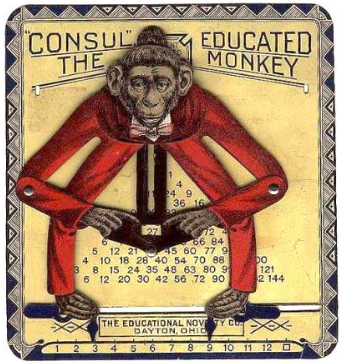
Einführung in die Programmierung



Tutorium 6

Wintersemester 2025/2026

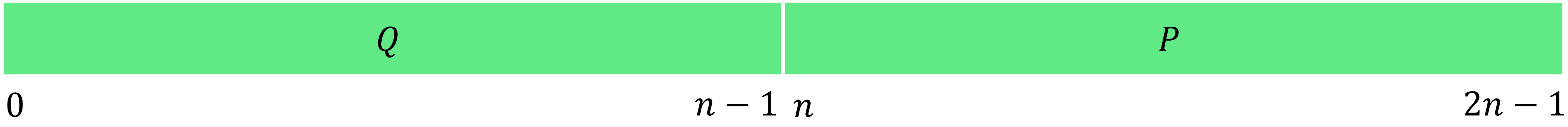
Arbeitsgruppe Systemsoftware
Angewandte Informatik 12

technische universität dortmund

Arbeitsgruppe Systemsoftware

Naive Darstellung rationaler Zahlen im Rechner:

- Zähler $q = \sum Q_i \cdot 2^i$ und Nenner $p = \sum P_i \cdot 2^i$ werden hintereinander als gleichgroße Datenworte $D = Q_0 \cdots Q_{n-1} P_0 \cdots P_{n-1}$ abgelegt
- Vor dem Ablegen sind Nenner und Zähler mithilfe ihres größten gemeinsamen Teilers (ggT) zu kürzen
- **Problem**: Rechnen in diesem Zahlensystem ist aufwendig und kann schnell zu Überläufen im Zähler oder Nenner führen



Darstellung von Zahlen im Festkommaformat (Naiv):

- Rationale Zahlen können wie gewohnt im Binärsystem mit zusätzlichem Nachkommateil dargestellt werden
- Wertigkeit des i-ten Bit im Vor- und Nachkommateil ergibt sich mit 2^{i-k}

$$z = (\text{-}1)^{\mathrm{v_z}} \cdot (V_{n-2} \cdots V_k . N_{k-1} \cdots N_0)_2 \equiv (\text{-}1)^{\mathrm{v_z}} \cdot 2^{\text{-}k} \cdot (V_{n-2} \cdots V_k N_{k-1} \cdots N_0)_2$$

$$= (\text{-}1)^{Z_{n-1}} \cdot 2^{\text{-}k} \cdot (Z_{n-2} Z_{n-3} \cdots Z_1 Z_0)_2 = (\text{-}1)^{Z_{n-1}} \frac{1}{2^k} \sum_{i=0}^{n-2} Z_i 2^i$$

$\vec{Z} \equiv$	v_z	V		N	
	$n-1$	$n-2$	k	$k-1$	0

0	0	0	0	0	1	1	0	0	.	1	1	0	0	0	0	0
1 : Negativ 0 : Positiv	128	64	32	16	8	4	2	1		0.5	0.25	0.125	⋯			

Darstellung von Zahlen im Festkommaformat (Zweierkomplement):

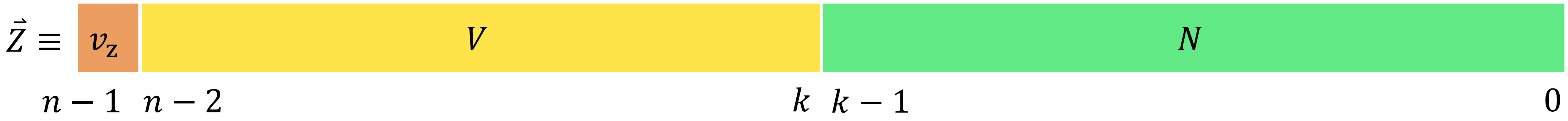
- Festkommazahlen werden in realen Systemen üblicherweise im Zweierkomplement mit zusätzlichem Nachkommateil dargestellt
- Man rechnet mit ihnen genau wie mit ganzen Zweierkomplement-Zahlen

$$z = (V_{n-2} \cdots V_k . N_{k-1} \cdots N_0)_2 - v_z 2^{n-1} \equiv 2^{-k} \cdot [(V_{n-2} \cdots V_k N_{k-1} \cdots N_0)_2 - v_z 2^{n-1}]$$

$$= 2^{-k} \cdot [(Z_{n-2} Z_{n-3} \cdots Z_1 Z_0)_2 - Z_{n-1} 2^{n-1}] = \frac{1}{2^k} \sum_{i=0}^{n-1} (-1)^{\delta_{i,n-1}} Z_i 2^i$$

Kronecker-δ

$$\delta_{i,j} \equiv \begin{cases} 1 \text{ if } i = j \\ 0 \quad \text{else.} \end{cases}$$



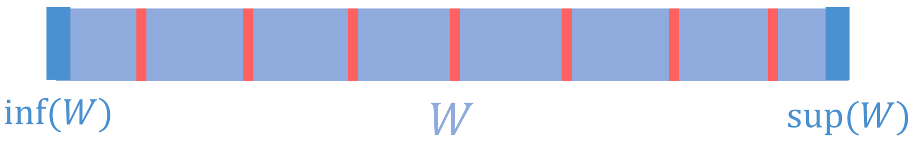
- **Fließkommazahlen** (**Gleitkommazahlen**, engl. **Floating-point numbers**) stellen rationale Zahlen durch die folgenden Bestandteile dar:
  - **Mantisse** M als normalisierte Binärzahl in gepackter, dh. impliziter Darstellung
  - **Kommaposition** als Charakteristik C in Exzess-Darstellung (**Exponent** E mit **Bias** k)
  - Vorzeichenbit v_z
- Bezeichnung für **Fließkommaformate**: SmEc

Verstecktes Bit (engl. **Hidden bit**)

$$z = (\text{-}1)^{v_z} \cdot 2^{\overbrace{C-k}^{E}} \cdot (\overset{\downarrow}{1}\,.\,M)_2 = (\text{-}1)^{Z_{n-1}} \cdot 2^{\sum_{j=m}^{n-2}[Z_j 2^j] - \overbrace{(2^{\overbrace{n-m-1}^{c}-1}-1)}^{k}} \cdot \underbrace{\frac{1}{2^m}\sum_{i=0}^{m-1} Z_j 2^j}_{M}$$

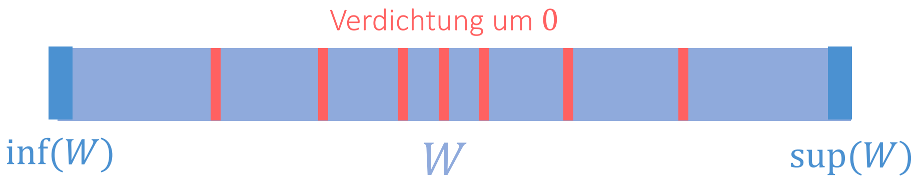
	Vorzeichenbit	Charakteristik	Mantisse
$\vec{Z} \equiv$	v_z	$c-1$ … C … 0	M
	$n-1$	$n-2$ … m	$m-1$ … 0

(Festkommazahlen)



Äquidistante Zerlegung

(Fließkommazahlen)



Zerlegung in uneinheitliche Abstände

- Die **IEEE-754-Norm** umfasst drei unterschiedliche Gleitkommaformate
  - **Single-Precision-Format** (S23E8) mit Vorzeichenbit: $n = 1 + 8 + 23 = 32$
  - **Double-Precision-Format** (S52E11) mit Vorzeichenbit: $n = 1 + 11 + 52 = 64$
  - **Extended-Precision-Format** (S64E15) mit Vorzeichenbit: $n = 1 + 15 + 64 = 80$
- Darüber hinaus bietet IEEE-754 Möglichkeiten, auch ± 0 und $\pm\infty$ darzustellen: Darstellung der Null, wenn $E = 0$ und $M = 0$ sind
- Die Norm IEEE-754 wurde 1985 durch das **Institute of Electrical and Electronics Engineers** festgelegt

IEEE-754 Einfaches Präzisionsformat S23E8

Vorzeichen	Charakteristik	Mantisse	Bedeutung	Mathematische Darstellung
0	$000 \cdots 000$	M	Fließkommazahl ohne 1	$(0.M)_2 \cdot 2^{-127+1}$
1	$000 \cdots 000$	M	Fließkommazahl ohne 1	$-(0.M)_2 \cdot 2^{-127+1}$
0	$0 < E < 2^{c-1} - 1$	M	Fließkommazahl	$(1.M)_2 \cdot 2^{C-127}$
1	$0 < E < 2^{c-1} - 1$	M	Fließkommazahl	$-(1.M)_2 \cdot 2^{C-127}$
0	$111 \cdots 111$	$000 \cdots 000$	Unendlich (positiv)	inf bzw. ∞
1	$111 \cdots 111$	$000 \cdots 000$	Unendlich (negativ)	-inf bzw. $-\infty$
0	$111 \cdots 111$	$M \neq 0$	Keine (rationale) Zahl	$\mathrm{N.a.N.}$ z.B.: $\underset{+}{\sqrt[2]{-1}} \cdot (1.M)$
1	$111 \cdots 111$	$M \neq 0$	Keine (rationale) Zahl	$\mathrm{N.a.N.}$ z.B.: $\underset{-}{\sqrt[2]{-1}} \cdot (1.M)$

v_z	7 C 0	M
31	30 … 23	22 … 0

IEEE-754 Doppeltes Präzisionsformat **S52E11**

Vorzeichen	Charakteristik	Mantisse	Bedeutung	Mathematische Darstellung
0	$000 \cdots 000$	M	Fließkommazahl ohne 1	$(0.M)_2 \cdot 2^{\text{-}1023+1}$
1	$000 \cdots 000$	M	Fließkommazahl ohne 1	$\text{-}(0.M)_2 \cdot 2^{\text{-}1023+1}$
0	$0 < E < 2^{c-1} - 1$	M	Fließkommazahl	$(1.M)_2 \cdot 2^{C-1023}$
1	$0 < E < 2^{c-1} - 1$	M	Fließkommazahl	$\text{-}(1.M)_2 \cdot 2^{C-1023}$
0	$111 \cdots 111$	$000 \cdots 000$	Unendlich (positiv)	inf bzw. ∞
1	$111 \cdots 111$	$000 \cdots 000$	Unendlich (negativ)	-inf bzw. $\text{-}\infty$
0	$111 \cdots 111$	$M \neq 0$	Keine (rationale) Zahl	N.a.N. z.B.: $\underset{+}{\sqrt[2]{\text{-}1}} \cdot (1.M)$
1	$111 \cdots 111$	$M \neq 0$	Keine (rationale) Zahl	N.a.N. z.B.: $\underset{-}{\sqrt[2]{\text{-}1}} \cdot (1.M)$

v_z	10 C 0	M
63	62 … 52	51 … 0

IEEE-754 Erweitertes Präzisionsformat **S64E15**

Vorzeichen	Charakteristik	Mantisse	Bedeutung	Mathematische Darstellung
0	$000 \cdots 000$	M	Fließkommazahl ohne 1	$(0.M)_2 \cdot 2^{-16383+1}$
1	$000 \cdots 000$	M	Fließkommazahl ohne 1	$-(0.M)_2 \cdot 2^{-16383+1}$
0	$0 < E < 2^{c-1} - 1$	M	Fließkommazahl	$(1.M)_2 \cdot 2^{C-16383}$
1	$0 < E < 2^{c-1} - 1$	M	Fließkommazahl	$-(1.M)_2 \cdot 2^{C-16383}$
0	$111 \cdots 111$	$000 \cdots 000$	Unendlich (positiv)	inf bzw. ∞
1	$111 \cdots 111$	$000 \cdots 000$	Unendlich (negativ)	-inf bzw. $-\infty$
0	$111 \cdots 111$	$M \neq 0$	Keine (rationale) Zahl	N. a. N. z.B.: $\underset{+}{\sqrt[2]{-1}} \cdot (1.M)$
1	$111 \cdots 111$	$M \neq 0$	Keine (rationale) Zahl	N. a. N. z.B.: $\underset{-}{\sqrt[2]{-1}} \cdot (1.M)$

v_z	C (14 … 0)	M
79	78 … 64	63 … 0

Datentypen für Fließkommazahlen in **C** und **C++**

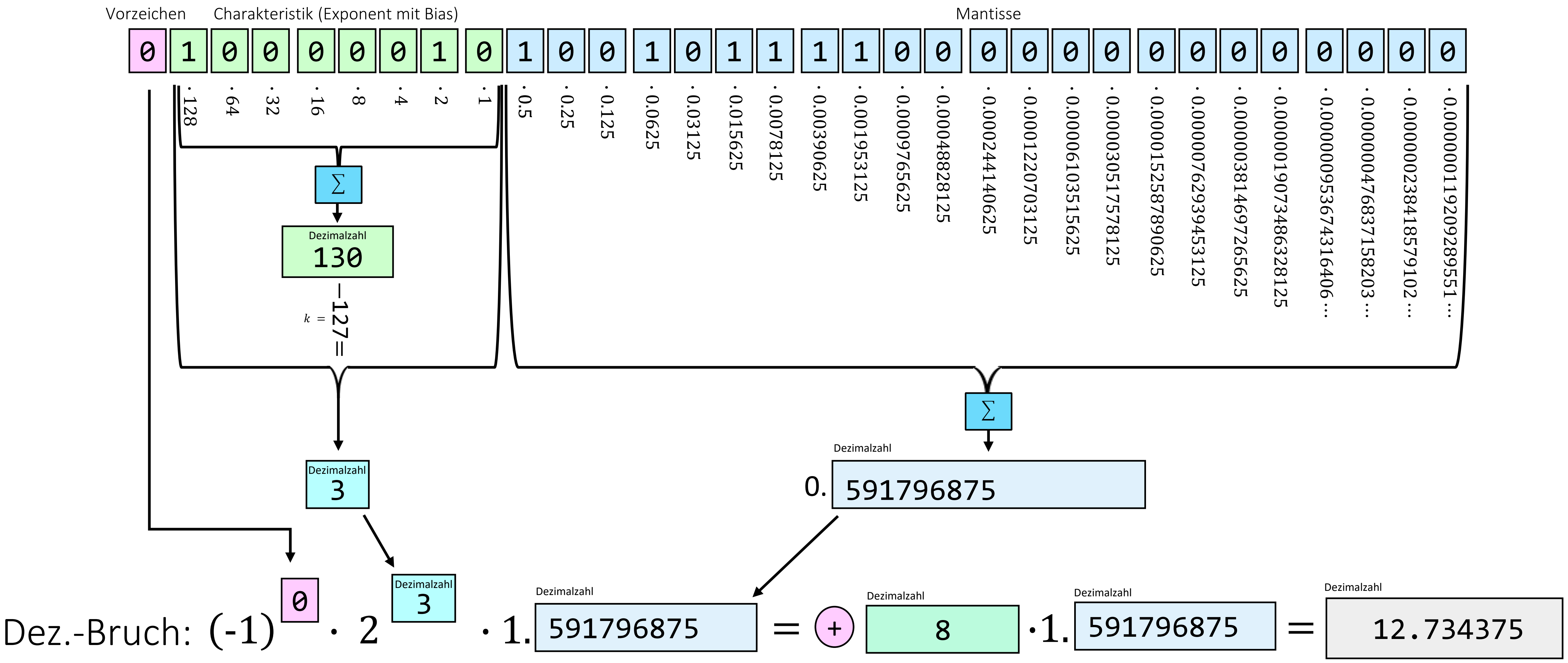
Theorieaufgabe f)

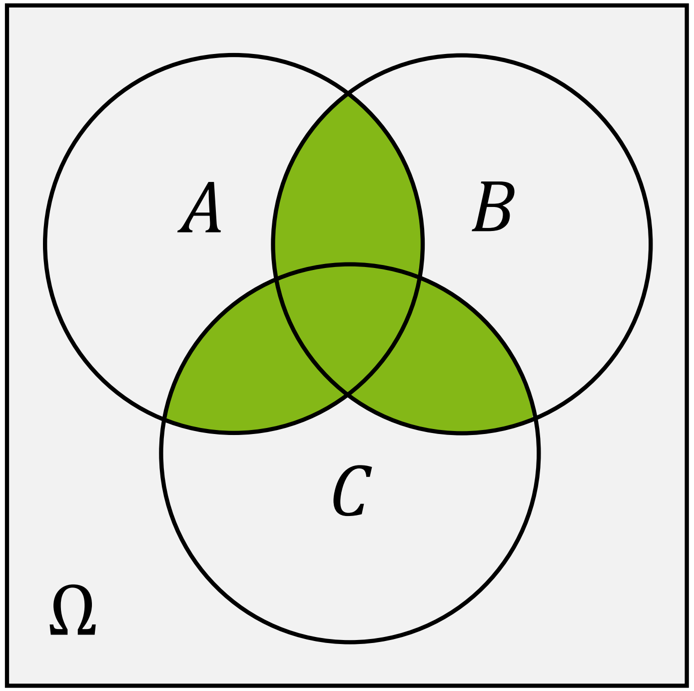
Art der Daten	Schlüsselwort	Format	Speicherplatz	Platzhalter
Rationale Zahlen $\mathbb{Q}$	`float`	Single-Precision-Format **S8E32**	32 Bit $\widehat{=}$ 8 Nibble $\widehat{=}$ 4 Byte	%f
	`double`	Double-Precision-Format **S11E52**	64 Bit $\widehat{=}$ 16 Nibble $\widehat{=}$ 8 Byte	%lf
	`long double`	Extended-Precision-Format **S15E64**	80 Bit $\widehat{=}$ 20 Nibble $\widehat{=}$ 10 Byte	%Lf

Fließkomma-Verarbeitungseinheit

- **Fließkomma-Verarbeitungseinheit** (engl. **Floating-Point-Unit**, kurz: **FPU**) ist auf **Fließkommazahlenarithmetik** spezialisierte Hardware
- Bereits 1980 nutzte die FPU 8087 von Intel® das **S15E64**-Format

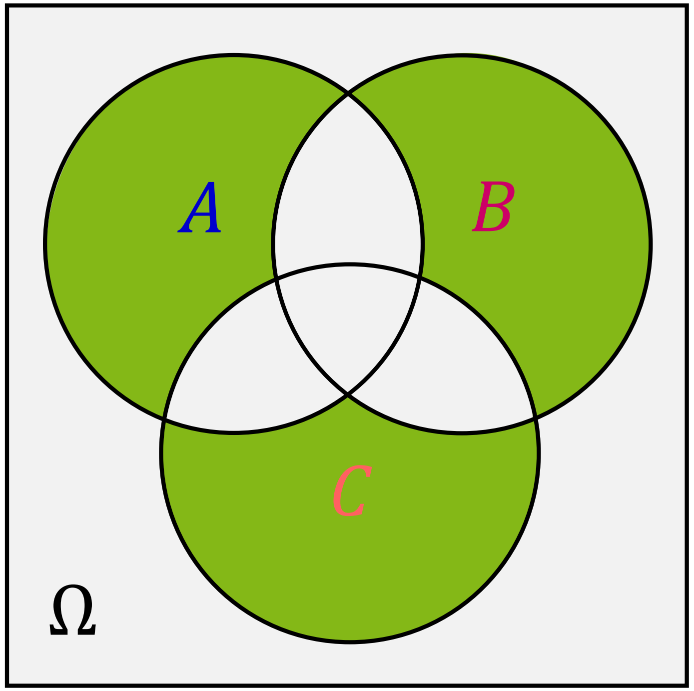
Lesen einer **S23E8**-Fließkommazahl (Lösung)



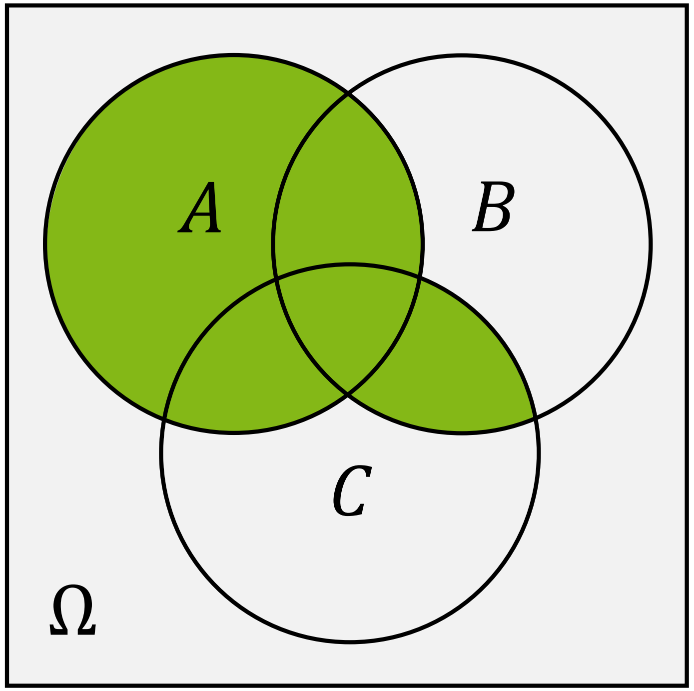


`(A&B)|(A&C)|(B&C)`

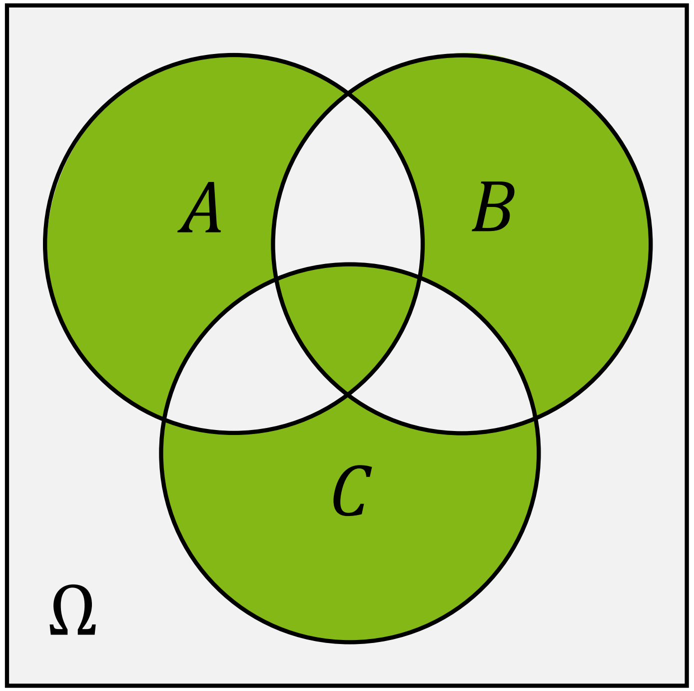
Venn-Diagramm $b)$



(A&~B&~C)|(~A&B&~C)|(~A&~B&C)




```
A | (B & C)
```



A ^ B ^ C

Rationale Zahlen sind im allgemeinen als **Brüche** bekannt und erlauben es, Zahlen miteinander zu dividieren, die als nicht teilbar gelten.

$\mathbb{Q} = \left\{ \frac{q}{p} \,\middle|\, q \in \mathbb{Z},\ p \in \mathbb{N}_{\neq 0} \right\} \cong \mathbb{N}^2 \cong \mathbb{Z}^2$ (Da ja $\mathbb{N} \cong \mathbb{Z}$ gilt)

q heißt **Zähler** (**Numerator**) und p heißt **Nenner** (**Denominator**)

$$+_{\mathbb{Q}}\colon \mathbb{Q}^2 \to \mathbb{Q};\quad \left(\frac{a}{b}, \frac{c}{d}\right) \mapsto \frac{a *_{\mathbb{Z}} d +_{\mathbb{Z}} c *_{\mathbb{Z}} b}{b *_{\mathbb{Z}} d} \qquad -_{\mathbb{Q}}\colon \mathbb{Q}^2 \to \mathbb{Q};\quad \left(\frac{a}{b}, \frac{c}{d}\right) \mapsto \frac{a *_{\mathbb{Z}} d -_{\mathbb{Z}} c *_{\mathbb{Z}} b}{b *_{\mathbb{Z}} d}$$

$$*_{\mathbb{Q}}\colon \mathbb{Q}^2 \to \mathbb{Q};\quad \left(\frac{a}{b}, \frac{c}{d}\right) \mapsto \frac{a *_{\mathbb{Z}} c}{b *_{\mathbb{Z}} d} \qquad \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a}{b} \cdot \frac{d}{c} = \frac{a \cdot d}{b \cdot c}$$

Eine Zahl $r^{-1} = \frac{1}{r}$ heißt **Reziproke**. Rationale Zahlen lassen sich ordnen. Es gilt die Regel: $0 < b < a \Rightarrow 0 < \frac{1}{a} < \frac{1}{b}$. **Dezimalbruch**: $a.\underbrace{b}_{pos} = a + \frac{b}{\underbrace{10 * \cdots * 10}_{pos}}$

Auch jede periodische Zahl ist eine rationale Zahl (dh. ein Bruch)!

$$x \coloneqq \left(c_{n-1}c_{n-2}\cdots c_0, c_{-1}c_{-2}\cdots c_{-m}\overline{a_{p-1}\cdots a_1 a_0}\right)_b$$

$$= \frac{C + b^{-p}(A - C)}{1 - b^{-p}} \cdot \frac{1}{b^m} = \frac{b^p C + b^p b^{-p}(A - C)}{b^p b^m - b^m b^{-p} b^p} \cdot \frac{\cancel{b^p}}{\cancel{b^p}} = \boxed{\frac{b^p C + A - C}{b^{p+m} - b^m}}$$

Beispiel für $b = 10$: $x \coloneqq 15.\underbrace{78}_{m=2}\underbrace{\overline{123}}_{p=3} \Rightarrow C \coloneqq 1578, A \coloneqq 123$

$$x = \frac{10^3 \cdot 1578 + 123 - 1578}{10^{3+2} - 10^2} = \frac{105103}{6660}$$

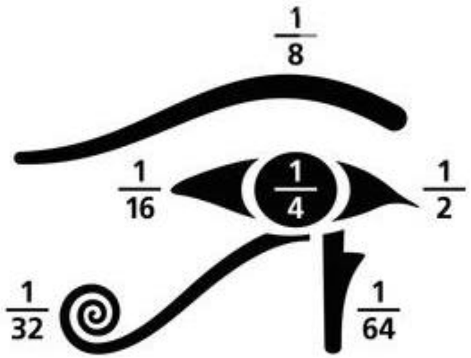
$$x := \left(c_{n-1}c_{n-2}\cdots c_0, c_{-1}c_{-2}\cdots c_{-m}\overline{a_{p-1}\cdots a_1 a_0}\right)_b$$

$$= \frac{C + b^{-p}(A - C)}{1 - b^{-p}} \cdot \frac{1}{b^m} = \frac{b^p C + b^p b^{-p}(A - C)}{b^p b^m - b^m b^{-p} b^p} \cdot \frac{\cancel{b^p}}{\cancel{b^p}} = \boxed{\frac{b^p C + A - C}{b^{p+m} - b^m}}$$

Beispiel für $b = 10$: $x := 0.\overline{999}$

$$0.\overline{9} = 0.9999\cdots = \sum_{i=1}^{\infty}\left(\frac{9}{10^i}\right) = 9\sum_{i=0}^{\infty}\left(0.1^i\right) - 9 = \frac{9}{1 - \frac{1}{10}} - 9 = 1$$

Anhang: Ägyptische Brüche



Dezimalbruch in Festkommazahl (Vorzeichenbitdarstellung) & IEEE-754 konvertieren

Theorie