Prof. Dr. Dirk Lebiedz
Dipl.-Math. oec. Klaus Stolle
Institut für Numerische Mathematik
Universität Ulm

Angewandte Numerik 1
SoSe 2020
22.05.2020

Übungsblatt 05

Besprechung in den Tutorien in der Woche vom 01.06.2020 bis 05.06.2020

Für dieses Übungsblatt gibt es 29 Theorie- und 25 Matlab-Punkte.
Die 70-Prozent-Grenzen liegen aktuell (inklusive Blatt 05) bei 96,6 Theorie- und 53,9 Matlabpunkten.

Aufgabe 16 *(Zahlendarstellung: Basiswechsel)* (8T+2T Punkte)

a) Für die Anwendung von Bedeutung sind bei der Zahlendarstellung die Basen $b = 10$ (Dezimalzahlen), $b = 2$ (Binärzahlen), $b = 8$ (Oktalzahlen) und $b = 16$ (Hexadezimalzahlen). Bei Hexadezimalzahlen treten Ziffern mit Werten zwischen 0 und 15 auf. Damit alle Ziffern einstellig notiert werden können, werden die Ziffern mit den Werten 10 bis 15 durch die Buchstaben A bis F dargestellt.

   Ergänzen Sie folgende Tabelle (geben Sie dabei alle Rechnungen an):

Dezimal	Dual	Oktal	Hexadezimal
30.125			
	11011.011111		
		75.21	
			A.BC

b) Welche Rechnungen sind aufwendig, welche nicht? Woran liegt das?

Aufgabe 17 *(*Programmieraufgabe: Darstellung natürlicher Zahlen) (6M+1M Punkte)

Wir beschränken uns zunächst auf natürliche Zahlen $n \in \mathbb{N}$ und suchen deren Darstellung zu einer gegebenen Basis $b \in \mathbb{N}$, $b \geq 2$. Gesucht sind also die Koeffizienten $a_k \in \{0, 1, ..., b-1\}$, $k = 0, ..., m$ der Darstellung $n = \sum_{k=0}^{m} a_k b^k$.

a) Schreiben Sie eine Matlab-Funktion `a = convert2basis(n, b)`, die für eine natürliche Zahl `n` und eine Basis `b` den Koeffizienten-Vektor `a` mit $a = (a_m, a_{m-1}, ..., a_1, a_0)$ berechnet.

b) Testen Sie die Funktion mit Hilfe des Skriptes `testConvert2basis.m`, welches auf der Homepage verfügbar ist.

Aufgabe 18 *(*Programmieraufgabe: Wert einer Gleitpunkt-Darstellung) (8M+1M+8M+1M Punkte)

Zu einer gegebenen Basis $b \in \mathbb{N}$, $b \geq 2$ lässt sich eine positive reelle Zahl $x \in [b^{-b^n}, b^{b^n - 1})$ eindeutig als

$$x = b^e \cdot f = b^{t \cdot e(\mathbf{v})} \cdot \sum_{j=1}^{\infty} d_j b^{-j} \qquad \text{mit} \qquad e(\mathbf{v}) = \sum_{j=0}^{n-1} v_j b^j,$$

und $1 \leq d_1 \leq b-1$, $0 \leq d_j \leq b-1$ für $j = 2, 3, \ldots$ und $d_j < b-1$ für unendlich viele j darstellen. Der Exponent e ist in dieser Darstellung eine n-stellige natürliche Zahl $e(\mathbf{v})$ zur Basis b mit den Ziffern $\mathbf{v} = (v_{n-1}, \ldots, v_0)$ und dem Vorzeichen $t \in \{-1, 1\}$.

In dieser Aufgabe beschränken wir uns auf positive reelle Zahlen x, die sich zur gegebenen Basis b mit endlicher Mantissenlänge m darstellen lassen.

a) Schreiben Sie eine Matlab-Funktion `x = value(b, d, v, t)`, die für eine gegebene Gleitpunktdarstellung, also eine Basis `b` $\in \mathbb{N}$, `b` ≥ 2, einen Zeilenvektor `d` der Koeffizienten der Mantisse, `d` $= (d_1, d_2, ..., d_m)$, einen Zeilenvektor `v` der Koeffizienten des Exponenten, `v` $= (v_{n-1}, ...v_1, v_0)$, sowie ein Vorzeichen `t` (`t` $\in \{-1, 1\}$), den Wert `x` der Gleitpunkt-Darstellung berechnet.

b) Testen Sie Ihre Funktion mit Hilfe des Skriptes `testValue.m`, welches auf der Homepage verfügbar ist.

c) Schreiben Sie eine Matlab-Funktion `[d, v, t] = flp(b, m, n, x)`, welche zu einer Basis `b` $\in \mathbb{N}, b \geq 2$, einer Mantissenlänge `m`, einer Exponetenlänge `n` und einer zu konvertierenden Zahl `x` deren normalisierte Gleitpunkt-Darstellung berechnet und `d`, `v` sowie `t` zurückliefert. Sie dürfen davon ausgehen, dass sich die zu konvertierende Zahl mit einer `m`-stelligen Mantisse und einem `n`-stelligen Exponent ohne Rundung darstellen lässt.

   Der Aufruf `[d, v, t] = flp(2, 3, 3, 0.0625)` sollte also die Werte `d = [1, 0, 0]`, `v = [0, 1, 1]` und `t = -1` zurückliefern.

d) Testen Sie Ihre Funktion mit Hilfe des Skriptes `testFlp.m`, welches auf der Homepage verfügbar ist.

Aufgabe 19 *(Maschinenzahlen)* (4T+4T Punkte)

In dieser Aufgabe betrachten wir die Menge der Maschinenzahlen $\mathbb{M}(b, m, r, R)$ in normalisierter Gleitpunktdarstellung mit Basis b, m-stelliger Mantisse und Exponenten $r \leq e \leq R$.

a) Mit x_{MIN} und x_{MAX} seien die betragsmäßig kleinste ($\neq 0$) bzw. grösste Zahl in $\mathbb{M}(b, m, r, R)$ bezeichnet. Zeigen Sie: $x_{\text{MIN}} = b^{r-1}$ und $x_{\text{MAX}} = (1 - b^{-m})\, b^R$.

b) Sei $x \in [x_{\text{MIN}}, x_{\text{MAX}}]$ und habe die Darstellung:

$$x = \left(\sum_{j=1}^{\infty} d_j b^{-j} \right) b^e, \qquad \text{mit } r \leq e \leq R,\ e \in \mathbb{Z}$$

   Sei $\text{fl}(x)$ die Maschinenzahl, die sich durch die Standardrundung ergibt. Zeigen Sie folgende Abschätzung:

$$|\text{fl}(x) - x| \leq \frac{b^{-m}}{2} b^e$$

Aufgabe 20 *(Umwandlung in und Operationen auf Maschinenzahlen)* (3T+3T+5T Punkte)

Gegeben seien die beiden Zahlen $x = (\frac{3}{5})_{10}$ und $y = (\frac{4}{7})_{10}$.

a) Bestimmen Sie für x, y und $x - y$ jeweils die Darstellung (2.42) der Form $x = \pm \left(\sum_{j=1}^{\infty} d_j b^{-j} \right) \cdot b^e$.

b) Bestimmen Sie die Darstellungen $\tilde{x}$ und $\tilde{y}$ von x und y jeweils in $\mathbb{M}(2, 5, -7, 8)$ und $\mathbb{M}(2, 3, -7, 8)$. Geben Sie auch die Differenz $x - y$ in $\mathbb{M}(2, 5, -7, 8)$ und $\mathbb{M}(2, 3, -7, 8)$ an. Runden Sie mit Standardrundung, falls die Zahlen nicht exakt darstellbar sind.

c) Berechnen Sie jeweils in $\mathbb{M}(2, 5, -7, 8)$ und in $\mathbb{M}(2, 3, -7, 8)$ $\tilde{x} \ominus \tilde{y}$. Wie groß ist der jeweilige absolute und der jeweilige relative Fehler? Vergleichen Sie jeweils $\tilde{x} \ominus \tilde{y}$ mit der Darstellung von $x - y$ in $\mathbb{M}(2, 5, -7, 8)$ bzw. $\mathbb{M}(2, 3, -7, 8)$. Wie heißt das beobachtete Phänomen?

Hinweise:
Die Lösungen der Theorieaufgaben und Ihren Text zu den Programmieraufgaben können Sie in LaTeX erstellen oder handschriftlich aufschreiben und einscannen. Die Programmieraufgaben sind in Matlab zu lösen. Der Source Code muss strukturiert und dokumentiert sein.
Falls Sie die Aufgaben im Team lösen, geben Sie bitte auf allen Lösungen alle an der Aufgabe beteiligten Teammitglieder an. Jedes Teammitglied muss eine Lösung abgeben.
Speichern Sie Ihre Lösungen und Ihre Ergebnisse in einem Directory mit dem Namen **Blatt05_*Vorname*_*Nachname*** und verpacken Sie dieses in eine `.zip`-Datei. Laden Sie **spätestens 48 Stunden vor Ihrem Tutorium** diese `.zip`-Datei in Moodle hoch.