Hidden Markov Model and ε-Machine Representation of 2D Ising Model

Yuchen Liu
Department of Chemistry
University of California, Davis
ycnliu@ucdavis.edu

Abstract

The 2D Ising model is a particular example of a thermodynamic system, and it's a model system for understanding phase transitions and has been generally learned and studied by scholars. Originally it was used for modeling the change of magnetization, but the model is also applicable to model various topics. The Ising model is intrinsically stochastic and can be understood as a Markov process. In this project, we modeled a 2D Ising model with a hidden Markov model and reconstructed the corresponding ε-Machine out of several tests. The block entropy and entropy rate are for the process are also obtained.

Introduction

The Ising problem was originally formalized and developed to simulate the phase change of a piece of a lattice with spin in each cell. In this situation, the total energy of the lattice is determined by the spin configuration of each neighborhood of cells, and the temporal evolution of our spin will follow the thermodynamic distribution. From a perspective of the information theory, the Ising model is intrinsically a Markov process, since it has the property that the next temporal state is only decided by the one previous state and irrelevant to further future, thus it is memoryless. The 2D Ising model is often simulated by Metropolis-Hastings Monte Carlo (MHMC) simulation, in which during each time step a spin is flipped and the change is kept by thermodynamic probability.

The behavior of our model or the system in each time step is easy to understand and track and the near future is predictable. But the reason that this model is so fascinating and has been studied by generations of scholars is its intrinsic stochasticity for a long future and capability of modeling such a complex system. Out of physics, the 2D model has also been used to describe some sociological system such as the vote, that each cell can represent an individual with its spin being the candidate he/she will vote to. Such expansions to the interpretation of the Ising model also demonstrates the model's capability of simulating more complicated systems.

The general Markov model (not Hidden Markov model) has been used to fully model the 2D MHMC simulation of the Ising problem by parsing the successfulness of flipping a spin at each time step to its word distribution. Previous work by Iglovikov [2], shows that the entropy rate and statistical complexity of the Monte Carlo simulation are equal to the thermodynamical entropy of the system that we simulate and the excess entropy of the system can be used as a natural order parameter to find phase transition temperature. In this work, I'm interested in finding a way to describe not the detailed change of configuration of each spin in the whole system but the collective property of the system as a whole by parsing the change of average magnetization (average spin) of the system per an arbitrary period to be the word distribution, then building a hidden Markov Model ε-Machine based on such setup and finally look at the power of predictability and transition probabilities of the machine.

Background

The 2D Ising Model

In simulation programs, a 2D Ising model is represented by a N x N matrix, with N being the dimension of the system to be modeled, a square matrix is often used by convention. The matrix elements of the N x N matrix can be either +1 or -1 representing two opposite kinds of spins. As a physical system, the Hamiltonian of the whole system is given by

$$\mathcal{H} = -h\mu_0 \sum_i \sigma_i - J \sum_{adj.} \sigma_i \sigma_j$$

σ is the direction of the ith particle's spin which can be either "up" (+1) or "down" (-1), h is the magnitude of the applied magnetic field (not Planck's constant), and J is the spin-spin coupling term between adjacent spins on the lattice. The first term is barely the sum of all spins times a constant, and the second term is an interaction term that makes things interesting.

Metropolis-Hastings Monte Carlo Update Algorithm

The update algorithms itself is simple at each time step

```
For i in number_of_steps:
    do:
    old_matrix = copy(matrix)
    old_energy = Hamiltonian(matrix)

    x, y = random(range = N, size = 2)
    matrix(x, y) = -1 * matrix(x, y)
    new_energy = Hamiltonian(matrix)

    if new_energy < old_energy:
        ## accept the change and move on
        continue
    else: ## keep the change by chance
       factor = 1 / (1 + e^((new_energy − old_energy) / kbT))
       r = random(range = (0, 1), size = 1)
       if r > factor:
           ## accept the change
           continue
       else: ## reject the change
           matrix = old_matrix
           continue
```

With such an algorithm, the system tends to find a lower energy state that is suggested by thermodynamics. For a large system, the global minimum is hard to solve, but after millions of iterations with this algorithm, we can see come local minimum and the change between states.

The visualization of a 2D Ising model after an evolution of 10000 steps shows below

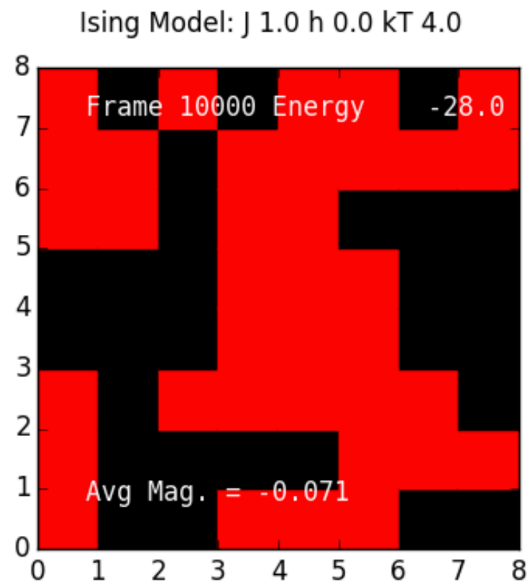

Figure 1. A 2D Ising modeling example, with red cells representing matrix elements with spin "up", and black cells being "down" spins. The first term in the Hamiltonian is ignored by this simulation, and the average magnetization (average spin) of this configuration is -0.071, which is almost neutral.

Methods

In this project, several simplifications have been applied to the model. First, the first term in the Hamiltonian is ignored by setting the magnitude of the external field to be zero. The reason is, with an applied external magnetic field and effective first term in the Hamiltonian, the system will tend to convert to have the same spin. That is to say that all spin "up" is a trivial solution and the system to be studied will become less chaotic and less interesting to look at. All cells are initialized by spin "up" (+1) as a convention.

The experiment is designed under a temperature of kT = 4 to gain more stochasticity and a nearly neutral average spin. The temperature dependence of the average spin (average magnetization) with respect to time steps is described below

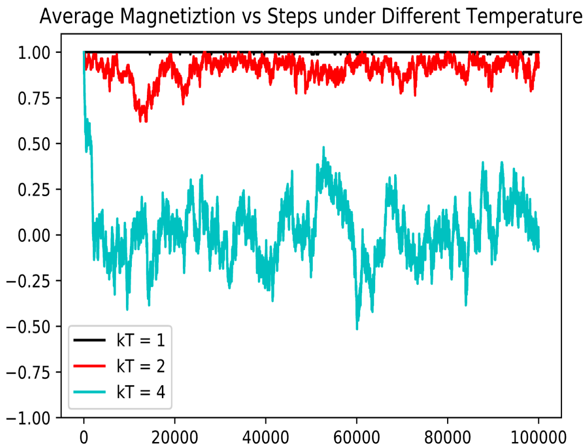

Figure 2. The change of average magnetization with respect to number of steps under different temperatures.

As is described in Figure 2. Under some low temperatures, the systems always have average spins near 1, all spin "up". As the temperature increases to kT = 4, the system starts to fluctuate around neutral total spin with some periodicity and patterns.

Unlike the Markov approach on understanding an Ising model, this project tries to interpret the system using a hidden Markov process and the MHMC process is parsed in the below way to generate the word distribution for reconstructing the corresponding ε-Machine. The MHMC modeling runs for 250,000 iterations in total for each test. The average spin of the system is uniformly sampled per 50000 steps that generate a word of length-5 per test run. The letters to be observed in each word are either "0" or "1", if the average spin at the current sampling step is greater than 0, it reads a letter of "1", otherwise, it reads a "0". For a length-5 output generated after a test run

“1” -> “1” -> “0” -> “1” -> “0”

It means at step 50000, 100000, 150000, 200000, and 250000, the average spin is greater than 0, greater than 0, less than 0, greater than 0, and less than 0, respectively. That is to say that the model is designed to collect and summarize the behaviors of the MHMC process for the first 250,000 steps. This range of iteration is chosen based on the balance of information to learn and computation time consumption. From Figure 2., we can see that this number is sufficient to see the turning point between all spin-up initial states and fluctuating states.

There are 100 test runs in total and therefore there are one hundred length-5 words that we collected to generate the word distribution. Afterward, the hidden states are identified with the depth-5 parse tree and depth-2 morphs. Then, the corresponding hidden Markov model with transition probabilities between states is established with its block entropy and entropy rate calculated by CMPy.

The block entropy is defined by

$$H(L) = - \sum_{s^L \in \mathcal{A}^L} \Pr(s^L) \log_2 \Pr(s^L) \ .$$

With L being the word length, that is 5 in this work, and $Pr(s^L)$ being the probability to observe each unique length-5 word. The excess entropy is generally defined by

$$h_\mu \equiv \lim_{L \to \infty} \frac{H(L)}{L} \ .$$

While we cannot possibly find the block entropy of infinity word length, $L = 5$ is used for calculating the asymptotic excess entropy instead.

Results

The parse tree of the 100 tests is successfully constructed and is shown in Figure 3. It is represented by the figure that the system always has to start from node-0 and move to node-2 by producing a "1", which is the initialization phase between 0 step and roughly 10,000 steps in Figure 2. Then the tree starts to split into nodes with close but nonequal probabilities to produce either "0" or "1", which suggests that our paring method and the choice of sampling period are proper that the parse tree successfully describe the oscillation of the average spin of the system by temporal evolution.

It’s a big system with 100 trail tests, so numerically equal splitting probabilities for all the nodes are not expected, which is to say that it is hard to find a way to transform *Figure 3.* exactly to a Markov diagram exactly unless each different node is a different state which will generate a huge number of causal states and is hardly helpful. Certain approximation is made to help the analysis: the splitting probabilities between nodes are approximated to 1 significant figure.

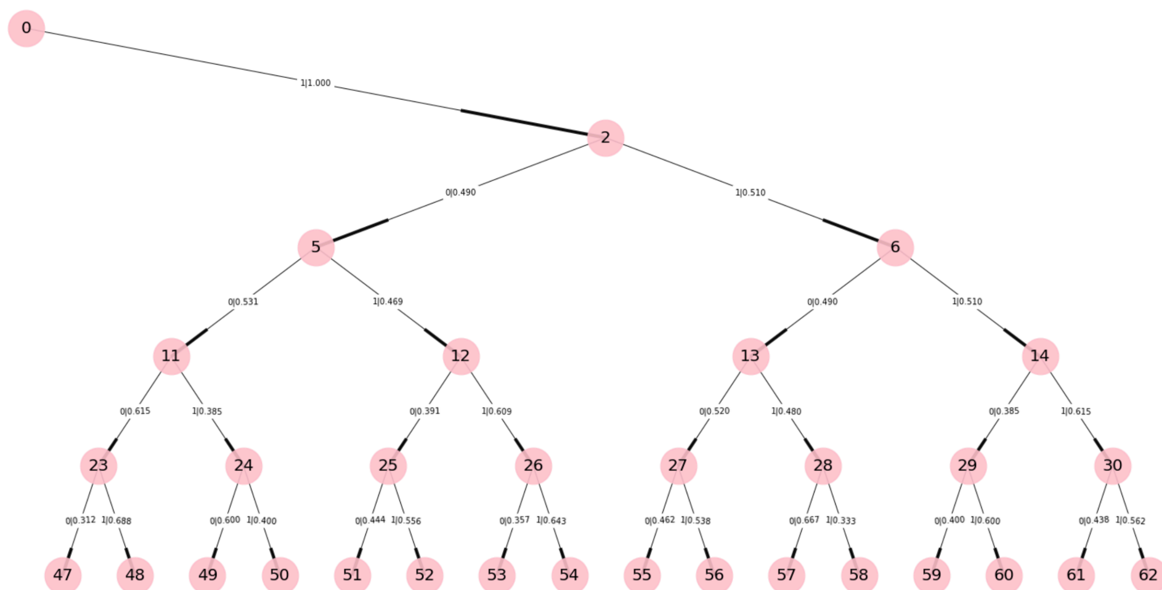

Figure 3. Parse tree of length 5 from 100 testing trials of the MHMC model. Nodes colored with pink has no explicit meaning with the numbering yet. The output letters on each arrow are separated from the conditional probability to observe the output ("0" or "1" meaning average magnetization greater or lower than zero).

With such approximation to splitting probabilities and combining nodes, a Markov diagram Figure 4. is inferred from Figure 3. In Figure 4., each node is labeled with a letter without explicit meaning since the model is a hidden Markov process, and thus states "A" to "K" are hidden states. The diagram is still not complete due to the limiting word length and diverges from the state "H", "I', "J", "K" that may or may not go back to existing states.

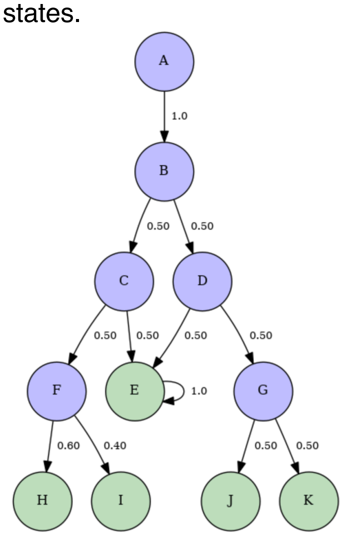

Figure 4. Inferred hidden Markov chain from MHMC simulation with 250,000 time steps.

The block entropy and entropy rate for this process are 3.93 and 0.785 respectively.

Discussion

In this project, the parse tree and its hidden Markov model for the 2D Ising model are established as expected. It is shown that the parse tree has a large level of randomness and looks like a fair-coin-shaped distribution at the first glance. However, with some approximation to the splitting/transition probabilities, we found not only many transient states but also one state, "E" that's recurrent to itself. State "E" in Figure 4. comes from nodes "12", "14", "25", "26", "29", and "30" on Figure 3. These findings are telling that the complexity of a simplified Ising model is much higher than other classical processes that have been studied in class, because most of the states are transient and don't come back, and each configuration is in a unique hidden state. It is also exciting that we found a recurrent state, "E". If more experiments can be performed with longer word length generated, we're able to see how persistent is the state "E" in the evolution of an Ising model and how is such state related to physical properties of a given Ising model. The excess entropy hμ is compared to reference thermodynamic entropy [2], 0.97 at kT = 4, the error is around 20%. The accuracy of the excess entropy is expected to be low because of the sampling method that we performed.

References

[1] Feldman and Crutchfield, *Structural Information in Two-Dimensional Patterns: Entropy Convergence and Excess Entropy*, PACS: 05.20.-y 05.45.-a 65.40.Gr 89.70.+c 89.75.Kd

[2] Iglovikov, *Information and Order Parameters in the Gauge Ising Model*, POCI class project.

[3] CMPy, http://cmpy.csc.ucdavis.edu/installation.html