Interdisciplinary Topics in Complex Systems: Cellular Automata, Self-Organized Criticality, Neural Networks and Spin Glasses

Branislav K. Nikolić
Department of Physics and Astronomy, University of Delaware, U.S.A.

PHYS 460/660: Computational Methods of Physics
Can few simple laws of CAM produce rich complex behavior?

- Cellular Automata: invented by von Neumann and Ulam in 1948 as an idealization of biological self-reproduction → this is why lattice site are called cells.
- More recently CAM have been applied to systems ranging from fluids to galaxies.
- CAM are example of a discrete dynamical system that can be simulated exactly on the computer.
- Discrete space, time, and physical quantities with integer values that are updated according to the local rules:
  1. Space is discrete and there is a regular array of sites (cells). Each site has a finite set of values.
  2. Time is discrete, and the value of each site is updated in a sequence of discrete time steps.
  3. The rule for the new value of a site depends only on the values of a local neighborhood of sites near it.
  4. The variables at each site are updated simultaneously ("synchronously") based on the values of the variables at the previous time step.

- CAM vs. discretized PDE: No accumulation of round-off errors.
Invented by J. Conway in 1970 to produce fascinating patterns via few simple rules:

1. Each cell in the grid has a neighborhood consisting of the eight cells in every direction including diagonals.
2. A live cell with two or three live neighbors stays alive (survival).
3. A dead cell with exactly three live neighbors becomes a live cell (birth).
4. In all other cases, a cell dies or remains dead (overcrowding or loneliness).

The Game of Life is not a game in the conventional sense. There are no players, and no winning or losing. Once the "pieces" are placed in the starting position, the rules determine everything that happens later.

Nevertheless, Life is full of surprises → In most cases, it is impossible to look at a starting position and see what will happen in the future. The only way to find out is to follow the rules of the game.

Life is one of the simplest examples of what is sometimes called "emergent complexity" or "self-organizing systems." It is the study of how elaborate patterns and behaviors can emerge from very simple rules.
Are There Universal Signatures of Complex Behavior?

- **Punctuated Equilibrium**: There are long periods of relative stasis punctuated by crises ("avalanched") of various sizes.

- **Power Laws**: The relationship between the sizes of these avalanches can be expressed in a simple exponential equation. There are no singular explanations for large events: the same forces that made the Dow Jones average drop five points yesterday also caused the crash of 1987.

- **Fractal Geometry**: Where a system exists in space, it is self-similar on all scales.

- **1/f Noise**: When a system evolves over a time, the record of evolution is also fractal.

- **Self-Organized Criticality**: explain all this phenomenological features by finding simple dynamics that spontaneously drives system into a critical state characterized by power laws.
Fluctuations in Physical Quantities: Noise and Its Power Spectrum

$I(t)$
Intensity

White Noise: Uncorrelated

Flicker Noise: Moderately Correlated

Brown[ian] Noise: Strongly Correlated

$S(f) \sim |I(f)|^2$
Power Spectrum

$1/f^0$

$1/f$

$1/f^2$

$\log S$

$\log f$
Main Idea: In many complex systems large events are part of a distribution of events and do not depend on special conditions or external forces.

If $s$ represents the magnitude of an event, such as the energy released in an earthquake or the amount of snow in an avalanche, then a system is said to be critical if the number of events follows power law (for $\alpha \approx 1$ there is one large event of size 1000 for every 1000 events of size 1).

$$N(s) \sim s^{-\alpha}$$

Power laws are scale invariant: $s \to bs \Rightarrow \tilde{N}(s) = \tilde{A}s^{-\alpha}$, $\tilde{A} = Ab^{-\alpha}$

Combining large number of independently acting random events gives Gaussian - no scale invariance and practically no large events:

$$N(s) \sim e^{-(s/s_0)^2} \iff s \to bs \Rightarrow \tilde{N}(s) \sim e^{-(s/s_1)^2}, s_1 = s_0 b^2$$
A model of dissipative dynamical system with local interacting degrees of freedom and many metastable states.

\[
\begin{align*}
    & z(x, y) \rightarrow z(x, y) - 4 \\
    \text{If } z > 3 \Rightarrow & \begin{cases} 
    z(x, y) \rightarrow z(x \pm 1, y) + 1 \\
    z(x, y + 1) \rightarrow z(x, y \pm 1) + 1
    \end{cases}
\end{align*}
\]

It self-organizes (without fine tuning of parameters, like temperature in the case of thermal critical phenomena, and independently of the initial conditions) into a critical state (scale-invariance) that is attractor of the dynamics and robust with respect to variations of parameters and the presence of quenched randomness.

<table>
<thead>
<tr>
<th><strong>SOC</strong></th>
<th><strong>Ising Magnet</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous spontaneous flow</td>
<td>Magnetization ( M )</td>
</tr>
<tr>
<td>Current of incoming particles</td>
<td>Magnetic Field ( H )</td>
</tr>
<tr>
<td>( \left&lt; z \right&gt;_c - \left&lt; z \right&gt; )</td>
<td>Reduced Temperature ( T / T_c )</td>
</tr>
</tbody>
</table>

Constant average height in the critical state \( \rightarrow \) the probability that activity will die is on the average balanced by the probability that activity will branch out.
Clusters in SOC state of Sandpile CAM

$s = 25 \ (a = 45)$

$s = 197 \ (a = 270)$
Power Laws in SOC state of Sandpile

Ensemble vs. Time Average

PHYS 460/660: Interdisciplinary Topics in Complex Systems
Size of the Cluster vs. Time

PHYS 460/660: Interdisciplinary Topics in Complex Systems
%program soc_relax.m

%PARAMETERS
size=100; %lattice size

%VARIABLES
flag=0; % logical variable

rand('twister',17);
z(2:N+1,2:N+1)=5+int8(5*rand(N,N));
time=1;
while 1 % infinite loop
    flag=0;
z(1,:) = 0; z(:,1) = 0;
z(size+2,:) = 0; z(:,size+2) = 0;
    for i=2:size+1
        for j=2:size+1
            if (z(i,j)>3)
                z(i,j)=z(i,j)-4;
                z(i+1,j)=z(i+1,j)+1;
                z(i-1,j)=z(i-1,j)+1;
                z(i,j+1)=z(i,j+1)+1;
                z(i,j-1)=z(i,j-1)+1;
                flag=1;
            end
        end
    end
    if (flag==0)
        break
    end
    zaverage(time)=sum(sum(z(2:size+1,2:size+1)))/(size*size);
    time=time+1;
end % infinite loop

for i=2:size+1
    for j=2:size+1
        if (z(i,j)>3)
            z(i,j)=z(i,j)-4;
            z(i+1,j)=z(i+1,j)+1;
            z(i-1,j)=z(i-1,j)+1;
            z(i,j+1)=z(i,j+1)+1;
            z(i,j-1)=z(i,j-1)+1;
            flag=1;
        end
    end
end
if (flag==0)
    break
end
zaverage(time)=sum(sum(z(2:size+1,2:size+1)))/(size*size);
time=time+1;
end % infinite loop
t=1:time-1;
plot(t,zaverage);
Neural Network vs. Human Brain

- **Neural Net** is an artificial representation of the human brain that tries to simulate its learning process.
- The term “artificial” means that neural nets are implemented in computer programs that are able to handle large number of necessary calculations during the learning process.

- **Human brain** consists of a large number (more than a billion) of neural cells that process information. Each cell works like a simple processor and only the massive interaction between all cells and their parallel processing makes the brain’s abilities possible:

  Neuron consists of a **core**, **dendrites** for incoming information and an **axon** with dendrites for outgoing information that is passed to connected neurons. Information is transported between neurons in form of electrical stimulations along the dendrites. Incoming information that reaches dendrites is added up and then delivered along axon to the dendrites at its end, where the information is passed to other neurons if the stimulation has exceeded a certain threshold. In this case, the neuron is said to be activated. If the incoming stimulation had been too low, the information will not be transported any further. In this case, the neuron is said to be inhibited. The connections between the neurons are adaptive, what means that the connection structure is changing dynamically. It is commonly acknowledged that the learning ability of the human brain is based on this adaptation.
Components of Neural Nets

Structure of a neuron of a Neural Net:

Three neuron layer Neural Net:

- Neural nets are being constructed to solve problems that cannot be solved using conventional algorithms, such as optimization and classification problems:
  - pattern association
  - pattern classification
  - regularity detection
  - image processing
  - speech analysis
  - optimization problems
  - robot steering
  - processing of inaccurate or incomplete inputs
  - quality assurance
  - stock market forecasting...
The Hopfield model consists of a single layer of processing elements where each unit is connected to every other unit in the network other than itself.

The connection weight matrix $W$ of this type of network is square and symmetric, i.e., $W_{ij} = W_{ji}$ for $i, j = 1, 2, ..., m$. Each unit has an extra external input $I_i$. This extra input leads to a modification in the computation of the net input to the units:

$$\text{Input}_j = \sum_{i=1}^{m} x_i W_{ij} + I_j$$

**Discrete Hopfield Model:**

$$x_i(t+1) = \begin{cases} 
+1 & \text{if Input}_i > \theta_i \\
x_i(t) & \text{if Input}_i = \theta_i \\
-1 & \text{if Input}_i < \theta_i 
\end{cases}$$

$E = -\frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} x_i W_{ij} x_j - \sum_{i=1}^{m} x_i I_i - \sum_{i=1}^{m} x_i \theta_i$

After many updates, discrete Hopfield Neural Net converges toward a local minimum of the energy function that corresponds to stored pattern. The stored pattern to which the network converges depends on the input pattern and the connection weight matrix.
Content Addressable Memory has to be able to: Store, Recall, and Display Patterns

To store $m$ patterns into the memory, choose couplings according to:

$$W_{ij} = \frac{1}{M} \sum_{m} x_i(m) x_j(m)$$

If the number of stored patterns exceeds $0.13N$, the energy landscapes change dramatically (stored patterns become unstable and the system ceases to function as a memory) → like phase transition in spin glasses

A new pattern $p$ can be learned by adding a small contribution to interactions:

$$W_{ij}^{\text{new}} = \beta W_{ij}^{\text{old}} + \alpha x_i(p) x_j(p)$$

where $x_i(p)$ is the new pattern, $\alpha$ is parameter that controls how fast the learning should occur, and the value of $\beta$ can be adjusted to allow for the fading of old memories.
Spin Models Revisited: Frustration

For $J>0$ we have a **ferromagnet** → energy is smallest when neighboring spins are aligned, either ↑↑ or ↓↓. Ground state ($T=0$ lowest energy state) is when all spins are aligned, or parallel (all +1 or all -1).

For $J<0$ we have an **antiferromagnet**. Energy is smallest when neighboring spins are opposite, either ↑↓ or ↓↑. Ground state is when all spins are antiparallel, that is, a checkerboard configuration for a square 2D lattice.

**FRUSTRATION:**

- Consider an Ising ferromagnet defined on a triangular lattice. This is similar to square lattice, except every site now has six nearest neighbors instead of four. Triangular lattice ferromagnet has a phase transition with the **same critical exponents** (the same universality class) as the square lattice ferromagnet.
- This is not the case for the **triangular lattice antiferromagnet**. Behavior is very different, since not all neighboring sites can simultaneously have the lowest energy. Links are unsatisfied. The model is said to be frustrated → Dealing with frustration is a difficult problem!
Spin Glasses

- Generalize the interactions in the Ising spin model so that interaction strength becomes a variable dependent on the lattice site:

\[ E = -\sum_{\langle ij \rangle} J_{ij} S_i S_j, \quad J_{ij} = +J \text{ or } -J, \quad J > 0 \]

- Suppose we choose the sign at random for every \( i,j \) link. This will introduce frustration since not all links around certain plaquettes can be satisfied.

- Randomness and frustration are hallmarks of real systems called spin glasses.

- Certain metallic alloys have magnetic interactions which oscillate in value as a function of the separation of the different atoms.

- These alloys can be created in an amorphous state - the positions of the atoms are random and disordered like a glass, rather than regular and ordered like a crystal.

- The magnetic interactions between spins are thus sometimes ferromagnetic, sometimes antiferromagnetic → this is what is known as a spin glass.
Transition to a “glassy” phase at low temperatures.

Susceptibility has a cusp rather than a divergence at $T_c$.

Very large number of metastable states (local minimum of energy) in spin glass phase.

Long equilibration times (logarithmic rather than exponential relaxation to equilibrium) - remanent magnetization $M \sim 1/\log t$.

Degeneracy (or near-degeneracy) of ground state, with many configurations having (nearly) the lowest energy value.

Finding ground states is a difficult problem with NP computational complexity!

Main properties of real spin glasses (e.g. amorphous alloys) can be described fairly well by $\pm J$ spin glass Ising model on a regular crystalline lattice.
Simulated Annealing

- In order to avoid the metastable states produced by quenching, metals are often cooled very slowly, which allows them time to order themselves into stable, structurally strong, low energy configurations. This is called annealing.
- Annealing gives the system the opportunity to jump out of local minima with a reasonable probability while the temperature is still relatively high.

- We can adopt the same approach in Monte Carlo simulations. We start with a random configuration at a very high temperature, and then reduce the temperature “very slowly” until we reach the desired low temperature. This should result in a thermalized configuration.

- In many regular models, it is possible to deduce the form of the ground state configurations. However, for disordered frustrated systems, such as spin glasses, finding a ground state is very difficult (in fact it is NP hard - there is no polynomial time algorithm for finding a solution).

- In practice, ground states are found using simulated annealing. This involves starting from a number of different random initial configurations, and slowly cooling them down to zero temperature. This produces local energy minima that are close to, and possibly equal to, the global minimum.
Simulated annealing works well for many combinatorial optimization problems.

Graph Partitioning: In designing computer chips, there are a number of connected elements. If there are too many elements to fit on one chip, they need to be partitioned onto (for example) two chips. Since wires between chips are expensive, the number of connections between chips should be minimized. Since silicon area is also expensive, the number of elements on each chip should be about the same. This is an example of a graph partitioning problem. If we consider the elements as vertices and the wires as edges, the problem is to partition the graph into two equal sets while minimizing the number of edges between each set (NP-hard problem).

- Introduce a variable $S_i$ which is +1 if the element is on the first chip, and -1 if it is on the second. Introduce a connectivity matrix $C_{ij}$ if $i$ and $j$ are connected, 0 otherwise. The penalty for connections between chips is $H$. We want roughly the same number of elements in chip 1 and chip 2, so we want to make $\sum_i S_i$ as close as possible to zero:

$$H = -\sum_{ij} C_{ij} S_i S_j + \mu \left( \sum_i S_i \right)^2 = \text{constant} - \sum_{ii} J_{ii} S_i S_j, \quad J_{ij} = C_{ij} - 2\mu$$

The Traveling Salesman Problem: A salesman has to visit $N$ cities, and wants to take the shortest possible route. Here, the cost function is the length of the tour. To change a configuration (a particular tour) is not as simple as a single spin flip in the spin glass problem, or moving an element from one side to the other in the graph partitioning problem. Each tour can be presented as a permutation of the numbers 1 to $N$, which represent the cities. The simplest change to the tour is to swap pairs of cities, and measure the change in the tour path.