

Přesnost a rychlost výpočtu

Jan Faigl

Katedra počítačů
Fakulta elektrotechnická
České vysoké učení technické v Praze

Přednáška 13

BOB36PRP – Procedurální programování

Přehled témat

- Část 1 – Přesnost výpočtu
Přesnost výpočtů a numerická stability
- Část 2 – Rychlost výpočtu (programu)
Maticové násobení
Rychlost výpočtu
Comparing C to Machine Code
Paralelní výpočet
- Část 3 – Implementace domácích úkolů
Diskutovaná témata

Přesnost výpočtů

Část I

Část 1 – Přesnost výpočtu

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 1 / 38

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 2 / 38

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 3 / 38

Přesnost výpočtů

Přesnost výpočtů

Přesnost výpočtů

Přesnost výpočtu - Příklad součtu 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 : %24.121f\n", a);
9     printf("b : %24.121f\n", b);
10    printf("a+b: %24.121f\n", a + b);
11
12    return 0;
13 }
14
15 clang sum.c && ./a.out
16 a : 10000000000.000000000000
17 b : 0.000000000000100
18 a+b: 10000000000.000000000000
```

lec13/sum.c

Přesnost výpočtu - 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
```

lec13/division.c

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 5 / 38

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 6 / 38

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 7 / 38

Přesnost výpočtů

Přesnost výpočtů

Přesnost výpočtů

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 $RE(x) = \frac{\hat{x} - x}{x}$.

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.

Možnosti zvýšení přesnosti

- Reprezentace racionálních čísel - podíl dvou celočíselných hodnot, např. *Homogenní souřadnice*.
- „Libovolná přesnost“ - speciální knihovny, např. *gmp* až do výše volné paměti.
<https://gmplib.org/manual/index>
 - Souřadnice x,y - 7511164176768 346868669952 3739567104 ~ 2008,57; 92,76.

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 8 / 38

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 9 / 38

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 10 / 38

Přesnost výpočtů

Součin dvou velkých čísel knihovnou gmp - 1/2

- V HW04B je uveden příklad (995663 - 995669)⁸ jako prvočíselný rozklad čísla
932865073719992059629773513614789388266580305083920591925740371392254317064584855785088915745761.
<https://cv.fel.cvut.cz/wiki/courses/b0b36prp/hw/hw04>
- Použijme knihovnu gmp pro mocninu a součin dvou čísel, #include<gmp.h>.
 - Typ celých čísel mpz_t, pomocné funkce mpz_init_set_str(), mpz_init(), gmp_printf() a mpz_clears() a operace mpz_pow_ui() a mpz_mul().
- Mocnina unsigned integer a násobení - multiplication. Knihovna nemusí být součástí operačního systému, proto může být nutně specifikovat cestu k hlavičkovému souboru a vlastní knihovně (-lgmp).
 - Můžeme zadat cestu ručně při kompilaci (nebo do Makefile).
 - Alternativně můžeme použít nástroj pkg-config (nebo pkgconf).
- Argumenty pro překlad (CFLAGS). Argumenty pro linkování (LDFLAGS).


```
$ pkgconf --cflags gmp -I/usr/local/include
$ pkgconf --libs gmp -L/usr/local/lib -lgmp
```

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 11 / 38

Přesnost výpočtů

Součin dvou velkých čísel knihovnou gmp - 2/2

```
1 #include <stdio.h>
2 #include <stdlib.h>
3
4 #include <gmp.h>
5
6 const char* resultSrc =
7 "932865073719992059629773513614789388266580305083"
8 "920591925740371392254317064584855785088915745761";
9
10 int main(int argc, char *argv[])
11 {
12     int ret = EXIT_SUCCESS;
13     mpz_t n1, n2, result;
14     mpz_init_set_str(n1, "995663", 10);
15     mpz_init_set_str(n2, "995669", 10);
16     mpz_init(result);
17
18     gmp_printf("n1: %Zd\n", n1);
19     gmp_printf("n2: %Zd\n", n2);
20
21     gmp_printf("%Zd * %Zd = %Zd\n", n1, n2, 8);
22
23     mpz_pow_ui(n1, n1, 8);
24     mpz_pow_ui(n2, n2, 8);
25 }
```

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 12 / 38

Přesnost výpočtů

Racionální čísla knihovny gmp - 1/3

- „Libovolné přesnosti“ reprezentace, např. souřadnic v rovině jako výsledek operací výpočetní geometrie, můžeme realizovat podílem dvou („libovolné velkých“) celých čísel.
 - Souřadnice x,y - 7511164176768 346868669952 3739567104 ~ 2008,57; 92,76.
- Knihovna gmp k tomuto účelu poskytuje typ mpq_t, kromě typu necelého čísla mpf_t, který využijeme pro převod mpq_t na celé číslo typu double.


```
49 double mpq2d(const mpq_t *op)
50 {
51     double ret;
52     mpf_t v;
53     mpf_init(v);
54     mpf_set_q(v, *op);
55     ret = mpf_get_d(v);
56     mpf_clear(v);
57     return ret;
58 }
```

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 13 / 38

Přesnost výpočtů

Racionální čísla knihovny gmp - 2/3

```
1 #include <stdio.h>
2 #include <stdlib.h>
3
4 #include <gmp.h>
5
6 double mpq2d(const mpq_t *op);
7
8 int main(int argc, char *argv[])
9 {
10     int ret = EXIT_SUCCESS;
11
12     unsigned long xl = 75111641767681;
13     unsigned long yl = 3468686699521;
14     unsigned long denl = 37395671041;
15     const unsigned int digits = 21;
16
17     double xd = 1. * xl;
18     double yd = 1. * yl;
19     double dend = 1. * denl;
20
21     printf("unsigned long: %Lu %Lu %Lu\n", xl, yl, denl);
22     printf("double: %f %f %f\n", xd, yd, dend);
23
24     printf("double x,y (.2): %f %f\n", xd/dend, yd/dend);
25     printf("double x,y (.46): %f %f\n", xd/dend, yd/dend);
26 }
```

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 14 / 38

Přesnost výpočtů

Racionální čísla knihovny gmp - 3/3

- Souřadnice x,y - 7511164176768 346868669952 3739567104 ~ 2008,57; 92,76.


```
$ make
clang -c -I/usr/local/include -g demo-gmp-mpq.c -o demo-gmp-mpq.o
clang demo-gmp-mpq.o -L/usr/local/lib -lgmp -o demo-gmp-mpq
clang -c -I/usr/local/include -g demo-gmp-mpz.c -o demo-gmp-mpz.o
clang demo-gmp-mpz.o -L/usr/local/lib -lgmp -o demo-gmp-mpz
```

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 15 / 38

Přesnost výpočtů

Makefile s pkg-config a gmp

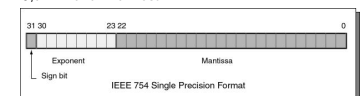
```
1 CFLAGS+=$(shell pkg-config --cflags gmp)
2 LDFLAGS+=$(shell pkg-config --libs gmp)
3
4 CFLAGS+=-g
5
6 DEMO_MPQ=demo-gmp-mpq
7 DEMO_MPZ=demo-gmp-mpz
8
9 TARGETS+=$(DEMO_MPQ) $(DEMO_MPZ)
10
11 bin: $(TARGETS)
12
13 info:
14 @echo $(CFLAGS)
15 @echo $(LDFLAGS)
```

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 16 / 38

Přesnost výpočtů

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 ≈ 16,7 milionu možností.

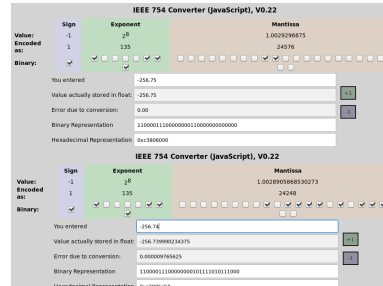


- double – 64 bitů (8 bajtů).
 - s – 1 bit znaménko (+ nebo -).
 - exponent – 11 bitů, tj. 2048 možností.
 - mantisa – 52 bitů ≈ 4,5 biliónu možností.
- bias umožňuje reprezentovat exponent vždy jako kladné číslo. Lze zvolit, např. bias = 2^{eb} - 1 - 1, kde eb je počet bitů exponentu. <http://www.root.cz/clanky/norma-ieee-754-a-pribuzni-formaty-plovouci-radove-tecky>

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 17 / 38

Přesnost výpočtů

Příklad reprezentace float hodnot dle IEEE 754



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

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 18 / 38

Přesnost výpočtů

Příklady reprezentace hodnot typu float

- Reprezentace čísla 85,125 (float)
 - 85 odpovídá 1010101(2).
 - 0,125 odpovídá 001
 - 0,125/2⁻¹ = 0,25 | 0
 - 0,125/2⁻² = 0,50 | 0
 - 0,125/2⁻³ = 1,00 | 1
 - 85,125 odpovídá 1010101,001(2) = 1,010101001(2) × 2⁸.
 - 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 0000.
 - 01000010 10101010 01000000 00000000.
 - V šestnáctkové soustavě to je 0x3d 0xxx 0xxx 0xxx, tedy 0x3dcccc.
- Reprezentace čísla 0,1 (float)
 - 0,1 má periodický rozvoj
 - 0,1 * 2 = 0,2 | 0
 - 0,2 * 2 = 0,4 | 0
 - 0,4 * 2 = 0,8 | 0
 - 0,8 * 2 = 1,6 | 1
 - 0,6 * 2 = 1,2 | 1
 - 0,2 * 2 = 0,4 | 0
 - Opakuje se 0011, 23-bitů tak reprezentuje menší hodnotu.
 - 0,1(10) ~ 0,0001 1001 1001 1001 1001 1001 100(2) = 1,1001 1001 1001 1001 1001 100(2) × 2⁻⁴.
 - Exponent je 127 - 4 = 123 odpovídá 0111 1011(2).
 - Normalizovaná mantisa ±100 1100 1100 1100 1100 1100.
 - 0-0111 1011-100 1100 1100 1100 1100 1100.
 - 0011101 11001100 11001100 11001100.
 - V šestnáctkové soustavě to je 0x3d 0xxx 0xxx 0xxx, tedy 0x3dcccc.
 - Prakticky je 0,1 převedeno na o něco větší číslo 0x3dcccc, protože absolutní chyba je menší.

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 19 / 38

Maticové násobení Rychlost výpočtu Comparing C to Machine Code Paralelní výpočet

Compiler Explorer – Analýza optimalizovaného kódu

- Vliv optimalizace -O2 na výsledný kód, který obsahuje nedefinované chování, přetečení celého císla. Příloha 3. přednášky.

<https://godbolt.org/z/G3GEz4vbv>

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 31 / 38

Maticové násobení Rychlost výpočtu Comparing C to Machine Code Paralelní výpočet

Comparing C to Machine Code

<https://www.youtube.com/watch?v=y0yaJXpAYZQ>

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 33 / 38

Maticové násobení Rychlost výpočtu Comparing C to Machine Code Paralelní výpočet

Příklad použití OpenMP - Maticové násobení 1/2

- Open Multi-Processing (OpenMP) - aplikační programové rozhraní (API) multiplatformních výpočtů se sdílenou pamětí. <http://www.openmp.org>
- Direktivou preprocesoru můžeme instruovat kompilátor k vytvoření kódu paralelního výpočtu, např. paralelizace přes vnější proměnnou `i`.

```

1 void multiply(int n, int a[n][n], int b[n][n], int c[n][n])
2 {
3     int i;
4     #pragma omp parallel private(i)
5     #pragma omp for schedule (dynamic, 1)
6     for (i = 0; i < n; ++i) {
7         for (int j = 0; j < n; ++j) {
8             c[i][j] = 0;
9             for (int k = 0; k < n; ++k) {
10                c[i][j] += a[i][k] * b[k][j];
11            }
12        }
13    }
14 }

```

`lec13/demo-omp-matrix.c`

Pro přehlednost uvažujeme čtvercové matice stejných rozměrů.

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 35 / 38

Maticové násobení Rychlost výpočtu Comparing C to Machine Code Paralelní výpočet

Příklad použití OpenMP - Maticové násobení 2/2

- Příklad násobení matic 1000 x 1000 s využitím OpenMP na iCore5 (2 jádra s HT ~ 4x výpočetní jednotky).
- Násobení matic 5000 x 5000 (Ryzen 9 5950X).

```

1 gcc -std=c99 -O2 -o demo-omp demo-omp-matrix.c -fopenmp
2 ./demo-omp 1000
3 Size of matrices 1000 x 1000 naive
4 multiplication with 0(n^3)
5 c1 == c2: 1
6 Multiplication single core 9.33 sec
7 Multiplication multi-core 4.73 sec
8
9 OMP_NUM_THREADS=2 ./demo-omp 1000
10 Size of matrices 1000 x 1000 naive
11 multiplication with 0(n^3)
12 c1 == c2: 1
13 Multiplication single core 9.48 sec
14 Multiplication multi-core 6.23 sec

```

```

1 OMP_NUM_THREADS=16 ./demo-omp 5000
2 Size of matrices 5000 x 5000 naive
3 multiplication with 0(n^3)
4 Multiplication single core 256.00 sec
5 Multiplication multi-core 18.05 sec

```

`lec13/demo-omp-matrix.c`

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 36 / 38

Diskutovaná témata

Shrnutí přednášky

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 37 / 38

Diskutovaná témata

- Numerická přesnost.
- Knihovna `gmp`.
- Maticové násobení a organizace paměti.
- Rychlost výpočtu a architektura procesoru.
- Paralelní výpočty OpenMP.
- Domácí úkol HW10B.

Jan Faigl, 2024 BOB36PRP – Přednáška 13: Přesnost a rychlost výpočtu 38 / 38