The Boltzmann Machine

- **Similarities to Hopfield Networks**
  1. **state values +1, -1**
  2. **weights symmetric**
  3. **unit selected at random**
  4. **no self-feedback**
- **Differences with Hopfield Networks**
  1. **Boltzmann permits hidden neurons**
  2. **Boltzmann uses stochastic neurons**
  3. **Hopfield unsupervised while Boltzmann may operate supervised**

- **Boltzmann Machine terminology**
  - **hidden, visible**
  - **clamped, operating freely**
  - **training: input output**
  - **thermal equilibrium**

- **Basic Operation**
      - **select neuron at random**
      - **update stochastically**
      - **prob(s_j -> -s_j) = 1/(1+exp(-delat_E_j/T)**
- **Clamped and free probabilities: P^+_{alpha} and P^-_{alpha}**
- **Energy for s_j -> -s_j)**
- **prob(s_j -> -s_j) = 1/(1+exp(2$s_j v_j$/T)**
- **Operates like a stochastic neuron**

1. **No self-feedback**
2. **external threshold theata$_j$ (s_0 = -1)**
3. **2^N states**

- **thermal equilibrium**
- **Boltzmann Distributi**
    - **P_{alpha} = 1/Z exp($-E_{alpha}$/T)**
    - **Z = ...**
    - **for large T all states are equiprobable**
    - **as T->0 only states with minimum energy level have non-zero probability**
    - **coarse search -> fine search**
    - **constraint- satisfaction: weak constraints**
- **The Boltzmann Learning Rule**
- **hidden neurons act as feature detectors**
- **state = visible/hidden (alpha / beta)**
- **2^K choices for alpha**
- **2^L choices for beta, where L = N-K**
- **N.B. text says alpha runs from 1 to 2^K**
- **Clamped Probability: P^+_{alpha} = ...**
- **Running free probability: P^-_{alpha} = ...**
- **actual /desired probabilities**
- **relative entropy**
- **gradient descent method**
- **Correlations**