Self-Managing Associative Memory for Dynamic Acquisition of Expertise in High-Level Domains

Jacob Beal
IJCAI 2009
SOMs as Associative Memory

- Fast, parallel content retrieval
- Generalize to representative models
  - Higher resolution for more frequent categories
- Unsupervised organization of inputs
  - Episodic memory, analogical retrieval, sensory maps
- Dynamics poorly understood
Self-Organizing Map

Set of models $M_i$, arranged in Euclidean space

(normally $k$-vectors in a grid)

Generalization of [Kohonen, '82]
Self-Organizing Map

Example $\xi_t$: find highest match quality $Q(M_p, \xi_t)$

(association = match w. incomplete example)

Generalization of [Kohonen, '82]
Self-Organizing Map

Blend into nearby models with $B(M_i, \xi_i, w(d(b,i)))$

(initial organization may use time-varying weight)

Generalization of [Kohonen, '82]
High-Level Domain Distributions

- Structured high-level models are very sparse
- Assumption: hierarchical clustering → “spikes”
  - $m_i$ = probability of draw from $i$th spike
Growth of a cluster

- Blend: linear, $w(d) = \max(0, \alpha(1-d/r))$
- Free parameters: $\alpha, r$
Analysis of Cluster growth

- Assume homogeneity
  - All growth on boundary
- Unconstrained, size $n$
  - boundary area = $O(\sqrt{n})$
  - $dn/dt = k/\sqrt{n} \rightarrow O(t^{2/3})$
  - Linear in $\alpha$, $r$, $m_0$
- Eventual equilibrium
  - Size based only on $m_i$
  - Converge exponentially
Experiment: Initial Growth

- 100x100 SOM, 40 trials, 5x10^4 examples
Experiment: Initial Growth

- 100x100 SOM, 40 trials, 5x10^4 examples
Experiment: Convergence

- 100x100 SOM, 40 trials, $10^7$ examples
Experiment: Convergence

- 100x100 SOM, 40 trials, $10^7$ examples

Effect of SOM size on saturated area

Effect of mass on saturated area

Mass is quadratic?
Change of distribution

- Experiment with three cases:
  - Join: green added to red
  - Shift: green instead of red
  - Decay: no spike
Dynamics: Time Response

- 40x40 SOM, 40 trials, $10^6$ old then $10^6$ new
Dynamics: Parameter Variation

- Large SOMs grow similarly; $m_2$ speeds growth
Dynamics: Parameter Variation

- Large SOMs decay slower; $m_2$ slows decay
Erosion of Prior Knowledge

- New knowledge erodes the old unevenly
  - More similar knowledge is more likely to be lost
Contributions

- Analytic and experimental measure of SOM dynamics for high-level associative memory:
  - Growth of expertise set by boundary interactions
  - Initial $O(t^{2/3})$ growth fast enough w. high sample rate
  - Growth/decay ratio can support long-term retention of expertise
- Learning erodes prior knowledge unevenly