Psychology 452
Week 11: Autoassociative Networks

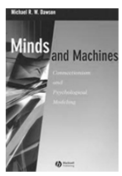
Physical basis of Hopfield networks
Energy minimization in Hopfield networks
Learning and attractors
Boltzmann machines

Course Structure

When	What
Weeks 1, 2, 3	Connectionist Building Blocks
Weeks 4, 5, 6	Case Studies of Connectionism
Week 7	Midterm Exam
Weeks 8, 9, 10	Interpreting Connectionist Networks
Weeks 11, 12	**Deep Learning Basics**
Week 13	Final Exam

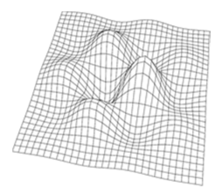
Chapter 7 Discussion

- **Questions?**
- **Important Terms**
  - **Feedback**
  - **Machine**
  - **Homeostat**
  - **Tortoise**
  - **Braitenberg vehicle**
  - **NETTalk**
  - **Cricket phonotaxis**
  - **Stigmergy**
  - **Law of uphill analysis and downhill synthesis**



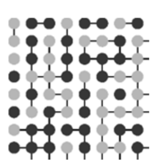
Least Energy

- **When some physical systems are disturbed, its components change their states to minimize the disturbance**
  - **Pressing a soap bubble produces a "least energy" configuration**
- **The system can be thought of as a point in state space**
  - **System properties give coordinates in space**
  - **System energy is height of state space surface at those coordinates**
- **Minimizing energy = moving along space to a low energy "sinkhole"**



The Spin Glass

- **A magnet can be thought of as a lattice; each component has a particular magnetic orientation**
- **In a typical magnet the orientation of poles is uniform**
- **In a spin glass, this orientation is random**
- **Changing the orientation in one component will affect the orientation of other spin glass components**

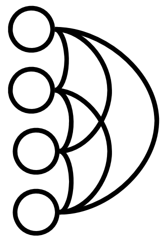


Magnets and Neural Nets

- **Physicist John J. Hopfield realized that the properties of a spin glass were similar to those of a neural network**
- **He used spin glass mathematics to derive a new kind of neural network architecture called a Hopfield network**

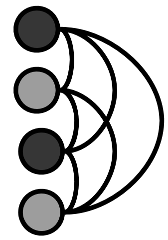
Hopfield Network

- A single set of processing units
- Units are often binary
  - (1, -1 frequently used)
- Units are linked by massively parallel connections
- Network is autoassociative!



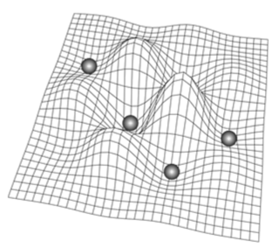
Learning

- Hopfield nets learn by being presented patterns
- The Hebb rule is used to store patterns in memory
- The point of learning is to establish “sinkholes” in an energy space



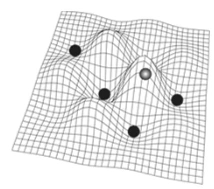
Attractors And Networks

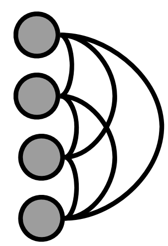
- An attractor is a stable state toward which a dynamic system evolves over time from initial conditions
- Once the attractor is reached, the system stays there until a disturbance occurs
- Memories in Hopfield networks are attractors that capture patterns of processor activities



Start Of Recall

- The network encounters a stimulus “object”
- The stimulus activates the processors
- This activity disturbs the current stable state of the network
- The network is “pushed” to a new, unstable, location in state space





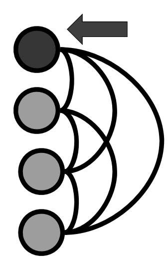
Net Input Function

- Net input for unit *i* is a function of the weighted signals from other units, environmental input (which may be present, but typically isn’t) and the unit’s threshold
- Threshold is typically equal to 0, but does not have to be

$$net_i = \Sigma w_{ij}a_j + i_i - T_i$$

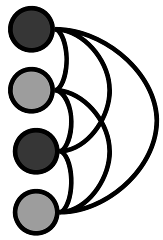
Activation Function

- Randomly choose one processing unit
- If net input > 0, the unit turns on
- If net input < 0, the unit turns off
- If net input = 0, it keeps current state
- The threshold is taken care of in the net input function



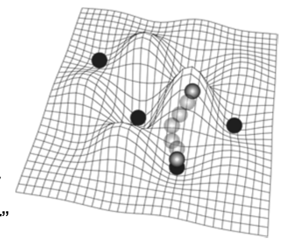
Convergence In Time

- The process is repeated
- Each "flip" of a unit increases stability
- Eventually the system will stabilize -- it will reach a constant state, and will not change unless disturbed



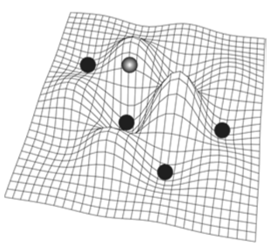
Recall Properties

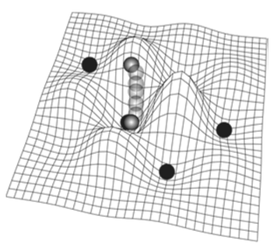
- In general, the network will retrieve the stored "name" that is most similar to the pattern that was input to the network
- The network – as a tiny "marble" – rolls the steepest path downhill to rest in the nearest sinkhole
- The more similar the input is to a stored pattern, the faster will the Hopfield network converge to a stable "answer"



Different Inputs, Different Results

- If the trained network is disturbed in a different way, then it can recall a different pattern
- It rolls into a different local minimum in the state space





Energy In The Network

- Using the spin glass analogy, Hopfield defined an energy (E) term for his network
- Let W be a matrix of weights, a a vector of activity, x an input vector, and t a vector of thresholds

$$E = - \tfrac{1}{2}aWa^T - xa^T - ta^T$$

- This term gets smaller as the network approaches the attractor!

The Effect Of Activity

- The activation of units is crucial for defining network energy
- How might the change in a unit's activity affect total network energy?
- Hopfield proved the following:

$$\Delta E = - (\Sigma w_{ij}a_j + i_i - T_i) (\Delta a_i) = -(net_i)(\Delta a_i)$$

Implications For Energy

- Consider the equation $\Delta E = -net_i \, \Delta a_i$
- When activity changes, energy decreases!

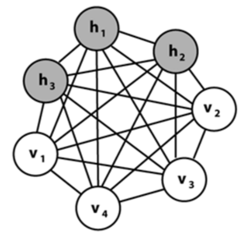
	$a_i = 1$	$a_i = -1$
$net_i > 0$ so $- net_i < 0$	Stays on $\Delta a_i = 0$ $\Delta E = 0$	Turns on $\Delta a_i = +2$ $\Delta E = -ve$
$net_i < 0$ so $- net_i > 0$	Turns off $\Delta a_i = -2$ $\Delta E = -ve$	Stays off $\Delta a_i = 0$ $\Delta E = 0$

Related Networks

- Other autoassociative networks have been used to solve problems in memory and vision
- Brainstate-in-a-box (Anderson, Silverstein, Ritz & Jones, 1977)
- Brainstate-in-a-sphere (Dawson, 1991)
- Various unsupervised networks
- Let’s explore a network that evolved into deep learning nets: the Boltzmann machine

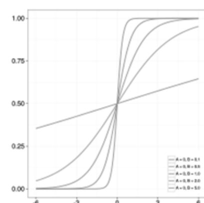
The Boltzmann Machine

- A Boltzmann machine is like a Hopfield network with hidden units
- The environment can only affect visible units (which are in essence input units)
- Hidden units are involved in processing, but cannot be directly changed by the environment
- Units adopt binary activity based on a probability that is computed from net input



Stochastic Units

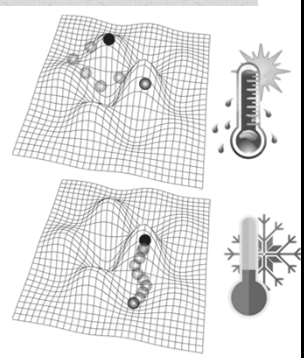
- In a Boltzmann machine units can turn on or off
- They do so stochastically
- They turn on or off as a function of the probability of their net input
- This probability is the logistic of the net input, with a temperature parameter added to the equation
- The temperature parameter affects the slope of the logistic



$$a_i = \frac{1}{1+e^{\left(-(net_i+\theta)/T\right)}}$$

Simulated Annealing

- Boltzmann machines are trained by a technique called simulated annealing
- Early in learning the temperature parameter of the logistic is set to a high value
- Network can move uphill on the energy surface
- As learning proceeds, the parameter is cooled
- <u>As a result global energy minima can be achieved</u>



Boltzmann Uses

- Using hidden units as model of environmental input:
- Fill in missing data with the right probability
- Generate sequences of data (modeled environment) with the right probability
- Solve optimization problems where units represent possible choices