

f e a t u r e

Applications of cellular automata: attractors and fractals in *analytical* chemistry?

Robert A. Lodder, Mark Selby and
Gary M. Hieftje
Bloomington, IN, U.S.A.

A cellular automaton is a discrete dynamic system of simple construction, yet capable of exhibiting complex self-organizing behavior. A cellular automaton can be used to model differential systems by assuming that time and space are quantized, and that the dependent variable takes on a finite set of possible values. Cellular-automaton behavior falls into four distinct universality classes, analogous to (1) limit points, (2) limit cycles, (3) chaotic attractors (fractals), and (4) 'universal computers'. The behavior of members of each of these four classes is explored in the context of digital spectral filtering. The utility of class 2 behavior in experimental data analysis is demonstrated with a laboratory example.

Most chemists, if they have heard of cellular automata at all, are probably familiar with them only as a means of synthesizing complex computer-graphic images. The public's view of automata has been shaped by cinematic applications like the Lucasfilm Computer Graphics Laboratory's creation of the 'Genesis bomb' explosion in *Star Trek II — The Wrath of Khan*. The perception of automata as an arcane mathematical construct has prevented their application to a number of problems for which they are naturally suited, and has left many researchers with the public's impression that cellular automata are best left in the hands of Hollywood's high-tech artists.

Nevertheless, cellular automata have been applied both to theoretical problems and experimental data analysis. For example, workers at the University of Toronto have used such constructs to theoretically model the formation of stars and galaxies', and cellular-automaton-based computers have been used to analyze data transmitted to earth from Landsat satellites². The purpose of the present review is to outline what a cellular automaton is and how it works, and to show how automata can be used to study real systems. The cellular automaton as a problem-solving machine will be compared to the

more common Turing-machine approach in the course of this discussion.

In essence, the comparison between Turing machines and cellular automata is a comparison between sequential and parallel methods of problem solving. Turing machines and cellular automata are both mathematical constructs that can be used to investigate the process of computation. Modern computing is based on theories developed using the Turing machine, and the use of parallel constructs like cellular automata represents a fairly recent development in computing.

The Turing machine and its operation have been described thoroughly", and only a brief discussion is necessary here. The Turing machine is usually thought of as a mechanical device, although few would bother to actually build one. (The operation of the machine is much more efficiently simulated on paper.) The 'machine' is composed of a tape (containing 'ones' and 'zeroes') and a tape scanner, and is really quite similar to a cassette recorder. The only real difference between the Turing machine and the tape recorder is that the Turing-machine head reads and writes in both directions on a single track. The direction of the head's movement and its read/write mode are determined by the state of the machine and the information configuration of the data tape. The head moves in discrete steps and operates on only one value on the tape at a time. The head's next step is determined by the current value on the tape (at the head) and the current state of the machine (a configuration of 'switches' that is itself determined by the previous value at the tape head). Despite its seemingly simple construction, the Turing machine has been shown to be capable of performing the most complex calculations, given enough time.

The simplest cellular automaton resembles the Turing machine in that it also starts with a one-dimensional array of data values. In cellular automata, however, each site can directly communicate with some of the others, and the values at all of the positions along the array are updated simultaneously instead of one at a time (sequentially). A cellular auto-

maton is perhaps the simplest architecture for a massively parallel processor.

In a typical cellular automaton, each cell in the array contains one member of a finite set of possible cell values. The values in all of the cells change according to the same set of rules (a program). These rules describe the state of a cell as a function of the previous value in the cell and the previous values of the cells in some defined neighborhood around the central site. The values in all of the cells are synchronously updated. In the simplest case, a cellular automaton is a one-dimensional array of cells arranged on a line so that each cell has only two nearest neighbors. A more common computational configuration has the cells in a 2-D square array, giving each cell eight nearest neighbors. Mathematical analyses of cellular automata treat the cells essentially as 'mailboxes', and describe the temporal evolution of structure in the cell-value configuration. The computational implementation of cellular automata is a bit more complex, with each cell representing a simple processor, some memory, and an I/O device. (As a matter of historical interest, the name 'automaton' came about as the result of John von Neumann's attempts to construct a machine that could reproduce itself. When Von Neumann adopted Stanislaw Ulam's suggestion to move the search for self-reproducing systems to the mathematical level, and Von Neumann discovered geometric configurations that spontaneously replicated in an array, cellular automata were born.)

The difference between conventional computing and computing in a cellular-automaton framework is essentially the difference between parallel and sequential methods of problem solving. In sequential problem solving, of course, a problem is broken into steps, each one of which must be solved before the next step can be attempted. In parallel problem solving, however, the problem is broken into pieces somewhat differently, so each piece can be solved independently of the others. (As mentioned above, cellular automaton sites generally communicate with a certain neighborhood of nearby cells, so the calculations are not totally independent.) Turing worked to prove the sequential method capable of universal computing, while Conway and Wolfram did the same for the parallel method using cellular automata.

Cellular-automaton behavior (in terms of cell-value configurations) falls into four distinct universality classes". Because cellular automata can be considered approximations of differential equations (with large numbers of discrete degrees of freedom), it is not too surprising that these configuration classes parallel the kinds of behavior of non-linear

ordinary differential equations. Each configuration class can be categorized with respect to its self-organizing behavior, and in particular to its attractor. A good appreciation of each of these classes of behavior can be obtained when one examines their performance in a digital-filtering context. To simplify the following discussion, a synthetic digitized 'spectrum' of 100 readings was created with full-scale peaks (given a value of 1) and a noiseless baseline (given a value of zero). The 100 synthetic spectral values that were created are: 011001101100111101100101111111011001010100010000101010000001000011100001100110111101011100101101010. The plots in Figs. 1-4 depict contours of a 100 x 100 element array denoting transitions between the 0 and 1 readings. The original 'spectrum' (which is the same for Figs. 1-4) occurs above the top row of each figure. The subsequent rows in the figures represent 100 successive synchronous steps in the temporal evolution of the 1-D automata. Each successive synchronous step is equivalent to a filtering pass on the previous spectrum, producing a new sequence of readings.

Fig. 1 depicts a cellular automaton rule acting on the spectrum to produce class 1 behavior. Class 1 behavior is characterized by a limit-point attractor: temporal evolution from all initial states tends toward a homogeneous final state. In other words, a class 1 filter acting on a spectrum tends to reduce all of the peaks to a flat baseline (with a non-zero value, perhaps). Running-means filters (among other rules) have this effect.

Fig. 2 depicts a class 2 cellular-automaton rule acting on the same initial spectrum used in Fig. 1. Class 2 rules are characterized by limit-cycle behavior: temporal evolution from all initial states tends to propagate the value-structures around some sites indefinitely while extinguishing the values at the remaining sites. The value of a particular cell after a

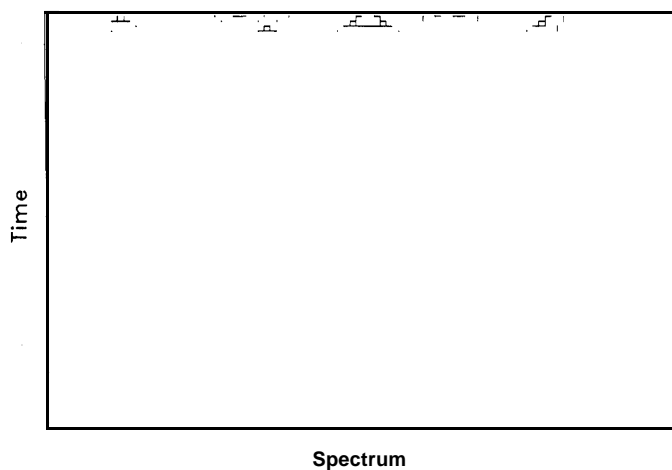


Fig. 1. A class 1 rule operating on a synthetic spectrum.

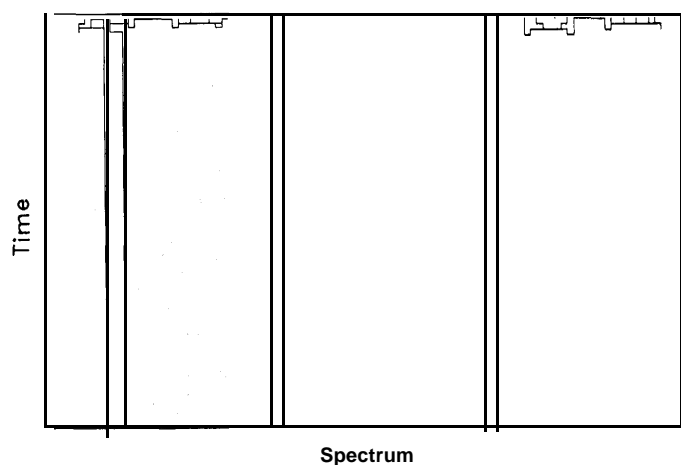


Fig. 2. A class 2 rule operating on a synthetic spectrum.

large number of time-iterations depends upon the initial values at a definite number of the 'original' cells. Running-median and certain other non-linear filters show this kind of behavior, which is desirable from a spectral-filtering or image-enhancement standpoint, because all values that are not part of a target structure are eventually extinguished.

Fig. 3 shows a class 3 rule acting on the same initial spectrum used in Figs. 1 and 2. Class 3 rules are characterized by chaotic (strange) attractors: temporal evolution from all initial states leads to reproduction of the original structures at seemingly random locations and scales. This random scaling is typical of fractals, and in fact, fractal dimensions can be calculated for cellular automata⁴. As the number of time-iterations increases under a class 3 rule, the value in a particular cell is determined by an ever-increasing number of the initial cells. The chaotic pattern that evolves has a fractal self-similarity that is undesirable from a spectral-filtering point of view,

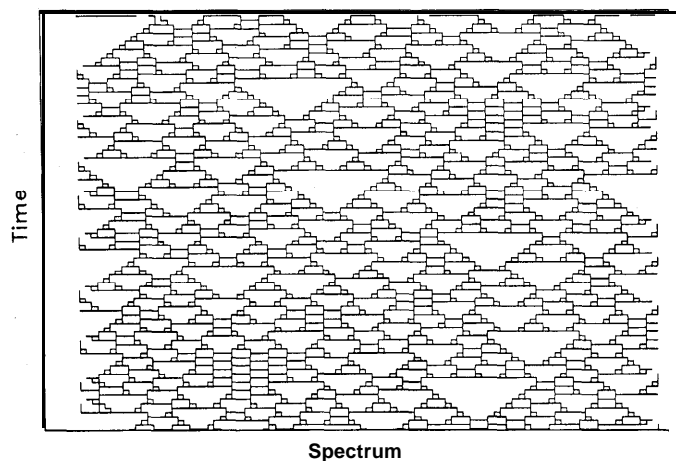


Fig. 4. A class 4 rule operating on a synthetic spectrum.

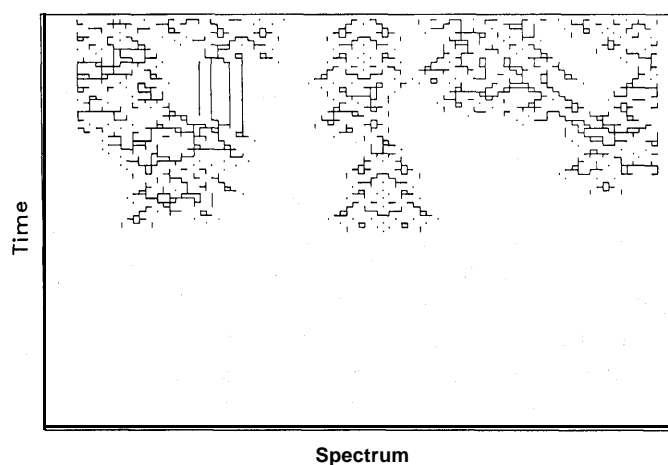


Fig. 3. A class 3 rule operating on a synthetic spectrum.

because the pattern means that peaks would spontaneously replicate in the spectrum during filtering, and would constantly change location and magnitude.

Fig. 4 illustrates the fourth class of cellular-automaton behavior, again using the same initial spectrum as the previous three figures. Class 4 rules are characterized by the lack of a definite attractor: temporal evolution from all initial states leads to configurations that may continue to change indefinitely, or may become extinct after the next time iteration. The value of a particular cell after a number of time iterations depends on the values of an indeterminate number of initial cell values. In other words, there is no general finite algorithm capable of predicting whether the cellular automaton will evolve to a homogeneous state after a certain amount of time, or whether some structures will continue to exist after an arbitrarily large number of time iterations. The final configuration of a class 4 cellular automaton is thus formally undecidable, except by the explicit simulation of every step in the evolution of the cellular automaton. The undecidability of the parallel-processing problem in cellular automata corresponds to the insolubility of the halting problem in universal Turing machines (the logical construct upon which modern CPU-based computers are predicated)⁴. Class 4 cellular automata can be considered as models for parallel processing in general, and some work indicates that cellular automata may be better models for computing at certain levels than the more conventional Turing-machine construct!

Parallel processing using cellular automata has many real applications, both inside and outside of chemistry. Mackay⁶ has proposed cellular automata as a way of developing a unified theory of crystal formation capable of describing the 230 space groups as well as the occurrence of pseudosymmetry and local

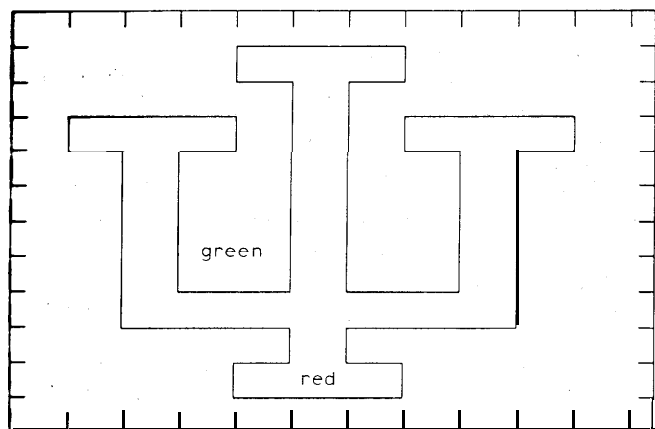


Fig. 5. The Indiana University logo was painted in this format (using red and green acrylic paints) on illustration board. The logo was then covered with one layer of white acrylic paint and one layer of white enamel paint.

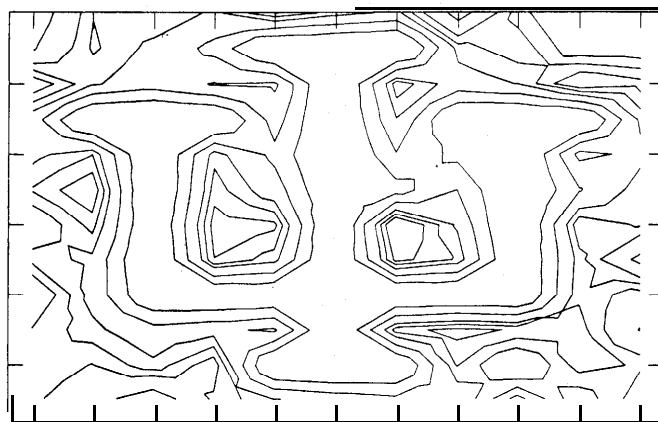


Fig. 6. The attractor for the cellular automaton formed by the near-infrared reflectance spectra of the overcoat-concealed painting in Fig. 5.

symmetry. Burks and Farmer⁷ at the Los Alamos National Laboratory have been working on modeling the evolution of DNA sequences using cellular automata. Axelrod and Hamilton have described the evolution of cooperation in a Darwinian world using elegant experiments involving cellular automata and a variety of rules contributed by invited theorists and by people responding to a magazine advertisement. At Indiana University we have recently been applying cellular automata to the three-dimensional imaging of surfaces by means of near-infrared reflectance spectrometry. The types of surfaces we have chosen to investigate are painted, specifically, the painted walls in historic buildings that are being restored⁸.

Over time, the accumulation of dirt and smoke on historic surfaces, as well as the aging of overcoat layers like varnishes, gradually obscures both the color and pattern of paintings, murals, and walls. Typically a general darkening of the surfaces gradually destroys their color and contrast. Later, restorers may repaint these surfaces to match their present (decadent) condition, and might consequently conceal entirely detail that was already becoming obscured by the aging process. Even very few such cycles of aging and restoration are sufficient to alter fundamentally the nature of a surface covered with decorative designs. Finally, political and aesthetic considerations often result in the deliberate and total concealment of original surfaces in historic buildings. In these cases the circumstances and even locations of such alterations are often long forgotten.

The usual method of investigating historic surfaces employs stereomicroscopic analysis and manual dissection of individual paint layers to detect the presence of concealed images and designs. This

method is tedious and expensive, and it unnecessarily damages the areas that do not conceal interesting subsurface patterns. A spectroscopic method is really simpler and less destructive.

In our work, a target surface was divided (not necessarily physically) into pixels (see Fig. 5), the near-infrared spectra of these pixels was loaded into a cellular automaton, and the cellular automaton rules were selected. The proper choice of (class 2 type) rules forces the temporal evolution of the automaton to converge on an attractor that is the image of the subsurface design (see Fig. 6). The use of this non-destructive spectroscopic method of reconstructing subsurface images might permit restorers to choose intelligently either to restore the original designs they find, or to duplicate them on a freshened surface with new stencils.

Cellular automata have contributed much to computer graphics, and they have much to contribute to chemistry and other sciences as well. Major changes in parallel processing and the implementation and role of pattern recognition are now underway⁹. The cellular-automaton model suggests that more than just the process sensors used in pattern-recognition methods can benefit from simplification: the computers, and even the calculations themselves, can benefit from a union of simplification and parallelism. Future work, particularly in the area of parallel algorithms and the design of instruments optimized for use with such algorithms, will open up a range of applications that have yet to be imagined.

Acknowledgements

This work has been supported in part by the National Science Foundation through Grant CHE 87-

22639, by the Office of Naval Research, and by the Upjohn Company.

References

- 1 W. L. Freedman and B. F. Madore, *Astrophys. J.*, 265 (1983) 140-147.
- 2 I. Peterson, *Science News*, 131 (1987) 28-30.
- 3 J. E. Hopcroft, *Sci. Am.*, 250 (1984) 86-98.
- 4 S. Wolfram, *Physica D*, 10 (1984) 1-35.
- 5 F. L. Carter, *Physica D*, 10 (1984) 175-194.
- 6 A. L. Mackay, *Phys. Bull.*, 27 (Nov) (1976) 495-497.
- 7 C. Burks and D. Farmer, *Physica D*, 10 (1984) 157-167.
- 8 R. Axelrod and D. Hamilton, *Science*, 211 (1981) 1390-1396.
- 9 R. A. Lodder and G. M. Hieftje, *Appl. Spectrosc.*, 42 (1988) 309-312.
- 10 R. Hecht-Nielsen, *IEEE Spectrum*, 25 (3) (1988) 36-41.

Robert A. Lodder and Gary M. Hieftje are at the Department of Chemistry, Indiana University, Bloomington, IN 47405-4001, U.S.A. Mark Selby's present address is Varian Techtron Pty. Ltd., P. O. Box 222, Springvale, Victoria 3171, Australia.

Robert A. Lodder received his B.S. degree cum laude in Natural Science from Xavier University (Cincinnati, OH, U.S.A.) in 1981. After deciding to pursue a career in chemistry he worked under Professor Richard T. O'Neill at Xavier, and received his M.S. in chemistry in 1983. He is currently a graduate student of analytical chemistry in Professor Hieftje's research group at Indiana University.

Mark Selby received his B.S. in chemistry from the University of New South Wales (Australia) in 1980 and his Ph.D. in chemistry from the same University in 1984 after which he served as a post-doctoral fellow for two years under the direction of Professor Hieftje.

Professor Gary M. Hieftje received his Ph.D. in chemistry in 1969 under the direction of H. V. Malmstadt. From 1964 to 1965 he served as a research associate in physical chemistry at the Illinois State Geological Survey in Urbana, IL, U.S.A. In 1969 he was appointed Assistant Professor at Indiana University. He now is a Distinguished Professor at the same University. His research interests include the investigation of basic mechanisms in atomic emission, absorption, and fluorescence spectrometry, and the development of atomic methods of analysis. He is also interested in the on-line computer control of chemical instrumentation and experiments, the use of time-resolved luminescence processes for analysis, the application of information theory to analytical chemistry, near-IR reflectance analysis and the use of stochastic processes to extract basic and kinetic chemical information.

interface

Developments in computer aided mass spectrometry: a networked workstation environment for data analysis

J. R. Chapman and P. A. Ryan
Manchester, U.K.

Introduction

Mass spectrometry, just as other sophisticated analytical techniques, has been profoundly affected by the impact of computer methods and has, in its turn, made considerable demands on these methods. The diversity of the applications of computers in mass spectrometry¹⁻³ can readily be appreciated by a consideration of three major areas: data acquisition, instrument control and data analysis. It is the purpose of this article to look at developments in these three areas from historical and technical viewpoints and then to consider, in more detail, the impact of the recent introduction of high-performance workstation facilities for the analysis of mass spectrometry data.

Data acquisition

Historically, the first requirement for computers in mass spectrometry was in data acquisition because of the increasing amount of experimental data that could be produced by this technique in a short period of time. This problem was particularly acute in two areas of mass spectrometry, *i.e.* the acquisition of complete high-resolution scans to give elemental composition data and the use of repetitive scanning techniques in gas chromatography-mass spectrometry (GC-MS).

An early approach to these difficulties was the use of off-line processing of recorded data. Off-line methods are, however, inherently time consuming and offer no possibility of feedback control. Again, the interposition of a data-logging device, such as magnetic tape, very often imposes technical limita-