

## Übungsblatt 5

### Aufgabe A – Gleitkommazahlen

Beantworten Sie folgende Fragen:

1. Warum werden i.d.R. Fließkommazahlen und nicht Festkommazahlen zur Darstellung von rationalen Zahlen verwendet? Anders formuliert: Welche Vorteile bringt diese Art der Darstellung?
2. Für eine Fließkommazahl-Darstellung werden die gegebenen Bits in drei Abschnitte unterteilt. Wie heißen diese Abschnitte?
3. Welchen Vorteil bringt es wenn dem Exponenten mehr Bits zugeteilt werden bzw. welchen Vorteil bringt es wenn der Mantisse mehr Bits zugeteilt werden?
4. Erklären sie Unterlauf (Underflow) und Überlauf (Overflow) für Fließkommazahlen.
5. Nennen sie Beispiele für Festlegungen, die der IEEE 754 Standard mitbringt. Warum ist eine solche Standardisierung sinnvoll?
6. Was ist die betragsmäßig größte bzw. kleinste darstellbare Zahl im IEEE-754 32bit Standard? Geben sie auch die Bits an.
7. Berechnen sie die Gleitkommazahlen:
  - a) 18,6125
  - b) 23,42

### Aufgabe B – Gleitkommazahlen in C

Gegeben sei folgender C Code (in Moodle auch als Sourcecode).

```
1 #include <stdio.h>
2 #include <ctype.h>
3 union zahl {
4     float float_zahl;
5     double double_zahl;
6     struct {
7         unsigned bitmuster1;
8         unsigned bitmuster2;
9     } long_format;
10 };
11 union zahl z1, z2;
12
13 void dual_ausgabe(unsigned long zahl, int strich1, int strich2) {
14     int i;
15     for (i=sizeof(long)*8-1; i>=0; i--)
```

```
16     printf("%ld%s", (zahl>>i) & 1, (i==strich1 || i==strich2) ? " | " : "");
17 }
18
19 int main(void)
20 {
21     printf("Gib Deine Gleitpunktzahl ein: ");
22     scanf("%lf", &z1.double_zahl);
23
24     z2.float_zahl = z1.double_zahl;
25     printf("double: ");
26     dual_ausgabe(z1.long_format.bitmuster2, 31, 20);
27     dual_ausgabe(z1.long_format.bitmuster1, -1, -1);
28     printf("\n float: ");
29     dual_ausgabe(z2.long_format.bitmuster1, 31, 23);
30     printf("\n");
31
32     return 0;
33 }
```

1. Kompilieren sie den obigen Code und Testen sie ihn mit den aus Aufgabenteil A umgerechneten Zahlen.
2. Lesen sie den code und versuchen sie den Code zunächst zu verstehen.
3. Erläutern sie die Umrechnung von Dezimal in die beiden Gleitkommaformate (float, double) nach IEEE 754.