

Beim Rechnen mit Kommazahlen am Computer können stets Fehler passieren!

auch in Excel etc.

Warum?

- Rundungsfehler: Wenn man nur auf 4 Nachkommastellen genau rechnet und etwa $\frac{1}{3} = 0.3333\dots = 0.\bar{3}$ auf 0.3333 rundet, so gilt exakt

$$0.3333 + 0.3333 + 0.3333 = 0.9999 \neq 1$$

aber genau natürlich $\frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$.

Fazit: Die Summe der Rundungen ist nicht die Rundung der Summe!

Dasselbe kann bei anderen Rechenoperationen, also Multiplikation, Subtraktion oder Division passieren. Das passiert genauso in anderen Stellenwertsystemen, etwa im Binärsystem, weshalb der Computer Rundungsfehler macht.

- subtile Rundungsfehler: Der Befehl `print((0.3-0.1) == 0.2)` in Python liefert als Ausgabe **FALSE!** Dies ist im Zehnersystem nicht mit Rundungsfehlern zu erklären.

Aber: Im Binärsystem, als binäre Kommazahl, sind 0.1 und 0.2 und 0.3 periodisch und brechen nie ab! Der Computer rechnet mit gerundeten binären Kommazahlen und macht deswegen manchmal Rundungsfehler, auch wenn man diese vom Dezimalsystem her keineswegs erwartet!

Mit den Befehlen

```
print("%.20f" % (0.3-0.1))
print("%.20f" % 0.2)
```

kann man sehen, dass 0.3-0.1 und 0.2 in Python verschieden sind! Die Ausgabe ist

```
0.19999999999999998335
0.20000000000000001110
```

1. CODIERUNG VON TEXTEN

1.1. ASCII- und Unicode-Codierung. Erklärungen im Video, vgl. auch [Wikipedia: ASCII](#).

Zu Unicode und UTF-8 (Unicode Transformation Format 8bit): Unicode ist wie ASCII eine Zuordnung zwischen gewissen Zeichen und Zahlen. Diese Zahlen schreibt man meist hexadezimal mit U+ als Präfix und nennt so etwas *Unicode code point*.

Zum Beispiel ist dem Eurozeichen € der Unicode code point U+20AC zugeordnet. Diesem ordnet die Kodierung UTF-8 die Folge 11100010 10000010 10101100 aus drei Bytes oder hexadezimal E2 82 AC zu (Beispiel aus [Wikipedia UTF-8; Examples](#)):

Zeichen: € $\xleftarrow{\text{Unicode}}$ Unicode code point: U+20AC $\xleftarrow{\text{UTF-8}}$ E2 82 AC

Warum speichert der Computer nicht direkt den Unicode code point? Wenn ich es recht verstehe, bräuchte er dann pro Zeichen 3 Bytes; mit weiterer Umwandlung per UTF-8 geht es oft sehr viel platzsparender, insbesondere bei englischen oder westlichen Texten, wo die meisten Zeichen mit einem oder zwei Byte codiert werden.

Aufgabe 1.1.1. (a) (ähnlich wie im Video) Was ergeben die folgenden hexadezimalen Zahlen, wenn man sie per ASCII (= American Standard Code for Information Interchange) als Zeichen interpretiert? Verwende eine ASCII-Tabelle aus dem Internet!

46 61 6c 73 63 68 3f

(b) Was ergeben die folgenden Dezimalzahlen, wenn man sie mit ASCII in Zeichen entschlüsselt?

71 117 116 32 103 101 109 97 99 104 116 33

(c) Für welche Zeichen stehen die folgenden Unicode code points?

U+0061 U+0430 U+1f99c U+1fbf9 U+1f939 U+1f3bc U+4e23

2. CODIERUNG VON BILDERN

Die Teile über RGB und Pixelgrafik-Teil folgen stark einer Vorlage von Barbara Pampel.

2.0.1. Die wesentlichen Verfahren zum Codieren von Bildern sind

- Rastergrafik (auch Pixelgrafik, *raster graphics*, *bitmap*, *pixmap*)
- Vektorgrafik (*vector graphics*)

Zuerst muss man aber wissen, wie man Farben kodiert.

2.1. RGB-Farbmodell.

2.1.1. Farben werden häufig mit dem RGB-Modell kodiert (RGB = red green blue = rot grün blau). Für jede der drei Farben (in der Reihenfolge RGB) speichert man den Farbanteil als eine Zahl zwischen 0 und 255, meist als zweistellige Hexadezimalzahl. Die Farbe ergibt sich dann durch **additive Farbmischung** (so wie sich die Lichtkegel von roten, grünen und blauen Scheinwerfern auf einer weissen Wand „addieren“).

Unterscheide dies von der **subtraktiven Farbmischung**, die etwa beim Drucken verwendet wird (CMYK-Farbmodell).

Aufgabe 2.1.2. (a) Welche Farbe kodiert 1E53DF?

(b) Gib die Codierung (in Hexadezimalziffern) eines Lila-Tons an, ohne die Farbwerte von Grün und Blau zu verändern.

(c) Wie viele Farben kann man mit dem RGB-Modell kodieren?

2.2. Rastergrafiken.

2.2.1. Wenn man ein Bild als Rastergraphik speichern will, legt man ein quadratisches oder rechteckiges Raster über das Bild, bestimmt für jedes Rechteck in diesem Raster eine Durchschnittsfarbe und speichert diese.

Aufgabe 2.2.2. Speichere den folgenden Code mit einem Text-Editor deiner Wahl (etwa mit Tigerjython) als ppm-Datei unter einem geeigneten Namen, etwa `bild.ppm`.

```
P3
# "P3" means this is a RGB color image in ASCII
5 6
255
# "255" is the maximum value for each color
# end of header, data below
255 255 255 255 255 255 255 0 0 255 255 255 255 255 255
255 255 255 255 0 0 255 0 0 255 0 0 255 255 255
255 0 0 255 0 0 255 0 0 255 0 0 255 0 0
0 0 255 0 0 255 0 0 255 0 0 255 0 0 255
0 0 255 255 255 255 0 0 255 255 255 255 0 0 255
0 0 255 255 255 255 0 0 255 0 0 255 0 0 255
```

Die Endung `ppm` steht für *portable pixmap* und bezeichnet ein sehr einfaches Grafikformat.

Öffne die Datei mit einem Bildanzeigeprogramm oder einem Bildbearbeitungsprogramm, beispielsweise [GIMP](#) (*Gnu Image Manipulation Program*).

Das Bild ist sehr klein, so dass du es stark vergrössern musst.

Schalte das Licht im Fenster auf Gelb und färbe die Tür braun oder manipulierte das Bild in einer anderen Weise.

Aufgabe 2.2.3. Schreibe eine ppm-Datei mit kleinen Abmessungen, so dass das entsprechende Bild alle 2^3 RGB-Farben anzeigt, bei denen jeder der drei RGB-Werte entweder 0 oder 255 ist, vgl. [Farbbalkentestbild](#).

Wenn du willst, kannst du auch alle 3^3 Farben mit RGB-Werten entweder 0 oder 127 oder 255 anzeigen.

2.3. Vektorgrafiken.

2.3.1. In der Vektorgrafik stellt man Bilder dar, indem man sie aus relativ einfachen Formen wie Linien, Kreisen, Polygonen, Ellipsen, Splines, Bezierkurven etc. zusammensetzt.

Ein gebräuchliches Format für Vektorgrafiken ist das SVG-Format (*scalable vector graphics*).

Aufgabe 2.3.2. Hier ist eine SVG-Datei.

```
<svg height="200" width="200">
  <polyline points="100,100 100,80
    120,80 120,120 80,120 80,60
    140,60 140,140 60,140"
    style="fill:green;stroke:black;stroke-width:2"/>
  <circle cx="80" cy="50" r="20" stroke="blue" stroke-width="5" fill="yellow" />
  <text x="30" y="130" fill="red" transform="rotate(-45 30,110)">SVG is great</text>
</svg>
```

Speichere sie etwa unter dem Namen `beispiel.svg` und zeige sie mit einem Grafikprogramm an, etwa mit [Inkscape](#).

Erstelle selbst eine SVG-Datei mit einer netten Grafik (etwa einer Kinderzeichnung). Wenn du zum Beispiel sehen willst, ob du den Befehl `polyline` verstehst: Zeichne damit das Haus vom Nikolaus.

Wenn du kreativ sein willst, lerne etwa von [W3Schools SVG-Tutorial: SVG <ellipse>](#), wie man ein Ellipse zeichnet, stöbere in diesem Tutorial und schau dir die Beispiele am Ende an.

2.4. Vergleich: Rastergrafik versus Vektorgrafik.

Aufgabe 2.4.1. (war schon im Video) Überlege dir Vor- und Nachteile von Rastergrafik versus Vektorgrafik.

Gerne darfst du im Internet stöbern, etwa auf den entsprechenden (englischen oder deutschen) Wikipedia-Seiten; dort gibt es nette Vergleiche mit Bildern.

2.4.2. Einige Links:

- [SVG-Tutorial Intro](#)

Am Ende der Seite sind einige Vorteile von SVG aufgelistet.

- [Wikipedia: Vector graphics; File formats](#)

Hier werden Standardformate für Rastergrafiken und Vektorgrafiken aufgezählt. Rechts ist auch ein nettes Bild, das Vorteile von SVG illustriert.

Aufgabe 2.4.3. (freiwillig) Wähle ein Foto in einem komprimierten Format wie beispielsweise [JPEG](#) (*Joint Photographic Experts Group*). Öffne es mit einem Bildbearbeitungsprogramm und speichere es wenn möglich als ppm-Datei (mit GIMP geht das jedenfalls).

Um welchen Faktor unterscheiden sich die Grössen der beiden Dateien?

Aufgabe 2.4.4. (freiwillig) Speichere eine SVG-Datei als ppm-Datei oder jpeg und führe einen Grössenvergleich durch.

Dieser Abschnitt ist freiwillig und vorrangig für die Schnellen/Interessierten.

3.0.1. Beim Übertragen von Daten können Fehler entstehen, etwa durch elektromagnetische Interferenz, aber auch ganz konkret durch fehlerhaftes Eintippen von Zahlen (alias Datenübertragung vom Papier in ein digitales Medium, z. B. Bankverbindung eingeben).

Eine Möglichkeit, Übertragungsfehler zu erkennen, sind sogenannte Prüfziffern.

Aufgabe 3.0.2. Das einfachste Beispiel einer Prüfziffer ist das sogenannte **Paritätsbit** (*parity bit*).

Stell dir vor, Alice möchte Bob eine aus 7 Bit bestehende Nachricht senden, etwa 1001100. Sie wählt ein weiteres Bit, das sogenannte Paritätsbit, so, dass die Summe aller 8 Bit gerade ist. Im Beispiel der Nachricht 1001100 ist das Paritätsbit 1, denn $1 + 0 + 0 + 1 + 1 + 0 + 0 + 1$ ist gerade (gleichbedeutend: $1 + 0 + 0 + 1 + 1 + 0 + 0 + 1 = 0$ im Zahlenring $\mathbb{Z}/2\mathbb{Z}$ alias der Uhr mit den beiden Uhrzeiten 0 und 1). Alice hängt an ihre 7-Bit-Nachricht dieses Paritätsbit an und sendet Bob die so erhaltenen 8 Bit über einen Kanal.

Wir nehmen an, dass bei der Übertragung maximal eines der 8 Bit fehlerhaft übertragen wird.

Bob erhält eine Folge von 8 Bit. Wie kann er feststellen, ob ein Übertragungsfehler vorliegt oder nicht? Bob kennt die ursprüngliche Nachricht natürlich nicht, sonst wäre es langweilig.

Aufgabe 3.0.3 (Prüfziffern bei ISBN). Die Internationale Standardbuchnummer ISBN ist eine dreizehnstellige Dezimalzahl $z_1 z_2 z_3 \dots z_{13}$, bestehend aus den Ziffern z_1, \dots, z_{13} . Die dreizehnte Ziffer z_{13} , die sogenannte Prüfziffer, wird so gewählt, dass

$$z_1 + z_3 + z_5 + \dots + z_{13} + 3 \cdot (z_2 + z_4 + \dots + z_{12}) = 0 \quad \text{in } \mathbb{Z}/10\mathbb{Z}$$

gilt, also auf der Uhr mit den zehn Uhrzeiten 0, 1, 2, ..., 9 (vgl. die Grafik etwas rechts oberhalb von [Wikipedia: ISBN; ISBN-13](#)).

- (a) Ermittle die Prüfziffer z_{13} für die ISBN 978382741926 z_{13} und finde per Internet heraus, um welches Buch es sich handelt.

(Ich empfehle Kapitel 14 in diesem Buch all denjenigen, die verstehen wollen, wie ein Computer funktioniert (was ich leider aus Zeitgründen nur bis zum 4-Bit-Addierer erklärt habe); dort wird bis zu einem gewissen Grad erklärt, wie der [Modellprozessor](#) funktioniert.)

- (b) Schreibe ein Programm in Python, das für eine 13-stellige Eingabe überprüft, ob es sich um eine gültige ISBN-Nummer handelt.
- (c) (freiwillig) Bei der Eingabe einer gültigen ISBN wurde eine einzige Ziffer falsch eingegeben bzw. zwei benachbarte Ziffern wurden vertauscht (das sind sehr häufige Tippfehler). Erkennt man dies an der Prüfziffer?

Was bisher erklärt wurde, heisst genauer ISBN-13 und ist das aktuelle Verfahren zur Bestimmung 13-stelligen ISBN. Bis Ende 2006 wurde ISBN-10 verwendet für 10-stellige ISBN.

- (d) Eine ISBN-10 $z_1 z_2 \dots z_{10}$ besteht aus neun Ziffern $z_1, \dots, z_9 \in \{0, 1, \dots, 9\}$ und einer Ziffer $z_{10} \in \{0, 1, 2, \dots, 9, X\}$, wobei X für 10 steht (wie bei den römischen Zahlen). Die Prüfziffer z_{10} ist so gewählt, dass

$$\underbrace{1 \cdot z_1 + 2 \cdot z_2 + \dots + 9 \cdot z_9}_{\text{„gewichtete Quersumme“}} = z_{10} \quad \text{in } \mathbb{Z}/11\mathbb{Z}$$

gilt.

Schreibe ein Python-Programm, das zu gegebenen neun Ziffern $z_1 z_2 \dots z_9$ die Prüfziffer z_{10} berechnet.

Aufgabe 3.0.4. Lies den Abschnitt Prüfsumme auf [Wikipedia: Internationale Bankkontonummer](#) (IBAN) um zu verstehen, wie in diesem Fall die Prüfsumme gebildet wird.

3.0.5. In den obigen Aufgaben wurden gewissen Zahlenfolgen Prüfziffern bzw. Prüfsummen zugeordnet. Diese Zuordnungen sind Beispiele von Hash-Funktionen, vgl. [Wikipedia: Hash-Funktion](#).

Unsere Prüfziffern sind „fehlererkennende Codes“: Wenn nicht zu viele Übertragungsfehler passieren, werden diese erkannt. Es gibt auch „fehlerkorrigierende Codes“; am bekanntesten sind wohl die Hamming-Codes, vgl. [Wikipedia: Hamming code](#).