有副大學

汇编语言与逆向技术课程实验报告

实验九: re_exercises_advanced



 学
 院
 网络空间安全学院

 专
 业
 信息安全

 学
 号
 2211044

 姓
 名
 陆皓喆

 班
 级
 信息安全

一、实验目的

- 1、进一步熟悉静态反汇编工具 IDA Freeware;
- 2、熟悉将反汇编代码进行反编译的过程:
- 3、掌握对于反编译伪代码的逆向分析;
- 4、运用熟悉的编程语言,实现简单的脚本编写。

二、实验原理

(一) task3

1、通过 IDA Freeware 得到 task3.exe 的反汇编代码,如图 1 和图 2 所示。

```
.text:00402A70
                                 sub
                                          esp, 0A4h
.text:00402A76
                                 mov
                                          eax, ___security_cookie
.text:00402A7B
                                 xor
                                          eax, esp
.text:00402A7D
                                          [esp+0A4h+var_4], eax
                                 mov
                                          [esp+0A4h+var_80], 42h; 'B'
.text:00402A84
                                 mov
.text:00402A89
                                 xor
                                          ecx, ecx
.text:00402A8B
                                          [esp+0A4h+var_7F], 7Eh;
                                 mov
.text:00402A90
                                          [esp+0A4h+var_7E], 77h;
                                          [esp+0A4h+var_7D], 73h;
.text:00402A95
                                 mov
                                          [esp+0A4h+var_7C], 61h; 'a'
.text:00402A9A
                                 mov
                                          [esp+0A4h+var_7B], 77h;
[esp+0A4h+var_7A], 32h;
.text:00402A9F
                                 mov
.text:00402AA4
                                 mov
                                          [esp+0A4h+var_79], 7Bh;
.text:00402AA9
                                 mov
                                          [esp+0A4h+var_78], 7Ch ;
.text:00402AAE
                                 mov
                                          [esp+0A4h+var_77], 62h;
[esp+0A4h+var_76], 67h;
.text:00402AB3
                                 mov
.text:00402AB8
                                 mov
.text:00402ABD
                                          [esp+0A4h+var_75], 66h;
                                 mov
                                          [esp+0A4h+var_74], 32h;
.text:00402AC2
                                 mov
                                          [esp+0A4h+var_73], 73h;
.text:00402AC7
                                          [esp+0A4h+var_72], 32h;
.text:00402ACC
                                 mov
.text:00402AD1
                                 mov
                                          [esp+0A4h+var_71], 61h;
                                          [esp+0A4h+var_70], 66h;
[esp+0A4h+var_6F], 60h;
.text:00402AD6
                                 mov
.text:00402ADB
                                 mov
                                          [esp+0A4h+var_6E], 7Bh;
.text:00402AE0
                                 mov
                                          [esp+0A4h+var_6D], 7Ch ;
.text:00402AE5
                                 mov
                                          [esp+0A4h+var_6C], 75h;
.text:00402AEA
                                 mov
                                          [esp+0A4h+var_6B], 28h;
.text:00402AEF
                                 mov
.text:00402AF4
                                          [esp+0A4h+var_6A], 18h
.text:00402AF9
                                          [esp+0A4h+var_69], 12h
                                 mov
.text:00402AFE
                                 xchg
                                          ax, ax
```

图 1 task3.exe 的反汇编代码

```
[esp+0A4h+var_80], 42h; 'B'
mov
xor
           ecx, ecx [esp+0A4h+var_7F],
mov
           [esp+0A4h+var_7E], 77h;
[esp+0A4h+var_7D], 73h;
[esp+0A4h+var_7C], 61h;
mov
                                               'w
mov
mov
mov
            [esp+0A4h+var_7B], 77h
           [esp+0A4h+var_7A], 32h;
[esp+0A4h+var_79], 7Bh;
[esp+0A4h+var_78], 7Ch;
mov
mov
mov
mov
            [esp+0A4h+var_77], 62h
           [esp+0A4h+var_76], 67h;
[esp+0A4h+var_75], 66h;
[esp+0A4h+var_74], 32h;
mov
mov
mov
mov
            [esp+0A4h+var_73], 73h
           [esp+0A4h+var_72], 32h;
[esp+0A4h+var_71], 61h;
mov
mov
            [esp+0A4h+var_70], 66h;
mov
            [esp+0A4h+var_6F], 60h ;
            [esp+0A4h+var_6E], 7Bh;
[esp+0A4h+var_6D], 7Ch;
mov
mov
mov
            [esp+0A4h+var_6C], 75h;
           [esp+0A4h+var_6B], 28h;
[esp+0A4h+var_6A], 18h
[esp+0A4h+var_69], 12h
mov
mov
mov
xchg
 a
  loc_402B00:
  mov
             al, [esp+ecx+0A4h+var_80]
al, 12h
  xor
  mov
             [esp+ecx+0A4h+var_80], a]
  inc
             ecx
             ecx, 18h
  cmp
  jb
              short loc_402B00
    <mark>⊒⊯</mark>⊞
lea
               eax, [esp+0A4h+var_80]
    push
               eax
               offset _Format ; "%s"
    push
     call
               j__printf
               eax, [esp+0ACh+var_54]
    lea
    push
               eax
    .
push
               offset _Format ; "%s
    call
               j__scanf
               ecx, [esp+0B4h+var_54]
    lea
     add
               esp, 10h
               edx, [ecx+1]
```

图 2 task3.exe 反汇编代码的图形化显示

2、使用 IDA 的反编译功能(F5 快捷键)得到伪代码,如图 3 所示。

```
1 int __cdecl main()
                   unsigned int v0; // ecx
unsigned int v1; // ecx
const char *v2; // eax
int v3; // edx
unsigned int v4; // ecx
unsigned int v4; // ecx
unsigned int v6; // ecx
char v7[8]; // [esp+30ch] [ebp-A4h] BYREF
char v8; // [esp+314h] [ebp-9Ch] BYREF
char v10; // [esp+315h] [ebp-9Bh] BYREF
char v10; // [esp+320h] [ebp-9Bh] BYREF
char v11[12]; // [esp+321h] [ebp-8Fh] BYREF
char v13[20]; // [esp+348h] [ebp-68h]
char v14[80]; // [esp+35Ch] [ebp-54h] BYREF
       10
       14
       16
                      qmemcpy(v12, "B~wsaw2{|bgf2s2af`{|u(", 22);
18
                       v0 = 0;
v12[22] = 24;
v12[23] = 18;
19
 20
21
                      do
      22
                     do

v12[v0++] ^= 0x12u;

while ( v0 < 0x18 );

j_printf("%s", v12);

j_scanf("%s", v14);

if ( strlen(v14) == 20 )
2324
26
27
      28
                            v13[0] = -15;
v3 = 0;
v13[1] = -55;
v13[2] = -31;
v13[3] = -1;
v13[4] = -25;
2930
32
```

图 3 task3.exe 的反编译伪代码

- 3、通过对反汇编命令及反编译伪代码的分析,逆向推理出待输入字符串的计算公式;
- 4、使用熟悉的编程语言(C++、Java、Python 等)对待输入字符串进行计算,完成逆向分析挑战。

Please input a string: Correct!

图 4 逆向分析, 完成 task3 练习

(二) task4

1、通过 IDA Freeware 得到 task4.exe 的反汇编代码,如图 5 和图 6 所示。

```
; CODE XREF: _main_01j
 .text:00401470 _main
                                   proc near
 .text:00401470
 .text:00401470 input
                                   = byte ptr -12Ch
 .text:00401470 target
.text:00401470 var_6C
                                  = byte ptr -0D8h
= dword ptr -6Ch
 .text:00401470 var_68
.text:00401470 var_64
                                  = dword ptr -68h
                                  = dword ptr -64h
 .text:00401470 var_60
                                  = dword ptr -60h
 .text:00401470 var_5C
                                  = dword ptr -5Ch
 .text:00401470 var_58
                                  = word ptr -58h
 .text:00401470 var_54
                                   = byte ptr -54h
 .text:00401470 var_4
                                  = dword ptr -4
 .text:00401470
v.text:00401470
                                            esp, 6Ch
                                   sub
                                            eax, ___security_cookie eax, esp
 .text:00401473
                                   mov
 .text:00401478
                                   xor
 .text:0040147A
                                            [esp+6Ch+var_4], eax
                                   mov
 .text:0040147E
                                            offset _Format ; "Please input a string:\n"
                                   push
 .text:00401483
                                   call
                                            j__printf
 .text:00401488
                                   lea
                                            eax, [esp+70h+var_54]
 .text:0040148C
                                   push
                                            eax
 .text:0040148D
                                            offset aS
                                   push
 .text:00401492
                                            j__scanf
                                   call
 .text:00401497
                                            ecx, [esp+78h+var_54]
                                   lea
                                            esp, OCh
 .text:0040149B
                                   \operatorname{\mathsf{add}}
 .text:0040149E
                                   lea
                                            edx, [ecx+1]
 .text:004014A1
 .text:004014A1 loc_4014A1:
                                                              ; CODE XREF: _main+36↓j
 .text:004014A1
                                   mov
                                            al, [ecx]
 .text:004014A3
                                   inc
                                            ecx
                                            al, al
 .text:004014A4
                                   test
 .text:004014A6
                                            short loc 4014A1
                                   jnz
```

图 5 task4.exe 的反汇编代码

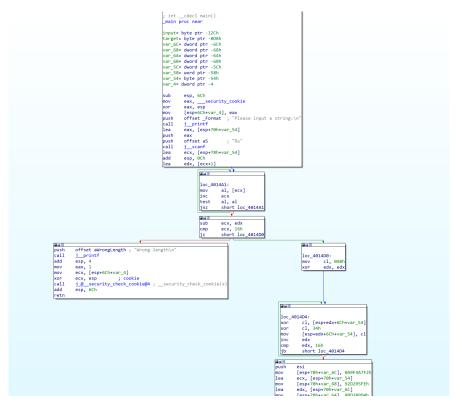


图 6 task4.exe 反汇编代码的图形化显示

2、使用 IDA 的反编译功能(F5 快捷键)得到伪代码,如图 7 所示。右键点击数字对象可实现数制转换。

```
char v1; // cl
unsigned int i; // edx
char *v3; // ecx
int *v4; // edx
unsigned int v5; // esi
         unsigned int vs; // esi
bool v6; // cf
int v7[5]; // [esp+C0h] [ebp-6Ch] BYREF
__int16 v8; // [esp+D4h] [ebp-58h]
char v9[80]; // [esp+D8h] [ebp-54h] BYREF
   11
          j__printf("Please input a string:\n");
j__scanf("%s", v9);
if ( strlen(v9) == 22 )
9 13
• 14
15
   16
17
             v1 = 0xAB;
18
             for ( i = 0; i < 0x16; ++i )
                v1 ^= v9[i] ^ 0x34;
20
21
                v9[i] = v1;
  22
23
             v7[0] = 0xA9F4A7F2;
             v3 = v9;
v7[1] = 0x92D295FE;
2425
2627
             v4 = v7;
v7[2] = 0x80D389D4;
28
29
             v5 = 0x12;
v7[3] = 0xB5E0BCEB;
v7[4] = 0xBEE4B5ED;
9 30
3132
             v8 = 0xBCED;
while ( *(_DWORD *)v3 == *v4 )
34
                v3 += 4;
               ++v4;
v6 = v5 < 4;
v5 -= 4;
3536
9 38
                if ( v6 )
   39
• 40
                   if ( *(_WORD *)v3 == *(_WORD *)v4 )
  41
                      j__printf("Correct");
43
                      return 0;
```

图 7 task4.exe 的反编译伪代码

- 3、通过对反汇编命令及反编译伪代码的分析,逆向推理出待输入字符串的 计算公式
- 4、使用熟悉的编程语言(C++、Java、Python等)对待输入字符串进行计算, 完成逆向分析挑战。

Please input a string: Correct

图 8 逆向分析, 完成 task4 练习

三、实验报告

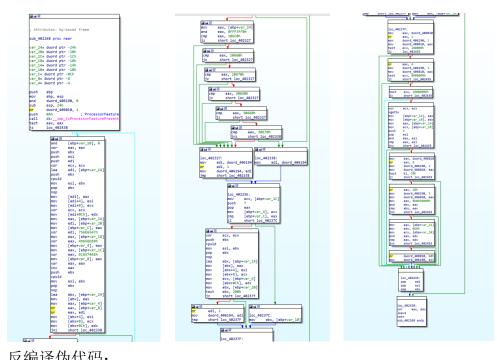
3.1 分别针对 task3、task4 使用 IDA Freeware,获得可执行文件的反汇编代码及反编译伪代码,提供截图。

3.1.1 task3

反汇编代码:

```
.text:0040:1275 or dword_406010, 1
.text:0040:127C push OAh ; ProcessorFeature
.text:0040:1284 test eax, eax
.text:0040:1286 jz loc_402:138 [ebp+var_10], 0
.text:0040:1290 xor eax, eax
.text:0040:1293 push esi
.text:0040:1293 push esi
.text:0040:1293 push esi
.text:0040:1295 xor ecx, ecx
.text:0040:1296 cpuid
.text:0040:1297 lea ebx
.text:0040:1298 push ebx
.text:0040:1298 push esi
.text:0040:1299 push esi
.text:0040:1299 push esi
.text:0040:1299 push ebx
.text:0040:1298 push ebx
.text:0040:
```

反汇编代码图形化显示:



反编译伪代码:

```
| vis[9] = -24; | vis[18] = -17; | vis[18] = -17; | vis[18] = -48; | vis[18] = -57; | vis[1
1 int __cdecl main_0(int argc, const char **argv, const char **envp)
                                    int __odecl main_0(int args, const char **args, of constigued sit v3; // ecc char v3; // ecc v3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                76 }
77 else
78 {
79  v13 = 4;
79  v4 = 0;
80  v4 = 0;
81  qame(py(v14, "!<=4s?6=4";S", sizeof(v14));
82  do
83     v14[v4++ - 1] ^= 0x53u;
84  while ( v4 < 0xD );
85  v5 = &v13;
86 }
87  sub_4011E5("%s\n", (char)v5);
88  return 1;
89 }
                                          do

Anglist[v3++] ^= 0x12u;

while ( v3 < 0x18 );

sub_4011E5("%s", (char)Anglist);

sub_4011Ef("%s", (char)Anglist);

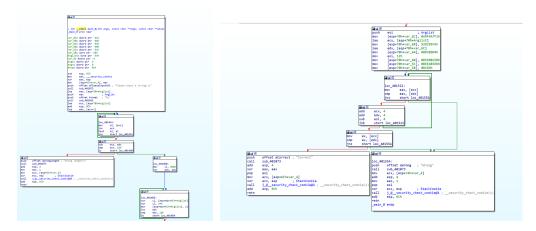
if ( strlen(Anglist) == 20 ) {
                                                             if ( strlen(Argii (
v16[0] = -15;
v6 = 0;
v16[1] = -55;
v16[1] = -31;
v16[3] = -1;
v16[3] = -25;
v16[5] = -109;
v16[6] = -12;
v16[6] = -12;
v16[6] = -12;
v16[8] = -44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           v10[v9++] ^= 0x90u;
while ( v9 < 7 );
v5 = v10:
```

3.1.2 task4

反汇编代码:

```
.text:0040144C
.text:00401450
.text:00401455
.text:00401458
.text:0040145A
.text:00401460
.text:00401463
.text:00401463
                                                              push dword ptr [esp+10h] sub_4011CC push word ptr [eax+4] push dword ptr [eax] call ds: _imp__stdio_core esp, 18h
                                                                align 10h
sub esp, 6Ch
mov eax, _security_cookie
xor eax, esp
mov [esp+Cch+var_4], eax
push offset apleaseInputASt; "Please input a string:\n"
call sb_4e1073
lea eax, [esp+78h+Arglist]
push eax
foffset Format; "%s"
call sb_4e1018
lea ecx. [esp+78h+Arglist]
```

反汇编代码图形化显示:



反编译伪代码:

```
int __cdecl main_0(int argc, const char **argv, const char **envp)
   sub_401073("Please input a string:\n", v12[0]);
sub_40101E("%s", (char)Arglist);
if ( strlen(Arglist) == 22 ) ;
                                                                                                                                      v10 = v9 < 4;
v9 -= 4;
if ( v10 )
       v5 = -85;
for ( i = 0; i < 0x16; ++i )
                                                                                                                                          if ( *(_WORD *)v7 == *(_WORD *)v8 )
          v5 ^= Arglist[i] ^ 0x34;
Arglist[i] = v5;
                                                                                                                    44
• 45
• 46
                                                                                                                                            sub_401073("Correct", v11);
return 0;
                                                                                                                                         break:
                                                                                                                    • 48
        v7 = Arglist;
v12[1] = -1831692802;
                                                                                                                        49
50
       V12[1] = -1831692602;

V8 = V12;

V12[2] = -2133620268;

V9 = 18;

V12[3] = -1243562773;

V12[4] = -1092307475;

V13 = -1717;

while ( *(_DWORD *)V7 == *V8 )
                                                                                                                    50

• 51

• 52

• 53

• 54

• 55

• 56

• 57

• 58

• 59

}
                                                                                                                                     sub 401073("Wrong", v11);
                                                                                                                                  sub_401073("Wrong length\n", v12[0]);
return 1;
```

3.2 分别针对 task3、task4 反编译伪代码的计算过程、数据结构、条件判断等信息进行逆向分析,列出正确输入字符串的计算公式。

3.2.1 task3

观察代码,发现有很多地方都出现了加密的过程,比如说开始时的qmemcpy(ArgList, "B~wsaw2{|bgf2s2af {|u(", 22)这句话就是对字母进行了加密,我们编写程序(具体程序全部放到后面一个版块),来求解字符串原本的含义。得出第一处的含义为"Please input a string"。

接下来是一个输入函数,将输入的字符串存入 Arglist 中,然后判断其长度是否为 20,对其进行相应处理,最后是结果是输出 v5 的字符串,我们先对长度不为 20 的部分(即错误部分)进行分析。

如果长度不为 20 的话,错误部分首先将 v13 赋值 4,将一个字符串赋给 v14,从 0-1 至 0xd-1 进行循环遍历 v14 字符串进行异或,注意到下图 v13 为 esp+14h,v14 起始为 esp+15h,因此,v14[-1]即为 v13,最后将 v5 指向 v13 的 地址,最后输出 v5 的值,所以 v5 输出的即为 v13 开始到 v14 末尾的值。

我们还是用前面的方法解密出字符串,此处字符串的含义为"Wrong length"。

现在对正确的处理逻辑进行分析,当输入的字符串长度为 20 时,先对 v16 进行了一串的赋值操作,然后对输入的数据每一位与 0xA5 进行异或和 v16 每一位进行比较,当遍历到 20 个字符(即输入的字符串已经遍历完毕)且全部匹配时,与之前的 Wrong length 逻辑相同,v11 和 v12 也是相邻的变量,首先将 v11 和 v12 初始化,然后对其每一位与 0x57 进行异或,最后输出结果。

我们先解密出最后的输出结果(猜测是输出 correct),最后解密完发现确实是"correct!"!那么我们就有了思路,发现这一块长度为 20 的代码就是我们要求的那一部分。

还有一处,就是当输出 while 时,那一处遍历出现问题时就会输出 "Wrong!"。

最后,写一个程序来计算出我们的求解结果:只需要用原来的一串数字和 0xA5 进行异或操作,公式如下所示:

$$for(i = 0; i < result().length; i + +) \{$$

$$result[i] = (l[i] \& 0xff) \land 0xA5;$$

$$\}$$

最后得出正确的结果"TLDZB6QJqMJekYGtXeTY"。

3.2.2 task4

先输出一个"Please input a string:",接着输入一个字符,如果字符的长度不等于 22 的话,就进不了 if 循环,直接跳转到 else 界面,输出"wrong length"。现在我们输入一个字符长度为 22 的字符,先给 v5 赋值为-85,接下来,进入for 循环,做 22 次操作,将输入值的每一位都与 0x34 进行异或运算,再跟 v5

进行异或运算,再将值赋值给输入的第几位。做完 22 次运算后,我们进入下一个环节,先将操作完的字符串赋值给 v7,然后将 v12 的五位数字都进行赋值。

所以在逆向编写 C++程序的时候,我们对于 20 次异或运算的操作需要进行 反向的编写,具体思路就是用本位与前一位进行异或,再与 0x34 进行异或。第 一位的话我们用第一位与 0x34 进行异或,然后与-85 进行异或。

最后进行比较环节,就是先初始化 v12, v13,然后让 v7 和 v8 的字节每四位比较一次,如果全部相同,那么就输出"correct",否则就输出"Wrong"。

输入字符串的计算公式为:

$$arg[n] = arg[n]'^{a} arg[n-1]^{a}(0x34), n \ge 2$$

 $arg[n] = arg[n]'^{a}(0x34)^{a}(-85), n = 1$

最后得出正确的结果: "magic string challenge"。

3.3 使用熟悉的编程语言,分别针对 task3、task4 编写脚本,计算出正确的字符串,提供脚本输出结果的截图

3.3.1 task3

计算第一部分的内容的 C++代码:

第一处的结果:

```
Microsoft Visual Studio 调试 × + v

the first answer:Please input a string:
C:\Users\Lenovo\Desktop\C++\??????\x64\Debug\???????.exe (?? 4664)???,??? 0?
????????. . . |
```

第二部分的代码:

第二部分的输出结果:

```
Microsoft Visual Studio 调试 × + | v

the second answer:Wrong length

C:\Users\Lenovo\Desktop\C++\???????\x64\Debug\???????.exe (?? 22028)???,??? 0?

????????. . . |
```

第三部分的代码:

第三部分的结果:

```
Microsoft Visual Studio 调读 × + マ

the third answer:Correct!
C:\Users\Lenovo\Desktop\C++\??????\x64\Debug\???????.exe (?? 16496)???,??? 0?
????????. . .
```

第四部分(while 结构)的代码:

```
| Timelide | Timelide
```

第四部分的结果:



最后的结果求解代码:

求解结果:

```
配 Microsoft Visual Studio 调试 × + v

TlDZB6QJqMJekYGtXeTY

C:\Users\Lenovo\Desktop\C++\??????\x64\Debug\???????.exe(?? 10560)???,??? 0?
????????. . .
```

3.3.2 task4

我们使用 C++程序进行逆向的编写, 求解出最后的结果。

最后得出结果:

```
magic_string_challenge
C:\Users\Lenovo\Desktop\C++\???????\x64\Debug\??????.exe (?? 32940)???,??? 0?
????????. . .
```

3.4 分别运行程序 task3、task4,输入计算得到的字符串进行验证,获得 "Correct"输出,提供截图。

3.4.1 task3

如图,我们输入TIDZB6QjqMJekYGtXeTY,得到结果"Correct!",实验成功!

```
C:\Users\Lenovo>D:\学学学\本科\大二上\汇编与逆向技术基础\2211044_陆皓喆_汇编实验\Lab9_re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises_advanced\Lab9-re_exercises
```

3.4.2 task4

如图,我们输入上面得到的字符串"magic_string_challenge",最后输出结果 "Correct",实验成功!



四、总结与反思

从本次实验中,我学到了更多的逆向分析的知识与方法,处理了两个带有 for 循环、while 循环等复杂结构的逆向分析处理,对逆向分析的思路以及 IDA 的使用更加熟练了,希望之后我还能继续用自己的汇编知识去解决一些更难的问题。