CSC 411

Computer Organization (Fall 2024)
Lecture 5: Casting, Byte ordering, Pointers

Prof. Marco Alvarez, University of Rhode Island

Quick notes

- For Windows users requiring Linux access:
  - install a Linux distribution
    - standalone or side-by-side with Windows (dual-boot)
  - use Windows subsystem for Linux (WSL)
    - https://docs.microsoft.com/en-us/windows/wsl/install-win10
  - use a virtual machine

Practice

- What decimal value is this bit sequence representing?

```
1 0 0 1 0
```

  - assume bit-width $n = 5$, unsigned
  - assume bit-width $n = 5$, two's complement
  - assume bit-width $n > 5$

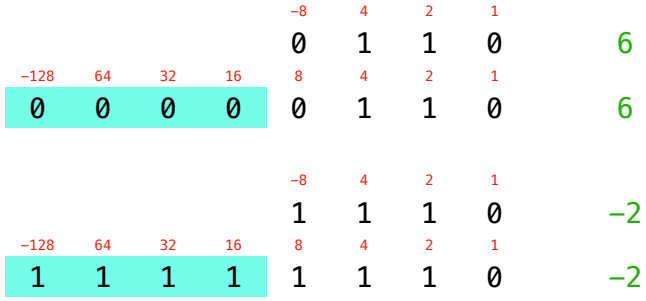
Casting

Casting in C

- Constants
  - considered signed integers by default, use U suffix for unsigned, e.g. `502123U`
- Casting
  - changes interpretation, the bit sequence is **maintained**
  - NOT the same as converting a positive value d into its negative $-d$
  - two's complement, conversions between signed and unsigned basically add/subtract 2^n
- Types of casting
  - **explicit casting:** requires parenthesized type specification
  - **implicit casting:** occurs automatically in assignments and function calls
    - e.g. assigning `unsigned int` to `signed int`

Casting to larger data types

- Sign extension
  - transform n-bit integer to m-bit integer ($m > n$), preserving value
  - method: make $m - d$ copies of the MSB



Practice

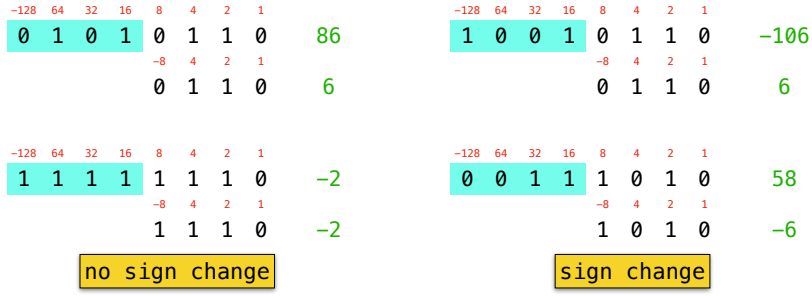
- What is the decimal value of this 64-bit number represented using two's complement?

`0xFFFFFFFFFFFFFFF8`

`11111111 11111111 11111111 11111111 11111111 11111111 11111111 11111000`

Casting to smaller data types

- Truncation
  - transform n-bit integer to m-bit integer ($m < n$), preserving rightmost m bits
  - method: drop top $n - m$ bits

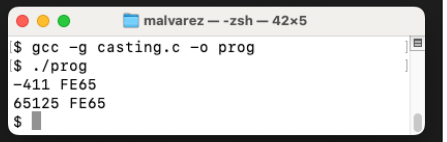


What is the output?

```c
#include <stdio.h>
#include <stdint.h>

int main() {
    // 411 is 0x019B
    // -411 is 0xFE65
    int16_t var1 = -411;
    uint16_t var2 = (uint16_t) var1;
    // %h (short integer)
    printf("%hd %x\n", var1, var1);
    printf("%hu %x\n", var2, var2);

    return 0;
}
```



Casting in expressions

- If an expression contains signed and unsigned integers …
  - signed values are **implicitly casted** to **unsigned**
    - **recall -2147483648 is the most negative number in 32-bit signed integers**

Expression	Type	Evaluation
0 == 0U	unsigned	true
-1 < 0	signed	true
-1 < 0U	unsigned	false
2147483647 > -2147483647 - 1	signed	true
2147483647U > -2147483647 - 1	unsigned	false
2147483647 > (int)2147483648U	signed	true
-1 > -2	signed	true
(unsigned)-1 > -2	unsigned	true

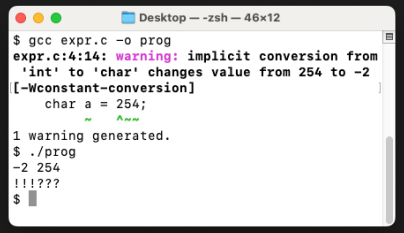
What is the output?

```c
#include <stdio.h>

int main() {
    char a = 254;
    unsigned char b = 254;
    unsigned int c = 0;

    printf("%d %d\n", a, b);

    if (-1 < c) {
        printf("yay\n");
    } else {
        printf("!!!???\n");
    }
}
```



Memory organization

Memory organization

- Memory as a **byte array**
  - used to store **data and instructions** for computer programs
  - contiguous sequence of bytes
    - each byte individually accessed via a **unique address**
- Memory address
  - **unique** numerical identifier for each byte in memory
  - **pointer** variables store memory addresses
    - provides indirect access to data stored at that location

Memory organization

- Data representation in memory
  - variables stored as byte sequences
  - interpretation depends on type
    - integers, floating-point numbers, characters, etc.
- OS provides private address space to each **"process"**
  - process: a program being executed
  - address space: enormous arrays of bytes visible to the process

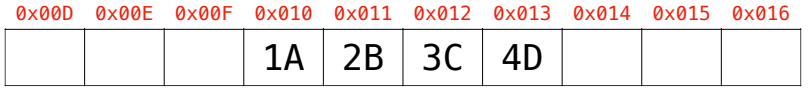
Machine words

- Computers have a **"word size"**
  - usually the size of integer-valued data and memory addresses
    - 32-bit word size: address range of $0 \ldots 2^{32} - 1$
      - 4294967296 bytes or ~4GB
    - 64-bit word size: address range of $0 \ldots 2^{64} - 1$
      - 18446744073709551616 bytes or ~18EB
- Machines support multiple data formats
  - fractions or multiples of word size

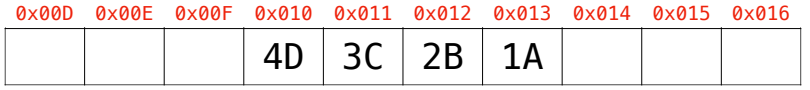
Byte ordering

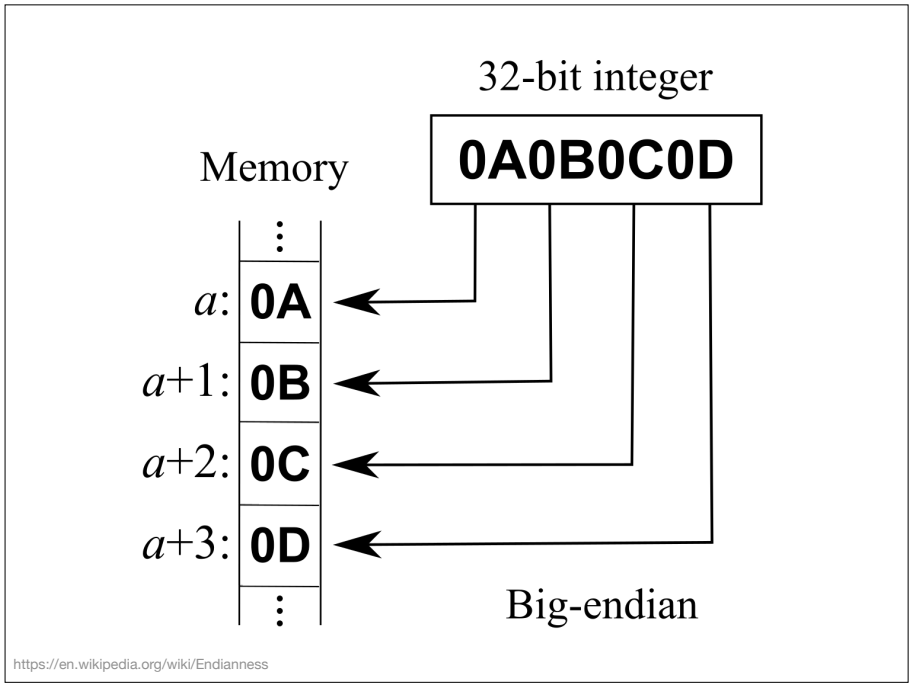
```
x = 0x1A2B3C4D
assume &x is 0x010
```

- Big Endian
  - stores the **most significant byte** at lowest memory address
  - IBM PowerPC, Motorola 68000, SPARC, used in network byte order

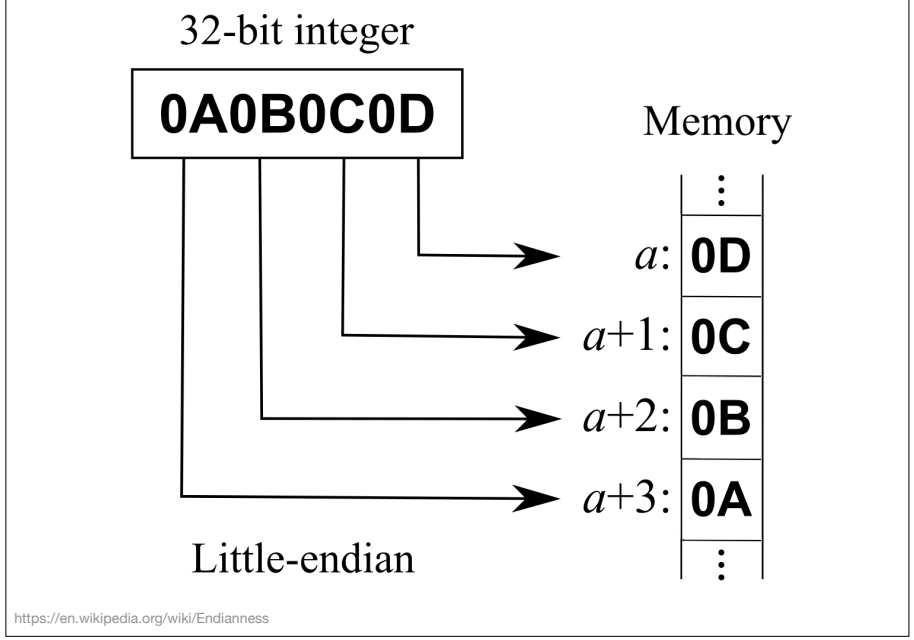


- Little Endian
  - stores the **least significant byte** at lowest memory address
  - intel x86, ARM, RISC-V, MIPS





https://en.wikipedia.org/wiki/Endianness



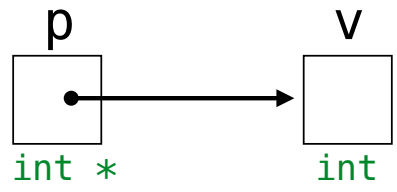
https://en.wikipedia.org/wiki/Endianness

Pointers

Variables and pointers

- Every variable exists at a **memory address** (regardless of **scope**)
  - memory address corresponds to a unique location
- The compiler translates names to addresses when generating machine code
  - C allows direct manipulation of variables and addresses

A **pointer** is a variable that stores the address of another variable



Pointers in C

- Must be declared before use
  - **pointer type** must be specified
- Pointer operators
  - **address-of operator**: get memory address of variable/ object

&

  - **dereference operator**: get value at given memory address

*

Declaring pointers

```c
// can declare a single
// pointer (preferred)
int *p;

// can declare multiple
// pointers of the same type
int *p1, *p2;

// can declare pointers
// and other variables too
double *p3, var, *p4;
```

Pointer operators

```c
int main() {
    int var = 10;
    int *ptr;
    ptr = &var;
    *ptr = 20;

    // ...

    return 0;
}
```

32-bit words

Address	Value	Variable
…		
0x91340A08		
0x91340A0C		
0x91340A10		
0x91340A14		
0x91340A18		
0x91340A1C		
0x91340A20		
0x91340A24		
0x91340A28		
0x91340A2C		
0x91340A30		
0x91340A34		
…		

Pointer operators

```c
int main() {
    int temp = 10;
    int value = 100;
    int *p1, *p2;

    p1 = &temp;
    *p1 += 10;

    p2 = &value;
    *p2 += 5;

    p2 = p1;
    *p2 += 5;

    return 0;
}
```

32-bit words

Address	Value	Variable
…		
0x91340A08		
0x91340A0C		
0x91340A10		
0x91340A14		
0x91340A18		
0x91340A1C		
0x91340A20		
0x91340A24		
0x91340A28		
0x91340A2C		
0x91340A30		
0x91340A34		
…		

Pointers and functions

```c
void increment(int *ptr) {
    (*ptr) ++;
}

int main() {
    int var = 10;

    increment(&var);
    increment(&var);

    // ...

    return 0;
}
```

32-bit words

Address	Value	Variable
…		
0x91340A08		
0x91340A0C		
0x91340A10		
0x91340A14		
0x91340A18		
0x91340A1C		
0x91340A20		
0x91340A24		
0x91340A28		
0x91340A2C		
0x91340A30		
0x91340A34		
…		

Null pointers and arrays

- The `null` pointer (`0x00000000`)
  - represents the absence of value
  - reading/writting with a null pointer can generate a **segmentation fault** signal
- Pointers and arrays
  - arrays treated as “constant pointers” to the first element
  - array names are NOT variables

Pointer arithmetic

- Can add values to pointers
  - treats addresses as unsigned integers
- Must be careful !
  - p+1 adds the size of pointed variable
    - p+1 does NOT add 1 “byte”
- Can use pointer arithmetic for array traversal

```
a[i]     is equivalent to     *(a+i)
```

Advanced pointer concepts

- Declaring a pointer allocates space for the pointer
  - 4 (32-bit architecture) or 8 bytes (64-bit architecture), NOT additional memory
- Generic pointers (`void *`)
  - use carefully to avoid bugs/vulnerabilities
- Pointer to functions
  - allow indirect function calls

```c
// declare and initialize
int (*func) (int, int) = &my_func;
// use
c = (*func)(a, b);
```

Changing a pointer inside a function

```c
#include <stdio.h>

void seek(int *array, int key, int n) {
    for (int i = 0 ; i < n; i++) {
        if (*array == key) {
            return;
        }
        array ++;
    }
}

int main() {
    int data[] = {1, 2, 3, 4, 5};
    int *p = data;

    seek(data, 3, 5);
    printf("%d\n", *p);

    return 0;
}
```

does it work?

Using double pointers

```c
// function to search for a key in an array
// - pointer to an array of integers
// - an integer key
// - an integer n, the number of elements

void seek(int **p, int key, int n) {
    for (int i = 0 ; i < n; i++) {
        if (**p == key) {
            return;
        }
        (*p) ++;
    }
}
```

Using double pointers

```c
int main() {
    int data[] = {1, 2, 3, 4, 5};
    int *p = data;

    seek(&p, 3, 5);
    printf("%d\n", *p);

    return 0;
}
```