A Cellular Automata Model of the Percolation Process

Lemont B. Kier,*,† Chao-Kun Cheng,† and Bernard Testa‡

Departments of Medicinal Chemistry and Mathematical Sciences, Virginia Commonwealth University, Richmond, Virginia 23298, and Institute of Medicinal Chemistry, University of Lausanne, CH-1015 Lausanne, Switzerland

Received September 16, 1998

A cellular automata model of a dynamic system has been created which predicts the concentration of onset and 50% probability of a spanning cluster existing which coincides with the percolation phenomenon. The valences of the cells at each concentration were monitored revealing patterns of diversity influenced by the joining and breaking rules of the simulation. The diversity of these cell valence types was quantified using the Shannon information content. The Shannon index curve versus the concentration of cells coincided almost exactly with the curve reflecting the fraction of the divalent cells at the same concentration. The simulation offers a useful solution to the difficult analysis of mobile or dynamic percolation characteristics.

I. INTRODUCTION

Percolation is a phenomenon associated with ingredients in a system reaching a critical state of association so that information may be transmitted from one ingredient to another across or through the system without interruption. This can be demonstrated by considering a system with a grid of spaces shown in Figure 1a. Some objects under study are randomly distributed throughout this grid. Because of the scarcity of these objects, little or no physical contact is encountered. No information is exchanged within the system. If enough additional objects are randomly added to the system (Figure 1b), a finite probability arises that some of these objects may be associated to form clusters. Some exchange of information occurs within the clusters, but the clusters are isolated and so the information exchange is confined within each cluster. If enough objects are randomly added to the system, the possibility arises that some clusters may appear as a single cluster which spans the entire length or width of the system. This spanning cluster produces a conduit through which an uninterrupted flow of information is possible across the system. This flow of information takes place within a process called *percolation* (Figure 1c). The minimum number of objects in the system necessary to have a finite probability of percolation occurring is called the *percolation threshold* or *percolation point*.

Percolation is widely observed in chemical systems. It was first recognized by Flory[1] and Stockmayer[2] as a method to describe how small, branched molecules react to form polymers, ultimately leading to an extensive network connected by chemical bonds. The term percolation and the mathematical concept were introduced by Broadbent and Hammersly,[3] who studied the spreading of hypothetical fluid particles through a random medium. Stauffer[4] described the application of percolation theory to conductivity and diffusivity. The relevance of percolation theory to the critical behavior of sols and gels was a major advance in understanding the physics and chemistry of these systems, contributed by De Gennes[5] and Stauffer.[6] Stanley and Teixeira[7] studied the structured continuum of water as a percolating phenomenon.

In biological systems the role of the connectivity of different elements is of great importance. Sahimi[8] has commented on several of these systems to illustrate how percolation theory may be relevant. Examples include self-assembly of tobacco mosaic virus, actin filaments, flagella, lymphocyte patch-and-cap formation, plus numerous precipitation and agglutination phenomena. Immune system function has been approached as discrete steps with the use of cellular automata approximations.[9]

The configuration of a system in which percolation may occur is classically treated as one in which the ingredients do not move. Considerable work has been devoted to these static models, leading to numerical solutions of the critical concentrations and cluster sizes associated with a percolation threshold. These have been worked out mathematically for several grid structures and dimensions.[4,10] In many situations such as the biological examples noted above, a static model is incomplete. In reality, the ingredients in a system often exhibit dynamic behavior within the system. This may give quite different results for the example in figure 1, where the percolation threshold is sought. One approach to the study of discrete, dynamic systems is through the use of computer simulations using cellular automata. In this study, we intend to model a two-dimensional dynamic system of varying concentrations to study the onset of percolation and large-cluster behavior, both visually and numerically.

II. CELLULAR AUTOMATA

Cellular automata are computer-generated dynamic systems that are discrete in space, time, and state, and whose behavior is specified completely by rules governing local relationships. This is an attempt to simplify the often numerically intractable dynamic simulations into a set of rules that mirror intuition and are easy to compute. As an approach to the modeling of emergent properties of complex systems, cellular automata are beneficial because they are

† Virginia Commonwealth University.
‡ University of Lausanne.

10.1021/ci980219l CCC: $18.00
Published on Web 01/22/1999

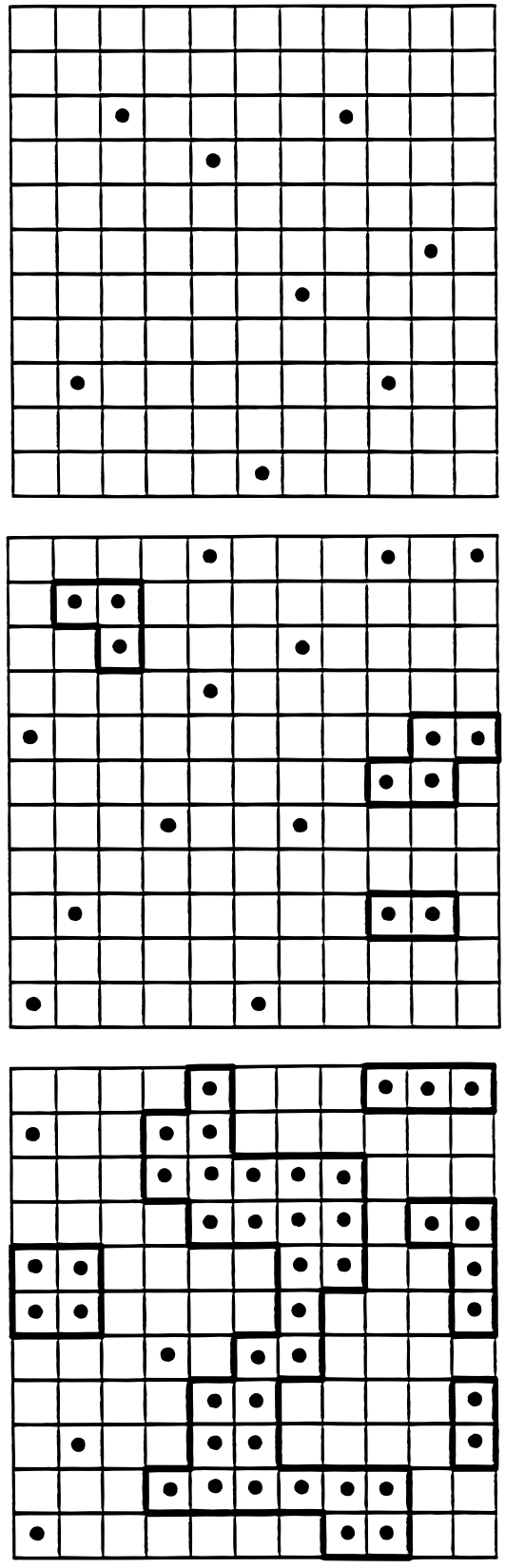


Figure 1. (a) A system in which a few objects are randomly distributed throughout the existing grid. (b) A system in which some objects are randomly distributed so that some clusters may form (encircled). (c) A system in which enough objects have been randomly added so that a probability exists of a cluster forming that spans the system (encircled).

visually informative of the progress of dynamic events at the molecular system level. From the development by von Neumann[11] a variety of applications ranging from gas phenomena to biological applications have been reviewed by Sahimi.[8] Kier and colleagues[13−17] have used cellular automata to advance our understanding of solution phenomena, including solubility, oil-water partitioning, micelle formation, aqueous diffusion, and acid dissociation.

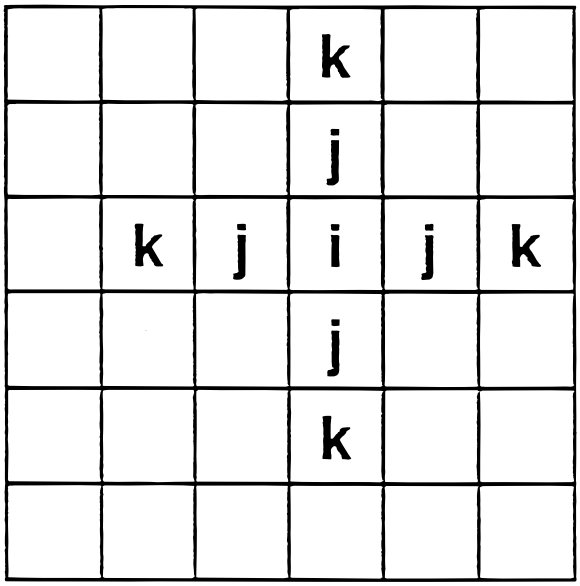
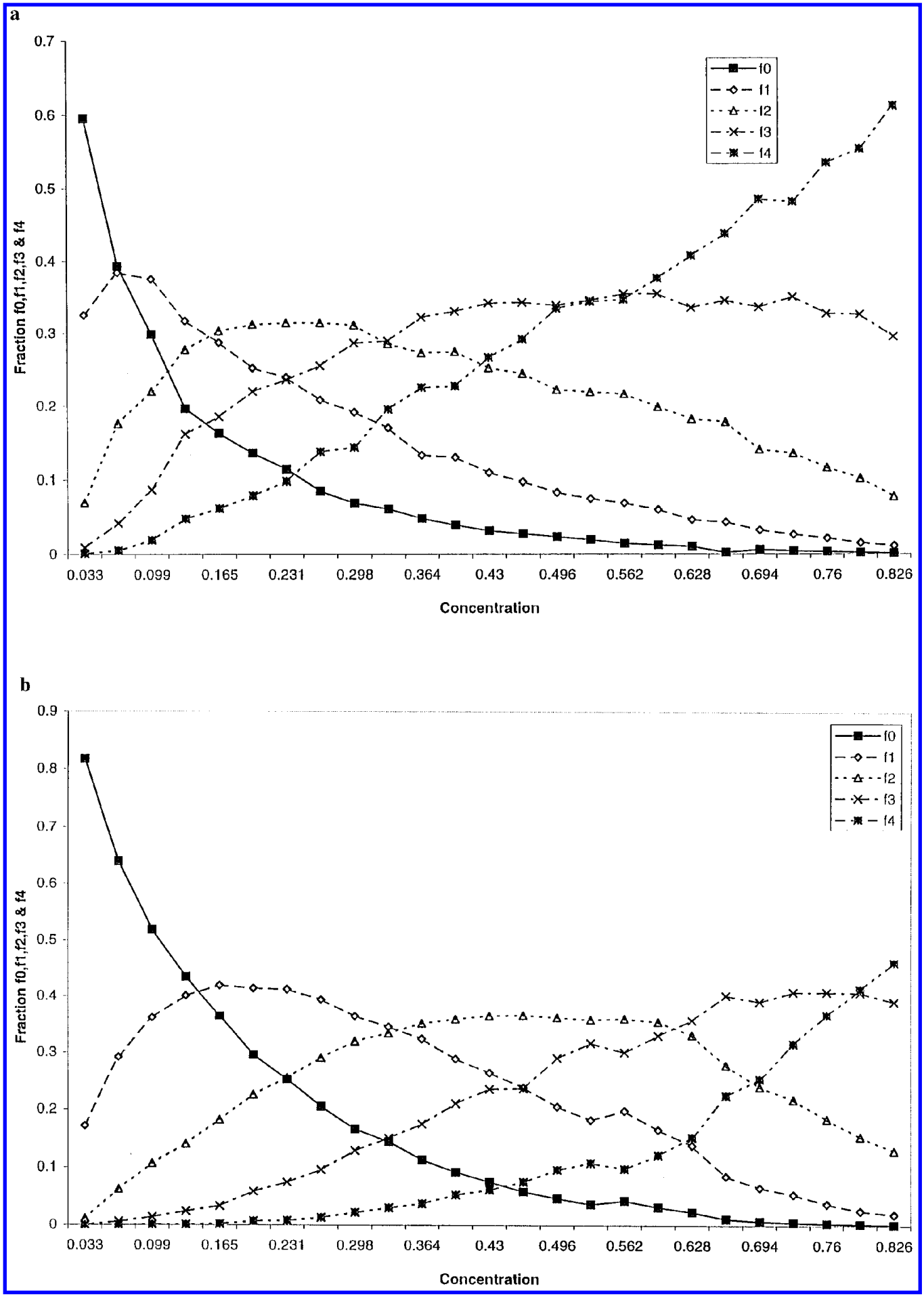


Figure 2. The extended von Neumann neighborhood centered on cell *i*.

III. EXPERIMENTS

A. The Model. Our model is composed of a grid of squares called cells (Figure 2). Each cell, *i*, has four adjoining neighbors, *j*, and four extended neighbors, *k*, beyond *j*. This is called an extended von Neumann neighborhood. Each cell is assigned a state governing whether it is empty or occupied, e.g., by a virtual molecule or other object. The contents of a cell may break away from an occupied neighboring cell or move to join a neighboring cell that is occupied. These movements are assigned as probabilistic rules at the beginning of the dynamics to reflect a relationship among the ingredients in the system. The rules then are applied randomly to each cell in turn until all cells have computed their state and trajectory. This is one iteration of time. The rules are applied uniformly to each cell of the same state. The initial configuration of the system is random, hence it does not determine subsequent configurations at any iteration. The same sets of rules do not yield the same configurations except in some average sense. After many iterations the configurations reach a collective organization that possesses a relative constancy in appearance and in reportable counts of cells, called attributes. These are the emergent characteristics of a complex system.

B. The Rules. The grid is the surface of a torus to eliminate boundary conditions. With a bounded region such as a square, the behavior of a cell at the boundary could influence whether the percolation might occur. Thus it is necessary to eliminate the boundary condition and to have a uniform minimum requirement for percolation to occur. For instance, on a sphere, a circle will disconnect the sphere however the equator and a small circle around the north pole do not share the same weight in the model for percolation.


a
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Fraction f0,f1,f2,f3 & f4
f0
f1
f2
f3
f4
0.033
0.099
0.165
0.231
0.298
0.364
0.43
0.496
0.562
0.628
0.694
0.76
0.826
Concentration
b
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Fraction f0,f1,f2,f3 & f4
f0
f1
f2
f3
f4
0.033
0.099
0.165
0.231
0.298
0.364
0.43
0.496
0.562
0.628
0.694
0.76
0.826
Concentration

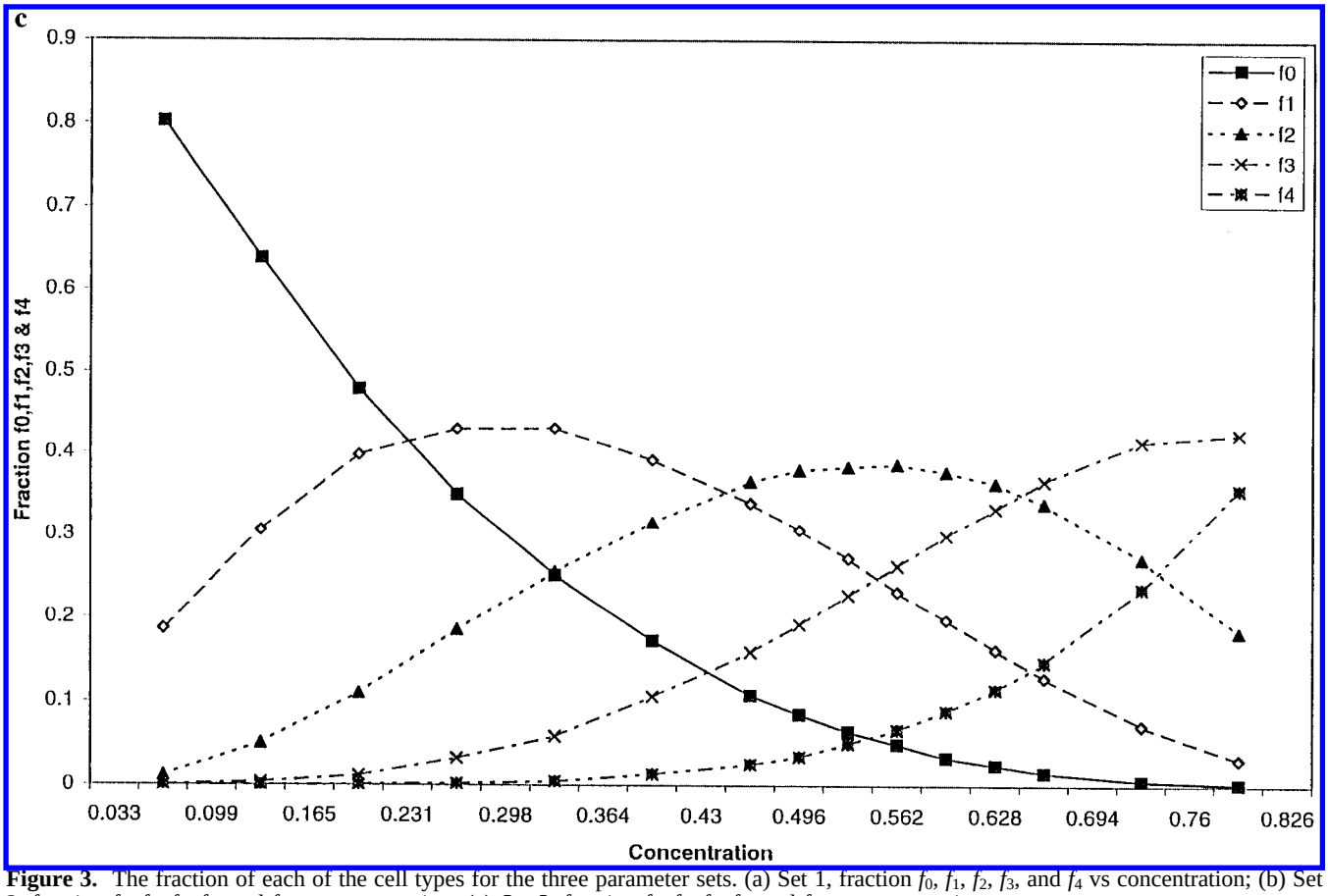


Figure 3. The fraction of each of the cell types for the three parameter sets. (a) Set 1, fraction f_0, f_1, f_2, f_3, and f_4 vs concentration; (b) Set 2, fraction f_0, f_1, f_2, f_3, and f_4 vs concentration; (c) Set 3, fraction f_0, f_1, f_2, f_3, and f_4 vs concentration.

The cross-section of the torus does support the uniform minimum requirement consideration, because the cross-sections must contain a minimum number of cells.

Two parameters were adopted in our model to govern the probabilities of "molecules" moving in the grid. The breaking probability, P_B, is the probability for a "molecule" at *i* to break away from another at *j* when it is occupied (see Figure 2). The value for P_B lies in the closed-unit interval. The second parameter, *J*, describes the movement of a "molecule" at *i* toward or away from the "molecule" at a *k* cell in the extended von Neumann neighborhood (Figure 2), when the intermediate *j* cell is vacant. The value of *J* is a positive real number. When $J = 1$, the molecule, *i*, has the same probability of movement toward or away as when *k* is empty. When $J > 1$, *i* has a greater probability of movement toward an occupied cell, *k*, than when *k* is empty. When $J < 1$, *i* has a lower probability of such movement.

In this series of studies we selected three sets of rules that cover a large range of parameter space. We intended to monitor the influence of the rules on the recorded attributes from the dynamics. The rule sets used are: Set 1, $P_B = 0.25$; $J = 3.0$. Set 2,$P_B = 0.50$; $J = 1.5$. Set 3, $P_B = 0.75$; $J = 0.5$. Set 1 describes a relationship between occupied cells in which there is a strong affinity to join and a low probability to separate. Set 3 describes an opposite relationship in which there is a weak affinity of occupied cells to join and a high probability of bonded cells to separate. Set 2 represents conditions midway between these two extreme conditions.

C. Attributes from Experiments. The studies are designed to reveal several attributes of the dynamic system. The first of these is the structure of the system as additional cells are added to the grid. In particular we are interested in the relative proportions of the five possible states of the cells described by the concentrations of each state, f_x: f_0, unbound; f_1, one face bound; f_2, two faces bound; f_3, three faces bound; f_4, all faces bound. The numerical values for each of these terms is the fraction of all the occupied cells present in that particular state.

The second attribute to be studied is the increase in size of the largest cluster as the number of cells in the system increases. Another important attribute is the concentration of occupied cells at which the percolation phenomenon occurs. In particular we are interested in the occupied cell concentration at which the onset of percolation begins and at which, about 50% of the time, percolation is occurring, averaged over multiple iterations of the dynamics. Finally we intend to evaluate the information content of the system as the concentration changes and to discover any possible relationships to attributes of the system.

D. Quantitation of Information Content. It is possible to quantify the attribute of diversity, uncertainty, or surprise in an ensemble of objects of different forms and concentrations. One method is based on the Shannon equation for the information content, *I*, of such a system. The value of the information content is:

$$I = -\sum P_i \log P_i$$

where *P* is the probability of the occurrence of any cell state. The terms are summed over all states. For example, if the concentrations (probabilities of selection) of ingredients in a system are: $f_0 = 0.5$; $f_1 = 0.2$; $f_2 = 0.2$; $f_3 = 0.1$; and f_4

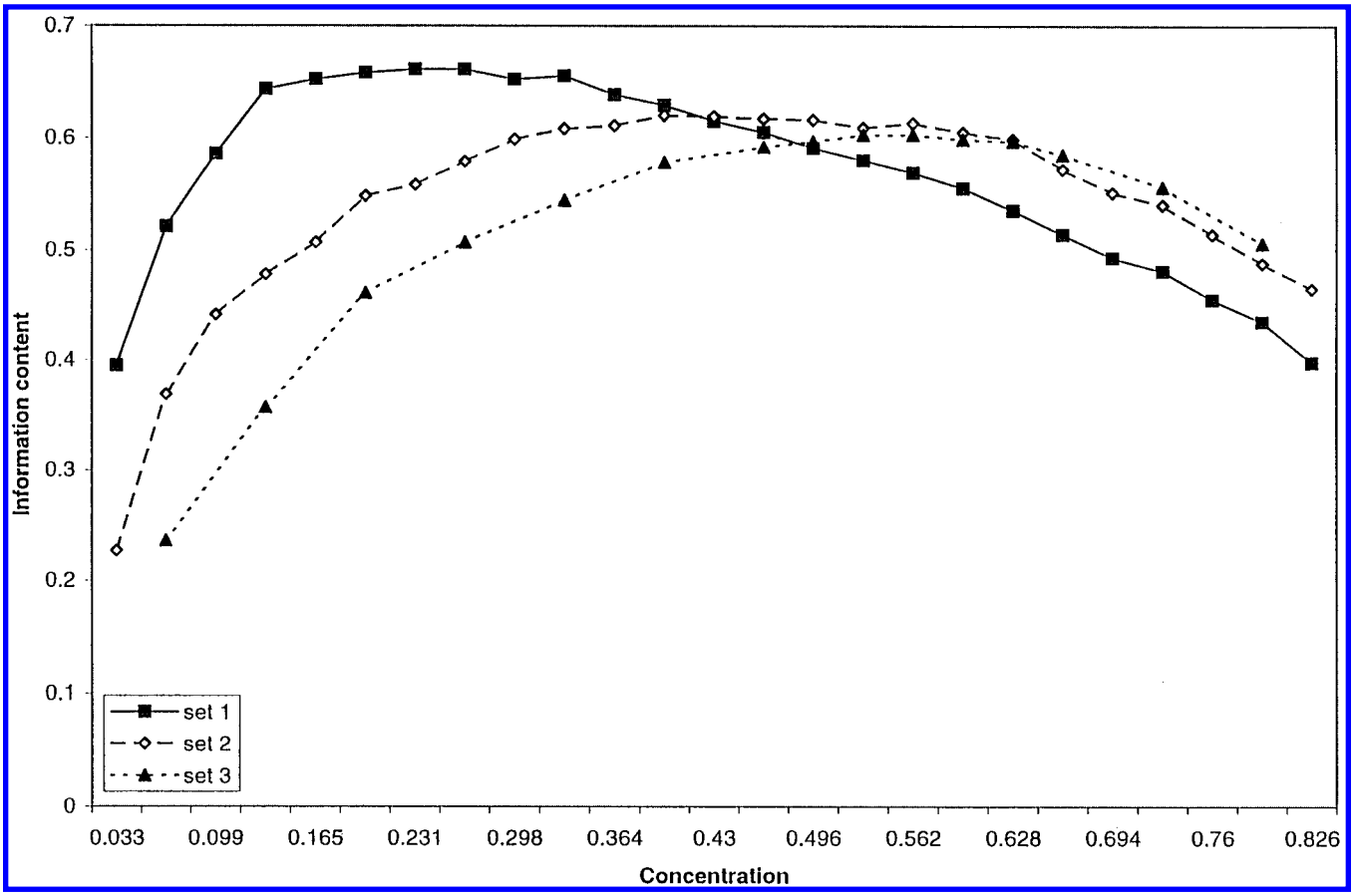


Figure 4. The diversity of the cell types, f_0 through f_4, at each concentration using the Shannon information content.

$= 0.0$, then the information content (diversity) of the system would be $I = 0.530$. If another system contained ingredients in the concentrations, $f_0 = 0.3$; $f_1 = 0.2$; $f_2 = 0.2$; $f_3 = 0.2$; $f_4 = 0.1$, then the calculated information content is $I = 0.676$. The higher information content reflects the greater diversity among the types of ingredients, a greater uncertainty upon random selection and a greater surprise upon sampling. Applying this analysis to the cell states, it is possible to calculate the information content or diversity at each concentration.

E. The Dynamic Simulation. The study was performed using 3025 cells in a 55×55 grid. The grid was the surface of a torus, hence there were no boundary conditions. A series of simulations were run starting with a concentration of 100 occupied cells. The dynamics were run for 1000 iterations with the last 100 iterations used to record attributes of the system. Successive increases in the occupied cell concentration were introduced in increments of 100. The concentrations of the five states, f_o through f_4, were recorded for each occupied cell concentration. The size of the largest cluster was also recorded at each occupied cell concentration. Finally, the occupied cell concentrations, at which the onset of percolation occurs and at which there is a 50% probability of percolation, were recorded.

IV. RESULTS

A. Occupied Cell States at Each Concentration. As the concentration of occupied cells is increased by 100 cell increments, the proportions of each of the five occupied cell states, f_0 through f_4, changed and was recorded. This profile for each of the three parameter sets is shown in Figure 3a−c. The evolution of the curves for each cell state shifts to higher concentrations with higher P_B values in the parameter set. Of particular interest is the distribution of cell states at each concentration. At lower concentrations, most of the cells are in the f_0 and f_1 states. At very high concentrations, the f_3 and f_4 cell states predominate. In the middle range of concentration, there is a significant diversity in the concentrations of the cell types.

By applying the Shannon equation to analysis of the cell states, it is possible to calculate the information content or diversity at each concentration. The I values from each parameter set for each concentration are shown in Figure 4. The information content or diversity apparently increases from low values at low concentrations to maximum values at different midconcentration values according to the influence of the parameter sets. An interesting observation comes from a comparison of the data in Figures 3 and 4. The shapes of the curves and the concentration positions of maximum I values in Figure 4 correspond very closely with the curves for the three sets of f_2 concentration values in Figure 3a−c.

B. Percolation Probability. The conditions associated with a finite probability of the onset of percolation and a 50% probability of its occurring in a dynamic system are of considerable interest. The percent percolation as a function of concentration for each parameter set have been estimated and are shown in Figure 5. These results reveal curves that have a sigmoidal shape. From these curves, we determined the concentration (for each parameter set) in which 50% of

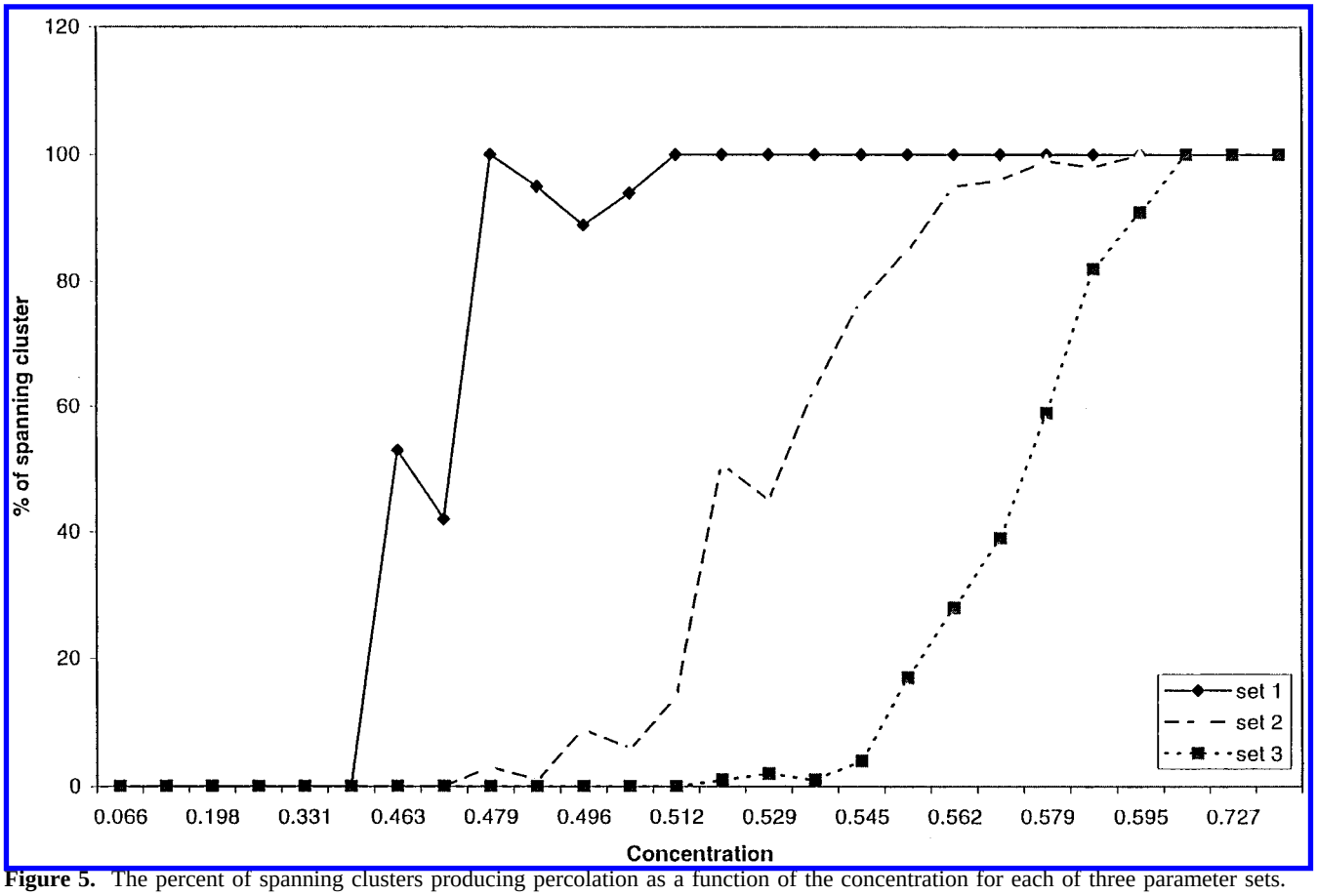


Figure 5. The percent of spanning clusters producing percolation as a function of the concentration for each of three parameter sets.

Table 1. Data from Dynamic Simulations

parameter sets		concentrations at			
P_B	J	percolation onset	50% percolation	f_2 (max)	I_{max}
0.25	3.0	0.375	0.476	0.264	0.329
0.50	1.5	0.484	0.530	0.496	0.477
0.75	0.5	0.548	0.573	0.562	0.533

the last 100 iterations in a run have a spanning cluster. The percolation point in our study of a dynamic system is therefore expressed as a concentration producing a 50% probability of percolation.

The concentration at which the onset of percolation occurs is an indefinite value within the range studied. This was estimated from studies with each parameter set and is shown in Table 1. A comparison between these percolation onset concentrations and the concentrations corresponding to the maximum information content, I_{max}, reveal a close correspondence. The I_{max} concentration appears to be very close to the percolation onset concentration for each of the three parameter sets.

C. The Largest Cluster Size at Critical Concentrations. The largest cluster size at a concentration producing 50% percolation was about the same for all parameter sets. This is a cluster size occupying close to 20% of the grid area in these studies. Stated another way, as the concentration increases for any parameter set study, there is a relatively common cluster size corresponding to a 50% probability that percolation will occur. This is an alternative definition of a percolation point when the system is dynamic.

The approximate concentrations at which there is any significant percentage of percolating clusters and the concentrations where the maximum information content of the system occurs are about the same for each parameter set as shown in Table 1. This implies that there is an inductive effect in the system whereby a maximum of diversity is necessary to create conditions where the percolation process may begin.

V. DISCUSSION

Currently it is a straightforward problem to determine the onset of percolation in several immobile systems with increasing concentrations.[8] Some of these values are exact; others are derived as approximations. In contrast, there has been considerable difficulty in assessing the percolation point in systems whose dynamic nature is taken explicitly into account, as is for "real-life" problems in chemistry, macromolecular sciences, biology, and many other fields.[4,8] Several models have been proposed for distinct situations; generality is a long way off in this realm of study.[4,8] The present study offers a convenient way of using the power of cellular automata to model the dynamic onset of percolation. Indeed we have shown here that a cellular automata model can reveal the critical concentration at which a system fluctuates into and out of a percolating state, expressing the percolation as a 50% probability. This is accomplished in a simple manner that captures the dynamic, evanescent character of the onset of percolation, as concentration is increased. We can only propose these models and the results obtained from them to the scientific groups studying these phenomena. Their value will be assessed in real situations where some experimental evidence is available. Having these models at hand will permit comparisons with other dynamic simulations which may be used.

Table 2. f_2 (max) and I_{max} Relationships

parameter set		density range (approx.)	
P_B	J	f_2 (max)	I_{max}
0.25	0.5	0.2−0.3	0.15−0.35
0.50	1.5	0.4−0.6	0.30−0.60
0.75	3.0	0.5−0.6	0.50−0.60

The analysis of the diversity of the occupied cell states in valences of f_0 through f_4 was accomplished by calculating the Shannon information content at several densities for each parameter set (Figure 3). The analysis reveals that the increase in diversity, as measured by the information content, corresponds to the increasing probability of a percolating cluster appearing in the system.

A very interesting observation is that the curves relating the information content versus density for each parameter set (Figure 4) are almost the same as the curves relating the f_2 concentration and density for the corresponding parameter sets (Figure 3). The interrelation of the f_2 and the I values for each parameter set are approximately $r^2 = 0.95$. The density range corresponding to the maximum values of each attribute are close as shown in Table 2. This finding indicates that the f_2 concentration is a surrogate for a physical manifestation of the diversity, as quantified by the Shannon information content.

What is the critical contribution of the f_2 valence state? This may be answered in a qualitative way by pointing out the roles of the various states in the dynamic events occurring as the system density increases. An f_0 valence state can never participate in a percolating cluster. An f_1 valence state may be a part of a cluster that percolates, but it is never a spanning part of that cluster. It is always an appendage, a fragment that is a dead end, not linking any two fragments in a cluster architecture. The f_2 cell is the lowest valence state that can be an active participant in the linking of two cluster fragments, ultimately leading to percolation. The f_3 cell has a significant probability of linking parts of a cluster to form a percolating system. The f_4 cell has a very good chance of being an integral part of a spanning cluster. Only the f_2, f_3, and f_4 valence state cells can play a direct structural role in the percolation process.

Considering again these five valence states, the f_0 cells have maximum movement, not being bound to any other cell. They can freely unite with any other cell or cells to create higher valence state structures. The f_1 cells have a lesser, but still significant probability of moving when the one shared bond to another cell is broken. The f_2 cell is bound to two other cells. It's probability of moving depends on the rupturing of those two bonds. This probability is less than that of a f_1 valence cell, but it is much greater than that of a f_3 and certainly a f_4 valence cell. The f_2 cell thus occupies an intermediate role in the spectrum of cell-type movements.

The conjuncture of the two attributes of structure, participation and movement probability, finds the f_2 cell type in an optimum position in the mixture of cell types to play a significant role in the formation of a spanning cluster. Further, the maximum concentration of the f_2 corresponds closely to a center of gravity of the five cell types. The f_2 cell type is thus reflected by the numerical value of the information content, which has its maximum value at the peak of diversity in the system. The structure contribution of the f_2 cells coupled with the functional movement potential of these cells is optimally critical in the percolation event.

Another finding from this study is that, in a grid of 55 cells, the onset of a 50% probability of percolation occurs at a relatively constant size of the largest cluster for a variety of parameter sets. This size is approximately a 20% concentration, and it is independent of the joining and breaking parameters used. The only difference occurring among the parameter sets is on the concentration scale where the critical largest cluster size occurs. The variation in parameters is thus important in determining the concentration at which the largest cluster size has a 50% chance of percolating in the system. Once a critical size of the largest cluster is achieved, the parameters have no further influence on the percolation. This occurs as an emergent property of the system. We infer from these studies that the onset of percolation is an emergent property in a new complex system. What happens to the behavior of the individual ingredients of the system in such a dynamic environment is another issue that may be approached with cellular automata.[18]

ACKNOWLEDGMENT

The author (L.B.K.) acknowledges the Fondation Herbette, Universite de Lausanne for generous support of this work.

REFERENCES AND NOTES

(1) Flory, P. J. Molecular Size Distribution in 3-Dimensional Polymers I. Gelation. *J. Am. Chem. Soc.* **1941**, *63*, 3083−3090.
(2) Stockmayer, W. H. Theory of Molecular Size Distribution and Gel Formation in Branched-Chain Polymers. *J. Chem. Phys.* **1943**, *11*, 45−55.
(3) Broadbent, S. R.; Hammersley, J. M. Percolation Process (I) Crystals and Mazes. *Proc. Cambridge Philos. Soc.* **1957**, *53*, 629−641.
(4) Stauffer, D. *Introduction to Percolation Theory*; Taylor and Francis Publ., London, 1985.
(5) de Gennes, P. G. Dynamics of Entangled Polymer Solutions II. Inclusion of Hydrodynamic Interaction. *Macromolecules* **1976**, *9*, 594−598.
(6) Stauffer, D. Kinetic Theory of Two-Component Hetero-Molecular Nucleation and Condensation. *J. Chem. Soc. Faraday Trans.* **1976**, *II*, 72.
(7) Stanley, H. E.; Teixeira, J. Interpretaion of the Unusual Behavior of H_2O and D_2O at Low Temperatures: Tests of a Percolation Model. *J. Chem. Phys.* **1980,** *73*, 3404−3422.
(8) Sahimi, M. *Applications of Percolation Theory*; Taylor and Francis Publ., London, 1994.
(9) Stauffer, D.; Sahimi, M. High Dimensional Simulation of Simple Immunological Models. *J. Theor. Biol.* **1994**, *166*, 289−297.
(10) Essam, J. W. Percolation Theory. *Rep. Prog. Phys.***1980**, *43*, 833−912.
(11) Von Neumann, J. In *Theory of Self-Reproducing Automata*, Burks, A., Ed.; University of Illinois Press, Urbana, IL, 1966.
(12) Sahimi, M.; Gavalas, G. R.; Tsotsis, T. T. Statistical and Continuum Models of Fluid-Solid Reactions in Porus Media. *Chem Eng. Sci.* **1990**, *45*, 1443−1452.
(13) Kier, L. B.; Cheng, C.-K. A Cellular Automata Model of an Aqueous Solution. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 1334−1337.
(14) Kier, L. B.; Cheng, C. K.; Testa, B.; Carrupt, P. A. A Cellular Automata Model of the Hydrophobic Effect. *Pharm. Res.* **1995**, *12*, 615−619.
(15) Kier, L. B.; Cheng, C.-K; Testa, B.; Carrupt, P.-A. A Cellular Automata Model of Micelle Formation. *Pharm. Res.* **1996**, *13*, 1419−1422.
(16) Kier, L. B.; Cheng, C.-K.; Testa, B.; Carrupt, P.-A. A Cellular Automata Model of Aqueous Diffusion. *J. Pharm. Sci.* **1997**, *87*, 774−778.
(17) Kier, L. B.; Cheng, C.-K.; Tute, M.; Seybold, P. G. A Cellular Automata Model of Acid Dissociation. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 271.
(18) Testa, B.; Kier, L. B. The Concept of Emergence-Dissolvence in Drug Research. *Adv. Drug Res.* **1997**, *30*, 1−14.

CI980219L