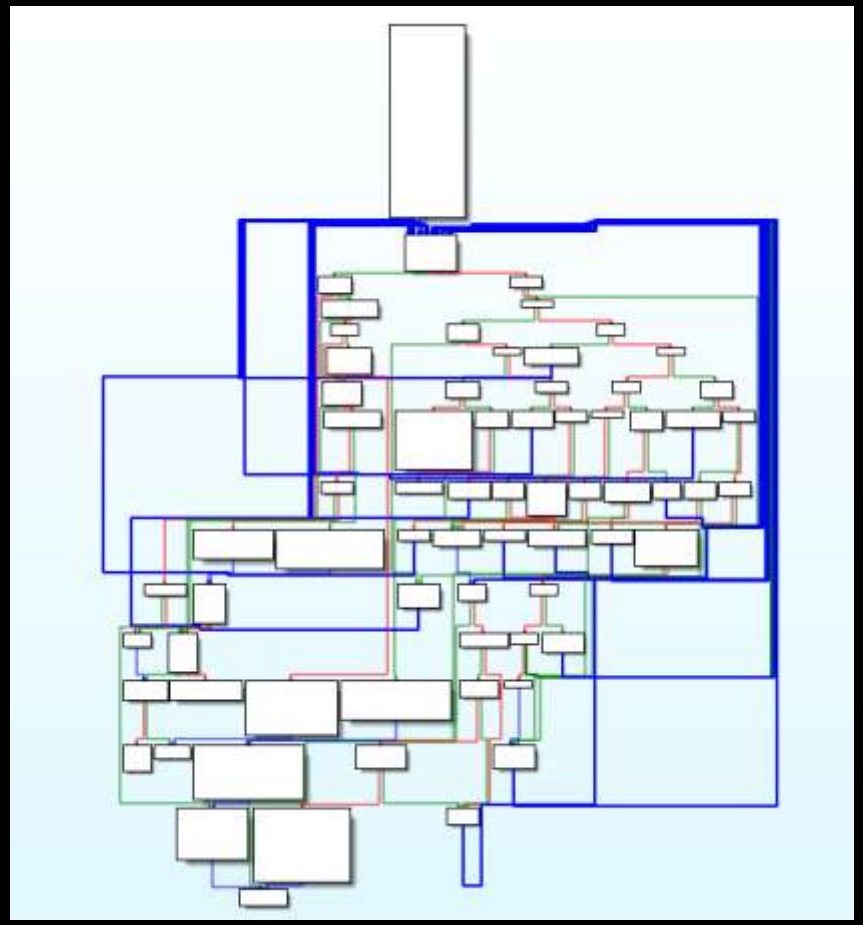
Dissassemblers
&
CFGs

Disassembly

Control Flow Graph (CFG)

- "Basic block"
- "Superblock"



Difficulties

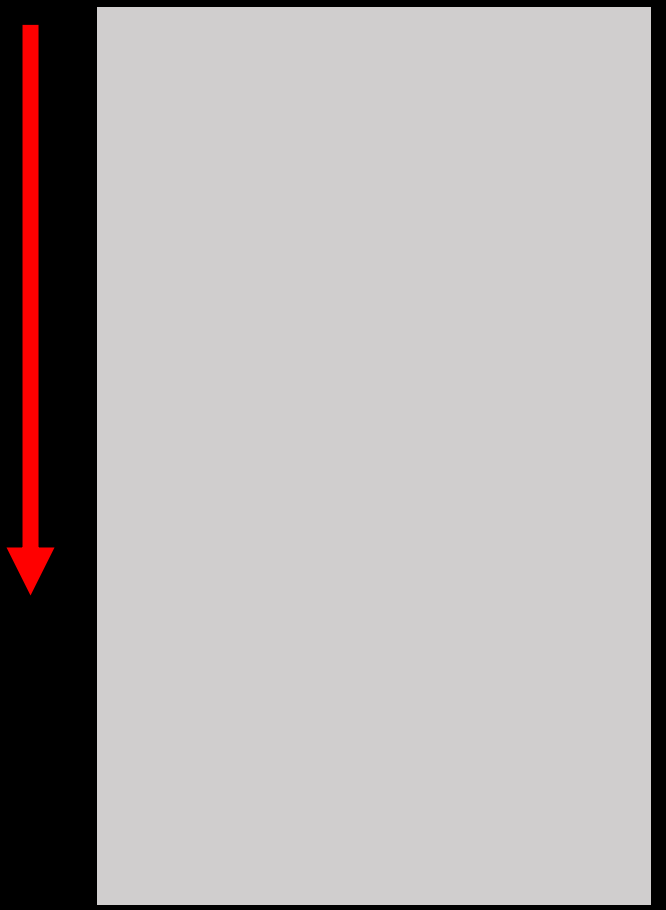
- Data in code
- Variable length instructions
- Indirect jumps
- Nonstandard control-flow constructions
- Nonstandard entrypoints

Methods

- Linear Disassembly
- Recursive Disassembly

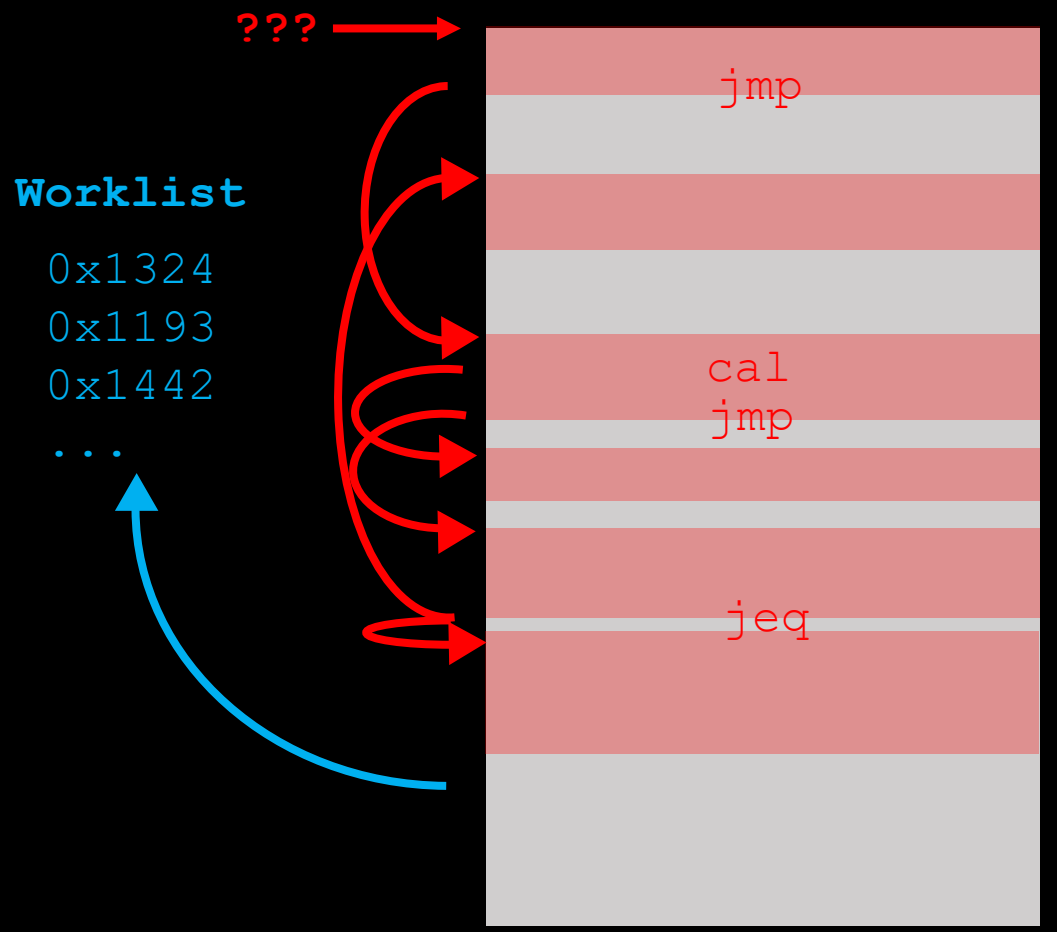
Linear Disassembly

- Start at an address
- Decode until you can’t
- Stop
- Advantages:
  - Simplicity
  - Speed
  - Will decode everything
- Disadvantages:
  - May decode more than everything



Recursive Disassembly

- Given set of start points
- Disassemble until control flow instruction
- Add destinations to worklist
- Repeat until worklist empty



Recursive Disassembly

- Identifying starting points
  - Function symbols
  - Section/segment permissions
  - Entry point
  - Function prologue detection
    - Hardcoded
    - Byteweight

Further Challenges

- Indirection/Jump Tables
  - Switch/case

Constant Propagation

```
lod r1, <roaddress>
add r1, 4
jmp r1
```

Backwards Slicing

```
lod r1, <roaddress>
…
…
…
add r1, 4
…
…
jmp r1
```