

Codes and Nomenclators

A Systematic View



Klaus Schmeh (www.schmeh.org)

HistoCrypt 2018 Uppsala



The Top 50 unsolved encrypted messages: 34. Unsolved nomenclator messages

For centuries nomenclators were the most popular kind of encryption. Still today messages enciphered with a well-designed nomenclator are hard or even impossible to break.

Von [Klaus Schmeh](#) / 24. April 2017 / [5 Kommentare](#) / [Weiterlesen](#)



Nomenclator encryptions: centuries old, but still hard to solve

The British National Archive has published an encrypted letter from the 17th century. The solution is known, but many other encrypted notes of the same kind still wait to be

broken.

Von [Klaus Schmeh](#) / 13. August 2016 / [4 Kommentare](#) / [Weiterlesen](#)

If you have an unsolved encrypted text, I will be happy to publish it on my blog.

Agenda

1. Basics: scope, terminology, classification
2. How to detect a code
3. How to cryptanalyze a code
4. Open research questions
5. A few solved code cryptograms
6. A few unsolved code cryptograms

Code facts

Please correct, if wrong

Codes are not used any more today. So, the topic is only of historical interest.

No specialist for deciphering code cryptograms currently exists.

Stephen Bellovin currently seems to be the most renowned codebook expert.



Terminology

Please check the following website for the current version of the terminology discussed at the workshop and afterwards:

<http://scienceblogs.de/klauser-krypto-kolumne/2018/07/09/a-terminology-for-codes-and-nomenclators/>

Terminology

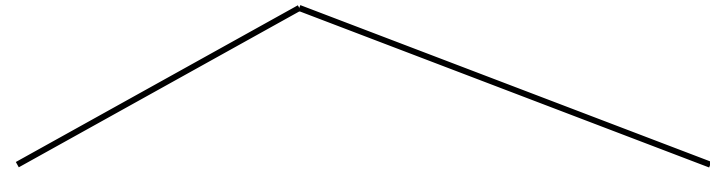
Please correct, if wrong

Cipher: encryption method that operates at the level of individual letters

Code: encryption method that operates at level of meaning

Nomenclator: small code

Codebook: big code



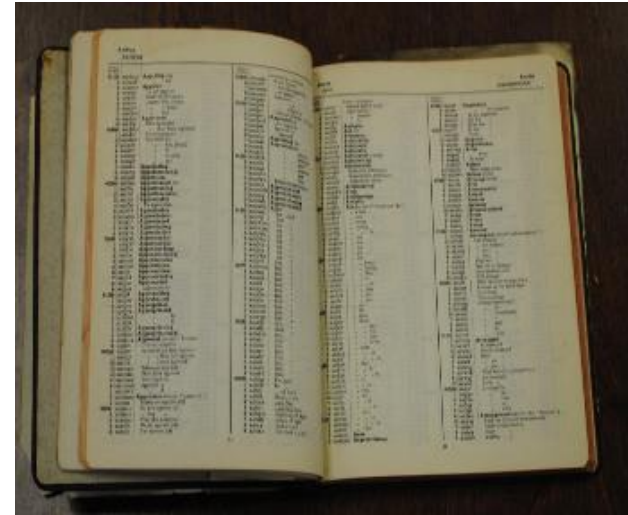
Terminology

Please correct, if wrong

- **Codegroup:** smallest unit of a code, may consist of several letters/numbers
- **Null:** meaningless codegroup
- **Nullifyer:** makes the previous codegroup meaningless
- **Homophones:** several codegroups with the same meaning
- **Wordgroup:** Codegroup that refers to a word
- **Lettergroup:** analog

Questions

What exactly is the difference between a nomenclator and a codebook?



Is there a better word for "code" ("code" has different meanings in cryptography and in other contexts)?

What about dictionary codes?

Dictionary code: a dictionary is used. Each word is referenced by its page and position.

- Easier to set up than codebook
- Less secure than a (well-made) codebook
- Different cryptanalysis approaches
- Many examples of usage known
- Related to book codes

Classifications of codes

Purpose

Codes used for encryption

Codes used for compression

Codes used for avoiding errors

Alphabet

Code numbers

Code words

Code symbols

Number of parts

One-part codes

Two-part codes

Agenda

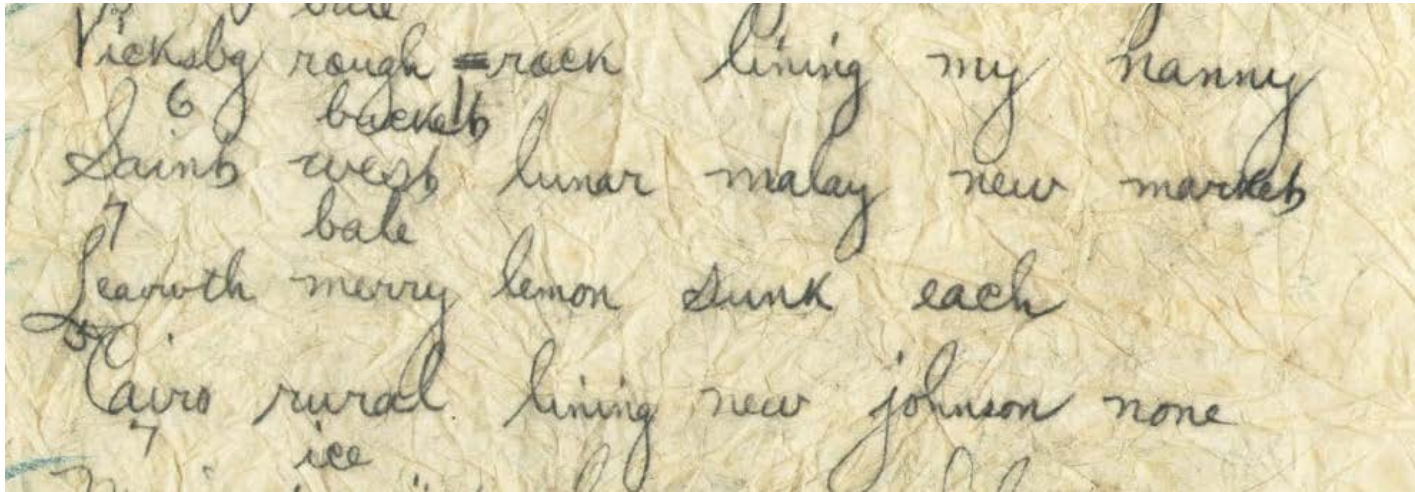
1. **Basics: scope, terminology, classification**
2. **How to detect a code**
3. **How to cryptanalyze a code**
4. **Open research questions**
5. **A few solved code cryptograms**
6. **A few unsolved code cryptograms**

How to detect a code?

I received your letter dated Sept. 2^o and should not have delayed so long sending an answer to it, had I any thing very material to communicate. 3693: 2517. 65: 3423. 576. 1100. 47. 1765. 3000. 259. 3032. 57. 66. 1795. 19. 211. 46. 1038. 1637. 970. 2609. 3369. 696. 3696. 427. 118. 3364. 1362. 456. 111. 566. 77. 1551. 2961. 1504. 1437. 3560. 1453. 2053. 1555. 1834. 1406. 9. 2044. 2694. 3423. 678. 1359. 493. 809. 1094. 956. 636. 1618. 61. 1437. 1369. 2316. 497. 314. 684. 1205. 193. 685. 2072. 65. 39. 3459. 3937. 2108. 2615. 1359. 766. 2450. 880. 1291.

An alphabet with hundreds of letters (e.g., numbers with several digits) is typical for a code.

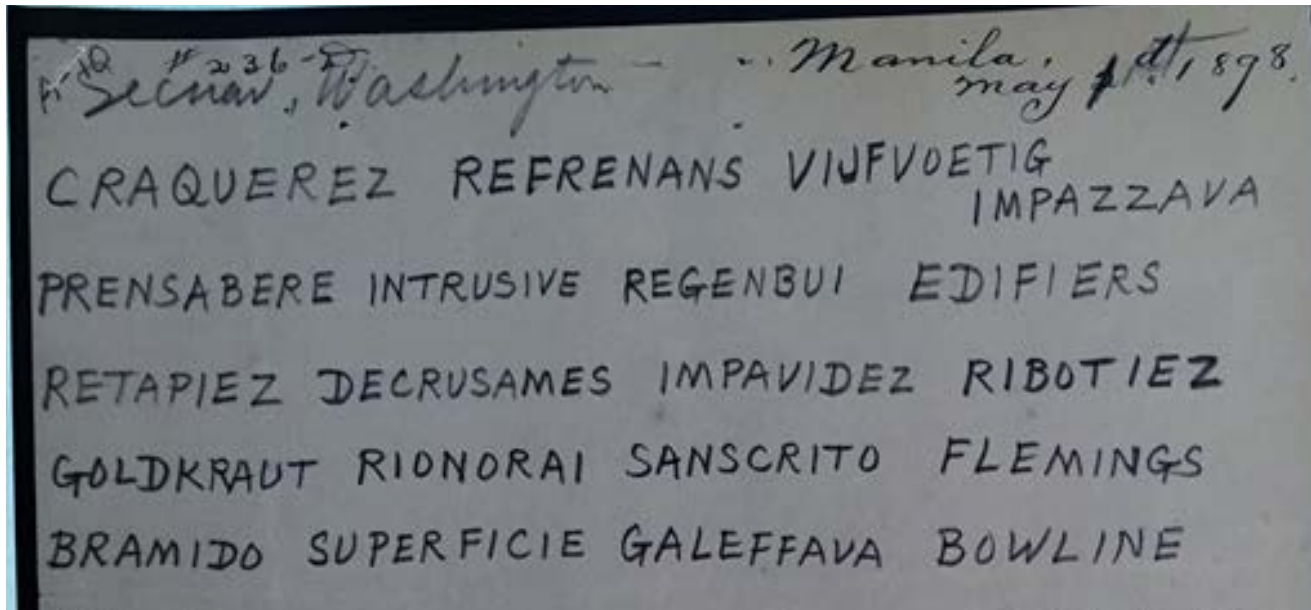
How to detect a code?



Handwritten text on crumpled paper, showing pseudo-meaningful words:

Vicksby rough = rock lining my hanny
6 baskets
Saints west lunar malay new market
7 bale
Leaveth merry lemon sunk each
Cairo rural lining new johnson none
ice

Pseudo-meaningful words are typical for a code.



Handwritten text on a document, showing pseudo-meaningful words:

Secnav, Washington - Manila, May 1st 1898.
CRAQUEREZ REFRENANS VIJFVOETIG
IMPAZZAVA
PRENSABERE INTRUSIVE REGENBUI EDIFIERS
RETAPIEZ DECRUSAMES IMPAVIDEZ RIBOTIEZ
GOLDKRAUT RIONORAI SANSKRITO FLEMINGS
BRAMIDO SUPERFICIE GALEFFAVA BOWLINE

How to detect a code?

Was nvlvaft by
aakat txpxsck upbk
txphn ohay ybtx cpt
mxhg wae sxfp zavfz
ack there first
txlk week wayx za
with thx

Cleartext words in an
encrypted message
are typical for a code.

How to detect a code?

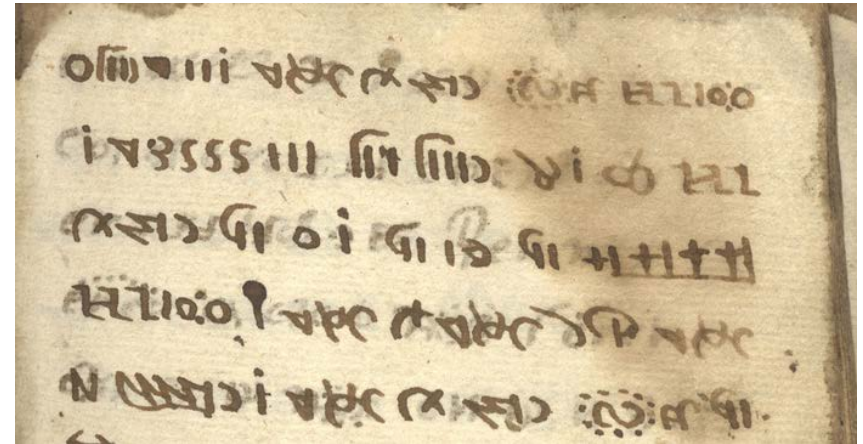
Sometimes it's easy

I received your letter dated Sept. 2^o and should not have delayed so long sending an answer to it, had I any thing very material to communicate. 3693: 2517. 65: 3423. 576. 1100. 97. 1765. 3000. 259. 3032. 57. 66. 1795. 19. 211. 46. 1038. 1637. 970. 2609. 3369. 696. 3696. 427. 118. 3364. 1362. 456. 111. 566. 77. 1551. 2961. 1504. 1437. 3560. 1453. 2053. 1555. 1834. 1406. 9. 2044. 2694. 3423. 678. 1359. 493. 809. 1094. 956. 636. 1618. 61. 1437. 1369. 2316. 497. 314. 684. 1205. 193. 685. 2072. 68. 39. 3459. 3937. 2108. 2615. 1359. 766. 2450. 880. 1291.

Was nvlvaft by aakat txpxsck upbk
txphn ohay ybtx cpt mxhg wae sxfp
zavfz ack there first txlk week
wayx za with thx

277. 409. 50. 24. 428. 82. 326. 13. 101. 308. 393. 409.
335. 107. 106. 251. 30. 3. 393. 291. 359. 262. 85. 160.
329. 136. 84. 43. 166. 50. 14. 184. 201. 226. 152. 129.
162. 350. 122. 13. 295. 409. 188. 168. 28. 43. 50. 78. 396.
326. 56. 138. 30. 22. 20. 50. 84. 66. 100. 344. 202. 369.
129. 91. 22. 8. 85. 286. 28. 324. 30. 56. 84. 50. 126.
188. 17. 14. 247. 35. 17. 200. 33. 84. 3. 8. 50. 275. 54.
8. 87. 69. 282. 260. 199. 393. 28. 109. 65. 50. 37. 39. 11.
152. 129. 204. 298. 320. 64. 44. 345. 161. 184. 279. 56.

Sometimes not



| | | | |
|------|------|------|------|
| EOCA | J2OL | BTOU | 2RFA |
| DYFM | NRLF | DUTX | DRPK |
| MYAT | DUNT | ZKYD | BYCY |
| PDUR | BOAX | RYEL | JLOT |

Agenda

1. **Basics: scope, terminology, classification**
2. **How to detect a code**
3. **How to cryptanalyze a code**
4. **Open research questions**
5. **A few solved code cryptograms**
6. **A few unsolved code cryptograms**

How to break a code?

Solution approaches:

- Find the codebook/nomenclator
- Frequency analysis, text statistics
- Word guessing

Weakly constructed codes may be helpful

Agenda

1. **Basics: scope, terminology, classification**
2. **How to detect a code**
3. **How to cryptanalyze a code**
4. **Open research questions**
5. **A few solved code cryptograms**
6. **A few unsolved code cryptograms**

Analysing tool applying text statistics

Number of
entries

Letters, syllables,
words, expressions

Homophones
Nulls
Nullifiers

Superencipherment

Code creation software

Software that creates a code based on parameters:

- language
- number of codegroups
- number of homophones
- number of nulls
- ...

Code database

**I would be helpful to have a
searchable database containing
all codes known.**

Agenda

1. **Basics: scope, terminology, classification**
2. **How to detect a code**
3. **How to cryptanalyze a code**
4. **Open research questions**
5. **A few solved code cryptograms**
6. **A few unsolved code cryptograms**

140

deyle

London

26

מברק TELEGRAM

1646 r4993 jw931 cde
New York 271/159

This form must accompany any enquiry respecting this telegram:
يجب ان يرسل هذا النموذج مع كل رسالة بخصوص هذه البرقية
את הטופס הזה צריך לצרף לכל השורה הנערכת בנידון המברק הזה.

Published by Karsten Hansky
John McVey, a design professor and codebook expert from Massachusetts recognized the codebook (Peterson's International Code (3rd edition, 1929)).

מקלט אדארית
הוראות משרדיות
קלטת פני
נמסרת
התאריך
ביום
وصلت في
نومبر 5



To ~~XXXXXXXXXXXX~~ vizl38 idz 1646 r4993 jw931 cde New York 271/159
4 1930
GOVT MEMISRAEL TELAVIV

الى
98

147 FOLLOWING TRANSMITTED FROM LONDON FOR GANEMANA QUOTE CABLE NUMBER 36 LE-RZY
YUBRB BIKUG NEYBH FIFOD KADAT OLUZA DYELN AWIAW LFYWI ANCRU KOKXE EGBFA IOGBU
ODVYE OJCEK KEWOF ONRIU AMIIM OIWXN ANJYZ ODVYE IOGVU ARLEJ DYRVY GIFYK FUPAR
VERTIALLY ITIYA FUPAR CYTUV OSYLO OTRVO ~~ERM~~ YUBRB BIKUG NEYBH FIFOD KADAT
OLUZA DYELN AWIAW LFYWI HCRU ~~KE~~ KOKXE KOKXE EGBFA IOGBU ODVYE OJCEK KEWOF
ONRIU AMIIM OIWXN ANJYZ ODVYE IOGVU ARLEJ DYRVY GIFYK FUPAR VERTICALLY ITIYA
FUPAR CYTUV OSYLO OTRVO ALHDE IOGBU FIFOD OYYSV ODHMU ~~NOIEX~~ HSOIG KFIYS
GIFYK NFIGU IUXBK OYHLI HDYIZ ITOXH IYVIR FLAAF GKATU IOXYJ ALFFM OUDNK
LRZIC AYBGV ODVYE FUPUK LMBOA AZWUV ALFFM EFZXU ITYHZ OWYRC AFGNE LZCOE
ANMNI PDYAV CYUWV ODVYE AWIAW JESSA IXDYN ANOEV ~~XXXX~~ ANOEV UPHY IOGBU

16th century nomenclator broken by Albert Leighton

Message sent from Poland to the Vatican in 1573 (excerpt):

608 53 17 11 75 17 55 25 77 75 29 97 41 77 13 79 11 77 15 59
19 79 15 79 17 39 19 79 15 59 13 79 99 58 99 11
17 59 13 67 79 15 77 17 99 15 15 83 54 97 41 57 15 77 75 15 59
26 99 15 37 15 38 34 17 37 57 19 79

...

Albert Leighton's assumptions:

- unmarked two-digit numbers: letter
- marked two-digit numbers, three-digit numbers: word
- numbers starting with 1 or 2 are vowels
- 77, the most frequent non-vowel stands for N

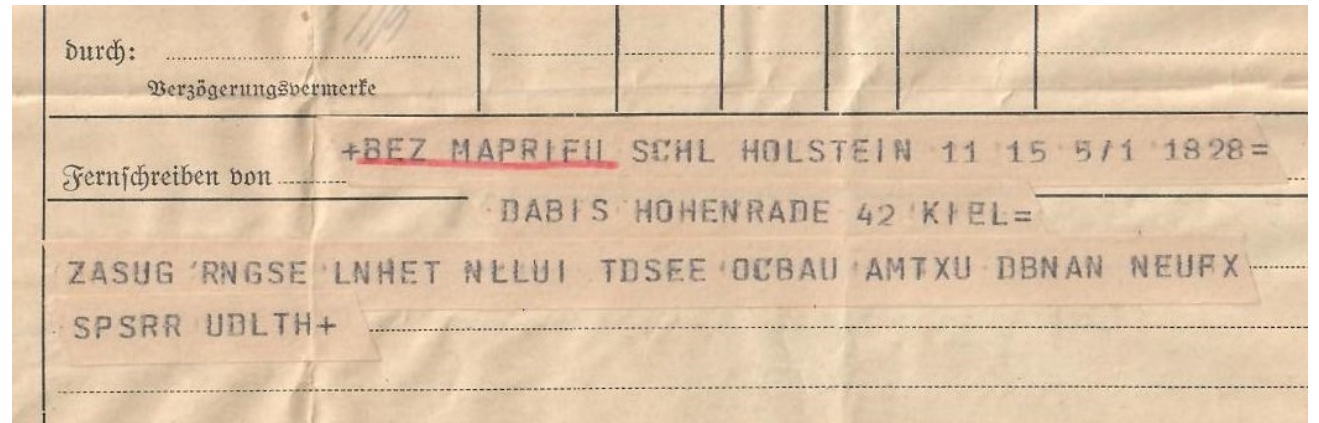
This information was sufficient to decipher most of the cryptogram.

Agenda

1. **Basics: scope, terminology, classification**
2. **How to detect a code**
3. **How to cryptanalyze a code**
4. **Open research questions**
5. **A few solved code cryptograms**
6. **A few unsolved code cryptograms**

A few unsolved code cryptograms

WW2 Telegram



A few unsolved code cryptograms

Manchester Cryptogram

Trenton N.J. 20. 1783

Sir,

I received your Letter dated Sept. 2^o and
should not have delayed so long sending an
answer to it, had I any thing very material
to communicate.

3693. 2517. 65. 3423. 576. 1100.
97. 1765. 3000. 259. 3032. 57. 66. 1795. 19. 211.
46. 1038. 1637. 970. 2609. 3369. 696. 3696. 427. 118.
3364. 1362. 456. 111. 566. 77. 1551. 2961. 1504. 1437.
3560. 1453. 2053. 1555. 1834. 1406. 9. 2044. 2694.
3423. 678. 1359. 493. 809. 1094. 956. 636. 1618. 61.
1437. 1369. 2316. 497. 314. 684. 1205. 193. 685. 2072.
65. 39. 3459. 3937. 2108. 2615. 1359. 766. 2450. 880. 1291.
647. 3339. 1175. 3714. 809. 184. 564. 2101. 1581. 566. 2323.
2066. 823. 665. 2401. 1692. 3560. 1444. 2794. 970. 330.
3601. 3263. 1612. 3000. 1291. 2000. 1936. 3056. 3278. 1618.
2894. 3498. 233. 2424. 3137. 3928. 1501. 3364. 434. 492.
566. 1998. 2450. 3560. 1603. 3905. 3082. 1504. 1242.
1624. 987. 2615. 1306. 350. 1245. 1504. 1145. 9. 3658.

J. John Heyman, Jr. & Co.

2622.

A few unsolved code cryptograms

Ohio Cryptogram

WAS NVKVAFT BY AAKAT TXPXSC UPBK TXPHN OHAY YBTX CPT
MXHG WAE SXFP ZAVFZ ACK THERE FIRST TXLK WEEK WAYX ZA
WITH THX.

A few unsolved code cryptograms

Van Gelder Cryptogram

1194 239 62 893 2395 1568 226 980 1382 929 800 75 2482 497
1195 211 560 511 475 524 680 228 1007 804 1696 35 331 4 39
356 2538 1049 564 934 416 748 601 577 179 2740 742 710 75
900 539 1341 805 420 561 2891 679 1632 1034 1171 8264
1294 332 508 707 1789 75 2655 632 189 473 1128 624 1194
2913 103 894 1193 850 227 93 657 1567 298 1300 948
255 1700 1576 816 735 609 547 189 2898 467 713 519
1567 2486 792 147 680 805 1843 21 2731 935 831 112 771 791
185 1343 1523 1413 2920 241 799 1097 699 613 795 15 1528
186 607 894 135 510 1388 1710 804 618 949 524 211 1413

END



A handwritten ledger with multiple columns and rows. The text is written in a cursive script. The columns appear to be organized into several sections, possibly representing different categories or accounts. The handwriting is dense and fills most of the page.



A printed ledger with multiple columns and rows. The text is printed in a clear, legible font. The columns are organized into several sections, similar to the handwritten ledger. The layout is clean and professional.

Klaus Schmeh (www.schmeh.org)