Sourlas (1989) "Spin-glass models as error-correcting codes."

Coding theory provides rigorous bounds on the cost-effectiveness of any code. The explicit codes proposed so far for practical applications do not saturate these bounds; that is, they do not achieve optimal cost-efficiency. Here we show that theoretical models of magnetically disordered materials (spin glasses) provide a new class of error-correction codes. Their cost performance can be calculated using the methods of statistical mechanics, and is found to be excellent. These models can, under certain circumstances, constitute the first known codes to saturate Shannon's well-known cost-performance bounds.

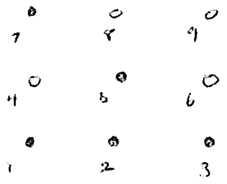
Spin-glass theory has already found applications to several fields outside of its original scope, such as combinatorial optimization, associative memories, categorization theory, and theoretical models of the immune system. Here I show that coding theory may be added to this list. Information theory, which forms the basis of coding theory, is also relevant to many of these other fields, and can in some sense be considered as their unifying element.

Ising Model [Baxter, Exactly Solved Models in Statistical Mechanics]

Consider a magnet as made up of molecules that are constrained to lie on the sites of a regular lattice. Suppose there are N such sites and molecules, $i=1,\dots,N$.

Regard each molecule as a microscopic magnet, each molecule i has two possible configurations labeled by a "spin" variable σ_i with values $+1$ or -1.

Let $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_N\}$ denote the set of all spins.



$\sigma = \{+,+,+,-,+,-,+,-,-\}$.

The Hamiltonian is a function $E(\sigma_1,\dots,\sigma_N)$ of the N spins. It has two parts

$$E(\sigma) = E_0(\sigma) + E_1(\sigma)$$

where E_0 is contribution from forces inside the magnet (interactions among ~~magnets~~ molecules) and E_1 is contribution from interaction between spins and an external magnetic field.

Since σ_i is effectively magnetic moment of molecule i, $E_1(\sigma)$ can be written as $E_1(\sigma) = -H \sum_i \sigma_i$

where H is proportional to component of field in direction of preferred axis

In a physical system, we expect interactions to be invariant under time reversal, so E is unchanged by reversing all fields and magnetizations, i.e. by negating H and $\sigma_1, \dots, \sigma_N$. It follows that E_0 must be an even function of σ, i.e.

$$E_0(\sigma_1, \dots, \sigma_N) = E_0(-\sigma_1, \dots, -\sigma_N).$$

We can consider the magnetization M as a function of the field H and temperature T, so $M(H,T)$.

How do we calculate thermodynamic functions such as $M(H,T)$ from microscopic forces between components of system. Consider Gibbs distribution. Consider system with states s and Hamiltonian $E(s)$. Form the partition function

$$Z = \sum_s \exp[-E(s)/kT]$$

where k is Boltzmann's constant and summation is over all allowed states s of system.

Then free energy F is

$$F = -kT \ln Z.$$

and probability of being in state s is

$$\frac{1}{Z} \exp[-E[s]/kT].$$

If X is some observable property of system like total energy or magnetization, with value $X(s)$ for state s, average thermodynamic value is

$$\langle X \rangle = \frac{1}{Z} \sum_s X(s) \exp[-E(s)/kT].$$

For the Ising model, the partition function is

$$Z_N(H,T) = \sum_{\sigma} \exp\left[-\left(E_0(\sigma) - H\sum_i \sigma_i\right)/kT\right].$$

Physically, we expect the free energy of a large system to be proportional to size of system, ie. we expect thermodynamic limit

$$f(H,T) = -kT \lim_{N\to\infty} \frac{1}{N} \ln Z_N(H,T)$$

to exist, where f is the free energy per site.

The magnetization is average of magnetic moment per site.

$$M(H,T) = \frac{1}{N}\cdot\frac{1}{Z_N}\sum_{\sigma}(\sigma_1+\dots+\sigma_N)\exp\left[-\left(+E_0(\sigma) - H\sum_i\sigma_i\right)/kT\right].$$

in the thermodynamic limit, when $N\to\infty$,

$$M(H,T) = \frac{-\partial}{\partial H} f(H,T).$$

The most common form of even Hamiltonian $E_0(\sigma)$ is nearest-neighbor:

$$E_0(\sigma) = -J\sum_{(i,j)}\sigma_i\sigma_j$$

where sum is over all nearest-neighbor pairs of sites in lattice.

$$Z_N = \sum_{\sigma}\exp\left[K\sum_{(i,j)}\sigma_i\sigma_j + h\sum_i\sigma_i\right]$$

where $K = \frac{J}{kT}$ and $h = H/kT$.

So magnetization per site is $M = \frac{\partial}{\partial h}\lim_{N\to\infty}\frac{1}{N}\ln Z_n(h,K)$.

Hop field Network is, a model of how neurons store and process memories: teach it to memorize patterns, and then recover them from noisy.

Neural network with neurons at vertices of graph and edges connecting neurons to one another. Hopfield networks have undirected edges.

Each neuron i has state s_i and a threshold b_i.

Each edge has weight W_{ij}.

Hopfield network is dynamical model, governed by ~~learning~~ update rule.

→ At step n in dynamics, we choose a random neuron i and update its state, s_i^n which is either +1 or −1.

To decide new state, ignore current state but compute incoming weight $T_i = \sum_{j \neq i} s_j^n W_{ji}$ and $\begin{cases} \text{if } T_i > b_i, & s_i^{n+1} = +1 \\ \text{if } T_i \leq b_i, & s_i^{n+1} = -1 \end{cases}$

Hopfield network is Ising model, with Hamiltonian

$$E = -\frac{1}{2} \sum_{ij} W_{ij} s_i s_j + \sum_i b_i s_i .$$

Notice each update either decreases energy or leaves it unchanged.

After enough updates, eventually reach local minimum.

→ Network converges to some output for a given input

hopefully memories are close to original stimulus, find way to remember particular memory

That was retrieval as a fixed point of dynamics.

What about storage.

To memorize set of m memories, set weights to be

$$W = \frac{1}{m} \sum_{k=1}^{m} v^{(k)} \left(v^{(k)}\right)^T$$

when $v^{(k)}$ is $n \times 1$ column vector representing kth memory.

Can also do Hebbian learning.

How many memories can network remember, and how robust is it at remembering a memory given a partial version of that memory.

See Abu-Mostafa and St. Jacques (1985)