Un-Plot-It Automated Digitizing System

DOUGLAS A. SMITH* and JAMES P. ADAMS

Department of Chemistry, The University of Toledo, Toledo, Ohio 43606-3390

Received January 13, 1992

In this age of computer graphics and inexpensive hardware, it is not uncommon to want as many gadgets hooked to your PC as possible, even if the demand for the item is limited. Scanners are such a device for many people, ourselves included. While it would be very nice to have a color, 600-dpi flatbed scanner, which could be used for any number of purposes, the number of times we would use it productively (rather than play with it) cannot justify the cost. There are, however, many times that we would like to be able to take graphical data generated on an instrument or machine not networked to our computers and import that data for manipulation or preparation of presentation-quality graphics in posters and papers.

Un-Plot-It is a hardware and software package for the PC[1] which allows for the digitization of graphs of many different sizes via a special sensor pen that is attached to your Hewlett Packard 7000 Series pen plotter and special software. The software not only allows for the digitization of graphs but also for the analysis and manipulation of the data through regression and smoothing programs. Graphs can be plotted to 8.5 in. × 11 in. paper only. Overall, this package provides a low-cost alternative to the purchasing of an expensive flatbed scanner and digitizing software.

Un-Plot-It runs on the IBM PC, XT, AT, and PS/2 or 100% compatible family of personal computers. System requirements include 256K or more of RAM and an asynchronous serial communications port. To effectively run the program one also needs a hard drive, a monitor with a graphics adapter, and a DB25 female parallel printer port. An HP 7000 Series or compatible plotter is also required. The hardware installation is fairly simple and well documented in the manual, which includes excellent figures and diagrams. Software installation was trivial; one simply copies the entire Un-Plot-It disk into a directory on your hard disk. The first time the program is run, a setup menu appears automatically which asks four questions concerning your hardware configuration. One nice feature of the software is the availability of self-test modes, which allows one to run diagnostics to determine whether or not the hardware is attached to the proper ports. The only drawback we found with the hardware was the lack of a simple on–off switch on the light pen which tracks the plot. In order to turn off the pen to prevent burnout, one must manually unplug it. This becomes burdensome after a while, particularly if your PC and plotter are in a crowded area such as a lab.

For graphs to be effectively scanned by Un-Plot-It they must first be prepared, a process which involves the photocopying of the graph to increase contrast differences and the removal of any stray marks which are close to the lines being scanned. The latter requirement presents some unique problems. Graphs done on grid paper must have the gridlines removed (at least in the immediate area of the lines to be digitized), and graphs which contain intersecting lines must have those intersections removed. This offers some problems for the experimentalist who wishes to use this to store data. Line symbols must be eliminated from all graphs. Many programs for spectroscopy data print peak information next to the respective peaks: these should be removed so that improper scanning will not occur. All of this may be avoided if one uses the manual scan procedure; however, the manual alternative is quite tedious and not as accurate as the automatic mode.

Once the graph has been properly prepared it may be scanned. Depending on the type of plotter used, the paper must be loaded in a specific manner relative to how the scanner defines the x- and y-axis. The manual includes very good figures to aid in this process. Two scan options, manual and automatic, are available to the user, each of which offers some unique challenges. Regardless of which mode is selected the user must manually tell the sensor equipment where the axes points are. This was not a difficult procedure but did take about 5–10 min to do accurately. Settings may be saved but one must remember that not every graph will have its axes placed exactly in the same position on the paper every time. It is advisable therefore to repeat this process for every new graph. The software does a nice job of adjusting for axes which are less than perfectly orthogonal to each other. The user is also required to adjust the sensitivity of the pen apparatus to allow the differentiation between light and dark portions of the graph. This is well documented in the manual and help screens of the software.

On our instrument the manual mode worked flawlessly in obtaining the data. The only drawback to this is that the user is responsible for which points are to be digitized by the apparatus. This means that for every point the user wants digitized he must manually select by hitting "pen down" on the plotter. The difficulty of setting the sensor pen by hand leads us to believe that there is some degree of inaccuracy in the manual procedure. If one uses the automatic mode this problem is avoided since the computer determines how and where to move the sensor pen. However, after following all of the directions religiously we were still unable to get the device to automatically follow the lines on our graph. The problem seemed to be a function of the sensitivity of the pen. We called the customer support number provided twice (this is *not* a toll free number) and received excellent assistance from friendly representatives but still had no luck with the automatic mode. Silk Scientific even offered to replace the apparatus which we had in hand.

One important consideration in purchasing a scanner system is the question of the time required to get usable data from the system. Depending on the complexity of the graph and including preparation time and the time required to scan in the axes, the user is looking at anywhere from 30 min for a simple graph to 1 h for a more complicated one using Un-Plot-It in the automatic mode. The time requirements using the manual mode depend on how many points are chosen, but from our experience would fall into the same range of 30–60 min.

Once the data had been scanned into the computer, we found the rest of the procedure flawless and easy to do. All of the digitized data can be saved to disk and may be called back at any time. The user has the option to perform any of an extensive list of regression analyses on the data or no analysis at all. The program provides options including simple least-squares fit (linear and nonlinear) up through fifth-derivative quintic sexic data smoothing. What was very good and surprising was that the program almost perfectly gave back the equations of the linear and quadratic functions originally plotted. The functions are plotted directly to the screen and may be dumped to the plotter or a separate printer. The smooth functions also worked very nicely on the digitized data.

Overall this part of the operation is by far the easiest and quickest to do. The requesters (pop-up menus which ask questions) are friendly and most helpful. This part of the program also allows for the preparation of the digitized data for future plotting or printing. The options are things such as frames; labels; solid, dashed, or dotted lines; titles; and tick marks. By adjusting the page locations of the graph, one has the capabilities to expand or compress the data on one or both axes. This allows for a nicer presentation. Overall this part of the program was very well thought out and is very powerful to use.

Once the data has been saved it may be then pulled back into memory and edited at any given time. This is a key feature to the program, since erroneous data may be scanned into the data set. A few askew points may be lethal in regression analysis. The recall and edit facility allows the user to remove inconsistent data or even add new data points and then call the data to the screen and view the change in tabular form. The program allows the user to export the tabular data either to a printer or the plotter. Menus are easy to follows, and data dumps are fairly simple and easy to use.

Overall we found Un-Plot-It to be worth the nominal price. For the person who has limited resources and access to a computer with a plotter hooked up to it, it is a worthwhile investment. The setup is fairly simple and the data acquisition, although somewhat lengthy, is not too bad if one has only an occasional graph to digitize. However, this package is not an efficient method for scanning numerous graphs. In our experience, to scan in and analyze 10 graphs would require a full day of work. Even though the program is very powerful, the time required to prepare the graphs, set the axes into memory, and scan the data is far too slow and tedious if one has multiple graphs to digitize.

The analysis program is exceptional and very powerful. For the non-serious digitizer this is an excellent device. If one does not have a plotter hooked up to his computer nor access to one, the cost of this device and a plotter is comparable to a decent flatbed scanner. Thus, our recommendation is positive for the novice but a strong no for the serious digitizer.

The manual well documents the procedures used, and the customer support is an added benefit. The representatives were courteous and knew their instrumentation quite well. Our calls were always handled expediently and satisfactorily. The strong point of this package is its regression analysis program, which offers a great deal of flexibility and power while not compromising its user-friendliness. Thus for an inexpensive alternative to expensive digitizers and software Un-Plot-It is a must.

REFERENCES AND NOTES

(1) Un-Plot-It is available from Silk Scientific, Inc., P.O. Box 533, Orem, UT 84059, for $385.00.

BOOK REVIEWS

Sequential Simplex Optimization. By Frederick H. Walters, Lloyd R. Parker, Jr., Stephen L. Morgan, and Stanley N. Deming. Chemometrics Series. CRC Press, Inc.: Boca Raton, FL. 1991. 325 pp. $49.95 (U.S.) and $59.95 (Foreign). ISBN 0-8493-5894-9.

My first reaction when I was asked to review this book was "Why would anyone want to write an entire book on an optimization method that we usually spend no more than one week discussing in a graduate-level course on optimization?" When I read the "Series Preface" by Steven D. Brown in the front pages, I had a better idea why someone might want to write such a book. Then when I read the "Preface" by the four authors, it became very clear to me that a book crafted around the central concept of sequential simplex optimization could, indeed, provide a very useful set of tools for its intended audience, namely, experimental chemists. Having read the book, my more informed reaction is that this is a well-conceived, well-written book that would be very useful and informative to anyone working in any experimental activity, especially chemistry.

The central thesis of this book is that the sequential simplex method, which was first published in 1962 by W. Spendley, G. R. Hext, and F. R. Himsworth of Imperial Chemical Industries Ltd in England, represents a simple, workable and efficient strategy for experimentation and optimization with chemical processes. The authors devote the first two chapters to motivating the need for optimization and establishing a framework for it. Chapters 3 and 4 describe the basic sequential simplex algorithm and a modified version called the "variable-size simplex algorithm". Chapters 5, 6, and 7 describe several strategies for applying the sequential simplex optimization approach. Chapter 8 discusses various formulations of so-called "desirability functions", which basically provide the experimenter with a rational basis for deciding what system responses to examine and which values of those responses are preferred. These chapters are all very thorough in describing sequential simplex optimization.

Chapter 9 examines the concepts of "experimental design". Although this chapter is well-written and nicely illustrated, its brevity (only 32 pages) is troubling. One would hope that the reader would go beyond the concepts treated here by examining several of the excellent references provided by the authors at the end of the chapter.

Chapter 10 is very brief but provides a long list of references to works which describe the use of sequential simplex optimization in various applications in chemistry. Chapter 11 gives an extensive bibliography of works on the subject of sequential simplex optimization, organized chronologically from 1962 to 1990.

In summary, *Sequential Simplex Optimization* is a well-conceived, thoroughly-researched, and superbly-written book that should provide chemists, as well as other scientists involved in experimental activity, with a highly useful desk reference on experimental optimization. It is clear that anyone applying the tools described in this book will achieve the results of their experimental efforts in less time and with lower costs. They will very likely get better results in the process.

William E. Biles, *University of Louisville*

Artificial Intelligence in Chemistry: Structure Elucidation and Simulation of Organic Reactions. By Zdzislaw Hippe. Translated by Andrzej Przyjazny. Studies in Physical and Theoretical Chemistry Vol. 73. Elsevier: New York. 1991. xii + 278 pp. Hardcover: $137.00. ISBN 0-444-987460-0.

This book is an important contribution to the field of artificial intelligence in chemistry. Even though it is written for the field of chemistry, some chapters are so general that can be used in the development of AI systems in any field.

The author starts with brief descriptions of natural and artificial intelligence and contemporary trends in AI studies. Part I of this book discusses the history of artificial intelligence, trends in general, and applications to chemistry in particular. Throughout these chapters the author discusses problem solving in AI and acknowledge representation and comparison to human memory. The author also discusses backward chaining and the creation of synthetic trees and different search methods in order to find the best (optimum) solution. Chapter 5 specifically deals with the technology of expert systems. The architecture of the expert systems, the tools needed in order to build them, and obviously their limitations are covered briefly.

Parts II and III of this book are more technical and contain practical applications of AI methods in the elucidation of the structure of organic