

本节内容

# 浮点数的表示

## 本节总览



# 定点数的局限性

钱包



我的财富：- 8540 ¥

2B 定点整数 short 即可表示



马云的财富：+302657264526 ¥

4B 定点整数 int.....都表示不了

8B long型也  
表示不了

如果换一种货币：1 人民币  $\approx$  10000000000 津巴布韦币

如何在位数不变  
的情况下增加数  
据表示范围？

定点数可表示的数字范围有限，但我们不能无限制地增加数据的长度

# 从科学计数法理解浮点数

普通计数法:

+302657264526

科学计数法:

+3.026 \* 10<sup>11</sup>

+11 +3.026

阶码反映  
数值大小

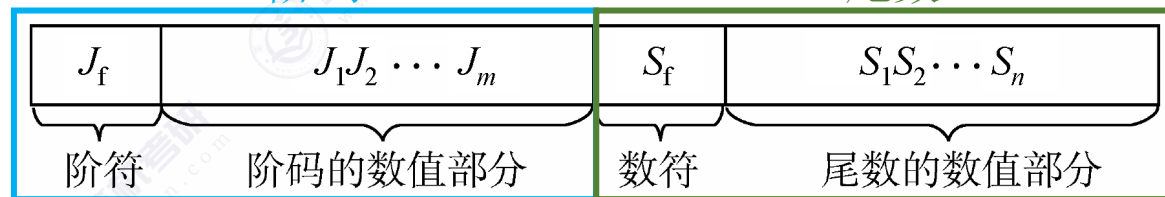
尾数反映  
精度

1 人民币  $\approx 10^{10}$  津巴布韦币

+21 +3.026

阶码

尾数



# 浮点数的表示

r 进制:  $K_n K_{n-1} \dots K_2 K_1 K_0 K_{-1} K_{-2} \dots K_{-m}$

$$= K_n \times r^n + K_{n-1} \times r^{n-1} + \dots + K_2 \times r^2 + K_1 \times r^1 + K_0 \times r^0 + K_{-1} \times r^{-1} + K_{-2} \times r^{-2} + \dots + K_{-m} \times r^{-m}$$

定点数: 如纯小数0.1011和纯整数11110

浮点数:

阶码

尾数

$J_f$	$J_1 J_2 \dots J_m$	$S_f$	$S_1 S_2 \dots S_n$
阶符	阶码的数值部分	数符	尾数的数值部分

阶码: 常用补码或移码表示的定点整数

尾数: 常用原码或补码表示的定点小数

浮点数的真值:  $N = r^E \times M$

阶码的底, 通常为2

类比十进制:  $+302657264526 = +3.026 * 10^{+11}$

可记为:  $+11 +3.026$

阶码E反映浮点数的表示范围及小数点的实际位置;

尾数M的数值部分的位数n反映浮点数的精度。

尾数给出一个小数, 阶码指明了小数点要向前/向后移动几位。

# 浮点数的表示

r 进制:  $K_n K_{n-1} \dots K_2 K_1 K_0 K_{-1} K_{-2} \dots K_{-m}$

$$= K_n \times r^n + K_{n-1} \times r^{n-1} + \dots + K_2 \times r^2 + K_1 \times r^1 + K_0 \times r^0 + K_{-1} \times r^{-1} + K_{-2} \times r^{-2} + \dots + K_{-m} \times r^{-m}$$

定点数: 如纯小数0.1011和纯整数11110

浮点数:

阶码

尾数

$J_f$	$J_1 J_2 \dots J_m$	$S_f$	$S_1 S_2 \dots S_n$
阶符	阶码的数值部分	数符	尾数的数值部分

浮点数的真值:  $N = r^E \times M$

阶码的底, 通常为2

阶码E反映浮点数的表示范围及小数点的实际位置;  
尾数M的数值部分的位数n反映浮点数的精度。

阶码: 常用补码或移码表示的定点整数

尾数: 常用原码或补码表示的定点小数

例: 阶码、尾数均用补码表示, 求a、b的真值

a = 0,01;1.1001

b = 0,10;0.01001

a: 阶码0,01对应真值+1

尾数1.1001对应真值-0.0111 =  $-(2^{-2} + 2^{-3} + 2^{-4})$

a的真值 =  $2^1 \times (-0.0111) = -0.111$

相当于尾数表示的定点小数算数  
左移一位, 或小数点右移一位

1B的存储空间

0	0	1	1	1	0	0	1
---	---	---	---	---	---	---	---

# 浮点数的表示

r 进制:  $K_n K_{n-1} \dots K_2 K_1 K_0 K_{-1} K_{-2} \dots K_{-m}$

$$= K_n \times r^n + K_{n-1} \times r^{n-1} + \dots + K_2 \times r^2 + K_1 \times r^1 + K_0 \times r^0 + K_{-1} \times r^{-1} + K_{-2} \times r^{-2} + \dots + K_{-m} \times r^{-m}$$

定点数: 如纯小数0.1011和纯整数11110

浮点数:

阶码

尾数

$J_f$	$J_1 J_2 \dots J_m$	$S_f$	$S_1 S_2 \dots S_n$
阶符	阶码的数值部分	数符	尾数的数值部分

浮点数的真值:  $N = r^E \times M$

阶码的底, 通常为2

阶码E反映浮点数的表示范围及小数点的实际位置;  
尾数M的数值部分的位数n反映浮点数的精度。

阶码: 常用补码或移码表示的定点整数

尾数: 常用原码或补码表示的定点小数

例: 阶码、尾数均用补码表示, 求a、b的真值

a = 0,01;1.1001

b = 0,10;0.01001

b: 阶码0,10对应真值+2

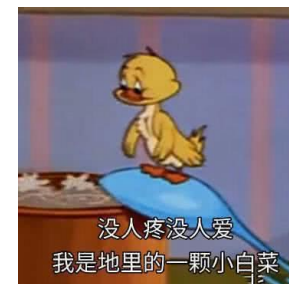
尾数0.01001对应真值+0.01001 = + (2<sup>-2</sup> + 2<sup>-5</sup>)

b的真值 = 2<sup>2</sup> × (+0.01001) = +1.001

相当于尾数表示的定点小数算数  
左移2位, 或小数点右移2位

1B的存储空间

0 1 0 0 0 1 0 0 1



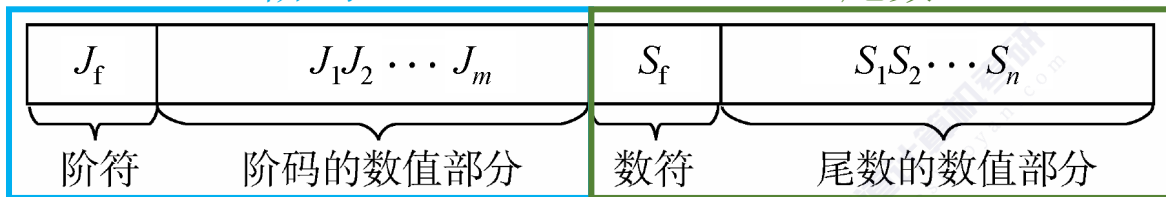


# 浮点数尾数的规格化

浮点数:

阶码

尾数



浮点数的真值:  $N = r^E \times M$

阶码的底, 通常为2

阶码E反映浮点数的表示范围及小数点的实际位置;  
尾数M的数值部分的位数n反映浮点数的精度。

阶码: 常用补码或移码表示的整数

尾数: 常用原码或补码表示的小数

例: 阶码、尾数均用补码表示, 求a、b的真值

a = 0,01;1.1001

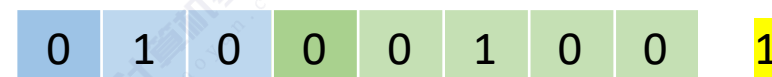
b = 0,10;0.01001

b: 阶码0,10对应真值+2

尾数0.01001对应真值+0.01001 =  $+(2^{-2} + 2^{-5})$

所以b =  $2^2 \times (+0.01001) = +1.001$

1B的存储空间



$$b = 2^2 \times (+0.01001) \\ = 2^1 \times (+0.10010)$$



尾数算数左移1位, 阶码减1。直到尾数最高位是有效值 (左规)

+302657264526 =  $+3.026 \times 10^{+11}$

可记为:  $+11 + 3.026$

也可记为:  $+14 + 0.003$

尾数的最高位是无效值, 会丧失精度



# 浮点数尾数的规格化

通过算数左移、  
阶码减1 来规格化

通过算数右移、  
阶码加1 来规格化

规格化浮点数：规定尾数的最高数值位必须是一个有效值。

左规：当浮点数运算的结果为非规格化时要进行规格化处理，  
将尾数算数左移一位，阶码减1。

右规：当浮点数运算的结果尾数出现溢出（双符号位为01或10）时，  
将尾数算数右移一位，阶码加1。

例：a = 010;00.1100, b = 010;00.1000, 求a+b

$$a = 2^2 \times 00.1100, b = 2^2 \times 00.1000$$

$$a+b = 2^2 \times 00.1100 + 2^2 \times 00.1000$$

$$= 2^2 \times (00.1100 + 00.1000)$$

$$= 2^2 \times 01.0100$$

右规

$$= 2^3 \times 00.1010$$

0	1	1	0	1	0	1	0
---	---	---	---	---	---	---	---

注：采用“双符号位”，当溢出发生时，可以挽救。更高的符号位是正确的符号位

## 规格化浮点数的特点

规格化的原码尾数，最高数值位一定是1

1. 用原码表示的尾数进行规格化：

正数为 $0.1 \times \times \dots \times$ 的形式，其最大值表示为 $0.11\dots 1$ ；最小值表示为 $0.10\dots 0$ 。

尾数的表示范围为 $1/2 \leq M \leq (1-2^{-n})$ 。

负数为 $1.1 \times \times \dots \times$ 的形式，其最大值表示为 $1.10\dots 0$ ；最小值表示为 $1.11\dots 1$ 。

尾数的表示范围为 $-(1-2^{-n}) \leq M \leq -1/2$ 。

规格化的补码尾数，符号位与最高数值位一定相反

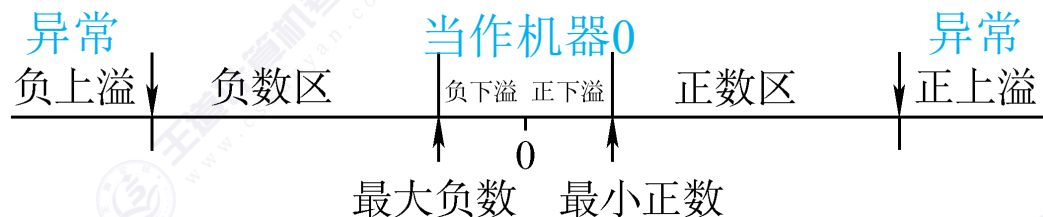
2. 用补码表示的尾数进行规格化：

正数为 $0.1 \times \times \dots \times$ 的形式，其最大值表示为 $0.11\dots 1$ ；最小值表示为 $0.10\dots 0$ 。

尾数的表示范围为 $1/2 \leq M \leq (1-2^{-n})$ 。

负数为 $1.0 \times \times \dots \times$ 的形式，其最大值表示为 $1.01\dots 1$ ；最小值表示为 $1.00\dots 0$ 。

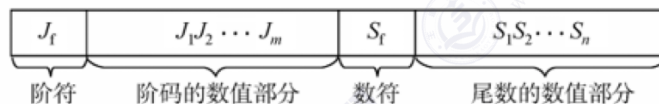
尾数的表示范围为 $-1 \leq M \leq -(1/2+2^{-n})$ 。



eg: 若某浮点数的阶码、尾数用补码表示，共4+8位：  
 $0.110$ ； $1.1110100$  如何规格化？

注：补码算数左移，低位补0；补码算数右移，高位补1。

# 知识点回顾



表示

阶码 + 尾数

尾数给出具体数值，阶码指明小数点前移、后移多少位

阶码通常是用补码、移码表示的定点整数

尾数通常是用补码、原码表示的定点小数

真值

$$N = r^E \times M$$

阶码、尾数用什么码？各取多少位比较合适？

尾数的最高数值位必须是一个有效值（类比十进制科学计数法，通常我们会让数值部分最高位为非0）

左规：数值位最高位无效时，通过尾数算数左移、阶码减1的方法处理，直到尾数最高数值位有效时停止

右规：若采用双符号位表示尾数，则当运算后尾数“假溢出”时，可以通过尾数右移、阶码加1的方法处理

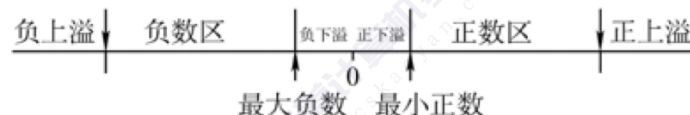
原码表示的尾数规格化：尾数的最高数值位必须是1

补码表示的尾数规格化：尾数最高数值位必须和尾数符号位相反

浮点数的表示

规格化

表示范围





公众号：王道在线



b站：王道计算机教育



抖音：王道计算机考研