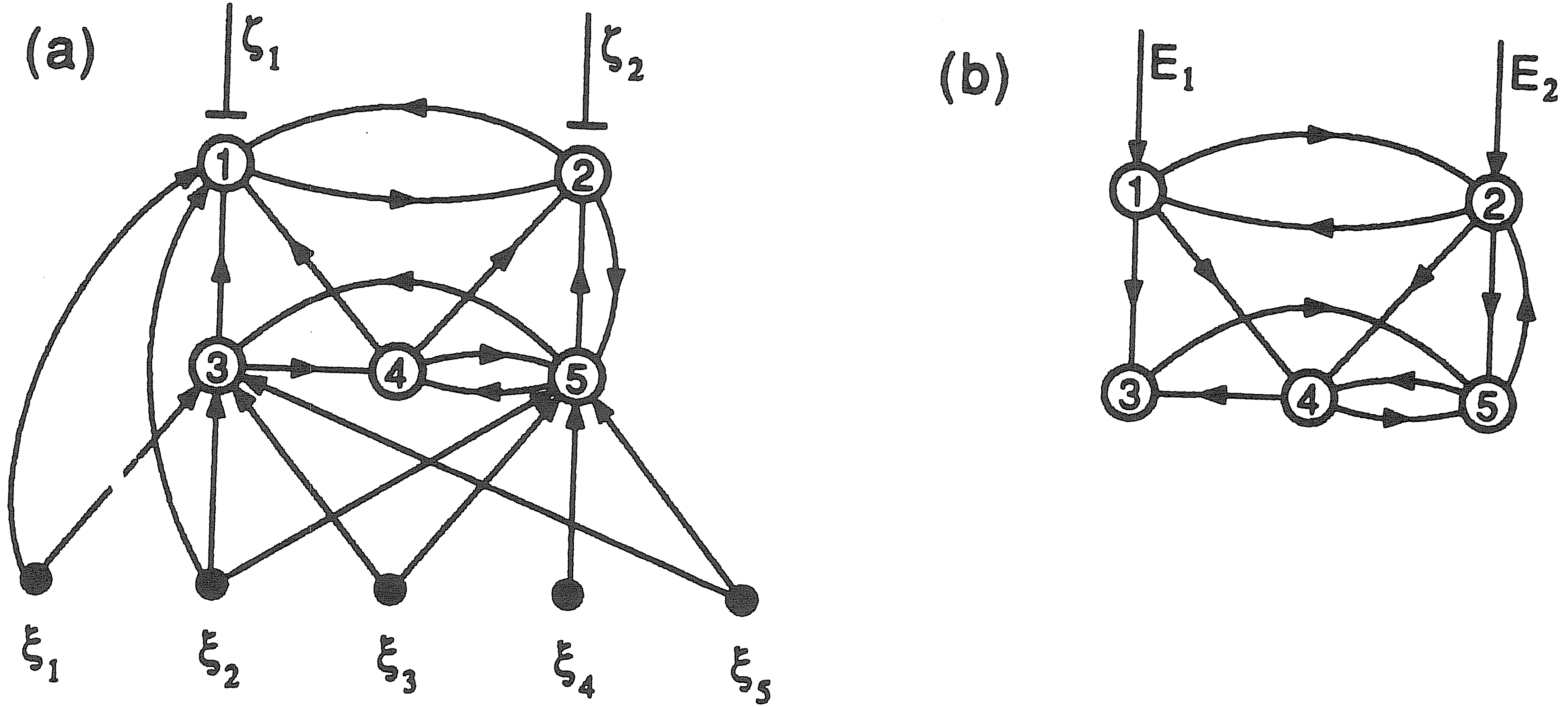
Artificial Neural Networks

- Static input/output
  - *Multi-layer feed-forward nets* (pattern classification and functional approximation)
  - *Hopfield nets* (associative memories, optimization problem solving)
  - *Kohonen's self-organizing maps* (vector quantization and clustering)
- Dynamic input/output
  - *Time-delay neural networks* (prediction and temporal association)
  - *Recurrent neural networks (RNNs)* (sequence classification, prediction and temporal association)


(a)
ζ1
ζ2
1
2
3
4
5
ξ1
ξ2
ξ3
ξ4
ξ5
(b)
E1
E2
1
2
3
4
5

Recurrent Backpropagation - I

- The backpropagation algorithm can be extended to recurrent nets that converge to a stable attractor (fixed point) to learn the association between static input and output patterns.
- For each pattern (***x***, ***t***) in the training set, the recurrent BP algorithm does the following:
  - 1. Relax the original network until convergence to stable activation values:

$$\tau \frac{\partial y_i}{\partial t} = -y_i + g\left(\sum_j w_{ij} z_j\right) \qquad y_i(t+1) = g\left(\sum_j w_{ij} z_j(t)\right)$$

$$\text{where} \quad z_j(t) = \begin{cases} x_j & \text{if } 1 \le j \le M \text{ (inputs)} \\ y_{j-M}(t) & \text{if } M+1 \le j \le M+N \end{cases}$$

  - 2. Determine the errors of the output neurons using the targets (assuming that a sum-of-squares error function is used):

$$e_k = \begin{cases} t_k - y_k & \text{if } k \text{ is an output unit} \\ 0 & \text{otherwise} \end{cases} \qquad E = \frac{1}{2}\sum_k e_k^2$$

Recurrent Backpropagation - II

– 3. Relax the error propagation dual network until convergence to stable activation values:

$$\tau \frac{\partial u_i}{\partial t} = -u_i + \sum_p g'(\sigma_p) w_{pi} u_p + e_i \qquad \text{where} \quad u_p = -\frac{\partial E}{\partial y_p}$$

$$u_i(t+1) = \sum_p g'(\sigma_p) w_{pi} u_p(t) + e_i \qquad \text{and} \quad \sigma_p = \sum_j w_{pj} z_j$$

– 4. Update the network weights by a gradient descent rule:

$$\Delta w_{pq} = -\eta \frac{\partial E}{\partial w_{pq}} = -\eta \frac{\partial E}{\partial \sigma_p} \frac{\partial \sigma_p}{\partial w_{pq}} = \eta \, \delta_p z_q$$

$$\text{where} \quad \delta_p = -\frac{\partial E}{\partial \sigma_p} = -\frac{\partial E}{\partial y_p} \frac{\partial y_p}{\partial \sigma_p} = u_p \, g'(\sigma_p)$$