

Introduction to Reverse Engineering

Gergely Erdélyi

Research Manager

F-SECURE®



BE SURE.

Agenda

- Reverse Engineering Intro
- Ethical and Legal Aspects
- Process of Reverse Engineering
- Tools of the Trade

What is Reverse Engineering? 1/2

What is Reverse Engineering? 1/2



Image Copyright © 2005 BAE Systems

What is Reverse Engineering? 1/2



Image Copyright © 2005 BAE Systems



Image Copyright © 2005 BAE Systems

What is Reverse Engineering? 2/2

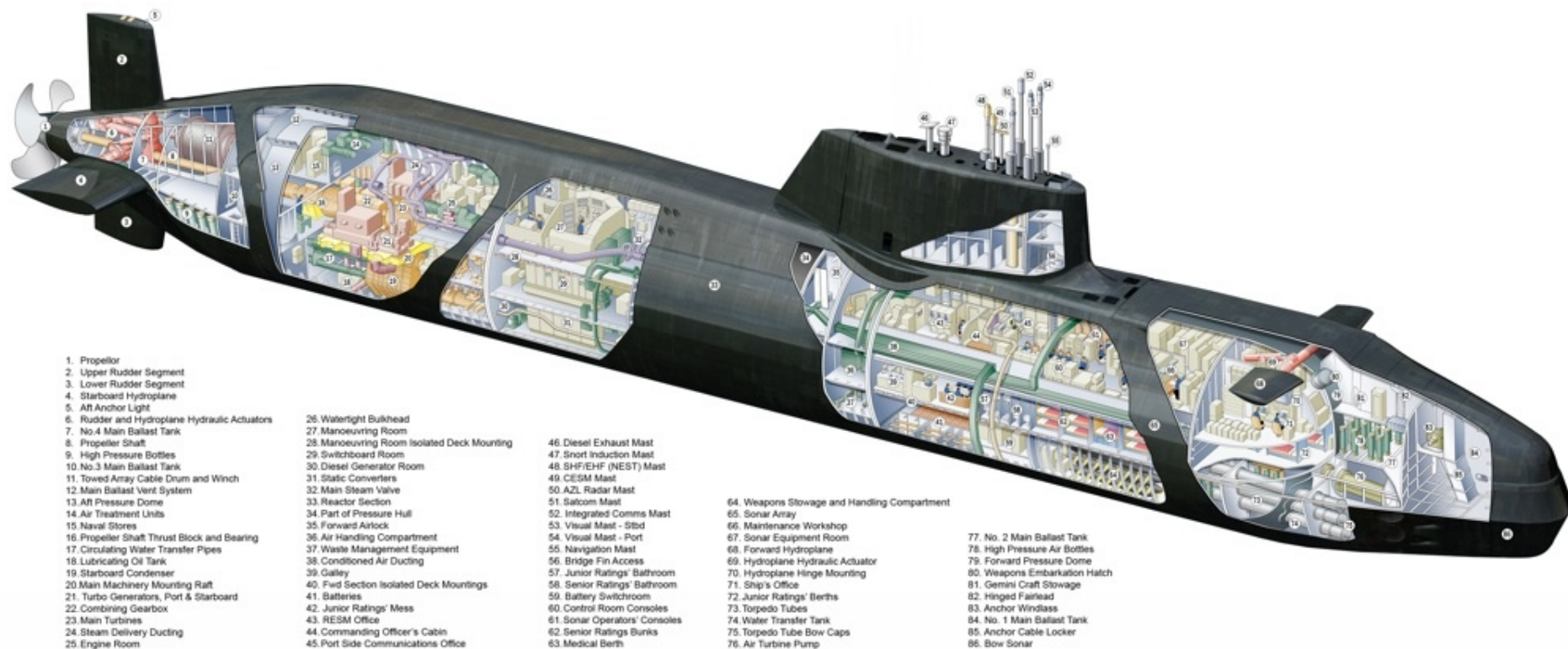
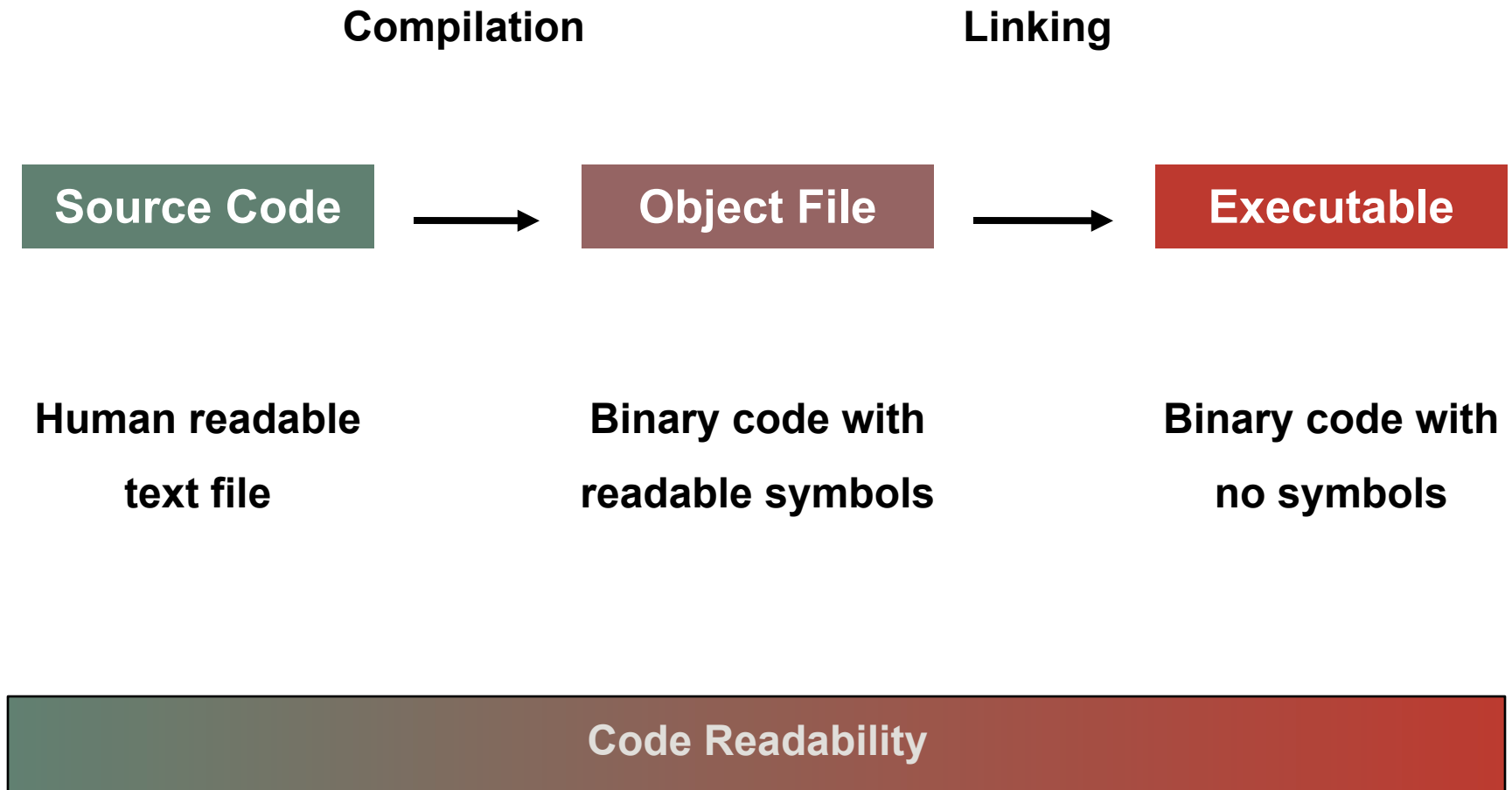


Image Copyright © 2005 BAE Systems

Reverse Code Engineering

- Reverse Engineering is also known as RE or RCE
 - **RE:** Reverse Engineering
 - **RCE:** Reverse Code Engineering
- RE is the process of understanding an existing product
- Malware analysis and security research often involves RE

Compilation Process



Compilation Results



Compilation Results

```
int ExecFile(char *FileName)
{
    PyObject* PyFileObject = PyFile_FromString(FileName, "r");

    if (!PyFileObject)
    {
        return 0;
    }

    if (PyRun_SimpleFile(PyFile_AsFile(PyFileObject), FileName) == 0)
    {
        Py_DECREF(PyFileObject);
        return 1;
    }
    else
    {
        Py_DECREF(PyFileObject);
        return 0;
    }
}
```

Compilation Results

```
int ExecFile(char *FileName)
{
```

```
    PyObject* PyFi
```

```
    if (!PyFileOb
```

```
    {
        return 0;
    }
```

```
    if (PyRun_Simp
```

```
    {
        Py_DECREF
        return 1;
    }
```

```
    else
```

```
    {
        Py_DECREF
        return 0;
    }
```

```
}
```

.text:00401250	E8 BB DA 0E 00 89 44 24	04 A1 2C A3 57 00 8B 40	F++aëD\$ í,úw.ï@
.text:00401260	10 89 04 24 E8 27 D5 0E	00 8B 15 2C A3 57 00 E9	ë \$F'+aï\$,úw.T
.text:00401270	4B FF FF FF 8D B6 00 00	00 00 8D BF 00 00 00 00	K ì ì+....
.text:00401280	55 89 E5 83 EC 08 C7 04	24 01 00 00 00 FF 15 18	Uësâ8 \$... \$
.text:00401290	A3 57 00 E8 B8 FE FF FF	90 8D B4 26 00 00 00 00	úw.F+ Éì &....
.text:004012A0	55 89 E5 83 EC 08 C7 04	24 02 00 00 00 FF 15 18	Uësâ8 \$... \$
.text:004012B0	A3 57 00 E8 98 FE FF FF	90 8D B4 26 00 00 00 00	úw.Fÿ Éì &....
.text:004012C0	55 8B 0D 54 A3 57 00 89	E5 5D FF E1 8D 74 26 00	UiTúw.ës] ßit&.
.text:004012D0	55 8B 0D 34 A3 57 00 89	E5 5D FF E1 90 90 90 90	Ui4úw.ës] ßÉÉÉÉ
.text:004012E0	83 EC 7C B8 70 B5 4E 00	89 44 24 34 B8 74 30 4F	â8 +p N.ëD\$4+t00
.text:004012F0	00 89 44 24 38 8D 44 24	60 89 44 24 3C B8 90 13	.ëD\$8iD\$`ëD\$<+É
.text:00401300	40 00 89 44 24 40 8D 44	24 1C 89 7C 24 74 89 5C	@.ëD\$@iD\$ ë \$të\
.text:00401310	24 6C 89 74 24 70 89 6C	24 78 89 64 24 44 89 04	\$lèt\$peìl\$xëd\$Dë
.text:00401320	24 E8 3A BE 0E 00 8B BC	24 80 00 00 00 85 FF 0F	\$F:+aï+\$Ç...à ¤
.text:00401330	84 8B 00 00 00 C7 04 24	10 20 57 00 8B 94 24 80	äi... \$ w.ïö\$Ç
.text:00401340	00 00 00 8D 44 24 50 89	44 24 04 BE 88 E1 56 00	...iD\$PëD\$ +ëßV.
.text:00401350	31 DB 89 74 24 50 B9 01	00 00 00 89 54 24 54 89	1 ët\$P ...ët\$Të
.text:00401360	5C 24 58 89 4C 24 20 E8	D4 59 00 00 89 44 24 04	\XëL\$ F+Y...ëD\$
.text:00401370	C7 04 24 10 20 57 00 E8	B4 5A 00 00 85 C0 74 2E	\$ w.F Z..à+t.
.text:00401380	8B 40 08 BA E8 EC 56 00	89 54 24 50 EB 34 66 90	ï@ F8V.ëT\$Pd4fé
.text:00401390	B8 E8 EC 56 00 89 44 24	50 8B 44 24 24 89 04 24	+F8V.ëD\$PïD\$ \$ë \$
.text:004013A0	B8 FF FF FF FF 89 44 24	20 E8 72 C4 0E 00 B8 E8	+ ëD\$ Fr-a+F
.text:004013B0	EC 56 00 89 44 24 50 89	F6 8D BC 27 00 00 00 00	8V.ëD\$Pë÷i+'....
.text:004013C0	31 C0 89 44 24 18 8D 44	24 1C 89 04 24 E8 6E BE	1+ëD\$ iD\$ ë \$Fn+
.text:004013D0	0E 00 8B 44 24 18 8B 5C	24 6C 8B 74 24 70 8B 7C	aïD\$ i\lît\$pi
.text:004013E0	24 74 8B 6C 24 78 83 C4	7C C3 8D B6 00 00 00 00	\$tìl\$xâ- +ì

Uses of Reverse Engineering

- Malware analysis
- Security / vulnerability research
- Driver development
- Compatibility fixes
- Legacy application support

Ethical and Legal Aspects



Disclaimer: I am not a lawyer, but here we go...

Image: Public Domain

Ethical and Legal Aspects

- Legality of reverse engineering is governed by copyright laws
- Copyright laws differ from country to country
- Reverse engineering is legal only in a few specific cases
- Black box testing does not constitute reverse engineering
- Reverse engineering for compatibility fixes is legal
- Reverse engineering spyware is illegal in most countries
- When in doubt, **do not** reverse engineer!

Legal Uses of Reverse Engineering

- Recovery of own lost source code
- Recovery of data from legacy formats
- Malware analysis and research
- Security and vulnerability research
- Copyright infringement investigations
- Finding out the contents of any database you legally purchased

Illegal Activities



Image Copyright © 2005 Klaus with K

Illegal Activities

- Illegal to reverse engineer and sell a competing product
- Illegal to crack copy protections
- Illegal to distribute a crack/registration for copyrighted software
- Illegal to gain unauthorized access to any computer system
- Copyright protected software is off-limits in most cases
- Spyware/Adware with companies behind them are included

Decompilation Process

Disassembly

Decompilation

Executable



Disassembly



Source Code

**Binary code with
no symbols**

**Reverse engineer
readable code**

**Human
readable code**

Code Readability

Disassembly Results

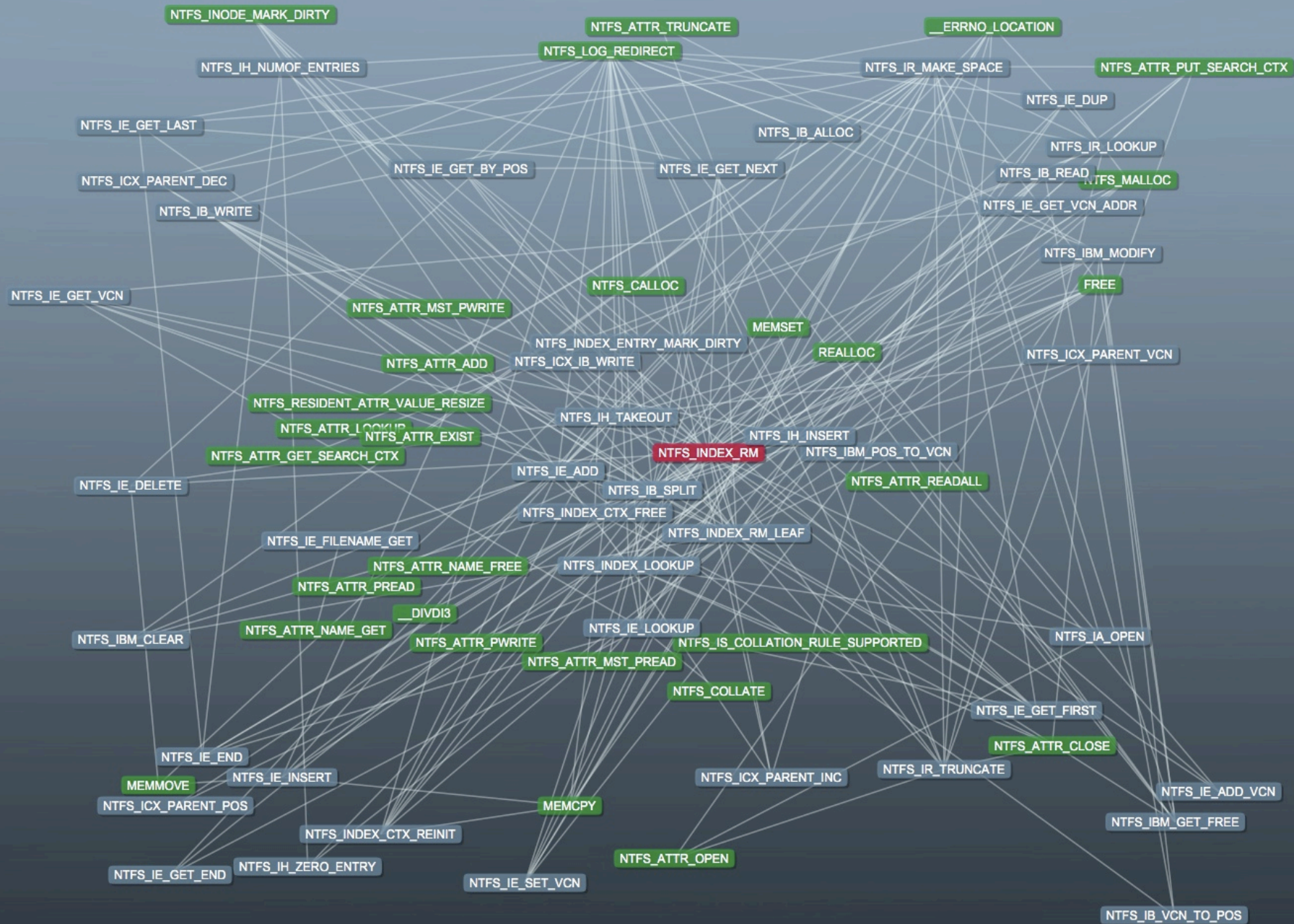


Disassembly Results

```
.text:00401250 E8 BB DA 0E 00 89 44 24 04 A1 2C A3 57 00 8B 40 F++aëD$ í,úw.ï@
.text:00401260 10 89 04 24 E8 27 D5 0E 00 8B 15 2C A3 57 00 E9 ë $F'+aï$,úw.T
.text:00401270 4B FF FF FF 8D B6 00 00 00 00 8D BF 00 00 00 00 K ì|....ì+....
.text:00401280 55 89 E5 83 EC 08 C7 04 24 01 00 00 00 FF 15 18 Uèsâ8| $...$
.text:00401290 A3 57 00 E8 B8 FE FF FF 90 8D B4 26 00 00 00 00 úw.F+| Éì|&....
.text:004012A0 55 89 E5 83 EC 08 C7 04 24 02 00 00 00 FF 15 18 Uèsâ8| $...$
.text:004012B0 A3 57 00 E8 98 FE FF FF 90 8D B4 26 00 00 00 00 úw.Fÿ| Éì|&....
.text:004012C0 55 8B 0D 54 A3 57 00 89 E5 5D FF E1 8D 74 26 00 UïTúw.ès] ßit&.
.text:004012D0 55 8B 0D 34 A3 57 00 89 E5 5D FF E1 90 90 90 90 Uï4úw.ès] ßÉÉÉÉÉ
.text:004012E0 83 EC 7C B8 70 B5 4E 00 89 44 24 34 B8 74 30 4F â8|+p|N.ëD$4+t00
.text:004012F0 00 89 44 24 38 8D 44 24 60 89 44 24 3C B8 90 13 .ëD$8iD$`ëD$<+É
.text:00401300 40 00 89 44 24 40 8D 44 24 1C 89 7C 24 74 89 5C @.ëD$@iD$ ë|$të\
.text:00401310 24 6C 89 74 24 70 89 6C 24 78 89 64 24 44 89 04 $lët$pël$xëD$Dë
.text:00401320 24 E8 3A BE 0E 00 8B BC 24 80 00 00 00 85 FF 0F $F:+aï+$Ç...à ¤
.text:00401330 84 8B 00 00 00 C7 04 24 10 20 57 00 8B 94 24 80 äï...| $ w.ïö$Ç
.text:00401340 00 00 00 8D 44 24 50 89 44 24 04 BE 88 E1 56 00 ...iD$PëD$ +ëßV.
.text:00401350 31 DB 89 74 24 50 B9 01 00 00 00 89 54 24 54 89 1|ët$P| ...ët$Të
.text:00401360 5C 24 58 89 4C 24 20 E8 D4 59 00 00 89 44 24 04 \$XëL$ F+Y..ëD$
.text:00401370 C7 04 24 10 20 57 00 E8 B4 5A 00 00 85 C0 74 2E | $ w.F|Z..à+t.
.text:00401380 8B 40 08 BA E8 EC 56 00 89 54 24 50 EB 34 66 90 ï@|F8V.ët$Pd4fÉ
.text:00401390 B8 E8 EC 56 00 89 44 24 50 8B 44 24 24 89 04 24 +F8V.ëD$PïD$$ë $
.text:004013A0 B8 FF FF FF FF 89 44 24 20 E8 72 C4 0E 00 B8 E8 + ëD$ Fr-a+F
.text:004013B0 EC 56 00 89 44 24 50 89 F6 8D BC 27 00 00 00 00 8V.ëD$Pë÷i+'....
.text:004013C0 31 C0 89 44 24 18 8D 44 24 1C 89 04 24 E8 6E BE 1+ëD$ iD$ ë $Fn+
.text:004013D0 0E 00 8B 44 24 18 8B 5C 24 6C 8B 74 24 70 8B 7C aïD$ ï\$lït$pï|
.text:004013E0 24 74 8B 6C 24 78 83 C4 7C C3 8D B6 00 00 00 00 $tïl$xâ-|+ì|....
```

Disassembly Results

```
.text:004013F0 sub_4013F0      proc near                ; CODE XREF: sub_406AB0+6F"p
.text:004013F0                                     ; sub_4601D0+5D"p
.text:004013F0
.text:004013F0 var_1C      = dword ptr -1Ch
.text:004013F0 var_18      = dword ptr -18h
.text:004013F0 arg_0       = dword ptr  4
.text:004013F0
.text:004013F0      push     edi
.text:004013F1      push     esi
.text:004013F2      push     ebx
.text:004013F3      sub      esp, 10h
.text:004013F6      mov      edi, [esp+1Ch+arg_0]
.text:004013FA      test     edi, edi
.text:004013FC      jz       short loc_40143D
.text:004013FE      mov      [esp+1Ch+var_1C], offset dword_572010
.text:00401405      call     sub_406F80
.text:0040140A      mov      ebx, eax
.text:0040140C      jmp      short loc_401439
.text:0040140C ; -----
.text:0040140E      align 10h
.text:00401410
.text:00401410 loc_401410:                ; CODE XREF: sub_4013F0+4B"j
.text:00401410      mov      [esp+1Ch+var_18], ebx
.text:00401414      mov      [esp+1Ch+var_1C], offset dword_572010
.text:0040141B      call     sub_406E30
.text:00401420      mov      [esp+1Ch+var_18], ebx
```



Required Skills

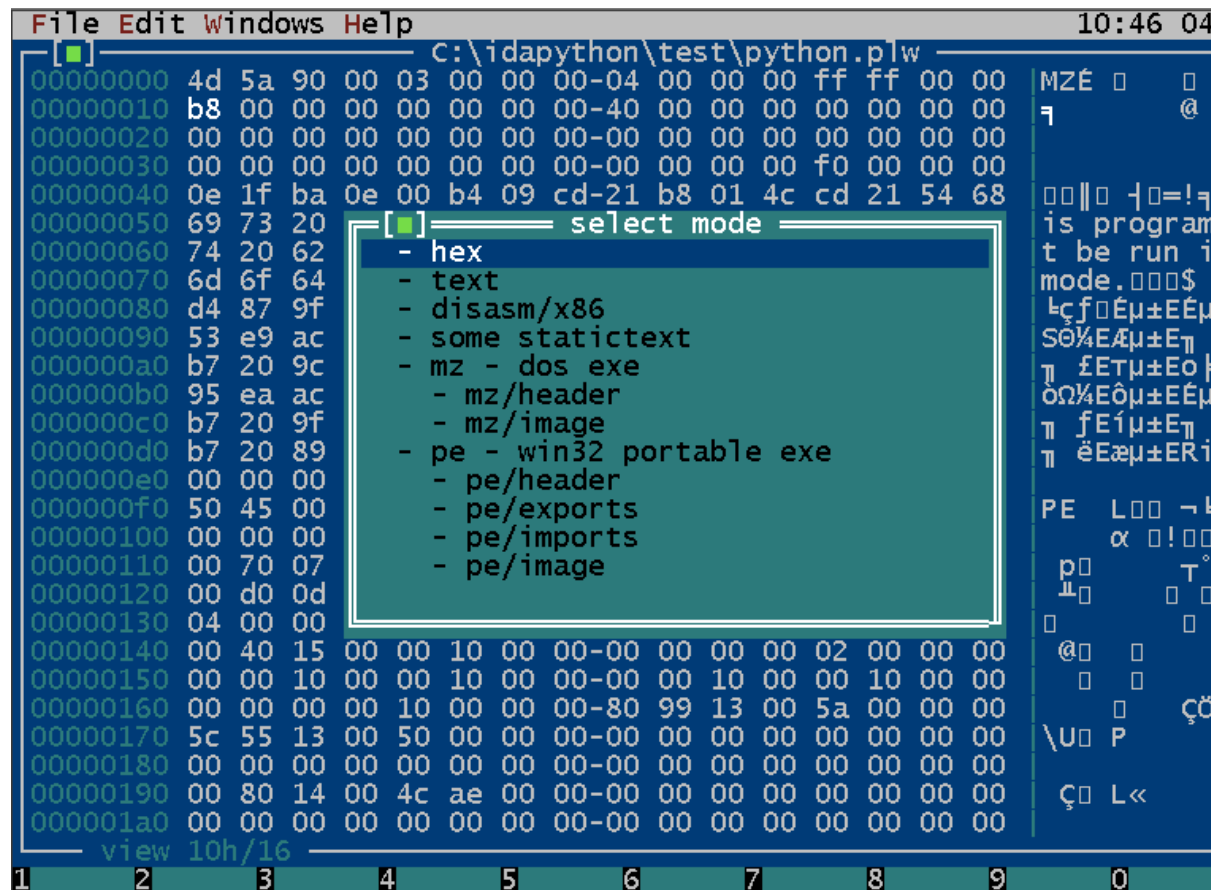
- General computer architecture knowledge
- Assembly programming of target processors
- Operating systems
- File formats
- Information search skills
- ...real persistence...

Most Commonly Used Tools

- Hex editor/viewer
- Disassembler
- Search engine
- Debugger
- Script language

Most Commonly Used Tools

- Hex editor/viewer
- Disassembler
- Search engine
- Debugger
- Script language

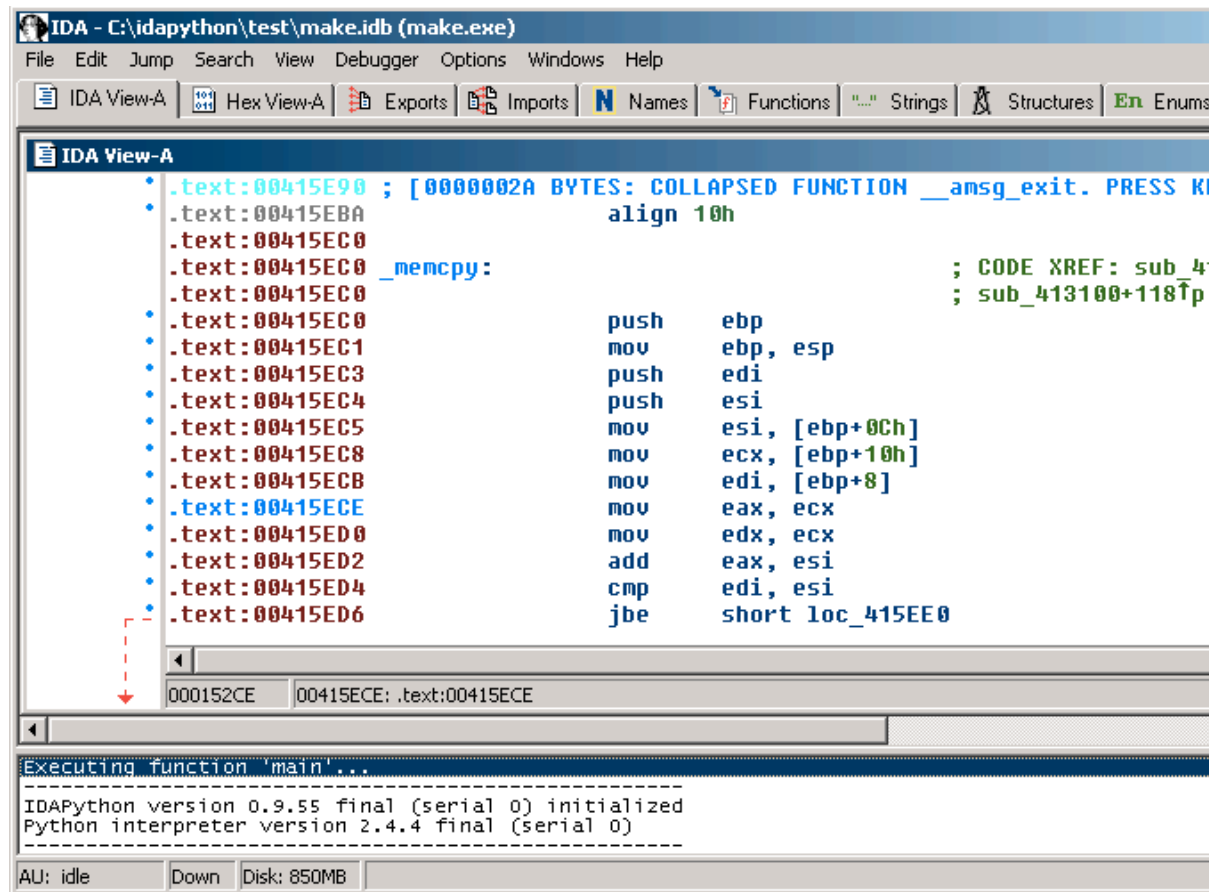


```

File Edit Windows Help
C:\idapython\test\python.plw
00000000 4d 5a 90 00 03 00 00 00-04 00 00 00 ff ff 00 00 MZÉ  @
00000010 b8 00 00 00 00 00 00 00-40 00 00 00 00 00 00 00 7
00000020 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00
00000030 00 00 00 00 00 00 00 00-00 00 00 00 f0 00 00 00
00000040 0e 1f ba 0e 00 b4 09 cd-21 b8 01 4c cd 21 54 68  []  |=!q
00000050 69 73 20 69 73 20 69 73-20 69 73 20 69 73 20 69  is program
00000060 74 20 62 74 20 62 74 20-62 74 20 62 74 20 62 74  t be run i
00000070 6d 6f 64 6d 6f 64 6d 6f-64 6d 6f 64 6d 6f 64 6d  mode.[]$
00000080 d4 87 9f d4 87 9f d4 87-9f d4 87 9f d4 87 9f d4 87 9f
00000090 53 e9 ac 53 e9 ac 53 e9-ac 53 e9 ac 53 e9 ac 53 e9 ac
000000a0 b7 20 9c b7 20 9c b7 20-9c b7 20 9c b7 20 9c b7 20 9c
000000b0 95 ea ac 95 ea ac 95 ea-ac 95 ea ac 95 ea ac 95 ea ac
000000c0 b7 20 9f b7 20 9f b7 20-9f b7 20 9f b7 20 9f b7 20 9f
000000d0 b7 20 89 b7 20 89 b7 20-89 b7 20 89 b7 20 89 b7 20 89
000000e0 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00
000000f0 50 45 00 50 45 00 50 45-00 50 45 00 50 45 00 50 45
00000100 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00
00000110 00 70 07 00 70 07 00 70-07 00 70 07 00 70 07 00 70
00000120 00 d0 0d 00 d0 0d 00 d0-0d 00 d0 0d 00 d0 0d 00 d0
00000130 04 00 00 04 00 00 04 00-00 04 00 00 04 00 00 04 00
00000140 00 40 15 00 40 15 00 40-15 00 40 15 00 40 15 00 40
00000150 00 00 10 00 00 10 00 00-10 00 00 10 00 00 10 00 00
00000160 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00
00000170 5c 55 13 5c 55 13 5c 55-13 5c 55 13 5c 55 13 5c 55
00000180 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00
00000190 00 80 14 00 80 14 00 80-14 00 80 14 00 80 14 00 80
000001a0 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00
view 10h/16
1 2 3 4 5 6 7 8 9 0
  
```

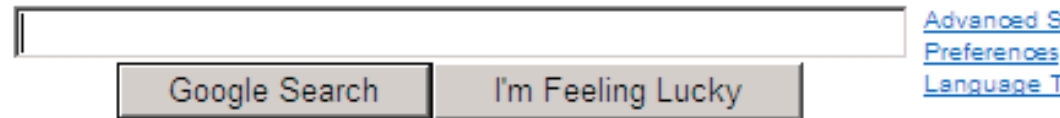
Most Commonly Used Tools

- Hex editor/viewer
- Disassembler
- Search engine
- Debugger
- Script language



Most Commonly Used Tools

- Hex editor/viewer
- Disassembler
- Search engine
- Debugger
- Script language



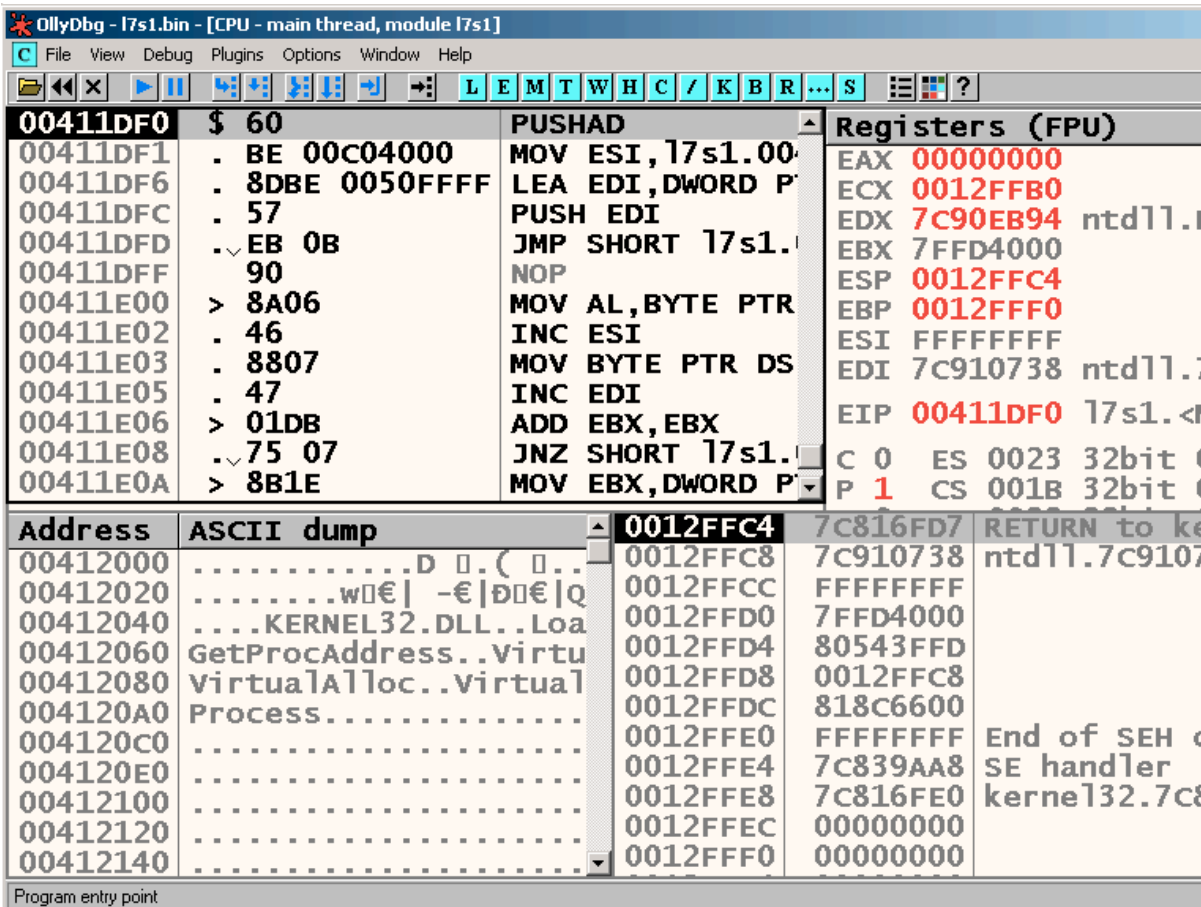
[Make Google Your Homepage!](#)

[Advertising Programs](#) - [Business Solutions](#) - [About Google](#) - [Go to Google Suomi](#)

©2008 Google

Most Commonly Used Tools

- Hex editor/viewer
- Disassembler
- Search engine
- Debugger
- Script language



Most Commonly Used Tools

- Hex editor/viewer
- Disassembler
- Search engine
- Debugger
- Script language

```
Python 2.5.1 (r251:54863, Oct 5 2007, 21:08)
Type "copyright", "credits" or "license" for
```

```
IPython 0.8.2 -- An enhanced Interactive Python
?          -> Introduction and overview of IP
%quickref  -> Quick reference.
help       -> Python's own help system.
object?    -> Details about 'object'. ?object
```

```
In [1]: █
```

Getting Started

- Master your tools
- Identify the target binary format
- Identify the target processor
- Identify the target operating system
- ...dig in and find out as much as you can...