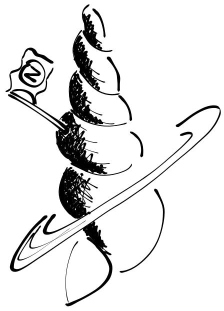
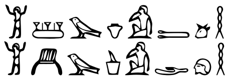
Schülerzirkel Mathematik
Fakultät für Mathematik. Universität Regensburg

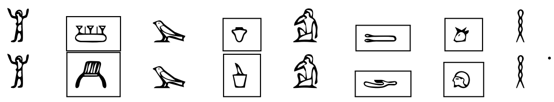


Planet Nuschel

Aufgabe 1 (Hieroglyphen* (nur für die Klassen 7/8) [4 Punkte]). Der Hamming-Abstand ist nicht nur auf Buchstaben beschränkt. Man kann diesen auch auf beliebige Zeichenfolgen der gleichen Länge anwenden. So ist der Hamming-Abstand von `1234` und `1226` zum Beispiel 2, da die letzten beiden Stellen verschieden sind. Was ist der Hamming-Abstand der beiden Zeichenfolgen:

 ?

Lösung. Der Hamming-Abstand der beiden Zeichenfolgen ist 4. Die verschiedenen Zeichen sind:



Aufgabe 2 (Hamming-Abstand 1* (nur für die Klassen 7/8) [4 Punkte]). Gib zwei deutsche Sätze an, die verschiedenen Inhalt haben, aber wo in jedem Wort maximal ein Buchstabe anders ist. Zum Beispiel: „Die Maus ist tot.“ und „Die Laus ist rot.“

Lösung. Hier gibt es natürlich sehr viele Lösungen. Eine Lösung für die volle Punktzahl ist zum Beispiel:

Die Hütte war sein Zuhause.
Sie hatte gar kein Zuhause.

Aufgabe 3 (Repetitionscode * (empfohlen ab Klasse 9) [4 Punkte]). Professor Hamming hat eine Idee, um trotz der Fehler bei der Übertragung die korrekten Wörter besser zu entschlüsseln. Er lässt jeden Buchstaben doppelt schicken. Dann empfängt er statt dem Wort `HALLO` nun die Zeichenfolge `HHAALLLLOO`. Wie verändert sich der Hamming-Abstand von zwei Wörtern, wenn man jeden Buchstaben verdoppelt?

Lösung. Der Hamming-Abstand verdoppelt sich. Denn durch die Verdoppelung aller Zeichen erscheint jedes unterschiedliche Zeichen auch zweimal.

Aufgabe 4 (Fehler erkennen und korrigieren** (empfohlen ab Klasse 9) [2+2 Punkte]). Wie in Aufgabe 3 lässt Professor Hamming die Nuschelwesen jedes Zeichen doppelt senden. Wenn er nun die Zeichenfolge `BBADLLLL` erhält, weiß er, dass es bei der Übertragung einen Fehler gab. Er kann die Zeichenfolge nämlich in Zweierblöcke zerlegen: `BB`, `AD`, `LL`, `LL`. Bei einer fehlerfreien Übertragung sind die Buchstaben in den Zweierblöcken gleich.

1. Angenommen pro übertragenem Wort gibt es maximal zwei Fehler. Kann Professor Hamming dann immer erkennen, ob es einen Fehler in der Übertragung gab? Begründe deine Antwort.

2. Professor Hamming lässt nun jeden Buchstaben fünfmal schicken. Angenommen pro übertragenem Wort gibt es immer noch nur maximal zwei Fehler. Zeige, dass Professor Hamming jeden Fehler erkennen und sogar korrigieren kann.

Lösung. 1) Professor Hamming kann nicht immer erkennen, ob es einen Fehler in der Übertragung gab. Wenn die Nuschelwesen zum Beispiel das Wort „Maus“ senden wollen, kann es bei zwei Fehlern in der Übertragung passieren, dass Hamming die Zeichenfolge `LL AA UU SS` erhählt. Diese Nachricht sieht für Hamming fehlerfrei aus, auch wenn es ein anderes Wort ist, das er erhalten hat.

2) Nach Vorraussetzung enthält jedes übertragende Wort maximal zwei Fehler. Wir starten unsere Überlegungen damit, dass beide Fehler im selben Fünferblock sind. Nehmen wir also an das Wort „Maus“ soll übertragen werden und im Fünferblock zu dem Buchstaben `A` sind die zwei Fehler. Professor Hamming erhält, dann fünf Zeichen. Da zwei Fehler darunter sind, wird er aber in jedem Fall dreimal das Zeichen `A` erhalten. Das liefert nun die allgemeine Strategie: Professor Hamming wählt aus jedem Fünferblock immer das Zeichen, dass sich mindestens dreimal wiederholt! Dies ist immer Möglich, da nach Vorraussetzung maximal zwei falsch sind. Mit dieser Strategie wird er immer die korrekte Nachricht erhalten.

Aufgabe 5 (ISBN** [1+1+2 Punkte]). Die meisten Bücher, die man kaufen kann, haben eine internationale Standardbuchnummer (kurz ISBN). In dieser Aufgabe arbeiten wir mit dem alten Standard ISBN-10. Eine ISB-Nummer besteht aus 10 Ziffern. Die ersten neun Ziffern können eine Zahl zwischen 0–9 sein. Die letzte Ziffer kann zusätzlich ein X sein, das für die Zahl 10 steht. Nun ist die Zeichenfolge

$$z_1 – z_2 z_3 z_4 z_5 z_6 – z_7 z_8 z_9 – z_{10}$$

ein gültiger ISBN-Code wenn die Zahl

$$1 \cdot z_1 + 2 \cdot z_2 + 3 \cdot z_3 + 4 \cdot z_4 + \ldots + 9 \cdot z_9 + 10 \cdot z_{10}$$

durch 11 teilbar ist.

1. Ist die folgende Zeichenfolge eine ISB-Nummer: 3–66248–955–4?
2. Welchen Wert muss $\boxed{?}$ haben, damit 1–54241–$\boxed{\mathbf{?}}$82–8 eine ISB-Nummer ist?
3. Was ist der minimale Hamming-Abstand, den zwei ISB-Nummern haben können?

Lösung. 1) Wir rechnen nach:

$$\begin{aligned}
&1 \cdot 3 + 2 \cdot 6 + 3 \cdot 6 + 4 \cdot 2 + 5 \cdot 4 + 6 \cdot 8 + 7 \cdot 9 + 8 \cdot 5 + 9 \cdot 5 + 10 \cdot 4 \\
&= 3 + 12 + 18 + 8 + 20 + 48 + 63 + 40 + 45 + 40 \\
&= 297 = 27 \cdot 11.
\end{aligned}$$

Damit ist 3–66248–955–4 eine gültige ISB-Nummer.

2) Wieder rechnen wir erstmal.

$$\begin{aligned}&1 \cdot 1 + 2 \cdot 5 + 3 \cdot 4 + 4 \cdot 2 + 5 \cdot 4 + 6 \cdot 1 + 7 \cdot x + 8 \cdot 8 + 9 \cdot 2 + 10 \cdot 8\\ &= 1 + 10 + 12 + 8 + 20 + 6 + 7 \cdot x + 64 + 18 + 80\\ &= 219 + 7 \cdot x.\end{aligned}$$

Wir suchen also eine Zahl x, so dass $219 + 7 \cdot x$ durch 11 teilbar ist. Wir können nun nacheinander die Zahlen $0, 1, \ldots, 9$ einsetzen und werden feststellen, dass $x = 8$ die gesuchte Zahl ist.

3) Der minimale Hamming-Abstand, den zwei ISB-Nummern haben können, ist 2. Zunächst zeigen wir, dass es zwei gültige ISB-Nummern gibt, die den Hamming-Abstand 2 haben. Aus Teil 1) wissen wir, dass 3–66248–955–4 eine gültige ISB-Nummer ist. Ohne viel Mühe rechnet man nach, dass dann auch 3–$\boxed{94}$248–955–4 eine gültige ISB-Nummer ist. Denn wir haben $2 \cdot 6 + 3 \cdot 6 = 2 \cdot 9 + 3 \cdot 4$.

Wir werden nun zeigen, dass der Hamming-Abstand nicht 1 sein kann. Das machen wir mit einem Widerspruchsbeweis. Im Beweis werden wir sehen, dass es entscheidend ist, dass 11 eine Primzahl ist. Genauer genommen benutzen wir folgende Eigenschaft von Primzahlen. Es sei p eine Primzahl und a, b beliebige ganze Zahlen mit $p \mid a \cdot b$, dann gilt schon $p \mid a$ oder $p \mid b$. Der Strich steht für „teilt". Wenn eine Primzahl ein Produkt teilt, dann teilt es schon einen der Faktoren.

Kommen wir nun zurück zum eigentlichen Beweis. Wir nehmen an, es gäbe zwei gültige ISB-Nummern:

$$z_1–z_2z_3z_4z_5z_6–z_7z_8z_9–z_{10}$$

und

$$x_1–x_2x_3x_4x_5x_6–x_7x_8x_9–x_{10}$$

mit Hamming-Abstand 1. Hamming-Abstand 1 bedeutet, dass es ein $1 \leq i \leq 10$ gibt, so dass $x_i \neq z_i$, aber $x_j = z_j$ für alle $j \neq i$.

Da beide Folgen gültige ISB-Nummern sind, wissen wir, dass die Summen

$$1 \cdot z_1 + 2 \cdot z_2 + 3 \cdot z_3 + 4 \cdot z_4 + \ldots + 9 \cdot z_9 + 10 \cdot z_{10}$$

und

$$1 \cdot x_1 + 2 \cdot x_2 + 3 \cdot x_3 + 4 \cdot x_4 + \ldots + 9 \cdot x_9 + 10 \cdot x_{10}$$

jeweils durch 11 teilbar sind.

Wenn zwei Zahlen durch 11 teilbar sind, dann ist auch deren Differenz durch 11 teilbar. Da x_i ungleich z_i ist, wird eine der beiden Zahlen größer sein. Wir nehmen nun an, dass z_i größer als x_i ist, und ziehen die zweite Summe von der ersten Summe ab. Weil $z_j = x_j$ für alle $j \neq i$ erhalten wir:

$$1 \cdot (z_1 - x_1) + 2 \cdot (z_2 - x_2) + \ldots + 9 \cdot (z_9 - x_9) + 10 \cdot (z_{10} - x_{10}) = i \cdot (z_i - x_i).$$

Als Differenz von zwei durch 11 teilbaren Zahlen ist der Ausdruck $i \cdot (z_i - x_i)$ ebenfalls durch 11 teilbar. Da 11 eine Primzahl ist, teilt 11 schon i oder $(z_i - x_i)$. Beides ist nicht möglich, da $0 < i < 11$ und $0 < z_i - x_i < 11$ gilt. Dies ist also ein Widerspruch und wir sehen, dass der Hamming-Abstand von zwei ISB-Nummern mindestens 2 ist.

Aufgabe 6 (Dreiecksungleichung*** (empfohlen ab Klasse 9) [4 Punkte]). Es seien $X = x_1 x_2 \dots x_n$, $Y = y_1 y_2 \dots y_n$ und $Z = z_1 z_2 \dots z_n$ Wörter der Länge n. Zeige, dass der Hamming-Abstand die sogenannte Dreiecksungleichung erfüllt, d.h. der Hamming-Abstand von X zu Z ist kleiner gleich als die Summe der Hamming-Abstände von X zu Y und von Y zu Z.

Zum Beispiel ist der Hamming-Abstand von `FRISST` zu `KRISQT` und der Hamming-Abstand von `KRISQT` zu `KRIEGT` jeweils 2. Der Hamming-Abstand von `FRISST` zu `KRIEGT` ist 3. Offensichtlich gilt $3 \leq 2 + 2$.

Lösung. Den Hamming-Abstand zwischen zwei Zeichenfolgen kann man als die minimale Anzahl der Zeichen auffassen, die man ändern muss, um von einer Zeichenfolge zur anderen zu gelangen.

Wenn der Hamming-Abstand von X zu Y gleich k ist, dann muss man k Zeichen ändern, um von X nach Y zu gelangen. Wenn der Hamming-Abstand von Y zu Z gleich m ist, dann muss man m Zeichen ändern, um von Y zu Z zu gelangen.

Wir können also durch $k + m$ viele Änderungen von X nach Z gelangen, in dem wir erst X zu Y in k vielen Schritten abändern und dann von Y zu Z in m vielen Schritten gelangen.

Damit ist $k + m$ eine obere Schranke für den Hamming-Abstand von X zu Z.

Exkurs: Kodierungstheorie

Im Alltag werden ständig Informationen übertragen. Das Smartphone überträgt im Mobilfunknetz, der Computer über die DSL-Leitung und der Fernseher empfängt sein Signal beispielsweise von einem Satelliten. Bei all diesen Übertragungen passieren Fehler. Vielleicht kennt ihr es von eurem Fernseher, dass die Bilder stocken, wenn es draußen stürmt.

Das Grundproblem der Kodierungstheorie ist es, einen Code zu entwerfen, bei dem der Hamming-Abstand zwischen den übertragenen Zeichenfolgen sehr groß ist, wobei die Datenmenge, die insgesamt übertragen wird, aber möglichst klein bleiben soll.

Vergleichen wir dazu mal den ISBN-Code mit dem Repetitionscode. Der ISBN-Code besteht im Wesentlichen aus neun Ziffern plus einer Kontrollzahl. Wir übermitteln also für eine ISBN zehn Zeichen. Würden wir hier einen Repetitionscode verwenden also jede Ziffer doppelt senden anstatt der Kontrollziffer, dann würden wir pro ISBN 18 Zeichen übertragen. Wir sparen uns also acht Zeichen ein bei ähnlich gutem Erkennen von Fehlern.

Wir enden diesen Exkurs mit einem Beispiel, dass ein optimaler Code sehr von der Situation abhängt. Vergleichen wir dazu mal das Beispiel des Fernsehers und den Abruf einer Internetseite.

Der Fernseher kann nur empfangen und selber nicht senden. Damit ist es sehr wichtig, dass, wenn Fehler auftreten, er diese nicht nur erkennen kann, sondern auch direkt korrigieren kann, damit er das richtige Bild darstellt. Hingegen reicht es beim Abruf einer Internetseite für den Computer aus einen Fehler zu erkennen. Denn wenn er einen Fehler erkannt hat, dann kann er die fehlerhaften Dateien erneut anfragen.

Weiterführende Links

https://de.wikipedia.org/wiki/Fehlerkorrekturverfahren