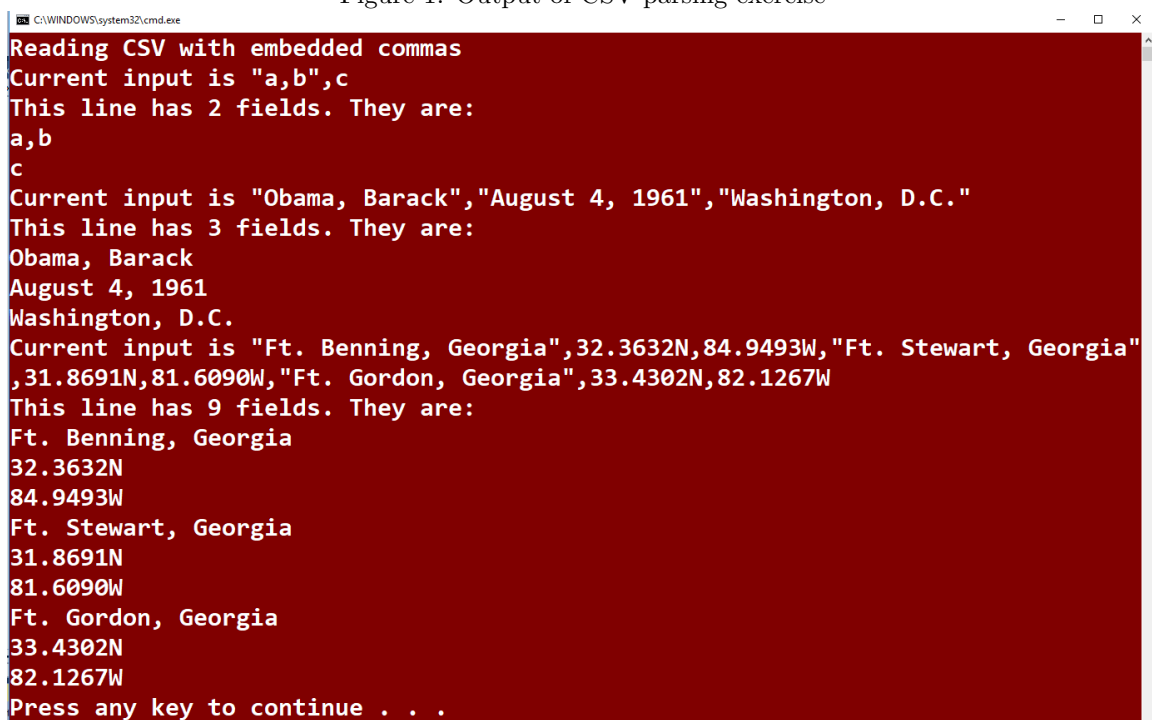# Programming Exercise 17
CSV parsing

C# Step by Step

Programs very commonly read data in from files over a network or contained in a text file in a local directory. Data files are almost always composed of a number of rows, and each row is composed of a number of fields. Lines are (usually) terminated by a *newline* character, usually represented as \n. Fields can be separated by a number of different characters, most commonly *commas* (",") or *tabs*, ("\t"). Sometimes, fields can be separated by *pipes*, ("|") or *colons*, (":"). In this exercise, our data will consist of fields with comma separated values and rows terminated by newlines.

Consider the following program. This contains three strings: input1, input2, input3, contained in a generic List specialized on the <string> datatype. If you try to compile this program, you will find that the method getCSV() has not been implemented. Your assignment is to implement this method. Your output should look similar to Figure 1.

```
class Program
{
    static void Main(string[] args)
    {
        /************************
         * read CSV with embedded commas
         * parse CSV into separate fields and
         * ignore commas within quoted string
         * **********************/
        Console.WriteLine("Reading CSV with embedded commas");
        List<string> myList = new List<string>();
        string input1 = "\"a,b\",c";
        myList.Add(input1);
        string input2 = "\"Obama, Barack\",\"August 4, 1961\",\"Washington, D.C.\"";
        myList.Add(input2);
        string input3 = "\"Ft. Benning, Georgia\",32.3632N,84.9493W," +
            "\"Ft. Stewart, Georgia\",31.8691N,81.6090W," +
            "\"Ft. Gordon, Georgia\",33.4302N,82.1267W";
        myList.Add(input3);

        foreach (string s in myList)
        {
            Console.WriteLine($"Current input is {s}");
            List<string> output = getCSV(s);
            int len = output.Count;
            Console.WriteLine($"This line has {len} fields. They are:");
            foreach (string s1 in output)
                Console.WriteLine(s1);
        }
    }
```

Figure 1: Output of CSV parsing exercise


```
C:\WINDOWS\system32\cmd.exe                                                    —    □    ×

Reading CSV with embedded commas
Current input is "a,b",c
This line has 2 fields. They are:
a,b
c
Current input is "Obama, Barack","August 4, 1961","Washington, D.C."
This line has 3 fields. They are:
Obama, Barack
August 4, 1961
Washington, D.C.
Current input is "Ft. Benning, Georgia",32.3632N,84.9493W,"Ft. Stewart, Georgia"
,31.8691N,81.6090W,"Ft. Gordon, Georgia",33.4302N,82.1267W
This line has 9 fields. They are:
Ft. Benning, Georgia
32.3632N
84.9493W
Ft. Stewart, Georgia
31.8691N
81.6090W
Ft. Gordon, Georgia
33.4302N
82.1267W
Press any key to continue . . .
```