

STUDIES ON THRESHOLD LOGIC

by

KAZUHIRO SUGATA

DEPARTMENT OF ELECTRICAL ENGINEERING

KYOTO UNIVERSITY

DECEMBER, 1967

STUDIES ON THRESHOLD LOGIC

by

KAZUHIRO SUGATA

DEPARTMENT OF ELECTRICAL ENGINEERING

KYOTO UNIVERSITY

DECEMBER, 1967

DOC

1968

2

電気系

PREFACE

Threshold Logic is the branch of Switching Theory that deals with the use of threshold elements as the fundamental building blocks for the synthesis of digital machines. Digital systems such as electronic digital computers, digital communication systems, and digital control systems ordinarily contain networks whose input and output variables have only two distinct values, namely binary-valued. The analysis and synthesis of such digital networks are the main material of Switching Theory.

A threshold element is a device with a single binary-valued output and a number of binary-valued inputs. The output value of the device is solely dependent on whether or not the weighted inputs sum exceeds a real number, called threshold. This mathematical model is applicable to certain areas such as Decision Theory and Adaptive Control. Physically, parametrons, magnetic cores, and Esaki diodes based on the threshold principle have been used for digital networks.

Recently, considerable attentions have been given on the threshold logic because of its logically powerful ability compared with the conventional logic gates such as AND, OR, NAND, or NOR gates. Therefore a given switching function can generally be implemented with fewer threshold elements by using this gates properly. Since new industrial techniques can now mechanize reliable, inexpensive and minute

threshold elements, it promises that this logic will be used more and more for digital systems in future.

On account of its increased logical complexity, however, the developments of a new switching theory about threshold gates are strongly required in order to use this gates efficiently.

This thesis deals with Threshold Logic mostly through the linear algebra aspects. The linear inequality system inherent to a threshold function will be treated elaborately. This approach will be easy to comprehend and achieve further insights and will give us an intuitive conception.

Chapter II consists of basic developments for this theory. Some fundamental relations and theorems which further advancements of our study base on are presented mainly in this chapter. Here the material is indicated as the autonomous network problem for some purpose. However this is hardly different from the ordinary problem handled in Threshold Logic.

Chapter III treats the physical weight values and threshold value which realize a given threshold function. These values' range is stated as the domain in Euclidean space.

The test for the linear separability is treated in Chapter IV, in Chapter V and in Chapter VI. Three distinct methods are proposed in each chapter. The first method is based on the solutions of the adjoint linear equation system. The second one is stated as the determinantal condition of certain matrices. The last one is based on the successive transformations

and eliminations of variables.

Chapter VII and Chapter VIII treat the synthesis procedures to realize any given transition of states for the autonomous network.

Chapter IX deals with the synthesis procedure to realize any given Boolean function through the successive decomposition.

Chapter X discusses the similar subjects, mentioned hitherto, by using the simplified inequality system based on the characteristic vector's properties.

Our study for the threshold logic started originally from the investigation for the nervous network. For this purpose, the electric analogue of the neuron was constructed at the first stage. This analogous device not only is one kind of threshold devices but also has other features such as the refractory period, the frequency modulation ability and so on. This analogue and its properties are presented in Chapter XI.

December, 1967

Kazuhiro Sugata

ACKNOWLEDGEMENT

I wish to express my deep thanks especially to Professor Toshiyuki Sakai who supervised this thesis and provided constant encouragements and guidances during the entire course of this work. I am also greatly indebted to Professor Chihiro Hayashi for his useful suggestions and generous supports.

I am very much indebted to Dr. Hideo Fukutome and Dr. Hiroshi Tamura who aroused my interest in this subject and initiated this study and gave me valuable discussions. Particularly, the content in Chapter XI was done under co-operation of them. I am grateful to Dr. Hidenosuke Nishio for his encouraging suggestions, comments and discussions.

I wish to thank Mr. Akibumi Makinouchi who made useful discussions and helped the study. I am also indebted to Mr. Muneo Iwamoto and Mr. Youichi Kobuchi for their fine co-operations, helps and discussions.

Finally it is a pleasure to thank the staff and students of Professor Toshiyuki Sakai's Laboratory, especially Dr. Shuji Doshita and Dr. Makoto Nagao for their daily helps during my study.

CONTENTS

CHAPTER I INTRODUCTION

1.1	Threshold Element	1
1.2	Boolean Function and Threshold Function . . .	5
1.3	Explanation of the Problem	8
1.4	Preliminary Consideration	11

CHAPTER II BASIC DEVELOPMENT

2.1	Outline of the Network	15
2.2	Notations and Definitions	16
2.3	Inherent Inequality System	23
2.4	Reformulation in the Equation Form	29
2.5	Fundamental Solution Matrix	32

CHAPTER III ARGUMENT ON WEIGHTS AND THRESHOLD SPACE

3.1	Weights and Threshold Values in R^{N+1}	34
3.2	Domain of Weights and Threshold	36
3.3	Introduction to Galois Field	45
3.4	Bounds for the Number of Divisions	50

CHAPTER IV LINEAR SEPARABILITY BY FUNDAMENTAL SOLUTION MATRIX

4.1	The Basic Background	56
4.2	Development	60
4.3	Example	66

CHAPTER V LINEAR SEPARABILITY BY VARIABLE TRANSFORMATION

5.1	Preliminary Consideration	75
5.2	Simplification Procedure by Elimination . .	77
5.3	Simplification Procedure by Transformation .	83
5.4	Expansion of the Concept of Curtailed Matrix	94
5.5	Practical Procedure for Linear Separability	97
5.6	Example	100

CHAPTER VI LINEAR SEPARABILITY BY DETERMINANTAL CONSISTENCY CONDITION

6.1	Basic Principle	105
6.2	Development	108
6.3	Example	111

CHAPTER VII SYNTHESIS PROCEDURE BY SUCCESSIVE SUPPLEMENT

7.1	Basic Principle	114
7.2	Practical Procedure for Synthesis	123
7.3	Example	126

CHAPTER VIII GENERAL SYNTHESIS PROCEDURE

8.1	Basic Principle	129
8.2	Weights and Threshold Values	133
8.3	Example	137

CHAPTER IX	MULTIGATE SYNTHESIS OF BOOLEAN FUNCTION	
9.1	Basic Principle	140
9.2	Synthesis Procedure	142
9.3	Example	148
CHAPTER X	LINEAR MAPPING OF FIRST QUADRANT	
10.1	Characteristic Vector and its Properties . .	150
10.2	Reduction of Inequalities	152
10.3	Development	158
10.4	Example	161
CHAPTER XI	AN ELECTRIC ANALOGUE OF THE NEURON	
11.1	Introduction	167
11.2	The Axon Analogue	169
11.3	The Synapse Analogues	177
BIBLIOGRAPHY		183

SPECIAL SYMBOLS AND ABBREVIATIONS

(The number indicates the page on
which the symbol first appears)

$\in$	is a member of	
$\notin$	is not a member of	
Φ	empty set	
∇	the set which contains all the elements	
$\cap$	intersection	
$\cup$	union	
$\subseteq$	is contained in	
$\supset$	contains	
$\gg$	sufficiently large	
$\mathbf{v} > 0$	positive vector	16
$\mathbf{v} \geq 0$	nonzero nonnegative vector	16
$\mathbf{v} \geq 0$	nonnegative vector	16
R^N	N-dimensional Euclidean space	
E_N	N-dimensional identity matrix	23
$\mathbf{e}_i$	i-th unit vector	23
$GF(p)$	Galois field composed of p elements	45
$K(H)$	polyhedral convex cone generated from matrix H	38
θ_i	threshold value	16
w_{ij}	weight value from j to i	16
w_i		17
$\hat{u}_j$		17

u_j		18
φ_1		18
$B(i)$		32
$c_{jj}(i)$		21
$c(i)$		21
$c_I(i)$		21
$c_{II}(i)$		21
P_1		20
Q_1		20
U		19
U_I		19
U_{II}		19
v_j		77
F	transition state diagram	20
$f \cdot \hat{u}_j$	the successor of $\hat{u}_j$ in F	20

CHAPTER I

INTRODUCTION

1.1 Threshold Element

Theoretical aspects of the logical device based on the threshold principle were first investigated by McCulloch and Pitts in 1943. Later von Neumann also considered this logic to discuss the reliable network composed of unreliable elements. They regarded the threshold element as a mathematical model of a neuron of living things rather than as a logical element for the engineering.

Since then, considerable attention has been given to the threshold logic. In particular, the availability of physical elements operating on the threshold principle seems to have stimulated further developments of the study of this logic. Some mathematical aspects of the threshold function were studied by Gilbert in 1954, however appreciable studies have been done since Karnaugh and Guterman⁽²⁷⁾ pointed out to use magnetic cores as threshold devices.

At present, most digital computers are constructed from the fundamental building blocks such as AND, OR, NAND and NOR gates. All of these conventional gates have some very excellent properties mentioned below.

- (1) All of them can be easily constructed from simple elements such as diodes, resistors and transistors.

Moreover they require not so much tolerance or reliability.

- (2) Their input-output relations are connected in the convenient forms for the Boolean expression. Besides, the input-output relation of a network synthesized is ordinarily described by these logical operations. Therefore the theory of the analysis and synthesis for digital networks using these gates will be tractable and unified. This fact is very desirable especially for the practical design engineer of digital systems.

These factors may suggest surely in some sense that it is entirely natural to choose these conventional gates for the construction of digital networks. However, there is no strict reason only such simple conventional gates must be adopted as the fundamental building blocks for digital networks.

In fact, as electronic digital computers are becoming more and more extensively used not only for arithmetic calculations but for various information processings, our requirements on their speed, ability and compactness increase more and more strongly. To aim at more speed and compactness, it needs at least a logically more powerful basic building block. In actual, the number of gates that are required to implement a given Boolean function can be practically decreased by using more powerful fundamental building blocks. Still more, this fact promises the possibility of constructing faster digital computers only if the switching time of the more powerful gate

is as small as the conventional gates.

Furthermore the promise of an intrinsic decrement in the number of gates will give rise to the decrement in the number of internal interconnections. This is particularly significant for the sake of decreasing the circuit complexity.

The threshold gate is certainly one of the logical devices that is endowed with these properties. In fact, since the Boolean function implemented by a single threshold gate is relatively complex compared with the conventional gates. In other words, any given Boolean function can generally be realized with less threshold gates. Therefore the proper use of threshold gates provides the possibility of increments in speed and savings of gates together with decrements of the circuit complexity.

Besides above all mentioned advantages, there exists another very principal advantage that the Boolean function implemented by threshold gates can be easily modified. Since the Boolean function realized by threshold gates is completely determined by weights values and threshold value setting, it is easy to change the Boolean function by adjusting weights values and threshold value.⁽⁴¹⁾

In order to modify the Boolean function of the network constructed from the conventional gates, it needs to change the internal interconnections. This is rather troublesome operations. Hence, this indicates that threshold devices are suitable for the network whose Boolean function has to be adjusted by the learn-^{(26) (36)}

ing process based on the past experiences.

The recent emergence of new technologies for the manufacture of reliable, minute and inexpensive threshold elements provides the possibility more surely that threshold gates can be used as the fundamental building blocks to realize any Boolean function in the same manner as AND, OR, NAND and NOR gates.

These factors seem to have aroused much attention in the threshold device. On the other hand, because of its logical complex performance, it requires further developments of the theory about Threshold Logic so that this device might be used efficiently. Considerable researches have been done by many investigators at various laboratories. These efforts have much contributed to develop the theory of Threshold Logic. However many interesting and challenging problems still remain.

Although the recent rapid growth of studies for Threshold Logic may be due to above mentioned factors, the positive attitude to grasp the information processing mechanism of living organisms is one of undeniable causes. ⁽⁵¹⁾⁽⁶¹⁾ Actually, appreciable works of Threshold Logic have been done in association with the neurons, since McCulloch and Pitts. It is not surprised that the attempt to analysing the nervous network's behavior leads to analysing the threshold devices network's behavior in the course of natural.

The properties of a threshold device have been surveyed.

To summarize, the advantages of a threshold device from the engineering point of view consist in

- (1) its logically powerful ability
- (2) its elastic behavior
- (3) its relative economy of physical realization.

1.2 Boolean Function and Threshold Function

In general, a Boolean function is defined as follows. Let X denote the set which consists of the two values of the binary logic. The two elements in X are usually denoted by the two integers 1 and 0. For any given positive integer n , consider the Cartesian power

$$X^n = X \otimes X \otimes \dots \otimes X$$

which is the Cartesian product of n copies of X .

Thus, the elements of X^n are given as the 2^n ordered n -tuples

$$(x_1, x_2, \dots, x_n)$$

where the k -th coordinate x_k is an element in X for every $k=1, 2, \dots, n$. Hereafter, X^n will be called the n -cube and its 2^n elements are called the input points or vertices. By a Boolean function $f(x_1, x_2, \dots, x_n)$ of n variables, we mean a function

$$f : X^n \rightarrow X$$

from the n -cube X^n into X . In other words, a Boolean function $f(x_1, x_2, \dots, x_n)$ of n variables is defined by assigning one

of two integers in X to each 2^n input point. Thus, there are 2^{2^n} distinct Boolean functions of n variables.

A threshold function is one class of Boolean functions which can be realized by a single threshold gate. A threshold gate is a device with a single binary-valued output and a number of binary-valued inputs. A real number is given in association with each input which is referred to as "weights".

The output of the device takes a constant value denoted by the logical value 0 unless the weighted sum of the inputs exceeds a real number, referred to as "threshold". On the other hand, if the weighted sum of the inputs exceeds the threshold, the output of the device takes a different constant value denoted by the logical value 1.

Therefore, a threshold device is defined by the following relations:

$$\left. \begin{aligned} f &= 1 \quad \text{if and only iff} \quad \sum_{j=1}^n w_j \cdot x_j > \theta \\ &= 0 \quad \text{if and only iff} \quad \sum_{j=1}^n w_j \cdot x_j \leq \theta \end{aligned} \right\} \quad (1.1)$$

where

- f = binary output of the device, 1 or 0
- x_j = j -th binary input to the device, 1 or 0
- w_j = weight of the j -th input, a real number
- n = total number of the inputs
- θ = threshold, a real number

Observe that $x_1, x_2, \dots, x_n$ are interpreted in Relation 1.1 as real numbers 0 or 1 rather than as Boolean values.

A given Boolean function $f(x_1, x_2, \dots, x_n)$ is said to be linearly separable if and only if there exist real numbers $w_1, w_2, \dots, w_n, \theta$ such that Relation 1.1 holds. A threshold function is referred to as various names such as a linear separable function, a majority or a voting function, and a linear input function. These names come from its behavior.

It is convenient to consider the geometrical interpretation of a Boolean function. This will serve to give an intuitive interpretation to some of the definitions and theorems given in the subsequent discussions.

Let us consider the n -cube in n -dimensional Euclidean space where each coordinate axis corresponds to one of the input variables. It is wholly immediate to establish a one-to-one correspondence between the set of vertices of the n -cube and the set of arguments of a Boolean function of n variables.

If the Boolean function takes the value 1 for a given argument, the corresponding vertex is said to be an element of the function. On the other hand if the opposite case is true, the vertex is said to be an element of the complement function. In this way, the value 0 or 1 which the Boolean function takes, divides the set of vertices into two classes.

If a given Boolean function is a threshold function, there exists an $(n-1)$ -dimensional hyperplane π which effects this

separation. That is to say, π separates the on-set $f^{-1}(1)$ from the off-set $f^{-1}(0)$. Therefore, the on-set of the threshold function lies on one side of π , and the off-set lies on the other side of π . The hyperplane π is called a separating hyperplane of the threshold function. The separating hyperplane can be expressed in the following form.

$$\pi : \omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n = \theta \quad (1.2)$$

For all those vertices lying on one side of the hyperplane π , the linear form of the left hand of Equation 1.2 takes the value greater or smaller than the threshold θ .

This geometrical interpretation is of great value when it is used to explain the properties that have been derived abstractly, even if it has the disadvantage of being nonvisual for the case of more than three variables.

1.3 Explanation of the Problem

With respect to threshold devices there are two basic problems. The first is the practical fabrication of reliable and inexpensive threshold elements which depends on the technical level at that time. The second is about Threshold Logic which does not depend on the technological level. Threshold Logic can be discussed apart from the physical elements. It is Threshold Logic that is dealt with henceforth.

Since only logical properties are concerned with, the theory of Threshold Logic is not necessarily limited to the logical

design, but can be applied to the system where the mathematical model of the threshold principle dominates such as Decision Theory.

In general, Switching Theory primarily consists of two problems;

- (1) specification of the input-output relation of the digital network often in the form of a truth table or a Boolean function.
- (2) the method of the synthesis of the digital network based on its input-output relation by using certain given fundamental building blocks.

The fundamental building blocks selected by the engineer are ordinarily determined by various criteria such as logical abilities, availabilities of physical elements, easiness of design procedures and reliabilities. At present AND, OR and NOT gates are usually used and its synthesis procedures are well known.

Once threshold gates are selected because of some requirements, let us consider the subject (2) mentioned concerning Switching Theory. This subject can be stated in details as follows.

- (1) Conditions which a Boolean function must satisfy to be a threshold function, that is, linear separability conditions. (9)(29)(54)
- (2) Actual procedures for finding the physical realization in terms of weights values and threshold value,

when it is a threshold function.

- (3) Practical algorithms for determining whether a given Boolean function is a threshold function or not.
(13)(39)(56)

- (4) Actual methods for synthesizing any given Boolean function with generally more than one threshold element.
(4)(7)(11)(34)

This thesis handles the materials mentioned above individually in each chapter.

These materials, however, will be reformulated in rather different forms from the subjects reviewed above. This is due to the motivation from which our study arised. However, there exists no essential difference between the original and reformulated materials. That is, the subjects will be treated as the problems of the autonomous network constructed from N threshold elements, instead of treating a single threshold element. This reformulation will be seen to be only the expansion of treating a single threshold element.
(3)

The problem of the autonomous network can be stated as follows. When a behavior of the autonomous network composed of N threshold elements is given as a transition state diagram, consider the conditions and practical procedures to testify whether or not such a given behavior is realizable with N threshold elements. This reformulation of the problem will correspond to the areas (1), (3) mentioned above.

If any given behavior of the autonomous network is seen to be realizable with N threshold elements, let us consider the systematic determination procedures of the actual weights values and threshold value which realize the given behavior. On the other hand, if the given behavior can not be realized with N threshold elements, consider the methods which realize the given behavior using some additional control threshold elements. Note here that the given behavior is realized with more than N threshold elements in the meaning that we are concerned only with its behavior of the network constructed from the originally given N threshold elements. In other words, it is only N threshold elements given from the beginning that their functions are observed. Additional attached threshold elements only serve to control the behavior of the autonomous network and their functions are not observed.

1.4 Preliminary Consideration

The motivation that the autonomous network of threshold elements was first treated, originates from the attempt to use the nervous network as an information processing filter. Here the information processing means the two-dimensional optical pattern recognition.

To explain the motivation more in details, consider a two-dimensional arrayed nervous network constructed from N neurons. The nervous network has 2^N internal states on the

whole if each neuron is assumed to have two states, namely, fire or rest. Attach to each neuron a receptor such as a photo-transistor which feels the optical pattern. The sensed input signal is amplified appropriately if necessary and transmitted to the corresponding neuron.

At some instant of time, a two-dimensional optical pattern comes into this nervous network system. Assume moreover that no input pattern comes successively for a while. Then the nervous network is initialized to one of 2^N internal states corresponding to the input pattern. Hence we are considering the autonomous network whose initial internal state is solely set by the external input pattern.

Furthermore suppose that the internal state of the network is transited autonomously and synchronously time to time from the initial state. After some lapse of time, the internal state of the network is changed periodically or absorbed into a certain (37) internal state, since there are only finite internal states and yet the following internal state is uniquely dependent on the previous internal state.

Let us try to determine the interconnection coefficients between the neurons and each threshold value to satisfy the following condition. The set of internal states of the network makes groups in the transition state diagram so that the grouped internal states may transit to each other within the internal states corresponding to input patterns of one symbol which are

deformed in various forms and may never transit between the internal states corresponding to the distinct symbols.

Thus all the internal states which belong to one group in the transition state diagram correspond to variously deformed patterns of one symbol. If such a synthesis of a nervous network is possible, the internal state always either passes over periodically or falls into the internal state corresponding to one of the representative symbol patterns after some lapse of time.

Therefore it will be possible to discriminate two-dimensional optical patterns by considering that some groups of the internal states in the transition state diagram make one symbol of the patterns or one category of the patterns. Since the internal state of the autonomous network exposed to an input pattern always becomes the internal state corresponding to one of the representative patterns, the output of to which category the input pattern belongs can be readily indicated.

Hitherto the autonomous network has been considered whose initial state is set by the external input. As long as we are concerned with the realization problem of such an autonomous network, this will be only the modification of Threshold Logic of a single threshold element. However, the synthesis problem of such a network will not necessarily become only the modification. In some phase the subject will be mainly discussed as the problem of the autonomous network and in another phase the subject will be discussed as the problem of Threshold Logic of one

threshold element like usual works investigated up to date.

However, the results obtained in the autonomous network problems can be immediately transformed to the results of Threshold Logic.

CHAPTER II

BASIC DEVELOPMENT

2.1 Outline of the Network

Although many electrical analogues for a neuron are proposed by various investigators, the essential behavior is based on the threshold principle. The following relations

$$u_i(t+\tau) = \Psi \left\{ \sum_{j,r} w_{ij}^{(r)} \cdot u_j(t-r\tau) - \theta_i \right\} \\ (i = 1, 2, \dots)$$

$$\Psi(x) = \begin{cases} 1, & \text{for } x > 0 \\ 0, & \text{for } x \leq 0 \end{cases}$$

are proposed as the fundamental behavior equations of the nervous network. ⁽⁸⁾ Here,

$w_{ij}^{(r)}$ = the coupling coefficient that transfers the pulse generating in neuron j at the instant of $(t-r\tau)$ to neuron i at the instant of t . τ is time delay.

θ_i = threshold of neuron i .

$u_j(t)$ = the state of neuron j at the instant of t , namely, 1 or 0 corresponding to fire or rest, respectively.

In actual, however, there is no necessity to treat Relation 2.1 itself. It is sufficient to deal with the case where the present internal state is completely dependent on the just

previous internal state. That is to say, the behavior equations of the nervous network are given as follows.

$$u_i(t+\tau) = \Psi \left\{ \sum_{j=1}^N w_{ij} \cdot u_j(t) - \theta_i \right\} \quad (i=1, 2, \dots, N) \quad (2.2)$$

$$\Psi(x) = \begin{cases} 1, & \text{for } x > 0 \\ 0, & \text{for } x \leq 0 \end{cases}$$

where $w_{ij} = w_{ij}^{(0)}$. This will be easily admitted by imaging additional neurons instead of taking account of the past effects $w_{ij}^{(r)}$ ($r=1, 2, \dots$). Therefore, there exists no substantial difference between Relation 2.1 and Relation 2.2.

Hereafter let us consider the autonomous network constructed from N threshold elements, $T_1, T_2, \dots, T_N$ whose behaviors are expressed by Relation 2.2. The network transits synchronously with time delay τ .

2.2 Notations and Definitions

If all the components v_i of a vector $\mathbf{v} = (v_1, v_2, \dots, v_n)$ are positive, let us denote $\mathbf{v} > 0$. If all the components v_i are nonnegative, let us denote $\mathbf{v} \geq 0$. Furthermore, if besides $\mathbf{v} > 0$, there is at least one positive component in $\mathbf{v}_i$ ($i=1, 2, \dots, N$), let us denote $\mathbf{v} \geq 0$. Hereafter, the expressions $\mathbf{v} > 0$, $\mathbf{v} \geq 0$ and $\mathbf{v} \geq 0$ are distinguished in such meanings. These will be called a positive vector, a nonzero nonnegative vector and a nonnega-

tive vector, respectively.

Likewise, the similar notations $v < 0$, $v \leq 0$, and $v \geq 0$ can be defined. With regard to a matrix H , the similar notations are also used such as $H > 0$, $H \geq 0$ and $H \leq 0$.

The letter "t" written at the left-hand upper corner of a vector or a matrix denotes the transpose of the vector or the matrix. The letter "-1" written at the right-hand upper corner of a matrix denotes the inverse of the matrix. By the terms of a (m, n) -type matrix, we mean the matrix has m rows and n columns.

Definition 2.1. Let $\hat{w}_1$ and w_1 denote the N -dimensional vector and the $(N+1)$ -dimensional vector, respectively, as follows.

$$\hat{w}_1 = (w_{11}, w_{12}, \dots, w_{1N})$$

$$w_1 = (-\theta_1, w_{11}, w_{12}, \dots, w_{1N})$$

These $\hat{w}_1$ and w_1 are called the weight vector and the threshold-weight vector.

Definition 2.2. Assume that each threshold element takes either of two states denoted by integer 1 or 0. Let u_j denote generally the state of element T_j . Then, the ordered set $(u_1, u_2, \dots, u_N)$ will be called the internal state vector. There exist 2^N distinct internal state vectors on the whole in a network constructed from N threshold elements. These 2^N distinct internal state vectors will be labelled by $\hat{u}_1, \hat{u}_2, \dots, \hat{u}_{2^N}$ in ascending order of these binary numbers. However, make an

exception in labeling $\hat{u}_{j+1}$ ($j=1, 2, \dots, N$). Namely, $\hat{u}_{j+1}$ represents the internal state vector such that only $u_j = 1$ and all the other $u_i = 0$ ($i=1, 2, \dots, N; i \neq j$). That is;

$$\begin{aligned}\hat{u}_1 &= (0, 0, 0, \dots, 0) \\ \hat{u}_2 &= (1, 0, 0, \dots, 0) \\ \hat{u}_3 &= (0, 1, 0, \dots, 0) \\ &\dots\dots\dots \\ \hat{u}_{N+1} &= (0, 0, 0, \dots, 1) \\ \hat{u}_{N+2} &= (1, 1, 0, \dots, 0) \\ &\dots\dots\dots \\ \hat{u}_{2^N} &= (1, 1, 1, \dots, 1)\end{aligned}$$

Observe that the labeling defined above for $\hat{u}_{N+2}, \hat{u}_{N+3}, \dots, \hat{u}_{2^N}$ bears no particular significance other than a convenience of notation. In Chapter III this labeling will be altered for a convenience of analysis by regarding the set of all the internal state vectors as Galois field $GF(2^N)$.

Definition 2.3. Consider the $(N+1)$ -dimensional vector u_j ($j=1, 2, \dots, 2^N$) whose first component is always 1 and the remainder components equal to the components of $\hat{u}_j$. That is:

$$u_j = (1, \hat{u}_j) \quad , \quad (j = 1, 2, \dots, 2^N)$$

This vector u_j is referred to as the content vector.

Definition 2.4. Consider the function ϕ_i defined over the set of internal state vectors that takes as its value the i -th component of the internal state vector.

That is:

$$\varphi_i(\hat{u}_j) = u_{ji} \quad , \quad (j, i=1, 2, \dots, N)$$

where $\hat{u}_j = (u_{j1}, u_{j2}, \dots, u_{jN})$.

Definition 2.5. Consider the $(2^N, N+1)$ -type matrix U whose j-th row ($j=1, 2, \dots, 2^N$) is the content vector u_j . This matrix is referred to as the universal matrix.

That is:

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & 0 & \dots & 1 \\ 1 & 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & 1 & \dots & 1 \end{pmatrix}$$

Furthermore consider the submatrix U_I that consists of the first $(N+1)$ rows of U. This U_I is referred to as the primary universal matrix.

That is:

$$U_I = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Consider the submatrix U_{II} that consists of the last $(2^N - N - 1)$ rows of U. This U_{II} is referred to as the secondary

universal matrix.

Proposition 2.1. The rank of the primary universal matrix is $(N+1)$.

The terms of the "internal state" or "state" will be used in place of the "internal state vector" for brevity, if it does not make any confusion from the context.

The behavior of the autonomous network can be, in general, expressed by assigning, for all the internal states, to what state the present one transits at the next time. The diagram which illustrates this behavior is called the transition state diagram and will be denoted by the notation "F".

Definition 2.6. Let the notation $\hat{u}_h = f \cdot \hat{u}_j$ denote the relation between $\hat{u}_h$ and $\hat{u}_j$ such that in a given transition state diagram "F", $\hat{u}_j$ is just followed by $\hat{u}_h$.

Definition 2.7. Consider two sets P_i and Q_i that consist of all such internal states as satisfy the condition, $\varphi_i(f \cdot \hat{u}_j) = 1$ or $\varphi_i(f \cdot \hat{u}_j) = 0$, respectively.

That is:

$$\begin{aligned} P_i &= \left\{ \hat{u}_j ; \varphi_i(f \cdot \hat{u}_j) = 1 \right\} \\ Q_i &= \left\{ \hat{u}_j ; \varphi_i(f \cdot \hat{u}_j) = 0 \right\} \end{aligned} \quad (i = 1, 2, \dots, N)$$

These two sets P_i and Q_i are called the positive set and the negative set, respectively.

Following two relations always hold as the consequence of

the previous definition.

Proposition 2.2.

$$\begin{aligned} P_i \cup Q_i &= \nabla \\ P_i \cap Q_i &= \Phi \end{aligned} \quad (i = 1, 2, \dots, N)$$

where ∇ and Φ denote the set composed of the whole internal states and an empty set, respectively.

Definition 2.8. Consider the $(2^N, 2^N)$ -type diagonal matrix $C(i)$. Let $c_{jj}(i)$ denote the j -th row and the j -th column component of $C(i)$ which takes a value either 1 or -1 as follows.

$$c_{jj}(i) = \begin{cases} 1, & \text{if and only if } \hat{u}_j \in P_i \\ -1, & \text{if and only if } \hat{u}_j \in Q_i \end{cases}$$

This diagonal matrix $C(i)$ is referred to as the characteristic matrix with respect to element T_i . (See page 16.)

Decompose $C(i)$ into the direct sum of two submatrices as follows.

$$C(i) = \begin{pmatrix} C_I(i) & 0 \\ 0 & C_{II}(i) \end{pmatrix}$$

where $C_I(i)$ is the $(N+1, N+1)$ -type diagonal matrix and $C_{II}(i)$ is the $(2^N - N - 1, 2^N - N - 1)$ -type diagonal matrix. $C_I(i)$ and $C_{II}(i)$ are called the primary characteristic matrix and the secondary characteristic matrix, respectively, with regard to element T_i .

These arguments result in the following proposition.

Proposition 2.3. The characteristic matrix $C(i)$ is another representation of a Boolean function.

Proof: There exist, by Definition 2.8, 2^{2^N} distinct characteristic matrices which have a one-to-one correspondence to each Boolean function of N variables. Thus, to give a characteristic matrix is equivalent to give a Boolean function.

Henceforth, let us denote the Boolean function by f_i which the characteristic matrix $C(i)$ represents.

Proposition 2.4. For a given transition state diagram of the network with N threshold elements, N characteristic matrices $C(i)$ ($i=1, 2, \dots, N$) are completely determined. Conversely, if N characteristic matrices are given, then a transition state diagram F can be drawn uniquely.

Proof: To give a transition state diagram is equivalent to give N Boolean functions $f_1, f_2, \dots, f_N$ of N variables.

From now, the relations mentioned in Proposition 2.4 are denoted as follows:

$$\begin{aligned} g(F) &= C(1) \oplus C(2) \oplus \dots \oplus C(N) \\ g^{-1}(C(1) \oplus C(2) \oplus \dots \oplus C(N)) &= F \end{aligned} \tag{2.3}$$

Proposition 2.5.

$$\begin{aligned} C(i)^{-1} &= C(i) \\ C_I(i)^{-1} &= C_I(i) \\ C_{II}(i)^{-1} &= C_{II}(i) \end{aligned}$$

Proof: $C(i)$ is a diagonal matrix whose diagonal components are all either 1 or -1.

Proposition 2.6. ${}^t C(i) = C(i)$

${}^t C_I(i) = C_I(i)$

${}^t C_{II}(i) = C_{II}(i)$

Definition 2.9. Let E_N denote the (N, N) -type identity matrix. Let e_i denote the i -th row of E_N , referred to as the i -th unit vector.

Proposition 2.7. The inverse of the primary universal matrix U_I^{-1} always takes the following form.

$$U_I^{-1} = \begin{pmatrix} 1 & 0 \\ -t_1 & E_N \end{pmatrix}$$

where -1 implies the N -dimensional vector $(-1, -1, \dots, -1)$.

2.3 Inherent Inequality System

Behavior Equation 2.2 implies Relation 2.4.

$$u_i(t+\tau) = \begin{cases} 1, & \text{iff } \sum_{j=1}^N \omega_{ij} \cdot u_j(t) > \theta_i \\ 0, & \text{iff } \sum_{j=1}^N \omega_{ij} \cdot u_j(t) \leq \theta_i \end{cases} \quad (2.4)$$

($i = 1, 2, \dots, N$)

where the terms "iff" implies if and only if. The terms will be

used hereafter for the simplicity of notation. Now, consider the function $v_i(t+\tau)$ defined as follows.

$$v_i(t+\tau) = 2 u_i(t+\tau) - 1 \quad (2.5)$$

Then, the function $v_i(t+\tau)$ has a value either 1 or -1 as the following manner.

$$v_i(t+\tau) = \begin{cases} 1, & \text{iff } \sum_{j=1}^N w_{ij} \cdot u_j(t) > \theta_i \\ -1, & \text{iff } \sum_{j=1}^N w_{ij} \cdot u_j(t) \leq \theta_i \end{cases} \quad (2.6)$$

(i=1, 2, ..., N)

Hence, the following equations

$$\sum_{j=1}^N w_{ij} \cdot u_j(t) \cdot v_i(t+\tau) \geq \theta_i \cdot v_i(t+\tau) \quad (2.7)$$

(i=1, 2, ..., N)

result from Relation 2.6 by the slight rearrangement. Observe that in Inequality 2.7 the relation $\geq$ holds iff $v_i(t+\tau) = -1$ and the relation $>$ holds iff $v_i(t+\tau) = 1$.

Inequality 2.7 is expressed in terms of the function of the time variable. Now regard Inequality 2.7 as the inequality system with respect to the internal state. For this purpose, let us denote the internal state at time t by $\hat{u}_k$, then the internal state at time $(t+\tau)$ can be known immediately as $f \cdot \hat{u}_k$

by the given transition state diagram. Thus, the following relations are obtained by the definitions of φ_i and $c_{kk}(i)$.

$$v_i(t+\tau) = 2\varphi_i(f \cdot \hat{u}_k) - 1 = c_{kk}(i)$$

and

$$u_j(t) = \varphi_j(\hat{u}_k)$$

(2.8)

Substitution of Relation 2.8 into Inequality 2.7 yields:

$$\sum_{j=1}^N \left\{ \omega_{ij} \cdot \varphi_j(\hat{u}_k) \cdot c_{kk}(i) \right\} - \theta_i \cdot c_{kk}(i) \geq 0 \quad (2.9)$$

$$(i=1, 2, \dots, N)$$

Now, consider the following replacement:

$$\omega_{i0} = -\theta_i \quad (2.10)$$

Furthermore, consider an imaginary threshold element which always remains state 1 regardless of the internal state to assume the following relation.

$$\varphi_0(\hat{u}_k) = 1 \quad (2.11)$$

Note that the function φ_i is used above in an extended meaning from Definition 2.4.

Substitution of Relation 2.10 and 2.11 into Inequality 2.9 yields the following inequality systems

$$\sum_{j=0}^N \omega_{ij} \cdot \varphi_j(\hat{u}_k) \cdot c_{kk}(i) \geq 0 \quad (2.12)$$

$$(k=1, 2, \dots, 2^N), \quad (i=1, 2, \dots, N)$$

With respect to each threshold element there are 2^N inequalities. The set of these 2^N inequalities is referred to as an inequality system inherent to a completely specified Boolean function. Hence, Inequality 2.12 contains N inequality systems.

Inequality Systems 2.12 can be rewritten by using the threshold-weight vector and the content vector as follows.

$$\left. \begin{aligned} c_{kk}(i) \cdot (u_k \cdot {}^t w_i) &> 0, & \text{for } \hat{u}_k \in P_i \\ c_{kk}(i) \cdot (u_k \cdot {}^t w_i) &\geq 0, & \text{for } \hat{u}_k \in Q_i \end{aligned} \right\} \quad (2.13)$$

$$(k=1, 2, \dots, 2^N), \quad (i=1, 2, \dots, N)$$

where the notation $(u_k \cdot {}^t w_i)$ implies the inner product of u_k and w_i .

Let us express Inequality Systems 2.13 in terms of the characteristic matrix $C(i)$ and the universal matrix U as follows.

$$C(i) \cdot U \cdot {}^t w_i \geq 0, \quad (i=1, 2, \dots, N) \quad (2.14)$$

The expression of each Inequality System 2.14 contains all the internal state vectors in the implicit form. Let us pay attention that in each Inequality System 2.14 the relation $\geq$ holds if and only if the internal state is involved in the negative set and the relation $>$ holds if and only if the internal state is involved in the positive set.

Each Inequality System 2.14 contains 2^N inequalities and $(N+1)$ unknown variables $w_{i0}, w_{i1}, \dots, w_{iN}$. Since the relation

$2^N > (N + 1)$ holds in general with exception of $N = 1$, it follows that we are faced to the inequality system where the number of inequalities is larger than the number of unknown variables. The difference between the number of inequalities and unknown variables in each Inequality System 2.14 increases monotonically. Almost every difficulty of our work consists in the fact that the number of inequalities becomes formidable as N increases a little.

By preceding arguments, it is concluded that the treatment of an autonomous threshold network is reduced to the treatment of each threshold element individually, in so far as we are concerned with the realization problem.

Hitherto discussions can be summarized in the following theorem.

Theorem 2.1. If a given transition state diagram "F" is realizable with N threshold elements, then Inequality Systems 2.14 are consistent for every characteristic matrix $C(i)$ ($i=1, 2, \dots, N$) determined completely by Relation 2.3.

Conversely, if a given transition state diagram is not realizable with N threshold elements, then there exists at least one characteristic matrix determined by Relation 2.3 for which Inequality System 2.14 is not consistent.

It is rather complicated to handle Inequality Systems 2.14 directly because the equality sign does not necessarily hold. That is, the equality sign holds only for the case where

the internal state is involved in the negative set. So as to avoid this complication, let us establish the following theorem stated in the absolute inequality form.

Theorem 2.2. The inequality system

$$C(i) \cdot U \cdot {}^t w_i \geq 0 \quad (2.15)$$

is consistent, if and only if the absolute inequality system

$$C(i) \cdot U \cdot {}^t w_i > 0 \quad (2.16)$$

is consistent.

Proof: Sufficiency is obvious. Thus, only necessity will be shown. If Inequality System 2.15 has a solution $\bar{w}_i$, then let us verify that Absolute Inequality System 2.16 always has a solution.

If for the solution $\bar{w}_i$, all the linear forms $C(i) \cdot U \cdot {}^t \bar{w}_i$ have positive values, there is no question. Thus, let us consider the case where some of the linear forms $C(i) \cdot U \cdot {}^t \bar{w}_i$ have zero values. That is, consider the case:

$$c_{jj}(i) \cdot u_j \cdot {}^t \bar{w}_i > 0 \quad , \quad \text{for } \hat{u}_j \in \{P_i \cup Q_i\} \quad (2.17)$$

$$c_{kk}(i) \cdot u_k \cdot {}^t \bar{w}_i = 0 \quad , \quad \text{for } \hat{u}_k \in Q_i \quad (2.18)$$

Since each left side of Inequality 2.17 is a continuous function with respect to $\bar{w}_i$, it is possible to change the vector $\bar{w}_i$ by a minute vector ξ within the range which does not alter the

direction of the inequality sign $>$ of Inequality 2.17. Besides it is possible to make the following relation

$$c_{kk}(i) \cdot u_k^t(\bar{w}_k + \epsilon) > 0$$

hold for the minute vector ϵ satisfying above conditions, since all the coefficients of the linear forms of Equation 2.18 are negative numbers. Q.E.D.

Therefore, Theorem 2.1 in association with Theorem 2.2 leads to the next theorem, which is a basic theorem for our further developments.

Theorem 2.3. If a given transition state diagram F is realizable with N threshold elements, then Absolute Inequality Systems

$$C(i) \cdot U \cdot t_{w_i} > 0, \quad (i=1, 2, \dots, N) \quad (2.19)$$

are consistent for every characteristic matrix $C(i)$ ($i=1, 2, \dots, N$) determined completely by Relation 2.3.

On the contrary, if a given transition state diagram is not realizable with N threshold elements, then there exists at least one characteristic matrix for which Absolute Inequality System 2.19 is not consistent.

2.4 Reformulation in the Equation Form

By theorem 2.3, it is concluded that we can deal with Absolute Inequality Systems 2.19 independently for each threshold element in order to treat an autonomous network. This

means there is no intrinsic difference between an autonomous network problem and a single threshold element problem.

Thus, hereafter let us focus our attention to only one Absolute Inequality System 2.16 with respect to a certain threshold element T_i . Consider the following theorem which is well known in the linear algebra so as to transform the absolute inequalities problem into the equations problem.

Theorem 2.4.⁽¹⁸⁾ Let H and x be an arbitrary (m, n) -type matrix and a n -dimensional column vector. Moreover, let y and a be m -dimensional row vectors. Then the equations

$$H \cdot x = a$$

are consistent if and only if the next equations

$$(y^{(k)} \cdot {}^t a) = 0, \quad (k=1, 2, \dots, m-r)$$

hold for all k . Where r implies the rank of the matrix H and $y^{(k)}$ implies the k -th fundamental independent solution of the adjoint linear equations:

$${}^t H \cdot y = 0$$

Moreover, $(y^{(k)} \cdot {}^t a)$ implies the inner product of the vectors $y^{(k)}$ and a .

Proposition 2.8. The rank of the matrix ${}^t U \cdot C(i)$ is always $(N+1)$, regardless of value i .

Therefore, instead of treating Absolute Inequality System

2.16 itself, consider the following equations.

$$C(i) \cdot U \cdot {}^t w_i = d_i \quad (2.20)$$

where d_i is a 2^N -dimensional column vector such that $d_i > 0$.

Now, let $b_i^{(k)}$ denote the k -th fundamental independent solution of the adjoint linear equations:

$${}^t U \cdot C(i) \cdot {}^t b_i = 0 \quad (2.21)$$

Before proceeding to Theorem 2.5, consider the following proposition.

Proposition 2.9. There exist $(2^N - N - 1)$ fundamental independent solutions for Equation System 2.21.

Then, the next theorem is obtained as the consequence of Theorem 2.4.

Theorem 2.5. Absolute Inequality System 2.16

$$C(i) \cdot U \cdot {}^t w_i > 0 \quad (2.16)$$

is consistent, if and only if the equations

$$(b_i^{(k)} \cdot d_i) = 0, \quad (k=1, 2, \dots, 2^N - N - 1) \quad (2.22)$$

(53)
have a positive solution $d_i > 0$.

Thus, the realization problem of a completely specified Boolean function by a single threshold element is reduced to the problem of whether or not certain linear equations have a positive solution. This material will be handled elaborately

in Chapter IV.

2.5 Fundamental Solution Matrix

In order to develop Theorem 2.5 further, it needs all the fundamental independent solutions of Equation 2.21. The purpose of this section is to introduce the fundamental solution matrix. This matrix will play an important role in Chapter IV.

Proposition 2.10. The $(N+1, N+1)$ -type matrix ${}^tU_I \cdot {}^tC_I(i)$ has rank $(N+1)$.

Theorem 2.6. Consider the $(2^N, 2^{N-N-1})$ -type matrix $\hat{B}(i)$ whose k -th column equals to the k -th fundamental independent solution of Equation 2.21. Then, $\hat{B}(i)$ can be expressed in the following form.

$$\hat{B}(i) = \begin{pmatrix} -C_I(i) \cdot U_I^{-1} \cdot {}^tU_{II} \cdot C_{II}(i) \\ E_{2^{N-N-1}} \end{pmatrix} \quad (2.23)$$

Proof: Each column of $\hat{B}(i)$ surely leads to zero vector, if it is substituted into the linear forms of Equation 2.21. Besides, it is linearly independent, since the identity matrix $E_{2^{N-N-1}}$ is contained as the submatrix of $\hat{B}(i)$. Q.E.D.

Definition 2.10. Let $B(i)$ denote the submatrix that consists of the first $(N+1)$ rows of $\hat{B}(i)$. This matrix $B(i)$ is referred to as the fundamental solution matrix.

Naturally, there exist infinite forms to express all the

fundamental solutions. It is desirable, however, that all the fundamental solutions are given in an easier form for the computation. In fact, it hardly requires complex computations with the exception of matrices' multiplication as shown in Matrix 2.23, since U_I^{-1} is given in Proposition 2.7. Thus, the expression by Matrix 2.23 is one of the most simplified forms to provide all the fundamental solutions.

CHAPTER III

ARGUMENT ON WEIGHTS AND THRESHOLD SPACE

Throughout this chapter, the materials are described as the problems of a single threshold element rather than as the autonomous network problems, because of the reason mentioned in the previous chapter. The results obtained here, however, are readily applicable to an autonomous network by trivial modifications.

In this chapter, the procedure of determining the practical values of weights and threshold which realize the given Boolean function will be presented at first. Furthermore it will be shown in the explicit form what domain these practical values occupy in $(N+1)$ -dimensional Euclidean space. Some knowledges about the Galois field are required for the advancement of our theory. Only the necessary knowledge will be listed up.

3.1 Weights and Threshold Values in R^{N+1}

By preceding arguments, it is concluded that we can testify by the following procedures whether or not a given transition state diagram is realizable.

- (1) For a given transition state diagram F , find each characteristic matrix $C(i)$ determined completely by Relation 2.3.
- (2) For each obtained characteristic matrix $C(i)$, produce

fundamental independent solutions $b_i^{(k)}$ ($k=1, 2, \dots, 2^{N-N-1}$) by Matrix 2.23.

- (3) For these fundamental independent solutions, ascertain whether or not Equation System 2.22 has a positive solution $d_i > 0$.
- (4) If for each i ($i=1, 2, \dots, N$), Equation System 2.22 has a positive solution, then the given transition state diagram is realizable by the autonomous network constructed with N threshold elements. Otherwise it is not realizable with N threshold elements.

Now, suppose that Equation System 2.22 is turned out to have a positive solution through the test stated in the subsequent chapter. Thus, suppose that the Boolean function represented by the characteristic matrix $C(i)$ can be realized by a single threshold element. Then, let us provide the procedure how to find the practical values of weights and threshold to realize the given Boolean function. Let $\overleftarrow{d}_1$ denote the first $(N+1)$ components of d_1 , then the following theorem is established.

Theorem 3.1. If a given Boolean function represented by $C(i)$ is realizable with a single threshold element, then the practical values of weights and threshold are given by the following vector w_1 .

$$w_1 = U_I^{-1} \cdot C_I(i) \cdot \overleftarrow{d}_1 \quad (3.1)$$

Proof: Substitution of w_1 given by Relation 3.1 into the linear forms of Absolute Inequality System 2.16 yields the vector d_1 . This is immediately followed by Theorem 2.4. Q.E.D.

The expression of d_1 generally contains $(N+1)$ parameters, since the number of Equations 2.22 is $(2^N - N - 1)$ and the number of unknown variables is 2^N . If d_1 is fixed to a certain positive vector by affording appropriate concrete real numbers into these parameters, the threshold-weight vector w_1 that realizes the given Boolean function can be obtained by Relation 3.1 for this fixed positive vector.

By this way, the practical values of weights and threshold can be easily obtained by a simple and unified computation, only if the positive solution of Equation 2.22 is found for the characteristic matrix $C(1)$.

3.2 Domain of Weights and Threshold ⁽⁴⁶⁾

In this section, the concrete correspondence between a given threshold function and the domain of $(N+1)$ -dimensional Euclidean space will be presented. Therefore, if the given transition state diagram can be realized with N threshold elements, then the actual threshold and weights values are given as an arbitrary element contained in the direct sum of such obtained N domains. To begin with, let us introduce two terms, a convex cone and a polyhedral convex cone, by Definition 3.1 and Definition 3.2, respectively.

These terms themselves are well known in the linear algebra but used here to represent a little different concepts from the familiar ones so as to handle the absolute inequalities rather than the inequalities where the relations hold always in the form $\geq$.

Definition 3.1. Let L be a subset of n -dimensional Euclidean space. Then L is called a convex cone, if the arbitrary linear combination with positive coefficients of any two elements of L is involved in L . That is, L is a convex cone if and only if

$$\alpha y_1 + \beta y_2 \in L \quad \text{holds for } \forall y_1, \forall y_2 \in L, \quad \forall \alpha > 0, \forall \beta > 0$$

Proposition 3.1. Let H and y denote an arbitrary (m, n) -type matrix and an n -dimensional column vector, respectively. Then the whole solutions R of the absolute inequalities

$$H \cdot y > 0$$

constitute a convex cone.

Proof: Since $H \cdot y_1 > 0$ and $H \cdot y_2 > 0$ hold for any two elements y_1, y_2 contained in R , the following relation is obtained.

$$H \cdot (\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 H \cdot y_1 + \alpha_2 H \cdot y_2 > 0$$

where $\alpha_1 > 0, \alpha_2 > 0$. Hence, $\alpha_1 y_1 + \alpha_2 y_2$ is an element of R .

Q.E.D.

Consequently, the whole solutions of any absolute inequalities can be given as a convex cone. This result itself is not of so much significance. However, it will be seen moreover that the whole solutions of any absolute inequalities constitute a polyhedral convex cone which has a more strongly defined structure. For this purpose, let us consider the following definition.

Definition 3.2. Let H and v denote an (m,n) -type matrix and an n -dimensional column vector respectively. Then the set of the whole vectors y which can be expressed in the following form:

$$y = H \cdot v, \quad v > 0$$

is called a polyhedral convex cone generated from the matrix H . This polyhedral convex cone is denoted by $K(H)$. That is:

$$K(H) = \left\{ y ; y = H \cdot v, v > 0 \right\}$$

Theorem 3.2. Any subspace L of the Euclidean space is a polyhedral convex cone.

Proof: If the dimension of a subspace L is k , then L equals to the set which consists of the whole vectors constructed from the linear combination of k linearly independent vectors, $e_1, e_2, \dots, e_k$, given by Definition 2.9.

Therefore, consider the matrix H with $2k$ columns such that

$$H = (e_1, e_2, \dots, e_k, -e_1, -e_2, \dots, -e_k)$$

Then, L equals to the polyhedral convex cone $K(H)$. Q.E.D.

Theorem 3.3. The set L^+ which consists of the whole positive vectors involved in a subspace L ;

$$L^+ = \left\{ y ; y > 0 , y \in L \right\}$$

is a polyhedral convex cone.

Proof: If the dimension of a subspace L is k , then L is the polyhedral convex cone generated from the identity matrix E_k . Namely $L^+ = K(E_k)$. Q.E.D.

Proposition 3.2. Let H and v be an arbitrary matrix and a column vector, respectively. Then the whole vectors that can be expressed in the form $H \cdot v$ make a subspace L .

Proof: Suppose that y_1 and y_2 can be expressed by $H \cdot v_1$ and $H \cdot v_2$, respectively. Then, the element expressed by $\alpha_1 y_1 + \alpha_2 y_2$ is involved in L , since $\alpha_1 y_1 + \alpha_2 y_2 = H \cdot (\alpha_1 v_1 + \alpha_2 v_2)$. Q.E.D.

Theorem 3.4. Let H be an (n, n) -type square matrix whose rank is n . Then, the whole solutions R of the absolute inequalities

$$H \cdot v > 0$$

can be expressed as the polyhedral convex cone $K(A)$.

Here, A is the (n,n) -type matrix.

Proof: Consider the subspace L that consists of the whole vectors expressed in the form $H \cdot v$. Then it follows from Theorem 3.3 that L^+ defined by the relation

$$L^+ = \{ y ; y = H \cdot v , y > 0 \} \quad (3.2)$$

is a polyhedral convex cone. Hence the relation

$$L^+ = K(E_n)$$

holds. Furthermore consider the vectors $q_1, q_2, \dots, q_n$ such that

$$e_j = H \cdot q_j \quad (j = 1, 2, \dots, n)$$

where e_j is the j -th unit vector. Note that each q_j is defined uniquely since $\det(H) \neq 0$ by the hypothesis.

Now suppose that the matrix A stated in this theorem has these vectors $q_1, q_2, \dots, q_n$ as its n columns. This is shown as follows. Let v be an arbitrary element involved in $K(A)$. Then, v can be expressed in the following form.

$$v = \sum_{j=1}^n \alpha_j \cdot q_j , \quad \alpha_j > 0 \quad (j = 1, 2, \dots, n)$$

Therefore the relation

$$H \cdot v = \sum_{j=1}^n \alpha_j \cdot H \cdot q_j = \sum_{j=1}^n \alpha_j \cdot e_j > 0$$

holds. Consequently the relation

$$K(A) \subseteq R$$

is obtained.

Conversely, let v be an arbitrary element involved in the set R of the whole solutions. Then the vector $H \cdot v$ is always contained in L^+ given by Relation 3.2. Thus, the vector $H \cdot v$ can be expressed in the following form.

$$H \cdot v = \sum_{j=1}^n \alpha_j \cdot e_j = \sum_{j=1}^n \alpha_j \cdot H \cdot q_j$$

where $\alpha_j > 0$ ($j = 1, 2, \dots, n$). This relation leads to the equation.

$$H \cdot (v - \sum_{j=1}^n \alpha_j \cdot q_j) = 0 \quad (3.3)$$

Consequently, since $\det(H) \neq 0$, the following equation

$$v - \sum_{j=1}^n \alpha_j \cdot q_j = 0$$

results from Relation 3.3. Thus, the relation

$$R \supseteq K(A)$$

is obtained.

Q.E.D.

The following theorem is well known in the linear algebra, which will be used for further developments. For the sake of

unity, let us mention it with a slight modification.

(38)

Theorem 3.5. Assume it is previously observed that there exist n linearly independent vectors in m vectors set $v_1, v_2, \dots, v_m$, however there never exist $(n+1)$ linearly independent vectors in this set even if we choose appropriately. Then, any set composed of k (where k is less than n) linearly independent vectors which are selected from $v_1, v_2, \dots, v_m$ can be expanded to contain n linearly independent vectors by adding adequately $(n-k)$ vectors in $v_1, v_2, \dots, v_m$.

Theorem 3.6. The whole solutions $R(i)$ of Absolute Inequality System 2.16

$$C(i) \cdot U \cdot {}^t w_i > 0 \quad (2.16)$$

can be given in the following form.

$$R(i) = \bigcap_{j=1}^s K(A_{ij}) \quad (3.4)$$

Namely, $R(i)$ is an intersection of finite number of polyhedral convex cones $K(A_{i1}), K(A_{i2}), \dots, K(A_{is})$. Where the bounds for the value " s " will be given in Theorem 3.14.

Proof: Let $\{W_i\}$ denote the set which consists of all the 2^N row vectors of the matrix $C(i) \cdot U$. Divide the set $\{W_i\}$ into " s " subsets $\{W_{i1}\}, \{W_{i2}\}, \dots, \{W_{is}\}$ so that the following conditions may be satisfied:

$$\{W_i\} = \bigcup_{j=1}^s \{W_{ij}\}$$

and

(3.5)

$$\det(W_{ij}) \neq 0, \quad (j=1, 2, \dots, s)$$

where each subset $\{W_{ij}\}$ contains $(N+1)$ elements and W_{ij} denotes the matrix whose each row is one of these $(N+1)$ elements. Observe that the relation

$$\{W_{ij}\} \cap \{W_{ik}\} = \Phi, \quad (j, k = 1, 2, \dots, s; j \neq k)$$

needs not necessarily hold. Theorem 3.5. guarantees that such a division exists.

Now, adopt such A_{ij} as is given by the relation;

$$A_{ij} = W_{ij}^{-1} \quad (j = 1, 2, \dots, s)$$

Then, the relation

$$R(i) = \bigcap_{j=1}^s K(A_{ij})$$

follows immediately from Theorem 3.4.

Q.E.D.

The next theorem is almost obvious from the above proof.

Theorem 3.7. The set $R(i)$ does not depend on how to divide the matrix $C(i) \cdot U$. Namely two distinct divisions, satisfying Relation 3.5, make the two intersections which represent the same set.

By these discussions, it is concluded that the whole solutions of Absolute Inequality System 2.16 can be given by the intersection of finite number of polyhedral convex cones. However, if any given Boolean function is a threshold function, we must treat Inequality System 2.15 itself inherent to the given Boolean function rather than Absolute Inequality System 2.16 so as to argue rigorously the whole practical solutions in terms of the weights and the threshold.

For this purpose, it only needs a minor modification. That is, according to Definition 3.2, a polyhedral convex cone equals to the set which consists of the whole linear combinations of the column vectors of a matrix H with positive coefficients. Now in order to argue about Inequality System 2.15, adopt nonnegative coefficients v_j for every j such that $\hat{u}_j \in Q_1$ instead of positive coefficients. In the actual situation, however, it hardly brings important effects with exception of whether or not the boundary of a polyhedral convex cone is contained. Therefore it will be admitted to restrict ourselves to the solutions of Absolute Inequality System 2.16 so as to avoid complications.

Thus, Absolute Inequality System 2.16 will be treated throughout the remainder instead of Inequality System 2.15. Then, if any given transition state diagram can be implemented by N threshold elements, the whole weights and threshold values realizing this behavior are given as a direct sum of such obtained N intersections

$$\bigcap_{j=1}^{s_1} K(A_{1j}) \oplus \bigcap_{j=1}^{s_2} K(A_{2j}) \oplus \dots \oplus \bigcap_{j=1}^{s_N} K(A_{Nj})$$

which is a subset of (N^2+N) -dimensional Euclidean space.

(38)

3.3 Introduction to Galois Field

To obtain the bounds for the number of divisions of the matrix $C(i) \cdot U$, some knowledges about Galois fields are used. The minimum background for the subsequent discussions is presented here for the sake of unity and continuity.

In the theory of rings, ideals play very important roles just like normal subgroups in groups. An ideal I is a subset of elements of a ring R with the following two properties.

- (1) I is a subgroup of the additive group of R .
- (2) For any element " a " of I and any element " r " of R , $a \cdot r$ and $r \cdot a$ are contained in I .

Since an ideal is a subgroup, cosets can be formed. In this case the cosets are called residue classes. Then, the residue classes of a ring with respect to an ideal form a ring. This ring is called the residue class ring.

Theorem 3.8. The residue class ring modulo p is a field if and only if p is a prime number. This field is called the Galois field of p elements, denoted by $GF(p)$.

In so far as we are concerned with the integer, the field which has p^n ($n > 1$) elements can not be generated. Now let us

consider the polynomial $f(X)$ of degree n with one variable X and with the coefficients from any field A :

$$f(X) = a_0 + a_1X + a_2X^2 + \dots + a_nX^n$$

Polynomials can be added and multiplied in the ordinary way, and thus they form a ring. A set of polynomials is an ideal if and only if it consists of all the multiples of a certain polynomial $f(X)$. This ideal is denoted by $\{f(X)\}$. The residue class ring formed by such an ideal $\{f(X)\}$ is called the ring of polynomials modulo $f(X)$.

Then, every residue class modulo a polynomial $f(X)$ of degree n contains either 0 or a polynomial of degree less than n . Zero is an element of the ideal, and every polynomial of degree less than n is in a distinct residue class.

Some symbol, usually α , is used to represent the residue class containing X . The residue class containing a field element is given the same name as the field element. This, however, should not cause any confusion. Then the coset that contains

$$a_0 + a_1X + \dots + a_{n-1}X^{n-1}$$

can be denoted by

$$a_0 + a_1\alpha + \dots + a_{n-1}\alpha^{n-1}$$

Thus every residue class can be represented by a polynomial of degree less than n with respect to α . Consequently the number of distinct residue classes is m^n where m denotes the number

of elements contained in a field A .

Theorem 3.9. In the residue class ring modulo a polynomial $f(X)$ of degree n , the relation $f(\alpha) = \{f(X)\} = \{0\}$ always holds. But no polynomial in α of degree less than n is the ideal $\{f(X)\}$ or $\{0\}$.

Theorem 3.10. Let $P(X)$ be a polynomial of degree n with the coefficients from a field A . If $P(X)$ is irreducible in A , that is, if $P(X)$ has no factors with the coefficients of A , then the residue class ring modulo $P(X)$ becomes a field.

It is shown that the ring of polynomials over any finite field has at least one irreducible polynomial of every degree. The field of polynomials over $GF(P)$ modulo an irreducible polynomial of degree n is called the Galois field of p^n elements, denoted by $GF(p^n)$. It is verified that every finite field is isomorphic to some Galois field. Furthermore it is also verified that any two finite fields with the same number of elements are isomorphic to each other. Hence, for any given positive integer β there exists a field $GF(\beta)$, if and only if β is a power of a prime number.

There is a natural correspondence between $GF(p^n)$ and the n -tuples composed from the elements of $GF(P)$. That is, the polynomial $a_0 + a_1\alpha + \dots + a_{n-1}\alpha^{n-1}$ corresponds to the n -tuple $(a_0, a_1, \dots, a_{n-1})$. The sum of two n -tuples corresponds to the sum of the two corresponding polynomials, and the multiplication by scalars carries over similarly. Thus

$GF(P^n)$ can be considered as a vector space of the dimension n over $GF(P)$. Consequently the set ∇ which consists of all the internal state vectors $\hat{u}_j$ ($j = 1, 2, \dots, 2^N$) can be regarded as $GF(2^N)$.

(48)
Theorem 3.11. In $GF(2^N)$ there is a primitive element α which is an element of order $(2^N - 1)$. Thus every nonzero element of $GF(2^N)$ can be expressed as a power of α . Therefore, the multiplicative group of $GF(2^N)$ is cyclic.

Example 3.1. The Galois field of 2^5 elements can be formed as the field of polynomials over $GF(2)$ modulo $X^5 + X^2 + 1$. Let α denote the residue class that contains the polynomial X . Then α is a root of $(X^5 + X^2 + 1 = 0)$, and this α happens to be a primitive element of the field. Then the 31 nonzero field elements of $GF(2^5)$ are given by Table 3.1.

α^0	$= 1$	$=$	$(1\ 0\ 0\ 0\ 0)$
α^1	$= \alpha$	$=$	$(0\ 1\ 0\ 0\ 0)$
α^2	$= \alpha^2$	$=$	$(0\ 0\ 1\ 0\ 0)$
α^3	$= \alpha^3$	$=$	$(0\ 0\ 0\ 1\ 0)$
α^4	$= \alpha^4$	$=$	$(0\ 0\ 0\ 0\ 1)$
α^5	$= 1 + \alpha^2$	$=$	$(1\ 0\ 1\ 0\ 0)$
α^6	$= \alpha + \alpha^3$	$=$	$(0\ 1\ 0\ 1\ 0)$
α^7	$= \alpha^2 + \alpha^4$	$=$	$(0\ 0\ 1\ 0\ 1)$
α^8	$= 1 + \alpha^2 + \alpha^3$	$=$	$(1\ 0\ 1\ 1\ 0)$
α^9	$= \alpha + \alpha^3 + \alpha^4$	$=$	$(0\ 1\ 0\ 1\ 1)$

α^{10}	$= 1$	$+ \alpha^4$	$= (1\ 0\ 0\ 0\ 1)$
α^{11}	$= 1 + \alpha + \alpha^2$		$= (1\ 1\ 1\ 0\ 0)$
α^{12}	$= \alpha + \alpha^2 + \alpha^3$		$= (0\ 1\ 1\ 1\ 0)$
α^{13}	$= \alpha^2 + \alpha^3 + \alpha^4$		$= (0\ 0\ 1\ 1\ 1)$
α^{14}	$= 1 + \alpha^2 + \alpha^3 + \alpha^4$		$= (1\ 0\ 1\ 1\ 1)$
α^{15}	$= 1 + \alpha + \alpha^2 + \alpha^3 + \alpha^4$		$= (1\ 1\ 1\ 1\ 1)$
α^{16}	$= 1 + \alpha + \alpha^3 + \alpha^4$		$= (1\ 1\ 0\ 1\ 1)$
α^{17}	$= 1 + \alpha + \alpha^4$		$= (1\ 1\ 0\ 0\ 1)$
α^{18}	$= 1 + \alpha$		$= (1\ 1\ 0\ 0\ 0)$
α^{19}	$= \alpha + \alpha^2$		$= (0\ 1\ 1\ 0\ 0)$
α^{20}	$= \alpha^2 + \alpha^3$		$= (0\ 0\ 1\ 1\ 0)$
α^{21}	$= \alpha^3 + \alpha^4$		$= (0\ 0\ 0\ 1\ 1)$
α^{22}	$= 1 + \alpha^2 + \alpha^4$		$= (1\ 0\ 1\ 0\ 1)$
α^{23}	$= 1 + \alpha + \alpha^2 + \alpha^3$		$= (1\ 1\ 1\ 1\ 0)$
α^{24}	$= \alpha + \alpha^2 + \alpha^3 + \alpha^4$		$= (0\ 1\ 1\ 1\ 1)$
α^{25}	$= 1 + \alpha^3 + \alpha^4$		$= (1\ 0\ 0\ 1\ 1)$
α^{26}	$= 1 + \alpha + \alpha^2 + \alpha^4$		$= (1\ 1\ 1\ 0\ 1)$
α^{27}	$= 1 + \alpha + \alpha^3$		$= (1\ 1\ 0\ 1\ 0)$
α^{28}	$= \alpha + \alpha^2 + \alpha^4$		$= (0\ 1\ 1\ 0\ 1)$
α^{29}	$= 1 + \alpha^3$		$= (1\ 0\ 0\ 1\ 0)$
α^{30}	$= \alpha + \alpha^4$		$= (0\ 1\ 0\ 0\ 1)$

Table 3.1. Ordered Representation of $GF(2^5)$

An irreducible polynomial of degree n over $GF(P)$ is called primitive if it has a primitive element of $GF(P^n)$ as a root. There exist the irreducible polynomials which are not necessarily primitive. Thus, the actual procedures to examine whether or not a polynomial is primitive are one of subjects to be studied. Moreover it is significant to find out a primitive polynomial for any given $GF(P^n)$.

With regard to such problems some studies have been done by various investigators. For the practical use they gave the primitive polynomials for $GF(2^n)$ where n is up to 34. Thus we can easily produce such an ordered representation of $GF(2^n)$ for any given positive integer n less than 34 as Table 3.1.

3.4. Bounds for the Number of Divisions

The bounds for the number of divisions of the matrix $C(i) \cdot U$ still remain unknown. The object of this section is to provide this bounds by utilizing the above stated preliminary knowledges about the Galois field. To begin with let us consider the following theorem which is almost evident from the definition of the linearly independence.

Theorem 3.12. If the given N vectors $v_1, v_2, \dots, v_N$ are linearly independent over $GF(2)$, then they are also linearly independent over the real number space.

Theorem 3.13. In the ordered representation of $GF(2^N)$, any successive N elements of $GF(2^N)$ are linearly independent.

Proof: Any successive N elements in the ordered representation of $GF(2^N)$ can be given as

$$\alpha^m, \alpha^{m+1}, \dots, \alpha^{m+N-1}$$

where α denotes the primitive element of $GF(2^N)$ and m denotes the integer such that $1 \leq m \leq 2^N - N$.

Therefore the linear combination $L(\alpha)$ of these N successive elements is expressed in the following form.

$$\begin{aligned} L(\alpha) &= a_0 \alpha^m + a_1 \alpha^{m+1} + \dots + a_{N-1} \alpha^{m+N-1} \\ &= \alpha^m (a_0 + a_1 \alpha + \dots + a_{N-1} \alpha^{N-1}) \end{aligned}$$

where $a_i (i=0, 1, \dots, N-1)$ is an element of $GF(2)$.

On the other hand, since α is a primitive element, α must be a root of an irreducible polynomial of degree N . Thus the following equation holds through Theorem 3.9.

$$a_0 + a_1 \alpha + \dots + a_{N-1} \alpha^{N-1} \neq \{0\}$$

Hence, $L(\alpha)$ does not equal to $\{0\}$ except for

$$a_0 = a_1 = \dots = a_{N-1} = 0$$

This implies $\alpha^m, \alpha^{m+1}, \dots, \alpha^{m+N-1}$ are linearly independent over $GF(2)$. Consequently any successive N elements in the ordered representation of $GF(2^N)$ are linearly independent over the real number space. Q.E.D.

In fact, the validity of Theorem 3.13 will be assured by

checking Table 3.1.

Theorem 3.14. When the whole solutions $R(i)$ of Absolute Inequality System 2.16 are expressed in the following form,

$$R(i) = \bigcap_{j=1}^s K(A_{ij})$$

the bounds for the number s are given by the next relation:

$$\left\lceil \frac{2^N}{N+1} \right\rceil \leq s \leq \left\lfloor \frac{2^N}{N} \right\rfloor$$

where the notation $\{x\}$ means the smallest integer which does not exceed the value x .

Proof: The number of divisions of the matrix $C(i) \cdot U$ is same as that of the matrix U , since $C(i)$ is a diagonal matrix and every diagonal component is not zero. Hence, it is sufficient to consider the matrix U .

The lower bound is obvious since each W_{ij} (defined in the proof of Theorem 3.6) is an $(N+1, N+1)$ -type square matrix.

As for the upper bound, divide all the internal state vectors into $\left\lceil 2^N/N \right\rceil$ sets by taking the internal state vectors N by N from top to bottom in the ordered representation of $GF(2^N)$. Permit, however, an overlapping extraction only for the last division unless 2^N is a multiple of N . Namely, the last division set may be constituted from the last N internal state vectors in the ordered representation of $GF(2^N)$.

Furthermore, add the vector $\hat{u}_1 = (0, 0, \dots, 0)$ to each division set obtained by the above stated process. Now replace the internal state vectors involved in each resulted division set by the corresponding content vectors. Then each matrix generated from the above mentioned $(N+1)$ content vectors has the rank $(N+1)$ through Theorem 3.13.

By such ways $\left[2^N/N\right]$ matrices, all of which have the rank $(N+1)$, can be produced. Therefore such $\left[2^N/N\right]$ matrices can be expressed as follows.

This completes the part of
the upper bound. Q.E.D.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & & \alpha^0 & & & \\ 1 & & \alpha^1 & & & \\ \dots & & \dots & & & \\ 1 & & \alpha^{N-1} & & & \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & & \alpha^N & & & \\ 1 & & \alpha^{N+1} & & & \\ \dots & & \dots & & & \\ 1 & & \alpha^{2N-2} & & & \end{pmatrix}$$

.

.

.

.

.

.

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & & \alpha^{2N-N-1} & & & \\ 1 & & \alpha^{2N-N} & & & \\ \dots & & \dots & & & \\ 1 & & \alpha^{2N-2} & & & \end{pmatrix}$$

Thus it is concluded that all the threshold and weights values realizing a threshold function represented by $C(i)$ can be given completely by Theorem 3.14. It is the well known fact that the whole solutions of the linear inequalities are generally expressed as a polyhedral convex cone. In this chapter, the practical procedures which determine the polyhedral convex cone have been provided. The terms of the polyhedral convex cone is modified here so as to handle the absolute inequalities.

Example 3.2. The submatrices $\bar{U}_j$ ($j = 1, 2, \dots, 7$) of the universal matrix U for $N = 5$, defined by the relation $A_{ij} = W_{ij}^{-1} = [\bar{C}_j(i) \cdot \bar{U}_j]^{-1}$ where $\bar{C}_j(i)$ denotes the submatrix of $C(i)$ corresponding to a division set expressed by $\bar{U}_j$, are given by Table 3.2. Thus $\bar{C}_j(i)$ is the (6,6)-type diagonal matrix which consists of the k 's-th rows and columns of $C(i)$, if $\bar{U}_j$ contains the k 's-th rows of U .

Therefore the polyhedral convex cone for the whole solutions of a threshold function of 5 variables can be constituted by using these matrices $\bar{U}_j$ in the form :

$$R(i) = \bigcap_{j=1}^7 K(\bar{U}_j^{-1} \cdot \bar{C}_j(i))$$

$$\bar{\bar{u}}_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\bar{\bar{u}}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

$$\bar{\bar{u}}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

$$\bar{\bar{u}}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

$$\bar{\bar{u}}_5 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\bar{\bar{u}}_6 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\bar{\bar{u}}_7 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

Table 3.2. The Basic Component Matrices for the Polyhedral Convex Cone of a Threshold Function of 5 Variables.

CHAPTER IV

LINEAR SEPARABILITY BY FUNDAMENTAL SOLUTION MATRIX

In this chapter the problem, generally known as the linear separability, is treated. Concerning this problem various approaches⁽¹⁾, practical procedures and conditions have been proposed such as Linear Programming⁽⁴⁵⁾, Function Tree⁽⁴⁰⁾ and Asummability. This problem is one of the most classical and yet the most difficult ones in Threshold Logic as the number of input variables increases a little. Here the subject is discussed with using the fundamental solution matrix. The linear separability condition is established with respect to the properties of this matrix.

4.1 The Basic Background

By the preceding argument, it is concluded that a given Boolean function represented by $C(i)$ is realizable by a single threshold element if and only if Equation System 2.22 has a positive solution d_i . Here the practical procedures which examine whether or not Equation System 2.22 has a positive solution are presented. For this purpose let us first consider the next theorem.

⁽¹⁸⁾
Theorem 4.1. Equation System 2.22 has a positive solution d_i if and only if the minimum convex body which involves

the 2^N points of the s -dimensional space ($s=2^N-N-1$) whose coordinates are given by

$$P_1 = \begin{pmatrix} (1) & (2) & \dots & (s) \\ b_{11} & b_{11} & \dots & b_{11} \end{pmatrix}$$

$$P_2 = \begin{pmatrix} (1) & (2) & & (s) \\ b_{12} & b_{12} & \dots & b_{12} \end{pmatrix}$$

• • • • •

$$P_{2N} = (b_{12N}^{(1)}, b_{12N}^{(2)}, \dots, b_{12N}^{(s)})$$

contains the origin. Here the k-th fundamental solution is expressed by the coordinate

$$\mathbf{b}_i^{(k)} = (b_{i1}^{(k)}, b_{i2}^{(k)}, \dots, b_{i2N}^{(k)})$$

However it is often difficult to analyse the Equation System 2.22 by the aid of Theorem 4.1. Now consider the next theorem established by Furtwängler and M. Fujiwara which provides the foundation of the subsequent argument.

(17)

Furtwängler's Theorem There exists a system of positive numbers $y_1, y_2, \dots, y_r$ such that the relations

$$M_k(y) = \sum_{j=1}^r \alpha_{kj} \cdot y_j < 0, \quad (k=1, 2, \dots, s) \quad (4.1)$$

hold, if and only if at least one of r linear forms

$$L_j(x) = \sum_{k=1}^s \alpha_{kj} \cdot x_k, \quad (j = 1, 2, \dots, r) \quad (4.2)$$

becomes a negative number for any nonzero nonnegative vector $x = (x_1, x_2, \dots, x_s) \geq 0$.

Since the linear forms of Relation 4.1 take negative values, let us represent these values by $y_{r+1}, y_{r+2}, \dots, y_{r+s}$. Then the next theorem is obtained immediately from Furtwängler's Theorem.

Theorem 4.2. The linear equations

$$\sum_{j=1}^r \alpha_{kj} \cdot y_j = -y_{r+k}, \quad (k=1, 2, \dots, s) \quad (4.3)$$

have a positive solution $(y_1, y_2, \dots, y_{r+s})$, if and only if at least one of r linear forms given in Relation 4.2 becomes a negative number for any nonzero nonnegative vector $x = (x_1, x_2, \dots, x_s) \geq 0$.

Therefore the fact that at least one of r linear forms given in Relation 4.2 becomes negative for any $x \geq 0$ is equivalent to the fact that the origin lies within the minimum convex body containing the following $s+r$ points in the s -dimensional space:

$$\begin{aligned} P_1 &= (\alpha_{11}, \alpha_{21}, \dots, \alpha_{s1}) \\ P_2 &= (\alpha_{12}, \alpha_{22}, \dots, \alpha_{s2}) \\ &\vdots \\ P_r &= (\alpha_{1r}, \alpha_{2r}, \dots, \alpha_{sr}) \\ P_{r+1} &= (1, 0, \dots, 0) \\ P_{r+2} &= (0, 1, \dots, 0) \\ &\vdots \\ P_{r+s} &= (0, 0, \dots, 1) \end{aligned}$$

Then the next theorem is established which is the basic theorem for the further development.

Theorem 4.3. Equation System 2.22 has a positive solution $d_i > 0$ if and only if one of $N+1$ linear forms

$$L_{1j}(x) = \sum_{k=1}^{2^N - N - 1} b_{1j}^{(k)} \cdot x_k, \quad (j = 1, 2, \dots, N+1) \quad (4.4)$$

becomes a negative number for any nonzero nonnegative vector $x = (x_1, x_2, \dots, x_{2^{N-N-1}}) \geq 0$.

Proof: The coefficients of Equation System 2.22 have the same forms as the coefficients of Equations 4.3. That is, the 2^N points stated in Theorem 4.1 have the following coordinates.

$$P_1 = (\overset{(1)}{b_{11}}, \overset{(2)}{b_{11}}, \dots, \overset{(s)}{b_{11}})$$

$$p_2 = (b_{12}^{(1)}, b_{12}^{(2)}, \dots, b_{12}^{(s)})$$

$$P_{N+1} = (b_{1N+1}^{(1)}, b_{1N+1}^{(2)}, \dots, b_{1N+1}^{(s)})$$

$$P_{N+2} = (\quad 1, \quad 0, \dots, \quad 0)$$

$$P_{N+3} = (0, 1, \dots, 0)$$

$$P_{2N} = (0, 0, \dots, 1)$$

Q. E. D.

The main purpose of this section is to develop Theorem 4.3 in order to make it more tractable. The conditions for the existence of a positive solution d_1 of Equation System 2.22 are stated as the properties of the fundamental solution matrix $B(i)$ defined by Definition 2.10.

- 60 -

Theorem 4.4. If there exists at least one nonnegative column in the fundamental solution matrix $B(i)$, then Equation System 2.22 has no positive solution.

Proof: Suppose that the k -th column of $B(i)$ is nonnegative. Consider a vector x such that only the k -th component is a sufficiently large number and all the other components are zero, then the linear form $L_{ij}(x)$ of Relation 4.4 does not become a negative number for any j . Q.E.D.

Theorem 4.5. If there exists at least one negative row in the fundamental solution matrix $B(i)$, then Equation System 2.22 has a positive solution.

Proof: Suppose that the k -th row of $B(i)$ is negative, then the linear form $L_{ik}(x)$ of Relation 4.4 becomes a negative number. Q.E.D.

Since Theorem 4.4 and Theorem 4.5 are not so applicable excepting the limited fundamental solution matrix, let us develop Theorem 4.3 in the easier form to apply. For this purpose consider the following two definitions.

Definition 4.1. If the l_1 -th row of $B(i)$ is nonzero nonpositive, then the submatrix $\bar{B}(i:l_1)$ of $B(i)$ is constructed in the following manner. That is, exclude all the columns from $B(i)$ such that the components of the l_1 -th row are negative. $\bar{B}(i:l_1)$ is called the reduced fundamental solution matrix of order l concerning the l_1 -th row. Similarly if the l_2 -th row

of $\overleftarrow{B}(i:l_1)$ is nonzero nonpositive, then the submatrix $\overleftarrow{B}(i:l_1 l_2)$ of $B(i)$ is constructed in the same manner. That is, exclude all the columns from $\overleftarrow{B}(i:l_1)$ such that the components of the l_2 -th row are negative. $\overleftarrow{B}(i:l_1 l_2)$ is called the reduced fundamental solution matrix of order 2 concerning the l_1 -th row and the l_2 -th row. Likewise the reduced fundamental solution matrix $\overleftarrow{B}(i:l_1 l_2 \dots l_n)$ of order n can be defined by the similar manner.

Definition 4.2. Consider the $(N+1, 2^{N-N-1})$ -type matrix $B(i:\text{row})$ which is constructed from $B(i)$ in the following manner. That is, first make a linear combination of some rows of $B(i)$ with positive coefficients, and then replace an arbitrary row which was joined in the linear combination with this linear combined row. $B(i:\text{row})$ is called the row-combined fundamental solution matrix. Similarly, the column-combined fundamental solution matrix $B(i:\text{column})$ can be defined.

Theorem 4.6. If $\overleftarrow{B}(i:l_1 l_2 \dots l_n)$ becomes the $(0, 0)$ -type matrix for an appropriate number n , namely, vanished, then Equation System 2.22 has a positive solution.

Proof: Let us represent the l_1 -th row, l_2 -th row, ..., and l_n -th row of $B(i)$ by $b_{i l_1}$, $b_{i l_2}$, ..., and $b_{i l_n}$, respectively. Then the linear combination

$$b_i = \sum_{j=1}^n \alpha_j \cdot b_{i l_j}, \quad (\alpha_1 \gg \alpha_2 \gg \dots \gg \alpha_n > 0)$$

becomes a negative vector. Hence the linear form of Relation 4.4 which takes the components of this negative vector b_i as its coefficients becomes

$$\sum_{k=1}^{2^N - N - 1} b_{ik} \cdot x_k = \sum_{j=1}^n \alpha_j \cdot L_{ij}(x) < 0$$

where $b_i = (b_{i1}, b_{i2}, \dots, b_{i2^N - N - 1})$. Therefore at least one of $L_{ij}(x)$ ($j=1, 2, \dots, n$) becomes a negative number. Q.E.D.

The case where $\overleftarrow{B}(i:1_1)$ becomes a $(0, 0)$ -type matrix corresponds to Theorem 4.5.

Theorem 4.7. If it is possible to make a column-combined fundamental solution matrix $B(i:\text{column})$ whose at least one of columns becomes nonnegative, then Equation System 2.22 has no positive solution.

Proof: Let $E(\text{column})$ denote the matrix which is constructed from the identity matrix $E_{2^N - N - 1}$ by performing the same linear combination and the same replacement as $B(i:\text{column})$. Then each column of the matrix

$$\begin{pmatrix} B(i:\text{column}) \\ E(\text{column}) \end{pmatrix}$$

becomes the fundamental solution of Equation 2.21. Q.E.D.

In Definition 4.1 the reduced fundamental solution matrix $\overleftarrow{B}(i:1_1 1_2 \dots 1_n)$ of order n is constructed for $B(i)$. Likewise, the reduced row-combined fundamental solution matrix

$\overleftarrow{B}(i:\text{row}:1_1 1_2 \dots 1_n)$ of order n can be defined for $B(i:\text{row})$ by the similar manner. Then, the next theorem is established.

Theorem 4.8. If $\overleftarrow{B}(i:\text{row}:1_1 1_2 \dots 1_n)$ becomes the $(0, 0)$ -type matrix for an appropriate number n , namely, vanished, Equation System 2.22 has a positive solution.

Proof: Let $\tilde{L}_{ij}(x)$ denote the linear form when $B(i:\text{row})$ is considered instead of $B(i)$ in Relation 4.4. Suppose that $B(i:\text{row})$ is constructed in the way such that the t -th row is replaced by the linear combination of m rows of $B(i)$ with positive coefficients $\beta_1, \beta_2, \dots, \beta_m$. Then the following relations hold.

$$\tilde{L}_{ij}(x) = L_{ij}(x), \quad (j \neq t)$$

$$\tilde{L}_{it}(x) = \sum_{k=1}^m \beta_k \cdot L_{ik}(x)$$

On the other hand let $\tilde{b}_{i1_1}, \tilde{b}_{i1_2}, \dots, \tilde{b}_{i1_n}$ denote the 1_1 -th row, 1_2 -th row, $\dots$, 1_n -th row of $B(i:\text{row})$ respectively. Then, the linear combination

$$\tilde{b}_i = \sum_{j=1}^n \alpha_j \cdot \tilde{b}_{i1_j} \quad (\alpha_1 \gg \alpha_2 \gg \dots \gg \alpha_n > 0)$$

becomes a negative vector. Hence the linear form of Relation 4.4 which takes the components of the negative vector $\tilde{b}_i$ as its coefficients becomes

$$\sum_{k=1}^{2^{N-N-1}} \tilde{b}_{ik} \cdot x_k = \sum_{j=1}^n \alpha_j \cdot \tilde{L}_{ij}(x) < 0$$

where $\tilde{b}_i = (\tilde{b}_{i1}, \tilde{b}_{i2}, \dots, \tilde{b}_{i2^{N-N-1}})$.

Therefore the following relation holds

$$\sum_{j=1}^n \alpha_j \cdot \tilde{L}_{ij}(x) = \alpha_t \sum_{k=1}^m \beta_k \cdot L_{ik}(x) + \sum_{j \neq t} \alpha_j \cdot L_{ij}(x) < 0$$

Thus at least one of $L_{ij}(x)$ ($j=1, 2, \dots, N+1$) becomes a negative number. Q.E.D.

Let us practise the arguments stated hitherto for the case $N=3$. The notations and terms will be used to represent the same conception defined in the preceding discussions.

4.3 Example

Example 4.1. Examine whether or not the transition state diagram given by Fig. 4.1 is realizable with three threshold elements. If it is realizable, then determine the actual weights values and the threshold value of each element. Furthermore let us determine what domain these weights values and threshold value occupy in the $(N+1)$ -dimensional Euclidean space.

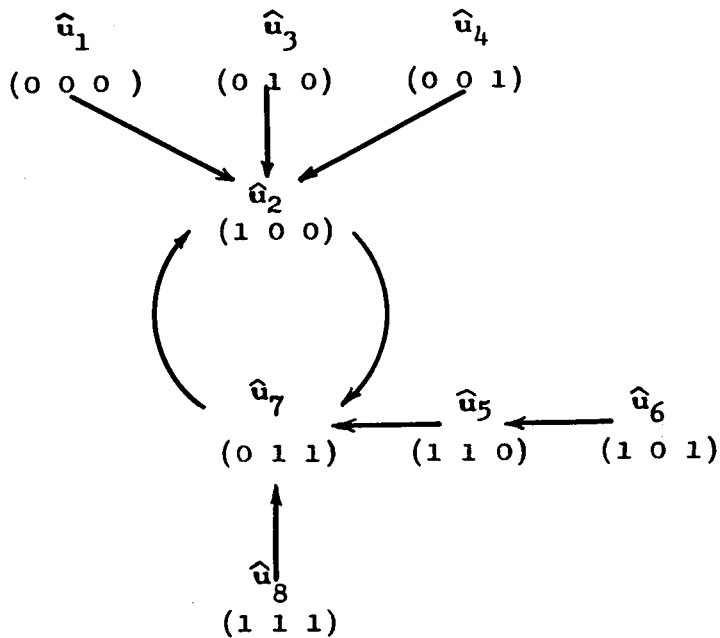


Fig. 4.1. Transition state diagram

Three characteristic matrices for this behavior are given by the following matrices.

$$C(1) = \begin{pmatrix} 1 & & & & \\ & -1 & & & \\ & & 1 & & \\ & & & -1 & \\ & 0 & & & 1 \\ & & & & & -1 \end{pmatrix} \quad C(2) = \begin{pmatrix} -1 & & & & \\ & 1 & & & \\ & & -1 & & \\ & & & -1 & \\ & & & & 1 \\ 0 & & & & & -1 \\ & & & & & & 1 \end{pmatrix}$$

$$C(3) = \begin{pmatrix} -1 & & & & \\ & 1 & & & \\ & & -1 & & \\ & & & -1 & \\ & & & & 1 \\ & 0 & & & & -1 \\ & & & & & & 1 \end{pmatrix}$$

Thus the corresponding fundamental solution matrices are given by the following three matrices.

$$B(1) = \begin{pmatrix} -1 & 1 & 1 & -2 \\ -1 & 1 & 0 & -1 \\ 1 & 0 & -1 & 1 \\ 0 & -1 & -1 & 1 \end{pmatrix} \quad B(2) = \begin{pmatrix} -1 & -1 & 1 & -2 \\ -1 & -1 & 0 & -1 \\ 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & 1 \end{pmatrix}$$

$$B(3) = \begin{pmatrix} -1 & 1 & 1 & -2 \\ -1 & 1 & 0 & -1 \\ 1 & 0 & -1 & 1 \\ 0 & -1 & -1 & 1 \end{pmatrix}$$

Since $B(3)$ is coincident with $B(1)$, only $B(1)$ and $B(2)$ are considered hereafter. Construct the row-combined fundamental solution matrix $B(1:\text{row})$ by adding the fourth row to the first row of $B(1)$. As the first row of $B(1:\text{row})$ is non-positive, construct the reduced row-combined fundamental solution matrix $\bar{B}(1:\text{row}:1)$ of order 1. Thus $\bar{B}(1:\text{row}:1,4)$ becomes vanished.

On the other hand since the second row of $B(2)$ is non-positive, construct $\bar{B}(2:2)$. Then $\bar{B}(2:2,3)$ becomes vanished. Hence Equation System 2.22 has a positive solution for any i . Thus the transition state diagram given by Fig. 4.1. is realized with three threshold elements. These processes are shown in Table 4.1.

$$B(1:\text{row}) = \begin{pmatrix} -1 & 0 & 0 & -1 \\ -1 & 1 & 0 & -1 \\ 1 & 0 & -1 & 1 \\ 0 & -1 & -1 & 1 \end{pmatrix} \Rightarrow B(1:\text{row}:1) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & -1 \\ -1 & -1 \end{pmatrix}$$

$$\Downarrow$$

$$B(1:\text{row}:1,4) = 0$$

$$B(2:2) = \begin{pmatrix} 1 \\ 0 \\ -1 \\ -1 \end{pmatrix} \Rightarrow B(2:2,3) = 0$$

Table 4.1. Actual processes for linear separability

The following three vectors d_1 , d_2 , d_3 can be positive solutions of Equation System 2.22.

$$d_1 = (4, 1, 2, 6, 3, 1, 4, 1)$$

$$d_2 = (1, 1, 1, 1, 1, 1, 1, 1)$$

$$d_3 = (4, 1, 2, 6, 3, 1, 4, 1)$$

Thus, by Theorem 3.1, we obtain

$$w_1 = (4, -5, -2, 2)$$

$$w_2 = (-1, 2, 0, 0)$$

$$w_3 = (-4, 5, 2, -2)$$

Therefore the autonomous network given by Fig. 4.2 shows the behavior given by Fig. 4.1.

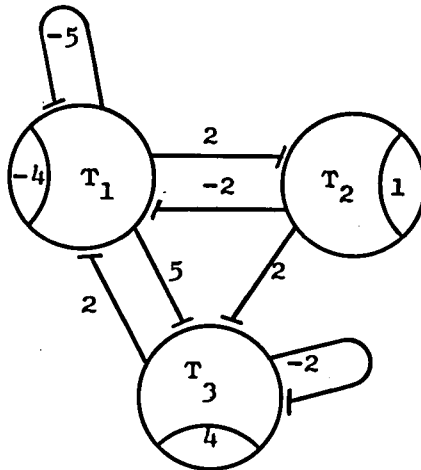


Fig. 4.2. The autonomous network

The domain of the weights values and the threshold value of each element is given as the intersection of the polyhedral convex cone, $R_i = \bigcap_{j=1}^2 K(A_{ij})$, where each matrix A_{ij} is provided as follows.

$$A_{11} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

$$A_{12} = \begin{pmatrix} -1 & 1 & 1 & 2 \\ 0 & 0 & -1 & -1 \\ 0 & -1 & 0 & -1 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

$$A_{21} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

$$A_{22} = \begin{pmatrix} 1 & 1 & -1 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

$$A_{31} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

$$A_{32} = \begin{pmatrix} 1 & -1 & -1 & -2 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

Example 4.2. Examine whether or not the transition state diagram given by Fig. 4.3 is realizable with three threshold elements.

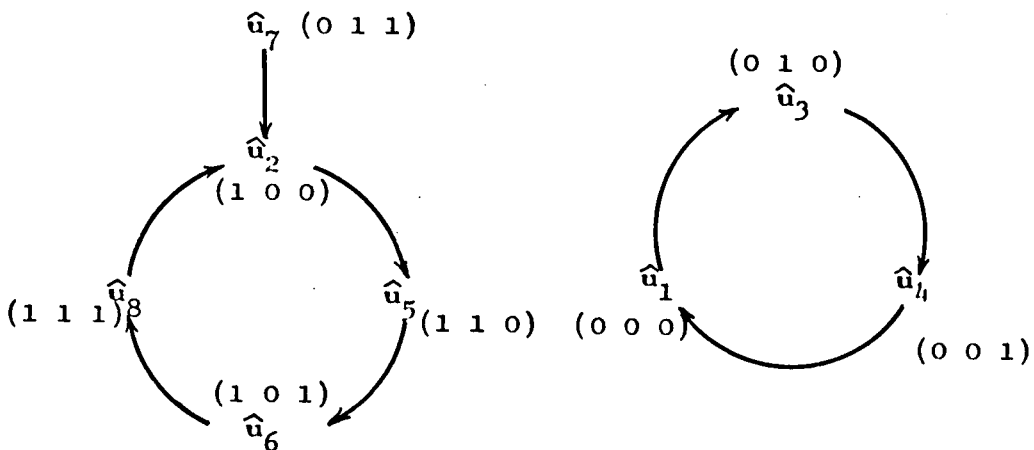


Fig. 4.3. Transition state diagram

Three characteristic matrices for this behavior are given as follows.

$$C(1) = \begin{pmatrix} -1 & & & & & & & \\ & 1 & & & & & & \\ & & -1 & & & & & \\ & & & -1 & & & & \\ & & & & 1 & & & \\ 0 & & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 1 \end{pmatrix} \quad C(2) = \begin{pmatrix} & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & -1 & & & \\ & & & & & -1 & & \\ 0 & & & & & & 1 & \\ & & & & & & & -1 & \\ & & & & & & & & -1 \end{pmatrix}$$

$$C(3) = \begin{pmatrix} -1 & & & & & & & \\ & -1 & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & 1 & & & \\ 0 & & & & & 1 & & \\ & & & & & & -1 & \\ & & & & & & & -1 \end{pmatrix}$$

Thus the corresponding fundamental solution matrices are given as follows.

$$B(1) = \begin{pmatrix} -1 & -1 & -1 & -2 \\ -1 & -1 & 0 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \quad B(2) = \begin{pmatrix} -1 & 1 & -1 & -2 \\ 1 & -1 & 0 & 1 \\ -1 & 0 & -1 & -1 \\ 0 & 1 & -1 & -1 \end{pmatrix}$$

$$B(3) = \begin{pmatrix} -1 & -1 & 1 & 2 \\ 1 & 1 & 0 & -1 \\ -1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1 \end{pmatrix}$$

Hence $\overleftarrow{B}(1:1)$ and $\overleftarrow{B}(2:3,1)$ become vanished through the processes shown below.

$$\overleftarrow{B}(2:3) = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \Rightarrow \overleftarrow{B}(2:3,1) = [0]$$

However the column-combined fundamental solution matrix $B(3:\text{column})$ constructed by adding the second column to the fourth column of $B(3)$ has a nonnegative column in the fourth column as follows.

$$B(3:\text{column}) = \begin{pmatrix} -1 & -1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix}$$

Hence Equation System 2.22 has no positive solution for $i=3$. Therefore the behavior given by Fig. 4.3 is not realizable with three threshold elements.

Note that if the component of a matrix is not given, that is, left blank, then it means the component is zero throughout the above representation of a matrix. Such a representation will also be used in the remainder of this paper.

CHAPTER V

LINEAR SEPARABILITY BY VARIABLE TRANSFORMATION

The tests for the linear separability so far proposed are, in general, not very practical when the function of more than five or six variables is to be tested. Since most of these proposed methods require a large amount of computation, the problem of determining whether or not a given function is a threshold function is commonly solved in actual by reducing the function to a canonical form and checking to see if it is contained in the entry of the tables prepared for this purpose. ^{(14) (62)}
^{(45) (64)}

Since the number of threshold functions grows very rapidly with an increase of the number of variables, it becomes necessary to enumerate only the representatives of equivalence classes. The number of representatives even in the most enlarged concept of equivalence becomes more than two thousands for the function of seven input variables. Thus this fact indicates that it is still necessary to develop a general test for the linear separability for the function of a large number of variables.

The main purpose of this chapter is to develop a general test to meet this requirement. This test is a method which examines whether or not a certain equation system has a non-zero nonnegative solution.

5.1 Preliminary Consideration

(57)
The practical procedures to be presented in this chapter consist of two fundamental operations, the transformation of input variables and the eliminations of columns and rows of matrices. This test procedures require at most N times of the transformation of input variables in order to examine whether or not the function of N variables is a threshold function. Hence this test is useful even if the number of input variables increases.

The fact that at most N times of the transformation are sufficient is due to the treatment of the adjoint linear equations instead of the inequalities directly. The following theorem promotes this idea and gives the foundation of the subsequent analysis. This theorem, brought out at first by Stiemke and settled later in the refined form by Tucker, plays a very important role not only in the theory of linear inequality but also in the theory of games and linear programming.

(46)
Stiemke's and Tucker's Theorem Let H denote an arbitrary (m,n) -type matrix. Then either of the following two propositions always holds and never holds simultaneously.

(1) The inequalities

$$H \cdot x > 0$$

have a solution $x = {}^t(x_1, x_2, \dots, x_n)$.

(2) The equations

$${}^tH \cdot y = 0$$

have a nonzero nonnegative solution $y = {}^t(y_1, y_2, \dots, y_m) \geq 0$.

By applying this theorem to Absolute Inequality System 2.16 it is concluded that we have to examine whether or not Equation System 2.21 has a nonzero nonnegative solution $b_i \geq 0$. That is, if Equation System 2.21 has no nonzero nonnegative solution $b_i \geq 0$, then Absolute Inequality System 2.16 has a solution and the reverse fact is also true.

This theorem states very concisely and explicitly the condition for the consistency of homogeneous strict inequalities, however it is not so surprised one. It states the entirely same fact as Theorem 4.3 which is reduced through Furtwängler's Theorem. Hence we can say that in Chapter IV we have rearranged this theorem into the more tractable form for the analysis.

The practical procedure to be discussed hereafter is a method which can test whether or not Equation System 2.21 has a nonzero nonnegative solution. There are $N+1$ equations in Equation System 2.21 whose coefficient matrix is ${}^tU \cdot C(i)$. The primary concept of this method is to diminish the equations one by one by the elimination of the row and columns of ${}^tU \cdot C(i)$ or by the transformation of input variables. Hence the necessary

the number of transformations to examine the existence for a nonzero nonnegative solution is decreased as many times as the order of the eliminations.

For the efficiency of procedures, the eliminations, in general, are performed before the transformations until further eliminations become to be unexcuted.

5.2 Simplification Procedure by Elimination

In this section let us present the simplification procedure of Equation System 2.21 by the eliminations of rows and columns of ${}^tU \cdot C(i)$.

Definition 5.1. Consider the set V_i of the internal state vectors defined as follows.

$$\begin{aligned} \hat{u}_j &\in V_i && \text{if } \varphi_i(\hat{u}_j) = 1, \\ &\notin V_i && \text{if } \varphi_i(\hat{u}_j) = 0 \end{aligned}$$

Proposition 5.1. If the relation

$$V_j \subseteq P_i \quad (j = 1, 2, \dots, N)$$

is satisfied, then the $(j+1)$ -th row of ${}^tU \cdot C(i)$ is a nonzero nonnegative vector whose each component is either 0 or 1. On the other hand if the relation

$$V_j \subseteq Q_i \quad (j = 1, 2, \dots, N)$$

is satisfied, then the $(j+1)$ -th row of ${}^tU \cdot C(i)$ is a nonzero nonpositive vector whose each component is either 0 or -1.

Proof: If $\hat{u}_k \in V_j \subseteq P_i$, then the k -th column of ${}^tU \cdot C(i)$ equals to $\hat{u}_k$. Hence the k -th component of the $(j+1)$ -th row of ${}^tU \cdot C(i)$ is 1 for $\hat{u}_k \in V_j$ and is 0 for $\hat{u}_k \notin V_j$. Thus in any way the component of $(j+1)$ -row is either 0 or 1.

If $\hat{u}_k \in V_j \subseteq Q_i$, then the k -th column of ${}^tU \cdot C(i)$ equals to $-\hat{u}_k$. Therefore by the similar argument the proposition follows. Q.E.D.

Proposition 5.2. If the relation

$$P_i = \nabla \quad (5.1)$$

is satisfied, all the components of ${}^tU \cdot C(i)$ are 1 and 0. On the other hand if the relation

$$P_i = \Phi \quad (5.2)$$

is satisfied, all the components of ${}^tU \cdot C(i)$ are -1 and 0.

Proof: If $P_i = \nabla$, then the characteristic matrix $C(i)$ becomes the identity matrix E_{2N} . Thus ${}^tU \cdot C(i) = {}^tU$. On the other hand if $P_i = \Phi$, then $C(i)$ becomes $-E_{2N}$. Therefore, ${}^tU \cdot C(i) = -{}^tU$. Q.E.D.

Proposition 5.3. If either Relation 5.1 or Relation 5.2 is satisfied, then Absolute Inequality System 2.16 is consistent.

Definition 5.2. If either of the two relations

$$V_{1_1} \subseteq P_i \quad \text{or} \quad V_{1_1} \subseteq Q_i$$

is satisfied, exclude all the k 's-th columns from the matrix ${}^tU \cdot C(i)$ such that $\hat{u}_k \in V_{1_1}$ and moreover exclude the 1_1 -th row. Such an obtained matrix is referred to as the curtailed matrix ${}^t\overleftarrow{U}(1_1) \cdot \overleftarrow{C}(i)$ of order 1 concerning the threshold element T_{1_1} .

Let us expand this concept and define the curtailed matrix ${}^t\overleftarrow{U}(1_1 1_2 \dots 1_n) \cdot \overleftarrow{C}(i)$ of order n concerning the threshold elements $T_{1_1}, T_{1_2}, \dots, T_{1_n}$ stated as follows. That is, if there exists the curtailed matrix ${}^t\overleftarrow{U}(1_1 1_2 \dots 1_{n-1}) \cdot \overleftarrow{C}(i)$ of order $n-1$ and moreover if either of the two relations

$$(V_{1_n} - \bigcup_{j=1}^{n-1} V_{1_j}) \subseteq P_i \quad (5.3)$$

or

$$(V_{1_n} - \bigcup_{j=1}^{n-1} V_{1_j}) \subseteq Q_i \quad (5.4)$$

is satisfied, exclude all the k 's-th columns from the matrix ${}^t\overleftarrow{U}(1_1 1_2 \dots 1_{n-1}) \cdot \overleftarrow{C}(i)$ such that $\hat{u}_k \in (V_{1_n} - \bigcup_{j=1}^{n-1} V_{1_j})$ and furthermore exclude the (1_n+1) -th row.

Proposition 5.4. The curtailed matrix of order n ($n=1, 2, \dots, N$) is an $(N-n+1, 2^{N-n})$ -type submatrix of ${}^tU \cdot C(i)$.

Proof: The number of the components contained in the set $\bigcup_{j=1}^n V_{1_j}$ is $(2^N - 2^{N-n})$. Hence the number of the columns which are to be excluded from the 2^N columns of ${}^tU \cdot C(i)$ is

$(2^N - 2^{N-n})$. On the other hand one row is always excluded as the order of the curtailed matrix is raised. Q.E.D.

Theorem 5.1. The equations whose coefficient matrix is the curtailed matrix of order $n-1$

$${}^t\overleftarrow{U}(1_1 1_2 \dots 1_{n-1}) \cdot \overleftarrow{C}(i) \cdot {}^t\overleftarrow{b}_{i,n-1} = 0 \quad (5.5)$$

have a nonzero nonnegative solution $\overleftarrow{b}_{i,n-1} \geq 0$, if and only if the equations whose coefficient matrix is the curtailed matrix of order n

$${}^t\overleftarrow{U}(1_1 1_2 \dots 1_n) \cdot \overleftarrow{C}(i) \cdot {}^t\overleftarrow{b}_{i,n} = 0 \quad (5.6)$$

have a nonzero nonnegative solution $\overleftarrow{b}_{i,n} \geq 0$.

Proof: First consider the "only if" part. The (1_n+1) -th row of ${}^t\overleftarrow{U}(1_1 1_2 \dots 1_{n-1})$ is a nonzero nonnegative vector which consists of 1 and 0, if Relation 5.3 holds. On the other hand, if Relation 5.4 holds, then this row is a nonzero nonpositive vector which consists of -1 and 0. Hence all the k 's-th components of the solution $\overleftarrow{b}_{i,n-1}$ of Equation 5.5 must be 0 such that $\hat{u}_k \in (V_{1_n} - \bigcup_{j=1}^{n-1} V_{1_j})$. Therefore from the definition of ${}^t\overleftarrow{U}(1_1 1_2 \dots 1_n) \cdot \overleftarrow{C}(i)$, it follows that the 2^{N-n} -dimensional vector $\overleftarrow{b}_{i,n}$ obtained by excluding all the k 's-th components of the 2^{N-n+1} -dimensional solution of Equation 5.5 such that $\hat{u}_k \in (V_{1_n} - \bigcup_{j=1}^{n-1} V_{1_j})$ can be surely a solution of Equation 5.6. This completes the proof of the "only if" part.

Let us construct a 2^{N-n+1} -dimensional vector $\overleftarrow{b}_{i,n-1} \geq 0$

from the solution $\overleftarrow{b}_{i,n} \geq 0$ of Equation 5.6 by attaching 2^{N-n} zeros as the k 's-th components such that $\hat{u}_k \in (v_{1_n} - \bigcup_{j=1}^{n-1} v_{1_j})$. Such an obtained vector $\overleftarrow{b}_{i,n-1}$ can be surely a solution of Equation 5.5 from the definition of ${}^t\overleftarrow{U}(1_1 1_2 \dots 1_n) \cdot \overleftarrow{C}(i)$. This completes the proof of the "if" part. Q.E.D.

For a convenience, the matrix ${}^tU \cdot C(i)$ itself will be called the curtailed matrix of order 0. The next corollary follows immediately from Stiemke's and Tucker's Theorem and Theorem 5.1.

Corollary 5.1.1. Suppose that there exists ${}^t\overleftarrow{U}(1_1) \cdot \overleftarrow{C}(i)$. Then Absolute Inequality System 2.16 is consistent if and only if the equations

$${}^t\overleftarrow{U}(1_1) \cdot \overleftarrow{C}(i) \cdot {}^t\overleftarrow{b}_{i,1} = 0$$

have no nonzero nonnegative solution $\overleftarrow{b}_{i,1} \geq 0$.

Theorem 5.2. For the case where there exists ${}^t\overleftarrow{U}(1_1 1_2 \dots 1_n) \cdot \overleftarrow{C}(i)$, the equations

$${}^tU \cdot C(i) \cdot {}^tb_i = 0 \quad (2.21)$$

have a nonzero nonnegative solution $b_i \geq 0$, if and only if the equations

$${}^t\overleftarrow{U}(1_1 1_2 \dots 1_n) \cdot \overleftarrow{C}(i) \cdot {}^t\overleftarrow{b}_{i,n} = 0 \quad (5.6)$$

have a nonzero nonnegative solution $\overleftarrow{b}_{i,n} \geq 0$.

Proof: This fact is immediately obtained by applying Theorem 5.1 successively by n times. Q.E.D.

The number of equations and unknown variables in Equation 5.6 is decreased by n and $2^N - 2^{N-n}$ respectively, compared with Equation 2.21. Hence Theorem 5.2 shows that the test for the linear separability becomes easier if we use Equation 5.6 rather than Equation 2.21. The nearer n approaches N , the easier the test with Equation 5.6 becomes.

Therefore in order to examine the consistency of Absolute Inequality System 2.16, Equation 5.6 where the curtailed matrix of the highest order is adopted is most convenient to investigate. Thus if ${}^t\vec{U}(1_1 1_2 \dots 1_M) \cdot \vec{C}(i)$ denotes the curtailed matrix of the highest order, then it is easiest to examine whether or not the equations

$${}^t\vec{U}(1_1 1_2 \dots 1_M) \cdot \vec{C}(i) \cdot {}^t\vec{b}_{i,M} = 0 \quad (5.7)$$

have a nonzero nonnegative solution $\vec{b}_{i,M} \geq 0$.

Theorem 5.3. If there exists the curtailed matrix of order N , then Absolute Inequality System 2.16 is consistent.

Proof: The curtailed matrix of order N always becomes a $(1,1)$ -type matrix from Proposition 5.4, and its component must be 1 or -1. Hence Equation 5.6 becomes ${}^t\vec{b}_{i,N} = 0$ for $n=N$. Thus it is concluded that there exists no nonzero nonnegative solution $b_i \geq 0$ of Equation 2.21 through Theorem 5.2. Q.E.D.

Theorem 5.4. Let $\overleftarrow{tU}(1_1 1_2 \dots 1_M) \cdot \overleftarrow{C}(i)$ denote the curtailed matrix of the highest order that exists really. Then neither nonnegative row nor nonpositive row exists in the rows of $\overleftarrow{tU}(1_1 1_2 \dots 1_M) \cdot \overleftarrow{C}(i)$. That is, each row of $\overleftarrow{tU}(1_1 1_2 \dots 1_M) \cdot \overleftarrow{C}(i)$ always contains 1, -1, and 0 as its components.

Proof: If the 1_{M+1} -th row of $\overleftarrow{tU}(1_1 1_2 \dots 1_M) \cdot \overleftarrow{C}(i)$ is a nonzero nonnegative (nonpositive) vector, then the relation

$$(v_{1_{M+1}} - \bigcup_{j=1}^M v_{1_j}) \subseteq P_i(Q_i)$$

holds. Hence the curtailed matrix of order $M+1$ can exist.

This contradicts the assumption.

Q.E.D.

5.3 Simplification Procedure by Transformation (24)

It is not necessarily so easy except for the case $M=N$ to investigate about Equation 5.7, since every equation contains both a positive coefficient term and a negative coefficient term. Thus in this section let us develop the practical procedure to examine the existence of a nonzero nonnegative solution of the equations like Equation 5.7.

The procedure to be discussed in this section is a method which decreases the number of equations by transforming the unknown variables. If the transformation is performed once, one equation always vanishes. Therefore if there are m equations in a given equation system, the transformation of $m-1$ times is sufficient to examine completely whether or not the

given equation system has a nonzero nonnegative solution.

Let H denote an arbitrary (m,n) -type matrix which has neither nonnegative row nor nonpositive row. Hence if $h_i = (h_{i1}, h_{i2}, \dots, h_{in})$ denotes the i -th row of H , then some of these components h_{ij} ($j=1, 2, \dots, n$) take positive numbers, others take negative numbers and the remainders may take zeros. Without loss of generality suppose that the first p components are positive, the successive q components are negative and the last r components are zeros. Then the i -th equation of the equation system $H \cdot x = 0$ can be written as follows.

$$h_{i1}x_1 + h_{i2}x_2 + \dots + h_{ip}x_p = -h_{ip+1}x_{p+1} - h_{ip+2}x_{p+2} - \dots - h_{ip+q}x_{p+q}$$

where $h_{ij} > 0$ ($j=1, 2, \dots, p$) and $-h_{ip+j} > 0$ ($j=1, 2, \dots, q$).

Now consider the transformation table given by Table 5.1 and make such a transformation of unknown variables as given by Relation 5.8.

z_{11}	z_{12}	$\dots$	z_{1q}	$h_{i1}x_1$
z_{21}	z_{22}	$\dots$	z_{2q}	$h_{i2}x_2$
z_{31}	z_{32}	$\dots$	z_{3q}	$h_{i3}x_3$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
z_{p1}	z_{p2}	$\dots$	z_{pq}	$h_{ip}x_p$

$$-h_{ip+1}x_{p+1} \quad -h_{ip+2}x_{p+2} \quad -h_{ip+q}x_{p+q}$$

Table 5.1. Transformation Table

$$\begin{aligned}
h_{11}x_1 &= z_{11} + z_{12} + \dots + z_{1q} \\
h_{12}x_2 &= z_{21} + z_{22} + \dots + z_{2q} \\
h_{13}x_3 &= z_{31} + z_{32} + \dots + z_{3q} \\
&\dots \\
&\dots \\
h_{ip}x_p &= z_{p1} + z_{p2} + \dots + z_{pq} \\
-h_{ip+1}x_{p+1} &= z_{11} + z_{21} + \dots + z_{p1} \\
-h_{ip+2}x_{p+2} &= z_{12} + z_{22} + \dots + z_{p2} \\
-h_{ip+3}x_{p+3} &= z_{13} + z_{23} + \dots + z_{p3} \\
&\dots \\
&\dots \\
-h_{ip+q}x_{p+q} &= z_{1q} + z_{2q} + \dots + z_{pq} \\
x_{p+q+1} &= z_{1p+q+1} \\
x_{p+q+2} &= z_{1p+q+2} \\
&\dots \\
x_n &= z_{1n}
\end{aligned} \tag{5.8}$$

This transformation can be expressed by the matrix shown in Table 5.2.

where $\alpha_j = h_{ij}^{-1} > 0$ ($j = 1, 2, \dots, p$), $\beta_j = -h_{ip+j}^{-1} > 0$ ($j = 1, 2, \dots, q$)

Table 5.2. Transformation Matrix

Definition 5.3. Let us represent by $\alpha = D_i Z$ the relation denoted by Table 5.2. The matrix D_i is called the transformation matrix of variables. Here the suffix i shows the i -th row of the matrix H .

Proposition 5.5. The transformation matrix D_i is always positive definite. That is, $D_i \geq 0$.

Theorem 5.5. Let H denote an (m,n) -type matrix which has neither nonpositive row nor nonnegative row. Moreover consider the transformation matrix D_i given by Table 5.2. Then the equation system

$$H \cdot D_i \cdot Z = 0$$

consists of $m-1$ equations and $p+q+r$ unknown variables where p , q and r denote the number of the positive components, the negative components and zero components contained in the i -th row of H , respectively.

Proof: The i -th equation of the equation system

$$H \cdot \alpha = 0$$

is transformed into the form

$$\sum_{i,j} z_{ij} - \sum_{i,j} z_{ij} = 0$$

by the transformation given by Relation 5.8. Then the i -th equation becomes vanished. The number of unknown variables is obviously $p+q+r$ through Table 5.2. Q.E.D.

Theorem 5.6. Let H denote an (m,n) -type matrix which has neither nonpositive row nor nonnegative row, and D_i denote the transformation matrix given by Table 5.2. Then the equation system

$$H \cdot x = 0 \quad (5.9)$$

has a nonzero nonnegative solution $x \geq 0$ if and only if the equation system

$$H \cdot D_i \cdot Z = 0 \quad (5.10)$$

has a nonzero nonnegative solution $Z \geq 0$.

Proof: If Equation 5.10 has a nonzero nonnegative solution Z , then x determined by the relation $x = D_i \cdot Z$ becomes a nonzero nonnegative vector through Proposition 5.5 and it can surely be a solution of Equation 5.9.

Converse fact is proved as follows. Here the proof is shown by giving the practical procedure to determine an actual nonzero nonnegative vector $Z \geq 0$ for a given nonzero nonnegative vector $x \geq 0$ connected by the relation $x = D_i \cdot Z$. The transformation $x = D_i \cdot Z$ is shown by Equation 5.8. Here with respect to the part

$$\begin{aligned} x_{p+q+1} &= z_{1p+q+1} \\ x_{p+q+2} &= z_{1p+q+2} \\ &\dots \dots \dots \\ x_n &= z_{1n} \end{aligned}$$

there is no problem, if it is determined that z_{ij} has the equal value to the given x_j ($j=p+q+1, p+q+2, \dots, n$). Hence it is sufficient to consider the remaining part.

$$\begin{aligned}
 h_{11}x_1(0) &= z_{11} + z_{12} + \dots + z_{1q} \\
 h_{12}x_2(0) &= z_{21} + z_{22} + \dots + z_{2q} \\
 . &. \\
 h_{ip}x_p(0) &= z_{p1} + z_{p2} + \dots + z_{pq} \\
 -h_{ip+1}x_{p+1}(0) &= z_{11} + z_{21} + \dots + z_{p1} \\
 -h_{ip+2}x_{p+2}(0) &= z_{12} + z_{22} + \dots + z_{p2} \\
 . &. \\
 -h_{ip+q}x_{p+q}(0) &= z_{1q} + z_{2q} + \dots + z_{pq}
 \end{aligned} \tag{5.11}$$

Here the index "(0)" attached to the term $h_{ij}x_j(0)$ represents the stage of the operation stated in the following paragraph. For any given $h_{ij}x_j(0) \geq 0$ ($j=1, 2, \dots, p$) and $-h_{ip+j}x_{p+j}(0) \geq 0$ ($j=1, 2, \dots, q$), let us determine concrete values $z_{ij} \geq 0$ ($i=1, 2, \dots, p, j=1, 2, \dots, q$) in Equation 5.11.

Stage 1

If the relation

$$h_{i1}x_1(0) + h_{ip+1}x_{p+1}(0) \leq 0$$

holds, then determine as follows

$$z_{11} = h_{11} x_1(0), \quad z_{12} = z_{13} = \dots z_{1q} = 0.$$

Moreover perform the replacement given as follows.

$$\begin{aligned}
 h_{i2}x_2(0) &= h_{i2}x_2(1) \\
 h_{i3}x_3(0) &= h_{i3}x_3(1) \\
 &\dots\dots\dots \\
 h_{ip}x_p(0) &= h_{ip}x_p(1) \\
 -h_{ip+1}x_{p+1}(0)-z_{11} &= -h_{ip+1}x_{p+1}(1) \\
 -h_{ip+2}x_{p+2}(0)-z_{12} &= -h_{ip+2}x_{p+2}(1) \\
 &\dots\dots\dots \\
 -h_{ip+q}x_{p+q}(0)-z_{1q} &= -h_{ip+q}x_{p+q}(1)
 \end{aligned}$$

Now rewrite Equation 5.11 by using such defined $h_{ij}x_j(1)$ ($j=2, 3, \dots, p$) and $-h_{ip+j}x_{p+j}(1)$ ($j=1, 2, \dots, q$) and enter the Stage 2.

On the other hand if the relation

$$h_{i1}x_1(0) + h_{ip+1}x_{p+1}(0) > 0$$

holds, then determine as follows.

$$z_{11} = -h_{ip+1}x_{p+1}(0) , \quad z_{21} = z_{31} = \dots z_{p1} = 0$$

Moreover perform the replacement given as follows.

$$h_{i1}x_1(0)-z_{i1} = h_{i1}x_1(1)$$

$$h_{i2}x_2(0)-z_{i2} = h_{i2}x_2(1)$$

.

$$h_{ip}x_p(0)-z_{ip} = h_{ip}x_p(1)$$

$$-h_{ip+2}x_{p+2}(0) = -h_{ip+2}x_{p+2}(1)$$

$$-h_{ip+3}x_{p+3}(0) = -h_{ip+3}x_{p+3}(1)$$

.

$$-h_{ip+q}x_{p+q}(0) = -h_{ip+q}x_{p+q}(1)$$

Now rewrite Equation 5.11 by using such defined $h_{ij}x_j(1)$ ($j=1, 2, \dots, p$) and $-h_{ip+j}x_{p+j}(1)$ ($j=2, 3, \dots, q$) and enter Stage 2.

After the similar operations as Stage 1 are performed by "t" times, Equation 5.11 is replaced into the following form.

$$h_{i\eta}x_{\eta}(t) = z_{\eta\mu} + z_{\eta\mu+1} + \dots + z_{\eta q}$$

$$h_{i\eta+1}x_{\eta+1}(t) = z_{\eta+1\mu} + z_{\eta+1\mu+1} + \dots + z_{\eta+1q}$$

.

$$h_{ip}x_p(t) = z_{p\mu} + z_{p\mu+1} + \dots + z_{pq}$$

$$-h_{ip+\mu}x_{p+\mu}(t) = z_{\eta\mu} + z_{\eta+1\mu} + \dots + z_{p\mu}$$

$$-h_{ip+\mu+1}x_{p+\mu+1}(t) = z_{\eta\mu+1} + z_{\eta+1\mu+1} + \dots + z_{p\mu+1}$$

.

$$-h_{ip+q}x_{p+q}(t) = z_{\eta q} + z_{\eta+1q} + \dots + z_{pq}$$

where η and μ satisfy the relation $\eta+\mu=t+2$. Now enter

Stage $t+1$ stated as follows.

Stage $t+1$

If the relation

$$h_{i\eta} x_{\eta}(t) + h_{ip+\mu} x_{p+\mu}(t) \leq 0$$

holds, then determine as follows.

$$z_{\eta\mu} = h_{i\eta} x_{\eta}(t), z_{\eta\mu+1} = z_{\eta\mu+2} = \dots = z_{\eta q} = 0$$

Moreover perform the replacement given as follows.

$$h_{i\eta+1} x_{\eta+1}(t) = h_{i\eta+1} x_{\eta+1}(t+1)$$

$$h_{i\eta+2} x_{\eta+2}(t) = h_{i\eta+2} x_{\eta+2}(t+1)$$

.....

$$h_{ip} x_p(t) = h_{ip} x_p(t+1)$$

$$-h_{ip+\mu} x_{p+\mu}(t) - z_{\eta\mu} = -h_{ip+\mu} x_{p+\mu}(t+1)$$

$$-h_{ip+\mu+1} x_{p+\mu+1}(t) - z_{\eta\mu+1} = -h_{ip+\mu+1} x_{p+\mu+1}(t+1)$$

.....

$$-h_{ip+q} x_{p+q}(t) - z_{\eta q} = -h_{ip+q} x_{p+q}(t+1)$$

Furthermore rewrite Equation 5.11 by using such defined

$h_{ij} x_j(t+1)$ ($j=\eta+1, \eta+2, \dots, p$) and $-h_{ip+j} x_{p+j}(t+1)$ ($j=\mu, \mu+1, \dots, q$) and enter Stage $t+2$.

On the other hand if the relation

$$h_{i\eta} x_{\eta}(t) + h_{i p+\mu} x_{p+\mu}(t) \geq 0$$

holds, then determine as follows.

$$z_{\eta\mu} = h_{ip+\mu} x_{p+\mu}(t), \quad z_{\eta+1\mu} = z_{\eta+2\mu} = \dots = z_{p\mu} = 0$$

Moreover perform the replacement given as follows.

[illegible]

Furthermore rewrite Equation 5.11 by using such defined $h_{ij}x_j(t+1)$ ($j=\eta, \eta+1, \dots, p$) and $-h_{ip+j}x_{p+j}(t+1)$ ($j=\mu+1, \mu+2, \dots, q$) and enter Stage $t+2$.

After similar operations are performed by $p+q$ times, the nonzero nonnegative vector $Z \geq 0$ has been already determined for any given nonzero nonnegative vector $x \geq 0$. Q.E.D.

5.4 Expansion of the Concept of Curtailed Matrix

In the preceding section the curtailed matrix is defined for the matrix ${}^tU \cdot C(i)$. Here, let us apply this concept to an arbitrary matrix H .

Definition 5.4. Let H be any given (m,n) -type matrix which has no row formed only by zero components. Suppose that the l_1 -th row h_{l_1} of H is a nonpositive (nonnegative) vector, then exclude all the columns from H such that the components of h_{l_1} are nonpositive (nonnegative) numbers and exclude, moreover, the l_1 -th row. Such an obtained matrix is called the curtailed matrix $\overleftarrow{H}(l_1)$ of order l .

By the similar manner, the curtailed matrix $\overleftarrow{H}(l_1 l_2 \dots l_n)$ of order n can be defined as follows. That is, if there exists the curtailed matrix $\overleftarrow{H}(l_1 l_2 \dots l_{n-1})$ of order $n-1$ which has no row formed only by zero components. Suppose, moreover, that the l_n -th row h_{l_n} of $\overleftarrow{H}(l_1 l_2 \dots l_{n-1})$ is a nonpositive (nonnegative) vector, then exclude all the columns from $\overleftarrow{H}(l_1 l_2 \dots l_{n-1})$ such that the components of h_{l_n} are nonpositive (nonnegative) numbers and exclude, furthermore, the l_n -th row. If there exists in $\overleftarrow{H}(l_1 l_2 \dots l_{n-1})$ such a row as is formed only by zero components, then at first exclude this zero vector and then perform the same process.

Then, the similar theorems as given in Section 5.2 can be established by the almost same arguments.

Theorem 5.7. The equations whose coefficient matrix is $\overleftarrow{H}(1_1 1_2 \dots 1_{n-1})$

$$\overleftarrow{H}(1_1 1_2 \dots 1_{n-1}) \cdot {}^t \overleftarrow{x}_{n-1} = 0$$

have a nonzero nonnegative solution $\overleftarrow{x}_{n-1} \geq 0$ if and only if the equations whose coefficient matrix is $\overleftarrow{H}(1_1 1_2 \dots 1_n)$

$$\overleftarrow{H}(1_1 1_2 \dots 1_n) \cdot {}^t \overleftarrow{x}_n = 0$$

have a nonzero nonnegative solution $\overleftarrow{x}_n \geq 0$

Theorem 5.8. Let $\overleftarrow{H}(1_1 1_2 \dots 1_M)$ denote the curtailed matrix of the highest order for H. Then the inequalities

$${}^t \overleftarrow{H} \cdot {}^t y > 0$$

are consistent if and only if the equations

$$\overleftarrow{H}(1_1 1_2 \dots 1_M) \cdot {}^t \overleftarrow{x}_M = 0$$

have no nonzero nonnegative solution $\overleftarrow{x}_M \geq 0$.

Theorem 5.9. If there exists a curtailed matrix $\overleftarrow{H}(1_1 1_2 \dots 1_{m-1})$ of order $m-1$ for any given (m,n) -type matrix H, then the inequalities

$${}^t \overleftarrow{H} \cdot {}^t y > 0$$

are consistent if and only if $\overleftarrow{H}(1_1 1_2 \dots 1_{m-1})$ is positive definite $\overleftarrow{H}(1_1 1_2 \dots 1_{m-1}) > 0$ or negative definite $\overleftarrow{H}(1_1 1_2 \dots 1_{m-1}) < 0$.

Proof: Since $\overleftarrow{H}(1_1 1_2 \dots 1_{m-1})$ is a row vector, the equations

$$\overleftarrow{H}(1_1 1_2 \dots 1_{m-1}) \cdot \overleftarrow{x}_{m-1} = 0 \quad (5.12)$$

have nonzero nonnegative solution $x_{m-1} \geq 0$ for any vector $\overleftarrow{H}(1_1 1_2 \dots 1_{m-1})$ except for $\overleftarrow{H}(1_1 1_2 \dots 1_{m-1}) > 0$ or $\overleftarrow{H}(1_1 1_2 \dots 1_{m-1}) < 0$. Q.E.D.

Observe here that Equation 5.12 can possess a nonzero nonnegative solution $x_{m-1} \geq 0$ even if $H(1_1 1_2 \dots 1_{m-1}) \leq 0$ or $H(1_1 1_2 \dots 1_{m-1}) \geq 0$.

Theorem 5.10. If there exists a curtailed matrix of order m for any given (m,n) -type matrix H , then the inequalities

$$t_H \cdot t_Y > 0$$

are consistent.

Proof: The curtailed matrix of order $m-1$ is always composed of one row h . Since there exists a curtailed matrix of order m by assumption, the row h has to be a nonnegative vector $h \geq 0$ or a nonpositive vector $h \leq 0$. However, if at least one zero component is contained in the components of h , all the components of the column corresponding to this zero component must be zero through the definition of the curtailed matrix. Consequently the vector h becomes positive or negative.

Q.E.D.

5.5 Practical Procedure for Linear Separability

The purpose of this section is to provide actually the procedures to examine the consistency of Absolute Inequality System 2.16 by using the basic principles so far discussed.

- (1) Construct the curtailed matrix ${}^t\vec{U}(1_1 1_2 \dots 1_{M_1})$. $\vec{C}(i)$ of the highest order for ${}^tU \cdot C(i)$. This process is done only by searching for a nonnegative row or a nonpositive row. Hence it is performed easily.
- (2) If $M_1 = N$, Absolute Inequality System 2.16 is consistent from Theorem 5.3.
- (3) If $M_1 < N$, compute the value $(p \cdot q + r)$ for each row of $\vec{U}(1_1 1_2 \dots 1_{M_1}) \cdot \vec{C}(i)$ and then determine the row for which this value becomes minimum where p , q and r represent the number of positive components, negative components and zero components of each row, respectively. This process is performed in order to make the dimension of the vector Z , introduced by Definition 5.3, minimum.
- (4) If the row searched by Process (3) is j_1 -th, then construct the transformation matrix D_{j_1} concerning the j_1 -th row.
- (5) Compute the product ${}^t\vec{U}(1_1 1_2 \dots 1_{M_1}) \cdot \vec{C}(i) \cdot D_{j_1}$. If $M_1 + 1 = N$, then $\vec{U}(1_1 1_2 \dots 1_{M_1}) \cdot \vec{C}(i) \cdot D_{j_1}$ becomes a matrix composed of one row. Thus it is readily

examined whether or not the equations

$${}^{\leftarrow}U(1_1 1_2 \dots 1_{M_1}) \cdot {}^{\leftarrow}C(i) \cdot D_{j_1} \cdot {}^{\leftarrow}x_{M_1+1} = 0$$

have a nonzero nonnegative solution by using Theorem 5.9.

- (6) If $M_1+1 < N$, then construct the curtailed matrix of the highest order for ${}^{\leftarrow}U(1_1 1_2 \dots 1_{M_1}) \cdot {}^{\leftarrow}C(i) \cdot D_{j_1}$. Let us denote this matrix by ${}^{\leftarrow}U(1_1 1_2 \dots 1_{M_1}) \cdot {}^{\leftarrow}C(i) \cdot \overleftarrow{D}_{j_1}(1_1 1_2 \dots 1_{M_2})$.
- (7) If $M_1+1+M_2=N$, then Absolute Inequality System 2.16 is consistent from Theorem 5.10.
- (8) If $M_1+1+M_2 < N$, then compute the value $(p \cdot q+r)$ for each row of ${}^{\leftarrow}U(1_1 1_2 \dots 1_{M_1}) \cdot {}^{\leftarrow}C(i) \cdot \overleftarrow{D}_{j_1}(1_1 1_2 \dots 1_{M_2})$ and then determine the row for which this value becomes minimum.
- (9) If the row searched by Process(8) is j_2 -th, then construct the transformation matrix D_{j_2} .
- (10) Compute the product ${}^{\leftarrow}U(1_1 1_2 \dots 1_{M_1}) \cdot {}^{\leftarrow}C(i) \cdot \overleftarrow{D}_{j_1}(1_1 1_2 \dots 1_{M_2}) \cdot D_{j_2}$.
- (11) If $M_1+M_2+2=N$, ${}^{\leftarrow}U(1_1 1_2 \dots 1_{M_1}) \cdot {}^{\leftarrow}C(i) \cdot \overleftarrow{D}_{j_1}(1_1 1_2 \dots 1_{M_2}) \cdot D_{j_2}$ becomes a matrix composed of one row. Hence it is readily examined whether or not the equations

$${}^{\leftarrow}U(1_1 1_2 \dots 1_{M_1}) \cdot {}^{\leftarrow}C(i) \cdot \overleftarrow{D}_{j_1}(1_1 1_2 \dots 1_{M_2}) \cdot D_{j_2} \cdot {}^{\leftarrow}x_{M_1+M_2+2} = 0$$

have a nonzero nonnegative solution by using Theorem 5.9.

- (12) If $M_1 + M_2 + 2 < N$, then construct the curtailed matrix of the highest order for $\overleftarrow{t_U}(1_1 1_2 \dots 1_{M_1}) \cdot \overleftarrow{C}(i) \cdot \overleftarrow{D}_{j_1}(1_1 1_2 \dots 1_{M_2}) \cdot D_{j_2}$. Let us denote this matrix by $\overleftarrow{t_U}(1_1 1_2 \dots 1_{M_1}) \cdot \overleftarrow{C}(i) \cdot \overleftarrow{D}_{j_1}(1_1 1_2 \dots 1_{M_2}) \cdot \overleftarrow{D}_{j_2}(1_1 1_2 \dots 1_{M_3})$
- (13) If the similar processes stated in (7), (8), (9), ... are repeated, the matrix $t_U \cdot C(i)$ is transformed into one row at last. Therefore since it can be readily examined whether or not Equation 2.21 has a nonzero nonnegative solution $x \geq 0$ by using Theorem 5.9, we can test very easily the consistency of Absolute Inequality System 2.16

By the above arguments it is concluded that the number of times of the transformation to be executed is decreased by $\sum_i M_i$. Therefore the necessary number of times of the transformation for the linear separability is given by

$$N - \sum_i M_i$$

Constructing a curtailed matrix is a much easier process than constructing a transformation matrix. Hence the procedure becomes easier as the value $\sum_i M_i$ becomes larger.

5.6 Example

Example 5.1. Examine whether or not the transition state diagram given by Fig. 5.1 is realizable with three threshold elements.

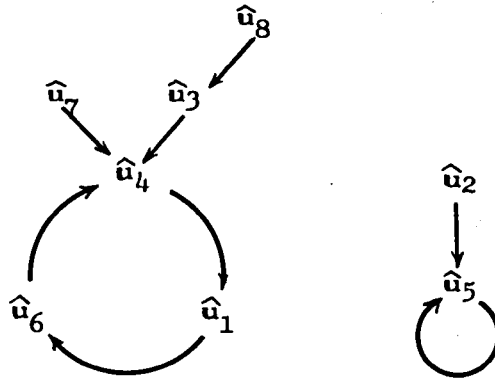


Fig. 5.1. The transition state diagram

The three characteristic matrices for this behavior are given as follows.

$$C(1) = \begin{bmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & -1 & & & & & \\ & & & -1 & & & & \\ & & & & 1 & & & \\ & 0 & & & & -1 & & \\ & & & & & & -1 & \\ & & & & & & & -1 \end{bmatrix} \quad C(2) = \begin{bmatrix} -1 & & & & & & & \\ & 1 & & & & & & \\ & & -1 & & & & & \\ & & & -1 & & & & \\ & & & & -1 & & & \\ & & & & & 1 & & \\ & 0 & & & & & -1 & \\ & & & & & & & -1 \\ & & & & & & & & 1 \end{bmatrix}$$

$$C(3) = \begin{bmatrix} 1 & & & & & \\ & -1 & & & & \\ & & 1 & & & 0 \\ & & & -1 & & \\ & & & & -1 & \\ 0 & & & & & 1 \\ & & & & & & 1 \\ & & & & & & & -1 \end{bmatrix}$$

There exists the curtailed matrix of order 3 for ${}^tU \cdot C(1)$ as shown in Table 5.3. Thus Absolute Inequality System 2.16 is consistent for $i=1$.

There does not exist any curtailed matrix for ${}^tU \cdot C(2)$. Thus construct the transformation matrix D_2 concerning the second row of ${}^tU \cdot C(2)$. There exists the curtailed matrix of order 2 for ${}^tU \cdot C(2) \cdot D_2$. Hence Absolute Inequality System 2.16 is consistent for $i=2$. These processes are shown in Table 5.4.

Since it is seen by the processes shown in Table 5.5 that Equation 2.21 has a nonzero nonnegative solution, Absolute Inequality System 2.16 is inconsistent for $i=3$. Therefore the behavior given by Fig. 5.1 is not realizable with three threshold elements.

$$t_{\overleftarrow{U}} \cdot \overleftarrow{C}(1) = \begin{pmatrix} 1 & 1 & -1 & -1 & 1 & -1 & -1 & -1 \\ 0 & 1 & 0 & 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & -1 & 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & -1 & 0 & -1 & -1 & -1 \end{pmatrix}$$

↓

$$t_{\overleftarrow{U}(3)} \cdot \overleftarrow{C}(1) = \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix}$$

↓

$$t_{\overleftarrow{U}(3,1)} \cdot \overleftarrow{C}(1) = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}$$

↓

$$t_{\overleftarrow{U}(3,2,1)} \cdot \overleftarrow{C}(1) = \begin{pmatrix} 1 \end{pmatrix}$$

Table 5.3. Simplification process by elimination

$$\begin{aligned}
 & \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & 0 & & & & \\ & 0 & & & 1 & & & \\ & & & & & 1 & & \\ 0 & 1 & 0 & 0 & 1 & 1 & & \\ & & & & & & 0 & 1 \\ & & 0 & & & & & 1 & 0 \end{pmatrix} = D_2 \quad \begin{pmatrix} -1 & 1 & -1 & -1 & 1 & -1 & -1 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 & -1 & -1 & 1 \end{pmatrix} = t_U \cdot c(2) \\
 & \quad \Downarrow \\
 & \quad \begin{pmatrix} -1 & 0 & -1 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 1 & -1 \\ 0 & -1 & 0 & -1 & -1 & 0 & -1 \end{pmatrix} = t_U \cdot c(2) \cdot D_2 \\
 & \quad \Downarrow \\
 & \quad \begin{pmatrix} 0 & 1 & 1 \\ -1 & -1 & 0 \end{pmatrix} = t_U \cdot \overleftarrow{c}(2) \cdot \overleftarrow{D}_2(1) \\
 & \quad \Downarrow \\
 & \quad \begin{pmatrix} +1 \end{pmatrix} = t_U \cdot \overleftarrow{c}(2) \cdot \overleftarrow{D}_2(1,3)
 \end{aligned}$$

Table 5.4. Simplification process by transformation and elimination

$$\begin{array}{l}
 \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & 0 & & 1 & & & \\ & & & & 1 & & \\ 0 & 1 & 0 & 0 & 1 & 1 & \\ & & 0 & & & 0 & 1 \\ & 0 & & & & 1 & 0 \end{pmatrix} & \begin{pmatrix} 1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 \\ 0 & -1 & 0 & 0 & -1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 0 & 1 & 1 & -1 \end{pmatrix} = t_U \cdot c(3) \\
 & = D_2 \quad \Downarrow \\
 & \begin{pmatrix} 1 & 0 & 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & -1 & 1 \\ 0 & 1 & 0 & -1 & 1 & 0 & 1 \end{pmatrix} = t_U \cdot c(3) \cdot D_2 \\
 \\
 \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ 0 & & 1 & & & & \\ 0 & 1 & 0 & 1 & 1 & 0 & \\ & & & 0 & 1 & 0 & \\ 0 & & & 0 & 0 & 1 & \\ & & 1 & 0 & 0 & & \end{pmatrix} & = D_4 \quad \Downarrow \\
 & \begin{pmatrix} 1 & -1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = t_U \cdot c(3) \cdot D_2 \cdot D_4 \\
 \\
 \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & & & & \\ & 1 & 1 & & 0 & \\ & & 0 & 1 & & \\ 0 & 1 & 0 & 1 & & \\ & 0 & 0 & 0 & 1 & \end{pmatrix} & = D_3 \quad \Downarrow \\
 & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = t_U \cdot c(3) \cdot D_2 \cdot D_4 \cdot D_3
 \end{array}$$

Table 5.5. Simplification process by transformation

CHAPTER VI

LINEAR SEPARABILITY BY DETERMINANTAL CONSISTENCY CONDITION

Concerning a system of inequalities many studies have been done from various aspects including the consistency condition. A system of strict inequalities has also been treated by many investigators such as W. B. Carver and T. Motzkin who studied the concept of an irreducibly inconsistent system. In the various proposed consistency conditions there exists such one that is grounded on the concept of "irreducibly inconsistency". The purpose of this chapter is to present the practical procedures to examine the linear separability based on this concept. This procedures consist of computation of the determinants of certain matrices.

6.1 Basic Principle

The method for the linear separability to be argued hereafter has two prominent features stated as follows.

- (1) The basic principle and the algorithm are very simple.
- (2) The condition for the linear separability can be written only in terms of the components of the characteristic matrix $C(i)$.

A system of inequalities is said to be irreducible inconsistent, if the system itself is inconsistent and if every

proper subsystem is consistent. Hence the system of linear inequalities

$$f_i(x) > \alpha_i \quad (i = 1, 2, \dots, m)$$

is irreducibly inconsistent, if and only if the following two conditions are simultaneously fulfilled.

- (1) Any $m-1$ of the linear functionals $f_1, f_2, \dots, f_m$ are linearly independent.
- (2) There exist m nonnegative numbers $\beta_i \geq 0$ ($i=1, 2, \dots, m$), not all zero, such that

$$\sum_{i=1}^m \beta_i \cdot f_i = 0 \quad , \quad \sum_{i=1}^m \beta_i \cdot \alpha_i \geq 0$$

The next theorem was established by Ky Fan which is the basic theorem for the subsequent argument.

(16)

Ky Fan's Theorem Let the rank of the matrix H be $m-1$.

Suppose that the first $m-1$ columns of H are linearly independent, and let

$$\tilde{H} = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1,m-1} \\ h_{21} & h_{22} & \dots & h_{2,m-1} \\ \cdot & \cdot & \dots & \cdot \\ h_{m1} & h_{m2} & \dots & h_{m,m-1} \end{pmatrix}$$

be submatrix of H formed by its first $m-1$ columns. For each

$i=1, 2, \dots, m$, let N_i denote the determinant of order $m-1$ obtained from $\tilde{H}$ by deleting its i -th row. Then the system of linear inequalities

$$H \cdot x > \alpha$$

is irreducibly inconsistent if and only if the following two conditions are both fulfilled.

$$N_i \cdot N_{i+1} < 0 \quad , \quad (i = 1, 2, \dots, m-1)$$

$$\sum_{i=1}^m \alpha_i \cdot |N_i| \geq 0$$

Clearly a system of linear inequalities is consistent, if and only if it does not contain any subsystem which is irreducibly inconsistent. Hence the next theorem and its corollary follow directly from Ky Fan's Theorem..

Theorem 6.1. Absolute Inequality System

$$C(i) \cdot U \cdot {}^t w_i > 0 \quad (2.16)$$

is inconsistent, if and only if certain q linearly dependent rows of the matrix $C(i) \cdot U$ contain a $(q, q-1)$ -type submatrix $\tilde{C}(i) \cdot \tilde{U}$ such that

$$N_k(i) \cdot N_{k+1}(i) < 0 \quad , \quad (k = 1, 2, \dots, q-1) \quad (6.1)$$

where $N_k(i)$ ($k=1, 2, \dots, q$) denotes the determinant of order $q-1$ obtained from $\tilde{C}(i) \cdot \tilde{U}$ by deleting its k -th row.

Corollary 6.1.1. Absolute Inequality System 2.16 is consistent if and only if there exists no submatrix such that Relation 6.1 is fulfilled.

6.2 Development

Theorem 6.1 states the inconsistency condition only as the property of the coefficient matrix of inequalities. Let us restate this theorem in the more developed form in order that it may be applied readily. Before advancing Theorem 6.2, consider the next definition.

Definition 6.1. Let $\tilde{U}(q)$ denote the $(q, q-1)$ -type matrix which consists of the q linearly dependent rows of the universal matrix U such that every $q-1$ rows of these q rows are independent. Hence the $q-1$ columns of $\tilde{U}(q)$ are linearly independent. Let us denote by $|\tilde{U}_k(q)|$ ($k=1, 2, \dots, q-1$) the determinant of the matrix obtained from $\tilde{U}(q)$ by deleting its k -th row.

Suppose that the rows of $\tilde{U}(q)$ correspond to the l_1 -th row, the l_2 -th row, $\dots$, and the l_q -th row of the universal matrix. Then the k -th row and the $k+1$ -th row of $\tilde{U}(q)$ become the content vector u_{l_k} and $u_{l_{k+1}}$ respectively. Hence the next proposition follows immediately.

Proposition 6.1. The determinants $N_k(i)$ and $N_{k+1}(i)$ defined in Theorem 6.1 are given by the relation

$$N_k(i) = \left\{ \prod_{\substack{j \neq 1_k \\ j \neq 1_{k+1}}} c_{1_j 1_j}(i) \right\} \cdot |\tilde{U}_k(q)|$$

$$N_{k+1}(i) = \left\{ \prod_{\substack{j \neq 1_{k+1} \\ j \neq 1_k}} c_{1_j 1_j}(i) \right\} \cdot |\tilde{U}_{k+1}(q)|$$

$$(k = 1, 2, \dots, q-1)$$

Theorem 6.2. Absolute Inequality System 2.16 has a solution if and only if there does not exist any submatrix $\tilde{U}(q)$ of the universal matrix U such that the condition

$$c_{1_k 1_k}(i) \cdot c_{1_{k+1} 1_{k+1}}(i) \cdot |\tilde{U}_k(q)| \cdot |\tilde{U}_{k+1}(q)| < 0 \quad (6.2)$$

$$(k = 1, 2, \dots, q-1)$$

is fulfilled.

Proof: The product of $N_k(i)$ and $N_{k+1}(i)$ can be written as follows.

$$N_k(i) \cdot N_{k+1}(i) = \left\{ \prod_{\substack{j \neq 1_k \\ j \neq 1_{k+1}}} c_{1_j 1_j}(i) \right\}^2 \cdot c_{1_k 1_k}(i) \cdot c_{1_{k+1} 1_{k+1}}(i) \cdot |\tilde{U}_k(q)| \cdot |\tilde{U}_{k+1}(q)|$$

Q.E.D.

In Relation 6.2 the part which will be influenced by a given transition state diagram is the product term $c_{1_k 1_k}(i) \cdot c_{1_{k+1} 1_{k+1}}(i)$. The product term $|\tilde{U}_k(q)| \cdot |\tilde{U}_{k+1}(q)|$ receives no influence and takes the fixed value notwithstanding the given

behavior. The number of matrices $\tilde{U}(q)$ and the values of the determinants $|\tilde{U}_k(q)|$ ($k=1, 2, \dots, q$) are fixed when the number of threshold elements of a network is once determined. If these matrices and the determinants have been preliminarily investigated in beforehand, the consistency of Absolute Inequality System 2.16 can be examined only by the sign of $c_{jj}(i)$.

The value of $c_{1_k 1_k}(i) \cdot c_{1_{k+1} 1_{k+1}}(i)$ takes 1 when the relations

$$\varphi_i(f \cdot \hat{u}_{1_k}) = 1$$

and

$$\varphi_i(f \cdot \hat{u}_{1_{k+1}}) = 1$$

hold simultaneously or when the relations

$$\varphi_i(f \cdot \hat{u}_{1_k}) = 0$$

and

$$\varphi_i(f \cdot \hat{u}_{1_{k+1}}) = 0$$

hold simultaneously.

On the other hand the value of $c_{1_k 1_k}(i) \cdot c_{1_{k+1} 1_{k+1}}(i)$ takes -1 when the relations

$$\varphi_i(f \cdot \hat{u}_{1_k}) = 1$$

and

$$\varphi_i(f \cdot \hat{u}_{1_{k+1}}) = 0$$

hold simultaneously or when the relations

$$\varphi_i(f \cdot \hat{u}_{1_k}) = 0$$

and

$$\varphi_i(f \cdot \hat{u}_{1_{k+1}}) = 1$$

hold simultaneously. Anyhow this value is immediately obtained when a transition state diagram is given.

6.3 Example

Example 6.1. Let us examine by Theorem 6.2 whether or not the behavior given by Fig. 4.3 is realizable with three threshold elements. In this case there are twelve matrices $\tilde{U}(q)$ which possess the properties prescribed in Definition 6.1 for $q=4$. These twelve matrices are given in Table 6.1 in the $(4,4)$ -type matrix rather than the $(4,3)$ -type matrix in order to show clearly the corresponding content vector. By this expression we can infer immediately what content vector each row of $U(q)$ consists of. The $(4,3)$ -type matrices required by Theorem 6.1 are constituted by excluding the column marked by "x" from each matrix.

$$\tilde{U}_1(4) = \begin{pmatrix} 1 & 0 & 0 & 0^* \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

$$\tilde{U}_2(4) = \begin{pmatrix} 1 & 0 & 0 & 0^* \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

$$\tilde{U}_3(4) = \begin{pmatrix} 1 & 0 & 0 & 0^* \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\tilde{U}_4(4) = \begin{pmatrix} 1 & 0 & 0 & 0^* \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

$$\tilde{U}_5(4) = \begin{pmatrix} 1 & 0 & 0 & 0^* \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\tilde{U}_6(4) = \begin{pmatrix} 1 & 0 & 0 & 0^* \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\tilde{U}_7(4) = \begin{pmatrix} 1 & 1 & 0 & 0^* \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}$$

$$\tilde{U}_8(4) = \begin{pmatrix} 1 & 0 & 1 & 0^* \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\tilde{U}_9(4) = \begin{pmatrix} 1 & 1 & 0 & 0^* \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\tilde{U}_{10}(4) = \begin{pmatrix} 1 & 0 & 1 & 0^* \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

$$\tilde{U}_{11}(4) = \begin{pmatrix} 1 & 0 & 1 & 0^* \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\tilde{U}_{12}(4) = \begin{pmatrix} 1 & 0 & 0 & 1^* \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

Table 6.1. The submatrices $\tilde{U}(q)$

By testing on these matrices it is easily observed that there is no submatrix $\tilde{U}(q)$ such that Relation 6.2 is fulfilled for $i=1$ and $i=2$. Hence Absolute Inequality System 2.16 is consistent for $i=1$ and $i=2$. However for $i=3$, the matrix $\tilde{U}_7(4)$ satisfies Relation 6.2. That is, for $\tilde{U}_7(4)$ each functional of Relation 6.2 takes the value

$$N_1(3) \cdot N_2(3) = -1 \cdot 1 \cdot -1 \cdot -1 = -1$$

$$N_2(3) \cdot N_3(3) = 1 \cdot 1 \cdot -1 \cdot 1 = -1$$

$$N_3(3) \cdot N_4(3) = 1 \cdot -1 \cdot 1 \cdot 1 = -1$$

Hence it is concluded that Absolute Inequality System 2.16 is inconsistent for $i=3$.

CHAPTER VII

SYNTHESIS PROCEDURE BY SUCCESSIVE SUPPLEMENT

The purpose of this chapter is to establish the procedures to realize any given transition state diagram by using more than N threshold elements in the case where the given transition state diagram is not realizable with N elements. The synthesis method to be discussed here consists of the procedures that are the successive supplements of threshold elements one by one until the given transition of states is realized. (55)

7.1 Basic Principle

In order to avail of threshold elements as the logic gates of digital systems, it is required that any given Boolean function can be realized by threshold gates. If a certain Boolean function is not realized by a single threshold element, this function has to be synthesized by more than one element.

Similarly, when a certain transition state diagram of the autonomous network is not realizable with N threshold elements since at least one of N elements does not satisfy the linear separability condition, we have to supplement some elements to achieve this transition. The supplemented threshold elements are called control elements henceforth.

In this case the transition is said to have been achieved

in the meaning that we are concerned only with the behaviors of the N threshold elements given from the beginning. These N threshold elements are called the primary elements so as to distinguish from the control elements. This interpretation corresponds to considering such an automaton where only the autonomous network given from the beginning is in contact with the external environment and we are concerned only with its inputs and outputs. On the other hand, the network supplemented later has no function except controlling the behavior of the network.

In this section the basic principle of how to supplement the control elements is discussed. Stiemke's and Tucker's Theorem becomes the fundamental theorem for our further developments.

Definition 7.1. Consider the replacements

$$\left. \begin{aligned} \vec{\hat{u}}_{j0}(0) &= \hat{u}_j \\ \vec{u}_{j0}(0) &= u_j \end{aligned} \right\} (j = 1, 2, \dots, 2^N)$$

$$\left. \begin{aligned} P_i(0) &= P_i \\ Q_i(0) &= Q_i \end{aligned} \right\} (i = 1, 2, \dots, N)$$

where the $\hat{u}_j$, u_j , P_i , and Q_i imply the internal state vector, the content vector, the positive set and the negative set defined in Chapter II, respectively.

Define the $(N+1)$ -dimensional vector $\vec{u}_{jr}(1)$ and the $(N+2)$ -dimensional vector $\vec{u}_{jr}(1)$ as follows.

$$\vec{u}_{jr}(1) = \begin{cases} (\vec{u}_{j0}(0), 1) & , \quad \text{if } \vec{u}_{j0}(0) \in P_r(0) \\ (\vec{u}_{j0}(0), 0) & , \quad \text{if } \vec{u}_{j0}(0) \in Q_r(0) \end{cases}$$

$$\vec{u}_{jr}(1) = (1, \vec{u}_{jr}(0)) \quad , \quad (j = 1, 2, \dots, 2^N).$$

Let $\vec{F}(1)$ denote the transition state diagram obtained by replacing $\vec{u}_j$ with $\vec{u}_{jr}(1)$ in the given transition state diagram F of the primary network.

Consider two sets $\vec{P}_i(1)$ and $\vec{Q}_i(1)$ of the vectors $\vec{u}_{jr}(1)$ defined as follows.

$$\begin{aligned} \vec{P}_i(1) &= \left\{ \vec{u}_{jr}(1) ; \quad \vec{\varphi}_i(\vec{f}(1) \cdot \vec{u}_{jr}(1)) = 1 \right\} \\ \vec{Q}_i(1) &= \left\{ \vec{u}_{jr}(1) ; \quad \vec{\varphi}_i(\vec{f}(1) \cdot \vec{u}_{jr}(1)) = 0 \right\} \\ &\quad (i = 1, 2, \dots, N+1) \end{aligned}$$

Here, the notation $\vec{f}(1) \cdot \vec{u}_{jr}(1)$ stands for the vector which is the successor of $\vec{u}_{jr}(1)$ in $\vec{F}(1)$ and $\vec{\varphi}_i$ implies the function defined over the set of vectors that takes as its value the i -th component of the vector.

Furthermore, define recursively the $(N+k)$ -dimensional vector $\vec{u}_{js}(k)$ and the $(N+k+1)$ -dimensional vector $\vec{u}_{js}(k)$ as follows.

$$\vec{u}_{jt}(k) = \begin{cases} (\vec{u}_{js}(k-1), 1) & , \quad \text{if } \vec{u}_{js}(k-1) \in \vec{P}_t(k-1) \\ (\vec{u}_{js}(k-1), 0) & , \quad \text{if } \vec{u}_{js}(k-1) \in \vec{Q}_t(k-1) \end{cases}$$

$$\vec{u}_{jt}(k) = (1, \vec{u}_{jt}(k)) \quad , \quad (j = 1, 2, \dots, 2^N)$$

The $\vec{u}_{jt}(k)$ and $\vec{u}_{jt}(k)$ are called the expanded internal state vector of order k concerning the threshold element T_t and the expanded content vector of order k concerning the threshold element T_t , respectively.

Let $\vec{F}(k)$ denote the transition state diagram obtained by replacing $\vec{u}_{js}(k-1)$ with $\vec{u}_{jt}(k)$ in the transition state diagram $\vec{F}(k-1)$. The $\vec{F}(k)$ is called the expanded transition state diagram of order k .

Moreover, consider two sets $\vec{P}_i(k)$ and $\vec{Q}_i(k)$ of the vectors $\vec{u}_{jt}(k)$ defined as follows.

$$\begin{aligned} \vec{P}_i(k) &= \left\{ \vec{u}_{jt}(k) \quad ; \quad \vec{\varphi}_i(\vec{f}(k) \cdot \vec{u}_{jt}(k)) = 1 \right\} \\ \vec{Q}_i(k) &= \left\{ \vec{u}_{jt}(k) \quad ; \quad \vec{\varphi}_i(\vec{f}(k) \cdot \vec{u}_{jt}(k)) = 0 \right\} \\ &\quad (i = 1, 2, \dots, N+k). \end{aligned}$$

Here, the notation $\vec{f}(k) \cdot \vec{u}_{jt}(k)$ stands for the successor vector just followed by $\vec{u}_{jt}(k)$ in $\vec{F}(k)$.

The next proposition is evident from the above definition.

Proposition 7.1. The transition of states in $\vec{F}(k)$ is completely same as the transition of states in $\vec{F}(k-1)$, if we

are concerned only with the behaviors of the elements $T_1, T_2, \dots, T_N, T_{N+1}, \dots, T_{N+k-1}$.

Therefore, the next proposition is an immediate consequence of this proposition.

Proposition 7.2.

$$\begin{aligned}\vec{P}_i(k) &= \vec{P}_i(k-1) \\ \vec{Q}_i(k) &= \vec{Q}_i(k-1)\end{aligned}\quad \text{for } i = 1, 2, \dots, N+k-1$$

Hence, only the sets $\vec{P}_{N+k}(k)$ and $\vec{Q}_{N+k}(k)$ are newly defined when $\vec{P}_i(k)$ and $\vec{Q}_i(k)$ are constituted.

Moreover, the following relations are clear from the definitions of $\vec{P}_i(k)$ and $\vec{Q}_i(k)$.

$$\begin{aligned}\vec{P}_i(k) \cup \vec{Q}_i(k) &= \nabla \\ \vec{P}_i(k) \cap \vec{Q}_i(k) &= \Phi\end{aligned}\quad \begin{aligned}\text{for } i &= 1, 2, \dots, N+k \\ k &= 1, 2, \dots\end{aligned}$$

Definition 7.2. Let $c_{jj}(N+i)$ denote the j -th row and the j -th column component of the characteristic matrix $C(N+i)$ of the control element T_{N+i} ($i = 1, 2, \dots, k$). Then $c_{jj}(N+i)$ takes the value given as follows.

$$c_{jj}(N+i) = \begin{cases} 1 & , \quad \text{for } \vec{\varphi}_{N+i}(\vec{f}(k) \cdot \vec{u}_j(k)) = 1 \\ -1 & , \quad \text{for } \vec{\varphi}_{N+i}(\vec{f}(k) \cdot \vec{u}_j(k)) = 0 \end{cases}$$

Definition 7.3. Consider the $(2^N, N+k+1)$ -type matrix $\vec{U}_i(k)$ whose j -th row equals to $\vec{u}_{ji}(k)$. This $\vec{U}_i(k)$ is called the universal matrix of order k concerning the element T_i .

$$\vec{U}_i(k) = \begin{pmatrix} \vec{u}_{1i}(k) \\ \vec{u}_{2i}(k) \\ \vec{u}_{3i}(k) \\ \vdots \\ \vec{u}_{2^{N_i}i}(k) \end{pmatrix}$$

Definition 7.4. Consider the $(N+k+1)$ -dimensional vector $\vec{w}_i(k)$ defined as follows.

$$\vec{w}_i(k) = (w_i, w_{iN+1}, w_{iN+2}, \dots, w_{iN+k})$$

This $\vec{w}_i(k)$ is called the expanded threshold-weight vector of order k .

Theorem 7.1. If there exist certain real numbers j and k such that the inequality systems

$$C(i) \cdot \vec{U}_j(k) \cdot \vec{w}_i(k) > 0 \quad (i = 1, 2, \dots, N+k)$$

are consistent for every i , then the network supplemented with k control elements realizes the transition of states defined by $\vec{F}(k)$.

Theorem 7.2. The equations

$$t_{U_1}^{\rightarrow}(k) \cdot C(i) \cdot t_{b_1}^{\rightarrow}(k) = 0 \quad (7.1)$$

have no nonzero nonnegative solution $\vec{b}_1(k) \geq 0$

Proof: Suppose that $v = (v_1, v_2, \dots, v_{2N})$ denotes the $(N+k+1)$ -th row vector of $t_{U_1}^{\rightarrow}(k) \cdot C(i)$. Then the following relation holds.

$$v_j = \begin{cases} 1 & , \quad \text{for } \vec{u}_{ji},(k-1) \in \vec{P}_1(k-1) \\ 0 & , \quad \text{for } \vec{u}_{ji},(k-1) \in \vec{Q}_1(k-1) \end{cases}$$

Therefore, there exists no solution $\vec{b}_1(k) = (b_{11}(k), b_{12}(k), \dots, b_{12N}(k)) \geq 0$ of Equation 7.1 such that all the j -th components $b_{1j}(k)$ satisfying $\vec{u}_{ji},(k-1) \in \vec{P}_1(k-1)$ are nonzero nonnegative. Hence, we can regard all the j -th components $b_{1j}(k)$ as zeros.

On the other hand, all the j -th components of the first row vector of $t_{U_1}^{\rightarrow}(k) \cdot C(i)$ satisfying $\vec{u}_{ji},(k-1) \in \vec{Q}_1(k-1)$ are -1. Therefore, there exists no solution $\vec{b}_1(k) \geq 0$ such that all the j -th components $b_{1j}(k)$ satisfying $\vec{u}_{ji},(k-1) \in \vec{Q}_1(k-1)$ are nonzero nonnegative.

Consequently, if we assume that Equation System 7.1 has a nonzero nonnegative solution, it contradicts either the first equation or the $(N+k+1)$ -th equation. Q.E.D.

Theorem 7.3. The inequality system

$$C(i) \cdot \vec{U}_1(k) \cdot t_{w_1}^{\rightarrow}(k) > 0$$

is always consistent regardless of the value k .

Theorem 7.3 guarantees that if the inequality system

$$C(i) \cdot \vec{U}_h(k-1) \cdot \vec{t}_{w_i}(k-1) > 0$$

is inconsistent, then this system can be made consistent by supplementing one control element T_{N+k} . This supplement gives rise to the construction of the expanded universal matrix $\vec{U}_1(k)$ concerning the element T_i .

Theorem 7.4. If the inequality systems

$$C(i) \cdot \vec{U}_\mu(k) \cdot \vec{t}_{w_i}(k) > 0 \quad (7.2)$$

$$(k = 0, 1, 2, \dots, n-1)$$

are consistent for at least one value of the order $k \leq n-1$, then the inequality system

$$C(i) \cdot \vec{U}_j(n) \cdot \vec{t}_{w_i}(n) > 0 \quad (7.3)$$

becomes always consistent.

Proof: Suppose that Inequality System 7.2 is consistent for $k=m < n$, then the equations

$$\vec{t}_{U_\eta}(m) \cdot C(i) \cdot \vec{t}_{b_i}(m) = 0 \quad (7.4)$$

have no nonzero nonnegative solution $b_i(m) \geq 0$. On the other hand, the first $N+m+1$ equations of the equation system

$$\vec{t}_{U_j}(n) \cdot C(i) \cdot \vec{t}_{b_i}(n) = 0 \quad (7.5)$$

are just same as Equation System 7.4 from the definition of $\vec{U}_j(n)$. Consequently, Equation System 7.5 has no nonzero non-negative solution $\vec{b}_i(n) \geq 0$. Q.E.D.

Theorem 7.4 guarantees that the consistency of the inequalities is not changed, even if some control elements are supplemented to the network.

Theorem 7.5. Assume that the expanded content vector $\vec{u}_{ji}(k)$ is generated if and only if the inequalities

$$C(i) \cdot \vec{U}_h(k-1) \cdot \vec{w}_i(k-1) > 0 \quad (7.6)$$

have no solution $\vec{w}_i(k-1)$. Then the rank of the expanded universal matrix $\vec{U}_i(k)$ becomes $N+k+1$.

Proof: The rank of the expanded universal matrix of order 1 is always $N+2$. If the rank of $\vec{U}_j(m)$ is $N+m+1$, then the rank of $\vec{U}_i(m+1)$ becomes $N+m+2$. Q.E.D.

Theorem 7.6. The inequality system

$$C(i) \cdot \vec{U}_j(2^N - N - 1) \cdot \vec{w}_i(2^N - N - 1) > 0 \quad (7.7)$$

$$(i = 1, 2, \dots, 2^N - 1)$$

is always consistent for every i .

Proof: Since the rank of $C(i) \cdot \vec{U}_j(2^N - N - 1)$ becomes 2^N from Theorem 7.5, all the row vectors of this coefficient matrix are linearly independent. Q.E.D.

Theorem 7.6 guarantees that the synthesis procedures so far discussed terminate at the $(2^N - N - 1)$ -th step without fail.

7.2 Practical Procedure for Synthesis

Since the basic principle to realize any given transition state diagram has been provided in the preceding section, here let us set forth the practical procedures in order.

- (1) Examine first of all whether or not Inequality System 2.16 is consistent for every characteristic matrix $C(i)$ determined by Relation 2.3. If it has turned out to be consistent for every $C(i)$, then determine the actual threshold-weight vector by Theorem 3.1.
- (2) If Inequality System 2.16 has turned out to be inconsistent for a certain $C(h)$, then supplement one control element T_{N+1} and make it consistent. This can be performed by constructing the expanded universal matrix $\vec{U}_h(1)$ in Theorem 7.3.
- (3) Moreover, if Inequality System 2.16 also turned out to be inconsistent for another $C(\eta)$, then supplement another control element T_{N+2} and make it consistent. This is possible by considering $\vec{U}_\eta(2)$ in Theorem 7.3.
- (4) Similarly, if there exist other inconsistent Inequality Systems 2.16, supplement control elements and make them consistent.
- (5) Even if control elements are supplemented to the network, such Inequality System 2.16 is still con-

sistent as is originally consistent. That is, even if the procedures mentioned in Step (1), (2), (3) are performed, the inequality system

$$C(i) \cdot \vec{U}_j(k) \cdot \vec{w}_i(k) > 0$$

is still consistent regardless of k and j , where $C(i)$ represents a threshold function.

Likewise, if a certain Inequality System 2.16 has been made consistent once, then further supplements of control elements do not change the consistency of this system from Theorem 7.4. Hence there is no necessity to deal with the primary elements again after the procedures stated (1), (2), (3) are performed once.

- (6) Now it is sufficient to consider only the control elements. Examine whether or not the inequality system for the control element T_{N+i}

$$C(N+i) \cdot \vec{U}_j(k) \cdot \vec{w}_{N+i}(k) > 0 \quad (7.8)$$

$$(i = 1, 2, \dots, k)$$

is consistent. If Inequality System 7.8 has turned out to be consistent for every i , then the given transition state diagram is realizable with the primary elements and the $N+k$ control elements so far supplemented.

- (7) On the other hand, if Inequality System 7.8 has turned out to be inconsistent for $C(N+\eta)$, supplement one more control element T_{N+k+1} . Then this inconsistent system becomes

$$C(N+\eta) \cdot \vec{U}_{N+\eta}(k+1) \cdot \vec{w}_{N+\eta}^t(k+1) > 0.$$

Thus, this inconsistent system is made consistent from Theorem 7.3.

- (8) Moreover, examine whether or not the inequality system for the control element T_{N+k+1} supplemented in Step (7)

$$C(N+k+1) \cdot \vec{U}_{N+\eta}(k+1) \cdot \vec{w}_{N+k+1}^t(k+1) > 0 \quad (7.9)$$

is consistent. If this system has turned out to be consistent, then search for the inconsistent system of Inequality System 7.8. Then, perform the similar procedure stated in Step (7).

- (9) If Inequality System 7.9 is inconsistent, then supplement one more control element T_{N+k+2} . Furthermore, examine whether or not the inequality system

$$C(N+k+2) \cdot \vec{U}_{N+k+1}(k+2) \cdot \vec{w}_{N+k+2}^t(k+2) > 0 \quad (7.10)$$

is consistent. Repeat the similar procedures.

- (10) The supplement of $2^N - N - 1$ control elements terminates these procedures at worst case.

7.3 Example

Example 7.1. Let us realize the transition state diagram shown in Fig. 4.3 by the successive supplement of control elements. It has already turned out that Inequality System 2.16 is inconsistent only for $C(3)$ through the argument mentioned in Example 4.2.

Hence, supplement one control element T_4 and construct the expanded content vector of order 1 concerning element T_3 as follows.

$$\vec{u}_{13}(1) = 1 \ 0 \ 0 \ 0 \ 0$$

$$\vec{u}_{23}(1) = 1 \ 1 \ 0 \ 0 \ 0$$

$$\vec{u}_{33}(1) = 1 \ 0 \ 1 \ 0 \ 1$$

$$\vec{u}_{43}(1) = 1 \ 0 \ 0 \ 1 \ 0$$

$$\vec{u}_{53}(1) = 1 \ 1 \ 1 \ 0 \ 1$$

$$\vec{u}_{63}(1) = 1 \ 1 \ 0 \ 1 \ 1$$

$$\vec{u}_{73}(1) = 1 \ 0 \ 1 \ 1 \ 0$$

$$\vec{u}_{83}(1) = 1 \ 1 \ 1 \ 1 \ 0$$

Thus, we obtain the expanded transition state diagram $\vec{F}(1)$ as follows.

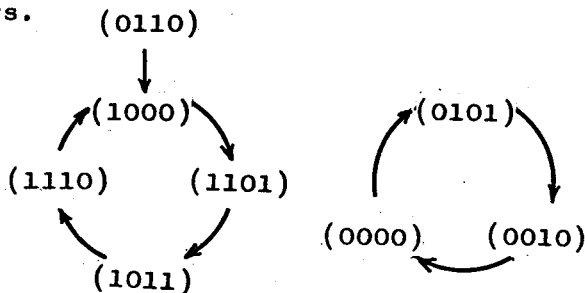


Fig. 7.1.
Transition state
diagram $\vec{F}(1)$

Therefore, the characteristic matrix $C(4)$ for the control element T_4 is given as follows

$$C(4) = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & -1 & & \\ & & & -1 & \\ & 0 & & & 1 \\ & & & & & -1 \\ & & & & & & -1 \end{pmatrix}$$

The inequality system

$$C(i) \cdot \vec{U}_3(1) \cdot {}^t\vec{w}_i(1) > 0$$

becomes consistent for $C(1)$, $C(2)$ and $C(3)$ through Theorem 7.3 and Theorem 7.4. Now, examine the consistency of the inequality system

$$C(4) \cdot \vec{U}_3(1) \cdot {}^t\vec{w}_4(1) > 0$$

Since this system turns out to be consistent, it is concluded that the transition state diagram shown in Fig. 4.3 is realized with one control element.

The expanded threshold-weight vectors $\vec{w}_i(1)$ ($i = 1, 2, 3, 4$) are obtained through the similar computation stated in Chapter III, for instance, as follows.

$$\vec{w}_1(1) = (-4, 8, 5, 3, -5)$$

$$\vec{w}_2(1) = (1, 0, -4, -2, 2)$$

$$\vec{w}_3(1) = (-1, 0, 0, 0, 2)$$

$$\vec{w}_4(1) = (1, 2, -1, -3, -1)$$

Consequently, the autonomous network shown in Fig. 7.2 realizes the transition state diagram given Fig. 4.3, if the transits of the elements T_1 , T_2 , T_3 are observed.

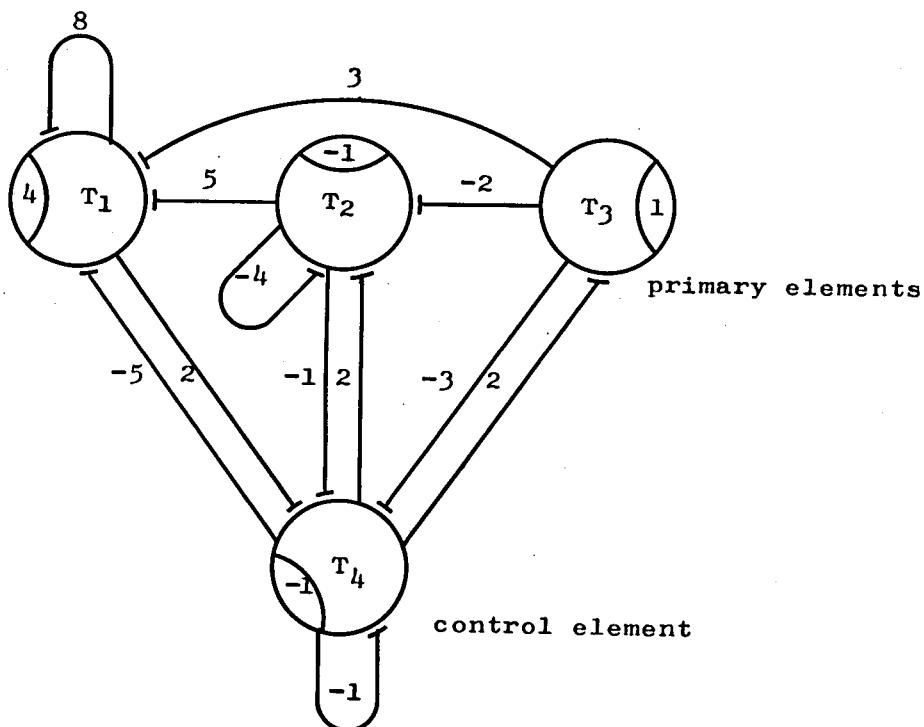


Fig. 7.2. The autonomous network

CHAPTER VIII

GENERAL SYNTHESIS PROCEDURE

The general method to realize any given transition of states is treated in this chapter, although concerning this problem an effective method has been already given in Chapter VII. The method to be discussed here is very intuitive and systematic, it is, however, forced to use much more control elements in compensation for its advantages. In fact, it requires $M-N-1$ control elements, where M denotes the number of states specified their successor states.

8.1 Basic Principle

The method presented here is in actual applicable to the incompletely specified transition state diagram, since the number of the necessary control elements becomes very large for the completely one. For the generality, however, the argument will be stated with respect to the latter case.

Consider the autonomous network which is constructed with N threshold elements. This network is called the primary network. The primary network has $(2^N)^{(2^N)}$ distinct transitions.

However some of them are realizable and others are not realizable by the primary network. If a certain given transition is not realizable, then supplement $2^N - N - 1$ threshold elements to the primary network. ⁽²⁴⁾ The network supplemented with $2^N - N - 1$ elements is referred to as the general network. Suppose, however, that only the inputs and the outputs of the primary network are observed from the outsides. Then let us show that the general network can realize all the $(2^N)^{(2^N)}$ transitions by adjusting the threshold values and the weights values. In this chapter the argument will be done for Inequality Systems 2.14 rather than Absolute Inequality Systems 2.19.

Definition 8.1. Let us define the $(2^N - 1)$ -dimensional vector $Ex(\hat{u}_i)$ by extending the internal state vector $\hat{u}_i$ as follows.

$$Ex(\hat{u}_i) = \begin{cases} (\hat{u}_i, \overbrace{0, 0, 0, \dots, 0}^{N - N - 1 \text{ zeros}}) & \text{for } i = 1, 2, \dots, N+1 \\ (\hat{u}_i, \overbrace{0, 0, \dots, 0, 1, 0, \dots, 0}^{(i-N-1)\text{-th position}}) & \text{for } i = N+2, N+3, \dots, 2^N \end{cases}$$

This vector is called the extended internal state vector.

Furthermore consider the 2^N -dimensional vector $Ex(u_i)$ attached 1 as the first component to $Ex(\hat{u}_i)$ as follows.

$$Ex(u_i) = (1, Ex(\hat{u}_i)), \text{ for } i = 1, 2, \dots, N+1, \dots, 2^N$$

This vector is called the extended content vector.

Definition 8.2. Consider the $(2^N, 2^N)$ -type matrix $\text{Ex}(U)$ whose i -th row equals to the extended content vector $\text{Ex}(u_i)$. This square matrix is called the extended universal matrix.

$$\text{Ex}(U) = \begin{bmatrix} U_I & 0 \\ U_{II} & E_{2^N-2^N-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & & & \\ 1 & 1 & 0 & 0 & \dots & 0 & & & \\ 1 & 0 & 1 & 0 & \dots & 0 & & & \\ 1 & 0 & 0 & 1 & \dots & 0 & & & \\ 1 & 1 & 1 & 0 & \dots & 0 & 1 & & \\ 1 & 1 & 0 & 1 & \dots & 0 & 1 & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \cdot & \\ 1 & 1 & 1 & 1 & \dots & 1 & & & 1 \end{bmatrix}$$

Definition 8.3. Consider the 2^N -dimensional vector $\text{Ex}(w_i)$ attached 2^N-2^N-1 weights values to the threshold-weight vector w_i as follows

$$\text{Ex}(w_i) = (w_i, w_{iN+1}, w_{iN+2}, \dots, w_{i2^N-1})$$

This vector is called the extended threshold-weight vector.

Definition 8.4. Let $\text{Ex}(F)$ denote the transition state diagram which is obtained by replacing $\hat{u}_i$ with $\text{Ex}(\hat{u}_i)$ for the given transition state diagram F .

Therefore $\text{Ex}(F)$ becomes an incompletely specified transition state diagram for the general network.

Then the next proposition is almost obvious from Defini-

tion 8.1 and Definition 8.4.

Proposition 8.1. If we take notice only the primary network, then the transition of states in $Ex(F)$ is completely same as the transition of states in F .

Definition 8.5. Let us define the characteristic matrices $C(i)$ for the control threshold elements $T_{N+1}, T_{N+2}, \dots, T_{2N-1}$ as follows.

$$c_{jj}(i) = \begin{cases} 1 \dots \text{for} & Ex(\hat{u}_{i+1}) = Ex(f) \cdot Ex(\hat{u}_j) \\ -1 \dots \text{for} & Ex(\hat{u}_{i+1}) \neq Ex(f) \cdot Ex(\hat{u}_j) \end{cases}$$

Here $Ex(\hat{u}_{i+1}) = Ex(f) \cdot Ex(\hat{u}_j)$ implies that $Ex(\hat{u}_j)$ is just followed by $Ex(\hat{u}_{i+1})$ in the extended transition state diagram $Ex(F)$.

Now consider the inequalities

$$c_{jj}(i) \cdot Ex(u_j) \cdot {}^tEx(w_i) \geq 0, \quad (j = 1, 2, \dots, 2^N)$$

where the relation " $\geq$ " holds for j satisfying $c_{jj}(i)=-1$ and the relation ">" holds for j satisfying $c_{jj}(i)=1$. Let us represent such defined inequalities by

$$C(i) \cdot Ex(U) \cdot {}^tEx(w_i) \geq 0 \quad (8.1)$$

$$(i = 1, 2, \dots, 2^{N-1})$$

Theorem 8.1. If Inequality System 8.1 is consistent for

every i , then the primary network performs the transition given by F in the general network.

Proof: If Inequality System 8.1 is consistent for every i , then the general network performs the transition given by $Ex(F)$. Q.E.D.

Proposition 8.2. The set of all row vectors of $Ex(U)$ is linearly independent.

Theorem 8.2. Inequality System 8.1 is consistent for every i .

Proof: It is sufficient to show that the set of all row vectors of $C(i) \cdot Ex(U)$ is linearly independent. In fact, the following relations hold.

$$\begin{aligned} \det \{ C(i) \cdot Ex(U) \} &= \det C(i) \cdot \det Ex(U) \\ &= \det C(i) \cdot \det U_I \cdot \det E_{2N-N-1} = (-1)^{|Q_i|} \end{aligned}$$

where $|Q_i|$ denotes the number of components contained in the negative set Q_i .

8.2 Weights and Threshold Values

The method how the general network realizes any transition of the primary network has become evident through the preceding discussion. In this section let us discuss the practical threshold values and weights values of the general network which realize the given transition.

Proposition 8.3. The inverse of the regular matrix

$C(i) \cdot \text{Ex}(U)$ is given as follows.

$$\begin{aligned} [C(i) \cdot \text{Ex}(U)]^{-1} &= \begin{pmatrix} C_I(i) \cdot U_I & 0 \\ C_{II}(i) \cdot U_{II} & C_{II}(i) \end{pmatrix}^{-1} \\ &= \begin{pmatrix} U_I^{-1} \cdot C_I(i) & 0 \\ -U_{II} \cdot U_I^{-1} \cdot C_I(i) & C_{II}(i) \end{pmatrix} \end{aligned} \quad (8.2)$$

Here the inverse U_I^{-1} is given by Proposition 2.7. Hence the troublesome computation which is ordinarily brought by the inverse calculation is excluded. The inverse of $C(i) \cdot \text{Ex}(U)$ can be readily obtained only by the multiplication of matrices.

Definition 8.6. Let ${}^t r_j(i)$ denote the j -th column vector of $[C(i) \cdot \text{Ex}(U)]^{-1}$.

Theorem 8.3. The whole solutions of Inequality System 8.1

$$C(i) \cdot \text{Ex}(U) \cdot {}^t \text{Ex}(w_i) \geq 0 \quad (i = 1, 2, \dots, 2^{N-1}) \quad (8.1)$$

are given by the following form.

$$\text{Ex}(w_i) = \sum_{j=1}^{2^N} \alpha_j(i) \cdot {}^t r_j(i) \quad (i = 1, 2, \dots, 2^{N-1}) \quad (8.3)$$

where $\alpha_j(i)$ is either an arbitrary positive number for $c_{jj}(i)=1$ or an arbitrary nonnegative number for $c_{jj}(i) = -1$.

Proof: Since the following relation

$$C(i) \cdot \text{Ex}(U) \cdot \sum_{j=1}^{2^N} \alpha_j(i) \cdot {}^t r_j(i) = {}^t (\alpha_1(i), \alpha_2(i), \dots, \alpha_{2^N}(i))$$

holds, $\text{Ex}(w_i)$ is surely a solution of Inequality System 8.1. Conversely, let y denote an arbitrary solution of Inequality System 8.1, then the relation

$$C(i) \cdot \text{Ex}(U) \cdot y \geq 0$$

holds. Hence we can rewrite by using the j -th unit column vector e_j as follows.

$$C(i) \cdot \text{Ex}(U) \cdot y = \sum_{j=1}^N \alpha_j(i) \cdot e_j$$

Therefore y can be expressed as follows.

$$\begin{aligned} y &= [C(i) \cdot \text{Ex}(U)]^{-1} \cdot \sum_{j=1}^{2^N} \alpha_j(i) \cdot e_j \\ &= \sum_{j=1}^{2^N} \alpha_j(i) \cdot {}^t r_j(i) \end{aligned}$$

Q.E.D.

It is not necessary to assign the internal state vector of the general network as given by Definition 8.1. Any assignment will do only if it makes the rank of the coefficient matrix of Inequality System 2.14

$$C(i) \cdot U \cdot w_i^t \geq 0$$

2^N . Thus there are many assignments to satisfy this requirement. However such an assignment is desirable as makes the further analysis easier. For instance, in order to determine the practical threshold values and weights values, such one is suitable that we can obtain the inverse of $[C(i) \cdot Ex(U)]^{-1}$ as easily as possible. In fact the inverse is given by Proposition 8.3, if we define the extended internal state vector by Definition 8.1.

This method requires $2^N - N - 1$ control elements for the completely specified transition. The number $2^N - N - 1$ becomes tremendous large if N increases slightly. Thus it is not necessarily applicable to this case. This method is rather appropriate for the incompletely specified transition where the number of states specified their following states is small.

One of the most advantageous points of this method is the fact that the practical threshold values and weights values can be determined very easily by Theorem 8.3.

There is a means to decrease the number $2^N - N - 1$ of control elements stated as follows. That is, at first we realize the

given transition by the general network. After that we search for the unnecessary control elements and exclude them.

8.3 Example

Example 8.1. The transition state diagram shown by Fig. 4.3 is not realizable with three threshold elements. Hence, supplement four control elements and consider the extended internal state vectors given as follows.

$$\text{Ex}(u_1) = (1 0 0 0 0 0 0 0)$$

$$\text{Ex}(u_2) = (1 1 0 0 0 0 0 0)$$

$$\text{Ex}(u_3) = (1 0 1 0 0 0 0 0)$$

$$\text{Ex}(u_4) = (1 0 0 1 0 0 0 0)$$

$$\text{Ex}(u_5) = (1 1 1 0 1 0 0 0)$$

$$\text{Ex}(u_6) = (1 1 0 1 0 1 0 0)$$

$$\text{Ex}(u_7) = (1 0 1 1 0 0 1 0)$$

$$\text{Ex}(u_8) = (1 1 1 1 0 0 0 1)$$

Then the characteristic matrices for this transition are given as follows.

$$C(1) = \begin{bmatrix} -1 & & & & & & & \\ & 1 & & & & & & \\ & & -1 & & & & & \\ & & & -1 & & & & \\ & & & & 1 & & & \\ & 0 & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 1 \end{bmatrix} \quad C(2) = \begin{bmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & -1 & & & & & \\ & & & -1 & & & & \\ & & & & -1 & & & \\ & 0 & & & & 1 & & \\ & & & & & & -1 & \\ & & & & & & & -1 \end{bmatrix}$$

$$c(3) = \begin{bmatrix} -1 & & & & & & & \\ & -1 & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & 1 & & & \\ & 0 & & & & 1 & & \\ & & & & & & -1 & \\ & & & & & & & -1 \end{bmatrix}$$

$$c(4) = \begin{bmatrix} -1 & & & & & & & \\ & 1 & & & & & & \\ & & -1 & & & & & \\ & & & -1 & & & & \\ & & & & -1 & & & \\ & 0 & & & & -1 & & \\ & & & & & & -1 & \\ & & & & & & & -1 \end{bmatrix}$$

$$c(5) = \begin{bmatrix} -1 & & & & & & & \\ & -1 & & & & & & \\ & & -1 & & & & & \\ & & & -1 & & & & \\ & & & & 1 & & & \\ & 0 & & & & -1 & & \\ & & & & & & -1 & \\ & & & & & & & -1 \end{bmatrix}$$

$$c(6) = \begin{bmatrix} -1 & & & & & & & \\ & -1 & & & & & & \\ & & -1 & & & & & \\ & & & -1 & & & & \\ & & & & -1 & & & \\ & 0 & & & & -1 & & \\ & & & & & & -1 & \\ & & & & & & & -1 \end{bmatrix}$$

$$c(7) = \begin{bmatrix} -1 & & & & & & & \\ & -1 & & & & & & \\ & & -1 & & & & & \\ & & & -1 & & & & \\ & & & & -1 & & & \\ & & & & & 1 & & \\ & 0 & & & & & -1 & \\ & & & & & & & -1 \\ & & & & & & & & -1 \end{bmatrix}$$

Now determine the extended threshold-weight vector by Relation 8.3. The following extended threshold-weight vectors are obtained by fixing $\alpha_j(i)=1$ ($j=1, 2, \dots, 8, i=1, 2, \dots, 7$).

$$Ex(w_1) = (-1 \ 2 \ 0 \ 0 \ 0 \ 0 \ 2 \ 0)$$

$$Ex(w_2) = (1 \ 0 \ -2 \ -2 \ 0 \ 2 \ 2 \ 2)$$

$$Ex(w_3) = (-1 \ 0 \ 2 \ 0 \ 0 \ 2 \ -2 \ -2)$$

$$Ex(w_4) = (-1 \ 2 \ 0 \ 0 \ -2 \ -2 \ 0 \ -2)$$

$$Ex(w_5) = (-1 \ 0 \ 0 \ 0 \ 2 \ 0 \ 0 \ 0)$$

$$\text{Ex}(\omega_6) = (-1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0)$$

$$\text{Ex}(\omega_7) = (-1 \ 0 \ 0 \ 0 \ 0 \ 2 \ 0 \ 0)$$

Therefore the autonomous network shown by Fig. 8.1 performs the behavior given by Fig. 4.3, if we observe only the primary network.

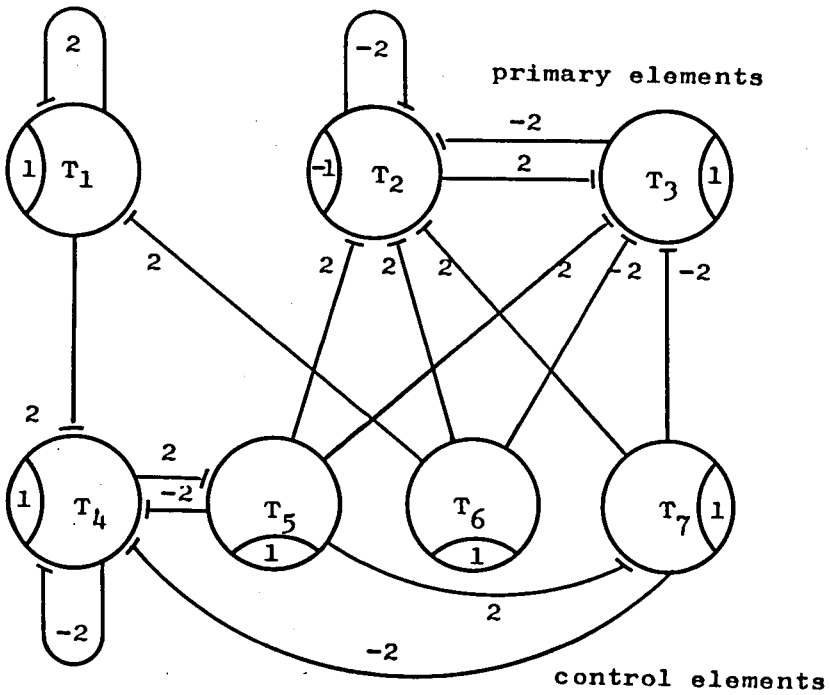


Fig. 8.1. The autonomous network

CHAPTER IX

MULTIGATE SYNTHESIS OF BOOLEAN FUNCTION

The problem of synthesizing threshold-element networks to realize any given Boolean function⁽⁴⁴⁾ is treated in this chapter. This problem has not received as much attention as the problem of a single threshold element. This certainly depends on the fact that sufficient knowledges of the latter seem to be essential for the successful solution of the former.

Although the network resulting from the synthesis procedures discussed here has a similar form to that obtained by Threshold-Or network synthesis or Threshold-Cascade network synthesis^{(14) (59)} so far proposed, the procedures to be presented are very straightforward. However, they do not give the most economic synthesis in the meaning of a network constructed with the least number of threshold gates.

9.1 Basic Principle

With a view to designing more efficient digital systems, it is desirable to obtain a most economic network of threshold gates realizing a given Boolean function. This leads us to the problem of finding for a given Boolean function f a minimal decomposition of f with threshold functions.⁽³⁵⁾

For this purpose, the classical network synthesis techni-

(6) (50)
 ques, such as the Quine-McCluskey procedure for AND-OR synthesis play an important role. If a minimal decomposition is obtained, the subsequent procedure is straightforward. Hence, the central subject consists in how to make the minimal decomposition.

As in many synthesis problems, the fact that a large number of the configurations can be employed for a general network entails many admissible decomposition forms. This often makes the synthesis procedure complex. In view of the complexity imposed by the unrestricted network, it is not necessarily unreasonable to postulate the network configurations. Hereafter, let us assume that the network consists of the two levels, an input level and an output level which is an OR gate.

If all the ONES of a Boolean function f_i are covered with $f_{i1}, f_{i2}, \dots, f_{im}$ and if there exists no other ONE covered with these functions, then f_i is said to be decomposed into $f_{i1}, f_{i2}, \dots, f_{im}$. Here, the connotation of "cover" implies the concept used in AND-OR synthesis.

As M. L. Dertouzos and other investigators have already pointed out, if a given Boolean function f_i can be decomposed into m threshold functions, f_i is realized with the two-level network whose m input level elements produce these m threshold functions and one output level element produces an OR function. Since an OR function is a threshold function, regardless of the number of input variables, these facts are restated in the fol-

lowing theorem which is the foundation of the practical procedures mentioned in the subsequent section.

Theorem 9.1. If a given Boolean function f_i can be decomposed into m threshold functions $f_{i1}, f_{i2}, \dots, f_{im}$, then f_i can be synthesized with the two-level network constituted by $m+1$ threshold gates.

However, the threshold element representing the output OR gate is not needed, since any two-level Threshold-Or network is equivalent to a cascade of the input level elements. This fact can be understood easily by considering the case where the weights values associated with all interconnecting leads are made sufficiently large to permit the propagation of a ONE generated by any threshold element toward the network output. (14)
These statements are formalized in the next theorem.

Theorem 9.2. If a given Boolean function f_i can be decomposed into m threshold functions $f_{i1}, f_{i2}, \dots, f_{im}$, then f_i can be synthesized with the Cascade-Network constituted by m threshold gates.

9.2 Synthesis Procedure

The practical procedures to synthesize any Boolean function are given without justification in this section. However, its justification will immediately follow from Theorem 9.1 and Theorem 9.2. Let us examine by the method presented in Chapter V or by any other available test whether or not a given Boolean

function is a threshold function. It is, of course, more immediate to consult for the same purpose the Table which lists up all the threshold functions of less than 8 variables.

The characteristic matrix $C(i)$ represents the Boolean function f_i as stated in Proposition 2.3. Hence, $C(i)$ is used in some case as the matrix and in another case as the Boolean function. However, it won't any confusion from the context.

Then, any given Boolean function $C(i)$ is synthesized by the following finite steps.

- (1) Examine whether or not $C(i)$ is a threshold function.

If $C(i)$ is a threshold function, then determine the actual weights values and the threshold value by Theorem 3.1. If $C(i)$ is not a threshold function, proceed to Step 2.

- (2) Let V_j denote the set given by Definition 5.1, and V_j^c denote the complement set of V_j . The notation $|X|$ implies the number of components contained in the set X .

Now, search for the suffix k_1 of the input variable such that

$$\max_j \left\{ |P_i \cap V_j| \times |P_i \cap V_j^c| \right\} = |P_i \cap V_{k_1}| \times |P_i \cap V_{k_1}^c|. \quad (9.1)$$

If there exist more than one suffixs satisfying Relation 9.1, select an arbitrary one of them. Moreover, construct two characteristic matrices $C(i; k_1^+)$ and $C(i; k_1^-)$ defined as follows.

$$c_{jj}(i; k_1^+) = \begin{cases} 1 & , \text{ for } \hat{u}_j \in (P_i \cap V_{k_1}) \\ -1 & , \text{ for } \hat{u}_j \notin (P_i \cap V_{k_1}) \end{cases}$$

and

$$c_{jj}(i; k_1^-) = \begin{cases} 1 & , \text{ for } \hat{u}_j \in (P_i \cap V_{k_1}^c) \\ -1 & , \text{ for } \hat{u}_j \notin (P_i \cap V_{k_1}^c) \end{cases}$$

where $c_{jj}(i; k_1^+)$ and $c_{jj}(i; k_1^-)$ denote the j -th row and j -th column of $C(i; k_1^+)$ and $C(i; k_1^-)$, respectively.

Examine whether or not $C(i; k_1^+)$ and $C(i; k_1^-)$ are threshold functions. If both of them are threshold functions, then $C(i)$ can be synthesized with the Cascade-Network of two elements. Determine the actual weights values and threshold value through Theorem 3.1 for $C(i; k_1^+)$ and $C(i; k_1^-)$. If only $C(i; k_1^-)$ is a threshold function, proceed to Step 3. On the other hand, if only $C(i; k_1^+)$ is a threshold function, proceed to Step 4. If neither $C(i; k_1^+)$ nor $C(i; k_1^-)$ is a threshold function, proceed to Step 5.

- (3) Search for the suffix k_2 of the input variable such that

$$\begin{aligned} & \max_j \left\{ |P_i \cap V_{k_1} \cap V_j| \times |P_i \cap V_{k_1} \cap V_j^c| \right\} \\ &= |P_i \cap V_{k_1} \cap V_{k_2}| \times |P_i \cap V_{k_1} \cap V_{k_2}^c| \end{aligned}$$

Furthermore, construct two characteristic matrices $C(i; k_1^+, k_2^+)$ and $C(i; k_1^+, k_2^-)$ defined as follows.

$$c_{jj}(i; k_1^+, k_2^+) = \begin{cases} 1 & , \text{ for } \hat{u}_j \in (P_i \cap V_{k_1} \cap V_{k_2}) \\ -1 & , \text{ for } \hat{u}_j \notin (P_i \cap V_{k_1} \cap V_{k_2}) \end{cases}$$

and

$$c_{jj}(i; k_1^+, k_2^-) = \begin{cases} 1 & , \text{ for } \hat{u}_j \in (P_i \cap V_{k_1} \cap V_{k_2}^c) \\ -1 & , \text{ for } \hat{u}_j \notin (P_i \cap V_{k_1} \cap V_{k_2}^c) \end{cases}$$

Examine whether or not $C(i; k_1^+, k_2^+)$ and $C(i; k_1^+, k_2^-)$ are threshold functions. If both of them are threshold functions, then $C(i)$ can be synthesized with the Cascade-Network of three elements. The actual weights values and threshold value are obtained through Theorem 3.1 for $C(i; k_1^-)$, $C(i; k_1^+, k_2^+)$, and $C(i; k_1^+, k_2^-)$.

If only $C(i; k_1^+, k_2^-)$ is a threshold function, construct two characteristic matrices $C(i; k_1^+, k_2^+, k_3^+)$ and $C(i; k_1^+, k_2^+, k_3^-)$ in the similar manner. On the other hand, if only $C(i; k_1^+, k_2^+)$ is a threshold function, construct $C(i; k_1^+, k_2^-, k_3^+)$ and $C(i; k_1^+, k_2^-, k_3^-)$ in the similar manner. If neither of them is a threshold function, construct similarly $C(i; k_1^+, k_2^+, k_3^+)$, $C(i; k_1^+, k_2^+, k_3^-)$, $C(i; k_1^+, k_2^-, k_3^+)$ and $C(i; k_1^+, k_2^-, k_3^-)$.

- (4) Search for the suffix k_2 of the input variable such that

$$\begin{aligned} & \max_j \left\{ \left| P_i \cap v_{k_1}^c \cap v_j \right| \times \left| P_i \cap v_{k_1}^c \cap v_j^c \right| \right\} \\ &= \left| P_i \cap v_{k_1}^c \cap v_{k_2} \right| \times \left| P_i \cap v_{k_1}^c \cap v_{k_2}^c \right| \end{aligned}$$

Moreover, construct two characteristic matrices $C(i; k_1^-, k_2^+)$ and $C(i; k_1^-, k_2^-)$ defined as follows.

$$\begin{aligned} c_{jj}(i; k_1^-, k_2^+) &= \begin{cases} 1 & , \text{ for } \hat{u}_j \in (P_i \cap v_{k_1}^c \cap v_{k_2}) \\ -1 & , \text{ for } \hat{u}_j \notin (P_i \cap v_{k_1}^c \cap v_{k_2}) \end{cases} \\ c_{jj}(i; k_1^-, k_2^-) &= \begin{cases} 1 & , \text{ for } \hat{u}_j \in (P_i \cap v_{k_1}^c \cap v_{k_2}^c) \\ -1 & , \text{ for } \hat{u}_j \notin (P_i \cap v_{k_1}^c \cap v_{k_2}^c) \end{cases} \end{aligned}$$

Henceforth, perform the similar procedures shown in Step 3.

- (5) Search for the suffixs k_2 and k_3 such that

$$\begin{aligned} & \max_j \left\{ \left| P_i \cap v_{k_1} \cap v_j \right| \times \left| P_i \cap v_{k_1} \cap v_j^c \right| \right\} \\ &= \left| P_i \cap v_{k_1} \cap v_{k_2} \right| \times \left| P_i \cap v_{k_1} \cap v_{k_2}^c \right| \end{aligned}$$

and

$$\begin{aligned} & \max_j \left\{ \left| P_i \cap v_{k_1}^c \cap v_j \right| \times \left| P_i \cap v_{k_1}^c \cap v_j^c \right| \right\} \\ &= \left| P_i \cap v_{k_1}^c \cap v_{k_3} \right| \times \left| P_i \cap v_{k_1}^c \cap v_{k_3}^c \right| \end{aligned}$$

Furthermore, construct four characteristic matrices $C(i; k_1^+, k_2^+)$, $C(i; k_1^+, k_2^-)$, $C(i; k_1^-, k_3^+)$ and $C(i; k_1^-, k_3^-)$

in the similar manner shown in Step 3 and Step 4.

Hereafter, perform the similar procedures shown in Step 3.

$C(i; k_1^*, k_2^*, \dots, k_N^*)$ becomes a threshold function, where "*" implies either "+" or "-", because $C(i; k_1^*, k_2^*, \dots, k_N^*)$ is a Boolean function which takes the value 1 for at most one input vertex. Hence the above synthesis procedure terminates, even at the worst case, after $C(i; k_1^*, k_2^*, \dots, k_N^*)$ is constructed.

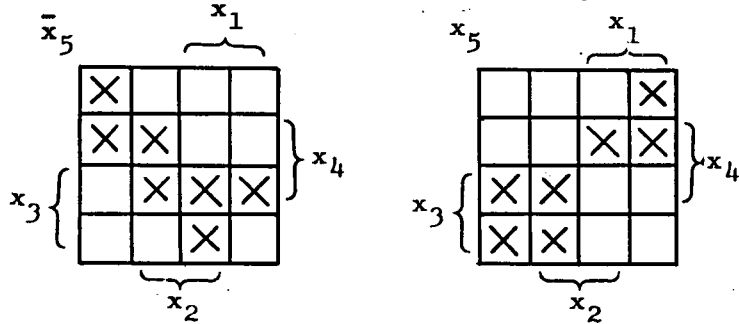
The synthesis procedure to have been presented here is not superior to the Threshold-Cascade synthesis or the Threshold-Or synthesis insofar as we are concerned only with the number of the required elements. However, this procedure does not need, in advance, the knowledge of whether the decomposed function is a threshold function. That is, at first the decomposition is performed and then the decomposed function is examined to be a threshold function. Thus, the procedure is straightforward and can be executed easily by the computer.

Some consideration to decrease the number of required elements is given by searching for the suffix of the input variable such that "max".
j

9.3 Example

Example 9.1. Synthesize the network of threshold elements realizing the Boolean function

$$f_i = \bar{x}_1 x_3 x_5 + x_1 \bar{x}_2 \bar{x}_3 x_5 + x_1 x_2 \bar{x}_3 x_4 x_5 + \bar{x}_1 \bar{x}_2 \bar{x}_3 \bar{x}_5 \\ + \bar{x}_1 x_2 x_4 \bar{x}_5 + x_1 x_2 x_3 \bar{x}_5 + x_1 x_3 x_4 \bar{x}_5$$



Since $C(i)$ is not a threshold function, decompose $C(i)$ into two functions $C(i; 5^+)$ and $C(i; 5^-)$ given as follows.

$$C(i; 5^+) = \bar{x}_1 x_3 x_5 + x_1 \bar{x}_2 \bar{x}_3 x_5 + x_1 x_2 \bar{x}_3 x_4 x_5$$

$$C(i; 5^-) = \bar{x}_1 \bar{x}_2 \bar{x}_3 \bar{x}_5 + \bar{x}_1 x_2 x_4 \bar{x}_5 + x_1 x_2 x_3 \bar{x}_5 + x_1 x_3 x_4 \bar{x}_5.$$

However, since $C(i; 5^+)$ is not a threshold function, decompose $C(i; 5^+)$ into two functions $C(i; 5^+, 1^+)$ and $C(i; 5^+, 1^-)$ given as follows.

$$C(i; 5^+, 1^-) = \bar{x}_1 x_3 x_5$$

$$C(i; 5^+, 1^+) = x_1 \bar{x}_2 \bar{x}_3 x_5 + x_1 x_2 \bar{x}_3 x_4 x_5$$

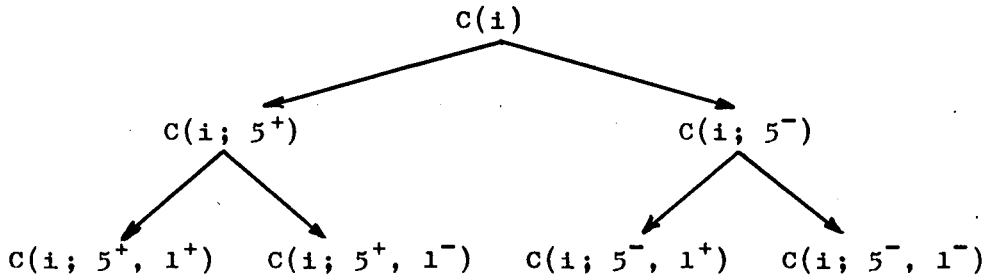
On the other hand, since $C(i; 5^-)$ is not a threshold function, either, decompose $C(i; 5^-)$ into $C(i; 5^-, 3^+)$ and

$C(i; 5^-, 3^-)$ given as follows.

$$C(i; 5^-, 3^+) = x_1 x_2 x_3 \bar{x}_5 + x_1 x_3 x_4 \bar{x}_5 + \bar{x}_1 x_2 x_3 x_4 \bar{x}_5$$

$$C(i; 5^-, 3^-) = \bar{x}_1 \bar{x}_2 \bar{x}_3 \bar{x}_5 + \bar{x}_1 \bar{x}_3 x_4 \bar{x}_5$$

Now, since $C(i; 5^+, 1^+)$, $C(i; 5^+, 1^-)$, $C(i; 5^-, 3^+)$ and $C(i; 5^-, 3^-)$ are threshold functions, $C(i)$ is composed of the four threshold functions shown as follows.



Hence, $C(i)$ can be synthesized by the cascade network of four elements shown in Fig. 9.1.

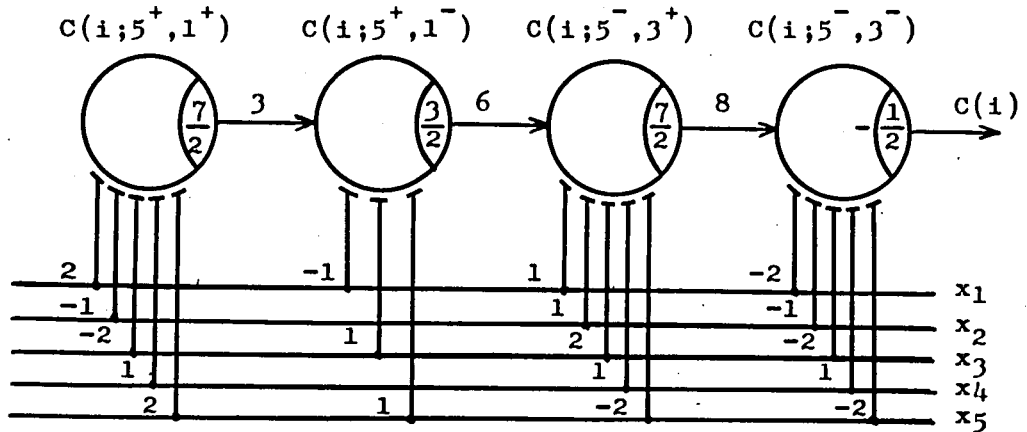


Fig. 9.1. The cascade network

CHAPTER X

LINEAR MAPPING OF FIRST QUADRANT

All the subjects, so far treated, have been investigated by solving the simultaneous inequalities directly. In this chapter, similar subjects are discussed from an aspect of the linear mapping which maps a first quadrant vector into a first quadrant vector. For this purpose, some properties of the characteristic vector are used.

Before proceeding to the treatment of the linear mapping, the procedure to reduce the number of Inequalities 2.16 inherent to the linear separability problem is presented. This reduction procedure is based on the idea that we can eliminate some inequalities which are the direct consequence of the other inequalities.

10.1 Characteristic Vector and its Properties

The characteristic vector or other similar parameters were pointed out by M. L. Dertouzos, C. K. Chow and other investigators⁽¹⁴⁾ for the purpose of identification of a Boolean function with a certain real number set. They also proved various properties of these parameters. However, let us list up only the necessary definitions and properties for the development of our theory.

The general expression of a Boolean function $f_i(x_1, x_2, \dots, x_N)$ of N variables is converted so that all the binary variables may have the value -1 or 1 rather than 0 or 1 , respectively. This is easily accomplished through the following algebraic transformations.

$$y_k = 2x_k - 1$$

$$g_i = 2f_i - 1$$

Hence, y_k and g_i take either 1 or -1 , depending on whether x_k and f_i are 1 or 0 , respectively.

The 2^N input points of a Boolean function $f_i(x_1, x_2, \dots, x_N)$ can be identified by the internal state vectors $\hat{u}_j$ ($j=1, 2, \dots, 2^N$) defined in Chapter II. Since the quantities y_k , and g_i are defined for every input point $\hat{u}_j$, they are denoted by $y_k(\hat{u}_j)$ and $g_i(\hat{u}_j)$, respectively.

Then the characteristic vector $\sigma_i = (\sigma_{i0}, \sigma_{i1}, \dots, \sigma_{iN})$ is defined as follows.

$$\sigma_{ik} = \sum_{j=1}^{2^N} y_k(\hat{u}_j) \cdot g_i(\hat{u}_j) \quad , \quad (k = 1, 2, \dots, N)$$

$$\sigma_{i0} = \sum_{j=1}^{2^N} g_i(\hat{u}_j)$$

Therefore, the characteristic vector σ_i is completely determined by the given Boolean function f_i .

With regard to the components of the characteristic vector,

some relations which involve the algebraic sign are proved. That is, if a given Boolean function f_i is a threshold function, then the following relations always hold.

- (1) $\text{sign } \omega_k = \text{sign } \sigma_{ik}, \text{ for } \sigma_{ik} \neq 0$
 - (2) If $\sigma_{ik} = 0$, then $\omega_k = 0$ is admissible. (10.1)
- ($k = 1, 2, \dots, N$)

where ω_k implies the weight value associated with the k -th input x_k . Since the component σ_{i0} is not used in the subsequent argument, the relation holding for σ_{i0} is eliminated here.

10.2 Reduction of Inequalities

A modified positive set and a modified negative set are first constructed by using the characteristic vector and then a partial order relation is introduced into these sets in order to find immediately such inequalities as the direct consequence of the other inequalities. The elimination of the inequalities is executed by constituting a minimal set and a maximal set in these partial ordered sets.

Definition 10.1. Consider the N -dimensional vector $\tilde{\sigma}_i = (\tilde{\sigma}_{i1}, \tilde{\sigma}_{i2}, \dots, \tilde{\sigma}_{iN})$ which is completely determined by the characteristic vector $\tilde{\sigma}_i$ as follows.

$$\tilde{\sigma}_{ik} = \begin{cases} 1 & , \quad \text{if } \sigma_{ik} > 0 \\ 0 & , \quad \text{if } \sigma_{ik} = 0 \\ -1 & , \quad \text{if } \sigma_{ik} < 0 \end{cases} \quad (k = 1, 2, \dots, N)$$

The $\tilde{\sigma}_i$ is called a sign vector.

If a given characteristic matrix $C(i)$ represents the threshold function f_i , $C(i)$ is identified with the characteristic vector σ_i which corresponds to only one threshold function f_i . Thus, $C(i)$ determines uniquely the sign vector $\tilde{\sigma}_i$. However, the $\tilde{\sigma}_i$ does not correspond to one characteristic matrix $C(i)$.

Definition 10.2. Consider the two sets $\tilde{P}_i$ and $\tilde{Q}_i$ generated from the positive set P_i and the negative set Q_i , respectively, as follows. That is, $\tilde{P}_i$ and $\tilde{Q}_i$ consist of such components as the bit-wise product between the sign vector $\tilde{\sigma}_i$ and every component contained in P_i and Q_i , respectively.

$$\begin{aligned} \tilde{P}_i &= \left\{ \tilde{u}_j = \left[\hat{u}_j \cdot \tilde{\sigma}_i \right] \quad ; \quad \text{for every } \hat{u}_j \in P_i \right\} \\ \tilde{Q}_i &= \left\{ \tilde{u}_j = \left[\hat{u}_j \cdot \tilde{\sigma}_i \right] \quad ; \quad \text{for every } \hat{u}_j \in Q_i \right\} \end{aligned}$$

where the notation $\left[\hat{u}_j \cdot \tilde{\sigma}_i \right]$ indicates the bit-wise product of $\hat{u}_j$ and $\tilde{\sigma}_i$. The $\tilde{P}_i$ and $\tilde{Q}_i$ are called the modified positive set and the modified negative set, respectively. The N-dimensional vector $\tilde{u}_j = \left[\hat{u}_j \cdot \tilde{\sigma}_i \right]$ is called the modified internal state vector.

We can regard $\tilde{P}_i$ and $\tilde{Q}_i$ as the partial ordered sets by

introducing a partial order relation defined as follows. That is, let $\tilde{u}_j = (\tilde{u}_{j1}, \tilde{u}_{j2}, \dots, \tilde{u}_{jN})$ and $\tilde{u}_h = (\tilde{u}_{h1}, \tilde{u}_{h2}, \dots, \tilde{u}_{hN})$ be arbitrary two elements contained in $\tilde{P}_i$. Then, it is said to be $\tilde{u}_j \geq \tilde{u}_h$ if and only if the relation

$$u_{jk} \geq u_{hk} \quad (k = 1, 2, \dots, N)$$

holds for every k . By the similar manner, the partial order relation can be introduced into $\tilde{Q}_i$.

Definition 10.3. A modified internal state vector $\tilde{u}_h$ is said to be minimal, if and only if, for an arbitrary modified internal state vector $\tilde{u}_j$, the relation

$$\tilde{u}_h \geq \tilde{u}_j$$

implies

$$\tilde{u}_h = \tilde{u}_j$$

Similarly, a modified internal state vector $\tilde{u}_h$ is said to be maximal, if and only if, for an arbitrary $\tilde{u}_j$, the relation

$$\tilde{u}_h \leq \tilde{u}_j$$

implies

$$\tilde{u}_h = \tilde{u}_j$$

Let $\text{Min}(\tilde{P}_i)$ denote the set which consists of all the minimal vectors contained in the set $\tilde{P}_i$ and let $\text{Max}(\tilde{Q}_i)$ denote the set which consists of all the maximal vectors contained in the set $\tilde{Q}_i$. The $\text{Min}(\tilde{P}_i)$ and $\text{Max}(\tilde{Q}_i)$ are called the minimal set and the maximal set.

Definition 10.4. Delete all the j -th rows and j -th columns from a characteristic matrix $C(i)$, if $\tilde{u}_j$ is not contained in the set-sum $\text{Min}(\tilde{P}_1) \cup \text{Max}(\tilde{Q}_1)$. Such an obtained matrix is denoted by $C_M(i)$.

Likewise, delete from the universal matrix U all the j -th rows, if $\tilde{u}_j$ is not contained in the set-sum $\text{Min}(\tilde{P}_1) \cup \text{Max}(\tilde{Q}_1)$. Such a deletion yields submatrix of U , denoted by U_M^* .

$$U_M^* = \begin{pmatrix} 1 & \hat{u}_{i_1} \\ 1 & \hat{u}_{i_2} \\ \dots & \dots \\ \dots & \dots \\ 1 & \hat{u}_{i_m} \end{pmatrix}$$

The subscript i_j appears in the above expression of U_M^* , only if $\tilde{u}_{i_j}$ is contained in the set-sum $\text{Min}(\tilde{P}_1) \cup \text{Max}(\tilde{Q}_1)$. Moreover, replace all the vectors $\hat{u}_{i_j}$ of U_M^* with the corresponding modified internal state vector $\tilde{u}_{i_j}$. Let us denote such an obtained matrix by U_M^+ .

$$U_M^+ = \begin{pmatrix} 1 & \tilde{u}_{i_1} \\ 1 & \tilde{u}_{i_2} \\ \dots & \dots \\ \dots & \dots \\ 1 & \tilde{u}_{i_m} \end{pmatrix} \quad \begin{matrix} \tilde{u}_{i_j} \in \text{Min}(\tilde{P}_1) \cup \text{Max}(\tilde{Q}_1) \\ (j = 1, 2, \dots, m) \end{matrix}$$

Furthermore, replace all the 1's of the first column of U_M^+ with

-1. The resulting matrix is denoted by U_M^- .

$$U_M^- = \begin{pmatrix} -1 & \tilde{u}_{i_1} \\ -1 & \tilde{u}_{i_2} \\ \cdot & \cdot \\ \cdot & \cdot \\ -1 & \tilde{u}_{i_m} \end{pmatrix} \quad \begin{aligned} \tilde{u}_{i_j} &\in \text{Min}(\tilde{P}_i) \cup \text{Max}(\tilde{Q}_i) \\ (j &= 1, 2, \dots, m) \end{aligned}$$

With respect to the matrices $C_M(i)$, U_M^+ and U_M^- defined above, the next theorem is established.

Theorem 10.1. Inequality System 2.16

$$C(i) \cdot U \cdot t_{w_i} > 0 \quad (2.16)$$

is consistent, if and only if either of the inequality system

$$\begin{aligned} C_M(i) \cdot U_M^+ \cdot t_{w_i}^+ &> 0 \\ w_i^+ &\geq 0 \end{aligned} \quad (10.2)$$

or the inequality system

$$\begin{aligned} C_M(i) \cdot U_M^- \cdot t_{w_i}^- &> 0 \\ w_i^- &\geq 0 \end{aligned} \quad (10.3)$$

is consistent.

Proof: If Inequality System 2.16 is consistent, then $C(i)$ is a threshold function. Hence, there exists the characteristic vector σ_i which has one to one correspondence with

C(i). The components of σ_i constrain the signs of the components of w_i as given by Relation 10.1. Hence, if we assume $w_i \geq 0$, then we must compensate this constrain by changing the signs of the components of the universal matrix U . This compensation is done by introducing U_M^+ and U_M^- . Here, note that U_M^+ and U_M^- are considered because we can not detect, in advance, the sign of w_{i0} , the threshold value. Therefore, either of Inequality System 10.2 or 10.3 becomes a subsystem of Inequality System 2.16.

Conversely, from the definitions of $\text{Min}(\tilde{P}_i)$ and $\text{Max}(\tilde{Q}_i)$, the following relations always hold.

$$\begin{aligned} ((1, \tilde{u}_a) \cdot w_i) &\leq ((1, \tilde{u}_b) \cdot w_i) \\ \text{for } \tilde{u}_a &\in \text{Min}(\tilde{P}_i) \\ u_b &\in P_i \text{ but } \tilde{u}_b \notin \text{Min}(\tilde{P}_i) \end{aligned}$$

and

$$\begin{aligned} ((1, \tilde{u}_c) \cdot w_i) &\geq ((1, \tilde{u}_d) \cdot w_i) \\ \text{for } \tilde{u}_c &\in \text{Max}(\tilde{Q}_i) \\ u_d &\in Q_i \text{ but } \tilde{u}_d \notin \text{Max}(\tilde{Q}_i) \end{aligned}$$

where $((1, \tilde{u}_j) \cdot w_i)$ implies the inner product of the vectors $(1, \tilde{u}_j)$ and w_i . Hence, if either of Inequality System 10.2 or 10.3 is consistent, then Inequality System 2.16 becomes consistent. Q.E.D.

If we regard $C_M(i) \cdot U_M^+$ and $C_M(i) \cdot U_M^-$ as the linear mappings of the $(N+1)$ -dimensional Euclidean space R^{N+1} , Theorem 10.1 can be restated as follows.

Theorem 10.2. Inequality System 2.16 is consistent, if and only if either of $C_M(i) \cdot U_M^+$ or $C_M(i) \cdot U_M^-$ is a linear mapping which maps a vector of the first quadrant to a vector of the first quadrant.

We can also rewrite Theorem 10.2 in the following form.

Corollary 10.2.1. Inequality System 2.16 is consistent, if and only if there exists a vector of the first quadrant of R^{N+1} which is mapped to a vector of the first quadrant by either of $C_M(i) \cdot U_M^+$ or $C_M(i) \cdot U_M^-$.

The less the number of components in the set-sum $\text{Min}(\tilde{P}_i) \cup \text{Max}(\tilde{Q}_i)$ becomes, the easier the analysis based on Inequality System 10.2 or 10.3 rather than Inequality System 2.16 becomes.

10.3 Development

Theorem 10.3. If there exists at least one nonpositive row in both the matrices $C_M(i) \cdot U_M^+$ and $C_M(i) \cdot U_M^-$, then $C(i)$ is not a threshold function.

Proof: Let the k -th row of $C_M(i) \cdot U_M^+$ is nonpositive, then the k -th inequality of System 10.2 is not satisfied.

Q.E.D.

Theorem 10.4. If there exists at least one positive column in either of $C_M(i) \cdot U_M^+$ or $C_M(i) \cdot U_M^-$, then $C(i)$ is a threshold function.

Proof: Let the k -th column of $C_M(i) \cdot U_M^+$ is positive, then consider a vector w_1 such that only the $(k+1)$ -th component w_{1k} is sufficiently large. Then, this w_1 becomes a solution of Inequality System 10.2. Q.E.D.

Theorem 10.5. If we can generate a nonpositive vector by the linear combination of the rows with positive coefficients for both $C_M(i) \cdot U_M^+$ and $C_M(i) \cdot U_M^-$, then $C(i)$ is not a threshold function.

Proof: The existence of nonpositive linear combinations shows that there exists at least one inequality which does not satisfy Relation 10.1 and 10.2. Q.E.D.

Definition 10.5. If the μ_1 -th column of $C_M(i) \cdot U_M^+$ is nonzero nonnegative, then eliminate all the rows from $C_M(i) \cdot U_M^+$ such that the components of the μ_1 -th column are positive. Such a resulting matrix is denoted by $\overleftarrow{C_M(i)} \cdot U_M^+(\mu_1)$. Similarly, if the μ_2 -th column of $\overleftarrow{C_M(i)} \cdot U_M^+(\mu_1)$ is nonzero nonnegative, then eliminate all the rows from $\overleftarrow{C_M(i)} \cdot U_M^+(\mu_1)$ such that the components of the μ_2 -th column are positive. Such an obtained matrix is denoted by $\overleftarrow{C_M(i)} \cdot U_M^+(\mu_1, \mu_2)$. Likewise, $\overleftarrow{C_M(i)} \cdot U_M^+(\mu_1, \mu_2, \dots, \mu_n)$ can be defined by the similar manner. For $C_M(i) \cdot U_M^-$, $\overleftarrow{C_M(i)} \cdot U_M^-(\eta_1, \eta_2, \dots, \eta_m)$ can be also defined by the similar way.

Theorem 10.6. If there exists either $\overleftarrow{C_M(i)} \cdot U_M^+(\mu_1, \mu_2, \dots, \mu_n)$ or $\overleftarrow{C_M(i)} \cdot U_M^-(\eta_1, \eta_2, \dots, \eta_m)$ which becomes the (0,0)-type matrix, namely, vanished, then $C(i)$ is a threshold function.

Proof: Consider the vectors w_i^+ or w_i^- such that

$$w_{i\mu_1}^+ \gg w_{i\mu_2}^+ \gg \dots \gg w_{i\mu_n}^+ \gg w_{ij}^+ > 0$$

$$(j \neq \mu_1, \mu_2, \dots, \mu_n)$$

or

$$w_{i\eta_1}^- \gg w_{i\eta_2}^- \gg \dots \gg w_{i\eta_m}^- \gg w_{ij}^- > 0$$

$$(j \neq \eta_1, \eta_2, \dots, \eta_m) .$$

The w_i^+ or w_i^- can be surely one of solutions of Inequality 10.2 or Inequality 10.3, respectively. Q.E.D.

Let us develop Theorem 10.6 in order to be more applicable form.

Definition 10.6. First, make a linear combination of some columns of $\overleftarrow{C_M(i)} \cdot U_M^+$ with positive coefficients so that the linear combined column may be nonzero nonnegative, then eliminate all the rows from $\overleftarrow{C_M(i)} \cdot U_M^+$ such that the components of this linear combined column are positive. Such an obtained matrix is denoted by $\overleftarrow{C_M(i)} \cdot U_M^+(\text{column:1})$. Similarly, make a linear combination of some columns of $\overleftarrow{C_M(i)} \cdot U_M^-(\text{column:1})$ with positive coefficients so that the linear combined column may be nonzero nonnegative, then eliminate all the rows from

$\overleftarrow{C_M(i)} \cdot U_M^+(\text{column}:1)$ such that the components of this linear combined column are positive. The resulting matrix is denoted by $\overleftarrow{C_M(i)} \cdot U_M^+(\text{column}:2)$. Likewise, $\overleftarrow{C_M(i)} \cdot U_M^+(\text{column}:n)$ and $\overleftarrow{C_M(i)} \cdot U_M^-(\text{column}:m)$ can be defined.

Then, the next theorem is established by the similar manner to Theorem 10.6.

Theorem 10.7. If there exists either $\overleftarrow{C_M(i)} \cdot U_M^+(\text{column}:n)$ or $\overleftarrow{C_M(i)} \cdot U_M^-(\text{column}:m)$ which becomes the (0,0)-type matrix, namely, vanished, then $C(i)$ is a threshold function.

10.4 Example

Example 10.1. Unless the unate Boolean function of more than 9 variables is tested, it is not instructive because various Tables have been already prepared for this purpose. However, for the illustration of the method mentioned above, test the following function.

$$f_1 = x_1 \bar{x}_3 x_4 + x_2 x_3 x_4 + \bar{x}_1 \bar{x}_4 \bar{x}_5 + \bar{x}_1 x_2 \bar{x}_4 x_5 + \bar{x}_1 \bar{x}_2 \bar{x}_3 x_4 x_5$$

The characteristic vector σ_1 is given as follows.

$$\sigma_1 = (-2, -2, 6, -2, 10, -6)$$

Hence, the sign vector $\tilde{\sigma}_1$ is given as follows.

$$\tilde{\sigma}_1 = (-1, 1, -1, 1, -1)$$

Therefore, the order relations in $\tilde{P}_1$ and in $\tilde{Q}_1$ are shown in Fig. 10.1 and in Fig. 10.2, respectively.

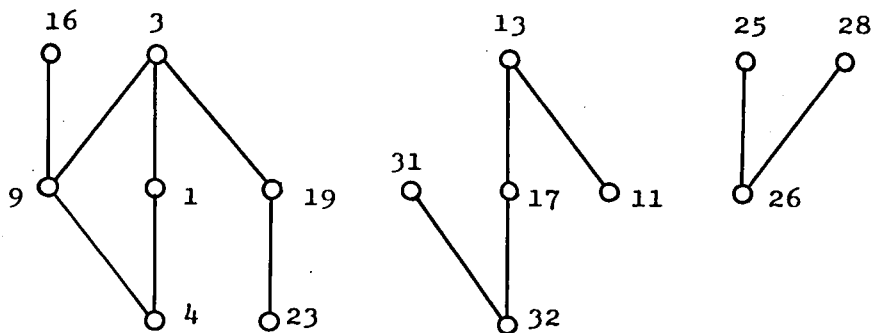


Fig. 10.1. The order relation in $\tilde{P}_1$

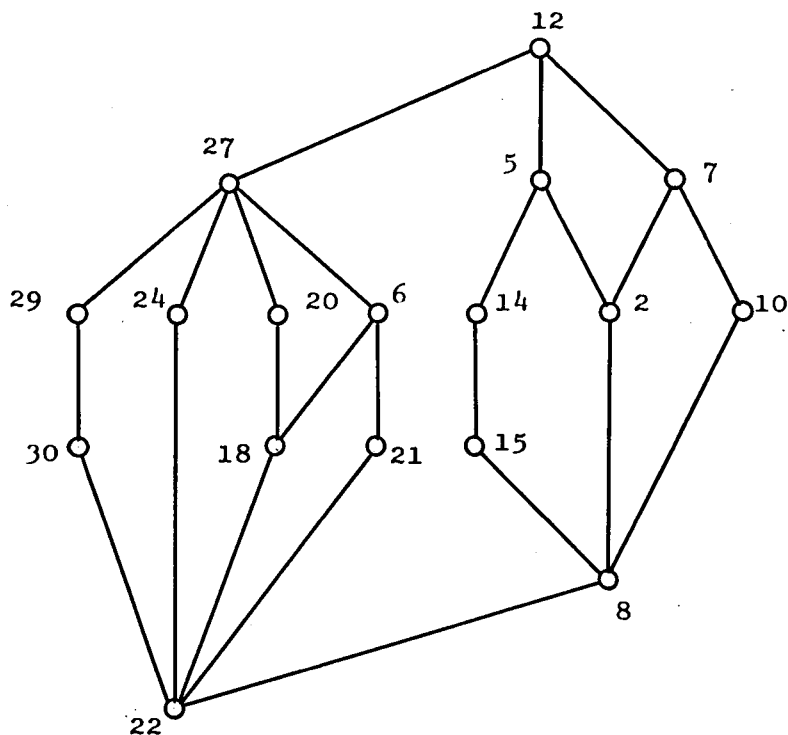


Fig. 10.2. The order relation in $\tilde{Q}_1$

Here, the number implies the suffix of the modified internal state vector.

Thus, we obtain $\text{Min}(\tilde{P}_1) = (\tilde{u}_4, \tilde{u}_{11}, \tilde{u}_{23}, \tilde{u}_{26}, \tilde{u}_{32})$ and $\text{Max}(\tilde{Q}_1) = (\tilde{u}_{12})$. Hence, $C(i) \cdot U_M^+$ and $C(i) \cdot U_M^-$ are given as follows.

$$C(i) \cdot U_M^+ = \begin{pmatrix} -1 & 0 & -1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 1 & -1 & 0 & 0 & 1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 \end{pmatrix}$$

$$C(i) \cdot U_M^- = \begin{pmatrix} 1 & 0 & -1 & 0 & -1 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 1 & -1 & 0 & -1 \\ -1 & -1 & 0 & 0 & 1 & -1 \\ -1 & -1 & 1 & -1 & 1 & -1 \end{pmatrix}$$

Since the first row of $C(i) \cdot U_M^+$ and the second row of $C(i) \cdot U_M^-$ are nonpositive, f_1 is not a threshold function.

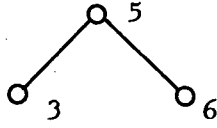
Example 10.2. Synthesize the network whose behavior is given by Fig. 4.3. Since $C(3)$ is not a threshold function, generate the characteristic vector σ_3 and the sign vector $\tilde{\sigma}_3$.

$$\sigma_3 = (-2, 2, 2, -2)$$

$$\tilde{\sigma}_3 = (1, 1, -1)$$

Hence, $\tilde{P}_1$ and $\tilde{Q}_1$ are given as follows.

$$\begin{aligned}\tilde{P}_3 \\ \tilde{u}_3 &= (0 \ 1 \ 0) \\ \tilde{u}_5 &= (1 \ 1 \ 0) \\ \tilde{u}_6 &= (1 \ 0 \ -1)\end{aligned}$$



$$\begin{aligned}\tilde{Q}_3 \\ \tilde{u}_1 &= (0 \ 0 \ 0) \\ \tilde{u}_2 &= (1 \ 0 \ 0) \\ \tilde{u}_4 &= (0 \ 0 \ -1) \\ \tilde{u}_7 &= (0 \ 1 \ -1) \\ \tilde{u}_8 &= (1 \ 1 \ -1)\end{aligned}$$

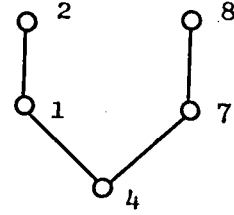


Fig. 10.3. The order relation in $\tilde{P}_3$ and $\tilde{Q}_3$

Hence, we obtain $\text{Min}(\tilde{P}_3) = (\tilde{u}_3, \tilde{u}_6)$ and $\text{Max}(\tilde{Q}_3) = (\tilde{u}_2, \tilde{u}_8)$.

Thus, $C(3) \cdot U_M^-$ is given as follows.

$$C(3) \cdot U_M^- = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & -1 & -1 & 1 \\ -1 & 0 & 1 & 0 \\ -1 & 1 & 0 & -1 \end{pmatrix}$$

Now, supplement one threshold element T_4 and convert $\tilde{P}_3$ and $\tilde{Q}_3$ into the following form.

$$\begin{aligned}\tilde{P}_3 \\ (1 \ 1 \ 0) &\longrightarrow (1 \ 1 \ 0 \ 1) \\ (0 \ 1 \ 0) &\longrightarrow (0 \ 1 \ 0 \ 1) \\ (1 \ 0 \ -1) &\longrightarrow (1 \ 0 \ -1 \ 1)\end{aligned}$$

$$\begin{aligned}\tilde{Q}_3 \\ (0 \ 0 \ 0) &\longrightarrow (0 \ 0 \ 0 \ 0) \\ (1 \ 0 \ 0) &\longrightarrow (1 \ 0 \ 0 \ 0) \\ (0 \ 0 \ -1) &\longrightarrow (0 \ 0 \ -1 \ 0) \\ (0 \ 1 \ -1) &\longrightarrow (0 \ 1 \ -1 \ 0) \\ (1 \ 1 \ -1) &\longrightarrow (1 \ 1 \ -1 \ 0)\end{aligned}$$

In accordance with this conversion, $C(3) \cdot U_M^-$ is altered as follows.

$$C(3) \cdot \vec{w}_3 = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 \\ -1 & 0 & 1 & 0 & 1 \\ -1 & 1 & 0 & -1 & 1 \end{pmatrix}$$

Hence, $\vec{w}_3 = (2, 1, 1, 1, 3)$ is one of the solutions of the inequalities

$$C(3) \cdot U_M^- \cdot \vec{w}_3 > 0.$$

The conversion mentioned above determines the characteristic matrix $C(4)$ as follows.

$$C(4) = \begin{pmatrix} 1 & & & & \\ & 1 & & & 0 \\ & & -1 & & \\ & & & -1 & \\ & & & & 1 \\ 0 & & & & -1 \\ & & & & & -1 \\ & & & & & & -1 \end{pmatrix}$$

Thus, we obtain $\sigma_4 = (-2, 2, -2, -6)$ and $\tilde{\sigma}_4 = (1, -1, -1)$.

Hence, the order relations in $\tilde{P}_4$ and $\tilde{Q}_4$ are shown in Fig. 10.4.

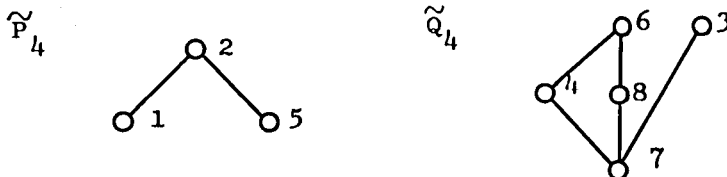


Fig. 10.4. The order relations in $\tilde{P}_4$ and $\tilde{Q}_4$

Thus, $C_M(4) \cdot U_M^+$ is obtained as follows.

$$C_M(4) \cdot U_M^+ = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 \\ -1 & -1 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

Therefore, $w_4^+ = (2, 2, 3, 5)$ is one of the solutions of the inequalities

$$C_M(4) U_M^+ t w_4^+ > 0 .$$

Consequently, the network shown in Fig. 10.5 performs the behavior given by Fig. 7.1.

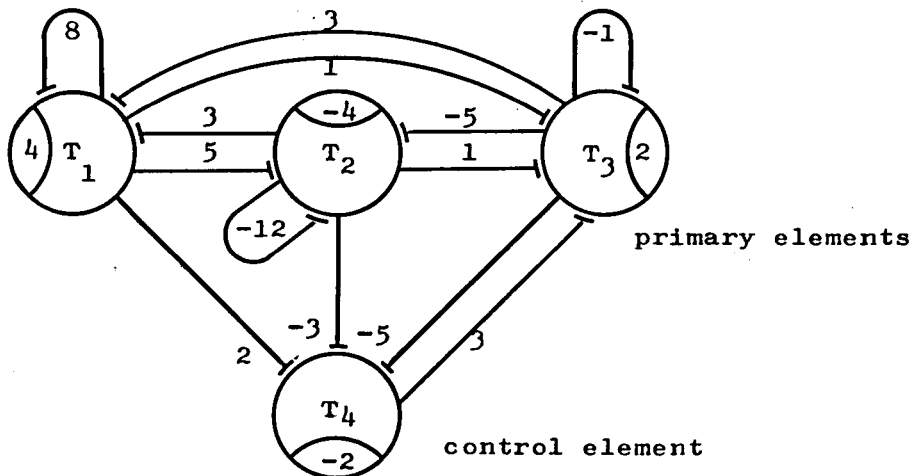


Fig. 10.5. The autonomous network

CHAPTER XI

AN ELECTRIC ANALOGUE OF THE NEURON

An electric analogue model of the neuron is described. ⁽¹⁹⁾⁽²⁰⁾

The model neuron is composed of the active units which simulate the electric behaviors of the active loci of the membrane of a neuron. The active units include the model axon and the model synapses of six different types (the ordinary, incremental and decremental ones each having the excitatory and the inhibitory types). The properties of the models are described.

11.1 Introduction

During the last decade there have been many attempts to ⁽³¹⁾⁽⁴⁷⁾ construct electric analogues of the neuron. It seems to have

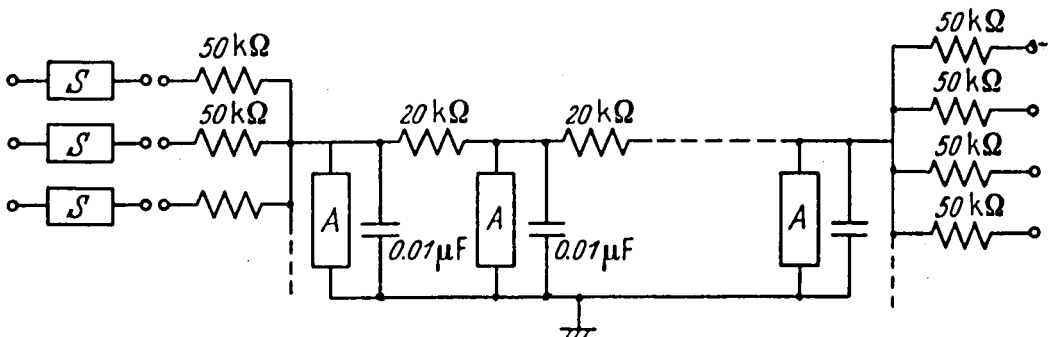


Fig. 11.1 The circuit of the model neuron

paramount importance to investigate the electric performance of the neuron and the nervous system constructing an analogue model. As the electrophysiological data show the electric properties of the neuron are fairly complex and it does not operate in a simple digital or analogue manner as the artificial data processing organs. In some respects it operates in the analogue way and in the other respects it does in the digital way. It seems fairly certain that such complex nature of the nervous function has some deep implications in the data processing in the nervous system. To understand the functions of the neuron in detail and how it works as a data processing unit in a nervous system the use of an analogue model seems to offer some advantages. On account of the complex nature of the neuron functioning the detail mathematical treatment becomes a formidable one or it needs gross simplification to make a mathematical treatment tractable. On the other hand the actual neuron does not yield to change the parameters of its electric functions arbitrarily or to construct an arbitrary nervous net. These difficulties may be circumbented by the use of the analogue model. By the use of the analogue model we may understand the functions of the neuron and its potentialities as the data processing unit in considerable detail. The understanding of the nervous functions and the small nervous net in their full complexity seems to be a necessary first step for the understanding of the nervous system.

Of course, we must be cautious on the limit of the model. It simulates only the electric behaviors of the neuron and not the chemical or the molecular processes. The chemical processes may be essential for some properties in the nervous system (especially in the problem of the memory).

In this paper we describe an electric analogue of the neuron and its properties. As the recent data of the electrophysiology shows the neuron membrane consists of the active loci of several different types. We intended to simulate the axon membrane and the synaptic loci of six different types (ordinary, incremental and decremental synapses each including the excitatory and the inhibitory ones).

Each active unit is constructed independently and they are assembled in a neuron as in Fig.11.1. We may assemble the synaptic units in many different combinations and the axonal units may be arranged in a long or short axon cable. This "active units" method has the advantage of the great flexibility.

11.2 The Axon Analogue

1) The circuit of the model axon.

The electric properties of the axon membrane were analyzed in detail by HODGKIN and HUXLEY.⁽³²⁾ To construct the model axon we intended to simulate qualitatively the essential features of the electric properties of the axon. The circuit of the model axon unit is shown in Fig.11.2. In this model the active ionic

currents across the axon membrane are represented by the current I through the capacitor C_2 . In the rest state the input voltage V is zero and the vacuum tube T_1 is on and T_2 is off if the grid voltage E_{g2} of T_2 is set below a critical negative voltage E_c . When V is decreased and reaches below the threshold value V_{th} which satisfies $V_{th} - E_{g2} = E_c$, T_1 is switched to the off state and T_2 to the off state. After a brief active period in which T_1 is off and T_2 is on, T_1 is again switched to the on state and T_2 to the off state. If V is set below the threshold V_{th} , the axon unit fires repetitively. In the active phase the current I consists of the two components; the current I_n which compensates the current i_2 flowing through the tube T_2 and the discharge current I_k of the capacitor C_2 . I_n is inward and I_k is outward. In the post active phase in which T_1 and T_2 are switched to the on and off states respectively, I consists only

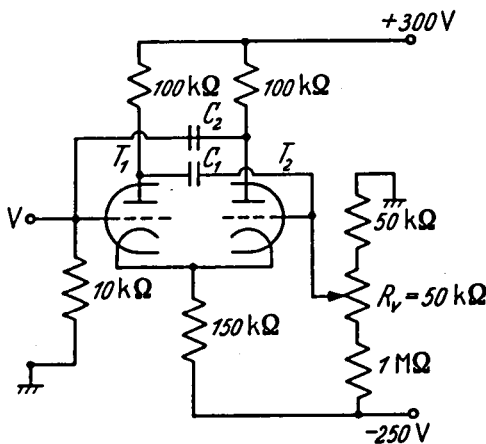


Fig. 11.2. The circuit of the model axon unit. A medium mu twin triode

12AU7 is used for T_1 and T_2 , $C_1 = 0.001\mu F$, $C_2 = 0.02\mu F$

of the discharge current I_k . I_n and I_k may be corresponded to the sodium and potassium currents respectively.

ii) The wave form of the axon spike.

In Fig.11.3 the wave form of the axon spike potential is shown. The wave form of the spike is fairly stable. It shows only about 10% change for the 10 fold changes of the threshold or the pulse density.

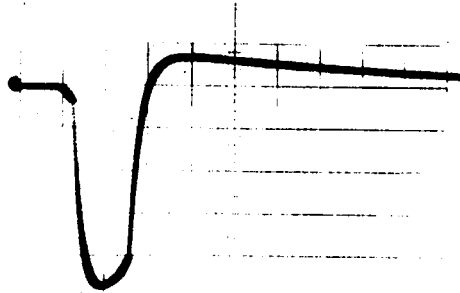


Fig. 11.3. The wave form of the axon spike potential. A scale gives 3 volts for the ordinate and 300 μ sec for the abscissa

iii) The threshold and the refractory period.

As was mentioned previously, the axon unit has the threshold. For the input which does not reach the threshold V_{th} the axon unit does not actively respond and the input voltage propagates passively over the axon cable and decays, while when the input reaches the threshold the axon unit fires actively and

the spike potential propagates regeneratively. The value of the threshold may be changed by adjusting the variable resistance R_v . The deeper the grid voltage E_{g2} than E_c , the larger the threshold V_{th} . If E_{g2} is set above E_c , then the axon unit fires spontaneously.

This axon unit has the refractory period. The time course of the change of the threshold after a firing is shown in Fig. 11.4. As seen from Fig. 11.4. the absolute refractory period is about 1 m sec.

iv) Pulse density modulation.

When the axon unit is fed a D.C. voltage exceeding the threshold, it fires repetitively and the frequency of the firing changes according to the magnitude of the input. In Fig. 11.5 and Fig. 11.6 the pulse frequency vs. voltage relations are shown.

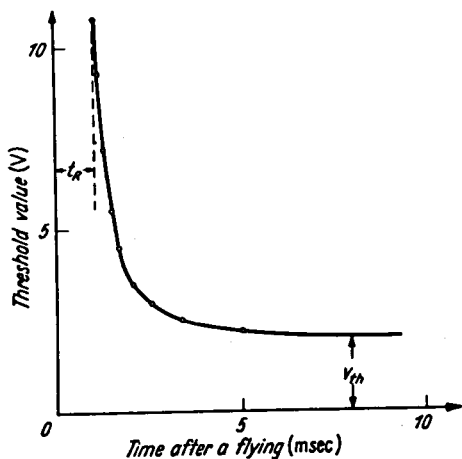


Fig. 11.4. The time course of the threshold change after a firing. V_{th} is sec at 2 volts

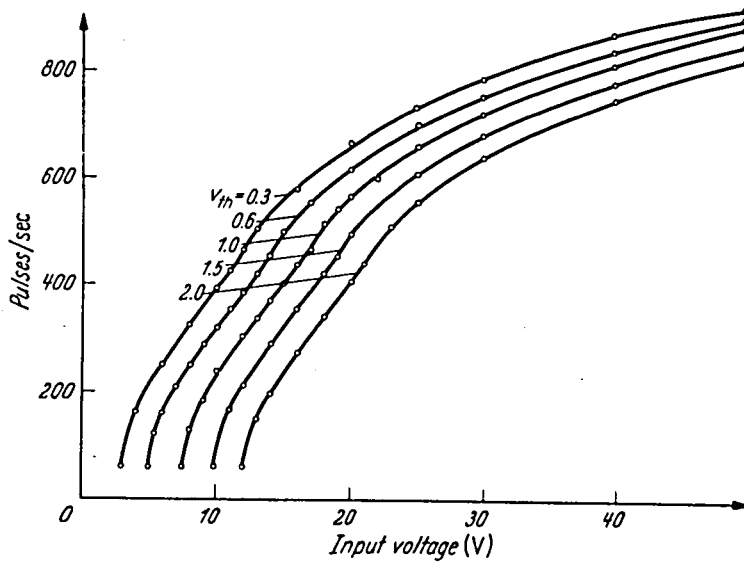


Fig. 11.5.

The relation between the applied d.c. voltage and the output

pulse frequency for the various values of the threshold

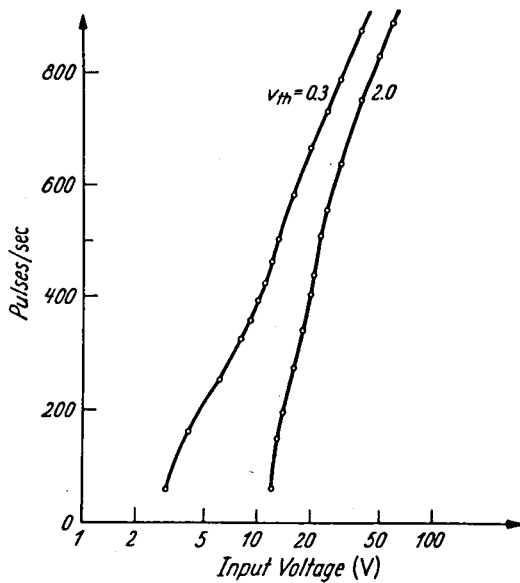


Fig. 11.6. The semi-log plot for the pulse

frequency vs. the input d.c. voltage relations

In Fig.11.5 the firing frequency vs. the input voltage relation is given in the linear scale for various values of the threshold and Fig.11.6 is the semilog plot of Fig.11.5. As seen from Fig. 11.6 the firing frequency is approximately linear with respect to $\log(V)$.
(33)(52)

In Fig.11.7 and Fig.11.8 the effects of the threshold change for the pulse density modulation are shown. In Fig.11.7 the frequency vs. the threshold relations for the fixed values of the input D.C. voltage are shown. In Fig.11.8 the input D.C. voltage vs. the threshold relations for the fixed frequencies are shown. As seen from Fig.11.7 and Fig.11.8 the pulse density for the fixed input voltage and the input voltage for the fixed pulse frequency vary linearly for the change of the threshold.

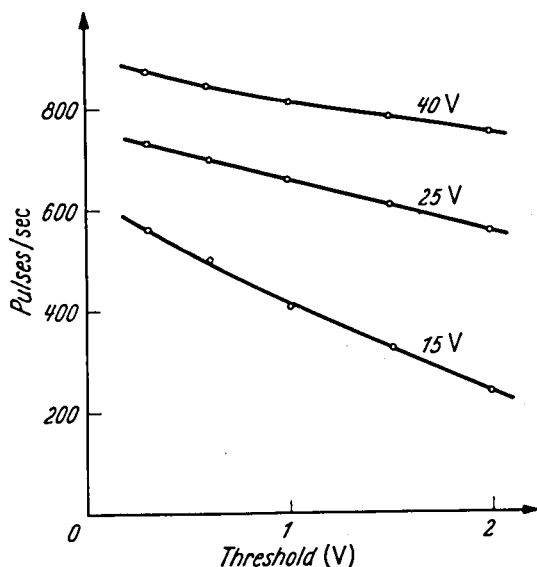


Fig. 11.7. The relations between the threshold and the pulse frequency for the fixed values of the input d.c. voltage

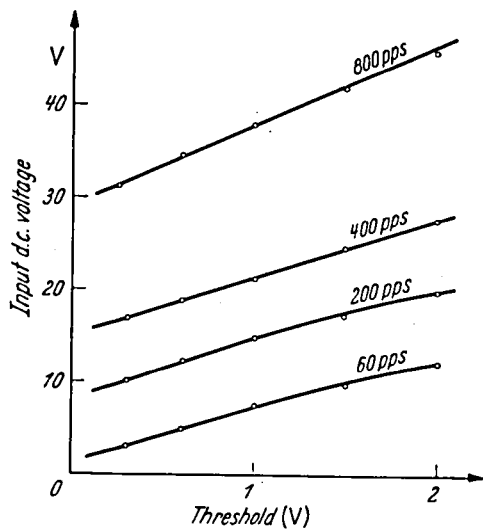


Fig. 11.8. The relations between the threshold and the input d.c. voltage for the fixed values of the

output pulse frequency

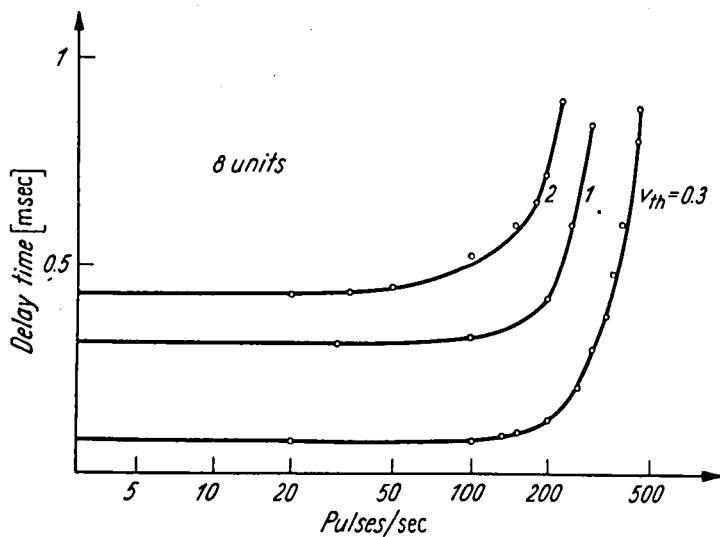


Fig. 11.9.

The relation between the delay time for the propagating spikes and the

pulse frequency. The delay time is measured for the axon cable consisting of 8 axon units and for several different values of the threshold

v) Propagation of the spike.

On account of the special regenerative nature of the propagation of the spike, there are some interesting properties for the propagation of the spike potential along the axon cable. When the first axon unit is fired a degraded potential change is induced at the site of the next axon. In order to fire the next axon unit it is necessary to reach the induced potential at the threshold value. So that there arise a delay to fire the next axon unit. The larger the threshold, the larger the delay. On account of the threshold change after the firing, the threshold of the axon units in the axon cable is raised when the axon is fired in the high frequency. So that the propagation velocity of the spike must be smaller for the high frequency firing than the low frequency firing.

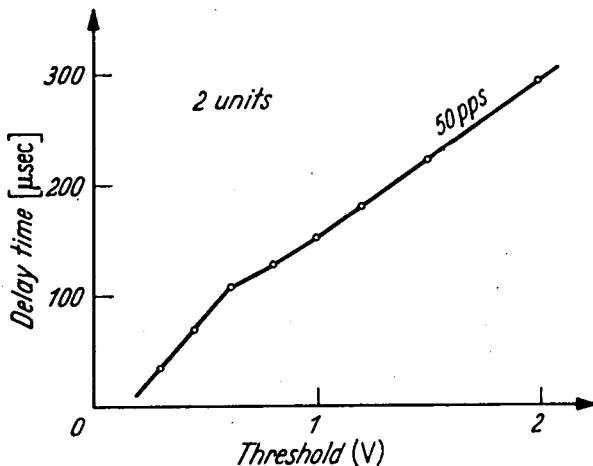


Fig. 11.10. The relation between the propagation delay of the spike and the threshold for a fixed value of the pulse frequency

(50 pulse per sec). The delay is measured for the axon cable consisting of 2 axon units

This is indeed the case. In Fig.11.9 the relations between the propagation delay and the firing frequency are plotted for the axon cable consisting of eight axon units for which the threshold is fixed at various values. In Fig.11.10 the relation between the delay time and the threshold at the

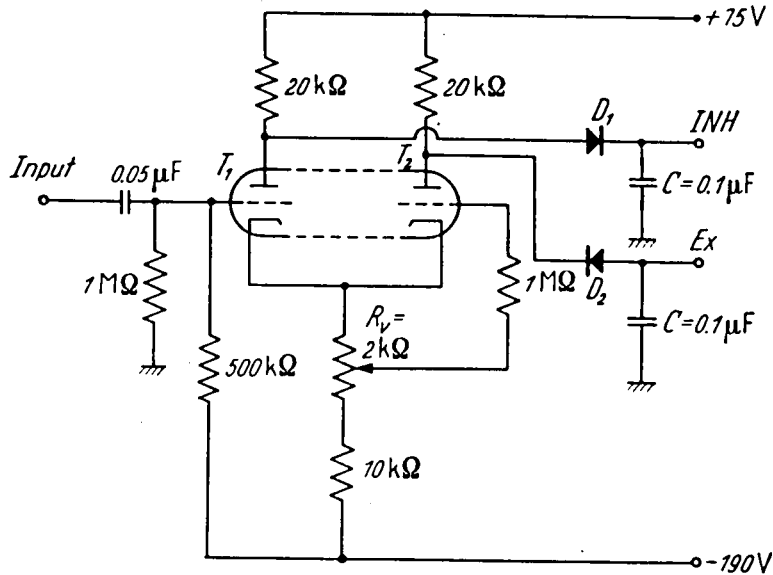


Fig. 11.11.
The circuit of
the ordinary
synapse model.
A twin triode
12AU7 is used
for T_1 and T_2 .

fixed firing rate (50 pulse per sec) for the axon cable consisting of 2 axon units is shown. As seen from Fig.11.10 the delay time is approximately linear for the threshold change.

11.3 The Synapse Analogues

According to the electrophysiological data we made three different types of the model synapse, the ordinary,

incremental and decremental ones. We made these synapse models in such a way that both the excitatory and the inhibitory outputs may be get from a synapse unit. The ordinary synapse is the most common synapse which responds to the spike inputs by the outputs of constant magnitude except for the close array of spikes for which the output synaptic potential overlaps (the time summation). The incremental (facilitatory) and the decremental (defacilitatory) synapses are such ones that for a series of the input spikes in succession the magnitudes of the output synaptic potentials gradually increase or decrease even when no overlapping of the output potential occurs.

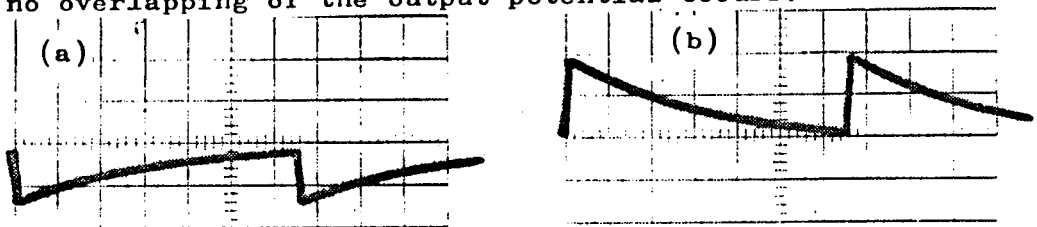


Fig.11.12. The wave forms of the ordinary synaptic potentials. a) excitatory, b) inhibitory. A scale gives 1 volt for the ordinate and 0.5 msec for the abscissa

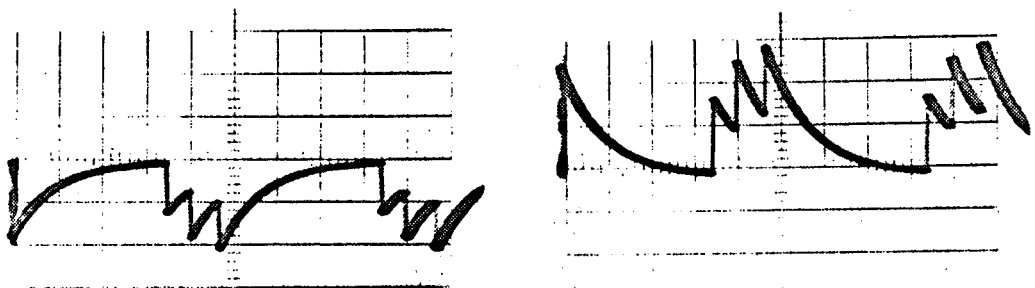


Fig.11.13. The time summed synaptic potentials for the input of three successive spikes. A scale gives 1 volt for the ordinate and 10 msec for the abscissa

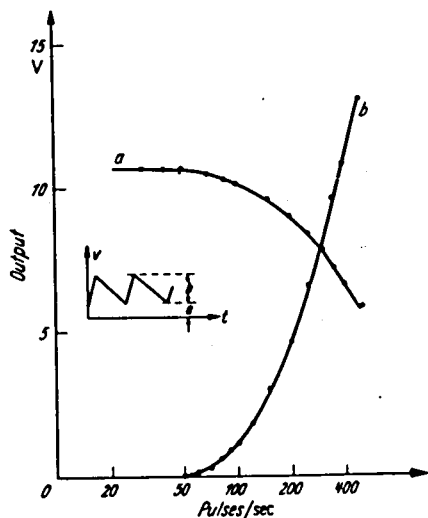


Fig. 11.14. The relation between the time summed synaptic potential and pulse frequency of the steady input spikes

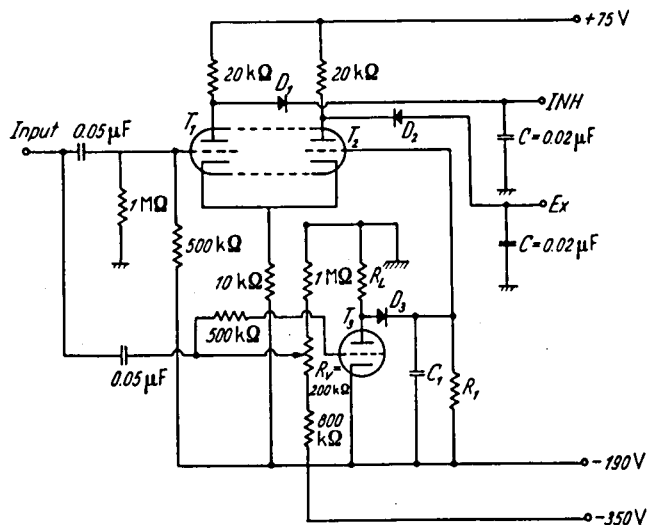


Fig. 11.15. The circuit of the incremental and decremental synapse models. The parameters are taken as follows; in the case of the

incremental synapse $R_L = 50K$, $R_1 = 500R$, $C_1 = 0.5\mu F$; in the case of the decremental synapse $R_L = 100K$, $R_1 = 200K$, $C_1 = 0.2\mu F$

i) The circuit of the ordinary model synapse.

The circuit of the ordinary type synapse unit is shown in Fig.11.11. From a synapse unit we may get both the excitatory and the inhibitory outputs (from EX and INH terminal respectively). The magnitude of the output potential may be changed by adjusting the volume R_v . The wave form of the synaptic potential is shown in Fig.11.12.

ii) The time summation.

For the closely spaced series of the input spikes the output synaptic potentials overlap and are summed up (the time summation). Fig.11.13 shows the time summed synaptic potential. The larger the pulse frequency of the input spikes, the larger the average level of the time summed synaptic potential. Fig. 11.14 shows how the D.C. voltage arises as the result of the time summation (it is given as the level of the bottoms of the valleys between the peaks of the synaptic potential) and the height of the synaptic potential (it is given as the height of the peaks of the synaptic potential from the bottom of the valley between the peaks of the synaptic potential to the top of the peak) change for the pulse frequency of the steady spike input. As seen from Fig.11.14 for the low frequency side each peak of the synaptic potential has clear individuality, but for the high frequency side the time summed D.C. potential becomes dominant. This suggests that the character of the integration of the inputs in the neuron is different for the low frequency

inputs and for the high frequency inputs.

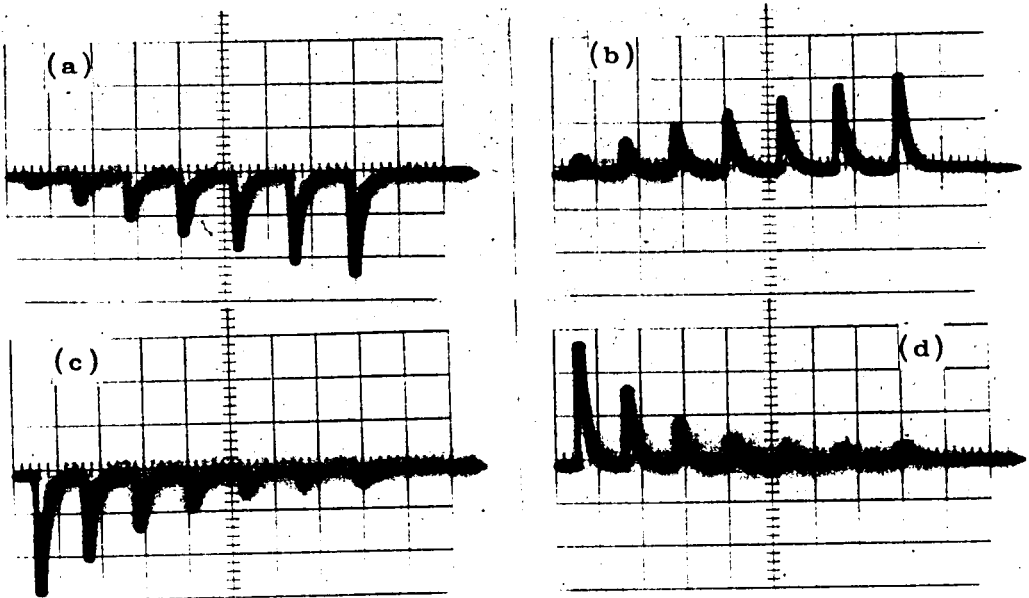


Fig.11.16. The wave forms of the incremental and decremental synaptic potentials for 7 successive input spikes. a) incremental excitatory, b) incremental inhibitory, c) decremental excitatory and d) decremental inhibitory. A scale gives 3 volts for the ordinate and 6 msec for the abscissa

iii) The incremental and the decremental synapses.

Fig.11.15 shows the circuit of the incremental synapse. The grid voltage of the tube T_3 is set in such a way that T_3 is on in the rest state. When a spike input comes T_3 becomes temporally off and the capacitor C_1 is charged up so that the grid voltage of T_3 is increased. When a series of spikes comes in, the grid voltage of T_2 increases gradually so that the magnitude of the output synaptic potentials increases, too.

The circuit of the decremental synapse is almost the

same as Fig.11.15 except that the grid voltage of T_3 is set in such a way that T_3 is off in the rest state. When a spike input comes, the aftershoot of the differentiated input to T_3 makes T_3 on and C_1 is discharged, so that the grid voltage of T_2 is decreased and the output for the next spike input is decreased. The wave form of the outputs of the incremental and the decremental synapses for a series of input spikes are shown in Fig.11.16.

The incremental and the decremental synapses seem to play important roles in the processing of the time pattern of the spike input. Roughly speaking, the incremental synapse may be regarded to perform the integration and the decremental synapse the differentiation. If we regard a neuron as a vacuum tube with many grids whose inputs and outputs are the continuous "information" coded by the spike density, then these synapses may play essential roles in the relaxation type performance of the nervous net.

BIBLIOGRAPHY

- (1) Akers, S. B., "On the algebraic manipulation of majority logic," IRE Trans. Electron. Computers, EC-10, 1961, pp. 779.
- (2) Amarel, S., Cooke, G. and Winder, R. O., "Majority gate networks," IEEE Trans. Electron. Computers, EC-13, 1964.
- (3) Arimoto, S., "The periodical transition of states realized by the threshold autonomous network," Technical Report of the Professional Group on Automaton and Automatic Control of IECEJ*, March, 1963.
- (4) Breuer, M. A., "Implementation of threshold nets by integer linear programming," IEEE Trans. Electron. Computers, EC-14, December 1965, pp.950-952.
- (5) Bullock, T. H., "Neuron doctrine and electrophysiology," Science 129, 997(1952).
- (6) Cardwell, S. H., "Switching circuits and logical design," John Wiley and Sons, New York, 1958.
- (7) Cameron, S. H., "The generation of minimal threshold nets by an integer program," IRE Trans. Electron. Computers EC-13, June 1964, pp. 229-302.
- (8) Caianiello, E. R., "Outline of a theory of thought-processes and thinking machines," Journal of Theoretical Biology, April 1961, pp. 204-235.
- (9) Chow, C. K., "Boolean functions realizable with single

- threshold devices," Proceedings of the IRE, vol. 49, number 1, January 1961, pp. 370-371.
- (10) Coates, C. L. and Lewis, P. M., "Linearly separable switching functions," Journal of the Franklin Institute, vol. 272, November 1961, pp. 360-410.
 - (11) Coates, C. L. and Lewis, P.M., "A realization procedure for threshold gate networks," IEEE Trans. Electron. Computers, EC-12, 1963, pp. 456-461.
 - (12) Cohn, M. and Linsdaman, R., "Axiomatic majority-decision logic," IRE Trans. Electron. Computers, EC-10, 1961 pp. 17-21.
 - (13) Dertouzos, M. L., "An approach to single-threshold-element synthesis," IEEE Trans. On Electron. Computers, vol. EC-13, number 5, October 1964, pp.519.
 - (14) Dertouzos, M. L., "Threshold logic: A synthesis approach," MIT Press, Cambridge, Mass., 1965.
 - (15) Eccles, J. C., "The physiology of nerve cells," Baltimore, 1957.
 - (16) Fan, K., "On systems of linear inequalities," Linear Inequalities and Related Systems, Annals of Mathematics Studies NO. 38, Princeton, 1956.
 - (17) Fujiwara, M., "On the system of linear inequalities," Report of Japan Academy, vol. 4, 1928.
 - (18) Fujiwara, M., "Matrix and determinant," Iwanami Press.

- (19) Fukutome, H., Tamura, H. and Sugata, K., "An electric analogue of the neuron," Kybernetik, Band 2, September 1963, pp. 28-32.
- (20) Fukutome, H., Tamura, H. and Sugata, K., "The model of the neuron," Science of Living Body, vol. 15, number 1, Feb. 1964, pp. 15-21.
- (21) Gabelman, I. J., "The synthesis of Boolean function using a single threshold element," IEEE Trans. Electron. Computers, EC-11, 1962, pp. 639- 642.
- (22) Gabelman, I. J., "An algorithm for threshold element analysis," IEEE Trans. Electron. Computers, EC-14, August 1965, pp. 623-625.
- (23) Gaston, C. A., "A simple test for linear separability," IRE Trans. Electron. Computers, EC-12, April 1963, pp. 134-135.
- (24) Gonzalez, R., "Synthesis problems in linear threshold logic," Doctoral Dissertation, Michigan University, June 1966.
- (25) Golomb, S. W., "On the classification of Boolean functions," IRE Trans. Circuit Theory, CT-6, 1959, pp. 176-186.
- (26) Goes, E. E., "An adaptive network for producing real function of binary inputs," Information and Control, vol. 8, number 11, 1965, pp. 111-123.
- (27) Guterman, S., Kodis, R. D. and Ruhman, S., "Logical and control functions performed with magnetic cores,"

Proceedings of the IRE, vol. 43, number 3, March 1955, pp. 291-298.

- (28) Hansen, W. H., "Threshold logic synthesis by algebraic methods," IRE Trans. Electron. Computers, EC-12, August 1963, pp. 401-402.
- (29) Highleyman, W. H., "A note on linear separation," IRE Trans. Electron. Computers, EC-10, number 4, December 1961, pp. 777-778.
- (30) Horna, O. A., "Majority-logic synthesis," IEEE Trans. Electron. Computers, EC-2, number 2, April 1963, pp. 131-132.
- (31) Harmon, L. D., "Studys with artificial neurons," Kybernetik, 89, 1961.
- (32) Hodgkin, A. L. and Huxley, A. F., "A quantitative description of membrane current and its application to conduction and excitation in nerve," Journal of Physiol., 1952.
- (33) Hartline, N. K. and Graham, C. H., "Nerve impulses from single receptor in the eye," Journal of Cell Comp. Physiol., 277(1932).
- (34) Hu, S. T., "Threshold logic," Berkeley: University of California Press, 1965.
- (35) Hopcraft, J. E. and Mattson, R. L., "Synthesis of minimal threshold logic network," IEEE Trans. Electron. Computers. EC-14, August 1965, pp. 552-560.

- (36) Ishida, H. and Stewart, R. M., "A learning network using threshold elements," IRE Trans. Electron. Computers, EC-14, number 3, June 1965, pp. 481-487.
- (37) Ishii, N. and others, "The periodical vibration mode in the threshold element network," Technical Report of Professional Group on Automaton and Automatic Control of IECEJ*, July, 1965.
- (38) Iyanaga, M. and Sugiura, M., "Algebra for the technical mathematician," Iwanami Press.
- (39) Kaszerman, P., "A geometric test-synthesis procedure for a threshold device," Information and Control, 6, 1963, pp. 381-398.
- (40) Lewis, P. M. and Coates, C. L., "Threshold logic," John Wiley and Sons, Inc., 1966.
- (41) Mays, C. H., "Effects of adaptation parameters on convergence time and tolerance of adaptive threshold elements," IEEE Trans. Electron. Computers (Short Note), EC-13, number 3, August 1964, pp. 465-468.
- (42) McNaughton, R., "unate truth functions," IRE Trans. Electron. Computers, EC-10, March 1961, pp. 1-6.
- (43) Minnick, R. C., "Linear input logic," IRE Trans. Electron. Computers, EC-10, number 1, March 1961, pp. 6-16.
- (44) Miyata, F., "Realization of arbitrary logical functions using majority elements," IEEE Trans. Electron. Computers, EC-12, 1963, pp. 183-190.

- (45) Muroga, S., Toda, I. and Takasu, S., "Theory of majority decision elements," Journal of the Franklin Institute, vol. 271, May 1961, pp. 376-418.
- (46) Nikaido, F., "Linear algebra for economists," Baifukan Press.
- (47) Nagumo, J. Arimoto, S. and Yoshizawa, S., "An active pulse transmission line simulating nerve axon," Proc. IRE No.10, 1962.
- (48) Peterson, W. W., "Error correcting codes," MIT Wiley.
- (49) Paull, M. C. and McCluskey Jr., E. J., "Boolean functions realizable with single threshold devices," Proc. IRE, 1960, pp. 1335-1337.
- (50) Quine, W. V., "A way to simplify truth functions," Am. Math. Monthly, 1955, pp. 627-631.
- (51) Rosenblatt, F., "Perceptron simulation experiments," Proc. IRE, 48, 1960, pp. 301-309.
- (52) Rushton, W. A., "Peripheral coding in the nervous system in sensory communication," MIT Tech. Press John Wiley and Sons, 1962.
- (53) Sakai, T. and Sugata K., "An analysis of the threshold element network," Technical Report of the Professional Group on Automaton and Automatic Control of IECEJ*, Nov. 1965.
- (54) Sakai, T. and Sugata K., "An analytical method for the threshold element network," Record of the 1965 Joint

Convention of the Kansai District of IECEJ*.

- (55) Sakai, T. and Sugata, K., "The synthesis method of the threshold element network," Technical Report of the Professional Group on Automaton of IECEJ*, May 1967.
- (56) Sakai, T. and Sugata, K., "The method of for the analysis of the threshold element network," Record of the 1967 Joint Convention of IECEJ*.
- (57) Sakai, T. and Sugata, K., "An analysis of the threshold element network emphasized on the linear separability," Technical Report of the Professional Group on Automaton of IECEJ*, July 1967.
- (58) Sheng, C. L., "A method for testing and realization of threshold functions," IEEE Trans. Electron. Computers, EC-13, 1964, pp. 232-239.
- (59) Sheng, C. L., "Compound synthesis of threshold-logic networks for the realization of general Boolean functions," IEEE Trans. Electron Computers, EC-14, December 1965, pp. 798-814.
- (60) Tohma, Y., "Decomposition of logical functions using majority decision elements," IEEE Trans. Electron. Computers, EC-13, December 1964, pp. 698-705.
- (61) Urbano, R., "Boolean matrices and the stability of neural nets," IEEE Trans. Electron. Computer, Ec-12, number 2, April 1963, pp. 61-66.

- (62) Winder, R. O., "Enumeration of seven-argument threshold functions," IEEE Trans. Electron. Computers, vol. 14, number 3, June 1964, pp.315.
- (63) Winder, R. o., "Bounds on threshold gate realizability," IEEE Trans. Electron. Computers, vol. EC- 12, No. 5, October 1963, pp. 561.
- (64) Yajima, S. and Ibaraki, T., "A lower bound on the number of threshold functions," IEEE Trans. Electron Computers, EC-14, December 1965, pp. 926.

*IECEJ : The Institute of Electrical Communication
Engineers of Japan.