

Lesson 10. Fundamental Data Type — Float & Double

Size of Float, Double & Long Double

For representing fractional numbers (用于表示小数)

Size

Float → 4 bytes = 32 bits

Double → 8 bytes = 64 bits

Long Double → 12 bytes = 96 bits

Size of these types data totally depends from system to system

Brief introduction to fixed and floating point

(定点和浮点简介)

Fixed Point Representation

Example:

Minimum value = -9.99
Maximum value = $+9.99$

Floating point representation:

Example:

Formula = $(0.M) * Base^{Expo}$
Minimum value = $-(0.9) * 10^{+9}$
Maximum value = $+(0.9) * 10^{+9}$

<exponent 指数><mantissa 尾数>

Float, Double & Long Double — difference between them

No more words. Let the code speak it out

```
#include <stdio.h>
int main()
{
    float var1 = 3.1415926535897932;
    double var2 = 3.1415926535897932;
    long double var3 = 3.141592653589793213456;

    printf("%d\n", sizeof(float));
    printf("%d\n", sizeof(double));
    printf("%d\n", sizeof(long double));
    printf("%.16f\n", var1);
    printf("%.16f\n", var2);
    printf("%.21Lf\n", var3);
}
```

"C:\Users\jaspr\Downloads\C programs of dennis ritchie\float_double

```
4
8
12
3.1415927410125732
3.1415926535897931
3.141592653589793213359

Process returned 0 (0x0)   execution time : 0.241 s
Press any key to continue.
```

```
#include <stdio.h>
int main()
{
    int var = 4/9;
    printf("%d\n", var);

    float var1 = 4/9;
    printf("%.2f\n", var1);

    float var2 = 4.0/9.0;
    printf("%.2f\n", var2);
}
```

```
0
0.00
0.44

Process returned 0 (0x0)   execution time : 0.282 s
Press any key to continue.
```