

# Struktury a uniony, přesnost výpočtů a vnitřní reprezentace číselných typů

Jan Faigl

Katedra počítačů  
Fakulta elektrotechnická  
České vysoké učení technické v Praze

Přednáška 06

B0B36PRP – Procedurální programování

## Struktura – struct

- Struktura je konečná množina prvků (proměnných), které nemusí být stejného typu.
- Skladba struktury je definovaná uživatelem jako nový typ sestavený z již definovaných typů.
- K prvkům struktury **přistupujeme tečkovou notací**, např. `struct_proměnná.prvek`.
- K prvkům můžeme přistupovat přes ukazatel operátorem `->`, např. `proměnná_typu_ukazatel_na_struct->prvek`.
- Pro struktury stejného typu je definován operátor přiřazení**, `var_struct1 = var_struct2;`
- Struktury (jako celek) **nelze** porovnávat relačním operátorem `==`.
- Struktura může být funkcí předávána hodnotou i ukazatelem.
- Struktura může být návratovou hodnotou funkce.

## Příklad struct – Inicializace

- Struktury:

```
1 struct record {          1 typedef struct {
2     int number;          2     int n;
3     double value;        3     double v;
4 };                       4 }; item;
```
- Proměnné typu struktura můžeme inicializovat prvek po prvků.

```
1 struct record r;
2 r.value = 21.4;
3 r.number = 7;
```
- Podobně jako pole lze inicializovat přímo při definici

```
1 item i = { 1, 2.3 };
```
- nebo pouze konkrétní položky (ostatní jsou nulovány).

```
1 struct record r2 = { .value = 10.4 };
```

## Přehled témat

- Část 1 – Struktury a uniony  
Struktury – struct  
Proměnné se sdílenou pamětí – union  
Příklad  
Základní číselné typy a jejich reprezentace v počítači  
Typové konverze  
Matematické funkce  
*S. G. Kochan: kapitola 9 a 17*
- Část 2 – Přesnost výpočtů a vnitřní reprezentace číselných typů  
*S. G. Kochan: kapitola 14 (typové konverze)*
- Část 3 – Zadání 6. domácího úkolu (HW06)  
*Appendix – Kódovací příklady*

## Příklad struct – Definice

- Bez zavedení nového typu (`typedef`) je nutné před identifikátor jména struktury uvádět klíčové slovo `struct`.
- Jméno struktury je ve jmenném prostoru složených typů (struktur).

```
1 struct record {          1 typedef struct {
2     int number;          2     int n;
3     double value;        3     double v;
4 };                       4 } item;
```
- Zavedením nového typu `typedef` používáme definovaný typ a nemusíme používat (a ani definovat) jméno struktury.  
*lec06/struct.c*

## Příklad struct jako parametr funkce

- Struktury můžeme předávat jako parametry funkcí hodnotou.

```
1 void print_record(struct record rec) {
2     printf("record: number(%d), value(%lf)\n",
3     rec.number, rec.value);
4 }
```
- Nebo hodnotou ukazatele

```
1 void print_item(item *v) {
2     printf("item: n(%d), v(%lf)\n", v->n, v->v);
3 }
```
- Při předávání parametru
  - hodnotou** se vytváří nová proměnná a původní obsah předávané struktury se kopíruje na zásobník (pro složený typ je definován operátor přiřazení);
  - hodnotou ukazatele** se kopíruje pouze hodnota ukazatele (adresa) a pracujeme tak s původní strukturou.

## Část I

### Část 1 – Struktury a uniony

## Definice jména struktury a typu struktury

- Uvedením `struct record` zavádíme nové jméno struktury `record`.

```
1 struct record {
2     int number;
3     double value;
4 };
```

  - Definujeme identifikátor `record` ve jmenném prostoru struktur.
- Definicí typu `typedef` zavádíme nové jméno typu `record`.

```
1 typedef struct record record;
```

  - Definujeme globální identifikátor `record` jako jméno typu `struct record`.
- Obojí můžeme kombinovat v jediné definici jména a typu struktury.

## Složený typ, operátor přiřazení a pole jako prvek složeného typu 1/2

- Velikost složeného typu musí být známa během překladu, proto můžeme mít definovaný operátor přiřazení.  
*Nebo naopak, abychom mohli jednoduše přiřazovat, tak potřebujeme znát velikost typu.*
- Prvek složeného typu může být pole (definované velikosti) nebo ukazatel.

```
1 void print(const char *str, int n, int *a); 10 for (int i = 0; i < n; ++i) {
2 #define N 10 // We need named literal.      20     s1.a[i] = n - i;
3 int main(void)                               21 }
4 {                                             22 print("s1.a", n, s1.a);
5     const int n = N;                          23 print("s2.a", n, s2.a);
6     struct { // Anonymous struct             24     return 0;
7         int a[N]; // Defined size, no VLA    25 } // end main()
8     int s1, s2; // Two struct variables      26
9     printf("s1 %p; s2 %p\n", &s1, &s2);      27 void print(const char *str, int n, int *a) {
10    s1, s2; // Two struct variables           28     printf("%s %p: ", str, a);
11    printf("s1 %p; s2 %p\n", &s1, &s2);      29     for (int i = 0; i < n; ++i) {
12    for (int i = 0; i < n; ++i) {             30         printf("%d%s", a[i], i < (n-1) ? ", " : "\n");
13        s1.a[i] = i;                           31     }
14    }                                           32 }
15    print("s1.a", n, s1.a);
16    s2 = s1; // Assignment
17    print("s2.a", n, s2.a);
18 }
```

  
*lec06/demo-struct\_array.c*

## Složený typ, operátor přiřazení a pole jako prvek složeného typu 2/2

- Příklad `lec06/demo-struct_array.c`
- Používáme anonymní složený typ - definice strukturu přímo v definici proměnných `s1` a `s2`.
  - Musíme použít textový literál pro definici velikosti položky `a` jako pole definované délky.
  - Ve funkci `print()` tiskneme hodnotu adresy, kde je alokované pole.  
*V našem případě se shoduje s adresou, kde je struktura uložena. Struktura je „organizovaný“ pohled na blok paměti důležitý zejména pro zpřehlední programu. Při běhu programu vlastně není nutné mít v paměti dílčí jména prvku složeného typu.*
- ```
s1 0x7fffffff80; s2 0x7fffffff818
s1.a 0x7fffffff80: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9
s2.a 0x7fffffff818: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9
s1.a 0x7fffffff80: 10, 9, 8, 7, 6, 5, 4, 3, 2, 1
s2.a 0x7fffffff818: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9
```
- V příkladu si vyzkoušejte chování překladu a programu v případě použití VLA nebo konstantní proměnné definující velikost pole.
  - Pole definované velikostí nahraďte dynamicky alokovaným polem.

## Struktura struct a velikost

- Vnitřní reprezentace struktury nutně nemusí odpovídat součtu velikostí jednotlivých prvků.
- ```
1 struct record {                1 typedef struct {
2   int number;                  2   int n;
3   double value;                3   double v;
4 };                             4 } item;

1 printf("Size of int: %lu size of double: %lu\n", sizeof(int), sizeof(double));
2 printf("Size of record: %lu\n", sizeof(struct record));
3 printf("Size of item: %lu\n", sizeof(item));
```
- ```
Size of int: 4 size of double: 8
Size of record: 16
Size of item: 16
```
- `lec06/struct.c`

## Proměnné se sdílenou pamětí – union

- **Union** je množina prvků (proměnných), které nemusí být stejného typu.
- Prvky unionu sdílejí společně stejná paměťová místa.  
*Překrývají se*
- Velikost unionu je dána velikostí největšího z jeho prvků.
- Skladba unionu je definována uživatelem jako nový typ sestavený z již definovaných typů.
- K prvkům unionu se přistupuje tečkovou notací.
- Pokud nedefinujeme nový typ, je nutné k identifikátoru proměnné unionu uvádět klíčové slovo `union`.  
*Podobně jako u struktury struct.*

```
1 union Nums {
2   char c;
3   int i;
4 };
5 Nums nums; /* THIS IS NOT ALLOWED! Type Nums is not known! */
6 union Nums nums;
```

## Příklad struct – Přiřazení

- Hodnoty proměnné **stejného typu** struktury můžeme přiřadit operátorem `=`.
- ```
1 struct record {                1 typedef struct {
2   int number;                  2   int n;
3   double value;                3   double v;
4 };                             4 } item;

1 struct record rec1 = { 10, 7.12 };
2 struct record rec2 = { 5, 13.1 };
3 item i;
4 print_record(rec1); /* number(10), value(7.120000) */
5 print_record(rec2); /* number(5), value(13.100000) */
6 rec1 = rec2;
7 i = rec1; /* IT IS NOT ALLOWED! */
8 // Different types, albeit with the same memory representation.
9 print_record(rec1); /* number(5), value(13.100000) */
```
- `lec06/struct.c`

## Struktura struct a velikost 1/2

- Při kompilaci zpravidla dochází k zarovnání prvků na velikost slova příslušné architektury.  
*Např. 8 bytů v případě 64-bitové architektury. Jednotlivé prvky jsou na adrese v násobNapř. 8 bytů v případě 64-bitové architektury.*
  - Můžeme explicitně předepsat kompaktní paměťovou reprezentaci, např. direktivou `__attribute__((packed))` překladací `clang` a `gcc`.
- ```
1 struct record_packed {
2   int n;
3   double v;
4 } __attribute__((packed));
```
- `lec06/struct.c`

## Příklad union 1/2

- Union složený z proměnných typu: `char`, `int` a `double`.
- ```
1 int main(int argc, char *argv[])
2 {
3   union Numbers {
4     char c;
5     int i;
6     double d;
7   };
8   printf("size of char %lu\n", sizeof(char));
9   printf("size of int %lu\n", sizeof(int));
10  printf("size of double %lu\n", sizeof(double));
11  printf("size of Numbers %lu\n", sizeof(union Numbers));
12
13  union Numbers numbers;
14  printf("Numbers c: %d i: %d d: %lf\n", numbers.c, numbers.i, numbers.d);
```
- Příklad výstupu.  
size of char 1  
size of int 4  
size of double 8  
size of Numbers 8  
Numbers c: 48 i: 740313136 d: 0.000000
- `lec06/union.c`

## Příklad struct – Přímá kopie paměti

- Jsou-li dvě struktury stejně veliké, můžeme přímo kopírovat obsah příslušné paměťové oblasti.  
*Například funkci `memcpy()` z knihovny `string.h`*
- ```
1 struct record r = { 7, 21.4};
2 item i = { 1, 2.3 };
3 print_record(r); /* number(7), value(21.400000) */
4 print_item(&i); /* n(1), v(2.300000) */
5 if (sizeof(i) == sizeof(r)) {
6   printf("i and r are of the same size\n");
7   memcpy(&i, &r, sizeof(i));
8   print_item(&i); /* n(7), v(21.400000) */
9 }
```
- V tomto případě je interpretace hodnot v obou strukturách identická, obecně tomu však být nemusí. Například v případě změny pořadí prvků typu `int` a `double`.
- `lec06/struct.c`

## Struktura struct a velikost 2/2

- Nebo
- ```
1 typedef struct __attribute__((packed)) {
2   int n;
3   double v;
4 } item_packed;
```
- Příklad výstupu:  
Size of int: 4 size of double: 8  
Size of record\_packed: 12  
Size of item\_packed: 12
- `lec06/struct.c`
- Zarovnání zpravidla přináší rychlejší přístup do paměti, ale zvyšuje paměťové nároky.  
<http://www.catb.org/esr/structure-packing>  
<https://stackoverflow.com/questions/4306186/structure-padding-and-packing>

## Příklad union 2/2

- Proměnné sdílejí paměťový prostor.
- ```
1 numbers.c = 'a';
2 printf("nSet the numbers.c to 'a'\n");
3 printf("Numbers c: %d i: %d d: %lf\n", numbers.c, numbers.i, numbers.d);
4 numbers.i = 5;
5 printf("nSet the numbers.i to 5\n");
6 printf("Numbers c: %d i: %d d: %lf\n", numbers.c, numbers.i, numbers.d);
7 numbers.d = 3.14;
8 printf("nSet the numbers.d to 3.14\n");
9 printf("Numbers c: %d i: %d d: %lf\n", numbers.c, numbers.i, numbers.d);
```
- Příklad výstupu  
Set the numbers.c to 'a'  
Numbers c: 97 i: 1374389601 d: 3.140000  
Set the numbers.i to 5  
Numbers c: 5 i: 5 d: 3.139999  
Set the numbers.d to 3.14  
Numbers c: 31 i: 1374389535 d: 3.140000
- `lec06/union.c`

### Inicializace union

- Proměnnou typu `union` můžeme inicializovat při definici.

```
1 union {
2   char c;
3   int i;
4   double d;
5 } numbers = { 'a' };
```

*Pouze první položka (proměnná) může být inicializována.*

- V C99 můžeme inicializovat konkrétní položku (proměnnou).

```
1 union {
2   char c;
3   int i;
4   double d;
5 } numbers = { .d = 10.3 };
```

### Příklad struktura, pole a výčtový typ 3/3

- Detekci národního prostředí provedeme podle hodnoty proměnné prostředí.

*Pro jednoduchost detekujeme češtinu na základě výskytu řetězce "cs" v hodnotě proměnné prostředí LC\_CTYPE.*

```
35 _Bool cz = 0;
36 while (*envp != NULL) {
37   if (strstr(*envp, "LC_CTYPE") && strstr(*envp, "cs")) {
38     cz = 1;
39     break;
40   }
41   envp++;
42 }
43 const week_day_s *days = cz ? days_cs : days_en;
44 printf("%d %s %s\n", day_of_week,
45        days[day_of_week].name,
46        days[day_of_week].abbr
47        );
48 return 0;
49 }
50 }
51 }
52 }
53 }
54 }
55 }
56 }
57 }
58 }
59 }
60 }
```

*V programu jsme využili koncept definování datových struktur, které následně programově přepínáme a využíváme. Alternativně můžeme data načítat ze souboru. V programu se snažíme obecně pracovat s datovými strukturami.*

### Reprezentace dat v počítači

- V počítači není u datové položky určeno jaký konkrétní datový typ je v paměti uložen.
- Proto musíme přidělení paměti **definovat** s jakými typy dat budeme pracovat.
- Překladač tuto definici hlídá a volí odpovídající strojové instrukce pro práci s daty, např. jako s odpovídajícími číselnými typy.

*Např. necelocíselné (float) typy a využití tzv. FPU (Floating Point Unit).*

#### Příklad zápisů stejného čísla v různých soustavách.

- 0100 0001<sub>(2)</sub> – binární zápis jednoho bajtu (8-mi bitů);
- 65<sub>(10)</sub> – odpovídající číslo v dekadické soustavě;
- 41<sub>(16)</sub> – odpovídající číslo v šestnáctkové soustavě;
- Obsah paměťového místa 0100 0001<sub>(2)</sub> o velikosti 1 byte může být interpretován jako znak A.

### Příklad struktura, pole a výčtový typ 1/3

- Hodnoty (konstanty) výčtového typu jsou celá čísla, která mohou být použita jako indexy (pole).
- Také je můžeme použít pro inicializaci pole struktur.

```
1 #include <stdio.h>
2 #include <stdlib.h>
3 #include <string.h>
4
5 enum weekdays { MONDAY, TUESDAY, WEDNESDAY, THURSDAY, FRIDAY };
6
7 typedef struct {
8   char *name;
9   char *abbr; // abbreviation
10 } week_day_s;
11
12 const week_day_s days_en[] = {
13   [MONDAY] = { "Monday", "mon" },
14   [TUESDAY] = { "Tuesday", "tue" },
15   [WEDNESDAY] = { "Wednesday", "wed" },
16   [THURSDAY] = { "Thursday", "thr" },
17   [FRIDAY] = { "Friday", "fri" },
18 };
19
20 int main(int argc, char *argv[], char **envp)
21 {
22   if (argc > 1 ? atoi(argv[1]) : 1;
23   if (day_of_week < 1 || day_of_week > 5) {
24     fprintf(stderr, "(EE) File: '%s' Line: %d -- Given day of week out of range\n", __FILE__,
25             __LINE__);
26     return 101;
27 }
28 }
29 }
30 }
31 }
32 }
33 }
34 }
35 }
36 }
37 }
38 }
39 }
40 }
```

lec06/demo-struct.c

### Datové typy

- Při návrhu algoritmu abstrahujeme od binární podoby paměti počítače.
- S daty pracujeme jako s hodnotami různých datových typů, které jsou uloženy v paměti předepsaným způsobem.
- Datový typ specifikuje
  - Množinu hodnot, které je možné v počítači uložit;
  - Množinu operací, které lze s hodnotami typu provádět.
- Jednoduchý typ** je takový typ, jehož hodnoty jsou atomické, tj. z hlediska operací dále nedělitelné.

*Záleží na způsobu reprezentace.*

### Číselné soustavy

- Číselné soustavy – poziční číselné soustavy (polyadické) jsou charakterizovány bázi udávající kolik číslic lze maximálně použít.
 
$$x_d = \sum_{i=-n}^{i=m} a_i \cdot z^i$$
, kde  $a_i$  je číslice a  $z$  je základ soustavy.
- Unární – např. počet vypitých pšlitrů.
- Binární soustava (bin) – 2 číslice 0 nebo 1.
 
$$11010, 01_{(2)} = 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 = 1 \cdot 2^4 + 0 \cdot 2^3 + 1 \cdot 2^1 + 0 \cdot 2^0 = 1 \cdot 16 + 1 \cdot 8 + 0 \cdot 4 + 1 \cdot 2 + 0 \cdot 1 = 1 \cdot 16 + 1 \cdot 8 + 2 = 26, 25$$
- Desítková soustava (dec) – 10 číslic, znaky 0 až 9.
 
$$138, 24_{(10)} = 1 \cdot 10^2 + 3 \cdot 10^1 + 8 \cdot 10^0 = 1 \cdot 10^2 + 3 \cdot 10^1 + 8 \cdot 10^0 = 1 \cdot 100 + 3 \cdot 10 + 8 \cdot 1 = 100 + 30 + 8 = 138$$
- Šestnáctková soustava (hex) – 16 číslic, znaky 0 až 9 a A až F.
 
$$0x7D_{(16)} = 7 \cdot 16^1 + D \cdot 16^0 = 7 \cdot 16 + 13 = 112 + 13 = 125$$

### Příklad struktura, pole a výčtový typ 2/3

- Připravíme si pole struktur pro konkrétní jazyk (angličtina a čeština).
- Program vytiskne jméno a zkratku dne v týdnu dle čísla dne v týdnu. *V programu používáme jednotné číslo dne bez ohledu na jazykovou mutaci.*

```
19 const week_day_s days_cs[] = {
20   [MONDAY] = { "Pondělí", "po" },
21   [TUESDAY] = { "Úterý", "ut" },
22   [WEDNESDAY] = { "Středa", "st" },
23   [THURSDAY] = { "Čtvrtek", "ct" },
24   [FRIDAY] = { "Pátek", "pa" },
25 };
26
27 int main(int argc, char *argv[], char **envp)
28 {
29   int day_of_week = argc > 1 ? atoi(argv[1]) : 1;
30   if (day_of_week < 1 || day_of_week > 5) {
31     fprintf(stderr, "(EE) File: '%s' Line: %d -- Given day of week out of range\n", __FILE__,
32             __LINE__);
33     return 101;
34 }
35 }
36 }
37 }
38 }
39 }
40 }
```

lec06/demo-struct.c

### Příklad číselných typů a vnitřní reprezentace

- 32-bitový typ `int` umožňuje uložit celá čísla v intervalu  $(-2147483648, 2147483647)$ , pro která můžeme použít:
  - aritmetické operace `+`, `-`, `*`, `/` s výsledkem hodnota typu `int`;
  - relační operace `==`, `!=`, `>`, `<`, `>=`, `<=`.
- Inicializovat hodnotou dekadického nebo hexadecimálního literálu.
 

```
1 int i; // definice promenne typu int
2 int decI = 120; // definice spolu s prirazenim
3 int hexI = 0x78; //pocatecni hodnota v 16-kove soustavě
4 int sum = 10 + decI + 0x13; //pocatecni hodnota je vyraz
```
- Vnitřní reprezentace typů (např. `int`, `short`, `double`) umožňuje uložit čísla z definovaného rozsahu s různou přesností.
- Číselné datové typy lze vzájemně převádět implicitní nebo explicitní typovou konverzí.
- Při konverzi nemusí být hodnota zachována** – viz `lec06/demo-types.c`.

### Kódování záporných čísel

- Přímý kód** – znaménko je určeno prvním bitem (zleva), snadné stanovení absolutní hodnoty. Reprezentace má dvě nuly.
 

|                      |                          |
|----------------------|--------------------------|
| 121 <sub>(10)</sub>  | 0111 1001 <sub>(2)</sub> |
| -121 <sub>(10)</sub> | 1111 1001 <sub>(2)</sub> |
| 0 <sub>(10)</sub>    | 0000 0000 <sub>(2)</sub> |
| -0 <sub>(10)</sub>   | 1000 0000 <sub>(2)</sub> |
- Inverzní kód** – záporné číslo odpovídá bitové negaci kladné hodnoty čísla. Reprezentace má dvě nuly.
 

|                      |                          |
|----------------------|--------------------------|
| 121 <sub>(10)</sub>  | 0111 1001 <sub>(2)</sub> |
| -121 <sub>(10)</sub> | 1000 0110 <sub>(2)</sub> |
| 0 <sub>(10)</sub>    | 0000 0000 <sub>(2)</sub> |
| -0 <sub>(10)</sub>   | 1111 1111 <sub>(2)</sub> |
- Doplňkový kód** – záporné číslo je uloženo jako hodnota kladného čísla po bitové negaci zvětšená o 1. Jediná reprezentace nuly.
 

|                      |                          |
|----------------------|--------------------------|
| 127 <sub>(10)</sub>  | 0111 1111 <sub>(2)</sub> |
| -128 <sub>(10)</sub> | 1000 0000 <sub>(2)</sub> |
| -1 <sub>(10)</sub>   | 1111 1111 <sub>(2)</sub> |

Struktury – struct    Uniony    Příklad    **Reprezentace číselných typů**    Typové konverze    Matematické funkce

## Více-bajtová reprezentace a pořadí bajtů

- Číselné typy s více-bajtovou reprezentací mohou mít bajty uloženy v různém pořadí.
  - little-endian** – **nejméně** významný bajt se ukládá na nejnižší adresu. x86, ARM
  - big-endian** – **nejvíce** významný bajt se ukládá na nejnižší adresu. Motorola, ARM
- Pořadí je důležité při přenosu hodnot z paměti jako posloupnosti bajtů a jejich následné interpretaci.
- Network byte order** – je definován pro síťový přenos a není tak nutné řešit konkrétní architekturu.
  - Tj. hodnoty z paměti jsou ukládány a přenášeny v tomto pořadí bajtů a na cílové stanici pak zpětně zapsány do konkrétního nativního pořadí. big-endian

*Informativní*

Jan Faigl, 2024    BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy    32 / 54

Struktury – struct    Uniony    Příklad    **Reprezentace číselných typů**    Typové konverze    Matematické funkce

## Příklad reprezentace celých čísel int

- Na 32-bitových a 64-bitových strojích je celočíselný typ **int** zpravidla reprezentován 32 bity (4 bajty).
- Typ **int** je znaménkový typ.
  - Znaménko je zakódováno v 1 bitu a vlastní číselná hodnota pak ve zbývajících 31 bitech.
    - Největší číslo je  $0111 \dots 111 = 2^{31} - 1 = 2\,147\,483\,647$ . *Reprezentujeme i nulu.*
    - Nejmenší číslo je  $-2^{31} = -2\,147\,483\,648$ . *0 už je zahrnuta.*
  - Pro zobrazení záporných čísel je použit **doplňkový kód**. *Nejmenší číslo v doplňkovém kódu 1000...000 je  $-2^{31}$ .*

Jan Faigl, 2024    BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy    33 / 54

Struktury – struct    Uniony    Příklad    **Reprezentace číselných typů**    Typové konverze    Matematické funkce

## Reprezentace záporných celých čísel

- Doplňkový kód –  $D(x)$ .
  - Pro 8-mi bitovou reprezentací čísel.
    - Můžeme reprezentovat  $2^8 = 256$  čísel.
    - Rozsah  $r = 256$ .

$$D(x) = \begin{cases} x & \text{pro } 0 \leq x < \frac{r}{2} \\ r + x & \text{pro } -\frac{r}{2} \leq x < 0 \end{cases} \quad (1)$$

- Příklady
 

| Desítkově | Doplňkový kód                                  |
|-----------|------------------------------------------------|
| 0-127     | 0000 0000 – 0111 1111                          |
| 128       | nelze zobrazit na 8 bitů v doplňkovém kódu     |
| -128      | $D(-128) = 256 + (-128) = 128$ to je 1000 0000 |
| -1        | $D(-1) = 256 + (-1) = 255$ to je 1111 1111     |
| -4        | $D(-4) = 256 + (-4) = 252$ to je 1111 1100     |

*Informativní*

Jan Faigl, 2024    BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy    34 / 54

Struktury – struct    Uniony    Příklad    **Reprezentace číselných typů**    Typové konverze    Matematické funkce

## Necelá čísla a přesnost výpočtu 1/2

- Ztráta přesnosti při aritmetických operacích.

**Příklad sčítání dvou čísel**

```

1 #include <stdio.h>
2 int main(void)
3 {
4     double a = 1e+10;
5     double b = 1e-10;
6     printf("a : %24.12lf\n", a);
7     printf("b : %24.12lf\n", b);
8     printf("a+b: %24.12lf\n", a + b);
9     return 0;
10 }
11 clang sum.c && ./a.out
12 a : 10000000000.000000000000
13 b : 0.00000000000100
14 a+b: 10000000000.000000000000
    
```

lec06/sum.c

Jan Faigl, 2024    BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy    35 / 54

Struktury – struct    Uniony    Příklad    **Reprezentace číselných typů**    Typové konverze    Matematické funkce

## Necelá čísla a přesnost výpočtu 2/2

**Příklad dělení dvou čísel**

```

1 #include <stdio.h>
2 int main(void)
3 {
4     const int number = 100;
5     double dV = 0.0;
6     float fV = 0.0f;
7     for (int i = 0; i < number; ++i) {
8         dV += 1.0 / 10.0;
9         fV += 1.0 / 10.0;
10    }
11    printf("double value: %lf ", dV);
12    printf("float value: %lf ", fV);
13    return 0;
14 }
15 clang division.c && ./a.out
16 double value: 10.000000 float value: 10.000002
    
```

lec06/division.c

Jan Faigl, 2024    BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy    36 / 54

Struktury – struct    Uniony    Příklad    **Reprezentace číselných typů**    Typové konverze    Matematické funkce

## Přesnost výpočtu - strojová přesnost

- Strojová přesnost  $\epsilon_m$  - nejmenší desetinné číslo, které přičtením k 1.0 dává výsledek různý od 1, pro  $|v| < \epsilon_m$ , platí
 
$$v + 1.0 == 1.0.$$

*Symbol == odpovídá porovnání dvou hodnot (test na ekvivalenci).*
- Zaokrouhlovací chyba - **nejméně**  $\epsilon_m$ .
- Přesnost výpočtu - aditivní chyba roste s počtem operací v řádu  $\sqrt{N} \cdot \epsilon_m$ .
  - Často se však kumuluje preferabilně v jedno směru v řádu  $N \cdot \epsilon_m$ .

Jan Faigl, 2024    BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy    37 / 54

Struktury – struct    Uniony    Příklad    **Reprezentace číselných typů**    Typové konverze    Matematické funkce

## Reprezentace reálných čísel

- Pro uložení čísla vyhradzujeme omezený paměťový prostor.

**Příklad – zápis čísla  $\frac{1}{3}$  v dekadické soustavě**

- $= 33333333 \dots 3333$
- $= 0,33$
- $\approx 0,33333333333333333333$
- $\approx 0,333$

*V trojkové soustavě:  $0 \cdot 3^1 + 0 \cdot 3^0 + 1 \cdot 3^{-1} = (0,1)_3$*

- Nepřesnosti v zobrazení reálných čísel v konečné posloupnosti bitů způsobují
  - Iracionální čísla, např.  $e, \pi, \sqrt{2}$ ;
  - Čísla, která mají v dané soustavě periodický rozvoj, např.  $\frac{1}{3}$ ;
  - Čísla, která mají příliš dlouhý zápis.

Jan Faigl, 2024    BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy    38 / 54

Struktury – struct    Uniony    Příklad    **Reprezentace číselných typů**    Typové konverze    Matematické funkce

## Model reprezentace reálných čísel

- Reálná čísla se zobrazují jako aproximace daným rozsahem paměťového místa.
- Reálné číslo  $x$  se zobrazuje ve tvaru
 
$$x = \text{mantisa} \cdot \text{základ}^{\text{exponent}}$$
- Pro jednoznačnost zobrazení musí být mantisa normalizována, např.  $0, 1 \leq m < 1$  nebo ve tvaru  $\pm 1.[\text{mantisa}] \cdot 2^{\text{exponent}}$
- Ve vyhrazeném paměťovém prostoru je pro zvolený základ uložen exponent a mantisa jako dvě celá čísla.

Jan Faigl, 2024    BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy    39 / 54

Struktury – struct    Uniony    Příklad    **Reprezentace číselných typů**    Typové konverze    Matematické funkce

## Příklad modelu reprezentace reálných čísel na 7 bajtů se základem 10

- Mantisa 3 pozice plus znaménko, délka exponentu 2 pozice plus znaménko, základ  $z = 10$ . *Reprezentace dle IEEE-754 používá dvojkový základ!*
- Reprezentace nuly.
 

? ?? + 000

+ 02 + 775
- Maximální zobrazitelné kladné číslo 0,999<sup>99</sup>.
 

+ 99 + 999
- Maximální zobrazitelné záporné číslo -0,100z<sup>-99</sup>.
 

- 99 - 100
- Minimální zobrazitelné kladné číslo 0,100z<sup>-99</sup>.
 

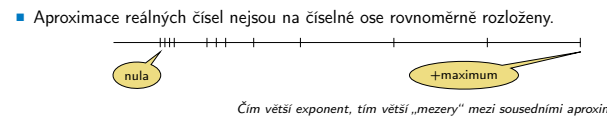
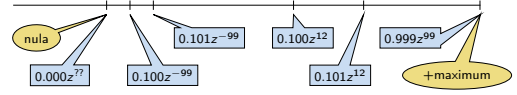
- 99 + 100
- Minimální zobrazitelné záporné číslo -0,999z<sup>+99</sup>.
 

+ 99 - 999

Jan Faigl, 2024    BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy    40 / 54

### Model reprezentace reálných čísel a vzdálenost mezi aproximacemi

- Rozsah hodnot pro konkrétní exponent je dán velikostí mantisy.
- Absolutní vzdálenost dvou aproximací tak záleží na exponentu.
  - Mezi hodnotou 0 a 1,0 je využit celý rozsah mantisy pro exponenty  $\{-99, -98, \dots, 0\}$ .



Čím větší exponent, tím větší „mezery“ mezi sousedními aproximacemi čísel.

### Přirázovací operátor a příkaz

- Slouží pro nastavení hodnoty proměnné.
  - Uložení číselné hodnoty do paměti, kterou proměnná reprezentuje.
- Tvar přirázovacího operátoru.

(proměnná) = (výraz)

Výraz je literál, proměnná, volání funkce, ...

- Zkrácený zápis (proměnná) (operátor) = (výraz)
- Přirazení je výraz asociativní zprava.
- Přirázovací příkaz – výraz zakončený středníkem ;

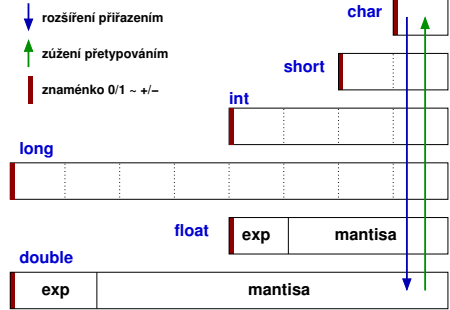
```

1 int x; //definice promenne x      1 int x, y; //definice promennych x a y
2 int y; //definice promenne y      3 x = 10;
4 x = 6;                             4 y = 7;
5 y = x = x + 6;                      6 y += x + 10;

```

### Konverze primitivních číselných typů

- Primitivní datové typy jsou vzájemně nekompatibilní, ale jejich hodnoty lze převádět.

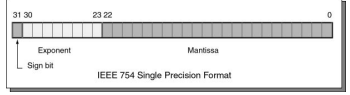


### Reprezentace necelých čísel – IEEE 754

- Reálné číslo  $x$  se zobrazuje ve tvaru  $x = (-1)^s \cdot \text{mantisa} \cdot 2^{\text{exponent} - \text{bias}}$ . Základ 2.

IEEE 754, ISO/IEC/IEEE 60559:2011

- Mantisa je **normalizována** na první jedničku vlevo (v soustavě o dvojkovém základu).
- float – 32 bitů (4 bajty):  $s$  – 1 bit znaménko (+ nebo -), **exponent** – 8 bitů, tj. 256 možností. **mantisa** – 23 bitů  $\approx 16,7$  milionu možností.



- double – 64 bitů (8 bajtů).
  - $s$  – 1 bit znaménko (+ nebo -).
  - exponent – 11 bitů, tj. 2048 možností.
  - mantisa – 52 bitů  $\approx 4,5$  biliardy možností (4 503 599 627 370 495).
- bias umožňuje reprezentovat exponent vždy jako kladné číslo.
  - Lze zvolit, např.  $\text{bias} = 2^{e-1} - 1$ , kde  $e$  je počet bitů exponentu.

<http://www.root.cz/clanky/norma-ieee-754-a-pribuzni-formaty-plovouci-radove-tecky>

### Typové konverze

- Typová konverze je operace převedení hodnoty nějakého typu na hodnotu typu jiného.
- Typová konverze může být
  - implicitní – vyvolá se automaticky;
  - explicitní – je nutné v programu explicitně uvést.
- Konverze typu **int** na **double** je implicitní.
  - Hodnota typu **int** může být použita ve výrazu, kde se očekává hodnota typu **double**, dojde k automatickému převodu na hodnotu typu **double**.

#### Příklad

```

1 double x;
2 int i = 1;
4 x = i; // hodnota 1 typu int se automaticky převede
5 // na hodnotu 1.0 typu double

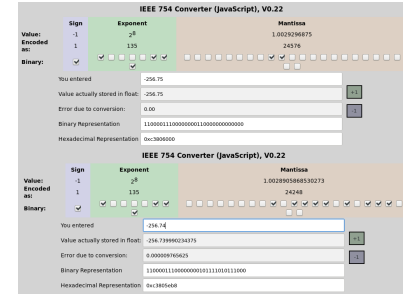
```

- Implicitní konverze je bezpečná.

### Matematické funkce

- <math.h> – základní funkce pro práci s „reálnými“ čísly.
  - Výpočet odmocniny necelého čísla  $x$ . `double sqrt(double x);`, `float sqrtf(float x);`
  - V C funkci nepřetěžujeme, proto jsou jména odlišena.*
  - `double pow(double x, double y);` – výpočet obecné mocniny.
  - `double atan2(double y, double x);` – výpočet  $\arctan y/x$  s určením kvadrantu.
  - Symbolické konstanty – `M_PI`, `M_PI_2`, `M_PI_4`, atd.
    - `#define M_PI 3.14159265358979323846`
    - `#define M_PI_2 1.57079632679489661923`
    - `#define M_PI_4 0.78539816339744830962`
  - `isfinite()`, `isnan()`, `isless()`, ... – makra pro porovnání reálných čísel.
  - `round()`, `ceil()`, `floor()` – zaokrouhlování, převod na celá čísla.
- <complex.h> – funkce pro počítání s komplexními čísly. ISO C99
- <fenv.h> – funkce pro řízení zaokrouhlování a reprezentaci dle IEEE 754. man math

### Příklad reprezentace float hodnot dle IEEE 754



- Chyba reprezentace -256.75 vs -256.74.
- Infinity (0x7f800000), -Infinity (0xff800000), a NaN (0xf7fffff).

<https://www.h-schmidt.net/FloatConverter/IEEE754.html>

### Explicitní typové konverze

- Převod hodnoty typu **double** na **int** je třeba **explicitně** předeepsat.
- Dojde k „odseknutí“ necelé části hodnoty **int**.

#### Příklad

```

1 double x = 1.2; // definice proměnné typu double
2 int i; // definice proměnné typu int
3 int i = (int)x; // hodnota 1.2 typu double se převede
4 // na hodnotu 1 typu int

```

- Explicitní konverze je potenciálně nebezpečná.

#### Příklady

```

1 double d = 1e30;                1 long l = 5000000000L;
2 int i = (int)d;                 2 int i = (int)l;
4 // i je -2147483648             4 // i je 705032704
5 // to je asi -2e9 místo 1e30   5 // (oříznuté 4 bajty)
                                   lec06/demo-type_conversion.c

```

## Část II

### Část 3 – Zadání 6. domácího úkolu (HW06)

## Zadání 6. domácího úkolu HW06

### Téma: Maticové počty

Povinné zadání: **2b**; Volitelné zadání: **3b**; Bonusové zadání: **3b**

- **Motivace:** Získání zkušenosti s dvojrozměrným polem.
- **Cíl:** Osvojit si práci s polem variabilní délky a předáváním ukazatelů.
- **Zadání:** <https://cw.fe1.cvut.cz/wiki/courses/b0b36prp/hw/hw06>
  - Načtení vstupních hodnot dvou matic a znaku operace ('\*' – násobení).
  - **Volitelné zadání** rozšiřuje úlohu o další operace s maticemi sčítání ('+') a odčítání ('-'), které mohou být zapsány ve výrazu.
  - **Bonusové zadání** pak řeší zpracování celého výrazu, ve kterém jsou však jednotlivé matice uvedeny jako symboly, které jsou nejdříve definovány načtením hodnot matic ze standardního vstupu.

Využití struct a dynamické alokace může být výhodnou, není však nutné.

- **Termin odevzdání:** 30.11.2024, 23:59:59 PST.
- **Bonusová úloha:** 11.01.2025, 23:59:59 CET.

PST – Pacific Standard Time

CET – Central European Time

## Shrnutí přednášky

## Diskutovaná témata

- Struktury, způsoby definování, inicializace a paměťové reprezentace
- Uniony
- Přesnost výpočtu
- Vnitřní paměťová reprezentace celočíselných i neceločíselných číselných typů
- Knihovna [math.h](#)

- **Příště:** Standardní knihovny C. Rekurse.

## Část IV

### Appendix

Jan Faigl, 2024 BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy 52 / 54

toupper() strev() strvc() streplit() Knihovna strings.h „String objekt“

## Kódovací příklad – Textové řetězce – toupper() 1/2

- Implementujeme funkci, která převede malá písmena na velká (ASCII znaky 'a'-'z'). Využijeme vlastní myMalloc().

```
1 #ifndef _MY_MALLOC_H_
2 #define _MY_MALLOC_H_
3 #include <stdlib.h>
4 #include <string.h>
5 void* myMalloc(size_t size, const char *filename,
6               int line);
7 #endif
8                               my_malloc.h
```

```
1 #include <stdio.h>
2 #include <stdlib.h>
3 #include "my_malloc.h"
4 void* myMalloc(size_t size, const char *filename,
5               int line)
6 {
7     void *ret = malloc(size);
8     if (ret) {
9         fprintf(stderr, "ERROR: Malloc failed called
10                at %s:%i\n", filename, line);
11         exit(-1);
12     }
13     return ret;
14 }
15                               my_malloc.c
```

```
1 #include <stdio.h>
2 #include <string.h>
3 #include "my_malloc.h"
4 int main(void)
5 {
6     const char *str = "I like prp!"; // Ukazatel na literál!
7     const size_t n = strlen(str); // Co se stane když str == NULL!
8     char *stru = myMalloc(
9         (n + 1) * sizeof(char), //+1 pro '\0'
10        _FILE_, _LINE_);
11    };
12    for (int i = 0; i < n; ++i) {
13        stru[i] = (stru[i] >= 'a' && stru[i] <= 'z') ?
14            stru[i] & 0xDF : stru[i]; // 0xDF viz ASCII tabulka!
15    };
16    stru[n] = '\0'; // zajištění textového řetězce
17    printf("%s\n", str);
18    free(stru); // Volání je ok i v případě, že stru == NULL.
19    return EXIT_SUCCESS;
20 }
```

- V našem případě je `str` platný řetězec, proto je řádek 9 v pořádku.
- Přesto převod prepíše do funkce `toupper()`, kde tomu tak být nemusí.

Jan Faigl, 2024 BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy 53 / 54

toupper() strev() strvc() streplit() Knihovna strings.h „String objekt“

## Kódovací příklad – Textové řetězce – toupper() 2/2

```
1 #include <stdio.h>
2 #include <string.h>
3 #include "my_malloc.h"
4 char* strtoupper(const char *str);
5 int main(void)
6 {
7     const char *cur = str; // kurzor vstupního řetězce
8     char *ret = ret; // kurzor výstupního řetězce
9     while (cur && *cur) {
10        *d = *cur++;
11        if (*d >= 'a' && *d <= 'z') {
12            *d = *d - 'a' + 'A';
13        }
14        d += 1;
15    }
16    return ret;
17 }
18 # clang strtoupper.c my_malloc.c && ./a.out
19 # I LIKE PRP!
```

```
1 char* strtoupper(const char *str)
2 {
3     char *ret = myMalloc( // Co se stane když malloc(0)?
4     // Uvěříme, zdali je str platný ukazatel
5     (str ? strlen(str) + 1 : 1) * sizeof(char),
6     _FILE_, _LINE_);
7     };
8     const char *cur = str; // kurzor vstupního řetězce
9     char *d = ret; // kurzor výstupního řetězce
10    while (cur && *cur) {
11        *d = *cur++;
12        if (*d >= 'a' && *d <= 'z') {
13            *d = *d - 'a' + 'A';
14        }
15        d += 1;
16    }
17    return ret;
18 }
19 # clang strtoupper.c my_malloc.c && ./a.out
20 # I LIKE PRP!
```

- Volání funkce `strtoupper()` může být předán neplatný ukazatel `NULL`.
- Explicitně ošetřujeme, ikdyž například funkce `strlen()` předpokládá validní vstup a volání `strlen(NULL)` může skončit chybou programu.
- V našem programu, alokujeme ve funkci `strtoupper()` paměť dynamicky a to vždy nejméně pro jeden znak ('\0').

Jan Faigl, 2024 BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy 55 / 54

toupper() strev() strvc() streplit() Knihovna strings.h „String objekt“

## Kódovací příklad – Textové řetězce – strev() 1/2

- Implementujeme funkci, která vrátí obrázený řetěz. Nejdříve však začneme pracovní verzí ve funkci `main()`.

```
1 #include <stdio.h>
2 #include <stdlib.h>
3 #include <string.h>
4 int main(void)
5 {
6     char *str = "I like prp!";
7     size_t j, n = strlen(str);
8     for (size_t i = 0; j = n-1; i < n/2; ++i, --j) {
9         char t = str[i];
10        str[i] = str[j];
11        str[j] = t;
12    }
13    printf("%s\n", str);
14    return EXIT_SUCCESS;
15 }
```

```
1 $ clang -g strev.c && ./a.out
2 I like prp!
3 Command terminated
4 Command: ./a.out
5 --10618--
6 --10618-- Process terminating with default action of signal 11 (
7                SIGSEGV)
8 --10618-- Bad permissions for mapped region at address 0x20065D
9 --10618-- at 0x2019F9: main (strev.c:13)
```

```
1 char str[] = "I like prp!";
```

- V cyklu využíváme operátor čárky k inicializaci a dekrementaci proměnné `j`.
- Opět v našem programu je řetězec `str` platný a můžeme tak bezpečně volat funkci `strlen(str)`.
- Nicméně po ovládní obrácení řetězce, program prepíše s implementací naší nové funkce `strev()`.

- Program však skončí chybou! Zapisujeme do paměti literál!
- Nahrazením ukazatele na literál pole, program funguje.
- 1 \$ clang -g strev.c && ./a.out
- 2 I like prp!
- 3 |prp ekil I

- Program prepíše s využitím `myMalloc()`.

Jan Faigl, 2024 BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy 57 / 54

toupper() strev() strvc() streplit() Knihovna strings.h „String objekt“

## Kódovací příklad – Textové řetězce – strev() 2/2

```
1 #include <stdio.h>
2 #include <stdlib.h>
3 #include "my_malloc.h"
4 char* strev(const char *str);
5 int main(void)
6 {
7     char *str = "I like prp!";
8     char *stru = strev(str);
9     printf("%s\n", str);
10    printf("%s\n", stru);
11    free(str);
12    return EXIT_SUCCESS;
13 }
```

```
20 char* strev(const char *str)
21 {
22     size_t n = str ? strlen(str) : 0;
23     char *ret = myMalloc((n + 1) * sizeof(char), _FILE_,
24        _LINE_);
25     const char *cur = str + n; // ukazatelová aritmetika
26     char *dst = ret;
27     while (str && cur != str) { // kontrola str!
28         *dst = *--cur;
29         dst += 1;
30     }
31     *dst = '\0'; //ret je vždy nejméně 1 byte dlouhý.
32     return ret;
33 }
```

- Funkce `strev()` vytváří nový řetězec, proto můžeme bezpečně předat ukazatel na textový literál.
- Volání `strev()` vrátí textový řetězec, nebo končí chybou (volání `myMalloc()`).
- Proměnná `str` tak vždy ukazuje na paměť, ve které je nejméně jeden znak a to '\0'.
- Program tak v rámci `main()` vždy skončí úspěšně `EXIT_SUCCESS`.
- Ve funkci `main()` tak vlastně ani explicitně neřešíme návratové hodnoty volání.
- Ve funkci explicitně ověřujeme, že vstupní řetězec není `NULL`.
- V naší implementaci je prázdný (`NULL`) řetězec ekvivalentní s řetězcem o délce nula.
- Pokud je `str == NULL`, není hodnota `cur` validní.
- Proto ve `while` cyklu explicitně testujeme `str`.
- Z hlediska efektivity bychom mohli volání funkce v případě `str == NULL` ukončit dříve.
- Nicméně volíme přehlednost, menší počet řádků a jediný `return` ve funkci.

Jan Faigl, 2024 BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy 58 / 54

toupper() strev() strvc() streplit() Knihovna strings.h „String objekt“

## Kódovací příklad – Textové řetězce – strvc() 1/2

- Implementujeme funkci, která vrátí počet slov v řetězci.
- Slovo interpretujeme jako souvislou sekvenci znaků vyhovující funkci `isalpha()` z knihovny `ctype.h`.

```
1 #include <stdio.h>
2 #include <stdlib.h>
3 #include <ctype.h>
4 int main(void)
5 {
6     int c, wc = 0;
7     bool inword = false;
8     while ((c = getchar()) != EOF) {
9         if (isalpha(c)) {
10            if (!inword) {
11                inword = true;
12                wc += 1;
13            }
14        } else {
15            inword = false;
16        }
17    }
18    printf("Input contains %d words.\n", wc);
19    return EXIT_SUCCESS;
20 }
```

```
1 $ cat in.txt
2 I like prp!
3 $ clang -g wc.c && ./a.out < in.txt
4 Input contains 3 words.
5
6 ■ Po počátečním ovládní implementujeme funkci strvc().
7
8 int strvc(const char *str)
9 {
10    int wc = 0;
11    bool inword = false;
12    const char *cur = str;
13    while (cur && *cur != '\0') {
14        if (isalpha(*cur)) {
15            if (!inword) {
16                inword = true;
17                wc += 1;
18            }
19        } else {
20            inword = false;
21        }
22        printf("Input contains %d words.\n", wc);
23        return EXIT_SUCCESS;
24    }
25 }
26
27 ■ Řádky 14–17 můžeme nahradit následujícím řádkem.
28 inword && (c >= 'a' && c <= 'z');
```

- Řádky 14–17 můžeme nahradit následujícím řádkem.

Jan Faigl, 2024 BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy 60 / 54

Jan Faigl, 2024 BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy 61 / 54

Jan Faigl, 2024 BOB36PRP – Přednáška 06: Struktury, uniony a číselné typy 63 / 54

### Kódovací příklad – Textové řetězce – strvc() 2/2

- Čtení znaků ze stdin funkcí `getchar()` nahradíme voláním `getline()` z `stdlib.h`. Viz man `getline`. `size_t getline(char ** restrict linep, size_t * restrict linecap, FILE * restrict stream);`

```

1 #include <stdio.h>
2 #include <stdlib.h>
3 #include <stdbool.h>
4 #include <ctype.h>
5
6 int strvc(const char *str);
7 int main(void)
8 {
9     char *line = NULL; // nezbytné k alokaci v getline()
10    size_t cap = 0; // alokovaná kapacita v getline()
11    // getline vrací -1 při chybě, proto size_t
12    ssize_t l = getline(&line, &cap, stdin);
13    int wc = strvc(line);
14    fprintf(stderr, "DEBUG: Read line \"%s\" that is %i\n",
15            long stored in %i bytes.\n", line, l, cap);
16    printf("Input contains %d words.\n", wc);
17    free(line); // proměnná je alokována dynamicky.
18    return EXIT_SUCCESS;
19 }

```

- Náčetý řetězec obsahuje 11 znaků, konec řádku, a '\0'.
- Celkem funkce `getline()` alokovala 16 bytů.
- Program můžeme upravit pro načítání souboru voláním `fopen()`.

```

1 #include <stdio.h>
2 #include <stdlib.h>
3 #include <stdbool.h>
4 #include <ctype.h>
5 #include <string.h>
6 #include <my_malloc.h>
7
8 int main(int argc, char *argv[])
9 {
10    char *line = NULL; // nezbytné k alokaci v getline()
11    FILE *fd = argc > 1 ? fopen(argv[1], "r") : NULL;
12    size_t cap = 0; // alokována kapacita v getline()
13    ssize_t l = getline(&line, &cap, fd ? fd : stdin);
14
15    $ clang -g wc-file.c && ./a.out in.txt
16    DEBUG: Read line "I like prp!"
17    * that is 12 long stored in 16 bytes.
18    Input contains 3 words.
19
20    V uvedeném příkladu ztrácíme informaci o chybě načtení souboru.
21    Je vhodné explicitně reagovat.
22    V programu netestujeme interpunkční znaménka, která jsou součástí
23    slov, ani předložky. Funkcionality implementujel!

```

- Funkce `getline()` načítá řádek ze souboru, argument `FILE * restrict stream`, používáme `stdin`.
- Funkce načte řádek včetně oddělovače řádků, tj. '\n'.

### Kódovací příklad – Knihovna – strings.h

- Implementované funkce `toupper()`, `strrev()`, `strvc()`, `strsplit()` vložíme do knihovny `strings.h` a `strings.c`.
- Do knihovny vložíme lokální verzi funkce `myMalloc()`, kterou definujeme jako `static` v souboru `strings.c`.

```

1 #ifndef _STRINGS_H_
2 #define _STRINGS_H_
3 #include <stdbool.h> // Protože bool v strsplit()
4 char* strtoupper(const char *str);
5 char* strrev(const char *str);
6 bool strsplit(const char *str, const char *delim, char
7             **s1, char **s2);
8 #endif
9
10 #include <stdio.h>
11 #include <stdlib.h>
12 #include <string.h>
13 #include <ctype.h>
14 #include <stdbool.h>
15 #include <strings.h>
16
17 static void* myMalloc(size_t size, const char *filename,
18 int line) { ... } // folded
19 char* strtoupper(const char *str) { ... } // folded
20 char* strrev(const char *str) { ... } // folded
21 int strvc(const char *str) { ... } // folded strings.c

```

```

1 #include <stdio.h>
2 #include <stdlib.h>
3 #include <string.h>
4 #include <stdbool.h>
5 #include "my_malloc.h"
6
7 int main(void)
8 {
9     char *line = NULL;
10    size_t cap = 0;
11    ssize_t l = getline(&line, &cap, stdin); //see getline
12    int wc = strvc(line);
13    fprintf(stderr, "DEBUG: Read line \"%s\" that is %i\n long stored in
14            %i bytes.\n", line, l, cap);
15    printf("Input contains %d words.\n", wc);
16    return EXIT_SUCCESS;
17 }

```

- `$ clang -Wall -c strings.c -o strings.o`
- `$ ar -rcs libstrings.a strings.o`
- `$ clang demo-wc.c -lstrings -L. -o demo-wc`
- `$ ./demo-wc < in.txt`
- DEBUG: Read line "I like prp!"
- \* that is 12 long stored in 16 bytes.
- Input contains 3 words.

### Kódovací příklad – Textové řetězce – strsplit() 1/2

- Implementujme funkci, která rozdělí daný řetězec na dva dle zadaného řetězce. **Všimněte si rozdílu ukazatelů!**

```

1 #include <stdio.h>
2 #include <stdlib.h>
3 #include <string.h>
4 #include "my_malloc.h"
5
6 int main(void)
7 {
8     const char *str = "I like programming and PRP
9     especially!";
10    char *s1, *s2;
11    char *delim = "and";
12    char *s = strstr(str, delim);
13    s1 = s2 = NULL;
14    if (!s) { // poděťezec (little) nalezen (v big)
15        fprintf(stderr, "D: str %i\n", strlen(str));
16        fprintf(stderr, "D: delim %i\n", strlen(delim));
17        fprintf(stderr, "D: s %i\n", strlen(s));
18        fprintf(stderr, "D: (s - str) %i\n", s - str);
19        // rozdíl ukazatelů. Oba odkazují do identického
20        // souvislého bloku paměti.
21        size_t n1 = strlen(str) - strlen(s);
22        size_t n2 = strlen(s);
23
24    }
25
26    s1 = myMalloc((n1 + 1) * sizeof(char), __FILE__, __LINE__);
27    s2 = myMalloc((n2 + 1) * sizeof(char), __FILE__, __LINE__);
28    strncpy(s1, str, n1); // Kopírujeme nejvýše n1 znaků
29    strncpy(s2, s, n2); // Kopírujeme nejvýše n2 znaků (a '\0')
30
31    printf("String: \"%s\"\n", str); // Vstupní řetězec
32    printf("s1: \"%s\"\n", s1); // 1. část
33    printf("s2: \"%s\"\n", s2); // 2. část
34
35    free(s1); // volání free(NULL) je v pořádku
36    free(s2); // program končí, nemusíme nastavovat s1 = s2 = NULL
37    return EXIT_SUCCESS;
38 }

```

- Při implementaci použijeme ladící výstupy na `stderr`.
- Program ovládáme a přepíšeme do funkce.

```

1 $ clang strsplit.c my_malloc.c && ./a.out
2 D: str 38
3 D: delim 3
4 D: s 19
5 D: (s - str): 19
6 String: "I like programming and PRP especially!"
7 s1: "I like programming "
8 s2: "and PRP especially!"

```

- Začátek řetězce v řetězci najdeme funkcí `strstr()`.

```

char* strstr(const char *big, const char *little)
    Viz man strstr.

```

### Kódovací příklad – „String objekt“

- S využitím složeného typu a ukazatele na funkci implementujeme variantu objektu textového řetězce.

```

1 #include <stdio.h>
2 #include <stdlib.h>
3 #include <string.h>
4 #include <stdbool.h>
5 #include "my_malloc.h"
6
7 typedef struct string {
8     char *str;
9     size_t len;
10    size_t (getLength)(const char *);
11 } string;
12
13 bool string_create(struct string *s, const char **);
14 void string_destroy(struct string *s);
15
16 int main(void)
17 {
18     string str = { .str = NULL, .len = 0, .getLength = &strlen };
19     string_create(&str, "I like PRP!");
20
21     string str2 = { .str = NULL, .len = 0, .getLength = &strlen };
22     string_create(&str2, "I like PRP!");
23     printf("String str: \"%s\"\n", string.str);
24     printf("String length is %i\n", string.getLength(string.str));
25     printf("strlen length is %i\n", strlen(string.str));
26     string_destroy(&str);
27     string_destroy(&str2);
28     return EXIT_SUCCESS;
29 }

```

```

1 bool string_create(struct string *s, const char **v)
2 {
3     if (!s) {
4         return false;
5     }
6     s->len = strlen(v);
7     s->str = myMalloc((s->len + 1) * sizeof(char),
8         __FILE__, __LINE__);
9     strncpy(s->str, v, s->len);
10    return true;
11 }
12
13 void string_destroy(struct string *s)
14 {
15     if (s) {
16         free(s->str);
17         s->len = 0;
18     }
19 }

```

- `$ clang strobj.c my_malloc.c && ./a.out`
- String str: "I like PRP!"
- String length is 11
- strlen length is 11

### Kódovací příklad – Textové řetězce – strsplit() 2/2

```

1 #include <stdio.h>
2 #include <stdlib.h>
3 #include <string.h>
4 #include <stdbool.h>
5 #include "my_malloc.h"
6
7 bool strsplit(const char *str, const char *delim, char
8             **s1, char **s2);
9
10 int main(void)
11 {
12     const char *str = "I like programming and PRP
13     especially!";
14    char *delim = "and";
15    char *s1, *s2;
16    strsplit(str, delim, &s1, &s2);
17    printf("String: \"%s\"\n", str);
18    printf("s1: \"%s\"\n", s1);
19    printf("s2: \"%s\"\n", s2);
20
21    free(s1); // it is ok to call free(NULL);
22    free(s2);
23    return EXIT_SUCCESS;
24 }

```

- Začátek řetězce v řetězci najdeme funkcí `strstr()`.
- Char\* `strstr(const char *big, const char *little)`. Viz man `strstr`.

```

1 bool strsplit(const char *str, const char *delim, char **s1, char **s2)
2 {
3     char *s = NULL;
4     if (
5         !str || !delim || !s1 || !s2 // Inverze podmínka na argumenty
6         || !(s = strstr(str, delim)) // Podřetězec nalezen.
7     ) {
8         return false;
9     }
10
11    size_t l2 = strlen(s); // Předpokládáme null-terminated řetězec.
12    size_t l1 = strlen(str) - l2; // strlen(str) >= 12
13    *s1 = myMalloc((l1 + 1) * sizeof(char), __FILE__, __LINE__);
14    *s2 = myMalloc((l2 + 1) * sizeof(char), __FILE__, __LINE__);
15    strncpy(*s1, str, l1);
16    strncpy(*s2, s, l2);
17    return true;
18 }

```

- `$ clang -g strsplit.c my_malloc.c && ./a.out`
- String: "I like programming and PRP especially!"
- s1: "I like programming "
- s2: "and PRP especially!"

- Funkci `strsplit()` můžeme dále doplnit, např. o rozdělení bez `delim`.