Neural Networks and Their Application to Associative Memory

Hiroshi Inaba
（稲葉　博）

Department of Information Sciences
Tokyo Denki University
（東京電機大学情報科学科）

Presented at Universiti Sains Malaysia
15 December 1997

Recent Fields of Research

1. Neural Network Theory with Application to Associative Memory

[1] H. Inaba and Y. Shoji, A Note on Construction Method of Dynamical Neural Networks, *Proc. SICE Symp. Dynamical System Theory,* pp.439-442, October 1997.

[2] Y. Shoji and H. Inaba, A Module Neural Network and Its Basic Behaviors, *Proc. IEEE International Conference on Neural Networks,* Houston, USA, vol. 2, pp. 1305-1310, June 1997.

[3] H. Inaba and Y. Shoji, Equilibrium Solutions Neural Networks with a Special Structure, *Proc. Int. Conf. Neural Networks and Signal Processing*, Nanjing, China, vol. 1, pp. 75-78, 1995.

[4] H. Inaba and H. Inaba, Muli-Module Neural Networks and Their Application to Associative Memories, *Proc. Int. Conf. Neural Networks and Signal Processing*, Nanjing, China, vol. 1, pp. 110-113, 1995.

[5] K. Ishii and H. Inaba, Associative Memories Using Dynamical Neural Networks with Limit Cycles, *Proc. ISCIE Int. Symp. Stochastic Systems Theory and Its Applications*, Osaka, Japan, pp.175-180, 1993.

2. Mathematical System Theory

[1] N. Ito and H. Inaba, Triangular decoupling over principal ideal domains, *SIAM J. Control and Opti-mization,* vol. 35, no. 3, pp. 744-765, 1997.

[2] N. Otsuka and H. Inaba, Parameter insensitive disturbance-rejection for infinite-dimensional sys-tems, *IMA J. Mathematical Control and Information,* vol.14, no.4, pp.401-413, 1997.

[3] N. Otsuka, H. Inaba and K. Toraichi, Parameter insensitive disturbance-rejection problems with incomplete-stae feedback for infinite-dimensional systems, Systems Science Journal, vol. 22,no.4, 1996.

[4] H. Inaba and W. Wang, Block-decoupling for linear sstems over rings, *Linear Algebra and Its Appli-cations,* vol. 241-242, pp. 619-634, 1996.

[5] N Otsuka, H. Inaba and K. Toraichi, Decoupling by incomplete state feedback for infinite-dimen-sional systems, *J. Japan Indus-trial and Applied Mathematics,* vol. 11. No. 3, pp. 363-377, 1994.

3. Perspective Systems Theory with Application to Machine Vision

[1] B. K. Ghosh, H. Inaba and S. Takahashi, Para-meter Identifiability of Riccati Dynamics under Perspective and Orthographic Projections, sub-mitted to *1998 International symposium on Math-ematical Theory of Networks and Systems.*

Outline

1. Introduction
2. Preliminaries
3. Module Neural Networks
4. Multi-Module Neural Networks
5. Concluding Remarks

1. Introduction

This talk presents my recent results on dynamical neural networks with McCullough-Pitts model. My talk consists of the following two parts:

In the first part:

(1) A method for constructing such a neural network so that a given set of vectors is assigned as its stable equilibrium points is discussed.

- **(i) One of basic problems in applying neural networks to associative memory is to show that when the number of information vectors to be stored approaches the number of neurons in the network the capability as associative memory suddenly decreases.**
- **(ii) To avoid this sudden decrease, a dynamical neural network having a special structure, called a module neural network, is introduced, and its basic properties are discussed.**
- **(iii) The effectiveness of this module neural network is demonstrated by computer simulation.**

In the second part:

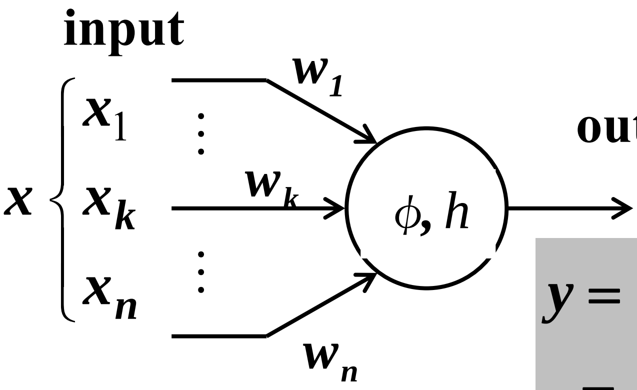
(2) Further, a different type of module neural net-works is introduced.

(i) Appropriately connecting such module neural networks, a multi-module neural network is constructed for associative memory to store more complicated information, such as English words instead of just alphabets.

(ii) Results of computer simulation for this type of networks are also presented to illustrate its effectiveness.

2. Preliminaries

2.1 A Mathematical Model of Neurons



$$y = \phi\left(\sum_{k=1}^{n} w_k x_k - h\right)$$
$$= \phi(wx - h)$$

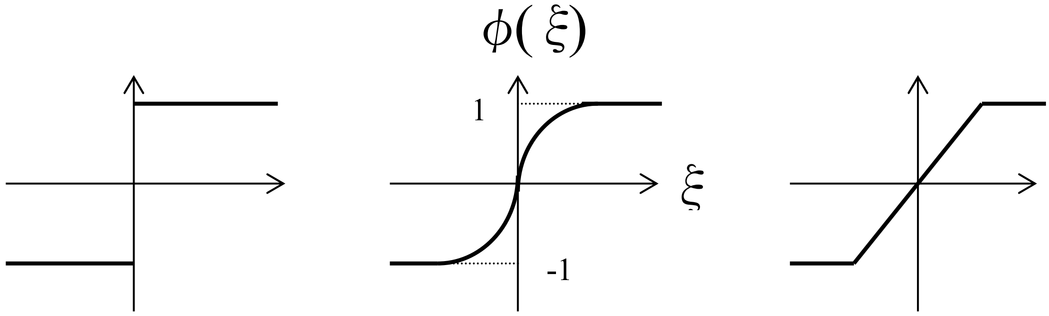
$x := [x_1 \ \cdots \ x_n]^T$ the input vector

$y \in R$ the output

$w := [w_1 \ \cdots \ w_n]$ the connection vector

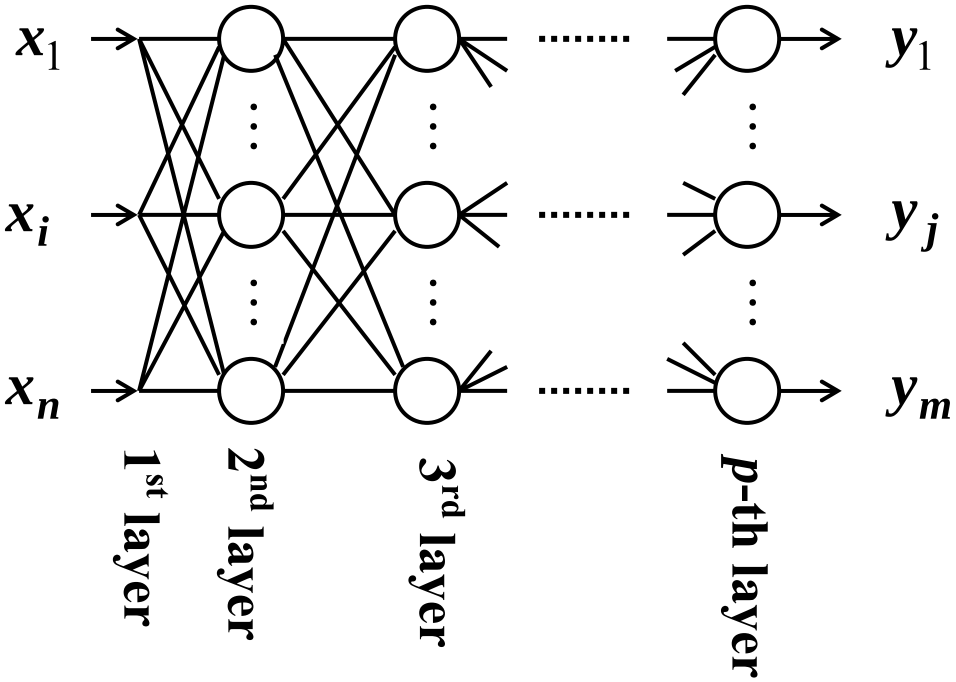
$h \in R$ the threshold

$\phi : R \to R$ the output function



Various Output Functions
(Sigmoidal Functions)

2.2 Multi-layerd Neural Networks



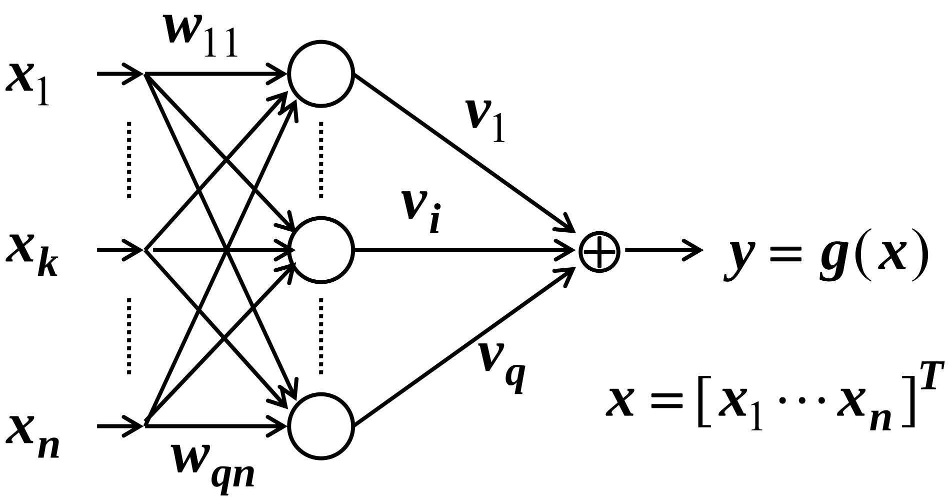
A multi-layered neural network determines a function:

$$y = g(x)$$

where

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in R^n, \qquad y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \in R^m.$$

2.3 Ability of 3-Layered Neural Networks



$$y = g(x) = \sum_{i=1}^{q} v_i\, \phi\{\sum_{k=1}^{n} w_{ik} x_k - h_i\}$$

Theorem(Cybenko, 1989)

Let $\phi: R \to R$ be any continuous sigmoidal func-tion, i.e.,

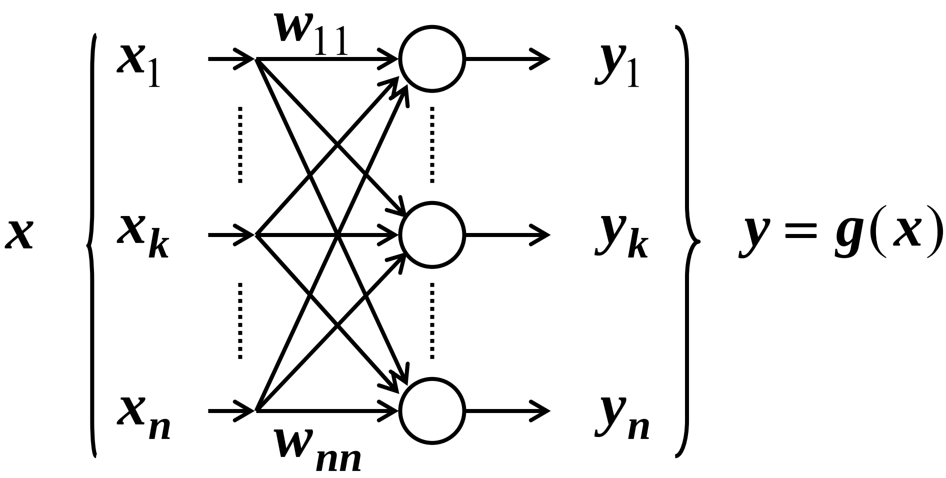
$$\phi(\xi) \to 1 \; (\xi \to \infty), \quad \phi(\xi) \to -1 \; (\xi \to -\infty),$$

and $C(D)$ denote the Banach space of all contin-uous functions $f: (D \subset R^n) \to R$ on a compact set $D \subset R^n$. Then, for any $f \in C(D)$ and any $\varepsilon > 0$ there exists a 3-layered neural network $g(x)$ such that

$$|f(x) - g(x)| < \varepsilon \quad \text{for } \forall x \in D.$$

That is to say that the set of all functions generated by 3-layered neural networks is dense in $C(D)$.

2.4 Dynamical Neural Networks



$$x := [x_1 \cdots x_n]^T \qquad y := [y_1 \cdots y_n]^T$$

$$W := \begin{bmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \cdots & w_{nn} \end{bmatrix} \qquad h := [h_1 \cdots h_n]^T$$

$$y = g(x) = \phi\{Wx - h\}$$

Then, the function $g(x) = \phi\{Wx - h\}$ defines a discrete-time dynamical neural network as

$$y(k+1) = \phi\{Wy(k) - h\}, \quad y(0) = x \qquad (1)$$

and the output sequence generated by (1) is given by

$$y(0) = x, \quad y(1), \quad y(2), \quad \cdots.$$

2.5 Stability of Dynamical Systems

Consider a discrete-time dynamical system on R^n of the form

$$y(k+1) = g(y(k)), \quad y(0) = \xi \in R^n \quad (2)$$

and denote the solution by $y(k;\xi)$.

Definition

(i) A vector $y^* \in R^n$ is called an equilibrium point or equilibrium solution of (2) if

$$g(y^*) = y^*.$$

(ii) An equilibrium point $y^* \in R^n$ is called asymptotically stable if there exists a $\delta > 0$ such that

$$\|y - y^*\| < \delta \Rightarrow y(k;y) \to y^* \quad (k \to \infty).$$

(iii) For an equilibrium point $y^* \in R^n$, the set defined by

$$D(y^*) := \{y \in R^n \,|\, y(k;y) \to y^* \, (k \to \infty)\}$$

is called the domain of attraction for y^*.

2.6 The McCullough-Pitts Model

$\boldsymbol{R} :=$ the set of real numbers

$\boldsymbol{B} := \{-1, 1\}$

$\boldsymbol{B}^n := \{\boldsymbol{x} = [x_1 \cdots x_n]^T \mid x_i \in \boldsymbol{B}, i = 1, \cdots, n\}$

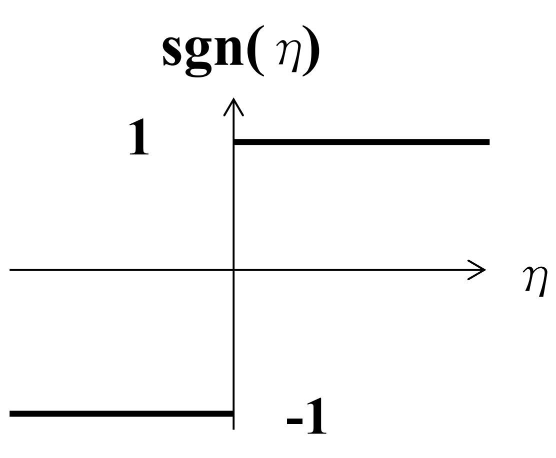
Definition. The **McCullough-Pitts model** is a dynamical system on $\boldsymbol{B}^n$ of the form

$$\text{MP}: \begin{cases} \boldsymbol{y}(k+1) = \text{Sgn}(\boldsymbol{W}\boldsymbol{y}(k) - \boldsymbol{h}) \\ \boldsymbol{y}(0) = \boldsymbol{x} \in \boldsymbol{B}^n \end{cases}$$

where Sgn is the **sign function** defined as

$$\text{Sgn}(\xi) = [\text{sgn}(\xi_1) \cdots \text{sgn}(\xi_n)]^T, \quad \xi \in \boldsymbol{R}^n$$

$$\text{sgn}(\xi_i) := \begin{cases} 1 & \text{if } \xi_i > 0 \\ -1 & \text{if } \xi_i \le 0 \end{cases}$$



2.7 Associative Memory

(1) Address Addressable Memory (AAMM)

In a digital computer, a desired set of information is recalled by giving the correct address of the memory storage in which the information is stored.

(2) Associative Memory (AM)

A full set of information in memory is recalled by giving a portion of the memory's information or an incomplete information about the desired information. (Content Addressable Memory)

The main problem involved in implementing AM's by neural networks includes:

(a) How to store each desired information as a vector corresponding to an asymptotically stable equilibrium point of the network,

(b) How to control the extent of the domains of attraction of the asymptotically stable equili-bria of the network, and

(c) **How to minimize the number of asymptotically stable equilibria in the network that are not used for storing information.**

3. Module Neural Networks

The objective of this section is to

(1) construct a dynamical neural network having a special structure, called a module neural network, such that

(a) a given set of vectors is assigned to its equilibrium points,

(b) such an equilibrium point is asymptotically stable, and

(c) there are no periodic solutions with period ≧2, and

(2) apply a module neural network to associative memory to illustrate the effectiveness of the proposed method.

4. Multi-Module Neural Networks

The objective of this section is to

(1) introduce a different type of module neural networks(MNN) and then a multi-module neural networks(MMNN) by mutually connecting a finite number of MNN's,

(2) propose a method for constructing an MMNN so that a given set of vectors is assigned to its equilibrium points, and

(3) finally apply the results obtained to con-structing associative memory to store English words and perform computer simulation to illustrate the effectiveness.

5. Concluding Remarks

In the first part:

(1) A neural network having a special structure, called a module neural network, was introduced, and a method was proposed for constructing connection matrices of such a network so that a prescribed set of vectors is assigned as its asymptotically stable equilib-rium points.

(2) A sufficient information about the domain of attraction for each equilibrium point was obtained.

(3) Some computer simulation for associative memory to store English alphabets was performed to illustrate the effectiveness of the proposed module neural network.

In the second part:

(4) First introducing a different type of module neural network from the first part, a muti-module neural network(MMNN) is proposed to store more complicated information such as English words instead of just alphabets.

(5) A method was also proposed for constructing the connection matrices of such an MMNN without showing asymptotical stability of the assigned equilibrium points.

(6) However, some computer simulation per-formed for associative memory to store English words strongly suggested that all the equilibrium points assigned by the proposed method are asymptotically stable.