Neural Networks

9. Hopfield Networks

Farkaš, Kuzma et al.

Center for Cognitive Science
Department of Applied Informatics
Faculty of Mathematics, Physics and Informatics UKBA

April 10th, 2018

Hopfield Network

Hopfield Network

- **network:** n neurons with ± 1-threshold activation
- **state:** one value for each neuron:

$$\mathbf{s} \in \{\pm 1\}^n$$

- **weights:** connections for each pair of neurons:

$$\mathbf{W} \in \mathbb{R}^{n \times n}$$

  - diagonal is empty (conventionally zero)

Hopfield Network: Dynamics

- energy of a state:

$$E_{\mathbf{W}}(\mathbf{s}) = -\frac{1}{2}\sum_{j}\left(\sum_{i\neq j} w_{i,j}s_i s_j\right)$$

- computation ~ relaxation to states with lower energy

Hopfield Network: Dynamics

- energy of a state:

$$E_{\mathbf{W}}(\mathbf{s}) = -\frac{1}{2}\sum_{j}\left(\sum_{i \neq j} w_{i,j} s_i s_j\right)$$

- computation ~ relaxation to states with lower energy

- only the current state influences the next state
- synchronous (parallel) dynamics: $\mathbf{s} \mapsto \mathbf{s}'$
  - all neurons change state at once
- asynchronous (sequential) dynamics: $\mathbf{s}_i \mapsto \mathbf{s}'_i$
  - one neuron "recomputes" at a time

Hopfield Network: Transitions

- conventional net input for the i-th neuron:

  $$net_i = \mathbf{w}_i \cdot \mathbf{s}$$

- deterministic transition:

  $$s_i' := \mathrm{sgn}^*(net_i)$$

- "$\pm$sign" function:

  $$\mathrm{sgn}^*(x) = \begin{cases} +1 & x \geq 0 \\ -1 & x < 0 \end{cases}$$

Hopfield Network: Transitions

- conventional net input for the i-th neuron:
  $$net_i = \mathbf{w}_i \cdot \mathbf{s}$$
- deterministic transition:
  $$s_i' := \operatorname{sgn}^*(net_i)$$
- "$\pm$sign" function:
  $$\operatorname{sgn}^*(x) = \begin{cases} +1 & x \geq 0 \\ -1 & x < 0 \end{cases}$$
- stochastic (probabilistic) transition:
  $$P[s_i' = 1] = \frac{1}{1+e^{-net_i/T}}$$
  - depends on temperature T
    - $\beta = 1/T$ is more convenient

Hopfield Auto-associative Memory

- **patterns**: P points in n-dimensional space:

$$\mathbf{x}^{(p)} \in \mathbb{R}^n; \quad \forall 1 \leq p \leq P$$

- we want the stored patterns to be energy minima

Hopfield Auto-associative Memory

- **patterns**: P points in n-dimensional space:

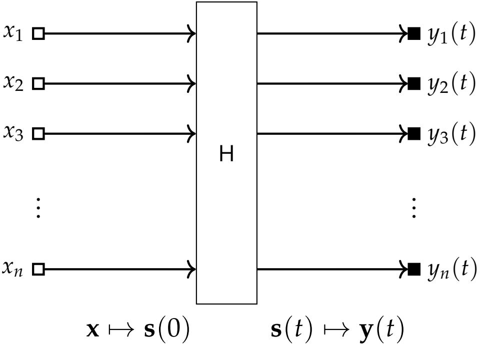
$$\mathbf{x}^{(p)} \in \mathbb{R}^n; \quad \forall\, 1 \leq p \leq P$$

- we want the stored patterns to be energy minima
- analytic training possible (~correlations):

$$w_{i,j} = \begin{cases} \frac{1}{P} \sum_p x_i^{(p)} x_j^{(p)} & i \neq j \\ 0 & i = j \end{cases}$$

Synchronous Network: Black Box

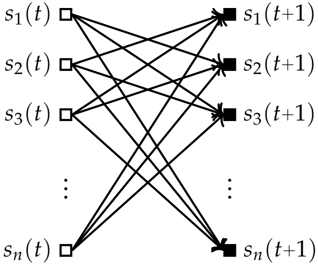
- one input produces a time-series of outputs:



- neuron activations are initialized to the inputs
- outputs at a time t are the neuron activations at t

Synchronous Network: Time Slice

- what happen inside – time step $t \mapsto t{+}1$



- deterministic Hopfield network – linear ± 1-threshold:

$$s_i(t{+}1) = \operatorname{sgn}^* \left(\sum_{i \neq j} w_{i,j} s_j(t) \right)$$

Addenda

- possible outcomes (deterministic synchronous):
  - fixed point (true or false attractor): $\mathbf{s}(t) = \mathbf{s}(t\text{–}1)$
  - limit cycle (even length): $\mathbf{s}(t) = \mathbf{s}(t\text{–}2k);\ \exists k \in \mathbb{N}$

Addenda

- possible outcomes (deterministic synchronous):
  - fixed point (true or false attractor): $\mathbf{s}(t) = \mathbf{s}(t\text{–}1)$
  - limit cycle (even length): $\mathbf{s}(t) = \mathbf{s}(t\text{–}2k);\ \exists k \in \mathbb{N}$
- geometric schedule: temperature value for epoch t is:

$$\beta_t = \beta_s \cdot \left(\frac{\beta_f}{\beta_s}\right)^{\frac{t-1}{t_{max}-1}}$$