TECHSIGHTING
SOFTWARE

Order From Chaos

BENOIT is a fractal analysis software product for Windows 95, Windows 98, or Windows NT used to find order and patterns in seemingly chaotic data, particularly where traditional statistical approaches to data analysis fail. It is widely used in disciplines as diverse as biology, chemistry, physics, economics, medicine, and geology. The U.S. Geological Survey, for example, employs fractal analysis to accurately predict the volume of undiscovered deposits of oil and natural gas, on the basis of data from known deposits.

BENOIT
TruSoft International, Inc.
St. Petersburg, FL.
$249; $189 (academic).
727-894-7426
www.trusoft-international.com

BENOIT measures user-supplied data by standard fractal methods. For a fractal, measures change in value as the scale decreases in size because ever-smaller pieces become included in the analysis. Measures are plotted as a function of ruler size on a log-log plot, and a fractal dimension is calculated from the slope of the resulting line. Users select one of 10 analytical measures with the software. Five of the available measures in the program act upon bitmap images in Windows BMP format. These are described as the "self-similar" or two-dimensional (2D) methods, while the remaining group of routines act upon time-series or 1D data. The latter group requires data to be in a simple but specific data format, such as is available in Excel. The program also features a data generator that produces files with a given fractal dimension. Users may find this useful for testing and control purposes.

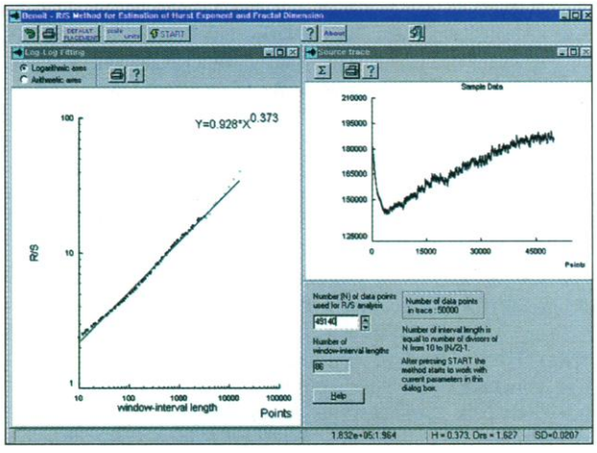

Fig. 1. Example of user data analyzed by a self-affine method.

The self-similar or image methods available in BENOIT measure different characteristics of bitmap objects in ways that should be scale-invariant. A real dataset normally has some fractal limit, and outside the limit, the fractal dimension will return a trivial value (1 for time-series or 2 for image data). Upper and lower fractal limits are controlled by the size of the dataset. Self-similar methods available in BENOIT are well known in fractal analysis: box dimension, perimeter-area dimension, information dimension, and ruler dimension. All methods are explained in standard Help files that contain several pages of information for each topic.

The 1D analysis routines use "self-affine" methods of analysis. Self-affine fractals differ from self-similiar fractals in that their parts need to be rescaled by different factors in different coordinates to resemble the original. In the roughness-length method, the root-mean-square variation or roughness of the data is calculated for a variety of horizontal scales (Fig. 1). The operation provides an estimate of the Hurst exponent, H, in a log-log plot, which is related to the fractal dimension. Standard self-affine methods available include R/S (Rescaled Range) analysis, power spectrum, roughness-length, variogram, and wavelets.

Printing of log-log figures is provided, but without many features that would be found in a spreadsheet. Documentation for the program is available online. BENOIT has a highly visual interface, complete with an animated grid or ruler for self-similar fractal methods, and it gives users control of all calculations that the program performs, unlike other fractal software.

Benoit is not without flaws. Some operations, such as name registration with the Windows NT 4.0 taskbar and the Open File requester, do not conform to standard Windows conventions. It would be of help to have an outline or flowchart of the operation of the different routines available in BENOIT for newcomers to fractal analysis.

In summary, the variety of fractal analysis methods available in BENOIT, together with generally detailed help files and significant user control of operations, make BENOIT a good resource for learning about and using fractal analysis methods.

—WILLIAM SEFFENS

Department of Biological Sciences and Center for Theoretical Study of Physical Systems, Clark Atlanta University, 223 Brawley Drive, SW, Atlanta, GA 30314, USA. E-mail: wseffens@cau.edu

TECHSIGHTING
SOFTWARE

Flash Your Data

How you present your scientific data will affect how well it is received by your audience. Whether the audience is a scientific journal reviewer, a symposium attendee, or the panel that will determine your tenure status, a good presentation will help and a poor one will hurt.

That said, most scientists are well versed in the use of basic graphics editing tools like Adobe Photoshop. With some simple Photoshop commands, one can crop and annotate figures, position collages of images, or design diagrams. However, with the growing use of the Web and lab Web sites, as well as increased reliance on computers for seminar presentations, the functionality of Photoshop has fallen behind the needs of the next generation of data presenters. The good news is that a relatively new technology called Flash from a company called Macromedia fills the gap and much more.

What is Flash? Simply put, Flash is a way to deliver the best and most compact graphics and animations through a Web browser or stand-alone application. It is a tool that all scientists should investigate because it will allow them to present their data in the largest variety of forms to the widest audience.

The first key to Flash is that the images are stored and moved in what is known as vector graphic form. This is in contrast to commonly used image formats (such as those of Photoshop) where the images are in bitmap form. Bitmap images are large because each pixel in the image must be described by several bits of information. Add up all of the bits, and file size grows rapidly. Vector graphics are described by mathematical formulas that are calculated on the fly as the image is loaded.

The second key to Flash is how well it integrates with the Web. The Flash player comes in the form of a plug-in to the browser architecture. It can be downloaded for free from the Macromedia Web site (www.macromedia.com) as a file less than 200K in size. However, since July 1998, all fourth-generation browsers from both Netscape and Microsoft have Flash capa-

bility built in. If you are using a 4.X browser, chances are you won't have to do anything special to see a Flash image, just point your browser to a "Flashed" site.

What can Flash do? The list is long, but some highlights of Flash functions are easy creation of animation, MP3 sound, form input, onion skinning, shape morphing, tweening, and freehand graphics. Onion skinning, for example, is a way to create see-through layers of art that can be superimposed to create complex animations. The effect is similar to flipping cards with images on them, but the images are digital. Tweening is a process whereby the starting and ending points of a movement are selected, but the intervening sequences of events are filled in by the computer. With tweening, you can easily make complex transitions between two points on the screen. Freehand graphics is a general term for turning hand-drawn (with the mouse) graphics into professional artwork. If you try to draw with a mouse, say in Photoshop, you will quickly find out that creating even simple curves is almost impossible. You end up with jagged edges. However, Flash will automatically smooth out these curves. It will also recognize that you are trying to draw shapes like circles, squares, triangles, and so forth, and fill them in with perfect angles.

What can Flash do for the average scientist? Nothing less than produce the next generation of digital presentations. We are moving into a new frontier that was pioneered by simple slide transitions in programs like Microsoft PowerPoint. With Flash, you can now design a presentation with sound, motion, rich text, and detailed graphics within one program. You could, for example, make a movie showing a cell dividing, transition to an SDS–polyacrylamide gel electroporesis gel showing the accumulation of a particular cyclin, return to a series of frames describing the technique of microinjection (complete with pictures of the equipment), and then morph to confocal slides of cells arrested in G_2 phase after injection of a mitotic inhibitor. The presentation could even be matched with sound narration from the lead author (who couldn't make the meeting). Because Flash can be run within a shell that does not include the browser, the presenter would not need an Internet connection or server.

Some more animated examples could include all kinds of training videos from safety seminars, to technique tutorials, to online classes. You don't even need to use animation to find a benefit for using Flash today. The ability to mix and match a variety of graphics formats to produce one image would be useful in itself.

How do you get this technology in your lab? First, ask around and, chances are, a student or co-worker may know the basics already. Next, download the software demo from Macromedia's Web site and take their digital tour. From there, it is up to you to Flash your colleagues with your new skills.

—ROBERT SIKORSKI AND RICHARD PETERS

Tech.Sight is published in the third issue of each month. Contributing editors: Robert Sikorski and Richard Peters, Mednav.com, Byfield, MA, and Kevin Ahern, Department of Biochemistry and Biophysics, Oregon State University, Corvallis, OR. Send your comments by e-mail to techsight@aaas.org.

TECHSIGHTING
PROTEIN EXPRESSION

Folding to Green

Anyone who has spent any significant amount of time behind the laboratory bench trying to express proteins in *Escherichia coli* knows that this process can often be frustrating. Overexpressed proteins have a tendency to misfold and aggregate: the result is that you end up with material that is neither biologically active nor useful for structural studies. It is assumed that thermodynamically or kinetically trapped intermediates, or improper disulfide bonds are major pathways to aggregation and hence insolubility. To get around this problem, students and postdoctoral fellows spend countless hours trying a number of approaches. These include changing the growth conditions (different media, different temperatures, different bacterial strains, and so forth), fusion of the protein with a soluble partner such as glutathione *S*-transferase, co-expression with folding catalysts such as chaperones, and unfolding and refolding experiments. In structural biology laboratories, for instance, laboratory personnel can spend many months trying to isolate large amounts of pure, soluble, and functional proteins.

A recent report published in *Nature Biotechnology* takes a step toward increasing the productivity of protein chemists (*1*). The authors used green fluorescent protein (GFP) as a reporter for protein folding: the protein of interest is expressed as an amino-terminal fusion with GFP. The scientists demonstrated that the correct folding of the downstream GFP and consequent formation of its chromophore are directly proportional to the folding of the upstream protein. Because correct folding of GFP leads to increased fluorescence, monitoring fluoresence of intact cells becomes an easy assay to quantify the solubility of expressed proteins. The system can also be used to isolate more-soluble variants of proteins that would normally aggregate.

How did they do it? They first expressed 20 cytoplasmic proteins from *Pyrobaculum aerophilum* in *E. coli* at 37°C as amino-terminal GFP fusions. Next, they measured the fluorescence of liquid cultures while the solubility of the protein expressed without the GFP tag was separately measured by SDS–polyacrylamide gel electrophoresis. There was a direct correlation between solubility and fluorescence, indicating that the fluorescence that emanates from the GFP folding reporter was an appropriate marker of the solubility of the upstream protein. Next, they decided to try to improve the solubility of known misfolded proteins and use the GFP assay as an indicator of success. They used the C33T mutant of gene V protein and the bullfrog H-subunit of ferritin. After putting these proteins through several rounds of directed evolution selecting for increased solubility, they assayed the intact cells for fluorescence and the nonfusion protein for solubility. The solubility of those proteins was increased by at least 10-fold, and fluorescence was increased at least 40-fold.

The authors also wanted to make sure that such a solubility increase could be achieved without loss of function: after all, a protein could end up being more soluble, but nonfunctional. To test this, they focused their efforts on the H-subunit of ferritin. Each of three mutants examined was highly soluble, but two of the three had a mutation that destroyed iron binding. The third mutant retained most of the wild-type iron oxidation activity, thus proving that directed evolution of solubility with GFP as an intramolecular folding reporter can be a successful approach.

There are a number of aspects of this report that are very attractive. The ability to obtain soluble functional variants of an otherwise insoluble protein without having to rely on functional screens other than GFP fluorescence should free up the hands of busy students and postdocs. Fluorescence-activated cell sorting can be used to select clones with high fluorescence. Finally, the approach could be coupled to high-throughput in vitro screening of protein folding from amplified genes. The technique needs to be adopted by a number of laboratories to see how applicable it can be across many other systems. Having spent a good amount of time at the bench (and cold room) trying to improve the solubility of expressed proteins ourselves, we believe that this relatively simple approach may have a great impact on the field of protein chemistry.

—RICHARD PETERS AND ROBERT SIKORSKI

References

1. G. S. Waldo *et al.*, *Nature Biotechnol.* **17**, 691 (1999).