Cellular automata as emergent systems and models of physical behavior

Jason Merritt

December 19, 2012

Abstract

Cellular automata provide a basic model for complex systems generated by simplistic rulesets. While each step in a simulation is dominated by local interactions, over time complex macroscopic behavior can emerge. Observation of this long-term emergent behavior due to simple, easily understood and computationally efficient rules has led to attempts to model physical systems within the framework of simple cellular automata. This paper aims to briefly review the behavior and properties of cellular automata, provide some specific examples of CA models for physical systems, and point out the advantages and disadvantages of approaching a problem with a CA-based simulation.

1. Introduction

Cellular automata have for decades held a foothold in the public consciousness thanks primarily to Conway's Game of Life. The Game of Life, consisting of a 2-dimensional grid whose cells are either “alive” or “dead,” evolves in timesteps as the same rules governing “life” and “death” are applied to every cell in the grid, using only the cell's knowledge of its eight nearest neighbors. While the rules are exceedingly simplistic, easily (if tediously) able to be applied to a finite grid by hand and trivially by a computer, the game is known for the complex animated structures it is able to create and its strong reliance on initial conditions. “Gliders,” repeating patterns of living cells which are able to move diagonally across the grid, may be infinitely spawned (on an infinite grid) from a single structure called a “glider gun,” whereas changing the value of a single cell in the “gun” may cause it to spontaneously die or collapse into stable configurations. Some other cellular automata, such as Paterson's worms, have limited recognition either for the visually interesting patterns they create when allowed to evolve for long time periods or as a mathematical curiosity, as many cellular automata are undecidable. However, cellular automata have since also gained recognition in science as a useful tool for physical simulations and for examining the evolution of complex systems.

The reasoning for using cellular automata as a modeling tool is based on direct analogy to physical systems. The local interactions in many physical systems, despite the extreme complexity of macroscopic outcomes, may be reduced to simple guiding principles such as kinematics for determining the outcome of a collision between two particles. In CA, such guiding principles are spelled out explicitly as the rule set for that particular automaton. The hope in these cases is that using a computer to allow the CA to evolve in accordance with these rules will result in a realistic – or at least insightful – picture of the physical system, without necessarily needing to know macroscopic theory for the system or needing to do complex math such as finding solutions to nonlinear differential equations. In cases where a good theoretical framework for macroscopic behavior already exists, the primary benefit of using CA is computational efficiency, as applying a simple ruleset over many timesteps is typically much faster than having the computer do complex calculations in accordance with results from theory. In some instances, such as studying traffic flow, CA may also be used to demonstrate the emergence of macroscopic behavior as directly resulting from local behavior and actors.

1.1 Cellular automata vs. lattice gas automata

Suppose there exists a grid (typically one- or two-dimensional, but may be n-dimensional) where each cell in the grid is assigned an element of some

set A. Then Toffoli et al.[1] define a cellular automata by any map $f : A^n \rightarrow A$ where f maps the n relevant neighbors of each cell to that cell's new value. Typically some other considerations must be made where a boundary exists (such as for a finite grid with non-periodic boundary conditions).

While simplistic (this is explicitly how Conway's Game of Life works), this is not typically the way a physical simulation is conceptualized. A simple lattice gas simulation, for example, could work with a grid where any cell is either filled with a particle or unfilled, and each timestep the algorithm might update the position of each particle, check for a collision, and redirect colliding particles. This is a seemingly more complex operation than simply looking at a position's neighboring cells and updating the position accordingly. Such a “lattice gas automata” can be defined by a map given in the form $g : A_1 x \ldots x A_n \rightarrow A_1 x \ldots x A_n$ where in this case the map would likely take the velocities and positions of particles as inputs and assign their new positions and velocities as outputs. Since this is easier to conceptualize, why should we care about the CA model at all?

First of all lattice gas automata are primarily useful for describing systems undergoing invertible processes, whereas CA are directly used more often for dissipative systems.[1] Second of all, it has been proven (by, for example,Toffoli et al.) that any lattice gas automata may be rewritten as a CA, whether or not such a rewriting is wholly intuitive. Therefore any general results proven for CA or classes of CA immediately apply to lattice gas automata that fall within those classes. Very often (but not always) CA may even be rewritten as lattice gases, although this does not concern us here.

1.2 Emergence and predictability in CA

Given the enormous complexity of many CA the fact that they can express emergent behavior should not be surprising. In fact, many CA – including Conway's Game of Life – can house universal Turing machines,[2] so any emergent behavior which can result from an algorithm at all can be expressed within the framework of CA given enough time. However, as running a CA in order to simulate a Turing machine is excessively inefficient, this is not a particularly useful result. The most interesting emergent behavior is that which evolves naturally due to the CA rules.

While there is no easy way of categorizing non-trivial CA in terms other than the lattice they act on and their number of inputs, Wolfram[3] proposed the existence of four general classes of cellular automata: those which rapidly tend to equilibrium regardless of initial conditions, those which settle into oscillations, those whose output appears to be random, and those which are able to propagate complex structures forward in time. The difference is most easily understood with reference to one-dimensional CA, as shown below.

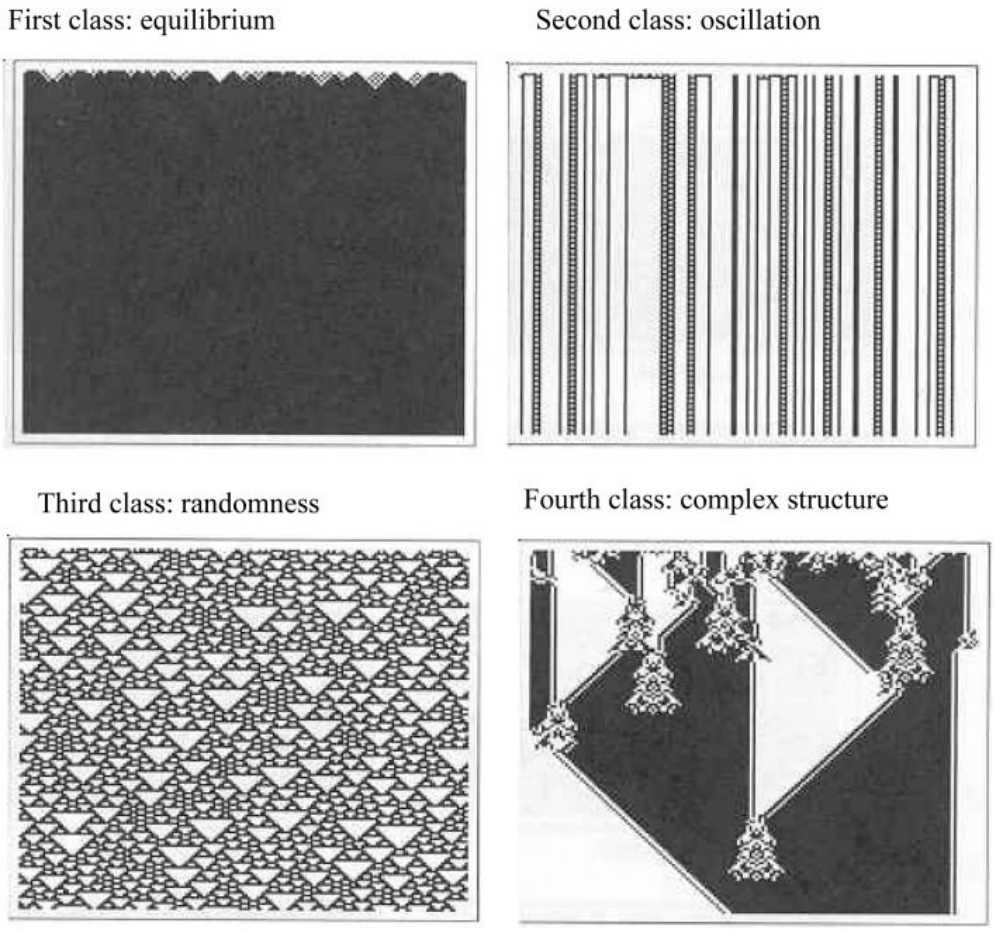

Fig 1. Four classes of cellular automata, where the vertical axis represents evolution in time. Reproduced from Wolfram[3] via Mainzer[4].

While these designations are subjective,[2] particularly the distinction between randomness and complex structure, hey demonstrate the broad difference in possible outcomes depending on the ruleset used even for simple CA (above, each cell's new value depends only on its old value and the values of its two nearest neighbors). In general, the “random” and “complex” classes are the main source of interest since they do not immediately collapse into easily predictable patterns.

Cellular automata are typically considered emergent in the sense that their long-term macroscopic behavior is (for non-trivial CA) very difficult to predict even given complete knowledge of the local behavior. For complex enough CA (such as the Game of Life) this has been likened to the behavior of biological systems, and the statistical study of such CA has been proposed to help develop realistic models for biological networks.[5] Depending on the CA, certainly any able to house a Turing machine, the long-term state of a random starting configuration may be algorithmically undecidable.

From this fundamental undecidability, emergent properties generally, and strong reliance on initial conditions for some CA, it is tempting to suggest that the behavior of such CA is fundamentally unpredictable. This, as Israeli and Goldenfeld have shown, is not necessarily the case.[2] Via coarse-graining the CA – losing some information by reducing the system size and lengthening the timesteps, in exchange for getting a new CA which in all cases examined was at less or at most equally complex – it may be possible to determine some long-term aspects of the CA's behavior.

Suppose that the system's initial configuration is a_0, that P Is a map that projects the old grid to the coarse-grained grid, that f_A and f_B are the initial CA map and the new, coarse-grained CA map, and finally that T is number of timesteps in the initial system per each step in the coarse-grained system. Then for the coarse-graining to be meaningful it must satisfy the commutativity condition $P \cdot (f_A)^T \cdot a_0 = f_B \cdot P \cdot a_0$ for all initial conditions a_0.[2]

Now consider site x_n in a one-dimensional grid of boolean values. Israeli and Goldenfeld determined, as one example, that Wolfram's[3] rule 105 (taking $x_n \to \overline{x_{n-1} \oplus x_n \oplus x_{n+1}}$) can be coarse-grained with timescale T=2 to rule 150 (taking $x_n \to x_{n-1} \oplus x_n \oplus x_{n+1}$) under the projection $(x_n, x_{n+1}) \to \overline{x_n \oplus x_{n+1}}$ where the bar represents logical NOT and $\oplus$ represents logical XOR. The result is that the interesting long-term behavior of rule 105 is preserved, even though information is lost in the coarse-graining:

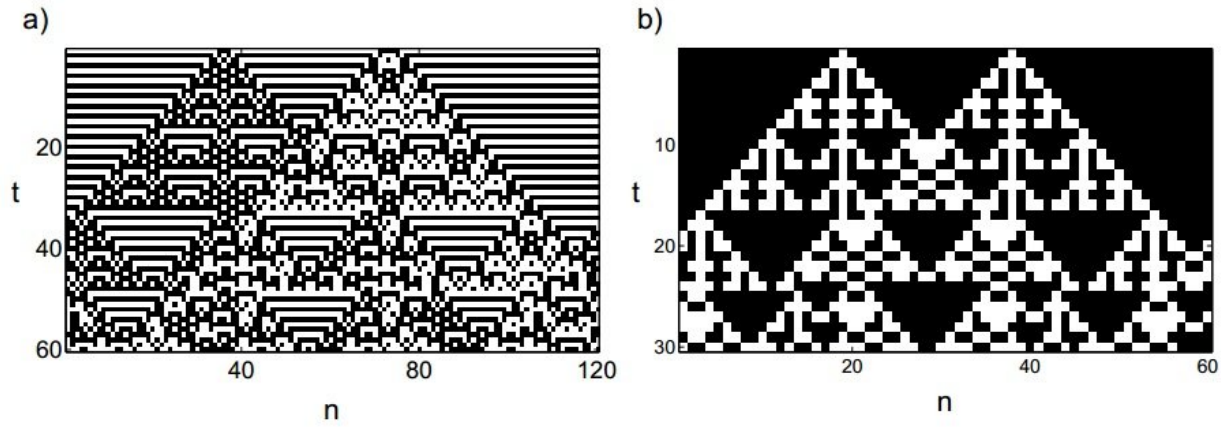

Fig 2. (a) Rule 105 and (b) Rule 105 coarse-grained to Rule 150[2]

While the possibility such a projection might have been assumed to be unlikely, such coarse-grainings were found for 240 of Wolfram's 256 simple CA rules[2], including a (trivial) case of an undecidable CA coarse-graining to a trivial decidable CA. These results imply that the emergent behavior of CA is not necessarily unpredictable, and that it may be possible to determine the interesting features or physical implications of some CA even if the CA itself is undecidable.

2. CA modeling in physical systems

2.1 The Nagel-Schreckenberg CA model for traffic flow

The basic Nagel-Schreckenberg cellular automata model, introduced in 1992, represents a one-lane road as a sequence of discrete sites with periodic boundary conditions occupied by cars with discrete velocity values. [6] Each “car” obeys simple and intuitive rules; it slows down to avoid hitting the car in front of it, and will accelerate whenever possible to reach a universal speed limit. To simulate the random slowdowns and stops that can cause traffic jams in real life, each car also had a fixed probability to randomly slow down during a timestep. Nagel and Schreckenberg showed that these simple CA rules yield results which closely resemble real freeway traffic data; below, a number represents the velocity of a car at that site:

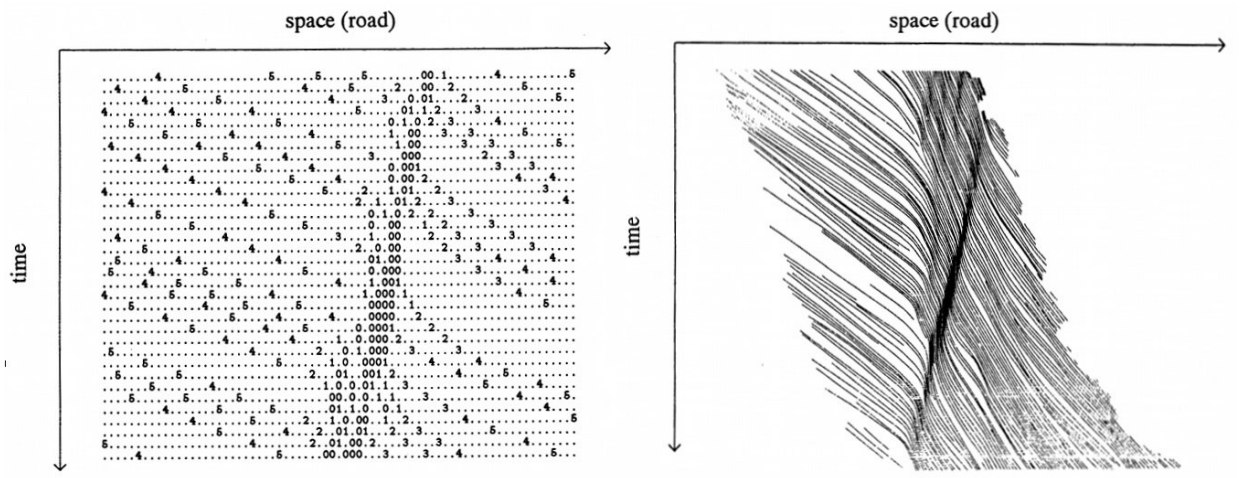

Fig. 3: Simulated traffic with a density of 0.1 cars per site (left) and trajectories of cars from aerial photography (left) [6]

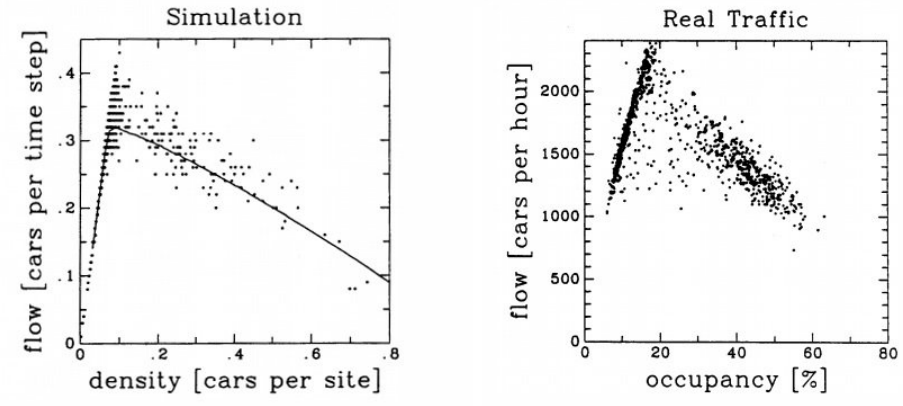

Fig. 4: Simulated and real traffic flow data, where on the right occupancy is defined by the percentage of road covered by vehicles [6]

The NS model has several advantages over approaches based on traffic flow theory, such as computational speed and the results being more easily understood from the perspective of a given individual driver. However, it by no means represents a complete picture of traffic. Other than the obvious limitations such as the initial 1992 model representing only a one-lane loop of road and failure to demonstrate some aspects of traffic flow such as metastability (which have been addressed by modified versions of the NS model), the model is itself inherently unphysical.

The most common accusation is that cars in the model come to a stop essentially instantly, decreasing their speed from its maximum value to zero in a few short timesteps when necessary to avoid a collision[7]. Whether such a rapid deceleration is physically possible or not, human drivers would not generally have the reaction time needed to pull it off – a failure, when one of the supposed benefits of the CA technique is to demonstrate the emergence of traffic jams from the point of view of the driver. More recently, other researchers such as Larranga and Alvarez-Icaza[7] have presented CA models with modified rulesets governing, for example, the idea of safe driving distances and emergency braking, which remain both computationally efficient and conceptually simple while still managing to reproduce most essential features of (single-lane) traffic flow.

2.2 Granular flows and CA

The dynamics of granular flows, whose particles can exhibit both liquid-like and solid-like behavior, are understood relatively poorly and remain an active area of research e.g. in soft matter physics. In theory, modified versions of the Navier-Stokes equations have been used to model the flows as a continuum,[8] and while simulations have been carried out attempting to model the interactions of the individual components of a flow such simulations rapidly become computationally expensive when dealing with, for example, hundreds of particles. Cellular automata have been considered as an alternative simulation model for granular flow primarily due to their computational efficiency; in fact, the original NS model paper for traffic flow itself drew an analogy to granular flow, in the case of sand falling through a narrow tube rather than traffic on a one-lane road.[6]

In the case of grains rotating in a shear cell, Jasti and Higgs III[8] attempted to simplify these simulations by using a CA or lattice gas approach for a shear cell experiment, discretizing space into lattice sites which may or may not be filled with particles, and discretizing the velocities of each particle such that they can move only to one of their eight neighboring sites in each timestep.

The rules are relatively simple, if not particularly realistic. Collisions between particles are handled elastically and modeled as well as possible

given the ability of particles to only move in 8 directions. The system itself, as a shear cell, is taken to have two boundaries, one which is stationary (which particles simply reflect off of) and one which is moving (which imparts some forward velocity to colliding particles not already moving with it). Finally, moving to an adjacent cell each timestep represents the maximum "velocity" for a particle; particles taken to be moving "slowly" may occupy the same lattice site for several timesteps before moving. Several other variables such as a roughness factor were also present, used during calculating the effects of a collision.

The results of this approach are more ambiguous than those for modeling traffic when compared to results from continuum-modeling theory.

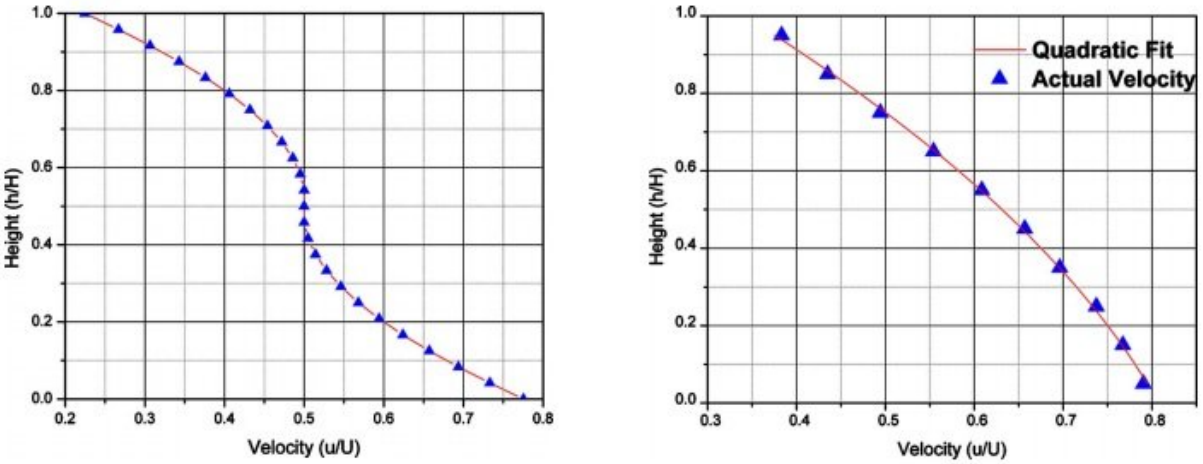

Fig 5. Height vs. velocity for continuum theory (left) and CA simulation (right). H and U are the height and velocity of the shear cell, respectively.[8]

Jasti and Higgs III note that the CA height vs. velocity graph lacks the non-shearing granular flow center predicted by theory, and that its near-linear profile more closely resembles that of a Couette flow for a Newtonian fluid. While success is claimed in other areas, such as the CA model producing, as in theory, slippage at the boundaries and a higher solid fraction near the center of the shear cell, the qualitatively different behavior inside the shear cell seems like it should be some cause for concern, as the two models are in effect predicting different physical behavior, with no clear indication of which is right or why.

The paper says such differences are the result of, for example, the CA simulation being "discrete in nature,"[8] which seems to beg the question. This represents a difficulty in working with CA; when all macroscopic behavior emerges from locally-defined interactions, it may be significantly less clear what causes a deviation from either experiment or theory. That being said, despite the discrepancy from continuum theory in one aspect, the

agreement of CA modeling with continuum modeling in others implies the results may be improved with more work. That there is any agreement at all is noteworthy given the unrealistic, idealized results for particle collisions, and it is likely that any CA model, if shown to be accurate, would be much faster computationally than continuum modeling.

3. Caveats on interpreting CA

While one of the greatest advantages of CA is the emergence of macroscopic behavior from local behavior, this same macroscopic behavior can sometimes be misleading, in particular when the reason it emerges is either not apparent or left unexamined. As an example, we briefly consider the density classification task (DCT). The goal of the DCT is to create a CA that accurately converges a system to more common of two boolean values in the initial state. That is, if the system is composed of 0s and 1s, and there are more 0s than 1s initially present in the system, the CA should ultimately result in every cell containing a 0 (likewise all cells should converge to 1 if there are initially more 1s than 0s). While an exact solution is impossible for a large enough system, there are a number of CA that exhibit high accuracy (~80% of random initial conditions or more).

Marques-Pita and Rocha performed a detailed analysis on two well known DCT CA (F_{GKL}', the mirror rule of a CA developed by Gács, Kurdyumov, and Levin, and F_{GP}, developed from genetic programming) that, while seemingly exhibiting drastically different macroscopic behavior during their evolution, had many similarities both in terms of accuracy and rulesets. While we will not go into the specifics of their analysis (the CA are much more complex than those discussed so far, each requiring information about both the original cell and its 6 nearest neighbors), they were able to show that the two CA had essentially identical rulesets, with the only difference being that F_{GP} causes cells to undergo state changes in several situations additional to those causing state changes in F_{GKL}'.[9] Therefore, they describe F_{GP} as being a more general case of F_{GKL}'; and, in fact, it has slightly higher accuracy.

However, the intermediate macroscopic behavior of the two CA is enormously different. As cells in F_{GP} undergo more state changes, FGP generates significantly more “domains,” defined as topologically distinct regions. However, in most cases none of these additional domains has much effect, implying that despite the increasingly complex macroscopic behavior, new information is rarely being carried through; instead, it is either discarded or simply duplicate information already present in an F_{GKL}' simulation or elsewhere in the system. In this sense, very little of the new complex emergent behavior is ultimately important.

Though this is a different approach, it yields a similar result to that of

Israeli and Goldenfeld, namely that the important long-term behavior of a CA could often be found even after eliminating “redundant” information and degrees of freedom through coarse-graining.[2] Not all of the behavior of a given CA is necessarily important or relevant to a system. Marques-Pita and Rocha suggest “too much attention [is paid] to the 'spots' and 'stripes'”[9] of CA, and Israeli and Goldenfeld advise researchers to focus their attention on the “physically relevant, coarse-grained degrees of freedom” when working with CA.[2]

4. Conclusions

Other than their inherent interest as mathematical objects and computers, CA show promise for modeling various physical systems and problems. They are capable of demonstrating rich emergent behavior from a handful of simple rules based on local information only, which is familiar to anyone who has worked with a system dominated at the microscopic level by local (especially nearest neighbor) effects. When modeled correctly, they are therefore capable of exhibiting emergent phenomena even when theory does not yet exist, is not fully understood, or is computationally expensive. All traditional lattice gas automata may be rewritten as CA, and many CA (or at least simple CA) exhibit the curious property of being able to be directly rewritten in terms of a different, often simpler CA via course graining.[2]

However, CA are not necessarily suited for all problems and all applications. Because the rules for CA are ideally generated without needing to know the results of theory – which will be necessary if CA are to be trusted and useful in the absence of theory – the local rules governing the evolution of the system are frequently far from perfect, as in the granular flow example of 2.2. In some applications, the emergent phenomena associated with CA tend may be surprisingly robust, as some of the qualitative behavior of the granular flow model (such as solid fraction) indicate.[8] However, in cases where the emergent phenomena does not match expectations (such as average velocity in the granular flow), the fact that macroscopic behavior emerges naturally from local processes implies figuring out what's “wrong” with a given a CA model may be very difficult.

Fixing the model may be as simple as making the rules more accurate to those in a real system, but in some cases being unphysical (such as unrealistically rapid deceleration in the NS model) may have little apparent impact at all on the macroscopic behavior. Worse, since CA are capable of generating patterns which seem complex at first glance but which simply carry redundant information, anyone looking too closely for emergent behavior and patterns in a given CA may be focusing on physically meaningless computational data if they are not careful in determining the most relevant physical degrees of freedom.

Therefore, while CA are computationally efficient and capable of demonstrating interesting emergent behavior with carefully constructed rules even in the absence of theory, writing CA to model a system is not an excuse to avoid careful consideration of the physical outcomes or the local properties of the rules chosen. That is, it may be expected to be rare to write a CA which competently models all interesting physical phenomena in a system without extensive modifications, and unless a given theory is incomplete the theory may in general be expected to yield more accurate results than the relatively simplistic discrete CA model. Still, the speed of CA modeling makes it a valuable tool, and even in cases where CA outcomes are not entirely realistic the qualitative results may yield novel supplementary data or counterpoints. Research is still ongoing to improve existing CA models, combine CA systems with other mathematical and conceptual models, and better understand the implications and nature of CA in general.

References

[1] “When - and how - can a cellular automaton be rewritten as a lattice gas?” T. Toffoli, S. Capobianco, P. Mentrasti. Theoretical Computer Science, 403 (2008)

[2] “Coarse-graining of cellular automata, emergence, and the predictability of complex systems.” N. Israeli, N. Goldenfeld. Phys. Rev. E, 73 (2006)

[3] “A New Kind of Science.” S. Wolfram (2002)

[4] “Symmetry and complexity in dynamical systems.” K. Mainzer. European Review, 13 (2005)

[5] “Emergence of System-Level Properties in Biological Networks from Cellular Automata Evolution.” B. Vescio, C. Cosentino, F. Amato. 18th Mediterranean Conference on Control & Automation (2010)

[6] “A cellular automaton model for freeway traffic.” K. Nagel, M. Schreckenberg. Journal de Physique I, Volume 2, Issue 12 (1992)

[7] “Towards a Realistic Description of Traffic Flow based on Cellular Automata.” M.E. Larranga. L. Alvarez-Icaza. 14th International IEEE Conference on Intelligent Transportation Systems (2011)

[8] “A Lattice-Based Cellular Automata Modeling Approach for Granular Flow Lubrication.” V. Jasti, C.F. Higgs III. World Tribology Congress III, 1 (2005)

[9] “Schema Redescription in Cellular Automata: Revisiting Emergence in Complex Systems.” M. Marques-Pita, L. Rocha. 2011 IEEE Symposium on Artificial Life (2011)