Chapter 2. The perception-action cycle
in Cognitive Dynamic Systems, Haykin, S.

Course: Autonomous Machine Learning

2016. 11. 04

Jiyoun Moon

Robotics & Intelligent Systems Laboratory
in Department of Electrical and Computer Engineering,
Seoul National University

Contents

Chapter 2. The perception-action cycle

2.7 Neural networks for parallel distributed cognitive information processing

2.8 Associative learning process for memory construction

2.9 Back-propagation algorithm

2.10 Recurrent multilayer perceptrons

2.11 Self-organized learning

2.12 Summary and discussion

2.7 Neural networks

How can we construct A MODEL OF A CONITIVE DYNAMIC SYSTEM based on the PERCEPTION-ACTION CYCLE?

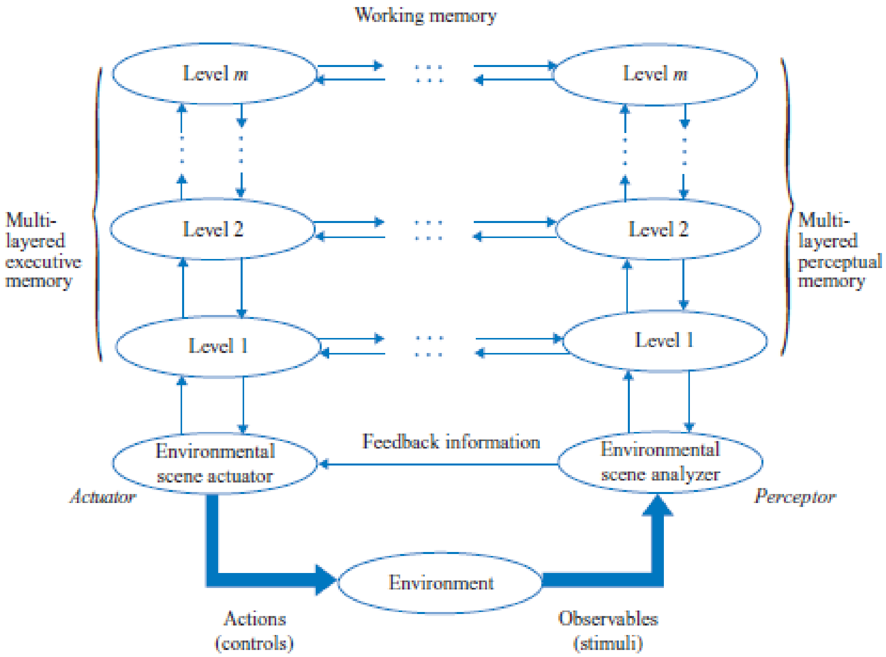


Fig1. Directed information-flow diagram in the perception–action cycle of a cognitive dynamic system with hierarchical memory.

Clearly, there is NO unique approach

Nonetheless, there is an approach inspired by the human brain, which is a complex, highly nonlinear, and distributed information-processing system.

"NEURAL NETWORKS"

2.7 Neural networks

NEURAL NETWORKS

- A machine that is designed to *model* the way in which the brain performs a particular task or function of interest
- A massively parallel distributed processor made up of simple but nonlinear processing units that has a natural propensity for storing experiential knowledge and making it available for use

Resemblance to brain

(1) Knowledge is acquired by the network from its environment through a learning process.
(2) Interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge.

Fig2. Brain

Fig3. Most neurons in the brain are connected to several thousand others

2.7 Neural networks

NEURAL NETWORKS

- **Massively parallel distributed structure**
- **Ability to learn and, therefore, can generalize**

Generalization: Production of reasonable outputs for inputs

CHARACTERISTIC of neural networks

(1) Nonlinearity: A neural network, made up of an interconnection of nonlinear neurons, is itself nonlinear

(2) Input–output mapping: A neural network consists of a unique input signal and a corresponding desired (target) response

(3) Adaptivity: Neural networks adapt their synaptic weights to changes in the surrounding environment

(4) Contextual information: Every neuron in the network is potentially affected by the global activity of all other neurons in the network

(5) Fault tolerance: If a neuron or its connecting links are damaged, a neural network exhibits a graceful degradation in performance rather than catastrophic failure

MODELS of neural networks

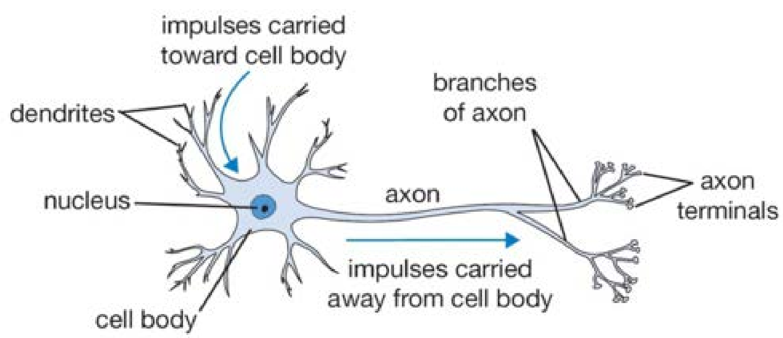


Fig4. Drawing of a biological neuron

- Dendrites: Receives input signals
- Cell body : Sum all signals
- Axon: Produces output signals
- Synapses: A small gap that connects other neurons across

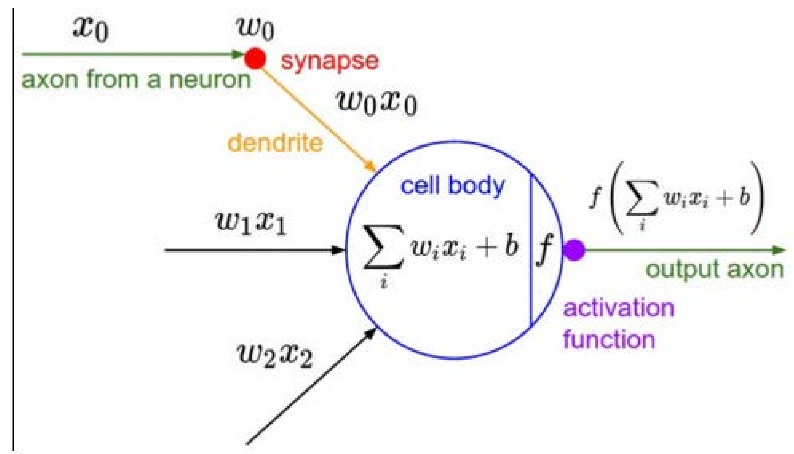


Fig5. Mathematical model of a biological neuron

- $x_0, x_1, \cdots, x_i$: Input
- $w_0, w_1, \cdots, w_i$: Synaptic weight
- b: Bias
- f: Activation function

2.7 Neural networks

MULTILAYER FEEDFORWARD NETWORKS

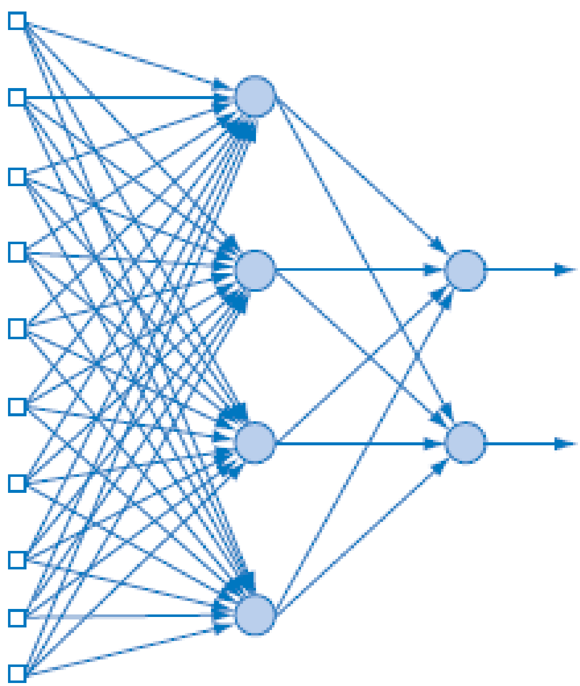


Fig6. Fully connected feedforward network with one hidden layer and one output layer

" Fully connected feedforward network with one hidden layer and one output layer "

- Fully connected: Every node in each layer of the network is connected to every other nodes

- Partially connected: Purposely missed from the network for the purpose of reduced computational complexity and, quite possibly, improved performance

2.8 Associative learning process

ASSOCIATIVE MEMORY: Brainlike distributed memory that learns by association, which has been known to be a prominent feature of human memory

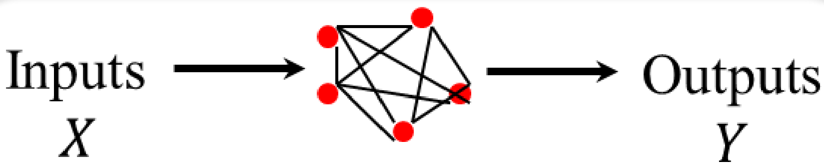


Fig7. Associative memory

Input: Pattern (often noisy/corrupted)

Output: Corresponding pattern (complete / relatively noise-free)

Process

1. Load input pattern onto group of highly-interconnected neurons.
2. Run neurons until they reach a steady state.
3. Read output of the states of the neurons.

- Autoassociation ($X = Y$)

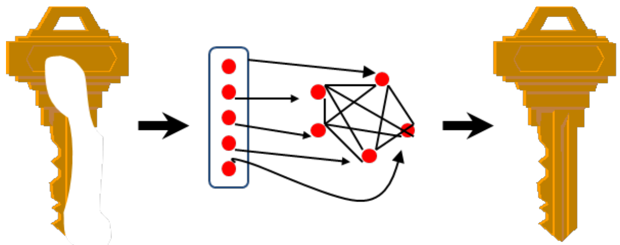


Fig8. Autoassociation

- Heteroassociation ($X \neq Y$)

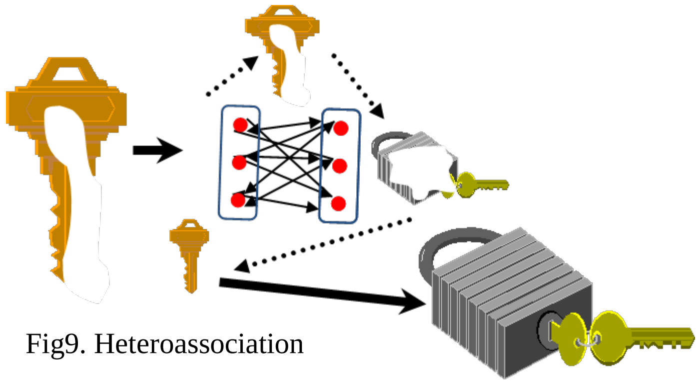


Fig9. Heteroassociation

2.8 Associative learning process

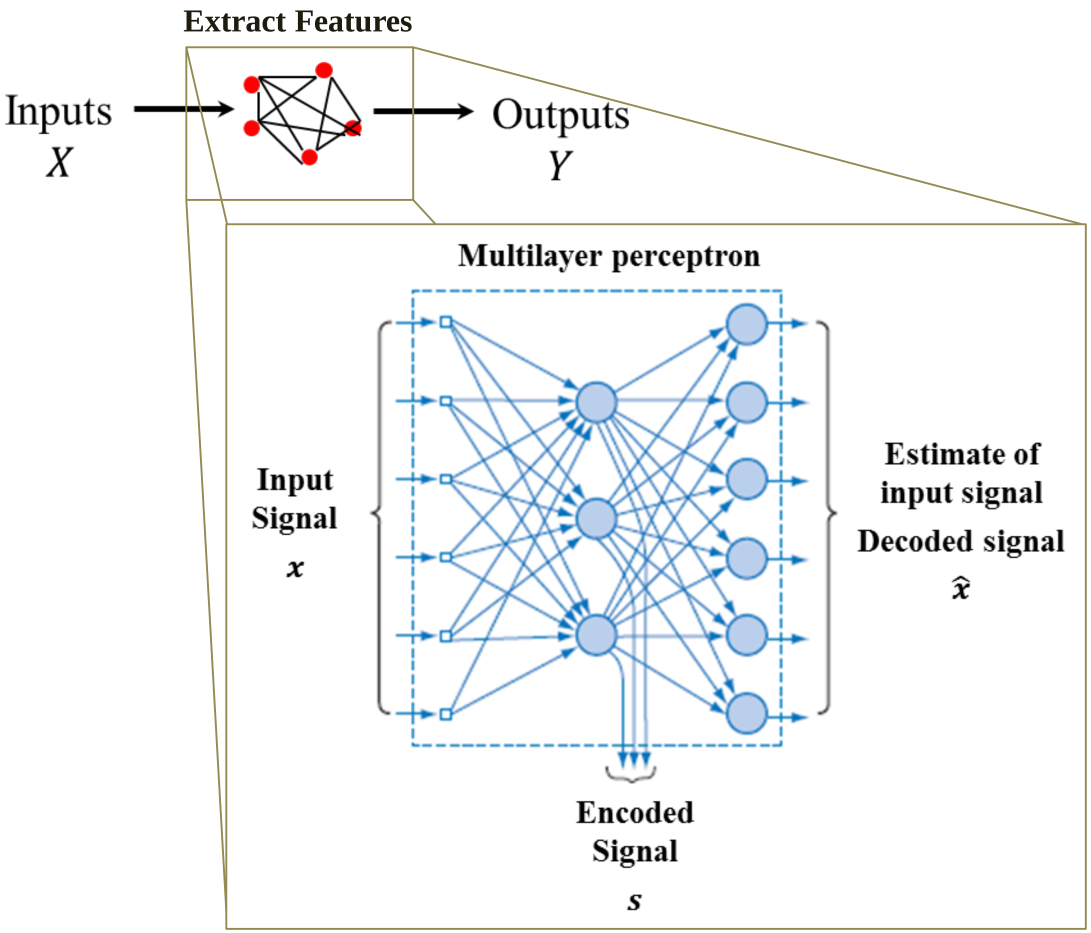


Fig10. Multilayer perceptron

Multilayer perceptron (MLP)

- Input and output layers have the same size, m
- Size of the hidden layer is smaller than m
- Network is fully connected

º ***Hidden neurons*** of MLP play a critical role as ***feature detectors***

º ***MLP*** as ***Identity mapping*** by encoding and decoding

2.9 Back-propagation

BACK-PROPAGATION: Method of training artificial neural networks

Requires ***a known, desired output***

- **Process**

Initialization
- No prior information
- Weights are picked from a uniform distribution

Forward computation
- Compute output signal by proceeding forward through the network, layer by layer

Backward computation
- Compute local gradients of the network, update weights
- **Stochastic gradient method**

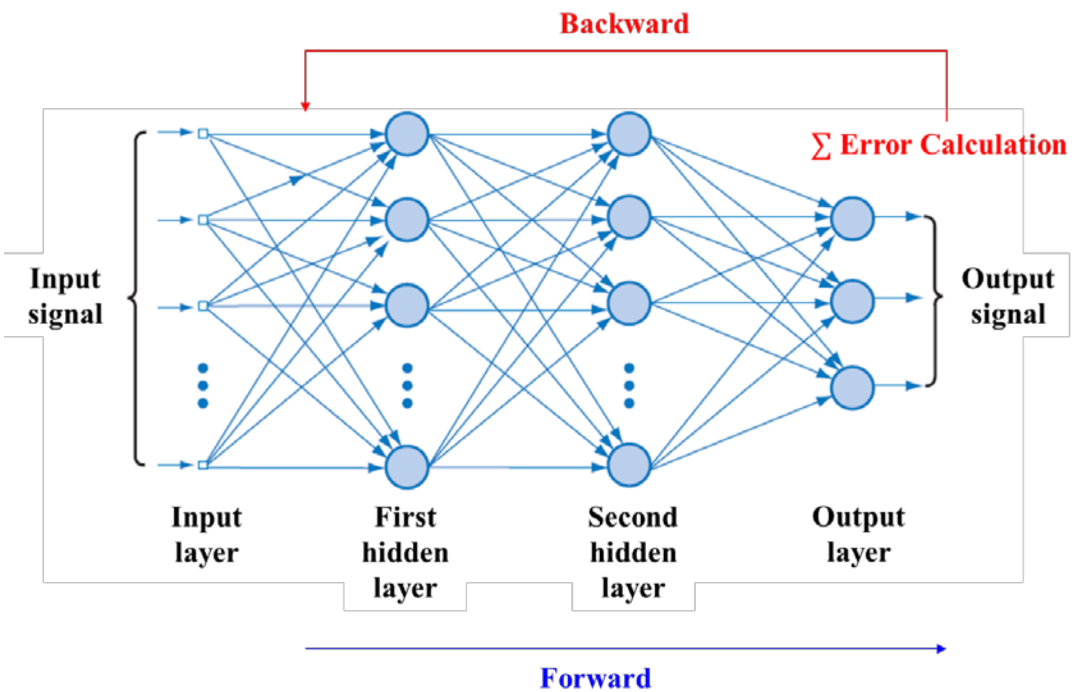


Fig11. Architectural graph of a multilayer perceptron with two hidden layers

2.9 Back-propagation

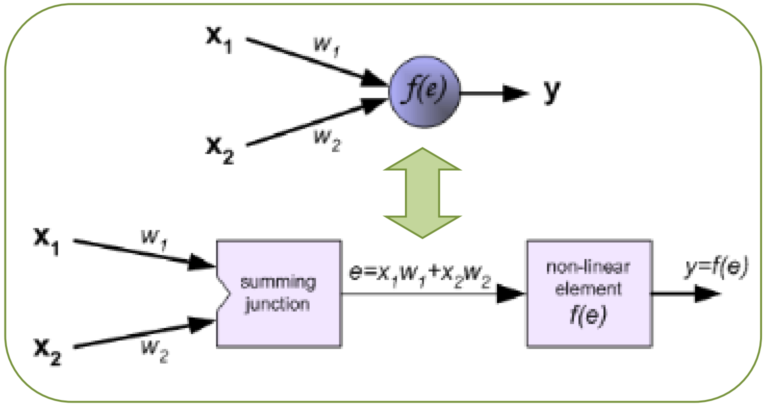


Fig12. Mathematical model of a neuron

Neuron $\boldsymbol{f(e)}$

- Input signal: x
- Summing junction : e
- Non-linear element : f (Activation function)
- Output signal: y

Example: 3 layer neural network with 2 inputs and 1 output

→ Look at teaching process of multi-layer neural network using ***backpropagation algorithm***

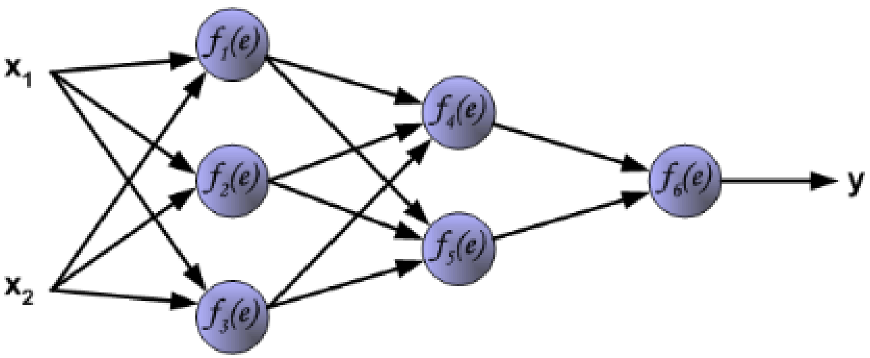
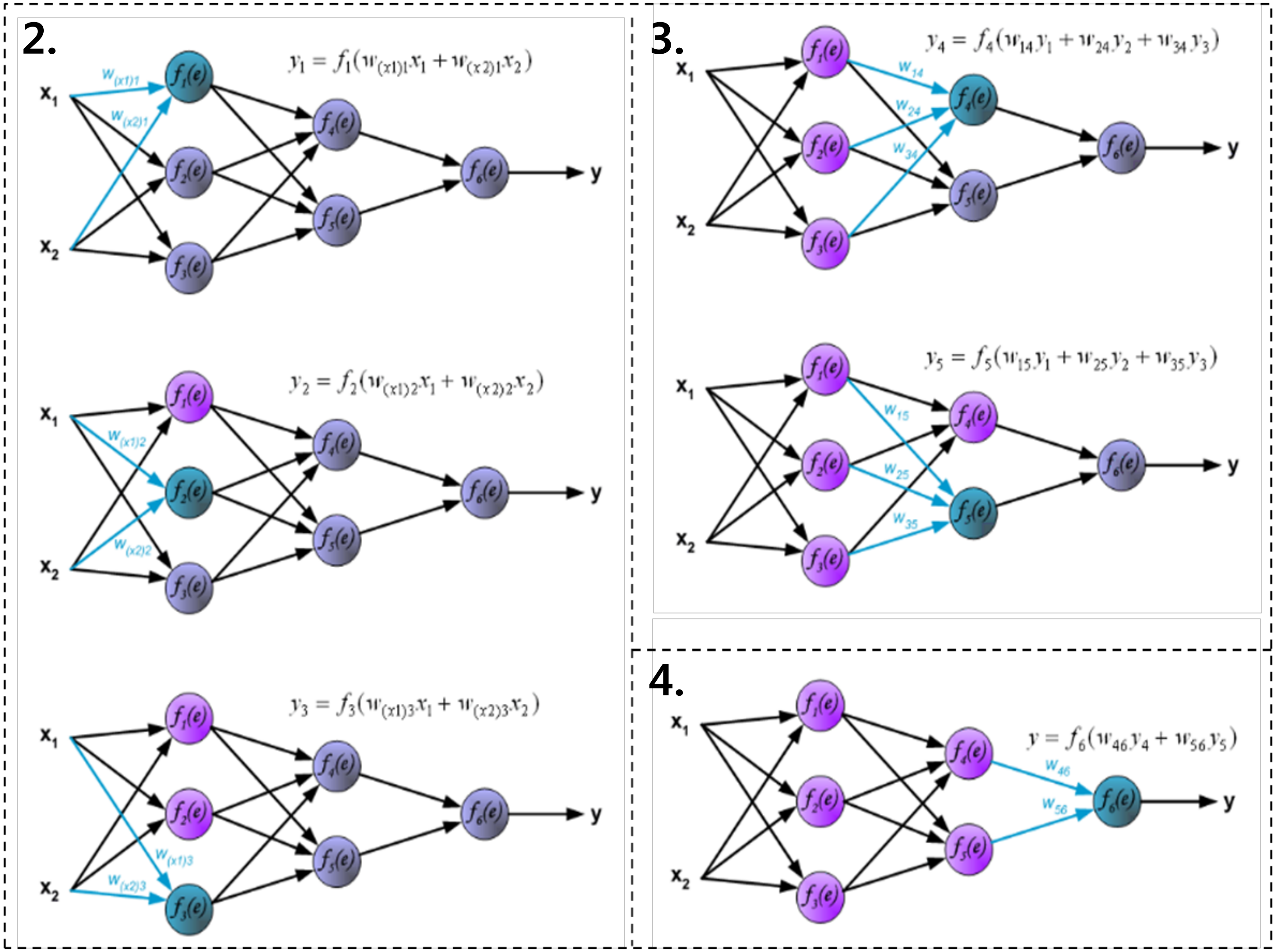


Fig13. 3 layer neural network

2.9 Back-propagation



1. Initialize

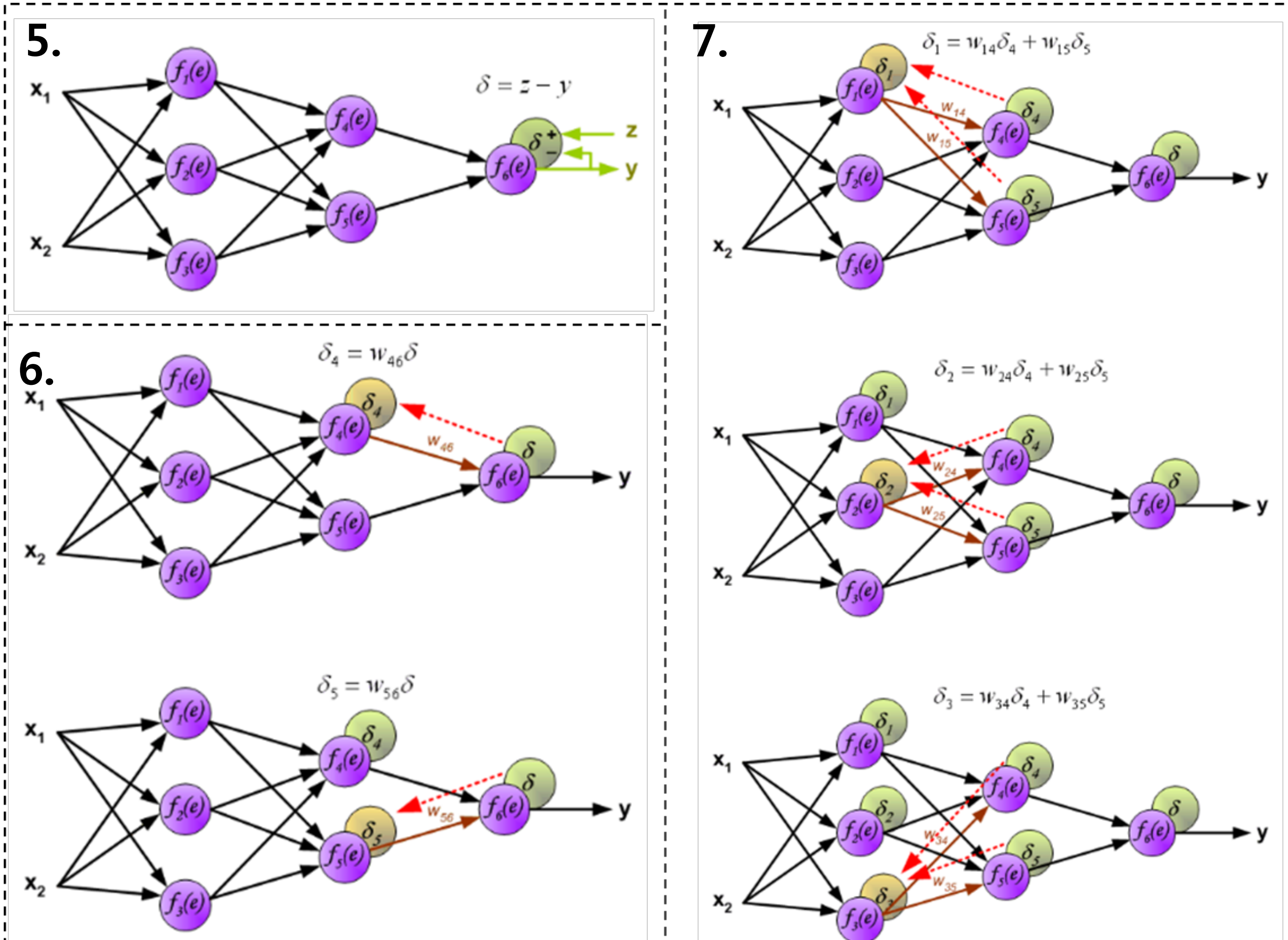
2~4. Forward computation

x_n: Input signal

$w_{(xm)n}$: weights of connection between input x_m and neuron n

y_n: Output signal of neuron n

2.9 Back-propagation

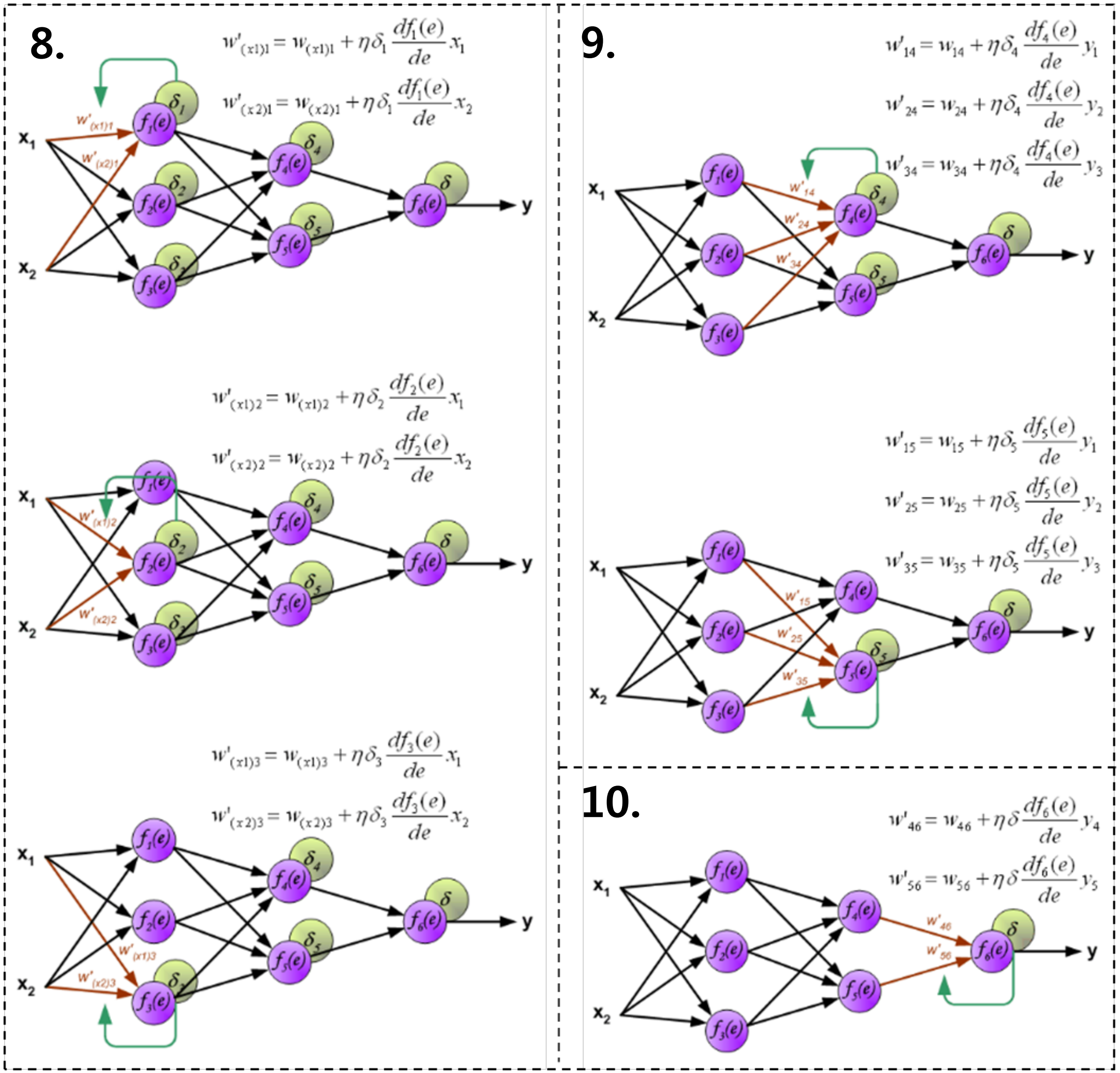


5~7. Backward computation – compute error signal for each neuron

z: Desired output

δ: Error signal

2.9 Back-propagation



8~10. Backward computation – compute weights coefficients/ modified weights of each neuron

$\frac{df(e)}{de}$: Derivative of neuron activation function

η: Teaching speed

11. Iterates 2~10 until the chosen stopping criterion

2.10 Recurrent multilayer perceptrons

Recurrent multilayer perceptrons (RMLP): A neural network with one or more hidden layers, and with each computation layer of the network having feedback around

- **Feedback loops** in the network makes the RMLP not only ***dynamic***, but also ***more computationally powerful*** compared with an ordinary MLP

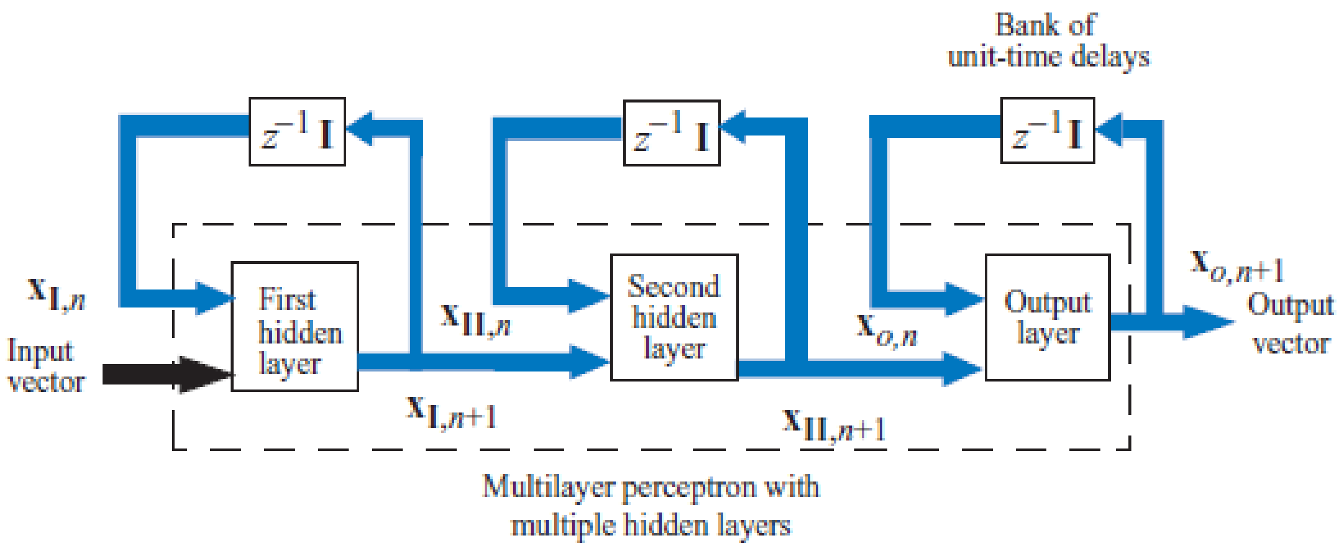


Fig14. RMLP; feedback paths in the network are printed in color

2.11 Self-organized learning

Supervised learning(back-propagation algorithm) requires a desired response vector
To overcome this limitation, **self-organized or unsupervised learning procedures**

> ***"When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place on one or both cells such that A's efficiency as one of the cells firing B, is increased."*** ***-Donald Hebb***

Generalized Hebb's algorithm (GHA)

① If two neurons on either side of a synapse (i.e. connecting link) are activated synchronously, then the weight of that synapse is significantly increased.

② If, on the other hand, the two neurons are activated asynchronously, then that synapse is selectively weakened in strength or eliminated altogether over the course of time.

$$\Delta w_{ji}(n) = ①\ \eta y_i(n) x_i(n) - ②\ \eta y_j(n) \sum_{k=1}^{j} w_{ki}(n) y_k(n)$$

$\Delta w_{ji}(n)$: **Weight change applied to synapse connecting node *j* to node i at time**

$x_i(n)$: **Input signal** $(i = 1, 2, \dots, m)$

$y_k(n)$: **Output signal** $(j = 1, 2, \dots, \mathrm{l})$

η: **Learning−rate parameter**

2.12 Summary and discussion

4 basic functions embodied in COGNITIVE DYNAMIC SYSTMES

- ***Perception:*** followed by action in the environment to feedback information
- ***Perceptual memory, executive memory, working memory:*** Memory is used to predict the consequences of actions in the system
- ***Attention:*** prioritize the available resources in the system
- ***Intelligence:*** the ability of the system to continually adjust itself through an adaptive process by responding to new changes in the environment

PERCEPTION is a probabilistic process

“ Given a set of stimuli received from the environment, estimate the hidden state of the environment *in the environmental scene analyzer as accurately as possible*”

(1) the ill-posed inverse problem → Chaper3: Power-spectrum estimation

(2) Bayesian inference problem → Chaper4: Bayesian filtering

THANK YOU