

A Relationship between the Cascaded Network and the Two-Layer Network of Threshold Logic Units

RYUZO TAKIYAMA

Kyushu Institute of Design, 226 Shiobaru, Minami-ku, Fukuoka, 815 Japan

It is proved that, given a cascaded network of K threshold logic units, there exists an equivalent two-layer network of $2K$ threshold logic units; furthermore, parameter values of the latter network are given in terms of those of the former network.

I. INTRODUCTION

Networks of threshold logic units (TLU) that may be used for information processing have been studied from various points of view. Especially they have been treated in conjunction with synthesizing logical functions (Winder, 1962; Dertouzos, 1965; Lewis and Coates, 1967; Muroga, 1971), building models of nervous systems (Amari, 1971, 1972, 1977; Rosenblatt, 1962; Kohonen, 1977; Nakano, 1972) and pattern recognition (Cover, 1965; Nilsson, 1965; Cadzow, 1968; Holdermann, 1971; Weaver, 1975; Osborne, 1977; Takiyama, 1978). Although considerable achievements have been obtained on various subjects in above regions, there remain many problems to be solved and relatively little is known about the networks. In order to improve the situation, it may be a great help to investigate in detail the functional properties of fundamental TLU networks, and clarify the relationships or differences between them. From this point of view, Takiyama (1976) has proved that the two-layer TLU network with direct inputs to its second layer can be equivalently transformed to one without direct inputs to its second layer.

The cascaded TLU network is a fundamental network, which is made use of in various information processing systems. In this paper it is proved that when a cascaded network with K threshold logic units is given, it can be equivalently transformed to a two-layer network with $2K$ units. Furthermore, parameter values of the two-layer network are given in terms of those of the cascaded network. This fact reveals the differences between two fundamental networks described above, and suggests a relationship between parallel and serial information processing.

II. THE CASCADED NETWORK OF THRESHOLD LOGIC UNITS

Let the input be represented by n -dimensional vector $x = (x_1, \dots, x_n)$, where x_i is assumed to take any real value. A cascaded TLU network which consists of K units is shown in Fig. 1. The weight vector to x and the threshold of the j th unit, T_j , are

$$b_j = (b_{j1}, \dots, b_{jn}), \quad j = 1, \dots, K, \quad (1)$$

and

$$-b_{j0}, \quad j = 1, \dots, K, \quad (2)$$

respectively. The connecting weight between T_j and T_{j-1} is

$$c_j, \quad j = 2, \dots, K. \quad (3)$$

Then the weighted sum of the inputs to T_j becomes as follows.

$$g_j(x) = b_j \cdot x + b_{j0} + c_{j+1} \Theta[g_{j+1}(x)], \quad j = 1, \dots, K-1, \quad (4)$$

$$g_K(x) = b_K \cdot x + b_{K0}, \quad (5)$$

where $b_j \cdot x$ denotes the inner product of b_j and x , and $\Theta[\]$ means the threshold mapping, i.e.,

$$\begin{aligned} \Theta[p] &= 1, & p &> 0, \\ &= -1, & p &\leq 0. \end{aligned} \quad (6)$$

The discriminant function $g(x)$ realized by the cascaded TLU network is $g_1(x)$, i.e.,

$$g(x) = g_1(x), \quad (7)$$

whose parameters are b_j , b_{j0} , $j = 1, \dots, K$, and c_j , $j = 2, \dots, K$.

Without loss of generality, let us assume

$$c_j > 0, \quad j = 2, \dots, K, \quad (8)$$

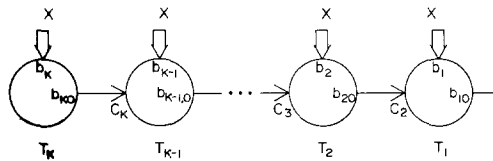


Fig. 1. Cascaded network of K threshold logic units.

and consider the following $2K - 1$ linear functions.

$$g_j^-(x) = b_j \cdot x + b_{j0} - c_{j+1}, \quad j = 1, \dots, K - 1, \quad (9)$$

$$g_j^+(x) = b_j \cdot x + b_{j0} + c_{j+1}, \quad j = 1, \dots, K - 1, \quad (10)$$

$$g_K(x) = b_K \cdot x + b_{K0}. \quad (11)$$

Denoting the n -dimensional space of x by R^n , we observe that R^n is divided into three disjoint subregions by parallel hyperplanes, $g_j^-(x) = 0$ and $g_j^+(x) = 0$. Any x which satisfies $g_j^-(x) > 0$ always satisfies $g_j^+(x) > 0$ by way of (8) (see Fig. 2). Let us consider the following $2K$ subregions of R^n .

$$\begin{aligned} V_1 &= \{x \mid g_1^-(x) > 0\} \\ V_2 &= \{x \mid g_1^+(x) \leq 0\}, \\ V_{2j-1} &= \{x \mid g_j^-(x) > 0, -c_{s+1} < b_s \cdot x + b_{s0} \leq c_{s+1}, s = 1, \dots, j-1\}, \\ V_{2j} &= \{x \mid g_j^+(x) \leq 0, -c_{s+1} < b_s \cdot x + b_{s0} \leq c_{s+1}, s = 1, \dots, j-1\}, \\ &\quad j = 2, \dots, K, \end{aligned} \quad (12)^1$$

where

$$g_K^-(x) = g_K^+(x) = g_K(x). \quad (13)$$

We put

$$X_1 = \bigcup_{j=1}^K V_{2j-1}, \quad (14)$$

$$X_2 = \bigcup_{j=1}^K V_{2j}, \quad (15)$$

where V_{2j-1} and V_{2j} are defined in (12).

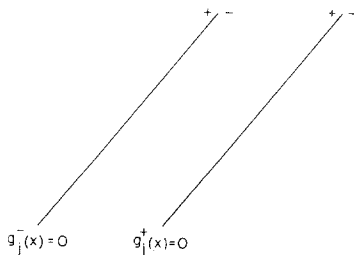
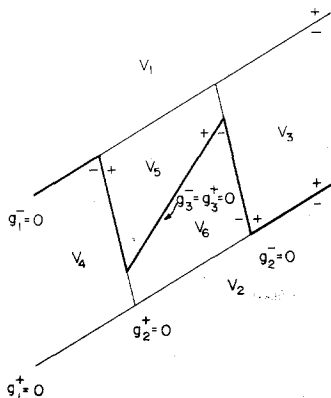


FIG. 2. Partition of R^n by $g_j^-(x) = 0$ and $g_j^+(x) = 0$.

¹ Note that $-c_{s+1} < b_s \cdot x + b_{s0} \leq c_{s+1}$ means $g_s^-(x) \leq 0$ and $g_s^+(x) > 0$.

FIG. 3. Discriminant regions by the cascaded network ($K=3$).

Noting that the inequality $-c_{s+1} < b_s \cdot x + b_{s0} \leq c_{s+1}$ in (12) means $g_s^-(x) \leq 0$ and $g_s^+(x) > 0$, and taking into account (9), (10), (11) and Fig. 2, we have the following lemma.

LEMMA 1. *It holds that*

$$X_1 \cup X_2 = R^n \quad (16)$$

and

$$X_1 \cap X_2 = \text{empty}. \quad (17)$$

An illustrative example for $K=3$ is given in Fig. 3. Further we have the following lemma.

LEMMA 2. *It holds that*

$$g(x) > 0 \quad \text{if } x \in X_1, \quad (18)$$

$$g(x) < 0 \quad \text{if } x \in X_2. \quad (19)$$

Proof. Since $g_j^-(x) = b_j \cdot x + b_{j0} - c_{j+1} > 0$ for $x \in V_{2j-1}$, it holds that

$$g_j(x) > 0.$$

Therefore we have

$$\begin{aligned} g_{j-1}(x) &= b_{j-1} \cdot x + b_{j-1,0} + c_j \Theta[g_j(x)] \\ &= b_{j-1} \cdot x + b_{j-1,0} + c_j. \end{aligned}$$

Further since

$$-c_{s+1} < b_s \cdot x + b_{s0} \leq c_{s+1}, \quad s = 1, \dots, j-1,$$

for $x \in V_{2j-1}$, we see that

$$g_{j-1} > 0.$$

In the same way, it can be said that $g_{j-2}(x) > 0, \dots, g_2(x) > 0$ and $g_1(x) > 0$, that is,

$$g(x) = g_1(x) > 0 \quad \text{if } x \in V_{2j-1}.$$

For $x \in V_{2j}$, it can be proved in the same way. Thus the lemma holds.

X_1 and X_2 are called the discriminant regions of the cascaded TLU network. In Fig. 3, an example of the boundary of discriminant regions is shown by thick lines.

III. AN EQUIVALENT TRANSFORMATION TO THE TWO-LAYER NETWORK

Now let us consider a two-layer network which consists of $L + 1$ threshold logic units (L units in the first layer) as shown in Fig. 4. This is a Type I network defined by Takiyama (1976), and is the committee machine-type network (Nilsson, 1965). Let the weight vector and the threshold value of the m th unit, T'_m , in the first layer be

$$a_m = (a_{m1}, \dots, a_{mn}), \quad m = 1, \dots, L, \quad (20)$$

$$-a_{m0}, \quad m = 1, \dots, L, \quad (21)$$

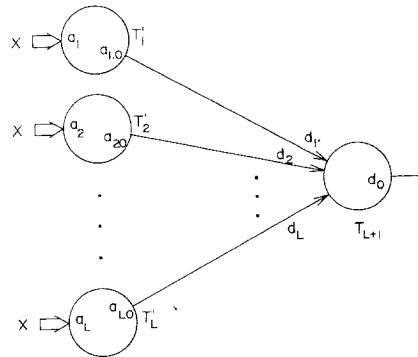


FIG. 4. Two-layer network of $L + 1$ threshold logic units.

respectively. Further let the connecting weight between T'_m and T'_{L+1} be

$$d_m, \quad m = 1, \dots, L, \quad (22)$$

and the threshold value of T'_{L+1} be

$$-d_0. \quad (23)$$

Then the discriminant function $\tilde{g}(x)$ realized by this two-layer network is described as follows.

$$\tilde{g}(x) = \sum_{m=1}^L d_m \Theta[a_m \cdot x + a_{m0}] + d_0, \quad (24)$$

whose parameters are a_m , a_{m0} , d_m , $m = 1, \dots, L$, and d_0 .

THEOREM. *The cascaded network with K threshold logic units can be equivalently transformed to the two-layer network with $2K$ threshold logic units.*

Proof. In the two-layer network shown in Fig. 4, let parameters be as follows.

$$L = 2K - 1, \quad (25)$$

$$\left. \begin{aligned} a_{2j-1} &= b_j, & j &= 1, \dots, K-1, \\ a_{2j} &= b_j, & j &= 1, \dots, K-1, \\ a_{2K-1} &= b_K, \\ a_{2j-1,0} &= b_{j0} - c_{j+1}, & j &= 1, \dots, K-1, \\ a_{2j,0} &= b_{j0} + c_{j+1}, & j &= 1, \dots, K-1, \\ a_{2K-1,0} &= b_{K0}, \end{aligned} \right\} \quad (26)$$

and

$$d_m = 2^{2K-1-m}, \quad m = 1, \dots, 2K-1, \quad (27)$$

$$d_0 = \frac{2^{2K-1} - 2}{3}, \quad (28)$$

where b_j , etc., on the right-hand side in (26) are parameters of the cascaded network discussed in the previous section. From (25) we see that this is a

two-layer network with $2K$ units ($2K - 1$ units in the first layer). We will show that the two-layer network given by (25)–(28) has the same discriminant regions X_1 and X_2 shown in Lemma 2, that is,

$$\tilde{g}(x) = \sum_{m=1}^{2K-1} d_m \Theta[a_m \cdot x + a_{m0}] + d_0 > 0 \quad \text{if } x \in V_{2j-1}, \quad (29)^2$$

$$\tilde{g}(x) = \sum_{m=1}^{2K-1} d_m \Theta[a_m \cdot x + a_{m0}] + d_0 < 0 \quad \text{if } x \in V_{2j}. \quad (30)^2$$

Defining $(2K - 1)$ -dimensional vectors

$$z = (\Theta[a_1 \cdot x + a_{10}], \dots, \Theta[a_{2K-1} \cdot x + a_{2K-1,0}]),$$

$$d = (d_1, \dots, d_{2K-1}),$$

then $\tilde{g}(x)$ can be written

$$\tilde{g}_{(x)} = d \cdot z + d_0.$$

If we denote z for $x \in V_{2j-1}$ by z^{2j-1} , it takes the following form.

$$z^{2j-1} = \underbrace{(\overbrace{-1, 1, -1, 1, \dots, -1, 1, 1, 1, *, \dots, *}^{2(j-1)})}_{2K-1}, \quad j = 1, \dots, K-1,$$

$$z^{2K-1} = (\underbrace{-1, 1, -1, 1, \dots, -1, 1, 1}_{2(K-1)}),$$

where $*$ designates 1 or -1 .

On the other hand, for $x \in V_{2j}$, it takes the following form.

$$z^{2j} = \underbrace{(\overbrace{-1, 1, -1, 1, \dots, -1, 1, -1, -1, *, \dots, *}^{2(j-1)})}_{2K-1}, \quad j = 1, \dots, K-1,$$

$$z^{2K} = (\underbrace{-1, 1, -1, 1, \dots, -1, 1, -1}_{2(K-1)}).$$

² Note that $a_{2j-1} \cdot x + a_{2j-1,0} = g_j^-(x)$, $a_{2j} \cdot x + a_{2j,0} = g_j^+(x)$ and $a_{2K-1} \cdot x + a_{2K-1,0} = g_K(x)$.

A simple calculation yields

$$\begin{aligned} d \cdot z^{2j-1} &> d \cdot z^{2K-1} & (j \neq K), \\ d \cdot z^{2j} &< d \cdot z^{2K} & (j \neq K), \end{aligned}$$

and

$$\begin{aligned} d \cdot z^{2K-1} &= -\frac{2^{2K-1} - 2}{3} + 1, \\ d \cdot z^{2K} &= -\frac{2^{2K-1} - 2}{3} - 1. \end{aligned}$$

Since

$$d_0 = \frac{2^{2K-1} - 2}{3} = -\frac{1}{2}(d \cdot z^{2K-1} + d \cdot z^{2K}),$$

it holds that

$$\begin{aligned} d \cdot z^{2j-1} + d_0 &> 0, & j = 1, \dots, K, \\ d \cdot z^{2j} + d_0 &< 0, & j = 1, \dots, K, \end{aligned}$$

hence (29) and (30).

VI. CONCLUSIONS

In order to make use of TLU networks for a variety of problems in information processing, it may be a great help to clarify in detail the functional properties of fundamental TLU networks. From this point of view, we have proved that the cascaded network with K units can be equivalently transformed to the two-layer network with $2K$ units. Parameter values of the two-layer network can be determined by those of the cascaded network. As the result, differences and relationships between two fundamental networks described above have fairly been made clear.

ACKNOWLEDGMENTS

The author wishes to express his gratitude to Professor Yosiro Oono of Kyushu University as well as Professor Shun-ichi Amari of the University of Tokyo for their encouragement.

RECEIVED: January 18, 1980; REVISED: August 1, 1980

REFERENCES

- AMARI, S. (1971), Characteristics of randomly connected threshold element networks and network systems, *Proc. IEEE* **59**, 35–47.
- AMARI, S. (1972), Learning patterns and pattern sequences by self-organizing nets of threshold elements, *IEEE Trans. C-21*, 1197–1206.
- AMARI, S. (1977), Neural theory of association and concept-formation, *Biol. Cybernet.* **20**, 175–185.
- CADZOW, J. A. (1968), Synthesis of nonlinear decision boundaries by cascaded gates, *IEEE Trans. C-17*, 1165–1172.
- COVER, T. M. (1965), Geometrical and statistical properties of linear inequalities with applications in pattern recognition, *IEEE Trans. Electron. Comput.* **EC-14**, 326–334.
- DETOUZOS, M. L. (1965), "Threshold Logic," MIT Press, Cambridge, Mass.
- HOLDERMANN, F. (1971), Classification by cascaded threshold elements, *Pattern Recognition* **3**, 243–251.
- KOHQENEN, T. (1977), "Associative Memory," Springer-Verlag, Berlin.
- LEWIS, P. M. AND COATES, C. L. (1967), "Threshold Logic," Wiley, New York.
- MUROGA, S. (1971), "Threshold Logic and Its Applications," Wiley-Interscience, New York.
- NAKANO, K. (1972), Association—a model of associative memory, *IEEE Trans. Systems Man Cybernet.* **SMC-2**, 381–388.
- NILSSON, N. J. (1965), "Learning Machines: Foundation of Trainable Pattern Classifying Systems," McGraw-Hill, New York.
- OSBORNE, M. (1977), Seniority logic: A logic for a committee machine, *IEEE Trans. Comput.* **C-26**, 1302–1306.
- ROSENBLATT, F. (1962), "Principles of Neurodynamics," Spartan Books, Washington, D.C.
- TAKIYAMA, R. (1976), A relationship between two types of two-layer threshold element pattern-recognition networks, *IEEE Trans. Comput.* **C-25**, 1136–1139.
- TAKIYAMA, R. (1978), A general method for training the committee machine, *Pattern Recognition* **10**, 255–259.
- WEAVER, C. S. (1975), Some properties of threshold logic unit pattern recognition networks, *IEEE Trans. Comput.* **C-24**, 290–298.
- WINDER, R. O. (1962), "Threshold Logic," Ph.D. dissertation, Princeton University, Princeton, N. J.