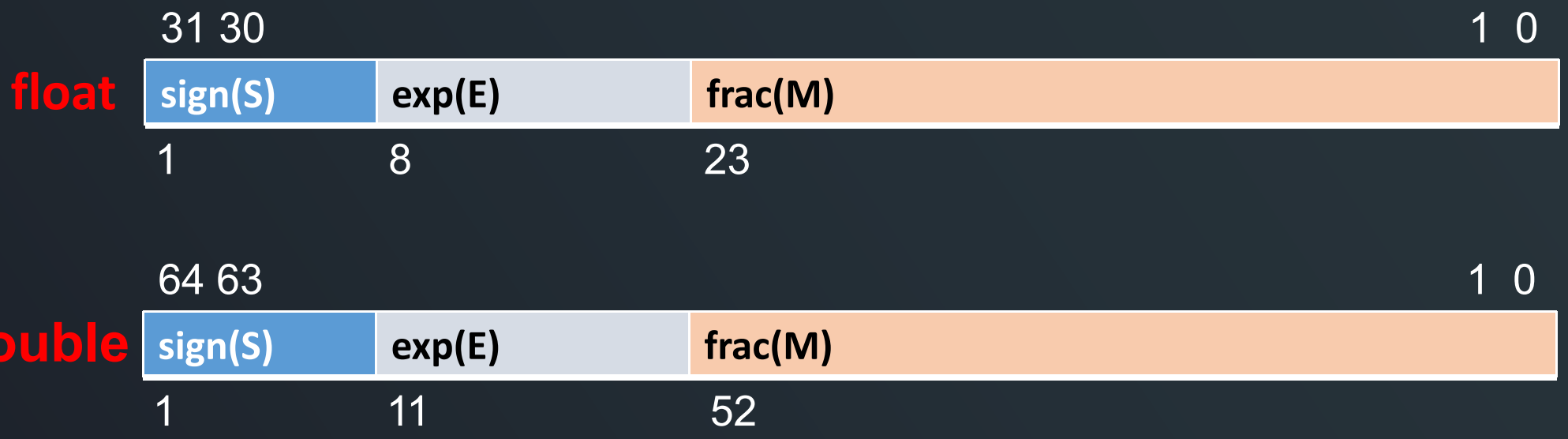
浮点型-存储

华清远见 | 创客学院 小美老师

IEEE754：计算机浮点数格式标准

$$f = (-1)^s * M * 2^E$$

S：符号（0为正 1为负） E：指数 M：尾数（有效数字）



IEEE-754标准规定

- 单精度浮点数的最高位为符号位，S为0，正数，S为1，负数；
- 后面跟8位经偏移的阶码 (移码)，偏移量为 127（科学计数法中，E可能为负数，所以规定，E的真实值需要再加上一个中间数，对于8位的E来说，中间数就是127
- 尾数用原码表示，且把尾数规格化为1.xxx，...x(x为0或1)，并将1去掉，尾数用23位表示。

IEEE-754标准规定

- 双精度浮点数的最高位为符号位
- 后面跟11位经偏移的阶码 (移码)，偏移量为 1023，
- 尾数用原码表示，且把尾数规格化为1.xxx，...x(x为0或1)，并将1去掉，尾数用52位表示。

举例：浮点数float 9.625 在内存中的存储

1.十进制转换为二进制

整数： 9=(1001)B

小数： 0.625 * 2 = 1.25
0.25 * 2 = 0.5
0.5 * 2 = 1.0
0.625 = (.101)B

乘2取整
小数部分为0时，运算结束，若不为0，比如0.14，只能保存近似值

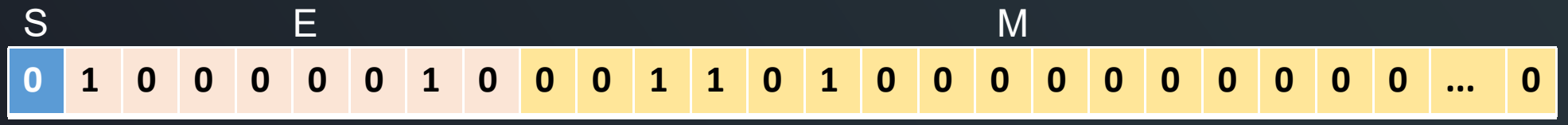
9.625=(1001.101)B

2.确定 S：符号 E:指数 M：尾数

$1001.101=(-1)^0 * 1.001101 * 2^3$

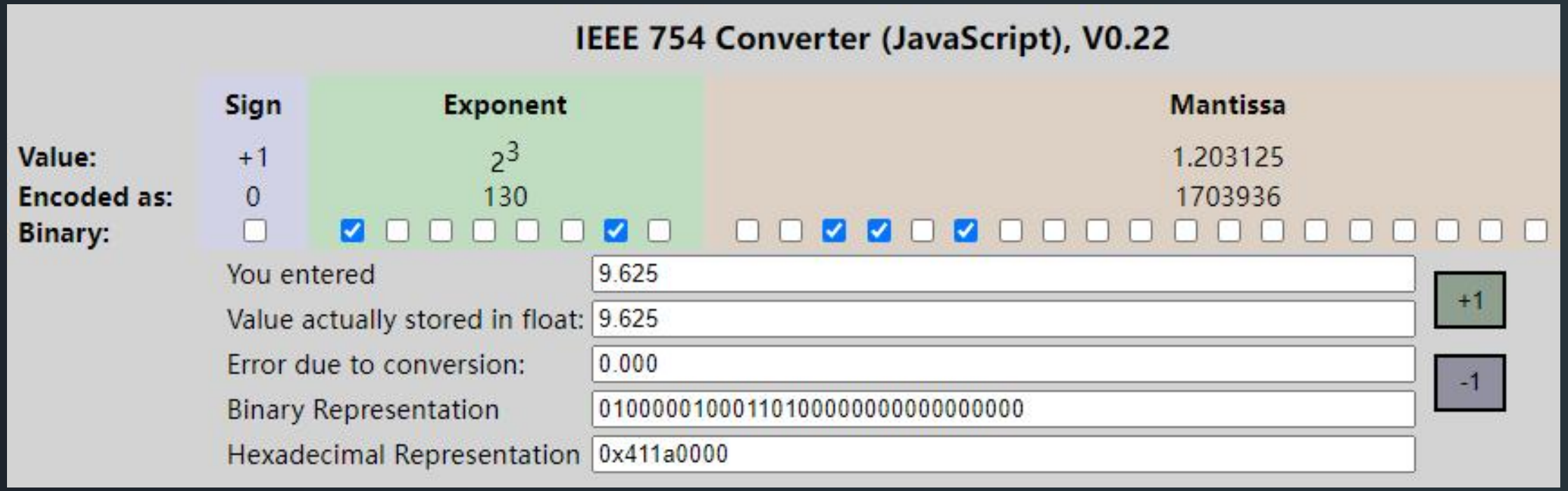
S:0 M:1.001101 E:3

实际存储 S:0 M:去掉整数1，001101 E:加上偏移127,130

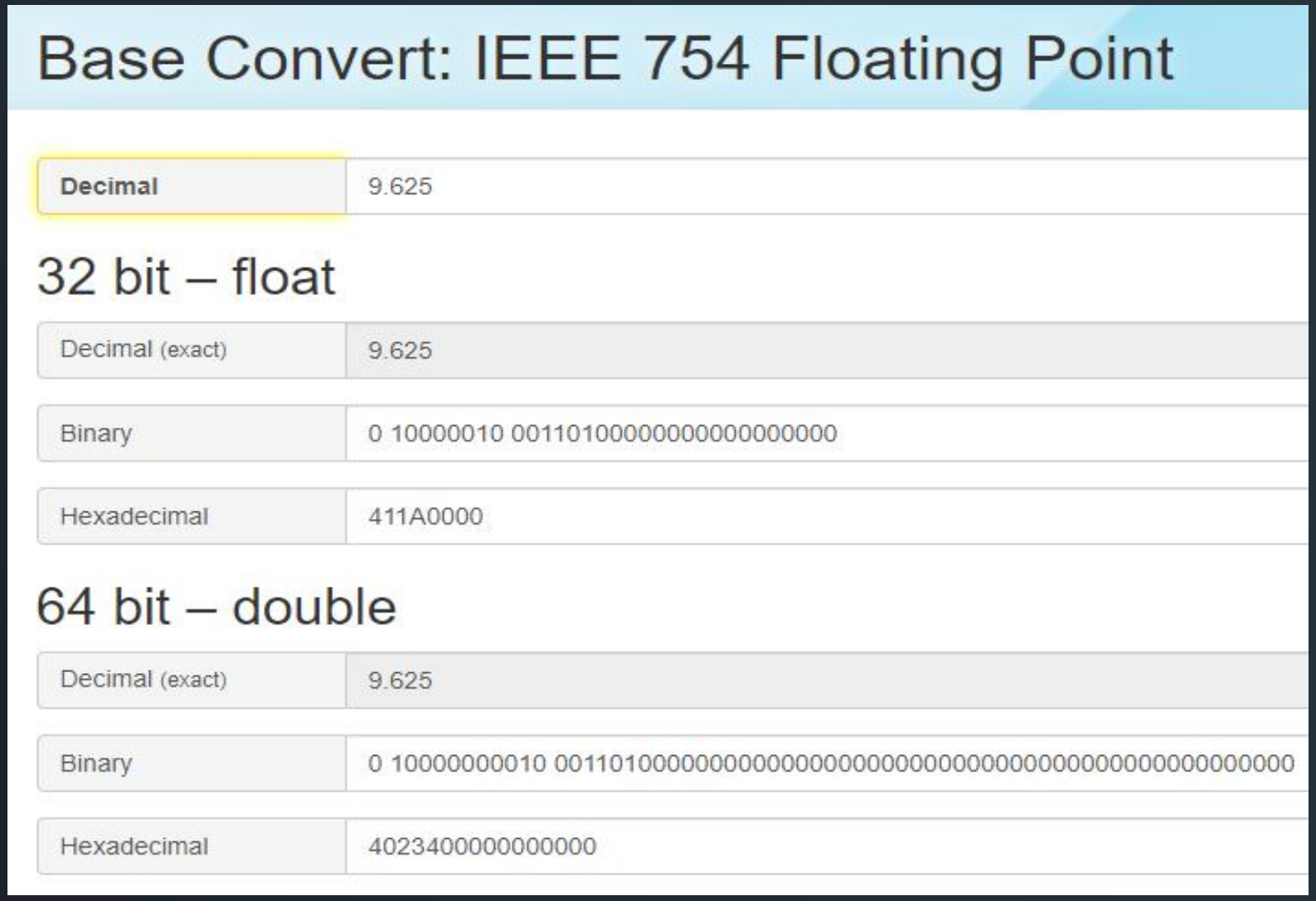


十六进制：0x411A0000

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



https://baseconvert.com/ieee-754-floating-point



程序验证

```
#include <stdio.h>

int main()
{
        float f = 9.625;
        printf("%#x\n", *(int *)&f);

        return 0;
}
```

```
$ gcc float_demo4.c -Wall
$ ./a.out
0x411a0000
```

扫一扫，获取更多信息

THANK YOU