3 Functions



Why Learning Functions

- 1. With Sequential, Branching and Looping, you will be able to build programs for simple applications. However, for more complex applications, your programs may be long and certain code may be repeated in the program.
- 2. Functions aim to group specific tasks, so that code will not be repeated. It also helps to improve your program readability and efficiency.
- 3. In this chapter, we discuss the concepts on functions.

Functions

- Function Definition
- Function Prototypes
- Function Flow
- Parameter Passing: Call by Value
- Storage Scope of Variables
- Functional Decomposition

2

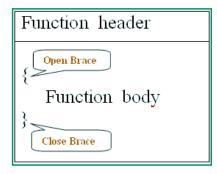


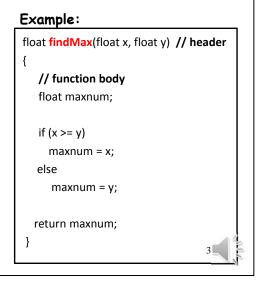
Functions

1. Here, we start by discussing function definition.

Function Definition

- A **function** is a <u>self-contained unit of code</u> to carry out a specific task, e.g. <u>printf()</u>, <u>sqrt()</u>.
- A function consists of
 - a header
 - an opening curly brace
 - a function body
 - a closing curly brace





Function Definition

- A function is a self-contained unit of code to carry out a specific task, e.g. printf(), sqrt().
- 2. Each function definition consists of a function header and a function body.
- 3. The function body contains the code, which specifies the actions of the function, and the local data used by the function.
- 4. An example is illustrated in the **findMax()** function, which has the function header and function body.

Function Header

Return_type Function_name (Parameter_list)

• Function_name

 specifies the name given to the function. Try to give a meaningful name to the function.

Parameter_list

- specifies a list of parameters which contain the data that are passed in by the calling function.

• Return_type

 specifies the type of the data to be returned to the calling function.

4



Function Header

- The function header has the following format: Return_type Function_name(Parameter_list)
- **2. Function_name** specifies the name given to the function.
- **3.** Parameter_list specifies a list of parameters which contain the data that are passed in by the calling function.
- **4. Return_type** specifies the type of the data to be returned to the calling function.

Function Header: Parameter List

- Parameters define the data passed into the function.
- A function can have <u>no</u> parameter, <u>one</u> parameter or <u>many</u> parameters.

type parameterName[, type parameterName]

Example: float **findMaximum**(float x, float y)

- Each parameter has:
 - parameter name
 - data type (such as int, char, etc.) of the parameter
- The function assumes that these parameter inputs will be supplied to the function when they are being called.

5



Function Header: Parameter List

- 1. Parameters define the data passed into the function.
- 2. The **Parameter_list** can be **void** or a list of declarations for variables called parameters: **type parameterName[, type parameterName]**
- 4. A function can have no parameter, one parameter or many parameters.
- 5. Each parameter has a parameter name and data type of the parameter (such as int, char, etc.).
- 6. The function assumes that these parameter inputs will be supplied to the function when they are being called.

Function Header: Return_type

- **Return Type** is the data type <u>returned from</u> the function, it can be int, float, char, void, or nothing.
 - int -- the function will <u>return</u> a value of the type int.
 - float -- the function will <u>return</u> a value of the type float
 - void -- the function will <u>not return</u> any value.

```
void hello_n_times(int n)
{
   int count;
   for (count = 0; count < n; count++)
        printf("hello\n");
   /* no return statement */
}</pre>
```

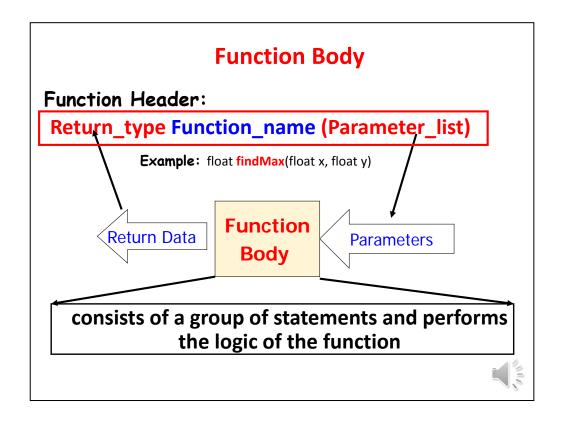
nothing – if defined with no type, the default type is int.

```
successor(int num) /* i.e. int successor(int num) */ {
    return num + 1; /* has a return statement */
}
```



Function Header: Return type

- Return_type is the data type of the value returned by the function, it can be int, float, char, void, etc.
- 2. The **return** statement is used for functions that return a value.
- 3. The syntax for the **return** statement is **return** (**expression**);
- 4. In the function **hello_n_times()**, when the return type is **void**, the function will not return any value. It prints a string "**Hello**" to the screen the number of times specified by the parameter **n**, which is defined to be of type **int**.
- 5. If nothing is specified for **Return_type** of a function header, i.e. when a function is defined with no type, for example, in the **successor**() function, then the default type **int** is used for that function. It means the function will return an integer value. Another example is the **main()** function.

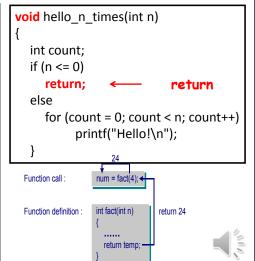


Function Body

- 1. The function body consists of a group of statements.
- 2. The statements are executed when the function is called.
- 3. The variables declared inside the function body are called **local** variables and are only known within the function.
- 4. The main purpose of function body is to perform the logic of the function.

Multiple Return Statements

 The return statement may appear in <u>any place</u> or in <u>more</u> <u>than one place</u> inside the function body.



Multiple Return Statements

- 1. The **return** statement terminates the execution of the function and passes the control to the calling function.
- 2. The **return** statement may appear in any place or in more than one place inside the function body.
- 3. In the program, the function **fact()** has **return** statements in various locations in the function body. If **n** is less than 0, then an error message is displayed, and the control is returned to the calling function. If **n** equals to 0, then the function returns a value 1. If **n** is greater than 0, then the factorial of **n** is evaluated using a **for** loop, and the result is returned.
- 4. A type void function may also have a return statement to terminate the function. This is illustrated in the function hello_n_times. However, it must not have a return expression. If the function does not have a return statement, then the control will be passed back to the calling function when the closing brace of the function body is encountered.
- 5. Sometimes, it is not necessary to use the value returned by a function. This is illustrated in the use of the functions such as printf() and scanf(). The printf() statement returns a value of type int that counts the number of characters printed, whereas the scanf() statement returns the number of items that are successfully read. However, if we do not require this information, we do not need to use the return value returned by these functions.

Function: Examples Compute Grade: [char findGrade(float marks) { char grade; // variable /* function body */ if (marks >= 50) grade = 'P'; else grade = 'F'; return grade; }

Function: findGrade()

- 1. The function **findGrade()** expects a parameter of type **float** and returns a value of type **char**.
- 2. The parameter marks is only accessible within the function findGrade().
- 3. There is one variable **grade** defined in the function **findGrade()**. The variable is a local variable and can only be accessed within the function.
- 4. The variable is created when the function is called, and destroyed when the function ends.

Function: Examples

Compute Grade:

```
char findGrade(float marks) {
    char grade; // variable

    /* function body */
    if (marks >= 50)
        grade = 'P';
    else
        grade = 'F';
    return grade;
}
```

Compute Circle Area:

```
float
areaOfCircle(float radius) {
  const float pi = 3.14;
  float area;

/* function body */
  area = pi*radius*radius;
  return area;
}
```



Function: areaOfCircle()

10

- 1. The function areaOfCircle() expects one parameter of type float. It returns a value of type float.
- 2. The parameter radius is only accessible within the function areaOfCircle().
- 3. A local variable **area** is also declared in the function. This variable is only accessible within the function **areaOfCircle()**.
- 4. It is also created when the function is called, and will be destroyed when the function exits.

Functions

- Function Definition
- Function Prototypes
- Function Flow
- Parameter Passing: Call by Value
- Storage Scope of Variables
- Functional Decomposition

11



Functions

1. Here, we discuss function prototypes.

Function Prototypes

- <u>function prototype</u> used to declare a function. It provides the information about
 - 1. the **return type** of the function

Example:

float findMax(float x, float y);

- 2. the **name** of the function
- 3. the **number** and **types** of the arguments
- The declaration may be the same as the function header but terminated by a **semicolon**.
- Two ways to declare parameters in the parameter list of function prototype:

(1) void hello_n_times(int n);

// with parameter name n

Or to be declared without giving the parameter names:

(2) double distance(double, double);

// no parameter names



Function Prototypes

- 1. We need to declare a function before using it in the **main()** function or other functions.
- A function declaration is called a function prototype. It provides the
 information about the type of the function, the name of the function, and the
 number and types of the arguments.
- 3. The declaration is the same as the function header but terminated by a semicolon. For example, **void hello_n_times(int n)**;
- 4. The function prototype can also be declared without giving the parameter names. For example, **double distance(double, double)**;
- 5. Function prototypes enable the compiler to ensure that functions are being called properly. The compiler will check whether the number of arguments and the type of the arguments of the function call match with the parameters used in the function definition. Warning messages will be given if the number of arguments is different.

Function Prototypes: Where to declare it?

- The declaration has to be done **before** the function is called:
 - (1) **before** the main() header
 - (2) inside the main() body or
 - (3) inside any function which uses it

13



Function Prototypes: Where to declare it?

- 1. A function must be declared before it is actually called.
- 2. It can be declared either before the **main()** header, inside the **main()** body or inside any function which uses it.

Function Prototypes: Before the main()

- The declaration has to be done **before** the function is called:
 - (1) before the main() header
 - (2) inside the main() body or
 - (3) inside any function which uses it

Before the main():

```
#include <stdio.h>
int factorial(int n); // function prototype

int main()
{    int x;
    x = factorial(5); // use factorial() here
}

int factorial(int n) /* function definition*/
{
....
}
```



Function Prototype: Before the main()

- 1. If the function prototype is placed before the **main()** function and at the beginning of the program, it makes the function available to all the functions in the program.
- 2. In this example program, the function **factorial()** is declared outside the **main()**. Therefore, it can be used by all the functions in the program.

Function Prototype: Inside the main() • The declaration has to be done **before** the function is called: (1) before the main() header (2) inside the main() body or (3) inside any function which uses it Before the main(): Inside the main(): #include <stdio.h> #include <stdio.h> int factorial(int n); // function prototype int main() int x; int main() int fact(int); // function prototype { int x; $x = \frac{fact}{5}$; // use fact() here x = factorial(5); // use factorial() here int factorial(int n) /* function definition* // function definition int fact(int n) 15

Function Prototype: Inside the main()

- In the second example program, the function fact() is declared inside the main() function.
- 2. This makes the function callable only within the main() function.

Functions

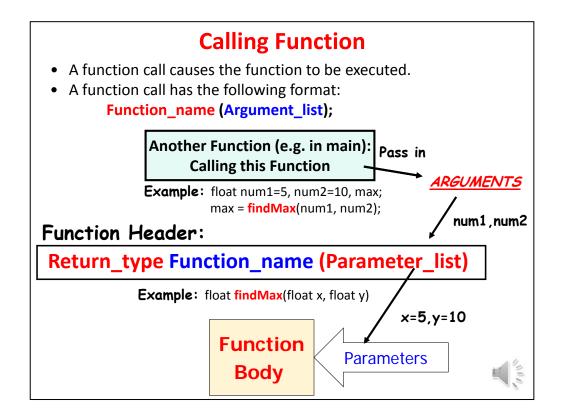
- Function Definition
- Function Prototypes
- Function Flow
- Parameter Passing: Call by Value
- Storage Scope of Variables
- Functional Decomposition

16



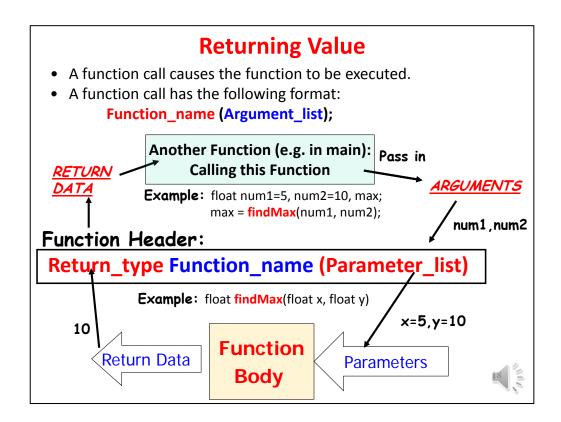
Functions

1. Here, we discuss function flow.



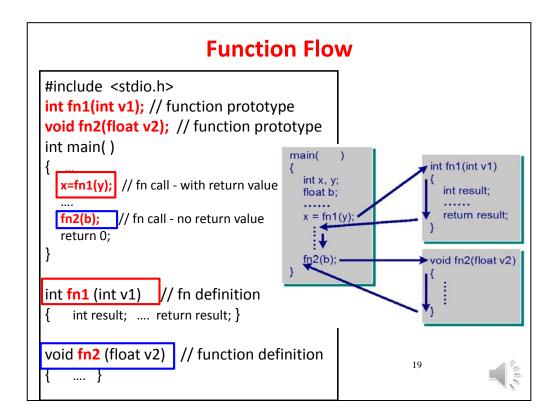
Calling Function

- 1. A function is executed when it is called.
- 2. A function call has the following format: function name(argument list);
- 3. A function can be called by using the function name followed by a list of arguments. For example, num1 and num2 in the findMax() function.
- 4. Function arguments can be constants, variables or expressions.
- 5. For the function **findMax()**, as **num1**=5, **num2**=10, the values 5 and 10 will be passed to the parameters **x** and **y** of the function respectively.



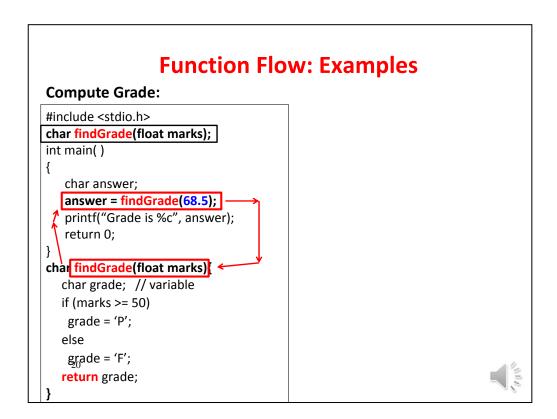
Returning Value

- 1. A function such as **void hello_n_times(int n)**; does not need to return a value.
- 2. However, a function can also return a value as shown in the function **findMax()** which computes the maximum value of **x** and **y** (i.e. 5 and 10 respectively), and returns the value of 10 to the calling function.
- 3. The returned value is then assigned to the variable **max**.



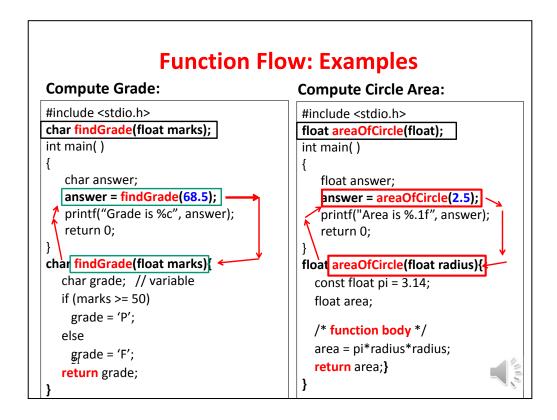
Function Flow

- 1. In the program, the **main()** function will start the execution.
- When the function fn1() is called, the program transfers the control to the fn1() function which then starts execution. As fn1() will return a value to the calling function, the statements in the function body of fn1() will be executed until a return statement is encountered.
- 3. Control is then transferred to the main() function. The value of the variable result will be assigned to the variable x in main(). The next statement after the function call then starts execution.
- 4. When the second function fn2() is called. The control is then transferred to fn2(). The function will execute until the end of the function body. Control will then be transferred to the main() function.



Function Flow: findGrade()

- In the program, the main() function calls the function findGrade(). When the statement: answer = findGrade(68.5); is executed, it calls the function findGrade().
- Control is then transferred to the function findGrade(). Information is passed between the calling function and the called function through the argument. In this case, the function receives one argument with the value of 68.5. It is assigned to the corresponding parameter in the function definition to compute the grade.
- 3. When the execution of statements in the function body encounters the **return** statement, the control is then transferred to the **main()** function, and the statement just after the function call in **main()** will continue to execute.
- 4. The name for parameter needs not be the same as function argument. However, the number of arguments and the data type of the arguments must be the same as parameters defined in function definition. In the program, the argument **68.5** must correspond to the parameter **marks** in the function call.
- 5. Note that the function prototype is declared as: **float findGrade(float marks)**; which is placed at the beginning of the program before the **main()** function.



Function Flow: areaOfCircle()

- In the program, the main() function calls the function areaOfCircle(). When the statement: answer = areaOfCircle(2.5); is executed, it calls the function areaOfCircle().
- Control is then transferred to the function areaOfCircle(). Information is passed between the calling function and the called function through the argument. In this case, the function receives one argument with the value of 2.5. It is assigned to the corresponding parameter in the function definition to compute the area of the circle.
- 3. When the execution of statements in the function body encounters the **return** statement, the control is then transferred to the **main()** function, and the statement just after the function call in **main()** will continue to execute.
- 4. Note that the function prototype is declared as: **float areaOfCircle(float)**; which is placed at the beginning of the program before the **main()** function.

Functions

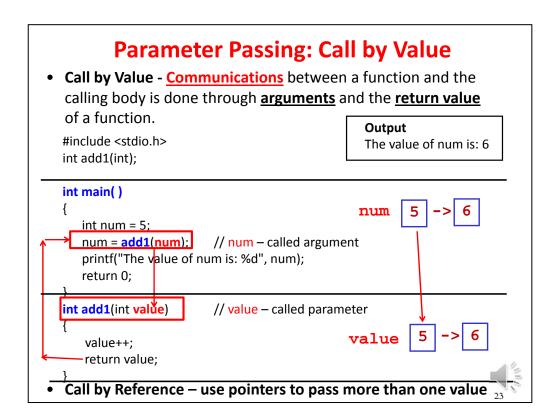
- Function Definition
- Function Prototypes
- Function Flow
- Parameter Passing: Call by Value
- Storage Scope of Variables
- Functional Decomposition

22



Functions

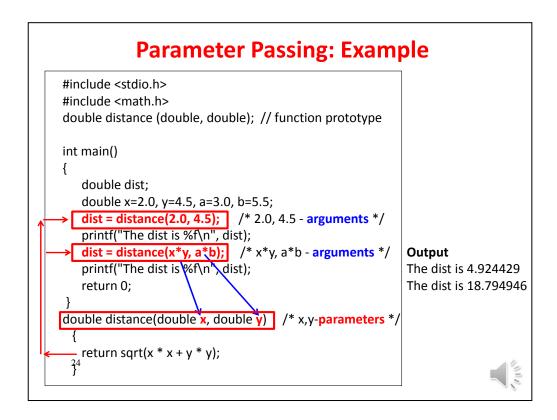
1. Here, we discuss parameter passing using call by value.



Parameter Passing: Call by Value

- Communications between a called function and the calling function is through arguments. The called function then performs the task based on the received argument values. The called function can also return a value to the calling function.
- Parameter passing between functions may be performed in two ways: call by
 value and call by reference. In call by value, the parameters must be declared in
 the function definition as regular variables. The arguments in function calls can
 be constants, variables or expressions.
- 3. When the function is called, the parameters hold a **copy** of the arguments locally. Therefore, any changes to the parameters in a function are done on the copy of the arguments.
- 4. This is illustrated in the function **add1()**. In the **main()** function, the variable **num** is assigned with 5. **num** is used as the argument when calling the function **add1()**.
- 5. The value in **num** is passed to the parameter value in the function **add1()**. The variable **value** is a local variable in the function. Then the variable **value** is incremented by 1. Finally, the value in the variable **value** is returned to the calling **main()** function which then assigns the returned value to the variable **num**.
- 6. In any programs, there are two ways for a called function to return values to the

- calling function. The first way is to use the **return** statement as shown in the function **add1()**. However, this can only be used when only **a single value** needs to be returned from a function.
- 7. If **two or more values** need to be passed back from a called function, we need to use another approach called **call by reference** using pointers. We will discuss call by reference in the chapter on Pointers.



Parameter Passing: Example

- 1. In the program, it calls the function **distance()**. When the statement: **dist = distance(2.0, 4.5)**; is executed, it calls the function **distance()** in **main()**. The arguments **2.0** and **4.5** in the **main()** function correspond respectively to the parameters **x** and **y** in the function call.
- 2. Control is then transferred to the function **distance()**. Information is passed between the calling function and the called function through arguments. In this case, the function receives two arguments with values of 2.0 and 4.5. They are assigned to the corresponding parameters in the function definition.
- 3. In addition, we can also use expression as an argument in the function as shown in the following statement: dist = distance(x*y, a*b); The arguments x*y and a*b in the main() function correspond respectively to the parameters x and y in the function call.
- 4. When the execution of statements in the function body encounters the **return** statement, the control is then transferred to the **main()** function, and the statement just after the function call in **main()** will continue to execute.
- 5. Note that the names for parameters need not be the same as function arguments. However, the number of arguments and the data type of the arguments must be the same as parameters defined in function definition.

Function Calling Another Function #include <stdio.h> int max3(int, int, int); /* function prototypes */ int max2(int, int); int main() Output int x, y, z; input three integers => 7 4 9 printf("input three integers => "); Find the max in 7, 4 and 9 scanf("%d %d %d", &x, &y, &z); Find the max of 7 and 4 printf("Maximum of the 3 is %d\n", max3(x, y, z)); Find the max of 4 and 9 return 0; Find the max of 7 and 9 int max3(int i, int j, int k) Maximum of the 3 is 9 printf("Find the max in %d, %d and \%d\n", i, j, k); return max2(max2(i, j), max2(j, k)); int max2(int h, int k) printf("Find the max of %d and %d\n", h, k); return h > k ? h : k;

Function Calling Another Function

- A function may be called by main() or another function through call by value. In the program, the function max2() specifies two parameters, h and k, of type int, and receives two function arguments from the calling function. The values of the arguments are then stored in the memory locations of the two parameters, h and k. The function then compares their values, and returns the larger value to the calling function.
- 2. The function max3() specifies three parameters, i, j and k, and receives the function arguments from the calling function, and compares their values to determine the largest value.
- The function max3() calls the function max2() to compare two values at a time and returns the maximum value: return max2(max2(i,j), max2(j,k));
- 4. Here, the function max2() is specified in the function max2() itself. The returned value from the called function max2() will be used again as arguments in the same function max2(). The maximum value is then returned to the calling function.

Functions

- Function Definition
- Function Prototypes
- Function Flow
- Parameter Passing: Call by Value
- Storage Scope of Variables
- Functional Decomposition

26



Functions

1. Here, we discuss the storage scope of variables.

Scope of Variables in a Function

- Scope of a variable
 - the sections of code that can use the variable. In other words, the variable is visible in that section.
- Variables declared in a function is ONLY visible within that function. We call it **block scope**.
- Example below: variables radius, pi and area are <u>NOT</u> visible outside this function.

Scope of Variables in a Function

- 1. The scope of a variable determines the sections of the code that can use the variable. In other words, the variable is visible in that section of code.
- 2. Variables declared in a function is ONLY visible within that function. We call it block scope.
- 3. In the function areaOfCircle(), the variables radius, pi and area are not visible outside the function areaOfCircle().

Local and Global Variables

- Local variables:
 - They are variables defined **inside** a function.
- Global variables:
 - They are variables defined **outside** the functions.
- Should global variables be used in your programs?
 - Advantages of using global variables:
 - simplest way of communication between functions
 - efficiency
 - **Disadvantages** of using global variables:
 - less readable program
 - · more difficult to debug and modify
- Strongly <u>discouraged</u> to use global variables instead you should use <u>parameter passing between functions</u> to achieve the same effect. So that <u>errors</u> will be <u>localized</u> within each function for easy debugging.

Local and Global Variables

- 1. Local variables are variables defined inside a function. They have block scope. They can be accessed only within the function. They cannot be accessed by other functions. Local variables are created when the function is invoked, and destroyed after the complete execution of the function.
- 2. Global variables are variables defined outside the functions. They have file (or program) scope. Thus, global variables are visible to all the functions that are defined following its declaration.
- 3. The advantages of global variables in programs are that global variables are the simplest way of communication between functions and they are efficient. The disadvantages of programs using global variables are that they are less readable and more difficult to debug and modify as any functions in the program can change the value of the global variables.
- 4. Therefore, it is a **good programming practice** to use local variables, and use parameter passing between functions for communication between functions. In this way, the value of each variable in the function is protected.
- 5. It is strongly discouraged to use global variables. Instead you should use parameter passing between functions to achieve the same effect. So that errors can be localized within each function for easy debugging.

```
Local and Global Variables: Example
#include <stdio.h>
int g_var = 5;
                            // global variable - has file scope
int fn1(int, int);
float expn(float);
int main() {
   char reply;
                            // local - these two variables are only
   int num;
                            // known inside main() function - block scope
}
int fn1(int x, int y) {
                            // local x,y - formal parameters are only
                            // known inside this function – block scope
   float fnum;
                            // local - these two variables are known
   int temp;
                            // in this function only – block scope
   g_var += 10;
float expn(float n) {
                            // local - this variable is known in expn()
   float temp;
                            // block scope
}
```

Local and Global Variables: Example

- 1. In the example program, the global variable **g_var** is declared outside the **main()** function. Global variables will have the file scope.
- 2. The variables **fnum** and **temp** are local variables which will have block scope and are only visible inside the function fn1.

- Using the static keyword
 Static Variables
 - The duration of a static variable is fixed.
 - Static variables are created at the <u>start</u> of the program and are destroyed only at the <u>end</u> of program execution. That is, they exist throughout program execution once they are created.

Note:

Automatic variable (local)

 the variable disappears after each function execution.

Static variable (like global)

 the variable stays until end of program execution.

```
Output
Static variable: 1
Automatic variable: 1
Static variable: 2
Automatic variable: 1
Static variable: 3
Automatic variable: 3
```

Static Variables

- 1. A **static** variable may be defined inside or outside a function's body. The duration of a static variable is fixed.
- 2. Static variables are created at the start of the program and are destroyed only at the end of program execution.
- 3. We can define static variables **inside** a function's body by changing an automatic variable using the keyword **static**.
- 4. If a static variable is defined and initialized, it is then initialized once when the storage is allocated. If a static variable is defined, but not initialized, it will be initialized to zero by the compiler. The initialization is done when the storage is allocated. If the static variable is defined inside a function's body, then the variable is only visible by the block containing the variable.
- 5. Static variables are very useful when we need to write functions that retain values between functions.
- 6. We may use global variables to achieve the same purpose. However, static variables are preferable as they are local variables to the functions, and the shortcomings of global variables can be avoided.
- 7. In the example program, the static variable **static_var** is declared and initialized only once when storage is allocated at the start of the program. The value of the static variable is retained for different calls to **function()**. The value stored in the static variable will remain until the end of program execution. This is

different from the auto_var variable as it is created and initialized every time when function() is called. However, since the static variable static_var is declared inside function(), it is only visible inside function().

Functions

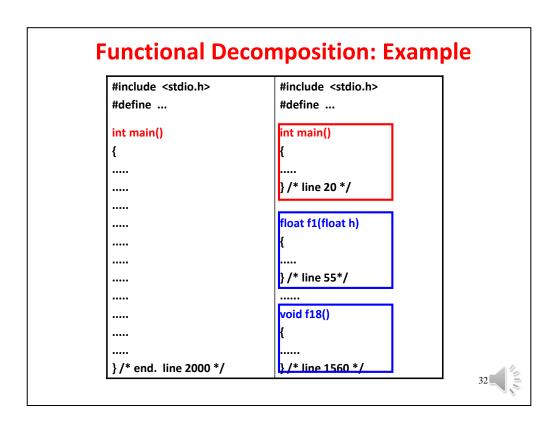
- Function Definition
- Function Prototypes
- Function Flow
- Parameter Passing: Call by Value
- Storage Scope of Variables
- Functional Decomposition

31



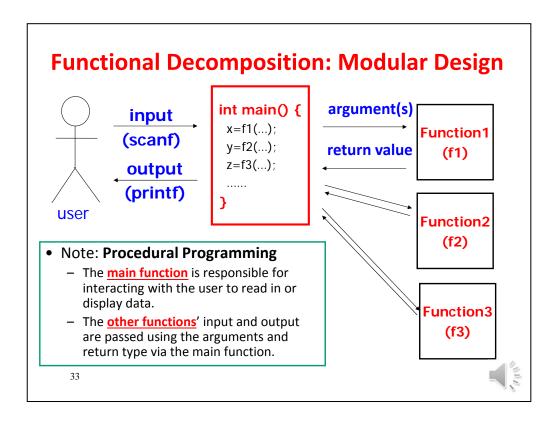
Functions

1. Here, we discuss functional decomposition.



Functional Decomposition: Example

- 1. In the original program, the **main()** function contains about 2000 lines of code which is difficult to read and debug.
- 2. Functional decomposition basically means the top-down stepwise refinement technique that uses the divide-and-conquer strategy. It starts with the high level description of the program and decomposes the program into successively smaller components until we arrive at a set of suitably sized functions (or algorithms). We design the code for the individual functions using stepwise refinement. At each level of refinement, we are only concerned with what the lower level functions will do.
- 3. Functional decomposition produces smaller functions that are easier to understand. Smaller functions promote software reusability. In general, functions should be small, so that they can be developed and tested separately. They should also be independent of each other.
- 4. In the example program, it is decomposed into a number of smaller functions. The **main()** function will start program execution and call other functions to perform different required operations.



Functional Decomposition: Modular Design

- 1. Using the functional decomposition and top-down stepwise refinement technique, a problem is broken up into a number of smaller subproblems or functions. We then develop the algorithms for the functions. These functions can then be implemented using a programming language such as C. These functions are also called **modules**. This approach of designing programs as functional modules is called **modular design**. The functions or modules should be small and self-contained, so that they can be developed and tested separately. They should also be independent of each other.
- 2. There are a number of advantages for modular design. Modular programs are easier to write and debug, since they can be developed and tested separately. Another advantage is that modular programs can be developed by different programmers as each programmer can work on a single module of the program independently. Moreover, a library of modules can be developed which can then be reused in other programs that require the same implementation. This can reduce program development time and enhances program reliability. Therefore, modular design can simplify program development significantly.
- 3. When writing C programs, we use procedural programming technique, which is different from object-oriented programming paradigm used in Python, Java and C++.
- 4. In procedural programming, a typical structure of a program consists of the

main function and other functions for solving a problem. Generally, if the functions are still quite complex, then they can be divided further into smaller functions. And each function should not be longer than a page.



Thank you

1. Thank you.