A.4 Installing and Running gawk

gawk is a high-level programming language that, among many other things, can be used to read, manipulate, reformat and output data contained in ASCII text files in a quick, convenient and relatively easy manner. There are several versions of gawk: the one used here is the GNU Project (http://www.gnu.org/) implementation, known as gawk (Dougerty 1996, Robbins 2001).

Version 3.1.3 of gawk, which was released in June 2004, is used throughout this book. Copies of this software, designed for use with either GNU/Linux or Microsoft Windows, are included on the CD-ROM; copies for these and other operating systems can also be obtained from the gawk homepage (http://www.gnu.org/software/gawk/gawk.html).

The remainder of this section describes how to download, install and run gawk on a personal computer running either the GNU/Linux operating system or a recent version of Microsoft Windows.

A.4.1 Instructions for Microsoft Windows

1. a) Installing from the GnuWin32 web site

   It is recommended that you download gawk from the GnuWin32 web site (http://gnuwin32.sourceforge.net/), if you have access to a reasonably fast network connection. To do so, point your web browser at http://gnuwin32.sourceforge.net/packages.html. This web page lists a number of software tools released under GPL and similar “open source” licenses, which are commonly available on GNU/Linux and UNIX systems, that have been ported for use on 32-bit computers running Microsoft Windows.

   Locate the entry for gawk (marked “Gawk 3.1.3 pattern scanning and processing”) and click on the link marked Setup on this line. This will direct you to a list of servers from which the software can be obtained. Choose the server closest to you geographically. You should then be presented with a pop-up window indicating the name of the file that you are about to download (e.g., gawk-3.1.3-2.exe), the size of the file (e.g., 6.35 MB), and three buttons that determine how or whether to proceed. Click on the button marked Run to start installing the software on your computer. You may be asked to confirm, via another pop-up window, that the installation procedure should indeed now run, in which case you should once again click the button marked Run.

   You should now follow the instructions presented by the installation wizard (Figure A.8 and Figure A.9). In most cases, all that is required is to accept the default suggestions, and to click the Next button to proceed to the subsequent step. There are, however, two important exceptions to this general rule. The first is that you will be asked to confirm whether or not you agree to the terms and conditions of the gnuplot software license (given in full in Appendix C). To proceed you must click on the I accept the agreement option (Figure A.8b). The second is that you will be asked to identify where on your system the
Installing and Running the Software

Figure A.9: Further steps involved in installing gawk on Microsoft Windows.

software should be installed; it is recommended that you click on the "Browse button and select the C:\Program Files\gnuWin32 folder (Figure A.8c) if you have permission to do so; otherwise, accept the default location suggested by the installation wizard or choose another folder to which you have access.

b) Installing from the CD-ROM

The program used to set up gawk on your computer (gawk-3.1.3-2.exe) is also provided in the software directory on the CD-ROM (d:\software\gawk -3.1.3-2.exe). Use a file explorer to locate this file, double-click on its icon, and follow the instructions presented by the installation wizard, described above.

2. gawk is a command-line tool (i.e., it does not have a GUI that makes use of point-and-click techniques). To use gawk, therefore, one must run the CLI. In Microsoft Windows XP, the CLI (known in the past as the DOS prompt) can be found via Start→All Programs→Accessories→Command Prompt.

3. Before using gawk, however, Microsoft Windows must be told where to find the gawk utility (gawk.exe). The most convenient way of doing this is to add the folder in which gawk.exe is installed to the environment variable known as the Path. This can be achieved by right-clicking on the My Computer icon in the Start menu (Figure A.10a), and from there selecting the Properties option. Select the Advanced tab in the Properties window, and then click on the Environment Variables button. A new window should appear (Figure A.10b). Search in the lower pane of this window, marked System variables, for the Path entry. Highlight this line in the textbox and then click on the Edit button. This process should produce a further pop-up window entitled Edit System Variable (Figure A.10c). In the text entry box that is marked Variable value, add the following text to the end of the current path: C:\Program Files\gnuWin32\bin (or another path name, if you have installed gawk elsewhere on your system) and then click the OK button.

4. The gawk utility should now be available on your system. To test this, open a command-line window (see above) and type the command gawk. You should be presented with a response similar to that shown in Figure A.11.
A.4.2 Instructions for GNU/Linux

1. A version of awk is included with most major distributions of GNU/Linux, in which case it is possible to use the awk utility by typing awk in a GNU/Linux
   console or xterm. Running awk in this way without further instructions causes
   it to print out basic information on how it should be used (Figure A.12). The
   version of awk available on the system can be checked by typing awk --version
   on the command line.

2. If awk is not already available on your system, you can install it in one of two
   ways described below:
   a) Installing from an RPM

   The procedure used to install awk from an RPM on GNU/Linux systems is
   much the same as that described for gnuplot in Section A.3.2. Search for the
   most recent version of awk for your system on one of the RPM repository
   web sites (e.g., http://rpmfind.net/linux/RPM/) and download this file onto your
   computer. Log onto your computer as super-user (if you don’t know what this
   is, then you probably don’t have the necessary permissions to install new
   software on your system; instead, consult the system administrator) and type

   rpm -i gawk-3.1.x.xxxx.rpm