Ghidra: https://ghidra-sre.org/

- An open-source reverse engineering tool from the NSA
- Functionality similar to Binary Ninja and IDA

Ghidra Basics

- Ghidra Project - a collection of binaries and libraries under analysis
- Ghidra Program - an individual binary under analysis
- Ghidra Tool - an analysis tool
  - Code Browser
- Ghidra Script - a “plugin” we write on top of Ghidra API, to check for overflows, leaks, other

Ghidra Basics

- A reverse engineering/decompilation tool
  - Builds high-level program representation from x86 binaries
- A program analysis framework
  - Exposes API into program representation
- All the “perks” of program analysis frameworks
  - API is huge!
  - Not documented!
  - Buggy!

Ghidra Example

crackme, Ghidra Decompiled

```c
local_1c = 0xf;
iVar2 = _strlen(buf);
if (iVar2 == buf) {
    local_3c = 0;
    pcStack24 = buf;
    while (local_3c < local_1c) {
        local_38[local_3c] = pcStack24[local_3c] ^ 0x21;
        local_3c = local_3c + 1;
    }
    local_c = _strcmp("ycbwkgyubgfajd", (char *)local_38);
} else {
    _printf("String of incorrect length! Returning...
    local_c = 0;
}
return (int)local_c;
```

After Copy Propagation and Some Obvious Renaming

```c
len = _strlen(buf);
if (len == buf) {
    i = 0;
    while (i < 0xf) {
        local_38[i] = buf[i] ^ 0x21;
        i = i + 1;
    }
    local_c = _strcmp("ycbwkgyubgfajd", (char *)local_38);
} else {
    _printf("String of incorrect length! Returning...
    local_c = 0;
}
return (int)local_c;
```
Program Analysis Basics

- Intermediate Representations
- Control Flow Graphs (CFG) and Basic Blocks
- Data-flow Analysis and Def-use Chains
- Static Single Assignment Form (SSA Form)
  - We’ll cover SSA; we already know the rest of these

Intermediate Representation (IR)

- Compiler (roughly):
  
  ![Diagram of Compiler Process]

- Ghidra is a Decompiler:
  
  ![Diagram of Ghidra Process]

Ghidra’s Pcode

- Pcode is a 3-address code IR
  - Reversed from binary, much lower level
- Varnode - a “variable”
  - a register, stack location, constant, other
  - E.g., `(stack, 0xfffffffffffffffe0, 4)`
- PcodeOp - a 3-address statement
  - E.g., `(register, 0x0, 4) INT_ADD
    (stack, 0xfffffffffffffffe0, 4),
    (const, 0x4, 4)`

Static Single Assignment (SSA Form)

- Multiple definitions for a given use
- E.g., def-use chains for `local_24`: (5,6) and (9,6)
- That’s inconvenient!

SSA Form

- SSA form transforms the 3-address code of a program so that each variable is defined exactly once (statically)
- Standard and universally applied technique in static analysis
- Cytron et al. “Efficiently Computing Static Single Assignment Form and the Control Dependence Graph”, TOPLAS 1991
SSA Form

- Easy case
  \[
  x = \text{INPUT}; \quad x_1 = \text{INPUT}; \\
  y = x + 10; \quad \text{becomes} \quad y = x_1 + 10; \\
  x = x + y; \quad x_2 = x_1 + y
  \]

  The 2 definitions of \( x \) become definitions of \( x_1 \) and \( x_2 \) respectively.

Phi-nodes

- If-then-else
  \[
  \text{if } (a>0) \quad \text{if } (a>0) \\
  x = 5; \quad x_1 = 5; \\
  \text{else} \quad \text{becomes} \quad \text{else} \\
  x = 10; \quad x_2 = 10; \\
  x_3 = \phi(x_1, x_2)
  \]

  Phi-node introduces a (static) definition of \( x \), \( x_3 \). If control took True arm, then \( x_3 \) is \( x_1 \), otherwise it is \( x_2 \).

- Loops
  \[
  x = 0; \quad x_1 = 0; \\
  x_2 = \phi(x_1, x_3) \\
  \text{while } (x \leq n) \quad \text{while } (x_2 \leq n) \\
  x = x + 1; \quad x_3 = x_2 + 1;
  \]

  If control took forward edge, \( x_2 \) is \( x_1 \); otherwise, i.e., control took back edge, it is \( x_3 \).

fill_bowl Example, in SSA Form

- Now there is a single def-use chain for \( \text{local} \_24 \_2 \): (6,7)

Putting it all Together in Ghidra

- Essential classes and methods:
  - PcodeOp, Varnode, PcodeBlockBasic
  - var.getDef() retrieves definition of \( \text{var} \)

- Phi-nodes: MULTIEQUAL
  \[
  (\text{stack}, \text{0xffffffffffffffe0}, 4) \\
  \text{MULTIEQUAL} \\
  (\text{stack}, \text{0xffffffffffffffe0}, 4); \\
  (\text{register}, \text{0x0}, 4)
  \]

Ghidra Example

- VC Script Demo