Lab Report on:

Reverse Engineering

SWE 4802: Software Maintenance Lab

Lutfun Nahar Lota

Assistant Professor

Department of Computer Science and Engineering

Islamic University of Technology

Submitted By:

Abrar Mahabub (200042103)

Mashrur Ahsan (200042115)

Nafisa Maliyat (200042133)

Shanta Maria (200042172)

Table of Contents

IDA Pro

IDA (Interactive Disassembler) is a comprehensive tool used for reverse engineering, which means understanding how a program works without having access to the source code. IDA offers a range of useful features such as disassembling binary files, displaying code flow visually through graphs and flowcharts, and running a debugger to see what each instruction is doing. It also allows use of Python scripts to automate the analysis process. In addition, it also works with many CPU types and platforms, making it suitable for analyzing different types of programs. The progress of any analysis is saved as .idb file so users can continue their work anytime.

Features of IDA Pro

Some prominent features of this tool are as follows:

- **Disassembly**: Converts executables into symbolic assembly code for many file formats and architectures
- **Decompiler (Pseudocode)**: Generates human-readable C-like pseudocode to ease understanding of binary logic (cloud-based in limited versions, local in Pro)
- **Cross-references & Type Reconstruction**: Automatically identifies cross-links, functions, stack frames, variables, and reconstructed data types
- **Interactive Editing**: Analysts can override auto-analysis results—rename, re-flag, retype—for precision reverse-engineering
- **Programmability & Automation**: Support scripting via IDC and IDAPython, plus plugin support to extend capabilities.
- **Plugin Architecture**: Open SDK enables writing GUI enhancements and analysis tools; large ecosystem of community plugins
- **FLIRT Signature Recognition**: Automatically identifies standard library functions to simplify disassembly
- **idalib (Headless Mode)**: Use IDA engine as a library for C++ or Python automation, running offline or in batch environments (Pro only)
- **Private Lumina Add-on**: Enterprise feature for sharing function metadata and analysis across teams (Pro only)
- **Teams Add-on**: Enables team collaboration, file version control, and syncing (Pro only)

C Code

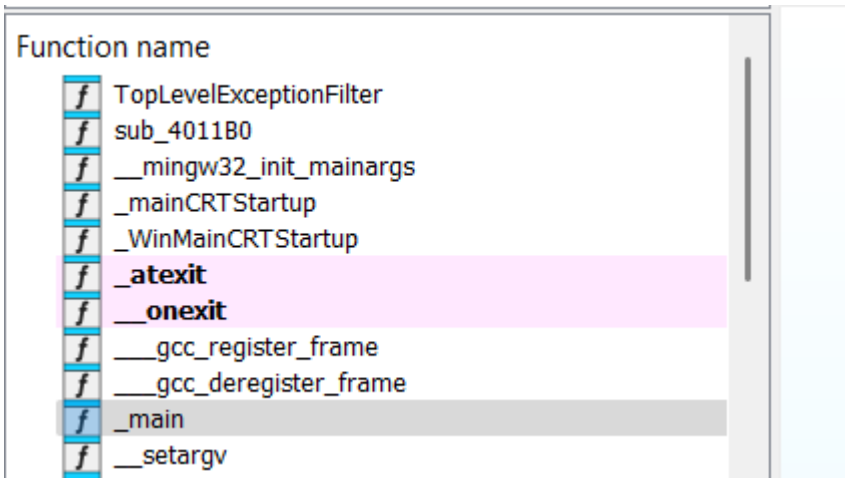
This program is designed to determine whether a given year is a leap year. It accepts a numerical input representing a year and applies a sequence of conditional logic. Specifically, the program checks if the year is divisible by 4 but not by 100, except when it is also divisible by 400. If these conditions are satisfied, the year is classified as a leap year, which comprises 366 days. Otherwise, the year is identified as a common year with 365 days.

```c
#include <stdio.h>
int main() {
   int year;
   printf("Enter a year: ");
   scanf("%d", &year);

   if (year % 400 == 0) {
      printf("%d is a leap year.", year);
   }

   else if (year % 100 == 0) {
      printf("%d is not a leap year.", year);
   }

   else if (year % 4 == 0) {
      printf("%d is a leap year.", year);
   }
   else {
      printf("%d is not a leap year.", year);
   }

   return 0;
}
```

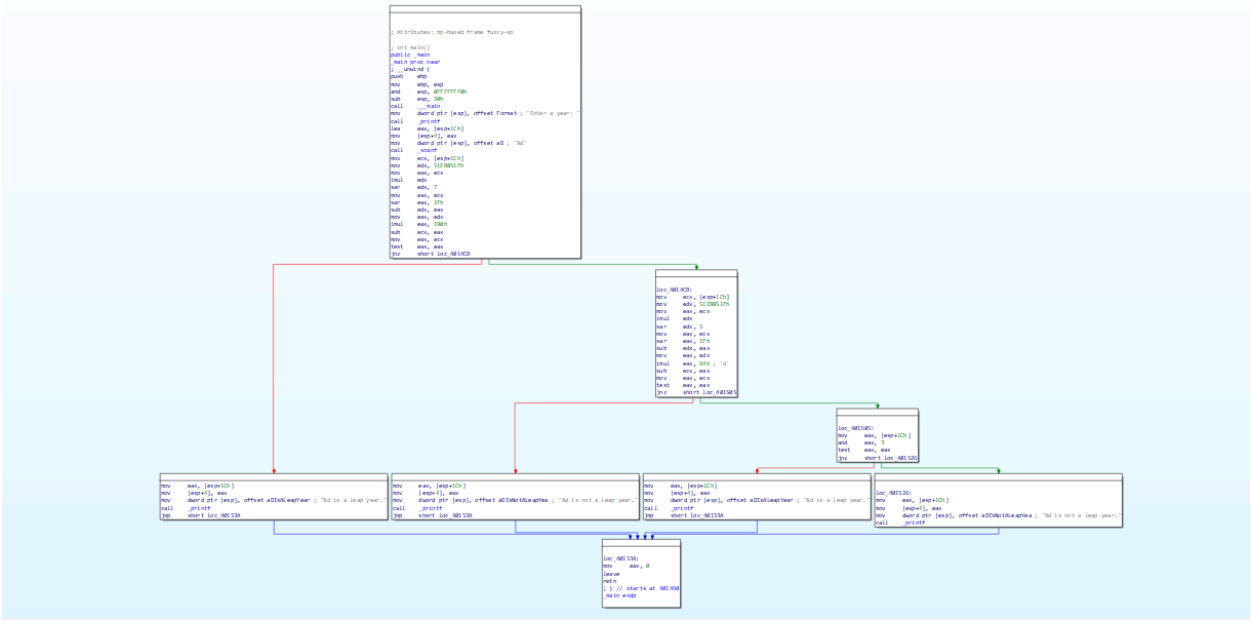
Analysis of our C code

Main function Identification

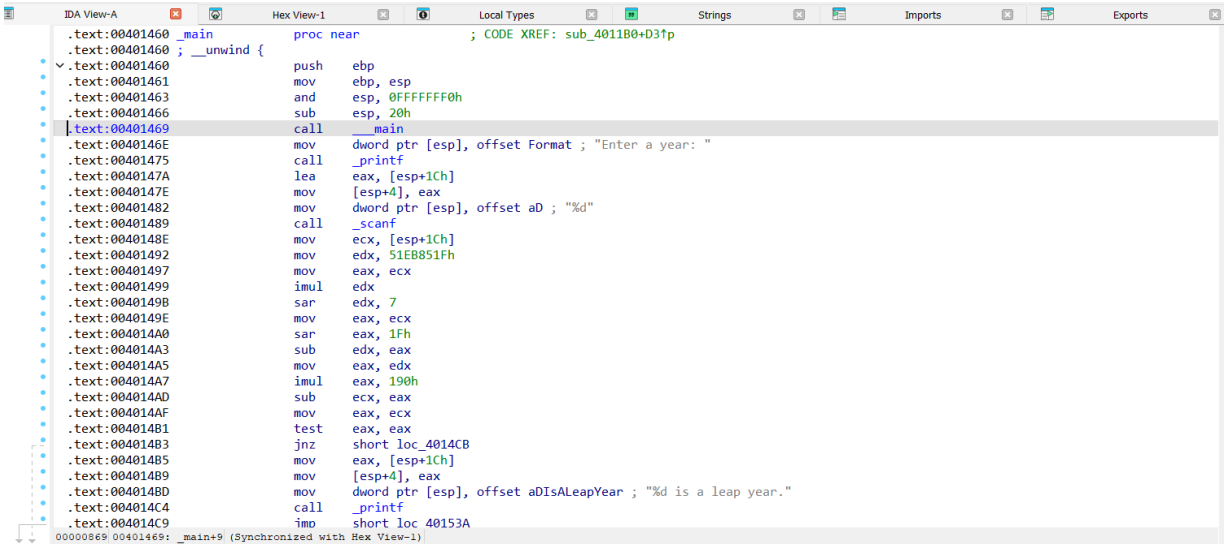
- In the left panel, we have the functions window
- After scrolling through the list of detected functions, we can find a function named “_main”.



- We can view IDA view-A by double-clicking the “_main” function.

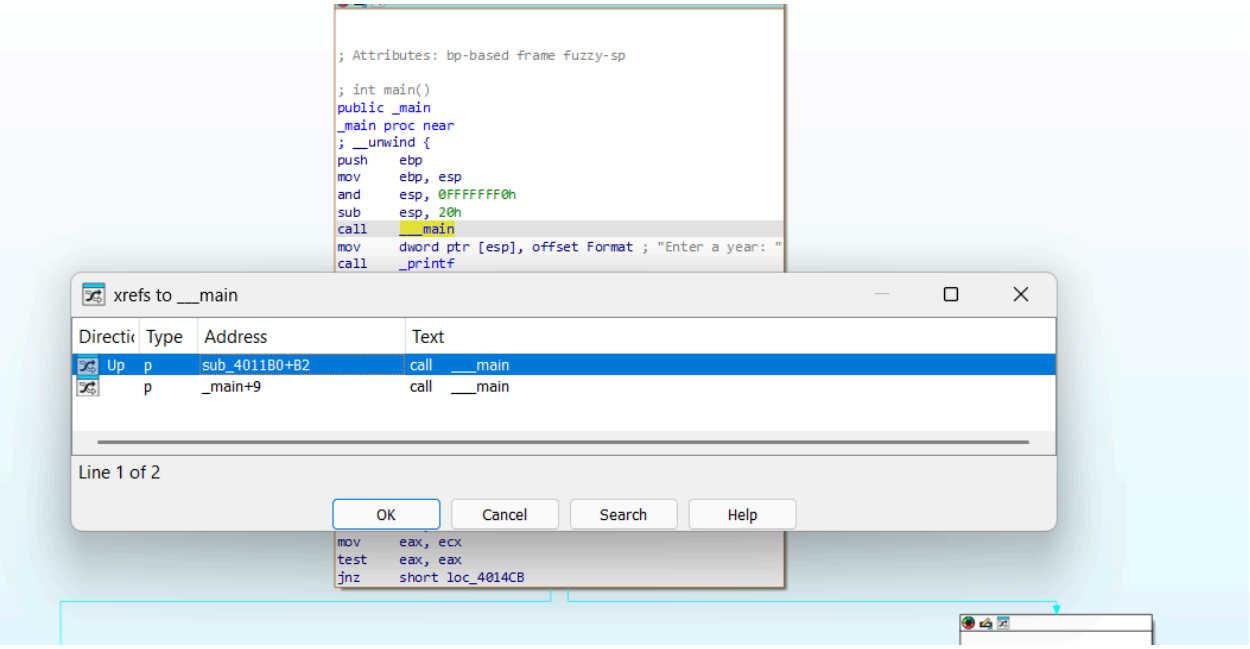


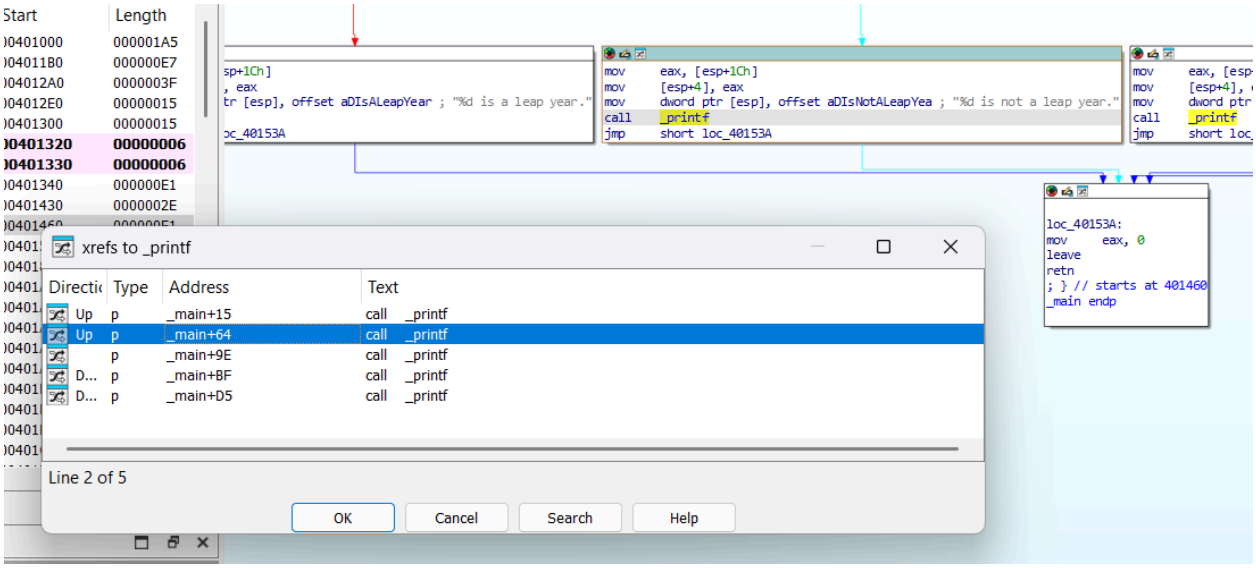
- We can toggle between Linear view and graph view using the space bar.



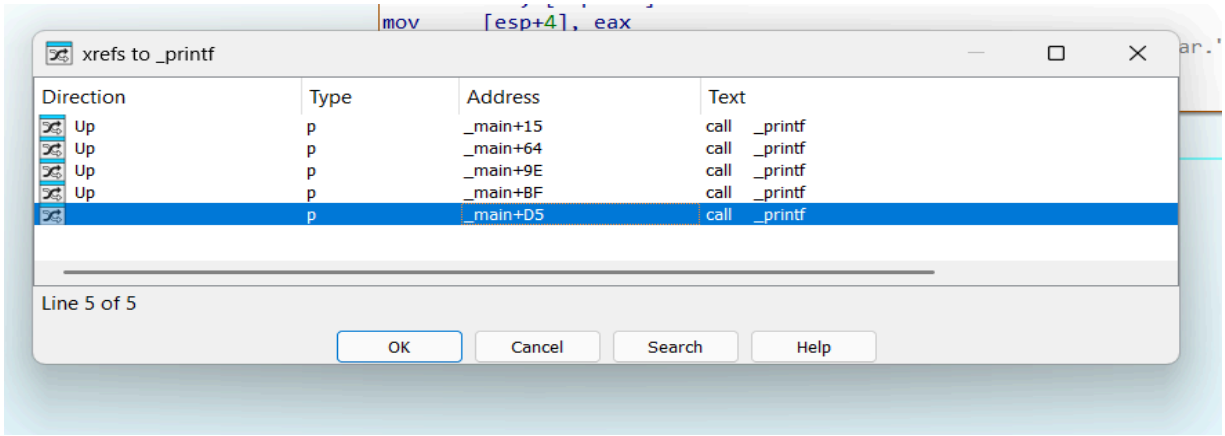
Cross-referencing Identification

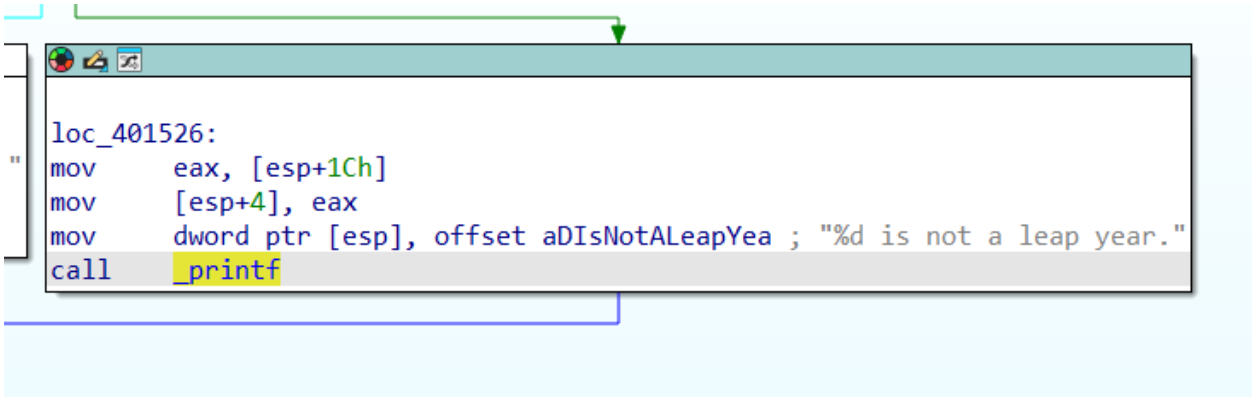
- In IDA view-A, we will hover the cursor on the function we wish to investigate.
- By pressing X, we can view the cross-references in the window.
- In the window, it will display the location of the function being called.



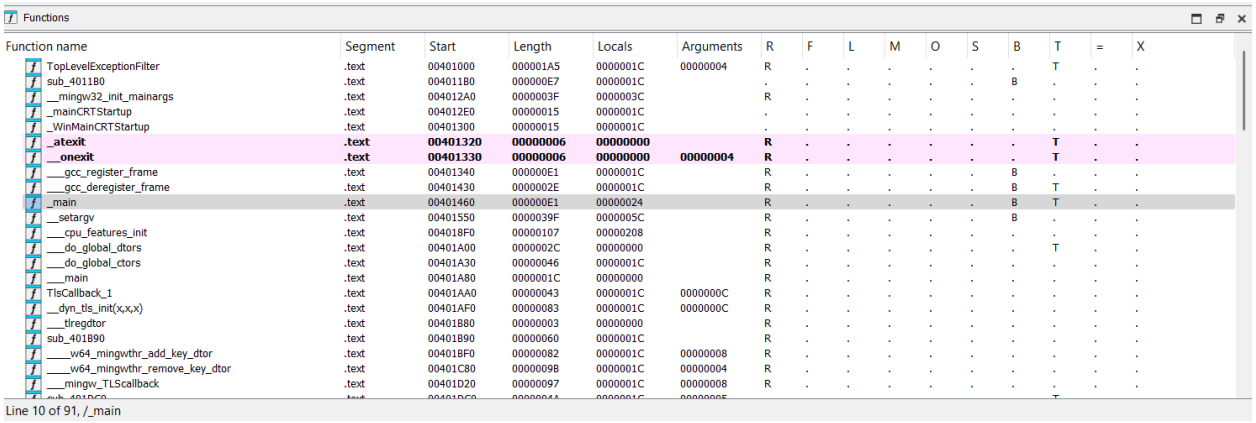


- Double-clicking any entry in the list to navigate directly to the usage.





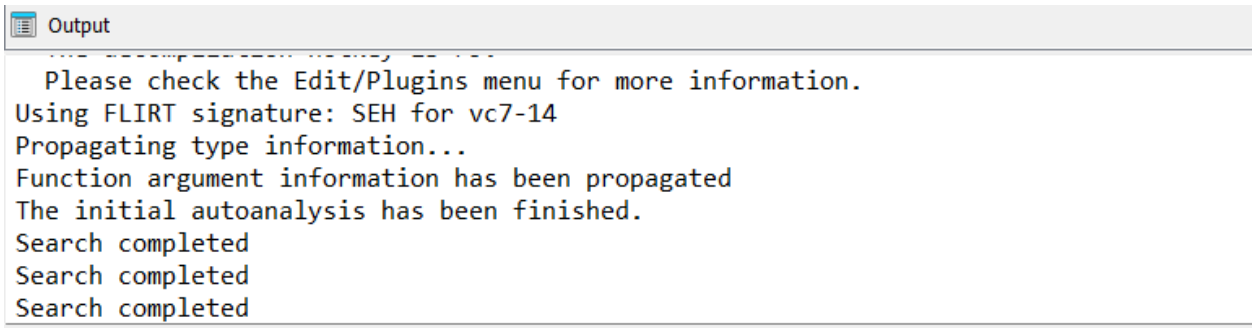
The Functions window



Here we can see several columns in the “Functions” window. Here is what the columns represent:

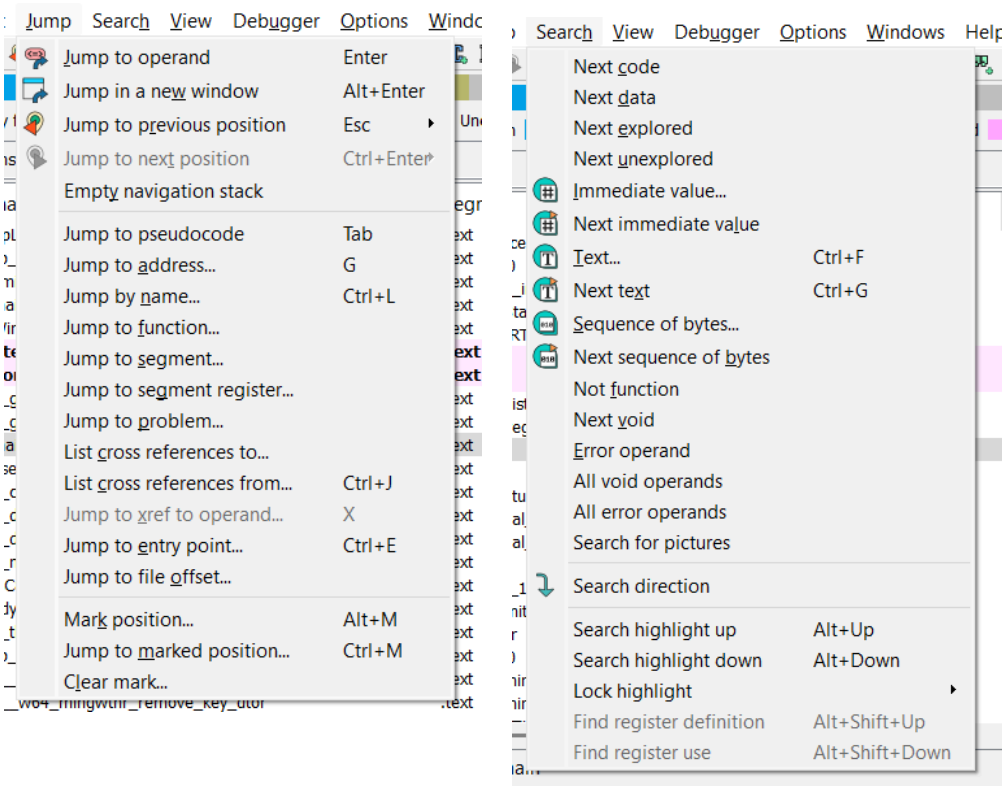
- The first column Function names
- Second column signifies the segment
- Third column shows the start Memory locations, fourth column shows the whole length
- Arguments column shows the number of arguments the functions that in
- Rest of the columns represent different status(s) like referred, library, manual etc.

Output window



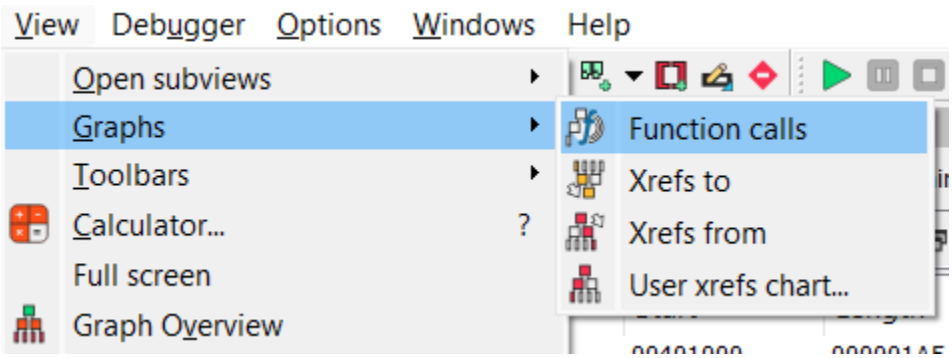
The output window shows all the logs during our analysis of the code. It acts as a logger. It shows the sequence of operations and logs confirmation messages accordingly.

Jump, Search, View options



Here we can see several columns in the “Functions” window. Here is what the columns represent:

- Jump option lets one jump directly to an operand, address, memory location, segment, function etc.
- The search operation lets one search for some data in code, memory locations
- The view option lets one look at function calls, cross references, references from or to, user charts



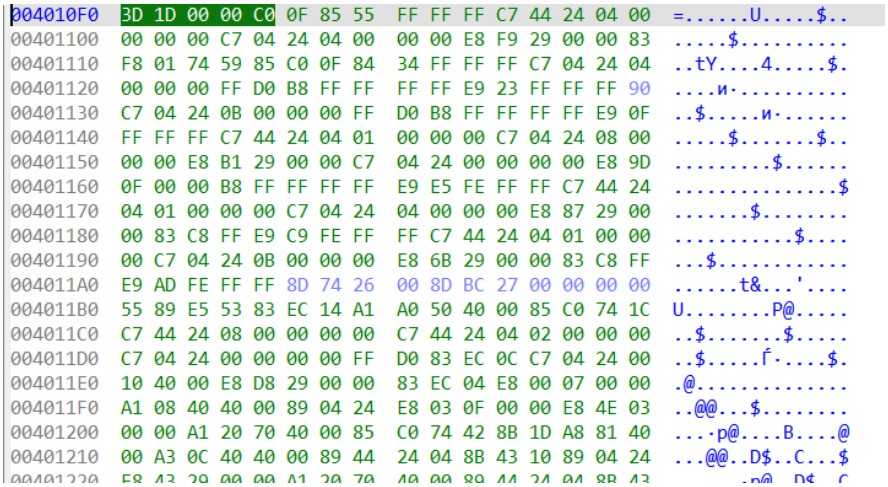
Hex View-1

Hex View-1 in IDA shows the raw bytes of the binary file. The format of the analysis of the compiled code in hex view-1 is as follows:

```
<Memory Address>   <Hex Bytes>   <ASCII equivalent (not always shown)>
```

The raw bytes represent the low-level assembly instructions, not the actual code. But these low-level instructions map to the high-level C code that the original code is written in.

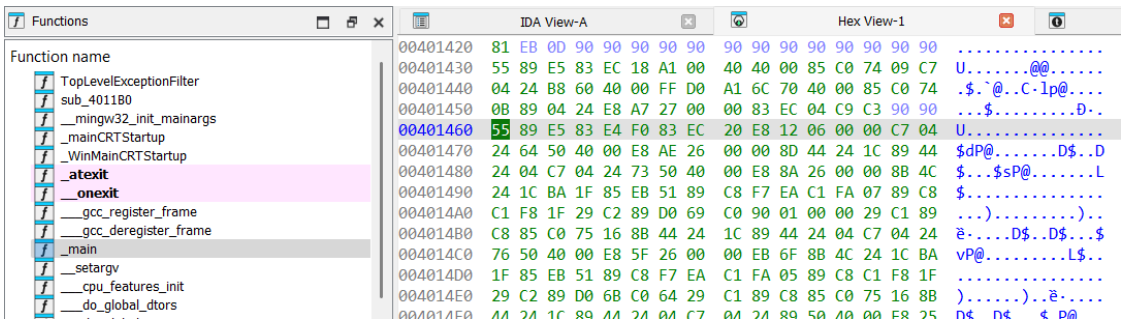
Below is a small part of the hex byte breakdown for our code:



From this breakdown, we can figure out some instructions, for example:

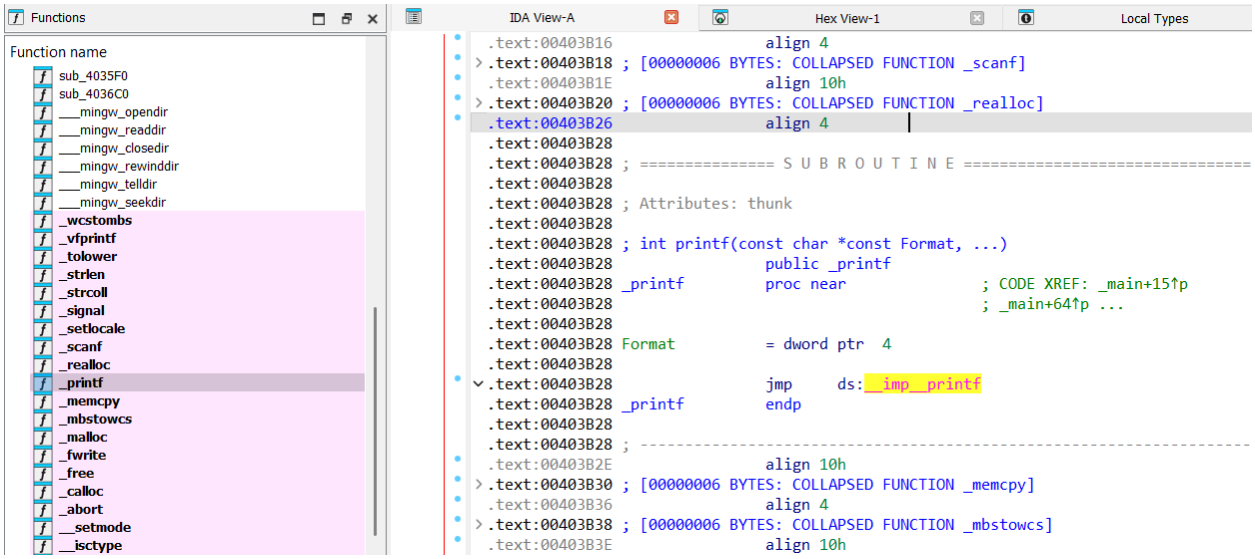
```
00401155: 00 00 E8 B1 29 00 00 C7
```

- **Instruction**: `jne 0x00401050`
- **Meaning**: Jump to address `00401155` if the comparison above was not equal.
- So this is a **conditional jump** that is `if` or `else if` logic.



If we click on a function name in the Functions panel, the respective address in the Hex View-1 panel would be shown. In the above example, we can see that upon clicking the `main` function we can see it is stored at `0x00401050`.

IDA View-A (Disassembled Code)



The code in the screenshot shows a tiny helper function called `_printf`. It's **not** the real `printf` function.

Instead, this helper just **jumps to the real `printf`**, which lives in a system file like a DLL (for example, `msvcrt.dll`).

This is how programs usually work when they use common functions like `printf` — they don't include the whole code for `printf` inside the program. They just **link to it** and call it when needed.

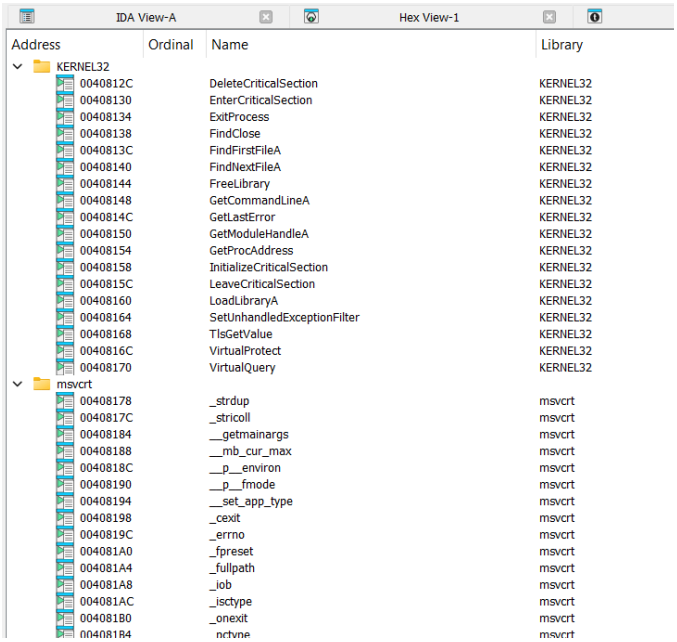
So, this `_printf` function is like a call to the `printf` function in the system library.

Imports

The imported functions can be viewed and analysed in the Imports tab. It shows all functions the program uses from external libraries in this case:

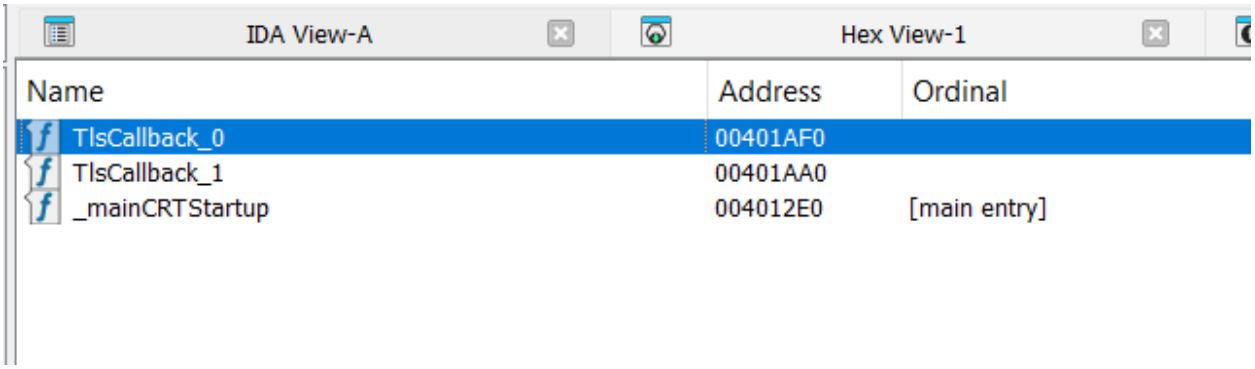
- `printf`
- `scanf`

The functions are linked from (like msvcrt.dll or kernel32.dll). The addresses or pointers used to jump to those functions. You can double-click any import to see where it's used in the disassembly.



Exports

Shows all the functions that we export from our code for use in other programs. We have no export function in our code as shown except main:



`_mainCRTStartup` is stored at the Address shown (`004012E0`) that is the main function of our C program.

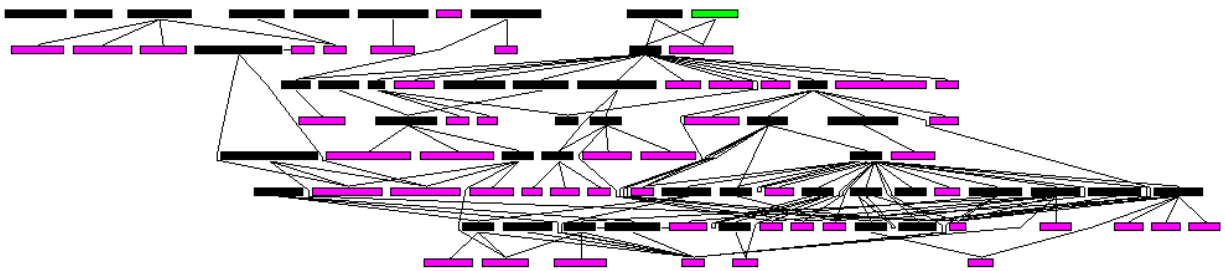
PseudoCode Generation

IDA Vie... | Pseudoco... | xrefs to .text:00401...

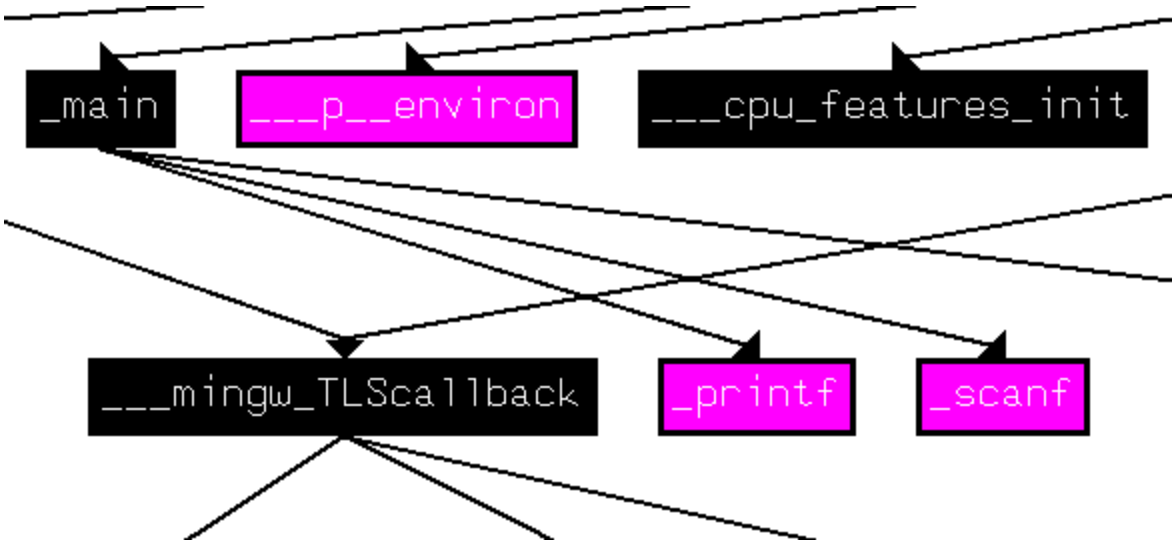
```
1  int main()
2  {
3    int v1; // [esp+1Ch] [ebp-4h] BYREF
4
5    __main();
6    printf("Enter a year: ");
7    scanf("%d", &v1);
8    if ( v1 % 400 && (!(v1 % 100) || (v1 & 3) != 0) )
9      printf("%d is not a leap year.", v1);
10   else
11     printf("%d is a leap year.", v1);
12   return 0;
13 }
```

IDA tool also provides an option to view the pseudocode of a binary through its **Hex-Rays Decompiler plugin**. The pseudocode, although not exactly the original source code, is much more readable and helps to quickly figure out what the program is doing — like how functions work, how data is being used, and what the overall logic looks like. The pseudocode helps analysts understand the logic and control flow of the program without having to interpret complex assembly instructions directly.

Graph View of Code Flow



The Graph View in IDA provides a clear visual representation of a function's control flow, displaying how the program executes from start to finish. It shows the function calls and branches in a top-to-bottom format, making it easier to follow the logic and structure. It also allows zooming in and out to explore the flow between different blocks and understand how one part of the function leads to another.



If the view is expanded, it shows `printf` and `scanf` function calls from main. This provides an understanding that the main function is directly calling these two functions, as well as other built-in functions. The view provides a clearer picture of the flow of execution within the program.