2

Wörter und Würmer oder Warum einfach, wenn's auch kompliziert geht?

Ein Wort, ein Satz –: aus Chiffren steigen erkanntes Leben, jäher Sinn (Gottfried Benn).

In diesem Kapitel stehen *polyalphabetischen* Chiffrierungen im Mittelpunkt. Das Ziel bei der Entwicklung solcher Verfahren ist, die Häufigkeiten der Geheimtextbuchstaben möglichst stark anzugleichen, um einem Angreifer die Arbeit möglichst schwer zu machen. Daher wird ein Klartextbuchstabe nicht stets zu demselben Geheimtextbuchstaben verschlüsselt. Eine polyalphabetische Chiffrierung kann also nicht einfach durch ein Klartextalphabet und ein darunter geschriebenes Geheimtextalphabet beschrieben werden.

Die Zuordnung eines Klartextbuchstabens zu einem Geheimtextbuchstaben darf aber auch nicht willkürlich erfolgen. Die *De*chiffrierung muss der strengen Regel der *Eindeutigkeit* genügen; sonst ist keine Dechiffrierung möglich. Anders gesagt: Wenn die Chiffrierung nicht eindeutig wäre, so befände sich der Empfänger prinzipiell in keiner besseren Lage als Mr. X!

Es gibt grundsätzlich zwei Arten, die Häufigkeiten der Geheimtextzeichen einander anzugleichen. Einerseits kann man jedem Klartextbuchstaben nicht nur ein Geheimtextzeichen, sondern eine ganze Menge davon zuordnen. Man spricht von einer *homophonen* Chiffre. Solche Algorithmen werden im folgenden Abschnitt dargestellt. Andererseits kann man aber auch die Geheimtextalphabete wechseln, also viele Geheimtexte im Wechsel verwenden. Die Untersuchung dieser im eigentlichen Sinne polyalphabetischen Verfahren ($\pi o \lambda \upsilon$ (griech.) = viel) wird den größten Teil des Kapitels ausmachen. Wir werden uns im Detail mit der Verschlüsselung nach Vigenère beschäftigen.

A. Beutelspacher, *Kryptologie*, DOI 10.1007/978-3-658-05976-7_2,

Tab. 2.1 Eine homophone Chiffre

Buchstabe	Zugeordnete Zeichen	**Buchstabe**	Zugeordnete Zeichen
a	10 21 52 59 71	**n**	30 35 43 62 63 67 68 72 77 79
b	20 34	**o**	02 05 82
c	28 06 80	**p**	31
d	04 19 70 81 87	**q**	25
e	09 18 33 38 40 42 53 54 55 60 66 75 85 86 92 93 99	**r**	17 36 51 69 74 78 83
f	00 41	**s**	15 26 45 56 61 73 96
g	08 12 97	**t**	13 32 90 91 95 98
h	07 24 47 89	**u**	29 01 58
i	14 39 46 50 65 76 88 94	**v**	37
j	57	**w**	22
k	23	**x**	44
l	16 03 84	**y**	48
m	27 11 49	**z**	64

2.1 Verschleierung der Häufigkeiten

Wie kann man erreichen, dass alle Geheimtextzeichen mit der gleichen Wahrscheinlichkeit auftreten? Ganz einfach: Bei einer *homophonen* Chiffre ordnet die Chiffriervorschrift jedem Buchstaben nicht nur ein Zeichen, sondern eine feste *Menge* von Zeichen (in unserem Beispiel: Ziffernpaare) zu, und zwar so, dass die folgenden Bedingungen erfüllt sind:

- Um das Dechiffrieren eindeutig zu machen, dürfen die Mengen, die zu verschiedenen Klartextbuchstaben gehören, kein gemeinsames Element besitzen. Man spricht von *disjunkten* Mengen.
- Die Anzahl der Geheimtextzeichen, die zu einem Klartextbuchstaben gehören, entspricht der Häufigkeit dieses Buchstabens. Wir verwenden hierzu die Häufigkeiten aus Tab. 1.2.

Im folgenden Beispiel einer homophonen Chiffre sind die Geheimtextzeichen die 100 Paare 00, . . . , 99 von Ziffern, die den Buchstaben gemäß Tab. 2.1 zugeordnet sind.

Beim *Chiffrieren* ordnet man einem Klartextbuchstaben zufällig ein dazugehöriges Geheimtextzeichen zu. Der Empfänger kann dann mit obiger Tabelle einfach *dechiffrieren*: 23520127 6429 97845929346663 04597396, 9945 5682 86886200712847 141513!

Algorithmus 2.1: Homophone Chiffrierung

Chiffrieren:	Um einen Buchstaben zu verschlüsseln, wählt man zufällig eines der Zeichen, das diesem Buchstaben zugeordnet ist.
Dechiffrieren:	Um ein Zeichen zu dechiffrieren, sucht man den Buchstaben, der zu dem Zeichen gehört.

Da die Zeichen beim Chiffrieren zufällig gewählt werden, kommt jedes Zeichen (in unserem Fall also jedes Ziffernpaar) gleich häufig vor (daher der Name „homophon" = gleich lautend). Ein potentieller Kryptoanalytiker sieht sich also vor eine wesentlich schwierigere Aufgabe gestellt als beim Brechen einer monoalphabetischen Chiffrierung.

Allerdings sollte der Systementwickler nicht zu früh jubeln, denn eine *Kryptoanalyse* ist auch hier möglich. Die Analyse basiert auf der Beobachtung, dass zwar die Häufigkeiten der Geheimtextzeichen, also der Ziffernpaare gleich sind, dass man aber aus der Betrachtung von *Paaren* von Geheimtextzeichen sehr wohl Information gewinnen kann. Vergleichen Sie dazu Tab. 1.4. Wir diskutieren zwei Beispiele.

Wenn Mr. X ein Geheimtextäquivalent des Buchstabens **c** betrachtet, also etwa die Zahl 28, so wird er feststellen, dass nur ganz bestimmte Geheimtextzeichen als unmittelbare Nachfolger von 28 in Frage kommen. Dies sind die Zahlen 07, 24, 23, 47, 89, also die Geheimtextäquivalente der Buchstaben **h** und **k.** Damit „weiß" er bereits, welche Zeichen den Buchstaben **h** oder **k** entsprechen.

Wenn Mr. X ein Geheimtextäquivalent des Buchstabens **e**, also etwa 99, ins Auge fasst, so stellt er fest, dass gewisse Geheimtextzeichen als Vorgänger und als Nachfolger von 99 vorkommen – und zwar praktisch gleich häufig. Dies müssen dann die Geheimtextäquivalente des Buchstabens **i** sein.

Diese Andeutungen sind natürlich noch längst keine Kryptoanalyse. Sie sollen Ihnen nur zeigen, dass Mr. X auch einem auf den ersten Blick scheinbar unknackbaren Geheimtext nicht völlig hilflos gegenübersteht.

2.2 Die Vigenère-Chiffre

Die Vigenère-Verschlüsselung (sprich: Wischen*ä*hr) wurde im Jahre 1586 von dem französischen Diplomaten Blaise de Vigenère (1523–1596) der Öffentlichkeit zugänglich gemacht. Die Grundidee ist, verschiedene monoalphabetische Chiffrierungen im Wechsel zu benützen. Diese Idee ist so natürlich, dass

Variationen der Vigenère-Chiffre mehrfach (wieder-)erfunden wurden. Zwei der wichtigeren Vorgänger waren Johannes Trithemius (1462–1516), dessen Bücher *Poligraphia* (1518) und *Steganographia* (1531) posthum veröffentlicht wurden, und Giovanni Battista Della Porta (1538–1615), der Erfinder der *Camera obscura*, der 1558 in seinem Buch *Magia naturalis* einen polyalphabetischen Algorithmus veröffentlichte, der große Ähnlichkeit mit der Vigenère-Chiffre aufweist.

In diesem Kapitel werden wir uns hauptsächlich mit der Vigenère-Verschlüsselung, der bekanntesten unter allen „periodischen" polyalphabetischen Algorithmen, beschäftigen, und zwar aus zwei Gründen.

- Die Vigenère-Verschlüsselung ist der Prototyp für viele Algorithmen, die bis heute praktisch eingesetzt werden.
- Bei der Kryptoanalyse werden wir zwei außerordentlich wichtige Methoden kennen lernen, den Kasiski-Test und den Friedman-Test.

Um nach dem *Algorithmus von Vigenère* chiffrieren zu können, braucht man zwei Dinge: Das *Vigenère-Quadrat* und ein *Schlüsselwort* (siehe Abb. 2.1).

Das *Vigenère-Quadrat* besteht aus 26 Alphabeten, die auf folgende Weise untereinander geschrieben sind. Das erste Alphabet ist das gewöhnliche Alphabet, das zweite das um einen Buchstaben nach links verschobene, das dritte das um zwei Buchstaben verschobene und so weiter. Mit anderen Worten: Das Vigenère-Quadrat besteht aus den 26 Verschiebechiffren in natürlicher Reihenfolge. Das erkennen wir auch daran, dass in der ersten Spalte das Alphabet steht. Das heißt: Zu jedem Buchstaben gibt es eine Zeile, die mit diesem Buchstaben beginnt.

Das Schlüsselwort kann jede beliebige Buchstabenfolge sein; für unser Demonstrationsbeispiel wählen wir das Wort VENUS. Wir schreiben dieses Schlüsselwort Buchstabe für Buchstabe über den Klartext (ohne Zwischenräume), und zwar so lange, bis die Länge des Klartexts erreicht ist:

Schlüsselwort:	V	E	N	U	S	V	E	N	U	S	V	E	N	U	S	V
Klartext:	**p**	**o**	**l**	**y**	**a**	**l**	**p**	**h**	**a**	**b**	**e**	**t**	**i**	**s**	**c**	**h**

Bei der *Chiffrierung* wird ein Klartextbuchstabe mit Hilfe des über ihm stehenden Schlüsselwortbuchstabens verschlüsselt. Genauer gesagt bestimmt der Schlüsselwortbuchstabe das Alphabet, d. h. die *Zeile* im Vigenère-Quadrat, mit dem dieser Klartextbuchstabe chiffriert wird.

Noch genauer: Um den ersten Geheimtextbuchstaben zu erhalten, müssen wir in dem Alphabet, das mit V beginnt, nachsehen, was in der *Spalte* **p** steht; dies ist der Buchstabe **K**.

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

Abb. 2.1 Das Vigenère-Quadrat

Noch ein Beispiel gefällig? Voilà: Um den zweiten Buchstaben zu verschlüsseln, suchen wir in der Zeile E den Buchstaben in der Spalte **o;** dieser ist **S**. Und so weiter. Insgesamt ergibt sich:

Schlüsselwort:	V	E	N	U	S	V	E	N	U	S	V	E	N	U	S	V
Klartext:	**p**	**o**	**l**	**y**	**a**	**l**	**p**	**h**	**a**	**b**	**e**	**t**	**i**	**s**	**c**	**h**
Geheimtext:	**K**	**S**	**Y**	**S**	**S**	**G**	**T**	**U**	**U**	**T**	**Z**	**X**	**V**	**M**	**U**	**C**

Es ist klar, dass die Vigenère-Chiffe Mr. X vor erheblich größere Probleme stellt als eine monoalphabetische Chiffrierung. Die Häufigkeit der Buchstaben ist viel gleichmäßiger verteilt. Dies erkennt man schon an unserem kurzen Beispiel. Die beiden Klartextbuchstaben **a** werden in verschiedene Geheimtextbuchstaben (**S** und **U**) verschlüsselt, während der Geheimtextbuchstabe **S** von verschiedenen Klartextbuchstaben, nämlich **o, y, a** herkommt.

Algorithmus 2.2: Vigenère-Chiffrierung

Chiffrieren:	Um einen Buchstaben zu verschlüsseln, bestimmt man den über ihm stehenden Schlüsselwortbuchstaben. Der Geheimtextbuchstabe ist der Schnittpunkt derjenigen Zeile, die mit dem Schlüsselwortbuchstaben beginnt, mit derjenigen Spalte, deren erstes Element der Klartextbuchstabe ist.
Dechiffrieren:	Um einen Geheimtextbuchstaben zu entschlüsseln, bestimmt man zunächst den über ihm stehenden Schlüsselwortbuchstaben. Dann sucht man das Geheimtextalphabet, das mit diesem Schlüsselwortbuchstaben beginnt. Mit diesem Alphabet entschlüsselt man; d. h. man sucht in diesem Alphabet den Geheimtextbuchstaben und geht zu dem darüber liegenden Klartextbuchstaben.

2.3 Kryptoanalyse

Mit heutigen Methoden kann auch ein Vigenère-chiffrierter Text geknackt werden. Denn ein genügend langer Geheimtext weist viele statistisch erfassbare Regelmäßigkeiten auf, die es einem ermöglichen, das Schlüsselwort zu erschließen. Der erste veröffentlichte Angriff stammt von dem preußischen Infanteriemajor Friedrich Wilhelm Kasiski (1805–1881), der diesen 1863 publiziert hat. Eine zweite Methode geht auf Colonel William Frederick Friedman (1891–1969) zurück. Beide Methoden dienen dazu, die Schlüsselwortlänge zu bestimmen. Da beide Tests auch über die spezielle Vigenère-Analyse hinaus weitreichende und grundlegende Bedeutung haben, sollen beide Methoden hier im Detail vorgestellt werden.

Angenommen, Mr. X hat den folgenden Text (Abb. 2.2) abgefangen, von dem er weiß oder vermutet, dass er Vigenère-chiffriert ist:

2.3.1 Der Kasiski-Test

Obwohl diese wirkungsvolle Methode zur Analyse polyalphabetischer Algorithmen zuerst von Kasiski *veröffentlicht* wurde, muss man erwähnen, dass der englische Mathematiker Charles Babbage (1792–1871), der unter anderem berühmt ist für seine Konzeption eines Vorgängers des modernen Computers, umfangreiche, allerdings unveröffentlichte Untersuchungen über Kryptographie durchgeführt hat. Insbesondere hatte er den Kasiski-Test bereits 1854 entwickelt, also neun Jahre vor Kasiski. Für eine detaillierte Darstellung siehe [Fra84].

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

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

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

Abb. 2.2 Ein Geheimtext, der mit der Methode von Vigenère verschlüsselt wurde

Wir illustrieren das Verfahren an einem mit der Methode von Vigenère verschlüsselten Geheimtext (siehe Abb. 2.2). Der Test beruht auf folgender Idee: Wenn im Klartext zwei Folgen aus gleichen Buchstaben auftreten (zum Beispiel zweimal das Wort **ein**), so werden im Allgemeinen die entsprechenden Folgen im Geheimtext verschieden ausfallen; denn schon der jeweils erste Buchstabe der beiden Folgen wird in der Regel verschieden verschlüsselt. Wenn aber die beiden Anfangsbuchstaben der Folgen mit Hilfe desselben Schlüsselwortbuchstabens verschlüsselt werden, so sind die beiden Geheimtextbuchstaben gleich. In diesem Fall werden auch die jeweils zweiten Buchstaben der Klartextfolgen mit demselben Schlüsselwortbuchstaben verschlüsselt; also ergeben sich auch im Geheimtext die gleichen Buchstaben. Das heißt also: Wenn die beiden Anfangsbuchstaben der Klartextfolgen mit demselben Schlüsselwortbuchstaben verschlüsselt werden, so bestehen die entsprechenden Geheimtextfolgen aus den gleichen Buchstaben. Siehe Abb. 2.3.

Wann tritt der Fall auf, dass zwei Buchstaben mit demselben Schlüsselwortbuchstaben verschlüsselt werden? Nun, genau dann, wenn das Schlüsselwort zwischen sie genau einmal, genau zweimal, genau dreimal, … „passt". Mit anderen Worten: Genau dann, wenn der *Abstand* der beiden Klartextbuchstaben ein Vielfaches der Schlüsselwortlänge ist (siehe Abb. 2.3).

Um den Abstand zwischen zwei Folgen zu bestimmen, geht man so vor: Man zählt die Anzahl der Buchstaben zwischen den jeweils ersten Buchstaben der Folgen, wobei man den ersten Buchstaben nicht mitzählt. Zum Beispiel haben die Folgen, die mit dem Buchstaben Nr. 11 beziehunsgsweise mit dem Buchstaben Nr. 26 beginnen, den Abstand 15.

Schlüsselwort: V E N U S V E N U S V E N U S V
Klartext: ... **e i n** ... ... **e i n** ...
Geheimtext: ... **R C F** ... ... **W D R** ...
Im Allgemeinen werden Klartextfolgen aus gleichen Buchstaben in Geheimtextfolgen aus verschiedenen Buchstaben chiffriert.

Schlüsselwort: V E N U S V E N U S V E N U S V
Klartext: ... **e i n** ... ... **e i n** ...
Geheimtext: ... **Y A I** ... ... **Y A I** ...

Wenn aber die Anfangsbuchstaben der beiden Folgen unter dem gleichen Schlüsselwortbuchstaben stehen, dann bestehen auch die entsprechenden Geheimtextfolgen aus gleichen Buchstaben.

Abb. 2.3 Der Kasiski-Test

Tab. 2.2 Auswertung der Folgen aus gleichen Buchstaben

Folge	Abstand	Primfaktorzerlegung des Abstands
TNE	50	$2 \cdot 5 \cdot 5$
FCRV	265	$5 \cdot 53$
NEVHJAOVWU	90	$2 \cdot 3 \cdot 3 \cdot 5$
VWU	75	$3 \cdot 5 \cdot 5$

Wir fassen zusammen: Wenn zwei Klartextfolgen aus gleichen Buchstaben einen Abstand haben, der ein Vielfaches der Schlüsselwortlänge ist, so entsprechen ihnen im Geheimtext Folgen aus gleichen Buchstaben.

Nun dreht Mr. X den Spieß um: Er sucht zunächst im Geheimtext Folgen aus gleichen Buchstaben (siehe Abb. 2.4). Er vermutet, dass ihr Abstand „wahrscheinlich“ ein Vielfaches der Schlüsselwortlänge ist. Diese Wahrscheinlichkeit folgt dem Gesetz „je länger, je lieber“: Gleiche Buchstaben sagen, wie wir wissen, gar nichts über die Schlüsselwortlänge aus, und auch Paare aus gleichen Buchstaben können sich „zufällig“ ergeben. Aber aus Folgen von drei oder mehr gleichen Buchstaben kann Mr. X schon ziemlich zuverlässig auf die Schlüsselwortlänge schließen. In unserem Beispiel erkennt er die Struktur, die in Abb. 2.4 zu erkennen ist.

Der größte gemeinsame Faktor ist 5. Also könnte ein (zu) optimistischer Kryptoanalytiker frohlocken und sagen, „Ich weiß, dass die Schlüsselwortlänge 5 ist“. In der Tat funktioniert der Kasiski-Test in der Praxis sehr gut.

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

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

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

Abb. 2.4 Folgen aus gleichen Buchstaben

Algorithmus 2.3: Kasiski-Test (Bestimmen der Schlüsselwortlänge)

Man sucht gleiche Folgen im Geheimtext und bestimmt deren Abstand. Dieser ist (vermutlich) ein Vielfaches der Schlüsselwortlänge.

Wenn der Kryptoanalytiker Mr. X allerdings vorsichtig ist, wird er nur sagen, „Dies ist ein starkes Indiz dafür, dass die Schlüsselwortlänge 5 ist." Warum tut Mr. X gut daran, vorsichtig zu sein? Es gibt dafür zwei Gründe.

1. Es könnte sein, dass zufällig (!) zwei Geheimtextfolgen aus drei oder mehr gleichen Buchstaben vorhanden sind, die einen nicht durch 5 teilbaren Abstand haben. Dann würde sich als größter gemeinsamer Teiler 1 ergeben! (In unserem Beispiel tritt dieser Fall tatsächlich auf: Die Folge **O I S** kommt zweimal vor, und zwar mit Abstand $26 = 2 \cdot 13$.) Das heißt, dass man den größten gemeinsamen Teiler nicht „blind" ausrechnen darf, sondern dass man ihn „mit Gefühl" bestimmen muss. Offensichtliche Ausreißer muss man unberücksichtigt lassen.
2. Gerade deswegen könnte man auf die Idee kommen, die Schlüsselwortlänge könnte nicht 5, sondern 10, 15 oder 30 sein, denn die Faktoren 2 und 3 kommen auch recht häufig vor. Mit anderen Worten: Der Kasiski-Test liefert einem die Schlüsselwortlänge bis auf *Vielfache (*oder *Teiler).*

Auch aus diesem Grund präsentieren wir noch eine zweite Methode; diese ergibt die *Größenordnung* der Schlüsselwortlänge. Eine Kombination beider Methoden lässt einen dann kaum mehr in die Irre gehen.

2.3.2 Der Friedman-Test

Dieses Verfahren wurde 1925 von William Friedman entwickelt, „der als größter Kryptologe aller Zeiten gilt“ [Fra82]. Bei diesem Test fragt man sich, *mit welcher Chance ein willkürlich aus einem Klartext herausgegriffenes Buchstabenpaar aus gleichen Buchstaben besteht*. Die Antwort darauf wird durch den Koinzidenzindex gegeben.

Stellen wir uns dazu zunächst eine beliebige Buchstabenfolge der Länge n vor. Sei n_1 die Anzahl der **a**'s, n_2 die Anzahl der **b**'s, … und n_{26} die Anzahl der **z**'s.

Wir bestimmen die Anzahl der Paare, bei dem beide Buchstaben gleich **a** sind. Wir verlangen *nicht*, dass es sich um Bigramme, also um aufeinanderfolgenden Buchstaben handelt. Für die Auswahl des ersten **a**'s gibt es nach Definition genau n_1 Möglichkeiten, für die Auswahl des zweiten **a**'s dann noch $n_1 - 1$ Möglichkeiten. Da es auf die Reihenfolge der Buchstaben nicht ankommt, ist die Anzahl der gesuchten Paare gleich $n_1(n_1 - 1)/2$.

Also ist die Anzahl der Paare, bei dem beide Buchstaben gleich sind, d. h. bei denen beide gleich **a** oder beide gleich **b** … oder beide gleich **z** sind, gleich

$$\frac{n_1(n_1-1)}{2} + \frac{n_2(n_2-1)}{2} + \ldots + \frac{n_{26}(n_{26}-1)}{2} = \sum_{i=1}^{26} \frac{n_i(n_i-1)}{2}.$$

Die *Chance*, ein Paar aus gleichen Buchstaben zu erwischen, lässt sich daraus nach der Melodie „Anzahl der günstigen Fälle durch Anzahl der möglichen Fälle“ wie folgt berechnen:

$$\frac{\sum_{i=1}^{26} \frac{n_i(n_i-1)}{2}}{n(n-1)/2} = \frac{\sum_{i=1}^{26} n_i(n_i-1)}{n(n-1)}.$$

Diese Zahl heißt der (Friedmansche) *Koinzidenzindex* und wird mit I bezeichnet:

$$I = \frac{\sum_{i=1}^{26} n_i(n_i-1)}{n(n-1)}.$$

Friedman selbst bezeichnete diese Zahl mit κ (griechisches Kappa); daher findet man für die Methode, die wir im Folgenden vorstellen, manchmal auch den Namen *Kappa-Test*.

Nun nähern wir uns diesem Koinzidenzindex von einer anderen Seite. In vielen Fällen kennt Mr. X nämlich nicht nur den Geheimtext, sondern weiß von vornherein auch etwas über die Verteilung der Buchstaben.

Nehmen wir an, er *wüsste*, dass im Text der Buchstabe **a** mit der Wahrscheinlichkeit p_1 auftritt, der Buchstabe **b** mit der Wahrscheinlichkeit p_2, ..., und schließlich der Buchstabe **z** mit der Wahrscheinlichkeit p_{26}. Mr. X weiß dies, wenn es sich um einen deutschen Text handelt. Dann kann er die konkreten Werte für die Wahrscheinlichkeiten p_i beispielsweise der Tab. 1.2 entnehmen.

Stellen wir uns nun zwei willkürlich herausgegriffene Buchstaben unseres Textes vor. Die Wahrscheinlichkeit, dass der erste dieser Buchstaben gleich **a** ist, ist p_1. Auch die Wahrscheinlichkeit dafür, dass der zweite Buchstabe gleich **a** ist, ist p_1. Da beide Buchstaben willkürlich gewählt wurden, ist die Wahrscheinlichkeit dafür, dass an beiden Stellen der Buchstabe **a** steht, gleich p_1^2.

Entsprechendes gilt auch für die anderen Buchstaben. Somit ist die Wahrscheinlichkeit dafür, dass an zwei beliebig herausgegriffenen Stellen der *gleiche* Buchstabe steht (also entweder der Buchstabe **a**, der Buchstabe **b** oder ...), gleich

$$p_1 \cdot p_1 + p_2 \cdot p_2 + \ldots + p_{26} \cdot p_{26} = \sum_{i=1}^{26} p_i^2.$$

Wir fassen zusammen: Wenn wir die Wahrscheinlichkeiten der Buchstaben kennen, können wir von vornherein den Koinzidenzindex ausrechnen. Es gilt:

$$I = \sum_{i=1}^{26} p_i^2.$$

Diese Zahl hängt natürlich von den Wahrscheinlichkeiten $p_1, \ldots, p_{26}$ ab. Wir betrachten zwei Beispiele.

- Man kann den *Koinzidenzindex der deutsche Sprache* bestimmen, indem man einfach die Wahrscheinlichkeiten der Buchstaben aus Tab. 1.2 in die Formel einsetzt:

$$\sum_{i=1}^{26} p_i^2 = 0{,}0651^2 + 0{,}0189^2 + 0{,}0306^2 + \ldots + 0{,}0113^2 = 0{,}0762.$$

Dies bedeutet, dass ein zufällig gewähltes Buchstabenpaar mit einer 7,62 %-igen Chance aus gleichen Buchstaben besteht. Mit anderen Worten: Etwa jedes 13-te Buchstabenpaar besteht aus gleichen Buchstaben.

- Stellen wir uns nun andererseits einen völlig zufälligen Text vor, also einen Text, in dem die Buchstaben bunt durcheinandergewürfelt sind. Darin sind die Buchstaben gleichverteilt, also kommt jeder Buchstabe mit derselben Wahrscheinlichkeit $p_i = 1/26 \approx 0{,}0385$ vor. In diesem Fall ergibt sich als *Koinzidenzindex einer zufälligen Buchstabenfolge*

$$\sum_{i=1}^{26} p_i^2 = \sum_{i=1}^{26} \frac{1}{26^2} = 26 \cdot \frac{1}{26^2} = \frac{1}{26} \approx 0{,}0385.$$

In einem solchen sinnlosen Buchstabensalat haben wir also nur eine etwa halb so große Chance, ein Paar aus gleichen Buchstaben zu treffen: Nur etwa jedes 26-te Buchstabenpaar besteht aus gleichen Buchstaben.

Was, zum Kuckuck, hat dies alles mit Chiffrieren und der Kryptoanalyse der Vigenère-Chiffre zu tun? Diese Frage soll sogleich beantwortet werden – allerdings werden wir zur Vigenère-Verschlüsselung erst später kommen.

Lassen Sie uns für einen Augenblick zu *mono*alphabetischen Chiffrierungen zurückkehren. Da eine monoalphabetische Chiffrierung im Grunde nur eine Permutation der Buchstaben ist, bleibt die Häufigkeitsverteilung der Buchstaben erhalten. Die Häufigkeiten der einzelnen Buchstaben werden zusammen mit den Buchstaben permutiert. Zum Beispiel gehört die Häufigkeit 0,174 nicht mehr zu dem Buchstaben **e**, sondern zu dem entsprechenden Geheimtextbuchstaben.

Allgemein kann man beweisen, dass der Koinzidenzindex (oder gleichwertig $\sum p_i^2$) *größer* wird, wenn der Text *un*regelmäßiger wird, und *kleiner* wird, je *gleichmäßiger* der Text ist. Der Wert 0,0385 ist das absolute Minimum für den Koinzidenzindex. Diese Behauptung soll in Übungsaufgabe 15 bewiesen werden.

Also bleibt bei einer monoalphabetischen Chiffrierung der Koinzidenzindex gleich, während er bei einer polyalphabetischen sinkt. Denn polyalphabetische Chiffrierungen sind ja gerade dazu gemacht, die Häufigkeiten der einzelnen Buchstaben einander anzugleichen.

Daraus leiten wir einen Test ab, der uns sagt, ob ein vorgelegter Geheimtext von einer monoalphabetischen Chiffrierung herkommt oder nicht.

Algorithmus 2.4: Test, ob ein Geheimtext monoalphabetisch verschlüsselt wurde

Man berechnet den Koinzidenzindex. Wenn dieser ungefähr 0,0762 ist, so ist die Chiffrierung wahrscheinlich monoalphabetisch. Wenn der Koinzidenzindex deutlich kleiner ist, dann ist der Text nicht monoalphabetisch chiffriert.

Nun verwenden wir den Koinzidenzindex, um die Schlüsselwortlänge eines Vigenère-chiffrierten Geheimtextes zu berechnen. Das Ziel ist es, den Koinzidenzindex dieses Textes zu berechnen – ohne den Text zu kennen.

Da ein polyalphabetischer Algorithmus verwendet wurde, ist der Koinzidenzindex kleiner als 0,0762. Aber um wie viel kleiner? Antwort: Das kommt darauf an. Genauer gesagt: Es kommt auf die Länge des Schlüsselworts an.

Um aus dem Koinzidenzindex die Länge des Schlüsselworts ausrechnen zu können, müssen wir diese Länge bezeichnen. Sei h die Länge des Schlüsselworts, das heißt die Anzahl seiner Buchstaben. Wir nehmen an, dass es aus lauter verschiedenen Buchstaben besteht.

Die Formel herzuleiten, ist eigentlich nicht schwierig, aber einigermaßen aufwändig. Wenn Sie im Augenblick keine Lust auf die Herleitung von Formeln haben, so bitte ich Sie, die folgende Seite getrost zu überblättern und erst zu Beginn des Abschn. 2.3.3 wieder aufzumerken.

Unsere Strategie ist die folgende: Wir tun so, als ob wir die Länge h des Schlüsselwortes kennen würden, berechnen daraus den Koinzidenzindex – und drehen dann die Formel um.

Wir stellen uns also vor, dass Mr. X einen Geheimtext hat, von dem er weiß, dass er Vigenère-verschlüsselt ist, und von dem er zumindest ahnt, dass das Schlüsselwort aus genau h Buchstaben besteht.

Daher kann Mr. X die Buchstaben des Geheimtexts isolieren, die mit dem ersten Schlüsselwortbuchstaben verschlüsselt wurden. Das sind die Buchstaben Nr. 1, Nr. h+1, Nr. 2h+1, … Entsprechend kann er die Buchstaben isolieren, die mit dem zweiten Schlüsselwortbuchstaben verschlüsselt wurden; es handelt sich um die Buchstaben Nr. 2, h+2, 2h+2, … Und so weiter.

Mr. X systematisiert dies, indem er den Geheimtext zeilenweise in h Spalten schreibt. Dann befinden sich in der ersten Spalte die Buchstaben Nr. 1, Nr. h+1, Nr. 2h+1 und so fort., also all diejenigen Buchstaben, die mit Hilfe des ersten Schlüsselwortbuchstabens chiffriert wurden. Entsprechend befinden sich in der zweiten Spalte diejenigen Buchstaben Nr. 2, h+2, 2h+2, …,

Erster Schlüsselwort-buchstabe	Zweiter Schlüsselwort-buchstabe	Dritter Schlüsselwort-buchstabe	...	h-ter Schlüsselwort-buchstabe
1	2	3	...	h
h+1	h+2	h+3	...	2h
2h+1	2h+2	2h+3	...	3h
3h+1	3h+2	3h+3	...	4h
...	...	...	...	...

Abb. 2.5 Nummern der Buchstaben eines Vigenère-verschlüsselten Textes

also diejenigen, die mit Hilfe des zweiten Schlüsselwortbuchstabens verschlüsselt wurden. Und so weiter. Aus Abb. 2.5 wird dieses Vorgehen deutlich.

Wenn man dieses Schema sorgfältig betrachtet, kann man den Koinzidenzindex ausrechnen.

Erste Beobachtung: Die Wahrscheinlichkeiten In jeder Spalte stehen Buchstaben, die durch dieselbe monoalphabetische Chiffrierung (sogar durch dieselbe Verschiebe-Chiffre) gewonnen wurden. Nach dem Algorithmus 2.4, dem Test, ob ein Text monoalphabetisch verschlüsselt wurde, ist die Chance, in einer Spalte ein Paar aus gleichen Buchstaben zu treffen, gleich 0,0762.

Nun betrachten wir die Paare aus Buchstaben, die in verschiedenen Spalten stehen. Wenn die Schlüsselwortlänge groß ist, stehen die Buchstaben eines solchen Paares in keinem inneren Zusammenhang. Da die zugehörigen Verschlüsselungsalphabete „zufällig" gewählt wurden (die Buchstaben des Schlüsselworts sind alle verschieden!), kann ein solches Paar nur zufällig aus gleichen Buchstaben bestehen.

Die Wahrscheinlichkeit dafür ist wesentlich niedriger als 0,0762, etwa 0,0385. Sie ist exakt 0,0385, wenn das Schlüsselwort eine zufällige Buchstabenfolge ist. Falls nicht, ist diese Wahrscheinlichkeit etwas höher.

Zweite Beobachtung: Die Anzahlen Wir zählen die Anzahl der Buchstabenpaare aus gleichen Spalten und die Anzahl der Paare aus verschiedenen Spalten.

Wenn der Geheimtext insgesamt n Buchstaben hat, so stehen in jeder Spalte genau n/h Buchstaben.

Bemerkung: Auf die Betrachtung von Rundungsfehlern verzichten wir hier grundsätzlich; der Text möge so lang sein, dass diese Fehler nicht ins Gewicht fallen.

Um den ersten Buchstaben zu wählen, gibt es genau n Möglichkeiten. Ist dieser Buchstabe gewählt, so liegt auch die Spalte, in der dieser sich befindet, fest. In dieser Spalte gibt es noch $n/h - 1$ andere Buchstaben, also $n/h - 1$ Möglichkeiten, den zweiten Buchstaben zu wählen. Also ist die Anzahl der

Paare von Buchstaben, die sich *in derselben Spalte* befinden, gleich

$$n \cdot \left(\frac{n}{h} - 1\right) / 2 = \frac{n\,(n-h)}{2h}.$$

Nun zu den Paaren aus Buchstaben aus verschiedenen Spalten. Wieder wählen wir zunächst den ersten Buchstaben (n Möglichkeiten). Da es genau $n - n/h$ Buchstaben außerhalb der Spalte gibt, die durch den ersten Buchstaben festgelegt ist, ergibt sich als die Anzahl der Paare von Buchstaben *aus verschiedenen Spalten*

$$n \cdot \left(n - \frac{n}{h}\right) / 2 = \frac{n^2 \cdot (h-1)}{2h}.$$

Nun fassen wir die Beobachtungen zusammen: Die erwartete *Anzahl* A von Paaren aus gleichen Buchstaben ist

$$A = \frac{n \cdot (n-h)}{2h} \cdot 0{,}0762 + \frac{n^2 \cdot (h-1)}{2h} \cdot 0{,}0385.$$

Also ist die *Wahrscheinlichkeit*, ein Paar aus gleichen Buchstaben zu treffen, gleich

$$\begin{aligned}\frac{A}{n\,(n-1)\,/\,2} &= \frac{n-h}{h\,(n-1)} \cdot 0{,}0762 + \frac{n\,(h-1)}{h\,(n-1)} \cdot 0{,}0385 \\ &= \frac{0{,}0377n + h\,(0{,}0385 - 0{,}0762)}{h\,(n-1)}.\end{aligned}$$

Wir wissen, dass der Koinzidenzindex eine sehr gute Annäherung an diese Zahl ist; daher gilt

$$I = \frac{0{,}0377n}{h\,(n-1)} + \frac{0{,}0385n - 0{,}0762}{n-1}.$$

Durch Umformen und Auflösen nach h ergibt sich daraus die wichtige, auf Friedman zurückgehende Formel für die Länge h des Schlüsselworts:

Algorithmus 2.5: Friedman-Test (Bestimmen der Schlüsselwortlänge)

Man bestimmt den Koinzidenzindex I und setzt ihn in folgende Formel ein:

$$h = \frac{0{,}0377n}{I \cdot (n-1) - 0{,}0385n + 0{,}0762}.$$

Diese Formel sieht vielleicht schwierig aus, und jedenfalls war ihre Herleitung langwierig. Es ist jedoch einfach, sie in einem konkreten Fall anzuwenden. Man braucht nämlich nur die Länge n des Textes und die Häufigkeiten n_i der Buchstaben – und mit diesen lächerlich wenigen Daten liefert uns die Zauberformel dann die Länge h des Schlüsselworts!

Nun ernten wir die Früchte unserer Arbeit, indem wir diese Theorie auf unser Beispiel vom Beginn dieses Abschnitts anwenden. Per Strichliste zählen wir die n_i's; es ergibt sich $n = 368$ und

$$\sum_{i=1}^{26} n_i\,(n_i - 1) = 6468.$$

Also ist

$$I = \frac{6468}{135.056} \approx 0{,}048.$$

Somit handelt es sich mit großer Wahrscheinlichkeit um eine polyalphabetische Chiffrierung. Mit der Formel aus Algorithmus 2.5 berechnen wir nun die Schlüsselwortlänge; es ergibt sich $h \approx 3{,}937$. Dies deutet zusammen mit den Ergebnissen des Kasiski-Tests darauf hin, dass die Schlüsselwortlänge tatsächlich 5 (und nicht 10, 15 oder 20) ist.

2.3.3 Bestimmung des Schlüsselworts

Nachdem nun die Schlüsselwortlänge bestimmt ist, geht es darum, das Schlüsselwort selbst zu erkennen. Das ist aber nicht mehr schwierig.

Wenn der Kryptoanalytiker Mr. X die Schlüsselwortlänge h kennt, so weiß er, dass die Buchstaben, die unter dem ersten Schlüsselwortbuchstaben stehen (das sind die Buchstaben Nr. 1, h+1, 2h+1, …) durch dieselbe monoalphabetische Chiffrierung, ja sogar durch dieselbe Verschiebe-Chiffrierung gewonnen worden sind. Es genügt in der Regel also, das Äquivalent des Buchstabens **e** zu finden.

Entsprechendes gilt für die Buchstaben, die unter dem zweiten Schlüsselwortbuchstaben stehen (also die Buchstaben Nr. 2, h+2, 2h+2, …). Und so weiter.

In unserem Beispiel ist $h = 5$. Von den 74 Buchstaben des „ersten" monoalphabetischen Teiltextes sind 13 gleich **F**. Daher entspricht **e** dem Buchstaben **F.** Ein Blick auf das Vigenère-Quadrat zeigt sofort, dass der erste Schlüsselwortbuchstabe **B** ist.

Auf diese Weise kann man leicht das Schlüsselwort erkennen und damit den Text entschlüsseln (vergleiche Übungsaufgabe 6).

2.4 Schlussbemerkungen

Wir haben gesehen, dass jede Vigenère-Chiffrierung zu knacken ist – sogar mit relativ einfachen Mitteln. Jede? – Nein, natürlich nur solche, bei denen das Schlüsselwort ziemlich kurz ist.

Konsequenterweise betrachten wir jetzt Vigenère-Chiffrierungen mit langem Schlüsselwort. Um uns nicht durch nebensächliche Diskussionen ablenken zu lassen (was heißt „lang"?), behandeln wir von vornherein die längstmöglichen Schlüsselwörter: Unsere Schlüsselwörter sollen so lang wie der Klartext sein. Wir stellen zwei Tricks vor, die beide dazu dienen, Mr. X die Freude an den oben beschriebenen Methoden zu vergällen.

1. Trick Man könnte versuchen, als Schlüssel„wort" den Text eines Buches zu verwenden. Ein solcher Schlüssel hat bestimmt den Vorteil, ohne große Probleme übermittelt werden zu können. Zum Beispiel muss der Empfänger nur die Information „Fontane: Irrungen, Wirrungen" haben, um den Geheimtext mit dem größten Vergnügen mittels des folgenden „Wortes" dechiffrieren zu können:

> An dem Schnittpunkte von Kurfürstendamm und Kurfürstenstraße, schräg gegenüber dem „Zoologischen", befand sich in der Mitte der siebziger Jahre noch eine große, feldeinwärts sich erstreckende Gärtnerei, deren kleines, dreifenstriges, in einem Vorgärtchen um etwa hundert Schritte zurückgelegenes Wohnhaus, trotz aller Kleinheit und Zurückgezogenheit, von der vorübergehenden Straße sehr wohl erkannt werden konnte. Was aber sonst noch zu dem Gesamtgewese der Gärtnerei gehörte, ja die recht eigentliche Hauptsache derselben ausmachte, war durch eben dies kleine Wohnhaus wie durch eine Kulisse versteckt …

Bei der Verwendung eines solchen Schlüssels laufen alle Methoden zur Bestimmung der Schlüsselwortlänge selbstverständlich ins Leere. Da aber der Schlüssel ein deutschsprachiger Text ist, schlagen statistisch signifikante Daten der Sprache auf den Geheimtext durch, so dass diese Chiffre nicht als sicher bezeichnet werden kann. Der erste, der diese Schwäche erkannte, war ebenfalls Friedman. Deshalb gehen wir noch einen Schritt weiter.

2. Trick Beim 1. Trick konnte Mr. X immer noch Statistik benutzen, weil das Schlüsselwort statistisch erfassbare Widerhaken aufwies. Deshalb wählen

wir nun als Schlüsselwort eine praktisch unendlich lange, völlig bunt gewürfelte Folge von Buchstaben, an der sämtliche statistische Tests widerstandslos abgleiten. Das ist eine (Buchstaben-)Zufallsfolge, die man sich z. B. erzeugen kann durch Werfen eines fairen 26-seitigen Buchstabenwürfels. Eine solche Folge nennen wir einen *Buchstabenwurm*.

Ein derartiger Wurm hat die Eigenschaft, dass man aus keinem noch so langen Stück auch nur einen einzigen weiteren Buchstaben vorhersagen kann. Verschlüsselt man einen Klartext mit Hilfe eines Buchstabenwurms, so hat offenbar auch der Geheimtext keine statistisch signifikanten Ansatzpunkte mehr, bei denen Mr. X für seine Kryptoanalyse einhaken könnte. Auch wenn er ein noch so langes Stück Geheimtext (ja sogar ein beliebig langes zusammengehöriges Stück Klartext-Geheimtext) kennt, kann er keinen einzigen (weiteren) Buchstaben bestimmen. Ein solches System ist sogar *theoretisch sicher*! Mit anderen Worten: Es bietet *perfekte Sicherheit*.

Diese perfekten Systeme werden wir im folgenden Kapitel genauer unter die Lupe nehmen.

2.5 Übungsaufgaben

1 Konstruieren Sie (auf dem Papier) eine „Maschine“, mit deren Hilfe Sie die in Abschn. 2.1 angegebene Chiffrierung vollziehen können. Konstruieren Sie entsprechend eine „Maschine“ zum Dechiffrieren.

2 @(a) Schreiben Sie ein Programm, das die in Abschn. 2.1 angegebenen homophone Chiffrierung realisiert.
   (b) Chiffrieren Sie damit einen relativ langen Text, und überprüfen Sie, ob (i) alle Buchstaben ungefähr mit der gleichen Häufigkeit auftreten, (ii) ob die Paare von Buchstaben im Geheimtext gleichmäßig verteilt sind.

3 Die Chiffriermethode aus Abschn. 2.1 ist nur ein Beispiel aus einer ganzen Klasse von Verfahren. Beschreiben Sie diese Klasse – und zwar so, dass klar wird, was der „Algorithmus“ ist und welches die „Schlüssel“ sind.

4 Angenommen, für jeden Buchstaben ist die Anzahl der Geheimtextsymbole die von Tab. 2.1. Das heißt, dass wir die Verteilung (5, 2, 3, 5, 17, 2, 3, 4, 8, 1, 1, 3, 3, 10, 3, 1, 1, 7, 7, 6, 3, 1, 1, 1, 1, 1) vorliegen haben.
   (a) Wie können die Symbole 00, …, 99 den Buchstaben zugeordnet werden, so dass
   - jeder Buchstabe die vorgeschriebene Anzahl von Geheimtextsymbolen erhält, und
   - diese Zuordnung „zufällig“ ist?

@(b) Schreiben Sie ein Programm, das jedem Buchstaben die geforderte Anzahl von Geheimtextsymbolen in zufälliger Weise zuordnet.

(c) Wie viele Zuordnungen der Symbole 00, …, 99 zu den Buchstaben gibt es, wenn jeder Buchstabe die geforderte Zahl von Geheimtextsymbolen erhalten soll?

5 Bestimmen Sie weitere Folgen aus gleichen Buchstaben im Beispiel aus Abschn. 2.3.

6 Bestimmen Sie alle Schlüsselwortbuchstaben des Beispiels aus Abschn. 2.3 und entziffern Sie den Text.

7 *Der folgende Text wurde mit dem Vigenère-Algorithmus verschlüsselt. Bestimmen Sie mit dem Kasiski-Test die Schlüsselwortlänge, bestimmen sie das Schlüsselwort und entschlüsseln Sie den Text:

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

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

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

Achtung: Bei so kurzen Texten ist der häufigste Buchstabe nicht immer e. Arbeiten Sie sorgfältig und mit Überlegung, sonst ergibt sich nur Murks.

8 Wenden Sie den Friedman-Test auf den Text der vorigen Übungsaufgabe an und vergleichen Sie das Ergebnis mit dem Kasiski-Test.

9 Ist die folgende Aussage richtig: „Es ist entscheidend, dass an jeder Stelle des Kryptogramms der Schlüssel eindeutig den Klartextbuchstaben zu jedem Geheimtextbuchstaben festlegt"?

10 Kann man durch statistische Untersuchungen eine Transpositions- von einer Substitutionschiffre unterscheiden?

11 Man kann das Verfahren von Vigenère auch dann anwenden, wenn man die Buchstaben durch die Zahlen 0, 1, 2, …, 25 darstellt. Beschreiben

Sie diese „modulo 26-Variante“ des Vigenère-Algorithmus (vergleichen Sie dazu Algorithmus 1.3).

12 @(a) Schreiben Sie ein Programm, das gleiche Folgen von Buchstaben in einem Text findet und den Abstand zwischen diesen Folgen bestimmt.

(b) Schreiben Sie ein Programm, das den Kasiski-Test benutzt, um die Schlüsselwortlänge eines Vigenère-verschlüsselten Textes zu bestimmen.

13 Chiffrieren Sie einen deutschen Text (der mindestens 40mal so viele Buchstaben enthält wie das Schlüsselwort) nach Vigenère, geben Sie diesen einer Freundin mit der Aufforderung, den Text zu entziffern.

14 Berechnen Sie den Koinzidenzindex des Geheimtextes aus Übungsaufgabe 20 von Kap. 1.

15 @Schreiben Sie ein Programm, das

(i) die Länge n eines Textes bestimmt,

(ii) die Häufigkeiten $n_1, \ldots, n_{26}$ der Buchstaben dieses Textes bestimmt, und

(iii) den Koinzidenzindex des Textes berechnet.

16 Seien $p_1, \ldots, p_{26}$ die Häufigkeiten der Buchstaben **a**, …, **z** eines Textes.

(a) Zeigen Sie

$$\sum_{i=1}^{26} p_i^2 = \frac{1}{26} + \sum_{i=1}^{26} \left(p_i - \frac{1}{26} \right)^2 .$$

(b) Erklären Sie, weshalb

$$\sum_{i=1}^{26} p_i^2$$

nicht kleiner als 1/26 (= 0,038) werden kann und diskutieren sie den Fall der Gleichheit.

17 @Entwerfen Sie ein Programm zur Vigenère-Chiffrierung, sowie eines, das Vigenère-chiffrierte Texte knackt (*).

18 ¿ M R U W N I E H C I L K R I W E G L O F E S E I D T S I

19 Der erste polyalphabetische (besser: nicht-monoalphabetische) Substitutionsalgorithmus wurde im Jahre 1470 von Leon Battista Alberti veröffentlicht. Er benutzt eine Maschine aus zwei konzentrischen Scheiben mit je 26 Zeichen (Alberti selbst verwendete 24 Zeichen). Auf der inneren Scheibe steht das Geheimtextalphabet in beliebiger, aber fester Anordnung; auf der äußeren Scheibe finden sich die Zahlen 1, 2, 3, und 4, sowie alle Buchstaben des Alphabets – mit Ausnahme der vier Buchstaben j, y, x, q, die die

geringsten Häufigkeiten aufweisen. Sender und Empfänger müssen sich auf einen Buchstaben, den sogenannten *Indikator* einigen.
Zum Verschlüsseln stellt der Sender die Scheibe beliebig ein. Er sucht den Indikator auf dem äußeren Ring und übermittelt als ersten Buchstaben denjenigen Geheimtextbuchstaben, der dem Indikator entspricht. Nun wird der Klartext verschlüsselt, als würde es sich um eine monoalphabetische Chiffrierung handeln. Falls zufällig einer der Buchstaben j, y, x oder q vorkommt, wird dieser einfach weggelassen.
Wann immer der Sender die Beziehung Klartext-Geheimtext verändern möchte, wählt er eine der Zahlen 1, 2, 3, oder 4, bestimmt den zugehörigen Geheimtextbuchstaben und übermittelt diesen als den nächsten Buchstaben. Dann dreht er die Scheibe so, dass dieser Geheimtextbuchstabe gegenüber dem Indikator liegt.
Aufgaben:
(a) Basteln Sie ein Modell für den *Alberti-Algorithmus* und verschlüsseln Sie eine Nachricht.
(b) Wie muss der Empfänger beim Entschlüsseln vorgehen?
(c) Welche Vorteile hat der Alberti-Algorithmus gegenüber dem Vigenère-Algorithmus?
(d) Welches sind die Schlüssel beim Alberti-Algorithmus?
(e) Wie sicher ist dieser Algorithmus?

Zusatzinformation: Eine „Modifikation" des Alberti-Algorithmus wurde in den italienischen Streitkräften noch nach dem zweiten Weltkrieg benutzt! Keine Angst: Der Alberti-Algorithmus wurde nur dazu verwendet, bereits verschlüsselte Nachrichten nochmals zu „überschlüsseln".

20 Machen Sie sich einen schönen Tag und lesen Sie Fontanes Roman *Irrungen, Wirrungen*.

http://www.springer.com/978-3-658-05975-0

Kryptologie
Eine Einführung in die Wissenschaft vom Verschlüsseln, Verbergen und Verheimlichen
Beutelspacher, A.
2015, XVII, 187 S. 52 Abb., Softcover
ISBN: 978-3-658-05975-0