Datenstrukturen in C

- Aus den einfachen Datentypen lassen sich komplexere konstruieren
- ► Array und struct: zusammengesetzte Datentypen Arvuy & Stvu & Stvu & Stvut &
- (Speicher-)Objekt: Speicherbereiche eines bestimmten Typs
 - Variablen, Parameter, etc.

sizeof

Der sizeof-Operator

- wird wie eine Funktion genutzt
- rmittelt die Größe des als Argument übergebenen Datentyps in Byte
- ► kann auch auf Variablen aufgerufen werden
- ermittelte Werte plattformabhängig (Ausnahme: sizeof(char) == 1)

```
1 size_t size;
2 size = sizeof(char);   // --> 1
3 size = sizeof(size_t);   // --> 8
4 size = sizeof(void *);   // --> 8
5
6 size_t a = sizeof(size_t);
7 size_t b = sizeof(a);
8 // Nun gilt: a == b, da 'a' vom Typ 'size_t' ist.
(die Größenangaben sind Beispielwerte auf einem üblichen LP64-System)
```

Alignment

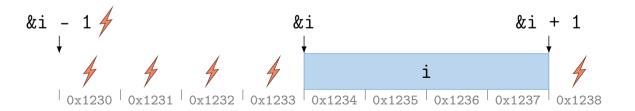
- Anforderung an Ausrichtung der Speicheradresse eines Objektes Speicheradresse muss ein Vielfaches vom Alignment sein
- Abhängig vom Datentyp
- ► Größe eines Datentyps ist Vielfaches des Alignments
- ► Beispiel: int/
 - ► Größe: 4 Byte
 - ► Alignment entweder 1, 2, oder 4 Byte
 - ▶ Üblicherweise 4 Byte Alignment

Pointer

- Pointer: Adresse eines Speicherobjektes, "zeigt" auf das Objekt
- Arrays: zusammenhängender Speicherbereich, enthält Datenobjekte gleichen Typs direkt aufeinanderfolgend
- ► Arrays können ebenfalls über Pointer genutzt werden!
- Operator &: nimmt Adresse eines Objekts
- Operator *: Dereferenzieren eines Pointers (Speicherzugriff)

```
1 int i = 0;
2 int* i_ptr = &i;
3 int i_new = *i_ptr;  // i_new == i
```

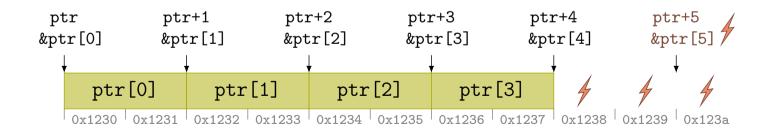
Pointerarithmetik



Pointerarithmetik: Arraysubscript

Valide Pointer zeigen immer auf...
...ein Objekt (z.B. eine Variable),
...eine Stelle in einem Array, oder
...an das Ende eines Arrays

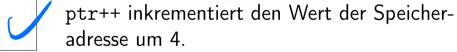
(dann nicht dereferenzierbar)
Andernfalls: Verhalten undefiniert!

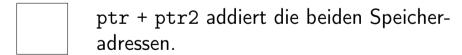


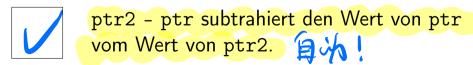
Quiz: Pointer und Pointerarithmetik (1)

Seien ptr und ptr2 vom Typ int* und sizeof(int) == 4. Welche Aussagen treffen zu?



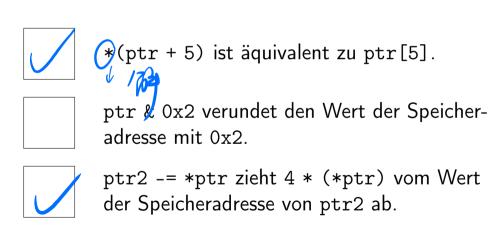






Quiz: Pointer und Pointerarithmetik (2)

Seien ptr und ptr2 vom Typ int* und sizeof(int)== 4. Welche Aussagen treffen zu?



void-Pointer und implizite Pointer-Casts

- ► void-Pointer nicht dereferenzierbar → 国本代本政府

 Dekrementieren oder International
- Dekrementieren oder Inkrementieren eines void-Pointers ist undefiniert
- Casting zw. void-Pointers und anderem Pointer-Typ ist implizit

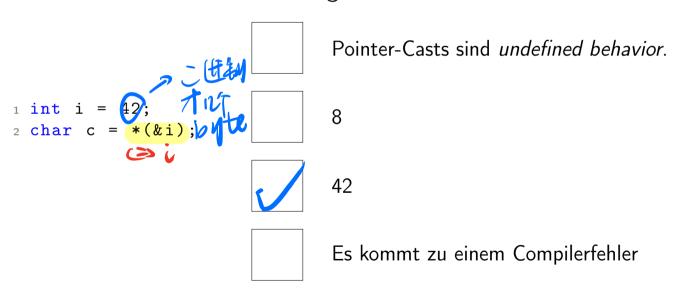
```
1 int* i_ptr = /* ... */;
void* v_ptr = i_ptr;
3 int* i_ptr2 = v_ptr;
```

Explizite Pointer-Casts

- Explizite Typumwandlung
- ► Neuer Datentyp darf keine strengeres Alignment fordern
- Derenferenzierung von umgewandelten Pointern: undefined behavior!
 (Einzige Ausnahme: char-Pointer)
- Explizite Casts daher vermeiden

Quiz: void-Pointer und Pointer-Casts (1)

Welchen Wert (dezimal) hat die Variable c am Ende des folgenden Codeauschnittes?



Quiz: void-Pointer und Pointer-Casts (2)

```
Welchen Wert (dezimal) hat die Variable c am Ende des folgenden Codeauschnittes?
        (Annahme: sizeof(int) == 4, CHAR_BIT == 8 und Little Endian)
                     int i = 123456789;
                     2 char* c_ptr = (char*) &i;
                     3 char c = c_ptr[2];
      Der Pointer-Cast ist undefined behavior.
      Es kommt zu einem Compilerfehler.
      Der Zugriff auf das dritte Element des Arrays, auf das c_ptr zeigt, ist un-
      defined behavior, da das Array nur ein Element enthält.
      Die Dereferenzierung von c_ptr ist undefined behavior, weil c_ptr aus
      dem Cast eines int-Pointers hervorging.
```

Quiz: void-Pointer und Pointer-Casts (3)

Welchen Wert (dezimal) hat die Variable i am Ende des folgenden Codeausschnittes? (Annahme: sizeof(int) == 4, CHAR_BIT == 8 und Little Endian)

```
char c = 42;
int* i_ptr = (int*) &c;
int i = *i_ptr;
```

- 8387882
- Es kommt zu einem Compilerfehler.
- 42
- Der Pointer-Cast ist *undefined behavior*.
- Die Pointer-Dereferenzierung ist *undefined behavior*.

Quiz: void-Pointer und Pointer-Casts (4)

Welchen Wert (dezimal) hat die Variable i am Ende des folgenden Codeausschnittes?

```
char c = 42;
void* v_ptr = &c;
int i = *v_ptr;
```

8387882



Es kommt zu einem Compilerfehler. 足怖性 名詞





42

Quiz: void-Pointer und Pointer-Casts (5)

Welchen Wert (dezimal) hat die Variable j am Ende des folgenden Codeausschnittes? (Annahme: sizeof(int) == 4, CHAR_BIT == 8 und Little Endian)

```
int i = 42;
void* v_ptr = &i;
int* i_ptr = v_ptr;
int j = *i_ptr;
```

- 8387882
- Es kommt zu einem Compilerfehler.
- Es fehlt ein expliziter Cast des int-Pointers zum void-Pointer.
- 42
- Die Pointer-Dereferenzierung ist *undefined behavior*.

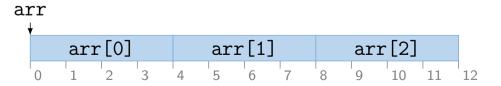
Arrays: Deklaration und Definition

- Speicherbereich mit Datenobjekten gleichen Typs direkt aufeinanderfolgend
- Zugriff: syntaktisch analog zu Pointern (Bsp: arr[0] oder *arr)
- Deklaration mit impliziter Größenangabe: int arr[3];
- ▶ Deklaration und Definition in einem Statement mittels "Compound Literal":

```
<u>int arr[3] = {1, 2, 3};</u>
```

Verzicht auf explizite Größenangabe:

```
int arr[] = {1, 2, 3};
```



Arrays: Größe

► Bestimmung der Größe durch den sizeof-Operator Achtung: gesamte Größe muss durch Elementgröße dividiert

```
1 int arr[3] = {1, 2, 3};

2 size_t arr_len = sizeof(arr) / sizeof(arr[0]); // --> 3

sizeof(arr[0]) = 4

sizeof(arr) = 12

arr[0] arr[1] arr[2]

0 1 2 3 4 5 6 7 8 9 10 11 12
```

variable Größe:

```
void func(size_t size) {
char buf[size];
// ...
}
```

Arrays als Parameter (1) int * arrs void func(int* arr, size_t arr_length) { for (size_t i = 0; i < arr_length; i++) {</pre> printf("%d\n", arr[i]); ₅ } 7 int main(void) { int arr $[3] = \{1,2,3\};$ func(arr, sizeof(arr) / sizeof(arr[0]));

10 // ...

11 }

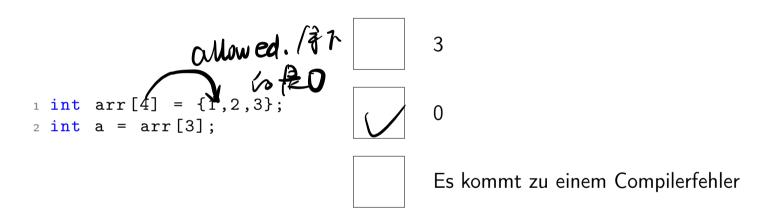
Achtung! sizeof von Array vs. sizeof von Pointer

Arrays als Parameter (2)

```
1 void func1(int arr[arr_length], size_t arr_length) {
      for (size_t i = 0; i < arr_length; i++) {</pre>
          printf("%d\n", arr[i]);
7 void func2(int arr[3]) {
      for (size_t i = 0; i < 3; i++) {</pre>
          printf("%d\n", arr[i]);
10
11 }
12 int main(void) {
      int arr [3] = \{1,2,3\};
13
      func1(arr, sizeof(arr) / sizeof(arr[0]));
14
15 func2(arr); // ...
16 }
```

Quiz: Arrays (1)

Was ist der Wert von a nach folgendem Codeauschnitt?



Quiz: Arrays (2)

Was ist der Wert von a nach folgendem Codeauschnitt?

```
1 int arr[3];
2 arr = (int [3]) {1,2,3};
3 int a = arr[1];
                                       Es kommt zu einem Compilerfehler
```

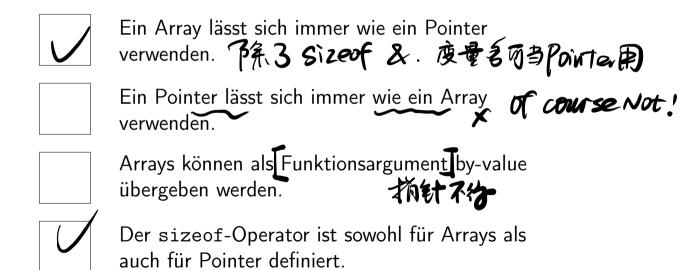
Quiz: Arrays (3)

Was ist der Wert von a nach folgendem Codeauschnitt?

```
1 int a = ((const int [3]) {1,2,3})[0];
                   newm感觉
        Es kommt zu einem Compilerfehler
```

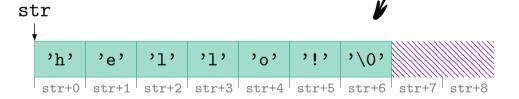
Quiz: Unterschiede zwischen Arrays und Pointern

Welche der folgenden Aussagen sind korrekt?



Strings

- Repräsentation als char-Array
- Länge implizit: 0-Byte als Terminal C: 0 かる
- Üblicherweise: ASCII-Zeichensatz bzw. -Kodierung (man 3 ascii)



Strings: Definitionsmöglichkeiten

Array mit explizitem null-Byte:
 char str[] = {'t', 'u', 'x', '\0'};

 mit Hexadezimalwerten (siehe man 3 ascii):
 char str[] = {0x68, 0x75, 0x68, 0x75, 0};

- Mittels eines String-Literals mit implizitem null-Byte:
- Char str[] = "tux";
 (äquivalent zur vorigen Variante)
- ▶ Über einen Pointer statt einem Array:
- (String hier nicht modifizierbar!)

Quiz: Strings (1)

Was ist der Unterschied zwischen den folgenden Möglichkeiten, einen String zu erstellen?

```
1 char str1[] = "tux";

2 char str2[4] = "tux";

3 char str3[] = {'t', 'u', 'x', '\0'};

4 char str4[4] = {'t', 'u', 'x', '\0'};
```



Mindestens eine Möglichkeit ist falsch.



Die ersten beiden Strings können nicht verändert werden.



Es gibt keinen Unterschied, das Resultat ist in allen Fällen identisch.

Quiz: Strings (2)

Welchen Wert haben len und size am Ende dieses Codeabschnittes?

```
len: 3, size: 3
1 char str[] = "tux";
                                            len: 3, size: 4
2 size_t len = strlen(str);
3 size_t size = sizeof(str);
                                            len: 4, size: 3
                                            len: 4, size: 4
```

Quiz: Strings (3)

Was könnte ein mögliches Problem bei folgendem String sein?

```
1 char str[3] = "tux";
```

Bei Strings darf es keine explizite Größenangabe für den Array geben.

Im resultierenden Array wird das terminierende null-byte fehlen.

Es gibt kein Problem.

Es kommt zum Compiler-Fehler.

Quiz: Strings (4)

Was könnte an folgendem Codeauschnitt problematisch sein?

```
1 char str[] = "tux"; √
2 char* str ptr = str; utux"
3 str_ptr[3] = 'i'; 7.76.
```

	Nichts,	alles	in	Ordnung.
--	---------	-------	----	----------

- Eine Modifikation eines Strings über einen Pointer ist immer *undefined behavior*.
- Der Zugriff auf das vierte Element des Arrays ist out of bounds, da der String nur aus drei Zeichen besteht.
- Das terminierende null-byte wird überschrieben.
- str_ptr zeigt auf die Array-Variable und nicht direkt auf den Array, somit ist der Zugriff semantisch falsch und zudem out of bounds.

Quiz: Strings (5)

Inwiefern könnte folgender Codeauschnitt problematisch sein?

```
char* str = "tux";
str[0] = '1';
```



Modifikationen von String Literals sind Undefined Behavior.



Gar nicht, alles in Ordnung.



Es kann zu einem Segmentation Fault kommen.



str sollte als const char* deklariert werden.



Auf einem Linux-System ist es nicht möglich, den String "tux" zu verändern.



Ein String kann nicht auf diese Weise erstellt werden, es wird ein Array benötigt, kein Pointer.

Quiz: Strings (6)

Welche Aussagen bezüglich des Codes sind korrekt?

```
1 void fn(/* ... */) {
2    char str1[] = "abcd";
3    const char* str2 = "abcd";
4    /* ... */
5 }
```

str2 zeigt auf den Stack.

str2 ermöglicht Speicherplatzoptimierungen.

Der Wert von str1 muss bei jedem Funktionsaufruf neu auf den Stack kopiert werden.

str2 kann auf str1 zeigen.

str2 ist zu bevorzugen, wenn der String nicht verändert wird.

Quiz: Strings (7)

Was ergibt der Ausdruck "1234" + 1? Einen Pointer auf den String "1235" (2) The Einen Pointer auf den String "12341" Einen Pointer auf den String "234" Den Integer 1235 Einen Compiler-Fehler

struct

- Zusammengesetzter Datentyp mit Elementen verschiedenen Typs
- ► Elemente liegen aufeinanderfolgend (ggf. mit Padding) im Speicher

Zugriff direkt oder über einen Pointer

```
penguin.age = 0;  // penguin ist vom Typ

typ

penguin_ptr

age = 0; // penguin_ptr ist vom Typ

// struct Penguin*
```

struct: Initialisierung

► Initialisierung über "Compound Literal":

```
this
this
struct Penguin penguin1 = { .age = 0, .name = "Tux" };

// Implizit: penguin2.name = NULL
struct Penguin penguin2 = { .age = 0 };

// Initialisierung mit festgelegter Reihenfolge
struct Penguin penguin3 = { 0, "Tux" };
```

Initialisierung aller Elemente zu 0:

```
1 struct Penguin penguin4 = { 0 };
```

struct als Parameter

```
1 struct Penguin {
      char* name;
      unsigned age;
4 };
5 void print_penguin_name1(struct
      printf("name: %s\n", penguin->name)
7 }
8 void print_penguin_name2(struct Penguin penguin
      printf("name: %s\n", penguin.name);
10 }
  int main(void) {
      struct Penguin penguin = { "tux", 5 };
12
      print_penguin_name1(&penguin);
13
     print_penguin_name2(penguin);
14
15 }
```

Quiz: struct (1)

Was ist der Wert (dezimal) von size nach folgendem Codeauschnitt? (Wir nehmen an, dass sizeof(int) == 4

```
und dass ein int ein 4-Byte Alignment erfordert.)
1 struct Penguin {
     char id_bytes[5];
                                                          15
      int id;
5 size_t size = sizeof(struct Pengual);
                                                          12
                                                          16
```

Quiz: struct (2)

Ist am Ende dieses Codeauschnitts der Wert von size1 gleich dem von size2? (Wir nehmen an, dass sizeof(int) == 4 upd dass ein int ein 4-Byte Alignment

```
图 4+2+21分了, $P3让Struct也满足)
1 struct Penguin1 {
      int id;
                                         Nein
    unsigned char age;
     char color;
                                          Ja
6 struct Penguin2 {
      unsigned char age;
      int id;
                                         Es kommt zu einem Compilerfehler
     char color;
size_t ize1 = sizeof(struct Penguin1);
12 size_t ($\frac{1}{2}ze2 = sizeof(struct Penguin2);
```

Code zu Quiz: struct (3)

```
1 struct Penguin {
      char *name; unsigned age;
3 };
4 struct Penguin * create_penguin1(void) {
      struct Penguin penguin = { .name = "tux", .age = 5 };
     return &penguin;
8 struct Penguin create_penguin2(void) {
      struct Penguin penguin = { .name = "tux", .age = 5 };
9
      return penguin;
10
11 }
12 int main(void) {
      struct Penguin *penguin1 = create_penguin1();
13
      struct Penguin penguin2 = create_penguin2();
14
      printf("name1: %s, name2: %s\n", penguin1->name, penguin2.
15
        name);
16 }
```

Quiz: struct (3)

Was ist der Unterschied zwischen den beiden auf der vorigen Folie zu sehenden Ansätzen, ein struct als Rückgabewert zu nutzen?



Es gibt keinen Unterschied.



Variante 1 führt zu einem Segmentation Fault.



Variante 2 führt zu einer Compilerwarnung, da ein struct nur über einen Pointer zurückgegeben werden kann.

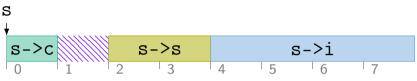


Variante 1 führt zu einer Compilerwarnung.



struct vs. union

struct speichert immer alle Elemente
struct { char c; short s; int i; }* s;
s



union speichert genau ein Element - das zuletzt geschriebene
union { char c; short s; int i; }* u;

struct mit union

Kombination mit struct mit Indikator für Gültigkeit der union-Elemente:

```
1 struct Dimension { /* ... */ };
2 struct Shape {
      int shape_kind; /* 1 = circle, 2 = rect */
4 union {
int circle_radius;
struct Dimension rect;
7 };
8 };
11
12 struct Shape my_circ =
          snape my_circ -
    { .shape_kind = 1, { .circle_radius = 10 }};
```

union als Parameter

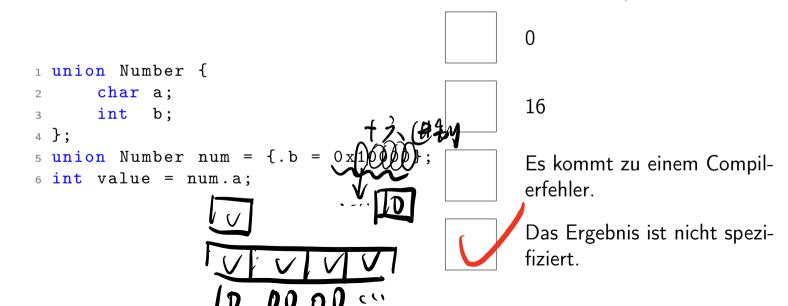
```
void f(union Number num) {
    // Zugriff auf das Attribut 'a' mit 'num.a' ...

void g(union Number *num) {
    // Zugriff auf das Attribut 'a' mit 'num->a' ...
}
```

Zugriff auf die Member einer union syntaktisch analog zu struct

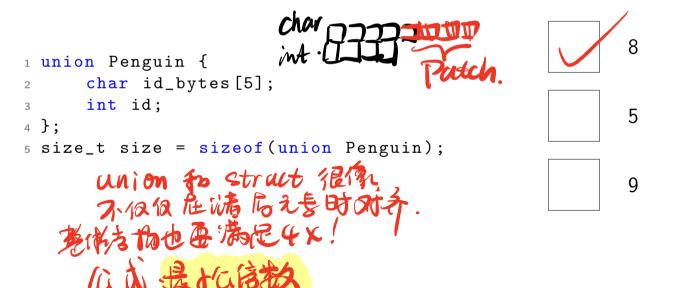
Quiz: union (1)

Was ist der Wert (dezimal) von value nach folgendem Codeauschnitt? (Wir nehmen an, dass sizeof(int) == 4 und CHAR_BIT == 8.)



Quiz: union (2)

Was ist der Wert (dezimal) von size nach folgendem Codeauschnitt? (Wir nehmen an, dass sizeof (int) == 4 und dass ein int ein 4-Byte Alignment erfordert.)



Quiz: union (3)

Was ist der Wert (dezimal) von size nach folgendem Codeauschnitt?

```
1 union Penguin {
     char id_short[5];
     char id_long[7];
4 };
5 size_t size = sizeof(union Penguin);
                                                         35
```

Verschachtelte Datentypen

Mehrdimensionale Arrays - Iterationsreihenfolge

```
1 unsigned matrix[3][4] = {
      { 1, 2, 3, 4 }, { 5, 6, 7, 8 }, { 9, 10, 11, 12 }
3 };
5 for (size_t i = 0; i < sizeof(matrix) / sizeof(*matrix); i++) {</pre>
      for (size_t j = 0;
            j < sizeof(matrix[i]) / sizeof(*matrix[i]);</pre>
            j++) {
           printf("%u ", matrix[i][j]);
9
10
11 }
12
13 printf("\n");
```

Verschachtelte Datentypen

Mehrdimensionale Arrays – implizite Größe

Auslassen der expliziten Größenangabe der "obersten Ebene":

Quiz: Mehrdimensionale Arrays (1)

Was ist der Wert von size nach folgendem Codeauschnitt?

```
1 char matrix[8][9];
2 size_t size = sizeof(matrix);
                                             Es kommt zu einem Compiler-
                                             fehler.
```

Quiz: Mehrdimensionale Arrays (2)

Was ist der Wert von size nach folgendem Codeauschnitt?

```
1 char matrix[5][];
2 size_t size = sizeof(matrix / *matrix);
          20 Précha matrix [][5]!
          5
          Es kommt zu einem Compilerfehler.
```

Code zu Quiz: Mehrdimensionale Arrays (3)

```
chaw matrix [433]
2 for (size_t i = 0; i < sizeof(matrix) / sizeof(*matrix); i++) {
      for (size_t j = 0;
           j < sizeof(matrix[i]) / sizeof(*matrix[</pre>
          j++) {
         matrix[i][j] <<= 1;
9 // Loop 2
10 for (size_t i = 0; i < sizeof(matrix) / sizeof(*matrix); i++) {</pre>
      for (size_t j = 0;
            j < sizeof(matrix[i]) / sizeof(*matrix[i]);</pre>
12
           j++) {
13
           matrix[j][i] <<= 1;
14
15
16 }
```

Quiz: Mehrdimensionale Arrays (3)

Welcher Unterschied folgt aus der unterschiedlichen Iterationsreihenfolge über das zweidimensionale Array matrix in den beiden for-Loops im folgenden Codeauschnitt?

Loop 2 führt zu einem Segmentation Fault.
Es gibt keinen funktionalen Unterschied.
Es kommt zu einem Compilerfehler.
Loop 1 wird im Allgemeinen etwas schneller sein.

Quiz: Verschachtelte Datenstrukturen (1)

Was könnte an folgendem Codeauschnitt problematisch sein?

```
1 struct Penguin {
      struct {
          unsigned height;
          unsigned age;
4
          unsigned id;
5
      };
      char id[256]:
8 };
```

Ein verschachteltes struct muss immer benannt sein.

unsigned alleine ist kein valider Datentyp.

Innerhalb eines structs darf ein weiteres struct nicht das erste Element sein.

Es kommt zu einem Compilerfehler.

```
Code zu Quiz: Verschachtelte Datenstrukturen (2)
   struct Penguin {
       char color;
       long id;
 4 };
 5 struct PenguinColony {
       char color[256];
       long ids [256];
 8 };
 9 struct Penguin penguin_colony1 [256]
                                               Array of Struct's
```

struct PenguinColony penguin_colony2; // Struct of Arrays

Quiz: Verschachtelte Datenstrukturen (2)

Was könnte im Allgemeinen ein wesentlicher Unterschied sein zwischen den auf der vorigen Folie zu sehenden Möglichkeiten, mit mehreren gleichartigen structs umzugehen? Sollte lieber ein separates struct mit Arrays der einzelnen Member gewählt werden oder sollte einfach ein Array des ursprünglichen structs Penguin gewählt werden?



Im Allgemeinen gibt es keinen Unterschied.



Je nach Zugriffsmuster kann es einen Performance-Unterschied geben.



Ein Struct of Arrays kann Speichereffizienter sein, da Padding innerhalb der structs entfällt.

Mini-"OOP" in C

```
O Struct 当类
O 用指针改加油
1 struct Penguin {
     char *name;
int age;
4 long position;
void (*walk)(struct Penguin* penguin);
6 };
7 void walk(struct Penguin* penguin) {
      penguin ->position++;
8
9 }
10 int main(void) {
      struct Penguin penguin = {
11
          .name = "tux", .age = 5, .position = 10, .walk = walk
12
      };
13
      penguin.walk(&penguin);
14
     printf("\frac{1}{n}, penguin.position); // -> 11
15
16 }
```

Quiz: Mini-"OOP" in C

Im vorigen Beispiel wurde der walk-Funktion der Penguin per Pointer übergeben. Wie unpraktisch! Man kann ein struct doch auch in diesem Beispiel direkt als Parameter spezifizieren, oder?

