Hopfield Networks

Vipul Venkataraman

Indian Institute of Technology Bombay
1 Motivation

2 What are Hopfield Networks?
   • Formal Definition
   • Characteristics

3 Memory
   • Humans vs Computers
   • Memories in Hopfield Networks
   • Why are Hopfield Networks interesting?
   • Why are Hopfield Networks Good Models for CAM?

4 Hopfield Memory Storage and Retrieval
   • Hopfield Update Process
   • Learning Rule for Hopfield Networks
   • Energy of a Hopfield Network
   • Capacity of a Hopfield Network
   • Intuition
Motivation

- Nature is filled with complex physical phenomena arising out of interactions between a large number of simple elementary components.
- An example is magnetic properties of a substance arising out of interactions between the fundamental particles that form the substance.
- The computational power of the brain arises out of the interactions between neurons.
- The Hopfield Network attempts to explain the phenomenon of memory through a network of perceptrons.
What are Hopfield Networks?

**Definition**

A Hopfield Network is defined by an undirected graph $G = \langle V, f \rangle$ and a function $f : V^2 \rightarrow \mathbb{R}$ where $f(a, b) = f(b, a)$ with the nodes of the Graph representing Binary Threshold Neurons and the function $f$ defines the strength of the connection between nodes.

**Remark**

*In this presentation we will assume that the possible values each neuron can take are -1 and +1 instead of 0 and 1.*
The Graph which defines the Hopfield Network is fully connected. This means that every neuron can be correlated with every other neuron. Weights encode the correlation between two nodes. Positive weights encourage both neurons to fire (or not) together.
Humans vs Computers

- Computers primarily use location addressable memory
  - A memory is accessed by giving a physical location on disk
- Content addressible memory on the other hand utilizes the (partial) content of the memory itself as an address
- For example, the phrases ”Harry” and ”J.K. Rowling” can be used to access memories of the books
Hopfield Networks are intriguing for many reasons

- They serve as models for understanding Content Addressible Memory (CAM)
- They do so despite the fact that they consist of simple binary threshold neurons
- They add credibility to the idea that important physical phenomena (memory in this case) can arise out of simple elementary components (neurons)
Why are Hopfield Networks Good Models for CAM?

- Hopfield Networks consist of a number of neurons which are structurally identical.
- They are asynchronous which makes them match them ideal for biological models.
- They provide error correction capabilities which humans demonstrate.
- For example, ”Barry Potter” and ”J.K.Howling” might still retrieve the same novel.
Hopfield Update Process

- Start from a corrupted memory (Part of the memory)
- Identify a random neuron whose output does not match with the other neurons
- Change its value as appropriate
- Repeat this procedure till convergence
Learning Rule for Hopfield Networks

Definition

The weights of a Hopfield Network are determined by the following equation

\[ W_{ij} = \sum_s (V^s_i)(V^s_j) \]

where

- \( V^s \), \( s = 1..n \) are the vectors needed to be stored
- \( W_{ij} = W_{ji} \) is the weight between neuron \( i \) and neuron \( j \)
- \( V^s_i \) refers to the value of the \( i^{th} \) bit of vector \( V^s \)
The weight between two neurons is the number of times they fire together minus the number of times they don’t.

Hopfield Networks store memories through associations between the neurons in the network.

This means that if two neurons or two bits in the memory repeatedly fire with the same value, the weight between them increases.

Similarly, the weight between two neurons is low if they repeatedly contradict each other (i.e. One is in state +1 while the other is in state -1).
Energy of a Hopfield Network

**Definition**

The energy of a particular bit vector $V$ is defined by

$$E = -\frac{1}{2} \sum_{i,j} W_{ij} V_i V_j$$

- Neurons with positive weights tend to agree with each other
- Similarly neurons with negative weights between them are encouraged to disagree
Field of a Neuron

Definition

The field $F_i$ at a neuron $i$ is defined by the following formula

$$F_i = \sum_{j \neq i} V_j W_{ij}$$

Remark

The field of a neuron is the influence of the other neurons on this neuron

Remark

The energy of a configuration $V$ is related to the fields by

$$Energy = -\frac{1}{2} \sum_i F_i V_i$$
A stable state in a Hopfield Network occurs when each neuron agrees with its field i.e. $F_i V_i > 0$

The terms in the field can be split into two

The first, $\sum_{j \neq i} \left[ V_j^{s'} \sum_{s=s'} V_i^s V_j^s \right] = (n-1) V_i^{s'}$, is the influence of a memory on itself

The second, $\sum_{j \neq i} \left[ V_j^{s'} \sum_{s \neq s'} V_i^s V_j^s \right]$, is the influence of the other memories onto this one

The second term is called Crosstalk and corrupts the network
An assumption here is that each memory is produced independently.

Each bit of a memory is also produced independently through a uniform distribution.

The mean of the Crosstalk term is 0.

Ideally this would be the case for every neuron when the current state is a memory.

Unfortunately, the variance of the Crosstalk term is not negligible.

The variance of the Crosstalk term is \((n - 1)(N - 1)\).
- Memory corruption gives rise to the notion of Memory Capacity
- Having more memories stored in the network leads to corruption
- The probability of an error grows with the number of memories
- This provides some mathematical intuition for the capacity of a network
- Empirically a capacity of 0.15n is shown from experiments
Well. Memories are nice but what else?
Neurons do more than memorise
Interestingly, so do Hopfield Networks
This problem requires some change
Previously, memories were provided and weights were inferred
Now, the challenge is to design weights to infer memories
5 Uses for Combinatorial Optimizations
- The Eight Rooks Problem
- The Eight Queens Problem
- Travelling Salesman Problem

6 Summary
The problem is fairly simple to state

How can you place 8 rooks on a chess board so that none of them can kill the other?

The important step is mapping the weights to provide solutions

We design the network to solve our problem

Here, we assume that a Hopfield Network can also handle energy functions with linear terms and i.e

\[ E = V^T W V + B^T V \]
The Eight Rooks Problem

One possible solution is as follows

- There are 64 neurons, one for each square on the board
- Each neuron has two indices (the row and the column)
- A neuron takes value 1 if it contains a rook and 0 if not
- The only constraints our solution must satisfy is that every row and column should have a single rook
- The energy function is defined as follows

$$E = \sum_{i=1}^{8} \left( \sum_{j=1}^{8} x_{ij} - 1 \right)^2 + \sum_{j=1}^{8} \left( \sum_{i=1}^{8} x_{ij} - 1 \right)^2$$

Vipul Venkataraman
Hopfield Networks
The Eight Queens Problem

- This problem is identical to the previous one except now we have diagonal constraints as well.
- Why not add terms for the diagonal similar to the rows and columns?
- We can’t because the diagonals don’t necessarily need to have a queen.
- We’d need to penalize for the case where there are 0 or 1 queens on a diagonal.
- This is doable but things are getting noticeably more complex.
- So are there limits to Hopfield Networks?
This is one of the most well studied problems in Computer Science. It is proved to be NP-Complete which makes it the ideal benchmark. So how do you use Hopfield Networks for TSP? An exact solution is too much to ask. We have to compromise on correctness for speed.
There are \( n^2 \) neurons (\( v_{ij} \)) indexed by

- \( i \) - The city
- \( j \) - The \( j^{th} \) city in the path

\( v_{ij} \) is 1 if city \( i \) is the \( j^{th} \) city on the path and 0 otherwise

In simple terms, \( v_{ij} \) tells you when each city was visited

Now, one simple objective function is as follows

\[
E = \sum_{i,j,k} v_{ik} v_{j(k+1)} d_{ij}
\]

where

\( d_{ij} \) is the distance between city \( i \) and city \( j \)

This is clearly wrong since \( v_{ij} = 0 \) is a trivial solution
Travelling Salesman Problem

- Well. That’s too bad but the TSP is NP-Complete after all
- We also need to add constraints to ensure each city is visited only once
- These are identical to the ones in the Eight Rooks Problem
- But now, the added complication is that we need to combine these two
- Whatever this combination is, we need to ensure that the resulting energy function is quadratic
- This seems a natural candidate

\[
E = \sum_{i,j,k} v_{ik}v_{i(k+1)d_{ij}} + \gamma \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{n} v_{ij} - 1 \right) \right)^2 + \sum_{j=1}^{n} \left( \sum_{i=1}^{n} v_{ij} - 1 \right)
\]
• Hopfield Networks are a model for Content Addressable Memory
• It models with some mathematical foundation the concepts of
  • Memory Capacity
  • Memory Corruption
  • Error Correction
• They show how complex phenomena can emerge from simple units
1. Neural Networks and Physical Systems with Emergent Collective Computational Abilities, 1982, John Hopfield
6. The Organization of Behavior - A Neuropsychological Theory, 1949, Donald Olding Hebb