even more C / bitwise (start)

quiz demo

last time

x86 condition codes

- set by most arithmetic instructions
- ZF (zero), SF (sign), OF (overflow), CF (carry)

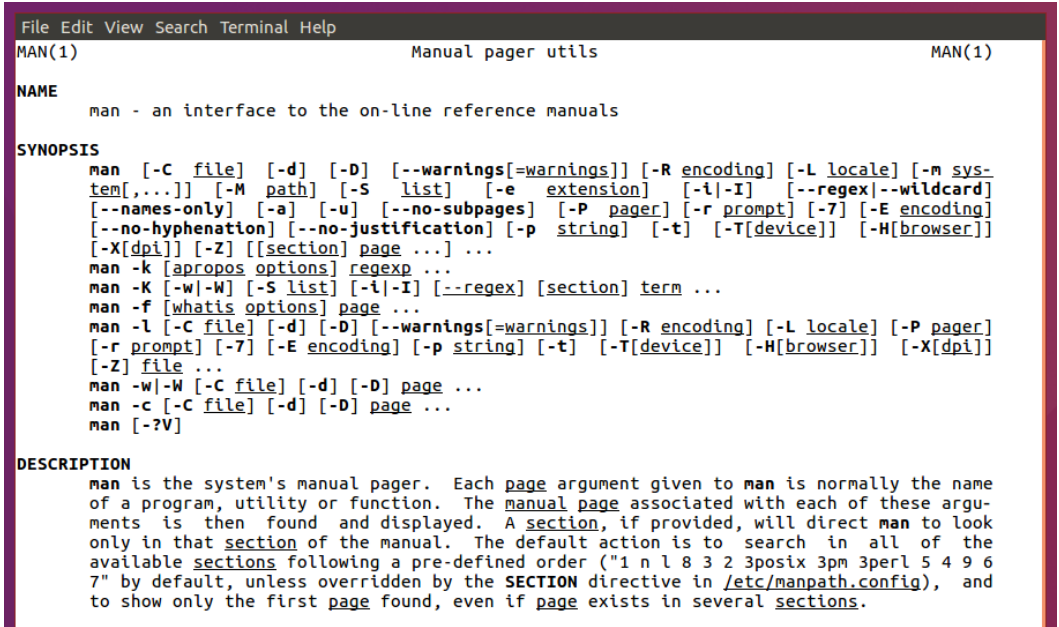
C types, booleans, pointers, pointer arithmetic

command line tips (start)

interlude: command line tips

```
cr4bd@reiss-lenovo:~$ man man
```

man man



File Edit View Search Terminal Help

MAN(1) Manual pager utils MAN(1)

NAME

man - an interface to the on-line reference manuals

SYNOPSIS

man [**-C** file] [**-d**] [**-D**] [**--warnings**[=warnings]] [**-R** encoding] [**-L** locale] [**-m** system[,...]] [**-M** path] [**-S** list] [**-e** extension] [**-i**|**-I**] [**--regex**|**--wildcard**] [**--names-only**] [**-a**] [**-u**] [**--no-subpages**] [**-P** pager] [**-r** prompt] [**-7**] [**-E** encoding] [**--no-hyphenation**] [**--no-justification**] [**-p** string] [**-t**] [**-T**[device]] [**-H**[browser]] [**-X**[dpi]] [**-Z**] [[section] page ...] ...

man -k [apropos options] regexp ...

man -K [**-w**|**-W**] [**-S** list] [**-i**|**-I**] [--regex] [section] term ...

man -f [whatis options] page ...

man -l [**-C** file] [**-d**] [**-D**] [**--warnings**[=warnings]] [**-R** encoding] [**-L** locale] [**-P** pager] [**-r** prompt] [**-7**] [**-E** encoding] [**-p** string] [**-t**] [**-T**[device]] [**-H**[browser]] [**-X**[dpi]] [**-Z**] file ...

man -w|**-W** [**-C** file] [**-d**] [**-D**] page ...

man -c [**-C** file] [**-d**] [**-D**] page ...

man [**-?V**]

DESCRIPTION

man is the system's manual pager. Each page argument given to **man** is normally the name of a program, utility or function. The manual page associated with each of these arguments is then found and displayed. A section, if provided, will direct **man** to look only in that section of the manual. The default action is to search in all of the available sections following a pre-defined order ("1 n l 8 3 2 3posix 3pm 3perl 5 4 9 6 7" by default, unless overridden by the **SECTION** directive in /etc/manpath.config), and to show only the first page found, even if page exists in several sections.

man man

File Edit View Search Terminal Help

EXAMPLES

man ls
Display the manual page for the item (program) ls.

man -a intro
Display, in succession, all of the available intro manual pages contained within the manual. It is possible to quit between successive displays or skip any of them.

man -t alias | lpr -Pps
Format the manual page referenced by `alias', usually a shell manual page, into the default **troff** or **groff** format and pipe it to the printer named ps. The default output for **groff** is usually PostScript. **man --help** should advise as to which processor is bound to the **-t** option.

man -l -Tdvi ./foo.1x.gz **>** ./foo.1x.dvi
This command will decompress and format the nroff source manual page ./foo.1x.gz into a **device independent (dvi)** file. The redirection is necessary as the **-T** flag causes output to be directed to **stdout** with no pager. The output could be viewed with a program such as **xdvi** or further processed into PostScript using a program such as **dvips.**

man -k printf
Search the short descriptions and manual page names for the keyword printf as regular expression. Print out any matches. Equivalent to **apropos** printf.

man -f smail
Lookup the manual pages referenced by smail and print out the short descriptions of any found. Equivalent to **whatis** smail.

man chmod

File Edit View Search Terminal Help

CHMOD(1) User Commands CHMOD(1)

NAME

chmod - change file mode bits

SYNOPSIS

chmod [OPTION]... MODE[,MODE]... FILE...
chmod [OPTION]... OCTAL-MODE FILE...
chmod [OPTION]... --reference=RFILE FILE...

DESCRIPTION

This manual page documents the GNU version of **chmod**. **chmod** changes the file mode bits of each given file according to mode, which can be either a symbolic representation of changes to make, or an octal number representing the bit pattern for the new mode bits.

The format of a symbolic mode is [**ugoa**...][[**-+=**][perms...]...], where perms is either zero or more letters from the set **rwxXst**, or a single letter from the set **ugo**. Multiple symbolic modes can be given, separated by commas.

A combination of the letters **ugoa** controls which users' access to the file will be changed: the user who owns it (**u**), other users in the file's group (**g**), other users not in the file's group (**o**), or all users (**a**). If none of these are given, the effect is as if (**a**) were given, but bits that are set in the umask are not affected.

The operator **+** causes the selected file mode bits to be added to the existing file mode bits of each file; **-** causes them to be removed; and **=** causes them to be added and causes unmentioned bits to be removed except that a directory's unmentioned set user and group ID bits are not affected.

The letters **rwxXst** select file mode bits for the affected users: read (**r**), write (**w**),

chmod

```
chmod --recursive og-r /home/USER
```

chmod

```
chmod --recursive og-r /home/USER
```

others and group (student)
− remove
read

chmod

```
chmod --recursive og-r /home/USER
```

user (yourself) / group / others
− remove / + add
read / write / execute or search

tar

the standard Linux/Unix file archive utility

Table of contents: `tar tf filename.tar`

eXtract: `tar xvf filename.tar`

Create: `tar cvf filename.tar directory`

(`v`: verbose; `f`: file — default is tape)

Tab completion and history

struct

```c
struct rational {
    int numerator;
    int denominator;
};
// ...
struct rational two_and_a_half;
two_and_a_half.numerator = 5;
two_and_a_half.denominator = 2;
struct rational *pointer = &two_and_a_half;
printf("%d/%d\n",
       pointer->numerator,
       pointer->denominator);
```

struct

```c
struct rational {
    int numerator;
    int denominator;
};
// ...
struct rational two_and_a_half;
two_and_a_half.numerator = 5;
two_and_a_half.denominator = 2;
struct rational *pointer = &two_and_a_half;
printf("%d/%d\n",
       pointer->numerator,
       pointer->denominator);
```

typedef

instead of writing:

```
...
unsigned int a;
unsigned int b;
unsigned int c;
```

can write:

```
typedef unsigned int uint;
...
uint a;
uint b;
uint c;
```

typedef struct (1)

```c
struct other_name_for_rational {
    int numerator;
    int denominator;
};
typedef struct other_name_for_rational rational;
// ...
rational two_and_a_half;
two_and_a_half.numerator = 5;
two_and_a_half.denominator = 2;
rational *pointer = &two_and_a_half;
printf("%d/%d\n",
       pointer->numerator,
       pointer->denominator);
```

typedef struct (1)

```c
struct other_name_for_rational {
    int numerator;
    int denominator;
};
typedef struct other_name_for_rational rational;
// ...
rational two_and_a_half;
two_and_a_half.numerator = 5;
two_and_a_half.denominator = 2;
rational *pointer = &two_and_a_half;
printf("%d/%d\n",
       pointer->numerator,
       pointer->denominator);
```

typedef struct (2)

```c
struct other_name_for_rational {
    int numerator;
    int denominator;
};
typedef struct other_name_for_rational rational;
// same as:
typedef struct other_name_for_rational {
    int numerator;
    int denominator;
} rational;
```

typedef struct (2)

```
struct other_name_for_rational {
    int numerator;
    int denominator;
};
typedef struct other_name_for_rational rational;
// same as:
typedef struct other_name_for_rational {
    int numerator;
    int denominator;
} rational;
```

typedef struct (2)

```c
struct other_name_for_rational {
    int numerator;
    int denominator;
};
typedef struct other_name_for_rational rational;
// same as:
typedef struct other_name_for_rational {
    int numerator;
    int denominator;
} rational;
// almost the same as:
typedef struct {
    int numerator;
    int denominator;
} rational;
```

typedef struct (3)

```
struct other_name_for_rational {
    int numerator;
    int denominator;
};
typedef struct other_name_for_rational rational;
```

valid ways to declare an instance:

```
struct other_name_for_rational some_variable;
rational some_variable;
```

INVALID ways:

```
/* INVALID: */ struct rational some_variable;
/* INVALID: */ other_name_for_rational some_variable;
```

typedef struct (3)

```
struct other_name_for_rational {
    int numerator;
    int denominator;
};
typedef struct other_name_for_rational rational;
```

valid ways to declare an instance:

```
struct other_name_for_rational some_variable;
rational some_variable;
```

INVALID ways:

```
/* INVALID: */ struct rational some_variable;
/* INVALID: */ other_name_for_rational some_variable;
```

typedef struct (3)

```
struct other_name_for_rational {
    int numerator;
    int denominator;
};
typedef struct other_name_for_rational rational;
```

valid ways to declare an instance:

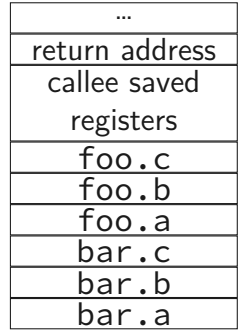
```
struct other_name_for_rational some_variable;
rational some_variable;
```

INVALID ways:

```
/* INVALID: */ struct rational some_variable;
/* INVALID: */ other_name_for_rational some_variable;
```

structs aren't references

```
typedef struct {
    long a; long b; long c;
} triple;
...

triple foo;
foo.a = foo.b = foo.c = 3;
triple bar = foo;
bar.a = 4;
// foo is 3, 3, 3
// bar is 4, 3, 3
```



linked lists / dynamic allocation

```
typedef struct list_t {
    int item;
    struct list_t *next;
} list;
// ...
```

linked lists / dynamic allocation

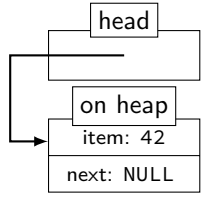
```
typedef struct list_t {
    int item;
    struct list_t *next;
} list;
// ...
```

linked lists / dynamic allocation

```
typedef struct list_t {
    int item;
    struct list_t *next;
} list;
// ...


  list* head = malloc(sizeof(list));
    /* C++: new list; */
  head->item = 42;
  head->next = NULL;
  // ...
  free(head);
    /* C++: delete list */
```

linked lists / dynamic allocation

```c
typedef struct list_t {
    int item;
    struct list_t *next;
} list;
// ...

  list* head = malloc(sizeof(list));
    /* C++: new list; */
  head->item = 42;
  head->next = NULL;
  // ...
  free(head);
    /* C++: delete list */
```

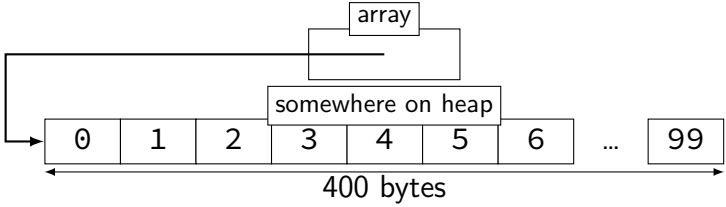


dynamic arrays

```
int *array = malloc(sizeof(int)*100);
  // C++: new int[100]
for (i = 0; i < 100; ++i) {
    array[i] = i;
}
// ...
free(array); // C++: delete[] array
```

dynamic arrays

```
int *array = malloc(sizeof(int)*100);
  // C++: new int[100]
for (i = 0; i < 100; ++i) {
    array[i] = i;
}
// ...
free(array); // C++: delete[] array
```



unsigned and signed types

type	min	max
`signed int` = `signed` = `int`	-2^{31}	$2^{31}-1$
`unsigned int` = `unsigned`	0	$2^{32}-1$
`signed long` = `long`	-2^{63}	$2^{63}-1$
`unsigned long`	0	$2^{64}-1$

unsigned/signed comparison trap (1)

```
int x = -1;
unsigned int y = 0;
printf("%d\n", x < y);
```

unsigned/signed comparison trap (1)

```
int x = -1;
unsigned int y = 0;
printf("%d\n", x < y);
```

result is 0

unsigned/signed comparison trap (1)

```
int x = -1;
unsigned int y = 0;
printf("%d\n", x < y);
```

result is `0`

short solution: don't compare signed to unsigned:

```
(long) x < (long) y
```

unsigned/sign comparison trap (2)

```
int x = -1;
unsigned int y = 0;
printf("%d\n", x < y);
```

compiler converts both to same type first

- `int` if all possible values fit
- otherwise: first operand (`x`, `y`) type from this list:
  - `unsigned long`
  - `long`
  - `unsigned int`
  - `int`

C evolution and standards

1978: Kernighan and Ritchie publish *The C Programming Language* — "K&R C"

very different from modern C

C evolution and standards

1978: Kernighan and Ritchie publish *The C Programming Language* — "K&R C"

- very different from modern C

1989: ANSI standardizes C — C89/C90/`-ansi`

- compiler option: `-ansi`, `-std=c90`
- looks mostly like modern C

C evolution and standards

1978: Kernighan and Ritchie publish *The C Programming Language* — "K&R C"

- very different from modern C

1989: ANSI standardizes C — C89/C90/`-ansi`

- compiler option: `-ansi`, `-std=c90`
- looks mostly like modern C

1999: ISO (and ANSI) update C standard — C99

- compiler option: `-std=c99`
- adds: declare variables in middle of block
- adds: `//` comments

C evolution and standards

1978: Kernighan and Ritchie publish *The C Programming Language* — "K&R C"

- very different from modern C

1989: ANSI standardizes C — C89/C90/`-ansi`

- compiler option: `-ansi`, `-std=c90`
- looks mostly like modern C

1999: ISO (and ANSI) update C standard — C99

- compiler option: `-std=c99`
- adds: declare variables in middle of block
- adds: `//` comments

2011: Second ISO update — C11

undefined behavior example (1)

```c
#include <stdio.h>
#include <limits.h>
int test(int number) {
    return (number + 1) > number;
}

int main(void) {
    printf("%d\n", test(INT_MAX));
}
```

undefined behavior example (1)

```c
#include <stdio.h>
#include <limits.h>
int test(int number) {
    return (number + 1) > number;
}

int main(void) {
    printf("%d\n", test(INT_MAX));
}
```

without optimizations: `0`

undefined behavior example (1)

```c
#include <stdio.h>
#include <limits.h>
int test(int number) {
    return (number + 1) > number;
}

int main(void) {
    printf("%d\n", test(INT_MAX));
}
```

without optimizations: `0`

with optimizations: `1`

undefined behavior example (2)

```
int test(int number) {
    return (number + 1) > number;
}
```

Optimized:

```
test:
    movl    $1, %eax        # eax ← 1
    ret
```

Less optimized:

```
test:
    leal    1(%rdi), %eax # eax ← rdi + 1
    cmpl    %eax, %edi
    setl    %al           # al ← eax < edi
    movzbl  %al, %eax     # eax ← al (pad with zeros)
    ret
```

undefined behavior

compilers can do whatever they want

- what you expect
- crash your program
- ...

common types:

- *signed* integer overflow/underflow
- out-of-bounds pointers
- integer divide-by-zero
- writing read-only data
- out-of-bounds shift

undefined behavior

why undefined behavior?

different architectures work differently

- allow compilers to expose whatever processor does "naturally"
- don't encode any particular machine in the standard

flexibility for optimizations

extracting hexadecimal nibble (1)

problem: given `0xAB` extract `0xA`

(hexadecimal digits called “nibbles”)

```
typedef unsigned char byte;
int get_top_nibble(byte value) {
    return ???;
}
```

extracing hexadecimal nibbles (2)

```
typedef unsigned char byte;
int get_top_nibble(byte value) {
    return value / 16;
}
```

aside: division

division is really slow

Intel "Skylake" microarchitecture:

- about six cycles per division
- ...and much worse for eight-byte division
- versus: four additions per cycle

aside: division

division is really slow

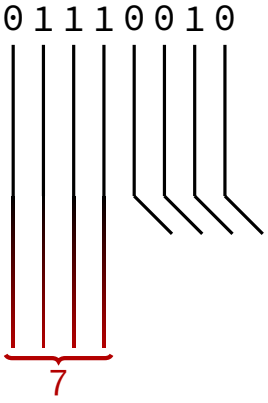
Intel "Skylake" microarchitecture:

- about six cycles per division
- ...and much worse for eight-byte division
- versus: four additions per cycle

but this case: it's just extracting 'top wires' — simpler?

extracting bits in hardware

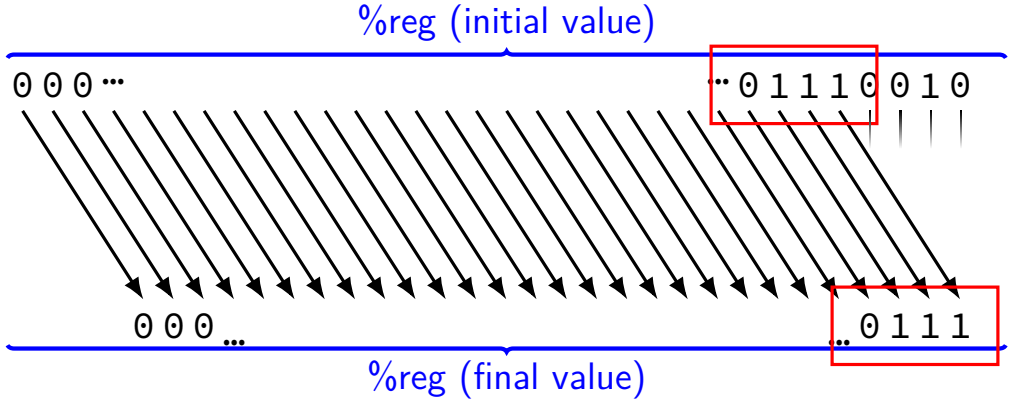
0111 0010 = 0x72



exposing wire selection

x86 instruction: `shr` — shift right

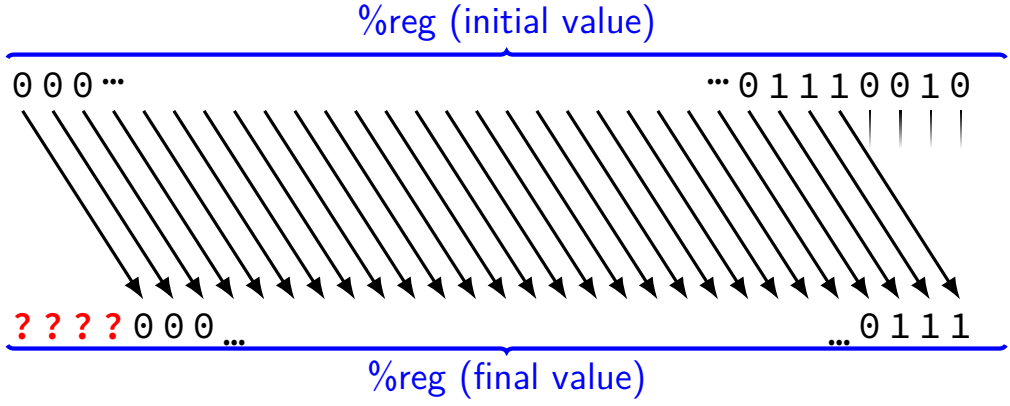
`shr $amount, %reg` (or variable: `shr %cl, %reg`)



exposing wire selection

x86 instruction: `shr` — shift right

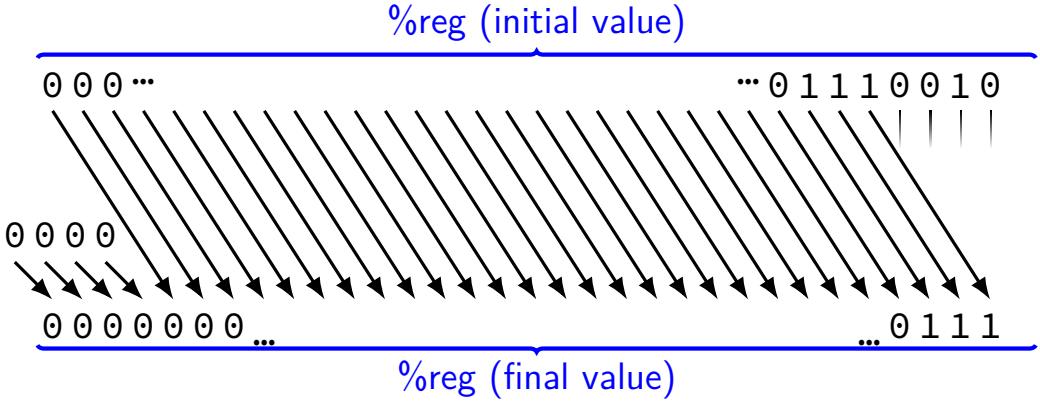
`shr $amount, %reg` (or variable: `shr %cl, %reg`)



exposing wire selection

x86 instruction: `shr` — shift right

`shr $amount, %reg` (or variable: `shr %cl, %reg`)



shift right

x86 instruction: **shr** — shift right

`shr $amount, %reg`

(or variable: `shr %cl, %reg`)

```
get_top_nibble:
 // eax ← dil (low byte of rdi) w/ zero padding
  movzbl %dil, %eax
  shrl $4, %eax
    ret
```

shift right

x86 instruction: **shr** — shift right

`shr $amount, %reg`

(or variable: `shr %cl, %reg`)

```
get_top_nibble:
 // eax ← dil (low byte of rdi) w/ zero padding
  movzbl %dil, %eax
  shrl $4, %eax
    ret
```

shift right

x86 instruction: **shr** — shift right

`shr $amount, %reg`

(or variable: `shr %cl, %reg`)

```
get_top_nibble:
 // eax ← dil (low byte of rdi) w/ zero padding
  movzbl %dil, %eax
  shrl $4, %eax
    ret
```

right shift in C

```
get_top_nibble:
  // eax ← dil (low byte of rdi) w/ zero padding
  movzbl %dil, %eax
  shrl $4, %eax
  ret
```

```
typedef unsigned char byte;
int get_top_nibble(byte value) {
    return value >> 4;
}
```

right shift in C

```
typedef unsigned char byte;
int get_top_nibble1(byte value) { return value >> 4; }
int get_top_nibble2(byte value) { return value / 16; }
```

right shift in C

```
typedef unsigned char byte;
int get_top_nibble1(byte value) { return value >> 4; }
int get_top_nibble2(byte value) { return value / 16; }
```

example output from optimizing compiler:

```
get_top_nibble1:
   shrb $4, %dil
   movzbl %dil, %eax
   ret

get_top_nibble2:
   shrb $4, %dil
   movzbl %dil, %eax
   ret
```

right shift in math

```
1 >> 0 == 1          0000 0001
1 >> 1 == 0          0000 0000
1 >> 2 == 0          0000 0000

10 >> 0 == 10        0000 1010
10 >> 1 == 5         0000 0101
10 >> 2 == 2         0000 0010
```

$$x \texttt{ >> } y = \lfloor x \times 2^{-y} \rfloor$$

exercise

```
int foo(int)
foo:
        movl %edi, %eax
        shrl $1, %eax
        ret
```

what is the value of `foo(-2)`?

A. -4 B. -2 C. -1 D. 0

E. a small positive number F. a large positive number

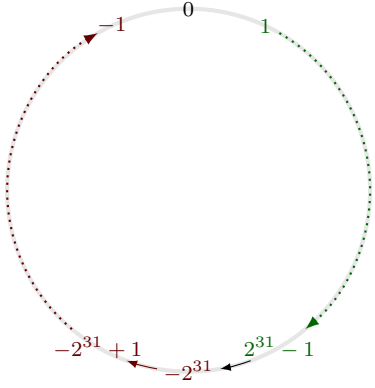
G. a large negative number H. something else

two's complement refresher

$$-1 = \begin{array}{cccccccc} -2^{31} & +2^{30} & +2^{29} & & +2^{2} & +2^{1} & +2^{0} \\ 1 & 1 & 1 & \ldots & 1 & 1 & 1 \end{array}$$

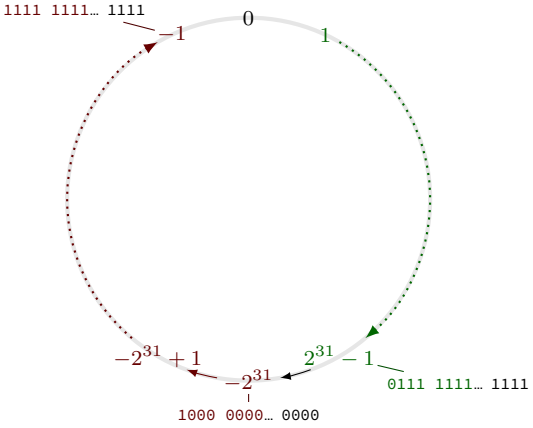
two's complement refresher

$$-1 = \begin{array}{cccccccc} -2^{31} & +2^{30} & +2^{29} & & +2^{2} & +2^{1} & +2^{0} \\ 1 & 1 & 1 & \ldots & 1 & 1 & 1 \end{array}$$



two's complement refresher

$$-1 = \begin{array}{cccccccc} & -2^{31} & +2^{30} & +2^{29} & & +2^{2} & +2^{1} & +2^{0} \\ & 1 & 1 & 1 & \ldots & 1 & 1 & 1 \end{array}$$



dividing negative by two

start with $-x$

flip all bits and add one to get x

right shift by one to get $x/2$

flip all bits and add one to get $-x/2$

dividing negative by two

start with $-x$

flip all bits and add one to get x

right shift by one to get $x/2$

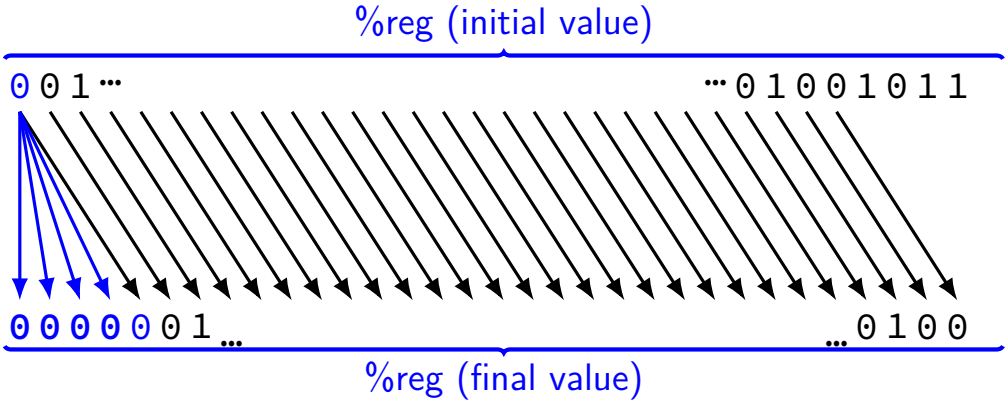
flip all bits and add one to get $-x/2$

same as right shift by one, adding `1`s instead of `0`s
(except for rounding)

arithmetic right shift

x86 instruction: `sar` — arithmetic shift right

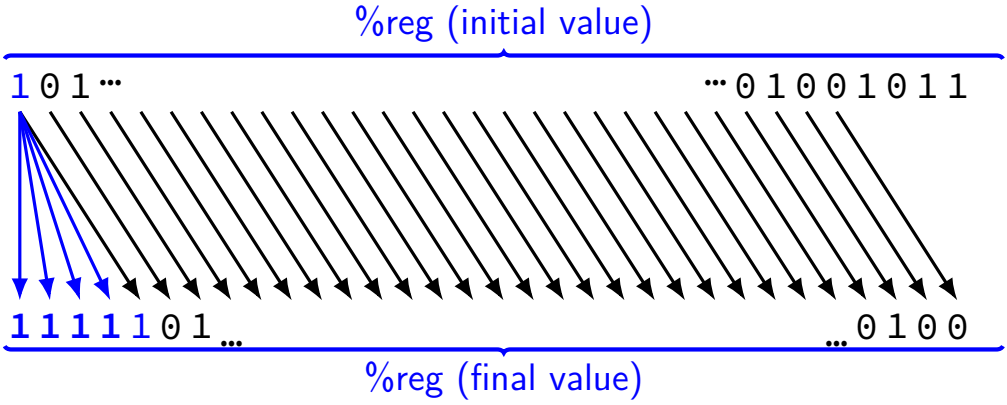
`sar $amount, %reg` (or variable: `sar %cl, %reg`)



arithmetic right shift

x86 instruction: `sar` — arithmetic shift right

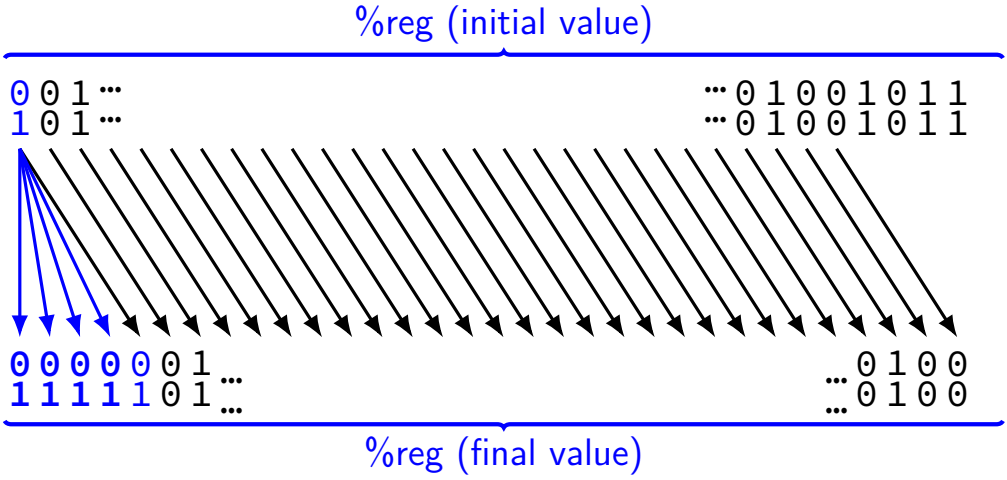
`sar $amount, %reg` (or variable: `sar %cl, %reg`)



arithmetic right shift

x86 instruction: `sar` — arithmetic shift right

`sar $amount, %reg` (or variable: `sar %cl, %reg`)



right shift in C

```
int shift_signed(int x) {
    return x >> 5;
}
unsigned shift_unsigned(unsigned x) {
    return x >> 5;
}
```

```
shift_signed:
    movl %edi, %eax
    sarl $5, %eax
    ret
```

```
shift_unsigned:
    movl %edi, %eax
    shrl $5, eax
    ret
```

standards and shifts in C

signed right shift is implementation-defined

- standard lets compilers choose which type of shift to do
- all x86 compilers I know of — arithmetic

shift amount $\geq$ width of type: undefined

- x86 assembly: only uses lower bits of shift amount

exercise

```
int shiftTwo(int x) {
    return x >> 2;
}
```

`shiftTwo(-6) = ???`

A. -4 B. -3 C. -2 D. -1 E. 0

E. some positive number F. something else

dividing negative by two

start with $-x$

flip all bits and add one to get x

right shift by one to get $x/2$

flip all bits and add one to get $-x/2$

same as right shift by one, adding `1`s instead of `0`s
(except for rounding)

divide with proper rounding

C division: rounds towards zero (truncate)

arithmetic shift: rounds towards negative infinity

solution: “bias” adjustments — described in textbook

divide with proper rounding

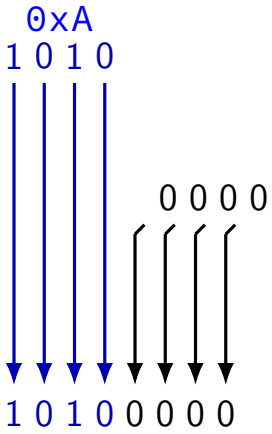
C division: rounds towards zero (truncate)

arithmetic shift: rounds towards negative infinity

solution: "bias" adjustments — described in textbook

```
divideBy8: // GCC generated code
    leal   7(%rdi), %eax // eax ← edi + 7
    testl  %edi, %edi    // set cond. codes based on %edi
    cmovns %edi, %eax    // if (SF = 0) eax ← edi
    sarl   $3, %eax      // arithmetic shift
```

multiplying by 16



$\texttt{0xA} \times 16 = \texttt{0xA0}$

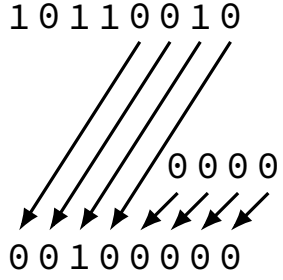
shift left

~~`shr $-4, %reg`~~

instead: `shl $4, %reg` ("**sh**ift **l**eft")

~~`value >> (-4)`~~

instead: `value << 4`



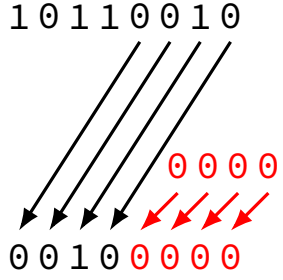
shift left

~~`shr $-4, %reg`~~

instead: `shl $4, %reg` (“**sh**ift **l**eft”)

~~`value >> (-4)`~~

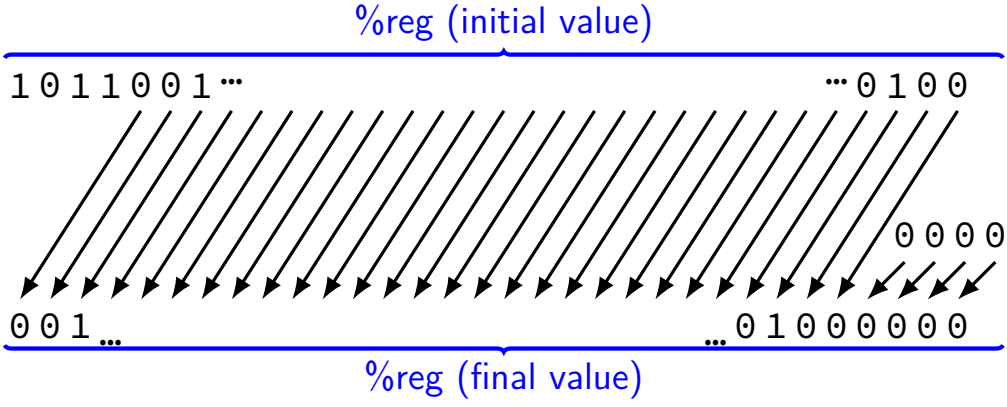
instead: `value << 4`



shift left

x86 instruction: **`shl`** — shift left

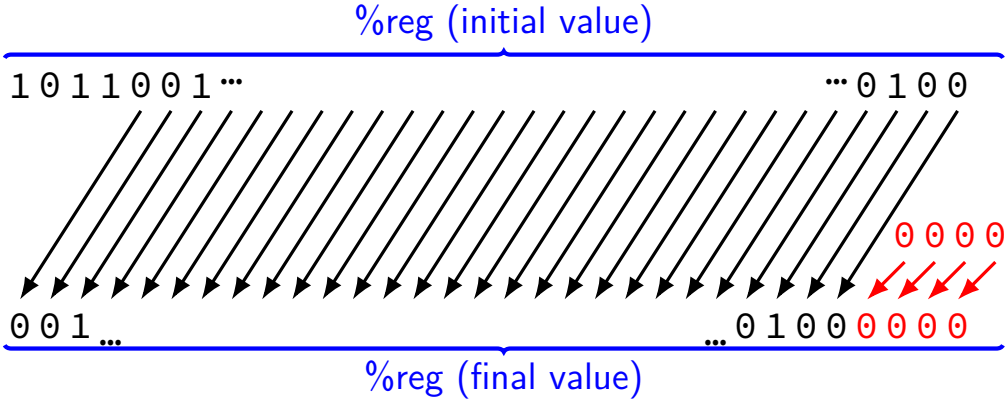
`shl` `$`*`amount`*`, %reg` (or variable: **`shl`** `%cl, %reg`)



shift left

x86 instruction: **`shl`** — shift left

`shl` `$`*`amount`*`, %reg` (or variable: **`shl`** `%cl, %reg`)



left shift in math

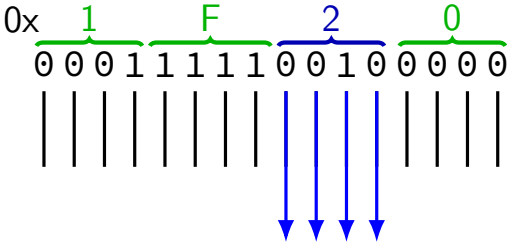
```
1 << 0 == 1                 0000 0001
1 << 1 == 2                 0000 0010
1 << 2 == 4                 0000 0100

10 << 0 == 10               0000 1010
10 << 1 == 20               0001 0100
10 << 2 == 40               0010 1000

-10 << 0 == -10             1111 0110
-10 << 1 == -20             1110 1100
-10 << 2 == -40             1101 1000
```

left shift in math

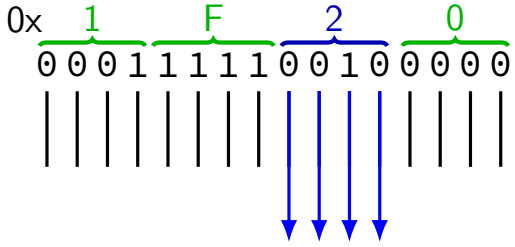
`1 << 0 == 1`	`0000 0001`
`1 << 1 == 2`	`0000 0010`
`1 << 2 == 4`	`0000 0100`
`10 << 0 == 10`	`0000 1010`
`10 << 1 == 20`	`0001 0100`
`10 << 2 == 40`	`0010 1000`
`−10 << 0 == −10`	`1111 0110`
`−10 << 1 == −20`	`1110 1100`
`−10 << 2 == −40`	`1101 1000`

$$x \texttt{ << } y = x \times 2^y$$

extracting nibble from more



extracting nibble from more



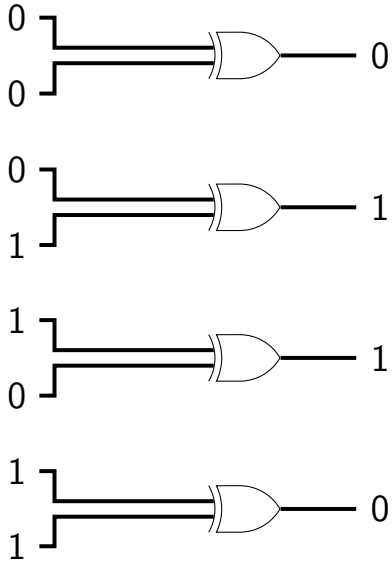

```
// % -- remainder
unsigned extract_second_nibble(unsigned value) {
    return (value / 16) % 16;
}

unsigned extract_second_nibble(unsigned value) {
    return (value % 256) / 16;
}
```

manipulating bits?

easy to manipulate individual bits in HW

how do we expose that to software?

circuits: gates



interlude: a truth table

AND	**0**	**1**
0	0	0
1	0	1

interlude: a truth table

AND	**0**	**1**
0	0	0
1	0	1

AND with 1: keep a bit the same

interlude: a truth table

AND	**0**	**1**
0	0	0
1	0	1

AND with 1: keep a bit the same

AND with 0: clear a bit

interlude: a truth table

AND	**0**	**1**
0	0	0
1	0	1

AND with 1: keep a bit the same

AND with 0: clear a bit

method: construct "mask" of what to keep/remove

bitwise AND — &

Treat value as array of bits

```
1 & 1 == 1

1 & 0 == 0

0 & 0 == 0

2 & 4 == 0

10 & 7 == 2
```

bitwise AND — &

Treat value as array of bits

```
1 & 1 == 1

1 & 0 == 0

0 & 0 == 0

2 & 4 == 0

10 & 7 == 2
```

```
    … 0 0 1 0
&   … 0 1 0 0
  ------------
    … 0 0 0 0
```

bitwise AND — &

Treat value as array of bits

```
1 & 1 == 1

1 & 0 == 0

0 & 0 == 0

2 & 4 == 0

10 & 7 == 2
```

```
      … 0 0 1 0
&     … 0 1 0 0
-------------
      … 0 0 0 0
```

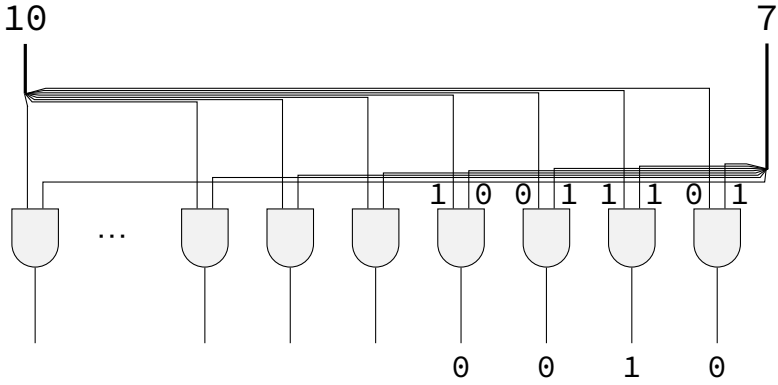
```
      … 1 0 1 0
&     … 0 1 1 1
-------------
      … 0 0 1 0
```

bitwise AND — C/assembly

x86: `and %reg, %reg`

C: `foo & bar`

bitwise hardware (`10 & 7 == 2`)



extract 0x3 from 0x1234

```
unsigned get_second_nibble1_bitwise(unsigned value)
    return (value >> 4) & 0xF; // 0xF: 00001111
    // like (value / 16) % 16
}

unsigned get_second_nibble2_bitwise(unsigned value)
    return (value & 0xF0) >> 4; // 0xF0: 11110000
    // like (value % 256) / 16;
}
```

extract 0x3 from 0x1234

```
get_second_nibble1_bitwise:
    movl %edi, %eax
    shrl $4, %eax
    andl $0xF, %eax
    ret

get_second_nibble2_bitwise:
    movl %edi, %eax
    andl $0xF0, %eax
    shrl $4, %eax
    ret
```

and/or/xor

AND	**0**	**1**
0	0	0
1	0	1

&

conditionally clear bit
conditionally keep bit

OR	**0**	**1**
0	0	1
1	1	1

|

conditionally set bit

XOR	**0**	**1**
0	0	1
1	1	0

^

conditionally flip bit

bitwise OR — |

```
1 | 1 == 1

1 | 0 == 1

0 | 0 == 0

2 | 4 == 6

10 | 7 == 15
```

```
    … 1 0 1 0
|   … 0 1 1 1
-------------
    … 1 1 1 1
```

bitwise xor — ^

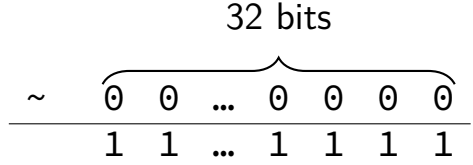
```
1 ^ 1 == 0

1 ^ 0 == 1

0 ^ 0 == 0

2 ^ 4 == 6

10 ^ 7 == 13
```

```
    … 1 0 1 0
^   … 0 1 1 1
-------------
    … 1 1 0 1
```

negation / not — ~

~ ('complement') is bitwise version of `!`:

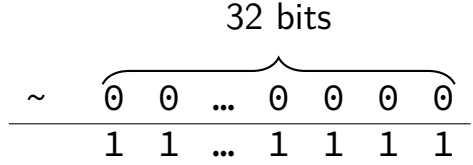
```
!0 == 1

!notZero == 0

~0 == (int) 0xFFFFFFFF (aka −1)
```



negation / not — ~

~ ('complement') is bitwise version of !:

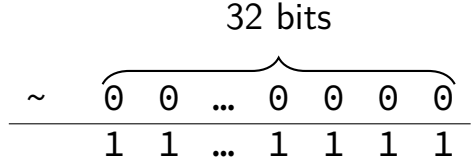
```
!0 == 1

!notZero == 0

~0 == (int) 0xFFFFFFFF (aka −1)

~2 == (int) 0xFFFFFFFD (aka −3)
```



negation / not — ~

~ ('complement') is bitwise version of !:

```
!0 == 1

!notZero == 0

~0 == (int) 0xFFFFFFFF (aka −1)

~2 == (int) 0xFFFFFFFD (aka −3)

~((unsigned) 2) == 0xFFFFFFFD
```



bit-puzzles

future assignment

bit manipulation puzzles

solve some problem with bitwise ops

- maybe that you could do with normal arithmetic, comparisons, etc.

why?

- good for thinking about HW design
- good for understanding bitwise ops
- unreasonably common interview question type

note: ternary operator

```
w = (x ? y : z)
```

```
if (x) { w = y; } else { w = z; }
```

one-bit ternary

`(x ? y : z)`

constraint: *x, y, and z are 0 or 1*

now: reimplement in C without if/else/||/etc.

(assembly: no jumps probably)

one-bit ternary

```
(x ? y : z)
```

constraint: *x, y, and z are 0 or 1*

now: reimplement in C without if/else/||/etc.

(assembly: no jumps probably)

divide-and-conquer:

```
(x ? y : 0)
(x ? 0 : z)
```

one-bit ternary parts (1)

constraint: *x, y, and z are 0 or 1*

```
(x ? y : 0)
```

one-bit ternary parts (1)

constraint: *x, y, and z are 0 or 1*

```
(x ? y : 0)
```

	y=0	**y=1**
x=0	0	0
x=1	0	1

→ `(x & y)`

one-bit ternary parts (2)

`(x ? y : 0)` = `(x & y)`

one-bit ternary parts (2)

`(x ? y : 0)` = `(x & y)`

`(x ? 0 : z)`

opposite `x`: `~x`

`((~x) & z)`

one-bit ternary

constraint: *x, y, and z are 0 or 1*

```
(x ? y : z)
```

```
(x ? y : 0) | (x ? 0 : z)
```

```
(x & y) | ((~x) & z)
```

multibit ternary

constraint: x *is 0 or 1*

old solution `((x & y) | (~x) & z)` only gets least sig. bit

`(x ? y : z)`

multibit ternary

constraint: x *is 0 or 1*

old solution `((x & y) | (~x) & z)` only gets least sig. bit

```
(x ? y : z)
(x ? y : 0) | (x ? 0 : z)
```

constructing masks

constraint: x *is 0 or 1*

`(x ? y : 0)`

if x = 1: want `1111111111...1` (keep y)

if x = 0: want `0000000000...0` (want `0`)

constructing masks

constraint: x *is 0 or 1*

`(x ? y : 0)`

if x = 1: want `1111111111…1` (keep y)

if x = 0: want `0000000000…0` (want `0`)

a trick: −x (`-1` is `1111…1`)

constructing masks

constraint: x *is 0 or 1*

`(x ? y : 0)`

if x = 1: want `1111111111...1` (keep y)

if x = 0: want `0000000000...0` (want `0`)

a trick: −x (`-1` is `1111...1`)

```
((-x) & y)
```

constructing other masks

constraint: x *is 0 or 1*

`(x ? 0 : z)`

if x = ~~1~~ 0: want `1111111111…1`

if x = ~~0~~ 1: want `0000000000…0`

mask: ~~>x~~

constructing other masks

constraint: x *is 0 or 1*

`(x ? 0 : z)`

if x = 1: 0: want `1111111111…1`

if x = 0: 1: want `0000000000…0`

mask: ~~>x~~ `-(x^1)`

multibit ternary

constraint: x *is 0 or 1*

old solution `((x & y) | (~x) & z)` only gets least sig. bit

```
(x ? y : z)
(x ? y : 0) | (x ? 0 : z)
((-x) & y) | ((-(x ^ 1)) & z)
```

fully multibit

~~constraint: x is 0 or 1~~

```
(x ? y : z)
```

fully multibit

~~constraint: x is 0 or 1~~

`(x ? y : z)`

easy C way: `!x` = 0 or 1, `!!x` = 0 or 1

- x86 assembly: `testq %rax, %rax` then `sete`/`setne` (copy from ZF)

fully multibit

~~constraint: x is 0 or 1~~

`(x ? y : z)`

easy C way: `!x` = 0 or 1, `!!x` = 0 or 1

x86 assembly: `testq %rax, %rax` then `sete`/`setne` (copy from ZF)

`(x ? y : 0) | (x ? 0 : z)`

`((-!!x) & y) | ((-!x) & z)`

simple operation performance

typical modern desktop processor:

bitwise and/or/xor, shift, add, subtract, compare — $\sim$ 1 cycle
integer multiply — $\sim$ 1-3 cycles
integer divide — $\sim$ 10-150 cycles

(smaller/simpler/lower-power processors are different)

simple operation performance

typical modern desktop processor:

bitwise and/or/xor, shift, add, subtract, compare — $\sim$ 1 cycle
integer multiply — $\sim$ 1-3 cycles
integer divide — $\sim$ 10-150 cycles

(smaller/simpler/lower-power processors are different)

add/subtract/compare are more complicated in hardware!

but *much* more important for typical applications

problem: any-bit

is any bit of x set?

goal: turn 0 into 0, not zero into 1

easy C solution: `!(!(x))`

another easy solution if you have − or + (lab exercise)

what if we don't have ! or − or +

problem: any-bit

is any bit of x set?

goal: turn 0 into 0, not zero into 1

easy C solution: `!(!(x))`

another easy solution if you have − or + (lab exercise)

what if we don't have `!` or − or +

how do we solve is x is two bits? four bits?

problem: any-bit

is any bit of x set?

goal: turn 0 into 0, not zero into 1

easy C solution: `!(!(x))`

another easy solution if you have − or + (lab exercise)

what if we don't have ! or − or +

how do we solve is x is two bits? four bits?

```
((x & 1) | ((x >> 1) & 1) | ((x >> 2) & 1) | ((x >> 3) & 1))
```

wasted work (1)

```
((x & 1) | ((x >> 1) & 1) | ((x >> 2) & 1) | ((x >> 3) & 1))
```

in general: `(x & 1) | (y & 1) == (x | y) & 1`

wasted work (1)

```
((x & 1) | ((x >> 1) & 1) | ((x >> 2) & 1) | ((x >> 3) & 1))
```

in general: `(x & 1) | (y & 1) == (x | y) & 1`

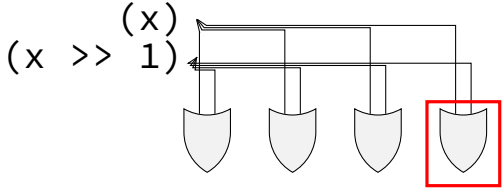
```
(x | (x >> 1) | (x >> 2) | (x >> 3)) & 1
```

wasted work (2)

4-bit any set: `(x | (x >> 1)| (x >> 2) | (x >> 3)) & 1`

performing 3 bitwise ors

...each bitwise or does 4 OR operations



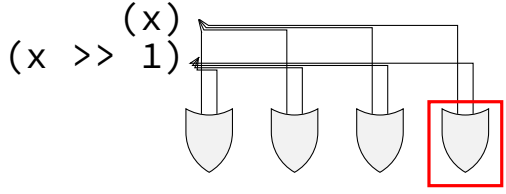
wasted work (2)

4-bit any set: `(x | (x >> 1)| (x >> 2) | (x >> 3)) & 1`

performing 3 bitwise ors

...each bitwise or does 4 OR operations

but only result of one of the 4!



any-bit: divide and conquer

four-bit input $x = x_1x_2x_3x_4$

`x | (x >> 1)` $= (x_1|0)(x_2|x_1)(x_3|x_2)(x_4|x_3) = y_1y_2y_3y_4$

any-bit: divide and conquer

four-bit input $x = x_1x_2x_3x_4$

`x | (x >> 1)` $= (x_1|0)(x_2|x_1)(x_3|x_2)(x_4|x_3) = y_1y_2y_3y_4$

`y | (y >> 2)` $= (y_1|0)(y_2|0)(y_3|y_1)(y_4|y_2) = z_1z_2z_3z_4$

$z_4 = (y_4|y_2) = ((x_2|x_1)|(x_4|x_3)) = x_4|x_3|x_2|x_1$ "is any bit set?"

any-bit: divide and conquer

four-bit input $x = x_1x_2x_3x_4$

`x | (x >> 1)` $= (x_1|0)(x_2|x_1)(x_3|x_2)(x_4|x_3) = y_1y_2y_3y_4$

`y | (y >> 2)` $= (y_1|0)(y_2|0)(y_3|y_1)(y_4|y_2) = z_1z_2z_3z_4$

$z_4 = (y_4|y_2) = ((x_2|x_1)|(x_4|x_3)) = x_4|x_3|x_2|x_1$ "is any bit set?"

```
unsigned int any_of_four(unsigned int x) {
    int part_bits = (x >> 1) | x;
    return ((part_bits >> 2) | part_bits) & 1;
}
```

any-bit-set: 32 bits

```
unsigned int any(unsigned int x) {
    x = (x >> 1) | x;
    x = (x >> 2) | x;
    x = (x >> 4) | x;
    x = (x >> 8) | x;
    x = (x >> 16) | x;
    return x & 1;
}
```

bitwise strategies

use paper, find subproblems, etc.

mask and shift

```
(x & 0xF0) >> 4
```

factor/distribute

```
(x & 1) | (y & 1) == (x | y) & 1
```

divide and conquer

common subexpression elimination

```
return ((-!!x) & y) | ((-!x) & z)
```

becomes

```
d = !x; return ((-!d) & y) | ((-d) & z)
```

exercise

Which of these will swap last and second-to-last bit of an `unsigned int` x? ($abcdef$ becomes $abcdfe$)

```
/* version A */
    return ((x >> 1) & 1) | (x & (~1));

/* version B */
    return ((x >> 1) & 1) | ((x << 1) & (~2)) | (x & (~3));

/* version C */
    return (x & (~3)) | ((x & 1) << 1) | ((x >> 1) & 1);

/* version D */
    return (((x & 1) << 1) | ((x & 3) >> 1)) ^ x;
```

version A

```
/* version A */
    return ((x >> 1) & 1) | (x & (~1));
    //      ^^^^^^^^^^^^^^
    //       abcdef --> 0abcde -> 00000e

    //                         ^^^^^^^^^^
    //       abcdef --> abcde0

    //      ^^^^^^^^^^^^^^^^^^^^^^^^^^^^
    //       00000e | abcde0 = abcdee
```

version B

```
/* version B */
    return ((x >> 1) & 1) | ((x << 1) & (~2)) | (x & (~3));
    //      ^^^^^^^^^^^^^^
    //       abcdef --> 0abcde --> 00000e

    //                          ^^^^^^^^^^^^^^^
    //       abcdef --> bcdef0 --> bcde00

    //                                              ^^^^^^^^^
    //       abcdef -->             abcd00
```

version C

```
/* version C */
    return (x & (~3)) | ((x & 1) << 1) | ((x >> 1) & 1);
    //      ^^^^^^^^^^
    //       abcdef -->             abcd00

    //                    ^^^^^^^^^^^^^^
    //       abcdef --> 00000f --> 0000f0

    //                                    ^^^^^^^^^^^^
    //       abcdef --> 0abcde --> 00000e
```

version D

```
/* version D */
    return (((x & 1) << 1) | ((x & 3) >> 1)) ^ x;
    //      ^^^^^^^^^^^^^^^
    //       abcdef --> 00000f --> 0000f0

    //                          ^^^^^^^^^^^^^^
    //       abcdef --> 0000ef --> 00000e

    //      ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^
    //      0000fe ^ abcdef --> abcd(f XOR e)(e XOR f)
```

expanded code

```
int lastBit = x & 1;
int secondToLastBit = x & 2;
int rest = x & ~3;
int lastBitInPlace = lastBit << 1;
int secondToLastBitInPlace = secondToLastBit >> 1;
return rest | lastBitInPlace | secondToLastBitInPlace;
```

backup slides

example: C that is not C++

valid C and invalid C++:

```
char *str = malloc(100);
```

valid C and valid C++:

```
char *str = (char *) malloc(100);
```

valid C and invalid C++:

```
int class = 1;
```