ASSIGNMENT 4

SETF_chall:

Sử dụng IDA Pro để xem mã giả.

Phần bắt đầu chương trình nằm ở sub_1080:

```
unsigned int v6; // er8@6
Function name
                                                    __int64 v7; // [sp+0h] [bp-158h]@1
                                               10
_init_proc
                                               11
                                                             118
                                                                      [sn+80h]
f _puts
f _strlen
                                                    __m128i v9; // [sp+90h] [bp-C8h]@1
__m128i v10; // [sp+A0h] [bp-B8h]@1
                                               12
                                               13
_printf
                                                    __m128i v11; // [sp+B0h] [bp-A8h]@1
                                               14
f _fgets
                                               15
                                                    __m128i v12; // [sp+C0h]
                                                                                 [bp-98h]@1
   _cxa_finalize
                                               16
                                                    __m128i v13; // [sp+D0h]
                                                    __m128i v14; // [sp+E0h]
                                               17
                                                                                 [bp-78h]@1
f start
                                                    __m128i v15; // [sp+F0h] [bp-68h]@1
                                               18
f sub_1250
                                               19
                                                     _m128i v16; // [sp+100h] [bp-58h]@1
f nullsub_1
                                                    __m128i v17; // [sp+110h]
                                               20
                                                                                  [bp-48h]@1
f term proc
                                                    __m128i v18; // [sp+120h] [bp-38h]@1
                                               21
f puts
                                               22
                                                    __m128i v19; // [sp+130h] [bp-28h]@1
f strlen
                                               23
                                                    __m128i v20; // [sp+140h] [bp-18h]@1
f printf
                                               24
__libc_start_main
                                             25
                                                    puts("Hello! Welcome to SEETF. Please enter the flag.");
f fgets
                                                    U8 = _mm_load_si128((const __m1281 ×)&xmmword_20F0);
U9 = _mm_load_si128((const __m128i ×)&xmmword_2100);
                                               26
   _imp__cxa_finalize
                                               27
_gmon_start_
                                                    v10 = _mm_load_si128((const __m128i ×)&xmmword_2110);
                                               29
                                                    v11 = _mm_load_si128((const __m128i *)&xmmword_2120);
                                                    U12 = _mm_load_si128((const __m128i *)&xmmword_2130);
U13 = _mm_load_si128((const __m128i *)&xmmword_2140);
                                               31
                                                    v14 = _mm_load_si128((const __m128i *)&xmmword_2150);
                                               33
                                                    v15 = _mm_load_si128((const __m128i *)&xmmword_2160);
                                                    v16 = _mm_load_si128((const __m128i *)&xmmword_2170);
                                                    v17 = _mm_load_si128((const __m128i *)&xmmword_2180);
                                               36
                                                    v18 = _mm_load_si128((const __m128i *)&xmmword_2190);
                                                            mm_load_si128((const _
                                                                                      _m128i ×)&xmmword_21A0);
                                                    U20 =
                                                           _mm_load_si128((const
                                                                                       _m128i ×)&xmmword_21B0);
                                               38
```

Nhận thấy v8 tới v20 là 1 mảng. Chúng mang những giá trị đặc biệt, có thể sẽ có liên quan tới chuỗi bí mật cần tìm (tạm gọi là flag).

Phân tích đoạn mã giả tiếp theo:

```
39
     fgets((char *)&u7, 128, stdin);
40
     if ( strlen((const char *)&v7) == 53 )
 41
142
       puts("Good work! Your flag is the correct size.")
143
       puts("On to the flag check itself...");
) 44
       v1 = strlen((const char *)&∪7);
        ∪2 = 0LL;
145
146
       03 = 01 - 1;
 47
       do
 48
         U6 = U2;
149
50
          if ( \cup 3 == \cup 2 )
 51
            puts("Success! Go get your points, champ.");
52
53
            return OLL;
 54
          }
55
          04 = 08.m128i_i32[02];
          v5 = v2 ^ (*((_BYTE *)&v7 + v2) + 69);
56
57
          ++v2;
 58
        }
        while ( (BYTE) \cup 4 == (BYTE) \cup 5 );
59
60
        printf("Flag check failed at index: %d", ∪6, ∪2);
61
       result = 1LL;
 62
     }
 63
     else
 64
        printf("Flag wrong. Try again.", 128LL);
65
       result = 1LL;
66
 67
68
     return result;
69}
```

v7 là chuỗi nhập vào. Nếu chuỗi này có chiều dài là 52 (không tính kí tự NULL) thì sẽ so sánh từng kí tự trong v7 với flag bằng vòng lặp do-while.

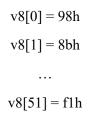
Nếu đã xét xong kí tự cuối cùng thì thông báo đúng flag:

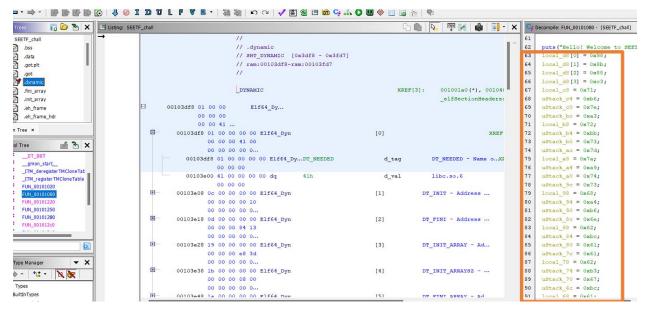
```
46
       v3 = v1 - 1;
47
       do
48
49
        U6 = U2;
50
        if ( U3 == U2 )
51
          puts("Success! Go get your points, champ.");
52
53
          return OLL;
54
         nii - no m1994 429[n9].
55
55
         04 = 08.m128i_i32[02];
56
         v5 = v2 ^ (*((_BYTE *)&v7 + v2) + 69);
57
         ++u2:
```

Trong từng vòng lặp, v2 là biến i, lấy ra từng kí tự của flag (v8) và thực hiện phép tính với v7 (input):

$$flag[i] = I \land (input[i] + 69)$$
 (I)

Sử dụng Ghidra để xem các giá trị của v8. Ta có:





Từ phương trình (I), suy ngược lại:

input
$$[i] = (v8[i] ^ i) - 69$$

Từ đó input cần nhập là: SEE {0n3 5m411 573p 81d215e8b81ae10f1c08168207fba396}

```
(kali kali) - [~/Downloads]
$ ./SEETF_chall
Hello! Welcome to SEETF. Please enter the flag.
SEE{0n3_5m411_573p_81d215e8b81ae10f1c08168207fba396}
Good work! Your flag is the correct size.
On to the flag check itself...
Success! Go get your points, champ.
```

Hard_chall:

Chạy thử chương trình:

```
(kali@ kali)-[~/Downloads]
$ ./hard_chal
[$] Enter your input in the form: words_with_underscores_and_letters: ddddddddd
[$] This won't do...
[$] Enter your input in the form: words_with_underscores_and_letters: dd_aa_bb
[$] Incorrect...
[$] Enter your input in the form: words_with_underscores_and_letters: \[
```

Ta mường tượng được logic của chương trình:

Bước 1: Nhập chuỗi từ bàn phím

Bước 2: Kiểm tra chuỗi có giống format yêu cầu (các ký tự cách nhau bằng dấu gạch dưới). Nếu khác, quay lại bước 1. Nếu giống format, kiểm tra chuỗi nhập có giống với 1 chuỗi cho trước:

Nếu khác, thông báo "không giống" và quay lại bước 1.

Nếu giống, ta quan sát mã giả để xem hành động của chương trình.

Sử dụng IDA Pro để xem mã giả ở hàm main:

```
while ( 020 )
       sub_10F0("[$] Enter your input in the form: words_with_underscores_and_letters: ");
56
57
       sub_1110("%s", &v36);
58
       U21 = 0;
59
       for (i = 0; ; i = s((unsigned int)i))
60
         09 = sub_{1000}(&036);
61
         if ( !(unsigned __int8)1((unsigned int)i, (unsigned int)v9) )
62
63
         if ( (unsigned __int8)eq(*((_BYTE *)&v43 + i - 1056), 95LL) ^ 1 )
v21 = s((unsigned int)v21);
64
65
66
       v10 = sub_10D0(&v36);
if ( (unsigned __int8)ev((unsigned int)v10) ^ 1
II (v11 = sub_10D0(&v36),
67
68
69
70
              v12 = su((unsigned int)v11, 1LL),
         (unsigned __int8)eq(*((_BYTE *)&v43 + v12 - 1056), 95LL))
|| (unsigned __int8)v(&v36) ^ 1
71
72
73
         II (unsigned __int8)ev((unsigned int)v21) ^ 1 )
74
75
         sub_10C0("[$] This won't do...");
76
       }
77
       else
78
         LODWORD(v13) = sub_10D0(&v36);
79
         U25 = U13 - 1;
U14 = 16 × ((U13 + 15) / 0x10uLL);
80
81
82
         while ( &U19 != &U19 )
83
         v15 = alloca(v14 & 0xFFF);
84
85
         if ( v14 & 0xFFF )
            *(__int64 *)((char *)&u19 + (u14 & 0xFFF) - 8) = *(__int64 *)((char *)&u19 + (u14 & 0xFFF) - 8);
86
         v26 = &v19;
         sub_10B0(&019, &036, &036, &019);
```

Chương trình có 1 vòng while để kiểm tra chuỗi nhập vào (biến v36)rồi lặp lại nếu không đúng. Ở lệnh if dòng 68:

```
bb
       v10 = sub_10D0(&∪36);
67
       if ( (unsigned __int8)ev((unsigned int)v10) ^ 1
II (v11 = sub_10D0(&v36),
68
69
70
              v12 = su((unsigned int)v11, 1LL),
         (unsigned __int8)eq(*((_BYTE *)&v43 + v12 - 1056), 95LL))
II (unsigned __int8)v(&v36) ^ 1
71
72
         II (unsigned __int8)ev((unsigned int)v21) ^ 1 )
73
74
75
         sub_10C0("[$] This won't do...");
76
77
       else
78
       {
79
         LODWORD(v13) = sub_10D0(&v36);
         U25 = U13 - 1;
80
         v14 = 16 \times ((v13 + 15) / 0x10uLL);
81
82
         while ( &U19 != &U19 )
83
84
         v15 = alloca(v14 & 0xFFF);
85
         if ( 014 & 0xFFF )
86
           x(__int64 x)((char x)&v19 + (v14 & 0xFFF) - 8) = x(__int64 x)((char x)&v19 + (v14 & 0xFFF) - 8);
87
         026 = &019:
         sub_10B0(&v19, &v36, &v36, &v19);
88
```

Chương trình sẽ in ra "This won't do" nếu chương trình đã so chuỗi nhập vào bằng hình thức nào đó. Nhớ lúc chạy thử chương trình, có thể hiểu đây là trường hợp nhập không đúng format.

Kiểm tra phần else:

Dòng 79 lưu kết quả trả về của hàm sub_10D0 có tham số là chuỗi nhập vào (v36) vào v13. Dòng 80 là v25 = v13 - 1. Điều này làm ta nhớ lại về hàm lấy chiều dài chuỗi, bỏ đi kí tự null cuối chuỗi. Có thể hiểu sub 10D0 là hàm strlen().

Dòng 81 có những con số liên quan tới thập lục phân (16, 15). Từ 84 tới 86 biến v14 được truyền cho hàm alloca, có thể đây là chiều dài của bộ nhớ cấp phát động cho v19.

```
82
         while ( & 113 != & 119 )
83
          v15 = alloca(v14 & 0xFFF);
84
85
          if ( ∪14 & 0xFFF )
86
            *(__int64 *)((char *)& + (v14 & 0xFFF) - 8) = *(__int64 *)((char *)& + (v14 & 0xFFF) - 8);
          v26 = (__int64)&<mark>v19</mark>;
87
          sub_10B0(&<mark>v| 9</mark>, &v36, &v36, &<mark>v19</mark>);
88
          for (j = 0; j = encode((_int64)&v43, v26, j))
89
90
91
            v16 = sub_1000(v26);
92
            if ( !(unsigned __int8)1(j, v16) )
93
94
95
          v17 = sub_1100(v26, "odt_sjtfnb_jc_c_fiajb_he_ciuh_nkn_atvfjp");
96
          if ( (unsigned __int8)eq((unsigned int)v17, OLL) )
97
98
            sub_10C0("[$] Correct!");
99
           U20 = 0:
100
          3
101
          else
102
103
            sub_10C0("[$] Incorrect...");
104
105
       }
106
     027 = 117;
107
108
    U28 = 105:
109 U29 = 117;
    U30 = 99;
110
```

/ UNIQUEE;

Cho tới khi lần cuối v19 xuất hiện, ta thấy nó chỉ có 1 tác động đến chương trình. Đó là ở dòng 87: v26 bằng địa chỉ v19, mà dòng v88 là 1 hàm được truyền 2 tham số là v19 và chuỗi nhập vào. Từ đó có thể hiểu v26 có liên quan tới chuỗi nhập vào qua hàm sub 10B0

```
)
 sub_10B0
                                               ; CODE XREF: main+23D1p
                   proc near
9
                   rep nop edx
ł
                   repne jmp cs:strcpy_ptr
      Nhân thấy sub 10B0 là hàm strcpy, liên kết lai và đúc kết rằng v26 = chuỗi nhập vào.
Dòng 89 – 100:
 89
           for (j = 0; j = encode((_int64)&v43, v26, j))
 90
           {
             v16 = sub_1000(v26);
 91
 92
             if ( !(unsigned __int8)1(j, v16) )
 93
               break;
 94
 95
          v17 = sub_1100(v26, "odt_sjtfnb_jc_c_fiajb_he_ciuh_nkn_atvfjp")
          if ( (unsigned __int8)eq((unsigned int)v17, OLL) )
 96
 97
           {
 98
             sub_10C0("[$] Correct!");
 99
             ∪20 = 0;
100
           }
sub 1100 là hàm stremp
0 sub_1100
                                               ; CODE XREF: main+2A01p
                    proc near
0
                    rep nop edx
                    repne jmp cs:strcmp_ptr
                                      ; DATA XREF: .got:__IIbc_start_main_ptrlo
 int strcmp(const char *s1, const char *s2)
                                      ; DATA XREF: .got:strcmp_ptrfo
              extrn strcmp:near
              extrn __isoc99_scanf:near
                                      ; DATA XREF: .got:__isoc99_scanf_ptrfo
              extrn __cxa_finalize:near ; weak
                                      ; DATA XREF: .got:__cxa_finalize_ptrfo
              extrn _ITM_deregisterTMCloneTable ; weak
                                      ; DATA XREF: .got:_ITM_deregisterTMCloneTable_ptrfo
              extrn __gmon_start__:near ; weak
                                      ; CODE XREF: _init_proc+14fp
                                      ; DATA XREF: .got:__gmon_start___ptrfo
              extrn _ITM_registerTMCloneTable ; weak
```

Dòng 95 (hàm main), strcmp sẽ so sánh 2 chuỗi đầu vào, trả về 0 nếu giống nhau, còn lại trả về 1. Kết quả trả về lưu vào v17.

end _start

; DATA XREF: .got:_ITM_registerTMCloneTable_ptrfo

Dòng 96, so sánh v17 với 0LL bằng hàm eq. Xem qua hàm eq:

```
1 int __fastcall eq(int a1, int a2)
 2 {
 3
    int result; // eax@3
 5
    __asm { rep nop edx } if ( a1 || a2 )
 6
 7
 8
       if ( a1 && a2 )
 9
         result = eq((unsigned int)(a1 - 1), (unsigned int)(a2 - 1));
10
         result = 0;
11
12
     }
13
    else
14
    {
15
       result = 1;
16
17
    return result;
18|}
```

Hàm eq này khá thú vị, sẽ giảm giá trị a và b dần để biết chúng bằng nhau hay không. Nếu a = b: trả về 1, ngược lại trả về 0.

Vậy có thể hiểu rằng v26 so sánh với chuỗi "odt_..._atvfjp". Nếu bằng sẽ in ra correct. Nhưng v26 đã được biến đổi trong hàm for trước đó. Ở đây ta có 3 hàm cần làm rõ là encode(), su() và hàm l ().

Đầu tiên là hàm su:

```
1 int __fastcall su(int a1, unsigned int a2)
2 {
 3
    int result; // eax@2
 5
    __asm { rep nop edx }
 6
    if ( (unsigned __int8)eq(a2, OLL) )
 7
 8
      result = a1;
 9
10
    else if ( (unsigned __int8)eq((unsigned int)a1, OLL) )
11
      result = 0;
12
13
    }
14
    else
15
      result = su((unsigned int)(a1 - 1), a2 - 1);
16
17
18
    return result;
19}
```

Logic hàm su:

Nếu số thứ nhất a1 > số thứ 2 a2: trả về a1 - a2

Nếu không: trả về 0.

Chạy thử chương trình mô phỏng hàm su bằng C (đổi eq bằng dấu ==):

Trong hàm encode, hàm su xuất hiện dạng su(biến a, 1), tương đương a-=1 trong C.

Tiếp theo là hàm l:

```
1 int __fastcall l(unsigned int a1, unsigned int a2)
 2 (
 3
    int result; // eax@3
 4
    int v3; // ebx@8
 5
     int v4; // eax@8
     __asm { rep nop edx }
 7
 8
    if ( (unsigned __int8)eq(a1, OLL) && (unsigned __int8)eq(a2, OLL) )
 9
     {
       result = 0;
10
11
    else if ( (unsigned __int8)eq(a2, OLL) )
12
13
14
       result = 0;
15
     else if ( (unsigned __int8)eq(a1, OLL) )
16
17
18
       result = 1;
19
     }
20
    else
21
22
       v3 = su(a2, 1LL);
23
       v4 = su(a1, 1LL);
24
       result = 1((unsigned int)v4, (unsigned int)v3);
25
26
    return result;
27|}
Logic hàm 1 (a1, a2):
      Nếu a1 = 0 và a2 = 0: trả về 0
      Ngược lại, nếu chỉ a2 = 0: trả về 0
      Ngược lại, nếu chỉ a1 = 0: trả về 1
      Ngược lại: gọi đệ quy hàm l (a1-1, a2-1).
```

Hàm này na ná hàm su() khi so sánh a1 và a2 bằng việc giảm dần giá trị cả 2 biến, hàm l() sẽ trả về 1 nếu a1 < a2, ngược lại là 0. (có thể hiểu hàm l là hàm less)

Kế đến là hàm s:

```
1 __int64 __fastcall s(int a1)
2 {
    __int64 v2; // [sp-8h] [bp-8h]@1
5    __asm { rep nop edx }
6    *((_DWORD *)&v2 - 1) = a1;
7    return (unsigned int)(*((_DWORD *)&v2 - 1) + 1);
8 }
```

Logic hàm s(a): tương đương a++ trong C.

Rồi hàm a(a1, a2):

```
1 int __fastcall a(int a1, unsigned int a2)
  2 (
  3
     int result; // eax@2
     unsigned int ∪3; // ebx@3
     int v4; // eax@3
 7
     __asm { rep nop edx }
 8
     if ( (unsigned __int8)eq(a2, OLL) )
  9
     {
10
       result = a1;
 11
     }
 12
     else
13
       v3 = su(a2, 1LL);
14
15
       v4 = s(a1);
16
       result = a(v4, v3);
17
18
     return result;
19}
```

Logic hàm a:

Nếu a2 = 0: trả về a1

Ngược lại, gọi đệ quy a(a1 + 1, a2 - 1) (như đã phân tích hàm su và hàm s ở trên)

Khi a2 giảm 1 lượng x, đồng nghĩa a1 tăng 1 lượng x. Khi a2 = 0 (tức x = a2) thì trả về a1 + a2. Vậy hàm a(a1, a2) sẽ trả về tổng a1 và a2. (hiểu hàm a là hàm add).

Cuối cùng là hàm m:

```
_int64 __fastcall m(int a1, unsigned int a2)
  2 {
  3
      __int64 result; // rax@2
  4
     unsigned int v3; // eax@3
      __asm { rep nop edx }
  6
  7
      if ( (unsigned __int8)eq(a2, OLL) )
  8
      {
  9
        result = OLL;
 10
      }
 11
      else
 12
13
        v3 = m((unsigned int)a1, a2 - 1);
14
        result = (unsigned int)a(a1, ∪3);
 15
      return result;
16
17|}
```

Ta đã có hàm su (trừ - substract), hàm a (cộng - add), và giờ là hàm m là multiply (phép nhân)

Logic hàm m:

```
Nếu a2 = 0: trả về 0.
```

Ngược lại, trả về tổng của a1 và m(a1, a2 - 1).

Cho tới khi a2 = 0 để m (a1, 0) trả về 0, hàm m sẽ có a2 lần cộng a1, tương đương tích của a1 và a2.

Như thế ta đã phân tích đủ các hàm xuất hiện trong hàm encode. Phân tích hàm encode:

Tại hàm main, encode được truyền 3 tham số là v43, v26 và j

```
89 for ( j = 0; ; j = encode((__int64)&v43, v26, j) )
90 {
v16 = sub_10D0(v26);
```

Xuyên suốt hàm encode, tham số đầu tiên không được sử dụng, tham số v26 là chuỗi nhập vào, j là inex.

Dòng 25 - 27:

2 vòng for trên thực chất sẽ so sánh từng kí tự với dấu gạch dưới (số 95 trong ASC II), bắt đầu từ bên trái, nếu giống thì xét kí tự tiếp theo bên phải cho tới khi gặp kí tự khác kí tự gạch dưới. Kí tự khác kí tự gạch dưới này (tạm gọi k1) sẽ lưu vào v11, rồi kí tự thứ 2 bên phải kí tự k1 mà khác với kí tự gạch dưới sẽ lưu vào v12. Ví dụ: a2 = "a_d" thì v11 lưu vị trí của a, v12 lưu vị trí của d.

```
v11 = *(_BYTE *)(i + a2);
    for (K = 0; (unsigned __int8)::1((unsigned int)k, 5LL); k = s((unsigned int)k))
30
31
    {
      for ( 1 = 0; (unsigned __int8)::1((unsigned int)1, 5LL); 1 = s((unsigned int)1))
32
33
        v3 = ::m((unsigned int)k, 5LL);
34
        v4 = a((unsigned int)v3, (unsigned int)1);
35
        if ( (unsigned __int8)eq((unsigned int)aAbcdefghijklmn[v4], (unsigned int)v11) )
36
37
          v14 = k;
38
          v15 = 1:
39
40
        }
41
      }
42
    v12 = *(_BYTE *)(j + a2);
```

Thoat đầu nhìn tưởng 2 vòng for từ dòng 30 đến 32 là ma trân 5x5, nhưng nhìn kĩ lai thì nó tính

```
index = k * 5 + 1
```

và như thế i thuộc [0;25] có thể biểu diễn 25 phần tử của biến aAbcdefghijklmn, vốn là 25 chữ cái alphabet (**không tính chữ q**).

```
aAbcdefghijklmn db 'abcdefghijklmnoprstuvwxyz',0
; DATA XREF: encode+D9fo
; encode+199fo
```

Nếu aAbcdefghijklmn[index] = kí tự đang xét (v11) thì lưu lại vị trí trong bảng chữ cái thông qua k (v14) và l (v15) (index = v14 * 5 + v15).

Tương tự với kí tự xét thứ 2 (v12), lưu vị trí thông qua v18 và v19.

Phần cuối cùng của hàm encode là thay đổi giá trị chuỗi đầu vào (a2, hay v26):

```
v7 = ::m((unsigned int)v14, 5LL);
*(_BYTE *)(a2 + i) = aVastbcdefghijk[a((unsigned int)v7, (unsigned int)v19)];
v8 = ::m((unsigned int)v18, 5LL);
*(_BYTE *)(a2 + 1) = aCornfieldsabgh[a((unsigned int)v8, (unsigned int)v15)];
return s((unsigned int) 1);
```

Kí tự đầu tiên không phải kí tự gạch dưới thay đổi thành aVastbcdefghijk[v14 * 5 + v19]

Kí tự thứ hai sau kí tự đầu tiên không phải kí tự gạch dưới thay đổi thành aCornfieldsabgh[v18 * 5 + v15]

Ở hàm main, vòng for dòng 89 sẽ chạy tới khi duyệt xong kí tự cuối cùng của v26:

rồi so sánh chuỗi của v26 sau khi encode có giống với chuỗi "odt sitfnb jc c fiajb he ciuh nkn atvfjp"

```
o trong aVast là 17 = 3 * 5 + 2 \Rightarrow v14 = 3, v19 = 2
```

```
d trong aCorn là 8 = 1 * 5 + 3 => v18 = 1, v15 = 3
```

```
=> v14 * 5 + v15 = 3 * 5 + 3 = 18, trong alphabet là t
```

Cứ tiếp tục dịch ngược, kết quả là đoạn v26 cần tìm là the_inside_of_a_field_of_corn_and_dreams

Chạy thử:

```
[$] Enter your input in the form: words_with_underscores_and_letters: the_ins
ide_of_a_field_of_corn_and_dreams
[$] Correct!
[$] uiuctf{the_inside_of_a_field_of_corn_and_dreams}
```