黄朝晖、刘念奇

Department of Mechanics and Engineering Science
Peking University

Ising model

Mean Field Multi-Agent Reinforcement Learning

December 9, 2021

Contents

Outline for Ising model

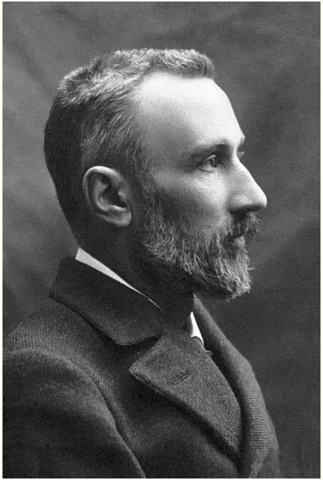
1. **Ising model**
   - Ising model in statistical mechanics
2. Mean field reinforcement learning methods

Phase Transitions

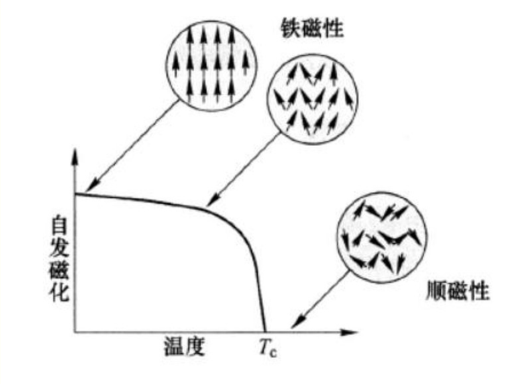
A major topic of interest in statistical mechanics (and in physics in general) is the understanding of phase transitions (e.g. freezing of water to form ice), which requires the study of interacting models.

water $\longrightarrow$ vapour

Ferromagnetic phase transition

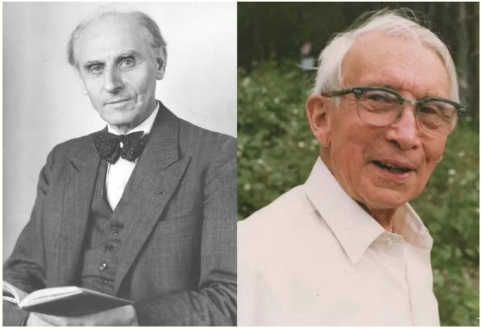


Pierre Curie (1895)



Why does the magnetism lose after heating the magnet beyond a certain temperature?

Background



Wilhelm Lenz and Ernst Ising
(Teacher-student relationship)

The Main ideas:

- Lenz considered the inside of the matter as a grid, where each node contains an atom represented by a simplified arrow.
- Ising applied the principle of the lowest energy. The magnetic forces tend to keep neighbouring arrow in the same direction, while disturbances will destroy this tendency.

Definition of Ising model

Ising model (1925)

Definition

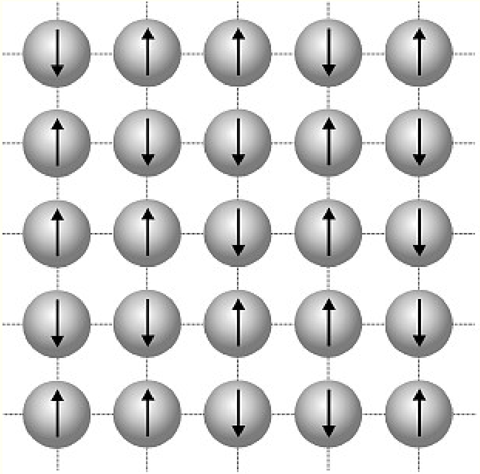
1. A periodic lattice;
2. With "spin" variables s_i, which can only take the values +1 ($\uparrow$) and -1 ($\downarrow$);
3. Every spin interacts with its nearest neighbors as well as with an external magnetic field h.

The Hamiltonian (system energy) of the Ising model is

$$H(\{s_i\}) = -J\sum_{\langle i,j\rangle} s_i s_j - h\sum_i s_i$$

The sum $< i,j >$ is over nearest neighbors. J is a constant specifying the strength of interaction.

Schemic



Ising model in 2D

It can be used to describe many physical phenomena, such as

- the order-disorder transition in alloys
- the transition from liquid helium to superfluid
- the freezing and evaporation of liquids
- the properties of glass substances
- forest fires
- urban traffic……

Statistical Mechanics

The Ising model is usually studied in the canonical ensemble. (It would be a nightmare to do it in the microcanonical ensemble.)

Theorem

In the canonical ensemble, the probability of finding a particular spin configuration $\{s_i\}$ is,

$$p\left(\{s_i\}\right) = \frac{1}{Z}\exp\left(-\beta H\left(\{s_i\}\right)\right), \quad \beta \equiv \frac{1}{k_B T}$$

where $Z = \sum_{\{s_i\}} \exp\left(-\beta H\left(\{s_i\}\right)\right)$ is the partition function.

Due to the Boltzmann factor, $\mathrm{e}^{-\beta H}$, spin configurations with lower energies will be favored.

Two parameters

We can now discuss the effect of J and h on the behavior of the spins.

- when $h > 0$, $si = +1$ is favored.
- when $h < 0$, $si = -1$ is favored.

This means that the spins wants to align with the direction of h.

- when $J > 0$, neighboring spins prefer to be parallel. (This is called the ferromagnetic model.)
- when $J < 0$, neighboring spins prefer to be anti-parallel. (This is called the anti-ferromagnetic model.)

Explanation of the transition

At low enough temperature:

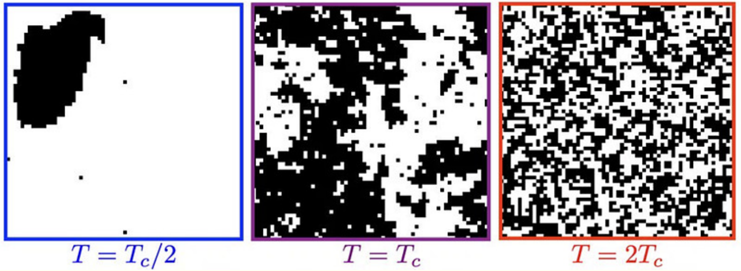
Spontaneous magnetization All spins in the Ising model will spontaneously align themselves even in the absence of the external field ($h = 0$).

At high enough temperature:

- The spontaneous magnetization is destroyed by thermal fluctuation.
- There is a critical temperature T_c, below which there is spontaneous magnetization and above which there isn't.
- There is a phase transition at T_c.

Traditional Implementations

steps	Monte Carlo simulation
0	Start with some spin configuration s_i.
1	Randomly choose a spin s_i
2	Attempt to flip it, i.e., $s_i := -s_i$ (trial)
3	Compute the energy change ΔE due to this flip.
4	If $\Delta E < 0$, accept the trial.
5	If $\Delta E > 0$, accept the trial with probability $p^{\text{acc}} = e^{-\beta \Delta E}$



Outline for Mean field reinforcement learning methods

Mean field approximation

The interactions within the population of agents are approximated by those between a single agent and the average effect from the overall population or neighboring agents.

Mean field approximation

The interactions within the population of agents are approximated by those between a single agent and the average effect from the overall population or neighboring agents.

Ising model

Let's recall the system energy in ising model,

$$E(s,h) = -\sum_i \left(h_i s_i + \frac{J}{2} \sum_{j \in \mathcal{N}(i)} s_i s_j \right)$$

where $\mathcal{N}(i)$ is the set of nearest neighbors of site i, h_i is the external field affecting site i. The probability is given by $P(s) = \frac{\exp(-E(s,h)/\tau)}{\sum_s \exp(-E(s,h)/\tau)}$, where τ is the system temperature.

Mean field approximation

$$E(s,h) = -\sum_i s_i \left(h_i + \frac{J}{2} \sum_{j \in \mathcal{N}(i)} s_j \right)$$

we can term $\left(h_i + \frac{J}{2} \sum_{j \in \mathcal{N}(i)} s_j \right)$ as the equivalent magnetic field B_i.
Due to thermal fluctuation, we compute the mean value:

$$\bar{B}_i = h_i + \frac{J}{2} \sum_j \bar{s}_j$$

If we completely ignoree the fluctuations, the average value of the spins of each spin variable is equal $\bar{s}_j = \bar{s}$ then

$$\bar{B}_i = h_i + \frac{zJ}{2} \bar{s} \equiv \bar{B}$$

Mean field approximation

The mean field theory provides an approximate solution to $\langle s_i \rangle = \sum_s s_i P(s)$ through a set of self-consistent mean field equations

$$\langle s_i \rangle = \frac{\exp\left(-\left[h_i s_i + J \sum_{j \in \mathcal{N}(i)} \langle s_j \rangle\right]/\tau\right)}{1 + \exp\left(-\left[h_i s_i + J \sum_{j \in \mathcal{N}(i)} \langle s_j \rangle\right]/\tau\right)}$$

which can be solved iteratively by

$$\langle s_i \rangle^{(t+1)} = \frac{\exp\left(-\left[h_i s_i + J \sum_{j \in \mathcal{N}(i)} \langle s_j \rangle^{(t)}\right]/\tau\right)}{1 + \exp\left(-\left[h_i s_i + J \sum_{j \in \mathcal{N}(i)} \langle s_j \rangle^{(t)}\right]/\tau\right)}$$

where t represents the number of iterations.

Stochastic Game

An N-agent (or, N-player) stochastic game Γ is formalized by the tuple

$$\Gamma \triangleq \left(\delta, A^1, \ldots, A^N, r^1, \ldots, r^N, p, \gamma\right)$$

Stochastic Game

An N-agent (or, N-player) stochastic game Γ is formalized by the tuple

$$\Gamma \triangleq \left(\delta, A^1, \ldots, A^N, r^1, \ldots, r^N, p, \gamma\right)$$

- δ the state space
- A^i the action space of agent i
- r^i the reward function for agent i
- p the transition probability
- γ the reward discount factor across time
- s an initial state
- a the joint action

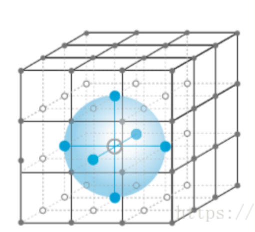
- policy: $\pi^j : \delta \to Q\left(A^j\right)$, where A^j is the collection of probability distributions over A^j.
- $\pi \triangleq \left[\pi^1, \ldots, \pi^N\right]$: the joint policy of all agents, then the value function $v_\pi^j(s) = v^j(s;\pi) = \sum_{t=0}^{\infty} \gamma^t \mathbb{E}_{\pi,p}\left[r_t^j \mid s_0 = s, \pi\right]$
- Q-function: $Q_\pi^j(s, \boldsymbol{a}) = r^j(s, \boldsymbol{a}) + \gamma \mathbb{E}_{s' \sim p}\left[v_\pi^j\left(s'\right)\right]$.

Mean Field MARL

As all agents act strategically and evaluate simultaneously their value functions based on the joint actions, we factorize the Q-function using only the pairwise local interactions:

$$Q^j(s, \boldsymbol{a}) = \frac{1}{N^j} \sum_{k \in \mathcal{N}(j)} Q^j\left(s, a^j, a^k\right)$$

where $N^j = |\mathcal{N}(j)|$.



Inference

$$Q^j\left(\mathbf{s}, \mathbf{a}\right) \approx Q^j\left(s, a^j, \bar{a}^j\right)$$

where $\bar{a}^j = \frac{1}{N^j} \sum_k a^k$.

Applications in ising model

We define the reward for each spin/agent as (to maximize its reward)

$$r^j = h^j a^j + \frac{J}{2} \sum_{k \in \mathcal{N}(j)} a^j a^k$$

To learn an optimal joint policy π^* for Ising model, we use the stateless Q-learning with mean field approximation (MF-Q), defined as

$$Q^j\left(a^j, \bar{a}^j\right) \leftarrow Q^j\left(a^j, \bar{a}^j\right) + \alpha\left[r^j - Q^j\left(a^j, \bar{a}^j\right)\right]$$

Order parameter

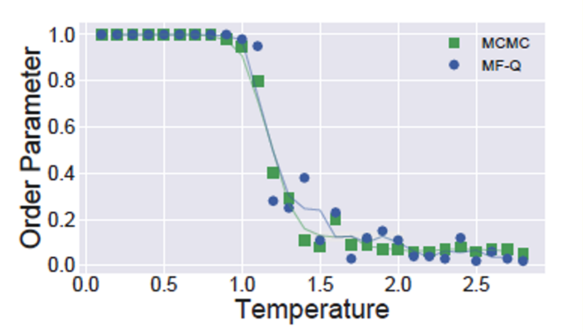
$$\xi = \frac{|N_\uparrow - N_\downarrow|}{N}$$

A traditional measure of purity for the Ising model.

Comparisons with MC

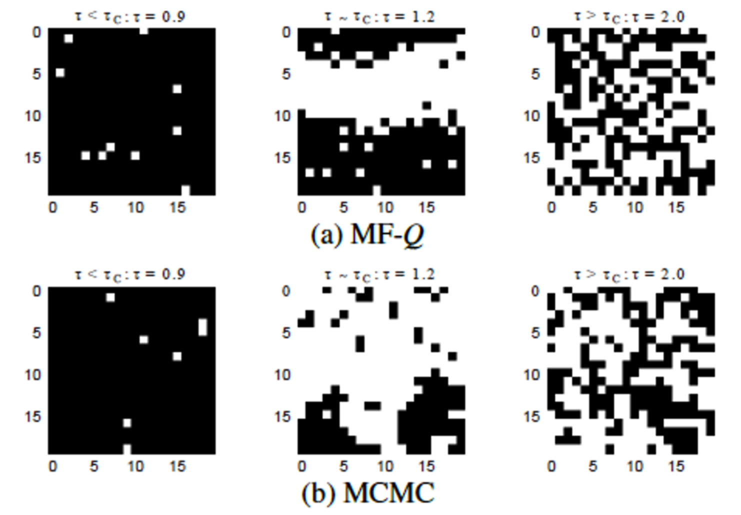
Attention

Notice that agents here do not know exactly the energy function, but rather use the temporal difference learning to approximate $\langle a^j \rangle$ during the learning procedure.



This is the first work that manages to solve the Ising model via model-free reinforcement learning methods.

Results



- As the temperature rises ($\tau = 1.2$, the Curie temperature), some spins become volatile and patches start to form as spontaneous magnetization.

References

[1.]

Yang Y, Luo R, Li M, et al. *Mean field multi-agent reinforcement learning*[C]//International Conference on Machine Learning. PMLR, 2018: 5571-5580.

Thank you for your listening!

Q & A

Homework

Pick one of two:

- How does ising model show the principle of the lowest energy?
- What is the main idea of the Mean Field Theory?