CENG 140 – C Programming

Introduction

Asst. Prof. Selim Temizer
C Programming Language - Remarks

• C is a simple language, very few commands, small and easy to learn
  • Just 5-6 pages of specification/grammar

• C is a powerful language
  • Enables you to directly access hardware
  • Language of Operating System development

• “Number of statements / Work done by CPU” ratio is high
  • You have to tell everything to the computer
  • A little lower level when compared to some other languages
  • Compare with Java, Python, Matlab, Scripting Languages
    • Most interpreted languages usually have higher ratios
Our Learning Goals for CENG 140

• C language is very helpful in understanding how computers work
• We will study how our C programs are processed by computers
  • Think like the compiler / computer
• At MIT, corresponding course is 6.001 SICP
  • Structure and Interpretation of Computer Programs
  • They used Lisp, Java, Python, ... as a tool to study structure & interpretation
• We will learn
  • C programming language, professionally
  • Structure of C programs
  • How computers execute our C programs
  • Compiled languages behave similarly in general, so our benefit is more than C
My Teaching Observations

• Taught C to high school students preparing for coding competitions
  • They liked visual descriptions

• Taught C to Food Engineering students (CENG 230)
  • They ranked 2\textsuperscript{nd} after EEE students (with good attendance to lectures)
  • They ranked 7\textsuperscript{th} or 8\textsuperscript{th} (with poor attendance to lectures)

• Taught C to co-workers at IT and Defence companies
  • They liked visual examples of how computers execute our programs
First Step – Let’s Talk about Computers
(From a Programmer’s Point)

• Everything is composed of 0’s and 1’s
  • Binary number system (cf. decimal, octal, hexadecimal number systems)
  • Bit, byte, word, double word
  • How to represent:
    • integers, large integers, negative numbers (sign bit, one’s complement, two’s complement),
      real numbers, letters, musical notes, machine language commands
  • How to store in memory, how to transmit
    • Big-endian
    • Little-endian
  • Reading: Appendix D & E

• Components of a computer
  • CPU
    • 32 bit, 64 bit
  • Memory (RAM)
    • Transistors (on/off)
    • Stream of bytes
  • Storage (Hard Disk)
    • Head, sector, track, ferromagnetic material
    • Stream of bytes
  • Keyboard/Terminal (Input)
    • Stream of bytes, buffered
  • Screen/Terminal (Output)
    • Stream of bytes, buffered/unbuffered
  • Other peripherals
Chapter 1, Part 1 – Computer Hardware

http://www.c-jump.com/CIS24/Slides/Hardware/lecture.html

https://www.amazon.com/PC-Intern-Encyclopedia-Programming-Developers/dp/1557553041
Computer Hardware – Physical Look

- **CPU**: Performs the calculations and does the comparisons needed for processing, as well as controls the other parts of the computer system.
- **Power Supply**: Converts standard electrical power into a form the computer can use.
- **Fan**: Cools the CPU.
- **Hard Drive**: Stores data and programs; the principal storage device for most computers.
- **Drive Bays**: Hold storage devices, such as the DVD and hard drives shown here.
- **DVD Drive**: Accesses data stored on CDs or DVDs.
- **Flash Memory Card Reader**: Accesses data stored on flash memory cards.
- **USB Ports**: Connect USB devices to the computer.
- **Expansion Card**: Connects peripheral devices or adds new capabilities to a computer.
- **Expansion Slots**: Connect expansion cards to the motherboard to add additional capabilities.
- **Motherboard**: Connects all components of the computer system; the computer's main circuit board.
- **Memory (RAM) Modules**: Store data temporarily while you are working with it.
- **Memory Slots**: Connect memory modules to the motherboard.
Computer Hardware - Notes

• Everything is stored as bits and bytes
  • CPU registers: specialized, fastest, electronic storage
  • Memory (RAM): specialized, fast, electronic storage
  • Floppy Disk, Hard Disk: slow, electro-mechanic, permanent storage
  • Flash Disk: slow, electronic, permanent storage

• Size of registers in CPU
  • Natural size of data that can be processed by CPU in a single time unit
  • 32-bit, 64-bit architectures
  • The size of a “word” usually corresponds to the natural size of an architecture

• Bus
  • Address Bus: Size is large enough to address all bytes in memory
  • Data Bus: Size is usually equal to the architecture’s natural size
Chapter 1, Part 2 - Software

• Computer Programs
  • Operating Systems (Linux, Unix, Windows, DOS, etc.)
  • Application Programs (Games, MS Office, Web Browsers, etc.)

• Programming Languages
  • Machine language (specific to the hardware)
  • Assembly language (using assemblers)
  • High level languages
    • Compiled languages (using compilers, for example C/C++)
    • Interpreted languages (using interpreters, for example PHP, Perl, Matlab)
    • Compiled halfway to bytecode, and then interpreted (for example Java, C#)
Levels of Programming Languages

High-level program

```java
class Triangle {
    ...
    float surface()
    return b*h/2;
}
```

Low-level program

```asm
LOAD r1,b
LOAD r2,h
MUL r1,r2
DIV r1,#2
RET
```

Executable Machine code

```
0001001001000101
0010010011101100
10101101001...
```

Machine Code in Hex

| 27BB0001 |
| 23BD8050 |
| 23DEFFF0 |
| A61D8018 |
| A77D8010 |
| 47E0F411 |
| B75E0000 |
| 6B5B4000 |
| 27BA0001 |
| A75E0000 |
| 23BD8050 |
| 47FF0400 |
| 23DE0010 |
| 6BFA8001 |

Assembly Code

```
ldah gp, main
lda gp, main
lda sp, -16(sp)
ldq r16, 8(gp)
ldq r27, printf
mov 7, r17
stq r26, (sp)
jsr r26, printf
ldah gp, main
ldq r26, (sp)
lda gp, main
clr r0
lda sp, 16(sp)
ret r26
```

High-Level Code

```c
main()
{
    int a, b, c;
    a = 3;
    b = 4;
    c = a + b;
    printf(“\n%d\n”, c);
}
```
C Language

• Designed by Dennis Ritchie, 1972
  • As the systems language for UNIX operating system
  • AT&T Bell Laboratories
  • Belongs to “Algol” family of programming languages

• History
  • Algol 60 (1960)
  • CPL (1963)
  • BCPL (1969)
  • B (1970)
  • C (1972)
  • ANSI C (1989)
Chapter 1, Part 3 – Programming Process

• Problem Definition
  • Purpose, Inputs, Outputs (Requirements Engineering)

• Program Design
  • Devise an algorithm (clearly defined steps to produce output from input)
  • Top-down design, stepwise refinement

• Program Coding
  • Express algorithm in a programming language
  • In C: each task can be implemented as a function
  • Structured programming: sequential, selective, repetitive control structures

• Program Compilation and Execution (*We will run examples together*)
  • Translate to machine code: (preprocess,) compile, link

• Program Testing and Debugging
  • Desk checking, compile-time errors, run-time errors, logical errors

• Program Documentation
  • Requirements, input/output, top-down design, algorithms, source, tests, user’s guide
  • Self-documenting, comments
GNU C Compiler: gcc

- **--help**: Display gcc help information on stdout (also: man gcc)
- **-E**: Stop after preprocessing (stdout: preprocessed c, extension: .i)
- **-S**: Stop after compiling (file: assembler code, extension: .s)
- **-c**: Stop after assembling (file: object code, extension: .o)
- **-o <file>**: place output in file
Chapter 2 – Sequential Structure

• Structure of a C program file
  • Building blocks: variables (memory storage, hold data) and functions (process data)
  • Standard Library and Third Party Libraries
  • Preprocessor directives and comments (handled by preprocessors)
  • Multiple-file C projects

• How to compile and run a C program
  • Preprocess (preprocessor, to pure C code)
  • Compile (compiler, to assembly language code)
  • Assemble (assembler, to object/machine code)
  • Link (linker, to executable file, with sections text, data, bss)
  • Load and start at main (loader and operating system)
    • Extra memory regions: stack, heap, command-line arguments and environment variables
Function Definitions - Syntax

ReturnType FunctionName ( Type1 Parameter1Name, Type2 Parameter2Name )
{
  Variable Definitions, External Variable/Function Declarations
     Type var1Name, var2Name, var3Name, ... ;
     extern Type globalVar1Name, globalVar2Name, ... ;
     extern ReturnType otherFunctionName ( Type1, Type2, Type3 ) ;
  Statements
     Expression1 ;
     Expression2 ;
     ...
}

Definition vs. Declaration

• Definition
  • Purpose: To *create building blocks* of applications (*variables* and *functions*)!
  • Variables: Storage space is allocated in memory, might also be initialized
  • Functions: Implementation is provided (to be compiled into machine code)

• Declaration
  • Purpose: To *let the compiler know* that a building block is created elsewhere
    (Otherwise, the compiler will be confused and may not be able to compile)
  • Variables: The storage is/will be created elsewhere, you may trust us to use it
  • Functions: The implementation is/will be created elsewhere, trust us to use it

• Attention:
  • Most people/books/articles use the two terms inconsistently!
Definition / Declaration Example

```
#include <stdio.h>

extern int result;
extern int add( int a , int b ) ;

int main ( void )
{
    result = add( 3 , 5 ) ;
    printf( "Sum is equal to : %d\n" , result ) ;
    return 0 ;
}
```

```
int result ;
int add ( int a , int b )
{
    return ( a + b ) ;
}
```

- Compile one by one, and then link
  - gcc –Wall –c myLibrary.c
  - gcc –Wall –c myApplication.c
  - gcc –Wall myLibrary.o myApplication.o

- Or, compile together in a single command
  - gcc –Wall myLibrary.c myApplication.c
Notes on Functions - 1

• Functions might be defined to have parameters
  • When calling them, we pass arguments
  • Sometimes people use parameter / argument inconsistently
• Arguments can be any valid expressions
  • They are evaluated first, before the called function is invoked
  • Warning: The order of evaluation might be compiler dependent
    • Example: Unexpected behavior might happen in a call like add( a++, ++a );
  • A copy of their values are passed to the function (call-by-value / pass-by-value)
    • C language does NOT support call-by-reference / pass-by-reference mechanism

• Function name is the address of the starting byte of the function’s machine code in memory
Notes on Functions - 2

• Function **prototype**
  • `ReturnType FunctionName ( Type1 Param1Name, Type2 Param2Name )`
  • Parameter names are actually not needed

• Function **signature**
  • `FunctionName ( Type1 Param1Name, Type2 Param2Name )`
  • `ReturnType` is not included in signature
  • Parameter names are actually not needed

• Sometimes people use prototype / signature inconsistently
Comments

• Start with character sequence  /*
• End with character sequence  */
• Can be placed wherever a whitespace can be placed
• Simply filtered out by preprocessor/compiler
• Cannot be nested (filtering out is not a smart operation)
  • Filtering starts with  /*  and ends with the first occurrence of  */
• Use comments to convey useful information to readers of your code
  • Nicely commented programs are easier to maintain by yourself and others
  • Don’t use non-informative comments, some useless examples:
    • int i ;  /* This is a variable */  → Yes, it is. Trust me, I know!
    • /* Now I will implement the main function below */  → We see it, no need to tell!
Character Set

- Set of characters that may appear in a legal C program
  - Textbook, Section 2.2, Page 27
- Graphic characters
- Non-graphic characters
  - Escape sequences
  - Backslash followed by a letter
- Whitespace characters
  - Space, horizontal tab, vertical tab, newline (and carriage return), formfeed
Data Types

- Textbook, Section 2.3, Page 28

- char (signed, unsigned)
- int (short, long) (signed, unsigned)
- float
- double (long)
- void

- char : 1 byte
- sizeof(short) ≤ sizeof(int) ≤ sizeof(long)
- Precision of float ≤ Precision of double ≤ Precision of long double
- Implementation specific bit widths in: limits.h, float.h
Data Types – Figure 1
Data Types – Figure 2

Data types

Primitive type
- char
- int
- float/double
- void

Derived type
- signed
  - signed char
  - signed short
  - signed int
  - signed long
- unsigned
  - unsigned char
  - unsigned short
  - unsigned int
  - unsigned long
- long double

User defined type
- Array
  - struct
  - union
  - enum
Data Types – Figure 3

Basic Datatypes (Primary Datatypes)

- Integer
  - Signed
    - int
    - short int
    - long int
  - Unsigned
    - int
    - short int
    - long int
- Floating Point
  - float
  - double
  - long double
- Character
  - char
  - signed char
  - Unsigned Char
## Data Types – Figure 4

<table>
<thead>
<tr>
<th>Type</th>
<th>Bits</th>
<th>Minimal Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>char</td>
<td>8</td>
<td>−127 to 127</td>
</tr>
<tr>
<td>unsigned char</td>
<td>8</td>
<td>0 to 255</td>
</tr>
<tr>
<td>signed char</td>
<td>8</td>
<td>−127 to 127</td>
</tr>
<tr>
<td>int</td>
<td>16 or 32</td>
<td>−32,767 to 32,767</td>
</tr>
<tr>
<td>unsigned int</td>
<td>16 or 32</td>
<td>0 to 65,535</td>
</tr>
<tr>
<td>signed int</td>
<td>16 or 32</td>
<td>Same as int</td>
</tr>
<tr>
<td>short int</td>
<td>16</td>
<td>−32,767 to 32,767</td>
</tr>
<tr>
<td>unsigned short int</td>
<td>16</td>
<td>0 to 65,535</td>
</tr>
<tr>
<td>signed short int</td>
<td>16</td>
<td>Same as short int</td>
</tr>
<tr>
<td>long int</td>
<td>32</td>
<td>−2,147,483,647 to 2,147,483,647</td>
</tr>
<tr>
<td>long long int</td>
<td>64</td>
<td>−(2(^63 ) − 1) to 2(^63 ) − 1 (Added by C99)</td>
</tr>
<tr>
<td>signed long int</td>
<td>32</td>
<td>Same as long int</td>
</tr>
<tr>
<td>unsigned long int</td>
<td>32</td>
<td>0 to 4,294,967,295</td>
</tr>
<tr>
<td>unsigned long long int</td>
<td>64</td>
<td>2(^{64} ) − 1 (Added by C99)</td>
</tr>
<tr>
<td>float</td>
<td>32</td>
<td>1E−37 to 1E+37 with six digits of precision</td>
</tr>
<tr>
<td>double</td>
<td>64</td>
<td>1E−37 to 1E+37 with ten digits of precision</td>
</tr>
<tr>
<td>long double</td>
<td>80</td>
<td>1E−37 to 1E+37 with ten digits of precision</td>
</tr>
</tbody>
</table>
Variables and Constants

• A computer program manipulates two kinds of data – variables and constants. (Textbook, Section 2.4, Page 29)

• Variables: Programs use variables to store values
  • Every variable has a type
  • Must be defined/declared before used
  • Variable names obey identifier naming rules (next slide)

• Constants: An entity whose value does not change during execution
  • 5 types: integer, floating-point, character, string, enumeration
  • Enumeration is discussed in Chapter 12 of textbook
Identifiers (User Given Names)

• Names of variables, functions, definitions
• Must begin with a letter or underscore ("_")
• May contain letters, underscore, decimal digits (0-9)
  • Corollary: Cannot contain whitespaces
• Case sensitive
• Significant length is compiler-dependent
  • ANSI C permits at least 31 significant characters
• Don’t use 32 standard keywords, standard library function names, etc.
  • List of standard keywords is in next slide
# 32 Standard Keywords in ANSI C

<table>
<thead>
<tr>
<th>auto</th>
<th>break</th>
<th>case</th>
<th>char</th>
<th>const</th>
<th>continue</th>
</tr>
</thead>
<tbody>
<tr>
<td>default</td>
<td>do</td>
<td>double</td>
<td>else</td>
<td>enum</td>
<td>, extern</td>
</tr>
<tr>
<td>float</td>
<td>for</td>
<td>goto</td>
<td>if</td>
<td>int</td>
<td>long</td>
</tr>
<tr>
<td>register</td>
<td>return</td>
<td>short</td>
<td>signed</td>
<td>sizeof</td>
<td>static</td>
</tr>
<tr>
<td>struct</td>
<td>switch</td>
<td>typedef</td>
<td>union</td>
<td>unsigned</td>
<td>void</td>
</tr>
<tr>
<td>volatile</td>
<td>while</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Constants

• Integer Constants
  • Decimal, Octal (0...), Hexadecimal (0x..., 0X...)
  • Type is normally \texttt{int}, if doesn’t fit then \texttt{long}
  • Possible suffix: u, U (unsigned), ul, UL (unsigned long)

• Floating-Point Constants
  • Written with a decimal point and/or in scientific notation (e, E)
  • Type is \texttt{double}
  • Possible suffixes: f, F (float), l, L (long double)
Constants

• Character Constants
  • Single character within apostrophes: ‘a’, ‘B’, ‘?’
  • Octal or Hexadecimal byte-sized bit pattern: ‘\ooo’, ‘\xhh’
  • Null character (zero): ‘\0’
  • Type is int (ASCII or EBCDIC code)
  • Escape character: backslash (\)

• String Constants
  • Zero or more characters (may also be non-graphic) within double quotation marks
  • String can be continued by putting a backslash at the end of the line
  • Adjacent string constants are concatenated at compile time
  • Escape character: backslash (\)
  • Compiler automatically puts ‘\0’ at the end of each string (storage = chars + 1 byte)
  • ‘Z’ is not the same as “Z”
  • Type is address (char *, pointer, address of the first byte of the string)
Arithmetic Operators

• Operator: A symbol that causes specific mathematical or logical manipulations to be performed

• Unary: +, -, increment (++) , decrement (––)

• Binary: +, -, *, /, %

• Integer division is truncated
  • 8 / 3 will give you an integral result → 2
  • 8.0 / 3 or 8 / 3.0 or 8.0 / 3.0 will give you a real number → 2.666667

• Remainder (%) operator takes only integral operands
Expressions and Statements

• Expression: A combination of
  • Constants
  • Variables
  • Operations (arithmetic, logic, ...)
  • Function calls
  • ...

• All expressions have a value
  • To compute the value, refer to precedence and association rules (Appendix B)
  • Or use balanced parentheses

• Almost everything in C is an expression

• Statement (roughly): An expression followed by a semicolon
Precedence and Associativity of Operators

- The C Programming Language
  - ANSI C, 2nd Edition
  - Kernighan & Ritchie
  - Page 53, Table 2-1

![Table 2-1. Precedence and Associativity of Operators](image)
Assignment Statements

• Assignment expression: variable = expression
• Assignment statement: assignment expression followed by semicolon
• Value of expression is assigned to the storage space (box) of variable
• Left side can only be an “L-Value”
  • A variable can be used as an l-value (box) or an r-value (the value in the box)
  • Constants are always r-values
• Value of assignment expression: whatever value is stored in variable
  • Assignment expressions can be chained and can be used in other expressions
• x = y has a different meaning/result than y = x
Increment (++) and Decrement (--) Operators

• Unary, Operates on a variable, that is, requires an L-Value (box)
  • For example the following usage is wrong (compile-time error): \( a = ++b++ \);

• Prefix usage (increment/decrement before using)
  • Example: \( \text{int } i = 4; \ \text{printf}( \text{"%d"}, ++i) ; \)
  • First increment/decrement the variable (and then use the resulting value)
  • The value of the expression is the incremented/decremented value

• Postfix usage (increment/decrement after using)
  • Example: \( \text{int } i = 4; \ \text{printf}( \text{"%d"}, i++ ) ; \)
  • The value of the expression is the original value (use the original value)
  • Then, as a side effect, increment/decrement the variable
Increment (++) and Decrement (--) Operators

• Examples
  • int i = 8; int k = ++i;
    • /* i becomes 9, k is assigned 9 */
  • int i = 8; int k = i++;
    • /* k is assigned value of expression (8), i becomes 9 */
  • int x = 3; int y = 5; int k = --x + y--;
    • /* x becomes 2, k is assigned 7 (2+5), y becomes 4 */

• Warning
  • Since the increment/decrement operator modifies its operand, use of such an operand more than once within the same expression can produce unexpected results.
  • int a = 10; int b = ++a + ++a + ++a; /* b might be assigned 39 or 37 */
    • /* ++ has higher precedence, a might be incremented 3 times first, then summed up */
Compound Assignment Operators

• C provides 10 compound assignment operators (for compact notation)
  • += , -= , *= , /= , %= , <<= , >>= , &= , |= , ^= 

• Semantics
  • variable operator= expression is equivalent to
  • variable = variable operator ( expression )

• Value of compound assignment is equal to whatever value is assigned to variable

• Example
  • i *= a + 1 ; is equivalent to i = i * (a + 1) ; 
Nested Assignments

• C permits multiple assignment expressions in a single statement
  • Since assignment expressions have values, they can be used like any r-value

• Assignment operators are right-associative
  • i = j = k = 0 ;  is interpreted as  i = ( j = ( k = 0 ) ) ;

• Examples from textbook
  • i += j = k ;  is treated as  i += ( j = k ) ;
  • i = j += k ;  is treated as  i = ( j += k ) ;
#include <stdio.h>

• Pure C language does not have language constructs for input/output

• ANSI C has a rich standard library
  • We will use `printf` and `scanf` functions from the standard library
  • Their definitions are in the compiled standard library
  • Their declarations are in `stdio.h` header file
**Printf**

- **Example**
  - `int a = 4; printf("a = %d\n", a);`

- `printf( control string , arg1 , arg2 , ... );`
  - Ordinary characters are sent to output without any modification
  - Escape sequences may be used (like ‘\n’, ‘\t’, ...)
  - Conversion specifications (placeholders) are replaced with argument values
  - Returns number of characters written (its value is an `int`)

---

- The integer argument is converted to decimal notation of the form `[-]ddd.`
- The `float` or `double` argument is converted to decimal notation of the form `[-]ddd.dddddd`. By default, six digits are printed after the decimal point.
- The `float` or `double` argument is converted to decimal notation of the form `[-]j.d.ddddddde[+]{j}.dd`. There is one digit before the decimal point, and the exponent contains at least two digits. By default, six digits are printed after the decimal point.
- The argument is taken to be a single character.
- The argument is taken to be a string.
Scanf

- Example
  - int a;
  - scanf( "%d", &a );

- scanf( control string, arg1, arg2, ... );
  - Space or whitespace within control string matches with whitespace in the input
  - Even if no space in control string, skips over whitespaces in the input to match next data item (unless reading a char and unless no corresponding whitespace in control string)
  - Ordinary characters are expected to be matched from the input (uncommon)
  - Requires addresses of (pointers to) variables (storage boxes) to store the value in
  - If the input buffer is empty, lets the user enter data, otherwise uses data in buffer
  - Returns number of arguments successfully read (its value is an int)

d, i A decimal integer is expected in the input. The corresponding argument should be a pointer to an int.

f, e A floating-point number is expected in the input. The corresponding argument should be a pointer to a float. The input can be in the standard decimal form or in the exponential form.

c A single character is expected in the input. The corresponding argument should be a pointer to a char. Only in this case, the normal skip over the whitespaces in input is suppressed.
Type Conversion

• An expression may contain variables and constants (and calls to functions that return values) of different types.
  • Example: 8 + 3.14 / 2.f (integer, double, float)

• Mechanisms of evaluating mixed type expressions
  • Automatic Type Conversion
  • Explicit Type Conversion
  • Type Conversion in Assignments
Automatic Type Conversion

- ANSI C performs all arithmetic operations with just 6 data types:
  - int, unsigned int, long int, float, double, long double
- char and short is implicitly converted to int before an operation
  - This is called automatic unary conversion (always performed)
  - char → int conversion is machine dependent (signed? or unsigned?)
- After unary conversions, automatic binary conversions are applied
  - Lower type is promoted to higher type before the operation proceeds
  - Performed when necessary

Textbook, Figure 2.4, Page 46
Explicit Type Conversion

• We can force a conversion that is different than automatic conversion:
  • Syntax: \((\text{cast-type})\) expression

• Cast is \textit{unary operator}, has same precedence as other unary operators

• Examples:
  • (int) 12.8 → 12
  • (int) 12.8 * 3.1 → 12 * 3.1 → 37.2
  • (int) (12.8 * 3.1) → (int) (39.68) → 39
  • 3 / 2 → 1
  • (float) 3 / 2 → 1.5
Type Conversion in Assignments

• variable = expression (or compound assignment operators)
• Type of expression’s value is automatically converted to type of variable

• Conversion of lower-order type to higher-order type is OK
  • Only the form of representation changes
• Conversion of higher-order type to lower-order type may cause
  • Truncation
  • Loss of information
Simple Macros

• Preprocessor directive  `#define`
• Syntax:  `#define macroName sequenceOfTokens`
• Macros can also have parameters (will be covered later in the course)
• Preprocessor simply finds and replaces `macroName` with `sequenceOfTokens`
  • Exception: in comments and in string constants
  • Be careful about the `final text after replacement`!
    • Example: `#define FIVE 2+3`  (What is the result? :  `FIVE * 2` )
  • Suggestion: Always put `sequenceOfTokens` inside a pair of parantheses
    • `#define macroName ( sequenceOfTokens )`
• Previously defined macros can be used in new macros
• Can appear anywhere in the program
  • Usually collected together at the beginning of the program
• Convention: Macro names are written in `uppercase letters`
Simple Macros - Example

• If we bury important values in code
  • Difficult to understand by others (what are those magic numbers?)

• Instead: define symbolic constants (like PI)

```c
#include <stdio.h>
#define PI 3.141592

void main (void)
{
    float radius ;

    printf( "Enter radius : " ) ;
    scanf ( " %f" , &radius ) ;
    printf( "Area of circle = " , PI * radius * radius ) ;
}
```
Chapter 3 – Selective Structure

• Selective structures allow us to test for a condition, and follow alternative paths depending on the result of the test
Example

#include <stdio.h>
#define PASSING_GRADE 60

void main (void)
{
  int grade ;

  printf( "Enter grade : " ) ;
  scanf ( " %d" , &grade ) ;

  if ( grade >= PASSING_GRADE )
      printf( "Congratulations! You passed the course!\n" ) ;
  else
      printf( "Sorry! You failed the course!\n" ) ;
}
Chapter 3 - Outline

• Relational Operators ( < , <= , > , >= , == , != )
• Logical Operators ( && , || , ! )
• Conditional Expression Operator ( ? : )
• Conditional Statements ( if , if else )
• Nested Conditional Statements ( if ... if else ... else ... if else )
• Multiway Conditional Statement ( if ... else if ... else if ... else )
• Constant Multiway Conditional Statement ( switch )
About Boolean Values

• The result of testing a condition may be true or false
• C language does NOT have Boolean values (true, false)
• Integers are used as substitutes for Boolean Values
  • Non-zero values (attention: positive and also negative) are interpreted as true
  • Zero is interpreted as false
• Example:
  • if ( 0 ) ... → false, the body of the if statement does NOT execute
  • if ( 2.5 ) ... → true, the body of the if statement executes
• Value of a true expression is 1, false expression is 0
  • printf( “value of expression = %d\n” , ( 2 > 3 ) );
About Compound Statements (Blocks)

• A **block** is a sequence of **variable declarations/definitions** (at the top) and **statements** (at the bottom) enclosed within braces.

• A block can be placed wherever a single statement can be placed
  • No semicolon after closing brace of block

• Some common usages:
  • Grouping statements
    • Fitting more than one statement where only a single statement is allowed
    • Example: Body of if statement, body of loops
  • Opening up a new scope that is enclosed within the current scope
    • Also enables us to define variables, even though we may have statements before in code

```c
if ( grade >= 60 )
{
    saveGradeInDatabase( grade         ) ;
    printf( "Congratulations!\n"       ) ;
    printf( "You passed the course.\n"  ) ;
}
```
Relational Operators

- There are 6 relational operators (Textbook, Table 3.1, Page 69)
- **Value** of a relational expression is of type `int` (0 or 1)
- Comparing characters means comparing their ASCII / EBCDIC codes
  - Example: ‘b’ > ‘a’
- Common programming error:
  - Confusing `==` with `=`
    - if ( x == 10 ) ... → compares
    - if ( x = 10 ) ... → always true!
  - Suggestion: Put r-value first
    - if ( 10 == x ) ... → compares
    - if ( 10 = x ) ... → compile-time error

<table>
<thead>
<tr>
<th>Relational Operator</th>
<th>Name</th>
<th>Relational Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;</td>
<td>Less than</td>
<td><code>exp1 &lt; exp2</code></td>
</tr>
<tr>
<td>&lt;=</td>
<td>Less than or equal to</td>
<td><code>exp1 &lt;= exp2</code></td>
</tr>
<tr>
<td>&gt;</td>
<td>Greater than</td>
<td><code>exp1 &gt; exp2</code></td>
</tr>
<tr>
<td>&gt;=</td>
<td>Greater than or equal to</td>
<td><code>exp1 &gt;= exp2</code></td>
</tr>
<tr>
<td>==</td>
<td>Equal to</td>
<td><code>exp1 == exp2</code></td>
</tr>
<tr>
<td>!=</td>
<td>Not equal to</td>
<td><code>exp1 != exp2</code></td>
</tr>
</tbody>
</table>
Logical Operators

• There are 3 logical operators (Textbook, Table 3.3, Page 71)

• Value of a logical expression is of type int ( 0 or 1 )

• Logical AND: expr1 && expr2
  • First expr1 is evaluated, if zero (false) then SHORT CIRCUITED, result is zero
  • Otherwise, expr2 is evaluated
  • Example: if ( (1>5) && (2/0 >= 3) ) ... → false, division by zero not executed

• Logical OR: expr1 || expr2
  • SHORT CIRCUIT evaluation like logical AND

• Logical NOT: ! expr
  • 0 if expr is nonzero (true), 1 if expr is zero (false)
About Evaluation of Logical Expressions

• For all expressions we use: Precedence and Associativity Table

• Attention (exception)
  • Logical AND and logical OR are always evaluated from left to right (and short circuited)
  • Be careful when expression contains a side effect
  • Good example: Textbook, Page 76
  • Suggestion: Always use parentheses to clearly express the exact order of operations you want

Consider, for example, the expression
\[ \neg a \lor \neg b \land \neg c \]
and let \( a = b = c = 10 \). Following the precedence rule and binding the operands to the operators, we get
\[ \left( \left( \neg a \right) \lor \left( \left( \neg b \right) \land \left( \neg c \right) \right) \right) \]
Hypothetically, if the evaluation order were to follow precedence, \& \& would be evaluated first and we would get
\[ \left( \left( \neg a \right) \lor \left( 9 \land \left( \neg c \right) \right) \right) \]
\[ \left( \left( \neg a \right) \lor \left( 9 \land 9 \right) \right) \]
\[ \left( \left( \neg a \right) \lor \text{true} \right) \]
\[ \left( \text{whatever} \lor \text{true} \right) \]
\[ \text{true, or 1} \]
that is, the expression would be true, and \( a = 10, b = 9, \) and \( c = 9 \). Actually, the evaluation proceeds from left to right, although \& \& has higher precedence than \| . Evaluating from left to right, we get
\[ \left( 9 \land \left( \left( \neg b \right) \land \left( \neg c \right) \right) \right) \]
\[ \left( \text{true} \lor \left( \left( \neg b \right) \land \left( \neg c \right) \right) \right) \]
\[ \left( \text{true} \lor \text{whatever} \right) \]
\[ \text{true, or 1} \]
That is, the expression is again true, but now \( a = 9, b = 10, \) and \( c = 10 \).
Conditional Expression Operator

• The only ternary operator in C language (takes 3 arguments)
• Syntax: \texttt{expr1 \ ? \ expr2 : expr3}

• \textbf{Value} of conditional expression:
  • First \texttt{expr1} is evaluated
  • If nonzero (true), \texttt{expr2} is evaluated, and also becomes the value of the whole
  • If zero (false), \texttt{expr3} is evaluated, and also becomes the value of the whole

• Example:
  • \texttt{larger = x > y \ ? \ x : y ; \rightarrow OK}
  • \texttt{c \? \ x=a : x=b \rightarrow will be interpreted as ( c \? \ x=a : x ) = b \ and is illegal (r-value)}

• Suggestion: Use proper parentheses to avoid unintended errors
Conditional Statements

• If statement
  • Syntax: if ( expression) statement
  • Expression is evaluated
    • If nonzero (true), execute statement
    • Otherwise, don’t execute statement
  • A block can be used as statement
  • Flowchart:

```c
#include <stdio.h>

int maximum ( int a , int b )
{
    if ( a >= b )  return a ;
    return b ;
}

void main ( void )
{
    int x , y ;
    scanf ( " %d %d" , &x , &y ) ;
    printf( "Maximum of them = %d\n" , maximum(x,y) ) ;
}
```
Conditional Statements

• If-else statement
  • Syntax: if ( expression) statement1 else statement2
  • Expression is evaluated
    • If nonzero (true), execute statement1 (if block)
    • Otherwise, execute statement2 (else block)
  • Block(s) can be used as statement(s)
  • Flowchart:

```c
if ( money >= 25.0 )
{
    printf( "You have enough money!\n" );
    printf( "You can buy a good meal.\n" );
}
else
    printf( "Sorry, not enough money!\n" );
```

Suggestion for if and else bodies (conditional statements):

It is good practice to use curly braces even if you have a single statement. Also easier to understand code when reading.
Nested Conditional Statements

• If-blocks and else-blocks can also contain other if, if-else statements
• There is no limit on the depth of nesting

```c
if ( undergraduate == 1 )
    if ( grade >= 60 )
        printf( "Passed!\n" );
    else
        printf( "Failed!\n" );
else
    if ( grade >= 70 )
        printf( "Passed!\n" );
    else
        printf( "Failed!\n" );
```
Sequence of Nested If Statements

- The nested conditional statement on the left is equivalent to the statement on the right
  - C language guarantees short-circuited left-to-right evaluation of logical expressions

```c
if ( undergraduate == 1 )
    if ( grade >= 60 )
        printf( "Passed!\n" );
```

```c
if ( expression-1 && expression-2 && ... && expression-n )
    statement
```
Dangling Else Problem

- Some if statements have *else* parts, some don’t. Which if statement does and *else* part belong to?
- Example: Intention (on left), reality (on right)
- How does the compiler resolve?
  - *Else* is associated with the closest previous *else-less* if
- Suggestion: Always use braces (blocks) to avoid mistakes
- Good examples on Textbook, Page 85-86
Multiway Conditional Statement

A special form of nested conditional statements occurs frequently in practice: cascaded if-else statements where each if-else statement but the last one has another if-else statement in its else block.

Conditional expressions are evaluated in order.

- If one is found to be true, the statement associated with it is executed, and whole chain terminates.
- If none is true, the final else-block is executed.
Constant Multiway Conditional Statement

• General form of the switch statement:

```java
switch ( expression )
{
    case value1 : statement1
                  break ;
    case value2 : statement2
                  break ;
    ...
    case valueN : statementN
                  break ;
    default     : statementD
                  break ;
}
```

Roughly equivalent to:

```java
if ( expression == value1 )
    statement1
else if ( expression == value2 )
    statement2
...
else if ( expression == valueN )
    statementN
else
    statement
```
Constant Multiway Conditional Statement

- Expression must be an **integral** expression
- Case values must be **constant integral** expressions
- The use of **default** is optional
- **Default** (if used) does NOT need to be at the end
- **How does it work?**
  - First, expression is evaluated
  - Its value is successively compared with case values
  - If a case value matches the expression value:
    - Execution continues from the point labelled with the corresponding case
    - We use `break` statement to break out of the switch statement (otherwise all statements until the end of the switch statement will be executed)
  - If value does not match with any case values
    - If there is a **default** case, execution continues from the point labelled with default
    - If there is no **default** case, execution continues with the statements coming after switch statement

```java
switch ( expression )
{
    case value1 : statement1     
                   break ;
    case value2 : statement2     
                   break ;
    ...                                          
    case valueN : statementN     
                   break ;
    default     : statementD     
                   break ;
}
```
Other Notes on Switch Statement

• Textbook, Page 91

No two case values should be the same in a switch statement. However, more than one case value can be associated with a set of statements by listing multiple case values before that particular set of program statements. It is not necessary to terminate each case with a break statement. Omitting the break statement from a particular case causes the execution to continue without regard for the case values and the default label. The break statement, although not logically necessary, should also be placed after the statement associated with default (or the last case value, if default is not at the end) as a matter of good programming practice. This practice avoids inadvertent introduction of a bug in the program when a case value and associated statements are added later on at the end of a switch statement.
```c
#include <stdio.h>

void printDay ( int dayNo )
{
    switch ( dayNo )
    {
        case 1   : printf( "Monday" ) ;
                   break ;
        case 2   : printf( "Tuesday" ) ;
                   break ;
        case 3   : printf( "Wednesday" ) ;
                   break ;
        case 4   : printf( "Thursday" ) ;
                   break ;
        case 5   : printf( "Friday" ) ;
                   break ;
        case 6   : printf( "Saturday" ) ;
                   break ;
        case 7   : printf( "Sunday" ) ;
                   break ;
        default  : printf( "INCORRECT VALUE" ) ;
    }
}

void main ( void )
{
    int day ;

    printf ( "Enter day no : " ) ;
    scanf ( " %d" , &day      ) ;
    printDay( day         ) ;
    printf ( "\n"         ) ;
}
```
About Math Library

#include <math.h>

• sin, cos, tan, asin, acos, atan, exp, log, log10, pow, sqrt, ceil, floor, fabs, ... (Complete list in Appendix A)

• Attention: trigonometric functions mostly use radians (not degrees)

• When compiling with gcc, don’t forget to add “–lm” as an argument
  • gcc –Wall –ansi –pedantic-errors –lm –o <outputFileName> yourApp.c
#include <math.h>
#include <stdio.h>

#define PI 3.141592

double degreeToRadian ( double degree ) { return degree * PI / 180.0 ; }
double radianToDegree ( double radian ) { return radian / PI * 180.0 ; }

void printLine ( double degree )
{
    double sinValue = sin( degreeToRadian( degree ) ) ;

    if ( sinValue <= -0.75 ) printf( "X........" ) ;
    else if ( sinValue <= -0.50 ) printf( ".X......" ) ;
    else if ( sinValue <= -0.25 ) printf( "..X....." ) ;
    else if ( sinValue <= 0.00 ) printf( "...X...." ) ;
    else if ( sinValue <= 0.25 ) printf( "....X..." ) ;
    else if ( sinValue <= 0.50 ) printf( ".....X.." ) ;
    else if ( sinValue <= 0.75 ) printf( "......X." ) ;
    else /*sinValue <= 1.00*/ printf( ".......X" ) ;
}

void main ( void )
{
    double degree ;

    for ( degree = 0 ; degree < 720 ; degree += 45 )
    {
        printLine( degree ) ; printf( "\n" ) ;
    }
Chapter 4 – Repetitive Structure – Outline

• Overview
• while Loop
• do-while Loop
• for Loop
• Nested Loops
• Loop Interruption (**break** and **continue** statements)
• Null Statement
• Comma Operator
Repetitive = Iterative = Loop

- Loop Continuation Condition
  - Controls the number of repetitions
  - Can be tested before or after the body
    - If before, loop body may never execute
      - while, for
    - If after, loop body executes at least once
      - do-while
  - Usually a loop control variable is used

- Loop Body
  - Single statement or a set of statements (compound statement; block)
While Loop

• Syntax: while ( expression )
  statement

• Semantics:
  • Expression is evaluated
  • If true (non-zero), statement is executed
  • If false (zero), the loop terminates

• Expression is NOT optional (it has to be provided)
• Block (compound statement) can be used if loop body (statement) needs to perform multiple actions
While Loop – Example

• Example (textbook, page 108)
• getchar returns int
  • char is not necessarily signed
  • EOF usually defined as -1
    • Terminal: CTRL-D (unix/linux), CTRL-Z (dos)
    • File: When file runs out of data

• putchar expects int
• Try variations:
  • Compact variable definitions
  • while ( ( ch = getchar() ) != EOF )
  • Use scanf/printf for echoing

```c
#include <stdio.h>
void main ( void )
{
  int ch    ;
  int count = 0  ;

  ch = getchar()  ;

  while ( ch != EOF )
  {
    putchar( ch )  ;
    count++        ;
    ch = getchar()  ;
  }

  printf( "Total characters echoed = %d\n" , count )  ;
}
```
Infinite Loop

• A loop that never terminates
  • Might be a logical mistake
  • Might be due to use of = instead of ==
  • Might be due to real numbers
  • Might be intentional
    • Example: A computer game’s main loop
      • while ( 1 )
        {
        getUserInput();
        applyToGame();
        sleep( 100 /* milliseconds */ );
        }
    • Will be broken by a break statement
    • Will be broken by a return statement

```
int n , sum = 0 ;
scanf( "%d" , &n ) ;  /* What if n < 0 */
while ( n != 0 )
{
    sum += n-- ;
}
printf( "Sum = %d\n" , sum ) ;
```
Do-While Loop

• Syntax:  
  \[
  \text{do} \\
  \text{statement} \\
  \text{while ( expression ) ;}
  \]

• Semantics:
  • Execute \text{statement}
  • Evaluate \text{expression}
    • If true (non-zero), execute \text{statement}
    • If false (zero), the loop terminates

• Expression is NOT optional (it has to be provided)
• \textbf{Block} (compound statement) can be used if loop body (\text{statement}) needs to perform multiple actions
Do-While Loop – Example

• Example (textbook, page 113)
• Also works properly for input: 0
  • Try solving with `while` loop
  • Will need special handling for input 0

```c
#include <stdio.h>

void main ( void )
{
    int number , digits = 0 , sum = 0 ;

    printf ( "Enter a non-negative integer : " ) ;
    scanf ( "%d" , &number ) ;

    do
    {
        sum    += number % 10 ;
        number /= 10          ;
        digits++              ;
    }
    while ( number > 0 ) ;

    printf( "Number of digits = %d\n" , digits ) ;
    printf( "Sum    of digits = %d\n" , sum    ) ;
}
```
For Loop

• Syntax: for ( expr1 ; expr2 ; expr3 )
  statement

• Semantics:
  • Evaluate expr1
  • Evaluate expr2
    • If true (non-zero), execute statement
    • Then evaluate expr3
    • If false (zero), the loop terminates

• expr1, expr2 and expr3 are optional

• Block (compound statement) can be used if loop body (statement) needs to perform multiple actions

/* Roughly equivalent to this: */
expr1 ;

while ( expr2 )
{
  statement
  expr3 ;
}
For Loop – Example

• Usually (but not necessarily) :
  • expr1 is for initialization(s)
  • expr2 is loop continuation condition
  • expr3 is re-initialization expression

• If expr2 is omitted
  • It is assumed to be true
  • Infinite loop
    • Might be broken by break, return

• Infinite loop examples
  • for ( ; ; ) ;
  • for ( ; ; ) {} 
  • while ( 1 ) ;
  • while ( 1 ) {} 

```c
#include <stdio.h>

int factorial ( int positiveNumber )
{
  int i , result = 1 ;

  for ( i = 1 ; i <= positiveNumber ; i++ )
  {
    result *= i ;
  }

  return result ;
}

void main ( void )
{
  printf( "5! = %d\n" , factorial( 5 ) ) ;
}
```
Examples

• The following are roughly equivalent
  (variable definitions are shown only in the first example)

```c
int i, sum = 0;
for (i = 1; i <= n; i++)
    sum += i;
printf("%d\n", sum);
```

```c
i = 1;
while (i <= n)
    {
        sum += i;
        i++;
    }
printf("%d\n", sum);
```

```c
i = 1;
for ( ; i <= n; )
    {
        sum += i;
        i++;
    }
printf("%d\n", sum);
```
Examples

• The following are roughly equivalent
  (the second example uses a dummy variable, x, of the same type as
  the expression)

```c
do
  {
    statement
  }
while (expression);
```

```c
for (x=1; x; x= (expression) )
  statement
```
Nested Loops

• The **statement** part of a loop might be / might contain any other loop

• There is no limit on the depth of nesting

```c
#include <stdio.h>

void main ( void )
{
    int row , col ;

    for ( row = 1 ; row < 10 ; row++ )
    {
        for ( col = 1 ; col < 5 ; col++ )
        {
            printf ( " X" ) ;
        }
        printf ( "\n" ) ;
    }

    printf( "\n" ) ;
}
```

Side Note: I wish I was sponsored by: SpaceX for the code example on this slide 😊
Loop Interruption

• Sometimes we may need to
  • Exit the loop using another mechanism than the loop termination condition
  • Skip just the current iteration, but continue with the loop execution

• Almost all languages provide break and continue statements
  • break (inside a loop body) exits the loop
  • continue (inside a loop body) skips the remaining statements in the loop body
Break Statement

• **break** statement in a loop body transfers the program control to the exit point of the loop

• If loops are nested, **break** statement applies to the **innermost** loop (whose body it appears in)

• **break** statement can only be used inside **loops** or **switch** statements

• In case of **nested** statements:
  • Example: loop inside another loop, switch inside loop, loop inside switch
  • **break** applies to the **innermost** statement (which directly contains break)
Break Statement – Example

- Finding sum of prime numbers between 10 and 100

- It is a good idea to pay attention to values of variables after the loop finishes

```c
for (i = 10; i <= 100; i++)
{
    /* check if i is prime */
    for (j = 2; j <= sqrt(i); j++)
        if (i % j == 0) /* i is not prime */
            break;
    if (j > sqrt(i)) /* i is prime */
        sum += i;
}
```
Continue Statement

• **continue** statement causes skipping of the remaining statements in a loop body
  • while, do-while loops: loop continuation condition is evaluated next
  • for loops: re-initialization expression is evaluated next
• If loops are nested, **continue** statement applies to the **innermost** loop (whose body it appears in)
• **continue** statement can only be used inside **loops**
• When **continue** statement is used, a **while** loop may not be equivalent to the **for** loop
Continue Statement – Examples

Sum of integers from 1 to n
(Integers divisible by 5 should not be included in the sum)

For loop is not equivalent to while loop when we use continue statement in the body

It is possible to avoid most uses of continue statement by using an appropriate if statement

```c
for (i = 1; i <= n; i++)
    if (i % 5 != 0)
        sum += i;
```

```c
i = 1;
while (i <= n)
    { 
        if (i % 5 == 0) continue;
        sum += i;
        i++;
    }
```
Break and Continue – Example

- Counting number of tab characters in user input

```c
count = 0;
while ( (c = getcharO) != EOF)
{
    if (c != '\t' && c != '\n') continue;
    if (c == '\n') break;
    if (c == '\t') count++;
}
```
Null Statement

• A solitary semicolon ;
• Can be placed wherever a regular statement can appear
• Execution of null statement has no effect
  • May seem useless
  • Sometimes necessary because C language syntax requires a statement
  • Example: Counting number of characters in input

```c
for ( count=0 ; getchar() != EOF ; count++ )
;
```

• Good programming practice: place null statement on a line by itself
Comma Operator

• Syntax: $\text{expr}_1, \text{expr}_2$ (can be extended: $e_1, e_2, e_3, \ldots, e_N$)

• Evaluation:
  • Expressions are evaluated from left to right
  • The type and value of whole expression is the type and value of right operand (left operand is evaluated only for the side effect, its value is discarded)

• Purpose: combining two related expressions into one expression, making programs more compact

• Comma operator has the lowest precedence of all operators
  • Can safely be used to turn any list of expressions into a single expression
/* Assume */
int i ;
float x ;

/* Expressions */
int i = 1
float x = i + 1

/* Combined into single expression */
i = 1, x = i + 1
/* type: float, value: 2 */

/* Swapping values of x and y */
temp = x ;
x = y ;
y = temp ;

/* Swapping values of x and y */
temp = x , x = y , y = temp ;
/* Evaluated left to right */

while ((c = getc()) != EOF) putc(c);

while (c =getc(), c != EOF) putc(c);

while ( printf( "Enter two integers : " ),
scanf ( "%d %d", &n1 , &n2 ) != EOF )
{
    printf( "Sum = %d\n", n1 + n2 ) ;
}
Chapter 5 – Functions

• Overview
• Function Definition
• Function Call
• Function Prototypes
• Block Structure
• External Variables
• Storage Classes
• Separate Compilation and Data Abstraction
• Recursion
Overview

- C applications usually consist of many functions in multiple files
- Functions (and logically grouped functions in files) can individually be
  - Developed, tested, debugged, compiled, maintained
- Sample function: cubesum
  - Type: int
  - Function name: cubesum
  - Parameters ("formal parameters")
  - Function body
    - Local variables
    - Function statements

```c
int cubesum(int number)
{
    int sum, residue, digit;
    residue = number;
    sum = 0;
    do
    {
        digit = residue % 10; /* rightmost digit */
        sum += digit * digit * digit;
        residue /= 10; /* after removing this digit */
    } while (residue != 0);
    return sum;
}
```
Narcissistic Cubes (uses “cubesum”)

```c
#include <stdio.h>
define MAX 5

int main(void)
{
    int i, count;
    int cubesum(int number);

    for(i=1, count=0; count < MAX; i++)
    {
        if (i == cubesum(i))
        {
            printf("%d\n", i);
            count++;
        }
        return 0;
    }
}
```
Function Definition – Syntax (Review)

```
ReturnType FunctionName ( Type1 Parameter1Name, Type2 Parameter2Name )
{
    Variable Definitions, External Variable/Function Declarations
    Type var1Name, var2Name, var3Name, ... ;
    extern Type globalVar1Name, globalVar2Name, ... ;
    extern ReturnType otherFunctionName ( Type1, Type2, Type3 ) ;

    Statements
    Expression1 ;
    Expression2 ;
    ...
}
```
Function Definition - Notes

• If return type is omitted, it is taken to be int
• Function names obey the identifier naming rules that we saw before
• Local variables only exist, and accessible, within their functions
• Local variable names shadow identically named variables outside
• Function parameters are also (initialized) local variables of the function
• Function statements are any valid C statements
• Function execution terminates when
  • Either, execution reaches the closing brace of the function body
  • Or, a return statement is encountered
Return Statement

• Syntax: \( \text{return expression;} \) or \( \text{return;} \)

• In first form; value of the expression is returned to the calling function

• If type of expression does not match type of function, it is converted

- If function return type is void:
  • Cannot use first form (error)

- If function return type is not void:
  • If second form is used, value returned is unpredictable

• More than one return statement can exist in a function body
Function Call (Expression)

• Syntax:  functionName ( argumentList )
• Arguments are also called “actual arguments”
• Use empty parentheses if function does not take any parameters
• Arguments are separated with commas
  • Note: This is not the comma operator! It is a syntactic entity
  • If you need to use comma operator, use parentheses

• Definitions:
  • Calling function (caller): Function which calls another function
  • Called function (callee): Function which is called
• The type/value of the function-call expression is the type/return value of the called function
• Calling function may choose to ignore the value returned by called function

f((a=0, a+=1), b)
A compiled and linked executable file stores:

- **TEXT** (machine code representation of the functions)
- **DATA** (initialized global/static data)
- **BSS** (uninitialized global/static data)
- Also dynamic library links, debugging info, symbol tables, etc.

When an executable file is started (double click or type its name):

- OS loader loads TEXT, DATA and BSS in their own memory segments
- Additional regions are allocated: **CL Arguments & Env. Variables**, **HEAP**, **STACK**

- **HEAP** segment houses dynamically allocated memory regions
- **STACK** segment is used to manage control flow of the application
Memory Layout of C Programs

https://www.geeksforgeeks.org/memory-layout-of-c-program/
Call Stack and Stack Frames (Activation Records)

• Each function call adds a stack frame to the STACK segment

• A stack frame (activation record) usually contains
  • Argument values passed to the function
  • Function’s own local variables
  • Return address
  • Storage for intermediate/temporary results (of long expressions)
  • Possibly some other housekeeping data, like
    • A pointer to the previous frame (and/or the size of the current frame)
    • Some saved register values that need to be restored after the function call
    • Some storage for return value (return value may also be returned through registers)
    • etc.

• Exact stack frame format is compiler/system dependent
Stack Frame – Sample Images from Internet

![Stack Frame Diagram]

- **Function Arguments**
  - Return address to the caller
  - Saved frame pointer (RBP)
  - Local variable storage
  - Buffer storage for temp data
  - Saved registers by the callee - RBX, RSI, RDI

- **Current stack pointer** - RSP (stack top)
- **Current frame pointer** - RBP

- **Stack growth**
  - %ebp
  - %esp

- **Stack Frame**
  - Argument 1
  - Argument 2
  - Return Address
  - Saved %ebp
  - Local Var 1
  - Local Var 2
  - Local Var 3
Stack Frame – Sample Images from Internet

```
Corresponding C code:

int function(int p1, int p2, int p3)
{
    int A, B, C;
    ...
}
```
### Memory Layout and Call Stack Example

<table>
<thead>
<tr>
<th>Entity</th>
<th>Hexadecimal</th>
<th>Decimal</th>
<th>Differences</th>
<th>Print Order</th>
<th>Memory Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Where argv[0] points to</td>
<td>0x7ffff0c3bb6f</td>
<td>140737232747375</td>
<td>503</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Address of argv[0]</td>
<td>0x7ffff0c3b978</td>
<td>140737232746872</td>
<td>248</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Address of heap2</td>
<td>0x7ffff0c3b880</td>
<td>140737232746624</td>
<td>8</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>Address of heap1</td>
<td>0x7ffff0c3b878</td>
<td>140737232746616</td>
<td>4</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Address of localMain</td>
<td>0x7ffff0c3b874</td>
<td>140737232746612</td>
<td>8</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Address of argc</td>
<td>0x7ffff0c3b886</td>
<td>140737232746604</td>
<td>12</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Address of argv</td>
<td>0x7ffff0c3b960</td>
<td>140737232745392</td>
<td>28</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Address of local1</td>
<td>0x7ffff0c3b844</td>
<td>140737232745656</td>
<td>8</td>
<td>20</td>
<td>8</td>
</tr>
<tr>
<td>Address of parameter1</td>
<td>0x7ffff0c3b83c</td>
<td>140737232745656</td>
<td>40</td>
<td>19</td>
<td>9</td>
</tr>
<tr>
<td>Address of local2</td>
<td>0x7ffff0c3b814</td>
<td>140737232745616</td>
<td>8</td>
<td>23</td>
<td>10</td>
</tr>
<tr>
<td>Address of parameter2</td>
<td>0x7ffff0c3b80c</td>
<td>140737232745608</td>
<td>40</td>
<td>22</td>
<td>11</td>
</tr>
<tr>
<td>Address of local3</td>
<td>0x7ffff0c3b800</td>
<td>140737232745584</td>
<td>8</td>
<td>26</td>
<td>12</td>
</tr>
<tr>
<td>Address of parameter3</td>
<td>0x7ffff0c3b7dc</td>
<td>140737232745460</td>
<td>140737208965036</td>
<td>25</td>
<td>13</td>
</tr>
</tbody>
</table>

Where heap2 points to | 0x16ae030 | 23781424 | 32 | 10 | 14 |
Where heap1 points to | 0x16ae010 | 23781392 | 1748160 | 9 | 15 |
Address of global1 | 0x602068 | 6299752 | 4 | 11 | 16 |
Address of staticLocalMain | 0x602064 | 6299748 | 4 | 6 | 17 |
Address of staticLocal3 | 0x602060 | 6299744 | 4 | 27 | 18 |
Address of staticLocal2 | 0x60205c | 6299740 | 4 | 24 | 19 |
Address of staticLocal1 | 0x602058 | 6299736 | 4 | 21 | 20 |
Address of staticGlobal2 | 0x602054 | 6299732 | 4 | 14 | 21 |
Address of staticGlobal1 | 0x602050 | 6299728 | 4 | 13 | 22 |
Address of global2 | 0x60204c | 6299724 | 2103519 | 12 | 23 |
Address of main | 0x40076d | 4196205 | 129 | 18 | 24 |
Address of function3 | 0x4006ec | 4196676 | 139 | 17 | 25 |
Address of function2 | 0x400661 | 4195937 | 139 | 16 | 26 |
Address of function1 | 0x4005d6 | 4195058 | 4195058 | 15 | 27 |

### Linux Prompt
```
Linux Prompt> size a.out
   text    data     bss     dec     hex filename
   3549     568      40    4157    103d a.out
```
Call by Value (Pass by Value)

- C only supports call-by-value mechanism for passing data to functions
- Some other languages also support call-by-reference (aliasing)
- Classical example:
  - Swapping values (not possible)
  - If programmer needs to swap:
    - Pass addresses (pointers), indirection
    - Or, use external variables

```c
#include <stdio.h>

int main(void)
{
    int i = 1, j = 2;
    void exchange (int, int);

    printf("main : i = %d j = %d\n", i, j);
    exchange(i, j);
    printf("main : i = %d j = %d\n", i, j);
    return 0;
}

void exchange(int i, int j)
{
    int t;

    t = i, i = j, j = t;
    printf("exchange: i = %d j = %d\n", i, j);
```
Function Prototypes (Declarations)

• Syntax: \texttt{returnType functionName ( listOfParameterTypes ) ;}

• Parameter names are not necessary
  • If used, they can also be different from the names used in definition

• The arguments are converted to the declared types of the parameters
  • As if by assignment
  • No need for explicit type casting

• Function \texttt{definition} also serves as a \texttt{declaration} for the \texttt{rest of the file}

• An undeclared function will be assumed to have the return type of \texttt{int}
  • Suggestion: Always declare functions before using them
  • If compiler is told the prototype, it can help programmer catch type mismatches
Block Structure

- Syntax is similar to the body of a function
  - Variable definitions/declarations and statements within curly braces
  - `{ variable definitions/declarations statements }`

- Blocks can be nested

- **Scope** of a variable defined in a block is until the end of the block

- **Block variable names** shadow identically named variables in outer blocks

- Technically, blocks do not act as stack frames
  - They allow to define new variables (coming after some statements)
  - They also provide **scope** for the variables defined within
  - From a programmer’s view, you may assume that they act as a stack frame
External Variables

• External variables
  • Defined outside of any function
  • Available to all functions below in the same file
  • If defined before any function definition: sometimes called “global variable”
    • Available to all function in the file

• Usages
  • Passing data between functions
  • A function that needs to return more than 1 values can use global variables

• Disadvantages
  • Hides relationship between different parts of a program (hard to understand)
  • Changes in one part of the program may affect some other part in unexpected ways
  • Using external variables in a function destroys generality of the function
Scope and Lifetime of Entities

• **Scope**: Part of the program where a name can be used/accessed
  • **External entities** (non-static global variables and non-static functions)
    • From the point of their *definition/declaration* until the end of the file
    • From the point of their *declaration* in a block, until the end of the block
    • Note: static external entities are NOT visible outside of the file they are defined in
  • **Function parameters**
    • Inside the function
  • **Local variables and static local variables** (in *functions* and/or *blocks*)
    • Inside the function/block

• **Lifetime**: The duration that the entity is kept in memory
  • **External entities** and all **static entities**: Equal to the lifetime of the *program*
  • **Function parameters** and **local variables**: Equal to the lifetime of the *stack frame*
  • **Dynamically allocated data**: Until they are *freed explicitly* (or *program termination*)
Storage Classes

• A variable belongs to one of two storage classes: **Automatic**, **Static**

• Do not confuse **static storage class** with the keyword **static** in C language

• **Storage class** is about **lifetime** of the variable

• **Automatic variables:**
  • Allocated upon entry to a code segment
  • Deallocated upon exit from that segment
  • They can only be defined within a block
  • Syntax: `auto type variableName ;`
  • If storage class specifier (**auto**) is not used: variables in blocks taken to be auto
  • Can also be initialized when defined: `auto type variableName = expression ;`
    • Expression is evaluated each time the block is entered
    • Function parameters, function calls, etc. can also be used in the expression (can be dynamic)
  • If not explicitly initialized, initial value of an automatic variable is **undefined**
Storage Classes

• **Static** variables:
  • Allocated at the beginning of the program execution
  • Storage remains allocated until the program terminates
  • **External variables** are always **static** (don’t confuse with **static** keyword!)
  • Syntax for static variables within blocks: `static type variableName ;`
  • Can also be **initialized** when defined: `static type variableName = expression ;`
    • Expression is evaluated **just for the first time (and once)** when the block is entered
    • Expression can only be a constant expression (should be compile-time-evaluatatable)
  • If not explicitly initialized, initial value of a static variable is **zero**
  • Values of static variables are retained across calls to the function they are defined in
    • Lifetime = Lifetime of program
    • Scope = To the end of the block from the point they are defined
Storage Classes - Example

```c
#include <stdio.h>

int main(void)
{
    int i;
    void incr(void);

    for (i = 0; i < 3; i++) incr();
    return 0;
}

void incr(void)
{
    int auto_i = 0;
    static int static_i = 0;

    printf("auto = %d static = %d\n", auto_i++, static_i++);
}
```

• Output:
  
  auto = 0  static = 0  
  auto = 0  static = 1  
  auto = 0  static = 2  

Separate Compilation and Data Abstraction

• Separate Compilation
  • Large programs are usually divided into logically related functions, called modules
  • Modules usually reside in separate files
  • Each file can be compiled, tested, maintained, etc. individually

• Abstract Data Objects
  • Equivalent to the concept of classes in object-oriented languages
  • Implementation details in a module/file can be hidden from the outside world
  • Only selected functions (interface) can be made available to the other files
  • Local changes in the module will not affect other files (other parts of application)
extern Declaration

• Remember difference between definition and declaration:
  • Definition allocates storage for the entity, also serves as a declaration
  • Declaration just lets the compiler know about the type and name of entity for the rest of the scope
• Every entity should be defined only once, somewhere
• Only external (top level) entities can be declared
  • Local variables cannot be declared, they can only be defined
• An external entity can be declared many times in same/different scopes
• For functions, use of extern keyword in declaration is optional
• External variables can be initialized only at definition (not declaration)
• Related declarations are usually grouped in header files for easy use
  • Other files just #include the header files, and then get linked with the implementations
Information Hiding

• Function and external variable definitions, prefixed with static keyword can only be accessed in the file they are defined
  • Their scope is limited from the point of their definition to the end of the file
  • Other files cannot access them (they are invisible to other files)

• Remember that:
  • static keyword applied to an external variable limits its scope to its file
  • static keyword applied to a local variable makes its lifetime equal to the lifetime of the application
Recursion

• A function may also call itself (same mechanism as calling any function)
  • This can be **direct**: call occurs inside the function itself
  • Or **indirect**: function calls other function, ..., last function calls first function

Recursion in computer science is a technique where **solution** to a **problem** depends on
  **solution(s)** to **smaller instance(s)** of the same **problem**

This approach can be applied to many types of problems, and it is one of the central ideas of computer science (used in Dynamic Programming, some AI algorithms)

Recursion - Example

```c
int factorial(int n)
{
    if (n == 0)
        return 1;  /* termination condition */
    else
        return n * factorial(n-1); /* recurse */
}
```

```c
int factorial(int n)
{
    return n == 0 ? 1 : n * factorial(n-1);
}
```

```c
int factorial(int n)
{
    if (n > 0)
        return n * factorial(n-1);
    else if (n == 0)
        return 1;
    else
    {
        printf("no factorial for negative values\n");
        return -1;
    }
}
```
Recursion – Notes from CENG 433

( Quick coverage of Recursion presentation from CENG 443 )
Recursion vs. Iteration

• In theory, any recursive function can be transformed into an iterative function
  • Recursive solutions are generally more understandable
  • Iterative solutions are generally faster
    • They eliminate function calls
    • They may eliminate duplicate computations

```c
int fib(int n)
{
    return n == 0 || n == 1 ? 1 : fib(n-1) + fib(n-2);
}
```
Chapter 6 - Arrays

• Basics of Arrays
  • Array Definition/Declaration
  • Accessing Array Elements
  • Array Initialization

• Arrays as Function Arguments
  • Passing Array Elements as Arguments
  • Passing Arrays as Arguments
Basics of Arrays

• C language provides a capability similar to **subscripting** in **mathematics**
  • We can process **collections of related data items** easily
  • There is no need to invent new variable names for each item
  • We can refer to each item by the same **group name** and a **subscript** (index)

• Example
  • In ***TicTacToe*** (HW1), we created separate variables for board cells: c0, c1, ... , c8
  • In ***DnDW*** (HW1), we created separate variables for each room: r0, r1, ... , r9
  • What would happen if we had 10000 cells/rooms in the game?
    • With arrays we can define/create all of the cells easily: `char cells[10000] ;`
  • Arrays provide us much practical ways to work with a set of related data items
    • With array notation we can access any cell/room by using an index: `rooms[395] = EMPTY ;`
    • Also, we can create **loops** to iterate over all cells/rooms instead of long lists of statements:
      `for ( i = 0 ; i < 10000 ; i++ ) { rooms[i] = EMPTY ; }`
Definitions and Array Basics

• **Array**: An ordered finite collection of data items (same type each)

• **Elements**: Individual data items in an array

• **Array indices (subscripts)** start at 0 (zero)!
  - Example: room[0] is the first element, room[9] is the 10th element

• **Properties of an array:**
  - **Type**: data type of its elements
  - **Location** (in memory): location of its first element
  - **Length**: number of elements in the array
  - **Size**: length of the array × size of an element (in Java, “size” refers to length)

• An array’s name is the address of the (1st byte of the) first element of the array

• Array elements are placed **contiguously** in the memory
  - There is no gap and/or padding between any two consecutive array elements
Examples

- One-dimensional (single-subscripted, linear) array (vector): elements
- Two-dimensional (double-subscripted) array (matrix): rows, columns
- Three-dimensional array: rows, columns, levels
- ... (C language allows any number of dimensions)
- Note: In computer’s memory, all data items are stored like a vector
  - See example code: VectorAsMatrix.c
Array Definition

• Syntax:  
  type arrayName [ expr_1 ] [ expr_2 ] ... [ expr_n ] ;

• Expressions should be compile-time computable and integral-valued
  • We can use constants and/or symbolic-constants in the expressions
  • Example:  int midtermGrades [ CENG_STUDENTS + ERASMUS_STUDENTS ] ;

• Arrays are just like regular variables
  • Arrays can be local or global
  • Arrays can be defined as auto or static
  • Global arrays can be declared using extern keyword
    • When declaring, you may omit the first dimension, but all other dimensions are necessary in order to calculate the correct position of an array entity
    • Example:  extern int testScores [ ] [ NUMBER_OF_TESTS ] ;
Accessing Array Elements

• `arrayName[index]` is equivalent to `*(arrayName + index)`
  • Basically, the box (of one of the elements) is calculated and returned to us
  • After that point, we can use this box just like any regular variable:
    Read the value in it, write some value in it, use it in an expression, etc.
    Example: `matrix[2][3] = --matrix[0][0] + matrix[1][5];`
  • See example code: `ArraysAndAddresses.c`
  • Multiple dimensions are handled similarly
    • Example: `array[index1][index2]` is equivalent to `*((*(array + index1) + index2))`

• `index` can be any integral-valued expression

• C does NOT check the `index` to lie within the array’s memory region
  • We need to be very careful to NOT access outside of array’s boundaries
Copying Arrays into Other Arrays

- We cannot assign arrays to other arrays directly
- We need to copy element-by-element using a loop, or we need to use a library function
- Later in the course, we will also see another way to copy arrays (put array in a struct, and assign structs)

```c
int from[5] = { 1, 2, 3, 4, 5 }; int to [5];
to = from;    /* Illegal */
to[] = from[];    /* Illegal */

for ( i = 0 ; i < 5 ; i++ )    /* OK */
{
  to[i] = from[i];
}

#include <string.h>
memcp
```
Array Initialization

• Syntax: List of initializers, enclosed in braces, separated by commas


• Multi-dimensional arrays
  • Inner braces are optional

```c
int test_score[5][4] =
{
    { 95, 98, 97, 96 },
    { 79, 89, 79, 85 },
    { 99, 98, 99, 99 },
    { 90, 89, 83, 86 },
    { 75, 72, 79, 69 }
};
```
Array Initialization – Special Rules

• If number of initializers < number of elements → remaining elements are set to 0 (zero)
  • Useful trick for local arrays: int array[1000] = {0};

```c
int score[5] = { 41, 97, 91 };
int test_score[5][4] = {
    { 0, 98 },
    { 79, 89, 79 },
    { 0, 0, 0, 99 },
    { 90 },
};
```

• If number of initializers > number of elements → ERROR

```c
int score[5] = { 41, 97, 91, 0, 0 };
int test_score[5][4] = {
    { 0, 98, 0, 0 },
    { 79, 89, 79, 0 },
    { 0, 0, 0, 99 },
    { 90, 0, 0, 0 },
    {0, 0, 0, 0 }
};
```
Array Initialization – Special Rules

• If there are initializers, you can omit the array length
  • Compiler can count the number of initializers to determine length

```c
float sqroot[] = { 0, 1.0, 1.414, 1.732, 2.0 };
```

```c
float sqroot[5] = { 0, 1.0, 1.414, 1.732, 2.0 };;
```

• Character arrays may also be initialized with string constants
  • Terminating '\0' is also copied to the array

```c
char os [4] = "AIX";
char computer[] = "sierra";
```

```c
char os [ 4 ] = { 'A', 'I', 'X', '\0' };
char computer[7] = { 's', 'i', 'e', 'r', 'r', 'a', '\0' };
```
Array Initialization – Special Rules

It is also possible to initialize a character array with an explicit length specification by a string constant with exactly as many characters. The array does not receive the terminating '\0' in that case. For example,

```c
char os [3] = "AIX";
```

is equivalent to

```c
char os[3] = { 'A', 'I', 'X' };
```

However, the string constant must not have more characters than the length of the character array being initialized. Thus, it is an error to initialize `os` as

```c
char os [3] = "ultrix";
```
Passing Array Elements as Arguments

- **Array elements** can be passed to functions just like regular variables
  - **Pass-by-value** mechanism (a *copy* of the value is passed to the function)
  - If type is different than parameter type, **conversion rules** apply as usual

```c
void printNumber ( long int x )
{
    printf( "%l\n" , x ) ;
}

void main ( void )
{
    int numbers[5] = { 1 , 2 , 3 , 4 , 5 } ;
    int i ;

    for ( i = 0 ; i < 5 ; i++ )
    {
        printNumber( numbers[i] ) ;
    }
}
```
Passing Arrays as Arguments

• Passing array’s name to a function = passing array’s starting address
  • Pass-by-value mechanism (a copy of the address is passed to the function)
  • In function prototype, indicating the first dimension is optional

```c
void add5 ( int array[3] )
{
    int i ;
    for ( i = 0 ; i < 3 ; i++ )
    {
        array[i] += 5 ;
    }
}

void main ( void )
{
    int numbers[3] = { 1 , 2 , 3 } ;
    add5( numbers ) ;
}
```

```c
void add5 ( int array[], int length )
{
    int i ;
    for ( i = 0 ; i < length ; i++ )
    {
        array[i] += 5 ;
    }
}

void main ( void )
{
    int num1[4] = { 1 , 2 , 3 , 4 } ;
    int num2[2] = { 8 , 7 } ;
    add5( num1 , 4 ) ; add5( num2 , 2 ) ;
}
```
Passing Arrays as Arguments

When declaring a multi-dimensional array as a parameter, it is necessary to specify the lengths of all but the first dimension. Thus, the prototype of a function to compute the cube root of every element of a two-dimensional array can be written as

```c
void matrix_cuberoott(double xx[10][5]);
```

or as

```c
void matrix_cuberoott(double xx[][5], int rows);
```

but not as

```c
void matrix_cuberoott(double xx[][],
                          int rows, int cols);
```
Array Practice – Self Study Suggestions

• Great and classical examples in textbook (pages 189 – 202), you really SHOULD study them

• It would be educative to rewrite TicTacToe and DnDW using arrays
  • Code will be much shorter and much easier to understand

• Coding vector/matrix operations (from linear algebra) would be very good practice, engineers always need linear algebra libraries

• Building grid/maze/board games would be fun, and would provide lots of array (especially 2D array) practice

• Coding various searching and sorting algorithms would provide good practice and will also help you in future courses (like CENG 213, 315)