Chapter - 16
File Input/ Output

I/O Packages

There are 3 different I/O packages available to the C++ programmer:

- The C++ streams package. This package is used for most I/O.
- The unbuffered I/O package. Used primarily for large file I/O and other special operations.
- The C `stdio` package. Many older, C-- programs use this package. (A C-- program is a C program that has been updated to compile with a C++ compiler, but uses none of the new features of the new language.)

  This package is useful for some special operations.

C++ I/O

C++ provides four standard class variables for standard I/O.

Variable	Use
`std::cin`	Console in (standard input)
`std::cout`	Console output (standard output)
`std::cerr`	Console error (standard error)
`std::clog`	Console log

Normally `std::cin` reads from the keyboard, and `std::cout`, `cerr`, and `clog` go to the screen.

Most operating systems allow I/O redirection

```
my_prog <file.in
my_prog >file.out
```

File I/O

File I/O is accomplished using the `<fstream.h>` package.
Input is done with the `ifstream` class and output with the `ofstream` class.

Example:

```
ifstream data_file;    // File for reading the data from
data_file.open("numbers.dat");
for (i = 0; i < 100; ++i)
    data_file >> data_array[i];
data_file.close();
```

Open can be consolidated with the constructor.

```
ifstream data_file("numbers.dat");
```

The close will automatically be done by the destructor.
To check for open errors:

```
if (data.file.bad()) {
    cerr << "Unable to open numbers.dat\n";
    exit (8);
}
```

Reading Number

```
/*****************************************************
 * read -- read in 100 numbers and sum them          *
 *                                                   *
 * Numbers are in the file "numbers.dat"             *
 *                                                   *
 * Warning: No check is made for a file with less than  *
 * 100 numbers in it.                                *
 *****************************************************/
#include <iostream>
#include <fstream>
#include <stdlib.h>
int main(){
    const int DATA_SIZE = 100;  // Number of items in the data
    int data_array[DATA_SIZE];  // The data
    ifstream data_file("numbers.dat"); // The input file
    int i;                      // Loop counter

    if (data_file.bad()) {
        std::cerr << "Error: Could not open numbers.dat\n";
        exit (8);
    }

    for (i = 0; i < DATA_SIZE; ++i)
        data_file >> data_array[i];


    int total;  // Total of the numbers


    total = 0;
    for (i = 0; i < DATA_SIZE; ++i)
        total += data_array[i];


    std::cout << "Total of all the numbers is " << total << '\n';
    return (0);
}
```

`getline` member function

```
istream &getline(char *buffer, int len, char delim = '\n')
```

Parameters:

`buffer` A buffer to store the data that has been read.

`len` Length of the buffer in bytes. The function reads up to len-1 bytes of data into the buffer. (One byte is reserved for the terminating null character: '\0') This parameter is usually `sizeof(buffer)`.

`delim` The character used to signal end of line.

The `sizeof` operator returns the size (in bytes) of a variable or type.

Output files

Example:

```
ofstream out_file("out.dat");
```

Full constructor:

```
ofstream::ofstream(const char *name, int mode=ios::out,
                   int prot = filebuf::openprot);
```

Parameters:

`name` The name of the file.

`mode` A set of flags or'ed together that determine the open mode. The flag `ios::out` is required for output files.

`prot` File protection. This is an operating system dependent value that determines the protection mode for the file. On UNIX the protection defaults to 0644 (read/write owner, group read, others read). For MS-DOS/Windows this defaults to 0 (Normal file).

Open Flag

Flag	Meaning
`std::ios::app`	Append data to the end of the output file.
`std::ios::ate`	Goto the end of file when opened.
`std::ios::in`	Open for input (must be supplied to opens for ifstream variables)
`std::ios::out`	Open file for output (must be supplied to ofstream opens).
`std::ios::binary`	Binary file (if not present, the file is opened as an ASCII file).
`std::ios::trunc`	Discard contents of existing file when opening for write.
`std::ios::nocreate`	Fail if the file does not exist (output files only. Input files always fail if there is no file.)
`std::ios::noreplace`	Do not overwrite existing file. If a file exists, cause the open to fail.

Example:

```
ofsteam out_file("data.new",
      ios::out|ios::binary|ios::nocreate|ios::app);
```

Conversion Routines

To print a number such as 567 you must turn it into three characters “5”, “6” and “7”.

The << operator converts numbers to characters and writes them. Conversion is controlled by a number of flags set by the `setf` and `unsetf` member function calls.

```
file_var.setf(flags);   // Set flags
file_var.unsetf(flags); // Clear flags
```

Conversion Flags

Flag	Meaning
`std::ios::skipws`	Skip leading whitespace characters on input.
`std::ios::left`	Output is left justified
`std::ios::right`	Output is right justified
`std::ios::internal`	Numeric output is padded by inserting a fill character between the sign or base character and the number itself.
`std::ios::dec`	Output numbers in base 10, decimal format
`std::ios::oct`	Output number in base 8, octal format.
`std::ios::hex`	Output numbers in base 16, hexadecimal format.
`std::ios::showbase`	Print out a base indicator at the beginning of each number. For example: hexadecimal numbers are proceeded with a “0x”.
`std::ios::showpoint`	Show a decimal point for all floating point numbers whether or not it’s needed
`std::ios::uppercase`	When converting hexadecimal numbers show the digits A-F as upper case.
`std::ios::showpos`	Put a plus sign before all positive numbers.
`std::ios::scientific`	Convert all floating point numbers to scientific notation on output.
`std::ios::fixed`	Convert all floating point numbers to fixed point on output.
`std::ios::unitbuf`	Buffer output. (More on this later)
`std::ios::stdio`	Flush stream after each output.

Example

```
number = 0x3FF;
std::cout << "Dec: " << number << '\n';
std::cout.setf(ios::hex);
std::cout << "Hex " << number << '\n';
std::cout.setf(ios::dec);
```

Output:

```
Dex 1023
Hex 3ff
```

Other Conversion Control Functions

Controlling the width of the output:

```
int file_var.width(int size);
```

Controlling the precision of floating point (number of digits after the point).

```
int file_var.precision(int digits);
```

Setting the fill character:

```
char file_var.fill(char pad);
```

I/O Manipulators

```
#include <iostream>
#include <iomanip.h>

number = 0x3FF;
std::cout << "Number is " << hex <<
                 number << dec << '\n';
```

I/O Manipulators

Manipulator	Description
`std::setiosflags(long flags)`	Set selected conversion flags
`std::resetiosflags(long flags)`	Reset selected flags.
`std::dec`	Output numbers in decimal format.
`std::hex`	Output numbers in hexadecimal format.
`std::oct`	Output numbers in octal format.
`std::setbase(int base)`	Set conversion base to 8, 10, or 16. Sort of a generalized `dec`, `hex`, `oct`.
`std::setw(int width)`	Set the width of the output.
`std::setprecision(int precision)`	Set the precision of floating point output
`std::setfill(char ch)`	Set the fill character
`std::ws`	Skip whitespace on input
`std::endl`	Output end of line (‘\n’)
`std::ends`	Output end of string (‘\0’)
`std::flush`	Force any buffered output out.

I/O Example

```
#include <iostream>
#include <iomanip>
int main()
{
    int number = 12;    // A number to output
    float real = 12.34; // A real number

    std::cout << "12345678901234567890123456789O\n"; //  ruler
    std::cout << number << "<-\n";
    std::cout << std::setw(5) << number << "<-\n";
    std::cout << std::setw(5) << std::setfill('*') << number << "<-\n";
    std::cout << std::setiosflags(std::ios::showpos|std::ios::left) << std::setw(5) <<
            number << "<-\n";

    std::cout << real << "<-\n";
    std::cout << std::setprecision(1) << std::setiosflags(std::ios::fixed) <<
            real << "<-\n";
    std::cout << std::setiosflags(std::ios::scientific) << real << "<-\n";
    return (0);
}
```

Output:

```
12345678901234567890123456789O
12<-
   12<-
***12<-
+12**<-
12.34<-
12.3<-
1e+01<-
```

Binary and ASCII files

ASCII Files

- Contain characters you can read
- Can be printed directly on the printer
- Take up lots of space
- Portable

Binary files

- Contain the “raw” data
- You can’t read them, they print garbage on the screen if you try to type them.
- Can not be printed directly on a printer.
- Relatively compact.
- Mostly machine dependent.

Binary vs. Character

In C++ we use the notation: ‘1’ to represent the character one. We represent the number as: 1.

The character ‘1’ has the numeric value 49.

To turn characters into numbers, we need to subtract 48 or the value of the char-acter ‘0’.

```
int integer;
char ch;


ch = '5';
integer = ch - 48;
std::cout << "Integer " << integer << '\n';
```

’0’ is 48, you can just subtract ’0’.

```
integer = ch - '0';
```

End of Line Puzzle

In the dark ages, BC (Before Computers) teletypes used <carriage return> <line feed> for end of line.
When computers came into existence storage cost $$$ so some people decided to cut the end of line to one character.

UNIX	Uses <line feed> only for end of line
Apple	Uses <carriage return> only for end of line
MS/DOS	Uses <carriage return><line feed> for end of line.

When reading ASCII files, the end-of-line must be translated into a new line character '\n'. Binary files do not need this translation. Translation of binary files causes problems.

Binary vs. ASCII opens

```
// open ASCII file for reading
ascii_file.open("name", ios::in);

// open binary file for reading
binary_file.open("name", ios::in|ios::binary);
```

We write 128 bytes. DOS gets 129. Why?

```
#include <iostream>
#include <fstream>
#include <stdlib.h>


int main()
{
    int cur_char;   // current character to write
    ofstream out_file; // output file


    out_file.open("test.out", ios::out);
    if (out_file.bad()) {
        (cerr << "Can not open output file\n");
        exit (8);
    }


    for (cur_char = 0; cur_char < 128; ++cur_char) {
        out_file << cur_char;
    }
    return (0);
}
```

Hint: Here is a hex dump of the MS-DOS/Windows file:

Dump of MS-DOS output

```
000:0001 0203 0405 0607 0809 0d0a 0b0c 0d0e
010:0f10 1112 1314 1516 1718 191a 1b1c 1d1e
020:1f20 2122 2324 2526 2728 292a 2b2c 2d2e
030:2f30 3132 3334 3536 3738 393a 3b3c 3d3e
040:3f40 4142 4344 4546 4748 494a 4b4c 4d4e
050:4f50 5152 5354 5556 5758 595a 5b5c 5d5e
060:5f60 6162 6364 6566 6768 696a 6b6c 6d6e
070:6f70 7172 7374 7576 7778 797a 7b7c 7d7e
080:7f
```

Binary Input

```
in_file.read(data_ptr, size);
```

data_ptr	Pointer to a place to put the data.
size	Number of bytes to be read.

Example:

```
struct {
    int     width;
    int     height;
} rectangle;

in_file.read((char *)(&rectangle), sizeof(rectangle));
if (in_file.bad()) {
    cerr << "Unable to read rectangle\n";
    exit (8);
}
if (in_file.gcount() != sizeof(rectangle)) {
    cerr << "Error: Unable to read full rectangle\n";
    cerr << "I/O error of EOF encounterd\n ";
}
```

Binary Output is similar:

```
out_file.write(data_ptr, size);
```

Buffering Problems

When will the output be printed?

```
std::cout << "Starting";
do_step_1();
std::cout << "Step 1 complete";
do_step_2();
std::cout << "Step 2 complete";
do_step_3();
std::cout << "Step 3 complete\n";
```

Print it now:

```
std::cout << "Starting" << std::flush;
do_step_1();
std::cout << "Step 1 complete" << std::flush;
do_step_2();
std::cout << "Step 2 complete" << std::flush;
do_step_3();
std::cout << "Step 3 complete\n" << std::flush;
```

Unbuffered I/O

How to pick up a bunch of paper clips. (Buffered input)

1. Pick up a paper clip in your left hand.
2. Put in your right hand.
3. Repeat the last two steps until the right hand (buffer) is full.
4. Dump the handful in the box.

How to pick up cannon balls (unbuffered I/O)

1. Pick up cannon ball using both hands.
2. Dump it in the box. Be careful to avoid dropping it on your feet.

Buffered I/O is useful for small things. Unbuffered works for larger reads and writes.

Unbuffered I/O routines

```
#include <sys/types.h>
#include <sys/stat.h>
#include <fcntl.h>


#ifdef __MSDOS__         // if we are MS-DOS
#include <io.h>          // Get the MS-
DOS include file for raw i/o
#else /* __MSDOS__ */
#include <unistd.h>      // Get the UNIX include file for raw i
/o
#endif /* __MSDOS__ */


int     file_descriptor;


file_descriptor = open(name, flags);    // existing file
file_descriptor = open(name, flags, mode);//new file
```

file_descriptor

An integer that is used to identify the file for the read, write and close calls. If file descriptor is less than 0 an error occurred.

name Name of the file.

flags Defined in the `fcntl.h` header file.

mode Protection mode for the file. Normally this is 0666 for most files.

Open Flags

Flag	Meaning
O_RDONLY	Open for reading only.
O_WRONL Y	Open for writing only.
O_RDWR	Open for reading and writing.
O_APPEND	Append new data at the end of the file.
O_CREAT	Create file (*mode* file required when this flag present).
O_TRUNC	If the file exists, truncate it to 0 length.
O_EXCL	Fail if file exists.
O_BINARY	Open in binary mode (Older UNIX systems may not have this flag).

Open examples:

```
data_fd = open("data.txt", O_RDONLY);
 out_fd = open("output.dat", O_CREAT|O_WRONLY, 0666);
```

Pre-opened files:

File Number	Description
0	Standard in
1	Standard out
2	Standard error

Read function

```
read_size = read(file_descriptor, buffer, size);
```

read_size The actual number of bytes read. A 0 indicates end of file and a negative number indicates an error.

file_descriptor

File descriptor of an open file.

buffer Pointer to the place to read the data.

size Size of the data to be read. This is the size of the request. The actual number of bytes read may be less that this. (For example, we may run out of data.)

Write and close functions

```
write_size = write(file_descriptor, buffer, size);
```

write_size

Actual number of bytes written. A negative number indicates an error.

file_descriptor

File descriptor of an open file.

buffer Pointer to the data to be written.

```
flag = close(file_descriptor)
```

flag 0 for success, negative for error.

file_descriptor

file_descriptor of an open file.

Copy program

```
/*****************************************
 * copy -- copy one file to another.     *
 *                                       *
 * Usage                                 *
 *      copy <from> <to>                 *
 *                                       *
 * <from> -- the file to copy from       *
 * <to>   -- the file to copy into       *
 *****************************************/
#include <iostream>
#include <stdlib.h>

#include <sys/types.h>
#include <sys/stat.h>
#include <fcntl.h>

#ifdef __MSDOS__         // if we are MS-DOS
#include <io.h>          // Get the MS-DOS include file for raw i/o
#else /* __MSDOS__ */
#include <unistd.h>      // Get the UNIX include file for raw i/o
#endif /* __MSDOS__ */

const int BUFFER_SIZE = (16 * 1024); // use 16k buffers
```

Copy Program

```
int main(int argc, char *argv[]) {
    char  buffer[BUFFER_SIZE];  // buffer for data
    int   in_file;              // input file descriptor
    int   out_file;             // output file descriptor
    int   read_size;            // number of bytes on last read

    if (argc != 3) {
        std::cerr << "Error:Wrong number of arguments\n";
        std::cerr << "Usage is: copy <from> <to>\n";
        exit(8);
    }
    in_file = open(argv[1], O_RDONLY);
    if (in_file < 0) {
        std::cerr << "Error:Unable to open " << argv[1] << '\n';
        exit(8);
    }
    out_file = open(argv[2], O_WRONLY | O_TRUNC | O_CREAT, 0666);
    if (out_file < 0) {
        std::cerr << "Error:Unable to open " << argv[2] << '\n';
        exit(8);
    }
    while (1) {
        read_size = read(in_file, buffer, sizeof(buffer));

        if (read_size == 0)
            break;              // end of file

        if (read_size < 0) {
            std::cerr << "Error:Read error\n";
            exit(8);
        }
        write(out_file, buffer, (unsigned int) read_size);
    }
    close(in_file);
    close(out_file);
    return (0);
}
```

Designing file formats

We need a configuration file for a graph program.
One layout:

height (in inches)
width (in inches)
x lower limit
x upper limit
y lower limit
y upper limit
x scale
y scale

Sample file:

```
10.0
7.0
0
100
30
300
0.5
2.0
```

C Style I/O Routines

File variables:

The declaration for a file variable is:

```
#include <stdio.h>
FILE *file-variable;        /* comment */
```

Open function:

```
file_variable = fopen(name, mode);
```

file-variable

A file variable.

name Actual name of the file (*data.txt*, *temp.dat*, etc.).

mode Indicates if the file is to be read or written. Mode is “w” for writing and “r” for reading.

Close function:

```
status = fclose(file-variable);
```

C's standard files

File	Description
`stdin`	Standard input (open for reading). Equivalent to C++'s `std::cin`
`stdout`	Standard output (open for writing). Equivalent to C++'s `std::cout`
`stderr`	Standard error (open for writing). Equivalent to C++'s `std::cerr`
	There is no C file equivalent to C++'s `std::clog`.

Counting Characters

```
#include <stdio.h>
#include <stdlib.h>     /* ANSI Standard C file */
#include <iostream>

const char FILE_NAME[] = "input.txt";   // Name of the input file

int main()
{
    int  count = 0;  // number of characters seen
    FILE *in_file;   // input file

    int ch;          // character or EOF flag from input

    in_file = fopen(FILE_NAME, "rb");
    if (in_file == NULL) {
        cerr << "Can not open " << FILE_NAME << '\n';
        exit(8);
    }

    while (1) {
        ch = fgetc(in_file);
        if (ch == EOF)
            break;
        ++count;
    }
    std::cout << "Number of characters in " << FILE_NAME << " is " << count << '\n';

    fclose(in_file);
    return (0);
}
```

Note: The function `fgetc` gets a single character or returns the *integer* `EOF` if there are none left.

Other functions

Writing a character:

```
fputc(character, file);
```

Getting a string:

```
string_ptr = fgets(string, size, file);
```

string_ptr Equal to `string` if the read was successful, or NULL if EOF or an error is detected.

string A character array where the function places the string.

size The size of the character array. Fgets reads until it gets a line (com-plete with ending \n) or it reads `size-1` characters. It then ends the string with a null ('\0').

Writing a string:

```
string_ptr = fputs(string, file);
```

C Conversion Routines

Printing

```
printf(format, parameter-1, parameter-2, ...);
```

Example:

```
printf("Hello World\n");
```

prints:

```
Hello World
```

Example:

```
printf(“The answer is %d\n”, answer);
```

Conversion Characters

Conversion	Variable Type
%d	int
%ld	long int
%d	short int
%f	float
%lf	double
%u	unsigned int
%lu	unsigned long int
%u	unsigned short int
%s	char * (string)
%c	char
%o	int (prints octal)
%x	int (prints in hexadecimal)
%e	float (in the form *d.ddd*E+*dd*)

Why does 2+2=5986?

```
#include <stdio.h>
int main() {
    int answer;

    answer = 2 + 2;

    printf("The answer is %d\n");
    return (0);
}
```

Why does 21/7 = 0

```
#include <stdio.h>

int main(){
    float result;

    result = 21.0 / 7.0;
    printf("The result is %d\n", result);
    return (0);
}
```

Printing to a file

```
fprintf(file, format, parameter-1,
   parameter-2, ...);
```

"Printing" to a string:

```
   sprintf(string, format, parameter-1,
                parameter-2, ...);
```

Example:

```
   char string[40];          /* the file name */

   /* current file number for this segment */
   int file_number = 0;

   sprintf(string, "file.%d", file_number);
   ++file_number;
   out_file = fopen(string, "w");
```

Reading data

`number = fscanf(`*`file, format, &parameter-1, . . .`*`);`

number Number of parameters successfully converted.
file A file opened for reading.
format Describes the data to be read.
parameter-1
First parameter to be read.

WARNING:

If you forget to put & in front of each variable for `scanf`, the result can be a "Segmentation violation core dumped" or "Illegal memory access" error. In some cases a random variable or instruction will be modified. This is not common on UNIX machines, but MS-DOS/Windows, with its lack of memory protection, cannot easily detect this problem. On MS-DOS/Windows, omitting & can cause a system crash.

Don't use `fscanf`

The end of line handling in `fscanf` is so weird that it's almost impossible to get the end of line right. To avoid the problems with `fscanf`, don't use it.

Instead use `fgets` / `sscanf`.

```
char line[100];    // Line for data

// Read numbers
fgets(line, sizeof(line), stdin);
sscanf(line, "%d %d", &number1, &number2);
```

C Style Binary I/O

```
read_size = fread(data_ptr, 1, size, file);
```

read_size	Size of the data that was read. If this is less than size, then an end of file or error occurred.
data_ptr	Pointer to the data to be read.
size	Number of bytes to be read.
file	Input file.

```
struct {
        int     width;
        int     height;
} rectangle;

if (fread((char *)(&rectangle), 1,
                  sizeof(rectangle), in_file) !=
    sizeof(rectangle)) {
        std::cerr << "Unable to read rectangle\n";
        exit (8);
}
```

Writing:

```
    write_size = fwrite(data_ptr, 1, size, file);
```