1 Introduction

In this article, I'm going to discuss a few topics which had their birth in the world of UNIX or are often found in UNIX environments, but are not part of the standard UNIX distribution. Even so, you're likely to find some or all of them on any given UNIX system. In addition, most of these programs are in the public domain, or protected by the GNU software license, so they can be obtained from archive sites on the Internet and used freely, as long as you don't charge a fee for their use.

2 Network News Readers

In the early 1980s (before the advent of such services as America Online and Compuserve), a small group of UNIX users in the Research Triangle of North Carolina experimented with the idea of passing messages among a network of computers, allowing users on any of the computers in the network to read the messages, and potentially contribute their own messages to ongoing discussions. From these small beginnings, the network news, or Usenet system of newsgroups has grown into an international network of machines and users with discussion groups for just about every imaginable (as well as some unimaginable) topics. Unlike bulletin boards, which are typically located on a single computer, there is no "home" for Usenet news; messages are transmitted from machine to machine all across the internet using the Network News Transmission Protocol (NNTP). To read the Usenet news, and to post articles of your own, you must have access to an internet machine which is running appropriate software; ask your system administrator if you have no success with the commands mentioned below.

To view the newsgroups, you'll probably want to use some kind of news reader software. One of the simplest (and oldest) news readers is `rn`; after choosing a newsgroup, each article is displayed and you either read it, or dismiss it and view the next. Alternatively, you can ask for a list of message headers, and pick the ones which interest you. Since `rn` uses a terminal interface, it can be somewhat awkward to pick and choose among the different articles. If you are running Xwindows, the `xrn` program will display a scrollable list of message headers, and will display the articles of your choice in a subwindow. Just about all the commands you could want to use are represented by buttons, so `xrn` is quite easy to use.

For a terminal interface, another alternative is `trn`, the threaded news reader. This program attempts to find articles discussing the same topic (by examining their subject lines), and to present them in a meaningful order. Since the volume of discussion in some news groups is very high, and since you probably won't be interested in every topic which arises, this program can be very helpful. There are other news reader programs available; ask a local guru about ones that might be on your system, if the ones I mentioned don't work out for you.

Even if you never read the network news (and be forewarned that it may be a very time-consuming activity), you may want to peruse the answers to the questions which arise again and again in some of the newsgroups. These questions are assembled into documents known as Frequently Asked Questions (FAQs), and they are posted regularly in each of the groups . If you have access to a World Wide Net browser like `mosaic` or `netscape`, you can view the FAQs for all the newsgroups by opening the URL `http://www.cis.ohio-state.edu/hypertext/faq/usenet`. Many of the FAQs are also available through anonymous ftp at `rtfm.mit.edu` in the directory `/pub/usenet`. It is strongly advised that you read the FAQ for a particular newsgroup before posting a message to the newsgroup; long-time devotees of the newsgroups get very upset when "newbies" ask the same questions over and over again.

3 Perl

Over the years, a variety of techniques have been developed to process files on UNIX systems, to reformat or extract data and produce reports. Until about 10 years ago, most of these techniques were carried out by writing shell scripts (that is, files containing instructions interpretable by a UNIX command shell) which in turn called such UNIX utilities as `awk`, `sed`, `grep`, `sort` and so on. While many programmers developed arcane skills in this enterprise, two major problems emerged. First, a variety of different shell languages and commands needed to be mastered before even fairly simple jobs could be attempted. This created a daunting barrier to most users, who were more interested in solving problems than learning yet another UNIX utility. But beyond that problem, there is a basic inefficiency in using shell scripts to coordinate the activities of different programs, because each time a new program was called, another shell had to be initiated to carry out the task, and then removed when the task was completed.

In response to these difficulties, Larry Wall of the Jet Propulsion Laboratory created `perl` (Practical Extraction and Report Language) in the early 1980s. The basic idea behind `perl` was to create a superset of the functionality of the most popular UNIX utilities, and to provide a coherent environment to carry the tasks out in an efficient manner. In addition, known problems and limitations of the existing tools were eliminated, and the end result is an extremely flexible tool, which can perform an unbelievably wide range of tasks. I routinely tell users at my site to forget about shell scripts and the other UNIX utilities I've mentioned above, and to focus on learning `perl`.

Although `perl` originated on UNIX systems, it has now been ported to a variety of environments. If you'd like to learn more about `perl`, a great place to start is to read the Frequently Asked Questions document (FAQ) from the Usenet newsgroup `comp.lang.perl`, archived on `rtfm.mit.edu` in the file `/pub/usenet/comp.lang.perl`.

4 TeX and LaTeX

Producing publication quality documents, especially with mathematical equations is a difficult task, made more difficult by the fact that often times such documents are produced by word processing programs which either no longer exist, or whose formats are trade secrets, making it very difficult to electronically share formatted documents with colleagues. Even if one particular word processor were adopted as a standard, most such programs internally use unprintable characters in their documents, and this would create difficulty when attempting to transmit the documents through electronic mail, or to make any sense of them without access to the program which created them. One solution to these problems is a text mark-up language know as TeX, and an extension of the language known as LaTeX.

TeX was developed in the 1980s by Donald Knuth of Stanford University as a system "intended for the creation of beautiful books". It is a markup language, as opposed to the more prevalent WYSIWYG (what you see is what you get) word processors which run on personal computers, which means that the object you operate on is an ordinary text file containing special directives (in the case of TeX the directives begin with the backslash character "\")

which instruct it to use different fonts or type sizes, to construct tables and mathematical displays, and to place figures and tables in particular locations. The downside of this approach is that you need to learn the language, as well as an editor to create the input files, and that your file must be processed each time you wish to see the changes and additions you make. The upside is that the document you produce is human readable, easily distributable, and, since it contains all the information needed to produce your document, you can exert a great degree of control over the appearance of the final product. TeX and LaTeX are quickly becoming standards for electronic distribution of manuscripts and many journals now accept or demand TeX input for their submissions.

LaTeX is actually a collection of TeX macros which simplify the production of a number of standard document types. Many, but not all TeX commands are acceptable to LaTeX, but most users find that it's easier to learn either TeX or LaTeX than to try to become an expert at both. Macros for formatting articles may be available from some journal or book publishers, so if you have a particular target in mind, you might want to find out whether they prefer TeX or LaTeX before starting to produce documents.

The FAQ for the `comp.text.tex` newsgroup can be found on `rtfm.mit.edu` in the file `pub/usenet/news.answers/tex-faq`.