ETH Institut für Verhaltenswissenschaft
Departement Informatik

Symmetrische Kryptographie

Leitprogramm in Informatik
Stufe Gymnasium
Dauer: 5 bis 6 Stunden

Adrian Schneider
Betreuer: Juraj Hromkovic
(Unerprobtes Leitprogramm)

Einführung

Um was geht es?
Wahrscheinlich mit oder zumindest sehr bald nach der Erfindung der Schrift, begannen manche Menschen ihre Nachrichten zu verschlüsseln, während andere versuchten, die Botschaften trotzdem zu lesen. Zwar waren die verwendeten Methoden am Anfang recht einfach, doch dachten sich die Geheimniskrämer immer komplexere Verfahren aus, um ihre Geheimnisse auch geheim zu halten, während die an den Geheimbotschaften interessierten (die oft ein und die selbe Person waren...), die Kunst des Codeknackens immer weiter vorantrieben. So wurden die verwendeten Methoden beider Seiten im Laufe der Jahrhunderte immer komplexer.

Bedeutung der Kryptografie heute
Im Internetzeitalter in dem wir heute leben, ist die Kryptographie nicht mehr aus unserem Leben weg zu denken. Ohne Datenverschlüsselung gäbe es weder Internetbanking noch online Bestellungen bei Internetshops. Sowohl in die Entwicklung neuer Algorithmen als auch in das Brechen derselben werden ungeheure Summen an Geld und Forschung investiert.

Was kannst du nach Bearbeitung dieses Leitprogramms?
Du verstehst einige einfache Verfahren der Verschlüsselung und kennst gängige Techniken, um solche zu brechen. Das Ziel ist aber nicht das Auswendig lernen von Algorithmen in grossen Mengen, sondern das Verständnis (und im Idealfall auch die Faszination) des ewigen Wettkampfs zwischen ’Gut’ und ’Böse’ in der Welt der Kryptographie, wobei überhaupt nicht klar ist, welche Seite jeweils die ’Guten’ sind...

Inhaltsverzeichnis

Arbeitsanleitung

Mit Hilfe dieses Leitprogramms wirst du das Thema „Symmetrische Kryptographie“ selbstständig erarbeiten. Das Programm ist in Kapitel gegliedert, die immer gleich strukturiert sind:

- **Übersicht**: Worum geht es in diesem Kapitel? Was ist zu tun?
- **Lernziel**: Was kann ich nach bearbeiten dieses Kapitels?
- Dann folgt der eigentliche **Lernstoff-Abschnitt**. Der besteht aus Wissensvermittlung und damit verbundenen Übungsaufgaben. Die wirst du selbstständig lösen und nachher mit den Lösungen am Ende des Kapitels vergleichen.
- **Lernkontrolle**: Habe ich alles verstanden.
- **Lösungen** zu den gestellten Aufgaben.

Nach jedem Kapitel wird den neu erlerntes Wissen in einem Kapiteltest geprüft. Die Testfragen holst du bei deiner Lehrerin oder deinem Lehrer.

Wenn du einmal nicht weiter weißt….
Solltest du einen Abschnitt oder eine Aufgabe auch nach längerem Überlegen (damit sind mindestens dreimal lesen und 5 Minuten nachdenken gemeint) nicht verstehen, so sind deine MitschülerInnen deine erste Anlaufstelle. Falls ihr auch gemeinsam nicht mehr weiterkommt, könnt ihr die Lehrperson um Hilfe fragen. Ausserdem gibt es als Anhang zum Leitprogramm eine Liste mit Erklärungen zu häufig verwendeten Begriffen.

Muss ich das ganze Programm durcharbeiten?
Jein. Der ganze Stoff aller Kapitel ist obligatorisch. Aber für die Schnellerlerner hat es in jedem Kapitel noch Zusatzaufgaben, die freiwillig sind. Meistens geht dabei darum, ein kleines Programm zu erstellen oder zu einem Thema noch mehr zu erfahren.

Kapitel 1 – Am Anfang war Caesar

Übersicht

Was lernst du hier?
Wir beginnen (beinahe) am Anfang der Kryptographie, nämlich bei Caesar und werden im Laufe des Kapitels erkennen, dass seine Geheimnisse nicht so geheim waren wie er dachte.

Was tust du?
Lies die Aufgaben durch und versuche die Aufgaben zu lösen. Manche sind kleine Verständnisfragen, manche erfordern ein wenig Überlegung und wieder andere erfordern ein klein wenig kryptographischen Spürsinn. Hilfsmittel brauchst du keine – ausser einer Zeitung oder einem Buch.

Lernziele:
Nachdem du dieses Kapitel durchgearbeitet hast,

- kannst du die Begriffe „Kryptografie“ und „symmetrisch“ erklären und von ähnlichen Begriffen abgrenzen.
- weisst du wie Caesar seine Nachrichten verschlüsselt hat.
- weisst du auch, wieso Caesars Geheimnachrichten nicht sehr sicher waren

Alles klar? Dann genug Vorbemerkungen. Los geht’s!

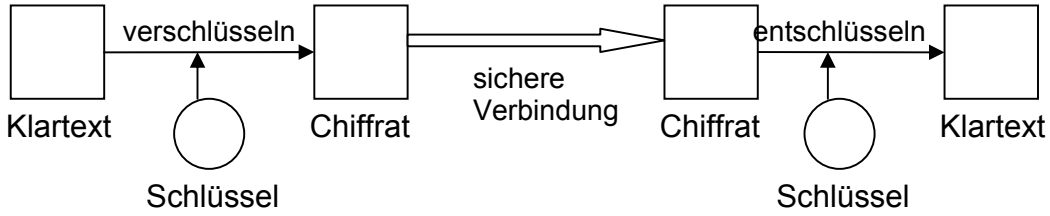
1.1. Begriffsklarifizierung

Wann spricht man von Kryptographie?
Auch wenn in der Vorbemerkung angedeutet wurde, dass der Anfang der Kryptographie bei Caesar liegt, so stimmt das natürlich nicht ganz. Wahrscheinlich schon immer haben Leute versucht, Nachrichten zu verschicken, so dass niemand ausser dem beabsichtigten Empfänger sie lesen kann. Seien das nun Diplomaten, heimlich Verliebte oder Kriminelle, fast alle haben ein Bedürfnis nach Diskretion.
Dazu gibt es unterschiedliche Möglichkeiten. Man kann z.B. einen vertrauenswürdigen Boten einstellen. Die Beispiele wo das Vertrauen aber nicht gerechtfertigt war, sind allerdings unzählig. Oder aber man kann das Vorhandensein der Nachricht selbst verheimlichen, beispielsweise indem man in einem unauffälligen Text bestimmte Buchstaben markiert (zum Beispiel durch ein kleines Loch unter dem Buchstaben, das nur sichtbar ist, wenn man das Papier gegen das Licht hält), die dann zusammen eine versteckte Botschaft bilden.
Die Wissenschaft der Kryptographie (oft wird auch der Begriff Kryptologie verwendet), befasst sich mit der Kunst, einen Text so zu verändern (zu verschlüsseln), dass ihn jeder, der möchte, sehen kann, aber nur der bestimmte Empfänger auch verstehen (also entschlüsseln) kann.

Aufgabe 1
Welche(r) der folgenden Tricks fällt oder fallen unter den Begriff „Kryptographie“?

a) die Nachricht im Innern einer Kokosnuss zu verstecken.
b) jeden Buchstaben nach einem fixen Schema durch ein chinesisches Schriftzeichen zu ersetzen.
c) die Nachricht mit unsichtbarer Tinte zu schreiben.

Im folgenden Diagramm siehst du den schematischen Ablauf einer typischen kryptographischen Anwendung. Ein Text, der geheim gehalten werden soll (Klartext genannt, weil es sich um einen lesbaren, z.B. deutschen Text handelt), soll an jemand anderes verschickt werden, ohne dass ein Fremder den Text lesen kann. Dazu wird der Klartext mittels eines Schlüssels verschlüsselt und das dadurch entstandene Chiffrat an den Empfänger geschickt. Die so entstandene Verbindung wird als ‚sicher’ bezeichnet, was bedeutet dass niemand die Nachricht lesen kann. (Zumindest wird das oft angenommen). Mittels desselben Schlüssels, kann der Empfänger das Chiffrat wieder in Klartext umwandeln (entschlüsseln). ‚Symmetrisch’ bedeutet in diesem Zusammenhang, dass der Empfänger denselben Schlüssel verwendet wie der Sender. Es gibt auch asymmetrische Verfahren, wo unterschiedliche Schlüssel zum Einsatz kommen. Diese Verfahren sind komplizierter, eröffnen aber auch ganz neue Anwendungen. In diesem Leitprogramm wollen wir uns aber auf die symmetrischen Verfahren konzentrieren.



Aufgabe 2
Wenn beide den Schlüssel kennen müssen, verschiebt sich doch das Problem nur. Anstatt die Nachricht sicher zu überbringen, muss nun einfach der Schlüssel auf geheimen Wegen ans Ziel kommen. Denke dir mindestens zwei Gründe aus, wieso Kryptographie trotzdem einen Vorteil darstellt.

Falls du zu einem späteren Zeitpunkt nachschlagen möchtest, was ein Begriff schon wieder genau bedeutet, so findest du im Anhang zu diesem Leitprogramm ein kleines Wörterbuch mit Begriffserklärungen.

1.2. Caesar: Verschiebechiffre

So, nachdem du nun weisst, was sich hinter dem Begriff „symmetrische Kryptographie“ versteckt, ist es Zeit, sich dessen ersten bekannten Anwender zu widmen: dem römischen Feldherr und Staatsmann Julius Caesar (100 bis 44 v. Chr.). Obschon er nicht der Erfinder der Kryptographie war, ist von ihm als erstem überliefert, dass und wie er seine Geheimnisse verschlüsselt hat.

Caesars Chiffre funktioniert auf folgende Weise: jeder Buchstabe vom Klartext (dem für jedermann lesbaren Originaltext) wird durch den Buchstaben ersetzt, der im Alphabet 3 Stellen weiter hinten ist. Also a wird zu D, b zu E und so weiter. Die Buchstaben am Ende des Alphabets werden einfach wieder vorne angehängt. Als x wird zu A etc. Beachte, dass die Kleinbuchstaben Klartext bedeuten während für das Chiffrat Grossbuchstaben verwendet werden. Diese Notation wird im ganzen Leitprogramm die so sein.

Aufgabe 3
Beschreibe in eigenen Worten die Caesarchiffre. Verwende dazu alle Begriffe, die im schematischen Diagramm der Kryptographie vorkommen.

Jetzt ist der Zeitpunkt gekommen, um die Caesarverschlüsselung einmal anzuwenden:

Aufgabe 4
Suche dir jemanden in der Klasse, der oder die ebenfalls bei dieser Aufgabe angelangt ist und schreibt euch gegenseitig eine mit Caesars Methode verschlüsselte Nachricht. Anschliessend sollt ihr versuchen, die empfangene Nachricht nach Klartext zu entschlüsseln.

1.3. Prinzip von Kerkhoff

Bevor wir nun die Idee von Caesar genauer betrachten und weiterentwickeln wollen, soll hier noch kurz das ‚Prinzip von Kerkhoff’ erwähnt werden: dieser niederländische Philologe mit dem schönen Namen Jean Guillaume Hubert Victor François Alexandre Auguste Kerckhoff von Nieuwendorf (1835 bis 1903) hat nämlich folgende Regel aufgestellt, die nicht nur immer noch gültig ist, sondern sogar mit der modernen Kryptographie erst recht an Bedeutung gewonnen hat. Es lautet:

> *Die Sicherheit eines Kryptosystems darf nicht von der Geheimhaltung des Algorithmus abhängen. Die Sicherheit gründet sich nur auf die Geheimhaltung des Schlüssels.*

Besonders heutzutage, wo Chiffrieralgorithmen auf komplizierter Mathematik beruhen und zum Teil jahrelang entwickelt und danach geprüft werden, ist es viel sicherer einen anerkannt und vielfach getesteten Algorithmus zu verwenden, als selber etwas zu erfinden und darauf zu hoffen, dass niemand herausfindet, was man genau gemacht hat. Die Sicherheit eines Chiffriersystems besteht dann darin, dass es viel zu viele Schlüssel gibt, um sie alle durchzutesten. Moderne Verfahren mit 128 Bit langen Schlüsseln haben zum Beispiel so viele mögliche Schlüssel, dass auch modernste Computer jahrzehntelang rechnen müssten, um alle auszuprobieren.

Aufgabe 5
Caesar als ‚Erfinder’ der Kryptographie kannte dieses Prinzip wahrscheinlich noch nicht. Wieso hat die Kryptographie bei ihm aber trotzdem gut funktioniert?

1.4. Caesar, fortgesetzt

Wenn wir nun Caesars Chiffre auf Schlüssel und Algorithmus untersuchen, so stellen wir fest, dass in der oben stehenden Formulierung gar kein Schlüssel erwähnt wird. Allerdings kann man die Caesarchiffre aber verallgemeinern:

Algorithmus: Ersetze jeden Klartextbuchstaben durch den Buchstaben des Alphabetes, des x Positionen weiter hinten steht.
Schlüssel: Die Zahl x, um viele Positionen das Alphabet verschoben wird.

Der original Algorithmus von Caesar verwendet also den Schlüssel x=3.

Aufgabe 6
Wie viele verschiedene Schlüssel hat nun eine allgemeine Caesarchiffre? Beachte dazu folgende Punkte:

a) Das Alphabet hat 26 Buchstaben. (Freiwillige Zusatzaufgabe: versuche herauszufinden, wie viele Buchstaben das Alphabet der Römer hatte.)
b) Ist x=87 sinnvoll? Oder gibt es ein kleineres x, dass genau dasselbe bewirkt?
c) Gibt es ein x, das keine sinnvolle Verschlüsselung darstellt?

Aufgabe 7
Nachdem du nun weißt, wie viele verschiedene Schlüssel möglich sind. Hältst du das System für sicher? Oder in anderen Worten: wenn du weisst, dass ein Text mit der Caesarchiffre verschlüsselt ist, wie würdest du vorgehen um einen abgefangenen Geheimtext zu entschlüsseln, auch wenn du keine Ahnung hast, was der Schlüssel war? Wie lange würde das etwa dauern?

Aufgabe 8
Und jetzt ist es soweit: dein kryptographischer Spürsinn ist ein erstes Mal gefragt. Du erhältst folgenden Geheimtext von dem du stark vermutest, dass er (naiverweise, siehe Aufgabe 7) mit der allgemeinen Caesarchiffre verschlüsselt wurde. Die Aufgabe ist klar, versuche den Klartext zu bestimmen.

```
GRRK CKMK LAKNXKT TGIN XUS
```

Welches x wurde als Schlüssel verwendet?

Lernkontrolle

1) Verschlüssle den Satz „Rom wurde auch nicht in einem Tag gebaut.“ Mit der allgemeinen Caesarchiffre und x=7.
2) Was ist der Grund, dass Texte, die mit der Caesarchiffre verschlüsselt wurden, sehr leicht entschlüsselt werden können, auch wenn der Schlüssel nicht bekannt ist?
3) (freiwillig) Finde (sinnvolle) Wörter, die durch eine Verschiebung der Alphabets ineinander übergehen. Hinweis: schon Wörter der Länge drei sind eine Leistung.

Wenn du das Gefühl hast, alles verstanden zu haben, so kannst du zu deiner Lehrerin oder deinem Lehrer gehen und den Kapiteltest abholen.

Lösungen zu den Aufgaben

Lösung 1
Nur bei b) wird die Nachricht tatsächlich verschlüsselt, Bei a) und c) wird sie nur versteckt.

Lösung 2
- Man kann sich eventuell einmal persönlich treffen und dabei den Schlüssel abmachen. Von dann an kann man sicherer (ganz sicher ist nie etwas) kommunizieren, ohne persönlichen Kontakt.
- Der Schlüssel ist typischerweise kürzer als die eigentliche Botschaft. Er kann damit unauffälliger transportiert, bzw. versteckt werden.
- Moderne Algorithmen erlauben es tatsächlich, dass zwei Parteien einen gemeinsamen Schlüssel abmachen können, und zwar so, dass jedermann die ganze Kommunikation abhören kann. Trotzdem wissen am Schluss nur die beteiligten Parteien, was der geheime Schlüssel ist. Das tönt sehr abenteuerlich, funktioniert aber tatsächlich!

Lösung 3 und 4
Dazu braucht ihr keine Lösung.

Lösung 5
Als einer der Ersten, die Kryptographie verwendeten, hatte Caesar den entscheidenden Vorteil, dass sich kaum jemand Gedanken zu dem Thema gemacht hat und vielleicht nur wenige überhaupt vermuten, dass Verschlüsselung am Werk sein könnte. Darauf verlassen konnte er sich allerdings auch nicht.

Lösung 6
25 Schlüssel sind sinnvoll. Einer ist die Identität, also a zu A und b zu B etc. Niemand wird einen Text damit verschlüsseln wollen. Ein x von 27 entspricht exakt x=1, man dreht sich einfach einmal im Kreis.

Lösung 7
Nein, das System ist nicht sicher. Man kann einfach alle Schlüssel ausprobieren. Bei 25 verschienen Schlüsseln geht das sogar von Hand ziemlich schnell. Tatsächlich basierte die Sicherheit dazumal wohl einzig darauf, dass der Algorithmus selber unbekannt war (oder geglaubt wurde).

Lösung 8
Alle Wege fuehren nach Rom. Als x wurde 6 verwendet.

Lösungen der Lernkontrolle

1) YVT DBYKL HBJO UPJOA PU LUTLS AHN NLEHBA.
2) Die sehr geringe Anzahl Schlüssel ermöglicht es jedem Angreifer, alle möglichen Schlüssel auszuprobieren.
3) Das war eine freiwillige Rätselaufgabe. Das Rätsel soll ruhig weiter bestehen…

Kapitel 2 – Mehr Schlüssel = mehr Sicherheit?

Um was geht es?
Im ersten Kapitel haben wir gesehen, wie Caesar vor ca. 2000 Jahren seine Geheimbotschaften verschlüsselt hat. Wir haben aber auch gesehen, dass sein System nur eine sehr beschränkte Anzahl Schlüssel zulässt. Daher ist es sehr einfach, alle möglichen Schlüssel durchzuprobieren, bis man den passenden gefunden hat. Im Falle von Caesar ist das sogar von Hand möglich. Heutzutage kann man mit Computern aber Millionen von Schlüsseln innerhalb von Sekunden testen. Es ist also notwendig, Caesars Kryptosystem so zu verbessern, dass die Anzahl Schlüssel viel, viel grösser wird. In diesem Kapitel wirst du sehen, wie die Anzahl Schlüssel massiv erhöht werden kann. Allerdings wirst du dann feststellen, dass das neu vorgeschlagene Kryptosystem immer noch nicht genügend sicher ist.

Lernziele
Nachdem du dieses Kapitel bearbeitet hast,
- kennst du das Prinzip der monoalphabetischen Verschlüsselung
- weisst du, wieso damit verschlüsselte Nachrichten immer noch nicht sicher sind.
- kannst du die Häufigkeitsanalyse anwenden, um einen monoalphabetisch verschlüsselten Text zu entschlüsseln (ohne dass du den Schlüssel kennst).

Dieses Kapitel beginnt gleich mit einer Frage:

Aufgabe 1
Wenn man bei der Caesarchiffre alle Schlüssel ausprobiert, an was erkennt man, dass man den richtigen gefunden hat? Und wie könnte man dies lösen, wenn der Computer die Schlüsselsuche automatisch durchführen soll?

2.1. Monoalphabetische Verschlüsselung

Die Zahl der Schlüssel kann mit folgendem Trick erhöhen. Anstatt das Alphabet nur um eine feste Anzahl stellen zu verschieben, kann man die Buchstaben beliebig permutieren (vertauschen). Konkret geht man folgendermassen vor: Man schreibt das Klartextalphabet in eine Reihe, danach verteilt man alle Buchstaben des Alphabets wild gemischt auf eine zweite Reihe. Das sieht dann so aus:

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

Ein Text wird nun verschlüsselt, indem jeder Buchstabe durch den darunter stehenden Buchstaben ersetzt wird. ‚hallo' wird also zu ‚XCJJS'.

Aufgabe 2
Verschlüssle mit Hilfe der oben stehenden Permutation das Wort ‚kryptologie' und entschlüssle das Wort ‚UOXOAZPAE'.

Aufgabe 3
Was ist nun der Schlüssel bei diesem System? Anders gefragt, was muss der Empfänger alles wissen, um eine verschlüsselte Botschaft zu entziffern?

Aufgabe 4
Wie viele Schlüssel hat nun dieses System? Hinweis: um das ‚a' zu verschlüsseln hat man 26 Möglichkeiten. Danach bleiben für das ‚b' noch 25 mögliche Buchstaben, und so weiter.

Aufgabe 5
Angenommen, ein Computer könnte pro Sekunde eine Million Schlüssel ausprobieren. Wie lange müssten 1000 Computer zusammen rechnen, um alle Schlüssel durchzuprobieren?

2.2. Unregelmässige Sprachen

Du hast also gesehen, dass dieses neue, ‚monoalphabetisch' genannte Kryptosystem eine unheimlich grosse Anzahl mögliche Schlüssel hat. Was bei Caesar noch sehr einfach funktionierte, nämlich alle Schlüssel durchzutesten, ist hier nun unmöglich.

Doch die Gegenseite, auch Kryptoanalyse genannt, hat auch ein paar Tricks herausgefunden. So hat man relativ schnell gemerkt, dass es gar nicht nötig ist, alle Schlüssel zu testen. Die Idee dahinter sind Unregelmässigkeiten, die in allen Sprachen vorkommen und die man ausnutzen kann. Die einfachste und wahrscheinlich offensichtlichste davon ist die Tatsache, dass nicht alle Buchstaben gleich häufig verwendet werden. Im Deutschen zum Beispiel kommen e, i oder n viel häufiger vor als q, x, und y.

Aufgabe 6
Nimm einen Text aus einem Buch oder einer Zeitung und zähle, wie oft jeder Buchstabe vorkommt. Am besten machst du eine Strichliste.
Wem das zu mühsam ist, der kann alternativ auch ein kurzes Programm schreiben, dass diese Aufgabe automatisch erledigt. Damit können natürlich viel längere Texte bearbeitet werden, was sich wiederum positiv auf die Verlässlichkeit der Resultate auswirken wird.

2.3. Monoalphabetische Kryptoanalyse

Wie können Unregelmässigkeiten in der Sprache nun ausgenützt werden, um einen monoalphabetisch verschlüsselten Text zu entschlüsseln? Die Idee ist, dass diese Unregelmässigkeiten auch im Chiffrat noch vorhanden sind. Wenn das e der häufigste Buchstabe des Klartexts ist, dann muss derjenige Buchstabe, durch den e ersetzt wurde, der häufigste Buchstabe im Chiffrat sein. Auf diese Weise kann man meistens die allerhäufigsten Buchstaben im Chiffrat entdecken und dazu verwenden, weitere Buchstaben zu erraten, so dass bald der ganze Text aufgedeckt wird.
Das tönt jetzt alles sehr kompliziert und aufwändig. Doch so schlimm ist es keinesfalls. Wir werden nun das ganze Verfahren nun anhand eine Beispiels demonstrieren.

Nehmen wir an, wir hätten folgenden verschlüsselten Text abgefangen, von dem wir vermuten, dass er monoalphabetisch verschlüsselt ist und ausserdem aus dem Deutschen stammt. Freundlicherweise sind sowohl Satzzeichen als auch Wortabstände erhalten geblieben, was die Sache natürlich einfacher macht:

```
CJ UAFFC CHZ WAZZ CHZCZ CJCN, LCT JDUOZ NAZGC SAUTC
LHC JACDVC XZRCTLTOJJCZ BXT WXCUNC GCFTAGCZ UAFFC,
LCJJCZ VTACPFC AECT ZXZ BX CZLC GHZGCZ, JO LAJJ CT
BXT ATECHF HWWCT XZFAXGNHDUCT KATL. LA LADUFC LCT
UCTT LATAZ, HUZ AXJ LCW PXFFCT BX JDUAPPCZ, AECT LCT
CJCN WCTVFC, LAJJ VCHZ GXFCT KHZL KCUFC, NHCP POTF XZL
WADUFC JHDU AXP LCZ KCG ZADU ETCWCZ: LOTF, WCHZFC CT,
VOCZZFC CT SA JFALFWXJHVAZF KCTLCZ.
```

Aufgabe 7
Wenn du das Chiffrat betrachtest, gibt es Hinweise, dass es sich tatsächlich um einen Deutschen Text handelt? Wieso ist der Text nicht englisch oder französisch? Nutze dazu die Tatsche, dass du sowohl die Satzzeichen als auch die Abstände sehen kannst.

Erster Schritt: Buchstabenhäufigkeit zählen
Durch die Analyse der Buchstabenhäufigkeit in diesem Chiffrat, können mehrere Sachen festgestellt werden. Erstens ob unsere Annahme – monoalphabetisch verschlüsselt - richtig war. Würden wir nämlich feststellen, dass die Buchstaben im Chiffrat praktisch gleichmässig verteilt sind, so würde es sich ziemlich sicher nicht um einen monoalphabetisch verschlüsselten Text handeln.
Ausserdem erlaubt uns das hoffentlich schon, ein paar häufig auftretende Buchstaben zu identifizieren. Betrachten wir dazu doch die 7 häufigsten Buchstaben im Text und vergleichen wir das mit den 7 häufigsten Buchstaben der deutschen Sprache:

Im Chiffrat		Im Deutschen	
C	62 mal	e	17,4%
Z	31 mal	n	9,8%
T	29 mal	i	7,5%
A	27 mal	s	7,3%
F	22 mal	r	7,0%
L	20 mal	a	6,5%
J	19 mal	t	6,2%

Leider kann man nun nicht einfach annehmen, dass diese Buchstaben eins zu eins zusammenpassen,

Aufgabe 8
Wieso nicht?

Zweiter Schritt: Häufige Buchstaben entschlüsseln
Aber man stellt fest, dass C genau doppelt so häufig vorkommt wie Z, der zweithäufigste Buchstabe. Im deutschen kommt das e, der häufigste Buchstabe auch beinahe doppelt so oft vor wie das n, das an zweiter Stelle steht. Man kann also guten Gewissens schlussfolgern, dass C dem e entspricht.
Die Frage ist nun, ob wir auch weiterhin Glück haben, und Z auch noch gerade n entsprechen würde. Im ersten Satz finden wir CHZ WAZZ CHZCZ und können feststellen, dass ZZ also zweimal derselbe Buchstabe vorkommt. Nur wenige Buchstaben können im Deutschen doppelt vorkommen. Das n ist einer davon. Das Auftreten von Doppelbuchstaben ist übrigens ein häufiger Trick der Kryptoanalytiker. Betrachten wir CHZ: C ist e, das könnte also ‚ein' heissen. Und CHZCZ stände dann für ‚einen'. Das soll uns als genügende Hinweise gelten, Z mit n und H mit i gleich zu setzten.

Somit wissen wir auch gleich, dass T, der dritthäufigste Buchstabe, nicht dem i entsprechen kann. Aber was denn sonst? Im Text finden wir mehrere CT und LCT. Mit C=e kann CT nur noch ‚er’ oder ‚es’ heissen (‚Ei’ geht nicht mehr, weil das i schon dem H zugeordnet wurde.). Auch finden wir dreimal LCT im Text. das würde in dem Fall für ‚der’ oder ‚des’ stehen, was durchaus Sinn macht. Nun ist ‚des’ aber wesentlich seltener als ‚der’, somit hätten wir zwei weitere Buchstaben: L als d und T als r. Schauen wir mal den teilentschlüsselten Text an:

```
eJ UAFFe ein WAnn einen eJeN, der JDUOn NAnGe SAUre
die JAeDVe XnRerdrOJJen BXr WXeUNe GeFrAGen UAFFe,
deJJen VrAePFe AEer nXn BX ende GinGen, JO dAJJ er
BXr ArEeHF HWWer XnFAXGNiDUer KArd. dA dADUFe der
Uerr dArAn, iUn AXJ deW PXFFer BX JDUAPPen, AEer der
eJeN WerVFe, dAJJ VeHn GXFer Kind KeUFe, NieP POrF Xnd
WADUFe JiDU AXP den KeG nADU EreWen: dOrF, WeinFe er,
VOennFe er SA JFAdFWXJiVAZF Kerden.
```

Dritter Schritt: Lücken schliessen

Dieser Text hat nun schon eine ziemliche Struktur, und man erkennt schnell, dass es sich tatsächlich um einen deutschen Text handelt. Viele Wörter sind beinahe ganz entschlüsselt und dienen uns nun als Anhaltspunkte, um die restlichen Buchstaben zu finden.

So finden wir zum Beispiel das Wort iUn. Das einzige deutsche Wort, das hier Sinn macht, ist ‚ihn’, also U=h. Dann ganz am Anfang: eJ. Weil das r schon belegt ist, kann das nur ‚es’ heissen. Mit dem Wissen um J=s und dem zweifachen vorkommen von dAJJ, liegt die Schlussfolgerung A=a sehr nahe. Das wiederum ermöglich uns das zweite Wort, das inzwischen zu haFFe wurde als ‚hatte’ zu identifizieren, also F=t.

Jetzt geht es sehr schnell weiter, der erste Teilsatz lautet ‚es hatte ein Wann einen eseN’, was den Schluss W=m und N=l zulässt. Weiter mit ‚der sDhOn lanGe Sahre die saeDVe’ kann mit ein wenig Raten zu D=c, O=o, G=g und V=k führen. Und so kann der Originaltext erstaunlich rasch gefunden werden:

> *Es hatte ein Mann einen Esel, der schon lange Jahre die Saecke unverdrossen zur Muehle getragen hatte, dessen Kraefte aber nun zu Ende gingen, so dass er zur Arbeit immer untauglicher ward. Da dachte der Herr daran, ihn aus dem Futter zu schaffen, aber der Esel merkte, dass kein guter Wind wehte, lief fort und machte sich auf den Weg nach Bremen: dort, meinte er, koennte er ja Stadtmusikant werden.*

Es handelt sich dabei um den Anfang der Bremer Stadtmusikanten, dem Märchen der Brüder Grimm.

Natürlich wird niemand die Satzzeichen so lassen, um möglichen Gegnern das Leben zu vereinfachen. Auch ein sehr beliebter Trick ist, die Wörter alle zusammen zu schreiben, so dass der Text möglichst wenig Struktur aufweist.

Doch das Prinzip der Kryptoanalyse wird dadurch jedoch nicht geändert, nur ein wenig erschwert.

Aufgabe 9
Funktioniert diese hier vorgestellte Analyse auch, wenn die Sprache des Originaltexts nicht bekannt ist? Müsste man das Vorgehen allenfalls anpassen?

Aufgabe 10
Welche Bedingungen müssen erfüllt sein, damit die Entschlüsselung funktioniert? (kannst du z.B. ‚KQPGMYSLYEHH' entschlüsseln?)

Aufgabe 11
So, nun ist dein ganzes Können gefordert: folgender Text soll von dir entschlüsselt werden. Gehe schrittweise vor, wie in diesem Kapitel gezeigt wurde. Der Originaltext ist deutsch und monoalphabetisch verschlüsselt. Die Umlaute wurden wie im Beispiel durch ae, ue und oe ersetzt. Viel Spass!

```
MTZ IXKIS WZTBBIK DHGLI DTRKPI IXK HZSIZ RTGJRHNQIZ SXP
BIXKIZ OZHU UKL BIXKIK JDIX QXKLIZK; LHB EUIENRIK RXIBB
RHIKBIG UKL LHB SHILNRIK WZIPIG. IZ RHPPI DIKXW JU EIXBBIK
UKL JU EZINRIK, UKL IXKSHG, HGB WZTBBI PIUIZUKW XKB GHKL
QHS, QTKKPI IZ LHB PHIWGXNRI EZTP KXNRP SIRZ BNRHOOIK. DXI
IZ BXNR KUK HEIKLB XS EIPPI WILHKQIK SHNRPI UKL BXNR MTZ
BTZWIK RIZUSDHIGJPI, BIUOJPI IZ UKL BCZHNR JU BIXKIZ OZHU:
"DHB BTGG HUB UKB DIZLIK? DXI QTIKKIK DXZ UKBIZI HZSIK
QXKLIZ IZKHIRZIK LH DXZ OUIZ UKB BIGEBP KXNRPB SIRZ RHEIK?"
"DIXBBP LU DHB, SHKK", HKPDTZPIPI LXI OZHU, "DXZ DTGGIK
STZWIK XK HGGIZ OZUIRI LXI QXKLIZ RXKHUB XK LIK DHGL UIRZIK,
DT IZ HS LXNQBPIK XBP. LH SHNRIK DXZ XRKIK IXK OIUIZ HK UKL
WIEIK VILIS KTNR IXK BPUINQNRIK EZTP, LHKK WIRIK DXZ HK
UKBIZI HZEIXP UKL GHBBIK BXI HGGIXK. BXI OXKLIK LIK DIW
KXNRP DXILIZ KHNR RHUB, UKL DXZ BXKL BXI GTB."
```

2.4. Schlussfolgerung

In diesem Kapitel hast du gesehen, dass es ein Leichtes ist, ein Kryptosystem zu entwerfen, das eine so riesige Anzahl Schlüssel hat, dass systematisches Durchprobieren nicht mehr möglich ist. Du hast aber auch gesehen, dass viele Schlüssel alleine noch lange keine Sicherheit bedeutet. Falls der verschlüsselte Text genügend statistische Unregelmässigkeiten aufweist, so dienen diese als Anhaltspunkte und ermöglichen einem geübten Kryptoanalysten, das Chiffrat trotzdem zu entschlüsseln.

Lernkontrolle

1) Beschreibe schematisch, wie eine monoalphabetische Verschlüsselung funktioniert. Was ist der Schlüssel, wie wird verschlüsselt, wie entschlüsselt?
2) Was sind weitere statistische Besonderheiten der deutschen Sprache, die ausgenützt werden können, um monoalphabetisch verschlüsselte Texte zu knacken?
3) Ist die Caesarchiffre auch eine monoalphabetische Verschlüsselung? Wieso?
4) (freiwillig) Finde einen deutschen Text mit mindestens 50 Buchstaben, in dem e nicht der häufigste Buchstabe ist.

Freiwillige Lektüre
In Edgar Allan Poe’s Kurzgeschichte *‚Der Goldkäfer’* steht ein monoalphabetisch verschlüsselter Text im Zentrum. Die Geschichte kannst du zum Beispiel hier lesen: http://gutenberg.spiegel.de/poe/kaefer/kaefe001.htm

Lösungen zu den Aufgaben

Lösung 1
Man kann das nur daran erkennen, das der entstanden Text auch Sinn macht. Die Chance ist extrem klein, dass zwei verschiedene Schlüssel zu einem Sinnvollen Text führen, siehe Lernkontrolle von Kapitel 1. Ein Computer kann einen beliebigen Schlüssel auf das Chiffrat anwenden und den entstandenen Text mit einem Wörterbuch vergleichen. Falls genügend bekannte Wörter vorkommen, kann er annehmen, dass das der richtige Schlüssel ist.

Lösung 2
Kryptologie wird zu GBWHSJSUAO und UOXOAZPAE bedeutet Geheimnis.

Lösung 3
Der Schlüssel ist die Permutation. Also im Prinzip das oben abgebildete Schema oder zumindest die untere Zeile davon. Das ist deutlich mehr als bei Caesar, wo der Schlüssel eine Zahl zwischen 1 und 25 war.

Lösung 4
Das System hat 26*25*…*1=26! Schlüssel. Das sind $4*10^{23}$.

Lösung 5
12.8 Milliarden Jahre.

Lösung 6
Das hängt natürlich von deinem Text ab. Bei Wikipedia findet man zum Beispiel folgende Tabelle:

Buchstabe	Relative Häufigkeit	Buchstabe	Relative Häufigkeit
e	17,40%	m	2,53 %
n	9,78 %	o	2,51 %
i	7,55 %	b	1,89 %
s	7,27 %	w	1,89 %
r	7,00 %	f	1,66 %
a	6,51 %	k	1,21 %
t	6,15 %	z	1,13 %
d	5,08 %	p	0,79 %
h	4,76 %	v	0,67 %
u	4,35 %	j	0,27 %
l	3,44 %	y	0,04 %
c	3,06 %	x	0,03 %
g	3,01 %	q	0,02 %

Lösung 7
Im englischen gibt es Wörter wie ‚I’ und ‚a’, die nur einen Buchstaben lang sind. Solche kommen in diesem Text nicht vor. Im französischen wiederum gibt es oft Apostrophs (z.B. j’ai) oder Bindestriche in Wörtern. Auch das findet man in diesem Text nicht.

Lösung 8
Die Buchstabenhäufigkeit hängt immer auch vom konkreten Text ab. Besonders bei kurzen oder extrem fachspezifischen Texten kann die Buchstabenhäufigkeit vom Soll abweichen. Die Wahrscheinlichkeiten für i, s und r sind sehr nahe beisammen und können leicht vertauscht sein in einem beliebigen Text.

Lösung 9
Wenn es sich um eine andere Sprache handelt, so können natürlich nicht die deutschen Buchstabenhäufigkeiten verwendet werden. Man kann aber mit Häufigkeitstabellen verschiedener Sprachen arbeiten und so eventuell herausfinden, welche am besten passt. Zum tatsächlichen Entschlüsseln sind Kenntnisse der jeweiligen Sprache aber Voraussetzung.

Lösung 10
Der Text muss genügend lang sein, damit die Buchstabenhäufigkeit einigermassen zuverlässig ermittelt werden kann. Auch muss der Text grossteils aus geläufigen Wörtern bestehen. Eine reine Auflistung von Fachbegriffen kann kaum entziffert werden.

Lösung 11
Es handelt sich bei dem Text um den Anfang von ‚Hänsel und Gretel’:

> *Vor einem grossen Walde wohnte ein armer Holzhacker mit seiner Frau und seinen zwei Kindern; das Buebchen hiess Haensel und das Maedchen Gretel. Er hatte wenig zu beissen und zu brechen, und einmal, als grosse Teuerung ins Land kam, konnte er das taegliche Brot nicht mehr schaffen. Wie er sich nun abends im Bette Gedanken machte und sich vor Sorgen herumwaelzte, seufzte er und sprach zu seiner Frau: "Was soll aus uns werden? Wie koennen wir unsere armen Kinder ernaehren da wir fuer uns selbst nichts mehr haben?" "Weisst du was, Mann", antwortete die Frau, "wir wollen morgen in aller Fruehe die Kinder hinaus in den Wald fuehren, wo er am dicksten ist. Da machen wir ihnen ein Feuer an und geben jedem noch ein Stueckchen Brot, dann gehen wir an unsere Arbeit und lassen sie allein. Sie finden den Weg nicht wieder nach Haus, und wir sind sie los."*

Der Schlüssel ist der folgende:

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

Lösungen zur Lernkontrolle

1) Der Schlüssel sind zwei Zeilen, in der ersten steht das normale Alphabet (das Original), und in der zweiten eine beliebige Permutation davon (der Code). Beim Verschlüsseln wird jeder Buchstaben des Klartextes durch denjenigen Codebuchstaben ersetzt, der unter dem Originalbuchstaben steht. Beim Entschlüsseln geht man genau umgekehrt vor: Der Chiffrebuchstaben wird im Code gesucht und durch den darüber stehenden Originalbuchstaben ersetzt.
2) Man kann zum Beispiel auf Kombinationen von aufeinander folgenden Buchstaben achten (ei und ie sind sehr häufig, qf oder ji werden sehr selten vorkommen). Ausserdem kann man die häufigsten Wörter mit zwei und drei Buchstaben betrachten.
3) Ja, eine Verschiebung ist ein Spezialfall der Permutation.
4) Selber ausdenken. Ist aber nicht sehr einfach,

Kapitel 3 – La chiffre indéchiffrable

Um was geht es?
Im ersten Kapitel haben wir die einfache Verschiebechiffre von Caesar kennen gelernt und herausgefunden, dass diese schon Aufgrund der geringen Anzahl Schlüssel sehr unsicher ist. Im zweiten Kapitel wurde die Schlüsselzahl drastisch erhöht, so dass eine blinde Attacke mittels Durchprobieren aller Schlüssel unmöglich wurde. Doch die Kryptoanalytiker haben wiederum auch neue Techniken entwickelt und konnten Unregelmässigkeiten in der Sprache ausnutzen, um auch monoalphabetisch verschlüsselte Botschaften zu entziffern. Also sind die Geheimniskrämer wieder gefragt, um sich ein neues Verfahren auszudenken, damit Geheimnisse auch wirklich geheim bleiben.

Was tust du?
In diesem Kapitel wirst du Verfahren kennen lernen, um gewisse Unregelmässigkeiten der Sprache zu verschleiern. Aber du wirst auch sehen, dass auch das noch nicht absolute Sicherheit bedeuten wird.

Lernziele
Nach der Bearbeitung dieses Kapitels,

- weisst du was eine polyalphabetische Chiffrierung ist und kennst die zwei Methoden, um diese zu Bewerkstelligen.
- kennst du die Ziele der homophonen Chiffre und weisst, worin ihre Schwachpunkte bestehen.
- kennst du die Verschlüsselung von Vigenère und weisst du auch, dass und warum sie doch ‚déchiffrable’ ist.

3.1. Polyalphabetische Chiffrierungen

Die Hauptschwäche der monoalphabetischen Chiffren bestand darin, dass sich die Buchstabenhäufigkeit vom Klartext exakt im Chiffrat widerspiegelte. Der Grund dafür war, dass jeder Klartextbuchstabe, immer auf ein und denselben Buchstaben verschlüsselt wird. Das muss natürlich nicht so sein! Die Idee der polyalphabetischen Chiffrierung ist, dass sich die Buchstabenhäufigkeiten des Chiffrats möglichst ausgleichen.
Dazu gibt es zwei Möglichkeiten:

- entweder wird ein Buchstabe nicht mehr mit einem, sondern mit mehreren Zeichen verschlüsselt.
- Oder man wechselt die Codierung während der Verschlüsselung.

Aufgabe 1
Was ist nun genau der Vorteil, wenn alle Buchstaben des Chiffrats ungefähr gleich häufig vorkommen?

3.2. Homophone Chiffren

Bei der ersten Variante werden die häufigen Buchstaben, also z.B. das ‚e' durch mehrere mögliche Zeichen verschlüsselt. Das nennt man *homophone Chiffren*. Allerdings muss man einen wichtigen Punkt dabei beachten:

Aufgabe 2
Betrachte folgende homophone Chiffrierung:

```
Original: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
Code :   H E N L * O W R * V Q G S * T C Y * B P U M D A F J
```

Wobei für die 4 häufigsten Buchstaben e, i, n und r folgende Spezialregeln gelten: e kann mit H, W oder B, i kann mit N, T oder D, n mit L oder H und r mit A oder J verschlüsselt werden. Dadurch können diese vier Buchstaben nicht mehr so einfach entdeckt werden.
Was ist das Problem bei dieser Chiffrierung? Was passiert, wenn du einen so verschlüsselten Text entschlüsseln willst?

Für die homophone Chiffrierung gelten also folgende Regeln: („Buchstaben' bezeichnet von nun an die Buchstaben des Klartexts, während „Zeichen' aus dem Alphabet des Chiffrats stammen.)

- Einem Klartext Buchstaben wird nicht mehr nur ein Zeichen, sondern manchmal gleich eine Menge von Zeichen zugeordnet.
- Die Zuordnung muss eindeutig sein. Dasselbe Zeichen darf nicht für mehrere Buchstaben verwendet werden.
- Das Alphabet des Chiffrats wird also mehr Zeichen enthalten als das Klartextalphabet.
- Die Häufigkeiten der Zeichen im Codealphabet soll möglichst ausgeglichen sein.

Aufgabe 3
Versuche jede der vier Regeln zu begründen.

3.3. Gleichmässige Verteilung

Der wichtigste Punkt bei der homophonen Chiffre ist die Verteilung der Chiffratszeichen. Die soll möglichst gleichmässig sein, damit die statistische Analyse aus Kapitel 2 ins Leere läuft. In diesem Abschnitt wollen wir dieses Ziel erreichen. Wir kennen die Buchstabenhäufigkeiten der deutschen Sprache (Tabelle im letzten Kapitel) und wollen ein Codealphabet finden, dass aus den Zeichen 00 bis 99 besteht, also aus 100 verschiedenen Zeichen. Die Frage ist nun, wie viele Zeichen jedem Buchstaben zugeordnet werden, damit alle Zeichen dann gleich häufig vorkommen.

Die gleichmässige Verteilung wird erreicht, indem jedem Buchstaben möglichst so viele Zeichen zugeordnet werden, wie es seinem Anteil in der Sprache entspricht. Da wir gerade 100 Zeichen haben, so ist die Anzahl Zeichen für einen Buchstaben ungefähr die Anzahl Prozent, mit der dieser Buchstabe gebraucht wird.

Aufgabe 4
Welchem Buchstaben werden am meisten Zeichen zugeordnet? Wie viele ungefähr? Und welchem am wenigsten? Wie viele aber mindestens?

Aufgabe 5
Versuche nun, eine homophone Verschlüsselung der 26 Buchstaben auf die 100 Zeichen 00 bis 99 zu kreieren.

Aufgabe 6
Was ist der Schlüssel bei der homophonen Chiffrierung? Findest du das praktisch oder eher kompliziert?

3.4. Schwachstellen

Die Frage ist nun natürlich, ob wir nun die perfekte Verschlüsselung gefunden haben. Sicher ist, dass die Analyse aus dem vorhergehenden Kapitel nicht mehr funktioniert, weil die Buchstaben nun praktisch keine Häufigkeitsunterschiede mehr aufweisen.

Aufgabe 7
Was ist mit Doppelbuchstaben? (also ss, ff, mm, pp, tt etc.) Sind die im Chiffrat noch erkennbar?

Ein Trick den man aber anwenden kann ist folgender: Man schaut, welche Zeichen vor oder nach einem bestimmten Zeichen stehen können. Das kann allenfalls dazu führen, gewisse Zeichen zu entschlüsseln. Der Grund dafür ist, dass manche Buchstaben von praktisch allen anderen Buchstaben gefolgt werden können, während manche nur eine sehr begrenzte Anzahl Folgebuchstaben erlauben.

Aufgabe 8
Finde einen Buchstaben, der im Deutschen nur von genau zwei verschiedenen anderen Buchstaben gefolgt werden kann. Es gibt sogar einen Buchstaben (der allerdings sehr selten ist), der immer vom selben Buchstaben gefolgt wird.

Konkret geht man nun so vor, dass man für jedes Zeichen eine Liste macht, für die Zeichen, die unmittelbar darauf folgen (bzw. unmittelbar davor stehen). Die meisten Zeichen werden nun von sehr vielen oder praktisch allen möglichen Zeichen gefolgt, Manche aber nur von sehr wenigen. Das sind nun Kandidaten für Buchstaben, wie du sie in Aufgabe 8 entdeckt hast.

Aufgabe 9
Auch wenn es nun also doch Schwachstellen gibt in homophonen Chiffren, inwiefern sind sie trotzdem sicherer geworden? Oder welche Art von Nachrichten hat die grösste Chance, unentschlüsselbar zu sein?

3.5. Die Vigenère-Chiffre

Als letztes wollen wir nun noch die Vigenère-Chiffre betrachten. Sie wurde im 16. Jahrhundert von einem französischen Diplomaten (namens Blaise de Vigenère, daher die Bezeichnung) erfunden und galt lange Zeit als ‚la chiffre indéchiffrable’, also als unentschlüsselbar. Obschon es mittlerweile Techniken gibt, dies Chiffre trotzdem zu knacken, so ist das Prinzip noch Grundlage vieler moderner Verschlüsselungsalgorithmen.

Die Grundidee ist die, dass man jedes Zeichen mit einer anderen Verschiebechiffre verschlüsselt. Man benötigt dazu ein Schlüsselwort (das wie der Name schon andeutet, der eigentliche Schlüssel ist), sowie das so genannte Vigenère-Quadrat, dass allgemein bekannt sein kann. Es sieht so aus:

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

Unter dem Klartext werden also alle 26 möglichen Verschiebechiffren hingeschrieben. Die konkrete Verschlüsselung mittels eines Schlüsselwortes soll nun anhand eines Beispiels gezeigt werden. Wir wollen ‚emmentaler’ verschlüsseln, und zwar mit dem Schlüsselwort SCHWEIZ, dazu schreiben wir über jeden Buchstaben der Nachricht einen Buchstaben des Schlüsselworts (und beginnen wenn nötig wieder vorne). Also so:

```
Schlüssel:  S C H W E I Z S C H
Klartext:   e m m e n t a l e r
```

Das erste e wird nun „mit dem S verschlüsselt“. Das bedautet, dass wir diejenige Zeile des Vigenère-Quadrats verwenden, die mir S beginnt. Diese Zeile verwenden wir nun wie eine gewöhnliche Caesarchiffre. In der Spalte unter dem e finden wir das W. Also wird das erste e mit W verschlüsselt. Für das erste m verwenden wir die Zeile, die mit C beginnt, also wird m zu O verschlüsselt. Am Ende sieht das dann so aus:

```
Schlüssel:  S C H W E I Z S C H
Klartext:   e m m e n t a l e r
Chiffrat:   W O T A R B Z D G Y
```

Aufgabe 10
Alles klar? Dann verschlüssle du nun das Wort ‚strandpromenade’ mit dem Schlüsselwort PALME.

In unserem Beispielwort kamen 3 ‚e’ vor. Die wurden zu W, Z und G verschlüsselt. Also polyalphabetisch, wie das gewünscht wurde. Die Vigenère-Chiffre sorgt also dafür, dass jeder Buchstabe mit gleicher Wahrscheinlichkeit zu mehreren anderen Buchstaben verschlüsselt werden kann.

Aufgabe 11
In dem gewählten Beispiel mit Schlüsselwort SCHWEIZ, zu wie vielen anderen Buchstaben kann jeder Buchstabe verschlüsselt werden?

3.6. Indéchiffrable?

Es scheint also auf den ersten Blick, dass diese Chiffre tatsächlich unknackbar ist. Auf einen zweiten Blick allerdings offenbaren sich auch hier einige Schwachstellen. Das sind Regelmässigkeiten, die ausgenutzt werden können, um mehr über den Klartext oder das Schlüsselwort zu erfahren.

Zum Beispiel wird das Schlüsselwort immer wiederholt. Das bedeutet, dass die Verschlüsselung auch eine gewisse Regelmässigkeit aufweisen wird. Für Wörter, die häufig vorkommen, so wie ‚ein’ oder ‚der’, besteht eine grosse Chance, dass sie mehrmals mit genau demselben Teil des Schlüsselwortes verschlüsselt werden. Falls man im Chiffrat also Folgen von Zeichen findet, die an mehreren Stellen vorkommen, so könnte dies auf einen solchen Fall hindeuten.

Aufgabe 12
In einem Chiffrat findet man die Zeichenfolge EDB an den Stellen 13, 28 und 43, und die Folge EVR an den Stellen 1, 36 und 56. Wie kann man daraus eine

Vermutung ableiten, wie lange das Schlüsselwort ist? Beachte, in welchen Abständen sich die Zeichenfolgen wiederholen.

Aufgabe 13
Dank einem anderen Hinweis vermuten wir, dass EVR tatsächlich für ‚der’ und EDB für ‚ein’ steht. Wie kannst du nun mittels dieser Annahme und dem Ergebnis von Aufgabe 12 das Schlüsselwort bestimmen?

Wenn das Schlüsselwort einmal gefunden ist, dann ist das Entziffern des ganzen Texts sehr einfach, schliesslich muss genau das auch der eigentliche Empfänger der Nachricht tun.

Natürlich war dieses Beispiel konstruiert und man hat auch nicht immer gleich viel Glück, aber es zeigt doch, dass auch die unentschlüsselbar geglaubte Chiffre von Vigenère Schwachstellen besitzt. Und so wie die ‚unsinkbare’ Titanic doch auf dem Meeresgrund endete, so hat noch mancher Entdecker eines ‚unknackbaren’ Verschlüsselungssystem feststellen müssen, dass seine Geheimnisse trotzdem in falsche Hände geraten sind.

Wenn du nach der Lernkontrolle das Gefühl hast, auch dieses Kapitel verstanden zu haben, kannst du nach dem Kapiteltest fragen. Danach hast du dieses Leitprogramm erfolgreich gemeistert. Gratulation!

Lernkontrolle

1) Beschreibe in eigenen Worten die Funktionsweise einer homophonen Chiffre und erkläre, warum sie sicherer ist als monoalphabetische Algorithmen.
2) Wie sollte man das Schlüsselwort einer Vigenère-Chiffre wählen, um eine möglichst hohe Sicherheit zu erreichen? Bei welchem anderen Mechanismus (den du eventuell täglich brauchst) gelten ähnliche Regeln?
3) Wie viele Schlüssel hat die Vigenère-Chiffre ungefähr?
4) (freiwillig). Schreibe ein Programm, das einen frei gewählten Text mittels eines eingegebenen Schlüsselwortes nach dem Schema von Vigenère verschlüsselt.

Lösungen zu den Aufgaben

Lösung 1
Die statistische Häufigkeitsanalyse aus Kapitel 2 funktioniert nicht mehr. Kein Buchstabe kann mehr zuverlässig als häufigster (und somit als e) identifiziert werden.

Lösung 2
Wenn in einem Chiffrat zum Beispiel ein B vorkommt, so kann das sowohl e als auch s bedeuten. Ähnliches gilt für H, W, N, T, D, L, H oder A. Man kann ein Chiffrat also nicht mehr eindeutig entschlüsseln.

Lösung 3
- damit werden die Häufigkeiten der Zeichen im Chiffrat verändert und im Idealfall sogar ausgeglichen.
- Siehe Aufgabe 2
- Eine direkte Folge von den ersten zwei Punkten
- Damit wird die Kryptoanalyse am schwierigsten. Siehe Aufgabe 1.

Lösung 4
Das ‚e' ist mit 17,4% mit Abstand der häufigste Buchstabe im Deutschen. Von 100 Zeichen werden wohl etwa 17 dem ‚e' zugeordnet. Die Buchstaben k, z ‚p ‚v, j, y, x und q kommen alle mit weniger als 1,5% vor und werden somit mit nur einem Zeichen codiert. Auch das q, das nur mit 0,03% Wahrscheinlichkeit vorkommt muss mit mindestens einem Zeichen codiert werden.

Lösung 5
Dieses Beispiel stammt aus dem Buch ‚Kryptologie' von Albrecht Beutelspacher. Die Zuordnung der Zeichen auf die Buchstaben ist natürlich willkürlich. Die Anzahl Zeichen pro Buchstabe sollte aber etwa mit dieser Lösung übereinstimmen.

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

Lösung 6
Der Schlüssel ist eine ganze Tabelle, zum Beispiel die von Aufgabe 5. Das ist eher viel Information. Der Schlüssel sollte aber möglichst kompakt sein, damit er einfach und unauffällig übermittelt werden kann.

Lösung 7
Nein, man kann (und soll) verschiedene Zeichen wählen, um aufeinander folgende doppelte Buchstaben zu verschlüsseln.

Lösung 8
Das c kann nur von h und k gefolgt werden. Das q sogar nur von u.

Lösung 9
Ein Chiffrat muss nun schon viel länger sein, um statistisch auswertbare Daten zu liefern. Wenn man zum Beispiel Statistiken über die Nachfolgezeichen machen will, so muss der Text genügend lang sein, dass jedes Zeichen oft (z. B.100 mal) vorkommt. Also sollte der Text mehrer Tausend Zeichen lang sein. Kurze Nachrichten sind auf diese Weise nicht mehr knackbar.

Lösung 10
```
Schlüssel:  P A L M E P A L M E P A L M E
Klartext:   s t r a n d p r o m e n a d e
Chiffrat:   H T C M R S P C A Q T N L P I
```

Lösung 11
Jeder Buchstabe wird mit einem der Buchstaben S, C, H, W, E, I oder Z verschlüsselt, also 7 Möglichkeiten.

Lösung 12
EDB wiederholt sich in Abständen von 15 Buchstaben. EVR nach 35 bzw. 20 Buchstaben. Die einzige Zahl (ausser 1), die solche Abstände bewirken könnte ist 5. Die Chancen sind also gross, dass das Schlüsselwort 5 Buchstaben lang ist.

Lösung 13
EVR kommt an Stelle eins vor, wird also mit den ersten drei Buchstaben des Schlüsselwortes verschlüsselt. Man sieht im Vigenère-Quadrat, dass das d zu E verschlüsselt wird, indem man Zeile B wählt. Also

```
Schlüssel:  B R A ? ?
Klartext:   d e r
Chiffrat:   E V R
```

EDB hingegen wird von den Buchstaben 3 bis 5 verschlüsselt (weil es zum Beispiel an Stelle 13 im text vorkommt.) Mit dem gleichen Vorgehen erhält man:

```
Schlüssel:  ? ? A V O
Klartext:       e i n
Chiffrat:       E D B
```

Man erkennt erstens, dass der mittlere Buchstabe in beiden Fällen ein A ist, was eine Bestätigung der Vermutung ist. Ausserdem kann man das Schlüsselwort zu BRAVO zusammensetzten, was ein sinnvolles Wort und somit wahrscheinlich richtig ist.

Lernkontrolle

1) Für einen Buchstaben stehen mehrer Zeichen zur Verfügung, von denen eines zufällig ausgewählt werden kann. Dadurch wird die Zeichenhäufigkeit im Chiffrat gleichmässiger und die statistische Analyse, die bei monoalphabetischen Systemen gut funktioniert, ist nicht mehr anwendbar.
2) Je länger das Schlüsselwort ist, desto kleiner wird die Chance, dass sich ein Wort oder ein Wortteil wiederholt. Auch sollte man nicht Schlüsselwörter verwenden, die leicht erratbar sind. Beides Bedingungen, die auch für die Wahl von Passwörtern gelten.
3) Jedes Wort kann ein Schlüssel sein. Weil das Schlüsselwort nicht auf eine Sprache beschränkt ist und auch Wortkombinationen möglich sind, kommen mehrere zehntausend Schlüssel in Frage. Wenn man als Schlüssel eine beliebige Buchstabenkombination nehmen will, so hat man eine praktisch unbeschränkte Anzahl Schlüssel. Es wird dann aber schwieriger, sich den Schlüssel zu merken!

Anhang A – Weiterführende Literatur

Vieles in diesem Leitprogramm ist inspiriert von folgendem Buch, das eine einfach lesbare und unterhaltsame Einführung in die Welt der Kryptographie ist:

Albert Beutelspacher, *Kryptologie – Eine Einführung in die Wissenschaft vom Verschlüsseln, Verbergen und Verheimlichen*, Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig/Wiesbaden, 2002.

Auch empfehlenswert und unterhaltsam ist dieses Buch, das eine Geschichte der Kryptographie von Caesar bis heute beinhaltet:

Simon Singh, *Geheime Botschaften*, Carl Hanser Verlag, 2000.

Die Faszination des Codeknackens kommt in diesem spannenden Roman gut zur Geltung. Er handelt von den englischen Wissenschaftlern, die im Zweiten Weltkrieg versuchten, die deutsche Verschlüsselungsmaschine ‚Enigma’ zu knacken.

Robert Harris, *Enigma*, Heyne Verlag, 1996.

Anhang B - Wörterbuch

Chiffrat: eine verschlüsselte Botschaft

Chiffre: ein →Kryptosystem

entschlüsseln: ein →Chiffrat mittels eines →Schlüssels wieder in einen →Klartext umzuwandeln.

homophone Chiffre: eine →polyalphabetische Verschlüsselung, wo jeder Klartextbuchstabe zu einer fixen Menge von möglichen Chiffratsbuchstaben verschlüsselt werden kann.

Klartext: ein allgemein lesbare Botschaft (also nicht verschlüsselt).

Kryptoanalyse: Der Vorgang, wenn man versucht, einen verschlüsselten Text zu →entschlüsseln, ohne dass man den →Schlüssel (oder sogar das verwendete →Kryptosystem) kennt.

Kryptographie: die Wissenschaft vom →Verschlüsseln und →Entschlüsseln von Botschaften.

Kryptologie: siehe →Kryptographie.

Kryptosystem: ein Verfahren (auch Algorithmus genannt) zum →Ver- und →Entschlüsseln von Botschaften.

monoalphabetische Chiffre: jedem Buchstaben vom →Klartext wird immer derselbe Buchstabe im →Chiffrat zugeordnet.

polyalphabetische Verschlüsselung: im Gegensatz zur →monoalphabetischen Kryptographie, wird ein Klartextbuchstabe nicht immer zum selben Chiffratsbuchstaben verschlüsselt.

Schlüssel: ein Geheimnis zwischen dem Sender und dem Empfänger einer Botschaft, das das ver- und entschlüsseln einer Botschaft ermöglicht.

verschlüsseln: einen →Klartext mittels eines →Schlüssels in ein nur für den Empfänger bestimmtes →Chiffrat umwandeln.

Anhang C – Kapiteltests

Kapitel 1

a) Beschreibe die Funktionsweise der Verschiebechiffre von Caesar. Definiere klar, was der Schlüssel ist, wie ver- und wie entschlüsselt wird.
b) Das Prinzip von Kerkoff sagt, dass man annehmen muss, dass nur der Schlüssel geheim ist, der Algorithmus selber aber allgemein bekannt sei. Nenne drei Gründe, warum diese Annahme sinnvoll ist.
c) Entschlüssle folgenden Satz, der mit einer allgemeinen Caesarchiffre verschlüsselt wurde: PUQ EUQNQZ TGQSQX HAZ DAY.

Hinweise zu den Lösungen

a) wurde im Kapitel genügend besprochen, bei b) sind folgende Gründe möglich:

- man sollte nicht darauf vertrauen, dass der Gegner das System nicht kennt. Da kann man sich nämlich arg verschätzen.
- ein öffentlich bekannter und von Experten analysierter Algorithmus ist sicherer als etwas selbst gebasteltes, das möglicherweise Schwachstellen aufweist.
- in einem ganzen Netzwerk (zum Beispiel in einer Firma), kann man immer denselben Algorithmus verwenden und muss nur individuelle Schlüssel generieren.

c) Die Sieben Huegel von Rom.

Kapitel 2

a) Beschreibe die monoalphabetische Verschlüsselung. Definiere klar, was der Schlüssel ist, wie ver- und wie entschlüsselt wird.
b) Was ist der Grund, dass auch monoalphabetisch verschlüsselte Texte nicht sicher sind?
c) Wie viele Schlüssel muss ein System haben, damit 1000 Computer, die je 1 Million Schlüssel pro Sekunde testen können, 10 Jahre brauchen, um das System zu knacken?

Hinweise zu den Lösungen

a) und b) wurden im Kapitel besprochen. Bei c) rechnet man 10 Jahre mal 1000 Computer mal eine Million Schlüssel/Sekunde und erhält $3{,}2*10^{17}$. Das entspricht einer Schlüssellänge von 58 Bits.

Kapitel 3

Die polyalphabetischen Algorithmen teilen sich in zwei Kategorien auf: homophone Chiffren und solche (wie die Vigenère-Chiffre), die die Verschlüsselung für jeden Buchstaben ändern.
Beschreibe für beide Verfahren, wie sie Funktionieren und was die jeweiligen Gründe sind, dass auch diese Verfahren nicht absolute Sicherheit garantieren.

Hinweis zur Lösung

Homophone Verschlüsselung: Jeder Buchstabe kann mit einem oder mehreren Zeichen verschlüsselt werden, das zufällig ausgewählt wird. Eine Schwachstelle dieses Verfahren sind die Häufigkeiten für die Folgezeichen jedes Zeichens, weil aufeinander folgende Buchstaben unterschiedliche Wahrscheinlichkeiten haben, die durch diese Chiffre nicht ausgeglichen werden.
Vigenère Chiffre: Mittels eines Schlüsselwortes wird für jede Position des Textes bestimmt, welche Verschiebechiffre verwendet wird. Durch das Vorhandensein von wiederholten Wortbestandteilen im Chiffrat können die Länge des Schlüsselwortes und damit eventuell auch das Schlüsselwort selber bestimmt werden.