Computer Software Reviews

RS/1. RESEARCH SYSTEM. BBN Software Products Corporation: Bolt Beranek and Newman, Inc., 10 Fawcett St., Cambridge, MA 02238. List price $995.00 ($495 academic price; maintenance price after 1 year $125).

RS/1 is software for data management and analysis, designed to provide IBM personal computer users with convenient means to archive, manipulate, and display tabular and graphic information. RS/1 is compatible with the IBM XT and AT, or an IBM PC upgraded to the equivalent of the IBM XT. The system requires a minimum of 512K main memory and a 10 Mbyte fixed disk. RS/1 incorporates English-like commands, which provide access to all the facilities of the RS/1, and an advanced PL/1-like language (RPL), which integrates the basic commands with a programming language. These components provide the power of the basic command language with the flexibility of a programming language.

The primary data storage means of the RS/1 are two-dimensional arrays incorporating free format numbers and strings called "Tables". The user can create and edit Tables directly by assigning values to a tabular form on the monitor, or indirectly, by a special command that assigns values to portions of the Table. The user may also copy DOS files (binary or ASCII) to a specific portion of the Table, a very convenient feature when processing large data files created by computerized data acquisition. The dimension of the array is limited only by the available disk space, and the user does not have to declare the dimensions of the table. Tables can be displayed on the screen, dumped to a printer or plotter, or saved by an appropriate one-sentence command. RS/1 provides special kinds of Tables called Models. The segments of a model can be expressed as a function of other segments of the same Table or as segments of another Table, or they can even be evaluated by an RPL procedure. Models can be very attractive in quality control labs and for those who must perform routine data evaluation.

RS/1 provides a package of statistical and curve-fitting algorithms. The package includes basic algorithms to obtain the mean, median, correlation coefficient, and variance. RS/1 does not support more elaborate techniques such as the Fast Fourier Transform or other frequency domain techniques. The program assists in data formatting for use by BMDP and SAS, but the user must provide these statistical packages separately. RS/1 provides useful curve-fitting commands, which enable fitting tabular data to user-specified linear, nonlinear, and polynomial functions. A special interactive algorithm may also provide a multidimensional fit of tabular data as a function of several variables. RS/1 does not support curve fitting by SPLINE techniques.

RS/1 supports four types of graphic techniques: graph charts, pie charts, bar graphs, and three-dimensional plots, called Threeds. A useful edit mode facilitates gradual alteration of parameters of a graphic object. The user may find this very useful, especially for complicated three-dimensional plots, where one may wish to rotate the object, scale the axes, or change some other parameter and view the same object again. The user may assign the graphs into a special rectangular area on the screen called a Window. The Window technique is very useful when one wants to get a complicated combination of a few graphs and tabular text on the same screen. RS/1 is rather conservative in the use of windows (compared to the SAS and LOTUS products) and does not support windows for the text editor.

Printing or plotting data objects (tables and graphs) on a dot printer or a pen plotter can be easily accomplished by a special printout command. The user may choose the size and location of the printout in a similar way to the partition of the screen. The selection of printing/plotting devices supported by RS/1 is rather limited: RS/1 Version 12.1 supports only IBM printers (or their equivalents) and IBM 7371 and IBM 7372 pen plotters (or their HP equivalents). Even Graphing Assistant by IBM supports more non-IBM devices than does RS/1.

The main advantage of RS/1 over other data management programs is the integration of a command language and a programming language (RPL) into a single unit. All the data management commands can be incorporated into algorithms written in RPL. RPL is basically an extended PL/1 and therefore exhibits the same advantages and drawbacks as its mother language. Thus, one may enjoy the free formatting and block structure of PL/1 but must learn a programming language that is less familiar than FORTRAN and BASIC.

The designers of RS/1 have made a special effort to make the program as user-friendly as possible; this is accomplished by providing an interactive form to most of the RS/1 commands. The system guides the novice user through a set of prompts to a full identification of the object, providing a default answer to every prompt and help facilities to most of them. Additionally the RS/1 provides a Help menu that displays a summary of the commands available by RS/1. The documentation of RS/1 is thorough and consists of three books, an installation manual, and a demonstration diskette, although the demonstration diskette does not represent the full capabilities of this powerful language.

RS/1 provides an extensive and friendly data management system, incorporating command and programming language. An unexperienced user should be able to use the basic RS/1 features after a few hours of practice; full exploitation of RS/1 will, however, require a much longer period.

Lev Ovadia, *The University of Texas at Austin*

ChemIntosh and ChemPanion. Versions 1.1. SoftShell Company: P.O. Box 632, Henrietta, NY 14467. Retail price: $295 and $150, respectively. (A demo version is available at no charge.)

ChemIntosh and ChemPanion are desk accessories designed for drawing chemical structures on the Macintosh computer with at least 512K of RAM memory. (A minimum of an 800K floppy drive is required to get the most from ChemIntosh and is recommended for ChemPanion although it will perform with a 400K drive. A hard disk drive is highly desirable although not necessary.) ChemIntosh is described as a desktop publishing desk accessory for drawing chemical structures while ChemPanion is touted as a desk accessory companion to applications for drawing chemical structures with the Macintosh computer. They are similar in operation to the MacDraw and ChemDraw application programs.

The major differences between ChemIntosh and ChemPanion are that ChemIntosh has the capability to print drawings from the desk accessory as well as save and retrieve documents prepared with the desk accessory (ChemPanion does not). Additionally, ChemIntosh has three added features: a ruler, a full page view for formatting, and a text palette to allow text to be added to the drawing.

ChemIntosh is described as a versatile chemical drawing desk accessory that functions much like the familiar MacDraw application program. One of the main advantages of ChemIntosh over other applications is that it *is* a desk accessory and as such may be utilized while other applications are in use. There is no need for a Switcher setup to allow easy rotation from, e.g., a word processing program to the drawing program and back. These desk accessories are similar to ChemDraw, a chemical drawing application distributed by Cambridge Scientific Computing Inc., Cambridge, MA.

Features: A series of palettes on the left side allow the selection of a variety of tools that are used to draw, add text, incorporate arrows, add labels, edit, and format the page. Rings of all sizes (3–8) can be drawn in a single motion. This includes two unsymmetrical cyclopentane rings, chair and boat cyclohexane rings, and a benzene ring in addition to the standard symmetrical rings. Adding rings to existing ones is somewhat confusing, but it gets easier with practice. Single bonds may be entered as normal, bold, bold wedges, hashed wedges, hashed bonds, wavy bonds (undesignated stereochemistry), or dotted lines. These are easily changed to double and triple bonds with a single operation. Bond lengths may be constrained or unconstrained as can the angle of a new bond, although once in place a bond angle or length cannot be changed. Labels are added to structures through a palette tool and a panel at the top where labels are typed in. One of the least attractive features is the difficulty with which existing labels within a structure can be changed. Straight arrows, dashed arrows, resonance arrows, equilibrium arrows, and especially curved arrows of any shape (for reaction mechanisms) are easily incorporated into the drawing. Solid and dashed circles and ovals are also available. The constraints on length, width, and angles of bonds and arrows are easily changed to each persons preference with a customizing palette or the constraint can be removed (length and angle only) by utilizing the shift key. Text is added through a palette tool with fonts, font sizes, and styles easily interchangeable. An eraser is available and a palette help tool for on screen help is also provided.

Page formatting is accomplished by utilizing the select tool (very similar to that in MacDraw) which allows structures to be moved and sized as well as aligned with other structures or text. A full page view can be used while working in the window in ChemIntosh although this is only for viewing and manipulations can only be carried out in the

normal window (not in the full page view). A ruler can also be placed at the top and left of the screen to help with alignment of structures.

The desk accessories are very compatible with word processing and other applications. ChemPanion is specifically designed to be used in conjunction with other applications since documents cannot be printed directly from the desk accessory but must be pasted into another document before printing. Structures or schemes can be cut from ChemIntosh/ChemPanion and pasted into other applications and then printed as a single document with no loss in resolution to the chemical structures (even with the LaserWriter). ChemIntosh documents can also be saved and opened during operation. One major drawback is that structures cut from one ChemIntosh document and pasted into another are treated as a "picture", i.e., a unit structure, and cannot be modified further. The structures drawn with this desk accessory can be printed with the ImageWriter or the LaserWriter, and the quality on the LaserWriter is excellent although not quite of the quality of MacDraw or ChemDraw.

The manuals are helpful and easy to follow. A quick start section is provided for experienced Macintosh users and a longer more detailed description is provided for the beginner. The on screen help is also quite useful. Overall, the programs are easy to learn and use, and if some of the problems with the early versions can be corrected in updated versions this will be a top notch drawing program.

Michael T. Crimmins, *University of North Carolina at Chapel Hill*

Molecular Editor. Version 1.0. Kinko's Graphics Academic Courseware Exchange: 4141 State Street, Santa Barbara, CA 93110. List price $30.55.

Purpose: Molecular Editor is a Macintosh software package for visualizing three-dimensional molecular structure. It is intended for use by students, chemistry teachers and professors, professional chemists, biochemists, materials scientists, and geologists. It illustrates several aspects of molecular and crystalline structures including basic shapes, interatomic distances, bond angles, isomerism, steric hindrance, and functional group substitution and is capable of representing proposed reaction mechanisms.

Features: Molecular Editor may be used on a 128 K Macintosh, but the 512 K Macintosh provides more memory for temporary structure storage and additional printing features.

Over 100 atoms may be displayed in three dimensions, as patterned circles with either covalent or ionic radii. Patterns and radii may be edited or adjusted, and bonds are filled rectangles or lines. Molecules may be rotated about any of three orthogonal axes in single steps or continuously. Up to 20 structures may be "open" and played in sequence so that simulation of reaction mechanisms is possible. Reflections, rotations, and inversions of the structures are possible. Regular solid shapes (line representation) may be displayed.

A small set of molecule files is included with the software to provide useful templates or examples for examining the various features of the software without first building a structure. The accompanying several pages of software description guide the user through many of the basic features, but the authors emphasize the importance of the hands-on feel of the program, so typical of the Macintosh environment.

The user must be aware that this program requires considerable knowledge of the actual atomic coordinates and bond arrangements for new molecules of interest. There are no default distances, angles, or valences that are assigned or checked. These features may disappoint the first-time user, but the software does not claim to be a "model builder". It is intended to facilitate the visualization of predetermined or hypothetical structures and has many excellent features that are implemented in a modest amount of code. New molecules are entered either from the screen or from a coordinate table.

Access to this coordinate table should delight inorganic chemists and crystallographers since the software does not constrain the user to an organic-based environment. The freedom to draw and manipulate unusual geometries is refreshing, and the ability to rotate small crystals can help the user to appreciate the projections of layers viewed from different angles.

The software is easy to use, particularly for those familiar with the Macintosh, and the basic features can be explored in a short time. While structure input is a bit cumbersome, the resulting access to the manipulation features is well worth the effort. Many sophisticated computational programs in quantum chemistry, molecular mechanics, and crystallography produce XYZ and connectivity information and it may be worth the scientists' effort to interface a standard tabular output (Multiplan template X-Ray is provided) to this Macintosh-based structure manipulator. The software measures distances between atoms, angles between bonds, and torsional angles which are important parameters accessible by the clicks of a mouse.

The software should be useful to chemists in a large number of fields. While many enhancements could be recommended, it is an excellent program at a very lost cost.

Richard L. Deming, *California State University, Fullerton*

STATA. Version 1.5. The Computing Resource Center: 10801 National Boulevard, Los Angeles, CA 90064. List Price: $395.00 (STATA alone); including optional STATA/Graphics, $590.00. Educational pricing: $265.00 (STATA alone); with STATA/Graphics, $395.00.

STATA is a sophisticated statistics/data analysis software package for the IBM PC and compatibles. It makes powerful statistical analysis tools readily available to the scientist with a simple comprehensive set of commands or menus. Minimum computer system requirements are the following: IBM PC, AT, or compatible, MS-DOS V2.0 and later, 256K memory (allowing 750 observations of 99 variables) up to 640K plus expanded memory, graphics card and monitor (CGA, EGA, Hercules), printer (many popular dot matrix printers and the HP line of LaserJet laser printers), or plotter (HP-GL pen plotters) for output. Support for the 8087 or 80287 math coprocessor (if present) is standard. Automatic expanded memory support, with data sets up to 8 megabytes of memory or 32754 observations, has been added since the 1.4 version. Limited graphics support is included in STATA, but with the optional STATA/Graphics program, some presentation quality plotting features are obtained for screen and printer/plotter output. STATA is not copy protected.

Features: All procedures are accessible through simple commands or through a menu system. Novice as well as experienced computer users can use the menu system to control an analysis session instead of manual command entry. Switching from menuing to manual command entry can be done while the program is running. On-line help is available at all times. User definable macros and batch files are supported. Simple single word commands are easier to understand and use than those found in more complex statistical packages like SAS or SPSS-X. The provided manuals give a good explanation on how to use the package and just as importantly what is being done to obtain the answers.

Data files can be read from ASCII files or input from the keyboard. With 640K, and the default 99 variables per observation, a data set can consist of roughly 2400 observations. A larger data set can be analyzed by reducing the variables per observation from the default. Output from STATA can be written to a disk file in ASCII format for use in other programs.

New analysis procedures include one way analysis of variance (ANOVA) with multiple comparison tests, *n*-way ANOVA and covariance, and maximum likelihood estimation. The one way ANOVA permits Bonferroni, Scheffé, and Sidak multiple comparison tests as options. The ANOVA procedure allows for balanced and unbalanced designs, including designs with missing cells. This procedure can also be used for regression analysis. Numerous commands allow easy inspection of a data set, both numerically in tables and graphically. Logit and probit models for individual (in STATA) or grouped (in other supplied routines) level data are possible.

The STAT.KIT and GRAPH.KIT files included with STATA are a collection of programs written in STATA. These supplement STATA with additional statistical tests (i.e., non-parametric tests) and convenient graph procedures (i.e., histograms and partial regression leverage plots).

STATA is a useful software package for chemists desiring to use most of the common statistical calculations.

Lance C. Nicolaysen, *Emory University*

SEQS. CET Research Group: P.O. Box 2029, Norman, OK 73070. List price $100 plus shipping; student version 10 copies for $250 plus shipping.

The personal microcomputer is becoming an important office tool for a growing number of chemists, fueled in part by an increasingly diverse mix of software tools on the market. Each of us collects various programs that we feel we can use, and we slowly build up a personal software library to support our particular interests. SEQS seems intended for that market. It is a general purpose program to solve algebraic equations.

The Good Side: The regular SEQS program can solve between 1 and 32 simultaneous equations. A lower cost student version can handle up to ten equations. There are versions for the Apple II, the MacIntosh, and MS-DOS machines. This review is based only on the MS-DOS version, and it should run on any IBM or clone that can support the MS-DOS operating system. SEQS is copyrighted but the diskette is not copy-protected. SEQS may be transferred to and run from a hard disk, and the sellers authorize production of one back-up copy for archival purposes by the registered user.

Equations are entered from the keyboard or from a previously stored disk file by using constants, variables, and mathematical operators. There is a full range of trigonometric, exponential, and hyperbolic functions

available. SEQS allows general expressions to be placed on both sides of the equal sign, so equations can be entered more closely to the way we might write them on paper. The expression syntax is the same as most computational languages.

SEQS uses an iterative approach to solve the equations. It starts from initial guesses of every variable, which must be supplied by the user. The system can have up to 50 symbolic parameters, but if there are more parameters than equations, some of those symbols must be designated as constants. SEQS will try to solve problems where there are more equations than unknowns, but it is conceivable there might not be a satisfactory solution in such an over-determined case. The program works by adjusting all of the variables to try to equate the two expressions on opposite sides of the equal sign. A residual is defined as the difference between the expressions on the left and right, and the sum of the squares of these residuals is used to test the convergence toward a solution. The residuals are given with the final solution, so the user can judge the quality of the answers. The algorithm used to make the new estimate for each variable on the next iteration is not explained in the documentation, nor is the criteria for acceptable convergence.

This is a useful program for the individual who often or even occasionally has to solve sets of simultaneous equations. It works on many types of equations, both linear and nonlinear. The system is not able to handle integral or differential equations, but it seems to be satisfactory on a number of the equilibria problems that were tried as a part of this review. It would appear equally valuable to the academic and industrial chemist.

Precautions: As with most iterative procedures, it is possible to encounter situations were the program does not converge or else finds an answer other than that desired. In some equation sets, there may be local minima in the residual surface that can trap the algorithm away from the desired solution. In others there may be multiple valid mathematical solutions to the equations, even if they are physically impossible. These difficulties are most likely when the initial estimates of the variables are not close to the actual solution.

To help the user find the desired answers, SEQS allows constraints to be applied to variables. It is possible to restrict the range of a parameter to positive or negative, or to be within minimum and maximum boundaries. In a triprotic acid equilibria problem, it was found that without constraining concentrations to be positive, the program did not converge to any solution, and its final estimate included several negative concentrations. With the constraints, however, the program reached a valid answer in 17 tries.

Because the convergence is tested by the sum of the squared residuals from all the equations, any equation that involves small numbers will not contribute much to the solution. For example, an equilibrium equation for the autoprotolysis reaction of water would have values around 10^{-14} on each side and may have little impact on the sum of the residuals. SEQS may produce a set of answers that is not correct. To get around this, SEQS has a mode in which the equations are scaled before they are solved, so they all contribute equally to the solution. This feature may also help achieve convergence in difficult cases.

In spite of these aids, there are some problems that are not easily solved. One equilibria problem was tried that involved a slightly soluble salt MX_2 in which the metal ion complexed with a ligand in two steps, ML and ML_2^{2-}. The ligand itself could form a diprotic acid and exist in HL^- and H_2L form. The anion of the metal salt was also a base and could form HX. When the dissociation of water was included, there were 7 equilibria involved, but there are 10 species with unknown concentrations. An 8th equation was written assuming the ligand L in all forms was fixed at a known concentration. A 9th was based on the constraint that M and X came only from the slightly soluble metal salt, and therefore the concentration of X in all of its forms had to be twice as large as the concentration of M in all of its forms. The 10th equation was obtained from charge balance. In spite of several tries, using both scaled and unscaled equations and several different sets of starting guesses, SEQS failed to give an accurate solution. Two times SEQS reported a solution, but the answers were not correct. In the other cases it did not converge.

The Bad Side: The PC-DOS version of SEQS falls short of contemporary concepts of "user friendly". The user selects options from menu screens, each of which is nothing more than directions for answering a fixed question. From these menus, the user can move into different modes that are needed to set up the problem, enter starting parameters, and solve the equations. The user must remember on which menu a particular function is to be found and must move to that menu before the function can be invoked.

The editing features are inconvenient. The mechanism for changing entered equations is reminiscent of the line editors we used before we had video screens. To change an equation, you first tell the program you want to edit equations, then indicate you want to modify an existing equation. Next you have to tell which equation you want to change. Then you must specify the characters to be changed and finally the replacement for those characters. You must be careful that the target string is sufficiently unique, because the program will always replace the first occurrence it finds. After all this, the entire equation is displayed again as altered. A more appropriate editor would allow the cursor to be moved to the desired place in the displayed equation and then allow insertion, deletion, or overstrikes on the equation.

The method of changing initial parameter values is also cumbersome. Here the variables must be examined one by one in order, and after each is displayed, you must indicate whether it should be changed or left alone or that you want to exit the correction process. Only after you locate the parameter you want, and explicitly tell the program you want to change that value, are you able to enter a new number. The keystrokes needed to do this seem inappropriate, and it is easy to do it wrong. Almost any mistake takes you away from the variable you want to change and there is no way to back up. You have to start the edit process over.

In a number of the editing functions, there is no way to quit. Once you tell the program you want to edit an equation or change the value of a parameter, you have to actually enter a change before you get the chance to stop editing.

There are a few annoyances in the display mechanism. The first keystroke in answer to a menu screen is echoed as the numerical value of the key reduced by 48. This makes sense when a number key is typed, but for nothing else. The second keystroke, no matter what it is, terminates the input. If the entry is not a legal menu selection followed by the return key, the program rejects it and shows the menu again. Since the 2nd keystroke cannot be used to edit the first one, there appears to be no reason for demanding that the return key be typed. Without it, you could move more easily between menus.

If a long equation is entered, and you reach column 80, the display scrolls horizontally to the left. This is a nice feature, but it is marred by the fact that the last character typed is echoed twice on the screen.

The brochure for the MacIntosh version suggests the window functions available with that operating system are used. It may therefore be easier to work with the MacIntosh version.

Conclusion: The mathematical concept behind the program appears to be sound, and it is useful to be able to enter equations and get a solution without having to code an algorithm for your particular equation set. The program works reasonably fast and produces answers to many of the kinds of problems we might struggle with if we tried to solve them manually. It can therefore be a useful addition to our personal computer software library. You have to be careful to check your answers because sometimes they are wrong. SEQS could be improved considerably in the way it interfaces to the human user. In that regard, it is simply not up to the standards we have come to expect. In spite of this, SEQS is not hard to learn, and those who are reasonably comfortable using their computers should be able to run this software after a few minutes of reading the manual.

Robert Megargle, *Cleveland State University*