CS 102
Introduction to Programming Using C++

Chapter 8

Files and Streams

Homework

- Written homework
- R8.3, 6, 7, 8, 9, 11 to 13

- Programs
- p. 379. Choose one of P8.1, 2, and 4.
- Also, choose one of P8.14 and 16.

Examining Storage

- Computers have primary storage
  - It is also called memory
  - Primary storage is temporary or volatile
  - Everything happens in primary storage
  - Where is my program while it's running?
- Computers also have secondary storage
  - It is also called permanent storage
  - Examples: Hard drive, CD, DVD, Flash drive

A Deep Look into Storage

- All data is stored in bits
  - This is true for memory, a CD or DVD, a flash drive, a hard drive
- A Bit (BInary digiT) is a switch
- It can be off (0) or on (1)
- A bit is too small to be useful
  - It's so small that the computer can't even access one bit in storage at a time!
- Bits are combined into bytes

Bytes-The Way We View Data

- A byte on a PC is made up of 8 bits
- A byte can hold a letter or a small number or any single character
  - Even a byte is too small for numbers in some cases
- A programmer thinks of storage as a long list of bytes
- Words
  - A word on a PC is made up of two bytes

File Types

- A file is data stored in secondary (permanent) storage
- There are two main types of files
  - Binary files
    - Data files are one type of binary file
    - Program files are another (a.exe, any app)
    - If you look at these files with a text editor, you only see symbols
  - Text files
    - You can look at these files with a text editor
    - A .cpp file is a text file

Working with Data

- We think of data as flowing in a stream
- This is like ducks in the children's pond game
- The ducks float around past a person
- The person can examine them as they come by

- Instead of saying "file", we will say "filestream"
  - The words are often used interchangeably

Using a File

- When you use a file, you have two choices
  - You can read the file (look at it)
    - This is also called input
  - You can write to the file (change it)
    - This is also called output
- We use file I/O to refer to file input and output
- In order to use a file, you have to
  - Open it before doing anything to it
  - Close it after you are finished with it

Opening a File Stream

- The #include for file I/O is #include <fstream>
- Opening a file stream allows you to use it
- You open a stream for reading by

```
ifstream data_file;
data_file.open ("\\cs102\\programs\\lab8-1.dat");
```

- You now can read the data stored in the file

- I will use data_file to refer to the file from now on

Reading the File

- Reading from a file is like reading from the keyboard using cin
- Instead of cin, use the file variable

```
string a_line;
data_file >> a_line;
```

- To read several variables, use

```
data_file >> line_1 >> line_2;
```

- Just like with cin, this reads a word at a time

Reading an Entire Line from a File

- Use the getline() function to read an entire line

```
string input_line;
getline (data_file, input_line);
```

- You can even use getline to cause cin to read an entire line

```
getline (cin, string_variable);
```

A Simple File Program: Echoing a File

```
#include <iostream>
#include <fstream>
using namespace std;
int main ()
{
   ifstream data_file;
   data_file.open ("\\cs102\\programs\\lab8-1.dat");
   string input_line;
   while (getline (data_file, input_line))
      cout << input_line  <<  endl;
   return 0;
}
```

An Example to Process a File

- This is Program chap08/babynames.cpp from p. 356
- Here is background information for the program
- There is a data file of 1000 records, each containing a boy's name and a girl's name
- The data was collected by the Social Security Administration and show names from the 1990s
- Each line (record) contains two identical sub-records
  - There is a rank
  - Then the line contains information about a boy's name
    - The name, the number and percentage of people given that name
  - Then the line contains the same information for girls' names

The Goal

- Looking through the list, someone noticed that there was a wider range of names given to girls than boys
- The person wondered if that were really true
- Or did it just seem to be true?
- If it was true, that would cause each girl's name to get a smaller percentage
- A program was written to test that idea

Implementation

- We will check that idea for the top 50% of the people
- One way to check for the top 50% would be to add up the frequencies (percentages)
  - You stop when they add up to 50%
- Another way would be to start at 50% and subtract the percents
  - In this case, you stop when you get to 0

Examining the Program

- Notice that the program uses the function process_name()
  - The function has a parameter that is a stream parameter
  - The stream parameter is a reference parameter (preceded by an &)
- You have to do this for any stream variable because it gets modified on a read or write
- Let's examine the program!

A Final Note: Typing in a File Name

- If you read a file name as a string, it gets read in as a C++ string
- You then need to convert it to a Cstring
  - You do that with the member function c_str

```
cout  <<  "Type the file name: "
string file_name;
cin >> file_name;
ifstream in_file;
in_file.open (file_name.c_str ());
```

Some Vocabulary

- The combination << in cout is called an insertion operator
- The combination >> in cin is called an extraction operator

Questions?

- Are there any questions?