | Dec | Hex | Symbol |
|----------|-----|-----|-----|--------|----------|-----|-----|-----|--------|
| 10100000 | 240 | 160 | A0 | NBSP | 11000000 | 280 | 192 | C0 | À |
| 10100001 | 241 | 161 | A1 | à | 11000001 | 281 | 193 | C1 | Á |
| 10100010 | 242 | 162 | A2 | ć | 11000010 | 282 | 194 | C2 | Â |
| 10100011 | 243 | 163 | A3 | č | 11000011 | 283 | 195 | C3 | Ã |
| 10100100 | 244 | 164 | A4 | | 11000100 | 284 | 196 | C4 | Ä |
| 10100101 | 245 | 165 | A5 | ě | 11000101 | 285 | 197 | C5 | Å |
| 10100110 | 246 | 166 | A6 | | 11000110 | 286 | 198 | C6 | Æ |
| 10100111 | 247 | 167 | A7 | ȝ | 11000111 | 287 | 199 | C7 | Ҫ |
| 10101000 | 250 | 168 | A8 | | 11001000 | 290 | 200 | C8 | È |
| 10101001 | 251 | 169 | A9 | © | 11001001 | 291 | 201 | C9 | É |
| 10101010 | 252 | 170 | AA | ƒ | 11001010 | 292 | 202 | CA | Ê |
| 10101011 | 253 | 171 | AB | ń | 11001011 | 293 | 203 | CB | Ë |
| 10101100 | 254 | 172 | AC | њ | 11001100 | 294 | 204 | CC | Ì |
| 10101101 | 255 | 173 | AD | SHY | 11001101 | 295 | 205 | CD | Í |
| 10101110 | 256 | 174 | AE | ő | 11001110 | 296 | 206 | CE | Î |
| 10101111 | 257 | 175 | AF | ŕ | 11001111 | 297 | 207 | CF | Ï |
| 10110000 | 260 | 176 | B0 | ř | 11010000 | 300 | 208 | D0 | Đ |
| 10010001 | 261 | 177 | B1 | ś | 11010001 | 301 | 209 | D1 | Ñ |
| 10110010 | 262 | 178 | B2 | š | 11010010 | 302 | 210 | D2 | Ӯ |
| 10110011 | 263 | 179 | B3 | § | 11010011 | 303 | 211 | D3 | Ӱ |
| 10110100 | 264 | 180 | B4 | | 11010100 | 304 | 212 | D4 | Ӯ |
| 10110101 | 265 | 181 | B5 | ڻ | 11010101 | 305 | 213 | D5 | Ӱ |
| 10110110 | 266 | 182 | B6 | ڻ | 11010110 | 306 | 214 | D6 | Ӱ |
| 10110111 | 267 | 183 | B7 | ڻ | 11010111 | 307 | 215 | D7 | Ӯ |
| 10111000 | 270 | 184 | B8 | | 11011000 | 310 | 216 | D8 | Ӯ |
| 10111001 | 271 | 185 | B9 | ܶ | 11011001 | 311 | 217 | D9 | ܶ |
| 10111010 | 272 | 186 | BA | ܷ | 11011010 | 312 | 218 | DA | ܶ |
| 10111011 | 273 | 187 | BB | ܷ | 11011011 | 313 | 219 | DB | ܶ |
| 10111100 | 274 | 188 | BC | | 11011100 | 314 | 220 | DC | ܶ |
| 10111101 | 275 | 189 | BD | | 11011101 | 315 | 221 | DD | ܶ |
| 10111110 | 276 | 190 | BE | | 11011110 | 316 | 222 | DE | ܶ |
| 10111111 | 277 | 191 | BF | ܶ | 11011111 | 317 | 223 | DF | ܶ |

| Bin | Oct | Dec | Hex | Symbol | Bin | Oct | Dec | Hex | Symbol |
|----------|-----|-----|-----|--------|----------|-----|-----|-----|--------|
| 11100000 | 320 | 224 | E0 | à | 11110000 | 340 | 240 | F0 | ð |
| 11100001 | 321 | 225 | E1 | á | 11110001 | 341 | 241 | F1 | ñ |
| 11100010 | 322 | 226 | E2 | â | 11110010 | 342 | 242 | F2 | ò |
| 11100011 | 323 | 227 | E3 | ã | 11110011 | 343 | 243 | F3 | ó |
| 11100100 | 324 | 228 | E4 | ä | 11110100 | 344 | 244 | F4 | ô |
| 11100101 | 325 | 229 | E5 | å | 11110101 | 345 | 245 | F5 | õ |
| 11100110 | 326 | 230 | E6 | æ | 11110110 | 346 | 246 | F6 | ö |
| 11100111 | 327 | 231 | E7 | ç | 11110111 | 347 | 247 | F7 | æ |
| 11101000 | 330 | 232 | E8 | è | 11111000 | 350 | 248 | F8 | ø |
| 11101001 | 331 | 233 | E9 | é | 11111001 | 351 | 249 | F9 | ù |
| 11101010 | 332 | 234 | EA | ê | 11111010 | 352 | 250 | FA | ú |
| 11101011 | 333 | 235 | EB | ë | 11111011 | 353 | 251 | FB | û |
| 11101100 | 334 | 236 | EC | ì | 11111100 | 354 | 252 | FC | ü |
| 11101101 | 335 | 237 | ED | í | 11111101 | 355 | 253 | FD | ý |
| 11101110 | 336 | 238 | EE | î | 11111110 | 356 | 254 | FE | þ |
| 11101111 | 337 | 239 | EF | ï | 11111111 | 357 | 255 | FF | þ |

Challenge 10.7 Decode: something in oct

Solution 10.7 Decoded: example of morse code

Challenge 10.8 Decode: something in hex

Solution 10.8 Decoded: example of morse code

10.5 DTMF (1963)

Since the invention of the telegraph in 1837 by Cooke, and next of the telephone in 1849 by Meucci, the main problem was to define an encoding system able to ensure communication between 2 communication devices. One of the first attempt, used since the second half of the 20th century was based on Pulse Dialing systems where a local loop circuit is interrupted according to a defined coding system for each signal transmitted. This system was invented in 1892 by Almon Brown Strowger and was able to send only digits. An example of this system can be found in rotary phones.

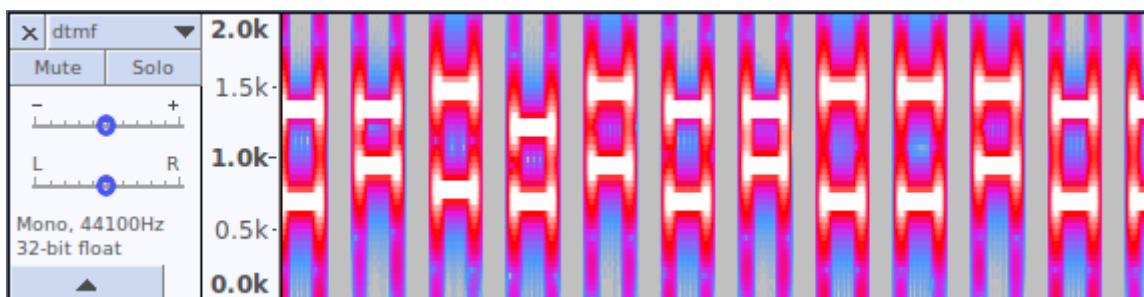
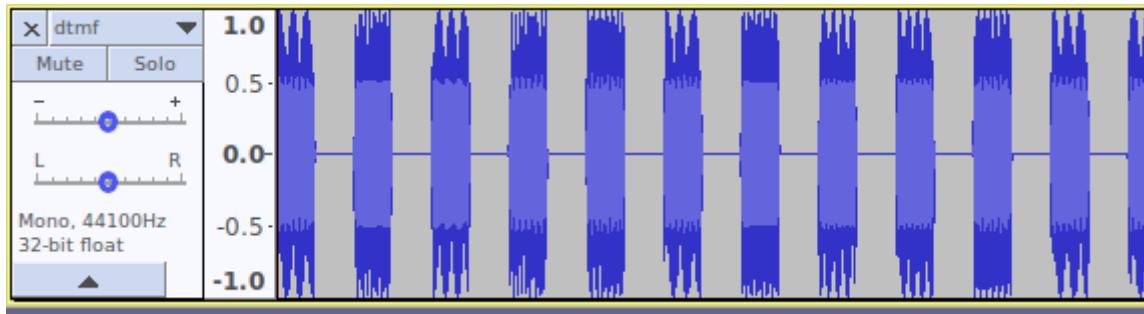


In 1963 engineers at Bell developed a new kind of technology "called Dual-Tone multi-frequency signaling (DTMF)" whose become the new communication standard. The new telephones had no rotors, but a pushable 16-keys keypad, friendly called Touch-Tone. This new system allowed to make long-range calls and for first time the communication channel was not restricted to copper cables but extended to microwave bridges and satellites.

The keypad hosts digits from 0 to 9, letters from A to D and two special characters: * and #. Each button produces a different DTMF signal, composed by a low frequency and an high frequency component. For example the key A produces a 697 Hz low tone and a 1633 Hz high tone. The next table will summarize the tone frequencies of each button:

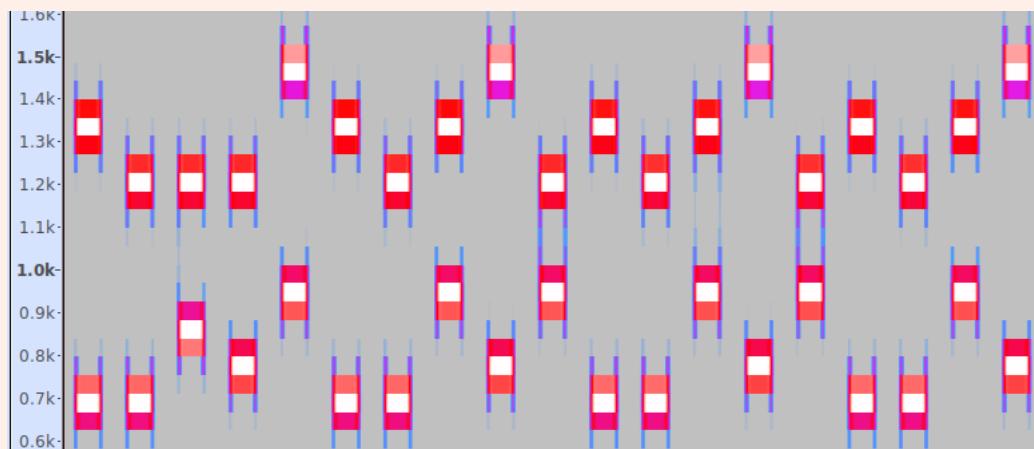
| | 1209 Hz | 1336 Hz | 1477 Hz | 1633 Hz |
|--------|---------|---------|---------|---------|
| 697 Hz | 1 | 2 | 3 | A |
| 770 Hz | 4 | 5 | 6 | B |
| 852 Hz | 7 | 8 | 9 | C |
| 941 Hz | * | 0 | # | D |

The decoding system was provided by filter banks first, and by digital signal processing then, usually using the Goertzel algorithm. Frequencies of tones of a DTMF communication can be visually seen by inspecting the spectrogram of the signal itself:



Decoding the signal graphically is quite simple, just use the encoding table. For example the signal in the picture can be decoded as "2064#2033#2...". DTMF signals are used in steganography to cover binary messages and hex messages even if with some limitations (E and F are not reproducible).

Challenge 10.9 Can you decode this DTMF signal?



Solution 10.9 Decoded: 2174#2106*2106*2106

10.6 BaseN Encodings (1993)

10.6.1 Base64

Base64 is an encoding schema that is used to represent binary data contained in an ASCII string in another format by translating it into a string composed by a sequence of symbols coming from an alphabet of cardinality 64. Each symbol of the generated base64 string represent 6 bits of the original data, so that every 3 original symbols are represented by 4 base64 symbols. This means that length of the base64 output is always longer with respect to the length of the original message. Due to its ductility and capacity to transform binary data, it is mainly used to carry information of different sources (images, videos, media, audio...) through the World Wide Web.

The encoding procedure is the following (if starting from binary, skip first 2 points):

- Convert each input symbol to its ASCII decimal representation;
- Convert each decimal into its 8-bit representation (octets);
- Group bits into buckets of size 24
- Threat every 6 bits of each group as sextets and convert back to decimal representation
- Following the encoding table, replace each number with the corresponding symbol
- If the last bucket contains less than 24 bits add one or two padding symbols (=) to fill the bucket

The encoding table is the following:

| Dec | Char | Dec | Char | Dec | Char | Dec | Char |
|-----|------|-----|------|-----|------|-----|------|
| 0 | A | 16 | Q | 32 | g | 48 | w |
| 1 | B | 17 | R | 33 | h | 49 | x |
| 2 | C | 18 | S | 34 | i | 50 | y |
| 3 | D | 19 | T | 35 | j | 51 | z |
| 4 | E | 20 | U | 36 | k | 52 | 0 |
| 5 | F | 21 | V | 37 | l | 53 | 1 |
| 6 | G | 22 | W | 38 | m | 54 | 2 |
| 7 | H | 23 | X | 39 | n | 55 | 3 |
| 8 | I | 24 | Y | 40 | o | 56 | 4 |
| 9 | J | 25 | Z | 41 | p | 57 | 5 |
| 10 | K | 26 | a | 42 | q | 58 | 6 |
| 11 | L | 27 | b | 43 | r | 59 | 7 |
| 12 | M | 28 | c | 44 | s | 60 | 8 |
| 13 | N | 29 | d | 45 | t | 61 | 9 |
| 14 | O | 30 | e | 46 | u | 62 | + |
| 15 | P | 31 | f | 47 | v | 63 | / |

Let's do an example: imagine we want to encode the word "strange" in base64. We have that:

| Word | s | t | r | a | n | g | e |
|--------|----------|----------|----------|----------|----------|----------|----------|
| ASCII | 115 | 116 | 114 | 97 | 110 | 103 | 101 |
| Binary | 01110011 | 01110100 | 01110010 | 01100001 | 01101110 | 01100111 | 01100101 |

Now the goal is to split into buckets of 24 bits:

| | | |
|--------------------------|--------------------------|----------------------------|
| 011100110111010001110010 | 011000010110111001100111 | 01100101xxxxxxxxxxxxxxxxxx |
|--------------------------|--------------------------|----------------------------|

It can be noticed that the last bucket contains only 8 bits and so will be filled in order to reach the 24 elements. Each bucket is now transformed according to the previous algorithm and to the encoding table. For each bucket we have:

| | | | | | | | | | |
|---------|--------|--------|--------|--------|---------|--------|--------|--------|--------|
| Sextets | 011100 | 110111 | 010001 | 110010 | Sextets | 011000 | 010110 | 111001 | 100111 |
| ASCII | 28 | 55 | 17 | 50 | ASCII | 24 | 22 | 57 | 39 |
| Base64 | c | 3 | R | y | Base64 | Y | W | 5 | n |

| | | | | |
|------------|--------|--------|--------|--------|
| Sextets | 011001 | 01xxxx | xxxxxx | xxxxxx |
| normalized | 011001 | 010000 | xxxxxx | xxxxxx |
| ASCII | 25 | 16 | ? | ? |
| Base64 | Z | Q | = | = |

All together we obtain that $\text{base64}(\text{"strange"}) = \text{"c3RyYW5nZQ=="}.$ The decoding process is extremely easy and will not be showed, just follow steps backward. There are some variants for the symbols at position 62 and 63 of the encoding table for different implementations that usually replace the '+' and the '/' symbols to others like '-' , '_' , '~' , '.' and ':' . Some implementations make the padding symbol optional.

Challenge 10.10 — superego. What have Freud to say?

FrQXQgdGhIHRpbWUgYXQgd2hpY2ggdGhlIE9lZGlwdXMgY2
9tcGxleCBnaXZlcYBwbGFjZSB0byB0aGUgc3VwZXItZWdvIH
RoZXkgYXJlIHNVbWV0aGluZyBxdWl0ZSBtYWduaWZpY2VudDs=

Solution 10.10 — superego. Such egocentric patients

At the time at which the Oedipus complex gives place to the super-ego they are something quite magnificent

10.6.2 Base32

Base64 has two huge limitation. The biggest one is that it cannot be used in case-insensitive systems due to the fact that the encryption table contains both a-z and A-Z symbols. Base64 cannot also be used to represent file names because the symbol '/' is included in the encryption table and for unix systems represents the path separator. A slightly simple variant of Base64 is Base32.

It is a very similar encoding scheme that encodes buckets of 5 bits instead of 6 bits producing an output, on average, 20% longer with respect to base64. The encoding algorithm is the following:

- Convert each input symbol to its ASCII decimal representation;
- Convert each decimal into its 8-bit representation (octets);
- Group bits into buckets of size 40
- Treat every 5 bits of each group as quintets and convert back to decimal representation
- Following the encoding table, replace each number with the corresponding symbol
- If the last bucket contains less than 40 bits add one or two padding symbols (=) to fill the bucket

As for base64 the most used encoding table for base32 is the one coming from RFC4648:

| Dec | Char | Dec | Char | Dec | Char | Dec | Char |
|-----|------|-----|------|-----|------|-----|------|
| 0 | A | 8 | I | 16 | Q | 24 | Y |
| 1 | B | 9 | J | 17 | R | 25 | Z |
| 2 | C | 10 | K | 18 | S | 26 | 2 |
| 3 | D | 11 | L | 19 | T | 27 | 3 |
| 4 | E | 12 | M | 20 | U | 28 | 4 |
| 5 | F | 13 | N | 21 | V | 29 | 5 |
| 6 | G | 14 | O | 22 | W | 30 | 6 |
| 7 | H | 15 | P | 23 | X | 31 | 7 |

RFC4648 suggests also another encryption table called "base32hex" with the following motivation:

"One property with this alphabet, which the base64 and base32 alphabets lack, is that encoded data maintains its sort order when the encoded data is compared bit-wise."

The new table is the following:

| Dec | Char | Dec | Char | Dec | Char | Dec | Char |
|-----|------|-----|------|-----|------|-----|------|
| 0 | 0 | 8 | 8 | 16 | G | 24 | O |
| 1 | 1 | 9 | 9 | 17 | H | 25 | P |
| 2 | 2 | 10 | A | 18 | I | 26 | Q |
| 3 | 3 | 11 | B | 19 | J | 27 | R |
| 4 | 4 | 12 | C | 20 | K | 28 | S |
| 5 | 5 | 13 | D | 21 | L | 29 | T |
| 6 | 6 | 14 | E | 22 | M | 30 | U |
| 7 | 7 | 15 | F | 23 | N | 31 | V |

Let's do the same example did before but this time with base32 encoding:

| Word | s | t | r | a | n | g | e |
|--------|----------|----------|----------|----------|----------|----------|----------|
| ASCII | 115 | 116 | 114 | 97 | 110 | 103 | 101 |
| Binary | 01110011 | 01110100 | 01110010 | 01100001 | 01101110 | 01100111 | 01100101 |

Now the goal is to split into buckets of 40 bits:

| | |
|--|--|
| 0111001101110100011100100110000101101110 | 0110011101100101xxxxxxxxxxxxxxxxxxxxxx |
|--|--|

It can be noticed that the last bucket contains only 8 bits and so will be filled in order to reach the 24 elements. Each bucket is now transformed according to the previous algorithm and to the encoding table. For each bucket we have:

| | | | | | | | | |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|
| Quintects | 01110 | 01101 | 11010 | 00111 | 00100 | 11000 | 01011 | 01110 |
| ASCII | 14 | 13 | 26 | 7 | 4 | 24 | 11 | 14 |
| Base32 | O | N | 2 | H | E | Y | L | O |

| | | | | | | | | |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Quintects | 01100 | 11101 | 10010 | 1xxxx | xxxxx | xxxxx | xxxxx | xxxxx |
| Normalized | 01100 | 11101 | 10010 | 10000 | xxxxx | xxxxx | xxxxx | xxxxx |
| ASCII | 12 | 29 | 18 | 16 | ? | ? | ? | ? |
| Base32 | M | 5 | S | Q | = | = | = | = |

All together we obtain that $\text{base32("strange")} = \text{"ON2HEYLOM5SQ===="}$. It is difficult to distinguish between base32 and base64 in some cases, but can be noticed that:

- Base32 does not contain lower case characters;
- Base32 contains only letters and numbers and not other symbols (excluding the padding one);
- Base64 has at most 2 padding symbols. Base32 can have up to 6 padding symbols.

Challenge 10.11 — ego. Tell me Sigmund!

KRUGKIDFM5XSA2LTEBXG65BAONUGC4TQNR4SA43FOBQXEYLU
MVSCAZTSN5WSA5DIMUQGSZB3EBUXI4ZANRXXOZLSEBYG64TU
NFXW4IDNMVZGOZLTEBUW45DPEBUXILROFYXCAQTVOQQHI2DF
EBZGK4DSMVZXGZLEEBWWK4THMVZSA2LOORXSA5DIMUQGSZBA
MFZSA53FNRWCYIDBNZSCA2LTEBWWK4TFNR4SAYJAQBQXE5BA
N5TCA2LUFYQFI2DFEBZGK4DSMVZXGZLEEBUXGIDPNZWHSIDD
OV2CA33GMYQHG2DBOJYGY6JAMZZG63JAORUGKIDFM5XSAYTZ
EB2GQZJAOJSXG2LTORQW4Y3FOMQG6ZRAOJSXA4TFONZWS33O
HMQGS5BAMNQW4IDDN5WW25LONFRWC5DFEB3WS5DIEB2GQZJA
MVTW6IDUNBZG65LHNAQHI2DFEBUWI==

Solution 10.11 — ego. So brilliant!

The ego is not sharply separated from the id; its lower portion merges into it.... But the repressed merges into the id as well, and is merely a part of it. The repressed is only cut off sharply from the ego by the resistances of repression; it can communicate with the ego through the id

Cryptography



- | | | |
|-----------|----------------------------------|-------------|
| 11 | Classical Cipher | *(.5pc).75 |
| 11.1 | Rot-n | |
| 11.2 | Substitution Ciphers | |
| 11.3 | Polygraphic Substitution Ciphers | |
| 11.4 | Transposition Ciphers | |
| | | |
| 12 | Modern Cryptography | *(.5pc).89 |
| 12.1 | RSA (1977) | |
| 12.2 | Rabin Cryptosystem (1979) | |
| 12.3 | Block Ciphers (1981) | |
| 12.4 | Stream Ciphers (1987) | |
| | | |
| 13 | Cryptanalysis | *(.5pc).101 |
| 13.1 | Frequency Analysis | |

11. Classical Cipher

11.1 Rot-n

One of the widest used cipher for beginning cryptography challenges, rot-n refers to a family of ciphers based on alphabet rotation. For the Rot-n ciphers family, when we talk about rotation, we mean that the second half of the alphabet represent the encryption of the first half of the alphabet. The set of symbols involved into the alphabet may differ according to different languages, and the complexity can be improved by adding also special symbols and digits to the alphabet. The most notable property of this kind of ciphers is that the encryption algorithm is the same as the decryption one:

$$Rot_n(Rot_n(x)) = x \quad (11.1)$$

Where Rot-n(x) is defined as:

$$Rot_n(x) = (x + n) \pmod{2n} \quad (11.2)$$

11.1.1 Rot-5

When the alphabet is only composed by digits from 0 to 9, the rotation cipher is called Rot-5. In this case the number 0 is encoded by the number 5, the number 1 by the number 6 and so on according to the following table:

| | | | | |
|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 |
| 5 | 6 | 7 | 8 | 9 |

Challenge 11.1 Decode: 297584

Solution 11.1 Decoded: 742039

```
static String rot5(String text) {
    StringBuilder sb = new StringBuilder();
    for (int i = 0; i < text.length(); i++) {
        sb.append((char) ((text.charAt(i) - '0' + 5) % 10 + '0'));
    }
    return sb.toString();
}
```

11.1.2 Rot-13

When the alphabet is only composed by english letters from a to z, the rotation cipher is called Rot-13. In this case the letter 'a' is encoded by the letter 'n', the letter 'b' by the letter 'o' and so on according to the following table:

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m |
| n | o | p | q | r | s | t | u | v | w | x | y | z |

Challenge 11.2 Decode: guvf zrffntr unf orra rapelcgrq

Solution 11.2 Decoded: this message has been encrypted

```
static String rot13(String text) {
    StringBuilder sb = new StringBuilder();
    for (int i = 0; i < text.length(); i++) {
        sb.append((char) ((text.charAt(i) - 'a' + 13) % 26 + 'a'));
    }
    return sb.toString();
}
```

11.1.3 Rot-13.5

Combining the Rot-13 with the Rot-5 we obtain the Rot-13.5.

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | 0 | 1 | 2 | 3 | 4 |
| n | o | p | q | r | s | t | u | v | w | x | y | z | 5 | 6 | 7 | 8 | 9 |



Rot 13.5 should result a strange notation. This is a special case in which 2 alphabets are involved. Both of them are encrypted and decrypted separately. According to the Rot-n definition using a single alphabet would have produced the following encryption/decryption schema:

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r |
| s | t | u | v | w | x | y | z | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

And would have been named something like Rot-18.

Challenge 11.3 Decode: guvf zrffntr unf orra rapelcgrq jvgu ebg-68

Solution 11.3 Decoded: this message has been encrypted with rot-13

```
static String rot13_5(String text) {
    StringBuilder sb = new StringBuilder();
    for (int i = 0; i < text.length(); i++) {
        char current = text.charAt(i);
        if (current >= '0' && current <= '9') {
            sb.append((char) ((current - '0' + 5) % 10 + '0'));
        } else if (current >= 'a' && current <= 'z') {
            sb.append((char) ((current - 'a' + 13) % 26 + 'a'));
        }
    }
    return sb.toString();
}
```

11.1.4 Rot-47

Rot-47 takes in consideration 94 consecutive symbols of the ASCII table, from '!' at position 33 to ' ' at position 126 of the table. This alphabet will cover all English upper case and lower case letters, all digits and lots of printable symbols:

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|----|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ! | " | # | \$ | % | & | , | (|) | * | + | , | - | . | / | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| P | Q | R | S | T | U | V | W | X | Y | Z | [| \ |] | ^ | _ | ' | a | b | c | d | e | f | g |
| 9 | : | ; | < | = | > | ? | @ | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | |
| h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z | { | | } | ~ | |

Challenge 11.4 Decode: %9:D >6DD286 92D 366? 6?4CJAE65 H:E9 #@E\cf



Solution 11.4 Decoded: This message has been encrypted with Rot-47

```
static String rot47(String text) {
    StringBuilder sb = new StringBuilder();
    for (int i = 0; i < text.length(); i++) {
        sb.append((char) ((text.charAt(i) - '!' + 47) % 94 + '!'));
    }
    return sb.toString();
}
```

11.2 Substitution Ciphers

A substitution cipher is a way to replace single units of the plaintext with the respective symbols of the ciphertext, according to a fixed schema; the "units" may be single letters (the most common), pairs of letters, triplets of letters, mixtures of the above, and so forth. The receiver deciphers the text by performing the inverse substitution.

11.2.1 Atbash, Albam, Atbah

In old Sacred texts, and in particular in the Old Testament, can be found 3 examples of substitution ciphers. The first one, called Atbash, was invented by Hebrew population and it was a matter of overturning the alphabet so that letter 'a' is encrypted with letter 'z', letter 'b' with letter 'y' and so on:

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
| z | y | x | w | v | u | t | s | r | q | p | o | n | m | l | k | j | i | h | g | f | e | d | c | b | a |

```
static String atbash(String text) {
    StringBuilder sb = new StringBuilder();
    for (int i = 0; i < text.length(); i++) {
        sb.append((char) ((25 - (text.charAt(i) - 'a')) + 'a'));
    }
    return sb.toString();
}
```

The second one, called Albam, is another name to identify the Rot-13 cipher where the first half of the alphabet is encrypted by the second half. The last one is quite more complex. The Atbah substitution have to satisfy a numeric relation. The first nine symbols are substituted in a way in which the sum of plaintext symbol and the ciphertext symbol gives 10 (Considering a=1, b=2, ..., z=26). For the next nine symbols the same rule is valid, but the sum must give 28. For the last 8 symbols, again, the same rule is valid and the sum must give 45:

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
| i | h | g | f | e | d | c | b | a | r | q | p | o | n | m | l | k | j | z | y | x | w | v | u | t | s |

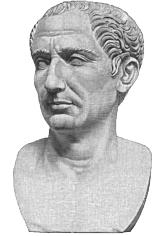
```
static String atbah(String text) {
    StringBuilder sb = new StringBuilder();
    for (int i = 0; i < text.length(); i++) {
        char current = text.charAt(i);
        if (current <= 'i') {
            sb.append((char) (8 + 2 * 'a' - current));
        } else if (current <= 'r') {
            sb.append((char) (8 + 2 * 'j' - current));
        } else {
            sb.append((char) (7 + 2 * 's' - current));
        }
    }
    return sb.toString();
}
```

It can be easily noticed that according to previous rules the encryption of symbols 'e' and 'n' are the symbols themselves. To avoid this some variants propose to encrypt symbol 'e' as 'n' and symbol 'n' as 'e':

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|----------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
| i | h | g | f | n | d | c | b | a | r | q | p | o | e | m | l | k | j | z | y | x | w | v | u | t | s |

11.2.2 Caesar Cipher (100 B.C. - 44 B.C.)

One of the widely known substitution cipher came to us from the "Da vita Caesarum", where the emperor was used to write secret messages replacing each letter by a letter some fixed number of positions down the alphabet. For example, with a shift of 3, letter 'a' becomes 'd', letter 'b' becomes 'e' and so on till letter 'x' returns to 'a', letter 'y' to 'b' and 'z' to 'c'. The encryption and the decryption of this simple cipher can be synthesize using modular arithmetic. For the English alphabet, which is composed by 26 letters, we have that, choosing an encryption key 'n' (the number of shifts), the encryption algorithm become:



$$E_n(x) = (x + n) \pmod{26} \quad (11.3)$$

And the decryption algorithm will be:

$$D_n(x) = (x - n) \pmod{26} \quad (11.4)$$



It is easy to notice that with a key of 13, we obtain the Rot-13 cipher (Recall from 3.2):

$$Rot_n(x) = (x + n) \pmod{2n}$$

with $n = 13$

$$Rot_n(x) = (x + n) \pmod{2 * 13}$$

$$Rot_n(x) = (x + n) \pmod{26} \longleftrightarrow E_n(x) = (x + n) \pmod{26}$$

Thus an example of encryption schema for key = 5 is:

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
| f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z | a | b | c | d | e |

Challenge 11.5 Decode: Ymnx mfx gjjs jshwduyji bnym Hjfxfw hnumjw

Solution 11.5 Decoded: This has been encrypted with Ceasar cipher

11.2.3 Affine Cipher

Affine cipher is generalization of the Caesar cipher, providing a bit more security but maintaining the flaws of a classical substitution cipher. As usual, letters of the English alphabet are represented by numbers from 0 to 25 but this time two different keys are used in order to encrypt and decrypt the message. The algorithms for encryption and decryption are the following:

$$E_{a,b}(x) = (a * x + b) \pmod{26} \quad (11.5)$$

$$D_{a,b}(x) = a^{-1} * (x - b) \pmod{26} \quad (11.6)$$

where a^{-1} is the modular multiplicative inverse of a modulus m (26 in this case). It can be easily computed by solving:

$$1 = a * a^{-1} \pmod{26} \quad (11.7)$$

The main problem is that the multiplicative inverse of a exists only if a and m are coprime. Choosing a that is not coprime with m makes the decryption impossible.

R With $a = 1$ and $b = n$ the affine cipher is actually the Caesar cipher:

$$E_n(x) = E_{1,b}(x)$$

$$(x + n) \pmod{m} = (1 * x + b) \pmod{m}$$

$(x + n) \pmod{m} = (x + b) \pmod{m}$ Since $b = n$ the both sides are equals. Since 1 is coprime with every possible modulus m , a is a valid key that makes the ciphertext decryptable.

We can now proceed showing an example. Suppose we want to encrypt the message "affinecipher" using keys $a = 3, b = 17$. First step is to encode the message as numbers so that "affinecipher" becomes 0 5 5 8 13 4 2 8 15 7 4 17. Then, for each number, the formula is applied resulting in 17 6 15 4 3 23 15 10 12 3 16. Converting back to chars we obtain "rggpedxpkmdq". The next table will summarize the process:

| plaintext | a | f | f | i | n | e | c | i | p | h | e | r |
|------------------------|----|----|----|----|----|----|----|----|----|----|----|----|
| numerical x | 0 | 5 | 5 | 8 | 13 | 4 | 2 | 8 | 15 | 7 | 4 | 17 |
| $3 * x + 17$ | 17 | 32 | 32 | 41 | 56 | 29 | 23 | 41 | 62 | 38 | 29 | 68 |
| $3 * x + 17 \pmod{26}$ | 17 | 6 | 6 | 15 | 4 | 3 | 23 | 15 | 10 | 12 | 3 | 16 |
| ciphertext | r | g | g | p | e | d | x | p | k | m | d | q |

Now to decrypt the ciphertext, as seen before, we have to compute a^{-1} . In our case 9 is the multiplicative inverse of 3 modulus 26. In fact $3 * 9 \pmod{26} = 1$. The decryption algorithm is so $D_{3,17} = 9 * (x - 17)$. Let's reconstruct the process table:

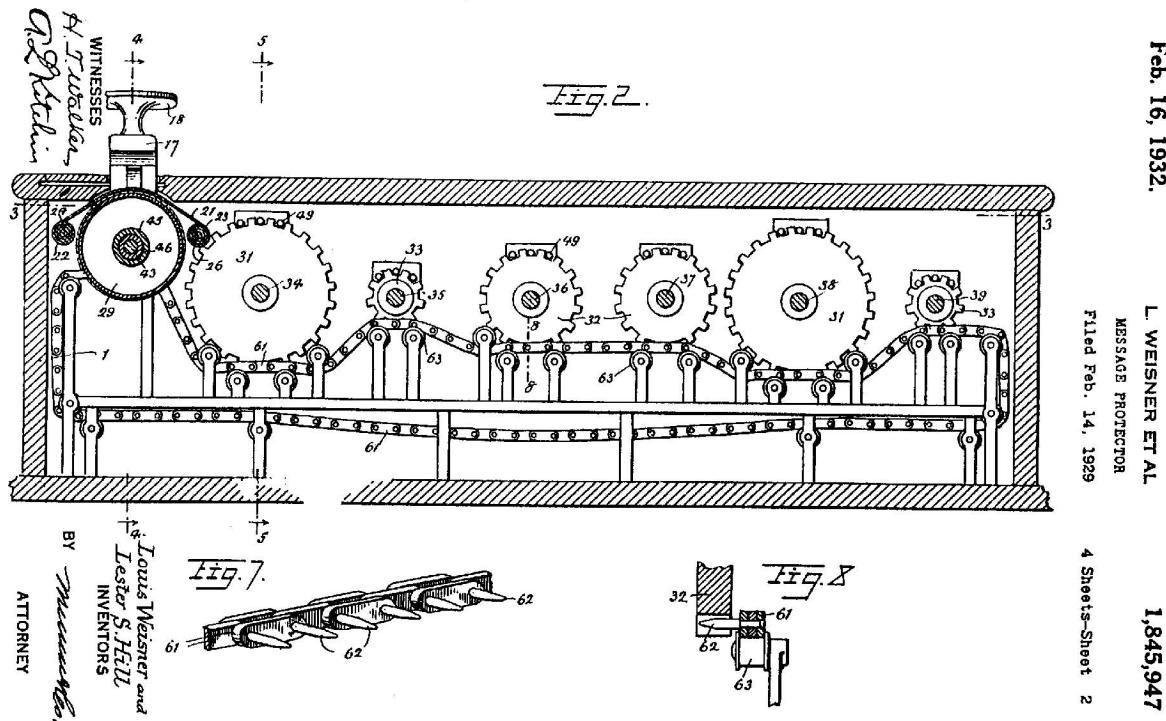
| ciphertext | r | g | g | p | e | d | x | p | k | m | d | q |
|--------------------------|----|-----|-----|-----|------|------|----|-----|-----|-----|------|----|
| numerical x | 17 | 6 | 6 | 15 | 4 | 3 | 23 | 15 | 10 | 12 | 3 | 16 |
| $9 * (x - 17)$ | 0 | -99 | -99 | -18 | -117 | -126 | 54 | -18 | -63 | -45 | -126 | -9 |
| $9 * (x - 17) \pmod{26}$ | 0 | 5 | 5 | 8 | 13 | 4 | 2 | 8 | 15 | 7 | 4 | 17 |
| plaintext | a | f | f | i | n | e | c | i | p | h | e | r |

11.3 Polygraphic Substitution Ciphers

what is, TODO

11.3.1 Hill's Cipher (1929)

The Hill's Cipher takes its name from its inventor Lester S. Hill that in 1929 developed an encryption system based on matrices. Each letter of the English alphabet is represented, with the usual convention, with numbers from 0 to 25. To encrypt the message, each block of n letters is transformed into a vector and then is multiplied by an invertible nxn matrix that represents the encryption key. To decrypt the message the same process is applied, but this time the vectors coming from the ciphertext are multiplied by the inverse of the key matrix. All operations are usually performed modulus 26 to maintain the same input notation.



Suppose we want to encrypt the word "cryptography" using the key "encodings". First we have to define the key matrix that in our case will be a 3×3 one ("encodings" has 9 letters and can be arranged into a 3×3 schema):

$$\begin{bmatrix} e & n & c \\ o & d & i \\ n & g & s \end{bmatrix} = \begin{bmatrix} 4 & 13 & 2 \\ 14 & 3 & 8 \\ 13 & 6 & 18 \end{bmatrix} \quad (11.8)$$

In the same way, the plaintext "cryptography" can be arranged into 4 vectors of size 3:

$$\text{cryptogrpahy} = \begin{bmatrix} c \\ r \\ y \end{bmatrix} \begin{bmatrix} p \\ t \\ o \end{bmatrix} \begin{bmatrix} g \\ r \\ a \end{bmatrix} \begin{bmatrix} p \\ h \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 17 \\ 24 \end{bmatrix} \begin{bmatrix} 15 \\ 19 \\ 14 \end{bmatrix} \begin{bmatrix} 6 \\ 17 \\ 0 \end{bmatrix} \begin{bmatrix} 15 \\ 7 \\ 24 \end{bmatrix} \quad (11.9)$$

Now we can start multiply each vector with the encryption key:

$$\begin{bmatrix} 4 & 13 & 2 \\ 14 & 3 & 8 \\ 13 & 6 & 18 \end{bmatrix} * \begin{bmatrix} 2 \\ 17 \\ 24 \end{bmatrix} \pmod{26} = \begin{bmatrix} 13 \\ 23 \\ 1 \end{bmatrix} = \begin{bmatrix} n \\ x \\ b \end{bmatrix} \quad (11.10)$$

$$\begin{bmatrix} 4 & 13 & 2 \\ 14 & 3 & 8 \\ 13 & 6 & 18 \end{bmatrix} * \begin{bmatrix} 15 \\ 19 \\ 14 \end{bmatrix} \pmod{26} = \begin{bmatrix} 23 \\ 15 \\ 15 \end{bmatrix} = \begin{bmatrix} x \\ p \\ p \end{bmatrix} \quad (11.11)$$

$$\begin{bmatrix} 4 & 13 & 2 \\ 14 & 3 & 8 \\ 13 & 6 & 18 \end{bmatrix} * \begin{bmatrix} 6 \\ 17 \\ 0 \end{bmatrix} \pmod{26} = \begin{bmatrix} 11 \\ 5 \\ 24 \end{bmatrix} = \begin{bmatrix} l \\ f \\ y \end{bmatrix} \quad (11.12)$$

$$\begin{bmatrix} 4 & 13 & 2 \\ 14 & 3 & 8 \\ 13 & 6 & 18 \end{bmatrix} * \begin{bmatrix} 15 \\ 7 \\ 24 \end{bmatrix} \pmod{26} = \begin{bmatrix} 17 \\ 7 \\ 24 \end{bmatrix} = \begin{bmatrix} r \\ h \\ y \end{bmatrix} \quad (11.13)$$

So the term "cryptography" has been encrypted as "xnbxpplfyryh". But now we have a huge problem. There is no way to recover the original message because the key matrix used is not invertible. A n -by- n matrix A is invertible if and only if there exists a n -by- n matrix B such that $AB = BA = I_n$ where I_n denotes the n -by- n identity matrix. This, sadly, is not our case. For our next example suppose that an unknown word has been encrypted using key "beer" obtaining the ciphertext "fics". This time the key matrix is invertible modulus 26:

$$\begin{bmatrix} b & e \\ e & r \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 4 & 17 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 4 \\ 4 & 17 \end{bmatrix}^{-1} = \begin{bmatrix} 17 & -4 \\ -4 & 1 \end{bmatrix} \pmod{26} = \begin{bmatrix} 17 & 22 \\ 22 & 1 \end{bmatrix} \quad (11.14)$$

We are now ready to decrypt the cipher and retrieve the original plaintext:

$$\begin{bmatrix} 17 & 22 \\ 22 & 1 \end{bmatrix} * \begin{bmatrix} f \\ i \end{bmatrix} \rightarrow \begin{bmatrix} 17 & 22 \\ 22 & 1 \end{bmatrix} * \begin{bmatrix} 5 \\ 8 \end{bmatrix} \pmod{26} = \begin{bmatrix} 1 \\ 14 \end{bmatrix} = \begin{bmatrix} b \\ o \end{bmatrix} \quad (11.15)$$

$$\begin{bmatrix} 17 & 22 \\ 22 & 1 \end{bmatrix} * \begin{bmatrix} c \\ s \end{bmatrix} \rightarrow \begin{bmatrix} 17 & 22 \\ 22 & 1 \end{bmatrix} * \begin{bmatrix} 2 \\ 18 \end{bmatrix} \pmod{26} = \begin{bmatrix} 14 \\ 10 \end{bmatrix} = \begin{bmatrix} o \\ k \end{bmatrix} \quad (11.16)$$

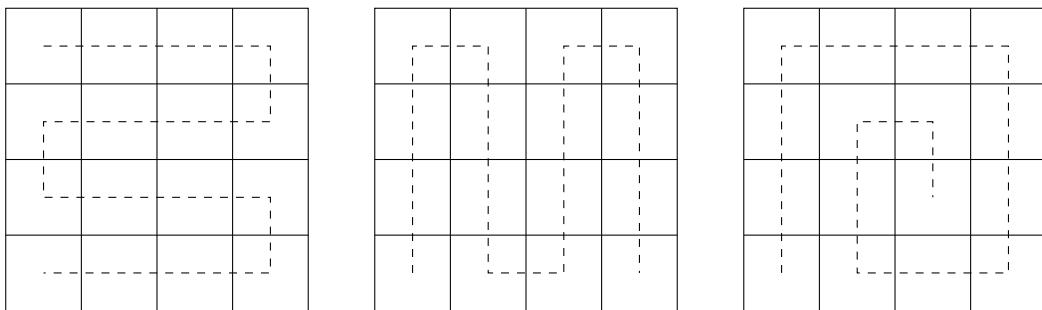
So the original message was "book". Well, book and beer, not really a good association...

11.4 Transposition Ciphers

So far we talk about ciphers that are build by substituting plaintext symbols with other symbols generated by an algorithm and an encryption key. On the other side, a transposition cipher does not substitute original plaintext symbols, but it scramble them, or group of them, moving from a position to another. In this way the ciphertext is exactly a permutation of the plaintext that is obtained by applying an algorithm and an encryption key. The decryption algorithm will unscramble the ciphertext in order to reassemble the plaintext.

11.4.1 Route Cipher

Route cipher, or Serpentine ciphers, usually refer to a family of transposition ciphers in which the ciphertext is obtained by following a continuous path on a $n \times m$ grid in which the plaintext has somehow been arranged. The decryption process consists in rebuilding the grid from the ciphertext stream and read back the plaintext message. Plaintext is usually arranged on the grid line by line or column by column starting from the upper left most cell, but more complex arrangements can be done in order to make the encryption stronger. The routes used to cover all cells of the grid for building the ciphertext can also be various, from linear ones to spiral ones. Here are showing some of the simplest possibilities:



Let's try to encrypt something using the "spiral route cipher". Usually the encryption and the decryption algorithm refers in the way the route is computed, in this case we use the "spiral" algorithm. But the algorithm alone is not enough to encrypt the plaintext neither to decrypt the generated ciphertext, we need to know the sizes of the grid. This will be our encryption and decryption key, let's say 4x4 grid. The plaintext we want to encrypt is just the word "RESPONSIBILITIES". Let's start arrange the letters in the grid column by column starting from the upper left most cell and read the spiral from the lower left most cell:

| | | | |
|---|---|---|---|
| R | O | B | T |
| E | N | I | I |
| S | S | L | E |
| P | I | I | S |

| | | | |
|----|----|---|---|
| R | O | B | T |
| E | N | I | I |
| \$ | \$ | L | E |
| P | I | I | S |

The ciphertext will be "PSEROBTIESIISNIL". How to recover the plaintext from the ciphertext? If the algorithm and the key are known it is easy to reverse the process and build the original grid. We know that the algorithm is "spiral" and the key is 4x4, so start rearrange the letters:

| | | | |
|---|---|---|---|
| - | - | - | - |
| E | - | - | - |
| S | - | - | - |
| P | - | - | - |

| | | | |
|---|---|---|---|
| R | O | B | - |
| E | - | - | - |
| S | - | - | - |
| P | - | - | - |

| | | | |
|---|---|---|---|
| R | O | B | T |
| E | - | - | I |
| S | - | - | E |
| P | - | - | - |

| | | | |
|---|---|---|---|
| R | O | B | T |
| E | - | - | I |
| S | - | - | E |
| P | I | I | S |

| | | | |
|---|---|---|---|
| R | O | B | T |
| E | N | I | I |
| S | S | L | E |
| P | I | I | S |

Reading the grid again, column by column starting from the upper right cell we obtain back the original message. Usually the length of the message cannot be arranged exactly in a chosen size $n \times m$ grid, specifically when the plaintext's length is a prime number. In these cases you can consider to keep or remove punctuation from the message or add some placeholder characters at the beginning or at the end to fill the grid completely.

Challenge 11.6 — The american road dream.

RIXTAERDHTXOUTHEOICFREYTESMRAAXMOOIS

Challenge 11.7 — The american road dream.

The encryption algorithm used is a column snake (similar of the one in example 2 but mirrored vertically), starting from the upper left most cell. The key is a 6x6 grid. The last 'X' symbols has been used to fill all the cells of the grid:

ROUTESIXTYSIXTHEMOTHERROADOFAMERICAX

ROUTE 66 THE MOTHER ROAD OF AMERICA

11.4.2 Columnar Transposition Cipher

Another transposition technique is provided by the columnar transposition cipher. As the route cipher, the plaintext is arranged into a grid, this time row by row starting from the upper left most cell. The number of the columns of the grid is given by the length of the key used to encrypt and decrypt the message. So if the key is "CAT" the message will be arranged into a $n \times 3$ grid (where n is the number of rows and 3 is the number of columns). Once the plaintext has been

arranged the order of the columns will be scrambled according to the key. The shuffle follows the lexicographical order of the letters in the key, so if the key is "CAT" the column corresponding to letter "A" will become the first of the grid, column corresponding to letter "C" will be the second and the last column remains there. Then the ciphertext is obtained by reading the grid column by column starting from the upper left most cell. Also in this cipher, like the previous one, there can be situation in which the grid is not fulfilled. In these cases some placeholders can be used, or just random text.

Consider the following example. We want to encrypt the message "meet at nine at park" using the key "MAZE". Start arranging the grid and shuffle according the lexicographical order of the letters composing the key:

| | | | |
|---|---|---|---|
| M | A | Z | E |
| 3 | 1 | 4 | 2 |

| | | | |
|---|---|---|---|
| M | E | E | T |
| A | T | N | I |
| N | E | A | T |
| P | A | R | K |

| | | | |
|---|---|---|---|
| A | E | M | Z |
| 1 | 2 | 3 | 4 |

| | | | |
|---|---|---|---|
| E | T | M | E |
| T | I | A | N |
| E | T | N | A |
| A | K | P | R |

Now we can proceed reading the ciphertext column by column obtaining "ETEATITKMANPENAR". Decrypting the message is very easy if you know the key. First we need to fill the grid that will have $\text{length}(\text{key})$ columns and $\text{length}(\text{ciphertext})/\text{length}(\text{key})$ rows. The filling starts from the upper left most cell and proceeds column by column. The key is put on the top of the grid as done in encryption mode but with the difference that it is ordered in reverse lexicographical order. Then we just need to shuffle the columns according to the defined order and read the plaintext message row by row:

| | | | |
|---|--|--|--|
| M | | | |
| 2 | | | |

| | | | |
|---|---|---|---|
| E | ⋮ | ⋮ | ⋮ |
| T | ⋮ | ⋮ | ⋮ |
| E | ⋮ | ⋮ | ⋮ |
| A | ⋮ | ⋮ | ⋮ |

| | | | |
|---|---|---|---|
| M | A | Z | E |
| 2 | 4 | 1 | 3 |

| | | | |
|---|---|---|---|
| E | T | M | E |
| T | I | A | N |
| E | T | N | A |
| A | K | P | R |

| | | | |
|---|---|---|---|
| Z | M | E | A |
| 1 | 2 | 3 | 4 |

| | | | |
|---|---|---|---|
| M | E | E | T |
| A | T | N | I |
| N | E | A | T |
| P | A | R | K |

We obtain back the message "MEET AT NINE AT PARK". Notice that every key with the same lexicographical order will produce the same result both on encryption and decryption phase.

Challenge 11.8 — Modern War.

EROSSYOOAONAHLWHRHWFTISLRNWCODTOAFUTRNTYUEAO

Hint: the length of the key is 5 ▀

Solution 11.6 — Modern War.

The key used for the cipher is PLANE. The decrypted message is so:

The war we confront today is thus solely a war of honour

Challenge 11.9 — Modern Warfare.

FVRAONINLMSITEYREIECUCFSEYOZWNTHSEETALKEDRAHLEAIIEFC

UDALEELIVTUGNLIONIEELOHDIGWOIIDIETOEDUAWEUALRNOVRHU

MTLBARCWDSFNIINBATAIFRNLKTLTCWDWDEOOOH

Hint: the length of the key is 9 ▀

Solution 11.7 — Modern Warfare.

The key used for the cipher is FACTORIES. The decrypted message is so:

if enough civilians were killed, factories could not function and if civilians were killed, the enemy would be so demoralized that it would have no ability to wage further war

11.4.3 Myszkowski Transposition (1902)

A variant of the classical columnar transposition cipher has been proposed by Émile Victor Théodore Myszkowski in his book "Cryptographie indéchiffrable" in 1902. The principle is the same as the original cipher but the idea is to make more confusion and ambiguity by choosing keys that have letters that appears more than once. Suppose we want to encrypt the sentence "The enemy in arriving in Paris" with the key "RADAR". We proceed filling the grid and shuffling columns as usual:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| R | A | D | A | R | A | A | D | R | R |
| 3 | 1 | 2 | 1 | 3 | 1 | 1 | 2 | 3 | 3 |
| T | H | E | E | N | H | E | E | T | N |
| E | M | Y | I | S | M | I | Y | E | S |
| A | R | R | I | V | R | I | R | A | V |
| I | N | G | I | N | N | I | G | I | N |
| P | A | R | I | S | A | I | R | P | S |

At this point, in the original columnar transposition cipher, each column will be read in order to generate the ciphertext "HMRNAEIIIIEYRGRTEAIPNSVNS". Myszkowski method works different and considers columns with the same lexicographical order as a single column where 2 or more letters are stored. In this way reading the first two columns we obtain "HEMIRINIAI",

reading the third is easy "EYRGR" and the last two columns encrypt the sequence "TNESAVINPS". The new ciphertext will be "HEMIRINIAIEYRGRTNESAVINPS" that, of course, is different from the previous one. In this case the decryption procedure described above doesn't work. the new procedure works as follows: Start writing the key over the grid in lexicographical order. If more letters inside the word have the same order fill the columns together, otherwise fill the single column. Last phase is to reassemble the key moving the columns in their original position and read back the plaintext row by row:

| A | A | D | R | R |
|-----|---|---|-----|-----|
| 1 | 1 | 2 | 3 | 3 |
| H | E | | --- | --- |
| M | I | | --- | --- |
| --- | | | --- | --- |
| --- | | | --- | --- |
| --- | | | --- | --- |

| A | A | D | R | R |
|---|---|---|-----|-----|
| 1 | 1 | 2 | 3 | 3 |
| H | E | E | --- | --- |
| M | I | Y | --- | --- |
| R | I | R | --- | --- |
| N | I | | --- | --- |
| A | I | | --- | --- |

| A | A | D | R | R |
|---|---|---|---|---|
| 1 | 1 | 2 | 3 | 3 |
| H | E | E | T | N |
| M | I | Y | E | S |
| R | I | R | A | V |
| N | I | G | I | N |
| A | I | R | P | S |

| R | A | D | A | R |
|---|---|---|---|---|
| 3 | 1 | 2 | 1 | 3 |
| T | H | E | E | N |
| E | M | Y | I | S |
| A | R | R | I | V |
| I | N | G | I | N |
| P | A | R | I | S |

12. Modern Cryptography

Algoritmi a chiave pubblica cosa è? idea...

12.1 RSA (1977)

RSA acronym derives from the first letter of surname of its creators: Ron Rivest, Adi Shamir and Leonard Adleman. It is one of the first public-key cryptosystem ever developed and it was widely used for secure data transmission. Generally speaking is a very good algorithm but with some flaws in parameter settings. Choosing a bunch of bad parameters makes RSA not secure at all and can be attacked in various ways as we will see later.

History says that in 1973 Clifford Cooks, while working at United Kingdom Government Communications Headquarters (GCHQ), invented a RSA equivalent cryptosystem. Because the idea was classified information, the fact remain hidden for 24 years till the work has been unclassified. Anyway the GCHQ never found a practical use of this cryptosystem. 4 years later a very similar algorithm has been independently invented by Rivest, Shamir and Adleman which were also able to find a practical use for it. It has never been found any evidence of information leak in GCHQ and so RSA is to be consider authentic.

But now let's focus on how RSA works. The encryption and the decryption formulas are the following:

$$c \equiv m^e \pmod{n} \tag{12.1}$$

$$m \equiv c^d \pmod{n} \tag{12.2}$$

Where m is the message to encrypt, e is the sender's public key, d is the receiver's private key and c is the encrypted message. The strength of RSA lies in choosing the right keys. The main idea is that it is very easy to compute $m^e \pmod{n}$ even with very large values of n and e , but at the same time is very difficult to infer d , reversing the encryption formula, even if the original message m is known due to the nature of n . N is so the most important parameter for RSA encryption but also e plays a fundamental role. What follows is the key generation algorithm:

- Choose two distinct prime numbers p and q
- Calculate $n = p * q$
- Compute $\lambda(n) = lcm(\phi(p), \phi(q)) = lcm(p - 1, q - 1)$
- Choose e such that $1 < e < \lambda(n)$ and $gcd(e, \lambda(n)) = 1$
- Calculate $d \equiv e^{-1} \pmod{\lambda(n)}$

With:

- $\lambda(n)$ is the Carmichael function defined as the smallest positive integer m such that $a^m \equiv 1 \pmod{n}$ for every integer between 1 and n that is coprime to n.
- $\phi(n)$ is the Euler's totient function defined as the number of positive integers up to n that are relatively prime to n.

The next table will summarize the Carmichael and Euler's totient values of first 20 integers:

| n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|--------------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|
| $\lambda(n)$ | 1 | 1 | 2 | 2 | 4 | 2 | 6 | 2 | 6 | 4 | 10 | 2 | 12 | 6 | 4 | 4 | 16 | 6 | 18 | 4 |
| $\phi(n)$ | 1 | 1 | 2 | 2 | 4 | 2 | 6 | 4 | 6 | 4 | 10 | 4 | 12 | 6 | 8 | 8 | 16 | 6 | 18 | 8 |



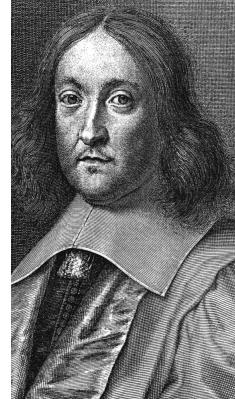
Let's do an example.

- We choose $p = 1009$ and $q = 797$ as our basic different prime numbers.
- Then we compute $n = p * q = 1009 * 797 = 804173$
- We calculate $\lambda(n) = lcm(p - 1, q - 1) = lcm(1008, 796) = 200592$
- We choose $e = 29$. Thus $gcd(29, 200592) = 1$
- We calculate $d \equiv 29^{-1} \pmod{200592} \rightarrow d = 608693$

Now we are ready to encrypt a message, for example the string "IT". First thing to do is to transform the string into numeric values. It can be done by transforming first to hex notation and then convert to dec base: "IT" \rightarrow 49 54 \rightarrow 0x4954 \rightarrow 18772. Now we can encrypt the message using the public key e: $c \equiv 18722^{29} \pmod{200592} = 382506$. To decrypt we can use the private key d: $m \equiv 382506^{608693} \pmod{200592} = 18722$

12.1.1 Breaking RSA

In its simplicity RSA is, in general, a good encryption algorithm, but if used superficially it can present some flaws and vulnerabilities. As we saw before n is a crucial parameter for guarantee high level security. If n can be easily factored, then the attacker have enough data to break the cipher. It is so important that p and q are big enough to guarantee that n cannot be factored easily. Online factorization DBs are quite common nowadays ad factordb.com is one of the biggest in the web. Looking for example for the number 435897296798372609834276094237603948760923737 it says that its factors are 4643 and 93882682920175018271435729967177245048659 that are both primes.



12.1.2 Fermat's Attack

Another common mistake is to use p and q too close each other. In this case, even if p and q are very large or even huge, the Fermat's factorization method is able to retrieve them easily. The theorem is based on the assumption that and odd number can be represented as a difference of two squares:

$$N = a^2 - b^2 \tag{12.3}$$

And it is well known that this difference is easily factorable as

$$a^2 - b^2 = (a + b) * (a - b) \tag{12.4}$$

Indeed if $n = c * d$ is a factorization of n , then $c * d$ can be rewritten as

$$N = \left(\frac{c+d}{2}\right)^2 - \left(\frac{c-d}{2}\right)^2 \quad (12.5)$$

$$a = \frac{c+d}{2} \quad (12.6)$$

$$b = \frac{c-d}{2} \quad (12.7)$$

Knowing this the Fermat's algorithm works as follows:

- Start with $a = \text{ceil}(\sqrt{N})$
- compute $b_2 = a^2 - N$
- until b_2 is not a perfect square do:
 $a = a + 1$
 $b_2 = \sqrt{a^2 - N}$

When the algorithm ends we know that $a = a$ and $b = \sqrt{b_2}$. Then we only need to solve the system provided by equations (12.6) and (12.7) to find d and c :

$$\begin{cases} a = \frac{c+d}{2} \\ b = \frac{c-d}{2} \end{cases} \quad \begin{cases} c = a + b \\ d = a - b \end{cases} \quad (12.8)$$

 Let's do an example. Imagine to have $N = 2151491$. Running the Fermat's Algorithm step by step we obtain:

| step: | 1 | 2 | 3 | 4 |
|-------|-------|-------|-------|------|
| a | 1467 | 1468 | 1469 | 1470 |
| b_2 | 598 | 3533 | 6470 | 9409 |
| b | 24.45 | 59.43 | 80.43 | 97 |

At step 4 b_2 is a perfect square and the algorithm ends. Let's proceed solving the equation system, knowing that $a = 1470$ and $b = 97$:

$$\begin{cases} c = 1470 + 97 \\ d = 1470 - 97 \end{cases} \quad \begin{cases} c = 1567 \\ d = 1373 \end{cases} \quad (12.9)$$

In this way we easily found $N = c * d \rightarrow 2151491 = 1567 * 1373$

12.1.3 Common modulus



Common modulus attack can be done when 2 messages encrypted with the same modulus N are intercepted by the attacker. The attacker needs to know also the public keys used to encrypt the messages, but since the keys should be public, this is a free step. To understand how the attack works let's recap how the intercepted messages are encrypted:

$$C_A = M_A^{e_A} \quad (12.10)$$

$$C_B = M_B^{e_B} \quad (12.11)$$

In this case, knowing that e_A and e_B must be primes, it is a pain fact. To explain why all this panic, let's introduce the Bézout's identity:

"Let a and b be integers with greatest common divisor d . Then, there exist integers x and y such that $ax + by = d$. More generally, the integers of the form $ax + by$ are exactly the multiples of d ."

With e_A and e_B primes it is always true that $\gcd(e_A, e_B) = 1$, and for the Bézout's identity follows that there must exist u and v such that:

$$e_A * u + e_B * v = \gcd(e_A, e_B) = 1 \quad (12.12)$$

To solve this equation, the Euclidean algorithm to compute $\gcd(a, b)$ comes in help. Given two integer numbers a and b , the algorithms works as following:

- $a = q_0 * b + r_0$
- $b = q_1 * r_0 + r_1$
- while r_k is not zero do:

$$r_{k-2} = q_k * r_{k-1} + r_k$$

In this way we obtain a sequence like this:

$$a = q_0 * b + r_0 \quad (12.13)$$

$$b = q_1 * r_0 + r_1 \quad (12.14)$$

$$r_0 = q_2 * r_1 + r_2 \quad (12.15)$$

$$\dots \quad (12.16)$$

$$r_{n-2} = q_n * r_{n-1} + r_n \quad (12.17)$$

$$r_{n-1} = q_{n+1} * r_n + 0 \quad (12.18)$$

Euclidean's algorithm is used to find gcd but we use it in another way. Once computed gcd we can rewrite the results of the single steps as follows:

$$r_n = r_{n-2} - q_n * r_{n-1} \quad (12.19)$$

$$\dots \quad (12.20)$$

$$r_2 = r_0 - q_2 * r_1 \quad (12.21)$$

$$r_1 = b - q_1 * r_0 \quad (12.22)$$

$$r_0 = a - q_0 * b \quad (12.23)$$

Substituting terms into an unique formula, will solve the equation, providing the terms u and v of the Bézout's identity.

R Suppose to have $e_A = 71$ and $e_B = 53$ and we want to solve the Bézout's identity $71 * u + 53 * v = 1$. We proceed computing $\gcd(71, 53)$ as follows using the Euclidean's algorithm:

$$71 = 1 * 53 + 18$$

$$53 = 2 * 18 + 17$$

$$18 = 1 * 17 + 1$$

$$17 = 17 * 1 + 0$$

Now we can rearrange the steps as follows:

$$(1) 1 = 18 - 17 * 1$$

$$(2) 17 = 53 - 18 * 2$$

$$(3) 18 = 71 - 53 * 1$$

We can now start the substitution process. Putting (2) in (1) we obtain $1 = 18 - 1 * (53 - 18 * 2)$.

Putting also (3) in this new equation it comes that $1 = (71 - 53 * 1) - 1 * (53 - 2 * (71 - 53 * 1))$.

Simplifying this last equation we get

$$71 - 53 - 1 * (53 - 2 * 71 + 2 * 53) = 1 \quad (12.24)$$

$$71 - 53 - 53 + 2 * 71 - 2 * 53 = 1 \quad (12.25)$$

$$3 * 71 - 4 * 53 = 1 \quad (12.26)$$

This solves the Bézout's identity with solutions $u = 3$ and $v = -4$.

Till now only a big bunch of boring concepts. But how this concepts can represent a problem for RSA? The fact is that (12.10) and (12.11) can be rewritten as:

$$C_A^u = M_A^{e_A^u} = M^{e_A * u} \quad (12.27)$$

$$C_B^v = M_B^{e_B^v} = M^{e_B * v} \quad (12.28)$$

Now, if we multiply both messages we obtain the current flaw:

$$M^{e_A * u} * M^{e_B * v} = M^{e_A * u + e_B * v} = M^1 = M \quad (12.29)$$

Since $C = M^e$, to decrypt RSA is sufficient to calculate

$$M = C_A^u * C_B^v \quad (12.30)$$

usually v is negative and so the inverse of $C_B \text{mod } N$ have to be computed:

$$M = C_A^u * C_{B,\text{inv}}^{-v} \quad (12.31)$$

12.1.4 Small Public Exponent

One of the worst practice in RSA is using small public exponent e . Even if N has been computed choosing strong primes p and q , a small exponent will make the whole encryption insecure. First, what is the main purpose of e ? Well, to make the message M bigger enough so that modulus N makes sense. Thus if $M^e < N$, then the encrypted message can be simply decoded using:

$$M = \sqrt[e]{C} \quad (12.32)$$

Even if this are extremely rare cases in which the message is very short and the public key is very low, it is a good situation to take care

- R A Free decryption is obtained when trivially $e = 1$. In this case the have that $C = M^1 \text{ mod } N$ and so $C = M$ and no decryption is needed at all (just convert number to ascii)

12.1.5 Hastad's Broadcast Attack

There is actually another problem in using small public exponents and it comes when the same message is sent to, at least, 3 different people. Suppose that the message is long enough that the previous technique cannot be applied, what is the problem in sending the same message to different people using a small exponent? Start showing the anatomy of the attack: Suppose that we were able to capture at least e ciphertexts corresponding to the same plaintext, then the following system of equations is true:

$$\begin{cases} C_1 \equiv M^e \pmod{N_1} \\ C_2 \equiv M^e \pmod{N_2} \\ \vdots \\ C_e \equiv M^e \pmod{N_e} \end{cases} \quad (12.33)$$

Solving this system is trivial because the chinese remainder theorem comes in help, telling us that for (12.33) exists a single module $N = \prod_{i=1}^e N_i$ in the case in which $\gcd(N_i, N_j) = 1 \forall i \neq j$. We can safely assume the this is true because otherwise we were able to find factors p and q by simply calculate $\gcd(N_i, N_j)$. Now we are able to rewrite the previous system in:

$$C \equiv M^e \pmod{N_1, N_2, \dots, N_e} \quad (12.34)$$

However, since for short messages $M < N_i \forall i$, then also $M^e < N_1 * N_2 * \dots * N_e$ and (12.41) can be rewritten as:

$$C \equiv M^e \quad (12.35)$$

That is simply solvable calculating $M = \sqrt[e]{C}$

12.1.6 Wiener's Attack

12.2 Rabin Cryptosystem (1979)

In 1979 Michael Oser Rabin, proposed a variant of RSA cryptosystem aiming to flat some flaws of this last technique. The key generation phase has been extremely simplified preserving a good level of security. In fact, Rabin cryptosystem is consider more secure than RSA because has been proven that decrypting a message encrypted by Rabin, without knowing private keys, lies in solving the integer "factorization problem" that is supposed to be NP-Complete since till now no algorithm has been provided to solve it in polynomial time. Sadly, even sharing a very similar encryption technique, has never been proved formally that RSA lies in the same decryption problem. Let's focus on Rabin cryptosystem. The encryption and the decryption formulas are the followings:

$$c \equiv m^2 \pmod{n} \quad (12.36)$$

$$\text{misoneof} \begin{cases} M_1 = a * p * m_q + b * q * m_p \\ M_2 = a * p * m_q - b * q * m_p \\ M_3 = -a * p * m_q + b * q * m_p \\ M_4 = -a * p * m_q - b * q * m_p \end{cases} \quad (12.37)$$

What does it mean? Analysing (12.37) it comes that in order to decrypt the message is necessary to solve 4 different equations and one of them will be the correct original plaintext. To better understand this, let's proceed with the key generation algorithm. The algorithm only requires, as RSA, to choose two different prime numbers p and q and then compute $N = p * q$. That's it: N will be the public encryption key and p, q will be the private decryption key. We can now encrypt any message using (12.36), but how the decryption algorithm works?

In order to decrypt Rabin, the following problem needs to be solved:

$$m^2 \equiv c \pmod{N} \quad (12.38)$$

For composite $N = p * q$ there is usually no solution for this problem, but since p and q are primes only one solution exists and we are able to compute the 4 square roots of (12.38). To do so we need to compute first:

$$m_p = \sqrt{c} \pmod{p} \quad (12.39)$$

$$m_q = \sqrt{c} \pmod{q} \quad (12.40)$$

$$a * p + b * q = 1 \quad (12.41)$$

In particular (12.41) can be computed using the extended Euclidean algorithm showed in (12.12). Finally the Chinese remainder algorithm is used to compute the final roots showed in (12.37). Which between M_1, M_2, M_3 and M_4 is the correct plaintext is usually unknown at prior.

12.3 Block Ciphers (1981)

Block ciphers differ from stream ciphers by the way in which the plaintext is encrypted. If in stream ciphers each symbol is transformed separately from each other, in block ciphers, as the name suggests, blocks of plaintext symbols are processed in order to generate the ciphertext. First specification of block cipher techniques comes from the Federal Information Processing Standards Publication 81 (FIPS81) where ECB, CBC, OFB and CFB are described.

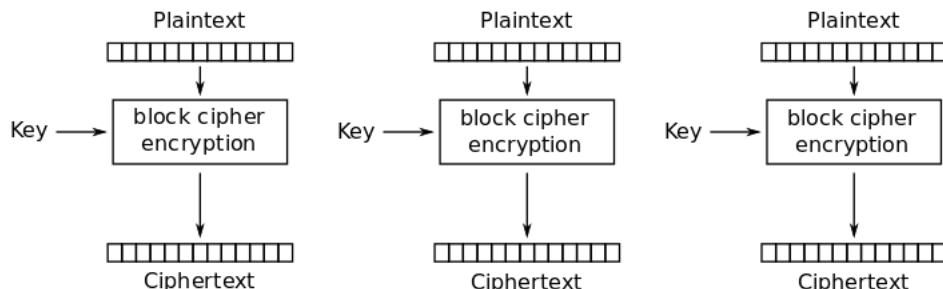
12.3.1 Electronic Codebook (ECB)

ECB is the simplest block based encryption schema, where the plaintext is divided into blocks of fixed length and each block is then encrypted separately using an encryption function. The encryption function takes the block to be encrypt (B_i) and the encryption key (K) as input and produces a block of ciphertext (C_i):

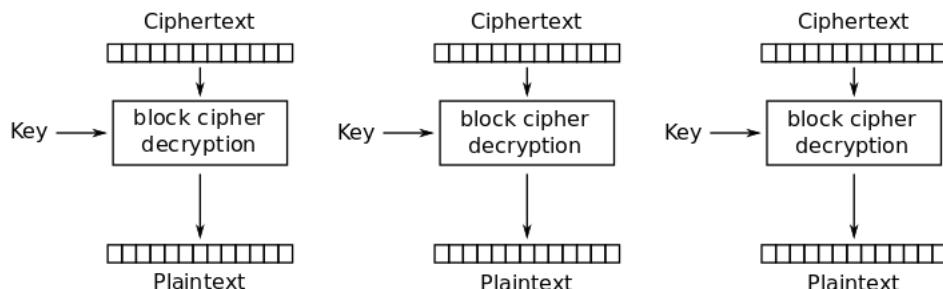
$$C_i = F_{enc}(B_i, K) \quad (12.42)$$

In the same way the decryption function takes in input each cipher block and reconstruct the original plaintext block:

$$B_i = F_{dec}(C_i, K) \quad (12.43)$$



Electronic Codebook (ECB) mode encryption



Electronic Codebook (ECB) mode decryption

Cracking ECB is so simple when known plaintext attack is used. This because each block is encrypted always in the same way, even in different positions in the original message.

- ④ Suppose that our encryption function simply reverse the block string, and we use ECB with block size 4. Then to encrypt "thisroomissoddark" we proceed as following:

| | | | | |
|-------------------|------|------|------|------|
| plaintext blocks: | this | room | isso | dark |
| ciphertext blocks | siht | moor | ossi | krad |

Obtaining "sihtmoorossikrad". Then using the same function we want to encrypt "thisismy-darkroom":

| | | | | |
|-------------------|------|------|------|------|
| plaintext blocks: | this | ismy | dark | room |
| ciphertext blocks | siht | ymsi | krad | moor |

That produces the ciphertext "sihtymsikradmoor", which it is easy to see the similarities with respect to the previous encrypted text. This property makes ECB not so secure.

12.4 Stream Ciphers (1987)

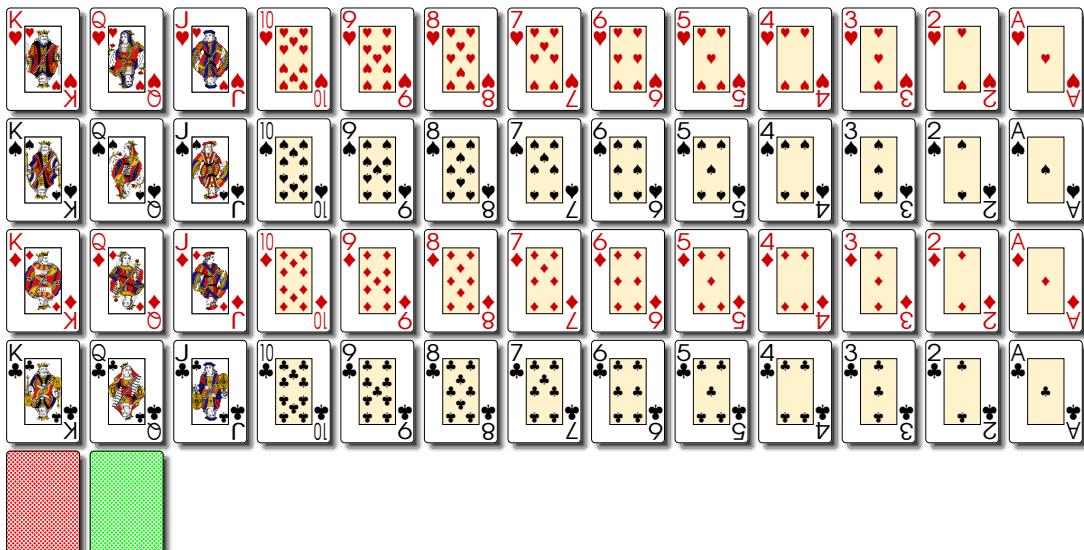
A stream cipher is an encryption system where each symbol of the plaintext is combined with a different pseudorandom generated symbol that composes the keystream. Keystreams are typically generated serially starting from a seed that serves as the key for the whole cryptosystem. Stream ciphers are inspired by the one-time pad (OTP) invented by Frank Miller in 1882 and patented to Gilbert Vernam in 1919. OTP is a theoretical unbreakable cipher where each plaintext symbol is combined with a TRUE random symbol. The proof of unbreakable has been first given by Vladimir Kotelnikov in 1941, but the document remained classified. In 1945 also Claude Shannon arrived at the same conclusion. The Shannon's results have been unclassified in 1949.

12.4.1 Solitaire/Pontifex Cipher (1999)

The Solitaire Cipher has been invented by Bruce Schneier in 1999 for the Neal Stephenson's novel "Cryptonomicon". The encryption algorithm works as follows:

- All characters that are not letters are removed from the plaintext message.
- All the characters are converted to same case (Upper or Lower no matter)
- Each character is converted in its numerical value (A=1, B=2, ..., Z=26)
- Add to each number the keystream value in the corresponding position, mod 26

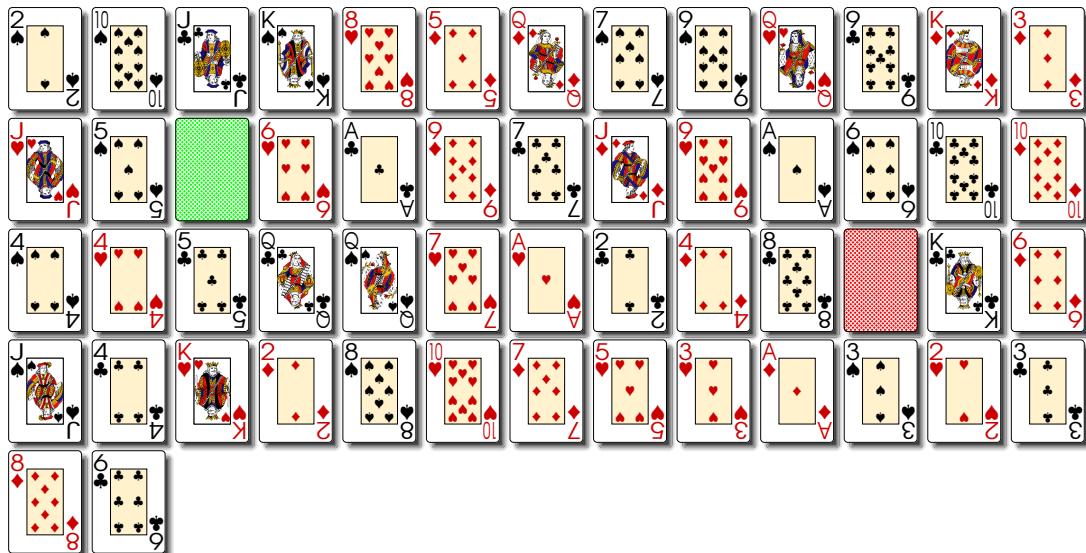
The decryption just work in the opposite way: subtract keystream values from ciphertext and convert numbers to letters. The tricky and fun part is that the every value of the keystream is generated starting from a deck of 52 card plus 2 jolly (the green and red cards), for a total of 54 cards.



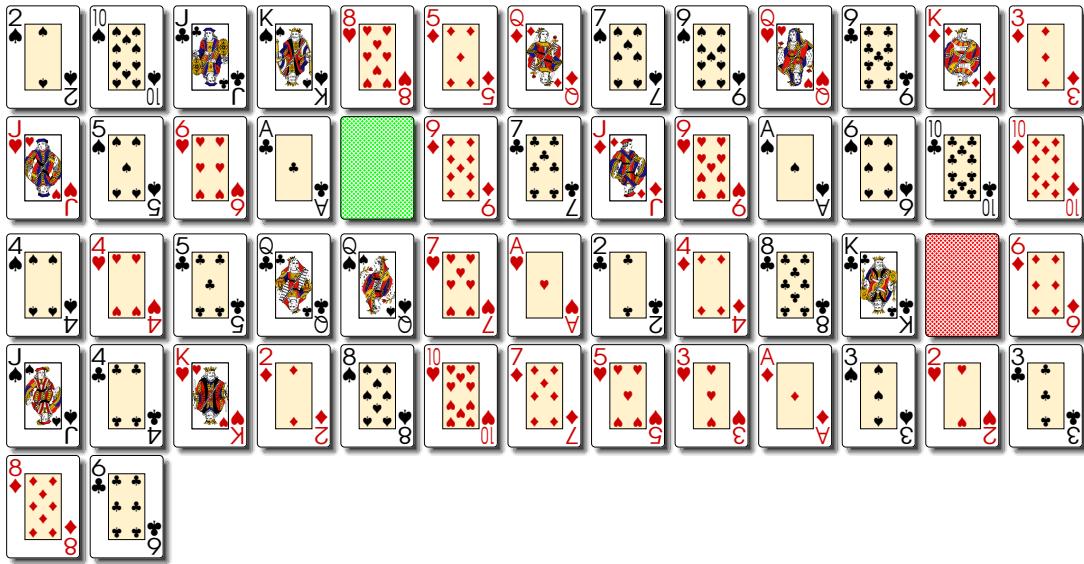
The keystream generation works as this:

- Shuffle the deck and put the cards face-up in sequence. This will be the key of the whole cipher. Anyone who knows the same key will be able to reproduce exactly the following steps. Notice that there are $54!$ or 23084369733924138047209274268302758108327856457180794 1132288000000000000000 possible different ways to arrange a deck of 54 cards. So every time you shuffle a deck it is extremely probable that that specific arrangement has never appeared in the history. This makes the initialization of the algorithm pretty safe.
- Locate the first joker in the arranged row of cards and move it one position right. If the joker is in the last position, move it to the second position of the deck (consider that the line is cyclic). In this way the
- Locate the second joker and move it two positions to the right with the same rules as the first joker.
- Split the deck into three sections. Everything above the first joker you encounter and everything below the second joker you encounter will be exchanged. The jokers and the cards between them, are left untouched.
- Now check the value of the last card of the deck. If it is a joker the value will be 53. Take that number of cards and insert them on the bottom of the deck but just before the last card.
- Look at the value of the top card. Let's say it is n . The next value of the keystream will be the value of the card at position $n + 1$. If a joker is encountered, don't add anything to the keystream.
- Repeat bullets from 2 to 6 to generate a number of keystream values equals to the length of the message to encode.

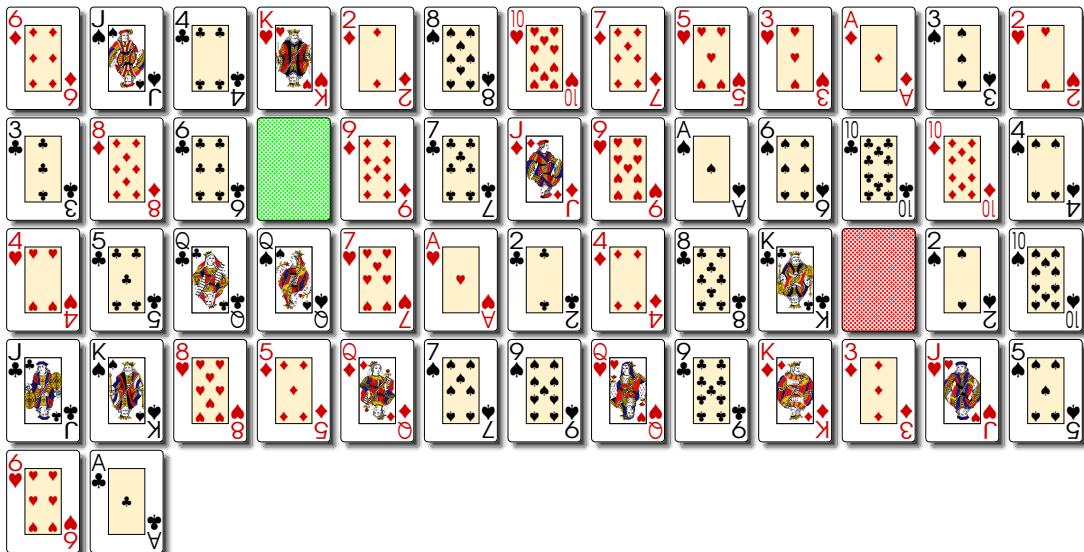
Let's make an example on how the procedure works. Before start we have to give a value to each card. Let's say that cards from Ace of Hearts to King of Hearts have values between 1 and 13. Cards from Ace of Spade to King of Spade have values between 14 and 26. Cards from Ace of Diamond to King of Diamond have values between 27 and 39. Cards from Ace of Club to King of Club have values between 40 and 52. Finally both jokers have value 53. Between the two jokers we need to specify which for us is the first and which is the second one. In our example red joker will be the first and green joker will be the second. Let's start arranging our deck randomly:



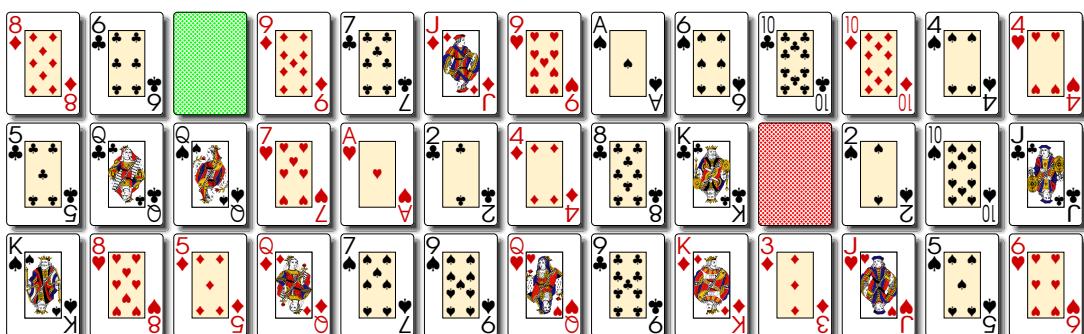
I will perform bullet 2 and bullet three together by moving our first joker (the red one) one position right, and the green joker two positions right:

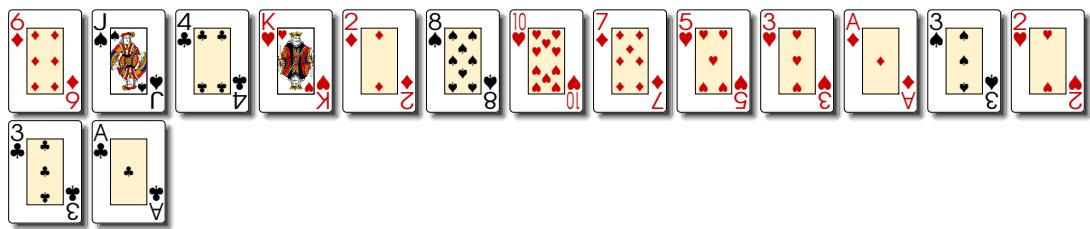


Now we can split the deck into three parts. The first 17 cards till the first jolly we encounter will replace the 16 cards from the second jolly we encounter to the end of the deck. Notice that the distinction between first and second jolly we made before it is irrelevant for this step.



Now we have to observe the value of the last card of the deck. It is the Ace of Spade. So for our convention it has the value 14. So we need to take the first 14 cards of the deck and put at the end of the deck, but before the last card, the Ace of Spade





Now we only have to read the value of the first card of the deck that is the eight of Diamond. Its value is 33. The first value of our keystream will be the value of the card at position 34, the nine of Club, that has value 48.

13. Cryptanalysis

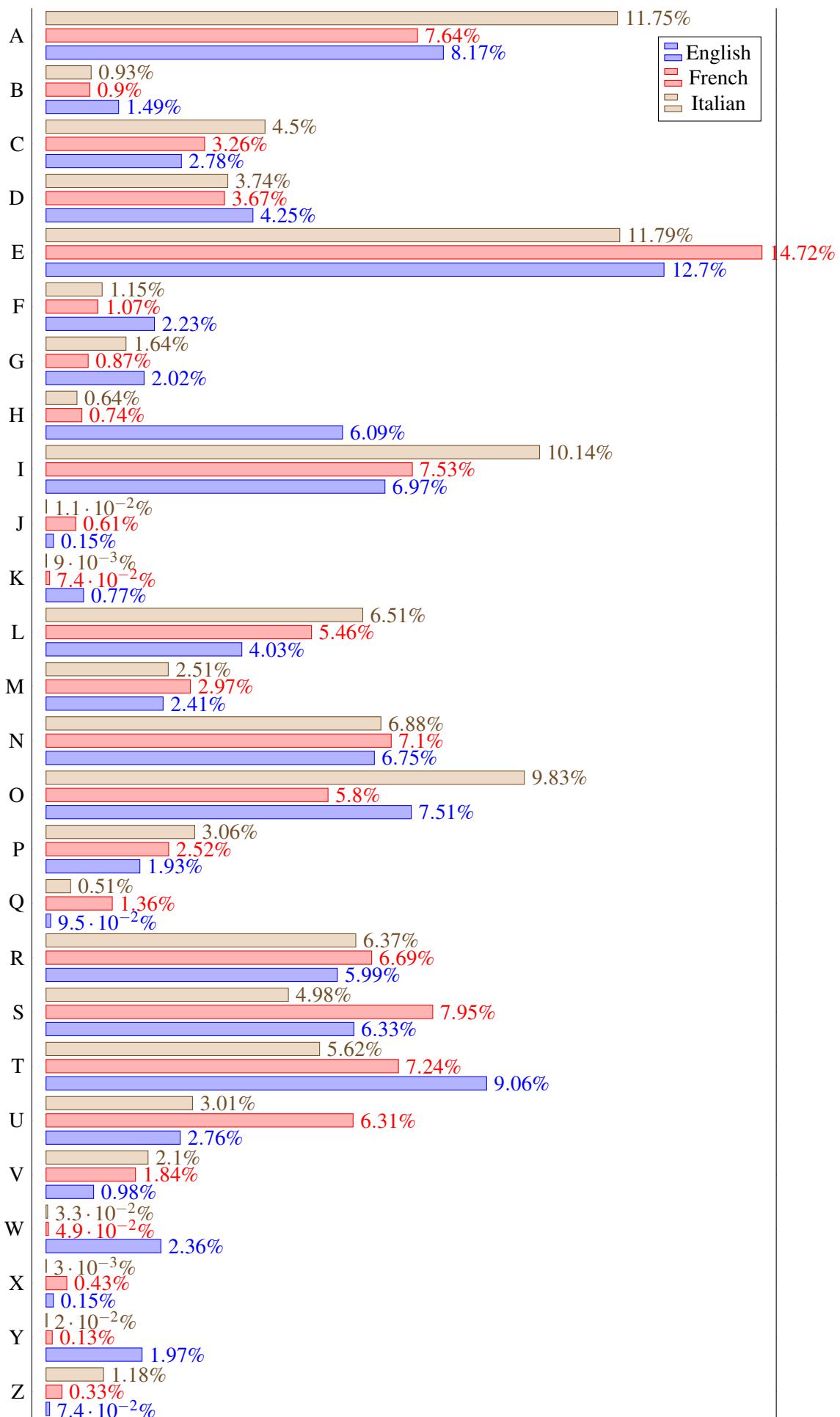
Cryptanalysis is the study of analyzing information systems in order to study the hidden aspects. It is mostly used to break cryptographic algorithms and gain access to the contents of encrypted messages, even if the cryptographic key is unknown. According to the amount of information available to the attacker, attacks can be classified as:

- **Ciphertext-only:** The cryptanalyst has access only to an encrypted set of messages.
- **Known-plaintext:** The cryptanalyst knows the set of original messages corresponding to encrypted ones.
- **Chosen-plaintext:** The cryptanalyst is able to obtain the encrypted message of an arbitrary chosen plaintext.
- **Adaptive chosen-plaintext:** As before, except the cryptanalyst can choose plaintext sequentially according to information obtained from previous computing.
- **Related-key attack:** As Chosen-plaintext, except the cryptanalyst obtain ciphertexts under different encryption keys.

13.1 Frequency Analysis

One of the easiest cryptoanalysis technique, classified as ciphertext-only, regards the study of the frequencies of single letters or group of them (bigrams or trigrams) that are involved in the encrypted text.

Frequency analysis is based on the fact that, in common text, certain letters and combinations of them occur with some frequencies. This means that when two texts written with the same language and with a moderate length are compared each other, a common pattern of frequencies is shared. For the English language, for example, it is easy to notice that letters E, T, A and O are the most common, while Z, Q and X are less frequent. In the same way bigrams TH, ER, ON, and AN are more likely than other couple of letters and SS, EE, TT, and FF are the most common repeats. The following chart summarize the unigram frequencies of three common languages: English, French and Italian.



The first known and recorded usage of frequency analysis was made by Al-Kindi, an Arab polymath of the 9th century, in A Manuscript on Deciphering Cryptographic Messages. In the Manuscript he studied the Qur'an discovering that Arabic has a characteristic letter frequency.

Later, in 1474, Cicco Simonetta wrote a manual on deciphering encryptions of Latin and Italian text.

A curiosity: Sherlock Holmes used it to solve the cryptogram contained in "The Adventure of the Dancing Man", shown previously in the book.

TODO example on how to decypt

13.1.1 Frequency steganography

Frequency analysis is a powerful technique to decrypt substitution ciphers, but it could be a steganography method itself, paradoxally... Try to solve the following challenge

Challenge 13.1 — The order matter. This text hides a secret, i am pretty sure...

FrFErFYFrFEFrqFEFrFUUrFEFrFqFrEFUFrEFrFqFreFrFEFqrFEFrFUUrFEFrFqcrFEFrFeEF
rFqrFUFrFErFqFrFEFrFcrFEFrUrFErFqFrFcEFrFqFrFErFNFrFEFqrFEFrFUUrFqFrFEeFr
FErFqFrFErFUFrqFrFEFNrFEFrFUFrEFqFrFErFrqrFErFUFErFqFrNFrFEFqrFEFrFUFrE
FeFrqFrFEFcFrEFrFqFrFEFrUFrFFqFFrEeFrFEFrFqFrEFUFrFqrErNErUFqFrFrFeFEFrq
FrEFYrFqFrEFUFrEFrFqrFEeFrqrEFUUrFEFrFYFrEFqFrFUFErFqFrFEFeFqFErFUFrFEFqr
FEFrFNFUrFEFrFqFEeFEFrFqrFcFrFEqFrFEFrUrFEFeFrFqFrEFUFrFEFrqNFqFrEFrFUF
EFcrFqFEFrFUUrFEFqFrEFrFeFrEFqFrFNrFEFrFUEFrFqFrFEeFrFqEFrFUFrFErFqFEFrYF
rFqFEFrFUFrFEqFrFEFrFqFrFEFrFqrFEFrFUEFrFqEFrFNeFEFrqFrFEFrUrEFrqFYFrFEFrqFrFEF
UFrEFrFeFrqFrFEFrUrFEFrFqrFNFEFrqFrFEFrFUEFrFeFrEFqFrFEFrFqrFEFrUrFqFrEFr
FNFrEFUFrFEFqFrFrFEqFrFEFrFUrFE

Solution 13.1 — The order matter. Checking the frequency of chars, and ordering them descending, follows that:

F: 287

r: 163

E: 98

q: 58

U: 33

e: 19

N: 10

c: 7

Y: 5

Sequences



14 Famous Sequences *(.5pc).107

- 14.1 Fibonacci Numbers
- 14.2 Lucas Numbers
- 14.3 Look-and-say Sequence

15 Figurate Numbers *(.5pc).111

- 15.1 Triangular Numbers
- 15.2 Square Numbers

14. Famous Sequences

Lots of riddles requires to find missing numbers, or number at a specific position, of given sequences. Unfortunately there is no such kind of sequence solver that works generally for every possible existing, or not existing, sequence. Thus a knowledge on at least the most famous ones is required in order to approach to this type of challenges.

14.1 Fibonacci Numbers

Challenge 14.1 Find the missing numbers of the following sequence separating by comma:
0,1,1,2,3,5,8,?,21,34,?,89 ... ▀

Fibonacci Numbers is one of the most popular and known sequence of integer numbers. The number at position n is simply the sum of the number at position $n - 1$ with the number at position $n - 2$ of the sequence, starting with 0 and 1 as the first two numbers of the sequence. More formally we have that:

$$Fib(n) = \begin{cases} 0 & n=0 \\ 1 & n=1 \\ Fib(n-1) + Fib(n-2) & n>1 \end{cases} \quad (14.1)$$

Solving the challenge is really simple, in this case. We just need to sum together the 2 numbers before the question marks to obtain its value:

Solution 14.1 Simply enough:

- For the first '?' we have that $5 + 8 = 13$
- For the second '?' we have that $21 + 34 = 55$

Thus the solution is 13,55.

Sometimes challenges involving sequences requires to calculate the value at a specific position. Most of times sequences are infinite and calculating the n th value by recursively apply the self representing formula is unsustainable, neither using a powerful calculator:

Challenge 14.2 Find the 150th number of the following sequence:
0,1,1,2,3,5,8,13,21,34,55,89 ... ▀

Fibonacci numbers have been well studied since its discovery, and now we are able to calculate the nth number of the sequence simply by applying the following formula:

$$Fib(n) = \frac{(1 + \sqrt{5})^n - (1 - \sqrt{5})^n}{2^n * \sqrt{5}} \quad (14.2)$$

Solution 14.2 Thus the solution at the previous challenge is:
 $Fib(150) = \frac{(1 + \sqrt{5})^{150} - (1 - \sqrt{5})^{150}}{2^{150} * \sqrt{5}} = 9969216677189303386214405760200$

14.2 Lucas Numbers

Similar to the Fibonacci numbers, each Lucas number is defined to be the sum of its two immediate previous terms. The first two Lucas numbers are 2 and 1 as opposed to the first two Fibonacci numbers 0 and 1:

$$Lucas(n) = \begin{cases} 2 & n=0 \\ 1 & n=1 \\ Lucas(n-1) + Lucas(n-2) & n>1 \end{cases} \quad (14.3)$$

Challenge 14.3 Find the missing numbers of the following sequence separating by comma:
2,1,3,4,7,?,18,29,?,76 ... ▀

Solution 14.3 Same as before:

- For the first '?' we have that $4 + 7 = 11$
- For the second '?' we have that $18 + 29 = 47$

Thus the solution is 11,47.

And, of course, we have also a fast calculation formula:

$$Lucas(n) = \left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \quad (14.4)$$

Challenge 14.4 Find the 150th number of the following sequence:
2,1,3,4,7,11,18,29,47,76 ... ▀

Solution 14.4 The solution is:

$$Lucas(150) = \left(\frac{1 + \sqrt{5}}{2} \right)^{150} - \left(\frac{1 - \sqrt{5}}{2} \right)^{150} = 22291846172619859445381409012498$$

14.3 Look-and-say Sequence

Sometimes sequences does not follow pure mathematical formulas or reasoning, but some kind of other logic. The look-and-say sequence is a wonderful example to show. The rule to build the sequence is very simple: To generate the next term, read the digits of the previous number, counting the number of digits in groups of the same digit that occurs sequentially. For example i will read the number 31131122211 as:

- one 3
- two 1
- one 3
- two 1
- three 2
- two 1

The next number of the sequence will be 132113213221. This sequence was first thought by John Horton Conway, more known as the inventor of the famous "Conway's Game of Life". Starting from digit 1 the sequence is so composed:

1,
11,
21,
1211,
111221,
312211,
13112221,
1113213211,
31131211131221,
13211311123113112211,
11131221133112132113212221,
3113112221232112111312211312113211,
1321132132111213122112311311222113112211312211,
11131221131211131231121113112221121321132132211331222113112211,
311311222113111231131112132112311321322112111312211312111322212311322113212221

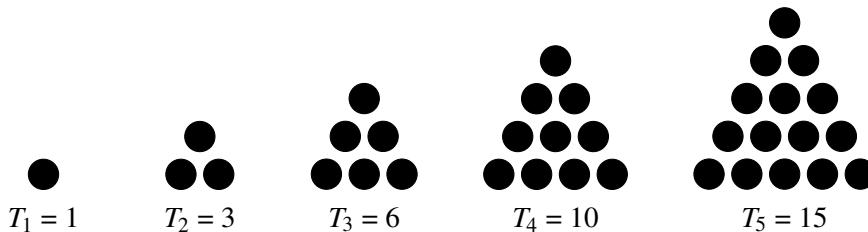
Challenge 14.5 — Just for fun. what is the next number of the sequence? ■

Solution 14.5 — Just for fun. 132113213221133112132113311211131221121321131
211132221123113112221131112311332111213211322211312113211

15. Figurate Numbers

Figurate numbers are special numbers that can be represented by some regular geometry, according to some kind of arrangement. The way to arrange them can vary and assume, at same time, different geometric forms in different geometrical dimensions. When the arrangement forms a polygon we refer to **polygonal numbers**.

15.1 Triangular Numbers



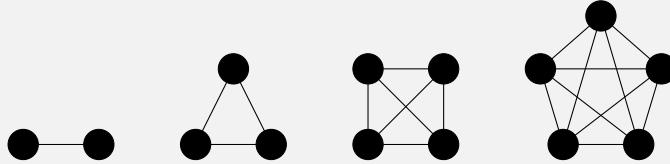
$$T_n = \sum_{k=1}^n k \quad (15.1)$$

$$T_n = \frac{n * (n + 1)}{2} \quad (15.2)$$

Challenge 15.1 — Galactic meeting. University of Trento is organizing a galactic conference where bilateral discussion each-to-each between participants are on the agenda. All participants come from different galaxies and no one knows any foreign language. Organizer have to guarantee one interpreter for every bilateral discussion. Each interpreter is able only to translate only one pair of languages. Because two participants refused the invitation, the university could reduce number of interpreters by 31. How many interpreters will be on the meeting after being

reduced? and how many delegates were invited originally? ■

Solution 15.1 — Galactic meeting. Starting guessing logically, for 2 delegates only 1 interpreter is required, for 3 delegates we need 3 interpreters, for 4 delegates we need 6 interpreters, for 5 delegates 10 interpreters are required. This schema can be seen graphically below, where dots are delegates and lines are interpreters:



It follows that the number of required interpreters follow the triangular numbers sequence:

$$\text{for } n \text{ delegates we need } T_{n-1} \text{ interpreters} \quad (15.3)$$

To find the number of originally invited delegates we need to find n such that

$$T_{n-1} - T_{n-3} = 31 \quad (15.4)$$

Combining (15.4) with (15.2) we have:

$$\frac{(n-1)*((n-1)+1)}{2} - \frac{(n-3)*((n-3)+1)}{2} = 31 \quad (15.5)$$

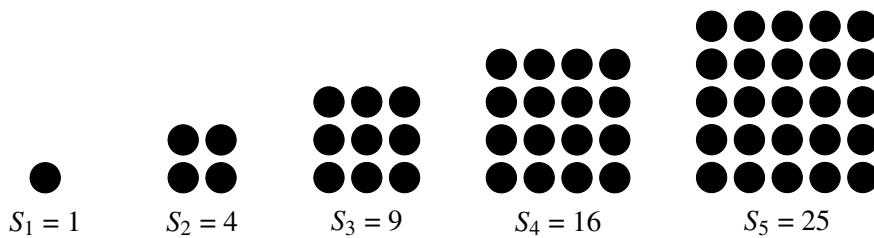
$$\frac{(n-1)*n}{2} - \frac{(n-3)*(n-2)}{2} = 31 \quad (15.6)$$

$$\frac{n^2 - n - (n^2 - 3n - 2n + 6)}{2} = 31 \quad (15.7)$$

$$\frac{4n - 6}{2} = 31 \rightarrow 2n - 3 = 31 \rightarrow 2n = 34 \rightarrow n = 17 \quad (15.8)$$

Originally 17 delegates were invited. 2 of them declined the invitation so only 15 delegates will come. From (15.3) follows that we need $T_{14} = 105$ interpreters.

15.2 Square Numbers



$$T_n = \sum_{k=1}^n k \quad (15.9)$$

Geometry

VI

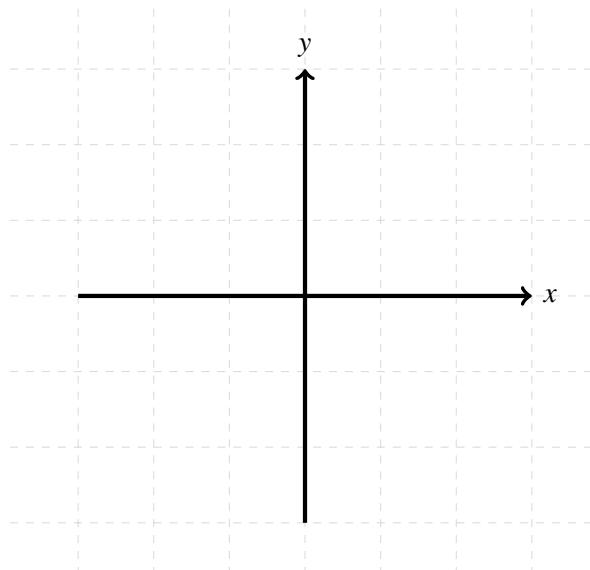
- 16 Coordinate Systems** *(.5pc).115
 - 16.1 Cartesian Coordinate System
 - 16.2 Geographic Coordinate System

- 17 Figures** *(.5pc).121
 - 17.1 Hexagon

16. Coordinate Systems

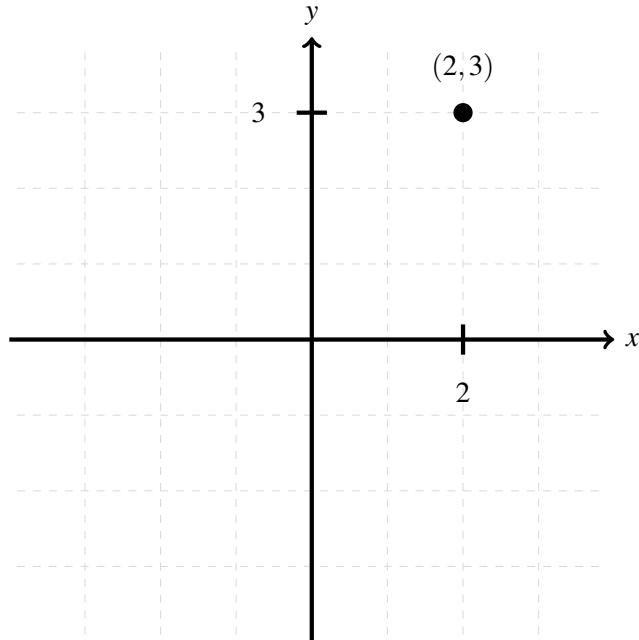
16.1 Cartesian Coordinate System

In the $XVII^{th}$ century, René Descartes provides for the first time in history a way to link Euclidean geometry and algebra through a coordinate system able to represent geometries described by Cartesian equations. The basic idea is that each point can be represented on the plain by a set of numerical coordinates. More formally these coordinates specify the signed distance of the point with respect to two fixed perpendicular lines. The point where these lines cross is called origin and has coordinates $(0,0)$. In this way is pretty easy to represent figure and functions in two or more dimensions. Moving from a 2-dimensional to a 3-dimensional or even more dimensions can be difficult to imagine but still easy to represent. A point in a 3-dimensional space will be represented by 3 variables and has the origin in $(0,0,0)$ and so on. The following figure represent a simple 2-dimensional plane:



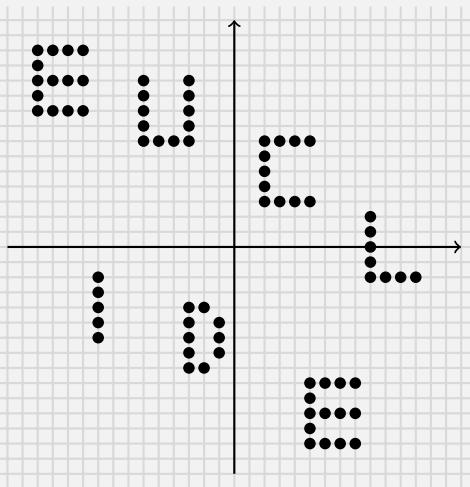
16.1.1 Points

Points are the most simple objects that can be represented in a coordinate system. As said before it is sufficient to specify a pair of coordinates to uniquely identify a point in the plane. As an example the point A(2,3) is represented as follows:



Challenge 16.1 — Points. (-13,13)(-13,12)(-13,11)(-13,10)(-13,9)(-12,13)(-11,13)(-10,13)
 (-12,11)(-11,11)(-10,11)(-12,9)(-11,9)(-10,9)(-6,11)(-6,10)(-6,9)(-6,8)(-6,7)(-3,11)(-3,10)(-3,9)
 (-3,8)(-5,7)(-4,7)(-3,7)(2,7)(2,6)(2,5)(2,4)(2,3)(3,7)(4,7)(5,7)(3,3)(4,3)(5,3)(9,-2)(9,-1)(9,0)(9,1)
 (9,2)(10,-2)(11,-2)(12,-2)(-9,-2)(-9,-3)(-9,-4)(-9,-5)(-9,-6)(-3,-4)(-3,-5)(-3,-6)(-3,-7)(-3,-8)(-2,-4)
 (-1,-5)(-1,-6)(-1,-7)(-2,-8)(5,-13)(5,-12)(5,-11)(5,-10)(5,-9)(6,-13)(7,-13)(8,-13)(6,-11)(7,-11)
 (8,-11)(6,-9)(7,-9)(8,-9) ■

Solution 16.1 — Points. Just draw coordinates:



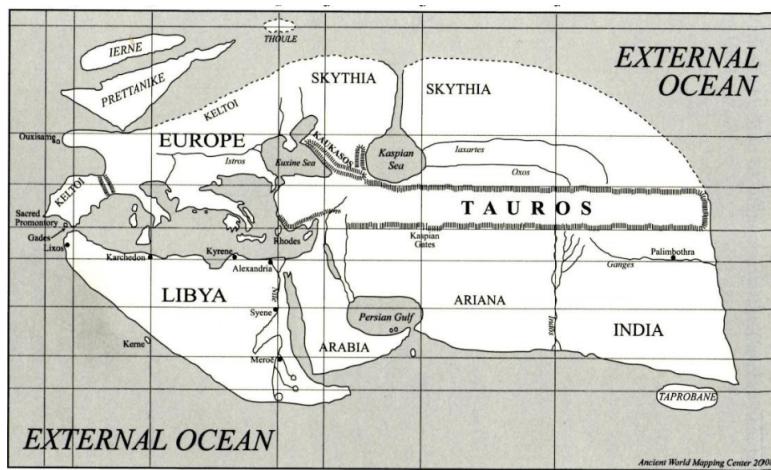
16.2 Geographic Coordinate System

16.2.1 Geo Points



When we talk about geographical points we immediately think about the most common and actual implementation in the world: Georeferentiation. With the advent of modern technologies, such as GPS devices, and their accessibility, almost all is able to identify a point on a geographic coordinate system, a coordinate system that identifies every location on Earth with a set of numbers, letters or symbols. Even if it can seem a modern technology, the first geographic coordinate system is credited to Eratosthenes of Cyrene in the far 3rd century BC (a reconstruction of the map is provided in the figure below). The 3 dimensions of the geographic coordinate system usually refers to latitude, longitude and elevation. To identify a point on a map it is first necessary to perform a map projection of latitudes and longitudes from a sphere or ellipsoid (the earth) to a plane, but how this is done is another story. No matters how projection is done or

which geodetic datum is used, neither the coordinate format and the symbols used to represent them. What cares is that, using the concept of latitude and longitude, we are able to represent the equivalent of a Cartesian point (the Geo point) to the equivalent of a Cartesian coordinate system (the Geographic coordinate system). This means that if Cartesian points can hide information, also Geo points can hide information in the same way as we seen in the previous subsection, but not only.



When we refer to Geo points we usually refer to specific location on the earth. Humans tend to identify things giving them at least a name, and locations have received, during years, the same treatment. Based on the level of granularity we choose a Geo point can belong to a continent, a state, a geographic region, a city, a route. This extends the possibility of hide information in a Geo point in different ways. Consider for example the following message containing also a list of Geo points:

we should kill:

- 54°37'02.8"N 7°52'37.1"W
- 53°34'30.5"N 7°23'33.6"W

- 54°36'42.4"N 6°24'38.4"W
- 54°07'10.9"N 8°03'18.1"W
- 53°39'45.6"N 7°13'43.5"W
- 53°39'23.6"N 7°22'29.4"W

All the points identify Irish lakes. The first one is the Lough Derg, the second is the Lough Owen, the third is Lough Neagh, the fourth is Lough Allen, the fifth is Lough Lene and the last one is Lough Derravaragh. Can you figure out what is hidden in these Geo points? The first letter of each lake give us a name: Donald.

Challenge 16.2 — Locations.

39°47'32.8"N 21°45'41.5"E
 58°50'49.9"N 25°48'12.7"E
 21°05'13.6"N 56°37'50.6"E
 39°10'33.3"N 8°04'22.5"W
 21°05'13.6"N 56°37'50.6"E
 42°33'52.1"N 12°38'51.0"E
 27°55'13.5"N 84°03'52.0"E
 8°30'55.4"S 179°06'22.4"E
 13°39'28.4"S 172°26'52.5"W



Solution 16.2 — Locations.

39°47'32.8"N 21°45'41.5"E = Greece
 58°50'49.9"N 25°48'12.7"E = Estonia
 21°05'13.6"N 56°37'50.6"E = Oman
 39°10'33.3"N 8°04'22.5"W = Portugal
 21°05'13.6"N 56°37'50.6"E = Oman
 42°33'52.1"N 12°38'51.0"E = Italy
 27°55'13.5"N 84°03'52.0"E = Nepal
 8°30'55.4"S 179°06'22.4"E = Tuvalu
 13°39'28.4"S 172°26'52.5"W = Samoa
 GEOPOINTS

Trolling with WGS 84

The system used in the previous section, the WGS 84, is the most used standard in coordinate systems used for cartography and GPS. Its boundaries go from 90N to 90S and from 180E to 180W. This means that most of the ASCII characters are covered when location are encoded in WGS 84. This can be used to hide information in a different way. Consider the following location:

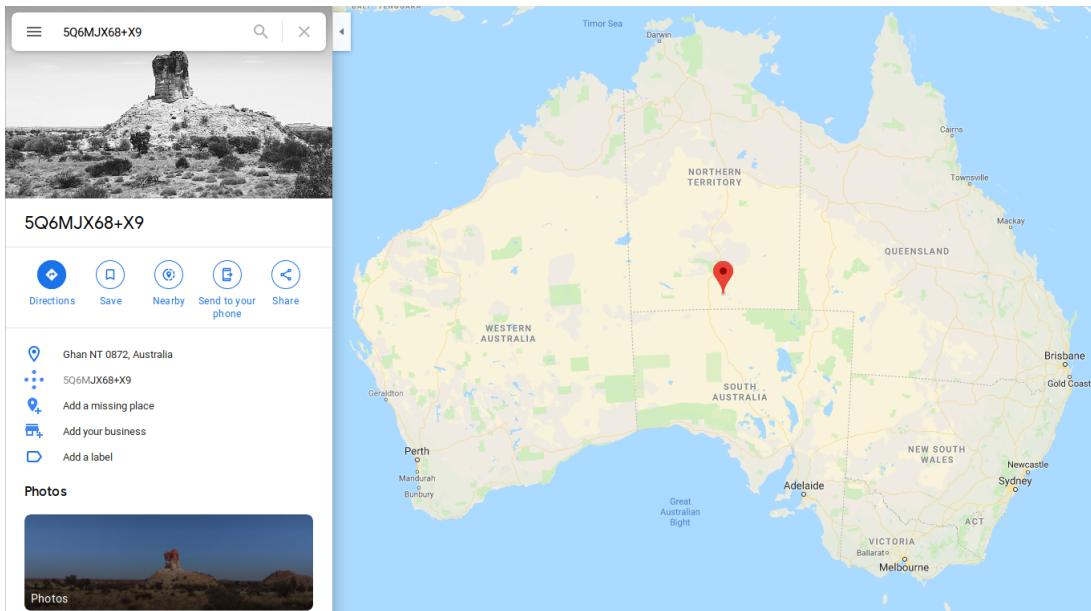
68°21'02.6"N 97°32'01.4"E

It can be refer to Ozero Lyuksina lake in Russia or, taking the degrees of latitude and longitude and converting them to ASCII characters, to the bi-gram "Da". This creates more ambiguity in decoding phase and opens coordinate information hiding to new fancy ways.

Open Location Code (OLC)

The open location code (OLC) born in the Google lab of Zurich in 2014 due to the necessity to encode locations in a more compact, easier and user friendly way than showing coordinates in the usual latitude-longitude format. As latitude-longitude system, near places have similar OLC codes

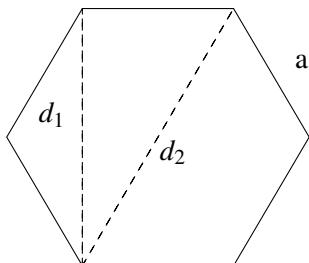
and they are starting to be supported by different applications like Google maps, but they are also accepted as postal addresses in some country like Cape Verde. OLC is only 10 characters long and it is composed by the first 4 characters representing the area code (a region of 100x100 kilometers), and the remaining characters representing the local code (a region of 14x14 meters inside the region specified by the area code). For more precision purposes and additional character can be added at the end of the OLC restricting the area to a 3x3 meters region.



17. Figures

17.1 Hexagon

Hexagon ...



$$d_1 = \sqrt{3} * a \quad (17.1)$$

$$d_2 = 2a \quad (17.2)$$

$$\text{Perimeter} = 6a \quad (17.3)$$

$$\text{Area} = \frac{3}{2} * \sqrt{3} * a^2 \quad (17.4)$$

(17.5)

Challenge 17.1 — Bees and honey. Mark assert that is able to calculate the liters of honey contained in a honeycomb just starting from its deep and from the product of its shortest and longest diagonals. Suppose that the honeycomb is equal to 5 cm deep and that the product of diagonals is equal to 240 cm. Are you able to tell me how many liters of honey fits into one honeycomb? ■

Solution 17.1 — Bees and honey. First of all we must recall that the volume of a regular solid is given by:

$$\text{Volume} = h * \text{Area} \quad (17.6)$$

So in our case:

$$\text{Volume} = 5 * \frac{3}{2} * \sqrt{3} * a^2 \quad (17.7)$$

From the riddle we know that the product between the long diagonal and the short one is 240:

$$d_1 * d_2 = 240 \quad (17.8)$$

$$(\sqrt{3} * a) * (2a) = 240 \quad (17.9)$$

$$2\sqrt{3} * a^2 = 240 \quad (17.10)$$

$$\sqrt{3} * a^2 = 120 \quad (17.11)$$

$$a^2 = \frac{120}{\sqrt{3}} \quad (17.12)$$

Now we can substitute (17.12) into (17.6):

$$\text{Volume} = 5 * \frac{3}{2} * \sqrt{3} * \frac{120}{\sqrt{3}} \quad (17.13)$$

$$= 5 * 3 * \sqrt{3} * \frac{60}{\sqrt{3}} \quad (17.14)$$

$$= 5 * 3 * 60 = 900 \text{ cm}^3 \quad (17.15)$$

Converting to liters we obtain 0.9 liters (1 liter == 1000cm³)

VI

Puzzles

- 18** **Japanese puzzles** *(.5pc).125
18.1 Nonogram
- 19** **Mathematical puzzles** *(.5pc).127
19.1 Petals Around the Rose

18. Japanese puzzles

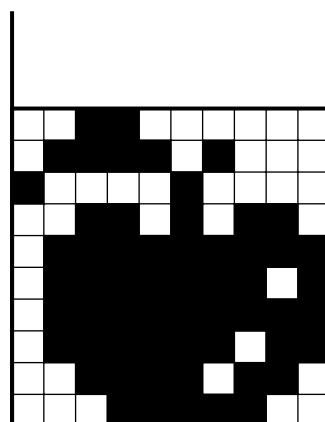
18.1 Nonogram

Nonogram puzzle born in 1987 from an idea of Non Ishida and Tetsuya Nishio whom in the same time, curiously, had the same idea. First painted nonograms appear in 1988 in Japan and in 1990 in the UK.

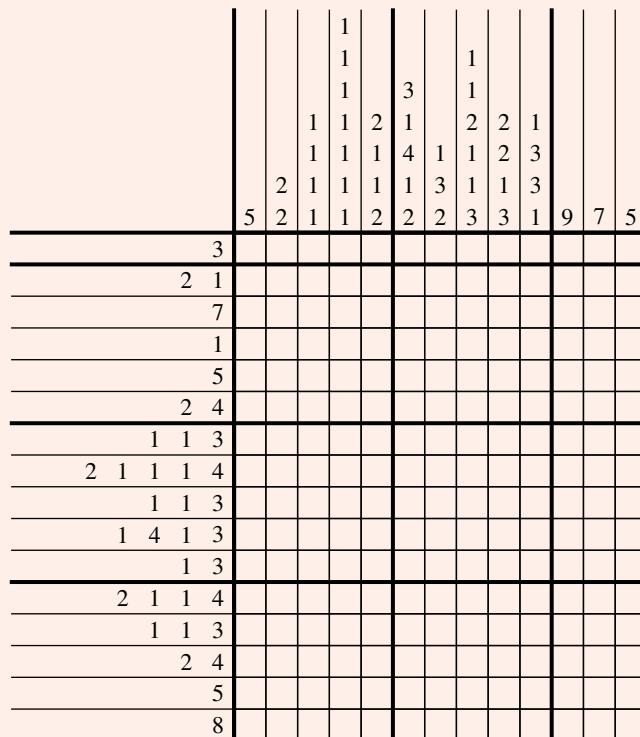
Nonograms are logic puzzles in which cells in a grid have to be filled or left blank according to the numbers at left and top of the grid. Filling the right cells will reveal the hidden picture. Clue numbers represent the numbers of consecutive filled cells are there in a row or column, spaces between numbers represent one or more spaces between 2 filled cells. For example the clue "2 4 1 3" means that there is a sequence of 2 filled cells, followed by an undefined sequence of non-filled cells, followed by a sequence of 4 filled cells, followed by an undefined sequence of non-filled cells, followed by 1 filled cell, followed by an undefined sequence of non-filled cells, followed by 3 consecutive filled cells. The order matter.

Nonograms have no theoretical limits on size, and are not restricted to square layouts.

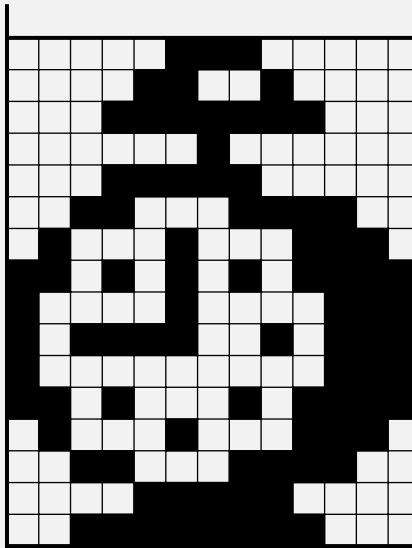
| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|--|
| | | 1 | 2 | 2 | 1 | | 1 | 4 | 4 | 2 | |
| 1 | 4 | 6 | 7 | 6 | 8 | 1 | 2 | 3 | 4 | | |
| 2 | | | | | | | | | | | |
| 4 | 1 | | | | | | | | | | |
| 1 | 1 | | | | | | | | | | |
| 2 | 1 | 2 | | | | | | | | | |
| 9 | | | | | | | | | | | |
| 7 | 1 | | | | | | | | | | |
| 9 | | | | | | | | | | | |
| 6 | 2 | | | | | | | | | | |
| 4 | 2 | | | | | | | | | | |
| 5 | | | | | | | | | | | |



Challenge 18.1 — One more minute please. Solve the following nonogram:



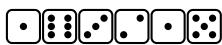
Solution 18.1 — One more minute please. Solved:



19. Mathematical puzzles

19.1 Petals Around the Rose

The first proposed mathematical challenging puzzle is called Petals Around the Rose where you are asked to guess a number that could be 0 or even. Number have to be guessed by looking at the result of a roll of some dice. The name of the game is a huge hint that helps to solve it, are you ready?



The answer is 6



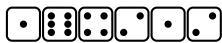
The answer is 12



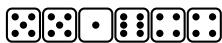
The answer is 2



The answer is 4



The answer is 0



The answer is 8



The answer is 4

19.1.1 Solving the Rose

As the name suggests we have to count how many petals are around the roses. In our case roses are simply die with upper face containing a center dot so that odd faces as 1, 3 and 5 are roses and even faces as 2, 4 and 6 are not.

The petals of a rose are represented by dots that surround the central dot. There are no petals on the 1 face, so it also counts as zero. There are two petals on the 3 face and four 5. Thus the solution is quite easy to compute: each face 3 we count two petals and each face 5 we count 4 petals. Dummy one when known.



Programming

20 **Esoteric Languages** *(.5pc).131
20.1 Brainf*ck Family (1993)

21 **Obfuscation** *(.5pc).135
21.1 Perl Obfuscation

Bibliography *(.5pc).137
Articles
Books



20. Esoteric Languages

During the computer age, lot of programming languages have been created. Some of them aims to solve some specific problem, others try to solve the limitations of other programming languages. There is a special category of programming languages in which belong those languages that were experimental, fast proof of concepts or simply designed to be jokes. These languages are called esoteric languages, esolangs for friends.

20.1 Brainf*ck Family (1993)

Brainf*ck is one of the most popular esolang. It has been created by Urban Muller in 1993 and it is very minimalist, composed only by 8 commands mapped to 8 different symbols, as explained in the following table:

| | |
|---|---|
| > | increment the data pointer (to point to the next cell to the right). |
| < | decrement the data pointer (to point to the next cell to the left). |
| + | increment (increase by one) the byte at the data pointer. |
| - | decrement (decrease by one) the byte at the data pointer. |
| . | output the byte at the data pointer. |
| , | accept one byte of input, storing its value in the byte at the data pointer. |
| [| if the byte at the data pointer is zero, then instead of moving the instruction pointer forward to the next command, jump it forward to the command after the matching] command. |
|] | if the byte at the data pointer is nonzero, then instead of moving the instruction pointer forward to the next command, jump it back to the command after the matching [command. |

The Brainf*ck language uses some components of a simple machine: a tape consisting of cells (all initialized to zero), a dynamic pointer (initialized at left most cell of the tape), an input tape and

an output tape. As an example the following program will output "brainfuck":

-
1. ++++++.
 2. +++++.
 3. -----.+++++.++++.
 4. -----.
 5. +++++++.
 6. -----.
 7. ++++++.
-

In this example, the pointer never moves from the cell zero, increasing it first 98 times, outputting its ASCII equivalent char, "b". Then increase till reach char "r", then decreasing till char "a" and so on. At the end of each line the output function is called. The classical "Hello World" is reached with the following string:

```
++++++[>++++[>++>+++>+++>+<<<-] >+>+>->>+[<]<-]>>. >---.+++++. .+++. >>. <-.<. +++. -----. -----. >>+. >++.
```

20.1.1 Brainf*ck Variants

The popularity of Brainf*ck, since its creation, increased rapidly and many variants has been developed. Simplest variants consists in some replacement of original symbology. Some of them can be seen in the following table:

| Brainf*ck | Alphuck | Blub | Ook! | Pikalang | Triplet | Ternary |
|-----------|---------|-------------|-----------|----------|---------|---------|
| > | a | Blub. Blub? | Ook. Ook? | pipi | 001 | 01 |
| < | c | Blub? Blub. | Ook? Ook. | pichu | 100 | 00 |
| + | e | Blub. Blub. | Ook. Ook. | pi | 111 | 11 |
| - | i | Blub! Blub! | Ook! Ook! | ka | 000 | 10 |
| . | j | Blub! Blub. | Ook! Ook. | pikachu | 010 | 20 |
| , | o | Blub. Blub! | Ook. Ook! | pikapi | 101 | 21 |
| [| p | Blub! Blub? | Ook! Ook? | pika | 110 | 02 |
|] | s | Blub? Blub! | Ook? Ook! | chu | 001 | 12 |

For completeness, here are the Hello World! scripts of these programming languages:

```
iiciccepepcceaiiaiaiaicccsascciijceeeeeejccijjccjcejaajaajcccjeeejaajaaijcccej
```

```
blub. blub? blub. blub.
blub. blub. blub. blub. blub. blub. blub. blub. blub? blub? blub. blub. blub. blub.
blub. blub. blub. blub. blub. blub. blub. blub. blub. blub. blub. blub. blub. blub.
blub. blub. blub. blub. blub? blub! blub! blub? blub! blub? blub. blub. blub!
blub. blub. blub? blub. blub.
blub. blub. blub. blub. blub! blub? blub? blub? blub. blub. blub. blub. blub. blub.
blub. blub. blub. blub. blub. blub. blub? blub! blub! blub? blub! blub? blub?
blub. blub. blub. blub! blub. blub. blub. blub. blub. blub. blub. blub. blub. blub.
blub. blub. blub. blub. blub. blub! blub. blub. blub! blub. blub! blub. blub. blub.
blub. blub. blub. blub. blub. blub! blub. blub? blub. blub? blub. blub? blub.
```

```

blub? blub. blub.
blub. blub. blub. blub. blub! blub? blub? blub. blub. blub. blub.
blub. blub. blub. blub. blub. blub? blub! blub! blub? blub! blub?
blub. blub! blub. blub. blub? blub. blub? blub. blub. blub. blub.
blub. blub. blub. blub. blub. blub. blub. blub. blub. blub. blub.
blub. blub. blub. blub. blub! blub? blub? blub. blub. blub. blub.
blub. blub. blub. blub. blub. blub. blub. blub. blub. blub. blub.
blub. blub. blub. blub. blub. blub. blub. blub. blub. blub. blub.
blub. blub. blub. blub. blub? blub! blub! blub? blub? blub. blub!
blub! blub! blub! blub! blub! blub. blub? blub. blub? blub. blub?
blub. blub? blub. blub! blub. blub. blub. blub. blub. blub. blub!
blub. blub! blub! blub! blub! blub! blub! blub! blub! blub! blub!
blub! blub! blub. blub! blub! blub! blub! blub! blub! blub! blub!
blub! blub! blub! blub! blub! blub! blub! blub! blub? blub.
blub? blub. blub! blub.
```

Ook. Ook? Ook. Ook.
Ook. Ook. Ook. Ook. Ook! Ook? Ook? Ook. Ook. Ook. Ook. Ook. Ook. Ook.
Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook? Ook! Ook! Ook? Ook!
Ook? Ook. Ook! Ook. Ook. Ook? Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook.
Ook. Ook. Ook! Ook? Ook? Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook.
Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook.
Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook! Ook. Ook. Ook? Ook. Ook?
Ook? Ook. Ook.
Ook. Ook. Ook! Ook? Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook.
Ook? Ook! Ook? Ook! Ook? Ook. Ook. Ook. Ook. Ook? Ook. Ook? Ook.
Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook.
Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook! Ook.
Ook! Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook! Ook. Ook. Ook? Ook. Ook?
Ook? Ook. Ook.
Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook.
Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook.
Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook.
Ook! Ook? Ook! Ook. Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook? Ook.
Ook? Ook. Ook? Ook. Ook. Ook! Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook. Ook!
Ook. Ook! Ook!
Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook!
Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook! Ook!
Ook! Ook! Ook. Ook? Ook. Ook. Ook! Ook.

pi pi pi pi pi pi pi pika pipi pi pi pi pi pi pipi pi pi pi pi
pi pi pi pi pipi pi pi pipi pi pichu pichu pichu ka chu pipi pi
pi pikachu pipi pi pikachu pi pi pi pi pi pikachu pikachu pi pi pi
pikachu pipi pi pi pikachu pichu pichu pi pi pi pi pi pi pi pi pi
pi pi pi pikachu pipi pikachu pi pi pi pikachu ka ka ka ka ka ka pikachu ka
ka ka ka ka ka ka pikachu pipi pi pikachu pipi pikachu

| | | | | | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 001 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 110 | 100 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 |
| 001 | 000 | 011 | 100 | 010 | 001 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 |
| 001 | 000 | 011 | 100 | 111 | 010 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 010 | 010 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 010 |
| 001 | 001 | 001 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 110 | 100 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 001 |
| 011 | 100 | 010 | 001 | 001 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 110 | 100 | 111 | 111 |
| 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 001 | 000 | 011 | 100 | 000 | 000 | 000 | 010 | 000 | 000 | 000 | 000 | 000 | 010 | 100 | 100 | 100 |
| 100 | 010 | 111 | 111 | 111 | 010 | 000 | 000 | 000 | 000 | 000 | 010 | 000 | 000 | 000 | 010 | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 | 000 |
| 000 | 000 | 010 | 001 | 001 | 111 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 | 010 |

111111111111111102011111111020111110111111101111111011100000000101201110111011
001011102001200101200110101020111111111111120201111120010120001020002011

11112010101010102010101010101010200101112001111120

20.1.2 Brainf*ck Extensions

Brainf*ck has been designed to be a very minimalist programming language but it has been widely extended with new functionality.

2D-Brainf*ck

The bi-dimensional view of brainf*ck can be achieved by using multi-tapes instead a single one. In this case the pointer is not only restricted to move right and left, but also up and down through the tapes. Usually 3 new symbols are added in the simplest extension:

| | |
|---|--|
| Λ | move the pointer up to the previous tape |
| ∨ | move the pointer down to the next tape |
| × | exit the program |

21. Obfuscation

Code obfuscation techniques are very popular nowadays.... What is obfuscation...

21.1 Perl Obfuscation

21.1.1 JAPH (Just Another Perl Hacker)

Simply speaking JAPH is a program written in perl that does only a single thing: it prints the string "Just Another Perl Hacker". JAPH programs are usually used to show how far obfuscation can go and how can be creative. The first JAPH program was written by Randal L. Schwartz in 1988 and was very simple and no obscure at all:

```
print "Just another Perl hacker,"
```

21.1.2 Other perl programs

Obfuscation in perl provides use some other funny examples. The next one will print "Just another Perl programmer"

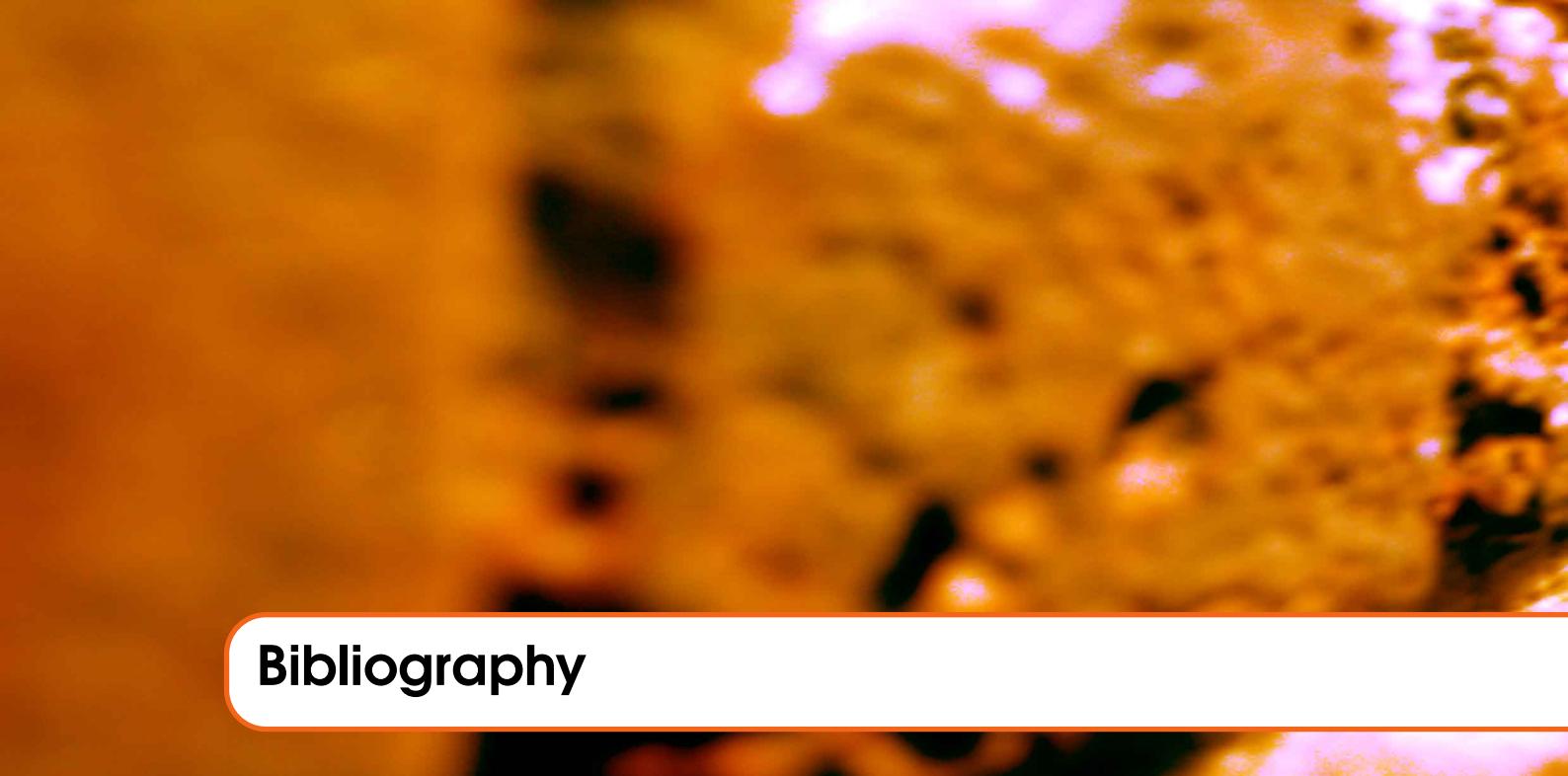
```
undef$/;$_=<DATA>;y/ODA\n / /ds;@yoda=map{length}split;print chr
oct join(' ',splice(@yoda,0,3))-111 while@yoda;
__DATA__
    0000000000000000    0000        DD000000000
    OD00000000000000    000000        000000000000
    0000    000        000  000        00D    0D0
    0000    000        00D    000        D00D00000D
    0000    D0D        0000000000000    000  0000
DD0000D00000    000        000000000D00D0    00D    D000000D0000
00000D00000    000        0000        00D0    00D    00000000D00
                                         0000000000000
                                         0000000000000
                                         0000000000000
                                         0000000000000
```

```
0000000000000000 000      D00      D00D000000      00D00
D0000 00000 0000000000D0 000 0000      00D00
000 000 00000D00000000 000 0D00000000000000000000
0      0 000D      0000 000 0000000000000000
```

or again, the following camel will produce 4 other camels, amzing!

```
use strict;
```

```
$_=`ev
al("seek\040D
ATA,0,
{<DATA>;}my
my$Camel ;while(
9s",$_);my@dromedary
_=<DATA>{@camel1hump;my$camel;
<DATA>){$_=sprintf("%-6
1=split//;if(defined($
ry1){my$camel1hump=0      ;my$CAMEL=3;if(defined($_=shif
t(@dromedary1 ))&&/\S/){$camel1hump+=1<<$CAMEL;}
$CAMEL--;if(d efined($_=shift(@dromedary1))&&/\S/){
$camel1hump+=1 <<$CAMEL;}$CAMEL--;if(defined($_=shift(
@camel1hump))&&/\S/){$camel1hump+=1<<$CAMEL;}$CAMEL--;if(
defined($_=shift(@camel1hump))&&/\S/){$camel1hump+=1<<$CAME
L;};$camel.==(split("//,".\040..m'{/J\047\134]L^7FX"))[$camel1h
ump];}$camel.=`\n";}@camel1hump=split(/\n/,$camel);foreach(@
camel1hump){chomp;$Camel=$_;y/LJF7\173\175'\047/\061\062\063\
064\065\066\067\070/;y/12345678/JL7F\175\173\047'/;$_=reverse;
print"$_\040$Camel\n";}foreach(@camel1hump){chomp;$Camel=$_;y
/LJF7\173\175'\047/12345678/;y/12345678/JL7F\175\173\0 47'/;
$_=reverse;print"\040$_$Camel\n";};$s/\s*//g;;eval; eval
("seek\040DATA,0,0");undef$/;$_=<DATA>;$s/\s*//g;();$s
;^.*_;;map{eval"print\"$_\"";}.{4}/g; __DATA__ \124
\1 50\145\040\165\163\145\040\157\1 46\040\1 41\0
40\143\141 \155\145\1 54\040\1 51\155\ 141
\147\145\0 40\151\156 \040\141 \163\16 3\
157\143\ 151\141\16 4\151\1 57\156
\040\167 \151\164\1 50\040\ 120\1
45\162\ 154\040\15 1\163\ 040\14
1\040\1 64\162\1 41\144 \145\
155\14 1\162\ 153\04 0\157
\146\ 040\11 7\047\ 122\1
45\15 1\154\1 54\171 \040
\046\ 012\101\16 3\16
3\15 7\143\15 1\14
1\16 4\145\163 \054
\040 \111\156\14 3\056
\040\ 125\163\145\14 4\040\
167\1 51\164\1 50\0 40\160\
145\162 \155\151
\163\163 \151\1
57\156\056
```



Bibliography

Articles

Books

