Lecture 23

Motivation

- Idea: with low probability, go against the local field
  - move up the energy surface
  - make the "wrong" microdecision
- Potential value for optimization: escape from local optima
- Potential value for associative memory: escape from spurious states
  - because they have higher energy than imprinted states

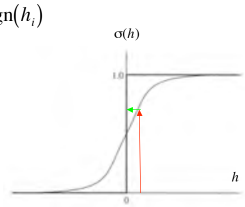
The Stochastic Neuron

Deterministic neuron: $s_i' = \mathrm{sgn}(h_i)$

$$\Pr\{s_i' = +1\} = \Theta(h_i)$$
$$\Pr\{s_i' = -1\} = 1 - \Theta(h_i)$$

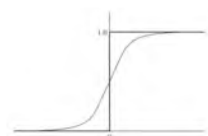
Stochastic neuron:

$$\Pr\{s_i' = +1\} = \sigma(h_i)$$
$$\Pr\{s_i' = -1\} = 1 - \sigma(h_i)$$



Logistic sigmoid: $\sigma(h) = \dfrac{1}{1 + \exp(-2h/T)}$

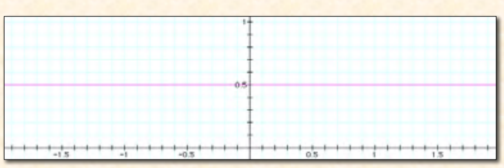
Properties of Logistic Sigmoid



$$\sigma(h) = \frac{1}{1 + e^{-2h/T}}$$

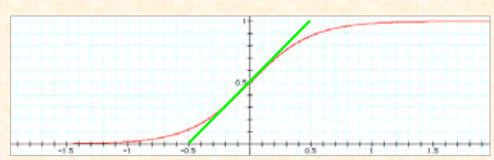
- As $h \to +\infty$, $\sigma(h) \to 1$
- As $h \to -\infty$, $\sigma(h) \to 0$
- $\sigma(0) = 1/2$

Logistic Sigmoid With Varying *T*



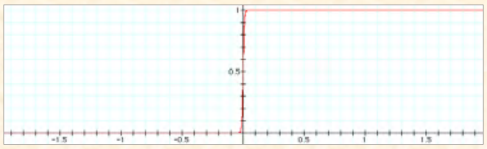
T varying from 0.05 to ∞ ($1/T = \beta = 0, 1, 2, \ldots, 20$)

Logistic Sigmoid $T = 0.5$



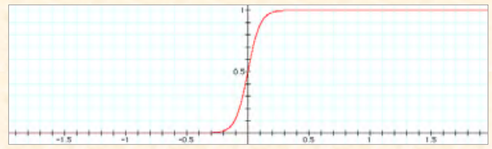
Slope at origin = $1 / 2T$

Logistic Sigmoid
$T = 0.01$



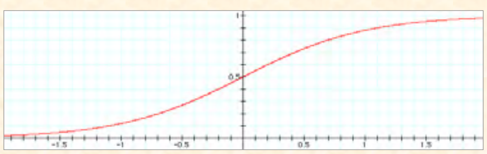
11/12/07 7

Logistic Sigmoid
$T = 0.1$



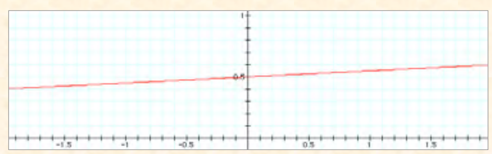
11/12/07 8

Logistic Sigmoid
$T = 1$



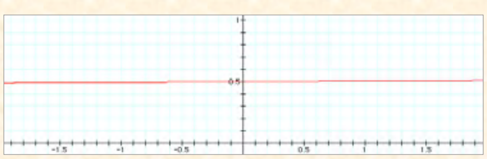
11/12/07 9

Logistic Sigmoid
$T = 10$



11/12/07 10

Logistic Sigmoid
$T = 100$



11/12/07 11

Pseudo-Temperature

- Temperature = measure of thermal energy (heat)
- Thermal energy = vibrational energy of molecules
- A source of random motion
- Pseudo-temperature = a measure of nondirected (random) change
- Logistic sigmoid gives same equilibrium probabilities as Boltzmann-Gibbs distribution

11/12/07 12

Transition Probability

Recall, change in energy $\Delta E = -\Delta s_k h_k$

$$= 2 s_k h_k$$

$$\Pr\{s'_k = \pm 1 \mid s_k = \mp 1\} = \sigma(\pm h_k) = \sigma(-s_k h_k)$$

$$\Pr\{s_k \rightarrow -s_k\} = \frac{1}{1 + \exp(2 s_k h_k / T)}$$

$$= \frac{1}{1 + \exp(\Delta E / T)}$$

11/12/07 13

Stability

- Are stochastic Hopfield nets stable?
- Thermal noise prevents absolute stability
- But with symmetric weights:

  average values $\langle s_i \rangle$ become time-invariant

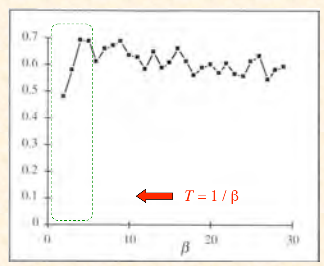
11/12/07 14

Does "Thermal Noise" Improve Memory Performance?

- Experiments by Bar-Yam (pp. 316-20):
  - $n = 100$
  - $p = 8$
- Random initial state
- To allow convergence, after 20 cycles set $T = 0$
- How often does it converge to an imprinted pattern?

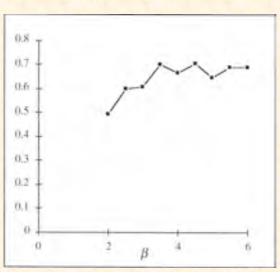
11/12/07 15

Probability of Random State Converging on Imprinted State (n=100, p=8)



11/12/07 (fig. from Bar-Yam) 16

Probability of Random State Converging on Imprinted State (n=100, p=8)



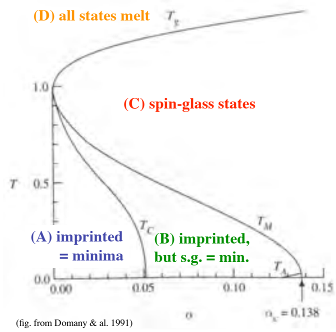
11/12/07 (fig. from Bar-Yam) 17

Analysis of Stochastic Hopfield Network

- Complete analysis by Daniel J. Amit & colleagues in mid-80s
- See D. J. Amit, *Modeling Brain Function: The World of Attractor Neural Networks*, Cambridge Univ. Press, 1989.
- The analysis is beyond the scope of this course

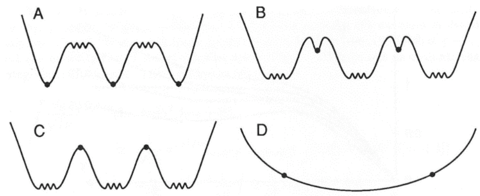
11/12/07 18

Phase Diagram



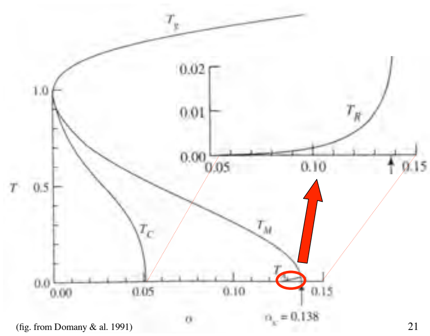
11/12/07 (fig. from Domany & al. 1991) 19

Conceptual Diagrams of Energy Landscape



11/12/07 (fig. from Hertz & al. *Intr. Theory Neur. Comp.*) 20

Phase Diagram Detail



11/12/07 (fig. from Domany & al. 1991) 21

Simulated Annealing

(Kirkpatrick, Gelatt & Vecchi, 1983)

11/12/07 22

Dilemma

- In the early stages of search, we want a high temperature, so that we will explore the space and find the basins of the global minimum
- In the later stages we want a low temperature, so that we will relax into the global minimum and not wander away from it
- Solution: decrease the temperature gradually during search

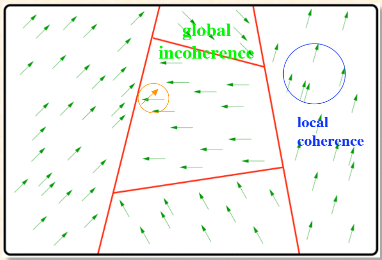
11/12/07 23

Quenching vs. Annealing

- Quenching:
  - rapid cooling of a hot material
  - may result in defects & brittleness
  - local order but global disorder
  - locally low-energy, globally frustrated
- Annealing:
  - slow cooling (or alternate heating & cooling)
  - reaches equilibrium at each temperature
  - allows global order to emerge
  - achieves global low-energy state

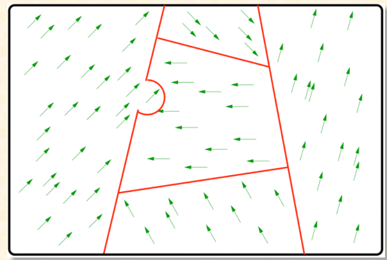
11/12/07 24

Multiple Domains



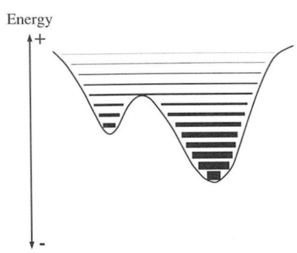
11/12/07 25

Moving Domain Boundaries



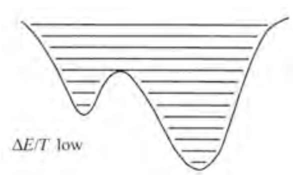
11/12/07 26

Effect of Moderate Temperature



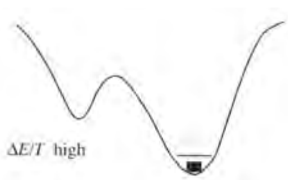
11/12/07 (fig. from Anderson *Intr. Neur. Comp.*) 27

Effect of High Temperature



11/12/07 (fig. from Anderson *Intr. Neur. Comp.*) 28

Effect of Low Temperature



11/12/07 (fig. from Anderson *Intr. Neur. Comp.*) 29

Annealing Schedule

- Controlled decrease of temperature
- Should be sufficiently slow to allow equilibrium to be reached at each temperature
- With sufficiently slow annealing, the global minimum will be found with probability 1
- Design of schedules is a topic of research

11/12/07 30

Typical Practical Annealing Schedule

- Initial temperature T_0 sufficiently high so all transitions allowed
- Exponential cooling: $T_{k+1} = \alpha T_k$
  - typical $0.8 < \alpha < 0.99$
  - at least 10 accepted transitions at each temp.
- Final temperature: three successive temperatures without required number of accepted transitions

11/12/07 31

Summary

- Non-directed change (random motion) permits escape from local optima and spurious states
- Pseudo-temperature can be controlled to adjust relative degree of exploration and exploitation

11/12/07 32

Additional Bibliography

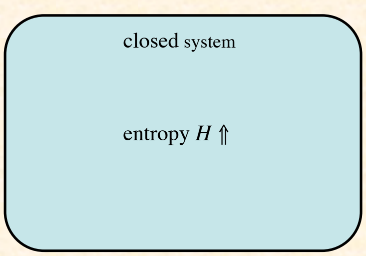
1. Anderson, J.A. *An Introduction to Neural Networks*, MIT, 1995.
2. Arbib, M. (ed.) *Handbook of Brain Theory & Neural Networks*, MIT, 1995.
3. Hertz, J., Krogh, A., & Palmer, R. G. *Introduction to the Theory of Neural Computation*, Addison-Wesley, 1991.

11/12/07 33

V. Genetics & Evolution

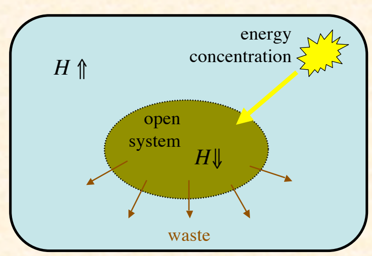
11/12/07 34

The Second Law of Thermodynamics



11/12/07 35

The Second Law and Open Systems



11/12/07 36

Nonequilibrium Thermodynamics

- Classical thermodynamics limited to systems in equilibrium
- Extended by thermodynamics of *transport processes*
  - i.e. accounting for entropy changes when matter/energy transported into or out of an *open system*
- Flow of matter/energy can maintain a *dissipative system* far from equilibrium for long periods
- Hence, *nonequilibrium thermodynamics*

11/12/07 37

An Energy Flow Can Create Structure

11/12/07 (photo from Camazine & al. *Self-Org. Bio. Sys.*) 38

Bénard Convection Cells

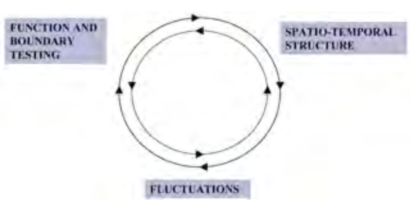
11/12/07 (photo from Camazine & al. *Self-Org. Bio. Sys.*) 39

Persistent Nonequilibrium Systems

- *If* flow creates system so structured to maintain flow
- *then* positive feedback causes nonequilibrium system to persist indefinitely
  - but not forever (2^{nd} law)
- Systems we tend to see are those most successful at maintaining nonequil. state
- Applies to species as well as organisms

11/12/07 40

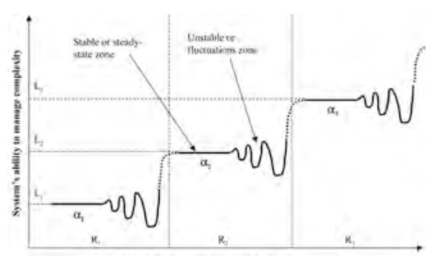
Order Through Fluctuations



- Fluctuations (esp. when system forced out of ordinary operating range) test boundaries & nonlinear effects
- May lead to stabilization of new structures

11/12/07 fig. < Hart & Gregor, *Inf. Sys. Found.* 41

Stabilization at Successively Higher Levels of Organization



11/12/07 fig. < Hart & Gregor, *Inf. Sys. Found.* 42

“Order for Free”

- Relatively simple sets of rules or equations can generate rich structures & behaviors
- Small changes can lead to qualitatively different structures & behaviors
- A diverse resource for selection
- A basis for later fine tuning (microevolution)
- See Kaufmann (*At Home in the Universe*, etc.) and Wolfram (*A New Kind of Science*)

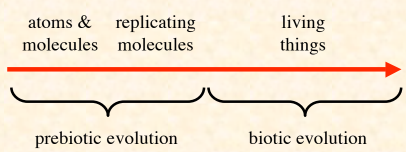
11/12/07 43

Evolution in Broad Sense

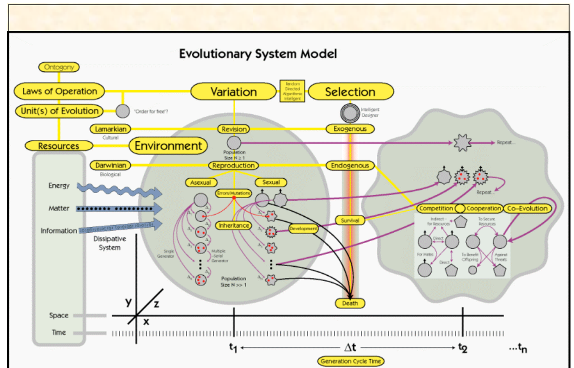
- Evolution in the broadest terms:
  - blind variation
  - selective retention
- Has been applied to nonbiological evolution
  - evolutionary epistemology
  - creativity
  - memes

11/12/07 44

Evolution



11/12/07 45



11/12/07 (from NECSI) 46