



Assembly Language: Part 2

Jennifer Rexford

1



Goals of this Lecture

Help you learn:

- Intermediate aspects of x86-64 assembly language...
- Control flow with signed integers
- Control flow with unsigned integers
- Arrays
- Structures

2



Agenda

Flattened C code

Control flow with signed integers
Control flow with unsigned integers
Arrays
Structures

3



Flattened C Code

Problem

- Translating from C to assembly language is difficult when the C code contains **nested statements**

Solution

- Flatten** the C code to eliminate all nesting

4



Flattened C Code

C	Flattened C
<pre>if (expr) { statement1; ... statementN; }</pre>	<pre>if (!expr) goto endif1; statement1; ... statementN; endif1:</pre>
<pre>if (expr) { statementT1; ... statementTN; } else { statementF1; ... statementFN; }</pre>	<pre>if (!expr) goto else1; statement1; ... statementN; goto endif1; else1: statementF1; ... statementFN; endif1:</pre>

5



Flattened C Code

C	Flattened C
<pre>while (expr) { statement1; ... statementN; }</pre>	<pre>loop1: if (!expr) goto endloop1; statement1; ... statementN; goto loop1; endloop1:</pre>
<pre>for (expr1; expr2; expr3) { statement1; ... statementN; }</pre>	<pre>expr1; loop1: if (!expr2) goto endloop1; statement1; ... statementN; expr3; goto loop1; endloop1:</pre>

See Bryant & O' Hallaron book for faster patterns

6

Agenda

Flattened C code

Control flow with signed integers

Control flow with unsigned integers

Arrays

Structures



7

if Example

C

```
int i;
...
if (i < 0)
    i = -i;
```

Flattened C

```
int i;
...
if (i >= 0) goto endif1;
i = -i;
endif1:
```



8

if Example



Flattened C

```
int i;
...
if (i >= 0) goto endif1;
i = -i;
endif1:
```

Assem Lang

```
.section ".bss"
i: .skip 4
...
.section ".text"
...
cmpl $0, i
jge endif1
negl i
endif1:
```

Note:

- cmp instruction (counterintuitive operand order)
- Sets CC bits in EFLAGS register
- jge instruction (conditional jump)
- Examines CC bits in EFLAGS register

9

if...else Example

C

```
int i;
int j;
int smaller;
...
if (i < j)
    smaller = i;
else
    smaller = j;
```

Flattened C

```
int i;
int j;
int smaller;
...
if (i >= j) goto else1;
smaller = i;
goto endif1;
else1:
    smaller = j;
endif1:
```



10

if...else Example



Flattened C

```
int i;
int j;
int smaller;
...
if (i >= j) goto else1;
smaller = i;
goto endif1;
else1:
    smaller = j;
endif1:
```

Assem Lang

```
.section ".bss"
i: .skip 4
j: .skip 4
smaller: .skip 4
...
.section ".text"
...
movl i, %eax
cmpl j, %eax
jge else1
movl i, %eax
movl %eax, smaller
jmp endif1
else1:
    movl j, %eax
    movl %eax, smaller
endif1:
```

Note:

- jmp instruction (unconditional jump)

11

while Example

C

Flattened C

```
int fact;
int n;
...
fact = 1;
while (n > 1)
{ fact *= n;
  n--;
}
loop1:
if (n <= 1) goto endloop1;
fact *= n;
n--;
goto loop1;
endloop1:
```



12

while Example

Flattened C

```
int fact;
int n;
...
fact = 1;
loop1:
if (n <= 1) goto endloop1;
fact *= n;
n--;
goto loop1;
endloop1:
```

Note:
jle instruction (conditional jump)
imul instruction

Assem Lang

```
.section ".bss"
fact: .skip 4
n: .skip 4
...
.section ".text"
...
movl $1, fact
loop1:
cmp $1, n
jle endloop1
movl fact, %eax
imull n
movl %eax, fact
decl n
jmp loop1
endloop1:
```

13

for Example

C

```
int power = 1;
int base;
int exp;
int i;
...
for (i = 0; i < exp; i++)
power *= base;
```

Flattened C

```
int power = 1;
int base;
int exp;
int i;
...
i = 0;
loop1:
if (i >= exp) goto endloop1;
power *= base;
i++;
goto loop1;
endloop1:
```

14

for Example

Flattened C

```
int power = 1;
int base;
int exp;
int i;
...
i = 0;
loop1:
if (i >= exp) goto endloop1;
power *= base;
i++;
goto loop1;
endloop1:
```

Assem Lang

```
.section ".data"
power: .long 1
.section ".bss"
base: .skip 4
exp: .skip 4
i: .skip 4
...
.section ".text"
...
movl $0, i
loop1:
movl i, %eax
cmpl exp, %eax
jge endloop1
movl power, %eax
imull base
movl %eax, power
incl i
jmp loop1
endloop1:
```

15

Control Flow with Signed Integers

Comparing signed integers

```
cmp{q,l,w,b} srcIRM, destRM Compare dest with src
```

- Sets CC bits in the EFLAGS register
- Beware: operands are in counterintuitive order
- Beware: many other instructions set CC bits
 - Conditional jump should **immediately** follow cmp

16

Control Flow with Signed Integers

Unconditional jump

```
jmp label Jump to label
```

Conditional jumps after comparing signed integers

```
je label Jump to label if equal
jne label Jump to label if not equal
jl label Jump to label if less
jle label Jump to label if less or equal
jg label Jump to label if greater
jge label Jump to label if greater or equal
```

- Examine CC bits in EFLAGS register

17

Agenda

Flattened C

Control flow with signed integers

Control flow with unsigned integers

Arrays

Structures

18

Signed vs. Unsigned Integers

In C

- Integers are signed or unsigned
- Compiler generates assembly language instructions accordingly

In assembly language

- Integers are neither signed nor unsigned
- Distinction is in the instructions used to manipulate them

Distinction matters for

- Multiplication and division
- Control flow



19

Handling Unsigned Integers

Multiplication and division

- Signed integers: `imul, idiv`
- Unsigned integers: `mul, div`

Control flow

- Signed integers: `cmp + {je, jne, jl, jle, jg, jge}`
- Unsigned integers: “unsigned cmp” + `{je, jne, jl, jle, jg, jge}` No!!!
- Unsigned integers: `cmp + {je, jne, jb, jbe, ja, jae}`



20

while Example

```
C
unsigned int fact;
unsigned int n;
...
fact = 1;
while (n > 1)
{ fact *= n;
  n--;
}
```

```
Flattened C
unsigned int fact;
unsigned int n;
...
fact = 1;
loop1:
  if (n <= 1) goto endloop1;
  fact *= n;
  n--;
  goto loop1;
endloop1:
```

21



while Example

```
Flattened C
unsigned int fact;
unsigned int n;
...
fact = 1;
loop1:
  if (n <= 1) goto endloop1;
  fact *= n;
  n--;
  goto loop1;
endloop1:
```

```
Assem Lang
.section ".bss"
fact: .skip 4
n: .skip 4
...
.movl $1, fact
loop1:
  cmpl $1, n
  jbe endloop1
  movl fact, %eax
  mull n
  movl %eax, fact
  decl n
  jmp loop1
endloop1:
```



22

for Example

```
C
unsigned int power = 1;
unsigned int base;
unsigned int exp;
unsigned int i;
...
for (i = 0; i < exp; i++)
  power *= base;
```

```
Flattened C
unsigned int power = 1;
unsigned int base;
unsigned int exp;
unsigned int i;
...
i = 0;
loop1:
  if (i >= exp) goto endloop1;
  power *= base;
  i++;
  goto loop1;
endloop1:
```

23



for Example

```
Flattened C
unsigned int power = 1;
unsigned int base;
unsigned int exp;
unsigned int i;
...
i = 0;
loop1:
  if (i >= exp) goto endloop1;
  power *= base;
  i++;
  goto loop1;
endloop1:
```

```
Assem Lang
.section ".data"
power: .long 1
.section ".bss"
base: .skip 4
exp: .skip 4
i: .skip 4
...
.movl $0, i
loop1:
  movl i, %eax
  cmpl exp, %eax
  jae endloop1
  movl power, %eax
  mull base
  movl %eax, power
  incl i
  jmp loop1
endloop1:
```



24

Control Flow with Unsigned Integers

Comparing unsigned integers

- Same as comparing signed integers

Conditional jumps after comparing unsigned integers

```
je label Jump to label if equal
jne label Jump to label if not equal
jb label Jump to label if below
jbe label Jump to label if below or equal
ja label Jump to label if above
jae label Jump to label if above or equal
```

- Examine CC bits in EFLAGS register



25

Agenda



Flattened C

Control flow with signed integers

Control flow with unsigned integers

Arrays

Structures

26

Arrays: Indirect Addressing

C

```
int a[100];
int i;
int n;
...
i = 3;
...
n = a[i];
```

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movslq i, %rax
salq $2, %rax
addq $a, %rax
movl (%rax), %r10d
movl %r10d, n
...
```

One step at a time...



27

Arrays: Indirect Addressing

Registers

RAX	[]
R10	[]
...	[]

Memory

a	[]	0	1000
	[]	1	1004
	[]	2	1008
	[]	3	1012
	[]
	[]	100	1396
	[]	i	1400
	[]	n	1404

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movslq i, %rax
salq $2, %rax
addq $a, %rax
movl (%rax), %r10d
movl %r10d, n
...
```



28

Arrays: Indirect Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movslq i, %rax
salq $2, %rax
addq $a, %rax
movl (%rax), %r10d
movl %r10d, n
...
```

Registers

RAX	[3]
R10	[]
...	[]

Memory

a	[]	0	1000
	[]	1	1004
	[]	2	1008
	[]	3	1012
	[]
	[]	100	1396
	[]	i	1400
	[]	n	1404



29

Arrays: Indirect Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movslq i, %rax
salq $2, %rax
addq $a, %rax
movl (%rax), %r10d
movl %r10d, n
...
```

Registers

RAX	[12]
R10	[]
...	[]

Memory

a	[]	0	1000
	[]	1	1004
	[]	2	1008
	[]	3	1012
	[]
	[]	100	1396
	[]	i	1400
	[]	n	1404



30

Arrays: Indirect Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movslq i, %rax
salq $2, %rax
addq $a, %rax
movl (%rax), %r10d
movl %r10d, n
...
```

Registers

RAX **1012**
R10

Memory

0	1000
1	1004
2	1008
3	1012
...	...
100	1396
i	1400
n	1404



31

Arrays: Indirect Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movslq i, %rax
salq $2, %rax
addq $a, %rax
movl (%rax), %r10d
movl %r10d, n
...
```

Registers

RAX **1012**
R10 **123**

Memory

0	1000
1	1004
2	1008
3	1012
...	...
100	1396
i	1400
n	1404



32

Note:
Indirect addressing

Arrays: Indirect Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movslq i, %rax
salq $2, %rax
addq $a, %rax
movl (%rax), %r10d
movl %r10d, n
...
```

Registers

RAX **1012**
R10 **123**

Memory

0	1000
1	1004
2	1008
3	1012
...	...
100	1396
i	1400
n	1404



33

Arrays: Base+Disp Addressing

Assem Lang

```
.section ".bss"
int a[100];
int i;
int n;
...
i = 3;
...
n = a[i]
```

C



34

One step at a time...

Arrays: Base+Disp Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movl i, %eax
salq $2, %eax
movl a(%eax), %r10d
movl %r10d, n
...
```

Registers

RAX
R10

Memory

0	1000
1	1004
2	1008
3	1012
...	...
100	1396
i	1400
n	1404



35

Arrays: Base+Disp Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movl i, %eax
salq $2, %eax
movl a(%eax), %r10d
movl %r10d, n
...
```

Registers

RAX **3**
R10

Memory

0	1000
1	1004
2	1008
3	1012
...	...
100	1396
i	1400
n	1404

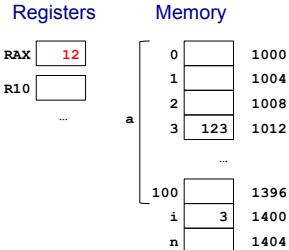


36

Arrays: Base+Disp Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movl i, %eax
call $2, %eax
movl a(%eax), %r10d
movl %r10d, n
...
```

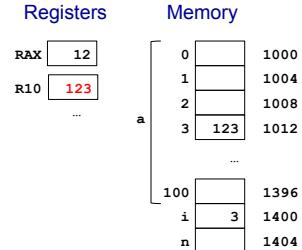


37

Arrays: Base+Disp Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movl i, %eax
call $2, %eax
movl a(%eax), %r10d
movl %r10d, n
...
```



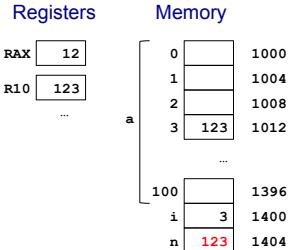
Note:
Base+displacement addressing

38

Arrays: Base+Disp Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movl i, %eax
call $2, %eax
movl a(%eax), %r10d
movl %r10d, n
...
```



39

Arrays: Scaled Indexed Addressing

C

```
int a[100];
int i;
int n;
...
i = 3;
...
n = a[i]
```

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movl i, %eax
movl a(%eax,4), %r10d
movl %r10d, n
...
```

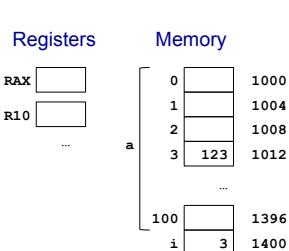
One step at a time...

40

Arrays: Scaled Indexed Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movl i, %eax
movl a(%eax,4), %r10d
movl %r10d, n
...
```

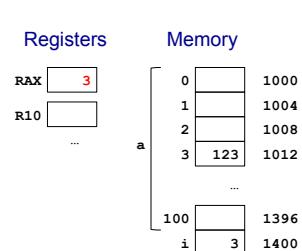


41

Arrays: Scaled Indexed Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movl i, %eax
movl a(%eax,4), %r10d
movl %r10d, n
...
```



42

Arrays: Scaled Indexed Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movl i, %eax
movl a(%eax,4), %r10d
movl %r10d, n
...
```

Registers	Memory
RAX [3]	0 [] 1000
R10 [123]	1 [] 1004
...	2 [] 1008
a	3 [123] 1012
...	...
100	100 [] 1396
i	3 [] 1400
n	1404 []

Note:
Scaled indexed addressing

Arrays: Scaled Indexed Addressing

Assem Lang

```
.section ".bss"
a: .skip 400
i: .skip 4
n: .skip 4
...
.section ".text"
...
movl $3, i
...
movl i, %eax
movl a(%eax,4), %r10d
movl %r10d, n
...
```

Registers	Memory
RAX [12]	0 [] 1000
R10 [123]	1 [] 1004
...	2 [] 1008
a	3 [123] 1012
...	...
100	100 [] 1396
i	3 [] 1400
n	1404 [123]

Generalization: Memory Operands

Full form of memory operands:

displacement(base, index, scale)

- **displacement** is an integer or a label (default = 0)
- **base** is a 4-byte or 8-byte register
- **index** is a 4-byte or 8-byte register
- **scale** is 1, 2, 4, or 8 (default = 1)

Meaning

- Compute the sum
(displacement) + (contents of base) + ((contents of index) * (scale))
- Consider the sum to be an address
- Load from (or store to) that address

Note:

- All other forms are subsets of the full form...

45

46

Operand Examples

Immediate operands

- `$5` => use the number 5 (i.e. the number that is available immediately within the instruction)
- `$i` => use the address denoted by `i` (i.e. the address that is available immediately within the instruction)

Register operands

- `%rax` => read from (or write to) register RAX

Memory operands: direct addressing

- `5` => load from (or store to) memory at address 5 (silly; seg fault)
- `i` => load from (or store to) memory at the address denoted by `i`

Memory operands: indirect addressing

- `(%rax)` => consider the contents of RAX to be an address; load from (or store to) that address

47

48

Operand Examples

Memory operands: scaled indexed addressing

- `5(%rax,%r10,4)` => compute the sum (5) + (contents of RAX) + ((contents of R10) * 4); consider the sum to be an address; load from (or store to) that address
- `i(%rax,%r10,4)` => compute the sum (address denoted by i) + (contents of RAX) + ((contents of R10) * 4); consider the sum to be an address; load from (or store to) that address

49

Aside: The lea Instruction

lea: load effective address

- Unique instruction: suppresses memory load/store

Example

- `movq 5(%rax), %r10`
 - Compute the sum (5) + (contents of RAX); consider the sum to be an address; load 8 bytes from that address into R10
- `leaq 5(%rax), %r10`
 - Compute the sum (5) + (contents of RAX); move that sum to R10

Useful for

- Computing an address, e.g. as a function argument
- See precept code that calls `scanf()`
- Some quick-and-dirty arithmetic

What is the effect of this?
`leaq (%rax,%rax,4),%rax`

50

Agenda

- Flattened C
- Control flow with signed integers
- Control flow with unsigned integers
- Arrays
- Structures

51

Structures: Indirect Addressing

C <pre>struct S { int i; int j; }; struct S myStruct; myStruct.i = 18; myStruct.j = 19;</pre>	Assem Lang <pre>.section ".bss" myStruct: .skip 8 .myStruct: .skip 8 .section ".text" movq \$myStruct, %rax movl \$18, (%rax) ... movq \$myStruct, %rax addq \$4, %rax movl \$19, (%rax)</pre>
---	--

Note:
Indirect addressing

52

Structures: Base+Disp Addressing

C <pre>struct S { int i; int j; }; struct S myStruct; myStruct.i = 18; myStruct.j = 19;</pre>	Assem Lang <pre>.section ".bss" myStruct: .skip 8myStruct: .skip 8 .section ".text" movl \$0, %eax movl \$18, myStruct(%eax) ... movl \$4, %eax movl \$19, myStruct(%eax)</pre>
---	---

Note:
Base+displacement addressing

53

Structures: Padding

C <pre>struct S { char c; int i; }; struct S myStruct; myStruct.c = 'A'; myStruct.i = 18;</pre>	Assem Lang <pre>.section ".bss" myStruct: .skip 8myStruct: .skip 8 .section ".text" movl \$0, %eax movb \$'A', myStruct(%eax) ... movl \$4, %eax movl \$18, myStruct(%eax)</pre>
---	--

Beware:
Compiler sometimes inserts padding after fields

54

Structures: Padding

x86-64/Linux rules

Data type	Within a struct, must begin at address that is evenly divisible by:
(unsigned) char	1
(unsigned) short	2
(unsigned) int	4
(unsigned) long	8
float	4
double	8
long double	16
any pointer	8

Compiler may add padding after last field if struct is within an array

55

Summary

Intermediate aspects of x86-64 assembly language...

Flattened C code

Control transfer with signed integers

Control transfer with unsigned integers

Arrays

- Full form of instruction operands

Structures

- Padding

56

Appendix

Setting and using CC bits in EFLAGS register

57

Setting Condition Code Bits

Question

- How does `cmpl` set condition code bits in EFLAGS register?

Answer

- (See following slides)

58

Condition Code Bits

Condition code bits

- ZF:** zero flag: set to 1 iff result is **zero**
- SF:** sign flag: set to 1 iff result is **negative**
- CF:** carry flag: set to 1 iff **unsigned overflow** occurred
- OF:** overflow flag: set to 1 iff **signed overflow** occurred

59

Condition Code Bits

Example: `addl src, dest`

- Compute sum (`dest+src`)
- Assign sum to `dest`
- ZF: set to 1 iff sum == 0
- SF: set to 1 iff sum < 0
- CF: set to 1 iff unsigned overflow
 - Set to 1 iff sum<`src`
- OF: set if signed overflow
 - Set to 1 iff
 $(src>0 \&\& dest>0 \&\& sum<0) || (src<0 \&\& dest<0 \&\& sum>0)$

60

Condition Code Bits

Example: `subl src, dest`

- Compute sum (`dest+(-src)`)
- Assign sum to `dest`
- ZF: set to 1 iff sum == 0
- SF: set to 1 iff sum < 0
- CF: set to 1 iff unsigned overflow
 - Set to 1 iff `dest<src`
- OF: set to 1 iff signed overflow
 - Set to 1 iff
 $(dest>0 \&& src<0 \&& sum<0) ||$
 $(dest<0 \&& src>0 \&& sum>=0)$

Example: `cmpl src, dest`

- Same as `subl`
- But does not affect `dest`



61

Using Condition Code Bits



Question

- How do conditional jump instructions use condition code bits in EFLAGS register?

Answer

- (See following slides)

62

Conditional Jumps: Unsigned



After comparing unsigned data

Jump Instruction	Use of CC Bits
<code>je label</code>	ZF
<code>jne label</code>	$\sim ZF$
<code>jb label</code>	CF
<code>jae label</code>	$\sim CF$
<code>jbe label</code>	CF ZF
<code>ja label</code>	$\sim(CF \mid ZF)$

Note:

- If you can understand why `jb` jumps iff CF
- ... then the others follow

63

Conditional Jumps: Unsigned



Why does `jb` jump iff CF? Informal explanation:

(1) largenum – smallnum (not below)

- Correct result
- => CF=0 => don't jump

(2) smallnum – largenum (below)

- Incorrect result
- => CF=1 => jump

64

Conditional Jumps: Signed



After comparing signed data

Jump Instruction	Use of CC Bits
<code>je label</code>	ZF
<code>jne label</code>	$\sim ZF$
<code>jl label</code>	OF ^ SF
<code>jge label</code>	$\sim(OF \wedge SF)$
<code>jle label</code>	$(OF \wedge SF) \mid ZF$
<code>jg label</code>	$\sim((OF \wedge SF) \mid ZF)$

Note:

- If you can understand why `jl` jumps iff OF^SF
- ... then the others follow

65

Conditional Jumps: Signed



Why does `jl` jump iff OF^SF? Informal explanation:

(1) largeposnum – smallposnum (not less than)

- Certainly correct result
- => OF=0, SF=0, OF^SF==0 => don't jump

(2) smallposnum – largeposnum (less than)

- Certainly correct result
- => OF=0, SF=1, OF^SF==1 => jump

(3) largenegnum – smallnegnum (less than)

- Certainly correct result
- => OF=0, SF=1 => (OF^SF)==1 => jump

(4) smallnegnum – largenegnum (not less than)

- Certainly correct result
- => OF=0, SF=0 => (OF^SF)==0 => don't jump

66

Conditional Jumps: Signed



(5) posnum – negnum (not less than)

- Suppose correct result
- => OF=0, SF=0 => $(OF \wedge SF) == 0$ => don't jump

(6) posnum – negnum (not less than)

- Suppose incorrect result
- => OF=1, SF=1 => $(OF \wedge SF) == 0$ => don't jump

(7) negnum – posnum (less than)

- Suppose correct result
- => OF=0, SF=1 => $(OF \wedge SF) == 1$ => jump

(8) negnum – posnum (less than)

- Suppose incorrect result
- => OF=1, SF=0 => $(OF \wedge SF) == 1$ => jump

67