

Struktury a uniony, přesnost výpočtů a vnitřní reprezentace číselných typů

Jan Faigl

Katedra počítačů
Fakulta elektrotechnická
České vysoké učení technické v Praze

Přednáška 07

BOB36PRP – Procedurální programování

Přehled témat

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

Část I

Část 1 – Struktury a uniony

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 **přístupujeme tečkovou notací**.
- K prvkům můžeme přistupovat přes ukazatel operátorem \rightarrow .
- Pro struktury stejného typu je definována operace přiřazení.

```
struct1 = struct2;
```

Pro proměnné typu pole není přímé přiřazení definováno, přiřazení pole je tak nutné realizovat po prvcích.
- Struktury (jako celek) **nelze** porovnávat relačním operátorem $=$.
- Struktura může být funkcí předávána hodnotou i ukazatelem.
- Struktura může být návratovou hodnotou funkce.

Příklad struct – Definice

- Bez zavedení nového typu (**typedef**) je nutné před identifikátor jména struktury uvádět klíčové slovo **struct**.

```
struct record {
    int number;
    double value;
};
typedef struct {
    int n;
    double v;
} item;
record r; /* THIS IS NOT ALLOWED! */
/* Type record is not known */
struct record r; /* Keyword struct is required */
item i; /* type item defined using typedef */
```
- Zavedením nového typu **typedef** můžeme používat typ struktury již bez uvádění klíčového slova **struct**.
lec07/struct.c

Definice jména struktury a typu struktury

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

```
struct record {
    int number;
    double value;
};
```

 - Definujeme identifikátor **record** ve jmeném prostoru struktur.
- Definicí typu **typedef** zavádíme nové jméno typu **record**.

```
typedef struct record record;
```

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

```
typedef struct record {
    int number;
    double value;
} record;
```

Příklad struct – Inicializace

- Struktury:

```
struct record {
    int number;
    double value;
};
typedef struct {
    int n;
    double v;
} item;
```
- Proměnné typu struktura můžeme inicializovat prvek po prvku.

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

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

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

lec07/struct.c

Příklad struct jako parametr funkce

- Struktury můžeme předávat jako parametry funkcí hodnotou.

```
void print_record(struct record rec) {
    printf("record: number(%d), value(%1f)\n",
        rec.number, rec.value);
}
```
- Nebo ukazatelem

```
void print_item(item *v) {
    printf("item: n(%d), v(%1f)\n", v->n, v->v);
}
```
- Při předávání parametru
 - **hodnotou** se vytváří nová proměnná a původní obsah předávané struktury se kopíruje na zásobník;
 - **ukazatelem** se kopíruje pouze hodnota ukazatele (adresa) a pracujeme tak s původní strukturou.lec07/struct.c

Příklad struct – Přiřazení

- Hodnoty proměnné stejného typu struktury můžeme přiřadit operátorem $=$.

```
struct record {
    int number;
    double value;
};
typedef struct {
    int n;
    double v;
} item;
struct record rec1 = { 10, 7.12 };
struct record rec2 = { 5, 13.1 };
item i;
print_record(rec1); /* number(10), value(7.120000) */
print_record(rec2); /* number(5), value(13.100000) */
rec1 = rec2;
i = rec1; /* THIS IS NOT ALLOWED! */
print_record(rec1); /* number(5), value(13.100000) */
```

lec07/struct.c

Struktury – struct Union Příklad

Příklad struct – Přímá kopie paměti

- Jsou-li dvě struktury stejně veliké, můžeme přímo kopírovat obsah příslušné paměťové oblasti.

Například funkci memcopy() z knihovny string.h

```

struct record r = { 7, 21.4};
item i = { 1, 2.3 };
print_record(r); /* number(7), value(21.400000) */
print_item(&i); /* n(1), v(2.300000) */
if (sizeof(i) == sizeof(r)) {
    printf("i and r are of the same size\n");
    memcopy(&i, &r, sizeof(i));
    print_item(&i); /* n(7), v(21.400000) */
}

```

- V tomto případě je interpretace hodnot v obou strukturách identická, obecně tomu však být nemusí.

lec07/struct.c

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, unieony a číselné typy 11 / 62

Struktury – struct Union Příklad

Struktura struct a velikost

- Vnitřní reprezentace struktury nutně nemusí odpovídat součtu velikostí jednotlivých prvků.

```

struct record {
    int number;
    double value;
};
typedef struct {
    int n;
    double v;
} item;
printf("Size of int: %lu size of double: %lu\n", sizeof(int), sizeof(double));
printf("Size of record: %lu\n", sizeof(struct record));
printf("Size of item: %lu\n", sizeof(item));

```

Size of int: 4 size of double: 8
 Size of record: 16
 Size of item: 16

lec07/struct.c

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, unieony a číselné typy 12 / 62

Struktury – struct Union Příklad

Struktura struct a velikost 1/2

- Při kompilaci zpravidla dochází k zarovnání prvků na velikost slova příslušné architektury. *Napr. 8 bytů v případě 64-bitové architektury.*
- Můžeme explicitně předepsat kompaktní paměťovou reprezentaci, např. direktivou `__attribute__((packed))` překladačů `clang` a `gcc`.

```

struct record_packed {
    int n;
    double v;
} __attribute__((packed));

```

lec07/struct.c

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, unieony a číselné typy 13 / 62

Struktury – struct Union Příklad

Struktura struct a velikost 2/2

- Nebo `typedef struct __attribute__((packed)) { int n; double v; } item_packed;`
- Příklad výstupu:


```

printf("Size of int: %lu size of double: %lu\n", sizeof(int), sizeof(double));
printf("record_packed: %lu\n", sizeof(struct record_packed));
printf("item_packed: %lu\n", sizeof(item_packed));

```

 Size of int: 4 size of double: 8
 Size of record_packed: 12
 Size of item_packed: 12
- Zarovnání zpravidla přináší rychlejší přístup do paměti, ale zvyšuje paměťové nároky. <http://www.catb.org/esz/structure-packing>

lec07/struct.c

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, unieony a číselné typy 14 / 62

Struktury – struct Union Příklad

Proměnné se sdílenou pamětí – union

- Union** je množina prvků (proměnných), které nemusí být stejného typu.
- Prvky unionu sdílejí společně stejná paměťová místa. *Prekrývají 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.*

```

union Nums {
    char c;
    int i;
};
Nums nums; /* THIS IS NOT ALLOWED! Type Nums is not known! */
union Nums nums;

```

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, unieony a číselné typy 16 / 62

Struktury – struct Union Příklad

Příklad union 1/2

- Union složený z proměnných typu: `char`, `int` a `double`

```

int main(int argc, char *argv[])
{
    union Numbers {
        char c;
        int i;
        double d;
    };
    printf("size of char %lu\n", sizeof(char));
    printf("size of int %lu\n", sizeof(int));
    printf("size of double %lu\n", sizeof(double));
    printf("size of Numbers %lu\n", sizeof(union Numbers));
    union Numbers numbers;
    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

```

lec07/union.c

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, unieony a číselné typy 17 / 62

Struktury – struct Union Příklad

Příklad union 2/2

- Proměnné sdílejí paměťový prostor.

```

numbers.c = 'a';
printf("\nSet the numbers.c to 'a'\n");
printf("Numbers c: %d i: %d d: %lf\n", numbers.c, numbers.i, numbers.d);
numbers.i = 5;
printf("\nSet the numbers.i to 5\n");
printf("Numbers c: %d i: %d d: %lf\n", numbers.c, numbers.i, numbers.d);
numbers.d = 3.14;
printf("\nSet the numbers.d to 3.14\n");
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

```

lec07/union.c

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, unieony a číselné typy 18 / 62

Struktury – struct Union Příklad

Inicializace union

- Proměnnou typu `union` můžeme inicializovat při definici.

```

union {
    char c;
    int i;
    double d;
} numbers = { 'a' };

```

Pouze první položka (proměnná) může být inicializována.

- V C99 můžeme inicializovat konkrétní položku (proměnnou).

```

union {
    char c;
    int i;
    double d;
} numbers = { .d = 10.3 };

```

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, unieony a číselné typy 19 / 62

Struktury – struct Union Příklad

Příklad struktura, pole a výčtový typ 1/3

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

```

#include <stdio.h>
#include <stdlib.h>
#include <string.h>
enum weekdays { MONDAY, TUESDAY, WEDNESDAY, THURSDAY, FRIDAY };
typedef struct {
    char *name;
    char *abbr; // abbreviation
} week_day_s;
const week_day_s days_en[] = {
    [MONDAY] = { "Monday", "mon" },
    [TUESDAY] = { "Tuesday", "tue" },
    [WEDNESDAY] = { "Wednesday", "wed" },
    [THURSDAY] = { "Thursday", "thr" },
    [FRIDAY] = { "Friday", "fri" },
};

```

lec07/demo-struct.c

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, unieony a číselné typy 21 / 62

Příklad struktura, pole a výčtový typ 2/3

- Připravíme si pole struktur pro konkrétní jazyk.
- 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] = { "Pondeli", "po" },
21     [TUESDAY] = { "Utery", "ut" },
22     [WEDNESDAY] = { "Streda", "st" },
23     [THURSDAY] = { "Cvrtrek", "ct" },
24     [FRIDAY] = { "Patek", "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",
32             _FILE_, _LINE_);
33         return 101;
34     }
35     day_of_week -= 1; // start from 0
36 }
37
38 // lec07/demo-struct.c
```

Přesnost výpočtu 1/2

- Ztráta přesnosti při aritmetických operacích.

Příklad sčítání dvou čísel

```
1 #include <stdio.h>
2
3 int main(void)
4 {
5     double a = 1e+10;
6     double b = 1e-10;
7
8     printf("a : %24.12lf\n", a);
9     printf("b : %24.12lf\n", b);
10    printf("a+b: %24.12lf\n", a + b);
11
12    return 0;
13 }
14
15 clang sum.c && ./a.out
16 a : 10000000000.0000000000000000
17 b : 0.00000000000000000000000000
18 a+b: 10000000000.0000000000000000
19
20 // lec07/sum.c
```

Zdroje a typy chyby

- Chyby matematického modelu - matematická aproximace fyzikální situace.
- Chyby vstupních dat.
- Chyby numerické metody.
- Chyby zaokrouhlovací.

- Absolutní chyba aproximace
 $E(x) = \hat{x} - x$, \hat{x} přesná hodnota, x aproximace.
- Relativní chyba $RE(x) = \frac{\hat{x}-x}{x}$.

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
45 printf("%d %s %s\n",
46     day_of_week,
47     days[day_of_week].name,
48     days[day_of_week].abbr);
49
50 return 0;
51 }
52
53 // lec07/demo-struct.c
```

Přesnost výpočtu 2/2

Příklad dělení dvou čísel

```
1 #include <stdio.h>
2
3 int main(void)
4 {
5     const int number = 100;
6     double dV = 0.0;
7     float fV = 0.0f;
8
9     for (int i = 0; i < number; ++i) {
10        dV += 1.0 / 10.0;
11        fV += 1.0 / 10.0;
12    }
13
14    printf("double value: %lf ", dV);
15    printf("float value: %lf ", fV);
16
17    return 0;
18 }
19
20 clang division.c && ./a.out
21 double value: 10.000000 float value: 10.000002
22
23 // lec07/division.c
```

Podmíněnost numerických úloh

- Podmíněnost úlohy $C_p = \frac{\text{relativní chyba výstupních údajů}}{\text{relativní chyba vstupních údajů}}$.
- Dobře podmíněná úloha $C_p \approx 1$.
- Výpočet je dobře podmíněný, je-li málo citlivý na poruchy ve vstupních datech.
- Numericky stabilní výpočet - vliv zaokrouhlovacích chyb na výsledek je malý.
- Výpočet je stabilní, je-li dobře podmíněný a numericky stabilní.

Část II

Část 2 – Vnitřní reprezentace číselných typů

Přesnost výpočtu - strojová přesnost

- Strojová přesnost ϵ_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 ϵ_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 $N \cdot \epsilon_m$.

Možnosti zvýšení přesnosti

- Reprezentace racionálních čísel - podíl dvou celočíselných hodnot, např. *Homogenní souřadnice*.
- „Libovolná přesnost“ - speciální knihovny, např. *gmp* až do výše volné paměti.
souřadnice x,y - 7511164176768 346868669952 3739567104 ~ 2008.57, 92.76

Informativní

Příklady chyb

- Ariane 5 - 4.6.1996**
 40 sekund po startu explodovala. Datová konverze z 64-bitového desetinné reprezentace na 16-ti bitový znaménkový integer.
http://www.esa.int/esaCP/Pr_33_1996_p_EN.html
- Systém Patriot - 25.2.1991**
 Systémový čas v desetinách sekundy, převod na sekundy realizován dělením 10, registry pouze 24 bitů.
<http://www.ima.umn.edu/~arnold/disasters/patriot.html>
<http://www5.informatik.tu-muenchen.de/~huckle/bugse.html>

Repräsentace 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 pak tuto definici hlídá a volí odpovídající strojové instrukce pro práci s datovými položkami například jako s odpovídajícími číselnými typy.
Např. neceločíselné (float) typy a využití tzv. FPU.

Příklad zápisů stejného čísla v různých soustavách.

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

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*

Datové typy

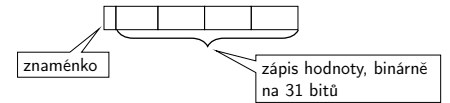
- Při návrhu algoritmu abstrahujeme od binární podoby paměti počítače.
- S daty pracujeme jako s hodnotami různých datových typů, které jsou uloženy v paměti předepsaným způsobem.
- Datový typ specifikuje
 - Množinu hodnot, které je možné v počítači uložit; *Záleží na způsobu reprezentace.*
 - Množinu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

Číselné soustavy

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

Příklad reprezentace celých čísel int

- Na 32-bitových a 64-bitových strojích je celočíselný typ **int** zpravidla reprezentován 32 bity (4 byty).



- Typ **int** je znaménkový typ.
- Znaménko je zakódováno v 1 bitu a vlastní číselná hodnota pak ve zbývajících 31 bitech.
 - Největší číslo je $0111 \dots 111 = 2^{31} - 1 = 2\,147\,483\,647$.
 - Nejmenší číslo je $-2^{31} = -2\,147\,483\,648$.
- Pro zobrazení záporných čísel je použit **doplňkový kód**.
 Nejmenší číslo v doplňkovém kódu 1000...000 je -2^{31} . *Repräsentujeme i 0!*
0 už je zahrnuta.

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 soustave
4
5 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 <lec07/demo-types.c>.

Kódování záporných čísel

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

121 ₍₁₀₎	0111 1001 ₍₂₎
-121 ₍₁₀₎	1111 1001 ₍₂₎
0 ₍₁₀₎	0000 0000 ₍₂₎
-0 ₍₁₀₎	1000 0000 ₍₂₎
- Inverzní kód** – záporné číslo odpovídá bitové negaci kladného čísla. Reprezentace má dvě nuly.

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

121 ₍₁₀₎	0111 1001 ₍₂₎
-121 ₍₁₀₎	1000 0110 ₍₂₎
-121 ₍₁₀₎	1000 0111 ₍₂₎
127 ₍₁₀₎	0111 1111 ₍₂₎
-128 ₍₁₀₎	1000 0000 ₍₂₎
-1 ₍₁₀₎	1111 1111 ₍₂₎

Repräsentace 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

Reprezentace reálných čísel

- Pro uložení čísla vyhrajujeme omezený paměťový prostor.

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

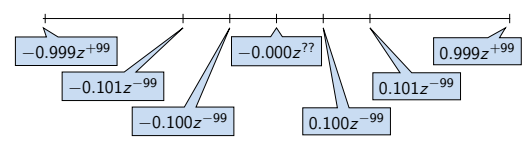
- $= 33333333 \dots 3333$
- $= 0,33$
- $\approx 0,33333333333333333333$
- $\approx 0,333$

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.

Příklad modelu reprezentace reálných čísel 2/2 – Limitní zobrazitelná čísla

- Maximální zobrazitelné kladné číslo $0,999z^{99}$. Maximální zobrazitelné záporné číslo $-0,100z^{-99}$.
- Minimální zobrazitelné kladné číslo $0,100z^{-99}$. Minimální zobrazitelné záporné číslo $-0,999z^{99}$.



Příklad reprezentace float hodnot dle IEEE 754

- Chyba reprezentace -256.75 vs -256.74.
- Infinity (0x7f800000), -Infinity (0xff800000), a NaN (0x7fffffff).

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

Model reprezentace reálných čísel

- Reálná čísla se zobrazují jako aproximace daným rozsahem paměťového místa.
- Reálné číslo x se zobrazuje ve tvaru

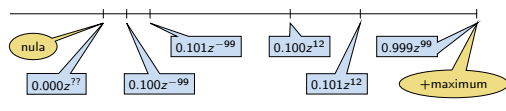
$$x = \text{mantisa} \cdot \text{základ}^{\text{exponent}}$$

- 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.

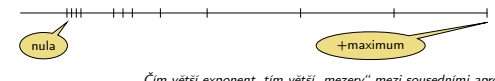


Model reprezentace reálných čísel a vzdálenost mezi aproximacemi

- 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 $\{-99, -98, \dots, 0\}$.



- Aproximace reálných čísel nejsou na číselné ose rovnoměrně rozloženy.



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

Příklady reprezentace hodnot typu float

- Reprezentace čísla 85,125 (float)
 - 85 odpovídá $1010101_{(2)}$.
 - 0,125 odpovídá 001
 - $0,125/2^{-1} = 0,25 \mid 0$
 - $0,125/2^{-2} = 0,50 \mid 0$
 - $0,125/2^{-3} = 1,00 \mid 1$
 - 85,125 odpovídá $1010101,001_{(2)} = 1,010101001_{(2)} \times 2^6$.
 - Bias pro float je 127.
 - Exponent je $127 + 6 = 133$
 - 133 odpovídá $10000101_{(2)}$.
 - Normalizovaná mantisa je $010101001_{(2)}$, kterou doplníme nulami na 23 bitů (zprava, je to desetinné číslo).
 - 0-1000 0101-0101 0100 1000 0000 0000 000.
 - 01000010 10101010 01000000 00000000.
 - V šestnáctkové soustavě $0x42\ 0xaa\ 0x40\ 0x00$, tedy $0x42aa4000$.
- Reprezentace čísla 0,1 (float)
 - 0,1 má periodický rozvoj
 - $0,1 \cdot 2 = 0,2 \mid 0$
 - $0,2 \cdot 2 = 0,4 \mid 0$
 - $0,4 \cdot 2 = 0,8 \mid 0$
 - $0,8 \cdot 2 = 1,6 \mid 1$
 - $0,6 \cdot 2 = 1,2 \mid 1$
 - $0,2 \cdot 2 = 0,4 \mid 0$
 - Opakuje se 0011, 23-bitů tak reprezentuje menší hodnotu.
 - $0,1_{(10)} \sim 0,0001\ 1001\ 1001\ 1001\ 1001\ 1001\ 1001_{(2)} = 1,1001\ 1001\ 1001\ 1001\ 1001\ 1001_{(2)} \times 2^{-4}$.
 - Exponent je $127 - 4 = 123$ odpovídá $0111\ 1011_{(2)}$.
 - Normalizovaná mantisa $\pm 100\ 1100\ 1100\ 1100\ 1100\ 1100$.
 - 0-0111 1011-100 1100 1100 1100 1100 1100.
 - 00111101 11001100 11001100 11001100.
 - V šestnáctkové soustavě to je $0x3d\ 0xcc\ 0xcc\ 0x0c$, tedy $0x3dcccc$.
 - Prakticky je 0,1 převedeno na o něco větší číslo $0x3dcccccd$, protože absolutní chyba je menší.

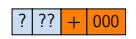
leoc07/floats.c

Příklad modelu reprezentace reálných čísel 1/2

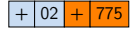
Reprezentace na 7 bajtů se základem 10

- Délka mantisy 3 pozice (bajtů) plus znaménko.
- Délka exponentu 2 pozice plus znaménko.
- Základ $z = 10$.
- Reprezentace nuly.

Reprezentace dle IEEE-754 používá dvojkový základ

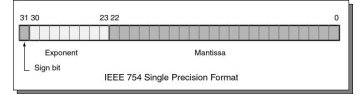


- Příklad $x = 77,5 = 0,775 \cdot z^{+02}$



Reprezentace necelých čísel – IEEE 754

- Reálné číslo x se zobrazuje ve tvaru $x = (-1)^s \text{mantisa} \cdot 2^{\text{exponent} - \text{bias}}$. 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 $\approx 4,5$ bilionů možností (4 503 599 627 370 495).
- bias umožňuje reprezentovat exponent vždy jako kladné číslo.

Lze zvolit, např. bias = $2^{eb-1} - 1$, kde eb je počet bitů exponentu.

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

Přirázovací operátor a příkaz

- Slouží pro nastavení hodnoty proměnné. *Uložení číselné hodnoty do paměti, kterou proměnná reprezentuje.*

Tvar přirázovacího operátoru. $\langle \text{proměnná} \rangle = \langle \text{výraz} \rangle$
Výraz je literál, proměnná, volání funkce, ...

- Zkrácený zápis $\langle \text{proměnná} \rangle (\text{operátor}) = \langle \text{výraz} \rangle$
- Přirazení je výraz **asociativní zprava**.
- Přirázovací příkaz – výraz zakončený středníkem ;

```
int x; //definice promenne x
int y; //definice promenne y
x = 10;
y = 7;
y = x = x + 6;
```

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

```
double x;
int i = 1;

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

- Implicitní konverze je bezpečná.

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, uniony a číselné typy 54 / 62

Explicitní typové konverze

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

Příklad

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

- Explicitní konverze je potenciálně nebezpečná.

Příklady

```
double d = 1e30; // i je -2147483648
int i = (int)d; // to je asi -2e9 místo 1e30

long l = 5000000000L; // i je 705032704
int i = (int)l; // (oříznuté 4 bajty)
```

1e07/demo-type_conversion.c

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, uniony a číselné typy 55 / 62

Konverze primitivních číselných typů

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

rozšíření přiřazením (downward arrow)

zúžení přetyčováním (upward arrow)

znaménko 0/1 ~ +/-

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, uniony a číselné typy 56 / 62

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 funkce 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`
 - `isnan()`, `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*

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, uniony a číselné typy 58 / 62

Část III

Část 3 – Zadání 6. domácího úkolu (HW06)

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, uniony a číselné typy 59 / 62

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 předávání ukazatelů.
- Zadání:** <https://cw.fel.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.

- Termín odevzdání:** 27.11.2021, 23:59:59 PST.
- Bonusová úloha:** 08.01.2022, 23:59:59 PST.

PST – Pacific Standard Time

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, uniony a číselné typy 60 / 62

Shrnutí přednášky

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, uniony a číselné typy 61 / 62

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. Rekurze.

Jan Faigl, 2021 B0B36PRP – Přednáška 07: Struktury, uniony a číselné typy 62 / 62