

RGHLXWERJFEBKJCTPQSRUATXSECVANAMVCINKFDITERSWSKYZGWPCFF
OFDOKNIXNQEKLERJOOHYHNRNLCRTFPGBLQSLRRIZGNPGNIRBCPEYFR
ZQIYEEFSGATAQUILXWVEKWKVKGLPXLSSXDDKEPLTBYFVEOIBVONXCZ
DNSDUBZSCVCYOPPWWZKWJEBNDZIRDTEPPOAGCMEAPVFQKAHLDBGHUVYF
PROLVATDHAZXUIBOIPQBYFKMEPPVVARYWVDWAXQNXDVPDJLRLCIZJZT
XTOHGEIJVRIZXKSRDPYVCTAKEGKQNLDERCEDRCLCCBZODUFQYKKGFVH
KIJQOATHAJDDBYJHVHYGJBPYDFCFSGBCVLXLMKPYHQVJNRXRTBXFEBZ
CJROUGTJTFCZRANOMSOELCISWAAZRFXRTATPUDSXSQAWFHVPUKGBGRGF
KDJJLCPAPMSCRLSNMVIYCJQNTYKDJYCFLZYGWYMIEKDEOKDGGHIUYVU
MPYAJLMIRHIQJUMHHPQFWDRCNTXFFZGKYEQQHANWZZGLXUBFVSOITY
KNMVEQIHPGWSXOHTRBUAPLIXFFDVVTUSBMXBIUFYDTGRWNRBLEKVUU
PQZCTHDXKMAFJAEBOOKYONGCLASSICALPCRYPTOGRAPHYPYRKBLPJZGA
EMXDVWKZZAEQSLZZQRJNPXBYLMADNESSCTDHTGNOVLJVYICAIYFYFHG
RYTARUVQRPKEJYGBEAEODQYCUTZKPGLVCLBVPQXSTCTQXEOQPJUJA
SKXJECBYOKMAOWPSKORUKVJALEVGUKCMCZHSLSVQABOFSRDZDEMCKXO
BLNPMVYWHNKKOKWPDHMFNRFEFRAWSLLLHDNRHULIIZINDCSTLSXIDA
CDHFSWEMAEEJYEUYKTGIIKBXKXBCDRGKLMBDQDKHTGUTVRQWVFKPW
VBZMMYYPKJDDOCYTXXUANBVRYQHLLIVHXUFOKHNIZTRIFAWIJYYBSPUWI
BSFPGRGUYOZXFJEGXCXDZCASLGATBUCNLFDEFQXQEDDLMOGNWBXILJZ
XNIEOAOPZCSMMSOHEAHKLJGZGBYPEOXBNXYRTBCGYHFKBABNPBNJQJS
ASAMDHDCRBCEUABZWELSYAAKHVWMIOXDZLWJATETKKBWURJZXOQADHK
RTWNOHUYABNJTZBLLUFCABCIVAZMFRQXOVXKAUNGIXHJPRDYUDDRZBH
CVFXWMYGEXBSVCXXGPPHDFGNIGEZCLEBTSQTZVJMXANBJTARBNSOEYV
CKGJPIAWDCOBHNHWDIBEZSAAPPYEEPIYRIIDGPNLPJKSMKMPYHGHINO
UKJWVLIGMJTWDLGSIKCYDZXQKTIUHSHFQIZGALHUKLJYWMLZKAHTKY
YSVTZSNCNYYMEKUFFLVFQVYZUVKLBIATHFIUCOCZXIBFKAYXHYKHVFZ
HZCCEQZMOWGNWTTQSTWWGUZBVDVKRCXDHBNQWUTPTMWSWMORMGCSB
UXVKDFPNXGBHIGEHQPLAMEEZAMRYGOVMXSHEIVBTWFZBCLZAMSZCMR
RYTNPZBCCXXRBTWSYLSHAYYOWIBMUPVDNPLKHFRVFBQYVLUHQRPRNBJ
MQGUKAWXDFOYUHFBNKLIWVIGPPVXMXJNVWGNJYKEPETARKUBCUSE
NVPLXHJXOZIDFJDTCMHXSFFQYLCZSZKTHGRYEGLKAYIKEDELBNHXYOW
HRNYTGTIEXQBBHRQNCIAVFECUPPQDGBDAKGUZCVRZKEFHSSIPXEQUZDV
YHLDBEFFJVDCEPBRQOXDPDBGSKJPENTXKXQKVHSPBJKXEFNFCDVHVB
TINYQWSWKHHXTQLSXFENSAUXSWCPONSMJQJRUTUFZFBIEQWURUGAR
GKKNJBBAJRKBKONNLOEHVLOSHZLBVSAPYBLZAQNIWFRERAYVLKAMHW
BRVBMWCWZAMJUIJAPXAORHSRJOXDCDCMMHAXLDJEXNKGUAHXHXUAD
USYNCWGJSUDCHILOARNSVCZTYNYAIRCDCEXYMVVJGLSCHCFCCAKRDM
VPWJMTDHDIZQIAUMASURABZGQXLLZQGDVONRWLTFXFCAXKIOXPEJEW
GRKRQJEFWLXECDAEBAIWIYBQYUGNWUYOMDDJAZAXEUAHGHJVCTN
RANPRATWUFWKHJRRPDYSPAUEHIOJLLOOPTJMJIPVVMRBSTSTQTKIOJHJQ
AWYUOOOUWOEMJMMADODRWSHTDZACWDUGNOMKAJJYKCNYSQMSOEF
MUGLSWNTTBMLJHBHGBZWBKLATDDCIWOSQNUJZAOUKJPBDWAEJSAPIW
LBZIIIRPIJYCTNYKSMCCWAVKNQUKTNYKMBFLPCKASFTVCVGALIJBPHMU

A Book on
Classical Cryptography
by madness

madness's book on classical cryptography
copyright page
last modified 2020-12-17
©2020 madness

Copyright 2020 by madness.
All rights reserved. And we mean ALL.

The ciphertexts, and corresponding plaintexts, for these exercises are from the British National Cipher Challenge and are copyright by the University of Southampton. They are used with permission.

Unit 34, Exercise 1
Unit 34, Exercise 2
Unit 44, Exercise 2
Unit 56, Exercise 1
Unit 59, Exercise 3
Unit 63, Exercise 4
Unit 65, Exercise 4
Unit 68, Exercise 2
Unit 69, Exercise 2
Unit 71, Exercise 2
Unit 79, Exercise 2
Unit 89, Exercise 6
Unit 97, Exercise 2
Unit 105, Exercise 3
Unit 117, Exercise 6
Unit 125, Exercise 2

Scrabble™ is a trademark of Hasbro in the United States and Canada, and of Mattel elsewhere.

Many of the example texts are from stories and novels that are no longer under copyright.

Photo credits

Unit 105

Exercise 4: a still frame from the TV show Futurama

Exercise 6: inside cover from Ozzy Osbourne's album *Speak of the Devil*

Unit 120

Wadsworth cipher disk: U.S. National Security Agency

Wheatstone Cryptograph: eBay

Urkryptografen: Museum of Cypher Equipment, Fife, Scotland

Unit 124

Bazeries cylinder: Étienne Bazeries

M-94: robbo@ev1.net

Unit 126

M-138: U.S. Department of Defense

Contents

Introduction

Part I: Linguistic data

- Unit 1: Textual corpora
- Unit 2: Word lists
- Unit 3: Monogram frequency tables
- Unit 4: Tetragram frequency tables
- Unit 5: The χ^2 statistic (optional)
- Unit 6: Monogram fitness based on the χ^2 statistic (optional)
- Unit 7: Angle between vectors
- Unit 8: Monogram fitness based on the angle between vectors
- Unit 9: Tetragram fitness
- Unit 10: Index of coincidence
- Unit 11: Entropy (optional)

Part II: Monoalphabetic substitution ciphers

- Unit 12: Monoalphabetic substitution cipher
- Unit 13: Atbash cipher
- Unit 14: Modular arithmetic: addition and subtraction
- Unit 15: Caesar shift cipher
- Unit 16: Brute-force attack on the Caesar cipher
- Unit 17: Attacking the Caesar cipher with cribs
- Unit 18: Attacking the Caesar cipher with monogram frequencies (χ^2) (optional)
- Unit 19: Attacking the Caesar cipher with monogram frequencies
- Unit 20: Greatest common divisor
- Unit 21: Modular arithmetic: multiplication and division
- Unit 22: Affine cipher
- Unit 23: Brute-force attack on the affine cipher
- Unit 24: Attacking the affine cipher with cribs
- Unit 25: Attacking the affine cipher with monogram frequencies
- Unit 26: Keyword substitution cipher
- Unit 27: Dictionary attack on the keyword substitution cipher
- Unit 28: Stochastic hill-climbing attack on monoalphabetic substitution ciphers

Part III: Periodic polyalphabetic substitution ciphers

- Unit 29: Periodic polyalphabetic substitution cipher
- Unit 30: Finding the period: Kasiski examination
- Unit 31: Finding the period with the index of coincidence
- Unit 32: Finding the period: twist method
- Unit 33: Vigenère cipher
- Unit 34: Brute-force attack on the Vigenère cipher
- Unit 35: Attacking the Vigenère cipher with cribs
- Unit 36: Dictionary attack on the Vigenère cipher
- Unit 37: Hill-climbing attack on the Vigenère cipher
- Unit 38: Attacking the Vigenère cipher as a periodic Caesar cipher
- Unit 39: Gronsfeld cipher (optional)
- Unit 40: Beaufort cipher
- Unit 41: Variant Beaufort cipher
- Unit 42: Porta cipher
- Unit 43: Periodic affine cipher
- Unit 44: Attacking the periodic affine cipher as a collection of affine ciphers
- Unit 45: Quagmire 1 cipher
- Unit 46: Two-stage attack on the quagmire 1 cipher
- Unit 47: Quagmire 2 cipher
- Unit 48: Quagmire 3 cipher
- Unit 49: Quagmire 4 cipher
- Unit 50: Hill-climbing attack on periodic polyalphabetic substitution ciphers

Part IV: Transposition ciphers

- Unit 51: Transposition ciphers
- Unit 52: Permutations
- Unit 53: Permutation cipher
- Unit 54: Heap's algorithm
- Unit 55: Factoradic numbers and permutations
- Unit 56: Brute-force attack on the permutation cipher
- Unit 57: Hill-climbing attack on the permutation cipher
- Unit 58: Matrix transposition
- Unit 59: Twisted scytale
- Unit 60: Columnar transposition cipher
- Unit 61: Double columnar transposition cipher
- Unit 62: Nihilist transposition cipher
- Unit 63: Railfence cipher
- Unit 64: Redefence cipher (optional)
- Unit 65: AMSCO cipher
- Unit 66: Myszkowsky cipher (optional)
- Unit 67: Cadenus cipher
- Unit 68: Hill-climbing attack on the Cadenus cipher

Part V: Grid-based ciphers

- Unit 69: Polybius cipher
- Unit 70: Playfair cipher
- Unit 71: Hill-climbing attack on the Playfair cipher
- Unit 72: Vertical two-square cipher
- Unit 73: Horizontal two-square cipher
- Unit 74: Hill-climbing attack on the two-square ciphers
- Unit 75: Four-square cipher
- Unit 76: Phillips cipher
- Unit 77: Hill-climbing attack on the Phillips cipher
- Unit 78: Phillips-RC cipher (optional)
- Unit 79: Double Playfair cipher
- Unit 80: Nihilist substitution cipher
- Unit 81: Bifid cipher
- Unit 82: Trifid cipher
- Unit 83: ADFGX cipher
- Unit 84: ADFGVX cipher

Part VI: Ciphers based on matrices

- Unit 85: Matrices and vectors
- Unit 86: Matrices over the set of residues
- Unit 87: Hill cipher
- Unit 88: Attacking the Hill cipher with cribs
- Unit 89: Affine Hill cipher

Part VII: Stream ciphers

- Unit 90: Stream ciphers
- Unit 91: Trithemius cipher
- Unit 92: Autokey cipher
- Unit 93: Hill-climbing attack on the autokey cipher
- Unit 94: Attacking the autokey cipher with monogram frequencies
- Unit 95: Running-key cipher
- Unit 96: Progressive Vigenère cipher
- Unit 97: Solitaire cipher
- Unit 98: Hill-climbing attack on the solitaire cipher with partially known key

Part VIII: Codes

- Unit 99: Codes
- Unit 100: Baconian cipher
- Unit 101: Triliteral cipher
- Unit 102: Morse code
- Unit 103: Monome-dinome cipher
- Unit 104: Straddling checkerboard cipher

Part IX: Miscellaneous ciphers

- Unit 105: Symbolic substitution
- Unit 106: One-time pad
- Unit 107: Slidefair cipher
- Unit 108: Nicodemus cipher
- Unit 109: Fractionated Morse
- Unit 110: Hutton cipher
- Unit 111: Scrabble cipher
- Unit 112: Homophonic substitution
- Unit 113: Polyphonic substitution
- Unit 114: Polyhomophonic substitution
- Unit 115: Pollux cipher
- Unit 116: Doubled-over substitution
- Unit 117: Duplicitous ciphers
- Unit 118: Combination-lock cipher
- Unit 119: Chase cipher

Part X: Proto-mechanical ciphers

- Unit 120: Disk ciphers (cipher clocks)
- Unit 121: Attacking cipher clocks with cribs
- Unit 122: Digram-counting attack on cipher clocks
- Unit 123: Hill-climbing attack on cipher clocks
- Unit 124: Cylinder ciphers
- Unit 125: Hill-climbing attack on cylinders
- Unit 126: Strip ciphers

Afterword

Index

Bibliography

Introduction

This book is an introduction to classical cryptography, with an emphasis on cryptanalysis. By *classical*, we mean cryptography that can be done with pen and paper. Historically, such ciphers were used for serious secret-keeping up to and into the Second World War, around which time mechanical ciphers came into use. Nevertheless, classical ciphers continue to be invented, even though the classical period has ended.

Let us begin with some definitions. *Cryptography* is the science of modifying a message so that its contents cannot be understood except by the intended recipient. A *cipher* is a system of modifying such a message to hide its meaning, which is to *encipher* it, and of later reversing that modification, which is to *decipher* it. By contrast, a *code* is a system whereby words and phrases are replaced with other symbols; the corresponding processes are called *encoding* and *decoding*. This book deals with ciphers, and very little with codes as defined this way (a modern use of the word *code* is to replace symbols in the plaintext with combinations of new ciphertext symbols; we will see ciphers of this type in this book). *Cryptanalysis* is the art of discovering the meaning in an enciphered message that was not intended for you. Doing so successfully is called *decryption*. Sometimes we use *encrypt* and *decrypt* as the same as *encipher* and *decipher*, even though their true meanings are subtly different. We may also will use *crack* or *break* for *decrypt*. The unmodified message is called the *plaintext* or the *cleartext*. The encrypted message is called the *ciphertext*.

Often, cryptographers will use the convention that plaintexts are written in lower-case letters, while ciphertexts are written in upper-case letters. Here, we will not be strict about this convention, especially since there are some ciphers that use both upper- and lower-case letters in their ciphertexts.

Everything in this book will be done in English (except for a rare word or two every now and then). Nevertheless, most if not all of what we learn here can be used for any language that uses the Latin alphabet. The only modifications necessary are in the linguistic data. For another language, Part I on linguistic data can be reworked in the new language to compile a new set of data for use by the methods of the latter parts of the book.

Although classical cryptography can be done with pen and paper, this does not mean that we should use pen and paper. An emphasis in this book is on the use of a computer to cryptanalyze a ciphertext. To that end, we recommend the Python language for its ease of manipulating text. (In modern cryptography, Python is also useful because it handles large integers seamlessly.) Since version 2 of the language is no longer maintained, any tips and examples in this book will use version 3 of Python. There are only three differences between these versions that are important for us: In version 3,

strings of characters and strings of bytes are different types of data. In version 3, the `print` statement has been replaced by a `print()` function, so that now parentheses are necessary. In version 2, the `/` operator behaved differently if it was dividing integers or floating-point numbers; in version 3 we will use two operators (`/` and `//`) for division. For this book, you should be able to write rudimentary Python scripts. We will provide tips as we go along, and you should expect to learn more about the language as we go. You will not need to be able to write object-oriented scripts.

To succeed in your study of classical cryptography, you should begin with Unit 1 and continue in order until the end of Part IV (with the possible exception of units marked as “optional”). Units build on each other, and if you skip any, you may find that you have missed something important. Once you reach the part on miscellaneous ciphers, you may do only the units that you wish to do. Along the way, you will be building your own library of functions and programs in Python for cracking ciphertexts.

The history of cryptography can be roughly divided into four eras. This book focuses on the first, the classical era. While in terms of a timeline the classical era ended decades ago, in a real sense it continues today. Classical ciphers continue to be invented. They continue to be studied. They continue to challenge cryptographers. So stop complaining that this book is useless.

Here is a short description of the four eras of cryptography:

- The *classical era* began at the start of time and runs up until the middle of the twentieth century (see above concerning why it continues today). It concerns ciphers that can be implemented with pen and paper. They can also be broken with pen and paper. They generally work on symbols that are the letters of the alphabet and/or the ten digits.
- The *mechanical era* ran from before World War II until the advent of computers. This era is characterized by the use of electric rotor machines. Each rotor contains a maze of wires, and each machine contains several rotors. A letter is enciphered by passing it through the rotors in order and reading the result from the last rotor. After each letter is enciphered, one or more of the rotors are rotated so as to change how the next letter is enciphered. The key for such a machine is the arrangement of rotors and their starting positions.
- The *modern era* is characterized by the use of computers. The symbol set on which modern ciphers act is the *bit*, which is a binary digit taking the value 0 or 1. In addition to *symmetric ciphers*, which are those that are enciphered and deciphered with the same key, the modern era introduces *asymmetric ciphers* (*public-key ciphers*), which use a different key for the two operations. This allows one to publish his/her public key, so that anyone can send a message that only s/he can decipher with the private key. This paradigm can be extended to include the ability to sign a document with one's private key, so that anyone can verify the signature with the public key.
- The *quantum era* exploits the laws of quantum physics. Its elementary unit is the *qubit*, which is a quantum state that can, when measured, take one of two values. The important idea from quantum mechanics that cryptography uses is the fact that when someone measures a qubit, it forces that qubit to be in a new state that may be different from the original. Quantum cryptography uses this principle to detect whether a stream of qubits has been intercepted. In

this way, it is possible to devise a scheme whereby the stream of qubits is used as a key for the one-time pad. The encrypted message is sent in some other, conventional, way.

Get to work!

Reading and references

Helen Fouché Gaines, *Cryptanalysis: a study of ciphers and their solution*, New York: Dover, 1956; previously titled *Elementary Cryptanalysis* and published by American Photographic in 1939; archive.org/details/cryptanalysis00gain; chapter I.

William F. Friedman, “Codes and Ciphers (Cryptology),” *Encyclopaedia Britannica*, 1956, www.nsa.gov/Portals/70/documents/news-features/declassified-documents/friedman-documents/reports-research/FOLDER_535/41772109081119.pdf

Whitfield Diffie and Martin E. Hellman, “New Directions in Cryptography,” *IEEE Transactions on Information Theory* 22 (1976) 644-654, ee.stanford.edu/~hellman/publications/24.pdf