Introduction to Cryptographic Engineering

Steven M. Bellovin
https://www.cs.columbia.edu/~smb

Cryptographic Engineering?

- There are lots of introductions to encryption
- But—using encryption in the real world requires more
- We have to *engineer* it
- If we get the *engineering* wrong, enemies can crack our systems

A Disclaimer

- I'll be talking about *classical* (and simple) encryption, because it's easier to see what's going on

- I don't have time to cover all of the issues even there

- Modern encryption systems also need engineering; many of the issues today are quite similar

Terminology

- Encryption is an *algorithm*
- It converts *plaintext*—the message we want to protect—and a *key to* ciphertext
- Decryption, of course, converts the ciphertext and the key to plaintext
- Design principle: the system should be secure even if you enemy knows the algorithm—the security should rest entirely on protecting the key (Kerckhoff, 1883)

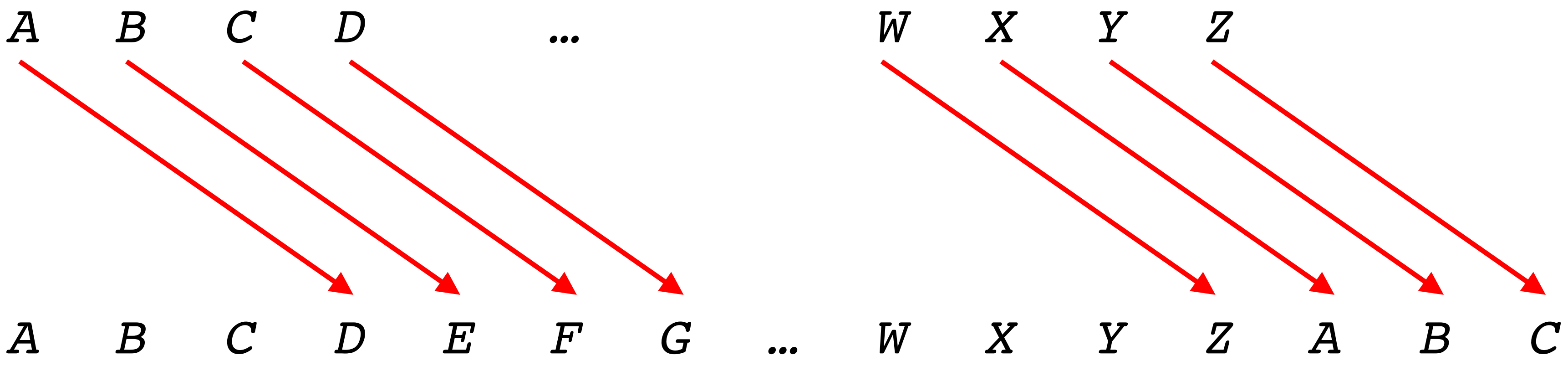
Codes and Ciphers

- Ciphers operate at the *syntactic* layer
- Replace a bit or a letter with a different bit or letter
- It doesn't matter what the language is

- Codes operate at the *semantic* layer
- Replace a word, phrase, or sentence with a *codeword*
- Language-dependent: you can't use an English language codebook to encode French

Caesar Cipher

- According to Suetonius (writing around 121 CE), Caesar used a cipher that shifted every letter by 3:



- We could say that the key is “*3*”—the amount of the shift—or we could say that it’s “*D*”—*A* becomes *D*

- (This cipher is very, very insecure, for lots of reasons, but it’s a simple example for now. Many of Caesar’s enemies were illiterate…)

Sample Encryption

- Winston Churchill:

  “This is the kind of tedious nonsense up
  with which I will not put”

  *WKLV LV WKH NLQG RI WHGLRXV QRQVHQVH XS
  ZLWK ZKLFK L ZLOO QRW SXW*

- What’s wrong?

Patterns Show Through

WKLV LV WKH NLQG RI WHGLRXV QRQVHQVH XS
ZLWK ZKLFK L ZLOO QRW SXW

- Word lengths: “L” can only be “A” or “I”
- Repeated letter patterns can show through

Patterns Show Through

WKLV LV WKH NLQG RI WHGLRXV QRQVHQVH XS
ZLWK ZKLFK L ZLOO QRW SXW

- Word lengths: “L” can only be “A” or “I”
- Repeated letter patterns can show through
- Solution: five-letter “groups”

WKLVL VWKHN LQGRI WHGLR XVQRQ VHQVH XSZLW
KZKLF KLZLO OQRWS XW

(How Many Words Have the Same Pattern as ‘*QRQVHQVH*’)?

- Look for letters 3-4 the same as 6-7
  - 132 such words, most rather uncommon, e.g., “obtected”
- Look for letters 1, 3, and 6 being the same
  - 45 such words, most rather uncommon, e.g., “anaplasm”
- Look for both patterns:
  - Only two, “cachucha” and “nonsense”
- Which do you think it is?

Multiple Keys

- Alice has to exchange secret messages with Bob, Carol, and David
- Bob and Carol are allowed read each other’s messages
- Bob and Carol must not see David’s messages; he must not see theirs

Multiple Keys

- Alice has to exchange secret messages with Bob, Carol, and David
- Bob and Carol are allowed read each other’s messages
- Bob and Carol must not see David’s messages; he must not see theirs
- Solution: Alice, Bob, and Carol share one key; Alice and David share a different one

Multiple Keys

- Alice has to exchange secret messages with Bob, Carol, and David

- Bob and Carol are allowed read each other's messages

- Bob and Carol must not see David's messages; he must not see theirs

- Solution: Alice, Bob, and Carol share one key; Alice and David share a different one

- Messages must show which key is being used

Indicators

- Messages must contain an *indicator*

```
KIBYZ WKLVL VWKHN LQGRI WHGLR XVQRQ VHQVH
XSZLW KZKLF KLZLO OQRWS XW
```

versus

```
ZSETK WKLVL VWKHN LQGRI WHGLR XVQRQ VHQVH
XSZLW KZKLF KLZLO OQRWS XW
```

- To the enemy, the indicator looks just like another code group

Message Lengths Matter

- Knowledge of message lengths matters
    - Why? Spot message importance, repeated messages, etc.
- We need to *pad* the real message with dummy stuff
- Also: recipient must be certain the entire message was received

Padding

```
This is the kind of tedious nonsense up with
which I will not put xxx blue red cat flower
rock
```

WKLVL VWKHN LQGRI WHGLR XVQRQ VHQVH XSZLW KZKLF KLZLO OQRWS XWAAA EOXHU HGFDW IORZH UURFN

The XXX is a Pattern

```
This is the kind of tedious nonsense up with
which I will not put the world wonders
```

`WKLVL VWKHN LQGRI WHGLR XVQRQ VHQVH XSZLW`
`KZKLF KLZLO OQRWS XWWKH ZRUOG ZRQGH UV`

- But now the recipient can be confused—and besides, we still have to worry about receiving the whole thing

Lengths

- The original message is 11 groups long, plus an indicator

```
KIBYZ 11 WKLVL VWKHN LQGRI WHGLR XVQRQ VHQVH
XSZLW KZKLF KLZLO OQRWS XWWKH ZRUOG ZRQGH UV
```

- But that’s no good—the attacker can see the message length, so the padding is useless

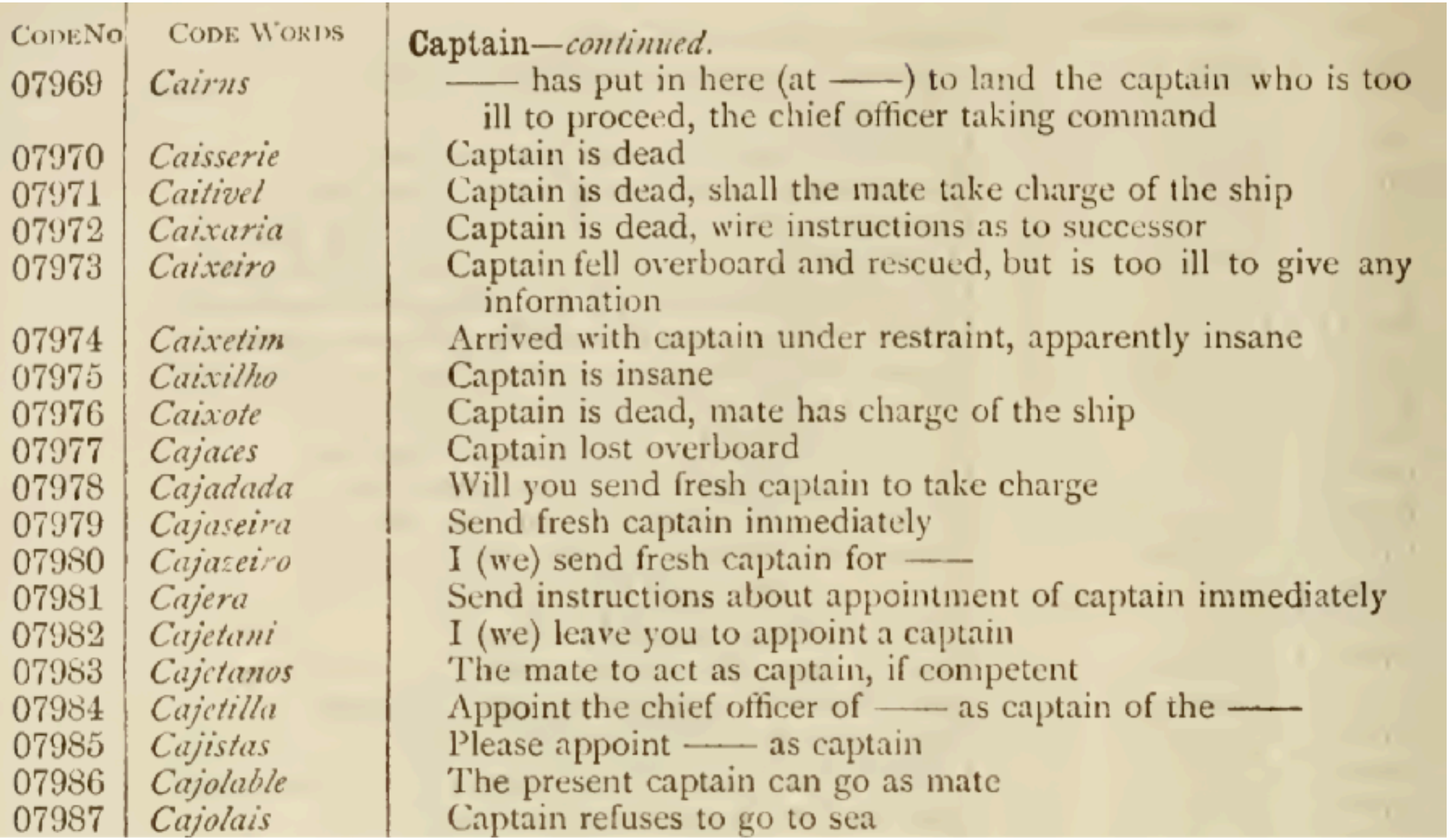
- Encrypt the length

```
KIBYZ ZNERL WKLVL VWKHN LQGRI WHGLR XVQRQ
VHQVH XSZLW KZKLF KLZLO OQRWS XWWKH ZRUOG
ZRQGH UV
```

ZNERL?

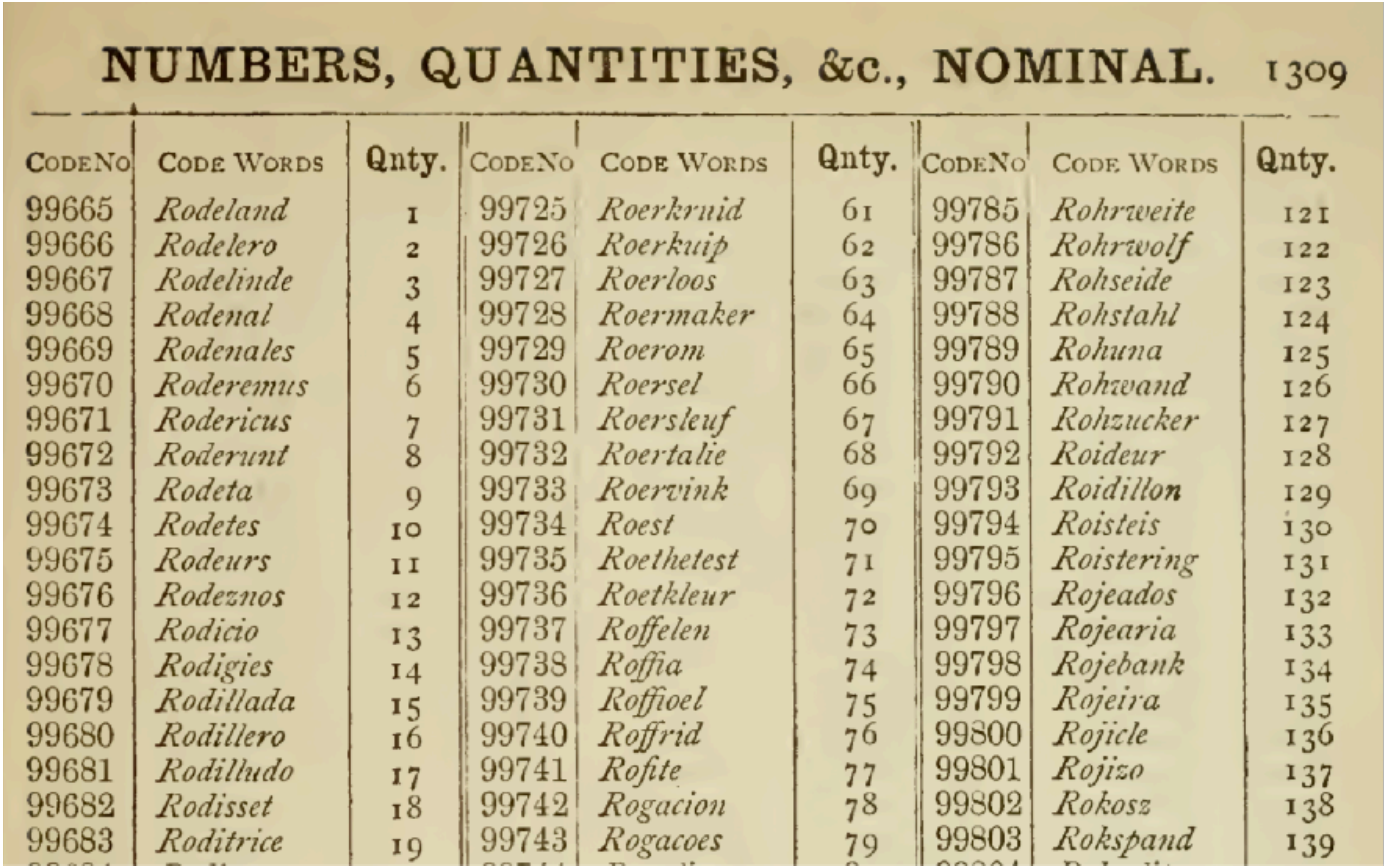
- Why does *`ZNERL`* mean 11?
- We're using a *code* for message lengths

A Commercial Codebook



CodeNo	Code Words	Captain—*continued.*
07969	*Cairns*	—— has put in here (at ——) to land the captain who is too ill to proceed, the chief officer taking command
07970	*Caisserie*	Captain is dead
07971	*Caitivel*	Captain is dead, shall the mate take charge of the ship
07972	*Caixaria*	Captain is dead, wire instructions as to successor
07973	*Caixeiro*	Captain fell overboard and rescued, but is too ill to give any information
07974	*Caixetim*	Arrived with captain under restraint, apparently insane
07975	*Caixilho*	Captain is insane
07976	*Caixote*	Captain is dead, mate has charge of the ship
07977	*Cajaces*	Captain lost overboard
07978	*Cajadada*	Will you send fresh captain to take charge
07979	*Cajaseira*	Send fresh captain immediately
07980	*Cajazeiro*	I (we) send fresh captain for ——
07981	*Cajera*	Send instructions about appointment of captain immediately
07982	*Cajetani*	I (we) leave you to appoint a captain
07983	*Cajetanos*	The mate to act as captain, if competent
07984	*Cajetilla*	Appoint the chief officer of —— as captain of the ——
07985	*Cajistas*	Please appoint —— as captain
07986	*Cajolable*	The present captain can go as mate
07987	*Cajolais*	Captain refuses to go to sea

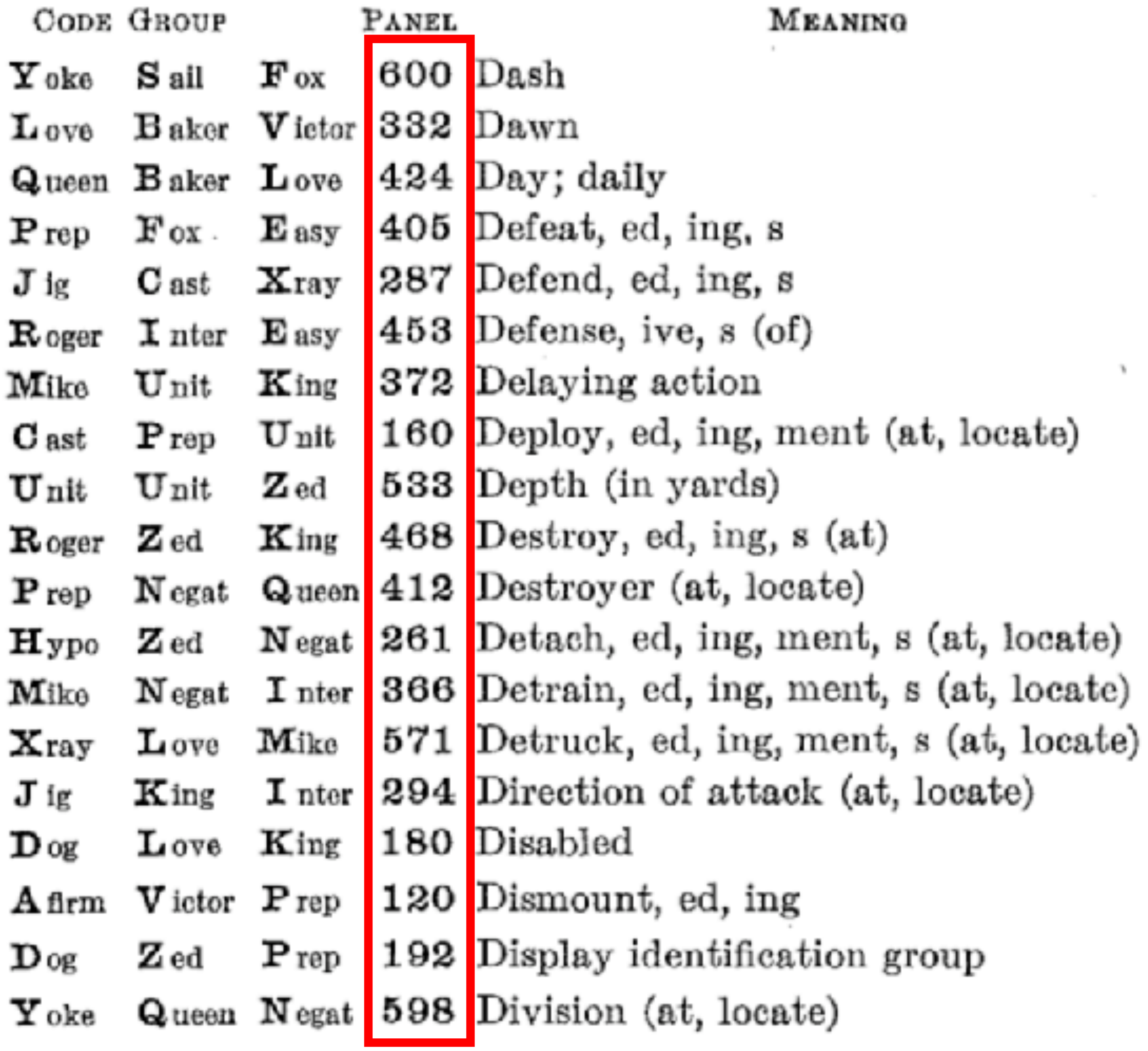
Encoding Numbers



NUMBERS, QUANTITIES, &c., NOMINAL. 1309

CODE NO	CODE WORDS	Qnty.	CODE NO	CODE WORDS	Qnty.	CODE NO	CODE WORDS	Qnty.
99665	*Rodeland*	1	99725	*Roerkruid*	61	99785	*Rohrweite*	121
99666	*Rodelero*	2	99726	*Roerkuip*	62	99786	*Rohrwolf*	122
99667	*Rodelinde*	3	99727	*Roerloos*	63	99787	*Rohseide*	123
99668	*Rodenal*	4	99728	*Roermaker*	64	99788	*Rohstahl*	124
99669	*Rodenales*	5	99729	*Roerom*	65	99789	*Rohuna*	125
99670	*Roderemus*	6	99730	*Roersel*	66	99790	*Rohwand*	126
99671	*Rodericus*	7	99731	*Roersleuf*	67	99791	*Rohzucker*	127
99672	*Roderunt*	8	99732	*Roertalie*	68	99792	*Roideur*	128
99673	*Rodeta*	9	99733	*Roervink*	69	99793	*Roidillon*	129
99674	*Rodetes*	10	99734	*Roest*	70	99794	*Roisteis*	130
99675	*Rodeurs*	11	99735	*Roethetest*	71	99795	*Roistering*	131
99676	*Rodeznos*	12	99736	*Roetkleur*	72	99796	*Rojeados*	132
99677	*Rodicio*	13	99737	*Roffelen*	73	99797	*Rojearia*	133
99678	*Rodigies*	14	99738	*Roffia*	74	99798	*Rojebank*	134
99679	*Rodillada*	15	99739	*Roffioel*	75	99799	*Rojeira*	135
99680	*Rodillero*	16	99740	*Roffrid*	76	99800	*Rojicle*	136
99681	*Rodilludo*	17	99741	*Rofite*	77	99801	*Rojizo*	137
99682	*Rodisset*	18	99742	*Rogacion*	78	99802	*Rokosz*	138
99683	*Roditrice*	19	99743	*Rogacoes*	79	99803	*Rokspand*	139

A World War II Military Codebook



Code Group			Panel	Meaning
Yoke	Sail	Fox	600	Dash
Love	Baker	Victor	332	Dawn
Queen	Baker	Love	424	Day; daily
Prep	Fox	Easy	405	Defeat, ed, ing, s
Jig	Cast	Xray	287	Defend, ed, ing, s
Roger	Inter	Easy	453	Defense, ive, s (of)
Mike	Unit	King	372	Delaying action
Cast	Prep	Unit	160	Deploy, ed, ing, ment (at, locate)
Unit	Unit	Zed	533	Depth (in yards)
Roger	Zed	King	468	Destroy, ed, ing, s (at)
Prep	Negat	Queen	412	Destroyer (at, locate)
Hypo	Zed	Negat	261	Detach, ed, ing, ment, s (at, locate)
Mike	Negat	Inter	366	Detrain, ed, ing, ment, s (at, locate)
Xray	Love	Mike	571	Detruck, ed, ing, ment, s (at, locate)
Jig	King	Inter	294	Direction of attack (at, locate)
Dog	Love	King	180	Disabled
Afirm	Victor	Prep	120	Dismount, ed, ing
Dog	Zed	Prep	192	Display identification group
Yoke	Queen	Negat	598	Division (at, locate)

Code Can Be Insecure

- The same codeword always means the same thing
- An enemy can recreate the codebook—which was routinely done by military cryptanalysts
- Solution: *superencipher* the code by using a book of *additives*

Codebook Additives

Page 137

00	50825	62424	63099	36442	52913
01	09688	88530	48525	98425	73807
02	47196	41570	82178	25272	12626
03	95697	22785	92911	04219	00369
04	26268	84115	02343	33874	21647
05	05516	28441	07963	14450	28494
06	77312	87426	50283	63730	70058
07	71124	62383	22000	54262	31432
08	72473	85872	88759	36150	58705
09	92346	74057	59815	71404	82269
10	96365	22045	09719	20053	81884
11	68321	16491	38622	65268	01214
12	95549	31926	64611	55481	48533
13	19566	98817	80809	33645	35048
14	53963	73491	02941	24300	36804

Additives

- Users had a book of additives—page upon page of random numbers
- Open the additive book to a random page; pick a random line
- Starting from there, use each number in turn and add it (without carrying!) to the *code number* from the codebook
- We now need an indicator for the additive: the page and line number

```
13705 05516 28441 16641 55329 17214, etc.
```

Additive Example

- You receive

```
13705 02480 25310
```

- The additives for that line are

```
      05516 28441
```

- Subtracting (but without borrowing!), we get

```
      07974 07979
```

- Turning to our codebook, we get….

Arrived with captain under restraint, apparently insane

Send fresh captain immediately

(This codebook, the *The A B C Universal Commercial Electric Telegraphic Code* from 1901, is available at https://books.google.com/books?id=CIDNAAAAMAAJ)

The Enigma Machine

- Used by the Germans during World War II
- Initially cracked by the Poles, who gave their insights to the British
- The British made major improvements and were able to read Enigma traffic constantly

(Photo: NSA)

Setting The Rotors

- The operator picked three random letters and encrypted them *twice*

- These encrypted letters were part of the indicators

Engineering and Usage Mistakes

- Encrypting one of the indicator fields twice was a fatal flaw
- Picking non-random letters for the indicator was a fatal flaw
- Sending the same, simple message every day was a fatal flaw
- Sending a message consisting of nothing but the letter “L” was a fatal flaw—this is partly usage, and partly a design weakness in the Enigma

The basic algorithm was decent—but it wasn’t <u>engineered</u> properly!

Questions?

(these slides at https://www.cs.columbia.edu/~smb/talks/intro-crypto-engineering.pdf)

Vigenère Cipher

- Write each letter of the key above the message, repeating as necessary
- Encrypt each *plaintext* letter with the key letter above it
- Note: because the key changes constantly, a single plaintext value can have a different ciphertext
- (Invented circa 1585; general solution found in 1863 by Kasiski)

Key:	SECRE	TSECR	ETSEC	RETSE	CRETS	ECRET
Message:	Thisi	sthek	indof	tedio	usnon	sense
Encrypted:	LLKJM	LLLGB	MGVSH	KIWAS	WJRHF	WGEWX

	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
A	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
B	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	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	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
F	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
G	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	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O