



ÜBUNG 6

ZUR EINFÜHRUNG IN DIE PROGRAMMIERUNG FÜR COMPUTERLINGUISTEN

ZAHLENSYSTEME

- Normalerweise schreibt man Zahlen im sog. Dezimalsystem

4 2 6 9



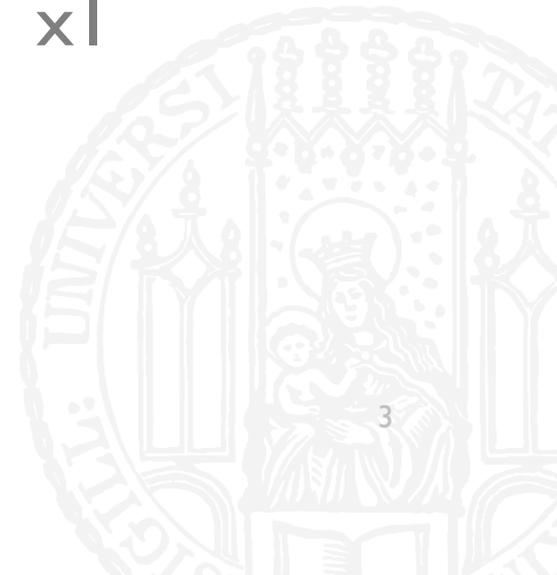
ZAHLENSYSTEME

4
x100

2
x100

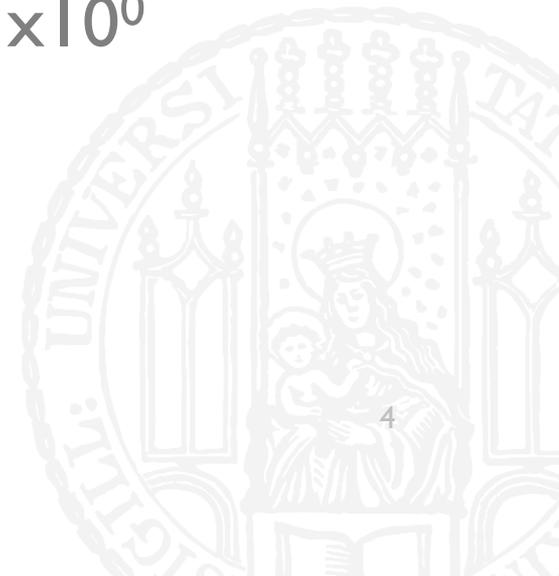
6
x10

9
x1



ZAHLENSYSTEME

$$\begin{array}{cccc} \mathbf{4} & \mathbf{2} & \mathbf{6} & \mathbf{9} \\ \times 10^3 & \times 10^2 & \times 10^1 & \times 10^0 \end{array}$$



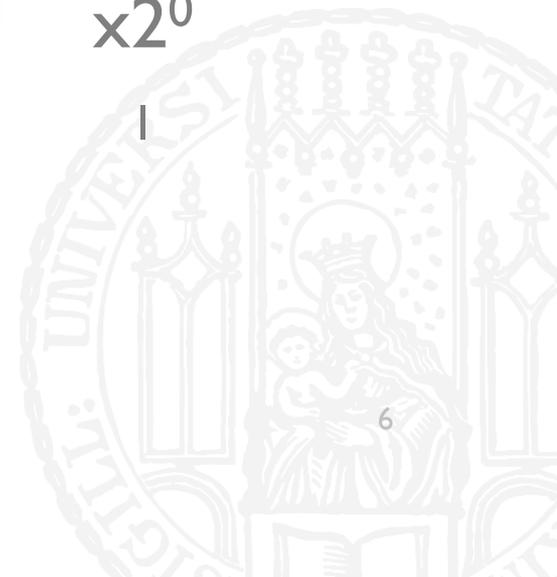
2 ANSTATT 10

- Computer können nur mit Booleans Rechnen
- True oder False
- Kekse oder keine Kekse
- Spam oder kein Spam
- Strom oder kein Strom
- Ladung oder keine Ladung
- 1 oder 0



BINÄRSYSTEM

0	1	0	1	0	1	0
$\times 2^6$	$\times 2^5$	$\times 2^4$	$\times 2^3$	$\times 2^2$	$\times 2^1$	$\times 2^0$
64	32	16	8	4	2	1



BINÄRSYSTEM

$$\begin{array}{ccccccc} \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \times 2^6 & \times 2^5 & \times 2^4 & \times 2^3 & \times 2^2 & \times 2^1 & \times 2^0 \end{array}$$

$$= \mathbf{0} * 2^6 + \mathbf{1} * 2^5 + \mathbf{0} * 2^4 + \mathbf{1} * 2^3 + \mathbf{0} * 2^2 + \mathbf{1} * 2^1 + \mathbf{0} * 2^0$$



BINÄRSYSTEM

$$\begin{array}{ccccccc} \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \times 2^6 & \times 2^5 & \times 2^4 & \times 2^3 & \times 2^2 & \times 2^1 & \times 2^0 \end{array}$$

$$\begin{aligned} &= \mathbf{0} * 2^6 + \mathbf{1} * 2^5 + \mathbf{0} * 2^4 + \mathbf{1} * 2^3 + \mathbf{0} * 2^2 + \mathbf{1} * 2^1 + \mathbf{0} * 2^0 \\ &= 2^5 + 2^3 + 2^1 = 32 + 8 + 2 = \mathbf{42} \end{aligned}$$

QUIZ

- Was ist **00110** im Dezimalsystem?
- a) 2
- b) 5
- c) 6
- d) 11



QUIZ

- Was ist **00110** im Dezimalsystem?
- a) 2
- b) 5
- c) 6
- d) 11



QUIZ

- Was ist **10010** im Dezimalsystem?
- a) 36
- b) 18
- c) 10
- d) 9



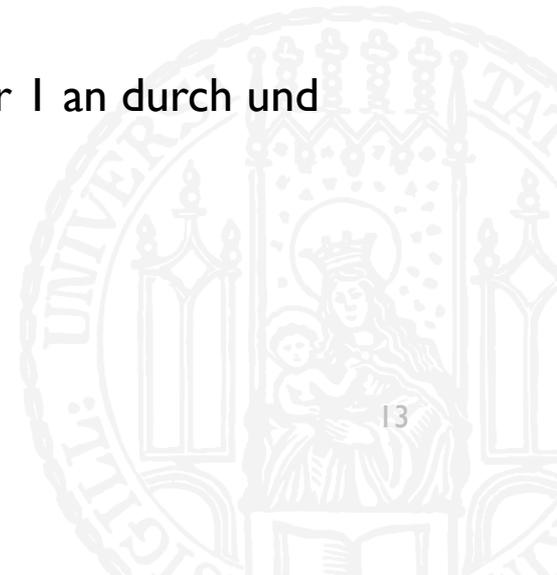
QUIZ

- Was ist **10010** im Dezimalsystem?
- a) 36
- b) 18
- c) 10
- d) 9



DEZIMAL NACH BINÄR

- Um eine Dezimalzahl in das Binärsystem umzuwandeln, muss man die enthaltenen Zweipotenzen finden
- Hierzu zieht man immer wieder die größte enthaltene Zweierpotenz ab und notiert diese
- Am Ende geht man alle existierenden Zweierpotenzen von der 1 an durch und notiert für jede gefundene eine 1 und für jede andere eine 0



DEZIMAL NACH BINÄR

- 54
- Größte Zweierpotenz in 54: **32**
- $54 - 32 = 22$
- Größte Zweierpotenz in 22: **16**
- $22 - 16 = 6$
- Größte Zweierpotenz in 6: **4**
- $6 - 4 = 2$
- Größte Zweierpotenz in 2: **2**
- $2 - 2 = 0$
- *Fertig*

64	32	16	8	4	2	1
0			0			0



QUIZ

■ Was ist 10 im Binärformat?

- a) 0010
- b) 0110
- c) 1010
- d) 0101



QUIZ

■ Was ist 10 im Binärformat?

- a) 0010
- b) 0110
- c) 1010
- d) 0101



QUIZ

■ Was ist 29 im Binärformat?

- a) 10110
- b) 10111
- c) 11101
- d) 11011



QUIZ

■ Was ist 29 im Binärformat?

- a) 10110
- b) 10111
- c) 11101
- d) 11011



OKTALSYSTEM

- Neben dem Binär (2) und dem Dezimalsystem (10) kann man auch beliebige andere Basen betrachten
- Im Oktalsystem rechnet man in Basis 8
- Dadurch sind die Ziffern 0 1 2 3 4 5 6 und 7 vorhanden



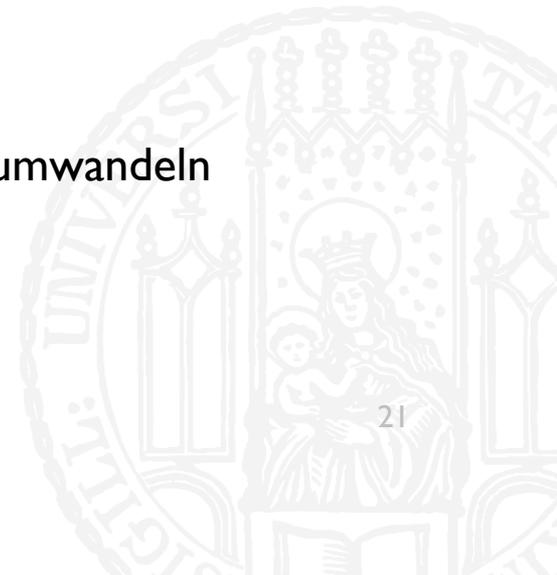
OKTALSYSTEM

3	0	6	4
$\times 8^3$	$\times 8^2$	$\times 8^1$	$\times 8^0$
256	64	8	1



OKTALSYSTEM

- Das Oktalsystem kann sehr einfach ins Binärsystem- und zurück umgewandelt werden
- Eine Ziffer im Oktalsystem hat den gleichen Wertebereich wie drei Ziffern im Binärsystem
- $0 - 7 \Leftrightarrow 000 - 111$
- Deswegen kann man jede Oktalziffer einzeln in 3 Binärziffern umwandeln



OKTALSYSTEM

- Oktal 74 = 111 100
- Oktal 412 = 100 001 010
- Oktal 36123 = 011 110 001 010 011
- Binär 010101 = Oktal 25
- Binär 111000 = Oktal 70
- Binär 101110 = Oktal 56

Oktal	Binär
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111

QUIZ

■ Was ist Oktal 29 im Binärsystem?

- a) 010 110
- b) 110 111
- c) 011 101
- d) 111 011



QUIZ

- Was ist Oktal 29 im Binärformat?



QUIZ

■ Was ist Oktal 22 im Binärsystem?

- a) 010 110
- b) 110 011
- c) 011 011
- d) 010 010



QUIZ

■ Was ist Oktal 22 im Binärsystem?

a) 010 110

b) 110 011

c) 011 011

d) 010 010



QUIZ

- Was ist Binär 101 011 010 im Oktalsystem?
 - a) 221
 - b) 532
 - c) 364
 - d) 521



QUIZ

- Was ist Binär 101 011 010 im Oktalsystem?
 - a) 221
 - b) 532
 - c) 364
 - d) 521



HEXADEZIMALSYSTEM

- Ein sehr beliebtes System zum Darstellen von Computerdaten ist das Hexadezimalsystem (Basis 16)
- Um die Basis 16 zu verwenden sind noch 6 weitere Ziffern nötig
- 0 1 2 3 4 5 6 7 8 9 A B C D E F
- Es wird gesetzt $A = 10$, $B = 11$ usw...
- Die Umwandlung funktioniert wie beim Oktalsystem, jedoch in Blöcken von vier
- Im Hexadezimalsystem kann man ein Byte (8 Bit/Binärstellen) in zwei Ziffern darstellen



HEXADEZIMALSYSTEM

- Hex A4 = Binär 1010 0100
- Hex FF = Binär 1111 1111
- Hex D00F = Binär 1101 0000 0000 1111
- Hex DIEB = Binär 1101 0001 1110 1011

Hex	Binär
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001
A (10)	1010
B (11)	1011
C (12)	1100
D (13)	1101
E (14)	1110
F (15)	1111

QUIZ

- Was ist Binär 101011010 im Hexadezimalsystem?
- a) 15A
- b) AD0
- c) 532
- d) 111AA



QUIZ

- Was ist Binär 101011010 im Hexadezimalsystem?
- a) 15A
- b) AD0
- c) 532
- d) 111AA



QUIZ

- Was ist Hexadezimal DEADAFFE in Binär?
 - a) 1101 1110 1010 1101 1010 1111 1111 1110
 - b) 1110 1101 1100 1110 1100 1111 1111 1101
 - c) 0011 1101 1010 0011 1010 1111 1111 1101
 - d) 0011 1110 1010 0011 1010 1111 1111 1110



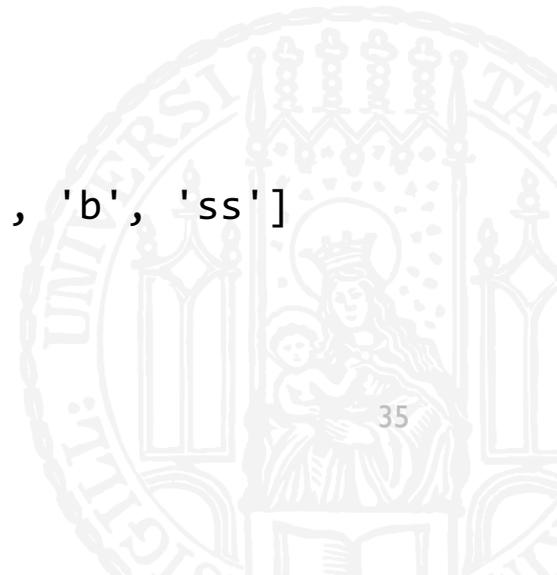
QUIZ

- Was ist Hexadezimal DEADAFFE in Binär?
 - 1101 1110 1010 1101 1010 1111 1111 1110
 - 1110 1101 1100 1110 1100 1111 1111 1101
 - 0011 1101 1010 0011 1010 1111 1111 1101
 - 0011 1110 1010 0011 1010 1111 1111 1110



SPLIT

- Mit `string.split('x')` kann man einen String in Einzelteile zerlegen
- Man übergibt ein Trennzeichen und erhält eine Liste aller Teile dazwischen zurück
- Das Trennzeichen verschwindet dabei
- `"Hallo Welt".split(' ')`
`["Hallo", "Welt"]`
- `"Dra Chanasa mat da Kantrabass".split('a')`
`['Dr', ' Ch', 'n', 's', ' m', 't d', ' K', 'ntr', 'b', 'ss']`



QUIZ

- Was ist `"Tux Müde. Tux Schlafen".split("u")` ?
- a) [`"Tu"`, `"x Müde. Tu"`, `"x Schlafen"`]
- b) [`"T"`, `"ux Müde. T"`, `"ux Schlafen"`]
- c) [`"T"`, `"ux M"`, `"de. T"`, `"x Schlafen"`]
- d) [`"T"`, `"x Müde. T"`, `"x Schlafen"`]



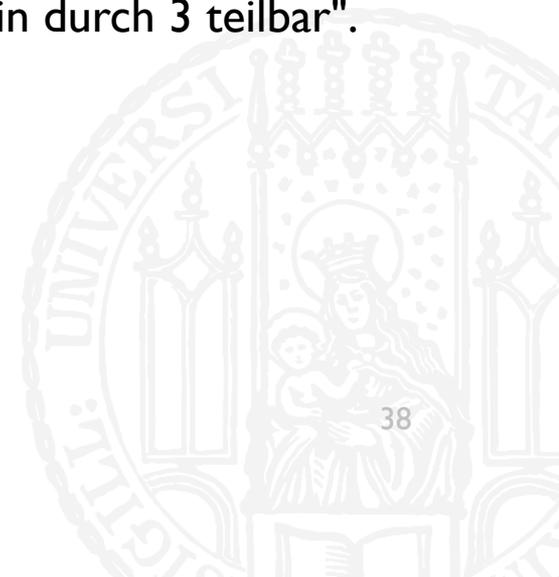
QUIZ

- Was ist `"Tux Müde. Tux Schlafen".split("u")` ?
 - a) `["Tu", "x Müde. Tu", "x Schlafen"]`
 - b) `["T", "ux Müde. T", "ux Schlafen"]`
 - c) `["T", "ux M", "de. T", "x Schlafen"]`
 - d) `["T", "x Müde. T", "x Schlafen"]`



MUSTERLÖSUNG 6-I

Schreiben Sie ein Programm, das mit Hilfe einer Schleife jede Zahl zwischen 0 und 30 ausgibt und bei jeder durch 3 teilbaren Zahl dazu schreibt: "Ich bin durch 3 teilbar".



MUSTERLÖSUNG 6-1

```
#!/usr/bin/python3
#Aufgabe 6-1
#Autorin: Leonie Weißweiler
for x in range(0,31):
    print (x)
    if x % 3 == 0:
        print('Ich bin durch drei teilbar')
```



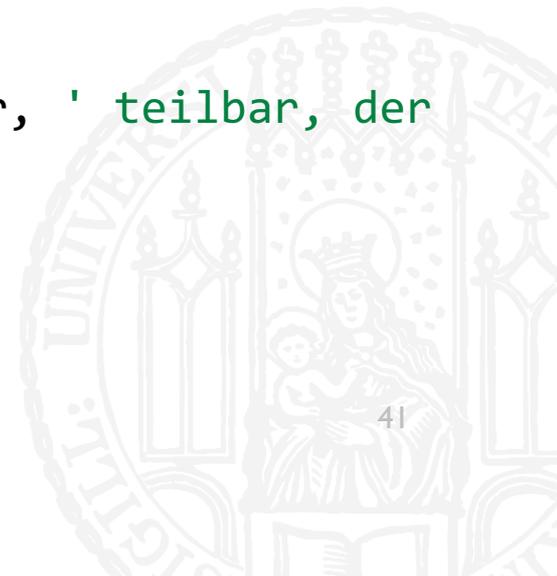
MUSTERLÖSUNG 6-2

Erweitern Sie die letzte Aufgabe dahingehend, dass der Teiler 3 nicht mehr fest vorgegeben ist, sondern der Benutzer nach einem beliebigen Teiler gefragt wird. Mit diesem Teiler soll die Teilbarkeit mit dem Multiplikationsfaktor nach folgendem Muster ausgegeben werden:

Ausgabe: Bitte geben Sie den Teiler ein Eingabe: 5 Ausgabe: 0 : ich bin durch
5 teilbar, der Multiplikator ist 0 1,2,3,4, 5 : ich bin durch 5 teilbar, der
Multiplikator ist 1 6,7,8,9, 10 : ich bin durch 5 teilbar, der Multiplikator ist 2 ...

MUSTERLÖSUNG 6-2

```
#!/usr/bin/python3
#Aufgabe 6-2
#Autorin: Leonie Weißweiler
teiler = int( input( 'Geben Sie einen Teiler ein\n'))
for x in range(0,31):
    if(x % teiler == 0):
        print('\n', x, ' : Ich bin durch ', teiler, ' teilbar, der
        Multiplikator ist ', x//teiler, sep='')
    else:
        print (x, end=',')
```



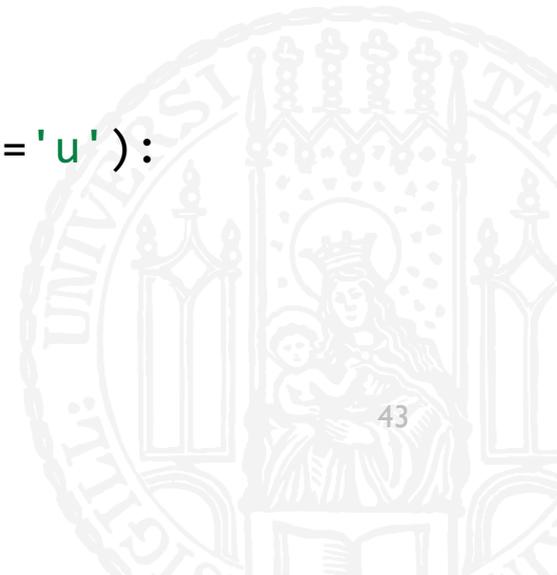
MUSTERLÖSUNG 6-3

Schreiben Sie ein Programm, dem Sie eine Zeile eingeben können, und geben Sie nur die Vokale (aeiou) der Zeile auf dem Terminal aus.



MUSTERLÖSUNG 6-3

```
#!/usr/bin/python3
#Aufgabe 6-3
#Autorin: Leonie Weißweiler
zeile = input('Geben Sie eine Zeile ein\n')
zeile = zeile.lower()
for x in zeile:
    if(x=='a' or x=='e' or x=='i' or x=='o' or x=='u'):
        print(x,end='')
print('\n', end='')
```



MUSTERLÖSUNG 6-4

Holen Sie den Text eines Interviews von Gerhard Polt in der Süddeutschen Zeitung mit folgendem Befehl:

```
lynx -dump http://sz-magazin.sueddeutsche.de/texte/anzeigen/36659
```

und sichern Sie im Verzeichnis Übungsblatt6 den Inhalt des Interviews in der Datei 'artikel.txt'.

MUSTERLÖSUNG 6-4

```
Leonie@Laptop $ cd Übungsblatt6
```

```
Leonie@Laptop $ lynx -dump http://sz-magazin.sueddeutsche.de/texte/anzeigen/36659 > artikel.txt
```

MUSTERLÖSUNG 6-5

Schreiben Sie ein Programm, das den Text aus der Datei 'artikel.txt' liest und das erste Wort jeder Zeile auf dem Terminal ausgibt.



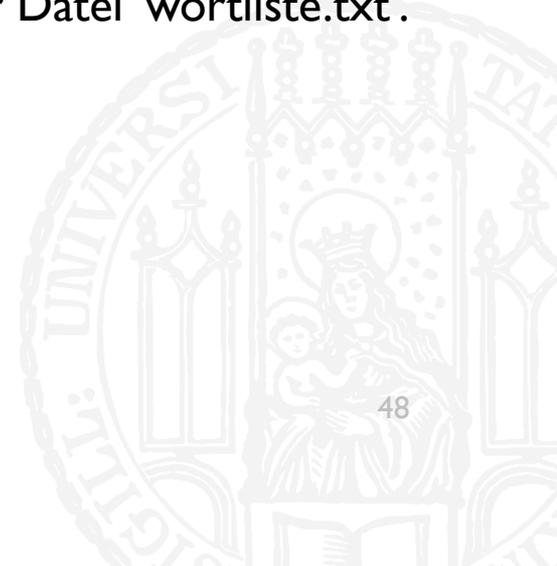
MUSTERLÖSUNG 6-5

```
#!/usr/bin/python3
#Aufgabe 6-5
#Autorin: Leonie Weißweiler
artikel = open('artikel.txt','r')
for line in artikel:
    print(line.split(' ')[0])
artikel.close()
```



MUSTERLÖSUNG 6-6

Schreiben Sie ein Programm, das den Text aus der Datei 'artikel.txt' liest und jedes Wort einzeln in einer Zeile ausgibt. Speichern Sie die Liste in der Datei 'wortliste.txt'.



MUSTERLÖSUNG 6-6

```
#!/usr/bin/python3
#Aufgabe 6-6
#Autorin: Leonie Weißweiler
artikel = open('artikel.txt', 'r')
wortliste = open('wortliste.txt', 'w')
for line in artikel:
    for word in line.split(' '):
        word = word.strip()
        wortliste.write(word)
        wortliste.write('\n')
artikel.close()
wortliste.close()
```



MUSTERLÖSUNG 6-7

Erweitern Sie das vorherige Programm dahingehend, das auch die Anzahl der Wörter, der Zeilen und die Anzahl der Zeichen (keine Spaces mitzählen!) in der Datei 'artikel.txt' gezählt werden.



MUSTERLÖSUNG 6-7

```
#!/usr/bin/python3
#Aufgabe 6-7
#Autorin: Leonie Weißweiler
artikel = open('artikel.txt', 'r')
wortliste = open('wortliste.txt', 'w')
zeilen = 0
wörter = 0
zeichen = 0
```



MUSTERLÖSUNG 6-7

```
for line in artikel:
    zeilen = zeilen + 1
    for word in line.split(' '):
        wörter = wörter + 1
        word = word.strip()
        zeichen = zeichen + len(word)
        wortliste.write(word)
        wortliste.write('\n')
print('Anzahl Zeilen:', zeilen)
print('Anzahl Wörter:', wörter)
print('Anzahl Zeichen:', zeichen)
artikel.close()
wortliste.close()
```



MUSTERLÖSUNG 6-8

Vergleichen Sie das Ergebnis der letzten Aufgabe mit der Ausgabe des UNIX-Befehls `wc`.



MUSTERLÖSUNG 6-8

Vergleichen Sie das Ergebnis der letzten Aufgabe mit der Ausgabe des UNIX-Befehls `wc`.

Das Ergebnis der Musterlösung beinhaltet weniger Worte und Zeichen. Mögliche Ursachen können beispielsweise sein, dass Newlines und Leerzeichen werden nicht mitgezählt werden.



MUSTERLÖSUNG 6-9

Ändern Sie das vorherige Programm dahingehend, dass Sie am Anfang des Programms nach einem Wort gefragt werden und ausgegeben wird, ob und wie oft das Wort in der Datei 'artikel.txt' vorkommt.



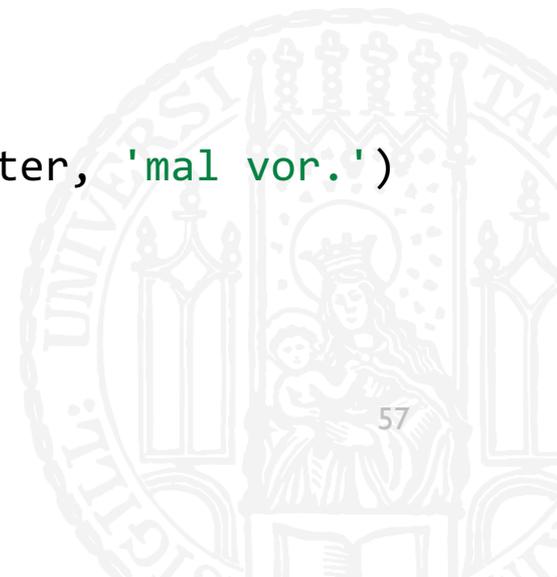
MUSTERLÖSUNG 6-9

```
#!/usr/bin/python3
#Aufgabe 6-9
#Autorin: Leonie Weißweiler
artikel = open('artikel.txt', 'r')
wortliste = open('wortliste.txt', 'w')
benutzerwort = input('Geben Sie ein Wort ein\n')
benutzerwörter = 0
```



MUSTERLÖSUNG 6-9

```
for line in artikel:
    for word in line.split(' '):
        word = word.strip()
        wortliste.write(word)
        wortliste.write('\n')
        if (benutzerwort.lower() == word.lower()):
            benutzerwörter = benutzerwörter + 1
print('Das Wort', benutzerwort, 'kam', benutzerwörter, 'mal vor.')
artikel.close()
wortliste.close()
```



MUSTERLÖSUNG 6-10

Vergleichen Sie das Ergebnis der letzten Aufgabe mit den Werten, die das Vorgehen aus Aufgabe 5-I liefert.



MUSTERLÖSUNG 6-10

Vergleichen Sie das Ergebnis der letzten Aufgabe mit den Werten, die das Vorgehen aus Aufgabe 5-1 liefert.

Die Ergebnisse sind identisch



MUSTERLÖSUNG 6-11

Wie sind folgende Zeichenketten in UTF-8 - und in ISO-Latin-1 Kodierung abgespeichert? (Tipp: Speichern Sie die Wörter in einer UTF-8 und dann in einer ISO-Latin-1 Datei und betrachten Sie, analog zur Vorlesung, den Hex-Dump der Datei mit dem unix-Befehl `hd`)

- a) Zeichenkette : 'Anton'
- b) Zeichenkette : 'Dez. 2010, 12:45,
- c) Zeichenkette : 'äää üüü ööö,

Worin unterscheiden sich die Codierungen?



MUSTERLÖSUNG 6-11

```
Leonie@Laptop $ hd worte-iso.txt
0000000 41 6e 74 6f 6e 0a 44 65 7a 2e 20 32 30 31 30 2c
0000010 20 31 32 3a 34 35 0a e4 e4 e4 20 fc fc fc 20 f6
0000020 f6 f6 0a
```

```
Leonie@Laptop $ hd worte-utf.txt
0000000 41 6e 74 6f 6e 0a 44 65 7a 2e 20 32 30 31 30 2c
0000010 20 31 32 3a 34 35 0a c3 a4 c3 a4 c3 a4 20 c3 bc
0000020 c3 bc c3 bc 20 c3 b6 c3 b6 c3 b6 0a
```

MUSTERLÖSUNG 6-12

Geben Sie Ziffernfolge der folgenden Zahlen im Binär-, Hexadezimal-, Dezimal- und Oktalsystem an

- a) Dezimalzahl 16
- b) Oktalzahl 777
- c) Hexadezimalzahl FF
- d) Dezimalzahl 512



MUSTERLÖSUNG 6-12

Basis	2	8	10	16
a)	10000	20	16	10
b)		777	511	1FF
c)		377	255	FF
d)	1000000000	1000	512	100

