\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Report: HW6\_1

Author: F74046349 梁祐承 <a135246a135246@gmail.com>

Class: 乙班

Description:

Compile Error:

hw6\_1.c: In function ‘main’:

hw6\_1.c:13:2: error: cannot convert to a pointer type

unsigned int \*ipt = &(unsigned int\*)fnum;//sizeof(float) == 4 == sizeof(int), use int pointer

^

put & in wrong place

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Code:

#include<stdio.h>

#include<stdlib.h>

void print\_bit\_float(unsigned int\*);

void print\_bit\_double(unsigned long\*);

int main()

{

int i;

char fc[33], dc[65];

float fnum;

double dnum;

unsigned int \*ipt = (unsigned int\*)&fnum;//sizeof(float) == 4 == sizeof(int), use int pointer

unsigned long \*lpt = (unsigned long\*)&dnum;//sizeof(double) == 8 == siezof(long), use long pointer

printf("Input the float number:");

scanf("%f",&fnum);

print\_bit\_float(ipt);

printf("Input binary to convert to float:");

\*ipt = 0;

scanf("%s",fc);

for(i = 31;i >= 0;i--)

{

if(fc[31 - i] == '1')

\*ipt |= ((unsigned int)1 << i);

}

printf("%e\n", fnum);

printf("Input the double number:");

scanf("%lf",&dnum);

print\_bit\_double(lpt);

printf("Input binary to convert to double:");

\*lpt = 0;

scanf("%s",dc);

for(i = 63;i >= 0;i--)

{

if(dc[63 - i] == '1')

\*lpt |= ((unsigned long)1 << i);

}

printf("%e\n", dnum);

return 0;

}

void print\_bit\_float(unsigned int \*ipt)

{

int i;

for(i = 31;i >= 0;i--)

{

printf("%d",(\*ipt >> i) & 1);

}

printf("\n");

}

void print\_bit\_double(unsigned long \*lpt)

{

int i;

for(i = 63;i >= 0;i--)

{

printf("%ld",(\*lpt >> i) & 1);

}

printf("\n");

}

Compilation:

gcc -o hw6\_1 hw6\_1.c

Execution:

./hw6\_1

Output:

Input the float number:-3.125

11000000010010000000000000000000

Input binary to convert to float:

11000000010010000000000000000000

-3.125000e+00

Input the double number:0.31525

0011111111010100001011010000111001010110000001000001100010010011

Input binary to convert to double:

0011111111010100001011010000111001010110000001000001100010010011

3.152500e-01

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Report: HW6\_2

Author: F74046349 梁祐承 <a135246a135246@gmail.com>

Class: 乙班

Description:

Compile Error:

no errors, but I got warning:

warning: format ‘%d’ expects argument of type ‘int’, but argument 2 has type ‘long unsigned int’ [-Wformat=]

printf("%d",(lnum >> i) & 1);

^

should use %ld

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Code:

#include<stdio.h>

#include<stdlib.h>

void print\_bit\_float(unsigned int);

void print\_bit\_double(unsigned long);

int main()

{

int i;

char fc[33], dc[65];

//use union contains two unsigned int and float

union{

unsigned int inum;

float Field;

}AT;

union{

unsigned long lnum;

double dnum;

}larger;

printf("Input the float number:");

scanf("%f",&AT.Field);

print\_bit\_float(AT.inum);

printf("Input binary to convert to float:");

AT.inum = 0;

scanf("%s",fc);

for(i = 31;i >= 0;i--)

{

if(fc[31 - i] == '1')

AT.inum |= (1 << i);

}

printf("%e\n", AT.Field); //A.T.Field Max Power!!!

printf("Input the double number:");

scanf("%lf",&larger.dnum);

print\_bit\_double(larger.lnum);

printf("Input binary to convert to double:");

larger.lnum = 0;

scanf("%s",dc);

for(i = 63;i >= 0;i--)

{

if(dc[63 - i] == '1')

larger.lnum |= ((unsigned long)1 << i);

}

printf("%e\n", larger.dnum);

return 0;

}

void print\_bit\_float(unsigned int inum)

{

int i;

for(i = 31;i >= 0;i--)

{

printf("%d",(inum >> i) & 1);

}

printf("\n");

}

void print\_bit\_double(unsigned long lnum)

{

int i;

for(i = 63;i >= 0;i--)

{

printf("%ld",(lnum >> i) & 1);

}

printf("\n");

}

Compilation:

gcc -o hw6\_2 hw6\_2.c

Execution:

./hw6\_2

Output:

Input the float number:-974652.031

11001001011011011111001111000000

Input binary to convert to float:

11001001011011011111001111000000

-9.746520e+05

Input the double number:0.1325468

0011111111000000111101110100101100100101100110000011111111111101

Input binary to convert to double:

0011111111000000111101110100101100100101100110000011111111111101

1.325468e-01

2-1:No,in IEEE 754,if underflow occurs,in will represented as denormalized numbers,which smallest is about 1.4e−45

2-2:00000000000000000000000000000000

2-3:No underflow occurs,and f2 < the smallest normalized floating number,so f2 will only store the smallest normalized floating number