Homework3

李俊霖

2021/2017/09

Homework3

- 题目1
  - Consider the following function
  - typedef unsigned char * byte_pointer;
  - void show_bytes(byte_pointer start, int len) {
    - int i;
    - for (i=0; i<len; i++)
      - printf("%.2x", start[i]);
    - }
  - int val = 0x140A0233;
  - byte_pointer valp = (byte_pointer) & val;

little endian: 小端：最低有效字节最先输出
↳(存在何地址)

Homework3

- 题目1 大端
  - What is the output of the following call to show_bytes on big-endian and little-endian machines respectively?

小端

	little-endian	big-endian
show_bytes(valp, 1);	33	14
show_bytes(valp, 2);	3302	140a
show_bytes(valp, 4);	33020a14	140a0233

Homework3

- 题目2
  - Fill in the missing information in the following table:

Fractional value	Binary representation	Decimal representation
1/8	0.001	0.125
3/4	0.11	0.75
43/16	10.1011	2.6875
25/16	1.1001	1.5625
51/16	11.0011	3.1875

Homework3

- 题目3
  - Given a floating-point format with a k-bit exponent and an n-bit fraction, write formulas for the exponent E, significand M, the fraction f, and the value V for the quantities that follow. In addition, describe the bit representation.
  - A. The number 5.0
  - B. The largest odd integer that can be represented exactly
  - C. The reciprocal of the smallest positive normalized value

倒数　　规范化

Problem 3

$bias = 2^{k-1} - 1$, $V = M \times 2^E$, $e = E + bias$

- For A:
  $5.0 = (0101.0)_2 = (1.01 \times 2^2)_2$
  Thus: $E = 2$, $M = 1.01$, $f = 0.01$, $V = 5.0$, $e = E + bias = 2^{k-1} + 1$.
  bits: `0 10……01 010……`
- For B:
  $M = 1.11111...$, $f = 0.11111...$(n bits of 1), $E = n$, $e = n + bias = n + 2^{k-1} - 1$,
  $V = 1111...$(n+1 bits of 1)
  bits: `0 (n+2^(k-1)-1) 111……`
- For C:
  The smallest positive normalized value is:
  `0 0……01 00……0`
  Thus, $e = 1, bias = 2^{k-1} - 1, M = 1.0, V = 2^{1-bias}$
  The reciprocal is: $V = 2^{bias-1}$
  $E = bias - 1$, $M = 1.0$, $f = 0.0, e = E + bias = 2bias - 1 = 2^k - 3$
  bits:
  `0 11...101 00...`

Homework3

- 题目4
  - Consider the following two 9-bit floating-point representations based on the IEEE floating-point format.
  - Format A
    - There is one sign bit.
    - There are k = 5 exponent bits. The exponent bias is 15.
    - There are n = 3 fraction bits.
  - Format B
    - There is one sign bit.
    - There are k = 4 exponent bits. The exponent bias is 7.
    - There are n = 4 fraction bits.

Homework3

- 题目4
  - Below, you are given some bit patterns in Format A, and your task is to convert them to the closest value in Format B. If rounding is necessary, you should **round toward** $+\infty$.
  - In addition, give the values of numbers given by the Format A and Format B bit patterns. Given these as whole numbers(eg.,17) or as fractions(eg.,17/64 or 17/26).

Homework3

- 题目4

	Format A		Format B	
	Bits	Value	Bits	Value
−	1 01110 001	$\frac{-9}{16}$	1 0110 0010	$\frac{-9}{16}$
+	0 10110 101	208	0 1110 1010	208
−	1 00111 110	$-7/1024$	0 0000 0111	$-7/1024$
+	0 00000 101	$5/2^{17}$	0 0000 0001	1×2^{-10}
−	1 11011 000	-2^{12}	1 1110 1111	−248
+	0 11000 100	$768=3\times2^{8}$	0 1111 0000	$+\infty$