FIRST EDITION – CHAPTER 6 REV 1

Kevin Thomas

Forward

This book is not associated or endorsed by the Rust Foundation. This is purely a FREE educational resource going step-by-step on how to build and reverse engineer Rust binaries.

Rust redefines how a binary is compiled. It does not operate with similar patterns like we see in older languages and is simply the most difficult language to reverse engineer in history.

Many of the C-style decompilers within the popular reversing tools are literally rendered useless with Rust.

Malware Authors are going the route of Rust as there are few tools that exist yet to properly statically analyze the compiled binaries which coupled with its speed and memory safety, it is very difficult to understand what it is actually doing.

The aim of this book is to teach basic Rust and step-by-step reverse engineer each simple binary to understand what is going on under the hood.

We will develop within the Windows architecture (Intel x64 CISC) as most malware targets this platform by orders of magnitude.

In later chapters we will within a Raspberry Pi 64-bit ARM OS so that you can get a perspective of what hacking that architecture looks like in Golang at the binary level.

Let’s begin...

Table Of Contents

Chapter 1: Hello Rust
Chapter 2: Debugging Hello Rust
Chapter 3: Hacking Hello Rust
Chapter 4: Scalar Data Types
Chapter 5: Debugging Scalar Data Types
Chapter 6: Hacking Scalar Data Types

Chapter 1: Hello Rust

We begin our journey with developing a simple hello world program in Rust on a Windows 64-bit OS.

We will then reverse engineer the binary in IDA Free.

Let’s first download Rust for Windows.

https://www.rust-lang.org/tools/install

Let’s download IDA Free.

https://hex-rays.com/ida-free/#download

Let’s download Visual Studio Code which we will use as our integrated development environment.

https://code.visualstudio.com/

Once installed, let’s add the Rust extension within VS Code.

https://marketplace.visualstudio.com/items?itemName=rust-lang.rust-analyzer

It is important to update Rust so I would encourage you to run this before every project.

```
rustup update
```

Let’s create a new project and get started by following the below steps.

```
cargo new one_hello
```

Now let’s populate our **main.rs** file with the following.

```
fn main() {
    println!("Hello, world!");
}
```

Let’s open up the terminal by click CTRL+SHIFT+` and type the following.

```
cargo build
```

Let’s run the binary!

```
.\target\debug\one_hello.exe
```

Output…

```
Hello, world!
```

Congratulations! You just created your first hello world code in Rust. Time for cake!

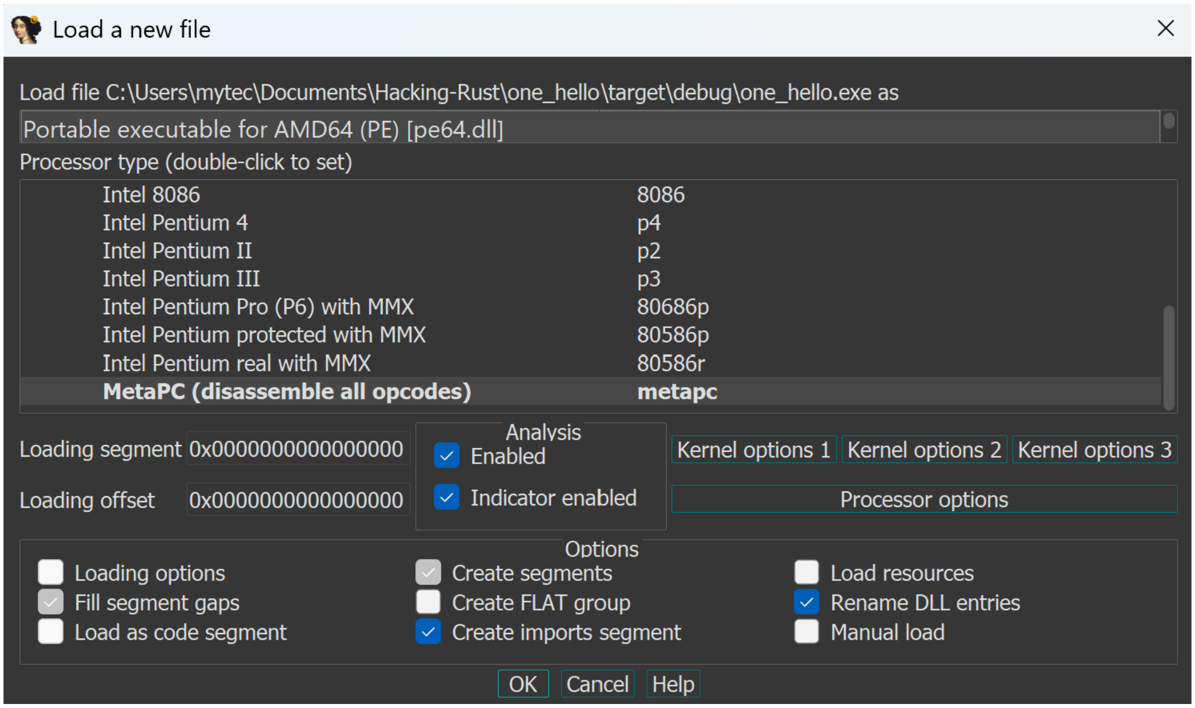
We simply created a hello world style example to get us started.

In our next lesson we will debug this in IDA Free!

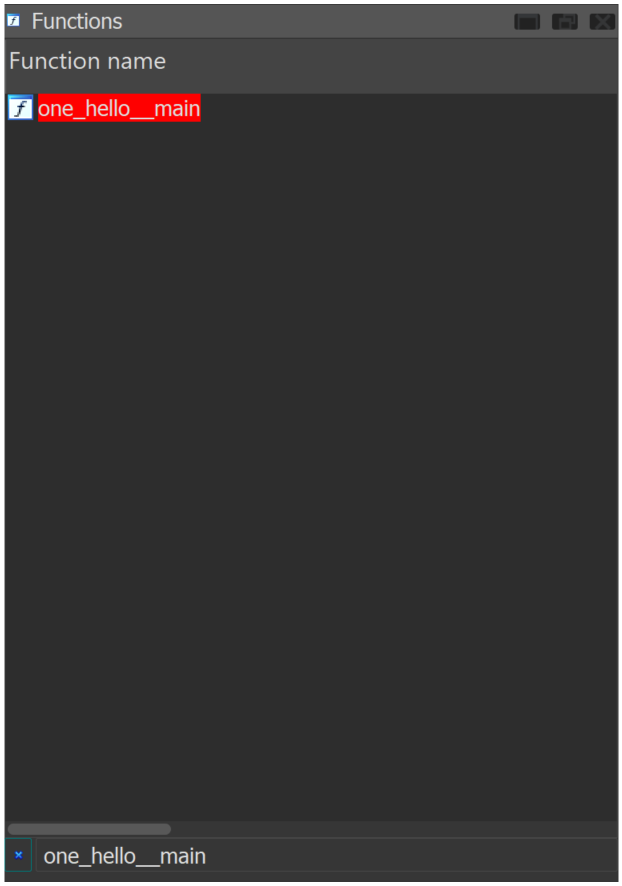
Chapter 2: Debugging Hello Rust

Let's debug our app within IDA Free.

Open IDA Free and we see the load screen. We can keep all the defaults and simply click OK.



In Rust at the assembler level we will need to search for the entry point of our app. This is the *one_hello__main* function. You can use CTRL+F to search.



Now we can double-click on the one_hello__main to launch the focus to this function and graph.

```
; void __fastcall one_hello::main()
one_hello__main proc near

var_38= qword ptr -38h
var_30= byte ptr -30h

sub     rsp, 58h
lea     rcx, [rsp+58h+var_30]
lea     rdx, off_14001E3A0 ; "Hello, world!\n"
mov     r8d, 1
lea     r9, aInvalidArgs ; "invalid args"
xor     eax, eax
mov     [rsp+58h+var_38], 0
call    _ZN4core3fmt9Arguments6new_v117ha9b71491ca997d8aE ; core::fmt::Arguments::new_v1::ha9b71491ca997d8a
lea     rcx, [rsp+58h+var_30]
call    _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
nop
add     rsp, 58h
retn
one_hello__main endp
```

We can see in the box our *"Hello, world!\n"* text.

If we double-click on *off_14001E3A0* it will take us to a new window where the string lives within the binary.

```
.rdata:000000014001E3A0 off_14001E3A0   dq offset aHelloWorld   ; DATA XREF: one_hello__main+9↑o
.rdata:000000014001E3A0                                         ; "Hello, world!\n"
.rdata:000000014001E3A8                 db  0Eh
```

Similar to Go, we can see the strings are in a string pool however we do see a 0Ah, 0 so it contains the newline and null terminator as Go handles it slightly different.

These lessons are designed to be short and digestible so that you can code and hack along.

In our next lesson we will learn how to hack this string and force the binary to print something else to the terminal of our choosing.

This will give us the first taste on hacking Rust!

Chapter 3: Hacking Hello Rust

Let’s hack our app with IDA Free.

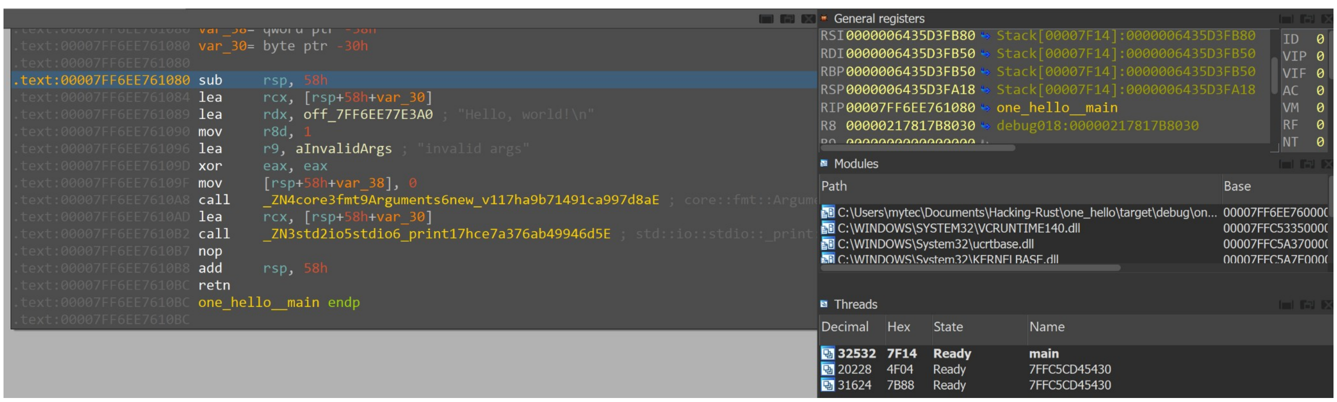
Let’s load up IDA and revisit the binary.

We start off with subtracting *0x58* bytes from *RSP*. We do not see the C-style prologue here. What we are seeing is what the Rust compiler chose to do at compile time.

This is a VERY simple example but I want to take the time to step through this step-by-step.

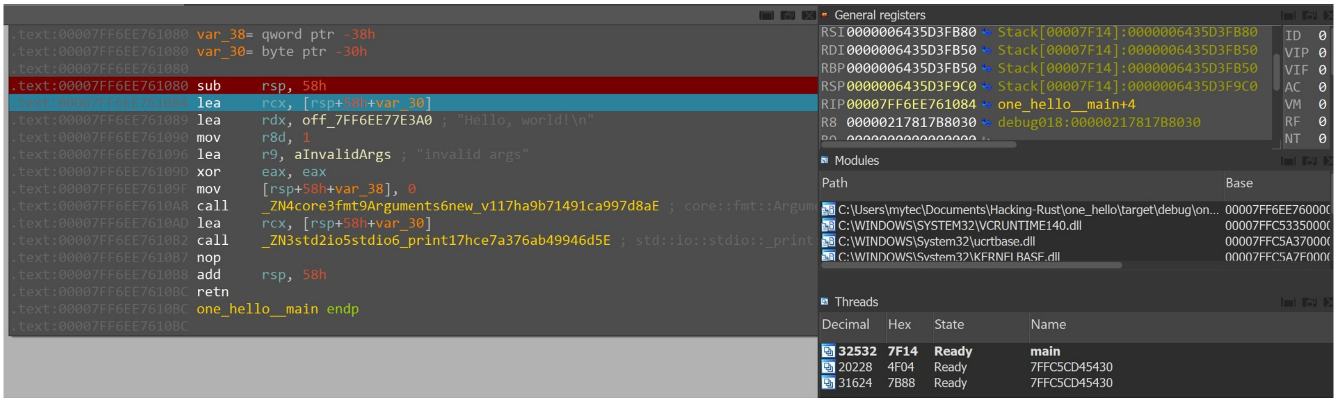
Let’s set a breakpoint at the entry.

Lets run.



Lets take note where *RSP* is *0x0000006435d3fa18*.

Lets step once.



We now see the value of *RSP* at *0x0000006435df9c0*. This makes sense as *0x0000006435d3fa18 - 0x58 = 0x0000006435d3f9c0*.

Next we see *lea rcx, [rsp+58h+var_30]* so let's break this down.

LEA stands for "Load Effective Address." It is a versatile instruction that can be used to perform arithmetic calculations on addresses without actually accessing memory.

RCX is the destination register where the effective address will be stored.

[rsp+58h+var_30] is the source operand, representing the address that needs to be calculated. It consists of multiple parts:
* *RSP* is the stack pointer register.
* *58h* is a hexadecimal constant value, which represents an offset from the stack pointer.
* *var_30* refers to a variable or memory location. In this case, it is *-30h*, indicating a byte-sized variable or memory location located at an offset of *-30h* (or *-48* in decimal) from the stack pointer.

Therefore, *lea rcx, [rsp+58h+var_30]* calculates the effective address by adding the stack pointer (*RSP*) with an offset of *58h* and the variable *var_30* located at an offset of *-30h* from the stack pointer. The resulting effective address is stored in the *RCX* register.

Next we see an offset value being loaded into *RDX*.

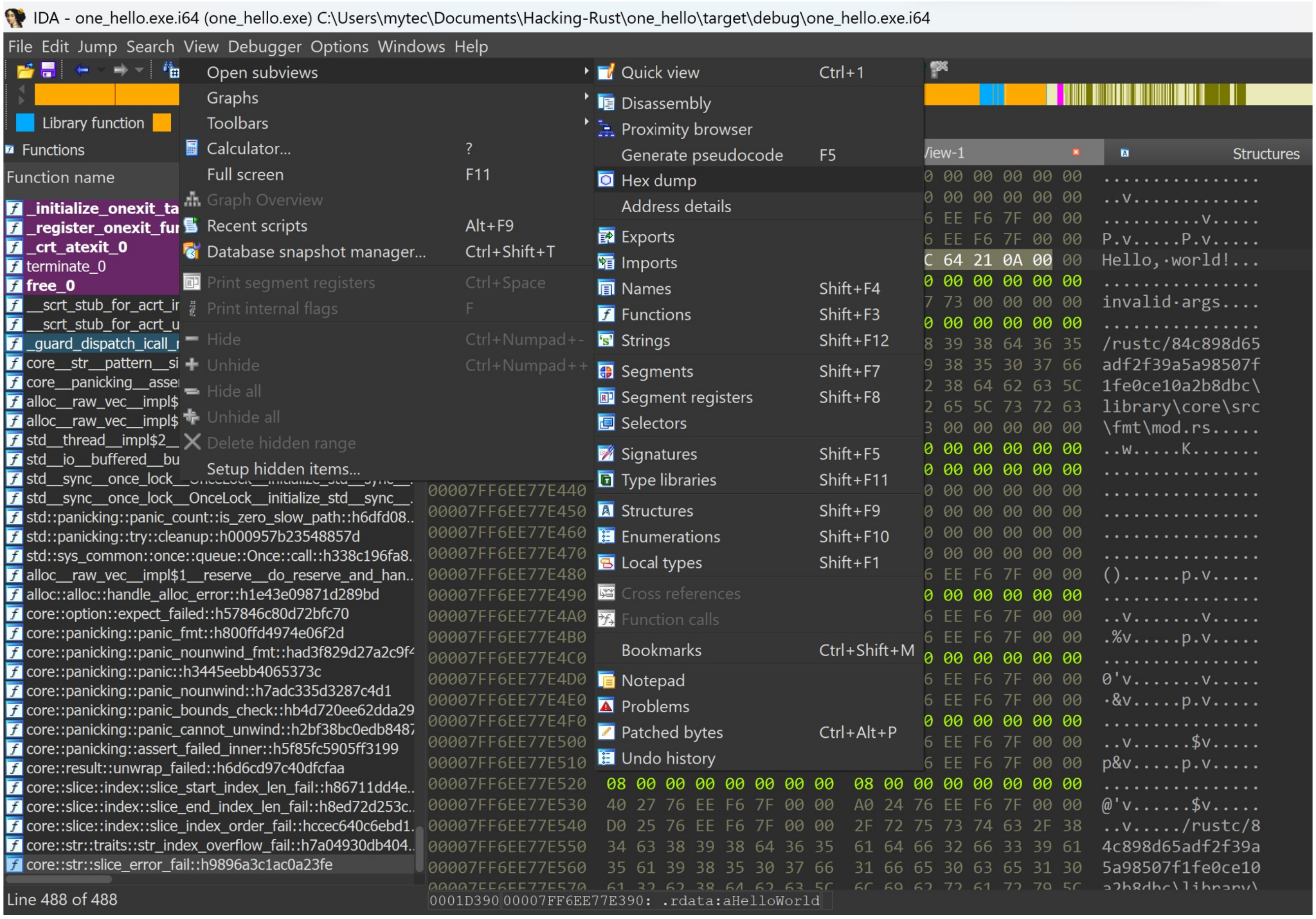
```
.text:00007FF6EE761080
.text:00007FF6EE761080
.text:00007FF6EE761080
.text:00007FF6EE761080 ; void __fastcall one_hello::main()
.text:00007FF6EE761080 one_hello__main proc near
.text:00007FF6EE761080
.text:00007FF6EE761080 var_38= qword ptr -38h
.text:00007FF6EE761080 var_30= byte ptr -30h
.text:00007FF6EE761080
.text:00007FF6EE761080 sub     rsp, 58h
.text:00007FF6EE761084 lea     rcx, [rsp+58h+var_30]
.text:00007FF6EE761089 lea     rdx, off_7FF6EE77E3A0 ; "Hello, world!\n"
.text:00007FF6EE761090 mov     r8d, 1
.text:00007FF6EE761096 lea     r9, aInvalidArgs ; "invalid args"
.text:00007FF6EE76109D xor     eax, eax
.text:00007FF6EE76109F mov     [rsp+58h+var_38], 0
.text:00007FF6EE7610A8 call    _ZN4core3fmt9Arguments6new_v117ha9b71491ca997d8aE ; core::fmt::Argu
.text:00007FF6EE7610AD lea     rcx, [rsp+58h+var_30]
.text:00007FF6EE7610B2 call    _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_prin
.text:00007FF6EE7610B7 nop
.text:00007FF6EE7610B8 add     rsp, 58h
.text:00007FF6EE7610BC retn
.text:00007FF6EE7610BC one_hello__main endp
```

This is our string. Let's double-click on the offset.

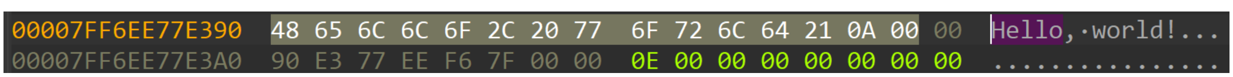
```
.rdata:00007FF6EE77E3A0 off_7FF6EE77E3A0 dq offset aHelloWorld  ; DATA XREF: one_hello__main+9↑o
.rdata:00007FF6EE77E3A0                                         ; "Hello, world!\n"
.rdata:00007FF6EE77E3A8 db  0Eh
```

We see 13 chars and the null terminator therefore *0x0e*.

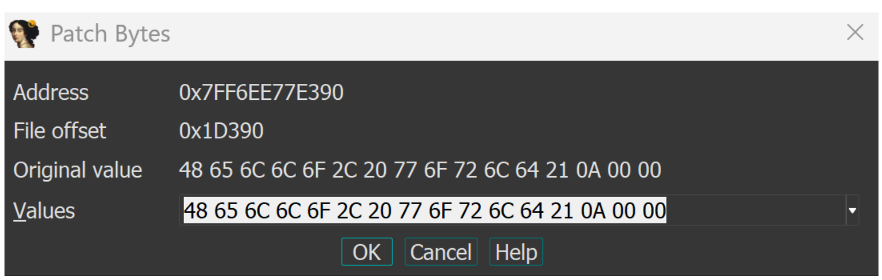
Let's stop the debug and click on the *"Hello, world!\n"* text and view the *Hex View-1*.



Here we see the hex view.

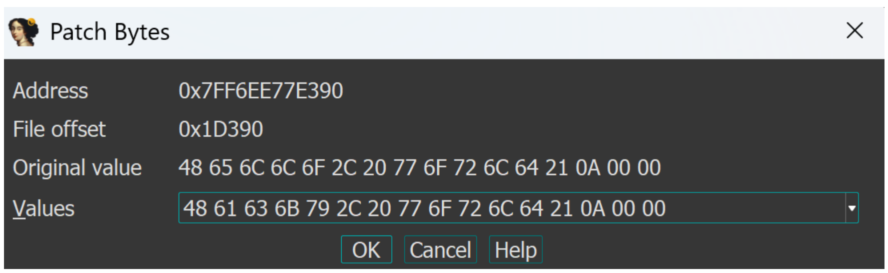


Click edit – Patch program – Change byte…

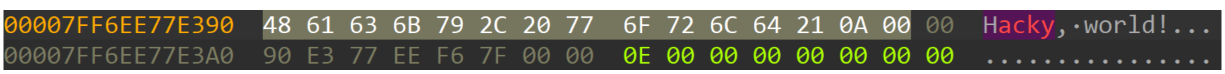


Let’s hack our binary to say “Hacky, world!\n” instead!

```
'H' -> 0x48
'a' -> 0x61
'c' -> 0x63
'k' -> 0x6B
'y' -> 0x79
',' -> 0x2C
' ' -> 0x20
'w' -> 0x77
'o' -> 0x6F
'r' -> 0x72
'l' -> 0x6C
'd' -> 0x64
'!' -> 0x21
```



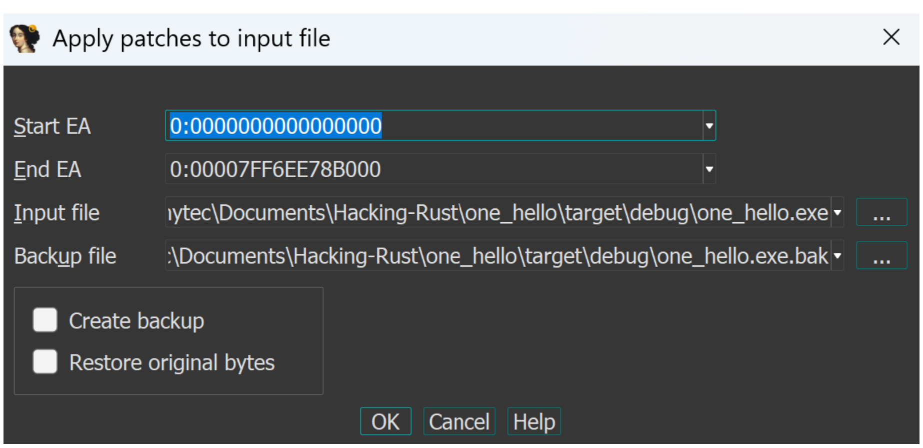
Click OK.



SUCCESS!

Press the red stop button to stop debugging.

Click edit – Patch program – Apply patches to input file…



Let’s run again…

Let’s F8, step over, until the *NOP*.

Observe the command window.

SUCCESS! We have successfully hacked our first Rust binary!

The rest of the Assembler is trivial as we are just calling std::io::stdio::_print to echo the line to STDOUT.

Stay tuned as this will get significantly more challenging. With all things we start small and build!

In our next chapter we will discuss the Scalar Data Types.

Chapter 4: Scalar Data Types

We continue our journey with a scalar data types example program in Rust on a Windows 64-bit OS.

Let’s create a new project and get started by following the below steps.

```
cargo new two_scalar-data-types
```

Now let’s populate our **main.rs** file with the following.

```
fn main() {
    let my_integer: i32 = 42;
    println!("integer: {}", my_integer);

    let my_float: f64 = 3.14;
    println!("float: {}", my_float);

    let my_char: char = 'A';
    println!("character: {}", my_char);
}
```

Let’s open up the terminal by click CTRL+SHIFT+` and type the following.

```
cargo build
```

Let’s run the binary!

```
.\target\debug\two_scalar-data-types.exe
```

Output…

```
integer: 42
float: 3.14
character: A
```

Congratulations! You just created another program in Rust. Time for cake!

We simply created an example of each of the scalar data types which are int, float and char.

In our next lesson we will debug this in IDA Free!

Chapter 5: Debugging Scalar Data Types

Let’s debug our app within IDA Free.

Open IDA Free and we see the load screen. We can keep all the defaults and simply click OK.

On the left-hand side we see the Function name window and we want to search for *two_scalar_data_types__main*. Double-click it and we will see the function.

```
; void __fastcall two_scalar_data_types::main()
two_scalar_data_types__main proc near

var_E8= qword ptr -0E8h
result= core::fmt::ArgumentV1 ptr -0D4h
var_A0= qword ptr -0A0h
var_98= qword ptr -98h
var_90= core::fmt::ArgumentV1 ptr -90h
var_58= qword ptr -58h
var_50= qword ptr -50h
var_44= core::fmt::ArgumentV1 ptr -44h
var_10= qword ptr -10h
var_8= qword ptr -8

sub     rsp, 108h
mov     dword ptr [rsp+108h+result.value], 2Ah ; '*'
lea     rcx, [rsp+108h+result] ; result
call    _ZN4core3fmt10ArgumentV111new_display17hda29a79969d0b5cbE ; core::fmt::ArgumentV1::new_display::hda29a79969d0b5cb
mov     [rsp+108h+var_A0], rax
mov     [rsp+108h+var_98], rdx
lea     rcx, [rsp+108h+result.value+4]
lea     rdx, off_140023430 ; "integer: \n"
mov     r8d, 2
lea     r9, [rsp+108h+var_A0]
mov     [rsp+108h+var_E8], 1
call    _ZN4core3fmt9Arguments6new_v117h8807bc63cce0ab3bE ; core::fmt::Arguments::new_v1::h8807bc63cce0ab3b
lea     rcx, [rsp+108h+result.value+4]
```

```
call    _ZN4core3fmt10ArgumentV111new_display17h0ac0f3d1bf4c49e3E ; core::fmt::ArgumentV1::new_display::h0ac0f3d1bf4c49e3
mov     [rsp+108h+var_58], rax
mov     [rsp+108h+var_50], rdx
lea     rcx, [rsp+108h+var_90.formatter]
lea     rdx, off_140023458 ; "float: "
mov     r8d, 2
lea     r9, [rsp+108h+var_58]
mov     [rsp+108h+var_E8], 1
call    _ZN4core3fmt9Arguments6new_v117h8807bc63cce0ab3bE ; core::fmt::Arguments::new_v1::h8807bc63cce0ab3b
lea     rcx, [rsp+108h+var_90.formatter]
call    _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
mov     dword ptr [rsp+108h+var_44.value], 41h ; 'A'
lea     rcx, [rsp+108h+var_44] ; result
call    _ZN4core3fmt10ArgumentV111new_display17h2cdf30f7c1587727E ; core::fmt::ArgumentV1::new_display::h2cdf30f7c1587727
mov     [rsp+108h+var_10], rax
mov     [rsp+108h+var_8], rdx
lea     rcx, [rsp+108h+var_44.value+4]
lea     rdx, off_140023488 ; "character: "
mov     r8d, 2
lea     r9, [rsp+108h+var_10]
mov     [rsp+108h+var_E8], 1
call    _ZN4core3fmt9Arguments6new_v117h8807bc63cce0ab3bE ; core::fmt::Arguments::new_v1::h8807bc63cce0ab3b
lea     rcx, [rsp+108h+var_44.value+4]
call    _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
nop
add     rsp, 108h
retn
two_scalar_data_types__main endp
```

We see the below instruction.

```
mov     dword ptr [rsp+108h+result.value], 2Ah ; '*'
```

The value of *2Ah* is *42* so we know that the literal value is being stored inside *RSP* + the offset value above.

We then see a call to the **println!** macro.

```
call    _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
```

Bare with me as I explain the floating point number. Let’s first review the actual code we are going to statically reverse.

```
    let my_float: f64 = 3.14;
    println!("float: {}", my_float);
```

The first instruction loads the floating-point value *3.14* from memory into the XMM0 register. The XMM registers are 128-bit registers that are used to store floating-point values and vectors.

```
movsd   xmm0, cs:__real@40091eb851eb851f
```

The next instruction stores the value in XMM0 to the memory location rsp+108h+var_90.value. This memory location is part of the stack frame that is allocated for the *println!()* function call.

```
movsd   [rsp+108h+var_90.value], xmm0
```

The next instruction sets the register *rcx* to point to the memory location *rsp+108h+var_90*. This memory location contains the *ArgumentV1* struct that will be passed to the new_display function.

```
lea     rcx, [rsp+108h+var_90] ; result
```

The next instruction calls the new_display function. The *new_display* function is responsible for converting the *ArgumentV1* struct into a string representation.

```
call    _ZN4core3fmt10ArgumentV111new_display17h0ac0f3d1bf4c49e3E ; core::fmt::ArgumentV1::new_display::h0ac0f3d1bf4c49e3
```

The *new_display* function returns the string representation in the *RAX* and *RDX* registers. The next instruction stores the value in *RAX* to the memory location *rsp+108h+var_58*. This memory location contains the *Formatter* struct that will be passed to the *println!()* function.

```
mov     [rsp+108h+var_58], rax
```

The next instruction stores the value in *RDX* to the memory location *rsp+108h+var_50*. This memory location contains the slice of bytes that represents the string representation of the *ArgumentV1* struct.

```
mov     [rsp+108h+var_50], rdx
```

The next instruction sets the register *RCX* to point to the memory location r*sp+108h+var_90.formatter*. This memory location contains the *Formatter* struct that will be passed to the *println!()* function.

```
lea     rcx, [rsp+108h+var_90.formatter]
```

The next instruction sets the register *RDX* to point to the string literal "*float:* ".

```
lea     rdx, off_140023458 ; "float: "
```

Finally we see the *println* macro print the value.

```
call    _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
```

The last item was the char ‘A’ which is simply moved into RSP and an offset below.

```
mov     dword ptr [rsp+108h+var_44.value], 41h ; 'A'
```

Then we print.

```
call    _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
```

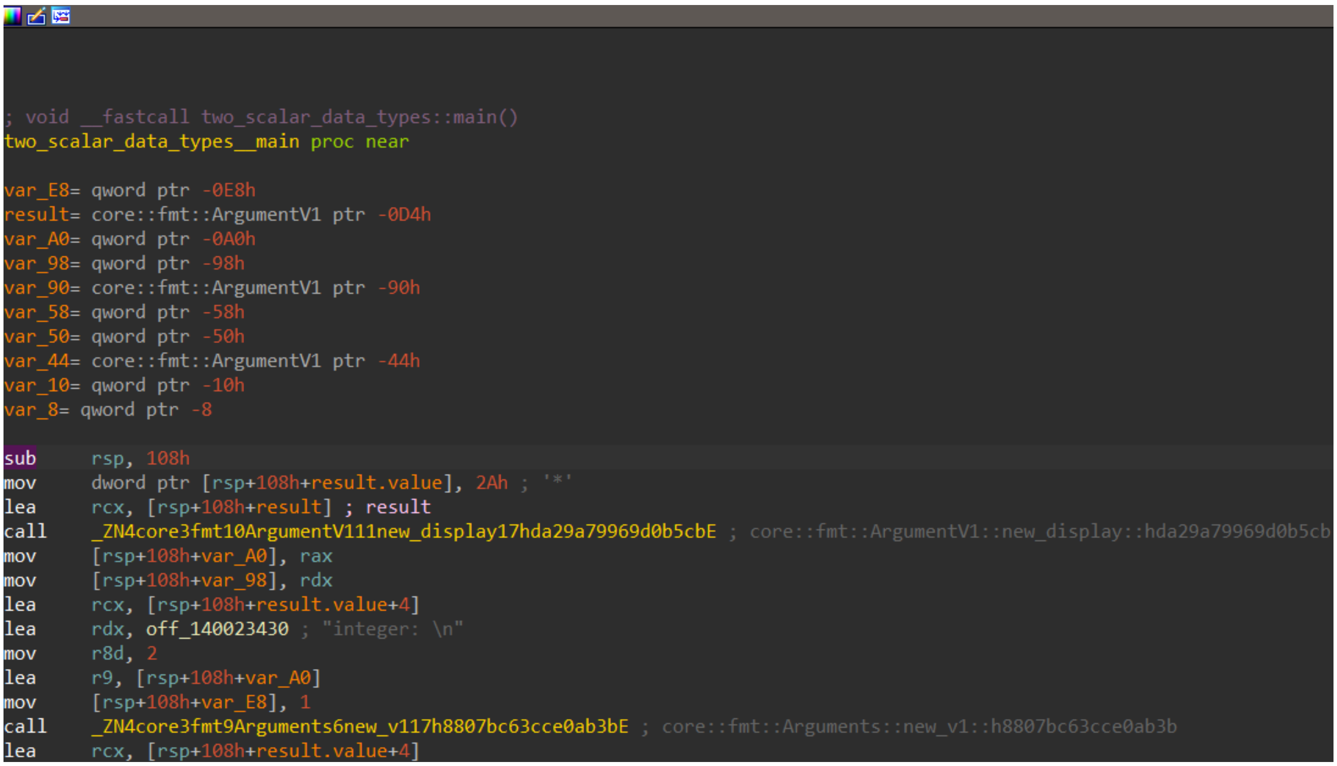
In our next lesson we will hack this!

Chapter 6: Hacking Scalar Data Types

Let's hack our app within IDA Free.

Open IDA Free and we see the load screen. We can keep all the defaults and simply click OK.

On the left-hand side we see the Function name window and we want to search for *two_scalar_data_types__main*. Double-click it and we will see the function.



```
call    _ZN4core3fmt10ArgumentV111new_display17h0ac0f3d1bf4c49e3E ; core::fmt::ArgumentV1::new_display::h0ac0f3d1bf4c49e3
mov     [rsp+108h+var_58], rax
mov     [rsp+108h+var_50], rdx
lea     rcx, [rsp+108h+var_90.formatter]
lea     rdx, off_140023458 ; "float: "
mov     r8d, 2
lea     r9, [rsp+108h+var_58]
mov     [rsp+108h+var_E8], 1
call    _ZN4core3fmt9Arguments6new_v117h8807bc63cce0ab3bE ; core::fmt::Arguments::new_v1::h8807bc63cce0ab3b
lea     rcx, [rsp+108h+var_90.formatter]
call    _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
mov     dword ptr [rsp+108h+var_44.value], 41h ; 'A'
lea     rcx, [rsp+108h+var_44] ; result
call    _ZN4core3fmt10ArgumentV111new_display17h2cdf30f7c1587727E ; core::fmt::ArgumentV1::new_display::h2cdf30f7c1587727
mov     [rsp+108h+var_10], rax
mov     [rsp+108h+var_8], rdx
lea     rcx, [rsp+108h+var_44.value+4]
lea     rdx, off_140023488 ; "character: "
mov     r8d, 2
lea     r9, [rsp+108h+var_10]
mov     [rsp+108h+var_E8], 1
call    _ZN4core3fmt9Arguments6new_v117h8807bc63cce0ab3bE ; core::fmt::Arguments::new_v1::h8807bc63cce0ab3b
lea     rcx, [rsp+108h+var_44.value+4]
call    _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
nop
add     rsp, 108h
retn
two_scalar_data_types__main endp
```

We see the below instruction.

```
mov     dword ptr [rsp+108h+result.value], 2Ah ; '*'
```

The value of *2Ah* is *42* so we know that the literal value is being stored inside *RSP* + the offset value above.

Let's set a breakpoint by highlighting the line and pressing F2.

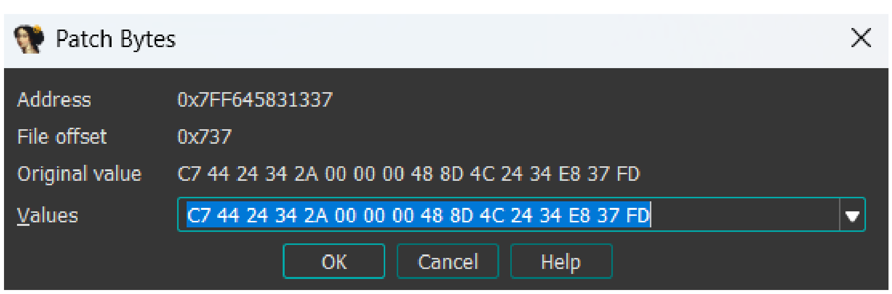
```
mov     dword ptr [rsp+108h+result.value], 2Ah ; '*'
```

Press play to debug.

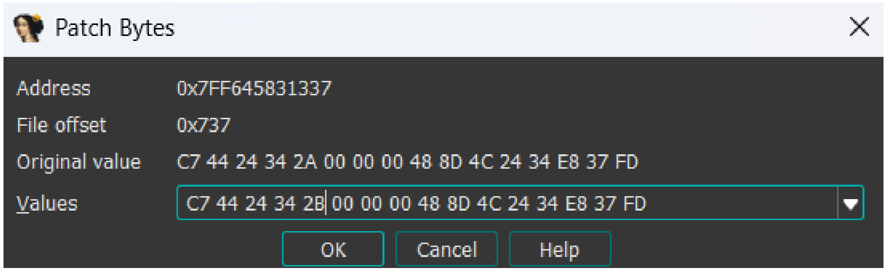
We now see the our breakpoint triggered.

```
.text:00007FF645831337 mov     dword ptr [rsp+108h+result.value], 2Ah
```

Click edit – Patch program – Change byte…



Change the 2A to 2B.



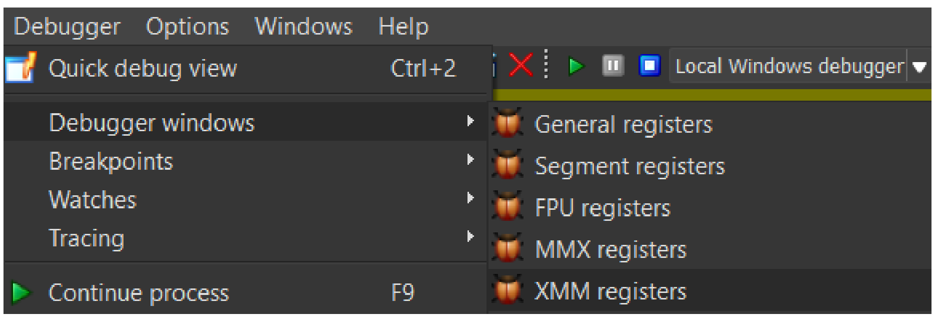
Click ok. We have now hacked our 42 value to 43!

Let’s put a breakpoint on the below line.

```
.text:00007FF73743138A movsd   [rsp+108h+var_90.value], xmm0
```

Lets press F9 to run to it.

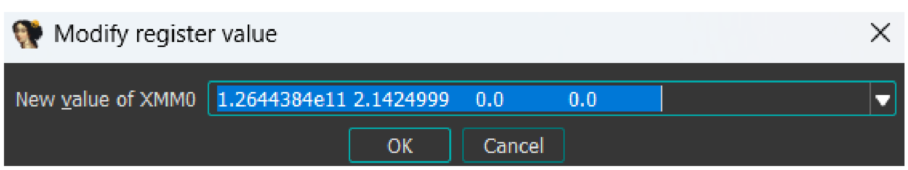
Click on Debugger – Debugger windows – XMM registers



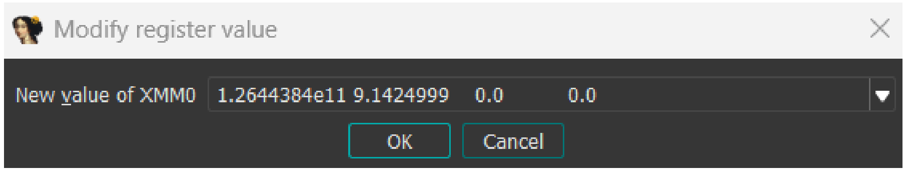
It will load the XMM registers.

```
XMM0  1.2644384e11  2.1424999     0.0           0.0
```

Double-click on the 2.1424999.



Change the value to 9 instead of 2.



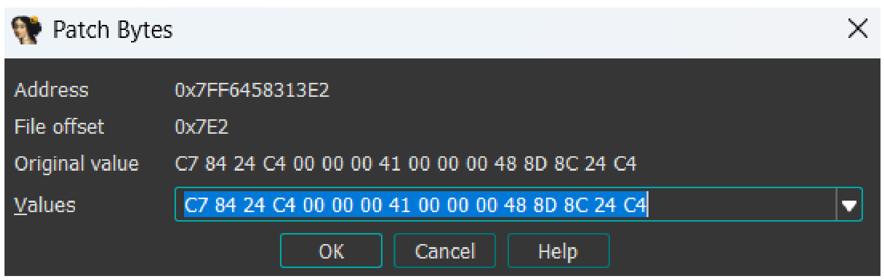
Press OK.

We have successfully hacked 3.14 to 299499.58!

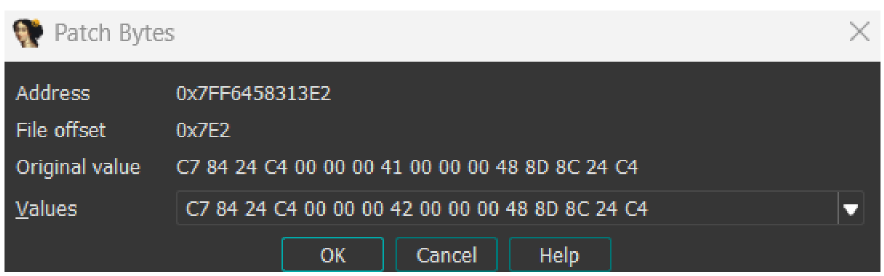
Put a breakpoint on the below line and press F9.

```
.text:00007FF6458313E2 mov     dword ptr [rsp+108h+var_44.value], 41h
```

Click edit – Patch program – Change byte…



Change 41 to 42.



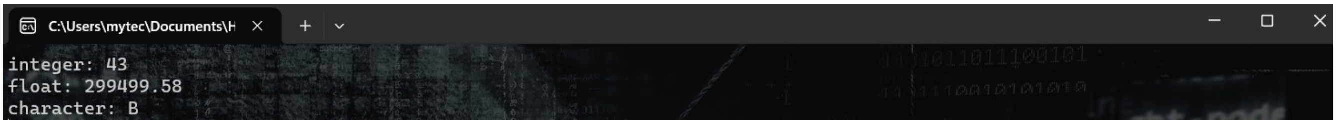
Press the red stop button to stop debugging.

Click edit – Patch program – Apply patches to input file…

Place a breakpoint on the NOP under the final call to println.

```
call    _ZN3std2io5stdio6_print17hce7a376ab49946d5E ; std::io::stdio::_print::hce7a376ab49946d5
nop
```

BOOM!



In our next lesson we will handle Compound Data Types.