Report on Chaotic Boltzmann Machines

Budurean Marius-Mihai, Nevezi-Strango David

January 2024

1 Boltzmann Machines

A **Boltzmann Machine (BM)** is a type of neural network model composed of stochastic elements. The machine is made from a network of interconnected binary nodes. Each connection between two nodes has associated weights that determine the strength of the connection. The state of each node can be either zero or one, representing the binary activation state.

Boltzmann machines are used in the following problems:

- optimization problems;
- repairing degraded images;
- learning interdependency among random variables.

These machines are relevant in the theoretical field due to their contribution to solving optimization problems and learning problems. Some downsides of Boltzmann machines are that they have a high computation cost and it is difficult to apply them in their original form to real-life problems.

The benefit of using Boltzmann machines to solve optimization problems is that convergence to global optimum is guaranteed. These problems can be solved without prior knowledge, but optimality is assured theoretically for an extremely slow annealing schedule.

In terms of learning problems, parameter values that represent data distribution can be learned by a simple algorithm. The only drawback is that the learning algorithm requires lengthy computation to obtain equilibrium statistics.

Since learning problem performance is crucial, there were many attempts to increase computation speeds. The first category of optimizations affects the learning capabilities by restricting the network structure and by using approximations in the learning algorithm. As a result of these performance increases, two types of Boltzmann machines are created:

- restricted Boltzmann machines,
- deep Boltzmann machines.

1.1 Restricted Boltzmann Machines

A **Restricted Boltzmann Machines (RBM)** is a variation of the original Boltzmann machine which addresses some of the computational challenges associated with training the original models. The term "restricted" refers to the connections between neurons in the network. Unlike a fully connected Boltzmann machine, an RBM is a two-layer neural network with a visible layer and a hidden layer, and there are no connections within the layers. Every neuron in the visible layer is connected to every neuron in the hidden layer, but there are no connections within the same layer. One of the key features of the RBMs is that, compared to the original BM, it is much more energy-based.

1.2 Deep Boltzmann Machines

A **Deep Boltzmann machine (DBM)** is another variation of the original BM, similar to the RBM. The difference between the 2 variations is that DBM has many hidden layers, while RBM has only one hidden layer. DBMs were used for tasks such as image generation, feature learning, and unsupervised representation learning. At this moment, the DBMs were replaced by more tractable deep learning architectures, such as deep neural networks. Nonetheless, DBMs represent an important concept in the history and understanding of deep generative models.

1.3 Hardware implementation

Hardware implementation is one approach to enhance the computation speed of Boltzmann machines without degrading their capability. One major difficulty in this approach is the generation of good random numbers. To solve this problem, different requirements appeared in the circuits, such as pseudorandom number generators or other controllable mechanisms for randomness (quantum effects). Other studies on hardware implementation avoid randomness by using mean-field approximations.

2 Chaotic Boltzmann Machines

The research paper **"Chaotic Boltzmann Machines"** presented a new approach to Boltzmann machines where the classic stochastic system is replaced with a deterministic system. In the first part of the results, there is presented the mathematical comparison of the two approaches. After that, the authors offer two practical examples of the problem addressed. The first one is related to optimization problems, and the second one presents an application of Chaotic Boltzmann machines in the Ising model. Both approaches had similar results, facts demonstrated through the plots displayed in the paper.

2.1 Stochastic System in Boltzmann Machines

$$z_i = b_i + \sum_j w_{ij} s_j, s_i \in \{0, 1\} \tag{1}$$

$$P[s_i = 1] = \frac{1}{1 + exp(\frac{-z_i}{T})} \tag{2}$$

$$P(s) = \frac{1}{Z} exp(\frac{-E(s)}{T}) \tag{3}$$

$$E(s) = -\sum_i b_i s_i - \sum_{i<j} w_{ij} s_i s_j \tag{4}$$

$$Z = \sum_s exp(\frac{-E(s)}{T}) \tag{5}$$

$s_i \rightarrow$ the value of i^{th} state, $i \in \{1, N\}$;
$z_i \rightarrow$ input;
$b_i \rightarrow$ constant bias;
$w_{ij} \rightarrow$ symmetric matrix which describes the interaction between i^{th} and j^{th} units;
$T \rightarrow$ temperature of the system;
$exp(x) \rightarrow e^x$;
$s = (s_1, s_2, \ldots s_N)$
$P(s) \rightarrow$ computed state based on Gibbs distribution;
$E(s) \rightarrow$ global energy;
$Z \rightarrow$ partition function;

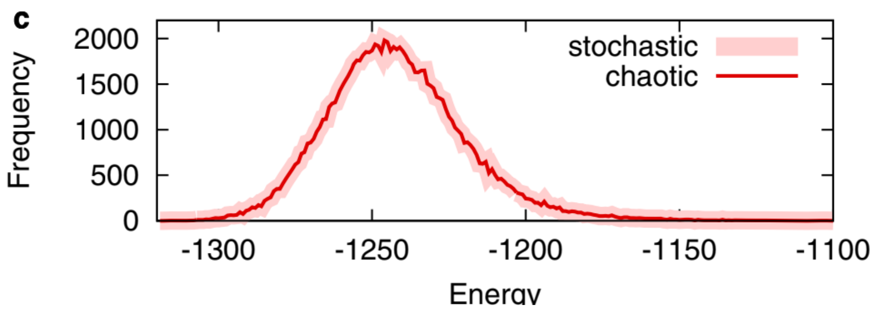
2.2 Deterministic System in Chaotic Boltzmann Machines

$$\frac{dx_i}{dt} = (1 - 2s_i)(1 + exp\frac{(1 - 2s_i)z_i}{T}) \tag{6}$$

$$s_i := 0 \text{ when } x_i = 0 \text{ and } s_i := 1 \text{ when } x_i = 1 \tag{7}$$

$x_i \rightarrow$ internal state of $i^{th} unit$;
The differential equation from equation (6) replaces the stochastic computation from equation (2).
Regardless of the states of other units, the right-hand side of the equation (6) is **positive** for $s_i = 0$ and **negative** for $s_i = 1$, so x_i continues to oscillate between 0 and 1.
The authors of the paper replaced the stochastic approach of computing the state with this deterministic approach based on the differential equation from above. However, since the states of other units in the system change, it is not theoretically assured that the probability of observing a certain state in the proposed system is the same as in the original Boltzmann machine. This

behavior is chaotic, in comparison to the stochastic system, but the numerical evidence provided in the paper shows that these two behaviors are producing similar results.



2.3 Application to combinatorial optimization problems

To prove the efficiency of the Chaotic Boltzmann machines, the first applications were tested on optimization problems, respectively maximum cut problems.

"Given an undirected network of N nodes whose edge weights are represented by a symmetric weight matrix (d_{ij}), the maximum cut problem is to find a subset $S \subset \{1, \ldots, N\}$ that maximises $\sum_{i \in S, j \notin S} d_{ij}$.".

To solve this problem, parameter values of the Boltzmann machines are set as $b_i = \sum_j d_{ij}$ and $w_{ij} = -2d_{ij}$. When the energy is minimized, the solution to the maximum cut problem is given by the set of nodes taking on the state $s_i = 1$.
The authors used the problem sets provided along with the Biq Mac solver, for which exact solutions are also provided, and compared the results with the ones obtained with stochastic Boltzmann machines. The results were similar.

Conclusion of the experiment:

"All of the problems with $N \leq 100$, optimal solutions are obtained. Overall, fairly good solutions (2–3% degraded from optima) are obtained without excessive tuning of the annealing schedule. The point to be noted here is that there is no significant difference between the chaotic and stochastic Boltzmann machines. A detailed comparison of the results has no meaning because time cannot be directly compared."

Table 1 | Solutions of maximum cut problems obtained by chaotic and stochastic Boltzmann machines. For each dataset, the network size and the value of maximum cut are shown. The first and second line for each dataset show the statistics of the results obtained by chaotic and stochastic Boltzmann machines, respectively

	Problem		Obtained results		
dataset	size	opt.	max.	avg.	s.d.
g05_100.0	100	1430	1430	1415.62	13.65
			1430	1414.97	13.69
pw09_100.0	100	13585	13585	13447.51	133.99
			13585	13440.37	134.19
ising3.0-300_5555	300	8493173	8454235	8304704.46	88561.68
			8430298	8215987.23	98257.26
t2g20_5555	400	24838942	24798931	24346513.46	266841.24
			24791603	24138133.96	300049.99
t3g7_5555	343	28302918	28302918	27910192.78	317104.08
			28302918	27550104.43	423472.58

2.4 Application to Ising Model

The second experiment presented in the paper is **an application to the ising model**. This model was applied to a simple two-dimensional lattice. The Ising model is a simple model for ferromagnetism, and it can be regarded as a Boltzmann machine whose connections are limited to neighboring nodes in the lattices. Despite the simplicity of the model, it exposes rich behavior including phase transition with critical behavior.

$$E(\sigma) = -\sum_{\langle i,j\rangle} \sigma_i \sigma_j, \sigma_i \in \{-1, 1\} \tag{8}$$

$$\frac{dx_i}{dt} = -\sigma_i exp(-\frac{1}{T}\sigma_i \sum_j \sigma_j) \tag{9}$$

Equation (8) is the Hamiltonian of the Ising model, and equation (9) is the differential equation for the chaotic Ising model, designed similarly to equation (6).

3 Conclusion

As a conclusion, the paper presents the differences between stochastic BM and chaotic BM. While chaotic Boltzmann machines show promise for efficient hardware implementation due to their parallel nature and lack of reliance on random numbers, they present challenges in numerical simulation speed and practicality on digital computers.
However, they offer potential advantages in parallel processing related to real neural networks and demonstrate similarities to neural network models with hysteresis neurons. The paper suggests further theoretical exploration of chaotic Boltzmann machines' dynamics and their applications in various fields, including machine learning, computing, and neuroscience.

Additionally, the text highlights the potential for chaotic Boltzmann machines to contribute to the development of future computing systems, particularly in microelectronic circuits and reversible computing.