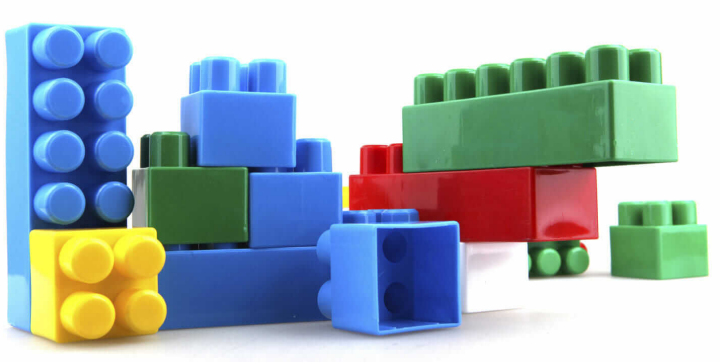
9. Modularisierung

Hochschule RheinMain



https://redshift.autodesk.com/the-benefits-of-modular-construction/

Notizen

Modularisierung in der Küche

Hochschule RheinMain

Die Herstellung von Apfelkuchen ist die eigentliche Aufgabe. Das ist das **Hauptprogramm**.
Die Herstellung von Hefeteig ist eine Teilaufgabe im Rahmen der Herstellung eines Apfelkuchens. Das ist eine **Funktion** oder ein **Unterprogramm**.
Das Starten der Aktivität „Hefeteig erstellen“ aus der Zubereitungsvorschrift von Apfelkuchen bezeichnen wir als einen Aufruf des Unterprogramms aus dem Hauptprogramm. Wir sprechen von einem **Unterprogrammaufruf** oder einem **Funktionsaufruf**.
Zwischen Haupt- und Unterprogramm müssen beim Aufruf ganz bestimmte Informationen fließen, z. B. darüber, wie viel Hefeteig zu erstellen ist und ob dem Teig Zucker zugesetzt werden soll.

Apfelkuchnrezept

600 g	Hefeteig
1,5 kg	Äpfel
...	

Zubereitung

Bereiten Sie den Hefeteig nach Rezept zu und rollen diesen dann auf einer bemehlten Arbeitsfläche quadratisch aus. Geben Sie den Teig dann auf ein mit Backpapier ausgelegtes Blech und ziehen den Rand an jeder Seite hoch. Der Teig kann dann mit einem Küchentuch abgedeckt noch ein wenig stehen bleiben. In der Zwischenzeit die Äpfel schälen, entkernen und in schmale Spalten schneiden. ...

Über den Austausch dieser Informationen muss zwischen Haupt- und Unterprogramm eine präzise Vereinbarung bestehen. Das Hauptprogramm muss wissen, welche Informationen das Unterprogramm benötigt und welche Ergebnisse es produziert. Eine solche Vereinbarung nennen wir eine **Schnittstelle**.
Eine im Rahmen der Schnittstelle vereinbarte Einzelinformation, wie z.B „Zuckerzugabe in Gramm“, nennen wir einen **Parameter**. Alle Parameter zusammen beschreiben die Schnittstelle. Ein Parameter, durch den Informationen vom Hauptprogramm zum Unterprogramm fließen, bezeichnen wir als **Eingabeparameter**. Einen Parameter, durch den Informationen vom Unterprogramm zum Hauptprogramm zurückfließen, bezeichnen wir als **Rückgabeparameter**.
Konkrete, durch die Parameter der Schnittstelle fließende Daten (z. B. 100 Gramm Zuckerzugabe) bezeichnen wir als **Parameterwerte**. Entsprechend der Flussrichtung bezeichnen wir die Parameterwerte auch als **Eingabewerte** oder **Rückgabewerte**.

Notizen

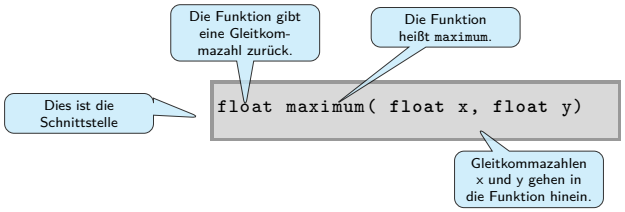
Funktionsschnittstelle

Hochschule RheinMain

An der **Schnittstelle** einer Funktion werden alle Informationen festgelegt, die zwischen Haupt- und Unterprogramm fließen.

Wenn wir eine Funktion erstellen wollen, die das Maximum von zwei Gleitkommazahlen bestimmen soll, so fließen an der Schnittstelle die folgenden Informationen:

- In die Funktion hinein fließen zwei `float`-Zahlen (Eingabeparameter), von denen die größere zu bestimmen ist.
- Aus der Funktion heraus fließt eine `float`-Zahl (Rückgabeparameter), nämlich die größere der beiden hinein geflossenen Zahlen.



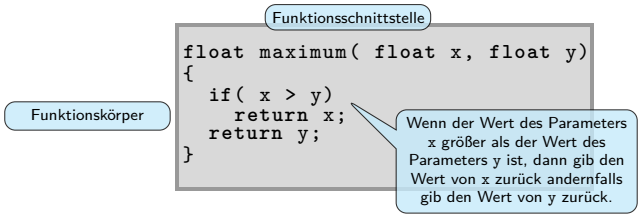
Wie die Funktion ihre Aufgabe erledigt, interessiert an der Schnittstelle nicht. Es geht nur um die Informationen, die sie benötigt , um ihre Aufgabe zu erledigen und die Informationen, die sie nach Erledigung der Aufgabe zurückgibt.

Notizen

Funktionskörper

Hochschule RheinMain

Im **Funktionskörper** wird festgelegt, wie die Funktion ihre Aufgabe erledigt. Der Funktionskörper enthält den Quellcode, der den Ablauf der Funktion festlegt:



Eine `return`-Anweisung beendet die Funktion und gibt den Wert des hinter `return` stehenden Ausdrucks an das rufende Programm zurück.

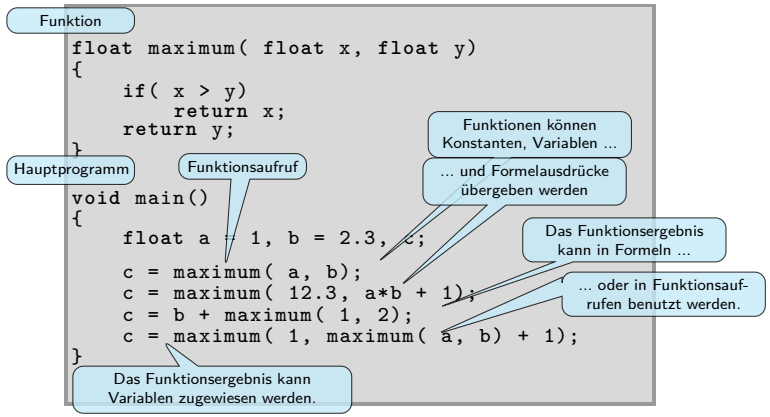
Der Ergebnistyp des Ausdrucks (hier `float`) muss zum Rückgabetyp der Funktionsschnittstelle passen.

Notizen

Funktionsaufruf

Hochschule RheinMain

Funktionen können aus dem Hauptprogramm oder aus anderen Funktionen gerufen werden. Die aufrufende Funktion übergibt dazu die erforderlichen Parameter und erhält den von der Funktion ermittelten Rückgabewert:




```
float maximum( float x, float y)
{
    if( x > y)
        return x;
    return y;
}

void main()
{
    float a = 1, b = 2.3, c;

    c = maximum( a, b);
    c = maximum( 12.3, a*b + 1);
    c = b + maximum( 1, 2);
    c = maximum( 1, maximum( a, b) + 1);
}
```

Anzahl und Typen der beim Aufruf übergebenen Parameter und die Verwendung des Rückgabetyps müssen mit der Schnittstellenvereinbarung übereinstimmen.

Notizen

Unterschiedliche Parameter und Returnwerte

Hochschule RheinMain

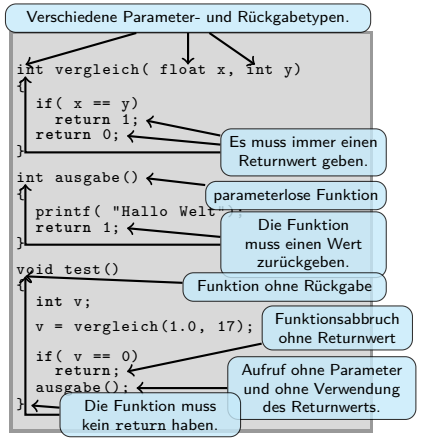
Eine Funktion kann unterschiedliche Parametertypen haben oder sogar parameterlos sein. Eine Funktion kann auch ohne Rückgabewert sein und erhält in diesem Fall den Rückgabetyp `void`:

Hat eine Funktion einen Returntyp, muss überall dort, wo der Kontrollfluss der Funktion endet, ein zum Typ passender Wert zurückgegeben werden.

Wert zurückgegeben werden.Hat eine Funktion keinen Returntyp, kann trotzdem eine `return`-Anweisung (ohne Returnwert) verwendet werden, um den Kontrollfluss zu beenden und unmittelbar zum rufenden Programm zurückzukehren.

An der Schnittstelle vereinbarte Parameter müssen vom aufrufenden Programm mit dem korrekten Typ übergeben werden.

Der Rückgabewert muss vom aufrufenden Programm nicht entgegen genommen und auch nicht unbedingt verwendet werden.



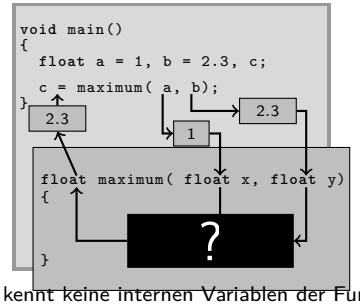

```
int vergleich( float x, int y)
{
  if( x == y)
    return 1;
  return 0;
}

int ausgabe()
{
  printf( "Hallo Welt");
  return 1;
}

void test()
{
  int v;

  v = vergleich(1.0, 17);

  if( v == 0)
    return;
  ausgabe();
}
```

Notizen

Das Blackbox-Prinzip

Hochschule RheinMain

Eine Funktion ist eine „doppelte Blackbox“, in die von außen niemand hineinsehen, aus der aber auch von innen niemand heraussehen kann. Von innen und außen sieht man nur die Schnittstelle.



Das aufrufende Programm kennt keine internen Variablen der Funktion. Umgekehrt kennt auch die Funktion keine Variablen des Hauptprogramms. Eine zufällige Gleichheit von Variablennamen ändert daran nichts. Weder das Hauptprogramm kann auf Variablen des Unterprogramms noch das Unterprogramm kann auf Variablen des Hauptprogramms zugreifen.

An der Schnittstelle werden **Kopien** der Parameterwerte übergeben. Änderungen dieser Werte haben keine Auswirkungen auf das rufende Programm.

Notizen

Technischer Ablauf: Funktionsaufruf

Hochschule RheinMain

Nötige Schritte

- Parameter**kopien** für aufgerufene Funktion hinterlegen
- Programmsteuerung an die Prozedur übergeben
- Arbeitsspeicher für Prozedur bereitstellen (lokale Variablen)
- Prozedur ausführen
- Ergebnis an den Aufrufer übergeben
- Arbeitsspeicher der Prozedur wieder freigeben
- zum Aufrufer zurückkehren

Beispiel: Funktionsaufruf

```
int funktion(int p1, int p2, ..)
{
  int lokal1, lokal2, ...;
  int ergebnis;
  .......
  return(ergebnis);
}

main()
{
  int a, b, c, ...;
  int resultat;
  ......
  resultat = funktion(a, b, ...);
  ......
}
```


Notizen

Technisch: Funktionsaufruf

Hochschule RheinMain

Beispiel Intel x86 (32-bit)

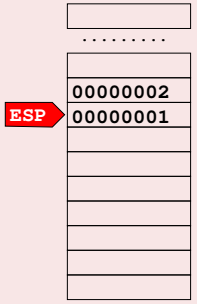
Argumentkopien auf den Stack

Call → Rücksprungadresse auf den Stack

Beispiel: Funktionsaufruf

```
int proz(int a, int b)
{
  return(a+b);
}
.....
rw = proz(1, 2);
.....
```

```
proz:
pushl ebp
movl esp,ebp
movl 12(ebp),eax
addl 8(ebp),eax
popl ebp
ret
.....
pushl #2
pushl #1
EPC call proz
rueckadr:
move eax,rw(ebp)
add#8,esp
```



Notizen

Technisch: Funktionsaufruf

Hochschule RheinMain

Beispiel Intel x86 (32-bit)

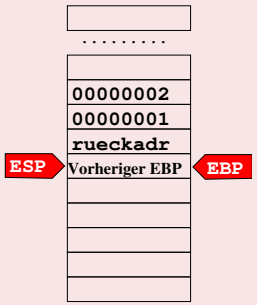
„Framepointer" (EBP) retten

Neuen Framepointer laden

Beispiel: Funktionsaufruf

```
int proz(int a, int b)
{
  return(a+b);
}
.....
rw = proz(1, 2);
.....
```

```
proz:
pushl ebp
EPC movl esp,ebp
movl 12(ebp),eax
addl 8(ebp),eax
popl ebp
ret
.....
pushl #2
pushl #1
call proz
rueckadr:
move eax,rw(ebp)
add#8,esp
```



Notizen

Technisch: Funktionsaufruf

Hochschule RheinMain

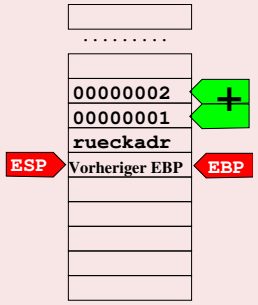
Beispiel Intel x86 (32-bit)

- Parameter addierennor
- Returnwert in Register (hier: EAX)

Beispiel: Funktionsaufruf

```
int proz(int a, int b)
{
  return(a+b);
}
.....
rw = proz(1, 2);
.....
```

```
proz:
pushl ebp
movl esp,ebp
movl 12(ebp),eax
addl 8(ebp),eax
popl ebp
ret
.....
pushl #2
pushl #1
call proz
rueckadr:
move eax,rw(ebp)
add#8,esp
```



Notizen

Technisch: Funktionsaufruf

Hochschule RheinMain

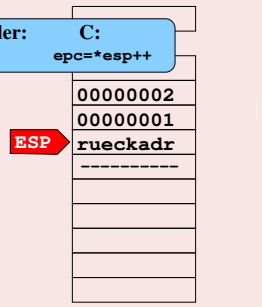
Beispiel Intel x86 (32-bit)

- Alten Framepointer wiederherstellen
- Rückkehr zum Aufrufer

Beispiel: Funktionsaufruf

```
int proz(int a, int b)
{
  return(a+b);
}
.....
rw = proz(1, 2);
.....
```

```
proz:
pushl ebp
movl esp,ebp
movl 12(ebp),eax
addl 8(ebp),eax
popl ebp
ret
.....
pushl #2
pushl #1
call proz
rueckadr:
move eax,rw(ebp)
add#8,esp
```



Notizen

Technisch: Funktionsaufruf

Hochschule RheinMain

Beispiel Intel x86 (32-bit)

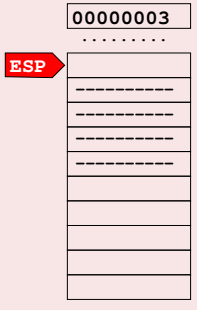
Ergebnis Speichern

Stack abräumen

Beispiel: Funktionsaufruf

```
int proz(int a, int b)
{
  return(a+b);
}
.....
rw = proz(1, 2);
.....
```

```
proz:
pushl ebp
movl esp,ebp
movl 12(ebp),eax
addl 8(ebp),eax
popl ebp
ret
.....
pushl #2
pushl #1
call proz
rueckadr:
move eax,rw(ebp)
add#8,esp
.....
```

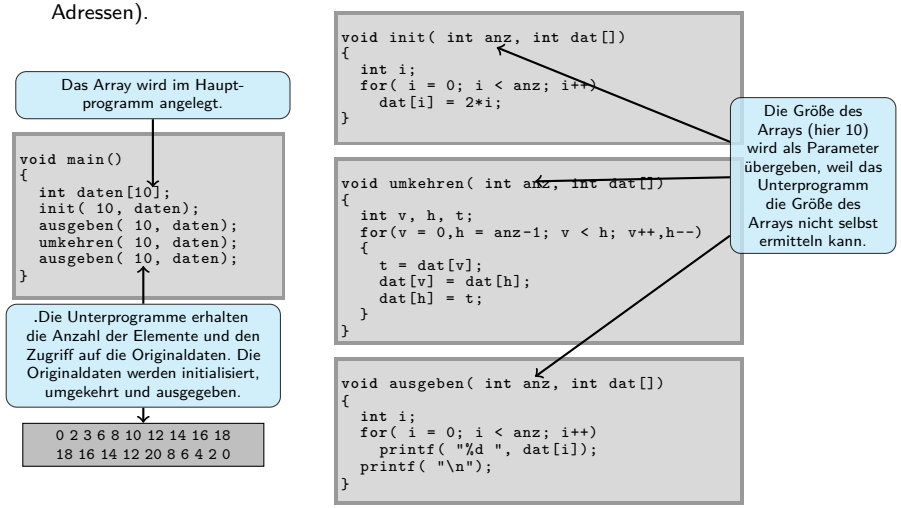


Notizen

Arrays als Parameter

Hochschule RheinMain

Wird ein Array an eine Funktion übergeben, so erhält die Funktion Zugriff auf die Originaldaten des Hauptprogramms. Es entsteht keine Kopie des gesamten Arrays sondern nur eine Kopie der Zugriffsinformation (Dazu mehr im Abschnitt über Zeiger und Adressen).



Notizen

Strings als Funktionsparameter

Hochschule RheinMain

Strings sind Arrays und werden daher an der Funktionsschnittstelle wie Arrays behandelt:

```
void main()
{
  int l, v;
  l = stringlaenge( "qwert");
  printf( "Laenge: %d\n", l);
  v = stringvergleich( "qwert", "qwerz");
  if( v == 1)
      printf( "gleich\n");
  else
    printf( "ungleich\n");
}
```

```
Laenge: 5
ungleich
```

```
int stringlaenge( char s[])
{
  int i;
  for( i = 0; s[i] != 0; i++)
    ;
  return i;
}
```

```
int stringvergleich( char s1[], char s2[])
{
  int i;
  for( i = 0; (s1[i]!=0)&&(s1[i]==s2[i]); i++)
    ;
  return s1[i] == s2[i];
}
```

Wegen des Terminatorzeichens kann das Unterprogramm das Ende des Strings ermitteln. Es muss daher keine Längeninformation mitgegeben werden. Verfügt das rufende Programm über die Länge des Strings, so kann diese mitgegeben werden, damit die Länge im Unterprogramm nicht erneut berechnet werden muß.

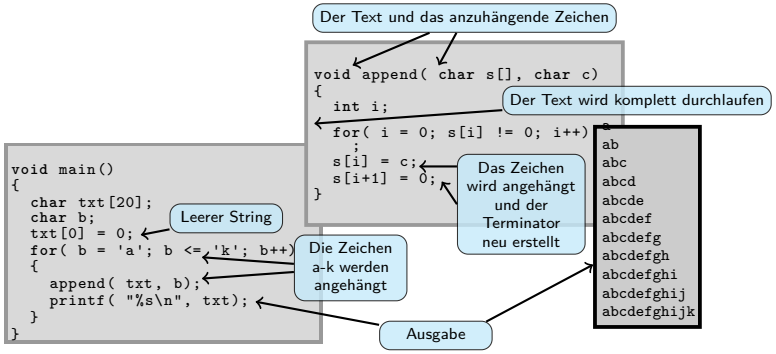
Wird der String im Unterprogramm verändert, sollte die Puffergröße als Parameter mitgegeben werden, um Pufferüberschreitungen (Buffer Overflow) vermeiden zu können.

Notizen

Strings als Funktionsparameter

Hochschule RheinMain

Strings können (wie Arrays) in Funktionen verändert werden. Zum Beispiel können Zeichen an einen String angehängt werden:



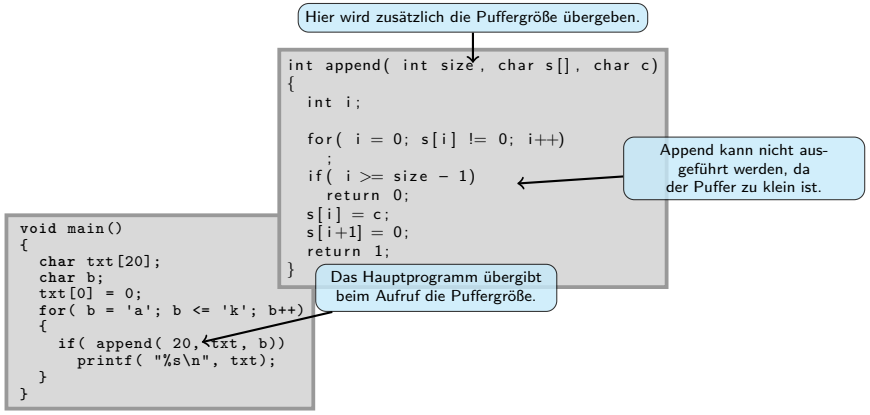
Vorsicht bei Veränderungen von Strings im Unterprogramm.
Programmierende sind dafür verantwortlich, dass keine Arraygrenzen überschritten werden und der String stets korrekt terminiert ist.
Wird im obigen Beispiel append() mehr als 19 mal gerufen, stürzt das Programm ab. Fehlende Sicherungen gegen Pufferüberschreitungen sind eine Hauptfehlerquelle in C-Programmen.

Notizen

Stringmanipulationen 2

Hochschule RheinMain

Um Pufferüberschreitungen abfangen zu können, sollte man Funktionen, die Strings verändern, die Puffergröße übergeben:



Nach wie vor kann append() nur 19 mal mit Erfolg gerufen werden, aber das Programm stürzt nicht mehr ab und meldet zurück, wenn der String nicht mehr vergrößert werden kann.

Notizen

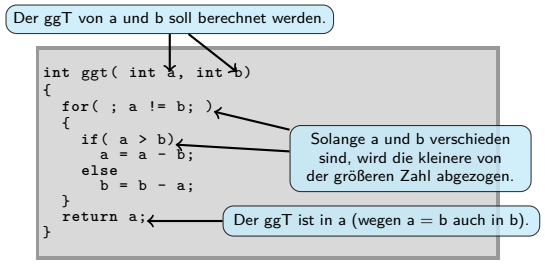
Bruchrechnung

Hochschule RheinMain

Im Folgenden werden Programmteile zum addieren und kürzen von Brüchen vorgestellt

Hier zunächst eine Hilfsfunktion zur Berechnung des größten gemeinsamen Teilers zweier Zahlen:

Den größten gemeinsamen Teiler (ggT) von zwei Zahlen erhält man, indem man solange die kleinere Zahl von der größeren abzieht bis beide Zahlen gleich sind.



Notizen

Bruchrechnung (2)

Hochschule RheinMain

Einen Bruch speichern wir mit Zähler und Nenner in einem Array mit zwei Elementen:

```
int bruch[2]; // bruch[0] ist der Zähler, bruch[1] ist der Nenner
```

Brüche werden gekürzt, indem Zähler und Nenner durch ihren ggT geteilt werden:

Der Bruch kommt als Array, Der Zähler hat den Index 0, der Nenner den Index 1.

```
void kuerzen( int b[])
{
  int gt;
  gt = ggt( b[0], b[1]);
  b[0] = b[0]/gt;
  b[1] = b[1]/gt;
}
```

Der ggT von Zähler und Nenner wird berechnet.

Zähler und Nenner werden durch den ggT dividiert.

Brüche werden nach der Formel $\frac{a}{b} + \frac{c}{d} = \frac{ad+cb}{bd}$ addiert, wobei anschließend gekürzt werden sollte:

Diese beiden Brüche sind zu addieren.

```
void addieren( int b1[], int b2[], int erg[])
{
  erg[0] = b1[0]*b2[1] + b2[0]*b1[1];
  erg[1] = b1[1]*b2[1];
  kuerzen( erg);
}
```

$\frac{a}{b} + \frac{c}{d} = \frac{ad+cb}{bd}$

Das Ergebnis wird gekürzt

Notizen

Bruchrechnung (3)

Hochschule RheinMain

Mit den Funktionen `kuerzen` und `addieren` kann ein kleines Programm zur Bruchrechnung geschrieben werden:

Drei Brüche

```
void main()
{
  int bruch1[2], bruch2[2], ergebnis[2];

  printf( "Bruch1: ");
  scanf( "%d/%d", &bruch1[0], &bruch1[1]);
  printf( "Bruch2: ");
  scanf( "%d/%d", &bruch2[0], &bruch2[1]);

  addieren(bruch1,bruch2,ergebnis);

  printf( "Ergebnis: %d/%d\n", ergebnis[0], ergebnis[1]);
}
```

```
Bruch1: 1/3
Bruch2: 1/6
Ergebnis: 1/2
```

ergebnis = bruch1 + bruch2

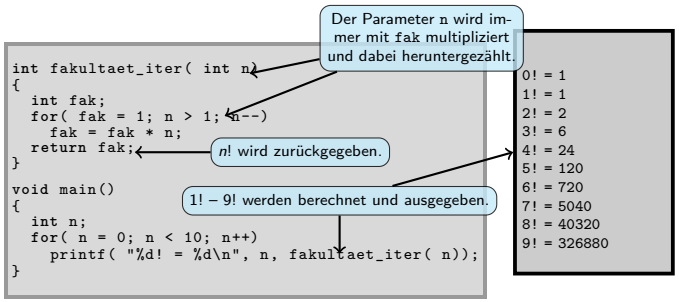
Notizen

Beispiel: Fakultätsfunktion

Hochschule RheinMain

Am Beispiel der Fakultätsfunktion wollen wir eines der wichtigsten Programmierprinzipien kennenlernen. Dazu implementieren wir diese Funktion zunächst in naheliegender Weise.

Die Fakultät $n! = 1 \cdot 2 \cdot 3 \ldots n$ kann **iterativ** in einer Schleife berechnet werden:



Notizen

Rekursion

Hochschule RheinMain

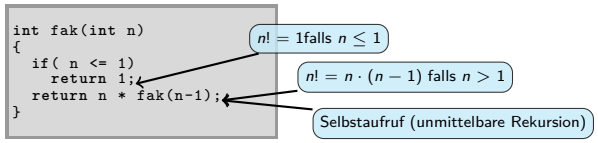
Funktionen können sich selbst unmittelbar oder mittelbar (über eine andere Funktion) aufrufen. Man nennt dies **Rekursion**. Rekursion ist eine einfache und elegante Lösungsstrategie, wenn man ein Problem auf ein oder mehrere strukturell gleiche, aber „kleinere“ Probleme zurückführen kann, bis man am Ende auf ein leicht zu lösendes Problem stößt.

Rekursive Definition der Fakultätsfunktion:

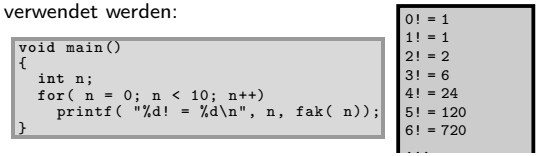
$$n! = \begin{cases} 1, & \text{falls } n \leq= 1 \\ n \cdot (n-1)!, & \text{falls } n > 1 \end{cases}$$

In dieser Definition führt man die Berechnung von $n!$ auf die Berechnung der kleineren Fakultät $(n-1)!$ zurück, bis man am Ende auf ein leicht zu lösendes Problem (Berechnung von 1!) stößt.

Die Definition führt unmittelbar zu einer rekursiven Implementierung der Fakultätsfunktion:



An der Schnittstelle ist diese Funktion identisch mit der nicht rekursiven Variante und kann daher exakt gleich verwendet werden:



Notizen

Rekursion aus Maschinensicht

Hochschule RheinMain

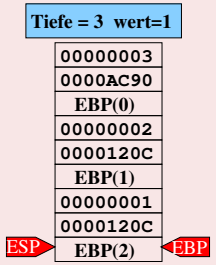
Im Beispiel hier: Intel x86 (32-bit)

Beispiel: Rekursiver Prozeduraufruf

(Gezeigt ist jeweils der Zustand bei Erreichen der mit STOP markierten Zeile)

```
int fak(int n)
{
if(n <= 1)
  return 1;
 else
 return n * fak(n - 1);
}
.....
f3 = fak(3);
....
```

```
0100 fak: pushl ebp
0101      movl esp, ebp
          ..........
1208      call fak
120C      move eax,..
          ..........
135       ret
          ..........
AC80      push #3
AC8E      call fak
AC9       move eax,..
```



Notizen

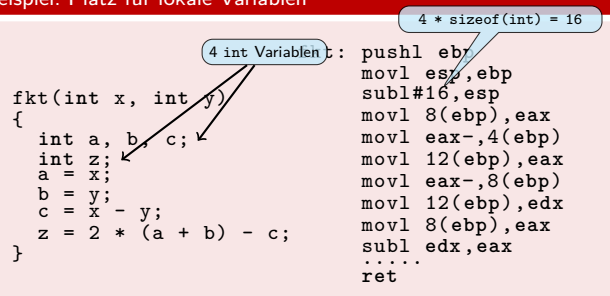
Lokale Variablen

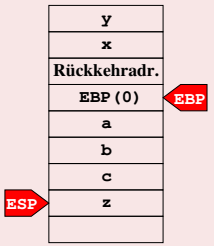
Hochschule RheinMain

Im vorigen Beispiel gab es außer den Funktionsparametern keine lokalen Variablen

Beobachtung: Stackpointer (ESP) und Framepointer (EBP) waren stets gleich
(Wozu dann überhaupt ein Framepointer?)

Beispiel: Platz für lokale Variablen





→ Lokale Variablen liegen zwischen EBP und ESP.

Jede Instanz der Funktion besitzt eigene lokale Variablen.

Lokale Variablen liegen im Stack-Speicher (gemeinsam mit Rückkehradresse, geretteten Registerinhalten, etc.).

Notizen

Dynamische lokale Variablen

Hochschule RheinMain

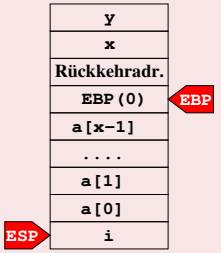
Im vorigen Beispiel war die Gesamtgröße der lokalen Daten konstant, deshalb konnte zur Allokation ein fester Betrages zum Stackpointer addiert werden

Seit dem C99 Standard gibt es jedoch auch lokale Felder *variabler* Größe:

Beispiel: variable Feldgröße

Die Array-Größe wird durch den Wert der Variablen x festgelegt

```
fkt(int x, int y)
{
  int a[x];
  int i;

  for(i = 0; i < x; i++)
    a[i] = ....;
  ....
}
```



→ Feldgröße kann von Instanz zu Instanz variieren

→ Adressabstand zwischen Framepointer und Variablen (hier z.B. `i`) ist *nicht* konstant.

Notizen

Pufferüberlauf-Attacke

Hochschule RheinMain

Vorsicht: Indizes werden in C nicht zur Laufzeit geprüft

Zum Beispiel kann über `a[x+1]` auf die Rückkehradresse der Funktion zugegriffen werden.

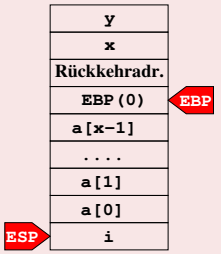
→ Mit `a[x+1] = Zieladresse` kann die Rückkehradresse verändert werden.

→ Statt zum Aufrufer zurückzukehrem, kann ein beliebiges Stück (womöglich in Form von Daten eingeschleusten) Schadcodes angesprungen werden.

Beispiel: Gefahr von Pufferüberlauf

```
fkt(int x, int y)
{
  int a[x];
  int i;

  for(i = 0; ; i++)
    a[i] = untrusted_data();
  ....
}
```

Fehler: Keine Begrenzung auf `i < x`



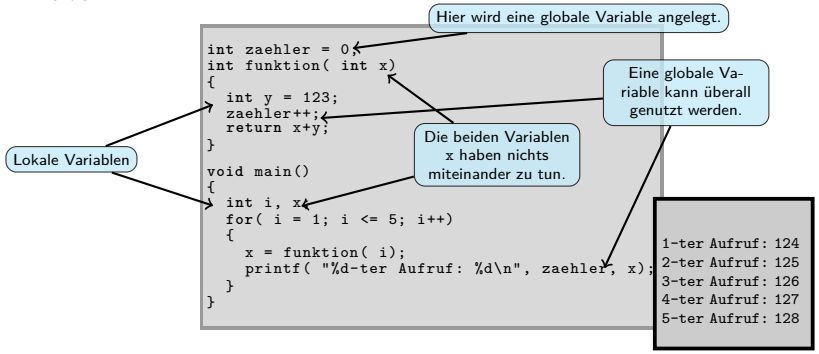
Keine spezielle Eigenschaft dynamischer Arrays: Diese Angriffsmöglichkeit besteht in gleicher Weise auch bei Arrays fester Größe.

Notizen

Globale Variablen

Hochschule RheinMain

Variablen außerhalb von Funktionen sind global und können funktionsübergreifend genutzt werden:



Globale Variablen sollten nur sparsam verwendet werden, da sie die Schnittstellen unterlaufen (Seiteneffekt). Verwenden Sie eine globale Variable nur für ein Datum, das programmweit überall mit der gleichen Bedeutung verwendet wird.

Während lokale Variablen mit jedem Funktionsaufruf eine eigene Instanz erhalten, existieren globale (statische) Variablen nur ein mal. Rekursive (und allgemeiner: *reentrante*) Funktionen sollten deshalb i.d.R. keine solchen Variablen benutzen.

Notizen

Notizen