ADMINISTRIVIA

• Will start a mailing list today
  – You will get a notice

• A blog page to discuss papers

• Labs are up
  – Ignore the dates and text on pdfs
  – Were created by Anders Oland last year

• Papers for weeks of 22 Sep and 29 Sep will be sent out
  – Please volunteer
  – 13 registered students who are yet to volunteer
The story so far

• Bain’s model:
  – The brain comprises nerve groupings
  – Learning = learning the interconnections between nerves
  – Co-occurrence ➔ co-excitation ➔ Increased connectivity between neurons
The story so far

• McCulloch and Pitts model
  – Neural connections can give you logic circuits
  – Can include excitation and inhibition
  – Threshold logic
  – Simple threshold logic can give you simple boolean relations
  – Complex threshold logic requires hierarchical threshold / McCulloch pitt neurons

• Did NOT specify learning rule
Story so far: Rosenblatt’s Perceptron

- Defining the modern Neural Network structure
  - Defining the perceptron: Rosenblatt 1958
  - Introduces the computational paradigm that the memory *is* the program
  - Biologically motivated, computational model
    - Based on McCulloch and Pitts
    - Also motivated by Hebb
  - Clearly delineates inputs, hidden units, output units
    - With different nomenclature..
  - Learns about *generalization* of learned patterns: one of the earliest use of NNet as a prediction/classification scheme
Today: Learning Rules

• Hebbian Learning..
  – The Hebbian learning rule 1949
  – The cell-assembly theory
  – Develops the theory of associationism
    • If X and Y fire together $\rightarrow$ Connection(X,Y) upgraded
    • Not presented as equations though
  – Unstable
  – Also one of the earliest uses of “connectionism”
  – *Srivaths Ranganathan*
Today: Hebbian Rule

• Hebbian rule
  – If X and Y fire together \( \rightarrow \) Connection(X,Y) upgraded
  – \( W_{k+1} = W_k + \alpha xy \)
  – \( W_k' = W_k + \alpha xy; \quad W_{k+1} = \lambda W_k' + (1-\lambda)W_k \)
  – Binary 0/1 values
  – Unstable
• The perceptron learning rule?

\[ W_{k+1} = W_k + \gamma e_k X_k \]

\[ e_k = d_k - y_k \]

– Takes value -1, 0 or 1 since \( d \) and \( y \) are binary

• Guaranteed to converge if inputs are linearly separable

– And to oscillate otherwise
More on learning rules: The LMS rule

- \( \alpha \)-LMS: \( W_{k+1} = W_k + \alpha e_k \frac{X_k}{\|X_k\|^2} \) “error correction”

- \( \mu \)-LMS: \( W_{k+1} = W_k + \mu e_k X_k \) “steepest descent”

Unlike perceptron rule, \( y \) is a continuous function of \( X \)
- May not give correct solution even if input are separable
Adalines and Madalines (Widrow)

A Madaline is a cascade of “Adaline”s
How do we train the Adalines in the Madalines?
  – Widrow and Lehr

Precursor to backprop
General principle of minimum interference
Zhiyong Chen will present
On Wednesday the 10th

• Continuing:
  – Oja: Shows Hebb’s rule relates to PCA
  – The “Generalized Hebbian” rule
  – Haohan Wang will present

– Addendum: Sanger’s rule
  • Multiple principal components
On Wednesday the 10th

• Back propagation
  – How to train multi-layer networks by “propagating” error
  – Tina Liu will present

– Addendum: RPROP
  • Faster version of back prop
Looking ahead

• Networks so far: feed forward
  – No backward or sideways flow of information
• Recurrence is required to model time series
  – Or temporal evolution
• Next class: SOMs and Hopfield networks
  – One for of feedback
  – Not strictly recurrence
Next Week: 15 September

• Other network types
  – Content addressible memories: Hopfield Nets
    • Anurag will present
  – Self-organizing maps
    • The self organizing map, Kohonen, Proc IEEE, 1990
    • Joseph Chang will present
Next week: 17th Sep

• Boltzmann machines
  – A generalization of Hopfield nets
  – A learning algorithm for Boltzmann machines
    • Ackley, Hinton, Sejnowski 1985
  – Improved simulated annealing, Boltzmann machine, and attributed graph matching
    • Lei Xu, Erkki Oja, 1990
  – Suyoung Kim will present

• Restricted Boltzmann machines
  – Or the “harmoninum” model
    • Paul Smolensky, 1986
  – Making the Boltzmann machine tractable
  – Dishan Gupta will present
Looking Further

• The restrictions of learning

• The power of neural networks

• The power of deep networks

• Alternate network types