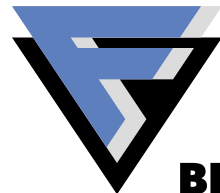


# Introduction to Reverse Engineering

Gergely Erdélyi

Research Manager

**F-SECURE<sup>®</sup>**



**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

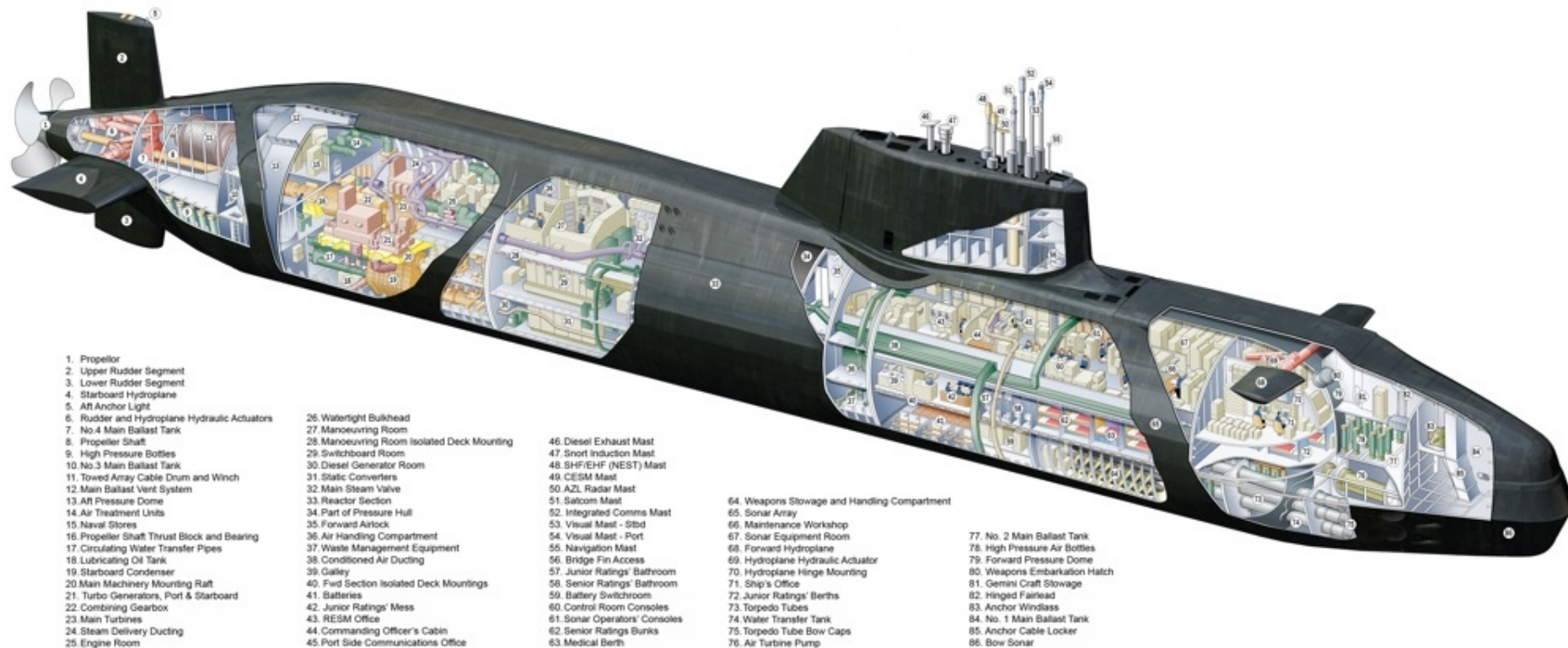


Image Copyright © 2005 BAE Systems



Image Copyright © 2005 BAE Systems

# What is Reverse Engineering? 2/2



1. Propeller
2. Upper Rudder Segment
3. Lower Rudder Segment
4. Starboard Hydroplane
5. Aft Anchor Light
6. Rudder and Hydroplane Hydraulic Actuators
7. No. 4 Main Ballast Tank
8. Propeller Shaft
9. High Pressure Bottles
10. No. 3 Main Ballast Tank
11. Towed Array Cable Drum and Winch
12. Main Ballast Vent System
13. Aft Pressure Dome
14. Air Treatment Units
15. Naval Stores
16. Propeller Shaft Thrust Block and Bearing
17. Circulating Water Transfer Pipes
18. Lubricating Oil Tank
19. Starboard Condenser
20. Main Machinery Mounting Raft
21. Turbo Generators, Port & Starboard
22. Combining Gearbox
23. Main Turbines
24. Steam Delivery Ducting
25. Engine Room

26. Watertight Bulkhead
27. Manoeuvring Room
28. Manoeuvring Room Isolated Deck Mounting
29. Switchboard Room
30. Diesel Generator Room
31. Static Converters
32. Main Steam Valve
33. Reactor Section
34. Part of Pressure Hull
35. Forward Airtock
36. Air Handling Compartment
37. Waste Management Equipment
38. Conditioned Air Ducting
39. Galley
40. Fwd Section Isolated Deck Mountings
41. Batteries
42. Junior Ratings' Mess
43. RESM Office
44. Commanding Officer's Cabin
45. Port Side Communications Office

46. Diesel Exhaust Mast
47. Snort Induction Mast
48. SHF/EHF (NEST) Mast
49. CESM Mast
50. AZL Radar Mast
51. Saloom Mast
52. Integrated Comms Mast
53. Visual Mast - Stbd
54. Visual Mast - Port
55. Navigation Mast
56. Bridge Fin Access
57. Junior Ratings' Bathroom
58. Senior Ratings' Bathroom
59. Battery Switchroom
60. Control Room Consoles
61. Senior Operators' Consoles
62. Senior Ratings Bunks
63. Medical Berth

64. Weapons Stowage and Handling Compartment
65. Sonar Array
66. Maintenance Workshop
67. Sonar Equipment Room
68. Forward Hydroplane
69. Hydroplane Hydraulic Actuator
70. Hydroplane Hinge Mounting
71. Ship's Office
72. Junior Ratings' Berths
73. Torpedo Tubes
74. Torpedo Transfer Tank
75. Torpedo Tube Bow Caps
76. Air Turbine Pump

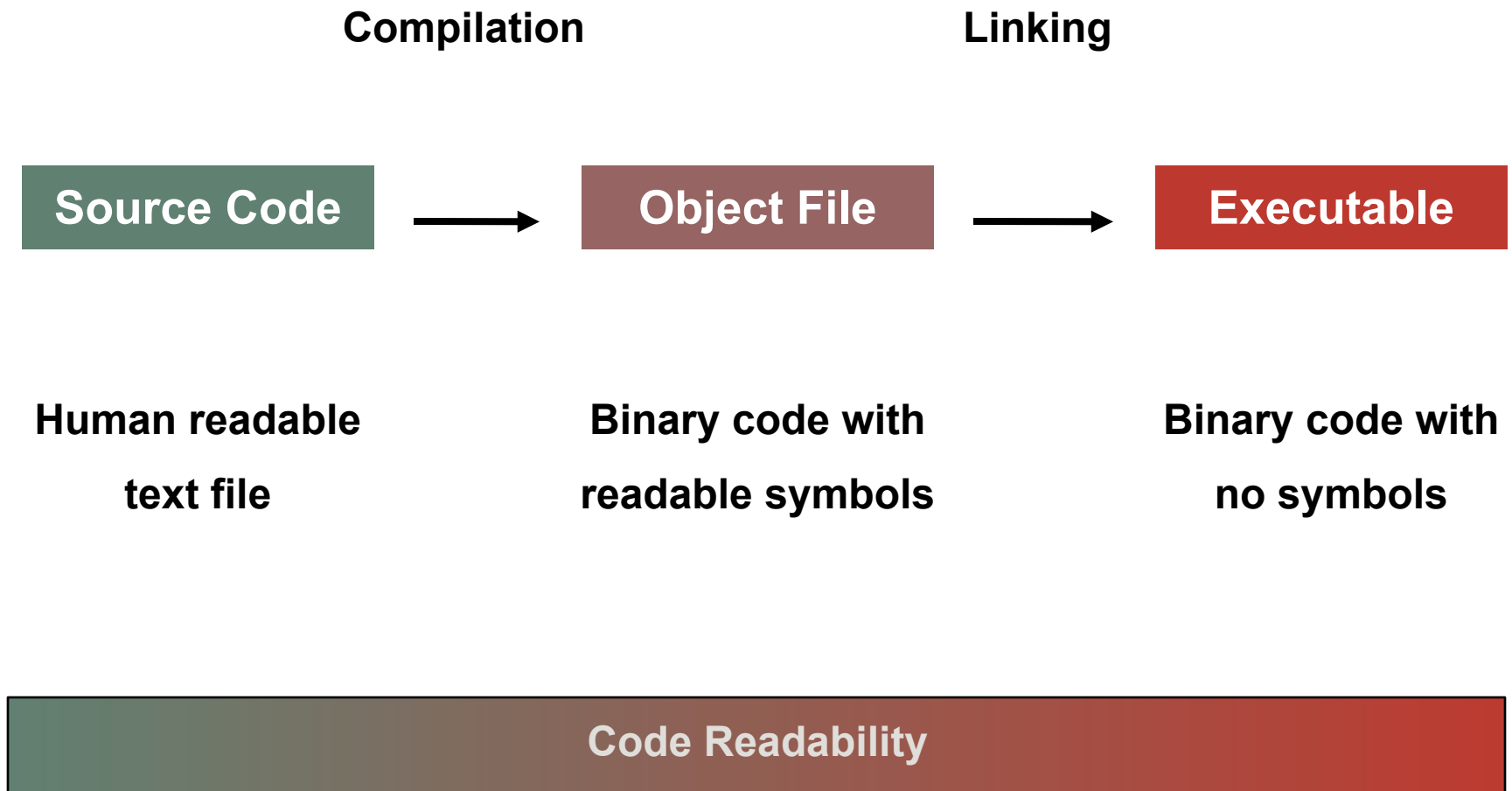
77. No. 2 Main Ballast Tank
78. High Pressure Air Bottles
79. Forward Pressure Dome
80. Weapons Embarkation Hatch
81. Gemini Craft Stowage
82. Hinged Fairlead
83. Anchor Windlass
84. No. 1 Main Ballast Tank
85. Anchor Cable Locker
86. Bow Sonar

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
    .text:00401250 E8 BB DA 0E 00 89 44 24 04 A1 2C A3 57 00 8B 40 F++aëD$ í,úw.î@
    .text:00401260 10 89 04 24 E8 27 D5 0E 00 8B 15 2C A3 57 00 E9 ë $F'+aî$,úw.T
    if (!PyFileOb
    .text:00401270 4B FF FF FF 8D B6 00 00 00 00 8D BF 00 00 00 00 K î|....î+....
    {
    .text:00401280 55 89 E5 83 EC 08 C7 04 24 01 00 00 00 FF 15 18 Uësâ8| $... $
        return 0;
    .text:00401290 A3 57 00 E8 B8 FE FF FF 90 8D B4 26 00 00 00 00 úw.F+| Éî!&....
    .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| Éî!&....
    .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 Uî4úw.ës] ßÉÉÉÉ
    if (PyRun_Simp
    .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$8îD$`ëD$<+É
        Py_DECREF
    .text:00401300 40 00 89 44 24 40 8D 44 24 1C 89 7C 24 74 89 5C @.ëD$@îD$ ë|të\
        return 1;
    .text:00401310 24 6C 89 74 24 70 89 6C 24 78 89 64 24 44 89 04 $lèt$peł$xëD$Dë
    .text:00401320 24 E8 3A BE 0E 00 8B BC 24 80 00 00 00 85 FF 0F $F:+aî+$Ç...à ¨
    .text:00401330 84 8B 00 00 00 C7 04 24 10 20 57 00 8B 94 24 80 äî...| $ w.îö$Ç
    }
    else
    .text:00401340 00 00 00 8D 44 24 50 89 44 24 04 BE 88 E1 56 00 ...îD$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
        Py_DECREF
    .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.
        return 0;
    .text:00401380 8B 40 08 BA E8 EC 56 00 89 54 24 50 EB 34 66 90 î@|F8V.ët$Pd4fé
    .text:00401390 B8 E8 EC 56 00 89 44 24 50 8B 44 24 24 89 04 24 +F8V.ëD$PîD$$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$Pë÷î+'....
    .text:004013C0 31 C0 89 44 24 18 8D 44 24 1C 89 04 24 E8 6E BE 1+ëD$ îD$ ë $Fn+
    .text:004013D0 0E 00 8B 44 24 18 8B 5C 24 6C 8B 74 24 70 8B 7C aîD$ î\lît$Pî|
    .text:004013E0 24 74 8B 6C 24 78 83 C4 7C C3 8D B6 00 00 00 00 $tîl$xâ-|+î|....
}

```

# 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 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**



**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] 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 $lët$pe|l$xëD$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 $Fn+
.text:004013D0 0E 00 8B 44 24 18 8B 5C 24 6C 8B 74 24 70 8B 7C aïD$ ì\l$it$pi|
.text:004013E0 24 74 8B 6C 24 78 83 C4 7C C3 8D B6 00 00 00 00 $tìl$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...

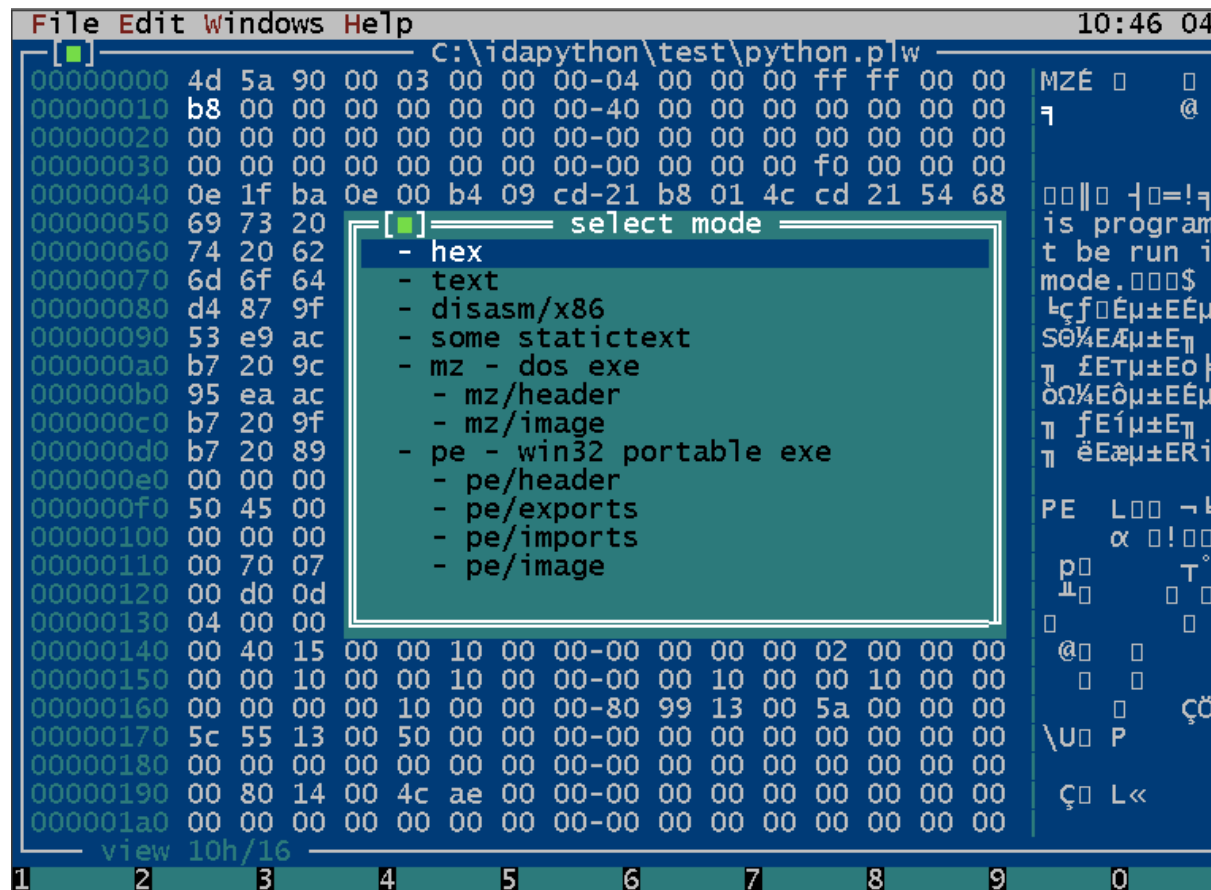
# Most Commonly Used Tools

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



# Most Commonly Used Tools

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



The screenshot shows a hex editor window with the following content:

```

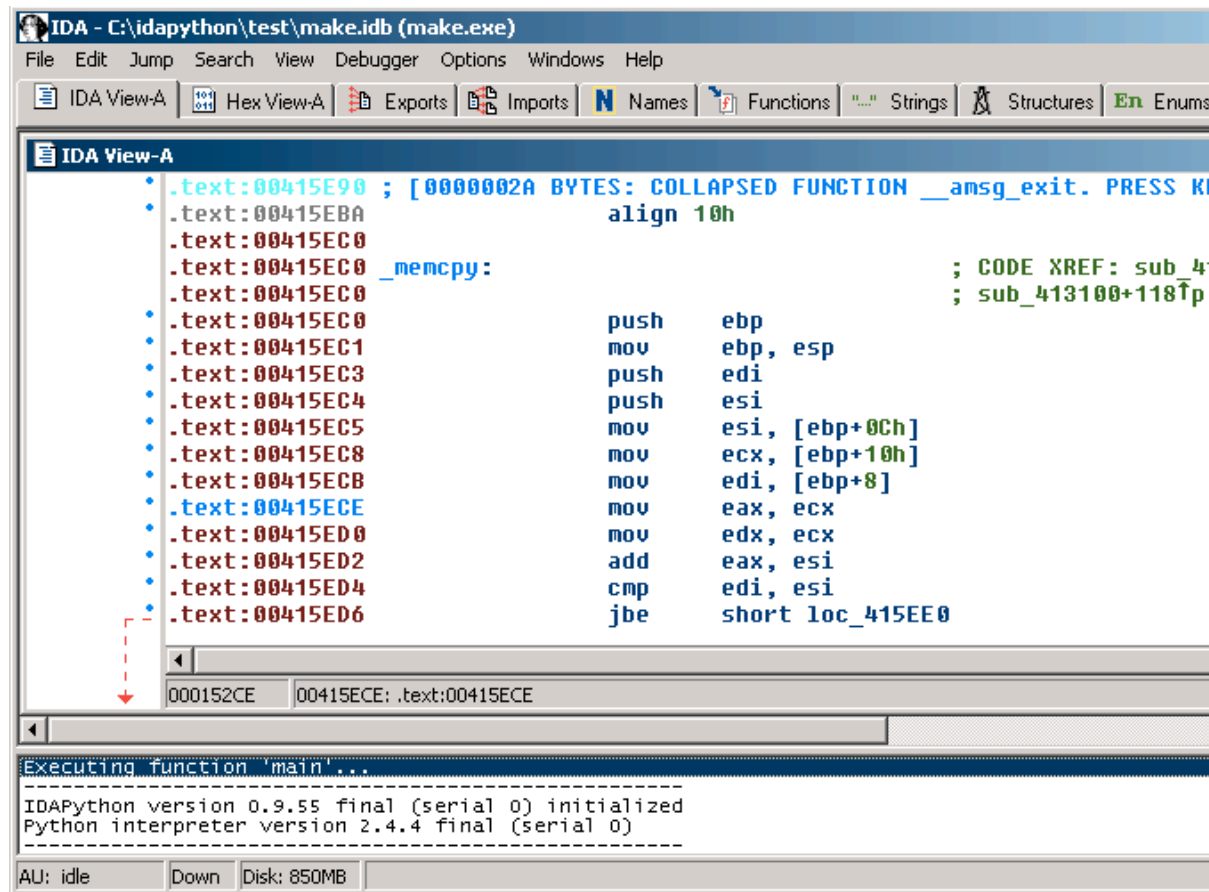
File Edit Windows Help
C:\idapython\test\python.plw 10:46 04
00000000 4d 5a 90 00 03 00 00 00-04 00 00 00 ff ff 00 00 MZÉ  @
00000010 b8 00 00 00 00 00 00 00-40 00 00 00 00 00 00 00 7
00000020 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00
00000030 00 00 00 00 00 00 00 00-00 00 00 00 f0 00 00 00
00000040 0e 1f ba 0e 00 b4 09 cd-21 b8 01 4c cd 21 54 68
00000050 69 73 20
00000060 74 20 62
00000070 6d 6f 64
00000080 d4 87 9f
00000090 53 e9 ac
000000a0 b7 20 9c
000000b0 95 ea ac
000000c0 b7 20 9f
000000d0 b7 20 89
000000e0 00 00 00
000000f0 50 45 00
00000100 00 00 00
00000110 00 70 07
00000120 00 d0 0d
00000130 04 00 00
00000140 00 40 15 00 00 10 00 00-00 00 00 00 02 00 00 00
00000150 00 00 10 00 00 10 00 00-00 00 10 00 00 10 00 00
00000160 00 00 00 00 10 00 00 00-80 99 13 00 5a 00 00 00
00000170 5c 55 13 00 50 00 00 00-00 00 00 00 00 00 00 00
00000180 00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00
00000190 00 80 14 00 4c ae 00 00-00 00 00 00 00 00 00 00
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
  
```

A dialog box titled "select mode" is overlaid on the hex editor, listing the following options:

- hex
- text
- disasm/x86
- some statictext
- mz - dos exe
  - mz/header
  - mz/image
- pe - win32 portable exe
  - pe/header
  - pe/exports
  - pe/imports
  - pe/image

# Most Commonly Used Tools

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



IDA - C:\idapython\test\make.idb (make.exe)

File Edit Jump Search View Debugger Options Windows Help

IDA View-A Hex View-A Exports Imports Names Functions Strings Structures Enums

IDA View-A

```

.text:00415E90 ; [0000002A BYTES: COLLAPSED FUNCTION __msg_exit. PRESS K
.text:00415EBA align 10h
.text:00415EC0
.text:00415EC0 _memcpy: ; CODE XREF: sub_4
.text:00415EC0 ; sub_413100+118Tp
.text:00415EC0 push ebp
.text:00415EC1 mov ebp, esp
.text:00415EC3 push edi
.text:00415EC4 push esi
.text:00415EC5 mov esi, [ebp+0Ch]
.text:00415EC8 mov ecx, [ebp+10h]
.text:00415ECB mov edi, [ebp+8]
.text:00415ECE mov eax, ecx
.text:00415ED0 mov edx, ecx
.text:00415ED2 add eax, esi
.text:00415ED4 cmp edi, esi
.text:00415ED6 jbe short loc_415EE0

```

000152CE 00415ECE: .text:00415ECE

Executing function 'main'...

-----  
IDAPython version 0.9.55 final (serial 0) initialized  
Python interpreter version 2.4.4 final (serial 0)  
-----

AU: idle Down Disk: 850MB

# Most Commonly Used Tools

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



Google<sup>™</sup>

  
  [Advanced S](#)  
[Preferences](#)  
[Language T](#)

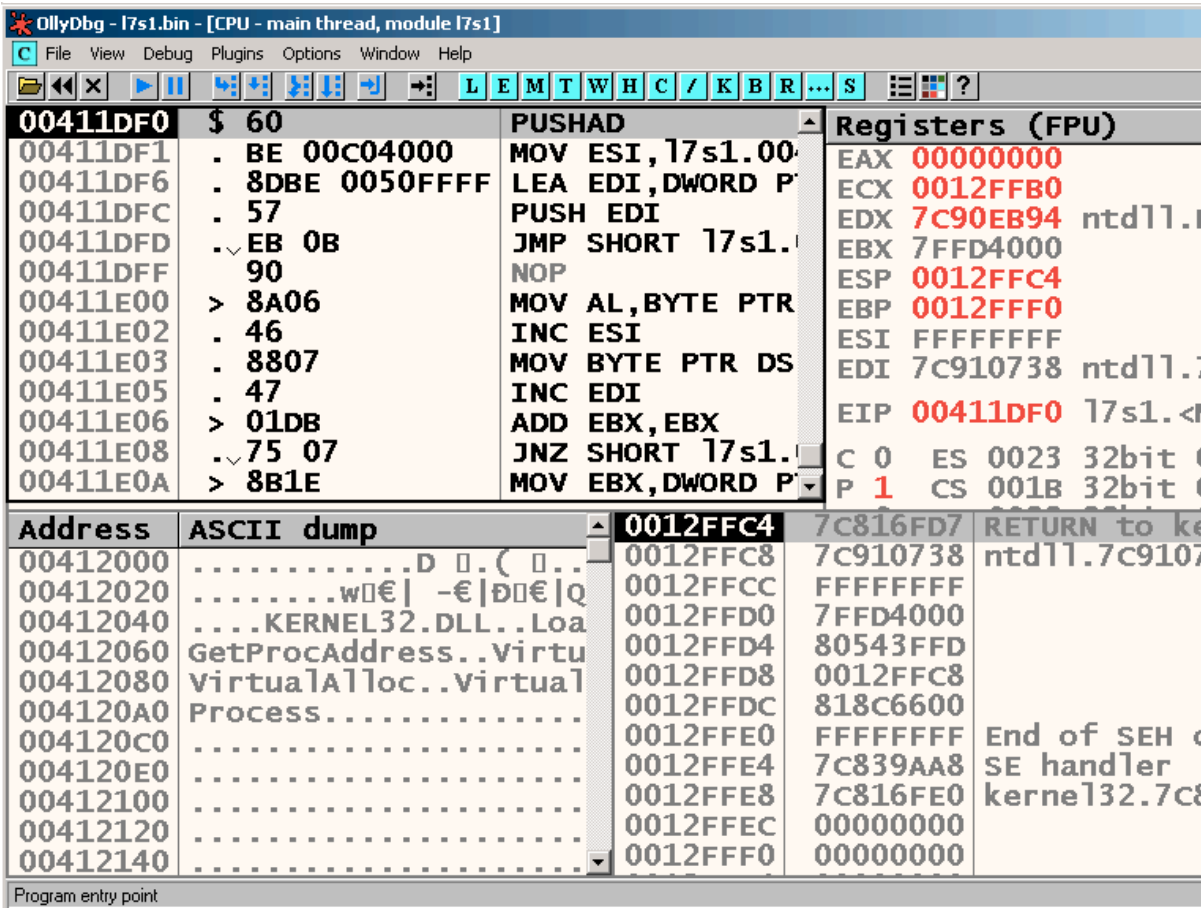
[Make Google Your Homepage!](#)

[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



The screenshot shows the OllyDbg interface with the following components:

- Assembly View:**

Address	Disassembly	Comment
00411DF0	PUSHAD	
00411DF1	MOV ESI, 17s1.00	
00411DF6	LEA EDI, DWORD P	
00411DFC	PUSH EDI	
00411DFD	JMP SHORT 17s1.	
00411DFF	NOP	
00411E00	MOV AL, BYTE PTR	
00411E02	INC ESI	
00411E03	MOV BYTE PTR DS	
00411E05	INC EDI	
00411E06	ADD EBX, EBX	
00411E08	JNZ SHORT 17s1.	
00411E0A	MOV EBX, DWORD P	
- Registers (FPU):**

Register	Value	Comment
EAX	00000000	
ECX	0012FFB0	
EDX	7C90EB94	ntd11.
EBX	7FFD4000	
ESP	0012FFC4	
EBP	0012FFFO	
ESI	FFFFFFFF	
EDI	7C910738	ntd11.
EIP	00411DF0	17s1.<
- Memory Dump (Address vs ASCII):**

Address	ASCII dump
00412000	.....D .( .
00412020	.....w  -€ D € Q
00412040	...KERNEL32.DLL..Loa
00412060	GetProcAddress..Virtu
00412080	VirtualAlloc..Virtual
004120A0	Process.....
004120C0	.....
004120E0	.....
00412100	.....
00412120	.....
00412140	.....
- Registers (FPU) - Lower Section:**

Register	Value	Comment
C 0	ES 0023	32bit
P 1	CS 001B	32bit

# 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 more details.

>>>

IPython 0.8.2 -- An enhanced Interactive Python
?          -> Introduction and overview of IPython
%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...