
学術情報センター紀要

Research Bulletin of
the National Center for Science Information System

第 3 号

1990年9月

学術情報センター

学術情報センター紀要 第3号

目 次

巻頭のことば	山田 尚勇 (学術情報センター研究開発部長)……………	i
研究論文		
韓国の文字コードについて	宮澤 彰 (学術情報センター)……………	1
韓国標準規格と日本工業規格の漢字について		
	李 春 澤 (韓国公州大学：富士大学)……………	21
文献の論理構造を考慮した全文検索システム		
	影浦 峽，大山敬三，宮澤 彰，根岸正光 (学術情報センター)	
	鳥居俊一，絹川博之 (日立製作所)……………	49
ドキュメント・デリバリーのための電子図書館		
～リソースシェアリングをめぐる制度的枠組み～		
	桂 英史，影浦 峽 (学術情報センター)……………	59
知識ベースシステムにおけるタスクの役割	小山照夫 (学術情報センター)……………	75
機械概念設計プロセスのモデリング	小山照夫 (学術情報センター)……………	87
大規模関係データベースのための並列処理マシンの一構成法		
	濱田 喬 (学術情報センター)，大久保一彦 (東京大学)……………	97
放送形トラヒックに対する高速パケット通信網のルート制御		
	浅野正一郎 (学術情報センター)，安藤史郎 (東京大学)……………	113
日米の研究開発体制を考える	山田尚勇 (学術情報センター)……………	139
テスト評価、均質的文化、独創性養成	山田尚勇 (学術情報センター)……………	161

**Research Bulletin
of
The National Center for Science Information System**

September 1990 Volume 3

Contents

<i>H. Yamada</i>	i	Preface
 Contributions		
<i>A. Miyazawa</i>	1	On The Korean Character Codes
<i>C. T. Lee</i>	21	Comparative Analysis of Character Set/Code between Korean and Japanese Standard
<i>K. Kageura and K. Oyama A. Miyazawa M. Negishi S. Torii H. Kinugawa</i>	49	A Full-text Retrieval System Reflecting the Logical Structure of Documents
<i>E. Katsura and K. Kageura</i>	59	A Conceptual Framework for Document Delivery Services Based on Electronic Library
<i>T. Koyama</i>	75	Roles of Tasks in Knowledge-based Systems
<i>T. Koyama</i>	87	Process Modeling of Mechanical Conceptual Design
<i>J. Hamada and K. Okubo</i>	97	A Multiprocessor Database Machine for Very Large Relational Database
<i>S. Asano and S. Ando</i>	113	Routing Control for Multi-destination Traffic
<i>H. Yamada</i>	139	On Research and Development Activities in U.S.A. and in Japan : a personal view
<i>H. Yamada</i>	161	Testing Procedures, Homogeneous Culture, and Creativity Cultivation : on selection procedures in education and in research

巻 頭 の こ と ば

学術情報センター研究開発部長

山田 尚勇

全国の国公立大学および大学共同利用機関を結ぶ、我が国の総合的な学術情報システムの中核機関として、1986年4月に文部省によって創設された学術情報センターは、人文、社会、自然科学の全分野にわたって、学術情報の収集、整理、および提供を行なう責務を課せられている。

それと並行して、現在実用化が行なわれつつあり、また近未来に実用化の可能性を予測できる情報処理や通信の最新技術をセンターの業務にすみやかに活用できるように、学術情報や情報システムに関する総合的な研究、開発を行なうことも、また本センターに課せられた責務である。

学術情報センターは、創設以来、センターの内外の頭脳を結集して、これらセンターの必要とする研究開発課題と取り組んできた。

学術情報センターのそうした研究開発活動の成果を報告する本紀要は、今回その第3号を発行する運びとなった。編集委員会を通し、各専門分野において指導的地位を占める専門家の査読を経てここに採録された10篇の論文は、学術情報の形成、表現、蓄積、伝送サービスなどに関する方法論、システム構成など、広い分野にわたって、最新の成果の一端を示しているものである。

今回新しく見られる特徴としては、国際的な問題を明からさまに取りあげた論文が数篇含まれていることである。現在の学術研究は、国際的視野に立つ情報交換と、それに立脚した研究手法の活用なくしては、遂行不可能になっている。したがって、わが学術情報センターにおいても、アメリカ、イギリスを手始めにして、相互の情報交換を実施に移してきた。これらの国際協力をさらに東アジアに拡大すべく、現在、中国、韓国との国際共同研究を実施しつつある。今回の発表のいくつかはその第一弾とも言うべく、今後も国際的視野に立つ研究が、ますます紀要に反映されてくることになるであろう。諸賢のご支援とご教示を切望する次第である。

研 究 論 文

韓国の文字コードについて

On The Korean Character Codes

学術情報センター 宮澤 彰

要 旨

韓国の文字コードについてコードスペースと文字セットの観点から紹介し、東アジアの文字データベース交換についての問題を考察する。コードスペースとは文字を表現するのに使用されるコード値の集合として定義する。文字コードは一般にコードスペースとそこに対応づけられる文字セットをあわせて言っている。この観点から文字コード拡張の基本的な標準である ISO 2022 を整理して紹介する。また、学術情報センターのサービスシステムの環境である N-1 の NVT コード、日立の VOS 3 KEIS また一般ユーザに最も普及しているシフト JIS コードについても、このコードスペースと文字セットの観点から整理する。韓国の文字セットについてはハングルの文字コードとして、N バイトコード、組合型、完成型の各タイプを、IBM のホストコード、標準コードの付録の例などを示して紹介する。また急速に普及しつつある標準コード KSC 5601-1987については、その構成や特徴、指摘されている問題点などを紹介して解説する。これらを使用するコードスペースについては、IBM ホスト、標準的な KSC 5601 の使用法、シフト JIS と似た形での拡張されたコードスペースなどの例を紹介する。いずれも文字セットの呼び出しは固定的である。しかし東アジアでの文字セットを扱わねばならず、コードスペースを広げるか、文字セットの動的呼び出しを行なうかをしなければならない。これは現行のシステムではできず解決しなければならない問題となる。

ABSTRACT

This paper introduces Korean character codes used in the Republic of Korea, and discusses interchange of the East Asian text databases. A concept of "code space" is introduced as a set of code values used for representation of graphic characters. Character sets are invoked into code space to form a character encoding system. ISO 2022, which is a basic standard for the character coding system, is explained from the viewpoint of code space. Character code environment of the NACSIS information services, which are N-1 NVT code, KEIS code of VOS3, and a widely used personal computer systems character code shift-JIS are analyzed. Then hangul codes of the Korean character sets are classified into three types, which are N-byte code, combination type

and complete type. Examples of these types are shown. Features of the rapidly disseminating standard code KS C 5601-1987 is explained, and some discussions about this character set are introduced. As the code space of these character sets, following examples are shown — IBM host code, the standard usage of the KS C 5601, and some extension methods influenced by the Japanese shift-JIS code. Invocations of character sets into these code spaces are all static and fixed.

However, for the interchange of the East Asian text database, more than four character sets are required, and larger code space or dynamic invocation of character sets is mandatory. It is new problem that existing operating systems can not handle such large code space or dynamic invocation.

1. はじめに

「NACSIS-CAT では文字の問題で中国語や韓国語の本が扱えない」ということになっている。しかし、中国語や韓国語を入出力できるシステムは日本にも他にあるし、パーソナルコンピュータでも中国語や韓国語のワープロが実現され、市販されている。このような状況からすれば NACSIS-CAT のような大きな資源を使用するシステムで中国語や韓国語の文字を扱えないわけではない筈である、と考えるのは自然なことであろう。

もちろん学術情報センターでもこのための検討は行なっているわけであるが、例えば韓国製の端末を買ってきてつなげば韓国語も扱えますとか、韓国語用ワープロのシステムを少し変更すれば韓国語用目録端末ができます、といったようなわけにはいかない。こういった入出力ハードウェアの問題に加え、コンピューターシステム（ホスト系、端末系）のオペレーティングシステムの問題、使用すべき文字コード、各国標準コードや国際標準コードの問題が絡んでくるためである。さらに NACSIS-CAT という環境に関して言えば、China-MARC や KOR-MARC といった書誌データベースをどのように利用できるかといった側面も上記の問題に絡んで来る。

科学研究費による国際共同研究「東アジア文字データベースの国際交換に関する実証研究」では特に書誌データベース交換の観点からこれらの問題点の解決方法を探っているが、この一環として 1989 年 12 月から 1990 年 1 月まで韓国を訪問し関連する状況を調査する機会を得た。

本稿はこの調査の結果の一部として韓国の文字コードの現況を報告するとともに、多言語システムの文字コードの問題点を探るものである。

2. コードスペースと文字セット

文字コードの状況について説明するために、本稿では「コードスペース」とその中に位置づけられる「文字セット」という概念を導入して整理を行なう。

なお、本稿で「文字」という語は一般的な意味での文字、すなわちアルファベットやカナ、漢字、

文中に用いられる句読点その他の記号類をさす。文字コードの用語としては、この一般的な意味での文字を「図形文字」graphic character、改行やバックスペースなどの制御機能を表現するものを「制御文字」と呼び、これらの総称として「文字」character の語を使うことが多いが、本稿では特に断らない限り、文字に制御文字を含めない。

また、以下文字コードの値（ビット組合せ）を表現する場合、16進数を使用する。また、{ } は集合を表わす。例えば {20, 21, 22} は文字コード値 20、21、22 の3つの要素からなる集合である。この略記として {20~22} を用いる。また、集合のユニオン U 及び直積×の記号を用い、 $A \times A$ を A^2 、 n 回の直積を A^n と略記する。注算子×は U に優先する。

2. 1 コードスペース

オペレーティングシステムまたはあるソフトウェアシステムで、文字を表現するのに使用されるコード値（ビット組合せ）の集合をコードスペースと呼ぶことにする。

漢字処理がオペレーティングシステムレベルの機能として取り込まれる以前は、1文字は1バイトで表現されるという方法が通常であった。但し1バイトの {00~FF} すべてがコードスペースだったわけではなく、低位のいくつかは制御文字であった。例えば IBM 系のメインフレーム OS では {40~FF} を文字コード空間として使用し、いわゆる ASCII コード使用の OS では {20~7E}、また JIS X 0201 を使用している場合、{20~7E, A1~FE} をコードスペースとして使用していた。

コードスペースは言わば容器であって、中に何が入るかは別の問題である。IBM 系の {40~FF} のコードスペースには EBCDIC と呼ばれる文字セットを割り当てて使用することが多いが、日本ではアルファベット小文字の代りにカタカナをいれた EBCDIK という別の文字セットも使用している。システムプログラムでもアプリケーションプログラムでも、かなりのものはあるデータ項目が文字であるかどうかだけが問題で、中の文字自体が何であるかは関知していない。このようなプログラムではコードスペースが同じであれば割当てられた文字セットが何であるかに関係なく動く。

逆に文字セットが同じであっても異なった文字コード空間に割当てられる場合もある。後に述べるように JIS X 0208 情報処理用漢字符号系はメインフレーム用 OS の2バイト文字コード空間や、パーソナルコンピュータ用のシフト JIS と呼ばれる文字コード空間で使用されている。

2. 2 文字セット

本稿では文字セットを、「文字の順序づけられた集合」と定義する。この順序づけられた集合の元には特定の文字に対応しない「未定義文字」を複数個含んでもよい。

JIS 或は ISO の中では「符号化文字集合」coded character set を「文字集合を定め、かつその集合内容の文字とビット組合せを1対1に関係づけるあいまいでない規則の集合⁽¹⁾」と定義している。この定義の符号化文字集合は文字と特定のビット組合せが対応づけられているが、本稿で使う文字セットは特定のビット組合せは前提にしない。また、順序を持っている点で JIS の定義中の文字集合（単に文字の集合の意味）とも異なる。

文字セットをこのような意味で使うのは、いわゆるシフト JIS コードのように、ビット組合せと

の対応を全く変えて使う例があるためである。シフト JIS の文字セットは未定義部分を含め順序を全く保存しているので JIS の文字セットであると言えるが、ビット組合わせは全く異なるので「符号文化文字集合」としては JIS と等しいとは言えない。

また、文字の集合としては等しくても、文字セットとして等しくない場合もある。学術情報センターの CAT、IR 用ホストシステムでは、1 バイトの内部コードスペースとして {40~FF} を使用し、文字セットとしては EBCDIK をベースに英小文字を入れたものを用いている。この文字セットは文字の集合としては JIS X 0201 と等しくなるようにしているが、順序は全く異り、文字セットとしては同じでない。

上記のように文字セットとコードスペースを分離する考え方は当然とも言えるが、実際のシステムでは別のものとして意識されていることは少い。文字セットは特定のコードスペースを意識して作られているし、また文字セットの対応は殆ど固定的で変更できない場合が多いためである。

2. 3 ISO 2022 コード拡張法

1 つのコードスペースに複数の文字セットを対応させるための標準規格として ISO 2022 がある。この規格は制御文字、図形文字の双方についての拡張方法として文字コードについての基本的な規格である。JIS では X 0202 情報交換用符号の拡張法となっている。もっとも内容はかなり複雑で、この規格を完全に満たすような現実のシステムはおそらく存在しない。

ここではコードスペースと文字セットという観点から、この規格の図形文字の拡張法について概観してみる。なお、ISO 2022 は 1986 年版を使用する⁽²⁾。これは JIS X 0202-1984 の対応している版より新しい。

2. 3. 1 文字セット

ISO 2022 では文字セットは 94 乃至 96 個の文字（未定義を含む）、又は 94^n 乃至 96^n 個 ($n \geq 2$) 文字からなり、{21~7E} (94 個文字セット) 乃至 {20~7F} (96 個文字セット) のビット組合わせに対応づけられる。複数バイト文字セット ($94^n, 96^n$) は、{21~7E}ⁿ、{20~7F}ⁿ に対応づけられる。

これらの文字セットは 1 バイト (94 乃至 96 文字) か複数バイト (94^n 乃至 96^n 文字) かと、終端文字と呼ばれる 1 バイト (30~7E、うち 30~3F は私用) で識別される。複数バイトセットでは終端文字 40~5F が 2 バイト、60~6F が 3 バイト、70~7F が 4 バイト以上と分けられている。

文字セットは登録制で、登録時に終端文字が割当てられる。例えば 1 バイト文字セットの終端文字 49 は JIS X 0201 カナ文字セット、複数バイト文字セットの終端文字 42 は JIS X 0208-1983 という具合になっている。

従って、私用の文字セットを別にすれば、ISO 2022 の文字セットは 1 バイト文字セットが 63 個、2 バイト文字セットが 32 個、3 バイト文字セットが 16 個、4 バイト (以上) 文字セットが 15 個というのが最大構成となる。(今後の拡張のない限り)。

2. 3. 2 コードスペース

ISO 2022 では 7 単位符号系と 8 単位符号系 (7, 8-bit environment) という 2 つのコードスパー

スが考えられているが、話を単純にするために以下の話を 8 単位符号系に限定し、さらに 3 バイト以上の文字セットはないことにする。

8 単位系のコードスペースは {20~7F, A0~FF} である。ただし、これはこのスペースに「呼び出し」to invoke された文字セットが 1 バイトの場合で、2 バイト文字セットが呼出されるとコードスペースも変化する。

「呼び出し」は {20~7F} (これを以降 G_L とする)、{A0~FF} (G_R とする) に独立に行なわれる。 G_L に呼び出された文字セットが 1 バイトで、 G_R に呼び出された文字セットが 2 バイトであれば、コードスペースは {20~7F} U {A0~FF}² となる。他の可能性をあげれば、{20~7F}² U {A0~FF} (G_L : 2 バイト、 G_R : 1 バイト)、{20~7F}² U {A0~FF}² (G_L 、 G_R とも 2 バイト)。

このように呼び出しの状況に応じてコードスペースが変化するのがこの規格の特徴である。また、{20~7F} × {A0~FF} や {A0~FF} × {20~7F} はコードスペースに含まれることはない。

2. 3. 3 $G_0 \sim G_3$ セット

さらにこの規格を複雑なものにしているのは、2. 3. 1 に述べた文字セットが直接 G_L 、 G_R に呼び出されるのではなく、一旦 $G_0 \sim G_3$ というセットに「指示」to designate され、 $G_0 \sim G_3$ が呼び出されるという二重構造になっている点である。

表 1 ISO 2022 の指示

$G_0 \sim G_3$	バイト数	文字数	エスケープ	シーケンス
G_0	1	94	ESC	28 F
	複	94	ESC	24 28 F
G_1	1	94	ESC	29 F
		96	ESC	2D F
	複	94	ESC	24 29 F
		96	ESC	24 2D F
G_2	1	94	ESC	2A F
		96	ESC	2E F
	複	94	ESC	24 2A F
		96	ESC	24 2E F
G_3	1	94	ESC	2B F
		96	ESC	24 2B F
	複	94	ESC	2F F
		96	ESC	24 2F F

F は終端文字。

指示は文字セットを終端文字で指定し、 $G_0 \sim G_3$ のどれとするか、1 バイトか複数バイトか、94 文字か 96 文字かによって定められたエスケープシーケンスを用いる。このうち 96 文字セットは G_0 としては指示できない。表 1 に指示シーケンスをまとめる。

呼び出しは $G_0 \sim G_3$ のどれを、 G_L に呼び出すか G_R に呼び出すかにより、定められたシフト機能を用いる。ただし G_0 は G_R には呼び出せない。また、 G_2 、 G_3 についてはシングルシフトと呼ぶ1文字(1バイトとは限らない)呼出しもある。SS 2 又は SS 3 シフト機能に続く1文字(G_2 又は G_3 文字セットのバイト数)が G_L に呼び出される。以上の呼び出しを表2にまとめる。

表2 ISO 2022のシフト機能(8単位符号)

ロック/シングル	セット	呼び出し先	名称	符号表現
ロ ッ キ ン グ	G_0	G_L	LS 0	0F
	G_1	G_L	LS 1	0E
		G_R	LS 1 R	ESC 7E
	G_2	G_L	LS 2	ESC 6E
		G_R	LS 2 R	ESC 7D
	G_3	G_L	LS 3	ESC 6F
G_R		LS 3 R	ESC 7C	
シ ン グ ル	G_2	G_L	SS 2	8E
	G_3	G_L	SS 3	8F

2. 3. 4 ISO 2022 の使用

上記のように図形文字コードの拡張だけを見ても ISO 2022 はかなり複雑な内容を持っている。これを完全に実現するシステムはないとは言え、通信の世界をはじめ、IBM 系メインフレーム以外のコンピュータの内部コードなど、多くのシステムはコード拡張についてこの規格のサブセットに準拠している。さらに文字セットも JIS X 0208 漢字符号や中国の GB 2312 文字セット等、多くはこの規格に従った構成をとっている。

この規格では指示、呼び出しが変更になるとコードスペースも変化する可能性があり、このことは多くのソフトウェアにとって厄介を引きおこす。このため多くのシステムでは指示・呼び出しを固定した状態で使用している。こうすれば、指示のためのエスケープシーケンスも呼び出しのシフト機能も省略できる。一般に G_0 としては ISO 646 或はこれとほぼ互換の文字セットを指示してこれを G_L に呼び出した状態を基本としている。

これに加え、 G_1 として、1バイトあるいは2バイトの文字セットを指示し G_R に呼び出した状態をデフォルトとする方法が多く使われている。JIS X 0201 のローマ字・カナを使う場合もこれで、 $G_0 = G_L$ 、 $G_1 = G_R$ が保たれ、指示・呼び出しの変更は行なわないですむ。

3つ或は4つの文字セットを使う場合は指示は固定できるが、呼び出しの変更が必要となる。JIS X 0201 の1バイトローマ字、カナに加え JIS X 0208 の漢字文字セットを使用したい場合はこれにあたる。 G_2 又は G_3 を使用せざるを得ないため、シフト機能を使ってコードスペースと文字セットの対応を動的に変更する必要がある。このようなシステムは実現されてはいるが、まだ数は多くない。

さらに4つをこえる文字セットを使用する場合、これは指示の変更が必要となってくる。仕様としてはともかく、現実そのようなシステムは動いていないと思う。ただし、漢字端末と呼ばれる世界では、ISO 2022の規格にG₂、G₃が導入される以前に製造されたこともあり、JIS X 0208に対して指示し呼び出す方法が使われており、通信ソフト等でも普及している。

3. 学術情報センターの文字コード環境

韓国の文字コードの現況を見る前に簡単に日本のシステムの文字セットとコードスペースを紹介する。データ交換を考える時、送り出し側としても受取側としても日本側、特に学術情報センターの状況は常に視野に入れておく必要があるためである。

ここでは学術情報センターのサービスに使用されている環境としてN-1の拡張NVTコード、またサービスを支えているセンター側のホストコンピュータの内部コードの例としてKEISコード、及びIR等のユーザ側の内部コードとして広く利用されているシフトJISコードについて取り上げる。

3.1 N-1 拡張NVT

学術情報センターのサービスでは日本のほとんどのコンピュータシステムと同じく、JIS X 0201のローマ字及びカナ文字セット、JIS X 0208の漢字文字セットを使用している。またNACSIS-CATではこれに加えてウムラウトやアクセントなどのついた拡張文字セット(EXC)も使用している⁽³⁾。

これらをサービスする通信上でのコードスペースは接続時の環境により異なるが、大学間ネットワークN-1のNVTコードにおける文字コード(基本コード、漢字文字コード)のコードスペースを見よう⁽⁴⁾。

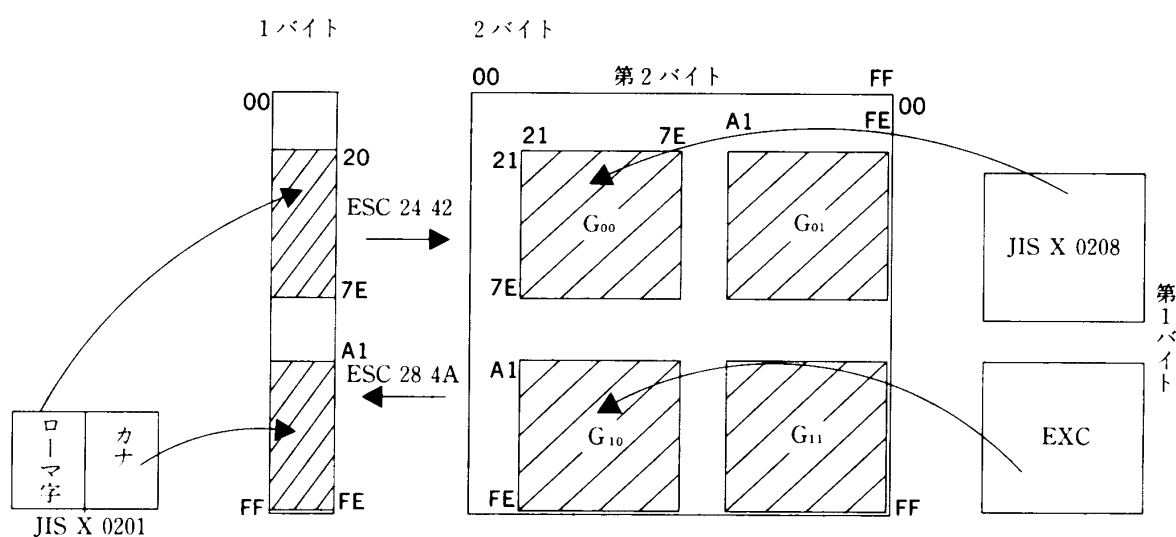


図1 N-1 NVT コードのコードスペースと文字セット

NVT のコードスペースは $\{20\sim 7E, A1\sim DF\} \cup \{21\sim 7F, A1\sim FE\}^2$ である (図 1)。1 バイトのコードスペースと 2 バイトのコードスペースの切換えはエスケープシーケンスを用い、ESC24 42 で 2 バイトに、ESC28 4A で 1 バイトになる。1 バイトのコードスペースには JIS X 0201 ローマ字、カナ文字セットを固定し、2 バイトでは図 1 の G_{00} に JIS X 0208-1983 を、 G_{10} に EXC を固定する。

この方法は ISO 2022 と異り、8 ビット 2 バイトを前提としている。このためコードスペースが ISO 2022 の場合より広くとれることになる。その代りに、文字セットは固定しているにもかかわらず、1 バイト 2 バイトの切替のための制御機能が必要となっている。この切替には指示のエスケープシーケンスを準用している (なお、N-1 のこの規格を決めた際に参照した JIS X 0202 は現在の ISO 2022 より古い版である)。

ところで、N-1 のこの切換シーケンスが、コードスペースを 1 バイト、2 バイトに切換える機能を持つのか、それとも文字セットを切換えてそれがたまたま 8 ビット 1 バイトや 8 ビット 2 バイトであったのかという考え方の相違がここで問題になってくる。もちろん、N-1 の規格では文字セットがこれしかないことになっているのだから、どちらでも結果は同じであり、明確には示されていない。ただ、もし将来、35344 字を超える文字が必要となるとすれば関係してくるであろう。

一般にソフトウェアにとっては、コードスペースが明示されている方が負担が少ないように思われるが、どうであろう。

なお、NACSIS-CAT で使用している VTSS は、NVT コードのコードスペースを 7 ビットの世界に押し込むための独自の方法を使用しているが⁽⁶⁾、ここでは触れない。

3. 2 VOS 3 KEIS コード

KEIS は日立のメインフレームコンピュータのオペレーティングシステム VOS 3 の漢字情報処理機能である。NACSIS-CAT や NACSIS-IR も内部コードとしてはこの KEIS コードで処理されている。

KEIS では VOS 3 の基本文字セットである EBCDIK に加え、JIS X 0208、日立が選定した「拡張文字セット 3」、及びユーザ定義文字セットが使用できる (KEIS 83)。これらの呼出しはすべて固定的で、1 つのアプリケーションプログラム内で動的に切換えることはできない⁽⁶⁾。図 2 に KEIS のコードスペースを示す。

図のようにコードスペースは $\{40\sim FF\} \cup \{41\sim FE\}^2$ である。切り換えコードとして OA41 と OA42 を用い、OA41 が来ると 1 バイト、OA42 が来ると 2 バイトとする。

2 バイトのコードスペースのうち、現在のところ文字セットと対応づけられているのは $\{59\sim FE\} \times \{A1\sim FE\}$ であり、残りの場所は保留領域として将来の拡張用となっている。

学情センターの運用では 1 バイトの EBCDIK を JIS の X 0201 ローマ字、カナ両文字セットと、文字の集合として同じものとし、ユーザ定義文字セットに EXC をわりあて、KEIS 拡張文字セット 3 は使用していない。NVT コードでの EXC は $\{A1\sim FE\} \times \{21\sim 7E\}$ のコードスペースにわりあてられているが、文字のわりあてを $\{DF\sim FE\} \times \{21\sim 7E\}$ から始めることにして、内部コードの EXC 用

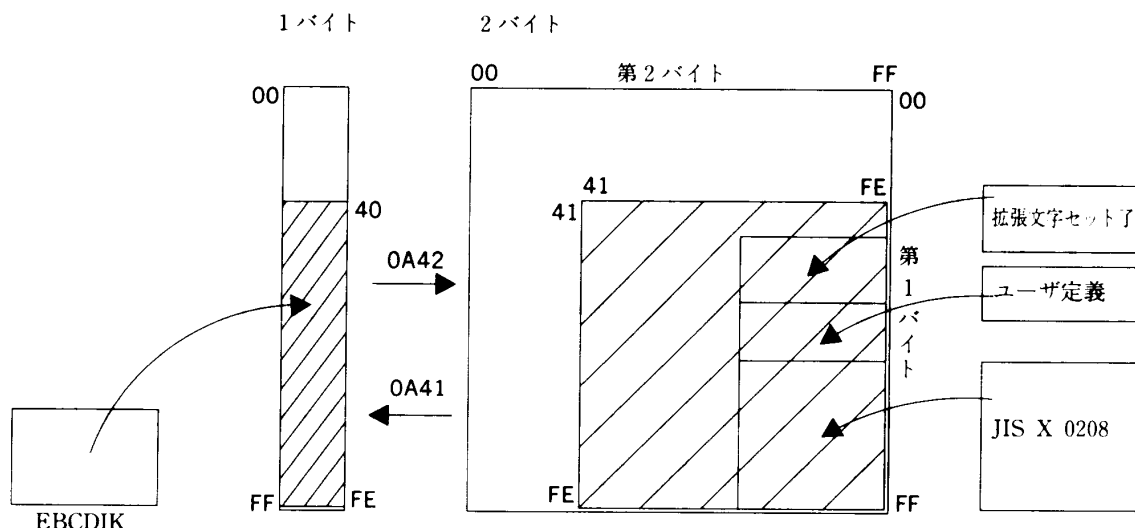


図2 VOS3 KEIS コードのコードスペースと文字セット

スペース {81~A0} × {A1~FE} と大きさをあわせてある。

3. 3 シフト JIS コード

シフト JIS はパーソナルコンピュータ用の OS として圧倒的なシェアを持つ日本語 MS-DOS で主として用いられている内部コードスペースである。MS-DOS 以外の OS でもシフト JIS コードを用いている例がある。

シフト JIS コードでも文字セットは JIS X 0201 ローマ字、カナ文字セットと JIS X 0208 漢字文字

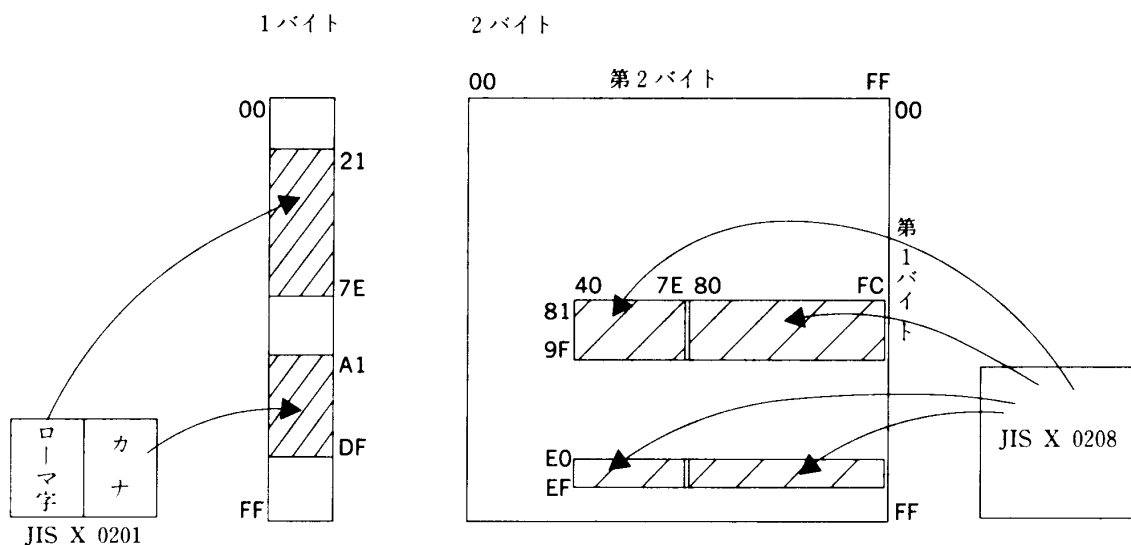


図3 シフト JIS のコードスペースと文字セット

セットを用いる。シフト JIS の特徴は、この3つの文字セットを用いながら呼び出しは固定で、しかも1バイト2バイトの切り換えが不要という点にある。

シフト JIS コードのコードスペースを図3に示す。図のように {21~7E, A1~DF} U {81~9F, E0~EF} × {40~7E, 80~FC} という変則的な形をしたコードスペースである。

2バイトスペースの第1バイト {81~9F, E0~EF} は47のポジション、第2バイト {40~7E, 80~FC} は188のポジションを持ち、ここに JIS X 0208 の94×94文字を納めることができるわけである。

このコードスペースでは1バイトのコードスペースと2バイトのコードスペースの第1バイトが分離されているため、切換制御文字なしで使用できる。

これらの利点に対し、欠点としては ISO 2022 の制御文字のコードスペースや JIS X 0201 カナ文字セットの未定義領域を文字コードスペースに使用したため、互換制の点で問題が出るという欠点もある。また、この方式ではコードスペースを大幅に拡張することはむづかしく、94×94の文字セットを2つ使用する事は不可能である。

4. 韓国の文字セット

韓国の文字セットは当然韓国語処理用としてハングル及びハンチャ（漢字）を含んでいる。漢字については常用漢字体のようなものはない点と、使用頻度が日本語より遥かに低い点を除けば基本的に日本語と同様の問題を持っている。韓国語固有の問題としてはハングルに関してであり、ここでは先にハングルの文字セットについて取り上げる。

4.1 ハングル

文字コードの観点からハングルの特徴をまとめておく。

ハングルは表音文字であり、現代の正書法ではその基本要素は24しかないが、1文字はこの要素が2~数個組み合わせられて成立つ。組み合わせは初声と呼ばれる子音と中声と呼ばれる母音、及びオプションな終声と呼ばれる子音の3部分からなる。初声、中声、終声いずれも単独の子音（母音）の場合と、複合子音（母音）の場合がある。

漢字ほどではないが多くの文字を持っており、理論的には1万以上の文字がある。個々の文字の使用頻度も漢字と同じように語の使用頻度に似た分布を示す。

1つの文字を要素に分解した時の並べ方は安定しており、また殆んど区切記号を使わなくとも要素の並びの列からもとの文字を構成するアルゴリズムが実現できる。この性質はハングルの入力方式を容易で効率的なものにしている。

漢字と同じく1つの文字に多くの要素が配置されるため、良質の印字のためには各文字毎のフォントを用意しておく必要がある。要素のフォントだけから文字を構成することも可能であるが、質の良い印字にはならない。このため出力機器にとっては漢字と同様の扱いが必要となる。

4.2 N バイトコード

N バイトコードは各要素を文字として扱う方法である。各要素を1バイトで表すと、ハングル1文字が2~Nバイトになるためこの名がある。文字セットとしてはハングルの要素24個ですむことになるが、複合子音や複合母音も1つの文字としてコード化する場合が多い。複合子音を含めると子音は29ほど、母音は21であり、いずれにしても1バイトにおさまる数である。

	4	5	6	7	8	9	A	B	C	D	E	F
0						ト						
1		ㄹ			ㄱ	ㅋ	ㄴ					
2	ㄱ	ㄷ			ㄴ	ㄷ	ㅌ					
3	ㄱ	ㄷ			ㅁ	ㅂ	ㅃ					
4	ㄱ	ㄷ			ㅅ	ㅆ	ㅈ					
5	ㄴ	ㄷ			ㅈ	ㅊ	ㅌ					
6	ㄴ	ㄷ			ㅈ	ㅊ	ㅌ					
7	ㄴ				ㅇ	ㅈ						
8	ㄷ	ㄷ			ㅈ	ㅊ						
9	ㄷ				ㅈ	ㅊ						
A					ㅈ	ㅊ						
B												
C					ㅋ							
D					ㅌ	ㅌ						
E					ㅍ	ㅍ						
F					ㅎ	ㅍ						

(ハングル以外略)

図4 IBM 3バイトコード

	4	5	6	7
0	Fill	ㄹ		
1	ㄱ	ㄱ		
2	ㄱ	ㄷ	ㅌ	ㅌ
3	ㄱ	ㅁ	ㅂ	ㅃ
4	ㄴ	ㅅ	ㅆ	ㅈ
5	ㄴ	ㅈ	ㅊ	ㅌ
6	ㄴ	ㅈ	ㅊ	ㅌ
7	ㄴ	ㅇ	ㅈ	ㅊ
8	ㄷ	ㅈ		
9	ㄷ	ㅈ		
A	ㄷ	ㅈ	ㅊ	ㅌ
B	ㄷ	ㅈ	ㅊ	ㅌ
C	ㄷ	ㅈ	ㅊ	ㅌ
D	ㄷ	ㅈ	ㅊ	ㅌ
E	ㄷ	ㅈ	ㅊ	ㅌ
F	ㄷ	ㅈ	ㅊ	ㅌ

図5 KS C5601-1987
付録4 ハングル要素7単位コード

N バイトコードの例として、韓国 IBM の「3-バイトコード」(図4)⁽⁷⁾と、後述の KSC 5601-1987 の付録4(図5)⁽⁸⁾をあげる。後者の方が複合子音が1つ多いが、これは通常の表記では用いられない字である。

N バイトコードは最も早く現れた形で、入力時のコードとしては都合がよいが、出力や文字列処理には向かない。既に処理コードとしては殆ど使われなくなっている方法である。

4.3 組合型コード

初声、中声、終声に用いられる要素を各々コード化し、これを組合せて文字のコードとするものがある。

現代のハングルでは初声のパターンが19、中声が21、終声が27ある。各々が欠けている場合を入

れると総数は 12320 となる（終声以外が欠けたものは普通の表記で使うわけではないが、丁度漢字のしんにょうだけといったような文脈で使われ得る）。

一般には各声を各々 5 ビットづくにわりあてて実質 15 ビットのコードとして 2 バイトの中で使用している。ただし、ビットのわりあては、コードスペースとの関係もあって各社まちまちとなっている。要素パターンが同じである限り、文字の集合としては同一のものとなるはずであるが、配列の異なった文字セットとなっている例である。

例として IBM のホストコードと呼ばれているもの (図 6)⁽⁷⁾ 及び KSC 5601 の付録 3 (図 7)⁽⁸⁾ を示す。

ビットパターン	初声	中声	終声
0 0 0 0 0		*	*
0 0 0 0 1	Fill	*	Fill
0 0 0 1 0	ㄱ	Fill	ㄱ
0 0 0 1 1	ㄴ	ㅏ	ㄴ
0 0 1 0 0	ㄷ	ㅑ	ㄷ
0 0 1 0 1	ㄸ	ㅓ	ㄷ
0 0 1 1 0	ㄹ	ㅕ	ㄹ
0 0 1 1 1	ㄺ	ㅗ	ㄹ
0 1 0 0 0	ㄱ	*	ㄷ
0 1 0 0 1	ㄴ	*	ㄸ
0 1 0 1 0	ㄷ	ㅑ	ㄹ
0 1 0 1 1	ㄸ	ㅓ	ㄺ
0 1 1 0 0	ㄹ	ㅕ	ㄻ
0 1 1 0 1	ㄺ	ㅗ	ㄼ
0 1 1 1 0	ㅓ	ㅕ	ㄽ
0 1 1 1 1	ㅗ	ㅑ	ㄾ
1 0 0 0 0	ㅑ	*	ㄻ
1 0 0 0 1	ㅓ	*	ㄼ
1 0 0 1 0	ㅕ	ㅑ	ㄽ
1 0 0 1 1	ㅗ	ㅓ	ㄾ
1 0 1 0 0	ㅑ	ㅕ	ㄽ
1 0 1 0 1	ㅓ	ㅗ	ㄾ
1 0 1 1 0	ㅕ	ㅑ	ㄽ
1 0 1 1 1	ㅗ	ㅓ	ㄾ
1 1 0 0 0		*	ㅓ
1 1 0 0 1		*	ㅕ
1 1 0 1 0		ㅑ	ㅕ
1 1 0 1 1		ㅓ	ㅗ
1 1 1 0 0		ㅕ	ㅑ
1 1 1 0 1		ㅗ	ㅓ
1 1 1 1 0		ㅓ	ㅕ
1 1 1 1 1	*		*

ビットパターン	初声	中声	終声
0 0 0 0 0	*	*	*
0 0 0 0 1	*	Fill	Fill
0 0 0 1 0	*	ㅏ	ㄱ
0 0 0 1 1	*	ㅑ	ㄴ
0 0 1 0 0	*	*	ㄷ
0 0 1 0 1	*	ㅓ	ㄷ
0 0 1 1 0	*	ㅕ	ㄹ
0 0 1 1 1	*	ㅗ	ㄹ
0 1 0 0 0	*	*	ㄷ
0 1 0 0 1	Fill	ㅑ	ㄸ
0 1 0 1 0	ㄱ	ㅑ	ㄹ
0 1 0 1 1	ㄴ	ㅓ	ㄺ
0 1 1 0 0	ㄷ	*	ㄻ
0 1 1 0 1	ㄸ	ㅗ	ㄼ
0 1 1 1 0	ㄹ	ㅕ	ㄽ
0 1 1 1 1	ㄺ	ㅑ	ㄾ
1 0 0 0 0	ㅑ	*	ㄻ
1 0 0 0 1	ㅓ	ㅑ	ㄼ
1 0 0 1 0	ㅕ	ㅑ	ㄽ
1 0 0 1 1	ㅗ	ㅓ	ㄾ
1 0 1 0 0	ㅑ	*	ㄽ
1 0 1 0 1	ㅓ	ㅑ	ㄾ
1 0 1 1 0	ㅕ	ㅑ	ㄽ
1 0 1 1 1	ㅗ	ㅓ	ㄾ
1 1 0 0 0	ㅑ	*	ㅓ
1 1 0 0 1	ㅓ	ㅑ	ㅕ
1 1 0 1 0	ㅕ	ㅑ	ㅕ
1 1 0 1 1	ㅗ	ㅓ	ㅗ
1 1 1 0 0	ㅑ	*	ㅕ
1 1 1 0 1	ㅓ	ㅑ	ㅕ
1 1 1 1 0	ㅕ	ㅑ	ㅕ
1 1 1 1 1	*		*

図 6 IBM「ホストコード」のビット構成 図 7 KS C5601-1987付録 3 16単位コード

KSC 5601 の付録 3、16 単位コードは交換用コードではなく、内部コードとして組合格型を使用する場合の標準を提供するものである。

なお、図 6 の IBM「ホストコード」の終声は図 7 のものより 1 種類多いが、これが何に使われるのかは今のところ不明である。

組合格型コードの長所は、理論的に可能なすべてのハングルを持った文字セットである点、及びキーボード入力からの変換規則が簡単になる点である。反面、文字セット中に未定義ポジションが多くなり、コードスペースの使用効率が落ちるといった欠点がある。特に、1 万を少し超える数になるため、ISO 2022 互換の文字セット 1 つ（最大 9216）では収まらないという点が問題となる。

4. 4 ^{ワンスンヒョン} 完成型コード

完成型は漢字等と同じく、1 つの完成したハングルを 1 つの文字とする。一般に組合格型のような理論的に可能な文字を総て収納するようなことはせず、実際に使用されている頻度の高いものを 2～3 千選定するのが普通である。

歴史的には N バイトコードから組合格型コード、完成型コードと動いて来たと言われる。IBM も PC 55 XX では完成型のハングルコードを用いている⁽⁹⁾。国家標準コードである KSC 5601 も完成型のコードである。

この様な方向に進んで来たのは、おそらく組合格型に要するコードスペースが、特にパーソナルコンピュータのオペレーティングシステムにとって負担になって来たためではないかと想像する。特に、ASCII 文字セット、ハングル文字セットに漢字文字セットも同時使用するためには、組合格型のためのコードスペースがとれなかったのであろう。この点については、コードスペースの節で検討する。

4. 5. KC S 5601-1987

KSC 5601 は漢国の国家標準の情報交換用符号である。何度かの改訂を行ない現在は 1987 年 8 月の版となっている。

KSC 5601 は JIS X 0208 と同様に ISO 2022 互換の 94×94 文字セットであり、その構成も JIS X 0208 から影響を受けていると見られる。以下に構成を示す。

- 1 区、2 区は記号で 163 文字ある。
- 3 区は ASCII と同じとしてある。
- 4 区はハングルの要素（複合を含む）。
- 5 区はローマ数字及びギリシャ文字。
- 6 区はけい線素片。
- 7 区は単位記号（cm、kg など）。
- 8 区は特殊アルファベットと丸つき数字等。
- 9 区は 8 区の小文字と括弧つき数字等。
- 10 区は日本語ひら仮名（JIS 4 区と同）。
- 11 区は日本語カタ仮名（JIS 5 区と同）。

12 区はロシア文字。

16～40 区がハングル 2350 文字。

41 区はハングル用ユーザ定義エリア。

42～93 区が漢字 4888 字。

94 区は漢字用ユーザ定義エリア。

漢字はハングル読みして配列しているが、異音同字の場合は 2 ヶ所（以上）に現れる。例えば「車」という文字は読みが 2 つあるために、43 区 71 点と 83 区 19 点の 2 ヶ所に現われる。

KS C 5601 に関してはハングルや漢字が十分でないという批判の他、技術的な点で「ハングル入力中の欠字」という欠点も指摘されてはいるが、概ね現代の韓国語表記には十分であるとの評価である。

「ハングル入力中の欠字」とここで呼んだのは、ハングルのキーボードからの入力途中で、コードのないものが出現する点である。図 8 にその例を示す。

入力したい文字	쌍 (区点 29-64)		
入力シーケンス	スクリーン	区点	
ㅍ (ss)	ㅍ (ss)	4-22	
ㅑ (ya)	ㅑ (ssya)	KSC 5601 になし	
ㅇ (ng)	쌍 (ssyang)	29-64	

図 8 ハングル入力中の欠字例

このような欠字自体は滅多に使われたいわけであるから、情報交換用コードとしてはやむを得ないのであるが、入力プロセッサを作成している者にとっては大きな問題ではある。こういったケースは KS C 5601 中に 6 つあると言われている。これらについては 6 つほどの文字をユーザ定義文字として追加することにより解決可能な問題ではある。

一方のハングル、或は漢字が十分でないという批判は、大規模なスクリプトを扱う際には不可避である。解決にはもっと大きな文字セットを使うなり、補助集的な文字セットを作り足して行く方法以外はない。韓国の標準の検討の中でこの議論はなされており、90 年中に、補助集を作る方向で動き出すということであった。

上記のような点はあるものの、KS C 5601-1987 は近年急速に普及しつつある。特に政府の行政電算網で使用するシステムでの KS C 5601 使用が義務づけられたことから、各メーカーが KS C 5601 への切替を急いだと言われている。このため KS C 5601-1987 は標準コードとして十分機能するものとなるであろうと思われる。ただし、シェアの大きい IBM は KS C 5601 を使用していない。

5. 韓国のコードスペース

前節では文字セットを中心として見た。これらハングル、漢字の他に当然基本的な文字セットとして ASCII 又は EBCDIC を 1 バイトの世界で使用している。日本と状況が異なるのは 1 バイトの世界で使われるカナ文字セットにあたるものがないことで、この点は韓国の方が有利な状況をしている。

以下、代表的なものとして、組成型を使った IBM メインフレームのコードスペース、ISO 2022 に沿った MS-DOS 等での KSC 5661 のコードスペースをとりあげる。また組成型のコードスペースの検討と、コードスペース拡張の例をとりあげる。

5. 1 IBM ホストコード

コードスペースは図 9 のように $\{40\sim FF\} \cup \{41\sim FE\}^2$ である。1 バイトと 2 バイトの切換は制御文字 SI、SO で行なう。

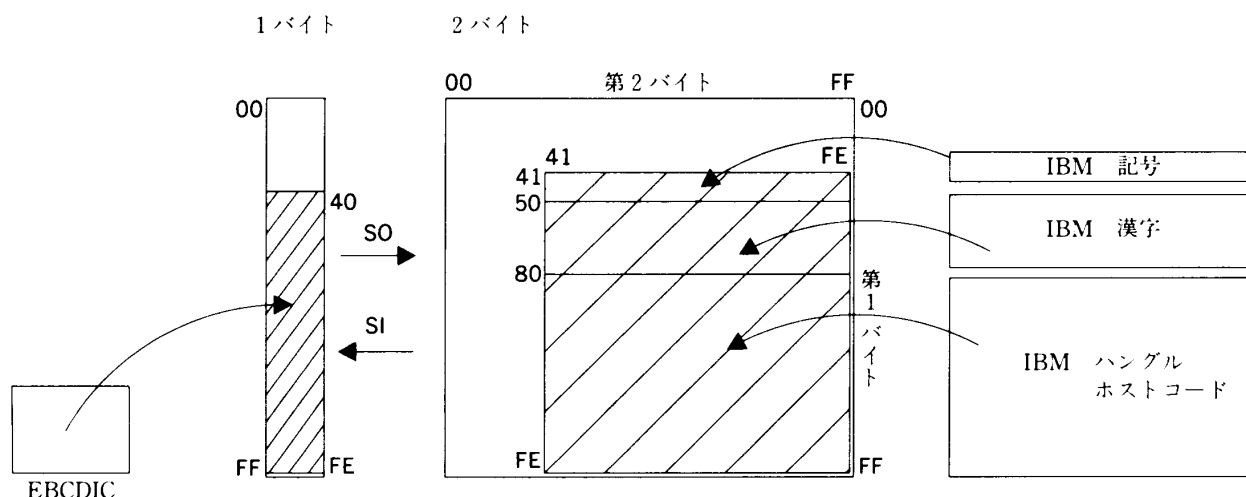


図 9 IBM ホストコードのコードスペースと文字セット

このコードスペースは、3. 2 節に述べた KEIS コードと同じである。しかし、文字セットの呼び出し領域は全く異っている。KEIS が ISO 2022 による 94 文字セットを使用しているのに対し、こちらは最初からこのコードスペースを前提とした文字セットを使用しているためである。

組成型のハングルコードがコードスペースの半分以上を占めている。図 6 の組成型のビット組合せは、最上位ビットを 1 にして $\{80\sim FE\} \times \{41\sim FE\}$ に呼び出すわけである。図 6 中の未使用コード（空白部）を除けば実際にハングル領域は $\{84\sim D3\} \times \{41\sim FD\}$ で、この中にも更に使用されない値がある。いずれにしても相当大きな部分を占めている。

5. 2 組成型とコードスペース

組成型でも 4. 3 節に見たようにいくつかの種類があるが、いずれも有効 15 ビットに最上位ビット 1 を加えて 16 ビットのコードとしている。漢字を使わないシステムでは、ASCII とこのコードだけと

いう方式もある。

組合型のコードスペースは、図 6 や図 7 のようなビット組合の配分によって変化して来るが、この際これらの図にある * 印（使用禁止）や空白（未使用）が意味を持って来る。

図 6 のコードスペースは、* 部分を除くことにより {80~FB} × {41~FE} となり、制御文字のエリアを避けているわけである。また、図 7 のコードスペースも * 部分を避けることにより、図 10 のように {A4~F3} × {21~7E, A1~FE} の範囲に入れることができています。これによって ISO 2022 に入れるような 94 × 94 の文字セット 2 つとすることは可能となっている。

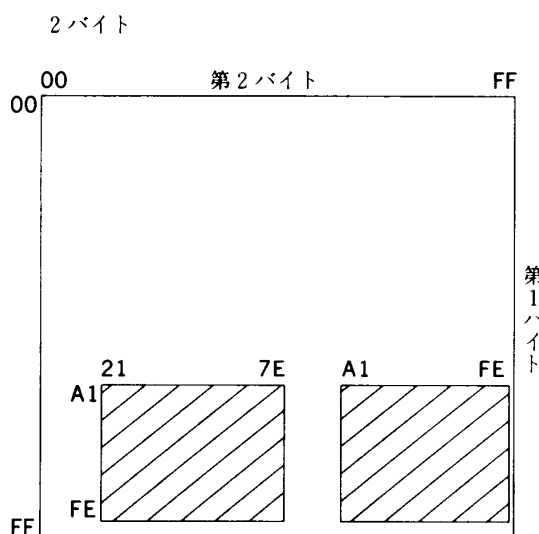


図10 KS C5601-1987付録 3 16単位コードのコードスペース

ただし、図 10 の 2 つはこのままでは ISO 2022 には適合しない。{A1~FE} × {21~7E} というコードスペースは ISO 2022 の 8 単位系では出てこないからである。もっとも、このままの形でも、ASCII との併用は可能で、シフト JIS のように、第 1 バイトの最上位ビットが 1 なら 2 バイトと見るというやり方をとればよい。

いずれにしても組合型では大きなコードスペースをハングルが占めてしまうため、漢字を収容することがむづかしくなる。図 10 のうち未使用文字部分を使えば 2 ~ 3 千の漢字を収容することも可能ではあるが、それも双方のコードスペースにまたがってとなってしまう。

こういった事情が組合型の標準文字セットとしての採用をむづかしくした理由ではないかと思われる。

5. 3 標準的な KSC 5601 の使用法

これは ISO 2022 の 8 単位符号系で、G₀ として ASCII、G₁ として KS C 5601 を指示し、G₀ を G_L に、G₁ を G_R に呼び出した固定的な状態で使用する使用方法である。コードスペースは {21~7E} U {A1~FE}² で、図 11 に示すように、各バイトの最上位ビットが 0 なら ASCII、1 ならば 2 バイトで KS というこ

となる。KS C 5601 は G_1 に指示せよということになっていて、この型で使用することを前提にしているようである。

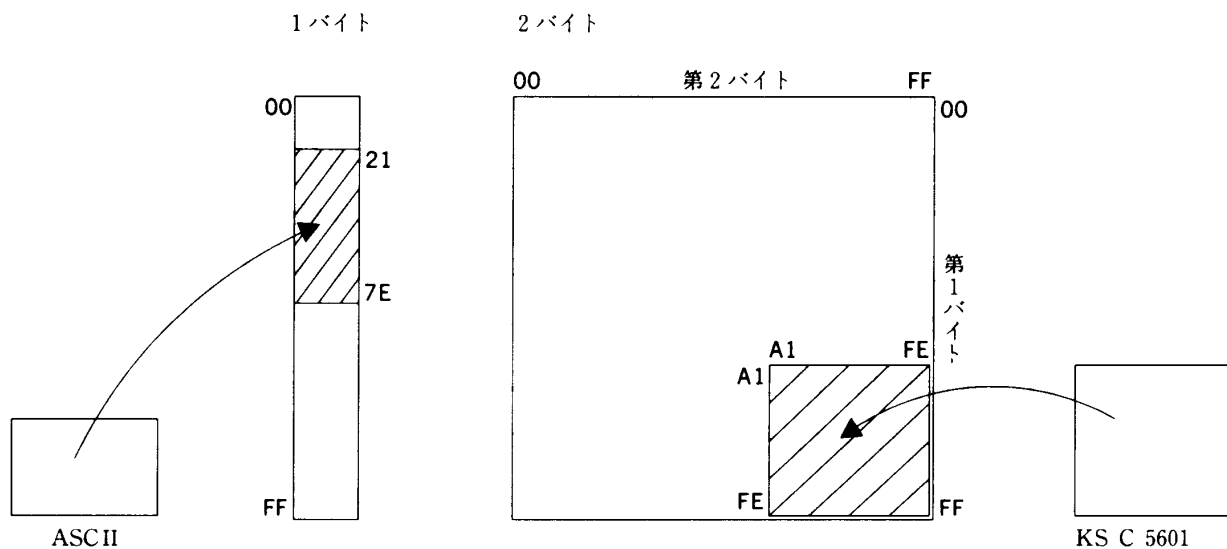


図11 KS C5601-1988の標準的なコードスペース

この方法は MS-DOS で使用されている他、マッキントッシュなど多くのシステムでも共通で⁽¹⁰⁾、KS C 5601 文字セットの導入とともに標準的方法となりつつある。ちなみに中国語用の MS-DOS である CC-DOS も、同様の方法で、 G_1 に GB 2312 を使用している。

この方法は標準との適合性や文字列処理の簡単さなどで優れた方法であるが、このままでは拡張できない点が問題である。例えば KS で新しい文字セットができて、このままでは対応できないことになる。

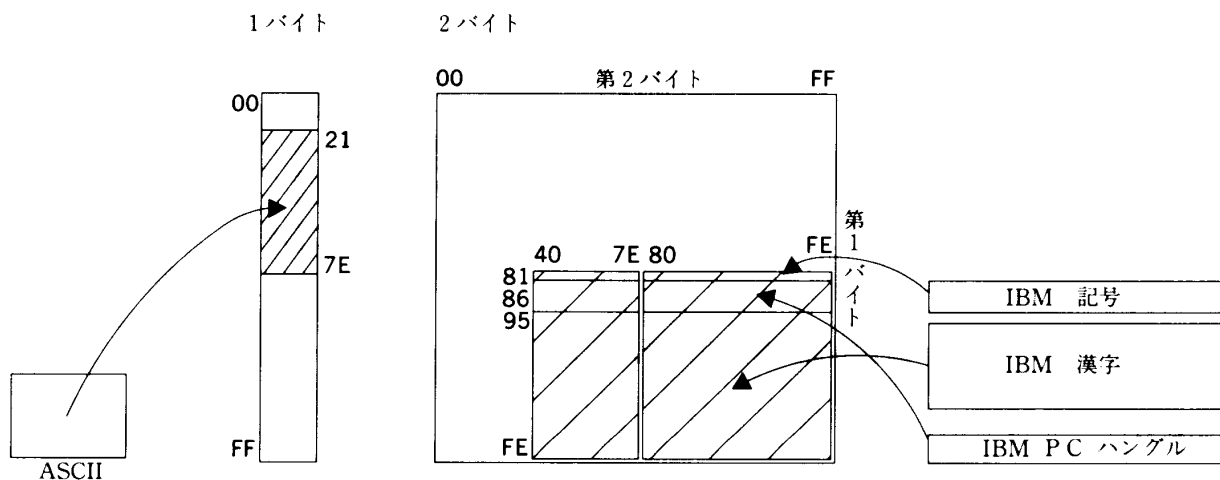


図12 IBM PC-55XX のコードスペースと文字セット

5. 4 コードスペースの拡張

図一に示したものは、IBM PC 55X で使用されているコードスペースである⁽⁹⁾。これは KS C 5601 は使用しておらず、ハングルについてはホストコードとも異なる独自の完成型文字セットを使用し、漢字はホストと同じものを異ったコードスペースに対応させている。

このコードスペースは、図3に示したシフト JIS のコードスペースを1バイト系のカナの部分まで拡張したものとなっている。シフト JIS と同様に ISO 2022 の制御文字の領域まで使うことにより、23688 文字分の大きなコードスペースを産み出している。MS-DOS では例え前節のような標準的な KS C 5601 の呼び出し方をしていたとしても、潜在的には図12と同じコードスペースがあると言える。同様の拡張は MS-DOS に限らず ASCII 系のコードを使用しているシステムであれば可能である。但し、制御文字領域を使用したことにより、プログラムの互換性は低下する。

制御文字領域の使用を避けて似たような拡張を行ったものに、国立中央図書館の TANDEM システムのものがある(図13)。コードスペースは図12のものから制御文字の領域を避けて{21~7E}U{A1~FE}×{21~7E, A1~FE}となっている。文字セットは{A1~FE}²に KS C 5601 (の空き領域に独自選定漢字を埋めたもの)を呼び出し、{A1~FE}×{21~7E}には独自の拡張漢字(8836 字)文字セットを割当てている。標準の立場から言えば、拡張漢字セットはともかくも、KSC の空き領域を使用した点はやや問題があると言わざるを得ない。後に KS C 5601 が改訂されて空き領域を使用する可能性があるためである。

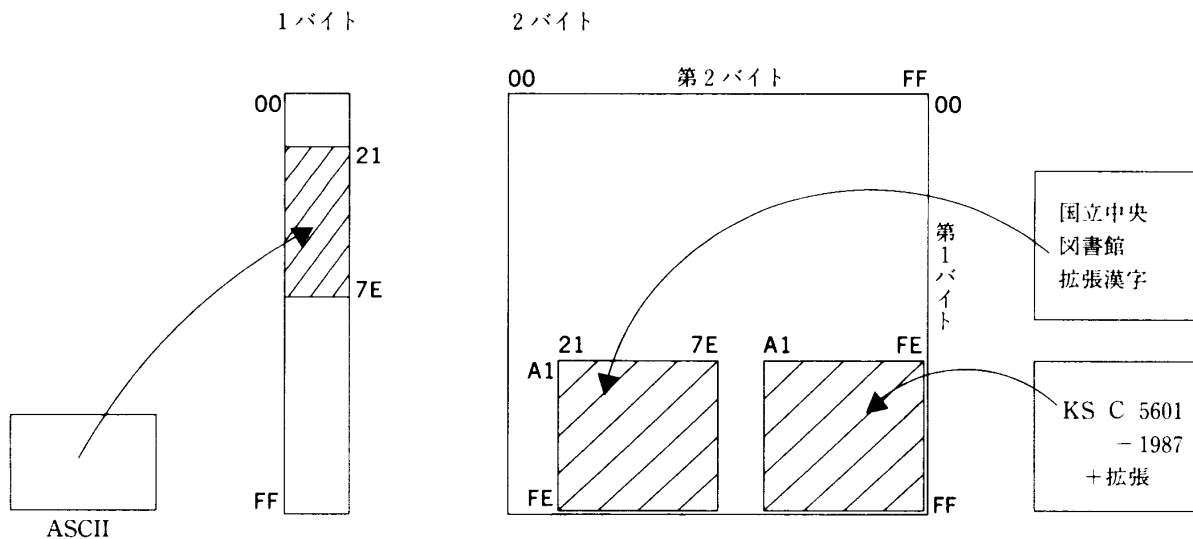


図13 国立中央図書館システムのコードスペースと文字セット

6. おわりに

コードスペースと文字セットという観点から韓国の文字コードについて概観した。もちろんここに紹介できなかったものもあることはお断りしておく。日韓、或は東アジアでの文字型データベース交換の問題点の詳細な検討は別稿に譲るが、以下簡単に問題点に触れてみたい。

本稿に見たように、現在のシステムの殆どはコードスペースに対する文字セットの呼び出しを固定にしている。呼び出し固定の状態では日本、中国、韓国の文字セットを扱うには、それらの文字セットが同時に入る大きさのコードスペースを持たなければならない。日本の JIS X 0208、韓国の KS C 5601、本稿には取り上げなかったが中国の GB 2312 がそれだとすれば、コードスペースには 94×94 文字セットを 3 つ以上含む大きさが必要である。

この条件を満たすものは、N-1 の NVT コードをはじめ、主としてメインフレーム系のシステムとなる。1 バイト、2 バイトの切替制御機能を持ち、1 バイトのコードスペースと独立に 2 バイトコードスペースのとれる方式でなければならない。

こういった世界であれば、例えば N-1 の NVT コード 2 バイト系の G_{01} に KS、 G_{11} 7 GB を固定的に呼び出し、内部コードとしては適当な対応規則を定め、という方法をとることは可能である。しかしながら、適用できるのが一部のメインフレームのみとなってしまう。また、韓国側を見ると、組合型をとっているメインフレームでは適用できない。さらに、各国標準での拡張文字セットに対応することが不可能になる、など問題点は多い。

本格的な解決のためには、コードスペースを 3 バイトや 4 バイトの世界に広げるか、あるいは、2 バイトのコードスペースに文字セットの動的呼び出しをするか、いずれかの方法をとる必要がある。3 節には述べなかったが、ISO 2022 に従って G_2 、 G_3 を動的呼び出し（指示は固定）するシステムは既にある。しかし、この方法では 2 バイト系で 94×94 文字セット 4 つ、1 バイト系で 94 文字セット 2 つを扱える N-1 の NVT コードすら完全には収容できないわけで、指示の動的切替など、より多くの文字セットを同時使用できるシステムが望まれるわけである。

コードスペースの拡張にせよ、文字セットの動的呼び出しにせよ、オペレーティングシステムの比較的低レベルまで影響を与える問題であり、短期間での解決はむづかしいと言わざるを得ない。

現在、ISO-IEC/JTC1/SC2 では DP 10646 マルチオクテット文字コードの検討が進められている。これは 4 オクテット（8 ビットバイト）のコードスペースに、世界中の文字を集めた唯一の文字セットを作ろうというものである。詳細はここには述べないが、こういった標準が確立されればオペレーティングシステム側でのサポートがされるのではないかという期待はできるだろう。

なお、韓国や中国の文字セットを完全に扱えなくとも、JIS プラスアルファの中で韓国語や中国語を扱う方法があるだろう、という意見もある。また、現在の KS 文字セットの中で日本語をかなり十分に扱えるという事実もある。これらは漢字の同一性やスクリプトとしての同一性という大きな問題を含み、ISO の DP 10646 の検討の中でも議論となっている点である。これらの点についてはむしろ言語の問題であると考えるが、稿を改めて検討したい。

いずれにしても、日韓のデータ交換を考える時の問題点は、標準の世界、オペレーティングシステムの世界、そして言語の世界にわたっていることは言えるだろう。これらの問題点の解決が今後の課題である。

謝辞

韓国滞在中調査のアレンジ等御助力下さった方々の名を記して感謝の意を表します。DACOM のユーキョンヒ氏、産業研究院のキムキューサー氏、クォンチュンファン氏、国立中央図書館のチョーウオンホ氏、イーチジュ氏、公州師範大学のヒョンキューソプ氏、イーチュンタク氏。

参考文献

- [1] JIS X 0202-1984、情報交換用符号の拡張法、第 2 刷、1988。
- [2] ISO 2022-1986, Information processing — ISO 7-bit and 8-bit coded character sets—Code extension techniques, 1986.
- [3] 宮澤 彰、拡張文字セットについて、東京大学文献情報センター紀要、第 1 号、pp. 29-62、1985。
- [4] 安達 淳、N-1 仮想端末プロトコルの拡張、東京大学文献情報センター紀要、第 1 号、pp. 130-170、1985。
- [5] 安達 淳他、TSS 接続による仮想画面転送(VTSS)方式。学術情報センター紀要、第 1 号、pp. 73-89、1987。
- [6] プログラムプロダクト VOS 3 漢字情報処理システム KEIS 概説、第 2 版、日立製作所、1987。
- [7] IBM ハングル使用指針書、第 1 版、韓国 IBM、1982。(韓国語)
- [8] KS C 5601-1987 Code for Information Interchange (Hangul and Chinese Character), in KS Handbook '87, Industrial Advancement Administration Republic of Korea, 1987.
- [9] IBM マルチステーション 5550 コード集、第 1 版、韓国 IBM、1984。(韓国語)
- [10] Macintosh ハングル Talk 使用者指針書、version 6.0, ELEX Computer Inc. (韓国語)

研 究 論 文

韓国標準規格と日本工業規格の漢字について

Comparative Analysis of Character Set/Code between Korean and Japanese Standard

韓国公州大学：富士大学 李 春 澤

緒 言

- 1 韓国標準規格の基本漢字の構成および性格
 - 2 韓国字
 - 3 日本工業規格との比較
 - 4 結論
- あとがき

要 旨

コンピュータで使用する漢字コードについては、韓国にも日本にも国家規格が KSC 5601 および JIS X 0208 として制定されている。いずれも数千字の漢字を収録しているが、同一ではない。本稿では、はじめに韓国国家規格の漢字の内容を、同一字が場合によって音が変わる字、異意異音同体字、同意異音同体字の 3 種類の基準から分析した。次に、約 500 字あるとされる韓国字の特色と国家規格の収録状況を述べた。さらに日本工業規格に収録された漢字との比較を体裁、字数、異体字の 3 種類の基準から分析した。二つの国家規格における漢字収録の状況を比較することにより、国際（地域）標準化の課題を提示した。

ABSTRACT

National standards on Hanja/Kanji character exist in Korea and Japan, KSC5601 and JISX0208 respectively. Both national standards include several thousands Hanja/Kanji characters, but are not identical. The paper discusses coverage and characteristics of Korean Hanja by KSC under three features such as different pronunciation; same form-different meaning & pronunciation; same form & meaning-different pronunciation. Characteristics of national or "Sino-Korean" characters, about 500 are known to exist, are described. A comparison between KSC and JIS is made by form, by number, by variations. Tasks are discussed for international (regional) standardization between two national standards.

Lee Choon Tack : Associate Professor, Kongju National University, Korea
Visiting Research Fellow (1988), Fuji University, Hanamaki City.

The paper is a revised and edited version of the presentation made at the "First Workshop on International Bibliographic Control in Japan", written and delivered in Japanese by the auther. (editor)

緒 言

韓国工業振興庁 (IAA: Industrial Advancement Administration) は、ハングル・コンピュータコードの標準化に続いて、1987年8月に漢字符号も2バイト完成型で確定し発表した。これが、いわゆる韓国標準情報交換用コード (The Korean Standard of Information Interchange Code)、すなわち KSC-5601 である。ここに選定された符号は、基本漢字 4,888 字をはじめとしてアルファベットと日本の仮名など外国文字および特殊文字 830 字と、使用者が選定して追加する漢字および特殊文字 94 字で構成されている。本稿では基本漢字 4,888 字に関して言及する。

漢字は中国で発生し、その後、東アジアに拡散し、いわゆる漢字文化圏を形成して、二千余年の間、韓国と日本の文字生活を築いてきた。韓国と日本は、ハングルと仮名という独特な音標文字を国字として持っているが、両国とも近世以来、今日まで、漢字廃止論がくりかえして台頭している現在も、漢字が依然として文字生活の重要な部分を占めているという事実を否認することは難しい。むしろ、最近においては、ワープロの出現とコンピュータの漢字処理能力の向上、一時の漢字廃止において生じた文字生活の不合理性に対する反省から、漢字使用が増えて、漢字教育が強化されるような印象もないではない。その例として、韓国では、一時、ハングル専用を唱える荒っぽく強い風のために学校で漢字教育を廃止したことがあった。しかし、1972年からはまた復活し、中・高等学校に漢文の独立教科を設定して今日にいたっている。日本でも当用漢字から常用漢字に移行したのは、そのような感を強く匂わしている。そして、ハングル専用が民族主体性回復であるとして、漢字を完全に廃止したことが知られてきた北朝鮮でも、漢字性単語に対する学生達の理解力が低下したため、今は基本漢字教育をさせていることが知られている。

しかし、これと同時に、長い期間、各国で独自に使用されている間に、漢字あるいは漢字から作られた単語の意味が変化した点、そして、中国と日本で急速に進行している漢字の簡体化問題、そして、いわゆる国字と呼ばれている各国で新しく作られた、その国の固有の漢字など、看過することができない多くの相違点を持つようになった。それにもかかわらず、漢字の特殊性、すなわち、大部分の漢字は未だに意味の変化なく昔のまま使われている。そのため、漢字がこれらの国家間の意志伝達や書誌情報の交換において、重要な部分を占めているのも事実である。

筆者は、これらの漢字の中で、韓国標準規格に選定された 4,888 の基本漢字を分析することによって、この基本漢字の性格と意味、日本工業規格との相互関連性について考察する。

1 韓国標準規格の基本漢字の構成および性格

まず、この基本漢字を分析してみると、おもしろい事実を発見することができる。この基本漢字すなわち 4,888 字のうちには、重出字がたくさんあることである。すなわち、2度以上重ねて出ている字が多いということで、同一字が2回出ている字は 257 字、3回出ている字が 4 字、甚だしいのは 4回重ねて出ている字もある。これらをすべて足してみると、字種として 262 字に、字数では 530 字に

表 1 韓国標準規格の重出字

2回重出字 (257種 514字)														
賈	42-25	가(カ)	降	43-29	강(カン)	豈	43-48	개(ケ)	更	43-54	갱(ケン)	車	43-71	거(コ)
	45-47	고(コ)		90-02	항(ハン)		49-34	기(キ)		44-58	경(キョン)		83-19	차(チャ)
見	44-24	견(キョン)	契	44-88	계(キエ)	滑	45-72	골(コル)	串	45-90	곳(コツ)	廓	46-09	곽(クワク)
	90-70	현(ヒョン)		48-48	글(クル)		92-33	활(フッル)		46-13	관(クワン)		92-09	화(フワク)
句	47-03	구(ク)	金	48-61	금(クム)	喇	49-52	나(ナ)	奈	49-53	나(ナ)	懶	49-56	나(ナ)
	47-91	귀(クイ)		49-49	김(キム)		52-90	라(ラ)		50-15	내(ネ)		52-91	라(ラ)
拏	49-57	나(ナ)	癩	49-59	나(ナ)	羅	49-60	나(ナ)	蘿	49-61	나(ナ)	螺	49-62	나(ナ)
	52-92	라(ラ)		52-93	라(ラ)		52-94	라(ラ)		53-01	라(ラ)		53-02	라(ラ)
裸	49-63	나(ナ)	邏	49-64	나(ナ)	洛	49-67	낙(ナク)	烙	49-68	낙(ナク)	珞	49-69	낙(ナク)
	53-03	라(ラ)		53-04	라(ラ)		53-06	락(ラク)		53-07	락(ラク)		53-08	락(ラク)
落	49-70	낙(ナク)	諾	49-71	낙(ナク)	酪	49-72	낙(ナク)	駱	49-73	낙(ナク)	亂	49-74	난(ナン)
	53-10	락(ラク)		53-11	락(ラク)		53-12	락(ラク)		53-13	락(ラク)		53-15	란(ラン)
卵	49-75	난(ナン)	欄	49-77	난(ナン)	爛	49-79	난(ナン)	蘭	49-80	난(ナン)	鸞	49-82	난(ナン)
	53-16	란(ラン)		53-17	란(ラン)		53-20	란(ラン)		53-21	란(ラン)		53-22	란(ラン)
嵐	49-86	남(ナム)	濫	49-90	남(ナム)	藍	49-92	남(ナム)	檻	49-93	남(ナム)	拉	49-94	납(ナブ)
	53-25	람(ラム)		53-29	람(ラム)		53-32	람(ラム)		53-33	람(ラム)		53-35	랍(ラブ)
臘	50-02	납(ナブ)	蠟	50-03	납(ナブ)	廊	50-07	낭(ナン)	朗	50-08	낭(ナン)	浪	50-09	낭(ナン)
	53-36	랍(ラブ)		53-37	랍(ラブ)		53-38	랑(ラン)		53-39	랑(ラン)		53-40	랑(ラン)
狼	50-10	낭(ナン)	郎	50-11	낭(ナン)	來	50-13	내(ネ)	冷	50-18	냉(ネン)	女	50-19	녀(ニョ)
	53-41	랑(ラン)		53-45	랑(ラン)		53-46	래(レ)		53-50	랭(レン)		69-92	여(ヨ)
年	50-20	년(ニョン)	撚	50-21	년(ニョン)	季	50-22	년(ニョン)	念	50-23	념(ニョム)	捻	50-26	념(ニョム)
	70-36	연(ヨン)		70-42	연(ヨン)		70-60	연(ヨン)		70-86	염(ヨム)		70-87	염(ヨム)
勞	50-30	노(ノ)	怒	50-33	노(ノ)	擣	50-34	노(ノ)	櫓	50-35	노(ノ)	爐	50-36	노(ノ)
	54-44	로(ロ)		54-45	로(ロ)		54-47	로(ロ)		54-48	로(ロ)		54-51	로(ロ)
盧	50-38	노(ノ)	老	50-39	노(ノ)	蘆	50-40	노(ノ)	虜	50-41	노(ノ)	路	50-42	노(ノ)
	54-52	로(ロ)		54-53	로(ロ)		54-54	로(ロ)		54-55	로(ロ)		54-56	로(ロ)
露	50-43	노(ノ)	魯	50-44	노(ノ)	鷺	50-46	노(ノ)	碌	50-47	녹(ノク)	祿	50-48	녹(ノク)
	54-58	로(ロ)		54-59	로(ロ)		54-60	로(ロ)		54-62	록(ロク)		54-63	록(ロク)
綠	50-49	녹(ノク)	菘	50-50	녹(ノク)	錄	50-51	녹(ノク)	鹿	50-52	녹(ノク)	論	50-53	논(ノン)
	54-64	록(ロク)		54-65	록(ロク)		54-66	록(ロク)		54-67	록(로)		54-69	론(ロン)
壟	50-54	농(ノン)	弄	50-55	농(ノン)	籠	50-57	농(ノン)	聾	50-58	농(ノン)	牢	50-62	뇌(ヌエ)
	54-70	롱(ロン)		54-71	롱(ロン)		54-75	롱(ロン)		54-76	롱(ロン)		54-79	뢰(ルエ)
磊	50-63	뇌(ヌエ)	賂	50-65	뇌(ヌエ)	雷	50-66	뇌(ヌエ)	尿	50-67	뇨(ニョ)	壘	50-68	누(ヌ)
	54-80	뢰(ルエ)		54-81	뢰(ルエ)		54-84	뢰(ルエ)		72-81	요(ヨ)		55-04	루(ル)
屢	50-69	누(ヌ)	樓	50-70	누(ヌ)	淚	50-71	누(ヌ)	漏	50-72	누(ヌ)	累	50-73	누(ヌ)
	55-06	루(ル)		55-07	루(ル)		55-08	루(ル)		55-09	루(ル)		55-11	루(ル)
縷	50-74	누(ヌ)	陋	50-75	누(ヌ)	杻	50-78	뉴(ニュ)	紐	50-79	뉴(ニュ)	勒	50-80	늑(ヌク)
	55-12	루(ル)		55-16	루(ル)		74-84	유(ユ)		75-10	유(ユ)		55-45	룩(ルク)
肋	50-81	늑(ヌク)	凜	50-82	늑(ヌム)	凌	50-83	능(ヌン)	稜	50-84	능(ヌン)	綾	50-85	능(ヌン)
	55-46	룩(ルク)		55-47	름(ルム)		55-48	릉(ルン)		55-50	릉(ルン)		55-51	릉(ルン)
菱	50-87	능(ヌン)	陵	50-88	능(ヌン)	泥	50-90	니(ニ)	匿	50-91	닉(ニク)	溺	50-92	닉(ニク)
	55-52	릉(ルン)		55-53	릉(ルン)		76-18	이(イ)		76-41	익(イク)		76-42	익(イク)
茶	50-94	다(タ)	丹	51-01	단(タン)	糖	51-56	당(タン)	宅	51-75	택(タク)	度	51-88	도(ト)
	83-17	차(チャ)		53-14	란(ラン)		87-24	탕(タン)		87-40	택(タク)		86-84	탁(タク)
讀	52-33	독(トク)	洞	52-55	동(トン)	掠	53-51	략(リャク)	略	53-52	략(リャク)	亮	53-53	량(リャン)
	52-70	두(トゥ)		87-51	통(トン)		69-17	약(ヤク)		69-18	약(ヤク)		69-25	양(ヤン)
兩	53-55	량(リャン)	涼	53-56	량(리ャン)	梁	53-57	량(리ャン)	糧	53-61	량(리ャン)	良	53-62	량(리ャン)
	69-27	양(ヤン)		69-28	양(ヤン)		69-36	양(ヤン)		69-46	양(ヤン)		69-48	양(ヤン)
諒	53-63	량(리ャン)	量	53-65	량(리ャン)	勵	53-68	려(リョ)	呂	53-69	려(리ョ)	廬	53-70	려(리ョ)
	69-50	양(ヤン)		69-54	양(ヤン)		69-90	여(ヨ)		69-91	여(ヨ)		69-94	여(ヨ)

表1 韓国標準規格の重出字(続き)

旅	53-73	려(リョ)	濾	53-75	려(リョ)	礪	53-76	려(リョ)	閭	53-79	려(リョ)	驪	53-81	려(リョ)
	70-01	여(ヨ)		70-04	여(ヨ)		70-07	여(ヨ)		70-13	여(ヨ)		70-15	여(ヨ)
麗	53-82	려(リョ)	黎	53-83	려(リョ)	力	53-84	력(リョク)	曆	53-85	력(リョク)	歷	53-86	력(リョク)
	70-16	여(ヨ)		70-17	여(ヨ)		70-19	역(ヨク)		70-23	역(ヨク)		70-24	역(ヨク)
轢	53-89	력(リョク)	憐	53-91	련(リョン)	戀	53-92	련(リョン)	漣	53-94	련(リョン)	煉	54-01	련(リョン)
	70-28	역(ヨク)		70-38	연(ヨン)		70-39	연(ヨン)		70-50	연(ヨン)		70-54	연(ヨン)
璉	54-02	련(リョン)	練	54-03	련(リョン)	聯	54-04	련(リョン)	蓮	54-05	련(리ョン)	輦	54-06	련(리ョン)
	70-57	연(ヨン)		70-63	연(ヨン)		70-65	연(ヨン)		70-69	연(ヨン)		70-68	연(ヨン)
連	54-07	련(리ョン)	鍊	54-08	련(리ョン)	列	54-10	렬(리ョ르)	劣	54-11	렬(리ョ르)	烈	54-13	렬(리ョ르)
	70-70	연(ヨン)		70-72	연(ヨン)		70-74	열(ヨル)		70-75	열(ヨル)		70-79	열(ヨル)
裂	54-14	렬(리ョ르)	廉	54-15	렴(리요ム)	殮	54-17	렴(리요ム)	簾	54-19	렴(리요ム)	獵	54-20	렵(리요브)
	70-81	열(ヨル)		70-85	염(요ム)		70-89	염(요ム)		71-01	염(요ム)		71-06	염(요브)
令	54-21	령(리ョン)	囹	54-23	령(리ョン)	嶺	54-26	령(리ョン)	伶	54-27	령(리ョン)	玲	54-28	령(리ョン)
	71-09	영(ヨン)		71-10	영(ヨン)		71-13	영(ヨン)		71-16	영(ヨン)		71-31	영(ヨン)
矜	54-30	령(리ョン)	聆	54-32	령(리ョン)	鈴	54-34	령(리ョン)	零	54-35	령(리ョン)	靈	54-36	령(리ョン)
	71-38	영(ヨン)		71-39	영(ヨン)		71-43	영(ヨン)		71-45	영(ヨン)		71-47	영(ヨン)
領	54-37	령(리ョン)	例	54-39	례(리예)	禮	54-41	례(리예)	醴	54-42	례(리예)	隸	54-43	례(리예)
	71-48	영(ヨン)		71-51	예(이예)		71-63	예(이예)		71-68	예(이예)		71-70	예(이예)
了	54-85	료(리요)	僚	54-86	료(리요)	寮	54-87	료(리요)	料	54-89	료(리요)	燎	54-90	료(리요)
	72-71	요(ヨ)		72-72	요(ヨ)		72-80	요(ヨ)		72-87	요(ヨ)		72-91	요(ヨ)
療	54-91	료(리요)	蓼	54-94	료(리요)	遼	55-01	료(리요)	龍	55-03	룡(리요ん)	劉	55-17	류(리유)
	72-94	요(ヨ)		73-07	요(ヨ)		73-12	요(ヨ)		73-44	용(ヨン)		74-69	유(ユ)
柳	55-19	류(리유)	流	55-21	류(리유)	溜	55-22	류(리유)	琉	55-24	류(리유)	留	55-26	류(리유)
	74-87	유(ユ)		74-92	유(ユ)		74-94	유(ユ)		75-04	유(ユ)		75-07	유(ユ)
硫	55-28	류(리유)	類	55-30	류(리유)	六	55-31	륙(리육)	戮	55-32	륙(리육)	陸	55-33	륙(리육)
	75-09	유(ユ)		75-26	유(ユ)		75-27	육(육)		75-29	육(육)		75-33	육(육)
倫	55-35	륜(리윤)	崙	55-36	륜(리윤)	淪	55-37	륜(리윤)	輪	55-39	륜(리윤)	律	55-40	륜(리윤)
	75-34	윤(윤)		75-38	윤(윤)		75-39	윤(윤)		75-44	윤(윤)		75-47	윤(윤)
慄	55-41	륜(리윤)	栗	55-42	륜(리윤)	隆	55-44	룽(리윤)	利	55-55	리(리)	吏	55-57	리(리)
	75-48	윤(윤)		75-49	윤(윤)		75-56	웅(윤)		76-06	이(이)		76-07	이(이)
履	55-59	리(리)	李	55-61	리(리)	梨	55-62	리(리)	理	55-66	리(리)	異	55-68	리(리)
	76-10	이(이)		76-16	이(이)		76-17	이(이)		76-21	이(이)		76-22	이(이)
痢	55-69	리(리)	罹	55-71	리(리)	裏	55-74	리(리)	裡	55-75	리(리)	里	55-76	리(리)
	76-24	이(이)		76-26	이(이)		76-32	이(이)		76-33	이(이)		76-37	이(이)
離	55-78	리(리)	吝	55-80	린(린)	憐	55-82	린(린)	璘	55-83	린(린)	藺	55-84	린(린)
	76-38	이(이)		76-53	이(인)		76-61	인(인)		76-62	인(인)		76-65	인(인)
隣	55-86	린(린)	鱗	55-87	린(린)	麟	55-88	린(린)	林	55-89	림(림)	淋	55-90	림(림)
	76-68	인(인)		76-71	인(인)		76-72	인(인)		76-87	임(임)		76-88	임(임)
臨	55-92	림(림)	立	56-01	립(립)	笠	56-02	립(립)	粒	56-03	립(립)	磻	58-82	반(반)
	76-90	임(임)		77-01	입(입)		77-02	입(입)		77-03	입(입)		59-68	번(번)
北	59-37	배(베)	便	60-05	변(비ョン)	復	60-54	복(복)	輻	60-63	복(복)	不	60-84	부(부)
	61-33	북(북)		88-21	편(비ョン)		61-05	부(부)		88-80	폭(복)		61-53	불(불)
泌	61-84	비(비)	數	62-92	삭(삭)	索	62-94	삭(삭)	殺	63-15	살(살)	參	63-19	삼(삼)
	89-18	필(필)		66-06	수(수)		63-67	색(섹)		65-77	쇄(스)		83-49	참(참)
狀	63-50	상(상)	塞	63-61	새(세)	省	63-72	생(센)	葉	64-81	삽(삽)	拾	67-06	습(습)
	77-78	장(찬)		63-65	색(섹)		63-93	성(송)		71-08	엽(요브)		68-09	십(십)
識	67-59	식(식)	辰	67-85	신(신)	沈	68-01	심(심)	什	68-07	십(십)	惡	68-34	악(악)
	82-29	지(지)		82-67	진(진)		86-56	침(침)		82-90	집(집)		71-87	오(오)
若	69-14	야(야)	易	70-22	역(역)	咽	70-76	열(열)	瑩	71-33	영(영)	阮	72-54	완(완)
	69-20	약(약)		76-15	이(이)		76-54	인(인)		91-09	형(히ョン)		74-33	원(원)
暈	73-87	운(운)	刺	77-09	자(차)	灸	77-19	자(차)	切	79-23	절(초르)	拓	84-12	척(초크)
	93-27	훈(훈)		84-07	착(차크)		78-59	작(초크)		84-78	체(체)		86-86	탁(탁)
暴	88-59	포(포)	行	90-01	항(항)									
	88-76	폭(복)		90-28	행(헝)									

表1 韓国標準規格の重出字(続き)

3回重出字(4種12字)														
龜	47-47	구(ク)	寧	50-27	녕(ニョン)	率	55-43	률(リュル)	說	64-67	설(ソル)			
	48-02	귀(クイ)		54-24	령(リョン)		65-67	솔(ソル)		65-13	세(セ)			
	48-24	균(キュン)		71-12	영(ヨン)		75-50	율(ユル)		70-82	열(ヨル)			
4回重出字(1種4字)														
樂	49-66	낙(ナク)												
	53-05	락(ラク)												
	68-37	악(アク)												
	72-89	요(ヨ)												
日本工業規格には出ていない重出字(8種16字)														
季	50-22	년(ニョン)	擄	50-34	노(ノ)	菘	50-50	녹(ノク)	杻	50-78	뉴(ニユ)	璉	54-02	련(リョン)
	70-60	연(ヨン)		54-47	로(ロ)		54-65	록(ロク)		74-84	유(ユ)		70-57	연(ヨン)
殮	54-17	림(リョム)	璘	55-83	린(リン)	磻	58-82	반(パン)						
	70-89	염(ヨム)		76-62	인(イン)		59-68	번(ボン)						

いたっている。もちろん、これらの大部分は日本工業規格にも含まれているが、8種16字は含まれていない。これは日本工業規格が各漢字ごとに代表音と訓を一個ずつ定めて配列して、字形を重視しているのに対して、韓国標準規格では音が重視されていることを意味する。したがって、日本工業規格には異体字がたくさんあるのに対して、韓国標準規格には重出字が多いのである。これらの重出字を抜き出して整理してみると表1「韓国標準規格の重出字」のとおりである。では、なぜこんなに多くの字が重出して出ているのか、表1を内容別に分類してみよう。

1-1 同一字が場合によって音が変わる字

第一は、同一字が場合によって音が変わる字で、習慣による場合と音韻の変化による場合がある。

①慣習による場合

これは、韓国人の慣習によって音が変わる場合である。この場合、ハングル「ㄴ」音が「ㄹ」音に変わる場合で、同一漢字がある場合は本音で、ある場合は俗音で発音される場合である。例を挙げると「諾」字は本音が「낙」(ナク)であるが、場合によっては「락」(ラク)とも発音するので、ハングル正書法では、その各々の音によって書くことを許容している(ハングル正書法第52項)。これに該当する字は4字である。

例	諾(ナク)	承諾(スンナク)	受諾(スラク)
	怒(ノ)	憤怒(ブンノ)	大怒(テロ)

②音韻法則による場合

これは韓国語音韻法則による場合である。同一字が音韻の変化によって発音が変わる場合であり、発音の変化は一字一字の漢字が持っている固有の音が二字以上の単語に組み合わせられるとき、その字の単語中に位置する所によって、すなわち、単語のはじめに来るときと終りに来るときに発音が変わる場合である。

周知のように、アルタイ語には語頭に「R」音(または「L」音)が立たない。したがって、韓国語

でも「ㄴ」音、「ㄹ」音が語頭にくると「ㄴ」音は「ㅇ」音に、「ㄹ」音は「ㄴ」または「ㅇ」音に変わって発音される現象がある。こういう現象を韓国語では頭音法則という。韓国標準規格の重出字の中には、この頭音法則が適用されている例がもっとも多く、次のように類型化できる。

ア) 「ㄴ」音から「ㅇ」音に変わる字 (12 字)

漢字音「녀」(ニョ)、「뇨」(ニョ)、「뉴」(ニユ)、「니」(ニ) が単語のはじめに立つと頭音法則が適用されて、各々、「여」(ヨ)、「요」(ヨ)、「유」(ユ)、「이」(イ)と発音される(ハングル正書法第 10 項)。

例 女子 「녀자」(ニョチャ) → 「여자」(ヨチャ)
 匿名 「닉명」(ニクミョン) → 「익명」(イクミョン)

しかし、単語の初め以外には頭音法則が適用されないため、本音のまま書く。

例 男女 「남녀」(ナムニョ) ○ 「남여」(ナムヨ) ×
 結紐 「결뉴」(キョルニユ) ○ 「결유」(キョルユ) ×

イ) 「ㄹ」音から「ㄴ」音に変わる字 (77 字)

漢字音「라」(ラ)、「래」(レ)、「로」(ロ)、「뢰」(ルエ)、「루」(ル)、「르」(ル) が単語のはじめに立つと、やはり頭音法則が適用され、「나」(ナ)、「내」(ネ)、「노」(ノ)、「뇌」(ヌエ)、「누」(ヌ)、「느」(ヌ)となる(ハングル正書法第 12 項)。

例 樂園 「락원」(ラクウォン) → 「낙원」(ナクウォン)
 来日 「래일」(レイル) → 「내일」(ネイル)

ウ) 「ㄹ」音から「ㅇ」に変わる字 (112 字)

漢字音「랴」(リャ)、「려」(リョ)、「례」(リエ)、「료」(リョ)、「류」(リュ)、「리」(リ) が単語の初めに立つと、やはり頭音法則が適用され、「야」(ヤ)、「여」(ヨ)、「예」(イエ)、「요」(ヨ)、「유」(ユ)、「이」(イ)と発音する(ハングル正書法第 11 項)。

例 良心 「량심」(リャンシム) → 「양심」(ヤンシム)
 歴史 「력사」(リョクサ) → 「역사」(ヨクサ)
 礼儀 「례의」(リエウイ) → 「예의」(イエウイ)
 龍宮 「룡궁」(リョンクン) → 「용궁」(ヨンクン)

その他

その他、特殊な例に「寧」と「不」がある。「寧」の場合は、その本音が「녕」(ニョン)であるが、場合によっては、本音と俗音で発音し、挨拶の言葉「安寧」の場合は本音「안녕」(アンニョン)と発音されるが、地名の「宜寧」の場合は俗音で「의령」(ウィリョン)と発音される。また、場合によっては頭音法則が適用されることもあって、「ㅇ」に変わる「寧日」の場合は「녕일」(ニョンイル)ではなくて「영일」(ヨンイル)になる。

「不」の場合は、その本音が「불」(プル)にもかかわらず、特殊な例として語彙構成上、「不」の後にハングル音「ㄷ、ㅌ」を初声とする字がくる場合だけ「불」(プル)が「부」(プ)に変わる現象がある。したがって「不」は時によって、「불」と「부」の二つの音で読む。「不当」の場合は「불당」(プ

表2 同一字で場合によって音の変わる字

字	KSC	音	JIS	字	KSC	音	JIS	字	KSC	音	JIS	字	KSC	音	JIS
慣習による場合 (ㄹ音→ㄷ音) (4字)															
喇	49-52	나(ナ)	51-41	拏	49-57	나(ナ)	57-28	諾	49-71	낙(ナク)	34-90	怒	50-33	노(ノ)	37-60
	52-90	라(ラ)			52-92	라(ラ)			53-11	락(ラク)			54-45	로(ロ)	
頭音法則による場合 (ㄹ音→ㅇ音) (12字)															
女	50-19	녀(ニョ)	29-87	年	50-20	년(ニョン)	39-15	撚	50-21	년(ニョン)	39-18	季	50-22	년(ニョン)	
	69-92	여(ヨ)			70-36	연(ヨン)			70-42	연(ヨン)			70-60	연(ヨン)	
念	50-23	념(ニョム)	39-16	捻	50-26	념(ニョム)	39-17	尿	50-67	뇨(ニョ)	39-02	扭	50-78	뉴(ニユ)	
	70-86	염(ヨム)			70-87	염(ヨム)			72-81	요(ヨ)			74-84	유(ユ)	
紐	50-79	뉴(ニユ)	41-19	泥	50-90	니(ニ)	37-05	匿	50-91	닉(ニク)	38-31	溺	50-92	닉(ニク)	37-14
	75-10	유(ユ)			76-18	이(イ)			76-41	익(イク)			76-42	닉(イク)	
頭音法則による場合 (ㄷ音→ㄹ音) (77字)															
懶	52-91	라(ラ)	56-81	癩	52-93	라(ラ)	65-90	羅	52-94	라(ラ)	45-69	蘿	53-01	라(ラ)	73-39
	49-56	나(ナ)			49-59	나(ナ)			49-60	나(ナ)			49-61	나(ナ)	
螺	53-02	라(ラ)	45-70	裸	53-03	라(ラ)	45-71	邏	53-04	라(ラ)	78-22	樂	53-05	락(ラク)	60-59
	49-62	나(ナ)			49-63	나(ナ)			49-64	나(ナ)			49-66	낙(ナク)	
洛	53-06	락(ラク)	45-76	烙	53-07	락(ラク)	63-64	珞	53-08	락(ラク)	64-68	落	53-10	락(ラク)	45-78
	49-67	낙(ナク)			49-68	낙(ナク)			49-69	낙(ナク)			49-70	낙(ナク)	
酪	53-12	락(ラク)	45-79	駱	53-13	락(ラク)	81-49	亂	53-15	란(ラン)	48-12	卵	53-16	란(ラン)	45-81
	49-72	낙(ナク)			49-73	낙(ナク)			49-74	난(ナン)			49-75	난(ナン)	
欄	53-17	란(ラン)	45-83	爛	53-20	란(ラン)	64-05	蘭	53-21	란(ラン)	45-86	鸞	53-22	란(ラン)	83-34
	49-77	난(ナン)			49-79	난(ナン)			49-80	난(ナン)			49-82	난(ナン)	
嵐	53-25	람(ラム)	45-82	濫	53-29	람(ラム)	45-84	藍	53-32	람(ラム)	45-85	檻	53-33	람(ラム)	75-02
	49-86	남(ナム)			49-90	남(ナム)			49-92	남(ナム)			49-93	남(ナム)	
拉	53-35	랍(ラブ)	57-39	臘	53-36	랍(ラブ)	71-36	蠟	53-37	랍(ラブ)	(47-25)	廊	53-38	랑(ラン)	47-13
	49-94	납(ナブ)			50-02	납(ナブ)			50-03	납(ナブ)			50-07	낭(ナン)	
朗	53-39	랑(ラン)	47-15	浪	53-40	랑(ラン)	47-18	狼	53-41	랑(ラン)	47-21	郎	53-45	랑(ラン)	47-26
	50-08	낭(ナン)			50-09	낭(ナン)			50-10	낭(ナン)			50-11	낭(ナン)	
來	53-46	래(レ)	48-52	冷	53-50	랭(レン)	46-68	勞	54-44	로(ロ)	50-09	携	54-47	로(ロ)	
	50-13	내(ネ)			50-18	냉(ネン)			50-30	노(ノ)			50-34	노(ノ)	
櫓	54-48	로(ロ)	47-06	爐	54-51	로(ロ)	64-04	盧	54-52	로(ロ)	66-26	老	54-53	로(ロ)	47-23
	50-35	노(ノ)			50-36	노(ノ)			50-38	노(ノ)			50-39	노(ノ)	
蘆	54-54	로(ロ)	73-35	虜	54-55	로(ロ)	46-26	路	54-56	로(ロ)	47-09	露	54-58	로(ロ)	47-10
	50-40	노(ノ)			50-41	노(ノ)			50-42	노(ノ)			50-43	노(ノ)	
魯	54-59	로(ロ)	47-05	鷺	54-60	로(ロ)	26-77	碌	54-62	록(ロク)	66-81	祿	54-63	록(ロク)	67-19
	50-44	노(ノ)			50-46	노(ノ)			50-47	녹(ノク)			50-48	녹(ノク)	
綠	54-64	록(ロク)	46-48	菘	54-65	록(ロク)		錄	54-66	록(ロク)	47-31	鹿	54-67	록(ロク)	28-15
	50-49	녹(ノク)			50-50	녹(ノク)			50-51	녹(ノク)			50-52	녹(ノク)	
論	54-69	론(ロン)	47-32	壘	54-70	롱(ロン)	52-66	弄	54-71	롱(ロン)	47-14	籠	54-75	롱(ロン)	68-38
	50-53	논(ノン)			50-54	농(ノン)			50-55	농(ノン)			50-57	농(ノン)	
聾	54-76	롱(ロン)	47-24	牢	54-79	뢰(ルエ)	47-20	磊	54-80	뢰(ルエ)	66-93	賂	54-81	뢰(ルエ)	47-08
	50-58	농(ノン)			50-62	뇌(ヌエ)			50-63	뇌(ヌエ)			50-65	뇌(ヌエ)	
雷	54-84	뢰(ルエ)	45-75	壘	55-04	루(ル)	52-62	屢	55-06	루(ル)	28-40	樓	55-07	루(ル)	60-76
	50-66	뇌(ヌエ)			50-68	누(ヌ)			50-69	누(ヌ)			50-70	누(ヌ)	
淚	55-08	루(ル)	46-62	漏	55-09	루(ル)	47-19	累	55-11	루(ル)	46-63	縷	55-12	루(ル)	69-63
	50-71	누(ヌ)			50-72	누(ヌ)			50-73	누(ヌ)			50-74	누(ヌ)	
陋	55-16	루(ル)	79-91	勒	55-45	룩(ルク)	80-53	肋	55-46	룩(ルク)	47-30	凜	55-47	름(ルム)	49-59
	50-75	누(ヌ)			50-80	늑(ヌク)			50-81	늑(ヌク)			50-82	늑(ヌム)	

表2 同一字で場合によって音の変わる字(続き)

凌	55-48 50-83	릉(ルン) 능(ヌン)	46-31	稜	55-50 50-84	릉(ルン) 능(ヌン)	46-39	綾	55-51 50-85	릉(ルン) 능(ヌン)	16-29	菱	55-52 50-87	릉(ルン) 능(ヌン)	41-09
陵	55-53 50-88	릉(ルン) 능(ヌン)	46-45												
頭音法則による場合(ㄷ音→ㅇ音)(112字)															
掠	53-51 69-17	략(リャク) 약(ヤク)	46-11	略	53-52 69-18	략(リャク) 약(ヤク)	46-12	亮	53-53 69-25	량(リャン) 양(ヤン)	46-28	兩	53-55 69-27	량(リャン) 양(ヤン)	49-32
涼	53-56 69-28	량(リャン) 양(ヤン)	49-58	梁	53-57 69-36	량(リャン) 양(ヤン)	46-34	糧	53-61 69-46	량(リャン) 양(ヤン)	46-40	良	53-62 69-48	량(リャン) 양(ヤン)	46-41
諒	53-63 69-50	량(リャン) 양(ヤン)	46-42	量	53-65 69-54	량(リャン) 양(ヤン)	46-44	勵	53-68 69-90	려(リョ) 여(ヨ)	50-15	呂	53-69 69-91	려(リョ) 여(ヨ)	47-04
廬	53-70 69-94	려(リョ) 여(ヨ)	55-10	旅	53-73 70-01	려(リョ) 여(ヨ)	46-25	濾	53-75 70-04	려(リョ) 여(ヨ)	63-41	礪	53-76 70-07	려(リョ) 여(ヨ)	66-74
閭	53-79 70-13	려(リョ) 여(ヨ)	79-67	驪	53-81 70-15	려(リョ) 여(ヨ)	81-75	麗	53-82 70-16	려(リョ) 여(ヨ)	46-79	黎	53-83 70-17	려(リョ) 여(ヨ)	83-53
力	53-84 70-19	력(リョク) 역(ヨク)	46-47	曆	53-85 70-23	력(リョク) 역(ヨク)	46-81	歷	53-86 70-24	력(リョク) 역(ヨク)	46-82	轆	53-89 70-28	력(リョク) 역(ヨク)	77-64
憐	53-91 70-38	련(リョン) 연(ヨン)	46-89	戀	53-92 70-39	련(리ョン) 연(ヨン)	56-88	漣	53-94 70-50	련(리ョン) 연(ヨン)	46-90	煉	54-01 70-54	련(리ョン) 연(ヨン)	46-91
璉	54-02 70-57	련(리ョン) 연(ヨン)		練	54-03 70-63	련(리ョン) 연(ヨン)	46-93	聯	54-04 70-65	련(리ョン) 연(ヨン)	46-94	蓮	54-05 70-69	련(리ョン) 연(ヨン)	47-01
輦	54-06 70-68	련(리ョン) 연(ヨン)	77-51	連	54-07 70-70	련(리운) 연(ユン)	47-02	鍊	54-08 70-72	련(리ョン) 연(ヨン)	47-03	列	54-10 70-74	렬(리ョ르) 열(ヨル)	46-83
劣	54-11 70-75	렬(리ョ르) 열(ヨル)	46-84	烈	54-13 70-79	렬(리ョ르) 열(ヨル)	46-85	裂	54-14 70-81	렬(리ョ르) 열(ヨル)	46-86	廉	54-15 70-85	렴(리요ム) 염(요ム)	46-87
殮	54-17 70-89	렴(리요ム) 염(요ム)		簾	54-19 71-01	렴(리요ム) 염(요ム)	46-92	獵	54-20 71-06	렵(리요브) 염(요브)	64-58	令	54-21 71-09	령(리요ン) 영(요ン)	46-65
囹	54-23 71-10	령(리요ン) 영(요ン)	51-90	嶺	54-26 71-13	령(리요ン) 영(요ン)	46-70	伶	54-27 71-16	령(리요ン) 영(요ン)	46-71	玲	54-28 71-31	령(리요ン) 영(요ン)	46-72
鈴	54-30 71-38	령(리요ン) 영(요ン)	70-25	聆	54-32 71-39	령(리요ン) 영(요ン)	70-57	鈴	54-34 71-43	령(리요ン) 영(요ン)	46-75	零	54-35 71-45	령(리요ン) 영(요ン)	46-77
靈	54-36 71-47	령(리요ン) 영(요ン)	80-45	領	54-37 71-48	령(리요ン) 영(요ン)	46-46	例	54-39 71-51	례(리예) 예(이예)	46-67	禮	54-41 71-63	례(리예) 예(이예)	67-25
醴	54-42 71-68	례(리예) 예(이예)	78-52	隸	54-43 71-70	례(리예) 예(이예)	46-76	了	54-85 72-71	료(리요) 요(요)	46-27	僚	54-86 72-72	료(리요) 요(요)	46-29
寮	54-87 72-80	료(리요) 요(요)	46-32	料	54-89 72-87	료(리요) 요(요)	46-33	燎	54-90 72-91	료(리요) 요(요)	63-89	療	54-91 72-94	료(리요) 요(요)	46-37
蓼	54-94 73-07	료(리요) 요(요)	72-90	遼	55-01 73-12	료(리요) 요(요)	46-43	龍	55-03 73-44	룡(리요ン) 용(요ン)	46-22	劉	55-17 74-69	류(리유) 유(유)	46-13
柳	55-19 74-87	류(리유) 유(유)	44-88	流	55-21 74-92	류(리유) 유(유)	46-14	溜	55-22 74-94	류(리유) 유(유)	46-15	琉	55-24 75-04	류(리유) 유(유)	46-16
留	55-26 75-07	류(리유) 유(유)	46-17	硫	55-28 75-09	류(리유) 유(유)	46-18	類	55-30 75-26	류(리유) 유(유)	46-64	六	55-31 75-27	륙(리육크) 육(육크)	47-27
戮	55-32 75-29	륙(리육크) 육(육크)	57-04	陸	55-33 75-33	륙(리육크) 육(육크)	46-06	倫	55-35 75-34	륌(리윤) 윤(윤)	46-49	崙	55-36 75-38	륌(리윤) 윤(윤)	54-38
淪	55-37 75-39	륌(리윤) 윤(윤)	62-45	輪	55-39 75-44	륌(리윤) 윤(윤)	46-56	律	55-40 75-47	륎(리윤르) 윤(윤르)	46-07	慄	55-41 75-48	륎(리윤르) 윤(윤르)	56-43
栗	55-42 75-49	륎(리윤르) 윤(윤르)	23-10	率	55-43 75-50	륎(리윤르) 윤(윤르)	46-08	隆	55-44 75-56	룽(리윤) 용(윤)	46-20	利	55-55 76-06	리(리) 이(이)	45-88
吏	55-57 76-07	리(리) 이(이)	45-89	履	55-59 76-10	리(리) 이(이)	45-90	李	55-61 76-16	리(리) 이(이)	45-91	梨	55-62 76-17	리(리) 이(이)	45-92
理	55-66 76-21	리(리) 이(이)	45-93	痢	55-69 76-24	리(리) 이(이)	46-01	罹	55-71 76-26	리(리) 이(이)	56-77	裏	55-74 76-32	리(리) 이(이)	46-02

表2 同一字で場合によって音の変わる字(続き)

裡	55-75	리(リ)	46-03	里	55-76	리(リ)	46-04	離	55-78	리(リ)	46-05	吝	55-80	린(リン)	50-71
	76-33	이(イ)			76-37	이(イ)			76-38	이(イ)			76-53	인(イン)	
隣	55-82	린(リン)	46-53	隣	55-83	린(リン)	-	蘭	55-84	린(リン)	73-34	隣	55-86	린(リン)	46-57
	76-61	인(イン)			76-62	인(イン)			76-65	인(イン)			76-68	인(イン)	
隣	55-87	린(リン)	46-58	隣	55-88	린(リン)	46-59	林	55-89	림(リム)	46-51	淋	55-90	림(リム)	46-52
	76-71	인(イン)			76-72	인(イン)			76-87	임(イム)			76-88	임(イム)	
臨	55-92	림(リム)	46-55	立	56-01	립(リップ)	46-09	笠	56-02	립(リップ)	19-62	粒	56-03	립(リップ)	46-19
	76-90	임(イム)			77-01	입(イブ)			77-02	입(イブ)			77-03	입(イブ)	
その他(3字)															
異	55-68	리(リ)	16-59	不	60-84	부(ブ)	41-52								
	76-22	이(イ)			61-53	불(ブル)									
寧	50-27	녕(ニョン)	39-11												
	54-24	령(リョン)													
	71-12	영(ヨン)													

ルトン)ではなくて「부당」(ブタン)になり、「不足」の場合も「불족」(ブルジョク)ではなく「부족」(ブジョク)となる。

最後に、「異」の場合は、「이」(イ)欄と「리」(リ)欄に二度出ているが、この字は確かに本音が「이」(イ)で、文法的に「리」(リ)と読まれる場合はない。しかし、ただ一つの例外として、山の名である「智異山」(韓国国立公園第1号)の場合は「리」(リ)と発音するが、これは取音である。本来、「チリサン」は漢字はなかったのに、音によった漢字表記からきた特殊な例である。これらを整理すると表2「同一字で場合によって音の変わる字」のようになる。

1-2 異意異音同体字

第二は、同一字が二つ以上の意味と同時に音を持っている場合である。すなわち、異意異音同体字である。これは総数43字で二つの意味に二つの音を持っている字が40字、三つの意味と三つの音を持っている場合が3字ある。もちろん、この他にも二つまたは三つ以上の意味または音を持っている字もたくさんあるが、ここには代表的なものだけを取めた。同一字が意味によって発音を異にするのは、日本語とほとんど同じである。これを日本の発音と一緒に比較して、表3「異意異音同体字」にまとめた。

表3 異意異音同体字

字	KSC 行列	韓 音	訓	日音	JIS 区点	字	KSC 行列	韓 音	訓	日音	JIS 区点
賈	42-25	가(カ)	うる	コ、ク	76-43	降	43-29	강(カン)	おりる	コウ	25-63
	45-47	고(コ)	あたい	ケ、カ			90-02	항(ハン)	降伏	コウ	
豈	43-48	개(ケ)	あに	カイ	76-17	更	43-54	갱(ケン)	また、ふたたび	コウ	25-25
	49-34	기(キ)	かちどき	キ、ケ			44-58	경(キョン)	あらためる	コウ	
見	44-24	견(キョン)	見る、見える	ケン	24-11	契	44-88	계(キエ)	ちぎる	ケイ	23-32
	90-70	현(ヒョン)	謁する	ケン			48-48	글(クル)	北夷の号	キツ	
滑	45-72	굴(コル)	すべる	カツ	19-74	串	45-90	곶(コツ)			22-90
	92-33	활(フッル)	にごす(濁す)	ニツ			46-13	관(クワン)	なれる	カン	
廓	46-09	곽(クワク)	くるわ	カク	19-39	金	48-61	금(クム)	かね	キン	22-66
	92-09	확(フワク)	ひろい	カク			49-49	김(キム)			
宅	51-75	택(タク)			34-80	度	51-88	도(ト)	のり、法制、法度	ト、ド	37-57
	87-40	택(タク)	すまい	タク			86-84	탁(タク)	はかる、謀也	タク	
讀	52-33	독(トク)	よむ	ドク	76-06	洞	52-55	동(トン)	ほら	ドウ	38-22
	52-70	두(トゥ)					87-51	통(トン)	抜き通す	ドウ	
率	55-43	률(リュル)	わりあい	リツ	46-08	北	59-37	배(ベ)	にげる、敗走する	ホク	43-44
	65-67	솔(ソル)	ひきいる	ソツ			61-33	북(ブク)	きた	ホク	
便	60-05	변(ビョン)	いばり	ベン	42-56	復	60-54	복(ボク)	かえる、かえす	フク	41-92
	88-21	편(ピョン)	たより	ベン			61-05	부(フ)	また、ふたたび	フツ	
數	62-92	삭(サク)	しばしば	サク	58-43	索	62-94	삭(サク)	なわ	サク	26-87
	66-06	수(ス)	かぞえる	スウ			63-67	색(セク)	もとめる	サク	
殺	63-15	살(サル)	ころす	サツ	27-06	參	63-19	삼(サム)	みつつ	サン	50-52
	65-77	쇄(スエ)	そげる	サイ			83-49	참(チャム)	まじる	サン	
状	63-50	상(サン)	かたち	ジョウ	30-85	塞	63-61	새(セ)	辺境	サイ	26-41
	77-78	장(チャン)	文書、便紙	ジョウ			63-65	색(セク)	ふさぐ	ソク	
省	63-72	생(セン)	はぶく	ショウ	30-42	葉	64-81	섭(ソブ)	姓氏	セフ	45-53
	63-93	성(ソン)	かえりみる、よくみる	セイ			71-08	엽(ヨブ)	は、草木のは	エウ	
拾	67-06	습(スブ)	ひろう	シュウ	29-06	識	67-59	식(シク)	しる	ショク、シキ	28-17
	68-09	십(シブ)	数の名	ジュウ			82-29	지(チ)	しるし、しるす	シ	
辰	67-85	신(シン)	とき、時也	シン	35-04	沈	68-01	심(シム)	姓氏	シン	36-32
	82-67	진(チン)	たつ(辰)、十二支の第五位	ジン			86-56	침(チム)	しずむ	チン	
什	68-07	십(シブ)	十人、十家	ジュウ	29-26	惡	68-34	악(アク)	わるい	アク	56-08
	82-90	집(チブ)	什器	ジュウ			71-87	오(オ)	にくむ	オ	
若	69-14	약(ヤ)	般若	ニヤ	28-67	易	70-22	역(ヨク)	かえる	エキ	16-55
	69-20	약(ヤク)	わかい、もしくは	ジャク			76-15	이(イ)	やさしい	イ	
咽	70-76	열(ヨル)	むせぶ	エツ	16-86	瑩	71-33	영(ヨン)	玉に似た美しい石	エイ	64-82
	76-54	인(イン)	のど	イン			91-09	형(ヒョン)	まとう	エイ	
切	79-23	절(チョル)	きる	セツ	32-58	拓	84-12	척(チョク)	ひろく	タク	34-83
	84-78	체(チェ)	すべて	サイ			86-86	탁(タク)	いしずり	タク	
暴	88-59	포(ポ)	あばれる	ポウ	43-29	行	90-01	항(ハン)	ならび	ギョウ	25-52
	88-76	폭(ボク)	あらわす	バク			90-28	행(ヘン)	ゆく	コウ(カウ)	
龜	47-47	구(ク)	昔の国名、山の名	キュウ	83-93	樂	53-05	락(ラク)	たのしい	ラク	60-59
	48-02	귀(クイ)	かめ	キ			68-37	악(アク)	たのしい、音楽	ガク	
	48-24	균(キュン)	亀裂	キ			72-89	요(ヨ)	このむ	ゴウ	
說	64-67	설(ソル)	とく	セツ	32-66	(43字)					
	65-13	세(セ)	あやす	ゼイ							
	70-82	열(ヨル)	うれしい	エツ							

1-3 同意異音同体字

第三は、一字が同じ意味で二つの音を持ち、場合によってその発音を異にする場合である。この場合は、本来二つの音を持っている字もあるし、後からもう一つの音を加えた字もあり、また、俗音で発音する字もある。これらは合計 13 字で、表 4 「同意異音同体字」に示した。

付記すべきやや特殊な例として、「句」、「丹」、「糖」、「磻」の 4 字がある。

「句」：常用「句」を付けてつくられた単語は、慣習上、「子」(ク)または「𠵼」(クィ)と読む傾向もあった。しかし、句の音は「子」(ク)であるため、「𠵼」(クィ)に発音されるわけがないから、今はすべて「子」(ク)で統一し、「𠵼」(クィ)と発音するのは間違いとなる。これは 1988 年 1 月 19 日に韓国文教部が新しく改正告示して 1989 年 3 月 1 日から施行した、「標準語規定」の中で、「標準語査定原則」(第 13 項)には「漢字句が付けて作られた単語は「𠵼」(クィ)と読むことを認定せず、「子」(ク)で統一する」と規定して、漢字音はすべて「子」(ク)に統一した。したがって、KSC 行列 47-91 の「句」はもう要らなくなった字である。

「丹」：「丹」は本音が「단」(タン)であるが、特殊な例として、「牡丹」の場合だけ、伝統的に「모단」(モタン)ではなくて、「모란」(モラン)と読む習慣を認めて、「모란」(モラン)を標準語にして

表 4 同意異音同体字

字	KSC 行列	韓 音	訓	日音	JIS 区点	字	KSC 行列	韓 音	訓	日音	JIS 区点
車	43-71	거(コ)	くるま	シャ	28-54	茶	50-94	다(タ)	おちゃ	サ	35-67
	83-19	차(チャ)		シャ			83-17	차(チャ)		チャ	
輻	60-63	복(ボク)	車のや轂と牙園 とを支える材	フク	77-53	泌	61-84	비(ビ)	早い流、 ほそい流	ヒ	40-71
	88-80	폭(ボク)		フウフ			89-18	필(ピル)		ヒツ	
阮	72-54	완(ワン)	古の国の名、 月琴	ケン	79-86	暈	73-87	운(ウン)	かさ、日月の周辺に 出来る淡水の電気	ウン	58-84
	74-33	원(ウォン)		グワン			93-27	훈(フン)		ウン	
刺	77-09	자(チャ)	さす	シ	27-41	灸	77-19	자(チャ)	あぶる	シャ	63-53
	84-07	작(チョク)		セキ			78-59	작(チョク)		セキ	
句	47-03	구(ク)	文章のくぎり	ク	22-71	丹	51-01	단(タン)	あか	タン	35-16
	47-91	귀(クィ)					53-14	란(ラン)			
糖	51-56	당(タン)	あめ、飴也、 砂糖	トウ	37-92	磻	58-82	반(パン)	水の名	ハン	
	87-24	탕(タン)					59-68	번(ボン)			
奈	49-53	나(ナ)	いかん	ダイ	38-64						
	50-15	내(ネ)		ナ							

いる。つまり国音字である。

「糖」：やはり本音は「당」(タン)であるが、特殊な例として「砂糖の場合だけ「탕」(タン)と読む習慣を認めて、「사탕」(サタン)、「설탕」(ソルタン)などを標準語にしている。なぜならば、これら二つの字は本音で読むと、その発音が自然らしくないからである。

「磻」：「반」(パン)と「번」(ボン)に重出されているが、この字の音はたしかに「반」(パン)で、「번」(ボン)と読む例はどの字典にも現われない。しかし、地名としてソウル(서울)市の洞名(恩平区磻磻洞)で 번(ボン)につかっている。日本工業規格には出ていない字である。

これらの重出字は、大部分、日本工業規格にも含まれているが、ただし、8字は含まれていない。含まれていない字の表は表1の下端に収めた。

2 韓国字

世に言われている「国字」とは、普通、昔から中国で作られて使われてきた字ではなくて、漢字を使っている各国において必要によって作られ、その国だけで通用している、その国の固有漢字を称する用語である。もちろん、日本にも少なくない国字があり、もう日本の範囲を越えて使用されている字もある。例えば、「秆」、「榧」、「呷」などの字は、日本字として知られているが、韓国字典にも収められている。そして、韓国の標準規格にも字典にも載っていない4字（「峠」、「畑」、「枳」、「笹」）の日本字が、形声による造音に含まれている。反対に、日本の字典にも韓国字で知られている字（「迭」など）が収められている。韓国の国字を、本稿では便宜上、韓国字と呼ぶことにする。

韓国字の場合、歴史が非常に古く三国時代までさかのぼる。韓国標準規格に収められている韓国字の場合、二つに分けて考えることができる。第一は、純粋に韓国で造られて韓国だけで使っている純粋韓国字である。第二は、すでに作られている字、すなわち既存の字に、特定の意味が追加されて使われている半漢半韓字がある。韓国字の全体的な数字は正確に知ることはできないが、概ね500字以上として知られている。しかし、韓国標準規格に含まれている韓国字はわずかなものである。先ず純粋韓国字の例を挙げると「芻」「迭」「詵」「沓」「笠」「芑」「楡」「泫」「豊」「邕」「釗」「簍」「鏑」「媿」「儂」「箕」「櫛」「岫」「曹」「轄」（20字）だけである。

この中で「泫と釗」は半漢半韓字で、本来は泫は「回流」の意味であったが、「灌漑用の堰」という韓国的意味が追加されて、二つの意味がある。実際の生活には主として後者の意味で使われている。そして、韓国標準規格にも後者の意味で使われているから、ここでは純粋韓国語として分類した。「回流」の意味で使うときは、その音が「峯」(ポク)であるが、灌漑用の堰の意味で使われるときは、その音が「阜」(ポ)である。韓国標準規格には「阜」(ポ)で使われている。釗もやはり本来は漢字として、みる「缶(ソ)」人名「丑(キョ)」などで使っているが、韓国的意味の鐵、又は韓国式の名又は地名が追加されて使っている。この時にはその音が斗(スエ)である。韓国標準規格には斗(スエ)で使われているから純粋韓国字として分類した。

そして「簍」は現在、全羅北道にある地名(箕簍島里)で使われているが、近世に形成されたものである。本来は「食」であったが、その地方の特産物が竹の製品のため、その地方性を表現するために「竹」部を加えたものである。

もうひとつの「儂」は、韓国の古代の国の名前前で、本来は「亻」部がないまま使われたが、後に「伽」と合わせるために「亻」部が追加されて「伽儂」になったものである。

第2番目は半漢半韓字である。これに該当する字は、「金」、「讀」、「串」、「宅」の4字である。これらの字は下記の表でみるように、上の欄の意味は本来の漢字の意味であって、下の欄の方は韓国字で韓国だけで使われている固有の意味と固有の音を持っている。これらはすべて日本工業規格にも含ま

表5 韓国字

純粹国字(20字)			
字	KSC 行列	音	意
芑	42-65	갈(カル)	人名又は地名、韓国語の終声表記から形成されたもので別の意味はない
迭	44-06	겹(コブ)	コブは類推音、地名
畱	51-44	답(タフ)	田、たんぼ、水田、水と田の二字の配合造字
垩	51-60	대(テ)	家をたてる土地、垩地 ※JISにも出ている(52-18)
芑	52-44	돌(トル)	人名又は地名、韓国語の終声表記から形成されたもので別の意味はない
朥	56-37	글(クッ)	人名又は地名、韓国語の終声表記から形成されたもので別の意味はない
椷	57-07	명(ミョン)	寛也、地名
洑	60-39	보(ボ)	灌漑用の堰、堰留め、いせき、ため池
芑	60-67	불(ボル)	人名又は地名、韓国語の終声表記から形成されたもので別の意味はない
芑	63-13	살(サル)	人名又は地名、韓国語の終声表記から形成されたもので別の意味はない
簞	67-57	식(シク)	地名
鐺	64-54	선(ソン)	油計量器、真鍮製のものが多い
釧	65-82	쇠(スエ)	鐵、人名、地名
媿	67-24	시(シ)	(名詞の前に付いて)夫の家の……、婚家の…… -父母：夫の母 舅と姑 -同生：夫の弟、義弟
倻	69-06	야(ヤ)	国名(伽倻)、地名(伽倻山)、樂器名(伽倻琴)
篔	72-04	오(オ)	柳行李(やなぎごうり)、地名
穢	77-75	장(チャン)	穢籠 1. 衣服を入れるやや小さめの箆笥、 2. 服などを入れておく箆笥や行李
岫	79-32	점(チョム)	地名、とうげ ※JISにも出ている(54-19)
曹	80-38	조(チョ)	韓国の姓氏
囍	93-78	희(フィ)	喜が重なるという意味
半漢半韓字として重出された字(4字)			
串	46-13	관(クワン)	なれる(慣)
	45-90	곶(コツ)	(地名に付けて用いて)……岬、長山-：チャンサン岬
金	48-61	금(クム)	かね
	49-49	김(キム)	韓国の姓氏
讀	52-33	독(トク)	よむ
	52-70	두(トゥ)	吏讀：三国時代から漢字の音と意味を借りて韓国語を表記す時に使った字
宅	87-40	택(タク)	すまい
	51-75	택(タク)	婦人の里の地名に付けて、その地方から嫁いで来た人であることを示す

れているが、それはもちろん表に示した上の欄の意味であると思われる。

以上のように、韓国標準規格に含まれている韓国字は、半漢半韓字を足しても、わずか 24 字にすぎない。この中で日本工業規格にも含まれている半漢半韓字と韓国字の中で日本工業規格にも含まれている 2 字堡 (52-18)、帖 (54-19) (この 2 字は韓国と日本で各々作られて両国の国字として使っているが、偶然その字形が一致する特異な字である。もちろんその意味は互いに違う) を除けば、韓国字の問題はささやかな問題だろう。これらの韓国字を表 5 「韓国字」としてまとめた。

3 日本工業規格との比較

前節までに韓国標準規格の性格をまとめたが、韓国標準規格と日本工業規格を比較してみると次のように対照することができる。

3-1 体 裁

先ず全体的な体裁面からみると、両者はいずれも 2 バイト型で、一行に 94 字ずつ配列されている点と同じである。

二番目は収録字の区分と配列である。日本工業規格は第 1 水準と第 2 水準で分けて、第 1 水準では基本的な性質の漢字 2,965 字を各漢字ごとに代表音訓 1 個を定め、その五十音順に配列し、同音訓のものは音→訓の順、同音または同訓のものは見出し形順・画数順に配列している。

そして、第 2 水準には第 1 水準以外の漢字 3,388 字を、その字形の見出し形によって康熙字典の 214 部に分類し、同部のものは原則として画数順に配列している。そして、同画数のものは、代表音の五十音順に配列して、若干のものを末尾に補充している。

これに比べ、韓国標準規格は水準の区分なく 4,888 字をすべてハングル字母順、すなわち順で配列して、同音のものは康熙字典の部に分類し、同部のものは画数順に配列している。これはハングルが日本語より音幅がひろく音が単純で、同音字が少ないから、音が重視されているからであると思われる。

3-2 字 数

両者の一番大きな特徴と同時に一番重要なものは、何と言っても字数ではないかと思われる。単純に収録字数を比べると、日本工業規格が 6,353 字、韓国標準規格が 4,888 字で、その差は 1,465 字に達している。しかし、詳しく調べると前述のごとく、韓国標準規格には字種 262 字に字数 530 字の重出字があるから、日本工業規格のように字形面からみると、実は 4,620 種の字が収録されているわけである。こういう現象は日本工業規格も同じことで、たとえ等しい字が重出された場合はないとしても、多くの異体字があつて、結果的には同じ現象であるとみることができる。特に異体字のうちまったく同じであると思われる略字だけでも、およそ 400 字近くに及んでいる。

これらを一字一字比べた結果、日本工業規格にはなく、韓国標準規格だけにある字は 426 種、434 字で、重出字を一字に計算すると、426 字になる。もちろん、この中には略字化によって少しずつ違った字が多くあるが、同一字として取り扱った結果である。同一字であるが、少しずつ違う形を持ってい

る字が12字あったが、単純に習慣によるものと思われるので、事実上、同一字とみなして取り扱った。これらを同一字でなく異体字として扱うと字数は438字に増える。同一字として扱った異体字を表6「同一字として取り扱った異体字」に掲げた。

表6 同一字として取り扱った異体字

KSC 字	KSC 行列	JIS 字	JIS 区点	KSC 字	KSC 行列	JIS 字	JIS 区点
慤	42-34	慤	56-34	癩	42-54	癩	65-82
髯	71-03	髯	81-89	勛	73-77	勛	50-08
廿	76-94	廿	38-91	纘	83-37	纘	69-83
恬	84-54	甜	37-28	恩	85-34	忽	55-68
愠	85-35	愠	56-27	蔥	85-39	葱	39-12
颯	88-92	颯	81-10	畫	91-94	畫	65-33

韓国字が24字含まれているのは前述したとおりである。特異なもので韓・日・中字典をいくら捜しても出ていない字が6字(「玆」(57-63)、「璵」(70-06)、「琿」(72-45)、「亏」(74-07)、「櫛」(90-43)、「滢」(91-05)があるが、この中で「玆」(ㅈ:ミョ)、「亏」(ㅇ:ウル)「璵」(ㅇ:ヨ)は地名で「琿」(ㅇ:ワン)、「櫛」(ㅇ:ホン)、「滢」(ㅇ:ヒョン)は人名で使っているがこれらの字の音と形でみると、たぶん形声字による造字ではないかと思われる。

全体的に見ると、これらの字はごく一部を除けば、ほとんど人名とか地名で使われているもので、その使用の頻度は幾字(例えば「畚」、「媿」、「曹」など)を除けば極めて少ない字ばかりである。これらを表で一瞥すれば次のとおりである。また検索の便宜のため部首順で整理した(表7「韓国標準規格だけ出ている字」、表8「表7を部首順で整理したもの」)。

表 7 韓国標準規格にのみ出ている字

表側は KSC 行列の行。表頭は KSC 行列の列。

行	列	426 種 434 字	重出字	
42	珏 ₃₆ 礪 ₅₆ 芑 ₆₅			3
43	鉀 ₅ 罈 ₈ 罍 ₁₇ 甗 ₂₄ 塏 ₃₅ 愷 ₃₆ 祛 ₆₈ 楗 ₈₁			8
44	鈴 ₂ 迭 ₆ 徑 ₄₁ 倥 ₄₂ 倣 ₄₄ 垆 ₄₈ 擎 ₅₄ 暎 ₅₇ 涇 ₆₀ 炅 ₆₁ 璟 ₆₃			12
	璣 ₆₄			
45	槩 ₁ 礩 ₅ 曷 ₂₈ 苾 ₄₁ 琨 ₆₈ 珙 ₈₅			6
46	藿 ₁₁ 瑄 ₂₁ 瓘 ₂₂ 愨 ₃₁ 适 ₃₃ 恍 ₃₄ 吹 ₄₁ 珖 ₄₃ 槐 ₅₇ 嶠 ₆₈			10
47	坵 ₆ 絛 ₂₉ 耆 ₃₀ 録 ₄₂ 芎 ₆₈ 港 ₇₈ 晷 ₉₂			7
48	勻 ₁₈ 昀 ₂₀ 筠 ₂₁ 勗 ₃₄ 姁 ₅₀ 吟 ₅₂ 芩 ₅₇ 倂 ₆₃ 夔 ₈₃			9
49	玘 ₁₂ 琦 ₁₃ 琪 ₁₄ 璠 ₁₅ 璣 ₁₆ 璩 ₃₀ 錡 ₃₆ 錫 ₃₇ 枅 ₈₇ 浦 ₈₉			10
50	柰 ₁₆ 季 ₂₂ 穉 ₂₈ 擣 ₃₄ 菘 ₅₀ 杻 ₇₈	季 擣 菘 杻		6
51	澆 ₂₂ 坍 ₂₇ 聃 ₃₆ 鏐 ₄₂ 畚 ₄₄ 逌 ₄₇ 戇 ₅₂ 圯 ₆₁			8
52	燾 ₇ 墩 ₃₄ 吨 ₃₇ 焯 ₄₀ 芑 ₄₄ 料 ₆₆ 芑 ₇₆ 鄧 ₈₈			8
53	擘 ₂₆ 唼 ₄₇ 樑 ₅₈			3
54	璉 ₂ 殮 ₁₇ 岑 ₂₅ 翎 ₃₁ 澧 ₄₀ 擣 ₄₇ 潞 ₄₉ 菘 ₆₅	璉 殮 擣 菘		8
55	萸 ₁₃ 唼 ₅₈ 潏 ₈₁ 璘 ₈₃ 砒 ₉₄	璘		5
56	邈 ₁₇ 忝 ₃₇ 輞 ₅₄ 邛 ₅₅ 沔 ₈₇			5
57	椹 ₇ 萸 ₁₂ 摹 ₂₅ 芑 ₃₇ 妙 ₆₃ 珉 ₇₈			6
58	汶 ₁ 雯 ₇ 沔 ₉ 楣 ₁₉ 漾 ₂₀ 湄 ₂₁ 旼 ₃₆ 玫 ₃₉ 珉 ₄₀ 泮 ₇₄ 盼 ₈₀	磻		12
	磻 ₈₂			
59	昉 ₁₈ 龐 ₃₄ 裴 ₄₉ 褱 ₅₀ 磻 ₆₈	磻		5
60	鰲 ₁₄ 晒 ₂₁ 曷 ₂₂ 榑 ₂₄ 駟 ₃₀ 洑 ₃₉ 淝 ₄₀ 潛 ₄₁ 瑤 ₄₂ 宓 ₅₃ 萑 ₆₇			13
	燧 ₇₆ 琫 ₇₇			
61	苧 ₂₀ 駟 ₃₁ 盼 ₄₃ 愨 ₈₀ 毗 ₈₁			5
62	嘖 ₁₃ 玘 ₂₃ 僊 ₃₉ 栖 ₅₈			4
63	迄 ₁₃ 煞 ₁₆ 鋈 ₂₈ 捷 ₈₂			4
64	奭 ₁₁ 敷 ₃₂ 渲 ₃₄ 琤 ₃₆ 瑄 ₃₇ 璇 ₃₈ 璿 ₃₉ 誦 ₅₀ 鐸 ₅₄ 饒 ₅₅ 高 ₅₇			14

表7 韓国標準規格にのみ出ている字(続き)

	剡 ⁷⁰ 戍 ⁸⁴ 城 ⁹¹		
65	炤 ³¹ 瘙 ³⁶ 涑 ⁵⁵ 蓀 ⁶⁴ 滄 ⁶⁶ 釗 ⁸² 岨 ⁹²		7
66	琇 ¹⁵ 璩 ¹⁶ 瀟 ⁵⁶ 琠 ⁵⁸ 璫 ⁵⁹ 枸 ⁶⁷ 檠 ⁶⁹ 珣 ⁷³ 錡 ⁸⁶ 鉤 ⁹² 崧 ⁹⁴		11
67	媿 ²⁴ 湜 ⁵⁵ 簋 ⁵⁷ 仇 ⁶⁴ 莘 ⁷⁸ 蓋 ⁸⁰		6
68	諶 ⁶ 唵 ⁵⁸ 厓 ⁷⁷		3
69	郇 ⁶ 敷 ³⁴ 痲 ⁶⁰ 擘 ⁷⁷ 窠 ⁸⁵		5
70	璵 ⁵ 礪 ⁶ 鮪 ⁹ 輦 ¹² 堧 ³² 挺 ⁴¹ 沅 ⁴⁴ 璉 ⁵⁷ 季 ⁶⁰ 績 ⁶⁴ 殮 ⁸⁹	季璉殮	12
	琰 ⁹²		
71	燁 ⁷ 嶸 ¹⁴ 漢 ²³ 滌 ²⁵ 潛 ²⁷ 煨 ²⁸ 鎡 ⁴⁴ 汭 ⁵⁵ 澨 ⁵⁶ 芮 ⁶⁰ 佻 ⁷⁵		12
	昨 ⁹⁰		
72	癸 ³ 箕 ⁴ 鈺 ¹³ 璠 ¹⁵ 緼 ¹⁸ 邕 ²⁷ 饗 ²⁹ 椀 ⁴¹ 琯 ⁴⁵ 琬 ⁴⁶ 腕 ⁵⁰		11
73	繇 ³ 埆 ²⁵ 墉 ²⁶ 瑑 ³⁵ 鏞 ⁴³ 盱 ⁵⁷ 玕 ⁵⁹ 瑀 ⁶⁰ 耦 ⁶³ 釭 ⁷³ 雱 ⁷⁶		21
	彘 ⁷⁸ 昱 ⁸⁰ 栒 ⁸¹ 煜 ⁸² 穢 ⁸³ 頊 ⁸⁵ 榑 ⁸⁸ 澮 ⁹⁰ 煨 ⁹¹ 蕓 ⁹⁴		
74	亏 ⁷ 嫫 ¹⁷ 沅 ²² 洄 ²³ 瑗 ²⁸ 暉 ⁴⁹ 璋 ⁵² 蔞 ⁵⁷ 禕 ⁶⁰ 庾 ⁷⁶ 杻 ⁸⁴	杻	12
	洧 ⁹¹		
75	埆 ²⁸ 齋 ³⁶ 玗 ⁴¹ 鈺 ⁴⁵ 融 ⁵³ 閭 ⁶¹ 蕙 ⁹¹		7
76	萑 ³¹ 瀾 ⁴³ 璘 ⁶² 綰 ⁶³ 靺 ⁷⁰ 佻 ⁷⁵ 駟 ⁸¹	璘	7
77	苈 ⁷ 玆 ²¹ 暉 ⁷¹ 穢 ⁷⁵ 獐 ⁷⁹		5
78	澨 ¹⁰ 菹 ⁴¹ 翟 ⁶⁵ 佺 ⁷⁷ 塿 ⁸⁴		5
79	琬 ² 癯 ²⁷ 妊 ⁴⁹ 畚 ⁶¹ 楨 ⁶³ 檉 ⁶⁴ 淳 ⁶⁹ 瀆 ⁷⁰ 炆 ⁷² 玎 ⁷³ 玼 ⁷⁴		14
	緹 ⁸² 諄 ⁸⁵ 鋌 ⁹¹		
80	曹 ³⁸ 璪 ⁴⁹ 侗 ⁸² 惊 ⁸⁵ 琮 ⁸⁹		5
81	姝 ¹⁵ 澍 ³¹ 炷 ³² 塿 ⁵⁶ 窩 ⁵⁷ 竣 ⁵⁹ 竣 ⁶⁴ 竣 ⁶⁵ 茁 ⁷² 繪 ⁸⁹		10
82	芷 ²⁶ 搢 ⁴² 振 ⁴⁵ 溱 ⁴⁹ 璫 ⁵¹ 璫 ⁵² 績 ⁶⁰ 藺 ⁶² 侄 ⁷³ 瓚 ⁷⁹ 灑 ⁹²		11
83	鑠 ¹ 灑 ²⁹ 璨 ³¹ 瓚 ³² 塚 ⁸¹ 窠 ⁸²		6
84	侗 ⁶ 憾 ¹⁰ 玗 ²⁹ 喆 ⁴² 澈 ⁴⁵ 詹 ⁵⁸ 堧 ⁶⁰		7
85	茗 ¹³ 醮 ²⁰ 摠 ³⁶ 騶 ⁶⁶ 琿 ⁸²		5
86	淄 ³⁰ 柒 ⁵¹ 琛 ⁵⁸ 拖 ⁷³ 坼 ⁸³ 皦 ⁸⁸ 瑯 ⁹³		7

表7 韓国標準規格にのみ出ている字(続き)

87	跽 ³⁷ 郅 ³⁸ 撐 ⁴³ 攄 ⁴⁴ 辦 ⁹³		5
88	涓 ¹⁰ 杵 ³⁴ 佈 ⁴⁷		3
89	秘 ¹⁹ 苾 ²³ 秘 ²⁴ 昆 ³² 閒 ⁵⁷ 姁 ⁸³ 沆 ⁹⁰		7
90	瀧 ¹⁴ 苻 ²⁷ 珣 ³² 櫛 ⁴³ 熨 ⁵⁰ 峴 ⁵⁴ 睨 ⁵⁷ 洙 ⁵⁸ 炫 ⁵⁹ 玆 ⁶¹ 睨 ⁶⁴		11
91	洞 ² 榮 ³ 滢 ⁴ 滢 ⁵ 熒 ⁷ 玗 ⁸ 邢 ¹⁴ 鎒 ¹⁵ 曄 ²¹ 蕙 ²² 岾 ³²		20
	溟 ⁴⁰ 澔 ⁴³ 濩 ⁴⁵ 灑 ⁴⁶ 祐 ⁵² 護 ⁶⁵ 顛 ⁶⁶ 烘 ⁸⁴ 燁 ⁹¹		
92	碯 ¹³ 皖 ²³ 紉 ²⁸ 愧 ⁴² 梲 ⁴⁶ 潢 ⁵⁰ 璜 ⁵² 澮 ⁷² 弘 ⁸³ 鑽 ⁸⁵ 敷 ⁹⁰		12
	洵 ⁹³		
93	幪 ¹² 翊 ¹⁶ 勛 ¹⁸ 塤 ²⁰ 堦 ²¹ 焜 ²² 煊 ³¹ 輝 ⁴⁰ 昕 ⁵⁸ 炘 ⁶⁰ 紇 ⁶⁴		16
	歆 ⁶⁸ 熙 ⁷⁵ 囍 ⁷⁸ 愜 ⁸³ 熿 ⁸⁹		

表8 表7を部首順で整理したもの

乙	芑 ⁴²⁻⁶⁵ 芑 ⁶⁰⁻⁶⁷ 芑 ⁶³⁻¹³	人	伋 ⁴⁸⁻⁶³ 佈 ⁸⁸⁻⁴⁷ 佻 ⁴⁶⁻³⁴ 佻 ⁶⁷⁻⁶⁴ 佻 ⁷⁶⁻⁷⁵ 佻 ⁷⁸⁻⁷⁷ 佻 ⁸²⁻⁷³ 佻 ⁴⁴⁻⁴¹ 佻 ⁷¹⁻⁷⁵ 倮 ⁴⁴⁻⁴² 倮 ⁸⁰⁻⁸² 倮 ⁸⁴⁻⁰⁶
亻	儻 ⁶⁹⁻⁰⁶ 儻 ⁶²⁻³⁹ 儻 ⁴⁴⁻⁴⁴	彡	飡 ⁶⁵⁻⁶⁶
刀	剡 ⁶⁴⁻⁷⁰	力	勛 ⁴⁸⁻³⁴ 勛 ⁹³⁻¹⁸ 辦 ⁸⁷⁻⁹³
勹	勻 ⁴⁸⁻¹⁸	口	高 ⁶⁴⁻⁵⁷ 唳 ⁵⁵⁻⁵⁸ 唳 ⁶⁸⁻⁵⁸ 唳 ⁶²⁻¹³
土	喆 ⁸⁴⁻⁴² 囍 ⁹³⁻⁷⁸ 坩 ⁵¹⁻²⁷ 坩 ⁴⁷⁻⁰⁶ 坩 ⁵¹⁻⁶¹ 坩 ⁸⁶⁻⁸³ 坩 ⁴⁴⁻⁴⁸ 坩 ⁷³⁻²⁵ 坩 ⁸¹⁻⁵⁶ 坩 ⁴³⁻⁰⁸ 坩 ⁷⁵⁻²⁸	大	奘 ⁶⁴⁻¹¹ 奘 ⁷⁵⁻³⁶
女	姁 ⁸⁹⁻⁸³ 媿 ⁶⁷⁻²⁴ 媿 ⁷⁴⁻¹⁷ 媿 ⁹¹⁻⁹¹	子	孳 ⁶⁹⁻⁷⁷
宀	宀 ⁶⁰⁻⁵³ 宀 ⁹²⁻⁸³ 宀 ⁶⁴⁻⁸⁴ 宀 ⁸³⁻⁸² 宀 ⁵⁰⁻²⁸ 宀 ⁸¹⁻⁵⁷	山	岑 ⁵⁴⁻²⁵
山	岾 ⁶⁵⁻⁹² 岾 ⁹¹⁻³² 岾 ⁹⁰⁻⁵⁴ 岾 ⁵³⁻⁴⁷ 岾 ⁶⁶⁻⁹⁴ 岾 ⁴⁶⁻⁶⁸ 岾 ⁶⁹⁻⁸⁵ 岾 ⁷¹⁻¹⁴	彔	彔 ⁷²⁻²⁷
巾	幪 ⁹³⁻¹²	廌	廌 ⁷⁴⁻⁷⁶ 廌 ⁵⁹⁻³⁴
心	愜 ⁴⁶⁻³¹ 愜 ⁵¹⁻⁵² 愜 ⁸⁰⁻⁸⁵ 愜 ⁹²⁻⁴² 愜 ⁴³⁻³⁶ 愜 ⁸⁴⁻¹⁰ 愜 ⁹³⁻⁸³	手	擎 ⁴⁴⁻⁵⁴ 擎 ⁵³⁻²⁶ 拖 ⁸⁶⁻⁷³ 掙 ⁶³⁻⁸² 掙 ⁷⁰⁻⁴¹ 掙 ⁸²⁻⁴² 掙 ⁸⁵⁻³⁶ 掙 ⁵⁴⁻⁴⁷
扌	攄 ⁸⁷⁻⁴⁴ 攄 ⁸⁷⁻⁴³	女	敷 ⁶⁹⁻³⁴ 敷 ⁶⁴⁻³² 敷 ⁹²⁻⁹⁰
日	曹 ⁸⁰⁻³⁸ 昃 ⁴⁴⁻⁶¹ 昃 ⁶⁰⁻²² 昃 ⁷³⁻⁸⁰ 昃 ⁸⁹⁻³² 昃 ⁴⁷⁻⁹² 昃 ⁷⁹⁻⁶¹ 昃 ⁴⁵⁻²⁸ 昃 ⁷³⁻⁵⁷		
目	眈 ⁴⁶⁻⁴¹ 眈 ⁴⁸⁻⁵² 眈 ⁵⁸⁻³⁶ 眈 ⁵⁹⁻¹⁸ 眈 ⁶¹⁻⁴³ 眈 ⁷¹⁻⁹⁰ 眈 ⁹³⁻⁵⁸ 眈 ⁵²⁻³⁷ 眈 ⁶⁰⁻²¹ 眈 ⁸¹⁻⁵⁹ 眈 ⁹⁰⁻⁵⁷ 眈 ⁹²⁻²³ 眈 ⁸⁶⁻⁸⁸ 眈 ⁷⁴⁻⁴⁹ 眈 ⁷⁷⁻⁷¹ 眈 ⁹¹⁻²¹ 眈 ⁴⁴⁻⁵⁷		
木	柰 ⁵⁰⁻¹⁶ 桼 ⁸⁶⁻⁵¹ 桼 ⁴⁵⁻⁰¹ 柰 ⁵⁰⁻⁷⁸ 柰 ⁵²⁻⁶⁶ 柰 ⁴⁹⁻⁸⁷ 柰 ⁸⁸⁻³⁴ 柰 ⁶²⁻⁵⁸ 柰 ⁶⁶⁻⁶⁷ 柰 ⁷²⁻⁴¹ 柰 ⁷³⁻⁸¹ 柰 ⁶⁰⁻²⁴ 柰 ⁵⁷⁻⁰⁷ 柰 ⁴³⁻⁸¹ 柰 ⁵⁸⁻¹⁹ 柰 ⁷⁹⁻⁶³ 柰 ⁸²⁻⁴⁵	歹	殮 ⁵⁴⁻¹⁷
水	汶 ⁵⁸⁻⁰¹ 汶 ⁵⁸⁻⁰⁹ 沆 ⁷⁰⁻⁴⁴ 沆 ⁷¹⁻⁵⁵ 沆 ⁷⁴⁻²²		

3-3 異体字

異体字の問題は大変に複雑な問題で、その歴史も非常に古く、漢の許慎までさかのぼる。すなわち、後漢の和帝に献上した許慎の「説文解字」にもすでに1,160字の異体字があったといわれている。本来、異体字の問題は本稿の目的ではなかった。しかし、いくつかの問題があるので簡単に触れておく。

第一は、略字化による字意の変質である。これは昔から存在しているにもかかわらず、その使用が少ないために、今は別字の略字で使われている字があった。つまり「缶」(KSC 61-14)、「予」(69-88)、「芸」(73-93)、「台」(87-27)、「芦」(91-56)、「欠」(93-66)の6字は、もともとから固有の意味と音を持っているものであるが、日本工業規格には、各々、「罐」(JIS 70-05)、「豫」(48-14)、「藝」(73-26)、「臺」(71-42)、「缺」(69-94)、「蘆」(73-15)の略字として使っていて、そのものの本来の意味と音が失われてしまった。(これらは日本工業規格と韓国標準規格ともに出ている字だけである。他にもあるがここでは除外した)。しかし、韓国標準規格には音も意も全然違う別字の本字の意味で使われている。そして、もうひとつ「円」(KSC 69-87)は、韓国でも「圓」(74-13)の略字として使うこともあるが、韓国標準規格では日本貨幣単位の「円」(エン)の意味で使われている。

第二は、似ている字との間の意味の混乱である。例として「唇」(JIS 31-16)と「脣」(70-92)、「諡」(75-74)と「諡」(75-75)とはまったく音も意も違う異字である。ところで、これらの字が似ている関係から俗に混用される場合がある。もちろん、こういう現象は韓国でも同じである。こういう現象が、日本工業規格では認定されるらしい異体字として取り扱っている。しかし、韓国標準規格には本字の意味そのまま別字として扱われている。

そして、第三に、ただ一字「全」(KSC 52-46)は、日本工業規格では漢字ではなく記号として扱って、記述記号(JIS 1-24)の同上記号として使っているが、韓国標準規格では漢字として使っているという点などを明らかにしておく。

最後に異体字の件数について触れてみよう。実際に調べると、思ったよりその件数が多いのに驚いた。日本工業規格では、異体字であるとして参照または印「←」で表示されている異体字だけでその数は555種、1,159字に達し、その他にも通字を除外した「全」字、「俗」字だけでも約70種、150字があって(もちろんこれらは字典に「全」字または「俗」字で用いるものである)これを足すと、その総数はおよそ約600余種、1,300余字におよび、全体の約1/5を占めていて、その数の多さを推量するのは難しくない。これに比べ韓国標準規格の状況はどうか、もちろん韓国標準規格にも異体字はある。しかし、それは日常的なものと人名に使っているものばかりで、日本工業規格に比べるとその数ははるかに少ない。すべてを足すと約88種、168にすぎない。これをすべて表で整理すれば次のとおりである(表9「日本工業規格の異体字」；表10「日本工業規格に異体字として表記されていない異体字」；表11「韓国標準規格の異体字」)。

表9 日本工業規格の異体字

2字異体字(517種 1,034字)											
亜亞 16 48 01 19	惡惡 16 56 13 08	芦蘆 16 73 18 35	鯪鯪 16 82 19 45	圧壓 16 52 21 58	庵菴 16 72 35 31	圉圍 16 52 47 03	為爲 16 64 57 10	医醫 16 78 69 48	育毓 16 61 73 58	一弌 16 48 76 01	壹壹 16 52 77 69
稻稻 16 67 80 43	飲飲 16 61 91 27	淫姪 16 53 92 21	隱隱 17 80 03 12	卯卯 17 50 12 41	鬱鬱 17 61 21 21	叡睿 17 66 35 47	營營 17 51 36 59	曳曳 17 59 40 10	榮榮 17 60 41 38	穎穎 17 17 47 48	衛衛 17 74 50 44
詠咏 17 50 51 73	馱驛 17 81 56 67	円圓 17 52 63 04	煙烟 17 63 76 61	艷艷 17 71 80 70	塩鹽 17 83 86 37	奧奧 17 52 92 92	往往 17 55 93 40	応應 17 56 94 70	欧歐 18 61 04 31	毆毆 18 61 05 56	鶯鶯 18 82 09 84
冲冲 18 49 13 53	穩穩 18 67 26 51	仮假 18 48 30 81	価價 18 49 33 11	嘩譁 18 75 62 86	峨峩 18 54 69 22	画畫 18 65 72 33	会會 18 48 81 82	解解 18 75 82 27	回回 18 49 83 37	壊壊 18 52 85 53	怪恠 18 55 88 63
懷懷 18 56 91 71	繪繪 19 69 08 73	蟹蟹 19 74 10 23	碍礙 19 67 23 08	蛎蠣 19 73 34 58	鈎鈎 19 78 35 76	扞擴 19 58 40 18	攪攪 19 57 41 88	殼殼 19 61 44 55	覚覺 19 75 48 20	岳嶽 19 54 57 54	樂樂 19 60 58 59
竈竈 19 67 86 62	釜釜 19 78 88 61	粥鬻 20 68 01 88	刈苻 20 20 02 03	勸勸 20 50 11 16	卷卷 20 50 12 43	歡歡 20 61 31 36	漙灌 20 62 35 85	缶罐 20 70 44 05	觀觀 20 75 49 23	諫諫 20 75 50 61	鑑鑒 20 79 53 40
関關 20 79 56 80	陥陥 20 79 57 92	館館 20 20 59 60	巖巖 20 54 64 62	顔顔 20 80 73 90	器器 20 51 79 58	奇奇 20 52 81 84	期朞 20 59 92 14	棋碁 20 59 93 87	棄弃 20 55 94 17	気氣 21 61 04 70	亀龜 21 83 21 93
偽偽 21 49 22 06	戲戲 21 57 26 06	犧犧 21 64 30 26	却卻 21 50 49 42	糾糺 21 68 74 93	旧舊 21 71 76 49	扞據 21 58 82 01	京京 21 48 94 23	峽峽 22 54 14 23	挾挾 22 57 20 49	況況 22 49 23 55	狹狹 22 64 25 37
堯堯 22 84 38 01	曉曉 22 58 39 92	区區 22 50 72 31	勲勳 23 50 14 14	群羣 23 70 18 26	恵恵 23 56 35 10	憩憩 23 56 38 60	携攜 23 58 40 24	経経 23 69 48 20	繼繼 23 69 49 75	莖莖 23 72 52 19	螢螢 23 74 54 05
輕輕 23 77 58 43	頸頸 23 80 59 84	鷄鷄 23 83 60 17	芸藝 23 73 61 26	欠缺 23 69 71 94	決決 23 49 72 51	儉儉 23 49 80 13	圈圈 23 52 87 01	榷檢 24 60 01 93	檢權 24 60 02 62	猷獻 24 64 05 59	県縣 24 69 09 49
險險 24 80 17 10	頭顯 24 80 18 93	驗驗 24 81 19 68	巖巖 24 51 23 78	虎帑 24 73 55 41	鼓鼓 24 83 61 73	効效 24 58 90 35	広廣 25 55 13 02	恒恆 25 55 17 81	晃晃 25 58 24 72	稿稟 25 67 38 44	劫劫 25 49 69 71
号號 25 73 70 43	轟聾 25 77 76 36	国國 25 52 81 02	嗟嗟 26 54 23 45	濟濟 26 63 49 27	碎碎 26 66 53 76	斎齋 26 67 56 23	劑劑 26 49 62 93	財賤 26 76 66 35	坂阪 26 26 68 69	桜櫻 26 61 89 15	冊冊 26 49 93 38
雜雜 27 80 08 24	臯臯 27 66 09 08	参參 27 50 18 52	慘慘 27 56 20 46	棧棧 27 60 23 02	蚕蠶 27 74 29 36	讚讚 27 76 30 13	贊贊 27 76 31 53	殘殘 27 61 36 44	糸絲 27 69 69 15	紙帑 27 54 70 67	齒齒 27 83 85 79

表9 日本工業規格の異体字(続き)

事 事 27 48 86 15	兒 兒 27 49 89 27	爾 尔 28 53 04 85	辞 辭 28 77 13 70	湿 濕 28 63 30 28	質 質 28 76 33 36	実 實 28 53 34 73	篠 筱 28 68 36 12	舍 舍 28 71 43 50	积 釋 28 78 65 57	呪 咒 28 50 86 80	寿 壽 28 52 87 72
収 收 28 58 93 32	讐 讎 29 76 18 08	従 從 29 55 30 47	獸 獸 29 64 35 57	縦 縦 29 69 36 52	肅 肅 29 70 45 73	準 準 29 50 64 37	処 處 29 49 72 61	嘗 嘗 30 65 08 19	将 將 30 53 13 82	廠 廠 30 50 19 50	松 恣 30 59 30 32
烧 燒 30 63 38 86	称 稱 30 67 46 42	証 證 30 75 58 90	乘 乘 30 48 72 11	剩 剩 30 49 74 84	場 場 30 52 76 39	壤 壤 30 52 77 65	嬢 嬢 30 53 78 48	条 條 30 59 82 74	浄 淨 30 62 84 38	穰 穰 30 67 87 53	讓 讓 30 76 89 10
釀 釀 30 78 90 54	囑 囑 30 51 92 86	触 觸 31 75 08 29	唇 脣 31 70 16 92	寢 寢 31 53 18 74	慎 慎 31 56 21 38	晋 晉 31 58 24 73	真 眞 31 66 31 35	刃 刃 31 49 47 67	尽 盡 31 66 52 24	酢 醋 31 78 61 44	凶 圖 31 52 62 06
厨 廚 31 55 63 04	垂 垂 31 52 66 17	粹 粹 31 68 72 79	翠 翠 31 70 73 35	醉 醉 31 78 76 45	随 隨 31 78 79 14	数 數 31 58 84 43	枢 樞 31 60 85 68	世 世 32 50 04 34	畝 畝 32 65 06 28	凄 淒 32 62 08 39	棲 栖 32 32 19 20
声 聲 32 70 28 65	静 靜 32 80 37 48	齐 齊 32 83 38 78	摄 攝 32 57 61 80	窃 竊 32 67 64 70	專 專 32 53 76 83	戰 戰 32 57 79 05	浅 淺 32 62 85 41	線 綫 32 69 94 32	船 船 33 71 05 53	賤 賤 33 76 08 45	踐 踐 33 76 09 88
錢 錢 33 79 12 02	善 譚 33 70 17 33	禅 禪 33 67 21 24	曾 曾 33 33 29 30	疎 疎 33 76 34 83	蘇 蘓 33 73 41 31	溯 溯 33 62 44 74	鼠 鼠 33 83 45 75	双 雙 33 50 48 54	壮 壯 33 52 52 67	搜 搜 33 57 60 51	插 插 33 57 62 71
争 爭 33 64 72 07	窓 窗 33 67 75 57	総 總 33 69 77 33	聡 聰 33 70 79 66	莊 莊 33 72 81 23	装 裝 33 74 85 70	走 走 33 76 86 65	騷 騷 33 81 91 59	臟 臟 34 71 01 39	蔵 藏 34 73 02 22	属 屬 34 54 16 04	続 續 34 69 19 84
卒 卒 34 50 20 32	多 彡 34 52 31 76	墮 墮 34 52 36 56	舵 舵 34 34 40 41	楫 楫 34 60 42 83	对 對 34 53 48 84	带 帶 34 54 51 72	滞 滯 34 62 58 92	台 臺 34 71 70 42	滝 瀧 34 34 76 77	択 擇 34 58 82 04	沢 澤 34 63 84 23
鐸 鉦 34 78 88 69	達 達 35 77 03 93	豎 豎 35 76 08 19	狸 狸 35 76 12 30	单 單 35 51 17 37	担 擔 35 57 20 31	胆 膽 35 71 32 28	团 團 35 52 36 05	弹 彈 35 55 38 28	断 斷 35 58 39 50	耻 耻 35 70 49 55	痴 癡 35 65 52 87
稚 穉 35 67 53 48	遲 遲 35 78 57 15	昼 晝 35 58 75 76	虫 蟲 35 74 78 21	鑄 鑄 35 79 82 41	瀦 瀦 35 63 85 44	猪 豬 35 76 86 23	町 埦 36 65 14 22	聽 聽 36 70 16 69	腸 腸 36 71 18 22	勅 敕 36 58 28 37	珍 珍 36 64 33 63
鎮 鎮 36 79 35 15	壺 壺 36 52 59 68	遞 遞 36 78 94 10	纏 纏 37 69 27 85	轉 轉 37 77 30 59	点 點 37 83 32 58	伝 傳 37 49 33 03	砺 礪 37 66 55 74	党 黨 37 83 62 62	梲 檣 37 59 78 77	盗 盜 37 61 80 25	涛 濤 37 62 83 25
灯 燈 37 37 84 85	当 當 37 65 86 36	逃 逃 38 77 08 77	鬪 鬪 38 82 14 12	働 働 38 48 15 30	独 獨 38 64 40 55	読 讀 38 76 41 06	朽 朽 38 59 42 29	届 届 38 53 47 92	繩 繩 38 69 76 74	二 式 38 48 83 17	迹 邇 38 77 86 78
韭 韭 39 80 03 76	妊 妊 39 53 05 12	襦 衤 39 39 09 10	迺 迺 39 77 22 82	埜 野 39 44 24 78	惱 惱 39 56 26 29	腦 腦 39 71 30 10	霸 霸 39 59 38 17	廢 廢 39 55 49 06	拜 拜 39 57 50 33	杯 盃 39 39 53 54	梅 棊 39 39 63 64

表9 日本工業規格の異体字(続き)

売賣 39 76 68 46	蠅蠅 39 74 72 04	麦麥 39 83 94 46	函函 40 49 01 66	発發 40 66 15 04	髮髮 40 81 17 91	罰罰 40 70 19 15	拔拔 40 57 20 22	蛮蠻 40 74 58 39	秘祕 40 67 75 16	桧檜 41 59 16 56	氷冰 41 49 25 54
浜濱 41 63 45 32	富富 41 41 57 58	扌拂 42 57 07 36	仏佛 42 48 09 39	幣幣 42 54 30 82	並竝 42 67 34 77	閉閉 42 79 36 58	篋篋 42 68 47 36	變變 42 58 49 46	舗舗 42 71 62 52	穂穂 42 67 70 47	峰峯 42 42 86 87
萌萌 43 72 08 46	褒褒 43 74 11 81	豊豊 43 76 13 20	冒冒 43 70 33 78	貌兒 43 66 38 06	没沒 43 61 55 83	本本 43 52 60 81	翻翻 43 70 61 44	楨楨 43 84 74 02	柀柀 43 59 81 38	俛儘 43 48 89 54	万萬 43 72 92 63
満満 43 62 94 64	脈脉 44 70 14 87	夢夢 44 52 20 77	麵麩 44 83 45 49	黙黙 44 64 59 52	餅餅 44 81 63 22	弥彌 44 55 79 29	薬薬 44 73 84 27	訳譯 44 76 85 03	癒癒 44 65 94 72	湧涌 45 45 15 16	予豫 45 48 29 14
余餘 45 81 30 17	与與 45 71 31 48	誉譽 45 76 32 05	揺揺 45 57 41 74	様様 45 60 45 75	窯窯 45 67 50 63	耀耀 45 64 52 02	謡謡 45 75 56 79	踊踊 45 76 57 93	遙遙 45 84 58 03	来來 45 48 72 52	乱亂 45 48 80 12
覧覽 45 75 87 21	略畧 46 65 12 32	溜溜 46 63 15 17	留留 46 65 17 23	竜龍 46 46 21 22	両兩 46 49 30 32	涼涼 46 49 35 58	獵獵 46 64 36 58	瞭瞭 46 59 38 02	糧糧 46 68 40 78	隣鄰 46 78 57 35	瑠璃 46 64 60 69
壘壘 46 52 61 62	涙泪 46 62 62 05	励勵 46 50 69 15	礼禮 46 67 73 25	隸隸 46 80 76 17	靈靈 46 80 78 45	齡齡 46 83 80 84	恋戀 46 56 88 88	聯聯 46 70 94 63	炉爐 47 64 07 04	勞勞 47 50 11 09	朗娘 47 59 15 13
楼樓 47 60 16 76	籠籠 47 68 22 38	禄禄 47 67 29 19	互互 47 47 42 43	湾灣 47 63 49 52	仞仞 48 48 32 33	俎俎 48 64 57 12	倅倅 48 48 70 71	僭僭 49 49 08 09	冤冤 49 53 45 67	寇寇 49 53 46 68	冪冪 49 70 49 18
冽冽 49 62 56 16	涸涸 49 62 57 33	剋剋 49 53 78 81	劊劊 49 68 85 20	匱奩 50 52 29 94	卮卮 50 54 40 66	厠厠 50 54 46 90	厦廈 50 54 47 92	廝廝 50 55 49 03	喀喀 51 51 29 30	噪噪 51 76 68 01	嚏嚏 51 51 73 74
嚙齧 51 83 87 86	址址 52 79 14 87	埒埒 52 52 31 32	壘壘 52 70 64 04	奘奘 52 55 89 18	佞佞 53 53 04 05	嫵嫵 53 53 38 39	嫵嫵 53 53 43 44	崙崙 54 54 38 39	嵌嵌 54 68 40 29	迪迪 55 77 15 76	彝彝 55 55 19 20
彎弯 55 55 30 31	彿髴 55 81 42 92	悴悴 56 56 12 13	慙慙 56 56 47 48	饑饑 56 56 82 83	戛戛 56 57 94 01	扌扌 57 57 13 14	擡抬 58 58 12 13	旛旛 58 58 57 58	曠昞 59 59 05 06	棕椶 60 60 03 04	櫛櫛 60 60 73 74
槩槩 61 61 01 02	櫟櫟 61 61 11 12	殲殲 61 61 51 52	炯炯 63 63 55 56	熙熙 63 63 70 71	熏燻 63 63 77 78	犁犁 64 64 20 21	豺豺 64 76 28 25	貉貉 64 76 34 27	猊貌 64 76 41 31	獾獾 64 76 51 34	琅琊 64 64 70 71
璫璫 64 66 79 85	瑶瑶 64 84 86 04	璵璵 64 64 93 94	疇疇 65 65 38 39	疣疣 65 70 47 79	鞞輝 66 66 16 17	盪盪 66 66 27 28	眦眦 66 66 36 37	稟稟 67 67 40 41	筐筐 67 68 94 01	筍笋 68 68 03 04	箏箏 68 68 23 24
籐籐 68 68 59 60	籤籤 68 68 62 63	濟濟 69 69 77 78	羹羹 70 70 29 30	脰脰 71 71 20 21	臘臘 71 71 36 37	臚舩 71 71 67 68	苒苒 71 71 85 86	萼萼 72 72 53 54	虱蝨 73 73 45 92	蟊蝥 74 74 17 18	蟒蟒 74 74 28 29

表9 日本工業規格の異体字(続き)

蠹 蠹 74 74 37 38	衄 衄 74 74 40 41	衽 衽 74 74 51 52	禪 禪 74 74 91 92	諛 諛 75 75 74 75	譖 譖 75 75 91 92	貌 貌 76 76 32 33	贓 贓 76 76 59 60	躑 躑 77 77 24 25	輒 輒 77 77 44 45	輻 輻 77 77 49 50	迺 迺 78 78 04 05
鑪 鈺 79 79 46 47	鑽 鑽 79 79 51 52	鬧 鬧 79 82 63 09	闊 潤 79 79 72 73	齋 齋 80 80 77 78	飄 颯 81 81 08 09	鯁 鯁 82 82 59 60	鳧 鳧 82 82 74 75	鴟 鴟 82 82 86 87	鵝 鵞 83 83 01 02	鶉 鶉 83 83 09 10	鷓 鷓 83 83 22 23
麩 麩 83 83 47 48											
3 字異体字 (30 種 90 字)											
穉 秋 穉 16 29 67 12 09 52	厩 廐 廐 17 54 54 25 93 94	蓋 盖 蓋 19 66 72 24 21 68	学 學 孛 19 53 53 56 60 61	雁 鴈 鴈 20 82 82 71 78 79	帰 歸 皈 21 61 66 02 37 07	拳 舉 舉 21 58 58 83 09 10	馭 駟 驅 22 22 81 78 79 60	径 徑 逕 23 55 77 34 45 84			
溪 溪 谿 23 62 76 44 68 16	冴 洳 洳 26 49 61 67 52 76	蕊 藥 蕊 28 73 73 41 02 03	写 寫 寫 28 49 53 44 48 77	洩 泄 澀 29 63 63 34 07 08	叙 敘 敘 29 58 58 86 38 39	獎 獎 獎 30 52 64 09 93 50	靱 鞞 靱 31 80 80 57 54 55	髓 髓 髓 31 71 81 81 27 82			
潜 潛 潛 32 63 63 88 10 11	織 織 織 33 69 69 01 89 90	庁 廳 廳 36 55 53 03 12 13	島 鳴 寫 37 37 54 71 72 26	式 貳 貳 38 76 76 85 40 41	淵 渊 渊 42 62 62 05 28 29	辺 邊 邊 42 78 78 53 20 21	宝 寶 寶 42 53 53 85 79 80	蓑 蓑 蓑 44 68 68 12 34 35			
婿 婿 聿 44 52 70 27 70 61	藪 藪 藪 44 68 73 89 56 14	羈 羈 羈 70 70 75 19 20 11									
4 字異体字・外											
鉦 砮 礦 鑛 25 25 66 79 59 60 72 42	崎 崎 碕 寄 26 26 26 54 74 75 76 31	疊 疊 疊 疊 30 65 65 65 86 40 41 42	体 躰 體 體 34 77 77 81 46 28 29 83	鉄 鉄 鐵 鐵 37 78 79 79 20 78 36 37	兔 菟 兔 菟 37 37 49 72 38 49 29 25	6 種 24 字					
弁 辨 辨 瓣 辯 42 49 50 65 77 59 94 01 02 71	* 韓国では 弁 が 瓣 の略字では使われない(1 種 5 字)										
劍 劍 劔 劔 劔 23 49 49 49 78 85 88 89 90 63	1 種 6 字										

総 555 種 1159 字

表10 日本工業規格で異体字として表示されていない異体字 (総70種145字)

2字異体字 (64種128字)												
井井 16 48 70 07	韻韻 17 80 04 81	映映 17 58 39 85	園園 17 17 64 82	歌詞 18 75 46 72	蝦蝦 18 82 60 49	我儂 19 49 70 15	嚙咬 19 50 90 91	喜僖 20 49 78 05	競競 22 49 05 31	斤鉞 22 78 52 67	傑杰 23 59 70 31	
倦券 23 50 81 05	耕畊 25 65 44 25	考攷 25 58 45 31	左𠂇 26 54 24 05	私亼 27 50 68 51	宐肉 28 38 21 89	繡綉 29 69 11 21	庄莊 30 72 17 23	飾飭 30 81 94 19	川𠂇 32 54 78 63	創𠂇 33 49 47 76	惣惣 33 56 58 27	
村邨 34 78 28 23	啄啗 34 51 79 22	弔吊 36 36 04 63	朝晷 36 58 11 74	德愆 38 55 33 60	附埒 41 52 77 20	鋪鋪 42 71 63 52	綿緜 44 69 42 36	茂楸 44 60 48 30	網罔 44 70 54 06	陽曷 45 58 59 70	裏裡 46 46 02 03	
櫓艦 47 71 06 66	从從 48 55 26 47	來徠 48 55 52 50	倣效 48 58 91 35	剿勦 49 50 86 11	匆忽 50 55 18 68	啣銜 51 78 18 82	瀏瀏 51 63 72 40	囹國 51 52 91 02	墻牆 52 64 54 15	忤牯 55 64 56 19	惧懼 55 56 92 86	
懼歡 56 61 85 36	拏拿 57 57 28 29	支攴 58 58 29 30	杆桿 59 59 24 69	椶檇 59 60 86 58	椹椹 60 79 27 12	洩涕 62 62 06 24	繼緹 69 69 08 18	繞遶 69 78 69 13	腮頤 71 80 08 91	膾鱠 71 82 26 70	訛譌 75 75 34 87	
锈縮 78 79 88 21	銚鏢 78 79 91 28	閔閔 79 82 66 10	鰲鰲 82 83 66 71									
3字異体字 (5種15字)												
岡垠崗 18 52 54 12 46 30	界堺畧 19 26 65 06 70 24	抵抵觝 36 64 75 81 18 26	鉞矛梓 43 44 59 40 23 66	無无毋 44 58 61 21 59 57								

表11 韓国標準規格の異字体

杆桿 42 42 50 52	潤礪 42 42 53 56	鑑鑿 42 42 92 93	甲鉀 43 43 03 05	岡崗 43 43 10 11	亶疆 43 43 17 18	盖蓋 43 43 44 47	杰傑 43 43 89 90	劍劔 43 43 92 93	堺界 44 45 87 03	攷考 45 45 25 37	埼崎 48 48 82 88
朞期 49 49 01 02	拏拿 49 49 57 58	年季 50 50 20 22	寧寧 50 50 27 28	德愆 51 51 76 77	島嶋 51 51 86 87	仝同 52 52 46 50	琅瑯 53 53 42 43	來徠 53 53 46 48	梁樑 53 53 57 58	粮糧 53 53 59 61	冽冽 54 54 09 12
裏裡 55 55 74 75	万萬 56 56 18 31	楸茂 57 57 74 82	玟珉 58 58 39 40	杯盃 59 59 42 45	槩藁 59 60 92 02	鰲鰲 60 60 14 15	峯峰 60 60 71 72	烽烽 60 60 75 76	浜濱 62 62 19 20	殺煞 63 63 15 16	膳饌 64 64 46 55
昭炤 65 65 25 31	溯溯 65 65 29 47	岳嶽 68 68 31 32	庵菴 68 68 61 64	揚敷 69 69 32 34	烟煙 70 70 51 53	映映 71 71 17 18	叡睿 71 71 53 58	鰲鰲 72 72 07 08	涌湧 73 73 31 32	妊妊 76 76 84 85	墻牆 77 77 63 77

庄 莊	吊 弔	点 點	晁 朝	晉 晋	窠 采	哲 喆	酢 醋	村 邨	摠 總	痴 癡	稚 穉
77 77 68 86	78 80 52 32	79 79 35 39	80 80 37 40	82 82 43 44	83 83 82 90	84 84 41 42	85 85 18 19	85 85 29 30	85 85 36 37	86 86 33 34	86 86 35 36
柒 漆	坂 阪	厦 廈	夏 晷	蝦 鰕	啣 銜	炯 炯	浩 濤	確 碯	晃 眇	勛 勳	埧 壙
86 86 51 52	87 88 88 01	89 89 29 31	89 89 30 32	89 89 36 40	89 89 65 71	91 44 06 62	91 91 39 43	92 92 12 13	92 92 44 45	93 93 18 19	93 93 20 21
熏 燻	僖 喜	愜 愜	熹 熿	丘 坵	坨 臺	寶 瑤	79 種 148 字				
93 93 23 24	93 93 74 76	93 93 82 83	93 93 88 89	46 47 88 06	51 51 61 70	60 60 36 42					
徑 徑 逕	溪 磧 谿	无 母 無	晒 曷 炳	弁 辨 辯	棲 栖 棲	6 種 18 字					
44 44 44 41 51 79	45 45 45 02 05 12	57 57 77 73 76 57	60 60 60 21 22 25	60 60 60 07 09 10	63 63 63 82 87 88						
日本工業規格では異体字として扱っているが韓国標準規格では別字として扱った字											
缺 欠	罐 缶	臺 台	蘆 芦	余 餘	豫 予	藝 芸	円 圓	脣 唇	諡 諡	10 種 20 字	
44 93 32 66	46 61 24 14	51 87 70 27	54 91 54 56	69 70 89 14	71 69 67 88	71 73 61 93	69 74 87 13	66 82 78 38	67 76 45 48		

4 結 論

前節までに韓国標準規格と日本工業規格の漢字について比較してみた。その概要を要約すれば次のとおりである。

第一に、韓国標準規格の基本漢字の中には、発音を異にすることから生ずる重出字が多い。すなわち2回以上重ねて出ている字が字種として262字、字数としては530字にいたっている。これは日本工業規格が字形を重視しているのに対して、韓国標準規格ではその音が重視されていることを意味する。これらの重出字の類型には次の3種類がある。

- 1) 同一字の音の場合によって異なる字：習慣による場合（4字）と韓国語の音韻法則による場合（201字）、その他（3字）がある。
- 2) 異意異音同体字：字体は同じであるが、その音と意味が違うので重出されている字で43字がある。
- 3) 同意異音同体字：字体も字意も同じであるが、場合によって二つの音で読まれている場合で13字がある。

第二は、韓国字の問題である。漢字を使っている国々が持っている固有の漢字の問題で、ここでは純粹韓国字と半漢半韓字に分けてみるができる。韓国には500字以上の固有の漢字があることが知られているが、韓国標準規格に納められている韓国字の数は多くなく、半漢半韓字を足してもわずか24字にすぎない。そして、韓・日・中の字典に出ていない6字すべてが韓国字であるとしても、そ

の数は 30 字だけである。

終りに、この結果を日本工業規格の漢字と対比してみると次のような特徴がある。

まず第一番目に、体裁面からみると、両者とも 2 バイト完成型になっている点と同じである。しかし、収録字の区分と配列は確然に違い、韓国標準規格がハングルの「가나다」順で配列しているのに比べ、日本工業規格は第 1 水準、第 2 水準に分けて、第 1 水準は仮名の五十音順、第 2 水準は康熙字典の部首順で配列している。

二番目は、字数面からみると、単純比較では、1,465 字の差異があるが、韓国標準規格には重出字、日本工業規格には異体字が多いので、単純な数値の比較は意味がないと思われる。これは体裁上の相違からくる結果で、例として韓国標準規格の基本漢字を日本工業規格の方式で配列すれば韓国標準規格の重出字はすべて要らなくなるからである。したがって、日本工業規格にはなく韓国標準規格にだけある字は、韓国字を含めても 426 字にすぎない。もちろん、両者の字の形が完全に一致するのではない。略字化などによって少しずつ違ったものはたくさんあるが、それはささやかなものと思われるので、同じ字として取り扱った。

三番目は、異体字の問題である。韓国標準規格には種数として約 80 種、字数として 160 余字にすぎない異体字が、日本工業規格には約 600 余種、1,300 余字に達する異体字がそれぞれ含まれている。そして、その数は全体の約 1/5 を占めている。その中には略字化によって字義が変わったものとして 6 字（「缶」、「予」、「芸」、「台」、「欠」、「芦」）があるが、韓国標準規格では昔のまま本字の意味で使われている。また、似ている字の混乱（「脣」「唇」、「謚」「諡」）なども指摘することができる。

あとがき

本稿を執筆するにあたり、富士大学の田辺広教授から多大な助言をいただいたことを付記して感謝の心をあげたい。そして、学術情報センターの将来の発展を祈る次第である。

注：本稿は 1989 年 7 月 18 日に学術情報センターにおいて開催された共同研究「日本における国際書誌調整」（研究代表者：学術情報センター研究開発部長 山田尚勇）による公開講演会「韓国・朝鮮語の資料と文字処理の現状－1」において、日本語によって発表された原稿をもとに加筆訂正したものである。

研 究 論 文

文献の論理構造を考慮した全文検索システム

A Full-text Retrieval System Reflecting the Logical Structure of Documents

学術情報センター 影浦 峡

学術情報センター 大山 敬三

学術情報センター 宮澤 彰

学術情報センター 根岸 正光

日立制作所 中央研究所 鳥居 俊一

日立制作所 システム開発研究所 絹川 博之

要 旨

全文データベースはデータベース・サービスの品揃えの一つとして発達してきたが、検索結果にノイズが多いとか利用が難しいといった問題点が指摘されている。学術情報センターでは、全文データベース・サービスを公開するとともに、こうした問題点を解決するような新しい全文検索システムについて研究を行ってきた。本稿ではこれに基づき、文献の論理構造を考慮した全文検索オペレーションを紹介する。また、こうした検索オペレーションを備えた全文検索システムの実現方式についても述べる。

ABSTRACT

Full-text databases have recently become popular as a type of database services. However, there are some problems in using these databases, e.g. they are difficult to use without producing great noises. NACSIS has been pursuing the research on a new full-text retrieval system to solve these problems, in parallel with servicing full-text databases. Based on the research pursued in NACSIS, this paper presents new full-text retrieval operations that reflect the logical structure of documents. System configuration of the new full-text retrieval system is also mentioned.

1. 始めに

全文データベースは「1つ以上の著作の本文(テキスト)全体をデータとするデータベース」¹⁾と定義されている。一般に全文データベースという場合、従来の書誌情報データベース(文献抄録型デー

データベース)の延長線上にあるものと捉えられる雑誌記事や論文などの全文データベースの他に、宗教学や文学の研究のための聖書や文学作品の機械可読全文が含まれる。本論文で想定しているのは前者のタイプの全文データベースであり、以下で全文データベースという場合には、聖書や文学作品の全文データベースは含まない。

全文データベースの提供は、まず判例やニュース記事のように比較的短い本文のものから始まったが、コンピュータの高速化と記憶装置の低廉化によって、雑誌論文などの全文を収録したデータベースも提供されるようになった。その後、機械可読形式で蓄えられた潜在的な全文データベース候補の増加とあいまって全文データベースの数は増加している。現在では Dialog や BRS といった商用のオンライン検索サービスを通してアクセスできる全文データベースはかなりの数にのぼり、学術情報センターでも、*Harvard Business Review* や化学系学会誌掲載論文の全文データベースを提供している。

全文データベースの検索サービスは、通常、本文を含むテキスト中の(機能語を除く)全単語が検索対象となっていること、オンラインで全文が入手可能であることを特徴とする。そして、面倒な統制語を気にすることなしに検索できる点や全文がすぐ入手できるといった点から、エンド・ユーザーの要求に答え得るものとして期待が集まっている。しかしながら現状のシステムを見る限りでは、利用者側で複雑な検索技法を用いないと不要な結果のみが検索されることになりがちで、むしろ欠点が目だつ場合が多い。

学術情報センターでは、現在の全文検索システムに見られる問題点を解決するために、全文という対象の捉えなおしとそれに基づく検索機能の拡張について検討してきた。本論文はそれについて述べるものである。以下、2章で現行の全文検索システムに見られる特徴を整理し、問題点と解決の方向とをまとめる。3章ではそれに対して学術情報センターで検討を進めてきた新しい検索方式を紹介する。次に4章で、3章で提案した検索機能を備えた検索システムの実現方式を紹介する。

2. 現行の全文検索システム

ここでは本論文で紹介する検索方式導入の背景を明確にするため、現行の全文検索システムの一般の特徴をまとめ、現行のシステムが持つ問題点を整理する。

2.1 一般的特徴²⁾

現在サービスされている多くの全文検索システムは、通常の本誌情報データベースとほぼ同じ構造を持っている。すなわち、データベースはテキスト・ファイルとインバーテッド・ファイルからなり、テキスト・ファイルは標題、著者名、抄録といったフィールドに分けられている。ここで本誌情報データベースと異なるのは、全文であるからパラグラフや見出し、表のタイトルなどの本誌情報データベースにはないフィールドが存在するという点だけである。インバーテッド・ファイルでは全文から切り出した語からなるインデックスとそれに対応するテキスト・ファイルの位置情報が管理されている。多くの場合、少数のストップワードを除くすべての単語がインデックスとして切り出される。

検索機能はシステムによって多少異なるが、それらをまとめると次のようになる。すなわち、検

索語の指定とフィールド指定を基本とし、検索語の指定の際に前方一致や後方一致、中間一致などの部分一致指定が許される場合がある。また、演算として AND、OR、NOT からなるブール演算のほか、近接演算が使えることが多い。近接演算としては、隣接指定や順序指定、語間距離指定のほかに同一センテンス、同一パラグラフといった指定がある。

全文検索における検索結果の表示は、基本的に従来の書誌情報検索システムと同様の機能を備えている。その他に、全文が表示対象となるため、表示機能に多少の工夫が凝らされている場合もある。最も一般的なのは、パラグラフ単位で、該当部分を、ヒットした検索語をハイライトさせて表示する機能である。また、ヒットした検索語とその前後一定数の語を KWIC 形式で表示する機能を持つ検索システムもある。さらに、検索された文献に対するブラウジング機能としてテキストの文字列検索や、フィールド単位またはパラグラフ単位のスキップ機能を備えたものもある。

2. 2 問題点

上述のような機能を持つ現行の全文検索システムを書誌情報データベースと比べた場合、一般に以下のような利点が指摘されている。

- ・全文がすぐに手に入る
- ・本文をブラウジングできるため、適合性の判断が容易にできる
- ・検索漏れが少なく、網羅的検索ができる
- ・固有名詞や実験手法などの特定のもしくは周辺的な情報が検索できる³⁾

ところが実際には、以下のような問題点も目立ち、全文検索システムに期待されている利点が発揮されているとは言い難い。

- ・表示出力が印刷体と比べて非常に劣るため、全文が入手できるメリットが少ない
- ・検索の手続きが難しくて使いこなせない
- ・大量の文献が検索されてしまいノイズばかりが増える

これらのうち、表示出力については出力用ハードウェアの問題を含む独立した問題領域であり、様々な点から改良が図られている。また、検索手続きの簡略化に関しては、自然言語処理手法の応用やグラフィカル・ユーザー・インタフェースの採用などが研究されている。

検索機能という点からも最も重要で本質的と考えられるのはノイズの増加の問題であり、これについてはいくつかの側面から改良の研究がなされている。第1は、検索方略の研究であり、これは現行の検索システムを最大限有効に利用する方法の研究である。第2に検索語彙の改良の研究がある。この研究では全文を利用した自由語検索と統制語の関係を問題としている。これらの研究は、現在ある全文データベースの構造を前提として検索上の問題を扱っている。

これに対して、全文という対象の捉え方から全文データベースを見直し、それに適した検索方式を考えようという立場もある。実際、現在の全文データベースサービスでは書誌情報データベースと同様に、全文を著者名や標題、パラグラフといった個々の独立したフィールドの集まりと考えているが、そうした前提が、全文が入力されていることの利点を引き出すことのできない重要な理由の1つであ

と思われる。したがって、全文をどのように捉えるのが適切かを考えた上で、検索オペレーションを再考していくことは適切でありかつ必要でもあろう。

3. 文献の構造を考慮した検索オペレーション

前章で現行の全文検索システムについてまとめ、問題点を整理した。本章では文献の論理構造と構造化された文献を対象とする検索オペレーションについて述べる。

3.1 文献の論理構造

文献を独立したフィールドの単なる集合と考えていることが現行の全文データベースの問題の1つであることは前述した。本来、1つ1つの文献は、様々なレベルの要素が複合的に結び付いたものとして構成されていると考えるのが自然である。ここで、文献がどのような要素からなりどのような構造を作り上げているかという点については様々な見方が存在する。

例えば言語学的観点から文献の構造を考えると、形態素を最小構成要素として語構成規則により単語という要素が構成され、単語というレベルの要素が文法規則に基づいて結び付くことによって句や節や文といった大きな単位を構成していると見ることができる。また、よりマクロなテキスト言語学的観点からは、例えば本文について、文間の相互関係や論述単位間の相互関係に基づく論述構造を認めることもできよう。それに対して、文献を紙の上に印刷されたもののイメージで見ると、いわゆる割付構造、すなわち字体の指定や字下げ、改行、改ページなどにより区別された要素から構成されるものと見なすことになる。

一方、SGMLやODAといった、文書ファイルの相互交換や再編集に関する標準化の流れでは、割付構造よりももう少し上の、論理構造と称されるレベルの構造を認めている。現在の機械可読文書の多くは写植機用データ・ファイルに基づくものであるため、印刷用の割付構造タグがつけられた形で流通している。これに対してSGMLやODAでの論理構造とは、標題や抄録、著者、所属、章見出しやパラグラフといった印刷上からも識別されるような論理的要素を認め、文献をそうした要素の相互関係からなっていると見なす場合に認められる文献の構造を指す。こうした論理構造は、機械可読文書に対して、受け手の側で容易に処理を加えられるようにするために導入されたものである。⁴⁾

本論文でいう文献の論理構造は、SGMLやODAで言われている論理構造に比較的近い。すなわち、意味的にまとまりを持った単位であって、かつ表示上一定のやり方で互いに区別されるような単位を論理要素とする。直感的には、標題や著者名、抄録、章と章見出し、パラグラフなどが論理要素に相当する。そして、1つ1つの文献においてこれらの論理要素の相互関係から構成される構造を論理構造とする。

論理構造をこのように考えると、それは文献を根とする木構造のイメージで表わされる。しかしながら、論理要素として表題や抄録、章、節、項といった単純なものと考えただけでも、その間に見られる関係は一意ではないし、認める構造が各人ごとに多少異なることもあろう。また、検索に関して言えば、構造の認め方に応じた様々な検索要求にできるだけ柔軟に対応できる必要もある。こうした

点を考慮し、文献の構造を次のような方法で表わすことにする。

- ・文献を構成する最も基本的な論理要素を定め、これをクラスと呼ぶ。各文献は、「文献」というクラスを根とし、適切なクラスをノードとする木構造で表わされる。
- ・各クラスに、そのクラスを特徴付ける性質を表わす様々なクラスカテゴリーを与える。クラスカテゴリーは、同様の性質を持つ異なるクラスをまとめるために用いられる。

典型的な雑誌論文に認められる論理構造を、クラスとその階層構造で表わすと、概略、図1のようになる。ここで、例えば「章」、「節」、「項」などのクラスに「見出し付き文書」というクラスカテゴリーを、また「標題」や「見出し」に「句」というクラスカテゴリーを与えることができよう。なお、当然のことながら、雑誌論文や会議録、単行書などといった文献の種類に応じて、または専門分野に応じて異なる論理構造が設定されることになる。

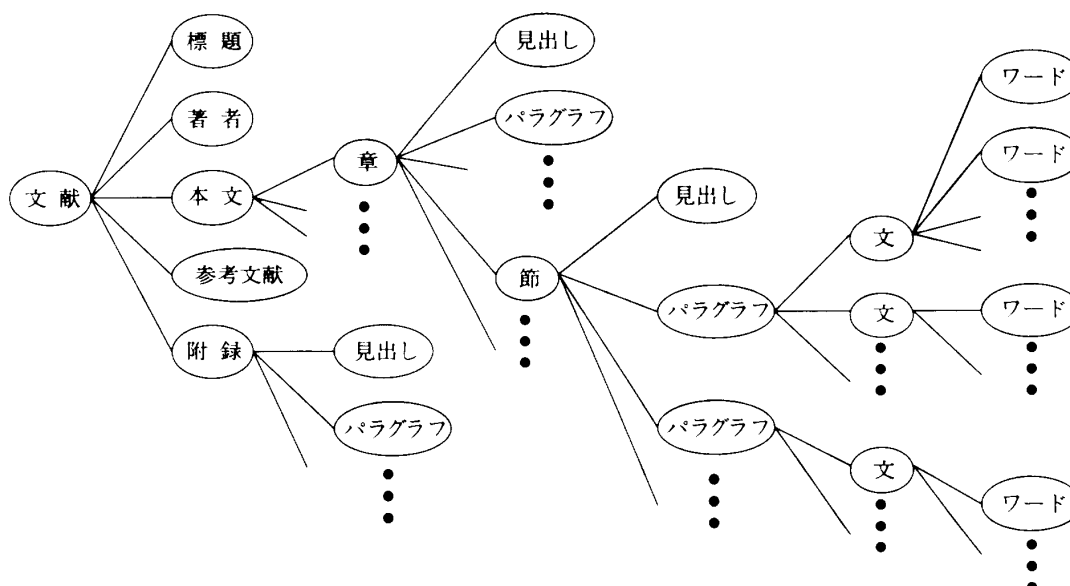


図1 文献の論理構造の例

3.2 検索オペレーション

前節で文献の論理構造について述べたが、ここではそれを考慮した検索オペレーションを提案する。全体の検索オペレーションは、基本検索、パスマッピング、集合演算という3つのオペレーションの組合せにより実現される。以下では、まず、基本検索、パスマッピング、集合演算それぞれの役割について述べ、次に、全体の検索オペレーションがどの様実現されるかを単純な例により示すことにする。

基本検索 基本検索は、検索語と検索対象とのマッチングを行う。論理構造を考慮した場合、基本検索の検索対象は木構造で表わされた文献の葉（図1ではワード）である。基本検索の結果として、検索語にマッチしたそれぞれの単位ワード（トークン）を要素とする集合が作られる。このとき、各ワードは文献の構造における根からのパス情報を持っており、パス中の各ノード（クラス）にはクラス

カテゴリーが付与されている。

パスマッピング 文献の論理構造に対応したオペレーションを行う。基本検索の結果できたワードからなる要素を適当なレベルのノードでくくって、それ自体集合であるような要素からなる集合を作る。またはパスマッピングの結果作られた集合の要素をあるノードのもとでくくったり分解したりして別の要素からなる集合を作る。またはノードに課された一定の条件により集合の要素を捨てる。

パスマッピングは次の2つの操作によって行われる。

- ・上下ノード指定 集合の要素が持つパスを指定されたノードのレベルでくくるまたは分解する。ここで、ノードをくくったときに、注目しているレベルでノードが同一である場合には集約操作がおこる。逆にノードを分解したときに指定されたレベルでノードが異なる場合には展開操作が行われる。根までさかのぼって適合しない場合その要素は捨てられる。

- ・要素選択 あるノード集合に対してさらに制約を加え、適合しない要素を捨てる。

パスマッピングにあたってノードを指定する方法は以下の4通りある。

- ・絶対パス指定 特定のノードを（クラスとその位置によって）指定する。
- ・相対パス指定 注目しているノードからの相対的位置を指定する。
- ・クラス指定 特定のクラスのノードを指定する。
- ・カテゴリー指定 特定のクラスカテゴリーを持つクラスのノードを指定する。

集合演算 パスマッピングによって定められた集合に対して和集合や積集合、補集合を取るといった集合演算を行う。

例えば、「標題中に a というワードがあり、ワード b と c が同一パラグラフ内にある文献を探せ」という要求は、以下のように処理されることになる。

基本検索：ワード a の集合 A と b の集合 B、c の集合 C が作られる。

パスマッピング：集合 B と C についてパラグラフのレベルに着目し、要素の集約を行う。このとき、一次検索の結果できた集合の要素（ワード）の中で、同一パラグラフにあるものは1つにまとめられる。

集合演算：その上で、パラグラフレベルでの共通部分からなる集合 D を作る。

パスマッピング：集合 A と D について文献のレベルに着目し、それぞれ要素の集約を行う。

集合演算：その上で A と D の共通部分を取り、検索結果とする。

この例は非常に単純なものであり、パスマッピングにおいてクラス指定による上下ノード指定しか行っていないが、より複雑な検索要求もこれに準じた方法で処理される。

以上紹介した検索オペレーションは、次のような利点を持っている。

- ・文献の構造を自然に反映した検索が可能で、ユーザーの多様な要求に柔軟に答えることができる。
- ・最終検索単位の指定や近接演算機能などをパスマッピング操作に還元するため、様々な検索要求を一貫したオペレーションの中で処理することが可能である。

ここでは検索オペレーションを主として手続き的な側面から紹介してきたが、今後、データモデル

の形式化及びその上での非手続き型言語を検討することにより、モデルと機能の完全性を考察していく必要がある。一方、検索の面からみると、検索機能が実際にうまく働くかどうかは、文献の論理構造をいかにうまく設定できるかに依存する面がある。前節で紹介した文献の構造は検索オペレーションの基本的な部分を示すために例としてあげたものであり、文献の論理構造についても今後検討していく必要がある。

4. 実現方式

学術情報センターでは、全文検索オペレーションの基本機能の検討と共に、その機能を備えた全文検索システムの全体構成と実現方式についても検討を進めてきた。本章では、そうした検索システムの実現方式に関する概略を述べることにする。

4. 1 全体構成

実現方式を考えるにあたっては、全文という大量のデータを扱わなくてはならないこと、論理構造を反映した複雑な検索に対応するユーザー・インタフェースを作らなくてはならないことという大きな問題がある。大量のデータを管理し処理するという点からはメインフレームが依然として有効であるから、全文データベースの構築と管理そして検索の基本処理を行うホスト的機能はメインフレームで担うのがよいであろう。

一方、前章で述べたような検索機能を考えた場合、ユーザーの検索要求は非常に多様なものとなり、そのモデル化は難しい。従って、検索のモデルをあらかじめ定めてそれを実現段階で反映させるのは困難である。むしろ、できるだけ検索の基本機能（すなわち基本検索とパスマッピング、集合演算）に即した低いレベルでのインタフェースを用意すればありうる利用に対応できるものと考えられる。そうしたインタフェースの一つとして、文献の構造と検索操作を直感的に理解できるようなグラフィカル・ユーザー・インタフェースを考えることができる。現在のところ、文献の論理構造を木のイメージで視覚的に表現したインタフェースを考えているが、その構築にはワークステーションが適している。

以上の点を勘案すると、データベースの管理と検索の基本部分をメインフレームで行い、検索インタフェースの構築をワークステーションで行うという方法が現実的な実現方式であると考えられる。このようにメインフレームとワークステーションとの分業を考えると、システムの実現イメージは図2のようになる。

さて、データとなる全文については、その文献の論理構造が表現されている必要がある。また、実際の側面を考えると、文献の清書出力のために印刷用タグを持っている必要があろう。こうした構造の表現及び清書出力用のタグとして、SGMLのマークアップに準じたタグを用い、タグづけされた状態で全文を蓄える方法が考えられる。

4. 2 ホストとワークステーション間で伝えられる情報

ここでは、ホストとワークステーションとの間でやり取りされるべき情報について簡単に整理する。

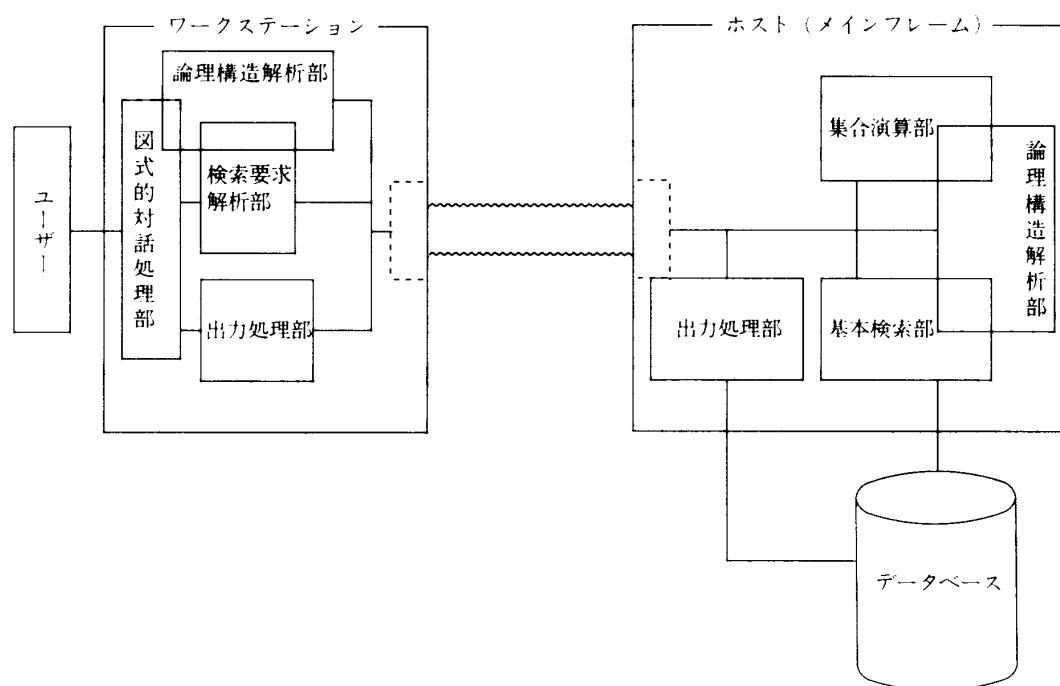


図2 システムの実現イメージ

まず、ワークステーションからホストに伝えるべき情報は、検索の指示と表示の指示という2つに区別することができる。

- ・ 検索の指示に関する情報 これは、前述の3つの検索オペレーションにそれぞれに対応する情報と、その相互関係に関する情報からなる。3つのオペレーションそれぞれに対する情報は次のようなものであろう。

基本検索に必要な情報 検索パターン (文字列)

パスマッピングに必要な情報 ノードの指定 (絶対位置、相対位置、クラス、クラスカテゴリー) とそのオペラント

集合演算に必要な情報 AND、OR、NOT などの論理演算子とそのオペラント

- ・ 表示の指示に関する情報 表示の指示はノードの指定によって行うのが普通であろう。表示の対象は葉であるから、指定されたノードのものと葉を順に出力することになる。

一方、ホストからワークステーションへ伝えるべき情報としては次のようなものがある。

- ・ 論理構造に関する情報 3章1項で述べたように、文献のタイプや分野によって論理構造は異なるであろうから、少なくともデータベースごとに論理構造の定義は異なる。従って、データベースに応じた文献の論理構造定義をワークステーションに送り、ユーザー・インタフェースに反映させなくてはならない。
- ・ ヒット件数に関する情報
- ・ 表示出力のための情報 表示出力のために、文献そのものを送る必要があるが、そのほかに、清書

出力のための出力制御用タグが必要になる。これについては、文献単位で全文を清書出力するためのタグと、部分表示用簡略タグを場合に依りて区別して送るといった必要があろう。表示出力の制御には、SGMLのような標準的なものを用いることが考えられる。

以上のような、ホストとワークステーションとの間の情報のやり取りに関しても、形式的な検討が必要であるが、この点についてはデータモデルとその上の非手続き言語の定式化を進めることにより自ずから決ってくるものと考えられる。

4. 3 ホストとワークステーションにおける各部の機能

次に、図2におけるホスト及びワークステーションの各部分について、その機能を簡単に整理することにしよう。

メインフレーム側では、基本的な検索操作と出力用の処理を行う必要がある。検索操作を行う部分は図2の基本検索部と論理構造解析部、集合演算部であり、それぞれ次のような役割を担う。

- ・基本検索部 検索文字列がデータベース中のどの文献のどこにあるかを見つけ出す部分で、検索オペレーションの基本検索機能を担当する。
- ・論理構造解析部 検索要求に基づいて、文献の論理構造におけるパスマッピング操作を行う。
- ・集合演算部 基本検索及びパスマッピングによって作られた集合に対して演算を行う。

表示出力のための処理は、出力部で行われる。これは、要求に応じて全文の中から出力に必要な部分を表示出力用の必要なタグと共にワークステーションに送る役割を果たす。出力の処理に関しては、ワークステーションとの間で様々な分担方法がありうる。例えば、常に文献単位で全文を表示出力用のタグ付きでワークステーションに渡し、要求に応じた出力処理をワークステーションで行うという方法を考えるとホスト上の出力処理部は不用となろう。しかしながら、通信の問題も考慮すると、ホスト上で出力範囲と出力形式に関してある程度の処理を行うのが適切と考えられる。この処理のためには、検索における論理構造解析部のような機能が必要となろう。

一方、ワークステーション側では、ユーザー・インタフェースの管理とユーザーからの検索要求の解釈、表示の処理を行う必要があり、図式的対話処理部と論理構造解析部、検索要求解析部、及び出力処理部でそうした処理を行う。各部分はそれぞれ次のような役割を担う。

- ・図式的対話処理部 論理構造解析部及び検索要求解析部、出力処理部と協調してユーザー・インタフェースを管理し、ユーザーからの要求を取り込む。
- ・論理構造解析部 ホストから送られてきたデータベースごとの文献の論理構造を解析し、図式的対話処理部でのインタフェース作成に必要な情報を引き渡す。また、検索要求解析部と協調してユーザーの要求をホストへ渡すかたちに変換する。
- ・検索要求解析部 論理構造を考慮した検索に対するユーザー・モデルを作るのが困難であることは前述したが、実際に円滑な検索を可能とするためには、基本的なモデルを持つ必要がある。検索要求解析部は、こうしたモデルに基づき、論理構造解析部と協調しながら対話処理部から渡された検索要求をホストへ渡すかたちに変換する機能を担う。

・出力処理部 ホストから引き渡された検索結果をもとに、表示出力のための処理を行う。

ただし、これらの構成は現在のところ非常に概略的なものである。現在この枠組みのもとで実現方式の細部を検討中であるが、今後、検索オペレーションの形式化や検索インタフェースイメージの明確化にともなって変更される点が出てくる可能性がある。

5. おわりに

ここで紹介した検索方式は現行の全文検索システムが持つ問題点に対する1つの解決案であるが、本文中でも述べたように、現在形式化を進めている段階で、まだ概略的なものである。この方式は、文献の自然な構造に応じた多様な検索ができること、それを一貫した方法で実現できることを特徴とするが、今後、検索の側面からのノイズや漏れに関する評価検討が必要になろう。

一方、システムの実現という点から見ると検索インタフェースやシステムの各構成部分とその相互関係などについてさらに詳細な検討を進めなくてはならない。また、それ以外でも、データベース構築や検索に関わる技術的な側面として、表示出力能力、図表や画像の扱い方など、考慮しなくてはならない問題は多い。

さらに、こうした全文検索システムをサービスとして実際に提供するたにあたっては、全文データベースや全文検索システム全体を、それを取り巻く広い環境の中で見直さなくてはならない。特に、原論文の作成から全文データベースへのデータの入力、利用者への提供という情報のフローについて考慮することがサービスの円滑な実現には不可欠である。これに関しては、研究者の論文執筆から投稿、査読、印刷、DB化までの一連のプロセスを考慮して、全文データベース構築プロセスにおけるタイムラグなどの問題を解決していく必要があるだろう。

注・引用文献

- 1) Young, H. ed. ALA 図書館情報学辞典. 丸善, 1988. [*The ALA Glossary of Library and Information Science*. Chicago, ALA, 1983.] 「フルテキスト・データベース」の項.
- 2) 本節は主として次の文献に基づいている。
全文データベースのシステムに関する現状調査報告書. 学術情報センター, 1987. 116 p.
- 3) Tenopir, C. "Full-Text Databases," *Annual Review of Information Science and Technology*. vol. 19, 1984. p. 215-246.
- 4) 根岸正光. 「学術分野における機械可読文書の作成と通信」 学術情報センター紀要. no. 2, 1989. p. 43-52.

研 究 論 文

ドキュメント・デリバリーのための電子図書館
～リソースシェアリングをめぐる制度的枠組み～

A Conceptual Framework for Document Delivery
Services Based on Electronic Library

学術情報センター 桂 英史

学術情報センター 影浦 峡

〈内容梗概〉

本論の目的は、リソースシェアリングという視点から電子図書館という考え方を制度的側面から再検討し、方法論的に位置付けることにある。電子図書館という考え方の多くは、フルテキストデータベース、マルチメディアドキュメントシステムあるいは分散処理などの技術革新に依存した将来展望であり、具体的な方法論や問題点に関して深い議論はなされていない。本論においては、まず電子図書館という技術的なアプローチが、従来の図書館機械化というアプローチとどのように異なるか、という問題に関して議論する。また、将来の電子図書館システムに基づくドキュメントデリバリーに内在する問題点に関して、とりわけ知的所有権の問題を踏まえながら検討を進める。その際、具体的にシステム構築の前提としてどのような必要条件を具備すればよいか、という問題に焦点を当てながら議論を展開する。

Eishi KATSURA, Kyo KAGEURA

[ABSTRACT]

A new conceptual framework for implementing document delivery services based on new technical methodologies such as fulltext database, multimedia document system and distributed communication environment is described. These new methodologies can be employed by library patron and have a profound potential benefits for document delivery services. The primary focus is on the relationship between copyrights and resource sharing in the feature of document delivery. And some important problems in the context of library systems in the future are also discussed.

1. 序

電子図書館という考え方が、1980年代に入ってから、さまざまな領域で本格的に言及され始められてきた⁽¹⁾。しかしながら、その言及の多くは技術革新の進展に依存した展望であり、具体的な方法論や問題点に関してあまり深い議論はなされていない。

本論の目的は、電子図書館システムの構築を前提として、リソースシェアリングという運用管理形態を方法論的に位置付けることにある。そのために、従来の枠組を整理し、再検討することから着手する。そして、これまでに議論されてきたドキュメント・デリバリーを著作権保護という側面から検討し、システム構築の前提としてどのような必要条件を具備すればよいか、という問題に焦点を当てながら議論を展開する。

以下、2章においては、電子図書館およびリソースシェアリングという概念についてまとめる。まず、本論において想定する電子図書館というアプローチが、従来の図書館機械化というアプローチとどのように異なるか、という問題に関して議論する。次に、リソースシェアリングがどのように具体化してきたかという背景を考察し、それを通して、リソースシェアリングがどのように問題となってくるかという点に関して明確にすることを試みる。

3章では、2章で議論した電子図書館システムに対するアプローチを具体化する上での方法論として、著作権保護という立場から、その制度的枠組みについて考察し、具体的に実システムに適用する場合に想定される諸問題について検討する。

さらに、4章では、3章までの議論をもとに、電子図書館システムの構築の前提となるリソースシェアリングの展望とその諸問題について総括する。

2. リソースシェアリングの背景とシステム構築の前提

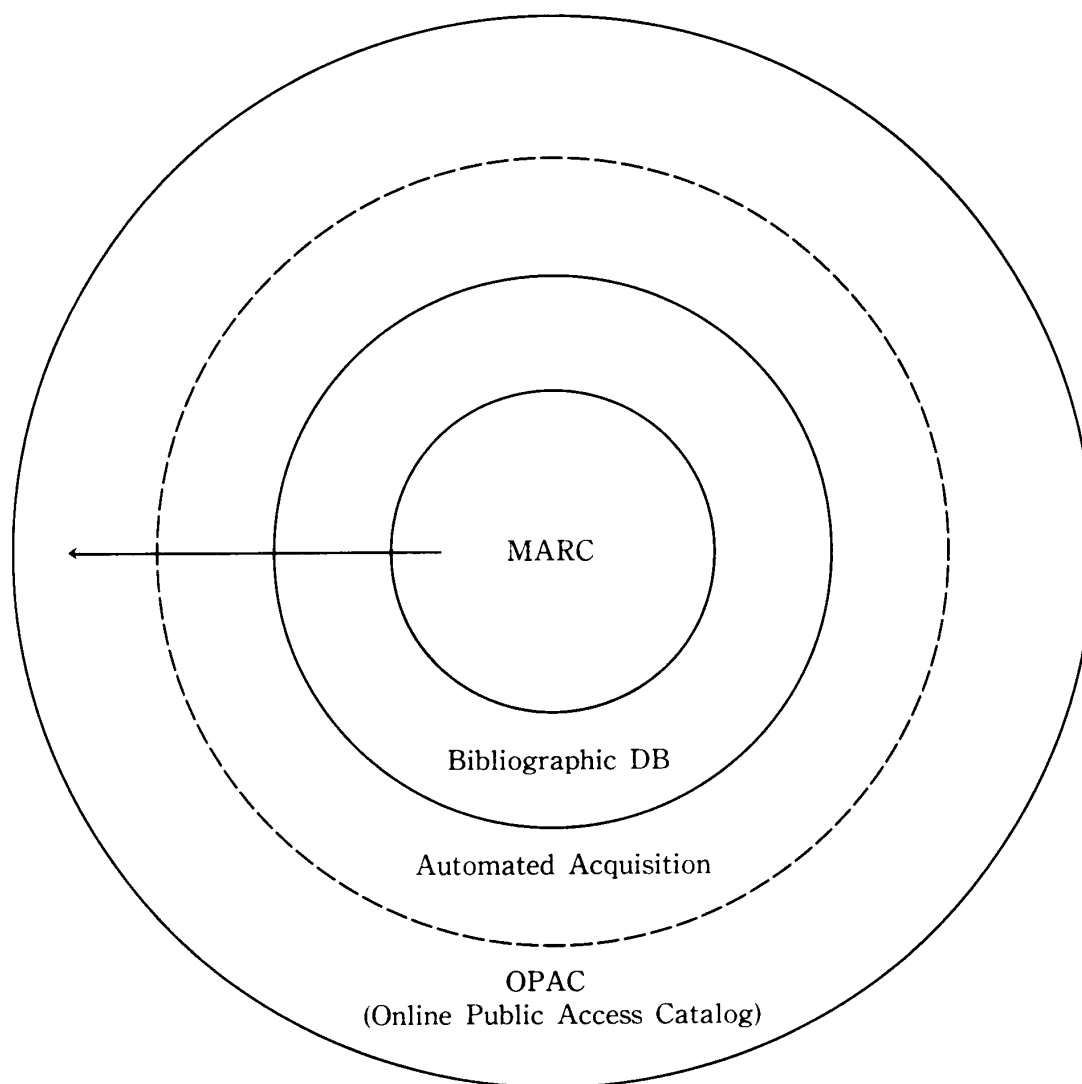
2.1 図書館機械化と電子図書館

図書館という情報管理と情報サービスの環境においては、その管理技術と処理技術に関して、さまざまなアプローチが試みられてきた。[図1-a, b]は、図書館をめぐる管理技術と処理技術に対するアプローチを簡単に示したものである。

[図1-a]においては、図書館機械化というアプローチを示している。図書館で収集される資料が膨大になるにつれて、その整理業務を中心とする各業務の合理化が必要不可欠となってきた。その合理化へのアプローチが、図書館機械化である。言うまでもなく、その合理化はさまざまなコストの低減を図ろうとするものである。

オンライン共同目録作業に代表される現在の整理業務においては、資料の膨大な増加量に対応するだけの労働コストを情報通信技術によって支援することを考え方の基本としている。また、その作業によって得られる書誌データを書誌作成の基本データとして再利用することが同時に図られる。

つまり、図書館機械化というアプローチは、基本的な書誌データの集積とその再利用とりわけ所在

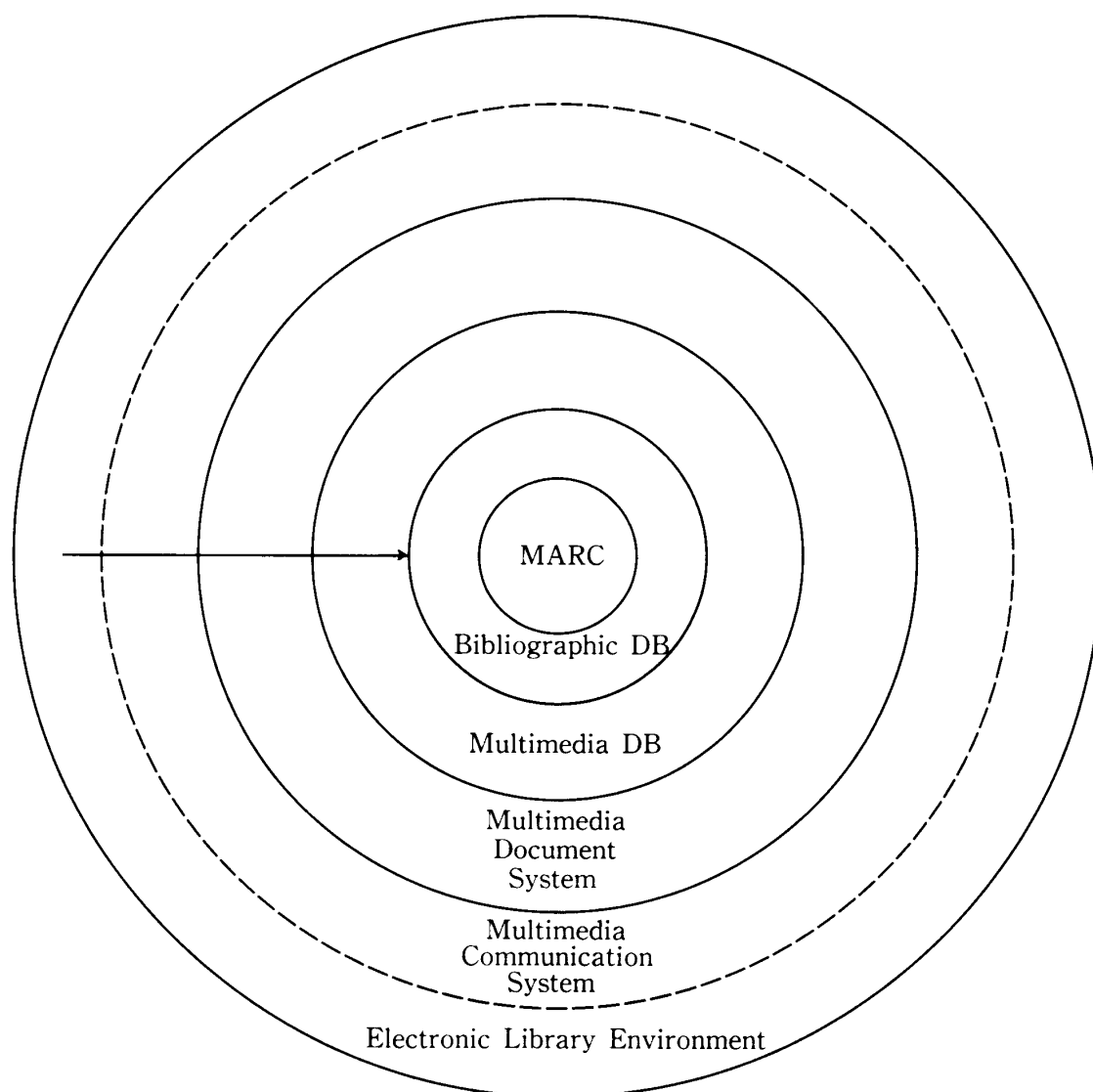


Approach I: "LIBRARY AUTOMATION"

図 1-a 図書館機械化

情報の確認を基本的な機能とする方法論である。その際、一義的に想定されるユーザは、図書館員などの整理業務担当者あるいは書誌作成者であり、エンドユーザへの情報検索サービスは、書誌データ再利用のひとつの効果として位置づけられる⁽²⁾。

現在でも、図書館機械化というアプローチを維持した研究開発は盛んである。例えば、エキスパートシステムや文字認識システムを利用した「自動目録システム」もそういった研究開発の流れを汲んでいる⁽³⁾。この限りにおいて、図書館機械化というアプローチにおける情報通信技術は、情報管理業務を支援するツールとして機能するものとして位置付けられる。このような研究開発は、図書館機械化というアプローチにとって自然な流れであるものの、問題はそれほど単純ではない。



Approach II: "ELECTRONIC LIBRARY"

図 1-b 電子図書館

現在、資料の爆発的な増加は言うに及ばず、資料の形態が記録媒体の多様化によって質的にも大きく変化しつつある。その管理の対象が、印刷メディアのみならず、大幅に広がりつつある。その量的かつ質的な資料の変容にともなって、エンドユーザーはますます資料の探索に大きなコストを強いられ、図書館業務従事者を始めとする資料の管理者にとっては、膨大なデータの処理と管理に追われることになる。そこで、図書館における情報サービスを総合的に見直す方法論として、電子図書館という考え方が出てくる^{(4),(5)}。

電子図書館というアプローチは、図書館機械化のそれとはシステム構築の方法論において大きな違いがある [図 1-b]。

その最も大きな違いは、エンドユーザに対する情報サービスに関する考え方である。

先にも述べたように、図書館機械化における問題解決は、資料管理の合理化である。しかしながら、電子図書館というアプローチにおいては、これに原資料の提供という側面が加味されることになる。それによって、システム構築の方法論は大きく異なってくる。いわゆるシステム・セマンティクスという形式的な仕様記述の方法論は、この問題解決へのアプローチという点に大きく依存する。

さらに具体的に述べると、ユーザのシステムに関わる関わり方が、資料を利用しようとする人と管理しようとする人では全く異なっており、電子図書館というアプローチにおいては、両者の問題を勘案することが必要となってくる。

すでに指摘したように、目録を始めとする二次資料をエンドユーザの側から見れば、原資料へアクセスするためのツールである。その一方で、図書館を始めとする資料の管理業務に携わるユーザにとっても、閲覧や参考などの業務を円滑に遂行する上で、書誌ファイルなどの資源共有は必要である。つまり、資源共有によって、整理業務がより効率的となることは重要な要求である。

一方、エンドユーザには、資料に記述されている内容あるいは原文書そのものを自らの要求に沿ってできる限り早く得たい、という要求がある。そのため、資料の所在を知るべく所蔵ファイルを参照するのはそのためである。したがって、多くの場合、書誌データベース上に表現されているデータを利用する動機が、資料の管理者と利用者とは異なる。

ところが、資料の管理者も、情報通信技術を用いて書誌データの共有や交換のみならず、参考業務や相互貸借などの情報サービスの側面において、より高度な利用をすることはできないか、という要求は存在する。この点で、原資料とその内容を何とか迅速にそしてわかりやすい形態で得たい、という目的を共有するわけである。

つまり、情報サービスの側面を十分に加味した図書館情報システムすなわち電子図書館を構築するにあたっては、こうした利害の共有も本質的な問題解決のひとつとして位置付けることができる。

この限りにおいて、本論における電子図書館というアプローチは、電子図書館というシステムが情報サービスという社会サービスを高度に媒介する公共財（公共メディア）として機能することを目指す研究開発であると言える。

そこで、次節以降においては、電子図書館を情報サービスの公共メディアとして位置付けるために、リソースシェアリングを支援するシステムに与えられるべき諸問題を、制度的側面と技術的側面との双方から検討する。その際、従来議論されてきた ILL サービスの枠組みに対する再検討を前提とする。

2. 2 リソースシェアリングの背景

前節で考察した電子図書館の背景を、技術的な方法論の側面からまとめると、以下のようなことがあげられる。

- ①複製技術
- ②通信伝達手段
- ③経営管理技術の変容

①に関しては、CD-ROMによるデータベース・サービスも実用化の段階を本格的に迎え、情報サービスのあり方そのものが大きく変わる可能性が高い。まだ、そのサービスの中心は書誌データにとどまっているものが多い。しかしながら、サービスの範囲の拡大は、さらに新しいメディアの登場とともにさまざまな方法論が確立しつつある。

研究開発のレベルにおいては、その処理あるいは管理の対象も、フォーマット化（コード化）された情報のみならず、図表や写真あるいは音など非フォーマットの情報にまで広がりつつある。メディアを横断的に管理し、サービスに供する今後の情報サービスを考える上での中核的なテーマのひとつとなるものと思われる。

前述したように、『図書館ネットワーク』あるいは『書誌調整』という理念を持ち出すまでもなく、図書館の情報サービスが、公共性の高い社会サービスであることは言うまでもない。これらの理念的枠組みは、制度的な側面と技術的な側面を相互に有機的に機能させ、整理業務や情報サービスの向上を図るという考えに基づく。書誌ユーティリティによるオンライン共同目録作業の確立は、そのひとつの具体的な成果と言える。また、その理念が国際的な展開を見せ始めており、その成果として、書誌データ交換や相互貸借などの諸制度の協力体制も徐々に整備されつつある。

つまり、リソースシェアリングという概念は、このような資源の共同利用を前提とする情報サービス形態として位置づけることができる。そして、このリソースシェアリングの技術的な側面を支えるインフラストラクチャとして位置付けられるのが、情報通信技術の基盤技術であることは言うまでもない。

電子図書館という考え方の背景には、情報通信技術の時間と空間を克服する技術的なポテンシャルが含意されている。つまり、②に代表されるネットワーク社会の技術革新は、「図書館ネットワーク」あるいは「書誌調整」といった理念の枠組みを具体的に実現する上でのインフラストラクチャとして位置付けられる。また、①は現在の情報サービスの多様性を生んだひとつの要因である。

そして、その背景として、③に該当する情報管理形態の確立があることも忘れてはならない。つまり、書誌ユーティリティに代表される共同利用的な情報サービスの形態は、資料管理のみならず図書館員を始めとする人的資源の管理と配分に大きな影響を及ぼしている。また、書誌データベースにおける書誌ファイルの品質管理に対する問題意識などは、そのひとつの大きな効用であるとも言えるかもしれない。共同利用の効用は、リスクに対する共同債務という考え方を部分的には派生させることを可能にしているとも言える。

本論においては、①から③までの技術的な背景を前提として、よりエンドユーザの利用を指向する電子図書館を実現するためのアプローチを議論をする。近年、『マルチメディアドキュメントシステム』⁽⁶⁾、『ハイパーテキスト』⁽⁷⁾あるいは『マルチメディア・データベース』⁽⁸⁾などと称した研究開発が、ここ数年情報通信技術における最も重要なテーマのひとつとして位置付けられ、さまざまなアプローチが試みられている。

つまり、電子図書館システムは、表現媒体を印刷媒体のみならず、電子メディアによる表現媒体を

もその管理対象としており、リソースシェアリングをめぐる制度的枠組みの考察にあたっては、そのような技術的基盤を背景として議論が必要となってくる。

2.3 リソースシェアリングの諸側面

前節でも考察したように、書誌ユーティリティーの一義的な役割期待である共同目録作業や書誌データの交換などは、大規模な情報通信システムを基盤としている。システム構成と参加するユーザの数が相乗的に大規模になればなるほどユーザの効用が大きくなることを称して、スケールメリットという。

〈図2〉は、そのスケールメリットを前提として、書誌ユーティリティーを始めとするリソースシェアリングの役割期待を「資源共有」、「コスト分担」、「リスク分散」という3つの機能的側面から抱括的に示したものである。

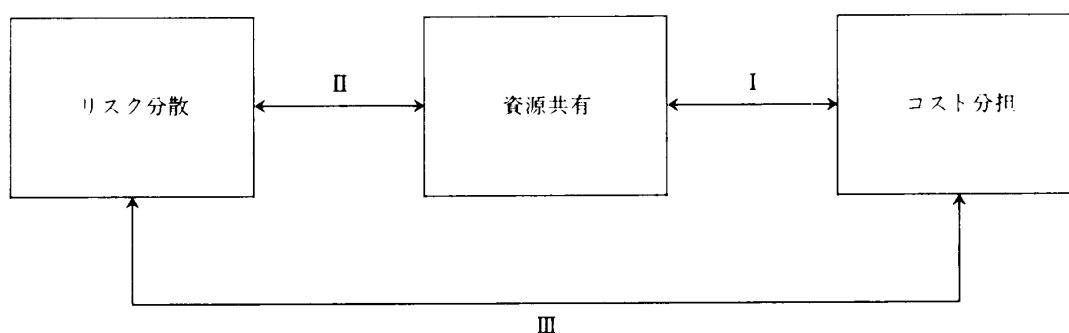


図2 リソースシェアリングの役割期待

「資源共有」は、情報通信システムや書誌データの共有などを指す。そして、電子図書館システムのように、原資料のサービスを前提にした共有の形態も含まれる。また「コスト分担」は、共同目録作業に代表される労働コストの分担を指している。さらに、「リスク分散」は、リソースシェアリングという情報サービスを行っていく上での債務やリスクを総称したものである。

そして、図中I、II、IIIすなわちその各機能間に生じるインタラクションは、以下のような諸問題を含んでいる。

I. 「資源の最適配分」

ここで言う資源には、情報通信システムなどのハードウェア、データやコンピュータ・プログラムなどのソフトウェアそして業務に従事するマンパワーなどが含まれる。つまり、資源の最適配分とは、これらの資源をユーザの要求に応じてどのように配分するかという問題である。

電子図書館というアプローチは、従来までの一極集中的な情報システムを再検討し、低いコストで原資料提供などの高い効用を実現する契機としても位置付け得る。

II. 「債務負荷の問題」

原資料のカレントなサービスを行う電子図書館システム構築にあたって、スケールメリットの側面から議論されなければならない問題が、資原共有とリスク分散とのインタラクションいわゆる債務負荷の問題である。

例えば、原資料をカレントに情報サービスとして提供しようとするれば、資料所蔵数の多い図書館等に債務負荷が高まるのではないかと、という問題が生じる。つまり、大規模な図書館による小規模な図書館への原資料提供の機会が多くなることが予想される。つまり、大規模な図書館の債務が大きくなる可能性が高いということである。また、原資料を現物貸借に供する機会が多くなると、所蔵管理や予算配分法の再検討などリスク管理が強いられることになる。

このような大規模図書館の負荷となる債務やリスクは、ひとつの社会的責務であるとする考えも成り立つ。しかしながら、そのリスク分散とコスト分担の問題は、できる限り電子図書館システムの仕様化技術に反映されなければならない。

III. 「業務フロー再編成の問題」

原資料を前提とするリソースシェアリングを図るためには、少なくとも複写サービスを行ったり、梱包したりする要員と機器を充当しなければならない。それによって、受入から整理そして閲覧といった業務のフローが変容する可能性があり、前述の「債務負荷の問題」も含めて相応の経営管理技術の確立が求められることになる。

これらの諸問題は、いずれもリソースシェアリングに伴う図書館経営管理形態の変容と密接に関連する問題である。1章でも述べたように、原資料そのものにアプローチすることができるようなシステムであることが電子図書館システムの基本的な仕様であるとすれば、以上3つの問題に関しては、相応の問題解決を想定できなければならない。

そこで、次章においては、原資料の提供を前提とする電子図書館システムにおける情報サービスの形態について討論する。ILL サービスというサービス形態を含めて、従来のサービス形態を原資料提供（ドキュメント・デリバリー）として包括的に位置付け、そのサービス形態を実現するための制度的枠組みに関して、著作権保護に焦点を当てながら議論する。

3. 情報サービス形態としてのドキュメント・デリバリー

3. 1 図書館ネットワークと ILL サービス

図書館ネットワークは、その発生時においては、かなり理念的なものであった。しかしながら、原資料の提供というサービスの側面から、改めて図書館ネットワークのあり方を議論する必要がある⁽⁹⁾。

例えば、『書誌コントロール』は、その実、制度的な枠組と技術的な諸問題が同居している。ただ、書誌ユーティリティの実現によって、所在情報を得ることに限ってネットワークが実現したことは、常に議論の前提とする必要がある。資料の所在情報を共有することによって、書誌ユーティリティが現物貸借を行う上での大きな基盤となったことも確かであり、その図書館間協力の形態もより強いも

の（コンソーシアム）となって現れている。事実、アメリカの主要大学間のコンソーシアムにおいては、外国資料の購入などかなり実際的な局面において協力体制が取られ、国際的な協力体制をめぐって制度的な確立も見られる⁽¹⁰⁾。

ただし、ILL サービスには基本的な限界がある。それは、書誌ユーティリティの基盤である情報通信システムと現物の輸送のシステムとが、本来全く異なるシステムであるに関わらず、二つのシステムを有機的かつ効率的にリンクさせなければならないことによる。そこには、コスト負担の問題（制度的側面）を考えても、異なるシステムを并存させる問題（技術的側面）から言っても、大きな問題が含まれている⁽¹¹⁾。

書誌ユーティリティやデータベースのサービス拡充に伴って、エンドユーザはその所在情報をますます迅速に入手できるようになる。とすれば、原資料の入手を支援するシステムを要求として持つことは、ごく自然である。そして、その要求についても、迅速かつ正確であることが求められてくる⁽¹²⁾。

複写機に端を発する複製物の大量消費や情報処理技術の高度化によって、従来の利用形態が大幅に変わりつつあることは言うまでもない。例えば、データベースの全文化という傾向は、今後より高度な技術的方法論の確立が予想される。複製技術と通信技術を高度に統合化した電子図書館システムの出現によって、記録媒体の形式が従来の印刷物とその複製物を大きく超えて、サービス形態もより多様化することが予想される。この限りにおいて、遡及的な書誌情報を提供する情報サービス形態に対して、原資料のカレントな提供を実現する情報サービス形態をドキュメントデリバリーとして包括的に位置付けることができる。また、そうしたサービスの多様化に伴って、印刷物という形態を取らない著者の表現媒体が増え、著者の権利を侵害する機会が増大してくる可能性が高いことも大きな問題である⁽¹³⁾。これは、技術革新と社会制度とのタイムラグが、大きくなってきたことのひとつの現れでもある。つまり、多様化しつつある情報サービスとそれを支える技術革新の動向にあって、著作権を始めとする著者の権利をどのように保護しつつ、情報サービスの機能的側面を高度化していくか、という問題が極めて重要となってくる。

実際、現物貸借による ILL サービスの形態は、著作権保護を前提とする図書館サイドのスケープゴードであった。もっとわかりやすく言えば、現物をユーザに渡してしまえば、ユーザがその現物すべての複製を取ろうが、その複製物をどう利用しようが、図書館にとっては関知しない、という立場を維持する極めて消極的な方法論である。それが、従来の ILL サービスにおいて、著作権を始めとする著作権保護の問題を事実上棚上げにしてきたひとつの大きな理由である。

もちろん、現物貸借を前提とするリソースシェアリングいわゆる ILL サービスは、今後も存続していくものと思われる。しかしながら、従来の枠組みでは議論できない問題が多くなっている。ILL サービスという考え方は、物理的な印刷物（印刷物そのものもとより、複写機による複写物やファクシミリによる出力なども含む）の検収を前提にしたものである。しかしながら、先に述べたドキュメントデリバリーを実現する上での方法論にあっては、資料形態に対する考え方そのものを変えていく必要がある。

図書館ネットワークという理念において、原資料提供を実現するドキュメントデリバリーが大きな役割期待であることは言うまでもない。そして、リソースシェアリングという情報サービスの形態は、共同利用という効用的な側面のみならず、コストの分担とリスクの分散という社会的な機能を前提としている。従って、電子図書館というシステムの構築にあたっては、その技術開発の側面を加味しつつ、コストの分担と同時に、著作者の権利保護などの共同債務の方法を同時に検討する必要がある。また、検討された課題は、技術的な側面を加味しながら電子図書館の仕様に反映されなければならない。

3. 2 『多段階著作者保護』の枠組み

リソースシェアリングという考え方は、図書館ネットワークあるいは学術情報活動をめぐる環境だけに与えられている課題ではない。近年とみにコンピュータ・ソフトウェアの資源共有を目的として一般化している PDS (パブリック・ドメイン・ソフト) というコミュニティのあり方も、そのひとつの顕著な現象と言える。

この PDS という考え方は、80 年代に入ってさまざまなものが登場してきたが、すでに 60 年代に登場した『リソース・ワン』⁽¹⁴⁾ を起源とする共同利用の形態と言える。その背景には、電子メディアが柔軟に大量の情報を表現したり、複製したりすることが可能である、という利点を肯定的に社会サービスとして共有しようとする考え方がある。

ところが、原資料の管理とサービスをめぐる制度的枠組みは、概ね情報通信技術の操作や機能を棚上げにした議論である場合が少なくない。そして、この限りにおいて、印刷媒体と刑法でいうところの電磁的記録⁽¹⁵⁾とは、少なくとも異なるものとして扱われている。刑法あるいは著作権法などの法律で、電磁的記録の著作者保護を図ることが必要であることは言うまでもない。と同時に、電磁的記録と印刷媒体との管理上の整合性を図る上でも、著者による表現媒体を区別して制度的枠組みの中で議論することそのものが問われなければならない。

とりわけ、ILL サービスというサービス形態を考慮すると、現在、資料を要求する側すなわち受益者であるユーザには、現物貸借自体のコストを債務として課していない。つまり、コストを負担する電磁的記録とコスト負担のない現物貸借が同じ表現媒体でありながら、並存することになってしまう。

例えば、全文データベースによる情報サービスが一般化したとしても、必ずしもユーザの要求する資料がデータベース上に電磁的記録として表現されているとは限らない。むしろ、そういったケースの方が多いためと考えられる。現在の制度的枠組みにおいては、現物貸借に関してコストを債務としないという姿勢であるため、所蔵情報を得て現物貸借に至るまでには回線使用量、複製代金あるいは郵送代金などのコストしか介在しておらず、著作権保護という考え方がコスト負担の原則に反映されていない。

しかしながら、全文データベースによるサービスと現物貸借による ILL サービスがリソースシェアリングの形態として並存してくると、事情は複雑になってくる。つまり、データベースのフルテキスト化によって、表現媒体が電磁的記録と印刷媒体にわたる可能性も十分に考えられるようになる。

現在の著作権保護をめぐる法律的な議論は、いわゆる電磁的記録を対象としたものが多くなっている⁽¹⁶⁾。もちろん、その背景には、情報通信技術の進展に伴う著作権保護に対する危機感があることは言うまでもない。しかしながら、そうした議論は大きく技術革新の動向に依存するため、著作権保護を図ろうとする議論としては、基本的な枠組みが欠落しているように思われる。

例えば、図書館における情報サービスにおいて、著作権者の権利を侵害する複写物が作成される可能性は、電磁的記録も印刷媒体も同様である。それにも関わらず、印刷物を物理的に貸借する場合には著作権保護の債務について議論せず、電磁的記録の側面には著作権保護という議論を持ち込むのは、少なからずかたよった議論であると言える。

電子図書館というシステムが原資料のエンドユーザへの提供を第一義とする限り、制度的枠組みの核心は、電磁的記録を制度としてどう扱うかという論点だけでは不毛である。著作権保護という立場から考えると、電磁的記録と印刷媒体をどのようにシステムとして共存させていくかというドキュメントデリバリーの方法論は、印刷媒体と電磁的記録をできる限り等価な著作権保護を図ることを基本的な枠組みとして位置付け、技術的な問題や制度的な運用方法を検討する必要がある。

Ted Nelson は、ソフトウェアやテキストなどの表現メディア全般に関して、段階的なローヤリティの導入を提唱している⁽¹⁷⁾。Nelson のアイディアにあつては、段階的な著作権保護を表現者(著作物の著者あるいはコンピュータプログラムの作成者。以下、これを「著作者」と総称する)に与える。利用者に対しては、表現形態(印刷媒体や電子メディア)に応じたアクセス方法を提供する。と同時に、そのアクセス方法に応じて著作者の権利とそれを保証するクリアリング機能を具備することによって、著作者の権利保護を包括的に実現しようという考え方である。

とりわけ、再三述べてきたように、ドキュメントデリバリーという概念は、資料の形態をさらに抽象化することによって情報サービスをより高度化する技術的な方法論である。その要求や仕様を定義する上でも、その制度的枠組みに関する詳細な議論が前提となることは言うまでもない。

3.3 電子図書館におけるドキュメントデリバリーの共同債務

これまでに述べてきたように、単なる現物貸借や一部の複写物だけをサービス形態とする ILL サービスの機能だけでは、議論としては不十分である。制度的側面から見ても、技術的な側面から見ても、ドキュメントデリバリーを前提とする情報管理の一般的な方法論を議論することはできない。

そこで、前説で議論した「多段階著作権保護」という枠組みを、全文データベースを用いたドキュメントデリバリーを想定して⁽¹⁸⁾、拡張することを試みる。〈図3〉では、ドキュメントデリバリーを情報サービスとして行う上で、著作権保護の監視を行うことを想定している。

この Nelson のアイディアを拡張した段階的な監視システムは、著作権保護のクリアリング機能を有するチャンネルとして位置付けることができる。さらに、この『多段階著作権保護』の考え方を各チャンネル単位で考えると、例えば以下のような制度諸問題を想定することができる。

〈チャンネル1：媒体への表現〉

著作物が何らかの媒体に表現された時点で、著作権保護が債務となるチャンネル。すなわち、印刷物

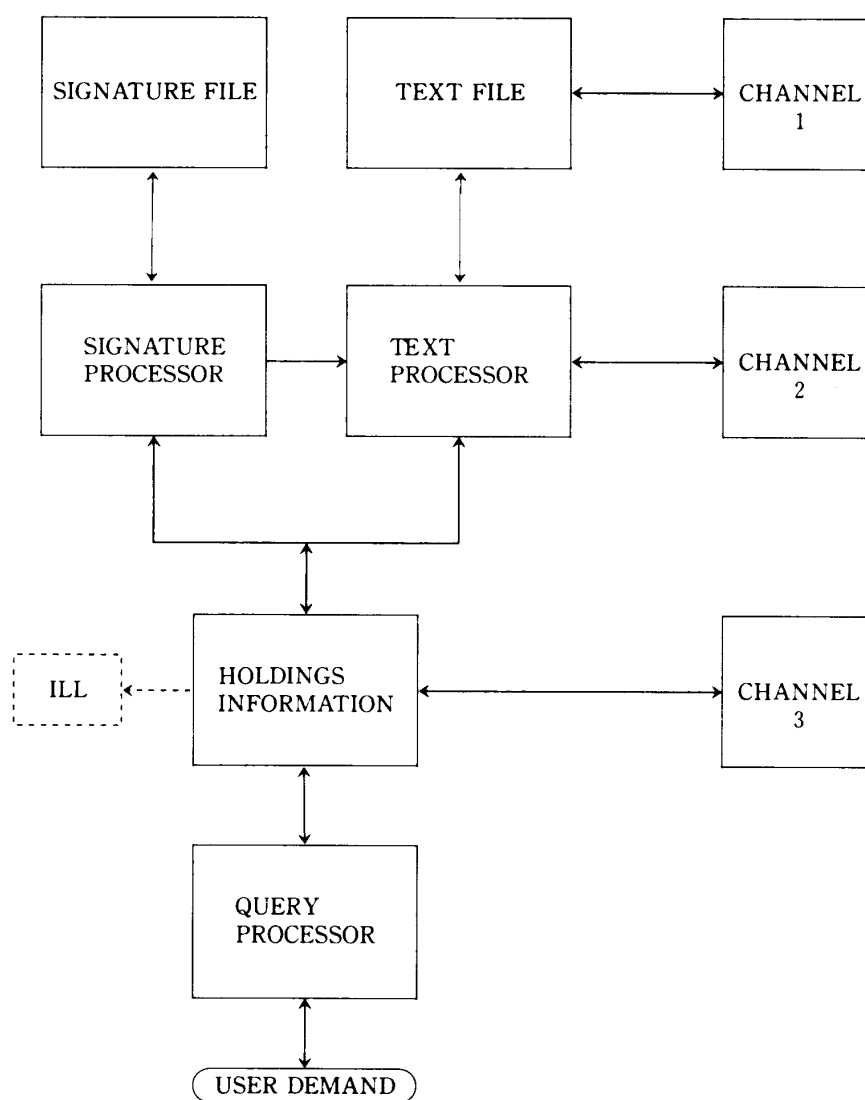


図3 「多段階著作権保護」の監視チャンネル

を発行する際に著作物の印税契約が成立するのと同様の枠組みを電磁的記録にも導入する必要がある。つまり、著作物がデータベース上に機械可読の形態になった時点で、著者に保護される権利が発生する、という考え方である。

〈チャンネル2：検収方法〉

従来のILLにおける検収にあたって、著作権保護が債務となるチャンネル。すなわち、書誌データのレベルまでは著者に対する著作権保護は考慮されず、それ以降すなわち原資料を入手するにあたっては、相応のコストを負担するという考え方である。つまり、要求に該当する原資料が所蔵する図書館の書架から搬出され、現物貸借の対象として検収される場合には、著作権に相当するコストを受益者が負担する。複写サービスの要求もこの中に含まれる。この時、受益者を図書館等の組織とするか、

エンドユーザとするかは、議論の余地があるものと思われる。

〈チャンネル3：ユーザアクセス形態〉

電磁的記録の形態となった表現媒体に対してアクセスするにあたって、著作権保護が債務となるチャンネル。

この点に関しては、あまりにも特定の機能に特化した議論は技術革新に制度が対応することができなくなる可能性があり、制度面での柔軟性を失わせる結果にもなりかねない。記録を知覚できる可能性のある形態をユーザが得た時には、常に著作権保護の債務が発生するという問題意識を図書館などのサービス機関とエンドユーザに定着させることが先決である。

いずれにしても、そうした著作権保護にあたっては、多段階の印税が、従来以上に課せられることになるため、そのコスト負担の問題が生じる。受益者であるエンドユーザのコスト負担としてどのように転化していくかという問題は、まさに制度的問題であり、政策的な意志決定の問題である。

それと同時に、各チャンネルにおける著作権保護を技術的にも保証することが必要となる。この著作権保護のクリアリング機能を、従来から運用されてきたデータベースの課金システムとして実現する技術的な方法論も並行して議論していく必要がある。

4. まとめ～『電子図書館』構築の枠組み

本論における考察からも明らかであるように、現物貸借を主眼とする ILL サービスを電子図書館システムの中で独立した情報サービスとして位置付けることは、今後の情報サービスの方向性から見ても無意味であり、システムとしての柔軟性に欠ける。そして、本論で考察した問題点を考慮して、今後電子図書館システムを構築していく上での前提としては、以下のような論点があげられる。

- ①ドキュメント・デリバリーにおける技術的課題
- ②リソースシェアリング支援と相互協力の枠組み
- ③知的所有権についての研究開発動向を加味した議論

システム構築の技術的方法論である①に関しては、いわゆる「マルチメディアデータベース」や「マルチメディア通信」などの研究開発において議論されている諸問題と基本的に整合する。しかしながら、本論でも述べたように、電子図書館システムにはリソースシェアリングという役割期待があることを考慮すると、②に関する公共メディアとしての役割を十分に議論する必要がある。

しかも、リソースシェアリングは国際的な規模で展開していきつつある。電子図書館システムが担うコストと債務に関する原則の確立が所望されることは言うまでもない。そして、制度的側面からは、表現媒体という意味で、印刷媒体と電磁的記録といった考え方を超えて、その電子図書館システムが直面する諸問題を包括的に議論することが必要となってくるものと考えられる。その際には、3章で述べたサービスのチャンネルに応じた著作権保護をシステム構築の前提としなければならない。

しかしながら、図書館あるいは情報サービスを行う機関がその公共性をどのように考え、技術革新をめぐるのさまざまな変化に対応していくか、という問題は、リソースシェアリングの今後の方向性を示す上でも大きな問題である。

例えば、3.3で述べた〈チャンネル2〉における債務は、「図書館の資料を有料化するのか」という基本的な問題に直面することになる。該当する資料は、いくつかのコストとリスクを背負って図書館のコレクションとなっている資料である。つまり、資料の価格に含まれる著者に対する著作権料やユーザによって権利が侵害された場合の責任が、コレクションの中には含まれている。図書館が著者に対して、その資料の購入にあたっては、個人のユーザーより高いコストを支払うのはごく自然である。著者の側から見れば、図書館という存在は自らの表現媒体や公共性を持つという利点と同時に、著作権者の権利が犯される可能性が実に高いことも、また確かである。

マルチメディアデータベースあるいはマルチメディア通信などの革新に伴って、印刷イメージそのままの資料が図書館で交換されるとすれば、現物貸借とほとんど同じ効果をユーザは得ることになる。資料形態の明確な区分がますます難しくなってくることは、技術革新の動向から言っても、不可避である。

電子メディア上の表現媒体の著作権保護に関する議論をより詳細化する必要があることは言うまでもない。しかしながら、それ以前に図書館で供される資料をめぐって、著作権保護のあり方についても一度包括的に議論し直す必要があるものと思われる。例えば、前述したような多段階の著作権保護に伴うコスト負担に関しては、情報サービス機関が積極的なイニシアティブを取ってしかるべき問題である。

電子図書館システムの構築は、情報サービスのあり方を大きく変える可能性を持っている。それと同時に、そのシステム構築がもたらす社会的な要因も同時に検討していかなければならない。もし、システム構築の方法論を誤ると、かえってシステムが社会的な要因に阻まれて柔軟性を失う結果にもなりかねない。そして、システム構築を前提とするこの検討課題は、急務な課題として詳細化する必要があることは言うまでもない。

付記

本論は、科学研究費総合研究(A)「密結合型図書館ネットワークにおける統合業務システムの研究」の研究結果に基づいている。

参考文献

- (1) F・W・ランカスター「紙なし情報システム」植村俊亮訳、共立出版、1984.
- (2) De Gennaro, R. "Library Automation and Networking: Perspectives on Three Decades," *Library Journal* 108, 1983.
- (3) Weibel, S., Oskins, M. and Vizine-Goetz, D. "Automated Title Page Cataloging," *Proceedings of the 50th ASIS Annual Meeting*, Boston, MA. 1987.
- (4) Gibbs, S., Tsihchritsis, D., Fitas, A., Konstantas, D., and Yeorgaroudakis, Y. "Muse: A multimedia filing system," *IEEE Software*, Mar. 1987.

- (5) Caplinger, M. "An Information System Based on Distributed Objects," ACM OOPSLA'87, 1987.
- (6) Postel, J. B., Finn, G. G., Katz, A. R., and Reynolds, J. K. "An Experimental Multimedia Mail System" ACM Transaction on Office Information Systems, Vol. 6, No. 1, Jan. 1988.
- (7) Yankelovich, N., Meyrowitz, N. and van Dam, A. "Reading and Writing the Electronic Book," IEEE Computer Vol. 18, No. 10, 1985.
- (8) Woelk, D., Kim, W. and Luther, W. "An Object-Oriented Approach to Multimedia Databases," Proceedings of the ACM SIGMOD Conference on the Management of Data, 1986.
- (9) Besemer, S. P. "Criteria for the Evaluation of Library Network, " Resource Sharing and Information Network, Vol. 4, No. 1, 1987.
- (10) 国立国会図書館図書館協力部国際協力課編訳「IFLA の国際貸出ガイドライン」びぶろす、Vol. 37, No. 7, 1886.
- (11) Chen, C. "Criteria of Effectiveness for Network Delivery of Citizens Information through Libraries," Washington, D. C., Department of Education, 1982.
- (12) Morrison, P. D., ed. "Public Lending right, " Library Trends. Vol. 29, No. 4, 1984.
- (13) R・デ・ジェナーロ「著作権、リソース・シェアリングと困難な時代／図書館の現場から」小林正訳、現代の図書館、Vol. 17, No. 1. 1978.
- (14) Roszack, T. "The Cult of Information, " Pantheon Books／Random House, 1987.
- (15) 通商産業省監修「情報六法～昭和 62 年版」通産資料調査会、1986.
- (16) 阿部浩二「知的所有権についての最近の動向」文化庁月報、1990 年 2 月号
- (17) Nelson, T. "Literary Machines／Computer Lib," (Self-Published), 1983.
- (18) Tsichritzis, D. C. "Office Automation," Springer-Verlag, 1985.

研 究 論 文

知識ベースシステムにおけるタスクの役割

Roles of Tasks in Knowledge-based Systems

学術情報センター 小山 照夫

1. はじめに

Chandrasekaran による generic task の概念の提唱以来、知識ベースシステムにおけるタスクの役割が注目を集めている。Chandrasekaran における Generic task は、さまざまな種類の問題解決作業に、ある種の部分的な作業が共通に出現するという事実に着目し、そのような部分的な作業を「タスク」と位置づけた上で、このような部分的作業に対応する問題解決機構を用意しておけば、知識ベースシステムの構築を効率化することができるのではないかという考え方に基づいている。この意味で Chandrasekaran 自身は "Generic" という用語を用いたのであるが、最近の議論ではむしろもう少し広い意味で、「タスク」という用語に触発された、全体としての推論作業をより基本的な作業単位に分割する方法論および、このような考え方に基づくシステム構築方法論が問題とされているということができよう。以下では知識ベースシステムにおけるタスクの役割について、主として情報モデリングの観点から考察を加える。

2. 知識ベースシステムと情報モデリング

知識ベースシステムに限らず、情報処理システム一般で行われているいわゆる情報処理は、現実の世界のさまざまな側面をコンピュータの中にモデルとして取り込むことにより、それに対してさまざまな仮想的な操作を加えることを可能とし、操作の結果を予測することによって、現実の世界における行動指針を明らかにしようとするものであると考えることができる。(図1)この意味で、現実世界をいかにモデル化するかは、情報処理システム全般における重要な問題であり、このモデル化の方法論を情報モデリングの問題と呼ぶことにする。

一般に、情報モデリングの問題には二つの主要な側面が考えられる。その一つは、問題に関係する対象世界のものごとをどのようなものとして捉えるかという、情報構造のモデル化であり、他の一つは、そのようなモデル化された情報構造に対して可能な情報操作のモデル化である。このことを最も端的に表している一つの例がデータベース管理システムであり、そこではデータ定義言語とデータ操作言語という形で、二つの側面からのモデル化の手段が提供されてきたということができよう。

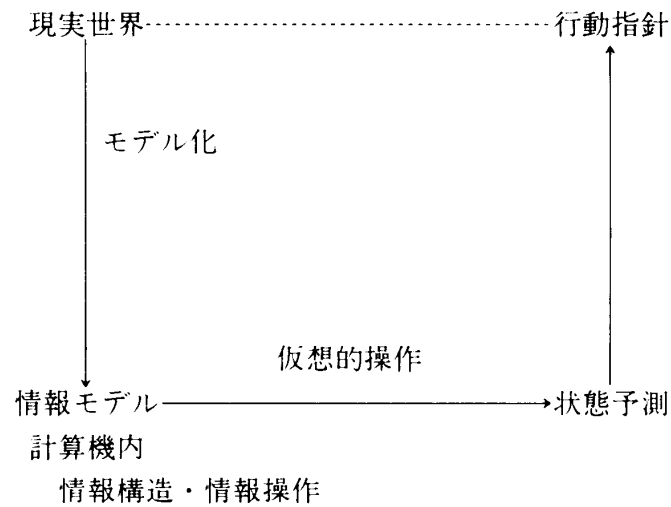


図1 情報モデリングの立場からみた情報処理

知識ベースシステムについていえば、通常知識表現と呼ばれるものが、情報モデリングの方法論を提供するものと考えられるのであり、これにしたがって記述された知識が、情報モデルを構成すると考えられる。

一般には知識ベースシステムにおいて知識はその中核をなす概念であるとされているが、実際には知識ベースシステムにおける知識とは何かという問題に関しては、きわめて曖昧な形で慣習的な了解が得られているにすぎない。多くの場合、システムが採用する知識表現に従って記述された、ルールやオブジェクトの総体を指して知識と呼んでいるが、これは情報モデリングの立場からすると、きわめて曖昧なものでしかないことは明らかであろう。以下ではまず、情報モデリングの立場から従来の知識表現の役割を考察し、ついでタスクの果たすべき役割について述べる。

3. 知識ベースシステムの知識表現と知識

情報モデリングの立場からすると、知識ベースシステムで扱われている知識もまた、対象領域の情報構造に係わるものと、情報操作に係わるものに分類することが可能である。これまでの知識表現の枠組みの中で、前者の代表的なものとしてフレームないしはオブジェクトを考えることができるし、後者の代表的なものとしてルールを挙げることができるであろう。さらには述語論理（整合論理式）についても、明確な意味付けがされているわけではないが、よりプリミティブな形で、情報構造とその上での操作を定義する手段が提供されていると考えることができる。

3.1 情報構造モデル

情報構造モデル化という観点からする限り、フレームないしはオブジェクトの枠組みは、従来のプログラミング言語で取り扱われてきた構造体ないしは抽象データ型の延長上にある。ここでは例えば、

特定の問題解決の局面において、対象とする一つの存在（例えば人間）について、どのような属性を想定すれば良いか、あるいは対象とするものごとにはどのような構造があり、どのような構成要素がどのような相互関係の下にまとめあげられているか、それらの属性ないしは関係に対してどのような関連情報や制約が存在するかなどが定義される。更に、各々の属性や関係の取り扱いや他のフレームないしオブジェクトとの情報交換の方法、制約条件のチェック、制約違反の場合の処理などにどのようなものがあるかが、手続き（付加手続きないしメソッド）の形で記述されることとなる。ここで特徴とされるのは、ものないしはことに関する情報構造の定義と、そのような構造の下での情報操作の枠組みが一体として定義されることである。場合によっては、手続きの代わりにルールなどの演繹推論関係が用いられることもあるが、このような場合には演繹推論関係を解釈実行する推論機構が必要となる。このような推論機構をシステムに取り込む方法としては、ルール集合が関連付けられたオブジェクトのメソッドとして推論機構を定義する方法や、ルールを管理するオブジェクトを別個に想定し、このオブジェクトの下にルールを管理するとともに、そのオブジェクトのメソッドとして推論機構を定義する方法などがある。

対象世界のものごとの情報構造の定義は、述語論理（整合論理式）を用いて行うことも可能である。述語論理の記述では、例えばホーン正節に見られるように、単なる事実の記述が可能であるが、これを利用することにより、特定のものごとに対して、どのような属性なり構成要素なりがどのような役割で関与するか、それらの属性や構成要素にどのような関連情報や制約などが存在するかを記述することが可能である。ただし、フレームないしはオブジェクトでは、このような情報構造の記述方法がある意味で定型化されており、選択の余地が少なかったのに対して、述語論理の記述でははるかに大きな自由度が残されている。これは、フレームないしはオブジェクトの場合、ものやことの情報構造を記述するという限定的な意味付けが明確であり、そのような意味付けのもとでの利用が前提とされているのに対して、述語論理の場合は、より汎用的な情報処理の枠組みを提供しようとするものであることによると考えられるであろう。また、述語論理の体系においては、フレームやオブジェクトと異なり、情報操作の枠組みは、情報構造の定義とは個別に定義された演繹推論関係として記述されることになる。ただし、このようにして定義された推論関係を、構造記述と一体のものとして見せるための環境を用意することは可能であるし、また、prolog などのように解釈系の特性が予めわかっている場合には、演繹推論関係を手続きの実行を指示するものとみなすことができる場合もある。

情報構造の定義においては、しばしば異なったものごとの間に関係が定義されるが、このようなものごとの間に関係が、対象領域の特殊性によって、さまざまな暗黙の推論関係を導くことがある。例えば複数の構成要素からなる機械の場合、構成要素の故障は通常機械全体の故障を引き起こすが、故障した要素に対して冗長要素が存在しており、その要素が故障していなければ全体の故障とはならないなどの推論が可能となる。しかしながら一般には、このようなものごとの間に関係に基づく暗黙的推論関係はきわめて多様かつ領域特殊性の高いものが多いところから、多くの場合概念階層に基づく属性の継承以外の推論関係は、必要となる都度、付加手続きの形で実現されている。

3. 2 情報操作モデルと演繹推論

知識ベースシステムでは、対象領域の情報構造の上での情報操作としては、フレームやオブジェクトの付加手続きやメソッドの形で、通常のプログラミング言語などによる手続きとして記述されるもののほか、フレームやオブジェクトとの間のメッセージ交換によるもの、および演繹推論によるものが考えられる。知識ベースシステムでは、一般にはグローバルな形での手続きが用いられることは少なく、大域的な情報操作のほとんどは、フレームやオブジェクトの間のメッセージ交換か、あるいは演繹推論関係を用いた推論によって行われることになる。これらの内で、メッセージ交換によるものは抽象データ型の枠組みにあらわれるものと類似のものであるとみなすことができるであろう。

演繹推論は、知識ベースシステムに特徴的な情報操作の枠組みであるということが出来る。ここでは通常のプログラミング言語で定義される手続きとは異なる、宣言された関係として演繹推論関係が定義され、関係の記述形式に対応する証明機構（推論機構）の働きにより、既知の事実からどのような未知の事実が推論可能であるかが示されることとなる。

演繹推論では、ごく単純な場合を除き、一般に実用となる程度に複雑な問題では、推論の単調性が完全に満足されている場合はむしろ希であると考えられる。多くの場合、ある種の仮説推論により、部分的な非単調性が生じてくる結果、演繹推論関係の適用順序に依存して、証明可能な結果が異なってくる事が予想される。この意味で演繹推論における推論関係適用の順序の制御は重要な問題である。プロダクションシステムに由来する演繹推論システムの場合、推論関係の適応は、基本的には作業記憶の内容の変更をもたらすと考えられる。このとき、ルールの適用順序が異なるならば、ある段階まで推論が進んだ段階で、作業記憶に記録される証明済みの事象はさまざまに異なることとなる。(図2)ここでは、1. 証明可能な作業記憶内の少なくとも一つが、当初の目標を満足するならば、そのような作業記憶内容を最も効率よく発見する推論関係適用順序はどのようなものになるか、2. 目的に適合する証明可能な作業記憶内容が複数存在する場合、そのような全ての作業記憶内容を最も効率よ

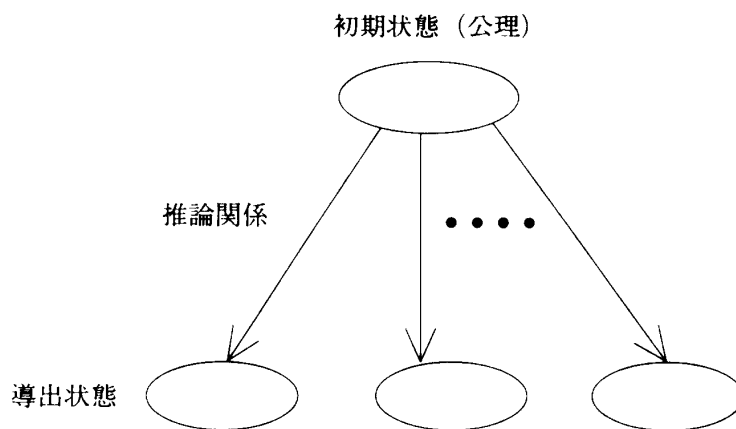


図2 推論関係の適用と状態の導出

く調べ挙げるための推論関係の適用順序はどのよなものになるのか、が問題となる。

演繹推論関係の適用順序に関しては、このほかに、必要となる情報のコストを最小化するためにはどのような適用順序が望ましいか、あるいは利用者に理解しやすい推論戦略を実現するためにどのような適用順序が望ましいかなどが問題とされる場合もある。

3.3 知識表現に関する問題点

以上、情報モデリングの観点から、現在の知識表現の果たす役割に関して考察を進めてきたが、ここで最も問題となるのが、問題解決のプロセスに係るモデリングの方法論が陽に定義されていないことである。この事が最も明確に現われるのが、演繹推論関係の適用順序の決定に関する問題である。現在の知識ベースシステムでは、システム全体の目的についての記述は存在するが、そのような目的を達成するために、どのような問題解決プロセスが想定されているかは、明確な形で示されていない。実際の問題解決は、局所的に定義された手続きや、フレーム、オブジェクト間のメッセージ交換、演繹推論などによって実行されるが、それぞれが全体問題解決プロセスに対してどのような役割を果たしているかは、一般には陽に記述されていない。特に演繹推論関係については、必要とされる処理全体に比較して、一つの推論関係で表現される情報操作単位が小さいこと、演繹推論関係自体、形式的な汎用の情報操作の枠組みを決定するものであっても、そのような形で記述された個々の情報操作の意味にまで立ち入るものではないことなどの事情から、個々の推論関係が問題解決全体の中で果たす役割は必ずしも明確でないことが多い。このように、問題解決プロセスを全体的に見渡す枠組みに欠けている点が、現在の知識ベースシステムの問題点の一つであると考えることができる。以下では、問題解決プロセスのモデリングの方法論の基礎として、全体作業の分割という観点からみたタスク概念の導入を提案し、その効果に関する考察を述べる。

4. 問題解決プロセスのモデルとタスク

これまでに述べてきたように、現在の知識ベースシステムが一般的に取り扱う知識には、対象世界の情報構造モデルを記述する知識と、この情報モデル上で情報操作を行うための知識とが存在するが、これらの内で情報操作に係わる知識の多くは、局所的な情報操作の形式的・汎用的枠組みだけを記述するものであり、かならずしも人間にとって意味の明確な情報処理作業単位を特定するものではなく、したがってシステムの達成すべき問題解決プロセスの構造を陽にモデル化するものでもなかった。見方を変えるならば、ここでは問題解決の形式は記述されているが、その意味は記述されていないと考えることもできるであろう。

問題解決プロセスをモデル化する一つの方法として、全体としてのシステムの目的と、局所的・形式的な情報操作記述との中間に、人間からみて意味のつかみ易い部分的問題解決作業を想定することが考えられる。このような中間段階の部分的問題解決作業を明示的に示す手段としてタスクを想定することができる。これはタスク概念を用いて問題解決プロセスのモデル化を試みるものであり、結果として得られる問題解決プロセスのモデルとは、全体的問題解決作業が、どのような部分的作業を実

行するタスクから成り立っており、それらのタスクがどのような順序で適用されていくかを明示的に記述したものであるということが出来る。ここでタスクの適用順序の決定は、従来の知識ベースシステムでいわれてきたところの、問題解決の戦略を定めるメタ知識に相当するとみることが出来るであろう。このように考えるならば、例えば HEARSAY などで採用された黒板モデルは、一つの知識源が実行する作業という形で、ある意味ではタスクに近い概念を実現する一つの枠組みとなっていると考えることも出来る。とりわけ HEARSAY の場合、異なったレベルの信号解釈の知識をレベルごとに明確に分離しているという意味で、人間にも理解し易い作業単位への分割を行っている。これは実質上、タスクと捉えることも可能であろう。

このように、タスクを全体作業を構成する部分的な作業であると位置づけるならば、次に問題となるのはそのようなタスクを決定するために何をすれば良いかである。ここで有効と考えられる一つの方法は、人間自身の問題解決プロセスの検討であろう。知識ベースシステムの場合、実際に問題とする領域の問題を最もうまく解決できるのは、その領域の専門家であると考えられる。したがって、領域の専門家の問題解決プロセスを理解し、それを適切な形で部分的作業に分解することができるならば、そのようにして求められた部分作業は、タスクとしても有効なものであると考えることができる。このようなものとして求められたタスクは、対象領域の問題解決作業を、人間にとって意味のある作業に分割したものであり、その全体に対する役割もまた明かなものとなるであろう。また、全体としての問題解決を適切に行うために、それらの起動の制御を考えることのできる作業単位ともなっている。

このような部分作業ないしはタスクへの分割は、かならずしも一段だけに限定されるわけではなく、場合によっては一旦分割された個々のタスクを、さらに詳細なタスクに分割することがより有効な場合も想定可能である。(図3)しかしながらこのような場合にも、分割されたタスクが全体的な問題解決とどのようなかかわりを持つかは、明確な形で整理されていると期待することが出来るであろう。

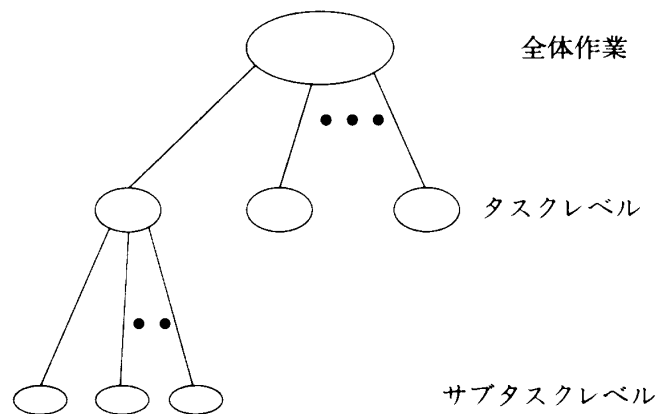


図3 タスクの階層性

以下ではこのような形で全体的問題解決をタスクに分割することによって、どのような利点が期待できるか、また、そのような枠組みを実現するためにどのような問題があるかについて考察を行うこととする。

5. タスクを想定する効果

これまでに述べてきたように、タスクの概念は、人間にとっても意味のあるレベルでの、全体作業の分割を考えることにより、問題解決プロセスの構造を明らかにしようとするものである。ここでは問題解決のための処理形式だけではなく、問題解決の意味までもが記述されることとなる。このような作業単位を想定する結果、システムを構築する上でのいくつかの効果が期待できるが、以下ではその内のいくつかについて考察する。

5.1 知識表現と知識獲得

タスクの考え方によれば、問題解決プロセスは必ずしも均一な作業ではなく、それぞれ一定の役割を持つ独立した部分的な作業の、組織された集合と考えられる。このように考えるならば、個々の部分的作業としてのタスクの実行に関わる知識は、他のタスクとは独立した、そのタスクに固有なものであると考えることができるのであり、その知識表現や推論方法も、タスクごとに異なったものであってもよいことになる。これまでの知識ベースシステムでは、システム全体にわたって知識表現と推論の方法は均一なものと考えられており、問題とする作業がどのようなものであっても、問題解決のための知識をその共通の枠組みの中で記述する必要があった。このことはともすれば、対象領域の専門家が知識を定義・保守していく上で大きな問題を引き起こしてきたといえるであろう。

実際、例えば Chandrasekaran の提唱する generic task の一つである hierarchical classification という形での問題解決が有効な場合には、概念の階層関係と、概念の分類に有効な属性の特徴とが明確にされていれば、分類タスクを実行する上では十分な情報が与えられるのであり、これを無理にルールの変換しなければならない必然性はほとんどないといえる。むしろ、専門家にその知識をルールの形で表現することを強制することが、専門家自身による知識ベースの保守を困難にしているという面すらある。同様に、概念の判定基準が決定可能な場合に、そのような判定基準を直接に記述することをゆるすクライテリアテーブルを知識表現とすることにより、専門家自身が知識ベースの保守を行うことが可能となった例が上野らによって報告されているし、また、いくつかの文献では、確信度に基づく仮説間の確からしさの順序づけを行う場合に、観測事象と仮説との関連の強さをテーブルの形で整理したものが、専門家に取って保守の容易な知識表現となることも報告されている。

知識獲得の観点からすると、タスクという形で作業が局所化されることにより、タスクの実行に必要なとされる情報の範囲もまた、より明確に限定されるという効果も期待できる。例えば、問題とする知識がどのような状況の下で推論を行うためのものであるかがわかれば、そのような知識について、どのような種類の情報を参照することにより、どのような種類の結論を導こうとしているかが決定できる可能性があるが、このような枠組みが決定できるならば、これに基づいて不足していると考えら

れる知識の指摘などが可能となることから、知識獲得に役立たせることのできる可能性もある。

このような、タスクに特殊化された知識を用いる推論の実現に関しては、二通りの考え方が存在する。一つの考え方は、推論システム自体もタスクに特殊化する事により、実行効率の面でも高い効率を得ようとする考え方であり、もう一つの考え方は、技術的にすでに確立ししているプロダクションシステムなどの上に、タスクに特殊化した推論系を構築しようとする考え方である。

5. 2 タスクの再利用可能性とシステム開発の効率化

全体的問題解決作業を部分的なタスクに分解しようとする場合、いくつかのタスクについては、さまざまな種類の問題に共通して利用可能なものが存在することに気づく。例えば機械の故障箇所を同定する上で、入出力の連鎖をたどることにより、入出力に不整合のある要素を発見するタスクとか、故障に関するさまざまな仮説を生成したり、目的とする機能を実現するための機構の候補を生成するタスク、故障仮説や、機構の候補について、システムの挙動を予測するタスクなどは、きわめて広い適用範囲を持つものであり、さまざまな領域ののさまざまな種類の問題に対して適用可能なタスクであると考えることができる。

このようなタスクを整理してその実装のためのノウハウを蓄積することにより、いわばシステム構築のための部品として用いることができる可能性が考えられる。これは広い意味でのソフトウェアの再利用の枠組みを与えるものであり、結果としてシステム開発の効率化が期待できるであろう。Chandrasekaran の提案した generic task は、このような、再利用可能なソフトウェア資源としてのタスクを想定していたと考えることができる。ただし、再利用可能なソフトウェア資源としてのタスクという観点から考えるならば、Chandrasekaran の提案したものよりはるかに多様なものが、有効なソフトウェア資源と考えられるであろう。

ソフトウェア再利用という観点からタスクを見る場合、タスクを実現するための推論機構そのものが、プログラムの一種として再利用可能かどうかに興味を引く。ここでは対象領域の情報構造の記述方法や、領域の物殊性などと独立に、同一の推論機構が適用できるかが問題となるが、このことに関してはさらに検討が必要であろう。

5. 3 システム構築指針の明確化

先にも述べたように、知識ベースの中で定義されたルールや局所的な手続きは、問題独立な情報操作記述の枠組みを提供するのみであり、それ自体ではシステム構築に関する指針を明らかに示すものではない。これに対してタスクは、具体的な問題解決のプロセスのモデルを提供するという意味で、より具体的なシステム構築指針を提供する可能性を持っている。

既に対象領域の問題解決の構造を明らかにし、個々のタスクに分解しようとする努力自体、システム構築指針を明らかにしていく努力の一つと考えることができるが、さらに、前節で述べたように、タスクの中にさまざまな問題解決に共通に利用可能なものがあるとするならば、与えられた問題に対するこのようなタスクの適用可能性を検討することにより、より明確なシステム構築指針が明かとなる可能性もある。この考え方をさらに進めれば、これまで診断型、設計型、制御型などと分類されて

いた知識ベースシステムについて、そのそれぞれを構成するタスクの面から再度分類をし直しすることも考えられるであろう。小林の提案する基本タスクは、おおむねこのような考え方に従うものとも考えることもできる。問題の型によって、それを構築するタスクの構造に共通性が見いだされるならば、そのようなタスク構造は、システム構築のための強力な指針を与えるものとなる可能性があると期待できるであろう。

5. 4 人間・機械協調のあり方に関する指針

先にも述べたように、有効なタスクを決定する一つの方法は、領域の専門家の問題解決プロセスの構造をしらべることであるが、これはまた人間の問題解決プロセスを、タスクを用いてモデル化する試みと考えることもできる。ところで、このような形で問題解決過程がモデル化されると、そこで要求されるタスクのすべてを機械に任せてしまうのではなく、人間と機械との間で作業を分担し合う人間・機械協調系によって、全体としての問題解決作業を実現する方法も考えられる。ここで、各々のタスクの性格を良く理解できるならば、それぞれのタスクについて、人間が行った方が有利であるか、あるいは機械にまかせた方が有利であるかの判断ができる可能性がある。このようなことが可能となるならば、人間・機械協調システムを構築する上で、そのあり方に関する貴重な指針が得られるであろう。

さらに、このような形で、人間と機械との間の情報処理の役割分担が明確になると、人間から機械へ、あるいは逆に機械から人間へ、どのような情報がどのように伝達され、それがどのように用いられるかも明らかとなってくる。このことは、ヒューマンインターフェースの設計においても有益な情報となるであろう。

5. 5 従来型システムとの統合化

タスクの概念の導入はまた、演繹的推論に基づかない情報処理作業を、自然な形で知識ベースシステムに取り込むことを可能とする。タスクは個々に独立したものであり、タスク間の関係を考える限りでは、各々のタスクがどのような手段を用いて実現されているかは問題とならない。知識ベースシステムに関して従来からいわれてきたことに、数値計算やデータベース検索など、これまでに蓄積されてきた従来からの手続き型のソフトウェア資産との統合の問題があるが、これら従来のソフトウェア資産を、タスクの形で呼び出して利用する枠組みを用意しておくことができるならば、知識ベースシステム全体として整合性の高い枠組みで従来のソフトウェア資産を統合することができる。この場合に、タスク呼び出しの環境設定を宣言的に定義しておき、実際の設定を演繹推論によって行なうようにできれば、手続きの呼び出しのための、操作性の良い手段が実現できるであろう。

6. タスクを考える上での課題

以上、知識ベースシステムに、問題解決プロセスのモデル記述手段としてタスク概念を持ち込むことにより、推論戦略までを含めたトータルな情報モデル化の可能性が開かれてきたことを述べた上で、タスク概念の導入によって期待される利点について考察してきたが、実際にタスク概念を生かしたシ

システム構築方法論を確立するためには、今後いくつかの問題点を解決して行く必要がある。

その第一は、そもそも与えられた課題を、常にタスクという形に分解できるかどうかである。この問題については、現段階では今後経験を積んでいくという以上のことはいえないが、少なくとも、問題を直接にルールやフレームなどの表現手段を用いて記述するよりは、タスクという中間的段階を想定する方が、問題の整理がやりやすい分野が相当程度存在すると思われることができる。

第二の問題として、タスクとしてどのようなレベルのものを想定すれば良いのかという問題がある。元々の問題をタスクに分解するしかたは、一義的に定まることが保証されているわけではないし、また、一旦タスクに分解されたとして、個々のタスクがさらに細かな処理単位としてのサブタスクに分解できることもあると考えられる。問題を理解しやすく、しかもタスクの再利用に有利な記述レベルがどのようなものであるかを検討する必要があるだろう。

別の問題として、タスクにどの程度の汎用性あるいは再利用可能性を期待できるかという問題もある。理想的には、generic task で主張されたような、少数のタスクがあれば、ほとんどの問題が記述できるという状況が実現できれば良いのであるが、実際には相当多様なタスクを用意しておかなければ、さまざまな問題に対処しきれないと考えられるし、また、推論プロセスという観点からは同種のタスクであっても、異なった情報構造モデルの上に適用しようとする場合、同一の知識表現と推論機構が適用できない場合が生じる可能性もある。この問題は、タスクに、問題整理とシステム設計指針を得るという利点だけを期待するか、実装面での効率化まで期待するかの問題とも考えられる。

これらの問題に加えて、実際にタスクという概念を実現するシステムとして、どのようなものかを考えるべきかという問題も残されている。以下では具体的なシステムを実装する上での課題について述べる。

7. タスク処理のためのシステム実装の枠組み

実際にタスク概念を導入したシステムを実現するには、その実装のための枠組みとして、どのようなものが要求されるかを明らかにしていく必要がある。先にも述べたように、タスクの概念を有効に生かすためには、同じルールに基づく演繹推論を行うタスクであっても、その目的によっては異なった推論系を採用する方が適切な場合もあると考えられる。例えば、仮説生成問題と、システムの構造を変更する際に変更部分を決定する問題とでは、演繹推論の方法を変えた方が良い可能性がある。ルール適用にあたっての競合解消戦略や探索の枝刈り、縦型探索か横型探索かの選択、分野依存のヒューリスティックの組み込みを許すかなど、さまざまな選択岐が存在している。システム実装の観点からは、タスクごとにこれらの選択が可能となる枠組みが要求される。

直接にルールを使用しない演繹推論タスクをどのように実現するかは、またもう一つの問題である。このようなタスクの実現には、大きく二つの方法が考えられるであろう。その第一は、タスクに相当する推論機構を独立して作成する方法であり、第二は、推論をなんらかの形でルール型の演繹推論に変換する方法である。独立した推論機構を採用する場合、タスクの性格に応じて最適な実行効率を達

成する推論手続きを採用できる可能性がある反面、推論機構の信頼性という面では疑問を残す場合もあるであろう。どちらの方法を採用するかについて、推論機構そのものの再利用可能性などを含めた、個別の判断が必要となろう。

全体作業をタスクに分解するという観点からは、タスク間の情報交換をどのように行うかの問題、およびタスクの起動に関するメタレベル推論をどのように実現するかが問題となる。既に黑板モデルでは、このようなタスク間の情報交換と、タスク起動に関するメタレベル推論の一部を実現していると考えて良い。しかしながら、作業記憶を唯一の情報交換の媒体とすることの可否、作業記憶が単一のものであって良いかどうか、作業記憶内容のインテグリティ保証をどのように考えるか、メタレベル推論の実現機構など、今後検討すべき多くの課題が残されている。特にメタレベル推論の実現に関しては、メタルールに代表される演繹推論系で実現するのか、あるいは手続きとして記述するのかという問題や、メタレベル推論自体を再びタスクの形で表現するかどうかの問題についても検討する必要があるであろう。

一方、タスクの実行に関しては、タスク間の並列処理可能性、入出力待ちなどを生じる可能性のあるタスクの取り扱いなどの観点から、OS レベルとも関連した効率の良いタスク実行機能の実現が望まれる所である。

8. まとめ

タスクの概念は、知識ベースシステムの中で定義される知識として、問題解決プロセスのモデルを陽に記述する方法論を導入する一つ的手段を提供するものと考えることができる。従来知識ベースシステムで意識されてきた、情報記述構造のモデル、演繹推論のモデルに問題解決プロセスモデルを加えることにより、広い範囲にわたる、トータルな情報システムのモデル化が可能となる。その結果として、知識ベースシステムの特徴が一層明かなものとなり、システム構築の方法論が整理されるとともに、効率の良いシステムの実現が可能となることが期待できるであろう。この意味でタスクは、知識ベースシステムのあり方に関する具体的指針を与えるものであり、いわばデータベースで議論されてきたスキーマの役割を果たすものであると考えることもできる。

もちろんタスクの考え方をういたシステム構築方法論の確立のためには、先に述べたようにさまざまな問題点も残されているのであり、今後、タスクの概念を意識したシステム構築の経験を積み、全体的な問題解決に関わるタスクの構造に関する事例を広く収集することにより、その適用の可能性を明らかにしていく必要がある。今後、これらの問題の検討が進み、新しい知識ベースシステム開発の方法論が確立されていくことを期待したい。

参考文献

1. Chandrasekaran, B., Generic tasks in knowledge-based reasoning : High level building blocks for expert system design., IEEE Expert., pp. 23-30, Fall, 1986.

- 2 . Ueno, H., and et. al., Design of a Criteria-based Rheumatology Consulting System, Proc. 3rd MEDINFO Tkyo, p. 1316-1320, 1980.
- 3 . 三宅浩之他、問診を入口とする医療支援コンサルテーションシステム DOCTORS の知識エディタの機能、第 6 回医療情報学連合大会予稿集、p. 65-68、1986。
- 4 . 小林重信、知識システム技術の現状と将来、計測と制御、vol. 27、no. 10、pp. 859-868、1988。
- 5 . Erman, L. D., and et. al., The HEARSAY-II speech understandin system: Integrating knowledge to resolve uncertainty., Computing Survays, vol. 12, no. 2, pp. 213-253, 1980.
- 6 . Ed. by Barr, A., and Feigenbaum, E. A., The Handbook of Artificial Intelligence, William Kaufmann, Inc. 1981-1982.

研 究 論 文

機械概念設計プロセスのモデリング

Process Modeling of Mechanical Conceptual Design

学術情報センター 小山 照夫

1. はじめに

機械設計における CAD のインテリジェント化に関連して、人間が行っている、概念設計を始めとする「上流」の設計プロセスの解明が一つのテーマとなっている。この問題は広い意味では設計論の問題として、また、やや狭い意味では機械の機能とその構造・動作とを関連づける問題として捉えることができるであろう。すなわち要求される機能を、どのような具体的構造とその動作によって実現できるかが問題となる。

実際に人間が設計作業を行う場合を考えるならば、訓練を積んだ設計者は、与えられた機能を実現し得る有力な構造の具体的候補を複数個、ほとんど即時に想定することができる。知的 CAD の立場からは、このように、達成すべき機能を指定された場合に、指定された機能に対する実現構造の候補を生成できるシステムが構成できるかどうか、また、そのようなシステムを構成しようとする場合、どのような情報をどのように取り扱う必要があるかを明らかにする必要がある。

実際の人間の設計過程を考えてみると、第一に設計経験の浅い設計者と、ある程度設計経験を積んだ設計者の間で、設計の能力に大きな違いがあること、ある分野の設計に訓練を積んだ設計者に別の分野の設計問題を与えた場合、本来の専門分野ほどには優れた能力を発揮できないが、それでもその分野の初心者よりは高い能力を示すこともあることなどから、機能と構造を関連付ける問題に関して、ある種の、学習によって獲得された知識としての情報が重要な役割を果たしていることはほぼ疑いのない事実であろう。問題は、設計プロセスの中でどのような種類の情報がどのように関与しているかである。この問題をシステムの観点からみると、機械 CAD を考えるにあたって、どのようなデータベースないしは知識ベースが、設計のどの段階でどのように関与し得るかの問題と考えることもできる。

本論文では、まず、機械設計の問題を複雑なものとしている、構造と機能との対応の不連続性について述べた後、設計者の情報処理モデルとして、説明による学習に見られる情報構造と類似した情報モデルを提唱し、このモデルに基づく概念設計における推論プロセス及び知識の学習に関する課題について考察を行う。

2. 構造と実現機能の不連続性

設計とは与えられた機能を実現する構造を決定する問題であるという観点から、機械の機能とその実現構造との関係を考えるにあたって、機械の機能と構造との間に関数論的対応関係を想定しようとする考え方がある。例えば吉川は、設計論の立場から、実現構造空間と機能空間をともにハウスドルフ空間と考え、この二つの空間の間に関数関係の存在することを示唆している。機械設計におけるこのような定式化は、確かに魅力のあるものであるが、実際には、構造の連続性が必ずしも機能の連続性を保証しないという問題がある。このことを言い替えるならば、構造空間に導入する位相と機能空間に導入する位相との間で準同型性を保つ対応関係を定めることが困難であるといつてよい。

例えば皿と湯呑と壺は、形状空間上のトポロジーとして同等であり、その部分的な寸法を連続的に変化させていくことにより、中間的な形状を経て相互に入れ替わり得るから、構造空間の位相の上ではある種の連続性が保たれているといえるであろう。にもかかわらず、例えば水を飲む道具というような特定の機能を想定する場合、皿では水をこぼしやすいし、壺では水を飲むことが不自由であるなど、ある部分で機能的な不連続性を生じていると考えざるをえない。同様に、絶対的寸法についても、片手で容易につかめる、あるいは一回に飲む水の量に適合する容量を持つなど、指定された機能を実現するための適切な範囲が存在しており、あまりにも大きすぎるものや小さすぎるものでは、水を飲むという機能を十分に発揮できない場合がある。機械設計の場合一般に、特定の機能に対する一つの実現構造が与えられた場合には、構造空間におけるその限られた近傍については類似の機能を実現することが保証されるが、ある所で機能的な不連続性が生じるのは避けられないと考えられる。

逆にこれと対照的な例として、全く異なった構造でありながら、きわめて近い機能を提供するものとも考えることもできる。たとえば機械式の時計と電子式の時計が、構造面では全く異なったものでありながら、機能としてはほぼ完全に等価な機能を実現している例や、一般の船舶とホバークラフトのように、部分的にかなり異なった構造を持つ機能が、結果として類似の機能を実現しているものもあ

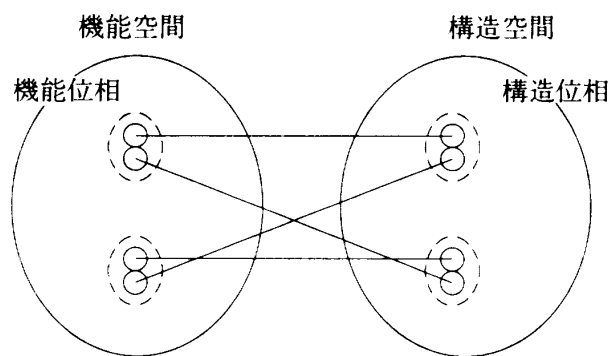


図1 機能空間と構造空間の非準同型性

る。図1は、このような構造と機能の対応の連続性を示すものである。

機械の概念設計の立場からすると、構造的に相当程度異なったものであっても、目的とする機能を実現する可能性のあるものならば、一つの構造の候補として検討する余地を残して置きたいというのが実状であろう。そこで問題は、可能性のある機能実現構造を幅広く想定するために、どのような情報構造が関係してくるかである。

3. 機能を実現する構造決定に関わる情報のモデル

機械の実現すべき機能を考える場合にまず考えなければならないこととして、多くの場合、機能そのものは決して操作し易い概念ではないという問題がある。例えば「輸送」という機能を純粹抽象的に考えることはしばしば困難を伴うであろう。一方で、機能は機械の振る舞いによって実現できると考えるとき、機械の振る舞い自体は、抽象的な機能よりは、はるかに操作し易い概念となり得る。さきに述べた「輸送」についていうならば、「輸送の対象を固定した物体を、指定された二つの空間座標の間を移動させること」と言い替えることにより、具体的な実現構造を想定することがはるかに容易なものになる。

このような振る舞いの記述では、全体としてのマクロな振る舞いを、より基本的なミクロの振る舞いの組合せとして達成する、いわば振る舞いの分解が可能となる場合がある。このことは、全体として当初の目的を達成する機械を、部分的な構成要素に分解する指針が与えられていると考えることもできるのであり、機械設計にあたって、機能とそれを実現する機械の振る舞いとの対応、また、そのような振る舞いがどのような部分的な振る舞いによって合成されるかに関する選択岐をあらかじめ用意することができるならば、目的とする機能を実現する機構の候補を生成できる可能性があることを示している。このような振る舞いの分解は、いくつかのステップに分かれていても良い。つまり、いくつかの振る舞いの合成として実現される相対的にマクロな振る舞いが、よりマクロな振る舞いの構成要素となることもありうる。(図2)

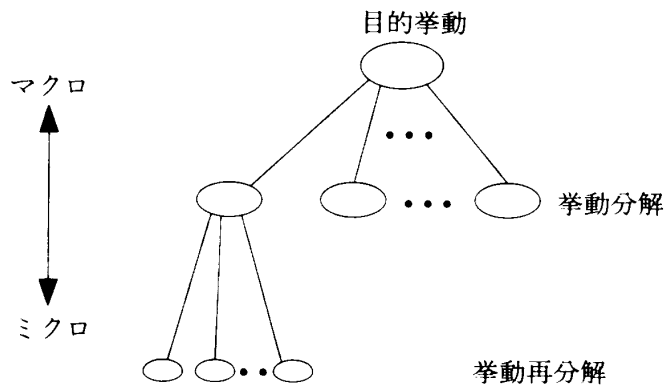


図2 目的挙動の分解

このように考えるならば、機能を実現する構造の決定に関しては、まず、指定された機能を、機械のどのようなマクロな振る舞いによって実現できるかを決定し、さらにこのようにして求められた振る舞いが、機械のさまざまな部分によって実現される相対的にマクロな振る舞いの集合としてどのように分解できるかを決定することにより、人間が行なっている概念設計に近い機能を実現できると考えられる。なお、実際には、抽象的なレベルでの機械の振る舞いの記述は必ずしも容易でない場合も考えられるのであり、多くの場合には、機能そのものが部分的な振る舞いの合成されたものとして実現されると考える方が实际的であろう。

3. 1 機能の部分的振る舞いへの分解

これまでに述べてきたことに従い、以下では指定された機能を部分的な振る舞いの組合せに還元する方法について考察する。これは一つの機能が与えられた場合に、それがいくつかの機械的要素の振る舞いに分解される場合、それぞれの振る舞いを実現しうる機械要素を総合するために考え得る構造を生成しようとするに相当する。ここでは、各部分構造なり機械要素なりについて、具体化された「もの」を想定するのではなく、全体の機能を達成するための部分的な振る舞いのみを与えようとするものであり、それぞれの要素の振る舞いがどのような具体的構造によって実現されるかは、別途管理されていると考えることにする。

この問題を考えるにあたっての一つの例として、「輸送」という一つの機能を考えてみよう。われわれは既に輸送手段として、列車や自動車、航空機など、いくつかの具体的機能実現の方法を知っている。先ほどから述べてきた論旨に従えば、これらの、一見異なった機能の実現方法の間に共通する、基本的な機械の振る舞いの組合せというものが存在し得るかどうかを問題にしなければならない。ここで輸送という機能を一連の部分的な振る舞いに分解する方法として例えば、「地面（固定面）との摩擦の軽減」、「駆動」、「運動方向制御」、「運搬物の輸送機械本体への固定」という4つの部分的な振る舞いを想定する事ができる。ここで「地面との摩擦の軽減」は、より一般化して「固体間の摩擦の軽減」とすることも考えられるし、「運搬物の輸送機械本体への固定」は「固体と固体の固定」としても良い。われわれ自身の経験からは、多くの輸送機械についてこのような振る舞いが、機械の部分構造によって実現されていると考えることができる。

ここで注意すべきは、機能を実現する振る舞いの記述が一意ではないことである。例えば「個体間の摩擦の軽減」という一般的な記述は「車輪の使用」というより具体的な記述に置き換えてみたとしても、依然としてある程度の有用性を保っている。「個体間の間の摩擦の軽減」という表現に対しては、それを実現する具体的な手段として、「ころがり」、「浮力」など、複数の候補を想定することができ、さらには「浮力」を実現する手段として「揚力」、「静圧」、「動圧」などを想起することができる反面、特定の問題に対して実際にどのような手段を用いるのが有利であるのか、詳細な実現機構をどのようにすれば良いのかについては、多くの検討を要するであろう。結局、想起する振る舞いの抽象度の相違は、抽象度の高い記述からはより多くの選択岐が導き得るのに対して、より具体的な記述からは、より簡単な推論によって具体的な構造決定が可能であるという点にあるといえる。

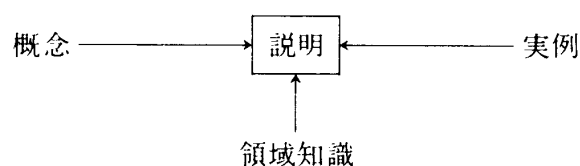
ここで、指定された機能に対してどのような抽象度のレベルでの部分的な振る舞いの組み合わせを想定するかは、興味深い問題を提示する。ここでは、熟練度の相違により、選択される抽象度に相違が現れてくるか、与えられた問題のどのような側面が抽象度に影響を及ぼすかなどの問題が興味を引く。いずれにせよ、ここでは振る舞いの組み合わせが目的とする機能に対して適切なものであること、および、要素として想定された各振る舞いが、具体的な機構として実現可能性を持ち、相互に組み合わせ可能であることが要求される。

この組み合わせの妥当性の問題と、実現可能性の問題を、説明による学習の枠組みと比較してみよう。ここでは、指定された機能を「学習すべき概念」に、機能とそれを実現するための最終的な振る舞いの組み合わせとの対応付けを「説明」に、実際の設計事例を「実例」にそれぞれ対応づけることが考えられる。説明による学習では、このほかに、学習すべき概念と実例とが説明によって正しく対応づけられており、しかも説明が操作可能性の条件を満足していることを保証するための「領域知識」を想定することになる。機械設計において操作可能性に相当するものは、提案された振る舞いの実現可能性と考えることができるが、これは、実際に提案された振る舞いの組み合わせについて、その各々の振る舞いを実現する機構を決定することが可能であり、かつそれらの機構の間に必要な連携を確保できる可能性が証明されることであると考えられる。この証明のためには、振る舞いと機構との間の対応、機構と機構との間の接続可能性に関する情報が存在している必要があるが、少なくともその一部はエンジニアリングデータベースとして実現が可能である。エンジニアリングデータベースの内容としては、データベースに蓄積される各種機構によって実現可能な振る舞いに加え、機構をどこまで変更することが可能であるか、各部分の形状・寸法の変更が最終的な振る舞いにどのように影響を及ぼすか、他の機構との連携の可能性を判定する条件、その機構を適用する上での各種制約、外的な環境に対する適合性などが管理されていると考えることができる。

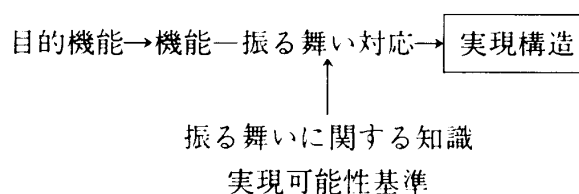
先に述べた、機構と振る舞い（によって達成される機能）との間に不連続性のあることを考えるならば、振る舞いと機構との対応は一对一ではなく、多義性があると考えられる。言い換えれば、同一の機構が複数の振る舞いに対応づけられていてもかまわないし、逆に一つの振る舞いに対応する、まったく異なった機構が複数存在していてもかまわない。

機械設計における領域知識には、エンジニアリングデータベースによって管理されるものの他に、指定された機能をどのような振る舞いの組み合わせによって実現することができるか、抽象的なレベルで記述された振る舞いを、どのような具体的なレベルの振る舞いに対応づけるか、相対的にマクロな振る舞いをどのようなマイクロな振る舞いの組み合わせに分解できるかなどの知識も含まれることになる。これらはすべて、領域知識と考えることができる。

一般の説明による学習では、領域知識の存在の仮定の下に、概念とその実例から適切な説明を生成することが目的とされるが、概念設計では、領域知識の存在の仮定の下に、概念とその説明から実例を生成する問題であると考えられるであろう（図3）。以下ではこのような枠組みに基づき、概念設計にかかわる情報の構造について考察する。



Explanation-Based Learning



Conceptual design

図3 説明による学習と概念設計

3.2 概念設計に関連する情報

前節で述べてきたことからして、機械の概念設計に関して少なくとも、目的機能、実現構造、説明、領域知識という4種類の情報が関与してることが予想される。概念設計のプロセスと知識の学習の問題を考えるにあたって、これらのうちで特に問題となるのが、説明および領域知識である。先にも述べたとおり、機械設計における説明とは、目標とする機能を機械の部分的な振る舞いの組み合わせに対応づけたものである。ここに記述される部分的な振る舞いは、それが機械的に実現される可能性を持たなければならないが、必ずしもそれ自身が具体的な機械要素と直接に関連を持つ必要はない。説明の中で参照される振る舞いは、いくつかの具体的な振る舞いに共通する、抽象的なレベルのものであっても良いし、いくつかのより基本的な振る舞いが合成されたマクロな振る舞いであってもかまわない。前者の例としては、「(個体間の摩擦の低減のために) 個体間に空間的ギャップを確保する」という抽象的な振る舞いを具体化するものとして、「空気浮上(=ホバリング)」、「磁気浮上」、「揚力」、「動圧利用」などのより具体的な振る舞いを挙げることができるし、後者の例としては「物体の、指定された二点間の移動」という振る舞いを、「物体と地面との摩擦の低減」、「駆動」、「方向制御」、「制動」という部分的な振る舞いに分解する例を挙げることができる。このようなさまざまな振る舞いの間の関係は、領域知識の一部を構成するものであり、機能の説明の中で参照される振る舞いは、これらの領域知識の適用によって、最終的に実現可能な部分構造ないしはその組み合わせに対応づけられる振

る舞いの集合に変換できるものでなければならない。

以上のことから、領域知識として要求されるものには、1. 機能を振る舞いの組み合わせに対応づける知識、2. 抽象的な振る舞いをより具体的な振る舞いに変換する知識、3. マクロな振る舞いを部分的な振る舞いに分解する知識、4. 最終的に得られた、具体的・基本的な振る舞いを具体的な機械の部分構造に対応づける知識が存在する必要がある。ここで最終的に求められた振る舞いを具体的な機械構造に対応づけられることは、すなわちそのような振る舞いを実現する可能性を持つことを示しているのであり、説明による学習でいうところの操作可能性基準を満たしていると考えることができる。この情報は、先に述べたとおり、少なくともその一部はエンジニアリングデータベースとして実現することが考えられる。

以上述べてきた情報の構造を図4に示す。

3. 3 概念設計における推語と学習

図4に示す情報モデルを想定した上で、このような情報モデルの上での推論及び学習の問題について考察する。

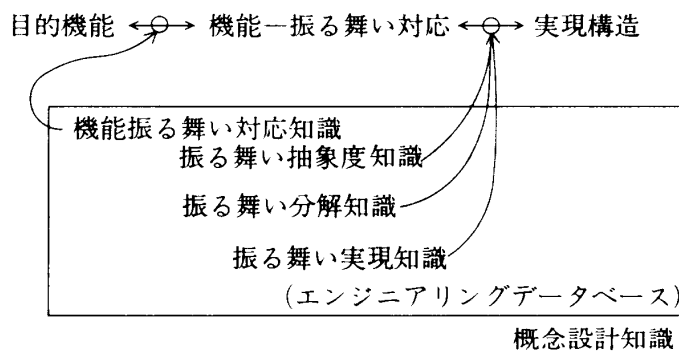


図4 概念設計に関わる知識情報

機械概念設計の目的は、指定された機能を実現する十分な可能性を持つ機械構造の候補を生成することであると考えられる。先に述べた情報構造を前提として考えるならば、このプロセスは一種の探索プロセスとみなすことができるのであり、機能に対する説明を求め、そこに指定される振る舞いを分解、具体化していくことにより、詳細化された振る舞いの列が生成される。このようにして求められた振る舞いの全てが、エンジニアリングデータベースを介して、具体的な構造に対応づけられ、かつ、それらの間の適切な連携が可能であることが証明されれば、機能を実現する一つの構造の候補が得られたことになる。ここでは1. 機能と振る舞いの対応付け(説明)、2. マクロな機能の分解、3. 抽象的な機能の具体化、4. 振る舞いとその具体的な実現機構との対応付けという、4つのレベルに

ついて、選択岐の存在が考えられるのであり、同時にその可能な組合せの数だけ、異なった代替案が存在することになる。

機能を振る舞いに対応付ける説明や、実際の振る舞いを実現するための機構はさまざまな抽象度のレベルで考えることが可能である。説明の段階で、極めて詳細な振る舞いが指定してあるならば、機能分解・具体化の選択岐は減少することとなり、最終的に生成可能な機構の総数は相対的に少なくなるであろう。また、説明の抽象度が高くても、エンジニアリングデータベースの中に、そのような振る舞いを実現する、高度の機構があらかじめ用意されているならば、煩雑な探索プロセスを省略して、一気に可能性のある機構の候補を求めることも可能である。逆に、機能に対応付けられる振る舞いの抽象度が高く、機能を分解・具体化するための領域知識が豊富に用意されており、さらに詳細化された振る舞いを実現するための具体的な機構の代替案も多数用意されているならば、システムが生成することのできる構造の候補は非常に多数のものとなりうる。

一般の探索問題と同様に、概念設計の問題においても生成可能な代替案のすべてが同様に望ましいことは希であると考えられるから、なんらかの探索のためのヒューリスティックが要求されることとなる。実際の人間の設計プロセスでは、まず一般に有力と考えられる少数の候補について検討を行い、これらの中に充分満足できるものが存在しない場合に限って、より広い範囲の探索を行うように見える。これを先の情報モデルに照らし合わせて考えるならば、まず第一には、抽象度の高い振る舞いを記述した説明と、そのような振る舞いを直接に実現する機構の組合せが求められるならば、そのような機構を優先的に検討することが考えられる。次には、より詳細度の高い説明が存在するならば、そのような説明を具体的な機構に対応付ける試みが行なわれ、最後に抽象度の高い説明を、さまざまな形に分解・具体化して、最終的に得られた振る舞いの列を具体的な機構に対応付ける試みが行なわれると考えられるであろう。このようにして、一般的に有効性が高いと期待されるものを優先的に検討し、そのような方法がうまく行かない場合に限り、より特殊な構造を幅広く検討することが可能となると期待できる。実際にはこの他に、実現すべき振る舞いに対応する具体的な機構の代替案を選択する上でのヒューリスティックも必要となると考えられるし、さらに機構の間の連携可能性など、各種の条件に依存する詳細なヒューリスティックの検討も必要となるであろう。

概念設計にプロセスを、説明による学習と対応づけることにより、設計知識の取得に関しても興味深い仮説を立てることができる。先の定式化では、設計者の知識は、大きく1. 機能と振る舞いとを関連付ける説明、2. 抽象的な振る舞いを具体的なものに変換する領域知識、3. マクロな振る舞いをミクロな振る舞いの列に対応付ける領域知識、4. 振る舞いとその実現機構を対応づけるためのエンジニアリングデータベースに分けられる。人間が設計過程を学習する場合、その多くは設計の目的とその実現機構との対応を多数経験することによると考えられる。これは基本的には説明による学習そのものであるが、領域知識が完全であることが必ずしも保証されないことからくるいくつかの問題がある。

このような場合には、説明の生成がうまく行かない場合に、エンジニアリングデータベースに対し

て、新しい振る舞いとその実現機構との対応を追加したり、領域知識に対して、振る舞いの具体化・分解の新しい方法を追加するなどの学習が必要となると考えることができる。しかし一般には、この種の学習は、非常に高い任意性を持つ結果、特殊な場合をのぞいては、実質的に意味のある学習を行うことが困難であると考えられる。そこで重要となるのが、目的とする機能と具体的機構の対応に加えて、その設計の背後にある説明までを与えられることにより、機構と振る舞いの対応及び振る舞いの詳細化・分解に関する知識を獲得する学習である。(図5) ここで説明は常に完全なものであることが保証されない場合も考えられるのであり、不完全な領域知識の存在の下で、必ずしも完全であることが保証されない与えられた説明から領域知識を獲得する学習がどのような条件の下でどの程度可能であるかを、今後明らかにしていく必要がある。

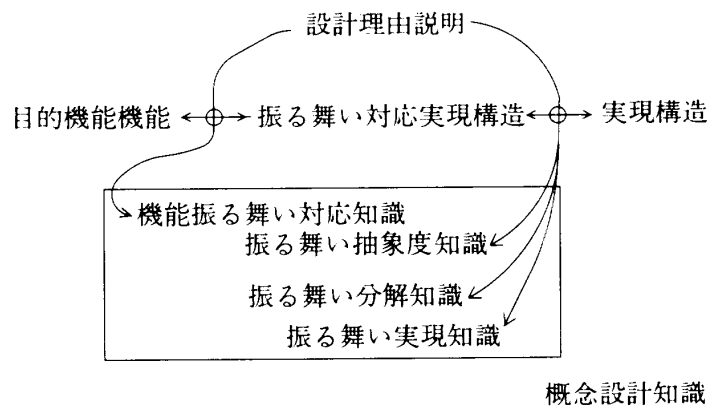


図5 設計理由の説明を与えられた学習

4. まとめ

機械設計における概念設計は、目的とする機能を実現する具体的機構の候補を生成することを目的としている。このような概念設計プロセスを説明による学習との類比で捉えることにより、概念設計にかかわってくる何種類かの知識を整理し、これらの知識を用いた推論プロセスと知識獲得の過程をモデル化することができた。この種の知識の獲得に関しては、目的機能、設計事例、部分的説明を与えられて、不完全な領域知識を補強する形の学習が要求されることが予想されるが、今後、この種の学習が、どのような条件の下でどの程度可能であるかを明らかにしていく必要がある。

参考文献

1. Yoshikawa, H., General Design Theory and a CAD System, Proc. IFIP Man Machine

- Communication in CAD/CAM, ed. by Sata, T. and Warman, E., p. 35-42, 1981.
- 2 . DeJong, G., and Mooney, R., Explanation-Based Learning ; An Alternative View, Machine Learning, vol. 1, pp. 145-176, 1986.
 - 3 . Segre, A. M., On the Operationality/Generality Trade-off in Explanation-Based Learning, Proc. IJCAI, p. 242-248, 1987.

研 究 論 文

大規模関係データベースのための並列処理マシンの一構成法

A Multiprocessor Database Machine for Very Large Relational Database

学術情報センター 濱田 喬*

東京大学大学院工学系研究科 大久保一彦**

要旨

本論文では、将来予想される大規模関係データベース環境に対処し得る、多モジュール構成のデータ処理向き並列計算機システムの一構成法を提案する。まず、処理性能のボトルネックの解消、コストパフォーマンス等を考慮に入れて、データベースマシンの設計理念を提言する。次に、計算機資源の利用率の向上を図ることにより、関係代数処理の並列度を増す処理アルゴリズムを提案する。具体的には、関係代数演算のうちで最も基本的な選択演算と結合演算の並列性の抽出を行い、それらの実装技法を述べる。また、モジュール間相互結合網の検討を行う。並列システムでは、複数モジュール間の通信オーバーヘッドがデータ転送量に比例して顕著となり、この問題を解消するため、負荷分散型バンヤン網における新たなデータ転送制御法を提案することにより、高性能な結合網を構成した。最後に、提案したシステムの性能をシミュレーションとトラフィック解析により評価し、本技法により優れたデータベースマシンを構成できることを確認した。

ABSTRACT

In this paper the authors discussed a database machine architecture for a very large database, which requires heavy processor load in relational database management. The proposed machine is highly-integrated parallel computer system and in such architecture a distribution scheme of processor load has a great influence on the performance of operations.

The authors proposed algorithms for parallel processing of operations and for routing control in multistage interconnection network. The performance of the proposed system is evaluated by simulation and traffic analysis and it is found that the proposed machine can cope with the problems for database environment in the future.

* Hamada, Takashi : National Center for the Science Information System.

** Okubo, Kazuhiko : University of Tokyo.

1. まえがき

関係モデルに基づく関係データベースは、データベースの定義と操作に関する高度なユーザインターフェースを提供することから急速に普及しており、これに伴いデータベースの発展は量的質的の両面において著しい。すなわち、メデータベース規模の増大、利用の高度化が今後一層進むものと予想される。しかしながら、関係データベースはデータ量の増加に伴ってその処理負荷が増大するため、大規模なデータベースに適用する場合には計算機の性能向上が重要な課題となる。この様な将来のデータベース環境に対して、従来のように汎用計算機と磁気ディスク、更にもその上で動作する DBMS で対処するのはもはや困難になってきている。従って、大容量のデータを能率良く蓄え、所要のデータを迅速に引き出したり、複雑な条件を満たすデータを捜し出したりすることに適した専用マシンが必要となるのである。現在までに様々なデータベースマシンが提案されているが、多くの場合、処理に内在する並列度を十分活用することができず、必ずしも十分な性能が得られていない。

本研究は、このような背景に基づき、多モジュール構成のデータ処理向き並列計算機システムにおいて、その計算機資源の利用率向上、すなわち資源に対する処理の負荷分散を図ることにより、処理の並列度を増す高性能なデータベースマシンを構築することを目的とする。そのために新しい並列性の抽出手法を提案するとともに、高速パケット交換網の導入によってこの手法を実現するシステム構成を示し、高速データベース処理を可能とした。

2. マシンアーキテクチャ

2. 1 システム設計の理念

現在または、将来の大規模関係データベースシステムの最も重要な特徴は、莫大なデータ量と多数のユーザであろう。これらの要求に加えて、現在の技術動向等を考慮すると以下のような従来技術に基づいたシステム設計の理念が挙げられる。

① 大記憶を可動ヘッドディスクにより構成する。このタイプのディスクは比較的安価であり、オンラインで大規模な記憶領域を提供し、かつパフォーマンスも悪くないため、最も適切である。

② 情報検索、特に選択処理高速化の一手段としてインデックスを用いるが、このインデックスレベルを最少ディスクアクセス単位、すなわちページに設定する。また、ページサイズを比較的大きくとり、1ページ分のデータを複数のプロセッサにより処理する。この様なページレベルのインデックスと静的クラスタリング技法の併用により処理の高速化が実現される。

③ データベースマシンは、オフ・ディスク方式をとるべきである。この方式ではオン・ディスク方式のものに比べて、比較的少量のデータを2次記憶側からプロセッサ側へ転送しなければならず、処理コストを増加させることになるが、大規模記憶にかかる費用を最小限に抑えることができる長所がある。

④ マシンは MIMD 方式で駆動する。計算機の性能向上、すなわち計算機資源の利用率向上を図る

ために並列処理が必然的であるとともに、この方式が複数ユーザの問い合わせを同時に処理するという問題の本質的な解となるからである。

本論文においては、上記の設計理念に基づいてデータベースマシンシステムの構築を試みる。

2. 2 システムの全体構成

高性能データベースマシンでは、複数台のプロセッサ、多バンク化されたステージング空間及び複数の二次記憶装置からなる並列処理システムの形態をとることによって、並列処理効率の向上を図ることができる。本方式のデータベースマシンも図1に示すように関係代数処理を行なうプロセッサモジュール群、計算の中間結果をステージングするメモリモジュール群、データベースを常時格納しておくディスクモジュール群、更にそれらを結合する相互結合網からなり、制御モジュールが全体の制

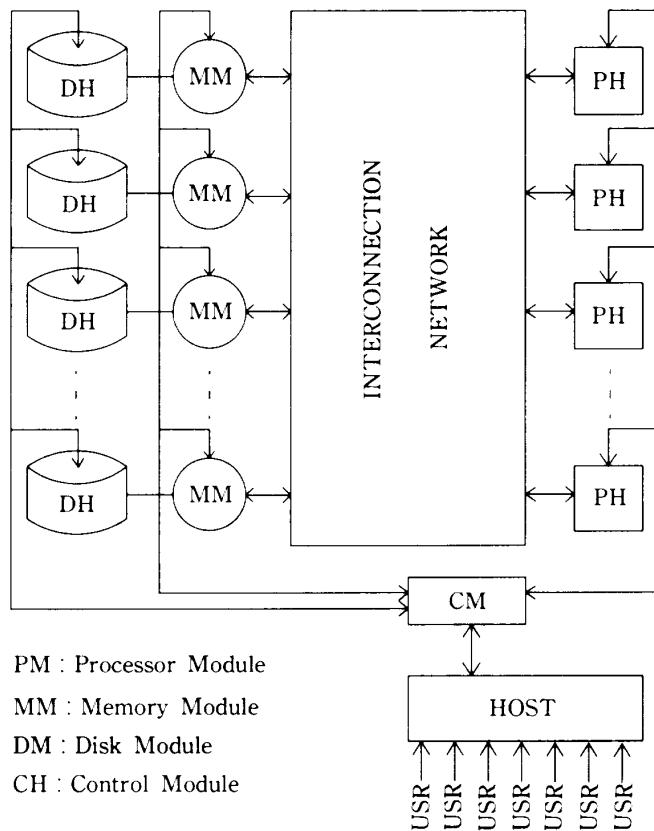


図1 マシンアーキテクチャ

御を司っている。すなわち、バック・エンド・マシンであるデータベースマシンは、大きく分けて上記5つの系から構成され、関係代数処理を高速に行う。さらに制御モジュールは、ユーザに対してインターフェースを提供するホストコンピュータへの窓口の役割を果たしている。以下、バックエンドマシンの5つの系について説明する。

① プロセッサモジュールは、内部に汎用のマイクロプロセッサと入力数のオーダ時間でソートを完了するハードウェア・ソータを備えており、関係代数処理を高速に行なうことができる構成をとっ

ている。

② メモリモジュールは、半導体の RAM で構成され、問い合わせの中間結果を保持しておくものである。また、各メモリモジュールでは、オペランドタプルが演算単位で別々に格納される。

③ ディスクモジュール、いわゆるこの系が大記憶であり、可動ヘッドディスクにより構成される。更に、ディスクモジュールは複数あるが、各々は独立に動作でき、またページ単位にデータにアクセスするものとする。こうして不要な探索動作、すなわち問い合わせに対するディスクアクセス回数の増加を抑えることができる。

④ 相互結合網におけるデータ交換方式としてはパケット交換を採用する。モジュール間でデータを転送する際には、1 タプルが見かけ上の転送単位であるが、実際パケットは、その転送制御の単純化から固定長として転送されるため、1 パケット長に対して大きなタプルは、転送側でパケットごとに分解されて転送され、受信側で組み立てられて再編成されることになる。

⑤ 制御モジュールは、ホストコンピュータと唯一連絡をとっているバックエンドマシン内部の系である。ホストコンピュータは、ユーザからの問い合わせをバックエンドマシンで実行できるように命令ストリーム、すなわち問い合わせ木等に変換し、そのを制御モジュールへ転送する。制御モジュールは、「どのプロセッサ群にどの問い合わせを処理させるか」を決定し、各モジュールへその情報を伝達する。実際、各モジュールとはバスネットワークで結線されていて、これを介して制御を行なうのである。

3. 並列処理の実装技法

一般に、関係代数演算の実行において実際に必要とされるオペランドタプルは全体の極一部である。従って、当該演算に関し、結果に寄与するタプル群のみに演算を施すことにより、不要な探索動作を省略でき、処理を高速化することができる。この様な技法は、当該演算に対してオペランドタプルを互いに独立な処理単位に分割するクラスタリングにより実現される。本研究では、このクラスタリング技法を用いて、関係代数演算のうちもっとも基本的な選択演算と結合演算の並列処理法について述べる。

またこれらの単一演算をシステム内で同時に実行する MIMD 方式の実装技法についても加えて説明する。

3. 1 選択演算

選択演算は、与えられた述語を満たすタプルの集合をディスク内のデータベースから得る作業であり、従って本演算に対してはディスクページを単位とした静的クラスタリングが適用される。選択演算に対するクラスタリングとは可能な選択述語各々に対し、関係をタプル方向に水平に分解し、述語を満たすタプル群を集めてなるべく少数のディスクページに収めること、すなわちディスクページのアクセス回数の最少化と等価である。

本システムでは、非常に良好なクラスタリング特性を与え第2次記憶において好ましい性質をもつ、

多次元クラスタリング技法を採用する。さらに複数あるうちのどのディスクモジュールも同程度にアクセスされることが望ましいため、関係をディスクページ単位にディスク群に対し分散格納する方法をとる。このとき、参照される頻度順に周期的にディスクに割り当てる。

以上は2次記憶系に関してであるが、次にプロセッサ群、メモリ群への負荷の割り当てについて説明する。選択演算のような各タプルを独立に評価できる操作において並列処理を実行する場合の各プロセッサに対する処理負荷の平坦化は比較的簡単で、図2のような転送アルゴリズムを各送信側で実行することで実現される。また、各メモリモジュールに対する問い合わせ処理の中間結果の平均的割り当ても、このアルゴリズムにより可能となり、計算機資源の利用率を向上している。

この様にデータベースマシンでは、ストリーム要素の大きさの単位、すなわちグラニュラリティを関係またはディスクページではなく、1タプルに設定することにより平均的な負荷分散を図っている。またこれによれば、プロセッサ側でストリームの一部、すなわち1タプルのデータが得られた時点で当該演算の評価を開始でき、処理と転送の重畳が可能となる。

```

転送アルゴリズム () {
  転送データ番号=0;
  while (転送するデータが存在する) {
    転送先アドレス = (転送データ番号 + 初期転送アドレス);
    send a tuple to 転送先アドレス mod 全モジュール数;
    転送データ番号++;
  }
}

```

図2 データ転送アルゴリズム

3.2 結合演算

結合演算のアルゴリズムのうち高速なものにソートマージ法があり、その処理時間は両関係の大きさの和のオーダーとなる。さらに、結合属性に関して両関係にハッシュを施し、全体を互いに独立な処理単位に分割（クラスタリング）すれば、異なるハッシュ値をもつのは、結合のとれる可能性がないため、これらを多数のプロセッサにより並列に処理することができる。

しかし、ハッシュ関数を用いたクラスタリングでは、データに応じて最適なハッシュ関数を選択することが不可能であるため、ハッシュ後のデータ分布は必ずしも均一にならず、特定のプロセッサに負荷の偏りが生じてしまい、性能劣化の要因となる。

本システムでは、以下のアルゴリズムに従って動的に適切なハッシュ関数を決定することにより、負荷分散を図っている。

step. 1 まず、初期ハッシュ関数により情報源空間をプロセッサ数の部分空間に分割する。

step. 2 さきにおけるハッシュ関数の選択が不適切であれば、特定のプロセッサに負荷が集中する。そこで、データの転送中に負荷の集中したプロセッサは、その負荷の集中した部分空間に対して更に別のハッシュ関数を施し負荷分散を行なう。

step. 3 step. 2 の操作を繰り返す。

更に、本研究では提案したアルゴリズムに対する評価を行った。もともと、問題の発端はハッシュ後のクラスタサイズが予測不可能なため負荷分散を実現できなくなることであった。そこで、評価においては負荷の偏り分布（位相 k のアーラン分布）を数種類設定し（図 3 参照）、これらに対する本技法の効果をシミュレーションにより調べた。なお、評価関数を次のように定義した。

$$\text{評価関数} = \frac{\text{最大負荷} - \text{平均負荷}}{\text{平均負荷}} \quad (1)$$

シミュレーション結果（図 4 参照）の示すものをまとめると、比較的プロセッサ数が多い場合でも評価関数が 10 % 程度であることから、提案したアルゴリズムにより並列度の高い処理が実現できること

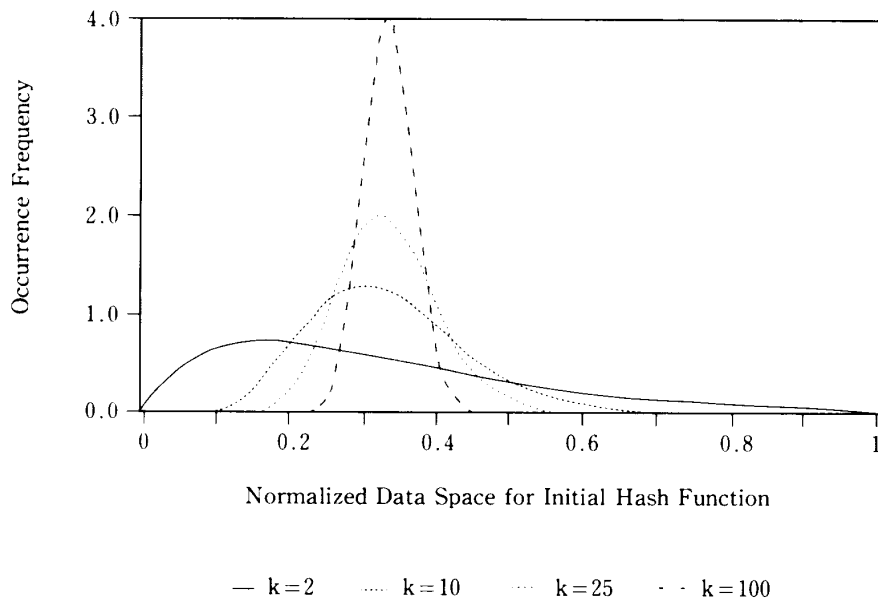


図 3 負荷の偏り分布

がわかる。また、データ流と共に動的クラスタリングを行う点で、負荷分散の平坦化に与えるマイナスの効果はほとんどない。

このようにアルゴリズムでは、適切なハッシュ関数を選択できなかったために特定のプロセッサに集中する負荷をデータ流とともに動的に分散することにより、最終的に負荷の平坦化を図るクラスタリングを実現し並列度を向上している。

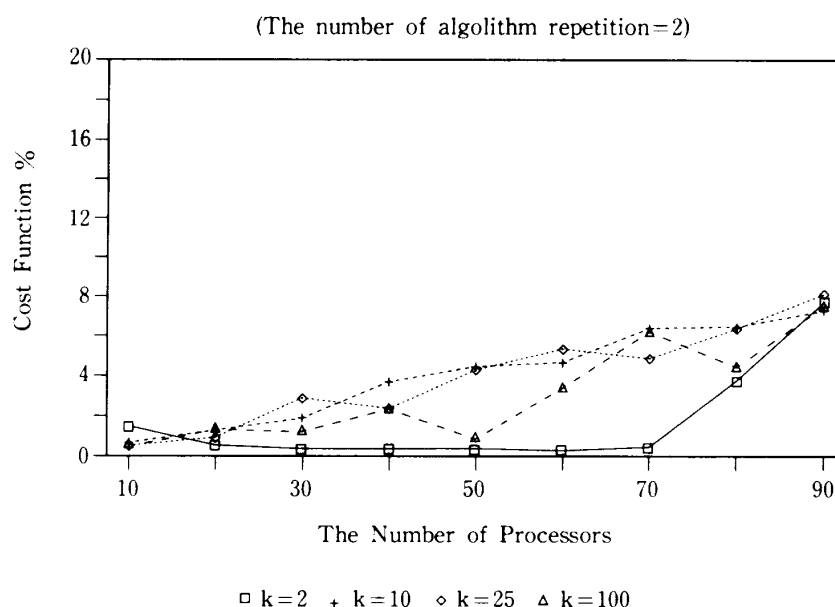


図4 プロセッサ数と負荷分散

3.3 MIMD方式の実装

MIMD方式に起因する制御オーバーヘッドの解消という観点から、なるべく制御モジュールにおける制御自体の削減、または他のモジュールへの負荷分散を考えなければならない。そこで、本システムでは、以下のような制御方策をとる。

① MIMD方式を実装するために、予めプロセッサ群をs分割する。この1分割のプロセッサ群をサーバとよび、基本的に1サーバに1ユーザの問い合わせを処理させる。問い合わせ処理が膨大な場合には、1問い合わせを複数サーバに処理させることも可能とし、これに対処する。このような方式の実装は各モジュールが、サーバ番号、サーバに割り当てられた問い合わせ番号、サーバに属するプロセッサ番号を要素とする履歴表を備えることにより可能となる。

② 従って、制御モジュールはホストコンピュータからの問い合わせをどのサーバに割り当てるか決定し、その旨を制御バスネットワークを介して各モジュールに伝達する。このため、制御モジュールは各サーバの稼働状況を監視しなければならない。

③ 一方、問い合わせ処理中の制御、すなわち「オペランドタプルの転送要求」等は、制御モジュールではなく、各サーバが分散して行うことにより、制御モジュールに対する制御オーバーヘッドを削減する。

4. モジュール間相互結合網

データベース処理において一回の転送量、すなわち1タプルの大きさは数十Byteから数KByte程度であり、従ってシステムにおける交換方式として、比較的制御が簡単である、セルフルーティング

型パケット交換方式を採用するのが得策である（文献(7)参照）。

セルフルーティング網の代表と考えられるバンヤン網は、ルーティングスキーマが単純であるが網内のアンバランスなトラフィック負荷に対しては呼損率が高いという欠点がある。そこで、本システムではこの問題点に対処できる負荷分散型バンヤン網である ASEN (Augmented Shuffle-Exchange Network) を結合網として適用した(図5 参照)。ASEN は、バンヤン網における各段のセル群がある

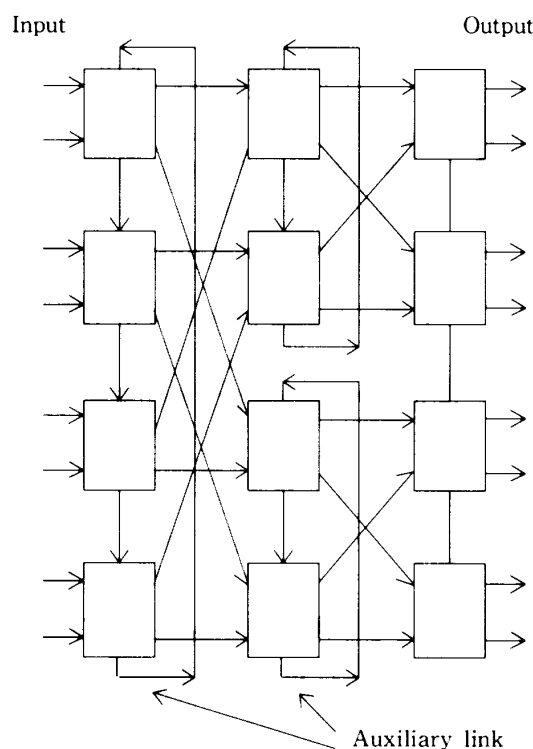


図5 8×8 ASEN の構成例

規則に従って連結された網であり、3×3スイッチセルが網の構成要素となっている。なお、本結合網では、ASEN の最終段のセルにおいても図5に示される補助リンク (Auxiliary link) を備えており、呼損率の低下を図っている。

4. 1 衝突検知再送方式

本研究で扱う対象は、音声、画像通信とは異なり、パケットの棄却は多少なりとも許されないデータ通信システムであるため、衝突検知再送方式の検討は重要となる。

このようなシステムでは、受信側がパケットを受信したことの確認信号、すなわち ACK 信号を返送することになるが、ASEN においては、行きのデータパケットと同様に帰りの ACK 信号も棄却される可能性があるため、結合網の性能劣化の要因となる。そこで本研究では、結合網において ACK 信号の衝突が起こらないルーティング制御を提案し、ネットワークの性能向上を図る。ただし結合網には、ACK 信号専用の経路が設けられているとする。

提案するルーティング制御は次の通りである。

[ルーティング制御] 行きのパケットは網内をセルフルーティングにより転送される。この時、パケットはそのヘッダに経路情報（どのセルのどのリンクを通過してきたか）を記憶する。この経路情報を基に、帰りの ACK 信号は行きの経路を逆戻りする。但し、次の規則に従って ACK 信号はルーティング制御される。

[逆戻りの規則] データパケットが行きにある段のセルの補助リンクを通らなかった場合は、それに対する ACK 信号は帰りにその段を通過するとき、セルで1タイムスロット待機する。そうでなければ、1タイムスロットで補助リンクを通り、同じ段の次のセルへ帰り道を進める（図6参照）。

この様なルーティング制御により ACK 信号の衝突が避けられる。また、上記のアルゴリズムに従うと、送信が成功した場合 ACK 信号は送信してからスイッチ段数×3タイムスロット後に必ず受信できる。従ってこのとき ACK 信号が戻らない場合、送信側は送信したはずのデータパケットが棄却されたことを検知することになる。

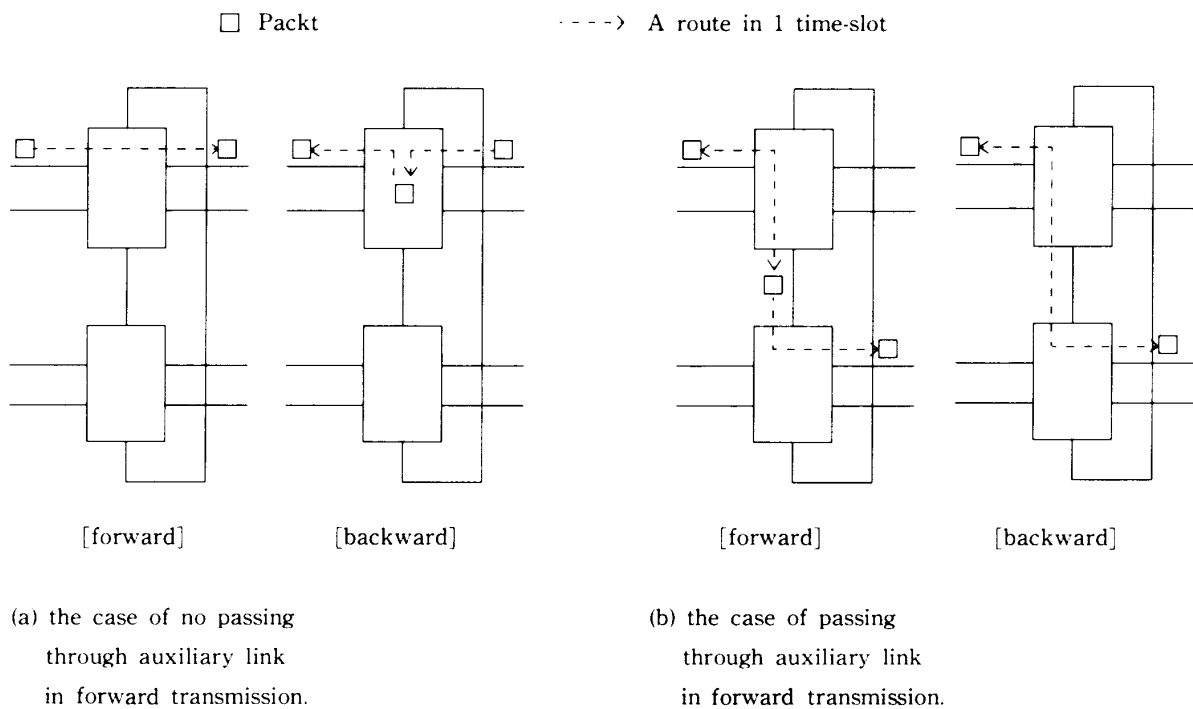
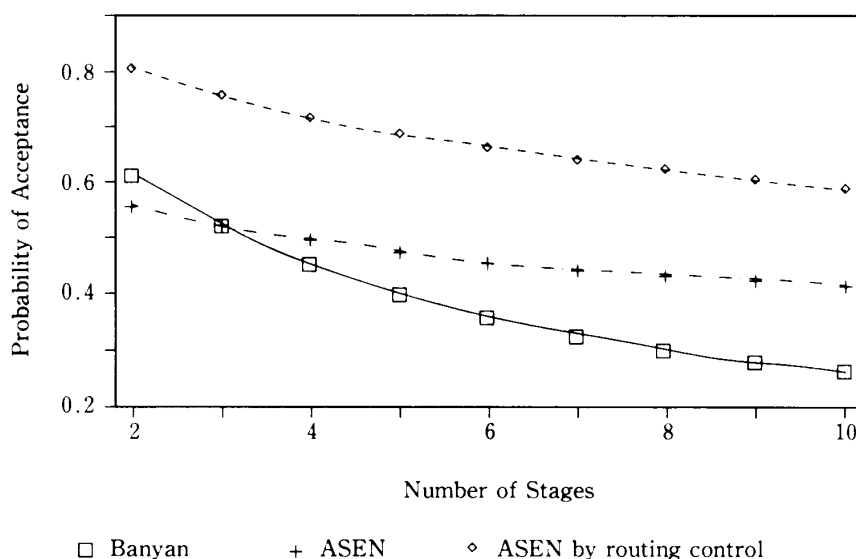


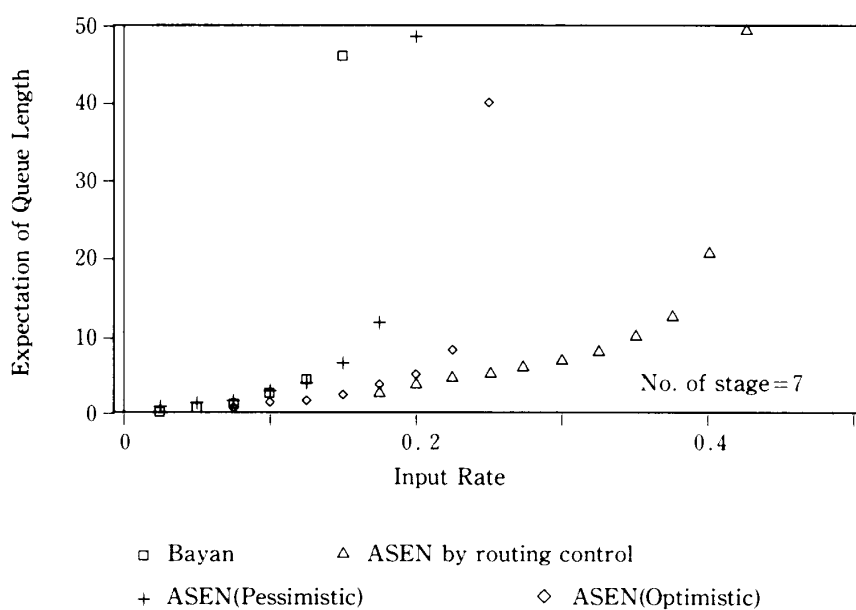
図6 逆戻りの規則

再送スキーマとしては、さまざまなものが提案されているが、本システムでは現実的で比較的性能が良く、理論解析もよくなされている Go-Back-N 方式を適用する。

[Go-Back-N 方式] 送信側は、パケットを連続的に送出し衝突を検知するとそれに対するパケットから再び順次再送する。



(a) Number of stages and probability of acceptance.



(b) Input rate and expectation of queue length.

図7 相互結合網の性能解析

4.2 ネットワークの性能解析

最後に、上述のことを考慮に入れて行なった解析結果を図7に示す。解析では、ACK信号の受信まで考慮に入れた通信成功率と平均待行列長について評価し、またバンヤン網、従来のASENと比較したが、提案した方式により相互結合網の性能向上が図られていることを確認できる。さらに、3で示した各送信側が従う転送アルゴリズムを考慮に入れたネットワークの性能についてシミュレーション

により評価し、転送アルゴリズムによって網の通信成功率は、転送をランダムに決めた場合に比べ、劣化しないことを確認した。

5. システムの性能評価

5.1 応答待ち型リアルタイムシステム

一般に検策システムにおいて、ユーザはあるコマンドに対する応答を待ってから次の作業にとりかかると考えて良い。この様なシステムではプロセッサ負荷が大きくなって応答時間が長くなると、コマンドの入力頻度も低下するため入力トラヒックを自動的に抑制する作用が働く。従って、輻輳システムとして評価する場合、M/M/sモデルを用いることは、このモデルが処理結果とは無条件に一定の入力があることを前提としているため、適当でない。以上の理由から、性能評価では図8に示される評価モデルを適用する。但し、このモデルではM/M/sと同様に処理時間、考慮時間が共に指数分布に従うものと仮定し、またフィードバックループは処理が終了して応答があってから考慮時間が始まることを意味する。

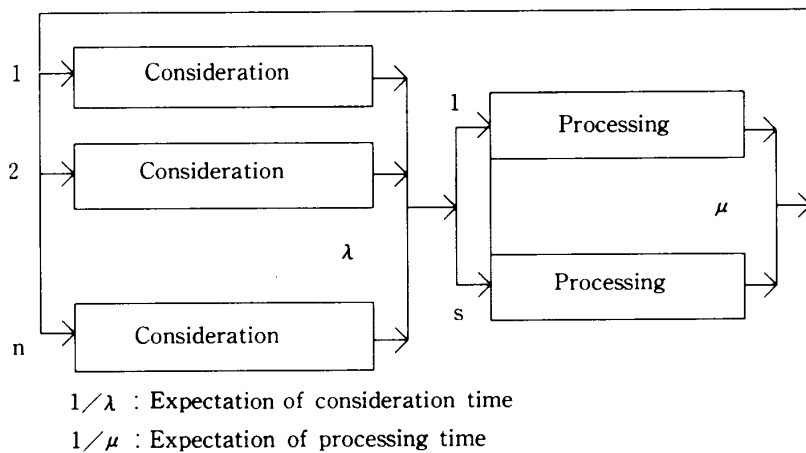


図8 応答待ち型リアルタイムシステム

更に、性能評価では、一般的問い合わせモデルとして図9に示されるモデルを仮定した。すなわち、ユーザが入力するコマンドはこの一般的問い合わせモデルの形式に従うことを示唆している。また、評価における応答待ち型モデル、バックエンドマシンのハードウェアモデル、対象とするデータベースのモデルに関する詳細をまとめて表1に示しておく。

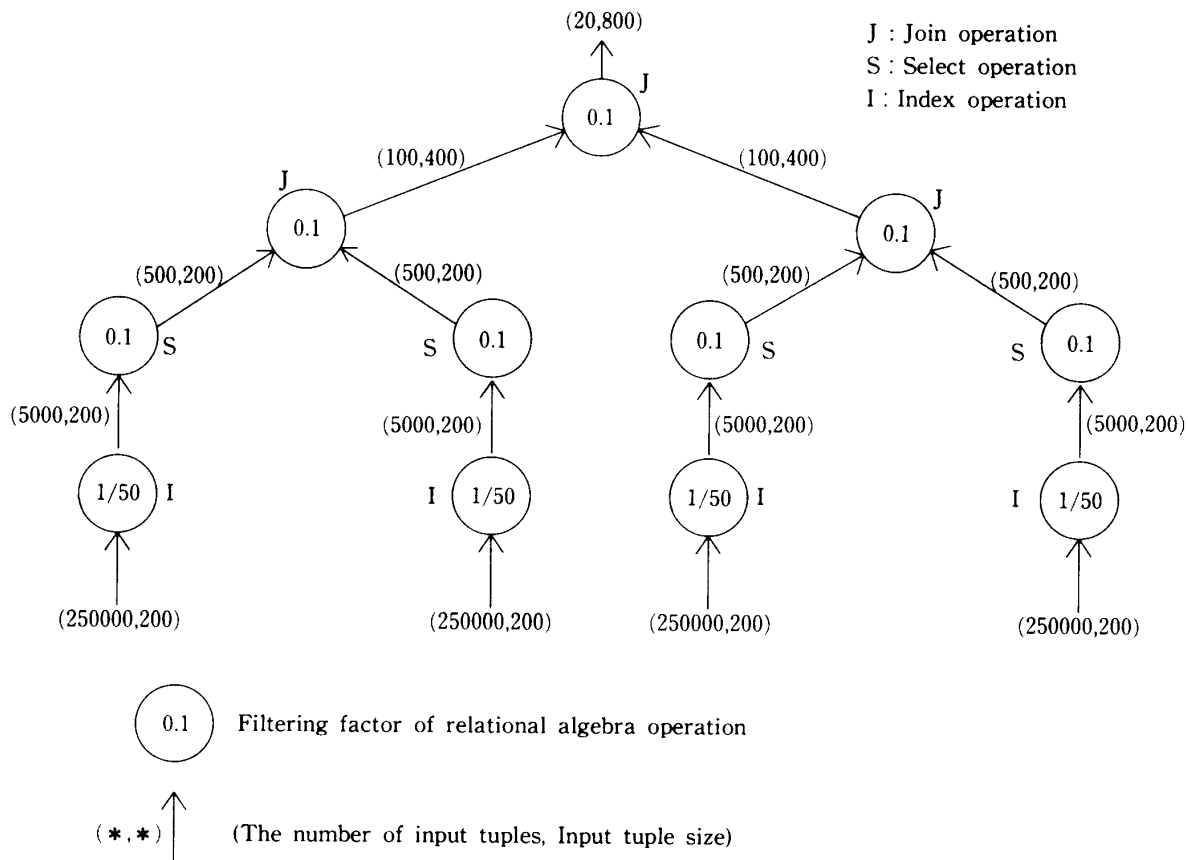
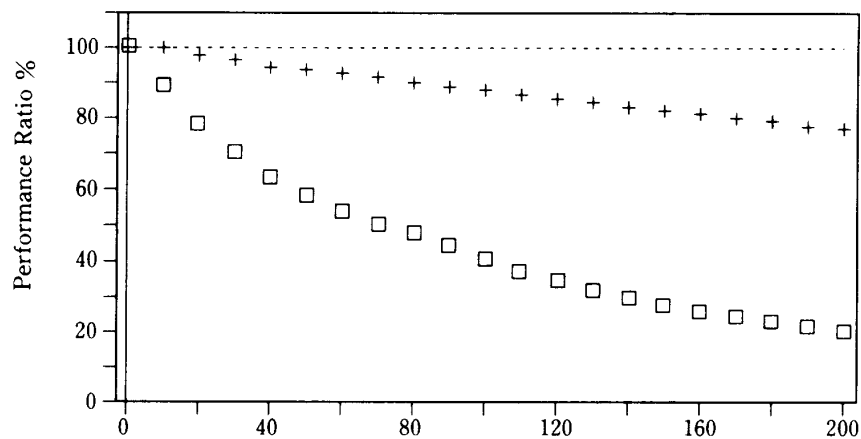


図9 一般的問い合わせモデル

表1 評価モデルの詳細

<ul style="list-style-type: none"> ・考慮時間の平均値 ・問い合わせモデル 	60秒 葉ノードへの入力タプル数の平均=25万タプル、標準偏差=500タプル、タプルサイズ平均=200Byte、標準偏差=100Byte、関係代数演算の選択率の平均=0.1、標準偏差=0.05
<ul style="list-style-type: none"> ・プロセッサ性能 ・メモリモジュールのデータ転送速度 ・相互結合網のデータ転送速度 ・結合網の通信成功率 ・ディスクのアクセスタイム 	8 MIPS(SUN 4程度) 8 MByte/sec 20MByte/sec(内部高速化比=2.5) 0.95(シミュレーション結果) 平均10msec、標準偏差 2 msec
<ul style="list-style-type: none"> ・最小ディスクアクセス単位 ・関係の大きさ ・データベースの大きさ 	20KByte 50MByte 関係数 1 万件(50GByte)



- The number of input tuples into a leaf node = 250,000
- + The number of input tuples into a leaf node = 2,500,000
- Ideal performance

図 10 一般的問い合わせモデルに対する性能

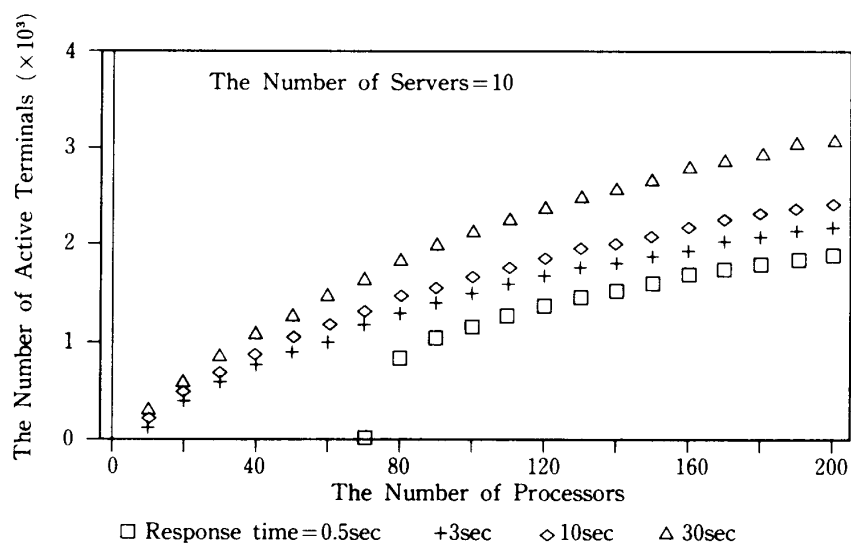


図 11 プロセッサ数と端末数の関係

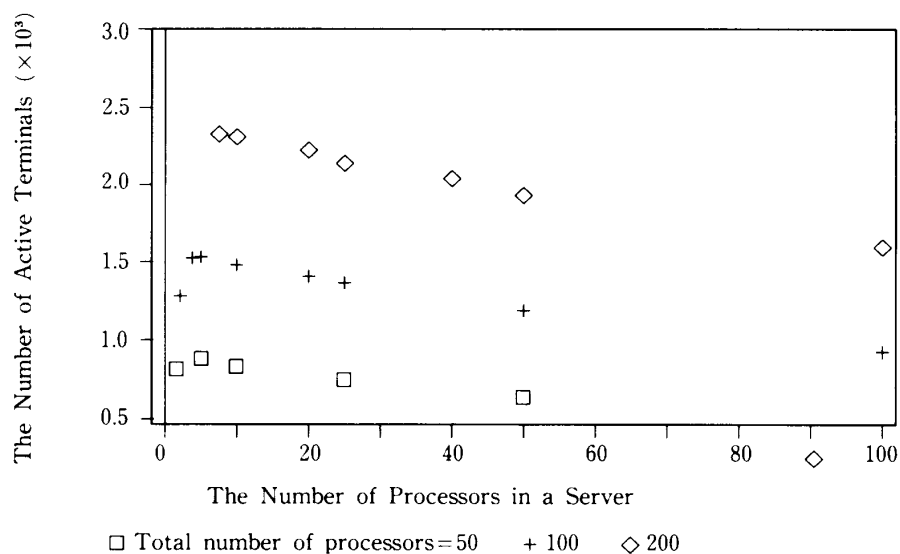


図 12 1サーバ当たりのプロセッサ数と端末数の関係

5. 2 システムの基本的性能

上述の仮定に基づき、評価を行った結果を図 10 から図 12 に示す。結果の示すものを以下にまとめて列挙する。

① まず、プロセッサ数 N と次式で示される性能比

$$\text{性能比} = \frac{1 \text{ 台のプロセッサによる処理時間} / N}{N \text{ 台のプロセッサによる実際の処理時間}} \quad (2)$$

の関係を調べたもの (図 10) では、ディスクアクセス時間が無視できなくなる等の理由で、並列度を増しても意味がなくなる限界があることがわかる。従って、性能劣化の少ない所で 1 問い合わせを処理し、同時に複数個の問い合わせを処理すべきであることがわかる。

② これに起因する結果は、プロセッサ数と端末数の関係 (図 11) から明らかである。また、1サーバ当たりのプロセッサ数とアクティブ端末数の関係を調べたもの (図 12) によれば、並列システムに最適なサーバ数が存在することがわかる。

結局、評価結果によれば、計算機資源を有効利用するために MIMD 方式をとることにより、性能向上を図ることができることがわかり、また同時にシステム構成の目安も得られた。

5. 3 制御モジュールに要求される性能

本データベースマシンシステムにおける制御モジュールは、システムに対し集中制御の形態をとるため、結果として処理のボトルネックとなり易い。また、5. 2 の評価では制御系のオーバーヘッドに

ついて全く考慮していない。従って、現実問題として、制御モジュールに対し、どの程度の性能が要求されるかについて検討した。

本研究では、問い合わせ処理の制御を大きく分けて、次の3つに分け、

- ① ホストコンピュータからの問い合わせが制御モジュールに到着し、サーバに割り当てられるまでの制御
- ② バックエンドマシンで、演算が実行されている間の制御
- ③ 問い合わせ処理が終って、プロセッサ側から、制御モジュールを介して、結果がホストコンピュータへ転送されるまでの制御

それぞれについて M/M/1 の待ち行列モデルを仮定し、制御オーバーヘッド、すなわち待ち行列の系待ち時間を前述の評価モデルに基づいて算出した。

この結果によると、特に処理のボトルネックとなるのは、問い合わせ処理結果をホスト側へ転送する際であり、これに携わるバスネットワークとしては、高性能なものが要求されることが分かった。また、既存のハードウェアを用いることにより、集中制御による制御モジュールに対する過負荷の問題が解決でき(制御オーバーヘッドの処理終了率に対する3%未満)、システムの制御系の構築が可能であることが分かった。

6. むすび

本論文では、将来予想される大規模な関係データベース環境に対処し得るシステムを構築するために、現在までに提案されてきた方式を概観し、処理のボトルネックの解消、低コストパフォーマンス等を克服するようなシステム構成を提言した。次に、関係代数演算のうち最も基本的な操作である、選択演算と結合演算の並列処理方法を述べ、それらのシステムへの実装方法を提案した。並列処理の抽出は、各モジュールに対する負荷分散が主な目的であった。さらに、並列マシンシステムにおいて、特に重要となるモジュール間相互結合網の検討を行った。提案するシステムでは、結合網として、負荷分散型バンヤン網を取り扱い、その構成の若干の改良と新たなルーティング制御を行うことにより、より高性能な特性を示す網の提案を行った。最後に、問い合わせ処理、データベースシステム全体、制御系に関する性能評価を行い、その結果、ここで提案したマシンアーキテクチャおよび、駆動方式による高性能なシステムを実現できることを確認した。

これからの高度情報社会では、ネットワーク技術の動向に伴い、データベースのもつ役割はより重要なものとなる。このような環境においては、大規模な情報をより高速に処理するマシンの開発技術が社会に及ぼす役割は非常に大きい。本論文ではデータベースマシン技術について様々な検討を加え、優れた性能をもつシステムを提案することができたが、まだまだ解決すべき課題も多く、特に同時更新処理に関する問題が重要であろう。

文 献

- (1) H. Boral, D.J. Dewitt: "Processor Allocation Strategy for Multiprocessor Database Machines", ACM Trans. Database System, 6,2, pp. 227-254.
- (2) T.Baba, S.B.Yao, A.Heuner: "Design of a Functionally Distributed Multiprocessor Database Machine Using Dataflow Analysis", IEEE Trans. Comput., C-36, 6, pp. 650-666.
- (3) D.J. Dewitt, "DIRECT-A multiprocessor organization for supporting relational database management systems", IEEE Trans. Comput., C-28, pp. 395-408.
- (4) G.Z. Qadah, K.B. Irani: "A Database Machine for Very Large Relational Database", IEEE Trans. Comput., C-34, 11, pp. 1015-1025.
- (5) 伏見伸也、喜連川優、田中英彦、元岡 達: "一般化 KD 木を用いたデータベースマシン GRACE の 2 次記憶系の設計", 信学論(D)、J68-D、 6、 pp. 1280-1287.
- (6) 浜田、大久保: "結合演算の高速化を図るデータベースマシンアーキテクチャ", 昭和 63 年電子情報通信学会秋期全国大会、D-79。
- (7) 坂井修一、喜連川優、田中英彦、元岡 達: "関係代数マシン GRACE におけるバケット分配網", 信学論(D)、J 68-D、 6。
- (8) P. Kumar, S.M. Reddy: "Augmented Shuffle-Exchange Multistage Interconnection Networks", IEEE COMPUTER, 6, pp. 30-40.
- (9) D.Towsley, J. Wolf: "On the Statistical Analysis of Queue Lengths and Waiting Times for Statistical Multiplexers with ARQ Retransmission Schemes", IEEE Trans. Commun., COM-27, 4, pp. 693-702.

研 究 論 文

放送形トラヒックに対する 高速パケット通信網のルート制御

Routing Control for Multi-destination Traffic

学術情報センター 浅野正一郎

東京大学大学院工学系研究科 安藤 史郎

1. 序 論

〈研究の背景〉

次世代のネットワークとして研究が行われている広帯域 ISDN では、動画像通信、テレビ会議、広帯域ビデオテックス、高品位テレビなども含め、通信サービスと放送サービスを統合的に提供することが考えられている。広帯域 ISDN を実現するための技術要件は高速伝送技術と高速交換技術であるが、伝送面は加入者線まで含めたすべての有線伝送系の光ファイバケーブル化によって対応できる。交換面ではパケット文換の流れをくむ ATM 交換方式が有望視され精力的な研究がなされている。ATM ではプロトコル手順の簡略化とハードウェアスイッチングによって最高数百 Mbit/s にいたる任意の速度の通信を統合的なインタフェースのもとに提供することが目標とされている。

広帯域 ISDN では 1 対 1 の接続を行う通信 (1 : 1 通信) のほかに TV 放送、CATV、多地点間 TV 会議といった 1 対多あるいは多対多の放送形の通信 (1 : N 通信) のサービスを提供することが考えられている。これらの放送形トラヒックを ATM 通信網のようなパケット通信網で扱う場合、1 : 1 通信の重ね合わせによって行くとトラヒックの増加が大きくなるためネットワーク資源を有効に利用することができない。そこで交換ノードにパケットの複製機能をもたせ、1 つの入力パケットを複数の宛先に送信することによってトラヒックの増加を抑えることが考えられるのだが、このような 1 : N 通信を実現するためには、単なるルート設定だけでなく、いつ、どこで、いくつの複製を作るかという制御、すなわち接続経路の枝分かれの仕方を決める制御が必要になる。また複数の端末に対する通信の発生、終了に伴う処理は、従来の 1 : 1 通信よりも複雑な処理を必要とする。

〈研究の内容〉

このような背景に基づき、本研究では高速パケット通信網による放送形トラヒックの制御方法について様々な側面から検討し、特にルーティングアルゴリズムとネットワークトポロジーの両面について新しい方式の提案を行いその性能を評価する。

2. 高速パケット通信網と放送形トラフィックの制御

2. 1 高速パケット通信網

〈広帯域 ISDN におけるサービス〉

表1 広帯域サービスの分類

区 分		具体的サービス例
対話型 サービス	会話型サービス	高品質 TV 電話／会議
	メッセージサービス	ビデオメール
	検索サービス	広帯域ビデオテックス
分配型 サービス	従来型放送サービス (プレゼンテーション 制御のないタイプ)	TV 放送、CATV
	個別制御型放送 サービス (プレゼンテーション 制御のあるタイプ)	CATV

広帯域 ISDN における広帯域サービスをその属性と網機能の特性に応じて分類すると、対話型サービスと分配型サービスに大きく分類できる(表1)。広帯域 ISDN ではその帯域の広さ、および通信速度の可変性を利用して、従来は制度的にも技術的にも放送という形で行われてきた通信も含めたサービスを提供することが大きな特徴である。

〈ATM 交換方式〉

広帯域 ISDN で目標とされている数十 Mbit/s 以上の通信を行うためには数百 Mbit/s 以上のスイッチング速度を持った交換機が必要になる。この実現を目指して、非同期転送モード(ATM: Asynchronous Transfer Mode)による情報伝送技術を用いた ATM 交換方式が研究されている。ATM 交換方式は原理的にはパケット交換技術の流れに沿うものであるが、交換ノードでは時間を要するプロトコル制御や通信処理などのソフト的な処理を一切省き、並列分散化したハードウェアスイッチングによって高速化を図り、個々の呼に対しては実効的に回線交換的な使い方をし、いわばパケット交換と回線交換の複合的な交換方式である。

ATM 交換方式の特徴の中で特に本研究に関連のあるものを以下に挙げる。

- (1)蓄積転送の過程でセルのコピーをとるなどの手段によって同一情報を複数の相手に伝達する放送分配形の交換接続が行える。
- (2)時間を要する経路選択制御をセルごとには行わず、また、順序制御を避けるためにも、パケット交換のバーチャルコール接続と同様に、情報転送に先立って接続経路をあらかじめ設定し、同一呼の一連のセルは同じ経路を通して接続するのが普通である。

2. 2 放送形トラヒックの制御

〈放送形トラヒックとは〉

本研究で扱う「放送形トラヒック」は通常の放送の他に、同報通信、ポイント・ツー・マルチポイントの接続を要求するあらゆるトラヒックを含む。これらはいずれも受信要求が起きてから情報の送出行う、すなわちデマンドリクエストで網内の一部の端末に対して通信経路を設定するという点で一般の放送と異なる。このような「時間的、空間的に有限の放送」をパケット通信網上で提供する制御を考えることが本研究の重要なポイントである。

〈放送形トラヒック制御の考え方〉

パケット通信網で放送形サービスを実現する方法はパケットの複製の作り方によって大きく分けて以下の2つの方式が考えられる。

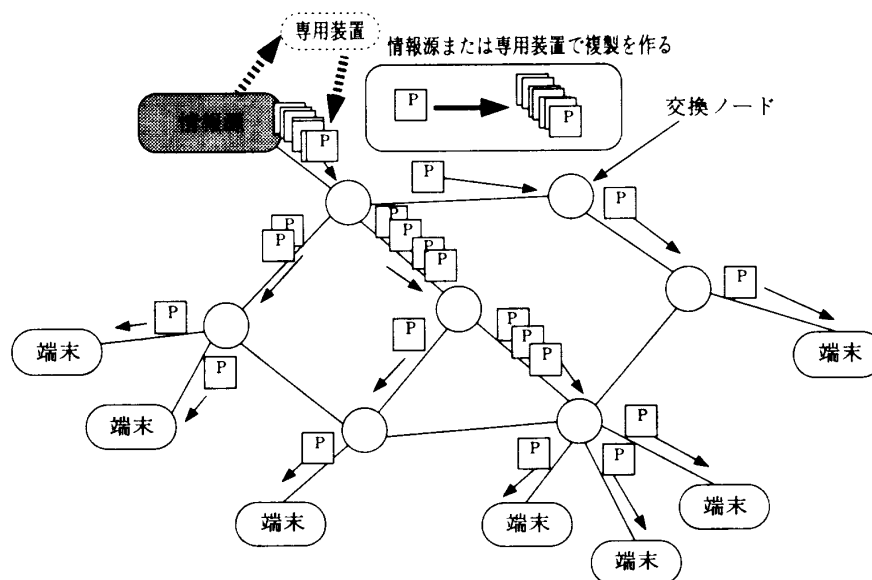


図1 [方式1] 複製を一回で集中的に行う方式

[方式1] パケットの複製を一回で集中的に行う方式 (図1)

放送形パケットが入った最初の交換ノードの出力段階までに宛先数の複製が作られ、あとは各々の宛先に対して別々の1:1通信として経路を設定する、1:1通信×Nの形。

[方式2] パケットの複製を数回で分散的に行う方式 (図2)

各交換ノードにパケットの同報接続機能をもたせ、1つの入力パケットが途中のノードで何度か分散的に複製され、経路分岐を行う。放送形の通信を複数のノードの協調動作によって実現する。

〈両方式の比較〉 (表2)

[方式1] 網管理における処理、交換ノードの機能の簡易性という点では有利だが、パケット転送量の増大にともなうトラヒックの圧迫や情報源のバースト性の影響が大きくなるという問題がある。また宛先端末数が増大してきた場合や、動画像情報のようにトラヒック量が大きい場合、必要数のパケ

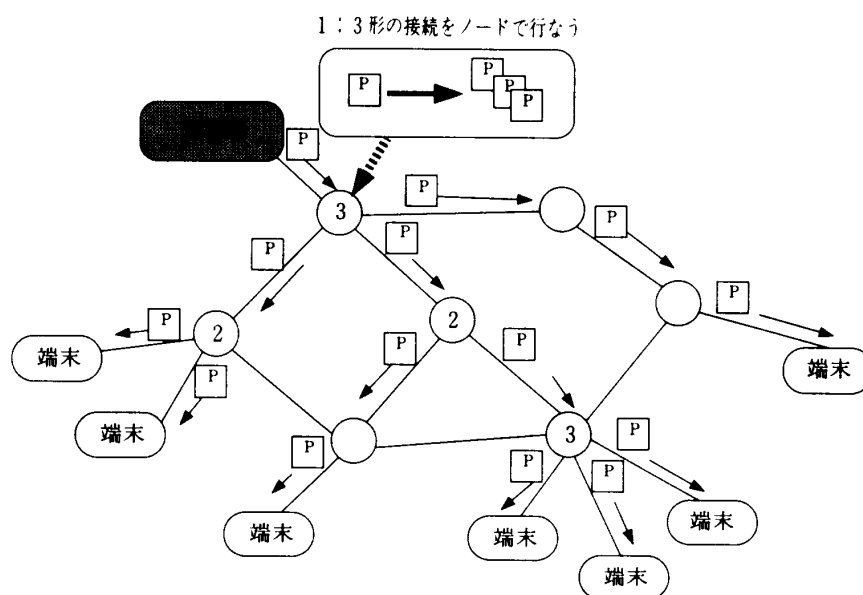


図2 [方式2] 複製を数回で分散的に行う方式

表2 方式1と方式2の比較

評価項目		方式1	方式2
機能配分	交換ノード	1 : 1 接続機能	1 : N 接続機能
	端末機能	N 出力の機能	1 出力の機能
網管理	経路決定	処理量 小	処理量 大
	接続状態管理	1 : 1 接続と同等	複雑ノードで分岐するため複雑
	宛先グループの変更に対する対応	容易	複雑
	ルーティング情報	1 : 1 接続と同等	複雑
通信品質	パケット転送量	大	小
	トラヒック圧迫	大	小
	遅延	大	小

ットの複製を短時間で生成して送出することは困難である。

[方式2] 端末ではなくネットワーク側で1 : N 接続を実現でき、ユーザに対して柔軟な接続形態を提供できる。また高速広帯域の通信ではトラヒック量が少なく済む点が非常に有利である。

以上の検討から、本研究では [方式2] を採用し、1 : 1 接続よりも複雑になる制御を合理的に実現することを考える。

3. ルーティングアルゴリズム

3.1 ネットワークモデル

ルーティングアルゴリズムを考える前提となるネットワークのモデルについて述べる。

〈ネットワークレベル〉

端末レベルまで扱わず、中継ネットワークレベルでの制御を考える。交換ノードは1:1から1:N (Nは出線リンクの数) までの任意の同報接続の機能を備えたブラックボックスとして扱い、内部の動作については言及しない。

〈ネットワークトポロジー〉

高速パケット通信網の交換ノード間の基本的な接続形態(ネットワークトポロジー)はスター形(ツリー形)、リング形、メッシュ形に分類できるが(図3)、スター形、リング形のトポロジーでは、あるノードから他のノードに至るツリー状の経路が一意に決まるため、特にルーティングアルゴリズムを考える必要がない。そこで本研究ではメッシュ形のトポロジーを想定しその上での制御を考える。

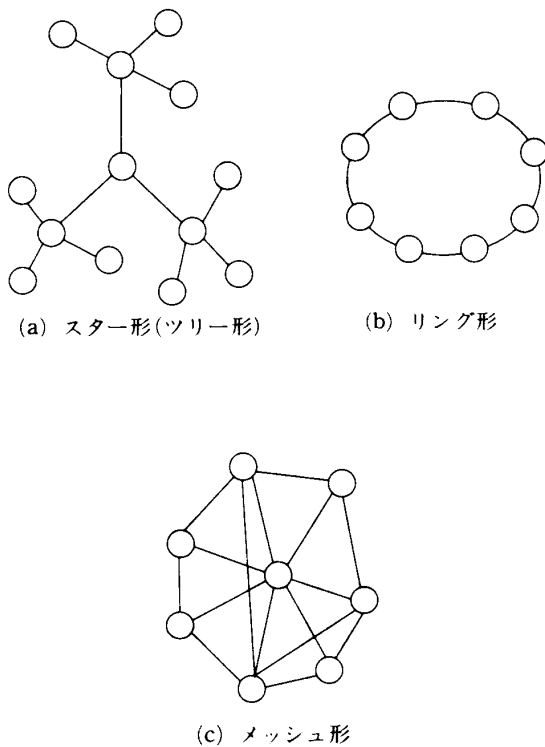


図3 種々のトポロジー

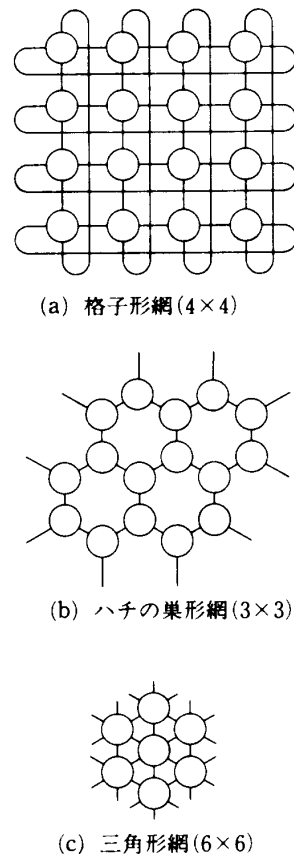


図4 メッシュ形の例

しかしランダムに接続されたメッシュ形ネットワークでは理論的な取扱いが困難なため、各ノード

が同数の入出力リンクを持ち、ネットワークが一様なものを考える(図4)。特にルーティングアルゴリズムを評価する段階では最も取扱いが簡単な格子形網を使う。

〈網内情報の収容形態〉

トポロジー情報、トラヒック情報などのルーティングに必要な情報を網内で管理、収容する方法は次のように分類できる。

(方式1) 各ノードが網内全体の情報を管理、収容する。

(方式2) 各ノードはその近傍の情報をのみを管理、収容する。

それぞれの方式でルーティングアルゴリズム、網内情報の管理がどのように変わるかを比較した結果を表3に示す。

表3 網内情報の収容形態による処理の比較

	方式1 各ノードが網全体の情報を持つ	方式2 各ノードが近傍の情報を持つ
ルーティング アルゴリズム	単一ノードで処理できるため → 簡単	複数ノードの協調動作必要 → 複雑
網内情報 の管理	全ノードの更新が必要 → 処理 大	自ノードだけ更新すればよい → 処理 小

本研究が目的としている放送形通信の効果的な制御手法が確立されれば、各ノードが近傍の情報を他の全ノードに対して放送すれば(方式1)の形態は容易に実現できるはずである。よって(方式1)を仮定してアルゴリズムを考える。

3.2 アルゴリズムに対する要求

放送形トラヒックのツリー状経路を決定するためのルーティングアルゴリズムに要求される機能について述べる。

〈アルゴリズムの目的〉

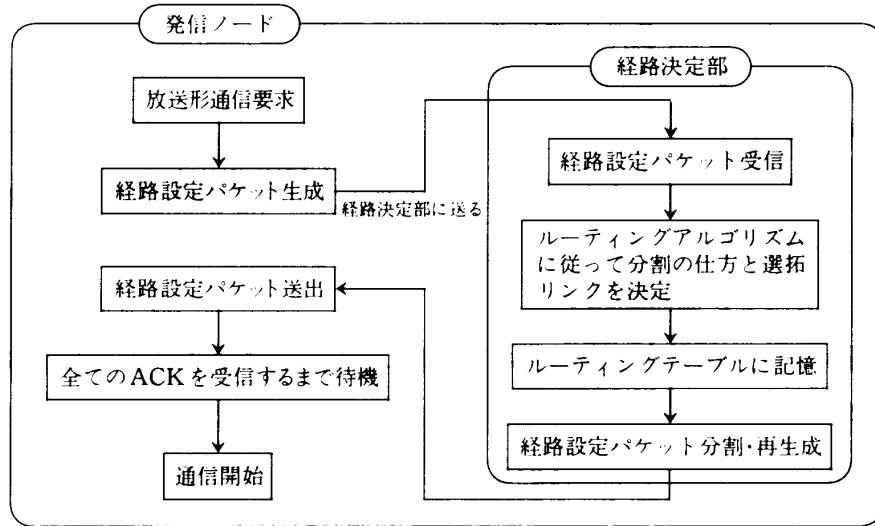
複雑な制御を必要とするツリー状経路の利点は、放送形通信によるパケットの増加が小さく、トラヒックに与える影響が少ないことにある。そこでアルゴリズムの目的は次のように定義された経路のホップ数を小さくすることになる。経路のホップ数は通信によって使用されるネットワーク資源に比例する指標と考えられる。

$$\begin{aligned}
 \text{経路のホップ数} &= \text{経路に含まれる中継リンクの数} \\
 &= \text{発信ノードを除く経路内の交換ノードの数}
 \end{aligned}$$

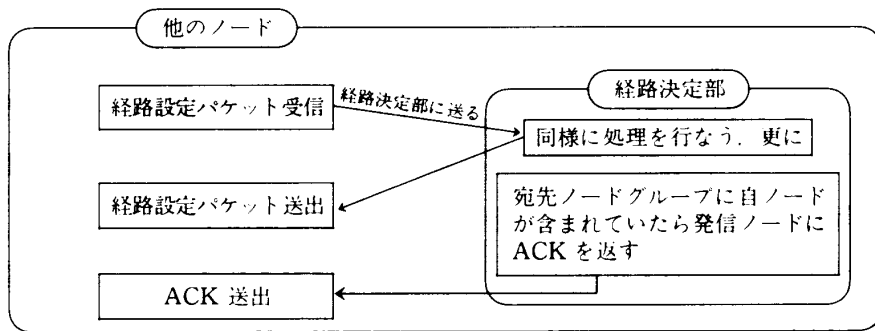
〈ルーティング情報〉

1:N通信のツリー状経路では各ノードで作る複製の数およびそれらが選択する中継リンクの情報がすべて必要である。このように多くの情報をヘッダに持たせることは処理上不利であり、CCITT勧告でほぼ決まっている5バイトのヘッダの中では情報だけでは制御が不可能である。

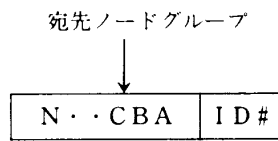
そのため通信を開始する前に経路設定を行うフェーズを設け、各ノードにこれらのルーティング情報を通信の ID とともに登録する方式が必要になる。このような処理は ATM の呼設定手順とも親和性があり妥当である。経路設定フェーズの手順をブロック図で示す (図 5)。



(a) 発信ノードの処理



(b) 他のノードの処理



(c) 経路設定パケット

ID#	DEST	LINK
ID 1	A, B	X
	C	Y
	⋮	⋮
	N	W
ID 2	⋮	⋮
	⋮	⋮

(d) ルーティングテーブル

図 5 経路設定フェーズの手順

〈宛先ノードグループの変化〉

放送形トラヒックの制御では、宛先ノードグループは通信中に自由に変更できなければならない。これには通信を継続したまま新たな経路設定パケットにより経路を再設定し、それが完了した時点で通信 ID を新しいものに変えて経路の切り換えを行なうことで対応できる。このような処理によって利用者は通信の切り換えを意識しないで済む。

3. 3 既存のアルゴリズム

最小スパニング木 (MST: Minimum Spanning Tree) とはネットワークの各ノードから他のすべてのノードへのパスが存在するグラフの中で、接続するリンクの数が最小のものである。MST を構成するアルゴリズムが多数提案されているが、これらを方式的な観的から 2 種類に分類して特徴を述べ、これらのアルゴリズムが本研究の要求を満たさないことを示す。

〈グループ 1: 単一ノードで処理を行なうタイプ〉

このグループの特徴は処理が単一のノード内で行われ、トライ&エラーの形の処理が多く、各ノードが網内全体の情報を保持していることである。例としてリンクの全体の長さが最小であるようなスパニング木を決定する Kruskal のアルゴリズムの動作例を図 6 に示す。

〈グループ 2: 複数のノードが協調動作するタイプ〉

グループ 2 の特徴は、ノード間でメッセージを交換しながら協調動作によって MST を構成し、各ノードがネットワーク全体の情報を持たなくてもよい点にある。ネットワークの規模が大きくなっても各ノードで行う処理があまり大きくならないという点で、コンピュータネットワークの分散処理化に適した手法である。

〈既存のアルゴリズムの問題点〉

[全ノードが対象] これらは全ノードに対する MST を構成することを目的としており、全ノードが対象であるメリットを十分に利用しているが、本研究で要求する制御は、ネットワーク内のノードの部分集合に対してもツリー状経路を構成することであり、単純なリンクの付け足しによるトライ&エラーの方法や、隣接ノード間でメッセージを交換する方法は適用できない。

[宛先ノードの変化] どちらのグループも一度経路を決定できればよい。あるいはある情報が全ノードに行き渡ればよいという考え方のため、処理の高速性や、簡易性をあまり追求していない。また宛先ノードの変更に対する対応もほとんど考えられていない。しかし本研究では、通信継続中でも宛先ノードグループの変化に対応するために、経路変更しやすさ、処理の簡易性、高速性を要求している。

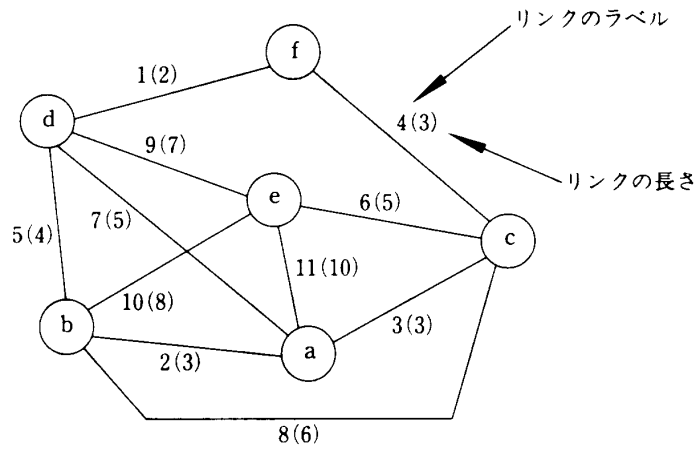
3. 4 重み付けアルゴリズムの提案

[基本的な方針]

- ・各ノードで分散的に処理を行う。
- ・厳密解ではなく近似的な最適経路を得る。
- ・トライ&エラーの手法は避け、トップダウンに経路を決定できる評価関数を使う。

[戦略]

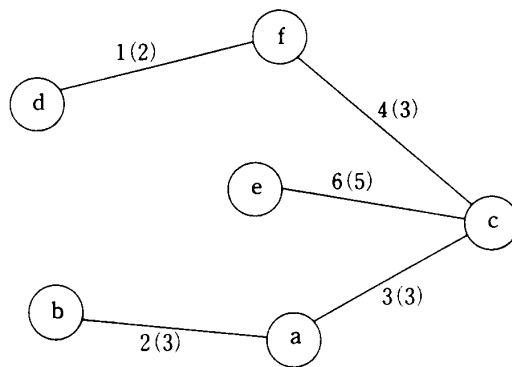
- ・各宛先ノードに対する最短パスを組み合わせることでツリー状の経路を構成する。
- ・経路の分岐をなるべく後回しにすることでネットワーク資源の節約を図る。
- ・経路決定に必要なとする情報は他の全ノードに対する最短経路のホップ数と最短経路に含まれる出リンクの方向情報である。



(a) 元のネットワークトポロジー

ノード	a	b	c	d	e	f
a		3	3	5	10	
b	3		6	4	8	
c	3	6			5	3
d	5	4			7	2
e	10	8	5	7		
f		3	2			

(b) ノード間のリンクの長さの表



(c) 結果として得られた最小スパニング木

図6 Kruskal のアルゴリズムの動作例

重み付けアルゴリズムの処理手順を図7に示す。

3.5 格子形網における性能評価

〈格子形網〉

格子形網は4方向にリンクを持つノードをm行m列に並べたトーラス状のネットワークである。これを以降 $m \times m$ と表記する。格子形網ではどのノードから見てもネットワーク全体のトポロジーは同じでありネットワークは一様である。

〈アルゴリズムの動作例〉

ここで重み付けアルゴリズムの動作例を示す。図8(a)のような 5×5 格子形網のノードsに宛先ノードグループ {abcde} を指定する経路設定 packets が左側のリンク (link 4) から到着すると、ノードsでは図9(a)のような表を参照して各リンクにホップ数による重み付けを行なう(図9(b))。この中でlink 1の重みが9で最大なので {abc} はlink 1に決定する。{abc} による他のリンクの重みを取り消すとlink 2の重みは0になり、次に大きいのは重み3のlink 3なので {de} はlink 3に決定する。これですべての宛先ノードが決定できたのでアルゴリズムは終了し、結果は図9(c)のようにlink 1とlink 3に分岐するルートとなる。

このように各ノードでルーティングを行なうと、結果的に図8(b)のような経路が得られる。

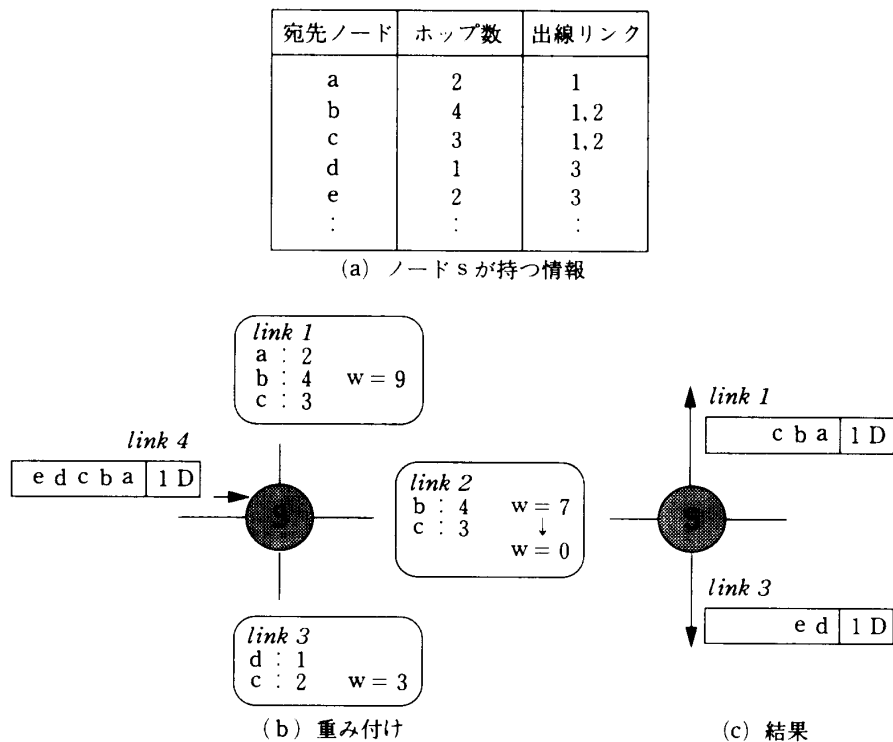


図9 重み付けアルゴリズムの動作例

〈シミュレーションの方法〉

1つの放送形通信によるツリー状経路のホップ数の平均値を得ることを目的としている。格子形網

(1) for all link i ($i=1,2,\dots,L$)

すべての出線リンク i に対して

weight_of (link i) = 0
 リンク i の重みを 0 にリセットする。

member_of (link i) = { }
 リンク i のメンバを空にする。

(2) for all dest j ($j \in \{\text{dest}1, \text{dest}2, \dots, \text{dest}N\}$)

経路設定にパケットに含まれるすべての宛先ノード j に対して

for all link k (link $k \in \{\text{shorteset route from nodeA to dest } j\}$)
 自ノード A から宛先ノード j に至る最短経路に含まれるすべての出線リンク k に対して

weight_of (link k) = weight_of (link k) + hop (nodeA, dest j)
 リンク k の重みにノード A から宛先ノード j に至るホップ数を加える。

member_of (link k) = member_of (link k) + {dest j }
 リンク k のメンバに dest j を加える。

(3) select link M (weight_of (link M) is max)

重みが最大のリンク M を選ぶ。

for all dest j ($j \in \text{member_of (link } M)$)
 リンク M のメンバに含まれるすべての宛先ノード j に対して

dest $j \rightarrow$ link M
 宛先ノード j についてはリンク M に送出することを決定する。

(4) for all link i ($i \neq M, \text{dest } j \in \text{member_of (link } i)$)
 宛先ノード j が選択できるその他のリンク i に対して

weight_of (link i) = weight_of (link i) - hop (nodeA, dest j)
 宛先ノード j によるリンク i の重みを取り消す。

member_of (link i) = member_of (link i) - {dest j }
 リンク i のメンバから j を除く。

(5) if all dest # are decided then end else goto (3)

すべての宛先ノードに対してリンクが決定されたら終了。
 そうでなければ(3)にもどる。

図7 重み付けアルゴリズムの処理手順

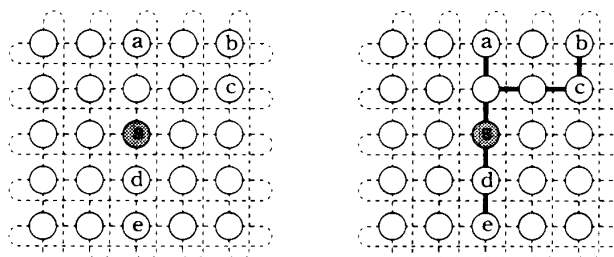


図8 (a)格子形網

(b)結果

は一様なので発信ノードは固定して考える。N ノードからなるネットワークにおいて放送数（宛先ノードの数）f の通信のパターンは $(N-1)C_f$ 通りある。また放送数 f は 1 から (N-1) まで考えられる。このすべてのパターンについて経路設定をシミュレートし、その平均ホップ数を得た。このようにして得られた平均値はトラヒックが一様に起こった場合の値と対応する。

〈放送数と平均ホップ数の関係〉

3 × 3、4 × 4、5 × 5、6 × 6 の格子形網について放送数（宛先ノード数）と平均ホップ数の関係を調べた。縦軸は経路のホップ数を放送数で割って 1 宛先ノード当りの平均ホップ数でプロットしている（図 10、図 11）。

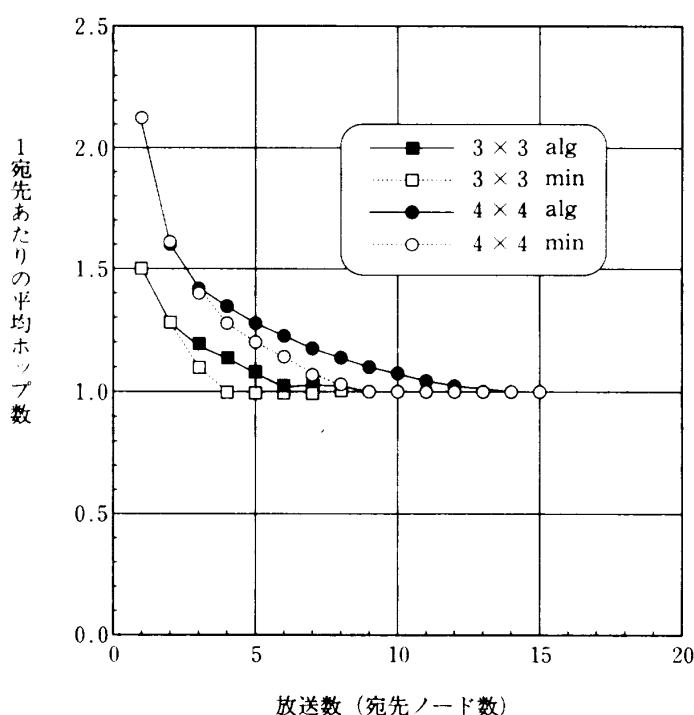


図 10 放送数と平均ホップ数の関係（その 1）
→簡単な計算によって厳密解に近い性能

[評価]

（図 10 から）

重み付けアルゴリズムでは比較的簡単な計算しか行わないにもかかわらず、厳密解に近い性能を発揮でき、その有用性が高いことが示された。1 宛先ノード当りの平均ホップ数が 1 になるのが若干遅れるのは、厳密解では網がトラスであることを活用しているのに対して、重み付けアルゴリズムでは、最短経路しか使用しないため、端と端がつながっていることが意識されないためである。

（図 11 から）

放送数と平均ホップ数の関係はネットワーク規模が変わってもほぼ同じ様な傾向で変化することが予想される。ネットワーク規模に対する特性の変化が緩やかな格子形網に対して、単純なアルゴリズム

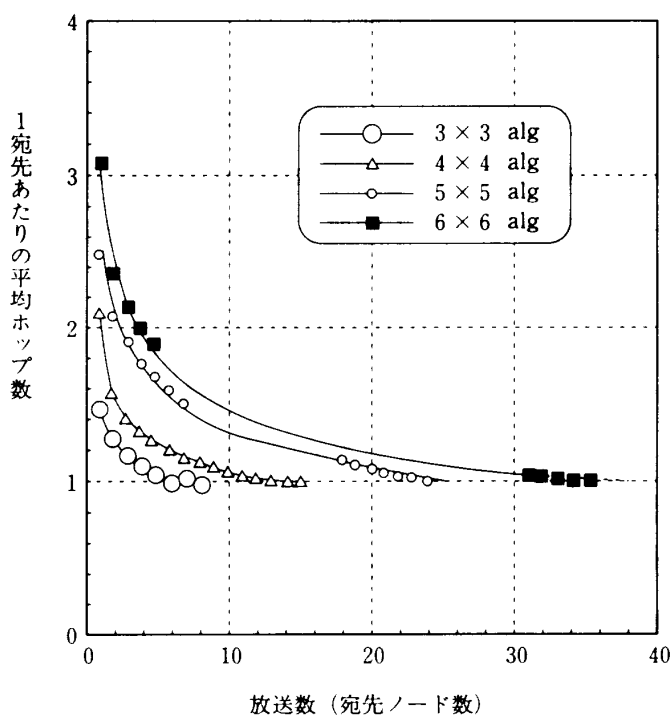


図 11 放送数と平均ホップ数の関係 (その 2)
→ ネットワーク規模によらず同じ傾向

ムを適用したのだから、網規模が変わっても特性曲線が同じ様な傾向を示すのは妥当である。

(全体から)

放送数の増加とともに 1 宛先ノード当たりの平均ホップ数は減少し、最終的には 1 になる。重み付けアルゴリズムによるツリー状経路の制御を行わず、すべて 1 : 1 通信で行った場合は、放送数が増加しても 1 宛先ノード当たりのホップ数は放送数 1 の場合と変わらないため、この曲線の下がり方は、ツリー状経路制御の有無によるトラヒック圧迫度の差をあらわすものと考えられる。したがってネットワークが大きいほど、また宛先ノードが多いほどツリー状経路の利用による網内資源の節約効果は高いことが示された。

4. ネットワークトポロジー

○平均ノード間距離の小さいネットワークトポロジーを目指して Loop 網の構成概念を提案し、平均ノード間距離が最小となる最適形態を求める。

○最適形態 Loop 網をハチの巣網、格子形網、三角形網の代表的な 3 つのネットワークトポロジーと比較し、その特性を評価する。

4. 1 比較対象と評価方法

〈比較対象とするトポロジー〉

(a)ハチの巣形網 (3x3 ノード)

(b)格子形網 (4 × 4 ノード)

(c)三角形網 (6 × 6 ノード)

これらのネットワークの大きさを表すために次のように漸化的に次数 n を定義する。

[1] 基本ネットワークとして1次、あるいは3次のネットワークを定義する。

[2] n 次のネットワークの周囲にノードを配置した形の、ちょうど一回り大きいネットワークの次数を $(n+2)$ 次とする。

[3] $(n+1)$ 次のネットワークは n 次のネットワークを包含し、 n 次のネットワークと $(n+2)$ 次のネットワークのちょうど中間的なネットワークとする。

これは格子形網の $n \times n$ 構成を n 次と考えたときの、次数とネットワーク構成の関係が他の2つの網でも同じようになるように定義したものである。格子形網、三角形網は1ノード構成を1次ネットワークとして定義し、ハチの巣形網は6ノード構成を3次ネットワークと定義したときの、それぞれの網の3、4、5次の形態を示す(図12)。

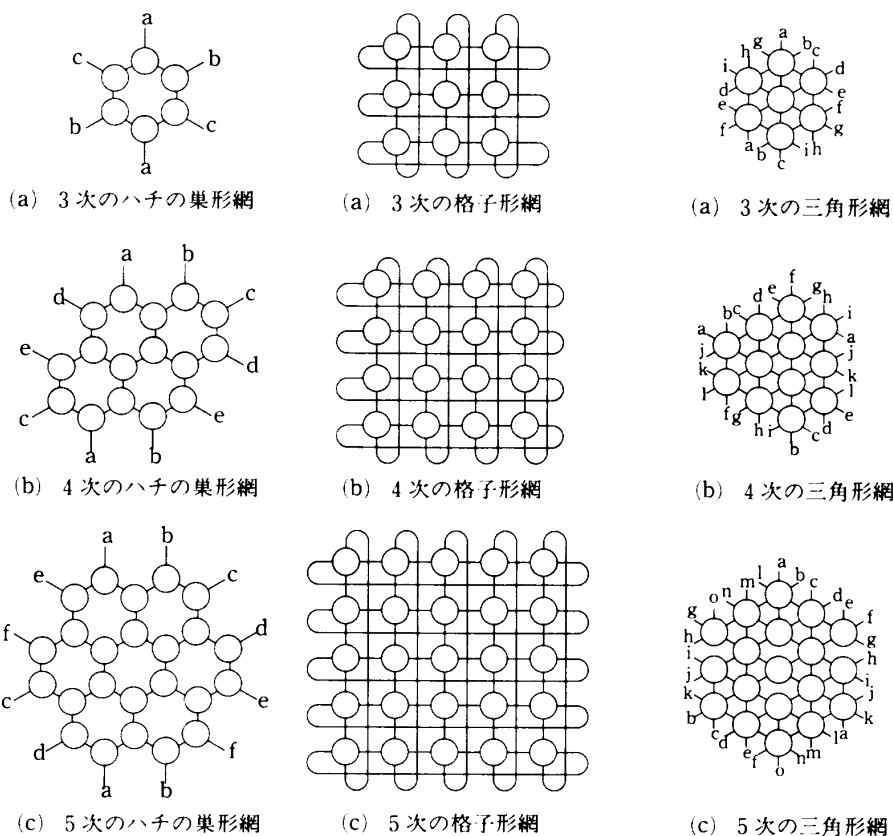


図12 比較対象とする網

〈評価方法〉

通信網への適用を考えたネットワークトポロジーを比較する場合、ノード数、リンク数、ノードの交換機能といったネットワーク資源をそろえて比較することにより、コストパフォーマンスを比較できる。また、ノード間のホップ数が小さいことを目標としているので、評価指標は平均ノード間距離

としてよい。したがって、網を構成するネットワーク資源をそろえたトポロジー同士を比較し、平均ノード間距離が小さいトポロジーが有利であるという観点で評価を行う。

比較対象とする3つのトポロジーのノード数(N)と平均ノード間距離(h)を次数(n)の関数として求めた計算式を表4に、また3次から20次までの実際の数値例を表5に示す。

表4 次数によるノード数、平均ノード間距離の式

網の種類		ノード数 (N)	平均ノード間距離 (h)
ハチの巣	n: 奇	$\frac{3}{2} \cdot (n-1)^2$	$\frac{2n^3 - 6n^2 + 5n - 1}{3n^2 - 6n + 1}$
	n: 偶	$(n-2) \cdot (\frac{3}{2}n+1) + 2$	$\frac{2n^3 - 4n^2 + n}{3n^2 - 4n - 2}$
格子形	n: 奇	n^2	$\frac{1}{2} \cdot n$
	n: 偶	n^2	$\frac{1}{2} \cdot \frac{n^3}{n^2 - 1}$
三角形	n: 奇	$\frac{1}{4} \cdot (3n^2 + 1)$	$\frac{1}{3} \cdot n$
	n: 偶	$\frac{3}{4} \cdot n^2$	$\frac{n^3}{3n^2 - 4}$

表5 3次から20次までの数値例

次数 n	ノード数 N			平均ノード間距離 h		
	ハチの巣	格子形網	三角形網	ハチの巣	格子形網	三角形網
3	6	9	7	1.400	1.500	1.000
4	16	16	12	2.267	2.133	1.454
5	24	25	19	2.696	2.500	1.667
6	42	36	27	3.585	3.086	2.077
7	54	49	37	4.019	3.500	2.333
8	80	64	48	4.911	4.064	2.723
9	96	81	61	5.347	4.500	3.000
10	130	100	75	6.240	5.051	3.378
11	150	121	91	6.678	5.500	3.667
12	192	144	108	7.571	6.042	4.037
13	216	169	127	8.009	6.500	4.333
14	266	196	147	8.902	7.035	4.699
15	294	225	169	9.341	7.500	5.000
16	352	256	192	10.234	8.031	5.361
17	384	289	217	10.674	8.500	5.667
18	450	324	243	11.566	9.028	6.025
19	486	361	271	12.006	9.500	6.333
20	560	400	300	12.898	10.025	6.689

4. 2 Loop 網の提案

Loop 網の構成条件は、各ノードの入出力数(m)が等しいことと、ノードを矩形 (r 行 \times c 列) に並べられることだけであり、与えられたネットワーク資源から網を簡単に構成することができる。以降 r 行 c 列の Loop 網を $r \times c$ と表記する。

<Loop 網の構成条件>

- (1)各ノードの交換機能 $m \times m$
- (2)ノードの総数 $T = rc$ (r, c は自然数) と表せる
- (3)リンクの総数 $L = mT$

処理の簡易性、網としての効率を考えると、次の条件も満たすことが望ましい。

- (4)行の数 $r = km$ (k は自然数)

さらに処理の簡易性を求める場合は次のより強い条件が加わる。

- (4)行の数 $r = m^k$ (k は自然数)

その理由は、Loop 網ではすべての処理が m を基数として行われるためである。

以降では少なくとも条件(4)も含めた構成条件を満たす Loop 網を対象として議論を進める。

<Loop 網のトポロジー>

ノード間のリンク接続方式は各列間で同様である。第 1 列の出リンクは第 2 列の入リンクに、第 2 列は第 3 列に、というように接続し、第 c 列は第 1 列に接続して全体として帯状のネットワークになる。情報はこのネットワークを一方向に流れる。

説明のために i 行 j 列のノードを $n_{i-1, j-1}$ と表記する (図 13)。

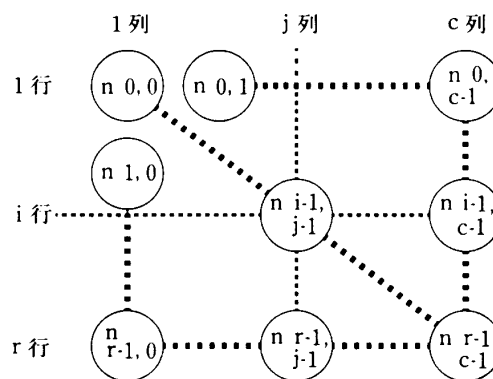


図 13 $r \times c$ の Loop 網

[Loop 網のリンク接続方式]

$n_{0,0}$ の m 本の出リンクは $n_{0,1}, n_{1,1}, \dots, n_{m-1,1}$ の入リンクに接続する。

$n_{1,0}$ の m 本の出リンクは $n_{m,1}, n_{m+1,1}, \dots, n_{2m-1,1}$ の入リンクに接続する。

...

と順に接続し、 $n_{r-1,1}$ まできたらまた $n_{0,1}$ に戻る。

一般に $n_{i,j}$ の m 本の出リンクは

$$n_{\{(i \cdot m + k) \bmod r\}, \{(j + 1) \bmod c\}} \quad (k = 0, 1, 2, \dots, m-1)$$

の m 個のノードの入リンクに接続する (図 14)。

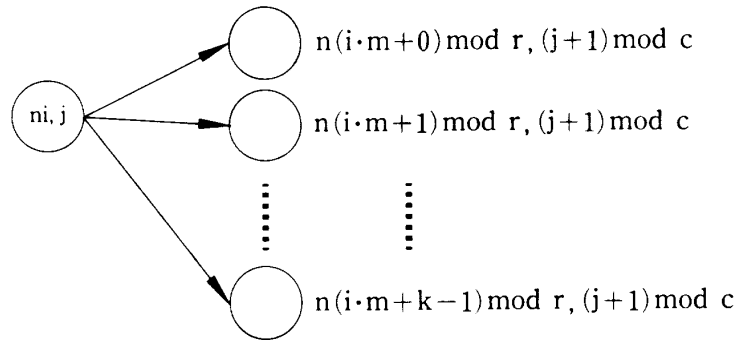


図 14 Loop 網のリンク接続方式

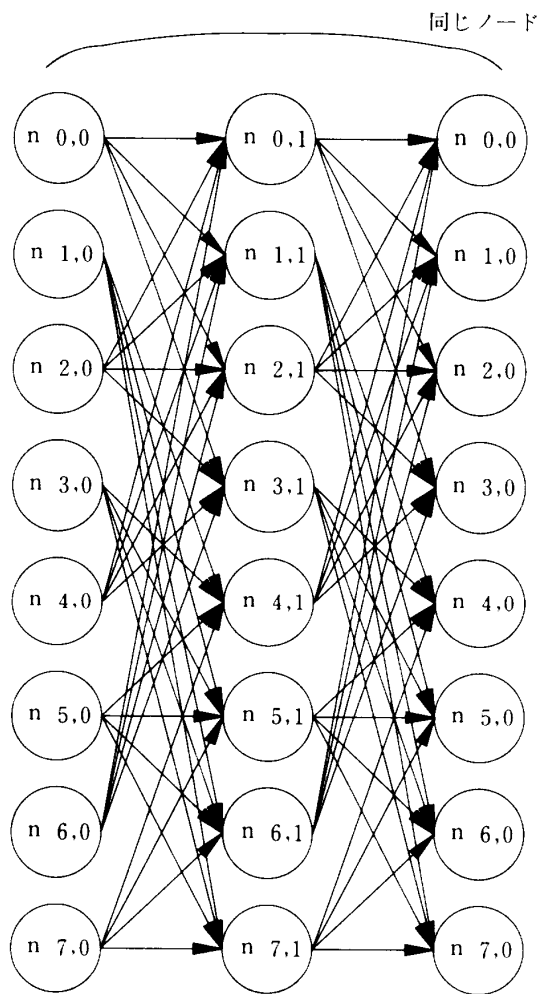
Loop 網の各ノード間の接続方式は Shuffle-Exchange 系の網とほとんど同じ思想をもっているため、1 ノード通過するごとに m 方路に拡散する経路を得られ、非常に効率よく経路が拡散し、平均ノード間距離が小さくなることが予想される。ここで 2 つの例を示す。

例 1 4×4 ノードからなる $8 \times 2 = 16$ ノードの Loop 網 (図 15 (a))

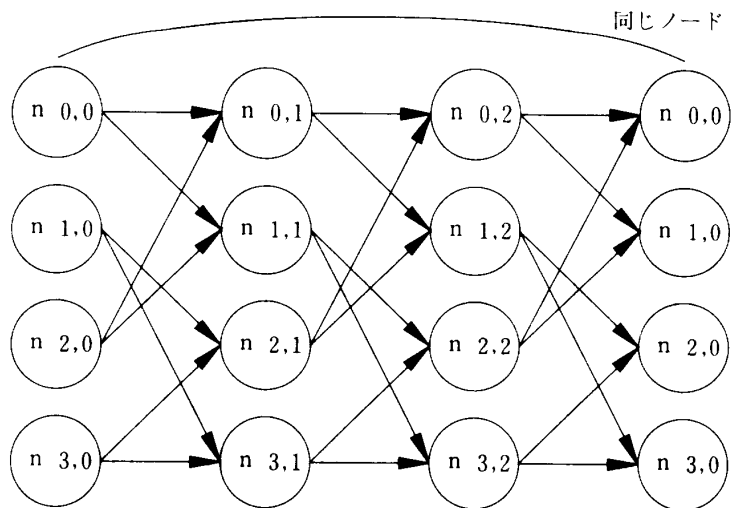
例 2 2×2 ノードからなる $4 \times 3 = 12$ ノードの Loop 網 (図 15 (b))

<最適形態 Loop 網>

Loop 網では $T=rc$ を満たす r と c の組み合わせが様々に考えられるが、どのような値にすれば平均ノード間距離が小さくなるかを調べた。 4×4 ノードから構成される Loop 網の、 $T = 2^k$ で表せるものについて、 c を 2、4、8、 \dots 、 2^{k-2} まで変化させ、それぞれの平均ノード間距離を求めた。 $c \geq \log_4 r$ の場合は計算で、 $c < \log_4 r$ の場合は経路設定をシミュレートするプログラムによって平均ノード



(a) 4×4 ノードからなる 8×2 の Loop 網



(b) 2×2 ノードからなる 4×3 の Loop 網

図 15 Loop 網の例

間距離を得た(ノード毎に平均ノード間距離が違うため)。結果を表6に、そのグラフを図16に示す。
 この結果、ノード数 T の値によらず $c=2$ 、すなわち縦長の2段構成にしたときに平均ノード間距離が最小となることが分かった。

表6 4×4ノードからなるLoop網の $r \times c$ の各構成における平均ノード間距離の値

ノード数	$r \times c$	平均ノード間距離
1024	512 × 2	4.813
	256 × 4	5.173
	128 × 8	6.843
	64 × 16	10.182
	32 × 32	17.861
	16 × 64	33.220
	8 × 128	64.938
	4 × 256	128.375
512	256 × 2	4.161
	128 × 4	4.845
	64 × 8	6.184
	32 × 16	9.863
	16 × 32	17.221
	8 × 64	32.939
	4 × 128	64.376
256	128 × 2	3.828
	64 × 4	4.188
	32 × 8	5.867
	16 × 16	9.224
	8 × 32	16.941
	4 × 64	32.377
128	64 × 2	3.185
	32 × 4	3.874
	16 × 8	5.228
	8 × 16	8.945
	4 × 32	16.378
64	32 × 2	2.873
	16 × 4	3.238
	8 × 8	4.952
	4 × 16	8.381
32	32 × 2	2.258
	16 × 4	2.968
	8 × 8	4.387
16	8 × 2	2.000
	4 × 4	2.400

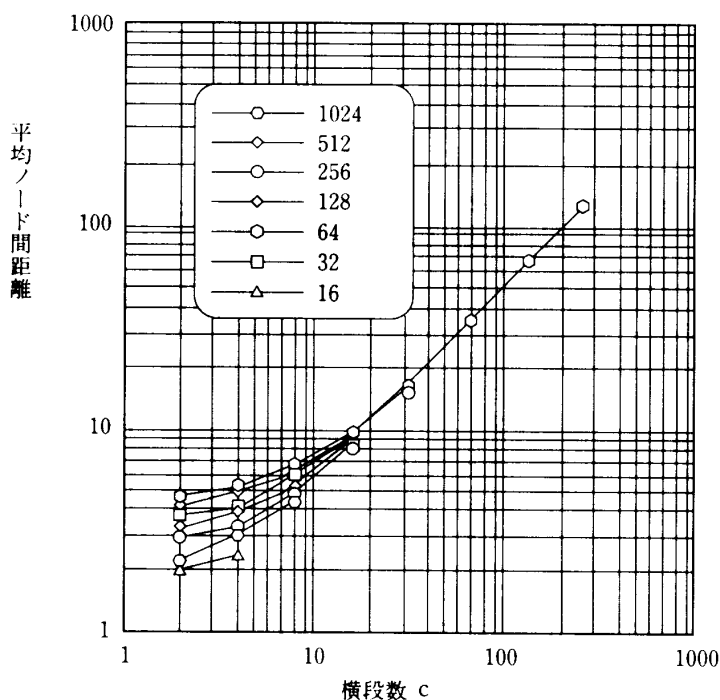


図 16 rXc の各構成における平均ノード間距離

一般に $m \times m$ のノードからなる Loop 網でも最適形態は 2 段構成の $r \times 2$ の形であることが予想される。そこで 3×3 、 6×6 のノードからなる様々なノード数の Loop 網について $c = 2, 3, 4$ と変化させたときの、ノード $n_{0,0}$ から見た平均ノード間距離を調べた。これは最悪値を示すものである。結果を表 7 に示すが、やはり 2 段構成の Loop 網が最も平均ノード間距離が小さくなっている。

以上から、任意の $m \times m$ のノードからなる Loop 網に対して、平均ノード間距離が最小となる最適形態は $c = 2$ の 2 段構成であるという結論を得た。

〈Loop 網の制御〉

Loop 網に対しても重み付けアルゴリズムは適用できるが、各ノードに用意する最短経路のホップ数とリンクの方向情報の表の作り方がやや複雑な手順を必要とする。

4.3 他のトポロジーとの比較

最適形態の Loop 網と 3 つの比較対象網の比較を行い、Loop 網の優位性を示す。

〈比較方法〉

各比較対象網に合わせて Loop 網も 3×3 、 4×4 、 6×6 を基本ノードとして 2 段構成の最適形態を構成し、経路設定をシミュレートするプログラムによって平均ノード間距離のデータを得た。同じネットワーク規模における比較網と Loop 網のネットワーク資源は等しい。

表7 3×3、6×6ノードからなるLoop網の
n 0、0から見た平均ノード間距離

ノード数	r × c	3×3ノード	6×6ノード
1 2	6 × 2	2. 0 0 0	1. 4 5 5
	4 × 3	2. 1 8 2	1. 9 0 9
	3 × 4	2. 3 6 4	2. 3 6 4
2 4	1 2 × 2	2. 5 2 2	2. 0 0 0
	8 × 3	2. 6 0 9	2. 2 1 7
	6 × 4	2. 9 5 7	2. 4 3 5
3 6	1 8 × 2	2. 8 5 7	2. 1 7 1
	1 2 × 3	2. 4 8 6	2. 4 8 6
	9 × 4	3. 1 4 3	2. 8 0 0
4 8	2 4 × 2	3. 0 2 1	2. 2 5 5
	1 6 × 3	3. 2 5 5	2. 6 1 7
	1 2 × 4	3. 4 8 9	2. 9 7 9
7 2	3 6 × 2	3. 4 3 7	2. 3 3 8
	2 4 × 3	3. 5 0 7	2. 7 4 6
	1 8 × 4	3. 8 3 1	3. 1 5 5
9 6	4 8 × 2	3. 7 0 5	2. 6 3 2
	3 2 × 3	3. 7 8 9	2. 8 1 1
	2 4 × 4	4. 0 0 0	3. 2 4 2
1 2 0	6 0 × 2	3. 8 6 6	2. 8 0 7
	4 0 × 3	4. 0 3 4	2. 9 5 0
	3 0 × 4	4. 2 0 2	3. 2 9 4
1 4 4	7 2 × 2	3. 9 7 2	2. 9 2 3
	4 8 × 3	4. 1 9 6	3. 1 2 6
	3 6 × 4	4. 4 2 0	3. 3 2 9
4 3 2	2 1 6 × 2	4. 9 5 1	3. 3 0 9
	1 4 4 × 3	5. 1 7 2	3. 7 1 0
	1 0 8 × 4	5. 3 9 2	4. 1 1 1

〈ネットワーク規模と平均ノード間距離の関係〉

Loop網と3つの比較網についてネットワーク規模と平均ノード間距離の関係を図17、図18、図19に示す。図20はそれらを1枚にまとめたものである。

凡例の中で $\langle \log_m N \rangle$ ($m=3, 4, 6$)はノード数Nに対して、底がmの対数をとった値である。

[評価]

結論的に各データは以下のようなノード数Nの関数曲線にほぼ一致して変化している。

(a) ハチの巢形網との比較 (3×3ノード)

$$\text{ハチの巢形網} \quad h = \frac{2\sqrt{6}}{9} \sqrt{N} = 0.54\sqrt{N}$$

$$\text{Loop網} \quad h = 0.9 \log_3 N$$

(b) 格子形網との比較 (4×4ノード)

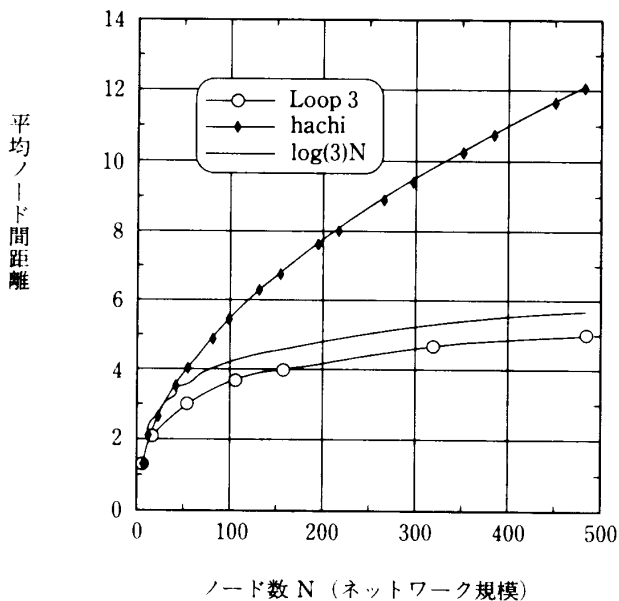


図 17 ハチの巢形網との比較

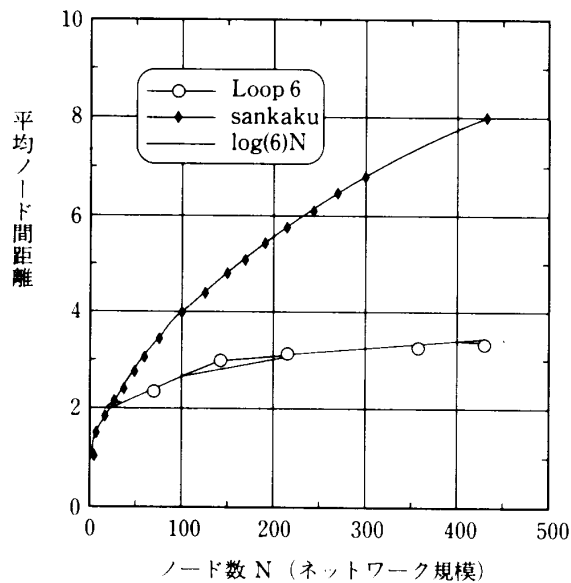


図 19 三角形網との比較

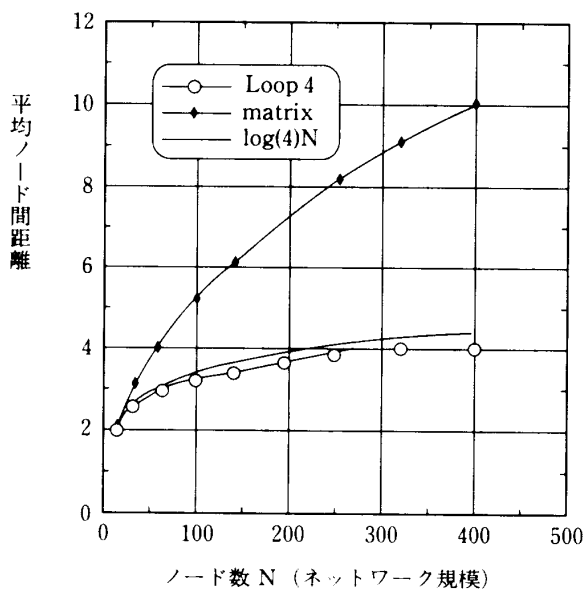


図 18 格子形網との比較

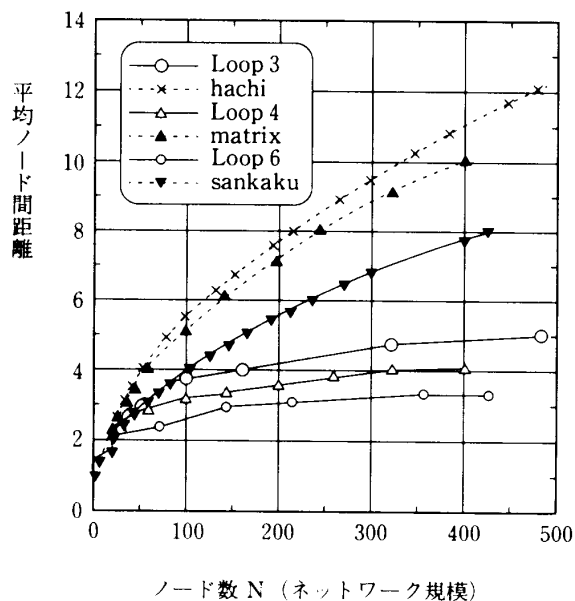


図 20 3つの比較網との比較

格子形網 $h = \frac{1}{2}\sqrt{N} = 0.5\sqrt{N}$
 Loop 網 $h = 0.95 \log_4 N$

(c) 三角形網との比較 (6 × 6 ノード)

三角形網 $h = \frac{2\sqrt{3}}{9} \sqrt{N} = 0.38\sqrt{N}$
 Loop 網 $h = \log_6 N$

いずれの比較網に対しても Loop 網の優位性が示されたばかりでなく、 $m \times m$ ノードからなる Loop 網の平均ノード間距離は $\log_m N$ の形で変化することが示された。

〈格子形網と Loop 網の放送形通信による比較〉

256 ノードの格子形網 (16×16) と Loop 網 (128×2) について、放送数 1 と 2 について重み付けアルゴリズムによってホップ数を計算し、あるホップ数をとる経路パターンの割合がわかるように分布図を示す (図 21)。図中矢印で示したところが平均値である。

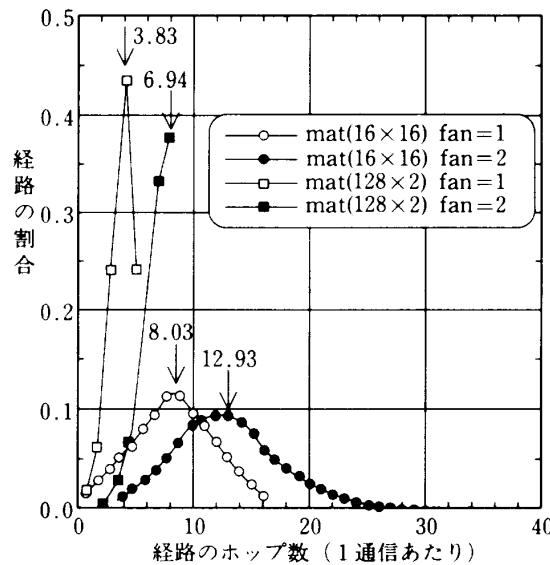


図 21 256 ノードの格子形と Loop 網の比較

[評価]

平均値での比較では Loop 網は格子形の半分程度のホップ数になり、256 ノードの Loop 網は格子形の 2 倍程度のトラヒックまで扱えることを示す。また分布についても Loop 網は局所的に固まっているのに対して、格子形は裾野の広い分布を示しており、これは通信パターンによってホップ数が大きく変わることになる。結局どちらからも Loop 網の優位性がいえる。

4.4 Loop 網のそのほかの特徴

〈不均一なネットワーク〉

2 段構成の Loop 網 (厳密には $c < \log_m r$ のとき) では、ノードによって他の全ノードにいたるツリ一状経路の形が異なるため、平均ホップ数が少し異なる。しかしその差は非常に小さいため、あまり大きな問題とはならない。

(例) 先の 256 ノード Loop 網の場合

平均ホップ数	ノードの割合
3.812	50.0 %

3.843 37.5 %

3.851 12.5 %

〈特殊なリンク接続方式〉

普通のネットワークではノードとノードの間は一方通行ではなく、入出力リンクがペアで接続されることが多い。これに対し Loop 網ではあるノードからは入力のみ、またあるノードへは出力のみという接続方式を持つ。ただし、 $r=m+1$ 、 $c=2$ の関係が成り立つ場合は、各ノード間に双方向のリンクが接続される。

〈ネットワークの拡張性〉

格子形（より一般的には矩形）ネットワークではノードの追加は縦 1 例あるいは横 1 例といった一部のリンクの付け換えで対応できるが、Loop 網に新しくノードを追加することはそれほど簡単に扱えない。この点については第 5 節で詳しく述べる。

5. 実用上の諸検討

本研究の議論を実際にネットワークに適用するための検討と、本研究では十分にふれることのできなかった問題について説明し、最後に本研究の適用領域について述べる。

5. 1 制御およびアルゴリズムについて

〈宛先メンバの変更に対する対応〉

（問題） CATV などを受信者が番組を探しているような時に、いちいち全経路の再設定を行うのは無駄であり、また時間的にも厳しい要求が出るため好ましくない。

（対応） 新たに加わったメンバに対しては通常の 1 : 1 接続を行って情報を流し、ある時間が経過するか、継続受信の意志表示があった場合のみ、そのメンバを 1 : N 通信に組み込み放送形経路を変更する。このような処理によって経路の再設定を余裕を持って行えるという利点も出る。

（課題） 組み込むタイミングを検討する必要がある。また $1 : N \rightarrow 1 : (N+1)$ と $1 : N \rightarrow (1 : N) + (1 : 1)$ のトラヒックの差も検討が必要だが、ホップ数については本研究のデータから計算できる。

〈重み付けアルゴリズムの発展性〉

（問題） 最短経路はトラヒック要求もトポロジーも一様な場合に、平均的に効率のよい経路である。実際にはトラヒックは一様ではないことが多く、分布も時間的に変動することが予想される。そのため、トラヒック状況を考慮したアルゴリズムが求められる。

（対応） リンクの混雑度が大きいほど、重みが小さくなるように係数付けを行うことで、混雑しているリンクを後回しにするアルゴリズムに発展できる。また次に短い経路 = 第 2 候補による重み付けを行う方法も考えられる。

（課題） どのような係数付けを行えば、トラヒックを分散できるかという問題と、速度可変の ATM 交換方式において、通信品質をどのように形式化するかという問題である。

5. 2 Loop 網の拡張について

〈列単位の拡張〉

(問題) 行単位の拡張はメリットがない(網全体のリンクの張り替えが必要、 $r=m$ の階乗の関係が崩れる可能性)。列単位の拡張はリンクの張り替えは増設を行った列だけだが、最適形態 = 2 段構成が崩れる。3 段、4 段にした時の質的劣化、最適形態との“ずれ”を調べる必要がある。

(対応) 最適形態を求めたときのデータから $\times 2$ 、 $\times 3$ 、 $\times 4$ の値を比べることで、増設に伴う平均ノード間距離の増加量、トラヒックの増加量を知ることができる。

最適形態 Loop 網に対しては 1 列、あるいは 2 列の増設まで、すなわち元の 2 倍まではそれほど品質を落とすことなく拡張が可能である。

〈網単位の拡張〉

(課題) 複数の Loop 網を結ぶことによる網の拡張では、どのように接続するかという問題が多くのパラメータを含んだ検討項目。また各網間のルーティング制御も考える必要があるため、先の列単位の増設に比べ格段に取扱が難しい。しかし Loop 網をリングネットワークとして捉えると、リングネットワークでは網単位 = プレーン単位の増設を考えることが多いため、このような増設方法についても検討が必要である。

5. 3 適用領域

本研究はすべて均一なトラヒック、トポロジーを前提とした平均値の議論を行った。また Loop 網の性能を十分に発揮させるためには、あらかじめ必要なノードを用意し 2 段構成で接続することが必要となる。

従って、本研究の議論を実際に適用する場合には、ある程度トラヒックが均一に発生し、はじめから必要なノードの数が決まっていることが望ましい。例えば、企業内通信網や、MAN (Metropolitan Area Network) などに対する適用が考えられる。ただしそれら同士を接続すると、ゲートに当たるノードを中心としてトラヒックの集中が起こることは必死であるから、やはり均一でないトラヒックに対する議論が必要となる。

6. 結論

本研究では、高速パケット通信網における放送形トラヒックの制御を様々な観点から検討した。特に、経路のホップ数はトラヒック量に比例するという観点から、これを小さくするための新たな方式をルーティングアルゴリズムとネットワークトポロジーについて提案し、その比較評価を行った。提案した“重み付けアルゴリズム”は最短経路を組み合わせるという簡単な処理により、真にホップ数最小となる厳密解に近い性能を実現できることを示した。また“Loop 網”は従来のトポロジーと同じネットワーク資源から、より平均ノード間距離が小さく、通信パターンによるバラつきも小さい網を構成できることを示した。最後に、本研究の議論を実際のネットワークに適用するための諸検討を行った。

高速パケット通信網による高品質動画像も含めた放送は、より高度な情報化社会のために、また広帯域 ISDN のフレキシブルな対応性を活かして魅力的な通信網を構築するためにも、是非とも実現が望まれる技術である。本研究がこれからの研究の一助となれば幸いである。

研 究 論 文

日米の研究開発体制を考える

On Research and Development Activities in U.S.A. and in Japan :
a personal view

学術情報センター 山田 尚勇

要旨	139
1. はじめに	141
2. 科学と技術	142
3. 製品化技術重点主義	143
4. 社会資本の不足と基礎研究	145
5. 均質社会と教育	146
6. 均質社会と独創性	148
7. 大学の環境	150
8. 若手の研究者をとりまく環境	151
9. 和の原理と個性の原理	152
10. 創造性養成への道	154
11. 学術情報センターに関連して	156
12. 執事の研究から王侯貴族の研究へ	158
13. おわりに	159
参考文献	160

要 旨

本稿においては、研究開発に関わる統計的分析を示すのではなく、日米両国において、大学および企業の研究所で長らく過したあいだの、個人的経験を述べる。

日本では科学と技術がほとんど区別されず、一体として理解されている。しかし、かりにその差を認めるときには、どちらかという、技術者のほうが科学者よりも地位が高いとする伝統がある。しかるに米国においては、科学は自然の法則を理解（分析）すること、工学はそれを実利に向けて応用（合成）することという、よりはっきりとした区分がなされている。しかも、基盤となる科学に関わる者のほうが、工学に関わる者より、常に高い評価を受けてきた。

こうした微妙な差異は、科学と技術に関わる2国の政策に、かなりの違いを生んでいる。米国が基礎研究と新しいアイデアの発見に力を入れてきたのに対し、日本では既成のアイデアと技術を導入し、それらにキメ細かい改良を加えた上での製品化技術に集中し、勤勉でレベルが高く、しかも質の揃った技術者、労働者の効果的活用により、高品質、高信頼性の製品を大量生産し、安価に世界市場に供給しつつ、工業立国の面目を発揮してきた。

そうした国策は、近年大幅な貿易黒字をまねくとともに、主として製品の信頼性と価格の差により、米国におけるいくつかの産業を極度に圧迫している。そうして起こった貿易摩擦、経済摩擦の結果、日本が基本的アイデアを生み出すこと少なく、もっぱら他国に頼りつつ、甘い汁を吸っているという、技術タダ乗り論が海外に台頭し、高度先端技術の日本への移転を制限する運動さえ起こりだした。

こうした国際状況の中であって、日本としては、貿易黒字減らしと基盤的創造性の養成に、いやおうなしに取り組まなければならなくなった。

改めて世界を見まわしてみると、日本製品の廉価供給を可能にした原因として、1人あたりの国民総生産が世界一であるにもかかわらず、日本における実質的生活水準の低さが目につくようになった。すなわち、世界において飛び抜けて多い年間労働時間数、流通機構の過保護による世界一の物価高、過少な社会資本投資の結果としての生活環境の貧困さなどである。そして、国外における日本製品の価格の低廉さは、こうした犠牲の上に可能となっているという国際的指摘がなされた。

さらに、もともと基礎研究というものは、豊かな「科学資本」つまり教育、設備、試験研究、知識の集積などに経費のかかるものである。日本は外国のアイデアにタダ乗りして、こうした資本の投資をもおこなっているというのである。

こうした国際的緊急状態に対処するために、日本は基礎研究に力を入れるとともに、創造性の組織的開発に取り組み始めた。

独創性を発揮するのは平均的思考能力を持つ人たちではなく、他人と変わった、独自の思考をする少数の人たちの集団であることが多い。

しかるに日本の社会では、単一化、画一化を陰に陽に奨励する文化が長らく定着しており、個性の強い者、変わり者が自由に伸びていくにはさまざまな障害が多い。

教育も、幼稚園から高校に至るまで、例外ではないから、大学にはいったから、にわかには個性を発揮しろと言われても、もともとそうした素質を持った人間は、それ以前の過程でかなり排除されてしまっている。

さらに日本の教育システムでは、小中学校の教育にかかる経費はかなり潤沢であっても、大学教育、特に大学院にかかる経費は、国際的に見てかなり少ない。かつ、教官に対する制限が強すぎて、かれらには自由に過ごせるまとまった期間が少なく、企業で働けず、またほとんどが所属大学の内部育ちであり、人事の交流も少ない。これらの環境条件は、またもや研究者の画一的思考を助長することになる。

若手の研究者についても、大学院生を教育、研究過程に積極的に組み込むことが法的にできなくなっており、また、若手の社会人に対する「生涯教育」も思うにまかせない。積極的に創造性を評価するのに臆病であるから、長老に対する功労賞は多くても、若手に向けた大きな功績賞はない。

幼いときから、自分の意見をはっきりと相手に主張し、相手と議論してでも意見を伝えることを教えこまれているアメリカ人に比べて、われわれの行動の主原理は和の精神であり、右顧左眄しつつものを言う会議は、とかく生産的でない。委員会なども初めからとかく同種意見の持ち主で構成されるから、画期的な結果が出にくく、討論会などでも、本来の「秩序ある対決」の精神が生かされない。

個性とアイデアとは切っても切れない関連があるから、個人の確立のないところでは、アイデアそのものを尊重する観念も希薄になる。それはアイデアを出した個人の報われかたにおける日米の差によく表われている。

こうした独創性の低さ、そしてそれに付随した、外国からの先端技術タダ乗り論の非難を解消するには、国として早急に創造性養成の施策を進めなければならない。それについても多くの提案があるが、どうも対症処置に終わっていて、原点に立ち戻って考えなおすことは極力避けているかに見える。それに、対症処置そのものが、また画一的、均質的で、別の硬直化を起こしそうなものが多いようだ。

ほんとうに創造性を上げようと思うのなら、国民全体の均質性、高水準性は多少犠牲になるかもしれないが、幼稚園教育から始めて、個性の自由な発露と、それに伴う多様化を、極力推し進めていくことが必要なのではなかろうか。

そうした教育制度の思いきった改革は、なかなかむずかしいと思うが、それなしには、ただでさえ多額な投資を必要とする基礎研究に研究費を注ぎ込んだとしても、独創的な成果は、なかなか思うようには出て来ないのではなかろうか。

わが国の研究の理念は、あまりにも実用的、工学的思考に傾きすぎている。たとえば気球との関わり方の歴史をみても、軍事的効用が認識されるまでは、わが国での反応は実に冷淡であったようだ。

わが学術情報センターにおいても、現実的、短期的な研究開発と並んで、将来、通信やコンピュータのコストがほとんどタダになったときに、たとえば、全世界から絶えず流されてくる科学技術情報から、研究者が、自分でキーワードで設定した、関心のあるテーマのプロファイルに合わせて、それに合致する文献、データ、画像などを常時拾い出しては、個人用のジャーナルを編集し、提供するようなシステムの開発などを行なえば、夢と実利があって、たのしいと思う。

科学における飛躍は、好奇心に駆られた、王侯貴族のお遊びの精神から始まることが多い。それに対しわれわれの実利主義は、執事、従僕、召し使いの立場に立つ感があり、他人のアイデアを磨き上げ、ものにしていくことに専念しすぎている。

1. はじめに

本稿の目的は、詳しい数値的調査結果に基づいた統計的分析など(たとえば柚木 1973、1975、1977)を示すことではない。むしろ、日本とアメリカとの2国で約半分ずつを、社会人として過した体験に

基づいて、あれこれと個人的な考えを述べてみることである。

アメリカで私は、教育や研究開発に関係しつつ、合わせて20年間過したので、むしろ日本の事情のほうが暗いかもしれないが、直接経験したことや見聞きしたことを主にして、この二つの国の研究開発事情について、もっぱら相違点を中心にして述べたい。

現在、日本は工業で飛躍的な進歩を遂げ、統計によれば、世界全体のGNP（国民総生産）の15%を引き受けている。人口は世界の40分の1ほどだから、非常にいい位置にあり、1人当たりでは、為替相場にもよるが、いま世界一になっている。

このGNPには、あらゆる生産活動が含まれているから、たとえば大地震や台風で壊れた橋などを架け直す費用も全部はいつてくる。だからGNPが大きいことが必ずしも、そのままわれわれの生活の豊かさを示す指標になるとは限らないにしても、それなりに反映していることは間違いない。

特許保有件数を見ると、アメリカが世界で一番たくさん持っているし、毎年の出願件数も一番多い。そのアメリカで、いま外国からの出願件数が著しく増えてきて、全体の40数%を占め、またそのうちの35%が日本からの出願である。これは、外国からの出願では、明らかに第1位である。

人口では、日本はアメリカの約2分の1だが、技術者・科学者は、学士の数でみるとアメリカの1.5倍ほどである。人口比にすれば、アメリカの3倍ぐらゐの技術者・科学者がいるわけである。また、現在アメリカでは毎年の大学卒業人口の1000人につき約7人しかエンジニアは出ていないが、日本では約40人もいる。

かつて日本は科学技術立国を宣言し、何十年か當々と努めてきた。やっとその念願を果たしたと言えるところに来たようだ。

1970年に第1回が行なわれ、その後、15年ぐらゐごとに1回行なわれることになっている国際理科教育調査の第2回が、1984～5年に行なわれた。22か国が参加したもので、その統計でみると日本の小学校5年生の理科教育は、世界第1位だという。それが中学校3年生になると第2位、高等学校3年生では第7位と、上に行くほど落ちてくる理由については、だんだんとあとで触れることにして、総合的に見れば、世界第1位だということである。

これらのこともあって、数年前、日本のある企業家たちが「もうアメリカに学ぶものはなくなった」という主旨のことを言ったというが、果たしてどうなのか、少し考えてみたい。

2. 科学と技術

日本では一般に科学技術と呼ばれ、この二つが異質のものだという自覚が少ない。それに反して、アメリカではこの二つはかなりはっきりと区別されるのがふつうである。

科学の目標は自然界における諸過程を理解し、そこに存在する法則を見いだすことにある。一方、技術、つまり工学の目標はこうした自然界の諸過程を、たとえ充分理解できないにしても、かなり良く予測し、制御することにより、人間の能力を延展することにあるとされる。

むかし寺田寅彦や橋田邦彦の時代には、日本でもこの二つは判然と区別して考えられていたという

話がある（飯田 1989）。そして、この二つが日本で区別されなくなったのは、輸入技術を脱却して、自主技術を根づかせてからのことだという。

その理由は、科学的原理がきちんと理解されなければ、ほんとうの工学的活動はできないという考え方が定着したからだという事らしい。もしそうだとすると、いまだにアメリカではエジソンの、うまいアイデアを何とか生かして製品にするという考え方で、工業が運営される傾向が強いということになる。

たとえアイデアは外国から導入したとしても、信頼性においてはぐっと優れた工業製品は日本が生産している現状を、この説はよく説明している。

だが上の理由付けには少しおかしいところがあるように思う。明治の開国以来、日本は実用技術の導入に熱心で、むしろその基礎となる科学は、ともすると忘れられがちであった。これではならじと科学の確立の重要性を説いたのが、寺田、橋田を含む、見通しのよい科学者たちであったというのが真相だと思う。

戦後になって、アメリカにおける科学優先と日本における技術優先の差が工業化の差となってあらわれたものに、トランジスタの応用がある。トランジスタは 1947 年の暮、アメリカのベル電話研究所で発明された、というよりは現象として発見され、1951 年には商品化されたが、まだ信頼性が低く、なかなか部品として組み込むわけにはいかなかった。わずかにその小型化能力が買われて、補聴器に使われた外は、例外的にトランジスタを使った脳波計が、1956 年になって、やっとオフナー・エレクトロニクス社から発売された。アメリカ工業技術の先端を行く MIT でもまたメーカーでも、そのころはまだトランジスタをコンピュータに用いようとする努力はほとんどされていなかった。したがって、トランジスタの実用化は 1958 年に日本の東京通信工業社（今のソニー）が小型ラジオに使用したことによって始まったと言える。

両国におけるこの考え方の違いは、それから三十数年経った 1987 年にも、超電導の研究において同じパタンを示している。アメリカでの研究が超電導の理論の解明に主力を注いでいるのに比べて、日本での研究は実用のための材料の発見と改良に主力が向けられ、アメリカをしのいでいる。

こうしたいくつかのことからも分かるように、日本ではいまでも依然として技術優先の哲学が社会の主力であるように思う。

3. 製品化技術重点主義

日本の技術は確かに優秀で、信頼性の高い製品を安く、大量に生産し、輸出し、しかも、アフターサービスもよろしいということで、国際的に日本は経済的、技術的に躍進している。そのお陰で、反面には貿易摩擦なども起き、問題になっている。

それでも、内部の事情を見ると、まだ喜ぶには、ちょっと早すぎるのではないかと、いうところがある。それは、研究開発の面である。英語では「research and development」といい、日本で研究開発というけれども、その内容にはだいぶ差があることが、よく指摘される。これは、私の実感でもあ

る。

IC（集積回路）チップが産業界の米と呼ばれるようになってから、もう久しいが、アメリカの太平洋株式市場では、1990年早々から動的RAMの先物取引の実行を計画していることを発表した。いよいよICが大豆や石油などと肩を並べて先物取引に名を連ねることになるわけだ。

このICチップのうち、日本は、メモリー・チップを世界中に集中豪雨的に輸出した。製品の品質は非常によろしいし、傾段も安いということで、アメリカの製品を追い抜いた。ところが、ICチップの内分けとして見ると、これはまさにメモリー・チップだけの集中豪雨的輸出である。

1985年の12月にアメリカで開かれたスーパーコンピューターの国際会議では、日本からの代表が、よそよそしい目で見られていた。かつてアメリカに長くいたときの知り合いの何人かと会場で顔を合わせたとき、昔と変わらない調子で、ザックバランな話し合いをすることができたが、彼らは口をそろえて、こう非難した。「日本はけしからん。ICを作って売るのはいいが、もうかりそうなところしか作らない。われわれアメリカのほうは、もうかるのも、もうからないものも、全体的な視野に立って作り、もうかるところでもうけて、もうからないところを補っている。日本は、もうかるところだけ食い荒してしまって、われわれの利益を奪い、しかも、もうからないところを放置している。」

最近また話題になったのが、やはりメモリーである。実用化そのものは、できたとしてもまだだいぶ将来の話らしいが、超低温で、非常に大量の記憶ができるメモリーの技術が、初めて日本で実用化に向けての実験に成功した。いままでのものに較べて、格段に記憶情報量が増えていて、1 cm²の中に1 ギガ (10⁹) ビットの情報が蓄えられるというものである。日本の新聞の4万ページ分の情報が、この1 cm²に載ってしまうことになる。4万ページでいうと、ざっと4年分の新聞である。

ちなみに、非常に大ザッパだが、日本で1年間に生産される情報量は、3 × 10¹⁷ 語だという推定がある。平均して、16ビットで漢字1字、漢字2字で1語（ワード）になるから、日本人1人当たりになると10⁹ 語、あるいは3 × 10¹⁰ ビットである。つまり、日本人1人当たりの生産量が10 cm²のチップにはいる。

実際にそんなに情報が生産されているのかと思うかもしれないが、これには、画像情報もはいつている。たとえば、テレビの画面の情報量はものすごく多いものである。とにかく、1年間に生産される情報を1人当たりで見ると、10 cm²、すなわち、3 cm角ほどのところに入ってしまうというしろものである。

これも、原理的には、ソビエト人が最初に考えたものだそうであり、IBMの研究所で理論化され、可能性が指摘されていたものを、日本で初めて実用化のメドをつけたわけである。

この辺のところに行くと、どこもやっていなかったことだから、大いに威張れるということになるかと思うが、やはり、基礎のところは外国でやられたあとに、それを実用化するという、わが国の技術立国の面目が躍如としている。だから、アメリカ人からは、これも「また、してやられた」と思われる材料になる可能性が大きい。

そういうわけだから、ある特定の分野の小手先技術については、ひょっとすると、いまのところも

う外国から何も学ぶものはないということもあり得よう。しかし、国際摩擦を回避するためには、広くみると、まだまだ日本の科学技術は片寄りがありすぎる。研究開発、とりわけ研究の分野においては、日本はまだこれから強化していかなければならない側面がたくさんあると思われる。

4. 社会資本の不足と基礎研究

まず、われわれの実感として、日本がGNPで世界1だとは思えないような現象がたくさんある。住宅事情がまだまだ悪いとか、労働時間が先進国の中では、飛び抜けて長いとか、いろいろある。欧米の労働時間が1年に1800時間ぐらいなのに、日本は2150~2200時間と、20%以上長い。最近、景気が上向いてきたので、かえって長くなっているところさえあるという。実感としては、必ずしも生活水準が上がっていない。

働いて何が悪いか、という開き直りもあるようだけれども、その辺が非常にむずかしい、生活哲学というのは、広い意味での宗教に属する。アメリカに暮らしてみると、まず、アメリカ人が生活を非常に大事にするということがわかる。彼らは、それを崩されることを極度に嫌う。これは、いい悪いの問題ではなくて、それを信念としているということである。だから、それを変えるような出来事は、何事であれ、宗教的確信を持って拒絶する。

「日本では、たとえば社会資本の投資が不十分で、基礎的な研究にもお金をあまり使わず、われわれが見つけたものをうまく利用するし、われわれの産業まで空洞化させつつある」ということで、かなり不満に思っている。事実、日本の社会資本額をアメリカ並みの水準に持ち上げるのには、1人あたり少なくとも2千万円以上の投資がいるという推定もあるほどだ。逆にいうと、それをしないことによって可能となっている安価な日本製品と競争することは、アメリカの生活の豊かさを下げることになると彼はら思っている。アメリカ人にとって、それはとうてい受け入れられないことだ。

基礎研究についても同じようなことが言えそうだ。だからわれわれは、これから社会資本の投下はもちろん、基礎研究のほうにも力を入れていかなければならないと思われる。

現在、日本では、とりわけ基礎研究や、そのもとになる大学院教育には、あまりお金を使っていない。その結果、企業にはいった人も含めて、アメリカへの留学生が毎年1万5000人ぐらいある。平均1人が2~3年は滞在するから、実際には数万人の大学院生が、常時アメリカで勉強している。

また、アメリカの主要大学で経営学修士(MBA)のコースに、たとえば1989年に入学した学生の4.6%、つまり22人に1人は日本人である。これは外国人学生の約4分の1で、諸外国の筆頭である。

広くみると、研究開発は経営学を離れて語ることができず、また経営学は研究開発を除外しては成り立たない。そうした意味でも、日本の研究開発はまだアメリカに負うところが大きい。

それに対して、日本はいま1年に1000人ぐらいしかアメリカから留学生を引き受けていない。いろいろな事情で引き受けられないということがある。私が併任の東京大学に外国人学生たちが留学に来ても、住むところがないとか、生活費が高いとかで、いろいろと苦勞しながら勉強している。ところが、アメリカに行った日本人留学生は割り合いに苦勞が少ない。そういう教育環境の格差から出る不

満から、また摩擦問題が出る。

にもかかわらず、日本の企業は、アメリカの大学のほうに多額の出資をしている。1984年の調査によると、1980年には日本の企業は国内の大学に93億円、アメリカの大学に108億円の委託研究費を出していた。それが1982年にはそれぞれ129億円と314億円、2年のあいだに額が1.4倍と3倍になったのはよいとしても、研究資金は大幅に海外へと流出した。(もっとも、こうした統計には不明な点も多く、別に総務庁統計局の調査によると、1986年に日本の企業約1万2千社が海外へ支出した研究費は225億円ほどで、8年のあいだに3倍あまりになっただけの数値となっている。)

このように、どの側面からとらえても、基盤的研究の学問では輸入超過にならざるを得ないから、日本がアメリカで小手先の技術を習ってきては、うまく企業化するということ、向こう側としてはにがにがしく思っているようである。

ほかにも、アメリカの学生が日本に来て、常にビックリすることの一つが、図書館のサービスが不備なことである。図書館員のサービスが悪いということではなくて、施設の数が非常に少ないこと、その上運営の仕方が閉鎖的であるなど、いろいろと文句を聞かされている。

たとえば米国IBMの研究所の図書館は24時間開いているし、スタンフォード大学に行くと、夜中まで開いていて能率よく使え、開架式でいろいろと文献を自由に見ることができるが、日本だと思いに任せないということを実感している。文献複写依頼についても同様である。一つには時代遅れの会計法の存在によって、本の管理に多くの制限がつくことにもよるといえることである。

データベースなどでも、いま日本でサービスしているものの主力は外国で作られたものであり、日本は大幅に出遅れている。日本文化の傾向として、「精神文明」を口にしながら、反面では形にならないものには、お金を払いたがらないところがあるから、データベースをはじめとしたソフト面の立ち後れが大きいのはその影響もあるようだ。

この辺も、学術情報センターとしては、ひとつごとではなく、急速に改善しなければならないと考えている。

5. 均質社会と教育

少し前から、日本でもこうした基盤的問題と国家的規模で取り組んで解決しようということになってきた。中でも、創造性を伸ばさなければいけないということで、政府、財界が音頭を取って真剣に取り組んでいるけれども、なかなか思うに任せない。その背後に何があるのかを、私の経験からちょっとお話ししよう。

まず、文化的背景の違いがある。前述のように、アメリカ人にとっては、生活のレベルを上げることが人生の第1の目標であって、そのために働くというところがある。日本人は、「ぶらぶらしては、おてんとう様に申し訳ない」などと言うように、とにかく働くのはいいことだ、というのがむかしからの信念であり、欧米での労働価値感とはだいぶ違うものを持っている。彼らの労働は、必要に迫られて行なうものであって、労働するのがいいことだからではない。彼らは、それをハッキリと言

う。これは時間の使い方、ひいては効率に大きく影響してくる。しかし、全く新しい考えというものは、自由な時間がたっぷりあるほうがずっと出やすいものであるから、時間のゆとりのあるほうが、独創性の発揮には有利である。

それから、日本人は、目立つことはできるだけしたくないし、仲間意識が強く、和の精神を尊び、調和を大事にする。これらに対しても、もともとアメリカは、ヨーロッパの宗教的な迫害などに反発して、新大陸に渡ってきた人たちが作った国だから、他人との調和などということあまり考えない。むしろ、自己を主張するところが非常に強い。そして、そうしたはっきりした個性を持った人が尊敬される。その辺から、他人とは異なる、より新しいアイデアで勝負することを尊ぶという、文化の差が出てくる。

したがって、教育制度も、日本ではできるだけ均質的な教育をする。うちの子供たちは、アメリカと日本と、両方の幼稚園、小学校に少しずつかよったが、日本では、服装までうるさく言う。アメリカでは、服装は個人の選ぶものだ、ということで、裸は困るとしても、何を着ていっても、適当なものであれば、構わない。

教材も、日本に帰ってきたときはビックリしたが、一々細かいものまで、全員が同じものを買って、持っている。アメリカでは、たいてい幼稚園や学校に備え付けのものが少しあり、必ずしもみんなが一度に同じものを使わなくても平気である。

そればかりか、アメリカでうちの子供が行っていた小学校では、1学級が15人で、生徒の能力に合わせて、先生が一人一人に別の課題を与えて、面倒を見てくれた。うちの子供たちは、日本で育ち、それからアメリカに行ったのだが、いきなり、向こうの学校に放り込んでみた。英語が全然わからない子供たちなのに、学校は一生懸命に世話を焼いてくれた。しかも、個人の個性、特長を伸ばしていくことが最優先になっていた。

日本では、目立たないことを重視するが、アメリカは逆であり、幼稚園から「ショー・アンド・テル(見せて、説明する)」というカリキュラムがあり、何でもいいから、身の回りのものを一つ、幼稚園に持っていく。うちの子供たちは、ゲタを持っていったり、着物を持っていったりした。そして、それを友だちに説明する。友だちが納得するまで説明するように指導してくれる。そのように、幼稚園、小学校のころから、自分の考えをちゃんと述べるのが非常に大事だとされている。

前述の国際教育調査でもわかるように、相対的に日本は成績がいいけれども、それにはそれなりの理由がある。日本では、とにかく公式や原理を覚えて、計算の練習をして、答えを出すということを重要視した教育をやっている。

ところが、子供がアメリカの小学校の3年のとき、授業参観に行ったら、理科で、爆撃機が爆弾を落とすときに、目標のどこの位置で落としたりいいか、ということをやっていた。先生は、ずっと何も答えを言わないので、子供たちの間でいろいろな意見が出てくる。

中には、行き過ぎてから落とさなければダメだ、という子供も出てくる。なぜならば、外は風が吹いていて、爆弾が流されるからだと言う。そんなときでも、先生は何も言わない。真上から落としたり

らしい、と言う子供の数が一番多く、手前で落としたほうがいい、と言うのは少数派だった。そうやって議論が出尽くしたところで、教室にヒモを張って模型を動かし、実際に弾を落としてみせる。そういう教育をやっているから、先へ進むピッチが非常に遅い。しかし、教育を体験としてやっているから、内容は本当によく理解できていると思われる。

そんなぐあいだから、教えている内容は、日本のほうが多いかもしれないけれども、向こうのほうが、感覚的に身につけているというところがある。もちろん、伸びない生徒は、そこで止まってダメになる。しかし伸びる子は、そこから自分で伸びていくのではないかと思った。それが国際教育調査でも、前述のように、上学年になるほど日米の差の縮まりとして出てくるのではないか。

6. 均質社会と独創性

アメリカでは小学生は、3年生から5年生ごろになると、自分で百科事典を引いて調べること、図書館を利用することを徹底的に教育される。日本では、最近は教育におけるテレビなどの利用も増えて、小校での受身の教育は減ってきたかもしれないが、全体としては、依然として、受身の教育、詰め込みの教育で、直接先生から教わる分量が多い。かつてはそれが小学校教育で身につける知識の80%以上とさえ言われたほどである。しかし、アメリカでは小学生から、徹底的に、自分で調べるということをやらされる。

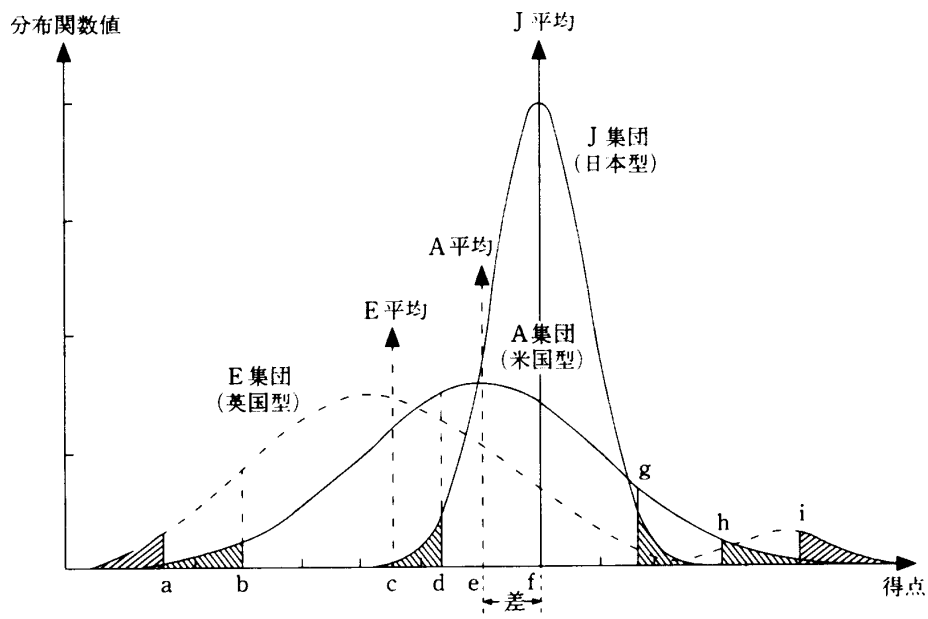
なぜ日本では、それができないのか。一つには、漢字教育ということに、非常に手間がかかる。これは両国での教育を見ていて、つくづくと感じた。彼らは、漢字教育に相当する部分が全然ない。スペリングという問題がないわけではないけれども、漢字教育に較べたら、ずっと楽である。だから、早くから自分で調べたり、いかに自分を表現するかというようなことに集中することができる。

日本に帰ってきてから、子供たちの日常を見ていると、漢字の学習と、計算を速くやるという練習に、とても多くの時間を取られていた。また、漢字が読めないから、一般の書物でものごとを調べることができない。この辺の事情があって、小学校のころから、ものの考え方に彼我の差がついてくるのではないかという気がする。

しかし日本では、教育には先生方も一生懸命だし、文部省もたくさん経費を出している。少なくとも小中学校の教育には、文部省は非常に多額の予算を使っているのだから、先生の質も国際的に高い。したがって、日本では、生徒が均質で、レベルが高くなっている。

だが、小中学生あるいは高校生の平均点が高く、均質的であるということで、無条件に喜んでいてよいわけではない。特に独創性との関わりを考えると、簡単に割り切れないものがある。付図は、何でもよいが、ある能力についての評価得点の分布のいくつかの型を示した模式図である（山田 1990-9から）。たとえばJ集団と記した曲線が、日本の生徒たちの能力の分布のつもりである。平均値を見ると、アメリカ型に比したA集団の分布のものよりも高い。ところが、アメリカでは、平均値は低くなっても、個性を伸ばすことに力を注いでいるので、広がった分布になっている。

独創性を発揮するのは平均値あたりにいる人ではない、ということが多い。すなわち、この図にお



平均と分散の異なる三つの分布の型

いて、各分布の右端の陰を付けたあたりにいるのが、創造性、独創性に関係がある人たちのグループである。こうやってみると、アメリカ的教育のほうが、人口の平均値は低くても、よくできる者はずっと高い得点を示している。

E 集団というのは、たとえばイギリスのような階級社会での模式図である。イギリスには、非常に傑出したカルチュラル・グループがある。それが右端にある小さな山である。もともと階級制度があって、貴族階級と労働階級とに分かれていて、貴族階級は徹底した英才教育を受けているので、こういう形になってもおかしくない。

今度は、左の端を見ると、いわゆる落ちこぼれというのは、日本ではほとんどない。つまり最低得点の者でも、A 集団の平均値にかなり近いところにいる。ところが、イギリス型になると、この分布は、たとえば労働者の質が悪いという事実を反映している。イギリスの産業界が落ちこんでしまった原因の一つは、やはり階級制度の存在ではないか。イギリスも、最近はこれに気がついて、教育に力を入れ始めている。その点は、非常に流動的だろう。

日本の教育は、言ってみれば、産業従事者の予備軍を養成するには向いているけれども、独創性のある人たちを生み出すには、もう少し工夫がいるのではないか。

ただし、この図はあくまでも模式図だから、あまり絶対化して受けとっていただいても困る。なお、この図の妥当性については、別のところ(山田 1990-9)において詳しく考察してあるので、関心をお持ちの方がたは、そちらを参照していただきたい。

7. 大学の環境

次に大学教育に移ると、統計的にハッキリしていることだが、外国と比べて日本では、大学教育、ことに大学院教育に当てている経費が少ないことが知られている。アメリカでの数%ないし、せいぜい10%程度しか大学院教育にかけていない。したがって、日本では大学院生の数も非常に少ない。研究者を養成する目的、独創性を養わせる目的のためには、これは問題ではないだろうか。

また、日本とアメリカとでは、大学の制度がかなり違っている。そのどっちがいいかという価値判断は主観になってしまうだろう。世界の中で、日本のほうが、技術的によりよく伸びているんだから、これでいいんだ、という考え方も当然できる。ただ、独創性を養おうという目的からすると、日本とアメリカとの間にある大学環境の違いが、日本に対してなかなかむずかしい問題を提供しているであろう。

たとえば、大学院の予算が少ないことは問題である。「そんなことはない」と言う人もあるだろう。確かに、国立大学などの予算を見ると、非常に大きい。しかし、それらの予算の全部が教育に向けられているわけではない。共同研究施設、附置施設などにかんがりのものが使われている。たとえば、今度独立するまで、東京天文台は、東京大学の予算でまかなわれていた。素粒子のための大型加速機の超大型予算も、教育行政にかかわっている文部省を通してつけられている。アメリカの統計では、こうしたものは教育予算にははいっていない。たとえば全米科学財団や国防総省などの予算から出ている。だから日本では、教官や大学院生1人あたりの研究教育に使われている実質金額は相対的にかなり少ない。これではなかなか学生の個性的なアイデアに研究経費をつぎこんで、リスクを負うことがむずかしい。

また、教官をとりまく、いろいろな制限が強すぎる。私は長らくアメリカの大学で准教授をしていたので、日本に帰ってきて、それをつくづく感じた。

アメリカの大学だと、州立大学でも、6年間勤務すると、有給の「サバティカル・リープ」という自由期間が1年間ある。週に6日働いて1日休むのと同じようなものである。その1年間は、たとえば本を書いたりして、自由に過ごせる。一番多いのは、ほかの大学に客員教授として行って、新しい刺激を受けてくるということであろう。一年間企業にはいって仕事をする教授もかなり多い。こういった環境の変化が、研究の新しいアイデアを生むのに大いに役立っている。

日本では、少なくとも国公立大学では、制度上そういうことはできない。しかも、ほとんどの教官が自分の所属している大学の出身者である。特に、新しい刺激を必要としている助教授の、少なくとも80%は自分の大学を出ている。アメリカでは、ごく少数の例外的な大学を除き、助教授の80~85%は、外から来ている。ずうっと自分の出身校にいた人は教授に上げないというのが、明文化した規約になっている大学はかなり多い。

このような制度は、オリジナリティを出すためには、非常に大事なことではないか。なぜならば、ある大学で育った人は、その大学のことしかわからない。ほかに行けば、全然考え方の違った人がい

るから、それらの人たちと意見を闘わすことによって、新しい考えが生まれる確率が高い。

この、違った考えの人たちと接触するという事は、アメリカの社会では、たいせつにされる。「意見の相違が世の中を動かして行く」という表現はよく聞かされたものである。大学の先生が、週に1日は大学外のところでコンサルタントとして活躍することは、容認されるどころか、むしろ積極的に奨励されている。大学の先生が社会の現場で働くことは、学問を実地に生かすと同時に、研究のために現実的問題を採り上げる、よい機会と考えられているからである。

週休2日だから、大学に4日来て、企業に1日行って、2日の休暇は自分で使う。そういう形が定着していて、極言すれば、コンサルティングの口もない先生は、能なしであるという評価さえなされることさえある。

それに対して、外国の人は、日本の大学教授を見て、よく「お気の毒だ」と言う。中には「ブラックホール・プロフェッサ」と評した人もいるとのことである。ブラックホールとは、宇宙空間で局所的に物質がたくさん集まっていて、引力が極度に強くなり、次つぎとまわりから物質を吸いこむが、何も出てこられないところである。光さえも出てこないのので、暗黒に見える。日本の大学の先生は、これと同じで、知識の吸収はするけれども、新しいものは何も出てこない、ということだそうだ。また、日本の大学教授を評して「お役所芸者」と書いてある英文のものを読んだことがある。有能な人ほど、官公庁関係のいろんな委員会に引っ張り出され、行政的なことに携わりすぎて、研究がおろそかになっている先生が多いということから来ている。

こうした酷評は言いすぎであろうが、それにしても日本の大学の現状は、アメリカに比べてなかなか研究に専念することができにくくなっていることは否定できないであろう。

8. 若手の研究者をとりまく環境

大学で苦勞しているのは教官だけではなくて、学生にとっても問題はある。アメリカの大学生、特に大学院の学生は、大学からちゃんとした月給などをもらって働くのがかなり当たり前になっている。その代わり、責任もあって、チームの一員として、大きなシステム作りなどに貢献している。

日本では法的に、学生は原則としてアルバイトができないことになっていて、大学が学生を雇用するわけにはいかない。これは大学にとっても学生にとってもかなり不利に働いているようである。特に、情報科学関係では、新しいソフトウェアを作るのには、頭の柔軟な、回転が速い、吸収力の盛んな大学院の学生は、ものすごい戦力になるが、日本では、かれらを正規に雇用することはむずかしい。これも日本の大学でなかなか世界的に通用する新しいソフトウェアが、アメリカでのように生まれてこない原因の一つとなっている。

学生の創造性の低さが問われるのなら、かれらを選抜する入学試験にも問題がある。現在の入学試験というものは、多くの問題の中から、いかに早くむずかしい問題を見抜いて、それらを避け、やさしい問題を優先的に解くか、ということがうまい人が受かりやすいようになっている。まじめに、非常にむずかしい問題とガッチリ取り組んで考えるというタイプの学生は、その出発点の、入学試験で

落ちてしまう。しかし、長期的には、そうした資質が独創性につながることが多いであろう。

独創性を発揮するには、今とは違ったタイプの学生も大学にはいってこることができるようにしなければならない。最近、そういう反省もあって、大学のほうもしかるべく対処するように努力しているようである。たとえば入学定員の1割りぐらいを、一芸に秀でた学生の選抜に向けるといったくふうである。

最近日本では、何となく老人を思わせる名前の、生涯教育ということが盛んに言われ出した。アメリカにおいて、これに対応しているのはコンティニュード・エデュケーションである。高度に発達した学問を4年間で十分修得しきれないまま、社会に出て就職した学生が、何年か会社で仕事をすると、実際的なことを身につけた上で、もっと勉強したいことがいろいろと出てくる。そうした若い人たちが、半年なり1年なり、大学に帰って、ミッチリ勉強するというのが、コンティニュード・エデュケーションであり、アメリカの社会では常識のように行なわれている。日本でも、制度的にはできないことはないが、いつも人手が足りないとかいった、会社や組織などの都合で、実行はなかなかむずかしい。この辺も、日本では、大学と産業界とがお互いに刺激し合って、若手研究者の知識交流の相乗効果を出すことをむずかしくしているという気がする。

若手が能力を十分発揮していない理由として、そのほかに、日本はまだ年功序列が強い社会だから、組織の中で自分だけが伸びても、あんまりいいことがないということもある。それに反してアメリカの社会では、自己を主張し、認められれば、地位がドンドンあがっていく。そうした制度には弊害もあるが、やはり独創性のある人は、はしごをのぼるのが速いから、よけい励みになる。日本では、なかなかそうはいかないから、良いアイデアが十分育てられない。

また日本では、業績の評価が、なかなかやりにくくなっているようである。アメリカでは、ピア・レビュー (peer review) と言って、かりに若くても、同じテーマで現在研究し、一流の成果を挙げている人たちが論文の審査をする。日本では、どちらかと言うと、功なり名とげて、忙しくてしょうがないような先生のところに論文が回ってきたり、どういうわけで選ばれたのかわからないけれども、自分で世界的な論文も書いたことがないような人が論文審査員になっていたりするから、若い人たちの独創的な研究がなかなか認めてもらえないことがある。

独創的なものがなかなか認めてもらえないというのは、世界的に共通の現象だが、日本では特に問題のようだ。たとえば、日本では〇〇賞というのが、必ずといっていいぐらい長老に与えられることになっている。20代、30代の人々が、大きな賞をもらうことはまずない。最近、日本にもだいたいふんい賞ができたけれども、もらうのはやはり功なり名とげて何十年も過ぎたというような人ばかりが多く、最近優れた仕事をした若い人のところには、なかなか回って行かない。これも、アメリカのような競争原理の社会と、和を尊ぶ日本の社会との差だと思う。

9. 和の原理と個性の原理

先年、日本がそういう和の原理で動いている社会であることを体験した。

日本で開かれたある国際会議で、テーマの都合で東洋人だけのパネリストからなる討論会の司会を勤めたが、一般に日本のパネル討論会というのは、パネリストがめいめい自分の意見を言うのが主で、相互の意見がかみ合ったディスカッションには、なかなかならない。そこで、本当のディスカッションへとうまく展開するように、あらかじめいろいろとお膳立てして、パネリストにお願いをしておいた。

会議のあと、打ち上げパーティーの席で、アメリカから来た古くからの友人たちの話を聞いたら、「日本ではめずらしく、よく議論が出て、非常によかった」と言ってくれた。ところが、そのすぐあとで、日本人の同僚に会ったら、「あれはかわいそうだった。日本人どうしでお互いにあんなに攻め合っで…」と言われた。

アメリカでは、パネル討論会というのは、節度を持った対決であるという考え方がある。しかし日本では、そういう意識は希薄で、パネル討論会が和の精神で運営されていることが多いようである。

そうした経験をした直後に、たまたまアメリカでのシンポジウムに行く機会があった。そこではブレイクアウト・セッションというのが準備されていた。これは割り合い新しい形式の討論会であり、現在まだ異論の多いテーマを取り上げて、いろいろな人の意見を取り入れながら、司会者が将来に対する見通しをまとめていくことをする。そうした性格上、次から次へともものすごい反論が続き、大変なセッションだったけれども、結果的には、みんな非常に満足したようである。こうした過程を通して新しいアイデアを発展させられるように、われわれ日本人にあっても、率直に意見を交換することができるようになることが、必要なのではないか。

先に付図で説明したように、独創とか発明とかいうのは、あくまでも平均的でない個人の問題である。新しいものごとを考えるのは、普通とは少し違った人が多い。これもアメリカでの経験だが、ある会社の研究所に在職していたとき、その関連会社で起こったことについて聞いたことがある。原子力潜水艦を作っているエレクトリック・ボートという会社で、第2次世界大戦中は輸送船を造っていた。船には昔から甲板があった。

その会社にある日個人から手紙が来た。それには、甲板を作らずに、垂直の隔壁を幾つか作って、石炭や穀物などをそこにバラ積みにするようにすれば、積み込み、積み下ろしが非常に早く済む、という提案がしてあった。しかし、そのころの造船技術では、まだそれができなかったのもので、手紙は無視されてしまった。

ところが、皮肉なことには、その手紙が来たすぐあとに、溶接技術が非常に進み、そういう船の作り方が可能になり、戦後には、そういう船がドンドン作られ始めた。そうってから、前に手紙を寄越した人から、再び「あれは、私の発明だから、アイデア料を寄越せ」と言ってきた。

結局、裁判になって、手紙を一度寄越しただけの街の発明家が会社から300万ドルだったかを受け取ることができた。いまにすると、おそらく1000万ドル以上の価値になるだろう。そのとき私は、アメリカの会社が個人から受け取った手紙を慎重に扱うことを厳しくしつけていること、受け取った事実を正直に認めたことなどと同時に、アメリカの社会が、たとえ大まかなものであっても、個

人のアイデアをいかに尊重するかということに、強い感銘を受けたことを覚えている。

アメリカ社会についてのこの種の話は、まだ外にもいろいろとある。原子核の研究に使うベトロンの発明についても、やはり同じような事実が知られている。

ところが、日本では、アイデアなんてものはタダだと思う人が非常に多い。情報というのは、アイデアに似たものだから、その情報を専門とする者が、他人の生産した情報に尊敬を払わずにいて、どうして自分たちでメシが食えるか、と私は学生にやかましく言っているが、それでも学生は、平気で本やプログラムをコピーしているようだ。

独創性を発揮するということには、一人一人の個性を伸ばしてやる社会にする必要がある。それには非常にお金がかかる。だから、今までの日本はそれを避け、もっぱら技術導入でうまくやってきた。それで、ここまで来た。しかし、これからもそれでいけるかどうか。これは、話が別である。

最先端技術の日本への流出にアメリカが神経を尖らせだした現在、この辺で、そろそろ、日本はみずから独創性のあることをやっていくようにしなければならないのではないか。前に披露したスーパーコンピュータの国際会議の場でも、「日本人には、アメリカのスーパーコンピュータを使わせるな」ということを公然と言って歩いているアメリカ人に会って、ショックだった。

最近、ニューラル・コンピュータ、つまり、人間の神経の働きに似た組織のコンピュータを作ろうということで、アメリカでも日本でもさかんに研究されている。そのニューラル・コンピュータの中で、最も進んだ製品のひとつに、アメリカのコネクション・マシンというのがある。これは、ニューロンに相当するプロセッサが5万何千個が使われている。そのアイデアをうまく使われてしまうのを嫌って、「これは、外国人には使わせない」と内部で決めてあるという話である。もっとも、使ってみることのできた日本人がいるそうではあるが。

いずれにしろ、これからは、自分のアイデアをどしどし生かしていかなければ、日本は伸びていけないのではないか。

幸いにして、いまコンピュータの分野で、たとえばTRONというプロジェクトが進んでいる。これは1990年代に普及する、最後のフォンノイマン・コンピュータになるだろうと言われている。本当にそうなるかどうかは、われわれのコンピュータ技術の関係者がどのくらい積極的に協力するかによるだろう。とくにメーカーがもっと一致団結してやっていかないと、またどこかの大手メーカーによって足をすくわれるようなことが起こるかもしれない。

10. 創造性養成への道

日本における没個性について、いろいろと厳しい文句を並べたが、それらは、独創性を育てることを前提にしての話である。独創性が発揮できる基盤の整備には、膨大な費用がかかる。だから独創性に、本当にそれだけの価値があるかどうかということとは、また別の問題である。しかし、諸般の国際情勢は、好むと好まざるとにかかわらず、いまわれわれがもっと創造的になることを要求している(House 1989)。

それでは、これからどうしたらいいのか。結局は、いままでの日本の考え方を少し変えていくことだろう。いままでうまく使ってきた考え方やり方の全てを捨てる必要はないし、また、捨てることは不可能だろう。しかし、同時に、もっと型破りの人間が伸びていくことができるような環境づくりを考えていかなければならない。

独創性をどうやって伸ばすかについて、いろいろと書かれているものを、ちょっとばかり読んでみた。企業あたりからもいろいろと提案がでていますが、それがまた画一的なのに驚いた。たとえば、大学の先生は、35歳までは契約にしようという。個人差を考えないで、全部そうしろという発想らしい。研究者は40歳以下でなければならないというのもあったが、これもまた画一的思考である。また、文部省の科研費は40歳以上には与えるべきではないというのもあった。これも年齢だけで区別している。

個人に属する独創性を伸ばすことを望んでいる企業の提案の中にも、このように個人差を尊重していないものがあることを非常に残念に思った。これでは、別の硬直化を提案しているにすぎないのではないか。

たびたび言うように、創造性と、個性を自由に発揮する発想とは、切っても切れない深い相関があるように思える。したがって、幼稚園から高校まで、徹底した均質主義の思想で設計され運営される学校教育を受けてきた者が、大学に至って初めて個性化を奨励されても、そうたやすくは創造性の発揮へとつながらないと思う。

この辺の事情については、別のところ（山田 1990-9）において詳しく考察してあるから、参照いただければ幸いである。

次に、研究というのは、一般に非常にお金がかかるものである。アメリカのある研究所にいたころ、マネジメントは、全研究活動数のどのくらいが成功すればいいというつもりで、研究者にお金を出しているのか、調べたことがある。

その会社の方針では、10のうち、二つか三つ成功すれば、大もうけだ、と言っていた。もう20数年まえのことだが、そのころたまたま、その会社から日本に出張するチャンスがあって、同じようなことをむかしの仲間聞いてみたら、日本の研究所では10のうち八つか九つまで成功しなければ、部長のクビが危ないんじゃないか、ということだった。研究に関して、日米ではこの辺の理念が違うと感じたことである。いまでは日本もそうとうリスクをとるよになってきたとは思うが。

大体、始めから80~90%も成功することがわかっているものは、もう基礎研究ではないだろう。20%ぐらいのところまで線を引くなら、これは相当のリスクがあるけれども、そのかわり、当たったときには、とんでもないものが出てくる可能性がある。

最近、ある人が「日本だってそういう基礎的な研究はやっているんだ。超電導を見なさい。まだ実用化のめどのない、あんな基礎研究に関する本が、もう20数冊も書かれているじゃないか」と言ったという。しかし、この20数冊というのは、どういう人が買っている本か。どうも基礎研究に興味を持っている人ではないようだ。むしろ、うまくいけば、もうかるんじゃないか、ということで、ビジネ

スの営業あたりの人が読んでいるのだそうである。また、そうした本を書く人びとが、あたかも研究者の代表のように、もてはやされる風潮にも問題があろう。

そういったわけで、いま日本は自由な発想で基礎研究に取り組める人材を急いで育て、かつ十分な研究資金投資をしないと、これからの日本は、国際競争の場で、なかなか厳しい情勢に立たされることになるのではないか。

11. 学術情報センターに関連して

だいぶ一般的、抽象的な話を、しかも主観的にしてきたが、少し具体的な研究開発の話に立ち戻って、それで、われわれ学術情報センターとして、これからどういう研究開発をやったらいと考えているのかについて、これも、あまり独創性がない話になるかもしれないが、一つ二つ述べたい。詳しいことは何度かの機会に述べておいた（たとえば山田 1998）。

もう数年も前に、マサチューセッツ工科大学（MIT）で個人向きの新聞編集システムというものを動かしたことがある。簡単に言うと、現在、ニュース・サービスというのは、AP、共同通信、ロイター、UP など、多くの通信社から、のべつ幕なしに情報がはいつてくる。新聞社は、それを取捨選択して、新聞に使う。

しかし、いまわれわれのところに配達される新聞を、隅から隅まで全部読む人はまずいない。自分の好きなところしか読まない。すなわち、いってみればめいめいが、自分の興味のあるところを、キーワードのリストによるプロフィールとして持っていて、記事の中からそれに合うものだけをすばやく見つけ出して、読んでいます。MIT のシステムはそれをコンピュータ化して、絶えず届くニュース情報を、あらかじめ指定したキーワードのプロファイルでふるい分け、関心のあるものだけを採り、ソースラに基づいて体係づけをし、個人用の紙面を組み上げるというものである。これはわれわれのところのサービスにとって将来かなりおもしろいシステムとして実用化できると思う。

研究者が自分の研究プロフィールをセンターに登録しておけば、世界中からはいつてくる研究文献情報のうち、自分に役立つものだけが選択され、電子メールを介して、自分用の目録として届けられる。そういう形にまで、将来は学術情報サービスを持っていけるのではないか。

もちろん、そうしたシステムが実際に普及するためには、まだまだ通信システム、コンピュータ・システムの進歩が必要だが、学術情報システムにおいても、10年、20年先にはやはりこのようなシステムが使われるようになるべきだという気がする。

それには、通信システムにしても、いまの何百倍も強力なものが普及しなければならないと思うけれども、原理的には、そんなに面倒なことではない。金に糸目をつけなかったら、技術的にはいまでも作れるものである。さらに、超広帯域幅をもつ光ケーブルで日本中の家庭が結ばれる時代がもしくれば、プロフィールの登録をする必要もなくなるだろう。通信線の中には全てのなま情報がそのまま流されていて、プロフィールによる選択は手もとの受信端末でやるようにすれば、サービスはもっとらくになる。

もう一つ(山田 1990-2 参照)。いまの図書館のサービスというのは、主として文献サービスである。しかし、すでに世間では、いろいろの数値情報とか、画像情報とかのデータベースが作られ始めている。さらに将来は、そういった情報のほかに、実際にもっと進んだセンサーを使って、それをスーパーコンピュータと組み合わせて、われわれが実際いかにも本物に触っているような、模擬空間の中の現実感というものを再現することができるようになる。

部品的には、そうした目的に合ういろいろなものが、すでにできている。たとえば「データ手袋」というものでは、手を動かすと、その位置情報がコンピュータにそのままはいるし、逆にピエゾ圧電配素子とか、形状記憶合金とかを使って、コンピュータからの反力を手にフィードバックして、触覚として伝えられる。そうしたものが実用化すると、コンピュータが作った3次元の模擬対象物の画像を目で見、手の触覚で確かめながら、研究開発活動をすることができるようになる。

そういう技術を使えば、学術情報サービスも、もっと多次元化することができるようになる。いまは主として言語と数式と画像に頼っている情報伝達が、五感の全ての活用にまで広げられる。たとえば、目に見えないような小さな分子を、目で見えるぐらいの大ききで示すことができるばかりか、手で触れてみることもできる。そうすると、二つの分子の組み合わせを、ちょうど子供がオモチャや積み木で遊んでいるような形で、感覚的に調べることもできるのではないか。

これは、ミクロの話だけれども、大きなほうでは、銀河系宇宙の中に手を突っ込んでこね回すといったこともできる。あるいは、地球のシミュレーションに指をつっこんで行くと、密度や温度に比例した触覚を手で感じるようなことができる。要するに、ものごとを論理的・数値的に把握するだけでなく、感覚的にも把握することができるような情報サービス・システムができるのではないか。

そんなことをして何の役に立つのか。研究効果はまだわからないけれども、少なくとも教育的効果はあると考えられる。いずれにしろ、すぐに役立てることばかり考えていたら、決して大きなことはできないだろう。これからは日本でも、もっとこういった研究をやっていないかなければならないだろう。

アメリカでは、この種のことは相当前からやられている。その最初の10年を、いわば冷やかし半分で見ていた日本の企業も、ここへ来て、実用的価値が見込まれる成果がチラホラ出始めると、途端に研究費を手みやげにして、MITへお百度参りをするようになった。それでも、こうしたことが学術情報の伝達のためにほんとうに有効に使えるようになるかどうか、まだやってみないとわからない。技術的には可能であるが。

こういう種類の話をする、日本のメーカーの人たちは、まず、「そんなことは…」と笑ってしまう。

しかし、MITの研究も、始めはお遊びのようなことから始めて、20年間積み上げてここまで来たのだ。基礎研究とは、そういった気長さも、また時にはむだも、必要だと思う。

だが、大学でこうした研究を立ちあげようとしても、企業からは、なかなか研究費が回ってこなくて、すでに述べたように、むしろアメリカのほうに流れている。たとえばMITのことなど、よく新聞に出ている。しかもそれは、基礎研究の成果が実用化される見通しがついてきた時点になってからのことが多い。

こうした事情にも反映されるように、日本の研究分野で割り合い優れているのは、即物的なところが多い。

12. 執事の研究から王侯貴族の研究へ

こうした現状にどう対応するかに示唆を与えるとともに、締めくくりとして、ちょっと調べたことのある気球の話をしてしよう。

歴史上初めて人間を乗せた水素ガス気球は 1783 年にパリで飛んでいる。これを目撃した駐仏アメリカ大使、嵐で雷の研究をしたという逸話で有名な、例のベンジャミン・フランクリンは、ただちにその将来の重要性を見抜いたという。

一方、日本人として初めて気球の飛行を見たのは、19 年後の 1802 年、カムチャッカに漂流して、モスクワに連れていかれた幸太夫（光太夫とも書かれている）という漁師だった。彼は、気球の飛行に大喜びをしている大衆を見て、「こんなもの、何のために用をなすものか」と無感動だった。

それからまた 60 年ほどたった 1860 年、日本の遣米使節が、アメリカのフィラデルフィアで気球の飛行を目撃したが、このときも、帰ってきての報告書には「空にのぼりて、何の益もなきこと」ということが書いてあるという。

だが、いったん軍事に役立つとわかると、途端に日本でもものすごい勢いで研究を始めた。そのように、日本は昔から実利的な面ですることへの反応は、非常に素速いけれども、お遊びの精神というか、王侯貴族的な楽しみというか、そういうものには、貧乏性のせい、反応の仕方が非常に鈍いというところがある。

国民性としてわれわれは、どちらかと言うと、王侯のまわりの世話役、イギリス的に言えば、執事、従僕、召し使いといった形で、外国人研究者という王侯貴族がやり散らかしたことを、実に丹念に整理して、家の中、すなわち科学技術の分野を整えるということをやってきた。

こうした態度はしばしば論文の書き方などにも顕われてくる。欧米の論文では、研究の基になったアイデアや考え方などを一番先に提出したのはだれかということ調べ、かつ失敗をも明示した上で、見通しを立て、先駆者の貢績を明らかにしてあることが多い。実はそれがリサーチ (re-search) ということばの起こりでもあるという。しかしわが国の論文では、先駆者への言及はしばしば省略するだけでなく、むしろそうした記述をうるさがり、その後のことだけを手柄として、事こまかに書きがちである。

これは研究者の世界だけのことではない。たとえば、あるテーマを追ったテレビのプログラムの構成などでも、アイデアや理念の取り扱いにおいて、しばしば彼我に見られる差である。

物理学においては、たとえば素粒子に関する論文など、世界的にみると、あとになって正しくなかったと分かったものが多く、その割り合いが、なんと 90%にもなるという話さえある（大石、小山 1988）。これは無責任な発表をしているということではなく、現在の知見の中で検討した結果、可能性がまだ一意に絞れていなくても発表されるということからくる帰結である。

科学の分野の中ではかなり厳密な理論立てが可能と思われており、またおそらくそれが正しいであろう物理学の分野においてさえ、そうなのである。まして、もっと不確定性の高い生理学や脳科学の関連している人間科学の分野では、理論が一意に絞れない現象が数多くあっても不思議はない。

そんな不確定性の高い思索に努力を費すのは無駄ではないかという、ともするとわれわれ日本人の落ち入りがちな考え方は当たっていない。実験によって事実を確かめることのできる科学の分野では、そうした複数個の理論を比較検討することにより、また新しい実験方法も明らかにされることが多いからである。

われわれ日本人は、とかく慎重になり過ぎて、大胆な仮説を立てることを躊躇しがちである。その結果、われわれの発想の形態は執事型に留まり勝ちになるといえよう。しかし、独創性を育てるためには、そうした知的冒険を受容する研究環境づくりが必要であろう。

この執事型の殻を破るには、幼いときからの教育の在り方を変えることも必要になろう。たとえば、小中学校で校舎の掃除を生徒にやらせるのは、アメリカはもちろん、その他の先進国にはないことで、わたしの知る限りでは、おそらく日本だけの慣習ではないかと思う。下級生の掃除ぐあいに上級生が注文をつけたりすることも、ときにはあるそうだから、知らず知らずのうちに年功序列の意識を養うことにも役立つ慣習であろう。また、集団掃除が、勉強は全くだめでも体を動かすことは好きな生徒のはげみになり、どうしようもない落ちこぼれになるのを救うという効果もあると聞いた。

それにしても、これは、やはり画一教育的色彩の強い慣習であり、かつ、王侯貴族的思考よりは、執事的思考の形成に向けた日常行事であろう。

また、敢えていえば、もともとお役所の仕事は、性格上執事的なものがその大部分を占めているはずである。したがって、国際的に見ると、最近の日本が管理社会と呼ばれるほど、官僚などが相対的にかなり強い権限を発揮しているのも、上に述べたような日本文化の特徴と無縁のことではないだろう。だから、社会の多様化のためには、レーガン前大統領の主張した、「小さな政府」のほうがよいだろう。

これから、われわれも、執事としての科学技術だけではなくて、王侯貴族の科学技術もやっていかなければならないのではないか。

13. おわりに

本稿は1988年7月29日、日本学術会議講堂で行なわれた「学術情報センター・シンポジウム」での「研究開発とR & D：創造的活動の日米比較」と題する講演の記録に手を入れ、1989年9月14日に横浜で開かれたInternational Symposium on New Technologies: International R & D Cooperation and the Development of New Technologies（新技術国際シンポジウム：研究開発の国際協調と新技術開発を考える）での講演で再度用い、その後にも少し手を入れたものである。文体には、講演の感じをいくらか残しておいた。

1988年の講演については本センターの井上如教授にお世話になった。その時の録音から文章に起こ

して初稿をまとめて下さったのは、記録業の竜岡博さんである。1989年の講演が実現したのは横浜国立大学工学部の仁木克己教授のご甚力による。本センターの大野公男教授は原稿を精読され、数か所のご意見をたまわった。たびたび手を入れた草稿を、そのつどタイプして下さったのは堤妙子さんである。これらの方がたに厚く感謝したい。

(1990年3月)

参考文献

(本稿は講演がもとになっているので、参考文献は最少限に留めてある。)

- ・ House, K. E., The '90s & beyond: though rich, Japan is poor in many elements of global leadership, Wall Street Journal, January 1, 1989.
- ・ 飯田庸太郎、我が国の科学技術、日本工学アカデミー講演、東京、1989年5月15日、EAJ ニュース、No. 9、pp. 3-4、1989年7月。
- ・ 大石進一、小山慶太、科学・技術が飛躍する余地は残されているか（対談）、ゑれきてる、No. 32、pp. 4-9、1989年冬。
- ・ 山田尚勇、学術情報システムにおける大学図書館の役割、大学と学生、No. 272、pp. 22-27、1988年8月。
- ・ 山田尚勇、21世紀の情報サービスへ向けて、第8回学術情報センターシンポジウム、京都、1990年2月21日。[学術情報センターニュース、No. 12、1990年6月に抄録]
- ・ 山田尚勇、テスト評価、均質的文化、独創性養成—教育・研究における選抜法について考える、学術情報センター紀要、No. 3、pp. 161-192、1990年9月。
- ・ 柚木久、アメリカ企業における研究開発システムについて、電子通信学会誌、56：1055-1065、1973。
- ・ 柚木久、企業における創造性管理システム、電子通信学会誌、58：744-753、1975。
- ・ 柚木久、テクノロジトランスファ、電子通信学会誌、60：11-18、1977。

研 究 論 文

テスト評価、均質的文化、独創性養成

—教育・研究における選抜法について考える—

Testing Procedures, Homogeneous Culture, and Creativity Cultivation :
on selection procedures in education and in research

学術情報センター 山田 尚勇

要旨	161
1. 才能・技能評価値の分布	162
2. テスト結果の総合評価方式	166
(A) 一芸方式	
(B) 総合点方式	
(C) 順序方式	
(D) 折衷方式	
3. 時間の経過にともなう動的变化	170
4. 得点分布の型と文化の型	173
5. 日本文化の均質性	177
6. 均質性の国際的特異性	178
7. 従来とは異なる尺度による評価法の可能性	181
(E) 評価点ベクトル方式	
(F) 一般化ベクトル方式	
(G) 平均点ベクトル方式	
(H) ベクトル方式の自由度	
8. 独創性の養成のためには	186
9. 創造性と研究の管理行政	188
10. おわりに	191
参考文献	191

要 旨

個性化、多様化と創造性が深く関わり合っている資質であることは、すでに多くの人びとの述べ

ているところであるが、本稿ではその定量的裏付けに向けて少し考えてみる。

その第一歩として、現在のテスト管理社会において、個人の能力を評価しているシステムを取りあげてみると、そうしたテストの連続で選抜されてきた集団では、質の向上とともに、時がたつにつれて、どうやら集団の均質化、平均化が起こるらしいことが推論される。

すなわち、テストによる管理の度合いと、それを実施している社会の持つ文化の型とには、ある種の相関が存在し、創造性の基となる、個性の自由な伸展を望むのなら、できるだけ均一化を引き起こさないような、テスト結果の総合評価法を考えていかななくてはならないようである。

いま広く使われているのは、各テストの得点の単純和をもって順位をつける評価法であるが、この方法には、計算がらくであるという以外に、大した根拠がない。むしろ各能力の得点を成分とし、その合成ベクトルを総合能力として評価するほうが自然で、かつ多様性の評価にも、より良く適合するようである。

多様化を考えるときに、よく心配されるのが、全体のレベルの低下ということであるが、このレベルという考え方は、実は一つの数値に頼った、画一化の概念である。したがって、多様化、個性化を必要とするこれからの情報社会においては、目標達成の基準として、全体のレベルという考えに代わるものを見つけ出さなくてはならない。

伝統的に基礎研究に多大の重み付けをしていた米国が、国際的環境の変化に押されて、実利主義へと向かいつつあるいま、日本がその逆向きの努力を余儀なくされていることには、なかなか意味深長なものがある。

1. 才能・技能評価値の分布

日本は現在世界にまれなテスト社会になっていると思われる。それと同時に日本は、国民の質の高さと均一さによって支えられた、大工業国としての国際的評価を受けている。

その反面、工業の基となる画期的な技術的アイデアにおいては、まだ諸外国に負うところが多く、その上、工業製品貿易のきわだった国際的優位が災いして、技術ただ乗り論の非難さえ受けるようになってきた。

それではならじと、ここへきて国内では、にわかに独創性一般の養成に強い関心が持たれるようになり、それに向けた種々のプログラムが提案され、いまそのいくつかは真剣に実行に移されつつある。

こうした一連の現象と、それらのあいだにあると思われる因果関係については、別のところで詳しく考察を試みておいた（山田 1990）。それに引き続き本稿では、そこで示唆しておいた、過重なテストによる管理が文化の均質性を誘発する可能性について考え、それと整合する特徴を日本文化の中にさぐり、実はそれが独創的個人の台頭とは相反した性質のものであるとの推論を試みる。

人間は数多くの才能、技能を持っている。説明を分かりやすくするため、ここでは何でもよいが、そのうちの二つをとり、それぞれに与える評価値を x 軸と y 軸方向にとり、各人の評価点をプロットする。図 1 はそうした値の分布に合わせて点をプロットしたものである。（A 型という呼称はあとで明

らかになる。)具体的には、第1近似として、おのおのの才能の評価値の分布が、その軸上でそれぞれガウス分布で与えられると仮定した上で、疑似乱数をもとにして、2000人分の評価値を計算し、図示してある。厳密には、各才能はそれぞれ他から必ずしも独立ではないし、またその評価値も負になり得ないから、明らかにガウス分布とはならない。しかし、ここでは定性的な第一近似理論を造ることを試みるから、そうした近似は本質的なことではない。より精密な議論は今後の課題としよう。

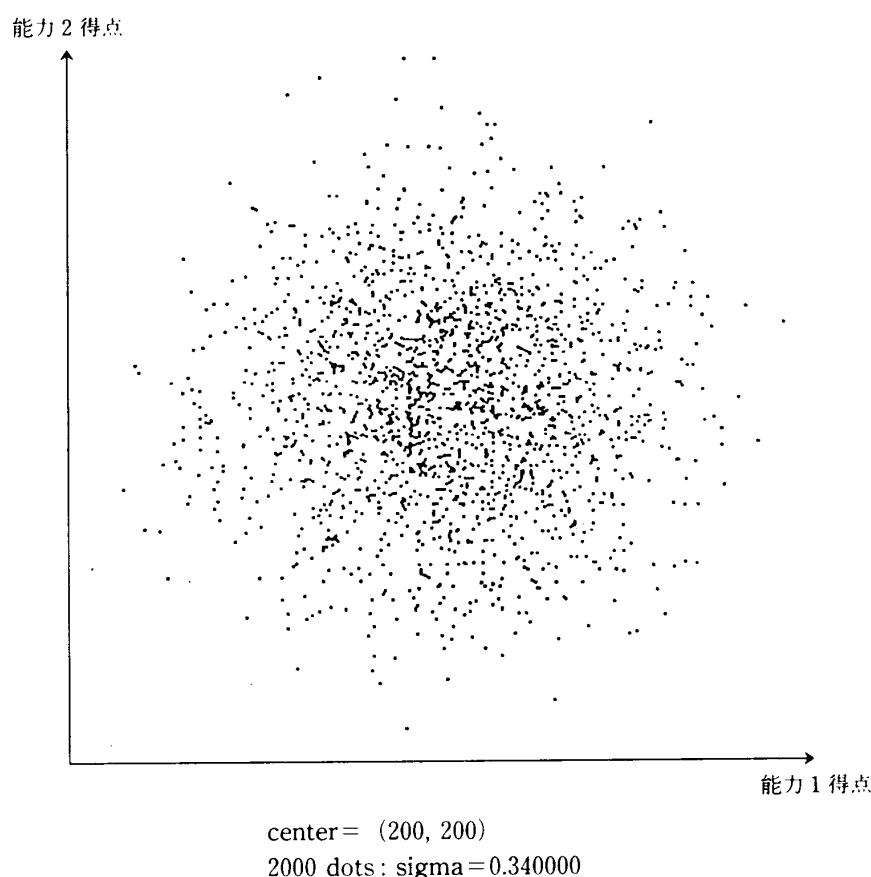


図1 A型母集団の得点分布

分布がガウス関数とすれば、平均値からいくら遠くても、値は完全に零にはならないが、図2では図1で与えられた2000点の収まる範囲と、別に約95%の点の収まる範囲とを同心円で示してある。この95%という値は、あとの節で独創性の議論を試みる際の都合に合わせて任意に選んだ値であって、ここではその値自体には特に意味はない。また外枠の円は、単に分布の分散の程度を示す目安である。

さて、日本人は、たとえばアメリカ人に比べて、大衆の教育レベルが高く、かつ粒がそろっているというような評価がよくなされる。詳しいことはあとにゆずって、ここではこれを才能・技能評価の平均値が高く、かつ、ばらつきが小さいことと解釈しておく。そうすると、たとえばアメリカ人の評価値の分布を図1で代表させたとき、日本人のは図3のように表現できよう。こちらもガウス分布に

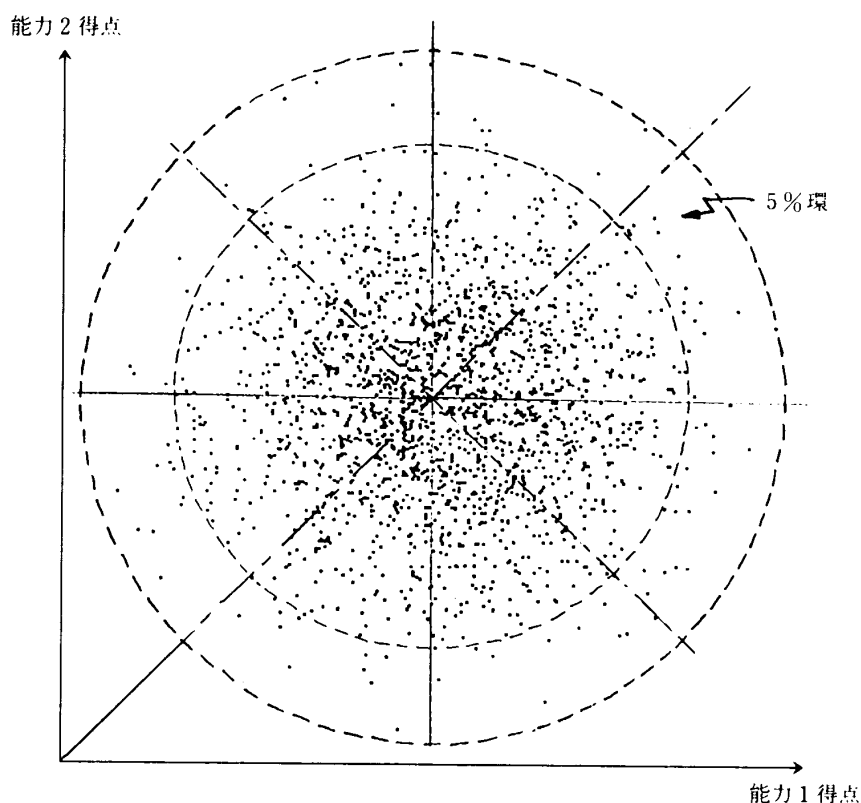


図2 A型得点分布の5%外環域

従った2000点で構成したが、中心、すなわち平均値を17.5%高くし、かつ、ばらつきを示す標準偏差を約4分の1に減らしてある。ただし、これらの数値の選択は図示が直観的に分かりやすい形になるように任意に選んだものであって、特別な意味のあるものではない。実際には億が単位となる国の人口を、わずか2000点で表現しているのだから、図示が論点を明瞭にしてくれるには、各種のパラメタの誇張が必要になる。そうすることにより、たとえば図1と図3とを重ねた4000点の分布が図4であるが、明らかにもとの二つの分布の重心がずれているさまが見てとれるようになっている。

図2のときと同じ要領で、枠を図3に書き加えたものが図5である。ただし、ここで注意して欲しいのは、たとえ確率は小さくても、統計的には枠外でも値が零でないということである。事実、図3や図5の点は、コンピュータによって発生させた疑似乱数を基に、簡単な統計的計算の結果全く機械的に求めたものであるが、図5の下方の矢印が示すように、丸枠から遠く離れたところにも、たまたま点が出現した。

これらの分布図はわれわれの直観に訴える目的で描かれたものだから、たった二つの因子、つまり二つの才能・技能しか含んでいない。一般に数の多い因子をn個を考慮するときには、これらの図はn次元に拡張しなければならない。理論上は何でもないことであるが、その図示はほとんど不可能となる。ここでは2次元図を見ながら、n次元構造への拡張を想像してもらうことにする。

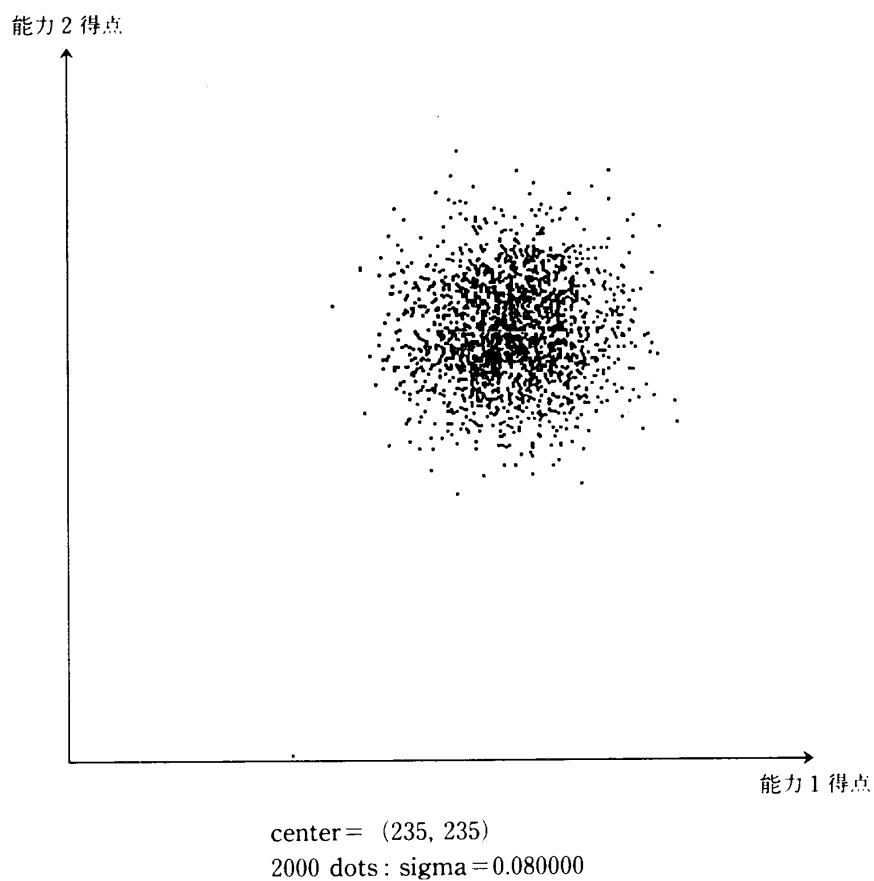


図 3 J 型母集団の得点分布

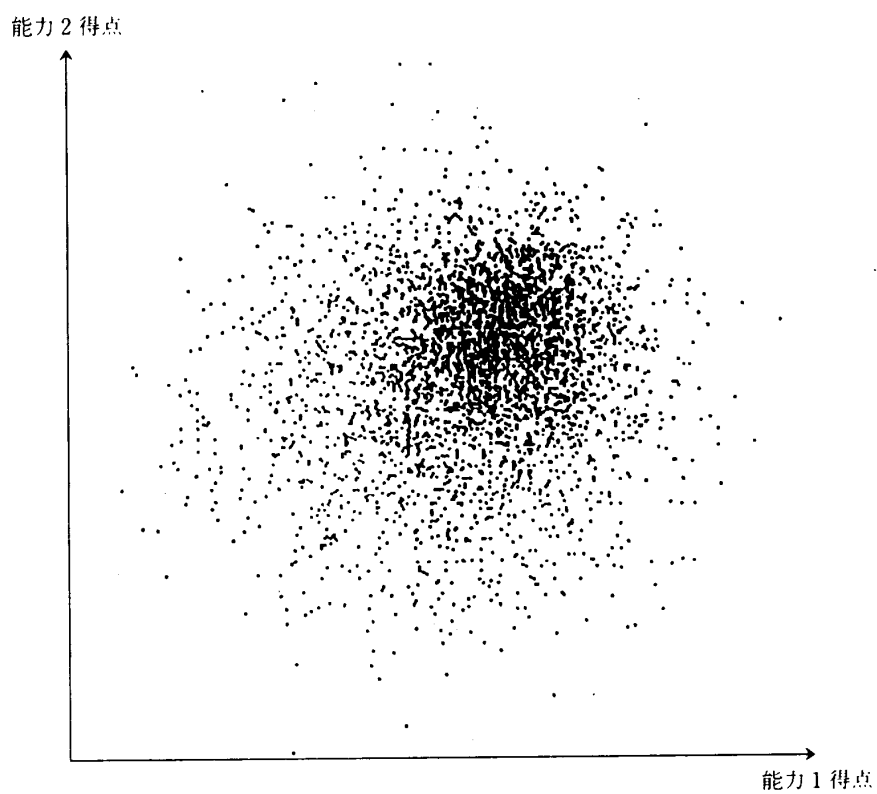


図 4 A 型分布と J 型分布の重ね合わせ

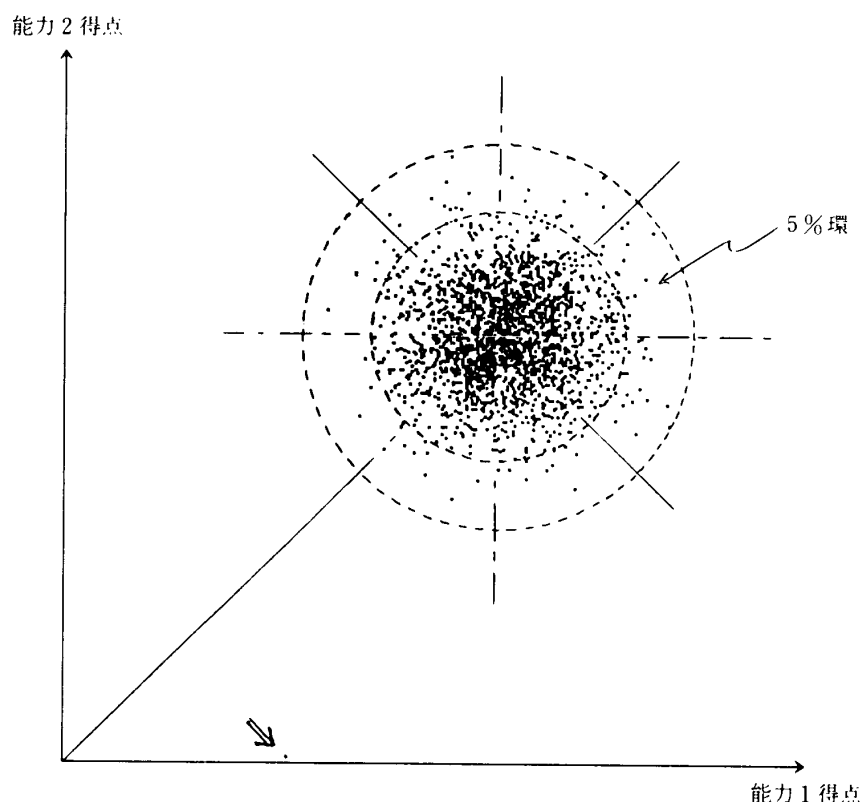


図5 J型得点分布の5%外環域

2. テスト結果の総合評価方式

入学試験、入社試験など、この世には各種のテストが氾濫している。それぞれのテストは、その道の専門家あるいは経験者が長い年月をかけ、いろいろと苦心をして築きあげた、各種才能・技能の評価方法である。ある目的で課される試験は、一般には種類の異なる複数個のテストの集合から成っている。また、一つのテストも、よく検討すると、一般には複数個の才能・技能の評価課題から成っている。今までのところでは、各テストの問題作成者の努力と工夫は、主としてこれら個々の出題に集中されてきたようである。

その問題作成の重要さは充分評価するが、ここではそうしたテストの内容の切り分けと出題はすでに与えられたものとして、その内容の構造には立ち入らない。ここで検討していくのは、そうした個々のテストによる評価が与えられたあと、それら複数個の評価値を基に、どうやって総合的な個人の評価を出し、順序系列を決めるのが良いかということに限っておく。

異なる背景を持った大勢の方がたに読んでもらえることを願って、以下の説明はできるだけ定性的に、やさしく書くようにしたい。

(A) 一芸方式

まず第1に考えられる評価法は、いろいろとある能力のうち、特に注目している一つの能力に限ってテストを行ない、他の能力についてはいっさい無視するという、一芸方式である。図6はこの一芸

方式によって集団をテストしたときに、合格する人たちを、二つの能力、つまり2次元に限って示している。当然ながら、領域Aにある者が合格し、領域Bにある者は不合格になる。一般にn種の能力がある場合には、n次元空間中に分布している点を、注目している能力の軸に垂直な一つの(n-1)次元平面で切り分けることになる。

この方式では、注目している能力において劣っている者は、いかに他の能力が優れていても合格することはない。反対に、注目している能力において優れている者は、いかに他の能力が欠如して

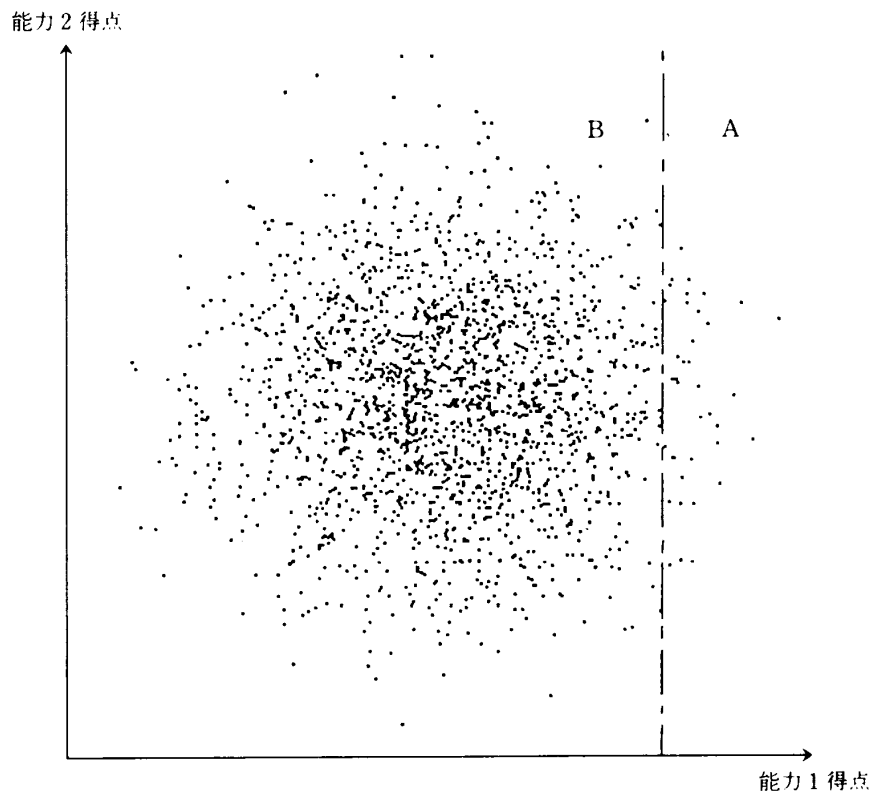


図6 一芸方式による選抜

も合格することになる。この一芸方式は、たとえば芸術家などを評価するときにはしばしば用いられている評価法であろう。

各個人はただ一芸によって評価されるわけであるが、目的によって別べつの能力に関してこの評価方式を関連した集団に適用すれば、全体としては、各能力について、おのおの最も秀でた者が選出されることになる。しかし、テストされたもの以外の能力については、なんらまとまってくるわけではなく、ばらばらのままであるから、合格者たち全体としては個性の強い者たちの集団となりやすい。

(B) 総合点方式

上の一芸方式と対照的なのが、おのおのの能力の評価の総合得点の上位から合格させていく、総合点方式である。まえと同じく、これを二つの能力評価の場合について例示したのが図7である。図7では、実線の右、領域Aにある者が合格し、領域Bにある者は不合格となる。ここで実線は、領域Aにある者の数が図6の場合の合格者数とほぼ同じになるように引いてある。ここでは、実線は第1能

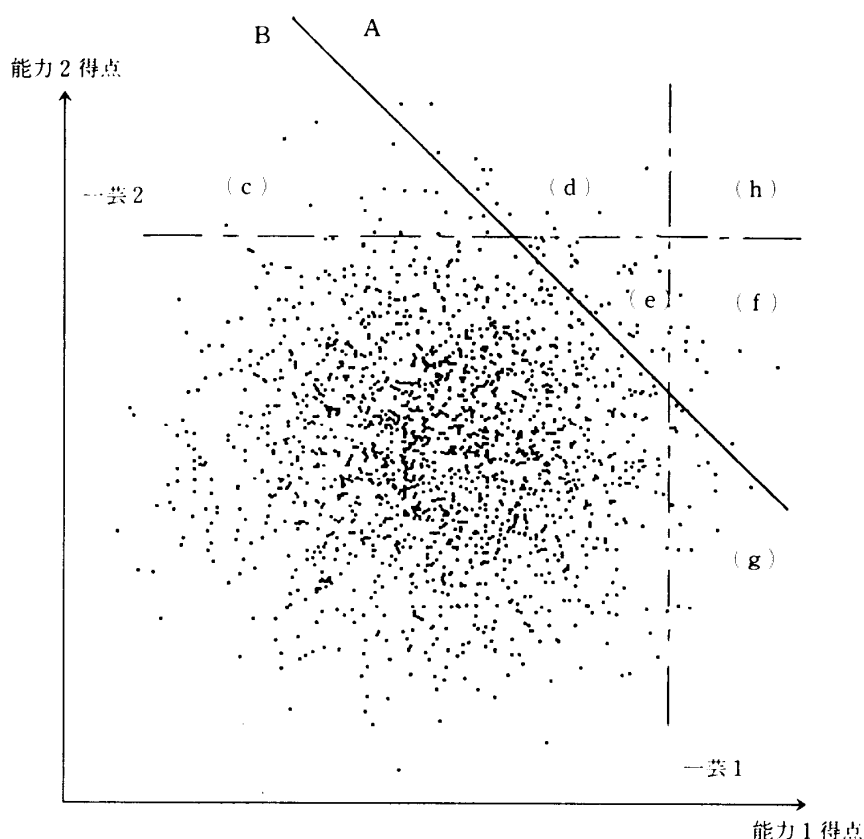


図7 総合点方式による選抜

力 (x 軸方向) と第 2 能力 (y 軸方向) との評価点に同じ重み付けを与えて総計する場合を示している。

この両者に与える重み付けの割り合いを変化させれば、実線の傾斜が変化する。その極限として、第 2 能力の重みを零にしたのが右側の垂直破線であり、第 1 能力の重みを零にしたのが上方にある水平破線になる。これらの極限は、図 6 に示した一芸方式の評価法に外ならない。

ここで注意して欲しいのは、この総合点方式による評価で合格することになる、三角領域(e)にある者は、一芸方式の評価では、第 1 能力による評価でも第 2 能力による評価でも不合格となる者であることである。また反対に、一芸方式によれば、第 1 能力による合格者である(g)領域中の者、あるいは、第 2 能力による(c)領域中の者は、ともに不合格となってしまふことである。

すなわち、一般に総合点方式では、評価の対象とする能力の数をふやせばふやすほど、それだけある特定の能力に傑出する者が相対的に淘汰され、領域(e)の一般化である万能選手、つまりそつのない者がその分余計に拾いあげられるということになる。

(C) 順次方式

この評価方式は、その名のとおり、複数個の評価基準について、独立に順次評価していく方式である。これを二つの評価基準に適用したものが図 8 の例示である。1 次評価の結果合格するのは、領域 A および B の中の者であり、第 2 次評価の結果、領域 A の者だけが最終的に合格する。各評価ごとに候補者を減らしていく過程は、俗に足切りと呼ばれている。この評価方式が異なった評価の順番付け

に対して同じ結果を与えることは明らかである。

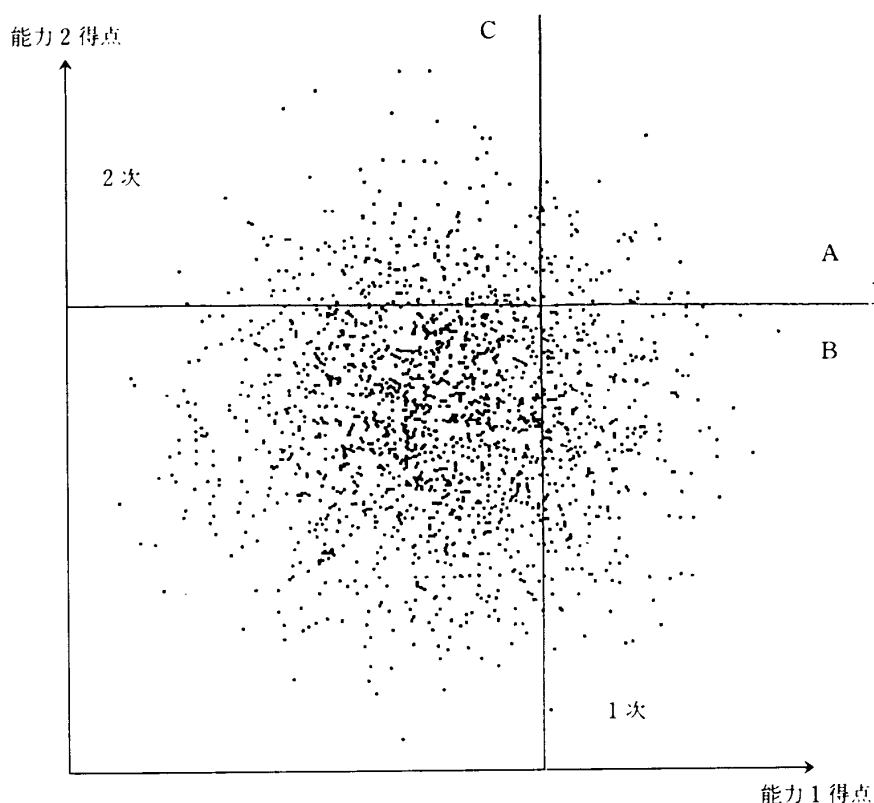


図8 順次方式による選抜

この方式は、現在大学入試などに大幅に取り入れられている。いまこの図8の評価線に図7と図6の評価線を重ねてみると、一般に図9のようになる。これで見ると明らかなように、順次方式では、総合点方式に比べても、領域(d)にあるそつのない者を余分に拾いあげる一方において、領域(c)、(e)にある、一芸に秀でた者をその分だけ排除することになる。

(D) 折衷方式

この方式は大学院の入試など、各専攻ごとに小人数の候補者の選考を行なうような場合にしばしば見られる、制度というよりはむしろ現象というべき方式である。

建て前としては、たとえば1次試験、2次試験、面接試験と、順次評価方式になってはいるものの、実状としては、1次試験による暫定的な足切り作業後も、2次試験の成績だけではなく、しばしば1次試験の成績を参考にした総合評価が行なわれ、また、ときにはそのあとでも厳しい面接点評価によって取り捨てが行なわれる。

これを図10によって見ると、1次試験によって、垂直線の左側にある者が足切りを受けたあと、2次試験の成績評価のための水平線に加えて、総合点による評価を表わす斜線が考慮され、その結果、領域(d)の者の代わりに、総合点の良いことを理由に、領域(c)の者が拾われたりすることがあるということである。また多次元のうちでも、ある特定の能力評価線(一般には面)、たとえば面接点の良いこ

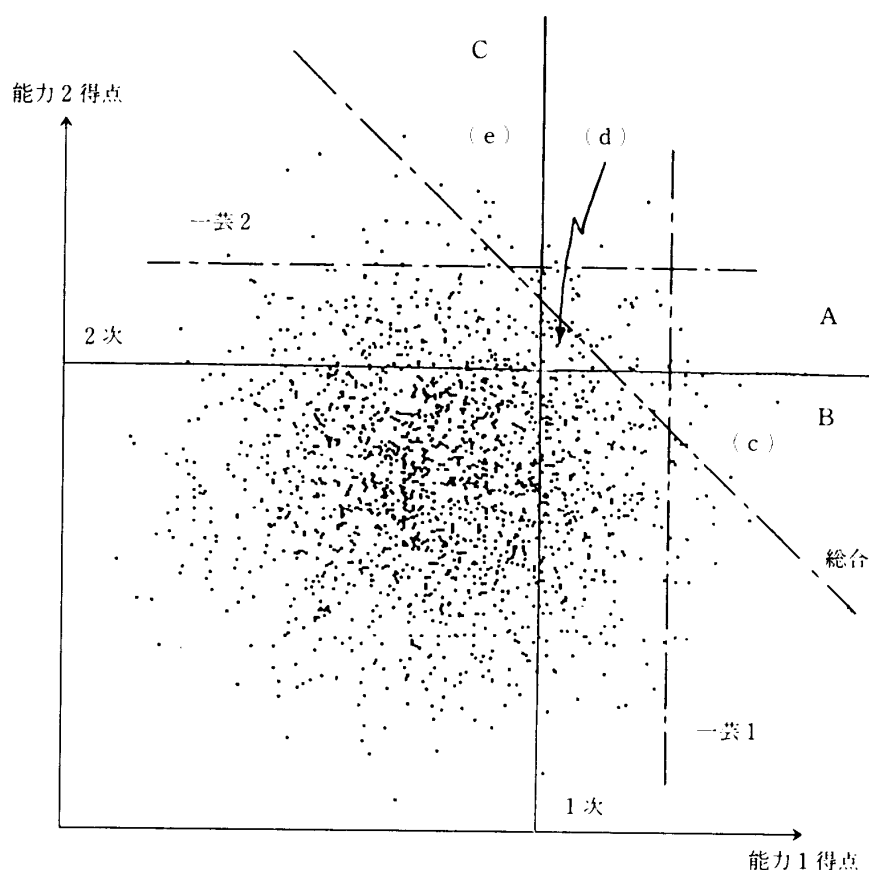


図9 3選抜方式の比較

とが特に厳しく要求されたりする。大学院入試は資格試験の要素が強く、必ずしも定員数だけ合格させることが要求されていないから、足切り後でも、そうした調整が可能になっている。

ここで注意して欲しいのは、順次方式、折衷方式とも、1次試験はしばしば基礎学力試験として、どちらかというとも基礎課目の一般の総合評価となっていることである。したがって、専門において特に秀でた者を含む、領域(a)あるいはその近傍にある者は、1次試験で早ばやと切り捨てられてしまっているから、2次試験以降では、これらの者は救うことはできないことになる。

3. 時間の経過にともなう動的变化

以上の試論においては、ある与えられた母集団に属する個人の持つ多種の能力をかなり良く反映している得点のデータが手に入っていることを前提としている。しかもそのデータは、ある与えられた時点における静的なものである。

しかし、人間社会は動的なものであり、その中において個人は、時とともに、たえず変化し続けている。そうした流れの中において、定期的に各種の能力テスト、特に入学試験などを受け、かつその結果によって社会の中における個人の進路が方向づけられているときには、ある方向にむかって選抜され続けた集団の持つ能力評価得点の分布は、その過程でどう変化していくであろうか。

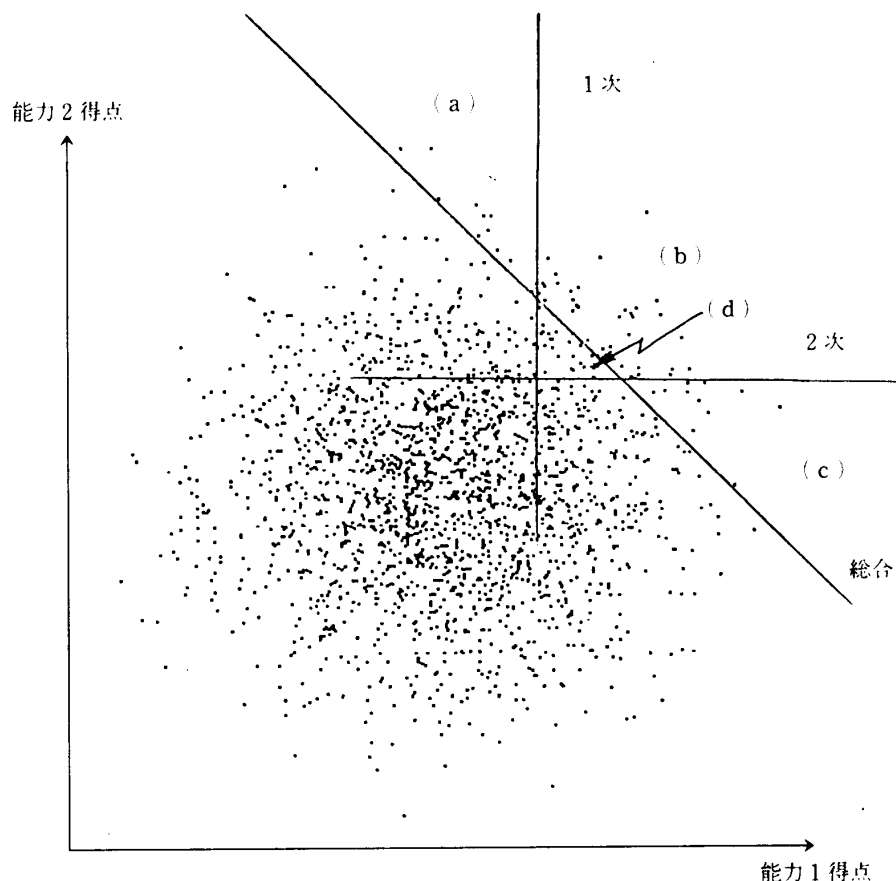


図10 折衷方式による選抜

これはなかなかむずかしい問題であり、確かな答をすぐさま出すだけの一般的知見は、おそらくまだ形成されてはいないであろう。ここでは、前節でみた各種の評価方式の性格を基にして、ごく大雑把な試論を述べてみたい。

第1に、世間における各種の選抜テストにおいて、いままでもっともよく使われてきた評価法は総合点方式であろうと思われる。しかし最近では、生活水準の向上と教育の普及にともない、高等教育機関など、各種の資格付与機関への入学志望者数が増大しており、資格審査試験への応募者数が増え、選抜試験が大型化し、順次試験の形態も増えてきた。

たとえば、旧制大学時代の入試といえば、学力の総合試験が1度あるだけで、2次試験は主として性格異状の可能性などをみる面接試験と、健康状態を調べる身体検査だけであり、一般にこれらの2次テストは、余程異状がない限り合格する、人数制限的な意味合いのないものであったろう。

しかし最近では、たとえば共通1次試験の採用など、順次選抜式の併用へと変わってきた。前節で指摘しておいたように、総合点選抜方式自体が、どちらかというとならで万能選手に有利に働く上に、順次選抜方式はその傾向を一層強化する。そこにおいては、たとえ見かけ上は一芸能力に秀でた者を拾いあげるかに見える折衷方式を加味したとしても、基礎課目能力の検定の名目のもとに行なわれる1次試験において、専門課目における真の一芸能力者の多くはすでに排除されたあとであるから、も

はやかれらを救うことはできない。

それだけでなく、わが国の世の中一般は、一芸能力主義から万能主義尊重の方向へ、時代とともにゆっくりと進んできたと考えられる。入試といえばむかしは中学校から始まったが、それでも全人口に対する進学志望者の割り合いは少なく、初めは大学でさえ、多くの学科では実質上無試験入学に近かったようだ。それが今では、幼稚園でさえかなりめんどうな入試を課するところが出現し、その後の各段階での入試はますます精緻になり、しかも国民の大多数がそうしたテストを順次受けさせられるようになってきた。

そうした状況のもとに置かれた集団には、全体としてどんな時間的変化が起こるであろうか。以下の議論は、今後まだ定量的理論化を必要とする定性的試論ではあるが、それでも、おおよそ次のようなことは言えるのではなかろうか。

- (1) まず第1に、ある特定のテストを受けるという人びとの集団によって、全人口は部分集団群に分けられる。ただし同じ個人が、一般には異なった複数個のテストに同時に関心があるから、1人がいくつかの部分集団に属しうる。つまり、これらの部分集団群は、排他的な分割 (partition) ではなく、重なり合う部分を持つ被覆 (cover) を形づくっている。
- (2) ある部分集団に属する個人たちは、たとえ評価法の意味することの細部をよく理解していないとしても、評価結果自体は強く意識するようになり、かつ得点の向上に努めるから、その部分集団の平均得点は時とともに向上する傾向がある。
- (3) ある部分集団では、それに属する資格をうる前に、まず別の集団に属しなければならないという前提条件 (prerequisite) がある。それによって、部分集団の集合はしばしば階層構造を持つ。たとえば入試を目的とした集団などがこれに当たる。このとき、もし、ある能力について同じテストが別べつの集団に課される場合には、同一の規準のもとでは上位の階層の集団における平均得点は下位のものより一般に高くなる。実際には、ふつうそのような場合には、階層上位の集団のテストとしては、よりむずかしいものが課せられる。たとえば同じ歴史のテストでも、高校のものは中学のものよりもむずかしい。このことは、のちに国際比較をするときに効いてくる要素となる。
- (4) テストの結果の評価に総合点方式、さらには順序方式を採用することによって、一芸能力者よりは万能者優先の傾向が助長される。したがって、そうしたテストをくり返し課されて数多い選抜を経験してきた上位階層の集団に所属する者は、単に平均得点が向上するだけでなく、万能型へと固まってくる、つまり各能力ごとの偏差が小さくなっていく。すなわち、どちらかというに一芸能力者よりは万能者を有利にするテストを次つぎに課されているテスト社会においては、全集団の得点水準が高くなるのはもちろん、その上に偏差の小さい、粒のそろった個人の集団へと移向する。各テストが一芸能力者の排除の程度を強めれば、それだけこの傾向は強まることになる。

以上のことを直観的にみるには、ふたたび図1と図3とを比べればよい。つまり、もし下位階層の集団の得点値が図1のような分布を示しているとしたら、上位の集団に施した、似かよったテストの得点値は、図3のように、平均値が高く、しかもより凝縮した分布を示すことになろう。

ここで、先に見た図7の分布に、図2で示した外円を重ねると、もし各集団が総合評価を繰り返して採れば、それは各集団がその円内の、直線の右上の切片内にある高成績者を切り出していることになるから、選抜後の平均得点は上がるとしても、分布のほうは、たとえば図3のような分布にはならず、右上に円弧部分を向けた円の切片の形になっていくのではないかと、という疑問が出る。しかし実際には、たとえば中学、高校と順次に課せられる入試の歴史の試験のように、後次段階でのテストが類似のものだとは言っても、詳しくみれば、その内容には充分変化があり、かつ、より高い水準のものになっているから、その得点は各回でけっこう変動し、分布が充分丸くなくなってしまふようである。事実このことは、入学試験を含め、何回ものテストを受けて、やっと大学入試にまであがってきた受験者の得点が、けっこう丸く分布していることから分かる。この辺の議論は、前節におけるテスト結果の評価方式についての試論とは、やや整合性を欠くものになっていることになる。

それでは、分布はなぜこのように凝縮していくのであろうか。

まず、こうした総合評価的選抜過程の繰り返しを経ると、分布が左下のほうから右上向きに縮んでくるのは分かる。同時にこれは一芸能力者の排除が、左上から右下へ、かつ右下から左上へと分布を圧縮してくるということにつながっていることに注意したい。

しかし、今までに見た評価法には、グラフの右上端に位置する万能型高得点者を排除する何ものもなかったにもかかわらず、あとに述べる国際比較の結果からみると、そうした評価法による選抜過程を経ることは、得点分布を右上のほうからも、左下に向けて圧縮するという傾向を、現象として起こしているようだ。

こうした凝縮が明白なかたちで実在するものかどうかは、さらに詳しい調査を必要とする研究課題であろうが、ここではひとまずそうした現象があることは認めた上で、次にはまずその理由について考え、さらにその文化的意味について考察してみることにする。

4. 得点分布の型と文化の型

上において、総合点選抜テストが繰り返して施されていると、集団としての得点分布は平均値が高く、かつ分散が小さくなっていく傾向を持つのではないかと推論した。そこで問題になったのは、果たして分布が得点の高いところから低いほうへと凝縮することまでも起こるのだろうか、また、もし起こるとしたら、その理由は何かということであった。

そうしたことの起こりうる理由の一つは、まず抽象的に言うと、諸能力の分布の型に個人差があるときには、ある型のテストの順次適用は、ある分布型を持つ者を優先的に選抜することになるからではないか、ということである。

もっと具体的に述べると、いま仮りに、たとえば細かく異なる100種の能力を考え、その内から無作為に半分を選び、それらについてある集団をテストしたとする。その結果、全てについて平均的な成績を示すような者の部分集団をとり、残る半分のテストを課すと、傾向としてかれらは、やはり今度も平均に近い成績を挙げるのではないだろうか。一方、さきの50種の1次テストについて上から下

まで大幅に変動する成績を示した者の集団は、残りの 50 種のテストについても、やはり大幅に変動する成績を示すのではないか。すなわち、万能型、一芸型というのは、個人について出現する型であって、テストに選んだ能力にはそれほど依存していないものなのではないか、ということである。

図を用いてこれを見ると、たとえば二つの能力についての得点分布を示す図 2 において、外縁環のなかにはいった者全体が、次にこれら二つ以外のいくつかの能力について他の者たちといっしょにテストされた場合、かれらの得点は、全員の得点の分布のなかでやはり（円環の一般化である）外縁殻のなかに、より集中する確率が高いのではなからうか。

これはなかなか興味ある疑問である。性格についての専門家によって、こうしたことはもう調べがっていることなのかもしれない。もしまだだとしたら、それは人間の能力についての本質的な研究課題になるのではなからうか。

とにかく、もしここに述べた推論が正しければ、総合点方式による評価の採用は、テストのたびに万能型を優先的に残すということと、回を重ねるとともに成績の平均が上がり偏差が減るという、二つの現象に対して、矛盾のない説明を与えてくれるようである。そればかりではない。以下に述べるように、ことはもっと重大な意味合いを持っているように思える。

まず、このような傾向が存在するかどうかを述べることができるのは、図 3 に対する図 1 といったぐあいに、実はほかの評価方式とその運用体系を持った異集団、つまり異文化集団における得点分布の型との比較においてのことである。一方、国際的に、日本人は教育のレベルが高く、かつ粒のそろった民族であるという評価が、しばしばなされている。そこで、いままでに述べてきた得点分布の型の形成過程がもし正しいとすれば、それは図式的には次のような説明によって定式化できるのではなからうか。

まず図 1 に示した一般的分布をとり、それに、それより平均値が高く分散の小さい、図 5 の分布を示す外円環だけを重ねてやったものが図 11 である。（図 5 の点そのもの、すなわち図 3 の分布の点は示されていない。）なお、図 11 の黒点は、図 1 の分布の中心、つまり平均点である。

ここで見て明らかなことは、分散の小さい分布のほうの平均値、つまり円の中心は、全ての能力において黒点よりも高い評価点を持っている。それでも、分散の大きい分布は円外の領域(a)にいくつかの点を含んでいる、つまり、より高得点の個人の存在を示している。

別の観点からそのことを示すために、図 11 の分布を、ある能力についての分布として眺めてみることにする。すなわち、この分布が紙面から手前に突き出している丘で表わされているとして、これを能力 1 の軸、つまり下のほうから眺める、言い換えると、能力 1 の軸への射影をとってみる。すると、それはたとえば一般的分布として図 12 の A 集団の曲線のような丘に見えるであろう。これに対して、分散の小さい分布のほうは図 12 の J 集団のように見えることになる。ここで、おのおのの曲線の下での面積は、それぞれの分布の正規化された総人口を示し、2 者は同じ面積を持っている。

ちなみに、いま例として能力 1 についての分布を見たが、一般にはどの能力についても同じような分布を示しそうである。さらに一般的には、総合得点に関する分布、つまり図 11 の原点と分布の中心

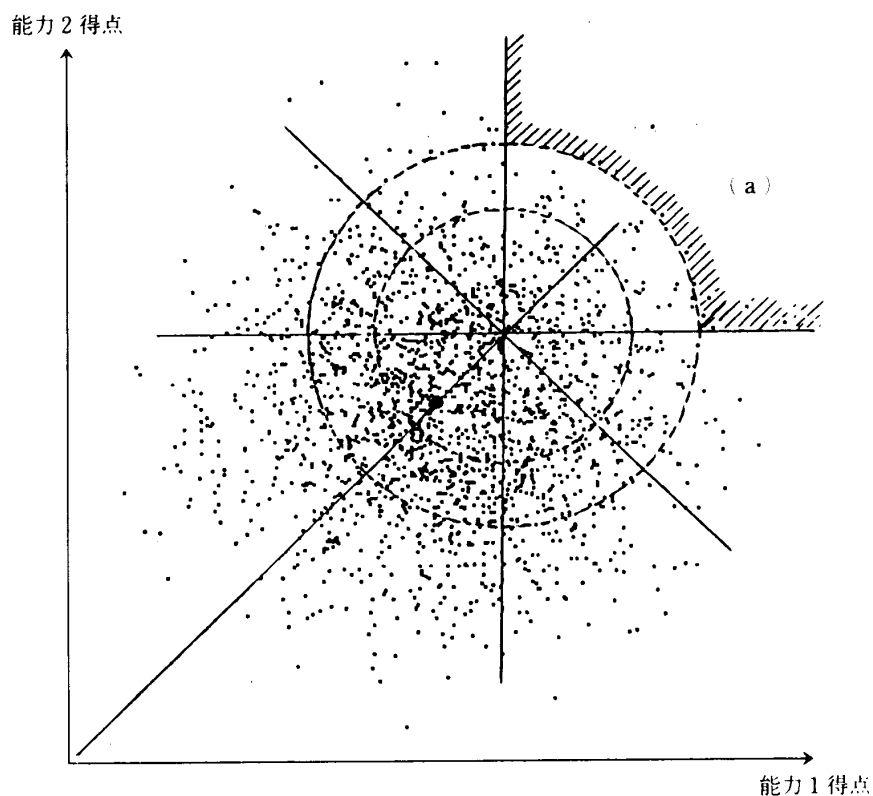


図 11 A 型母集団と J 型分布域

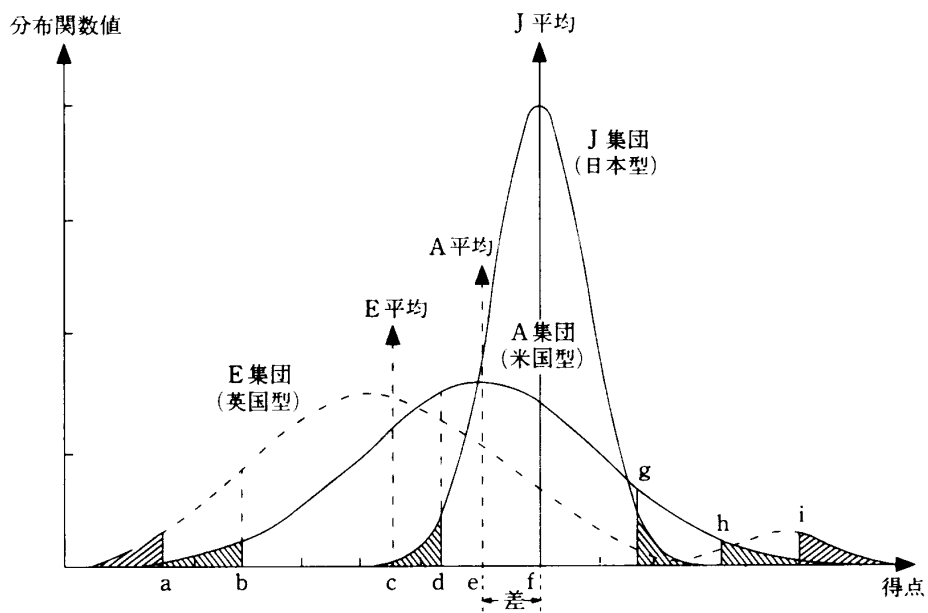


図 12 平均と分散の異なる三つの分布の型

とを結ぶ斜めの線に沿っての分布も、得点そのものは増加するが、分布の形は同じようになろう。

こうして作られた、なにかの能力についての得点の分布を抽象的に示している図 12 を眺めると、い

くつかのことを直観的に読みとることができる。いまここで、文化の型の国際的比較という視点をとることにし、図1に示した一般的分布の集団Aを、たとえばアメリカ型文化に対比し、図3に示した分散の小さい分布の集団Jを日本型に対比して考えてみる。すると集団Jの平均値は集団Aの平均値よりも高く、かつ集団Jの分散は集団Aの分散よりも小さい、つまり峻しく高い丘になっている。すでに図1の説明で述べたように、ここで曲線の左右が対称になっているのは能力の分布をガウス分布で近似したからであるが、能力の分布のようなものは厳密にはガウス分布にはならず、一般には左右非対称になる。しかし、ここにおけるわれわれの目的のためには、ガウス分布による近似で充分まにあう。

さて、分散が小さいということは、いわゆる粒がそろっているということである。たとえば、あるものごとに対する考え方が似通っているわけだから、お互いに知識の交換や考えの伝達がやさしい、つまり意志の疎通が相対的にらくであるということである。これはたとえば大量生産のための大工場などで、作業員を組織するのが相対的にらくになることを意味する。こうした粒のそろった従業員集団の存在によって、日本の産業界が大いに利益を受けていることはよく知られている。

その反面、いま独創性という問題を取りあげてみると、そもそも独創性というのは、今までの常識を越えたところにある、新しい知見の形成能力のことである。つまりそれは、平均値を中心とするところに密集している人びとよりは、平均値から大きく離れた人びとのほうが貢献の度が大きくなることが期待される素質の一つであると考えられる。図11について言えば、一般にそれは上端、右上、右端に位置する人びとに、より期待されるところの素質である。また図12について言えば、一般には各曲線の右端のgやhの右側の、影つけをした部分にいる人たちに大いに期待されてよい素質ということになる。

こうした観点から図11を見ると、たとえ全体の平均値は低くても、分散の大きいアメリカ型の分布では、日本型の分布の限界を示す枠を越えた領域(a)にも、なおかつ個人が存在していることが分かる。見やすさのために、図12では少し誇張してそれを描いてあるが、アメリカ型分布中の、hの右に位置する創造性集団の得点は、日本型分布でgの右にいる集団を大きく越えたところに位置している。

現在日本は、アメリカなどに比べて、訓練度が高く、粒のそろった、勤勉な労働力を活用し、信頼性の高い先端技術製品を量産し、国際経済摩擦を引きおこすほどの安価で世界に供給している。反面、独創性に乏しく、製品の基となっている基礎概念の多くが、外国からの導入ものであるという国際的指摘を受けている。ここでそれを全面的に肯定するつもりはないが、しかし図12はそうした指摘とはよく整合した、二国間の文化の型の相違を表わしていると見ることができる。事実、人口が日本の約2倍のアメリカは、独創性の一つの尺度となっているノーベル賞を、科学の分野で日本の30倍以上も受賞しているのである。

ちなみに、図12における集団Eの曲線は、社会において実質上の階級制度が存在し、階級によって教育などの扱われかたが異なるときの能力点分布の型を、いささか誇張して描いたものである。たとえば、ノーベル賞の受賞者数は、人口の割りに多いのに、工業製品においてはあまり振るわないとい

った、今日のイギリスの状況などを、この図は示唆しているとみることができる。

5. 日本文化の均質性

上には、日本文化の特徴の一つとして、民衆の教育水準が高く、しかも粒がそろっているということを取りあげ、かつそれが、先に述べた、高度に組織されたテスト管理社会がもたらす結果の推定とよく整合していることを述べた。当然ながら、そうした競争の繰りかえしは、副産物として、民衆の勤勉さを助長するのにも大いにあずかっていることであろう。

ただし、ここではそうした文化的特徴が全てテスト制度の産物であるという、単純な因果関係の存在を主張するつもりはない。日本の均質的文化の発生には、たとえば使用文字の影響など(手塚 1987、山田 1987)、中にはちょっと思いがけないものも含んだ、数多くの要因がありうるから、社会のテスト志向は、ただその一つになっているのかもしれないということである。あるいは、そうした因果関係さえ全くなく、日本社会におけるこの二つの現象の存在は、たまたま偶然の一致であるのかもしれない。さらには、むしろ逆に、日本文化の中に流れている、均質性への志向が、徐々にテスト管理社会の成立を助長してきたのかもしれない。

日本文化の特質全般を論ずるのは本稿の目的ではないから、以下では均質性志向に限って話を進めることにする。一般的にいうと、諸外国に比べて、日本の社会が全ての面で質の均一性を特に重要視していることは疑いのないことである。この均質性、画一性への努力は、早くも幼児のしつけから始まることは、以前から指摘されていることである(たとえば Benedict 1946)。

アメリカと日本と二つの国で、どちらもそれぞれの幼稚園と小学校の両方の教育を経験したわが家の2児の日常を観察した範囲内においても、この事実は随処に散見できる。日本における制服の着用(朝日新聞 1989-5-18)、教材の画一化(朝日新聞 1989-6-22)、全国共通の検定教科書と指導要項の存在(木村 1989)などが、その手はじめである。

アメリカの小学校では、日本に比べてかなり小人数のひとクラスの中でさえ、(そして小人数のクラスだからできることでもあるが、)一人ひとり生徒の能力に合わせて、教材や参考書を変えてくれることさえある。それは大幅に先生たちの裁量にまかされているし、もちろん全国的統一の検定教科書や指導要項などはない。そうした在り方の良し悪しは、別に議論しなければならないことであるが、これがアメリカ人のあいだにおいて、知識の大幅な個人差と多様性に大いに貢献している。

画一的に管理された教育と社会生活を経てきた日本人が、均質性を通りこして、画一性へと向うのは当然の勢いなのかもしれない。よく言われることだが、海外旅行ブームとはいっても、同じ髪型、同じ服装の若者たちが団体として、一斉にガイドブックの指示どおりに歩き、おみやげ店でブランド商品を大量に買って帰る日本人は、外国人の眼から見ると、こっけいでもあり、恐ろしくも見えるようだ。

日本におけるこうした均質性志向の文化的傾向には、本稿に述べてきた、教育過程における各種テストの連続が与える均質化の影響が、かなりの拍車をかけていることと思われる。この国民のあいだ

の均質性は、すでにいくつか述べたような、いろいろの利点をわれわれの社会にもたらす。しかし、同時にそれは不利に作用する面のあることも忘れてはならない。

アメリカに比べて日本のコンピュータ・ソフトウェアが遅れている理由の一つとして、漢字の使用が挙げられることがある。すなわち、漢字の入力や記憶が重荷となって、現在の世界の先進国の中で、日常の使用言語以外の言語系に基づくプログラム言語を常用していることでは、日本と中国がきわだっているという。

しかし、プログラム作成に関わる思考の形態は、しばしば日常言語の構造とかなりかけ離れているから、そのこと自体が、それほど彼我の差を生み出しているとは思えない。むしろ、日本においてソフトウェア会社を設立したアメリカ人サリバン (Bonnie Sullivan) の指摘するように、優れたプログラム作成には傑出した自由な発想が不可欠であるのに、他と異なることに後めたささえ覚えるという日本文化の性格のほうが問題なのであろう (Ward 1989)。

その反面、日本で作成されたソフトウェアはキメが細かく、信頼性が高いという国際的評価がある。全米コンピュータ・グラフィクス協会 (NCGA) のオア (Joel Orr) の指摘によると、個性をむき出しにしないプログラムは、一般に誤りが少ないということだから、こちらのほうも、やはり日本文化の性格がプログラムの質に反映されているのだと言えよう。

この粒をそろえるという信念が国家的規模で発揮された1例が、日本の経済政策であろう。すなわち、家電製品、特に先端技術エレクトロニクスのように世界の先頭を走っている産業がその先導力を失わないように勉めるかたわら、靴や航空機のように、まだ国際的競争力のない産業においては、輸入制限などを適用して、国内産業の成長を助成する。そればかりか、農業や畜産業、はたまた兵器産業のように、どう逆立ちしても、まず永久に国際的な競争力を獲得することが不可能な立地条件にある産業でさえ、国民全体の犠牲において手厚く保護を加えている (Fallows 1989, 松井 1989)。

同じような哲学は個々の企業体にも見られるところである。エレクトロニクスを例にとれば、全てのメーカーが、ラジオ、テレビをはじめ、ICチップ、コンピュータ、ファックスと、およそ世界で作られている物は全て、アメリカでなら独占禁止法違反になりかねないほど制限的な、自系列の会社の枠内での製造、流通を始める。

こうした方策は、日本では鉄のような素材産業についてさえもあたりまえと考えられているが、産業界におけるこれらの慣行が、いま日本を巻き込んでいる国際経済摩擦の原因の一つになっている (朝日新聞 1989-8-11, -8-12(a))。

6. 均質性の国際的特異性

こうした均質性志向は、国際的見地に立って眺めるとき、世界の中で日本の文化の特異性をきわ立たせている。以下には、外国人の眼にうつる、日本文化の評価をいくつか拾ってみる。われわれがこれらの評価を異様に感じるとすれば、それは逆に、われわれの文化が外国人に異様に見えていることの反映である。

たとえば、自由圏の諸国に比べて、日本の大新聞のおおのの発行部数の極端に多いこと、しかも各紙とも傾向や主張がはっきりせず、紙題を外せば、ほとんど区別ができないことなどが外国人を驚かせる。これもわれわれ読者層の均一性を反映してのことである。

こうした均質性を尊ぶ心情の者は、内に対しては他と異なることを恐れ、外に対しては自分と異なる者を排除する動きを誘う。

異質の思想と価値感の中から、共通の行動目標を求めてゆくべき使命をもつ政治をとってみても、日本には思想と哲学を持った議論ができる政治家がいなく、世界の場において政治的指導性に欠けるばかりか、多くの場合われわれは、他民族との会話や議論を初めから避けることで、しばしばあわれみの目で見られている (House 1989, 朝日新聞 1989-5-10, 石田 1989)。

さらに、他と異なるということが異端視される結果として、外国人の眼から見ると、数多くの異質な現象が浮かびあがってくる。たとえば、調和をみだすことの象徴である訴訟の件数が日本では少ないこと、転職するような異端者は新しい職場で人並みに受け入れられないために、半ば仕方なく続いている終身雇傭制などは、日本人論者によってはしばしば美点として論じられているものの、実は日本文化の単なる特異性を反映しているに過ぎないと言える。その上、たとえば建築業界に見られる海外労働者に対する大幅な受け入れ制限、日本の専門家によってさえ、たびたび否定されているにもかかわらず (たとえば田中 1989)、われわれのあいだにいる少数民族の存在を否定する単一民族論や、生まれつき髪の毛の赤い者に髪を黒く染めるのを強要する (読売新聞 1989) という、広く世界的視野に立てば犯罪的な思考形態が存在するのも、この粒をそろえるという情動の産物であろう。

裏を返すと、これはわれわれと異なる者のあいだ、つまり外国のことなら、何であっても、大して強い関心を示さないということにもなるのであろう。たとえば、人種差別のゆえに世界から受けている、南アフリカ連邦の経済制裁に対して日本が消極的であるとか、今回の中国北京市における動乱の推移に対して、わが国の政府の発言が他の諸国に比べて格段に遅い、という国際的指摘などがその顕れであろう。

大幅に自分と異なるものの存在を当然とする社会においては、同時に絶対的価値というような抽象的思索を刺激する。それに対して、実体は何であれ、似たもの同士の集まりと共にあれば安心するという心情の持ち主たちのあいだでは、その中心に祭りあげるものの絶対的価値ということを考える必要はあまりない。

こうした心情は、日本人に、名誉とか規律とか義理とか、仲間うちでの忠誠心をもたらず反面、その中心的存在である、天皇制のように重要なことの本質に関する論議は一向に活性化しないという状況や、三権分立の大原則を放棄して、たとえば自衛隊の存在や教科書検定などに対する合否の憲法判断を下さず、逃げに徹している裁判所の行動は、国際的に特異なもので見られている。

絶対的価値観の尊重の度が大きく影響することの一つに、たとえば日本と外国とにおける会議の在り方の違いがある。少しばかり類型化した言い方になるが、全員に共通な絶対的価値というものがあることを、始めから確信している者の集まりにおいては、たとえ当初には相当な意見のくい違い

があっても、いや、くい違いがあればある程、絶対的価値を持つものが何であるかについて各人が主張し、相互に比較されつつ、意見の相違が修正され調整され、最後には一致した見解へと到着させることに勉める。これがたとえばアメリカで各種の委員会を構成するにあたり、おのおのそれなりに妥当ではあるが、しかし異なった意見の持ち主を集めて作る、論点 (contention) システムなどと呼ばれる、委員会の運用の在り方である。

それに対して日本では、委員会の各委員が、全く個人の立場で、絶対的価値を求めて討議するという観念が稀薄であり、会議においても各委員はむしろ自分の所属する集団に忠実で、その利害から大幅に離脱しないことを念頭に置きがちであるから、会議はえてして異集団の利益代表の意見調整になりがちである。したがって、そもそも委員会の編成にあたって、あまりに対立する意見、つまり利害の持ち主で構成したのでは、意見の収束が望めないから、始めからある程度意見のそろった者を委員として構成する、いわゆる合議 (consensus) システムになりがちである。

現在、国際経済摩擦にかかわる日米間の各種の政治的調整の会議において、その進み方を検討してみると、しばしば会議が難航するのは、この合議主義と論点主義という、日米間における会議の在り方についての理解の差によるところが大きいように思える (Fallows 1989)。

たとえば、日本の電気通信市場の開放をめぐる、先日の日米交渉においても、モトローラ社型自動車電話の首都圏参入について、日本側には確固たる理念に立つ政策がなく、最初はゼロの回答をしながら、自由貿易の理念に立つアメリカ側が強く押せば押すほど、日本側が譲歩を重ねて行った。こうした日本の態度では、ますます国際的不信感が助長されて行くことになろう。すでに最近のアメリカの世論調査では、米国との商取引について、日本が不公正であるとするものが60%もあり、しかもよりよく日本を理解しているはずの高学歴、高収入者ほど、この率が高まっているという (朝日新聞 1989-4-12)。

このような、自分の所属する集団に対する同質化を基調にして、理念抜きに自分の行動を律するという考え方は、よく東洋的思想と言われたりする。しかし、必ずしもそれは正しいとは言えず、むしろかなり日本独特のものである。

たとえば、アメリカにおいて日本からの情報が入手しにくいという不満が、最近の国際摩擦などからんで、しばしば欧米から聞かれるようになってきた。その原因の一端は、われわれの文章の表記法が漢字という、外国人にとっては特にやっかいな文字を用いることにあるだろう。

しかし台湾や韓国のような、漢字圏の新興工業国 (NIES) の諸国が国際経済に果たしている役割りも、最近では相当なものであるが、これらの国々について、情報の疏通が悪いという不満が欧米から起こっていない。その理由として、これらの国から欧米に移住し帰化した人びとの二世三世は、依然としてルーツ文化の言語能力を失わず、かれらを通して情報が充分伝わっていることが挙げられている。独り日本人だけが、新しく帰属した国の文化に一辺倒に忠誠となり、日本文化を急速に捨ててしまうので、日本からの情報の媒介者が特に少ないことが指摘されているそうである (斎藤 1989)。こうした日本人だけに特異な文化的行動も、均質性志向の現象とよく整合している。

とにかく、世界の政治経済がいよいよ国際化しつつある今日、日本だけが文化の特異性に固執し、導入アイデアを基とした技術立国の政策をとり続けていくことは、もはや不可能になりつつある。したがって、なによりもまずわれわれは、独創力を発揮する人材の養成と、独創性を発揮できる諸般の環境づくりとを、すみやかに実行に移さなければならない。

その方略を論ずるまえに、まず下準備として、ここでもう一度、第2節のテストの評価法に戻り、新しい評価尺度の在り方について考察する。

7. 従来とは異なる尺度による評価法の可能性

第2節に述べた四つの評価方式の全ては、各能力の得点を、その値だけに注目するスカラー量として扱い、最終評価を求めるものである。しかし、人間の能力の総合評価を求めるに当たり、各能力をスカラー量と考え、その算術和を用いる計算法は、能力にとって特に本質的なものではない。にもかかわらず、永年そうした評価方式が用いられてきたのは、加算が最も基礎的な演算であり、かつ易しいものだから、入学試験のように、大量の計算をきわめて短時間に実行しなければならない場合に、大いに都合が良かったからでと考えられる。

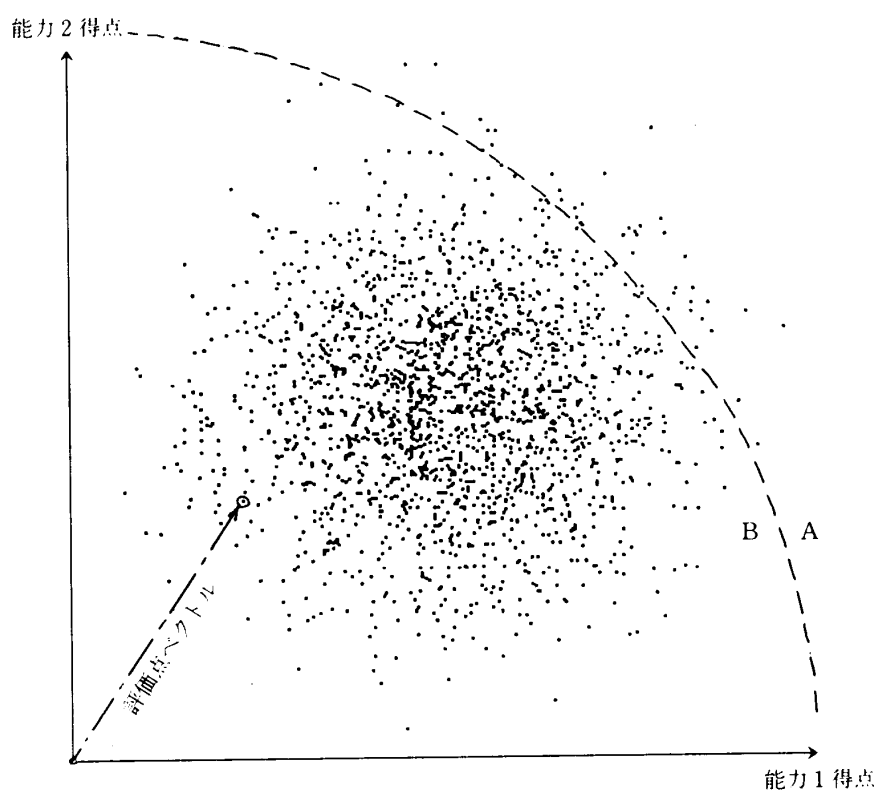


図 13 評価点ベクトル方式による選抜

しかし、いまはコンピュータの世の中であり、大量の計算の高速遂行は、もはや何でもないことである。だからこの辺で原点に立ち帰り、人間の能力そのものということを中心課題として、改めて個々

の評価点の総合評価の方法について考えてみてもよいのではなからうか。

その第一歩として、個々の能力の測定値が構成するベクトルとして、総合能力をとらえる評価方式を考えてみることにする。

(E) 評価点ベクトル方式

いま異なる n 個の能力の評価点が、 n 次元のユークリッド空間中の直交座標上での座標値を形成していると考えれば、この座標の原点からの距離の大きい者から順に採っていけば、各人の総合能力は原点からのベクトル長で評価していることになる。図 13 は、これを 2 能力について 2 次元上で例示したものである。この図から、各能力のベクトルの合成として、総合能力を評価する本方式は、各能力の得点の単純和という総合点方式に比べて、直観的にはもっと自然な数値化法に立脚していることがわかる。しかもこの図 13 と図 7 と比較してみれば、このベクトル方式は、一芸に秀でた者も拾い、万能選手も拾うという、一芸方式と総合点方式の両者の特徴を合わせ持つようなものになっている。

(F) 一般化ベクトル方式

だが、上に示したベクトル方式には、実は一つの問題がある。それは、ベクトルの原点として、全ての能力評価値が零の点を採用していることである。しかし、そもそもこの零点は各能力に対する出題の難易によってかなり人為的に定まってくるものであり、たとえある能力の得点が零であっても、そ

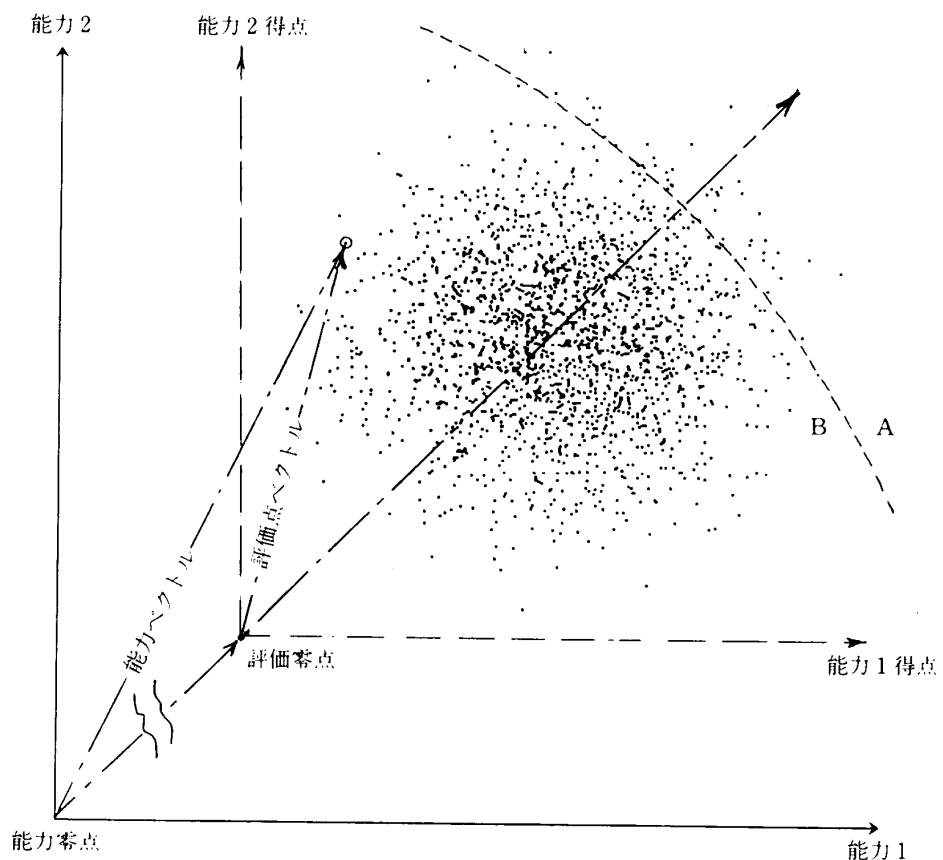


図 14 一般化ベクトル方式による選抜

の能力が絶対的に欠けていることを意味していないことは明らかである。ということは、もし能力絶対零の点を原点として考えた場合には、図 14 として一般化された能力ベクトルの方式において、右上の、ある適当な部分を表現しているのが図 13 だということになる。逆に言うと、図 13 におけるベクトルの原点は、一般には必ずしも評価点座標の原点である必要はなく、ある規準のもとに、どこに定めてもよいことになる。たとえばこのベクトル原点を、すべての座標軸上で同値を保ちつつ、負の無限大に移していった極限が、図 7 に例示した総合点方式に対応する。また、その総合点方式において能力ごとに重みを変えるとすることは、負の無限大に向う方向を変化させることに対応している。

この方式において、各能力ごとに評価の重み付けを変えることは、一般には高次元球体殻が高次元楕円体殻に変わることを意味する。以下では、重み付けが一様な球体殻の場合のみを考えることにする。

(G) 平均点ベクトル方式

今度はベクトルの原点を、得点の分布の中心、すなわち全項目の平均点を座標とする点に置いた場合の 2 次元例が、図 15 である。ここで合格するのは、たとえば領域 A にある者である。合格者の数は図示した内円の半径を変化させることによって調整することができる。(ここでベクトルを平均点ベクトルと呼ぶのは、ほかの場合と軌を異にするが、特に良い名前を考えることができないので、こう呼んでおく。しかし、これは平均得点から立てたベクトルであることには注意しよう。)

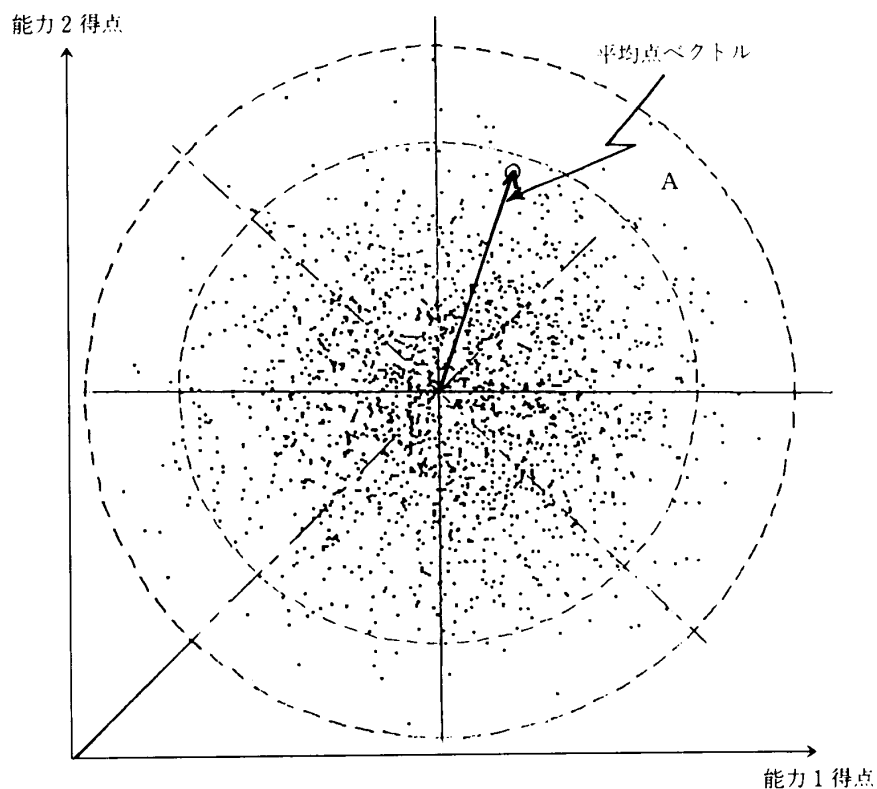


図 15 平均点ベクトル方式による選抜

ここで領域 A から合格者を採るということは、それらの者が全ての能力において少なくとも全体の平均値には達していなければならないという制約がついていることを意味する。

この方式を、現在よく使われている、図 7 の総合点方式と比較すると、万遍ない能力の持ち主よりも、一つの能力において突出している者の一部が、より優位に立って選ばれることになるのがわかる。すなわち、図 16 において、領域(a)および(b)中にある、一芸に秀でた者を余分に拾うとともに、その分だけ領域(c)にある万能選手は排除している。

(H) ベクトル方式の自由度

このことは、図 14 に示した一般化ベクトル方式において、ベクトルの原点を負の無限大から右上に動かせば動かすほど、より顕著になってくる性質である。こうしたぐあいに原点を選べる一般ベクトル方式においては、運用に当たっていくつかのパラメタの設定が必要になり、しかもその上に、種々の折衷案も考えられるであろう。まず第 1 に、ベクトル原点をどこに置くかを定めなければならない。たとえば原点の設定にあたっては、各能力の得点は、少なくともこれだけの得点は欲しいという要請に応じて原点を決めるようにすればよかろう。さらに、この方式を基とした折衷法の採用に当たっても、たとえば各能力のおのおのについて最低得点を要求するとか、あるいはそれら能力の総合点の最低を確保するとかが考えられる。

また、あまり一般的ではないが、特殊な目的のために、特に万能選手だけの選択に限定したいとき

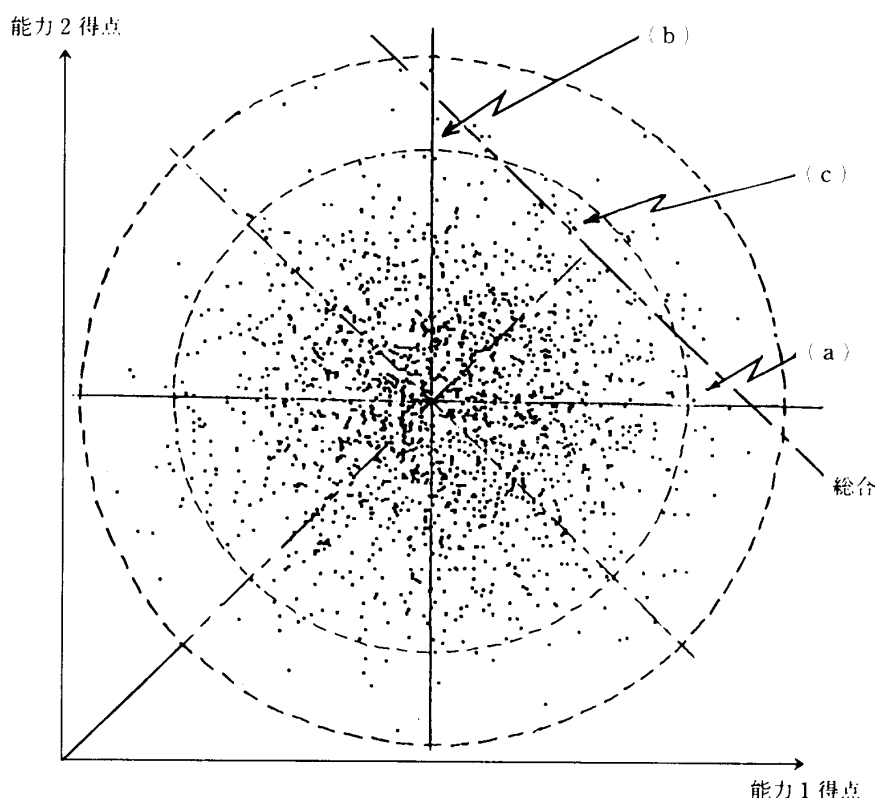


図 16 平均点ベクトル方式と総合点方式の比較

には、図 14 に帰って、能力零点を頂点とし、これと評価零点とを結ぶ直線を軸とする適当な(高次元)円錐を設定し、この円錐内にある者を高得点者から順に拾うことも考えられる。そのような選択方法は、図 9 でみた順次方式の場合以上に、個性のある一芸能力者を厳しく排除することになるのは明らかである。

以上、第 2 節では、現在実用にされている評価方式、また本節では、おそらく個人の総合能力をより忠実に数値化するであろう、新しい評価方式のいくつかずつについて、その特質を述べた。いま、個性化対平均化ということだけに観点を絞って、これらを比較してみると、個性的な者を最も多く拾いあげるのは一芸方式(の集合)であり、以下は、ベクトル方式、総合点方式、順次方式の順になると考えられる。ただし、折衷案としていろいろな制限条件をつけた場合には、これらの順序付けが変わることは、たとえば上に示した高次元円錐内に制限されたベクトル方式が、最も均一な万能者を選び出すことから明らかである。

ちなみに、図 12 に示した、ある一つの能力についての得点分布関数は、単純加算による総合得点をとった場合であるが、たとえばここに述べたベクトル方式を用いて総合評価をする場合には、この分布を出すのに単純な射影よりももう少し複雑な計算を必要とする。しかし、結果として得られる分布図は、形としてそれほど変わってはいない。

ここまでの議論は、直観的理解を重んじて、主として 2 次元の空間について述べてきた。しかし対象要素数を増して、議論を高次元に拡張するとき、高次元の空間は、ときには思いがけない性質を示すことがあるので、いままでの議論にはそうした落とし穴のないことを確かめておく必要がある。

そうした予想外の性質として、たとえば、(1)半径 1 の球体に外接する立方体の頂点の高さ(すなわち中心からの距離)は、3 次元では、1.73 であるが、100 次元では 10 となる。だから、(2)その外接立方体の体積となると、3 次元では球体の体積の 1.91 倍であるが、100 次元では、なんと 5.35×10^{69} 倍にもなる。しかし、(3)球体と同じ体積の立方体の辺の長さは、3 次元では 1.61 であり、100 次元では 0.402 となる。また(4)この同体積立方体の頂点の高さは、3 次元では 1.40 であるが、100 次元では 2.01 となる。

この(1)、あるいは(3)と(4)、から、立方体の頂点の高さは、3 次元では辺長の 0.866 倍であるが、100 次元では 5.00 倍となり、直観的には、一般に高次元の立方体は頂点がクリのいがのよりもきわだってくる。100 次元ではその数が 1.3×10^{30} 本もある。事実、(2)によると、一般に高次元の立方体では体積のほとんどが頂点に集中しているといえる。しかし、(4)にみられるように、球体と同体積の立方体を比較すると、当然ながら、頂点の高さは次元によってそう変わらない。

したがって、たとえば図 16 について述べた議論は、本質的には体積を同じに保ちつつ、高い n 次元球体と (n-1) 次元立方体(高次平面)との相互関係についての定性的なものであるから、詳しい議論は省略するが、上の(3)や(4)の部類に近く、本質において見当外れとはなっていないはずである。

8. 独創性の養成のためには

先端技術分野をめぐる、日米間の摩擦がいよいよ激しくなっているいま（朝日新聞 1989-5-19）、これまで日本の採ってきたアイデア導入政策は、技術ただ乗り論として、これからますます厳しい評価を受け、かつ、締めつけられていくことになる（Levine 1989）。

その原因の一端が、日本のかかえこんでいる大幅な貿易黒字であるのは事実である。しかし、新興の工業国（NIES）がどんどん力をつけてきている今日、この黒字状態は永久に安定していると考えるのは危険である。先端技術における優位をあくまでも維持しなければ、昨今のアメリカに見られだしたように、いつかは日本も NIES によって経済的基盤をうばわれてしまうことになる。

この両面からの圧力をはね返すには、われわれ自身の創造性を育てていくよりほかはない。そのためには、本稿でみてきたように、まず、個性を自由に伸ばしてやれる環境づくりが不可欠であると思われる。

だが、社会の構造は複雑で、いろいろな要素が密に結合されて全体を構成しており、他の要素に全く影響を与えることなく、特定の要素だけを変えることはなかなかできない。

早い話が、第2節、第3節で考察したように、総合点方式のテストによる選抜を繰り返して受けて成長すると、能力得点の平均は高いが、分散が小さい集団になる可能性がある。これは傑出した能力を持つ者を抑えこんでしまうことを意味する。

逆に、独創性を持った者を自由に伸ばしてやる環境を作ると、分散が大きくなることは避けられまい。同時に、おそらくそれは、必然的に平均値の低下を招く。その上、学生、生徒の多様化に対処するためには、教師の手間や教育のコストが大幅に増大することを覚悟しなければなるまい。

それでも、われわれが創造性の大幅な向上を図ることは、現在、国際的な要請となっている。それがなくても、過去の農業社会、工業社会に比べて、現在の情報社会は、価値観が多様化し、個性を自由に発揮することを希求する社会であるから、もしそれらを意図的に抑えれば、萎縮してしまう社会であることは、最近の東ヨーロッパの激動の示す、貴重な教訓である。

したがって、好むと好まざるとにかかわらず、これからの世界で日本が活力を維持し続けていくためには、個人の能力を自由に伸ばしてやれるように、社会構造の変革を進めていかななくてはなるまい。当然それはわれわれの社会の均質性を、大なり小なり手離すことから始めなければなるまい。

その手はじめとなるのが、小学校からの教育の多様化である。まずそれは、検定教科書の多様化、学習指導要項の多様化であろう（木村 1988）。しかし、この問題の性格と過去の経緯を考え合わせると、そうした前進が文部省内部で立案されることはなかなかむずかしそうである。それは国として、もっと高いレベルで処置されるべきことかもしれない。

そのほかにも、もろもろの施策について大いに検討されるべきであるが、それは本稿の目的とするところでない。

しかし、ここで一つだけ指摘しておきたいのは、第2節および第7節でとりあげた、テストの結果

の総合評価方式についてである。

この2節の議論ですでに明らかにされているように、個性を伸ばし、社会を多様化することを至上目標とするならば、分野ごとの一芸方式の集合を用いるのがもっとも効果的であろう。しかし、極端に非常識な個人の出現はある程度抑えることが必要であるから、全体的評価にあたっては、ある程度総合得点による成績評価を加味する必要があるだろう。

そうした観点からすると、一般的な選抜には、図13の評価点ベクトル方式によるのが良いように思える。しかし、特に選抜度が厳しく、少なくとも一般常識は身につけていることという要求があれば、図15の平均点ベクトル方式によるのがよいだろう。

特に、多くのポストを次つぎと移動する、万能選手的能力を期待される官僚を志す者の選抜には、この後者がよかろう。反対に、独創性を極度に重んじる分野においては、必要不可欠の能力だけから成る、低次元数の空間での評価点ベクトル方式が良いのではなかろうか。

いずれにしろ、現在主として用いられている、順次方式や、あるいは総合点方式の積み重ねは個性を抑圧する傾向が強過ぎよう。この辺は、専門家による、さらなる研究、調査を期待したい。

そうした検討の結果、ある方式が創造性の養成の目的に合致すると分かっても、全ての大学で一斉に同じ方式を採用するということは、これまた画一的に外ならない。むしろ全体としては、各大学が各専攻についてもっとも適した評価方式を選び、その方式をあらかじめ公表した上で入学試験を実施すれば、高等教育の多様化はより理想的になるであろう。

なお、大学入試の共通一次試験において、科目に対する配点を変えても、ある大学のある学部の2次試験の受験者の、得手による分類ごとの合格率が大きく変動しないというシミュレーションの結果が報告された(朝日新聞 1989-8-16)。その細部については調べていないが、まとめを見たところでは、その高い低いにかかわらず、総合力がバランスしている者については、それは当然予想される結果である。しかし、能力が文系あるいは理系に片寄っている者については、予想されることであるが、記事の示すグラフで見ても、やはりかなり変動しているように見え、記事の記述は誤解を招くものであるといえる。

さらに、本稿に述べた線に沿って総合評価法を変えたとしたら、このシミュレーションの結果がどうなるかは、興味のあるところだし、また、もしそうした評価法が学校や専門ごとに入試前に公表されたならば、受験志望者、ひいては合格者の分布はかなり変動するものと思う。

1990年春の国公立大学の入学試験で、新テストの科目数を減らしたり、分離分割テストを増やしたのは(朝日新聞 1989-8-13)、選抜方法の多様化に向けての第一歩として歓迎すべきことと思う。この上は、総合評価にあたっては、将来、本稿に述べたような検討が加えられるとよいであろう。

また、今までは、ほとんど文部省の取り仕切っていた高等教育政策が、通産省などにも開かれつつあるのも(朝日新聞 1989-8-12(b))、多様化のためには喜ぶべき傾向であろう。

ちなみに、いま学術会議では大学院の充実に関して審議がなされているが、その中で、外国人の博士号取得と関連した、教科の多様化も検討されている。しかしそこでは、ある種の多様化を実施する

と、博士号のレベルが落ちるのではないかという心配が表明されているそうである。(学術会議 1989)。

このレベル低下ということは、たとえば小学校のカリキュラムや教科書を多様化したときに、文部省あたりをはじめとして、多くのひとびとの心配するところでもあろう。

簡単に私見を述べさせてもらおうと、レベルという発想は、いまだに一つの数値に頼る価値観、すなわち、単一化、画一化の概念ではなからうか。多様化し、ベクトル化した能力に満ちた社会においては、もはやレベルという概念は、ほとんど定義不可能となり、そこでは、個性がいかに伸ばされているか、個人がいかに社会に貢献できるか、といった、多価値化した評価が求められてくるのではなからうか。

パソコンを導入し、生徒の個人情報の管理化を図ることについての是非の論争なども(読売新聞 1989-5-11)、既成の価値観を離れ、多様化した価値観と積極的に取り組むという観点からの評価に持って行けば、複雑化する評価法は、パソコンの助けなしには実行不可能という結論が出るのではなからうか。

9. 創造性と研究の管理行政

以上では、これからの日本に求められている、創造性の開発と文化形態一般、中でも幼時からの教育の多様化が、かなり密接に関係しているのではないかとすることを考察してみた。

そうした議論とは別に、ここ 10 数年、研究が盛んになってきた動的 (dynamical) システムや相転移の理論は、本稿に述べてきたような多様化の努力が、社会一般において必然的にもたらず変化に理論的根拠を与えてくれるのではないかと期待される。

詳しくは最近のある国際会議の論文集(東大シンポジウム 1989)にゆずるとして、たとえばフリーマン (Walter J. Freeman) 教授は、生物系が無構造活動状態 (state of unstructured activities) にあるときには、外からの刺激に対する反応が速くなるという事実の理論付けを試みているし、イエーツ (F. Eugene Yates) 教授は、「あらかじめ決められた道筋によって管理された (program-driven)」システムよりも、「作業の実施結果によって制御を決めていく (execution-driven)」システムのほうが能力が高いことを示している。村上陽一郎教授は、研究手法が既知の体系によって整理され、透明になり過ぎること (trivialization, フリーマン教授は classicalization, 古典化とも呼んでいる) は研究の活性をそぐことになるかと推論している。

当然のことといえばそれまでであるが、生理学的にも、ある機能をくり返し使った場合には、脳の中でその機能に対応する部分が分量的により発達していくことが知られている。逆にいうと、機能の均質化によって疎外された部分は発達しないから、それを必要とする機会が巡ってきたときに、十分に活用することができない。同時にこれは、社会が均質化するということが、集団としての盲点を作りあげてしまうということにつらなっている。

同会議におけるカーナー (Edgar Koerner) 教授の発表によるまでもなく、創造性とは心の内外における不整合 (inconsistency) によって発動されるものであるから、まわりに自分と似通った者しか

いない均質的社会においては、創造性のひらめきの機会も減ってしまうのは明らかである。鈴木増雄教授も、創造的活動が無秩序状態から秩序状態への相転移としてとらえられると考え、未解決で混沌とした現状を知り、不安定な状態に自分を置くことは、創作活動の出発点であると論じている。

一方、古川俊之教授は、主として保健医学の観点から、均質性の高い日本の社会が、いかに国民一般の健康の質を高めているかを、各種の統計資料を駆使して明らかにしている。

このように、均質性のもたらす高水準と、混沌のもたらす創造性とは、本質的に相反するところがあるから、高水準と創造性の二兎を追うことは、至難のわざである。いかにすれば均質性社会の利点の喪失を最小限に抑えつつ、多様化社会の利点を手中のものとしてできるような教育体系を編みだすかは、むしのいい話かもしれないが、今後、われわれが真剣に取り組まなければならない重要課題であろう。それなしには、人口が日本の2倍でしかないアメリカが、創造性の指標の一つである科学分野のノーベル賞を30倍ほども受賞しているという、彼我の格差を縮めていくことはできないであろう。

斎藤栄三郎科学技術庁長官は、日本の受賞者が少ないのは、足の引っ張り合いのせいであるという説を支持している(朝日新聞 1989-9-12)。本稿に述べてきたように、均質化の圧力は、下部を見ると引き上げる力となって働くが、上部をとれば、逆に引き下げる力となって働くことになるから、斎藤の説は本稿に述べてきた分析の一端と、現象としてはよく整合している。

さて次に、創造性の強化のために、教育と並んで重要なのが、研究の管理、特に研究費配分にかかわる行政であろう。

今日のように、科学の研究が広域化、巨大化し、進歩が高速化すると、研究費も膨大なものになり、要求の全てを満足させるわけにはいかない。したがって、国からの研究助成金への要求は、何らかの方針に基づいて取捨選択をすることになる。

アメリカにおける基礎研究の助成機関の最たる全米科学財団(NSF)の政策についても、たえず問題になるのは、微小(個人的)科学と巨大科学とのバランス、科学と工学とのバランスである。このどちらもが、それぞれほぼ基礎と応用のあいだの区分けとも考えられる。

6年の任期を持つNSF長官に、1984年に就任したブロック(Erich Bloch)は、大まかに言ってこのそれぞれの後者、すなわち巨大科学、工学、応用に関わる研究を集めた科学技術センターを数か所新設して、目的意識の強い研究を推進しようとしてきた。しかし、これらセンターのほかに、既設の工学センター、スーパーコンピュータ・センターを入れても、まだその経費はNSFの研究総予算のわずか8%でしかないにもかかわらず、すでに全米科学アカデミー(NAS)は、研究分野に対して優先度を持たせた、こうしたNSFの研究費配分に公式に反対を表明している。また生物系の研究者たちも、かれらの研究費の主たる支出者である国立衛生院(NIH)に対して、同様の決議を提出している。

こうした、政府当局による研究の階層化に対し、ダットンとクロウの両教授は、むしろ研究者自身による、そうした研究優先度付けの具体的な階層化手続きを提案している(Dutton and Crowe 1988)。しかしそれに対しても、一部の研究者からは、やはり反論がでている(Stout and Brankenridge 1988)。

こうした、基礎研究活動のプロジェクト化に反対する思想の中心には、基礎研究がどのような進路を

とるかは実行中に決まることが多く、あらかじめ研究計画を立てられないところに基礎研究の価値があるということがある。したがって、研究に目標を持たせるプロジェクト化が国家的規模で行なわれるならば、基礎研究への研究費の支出が正当化しにくくなり、ひいては枯渇することさえも恐れられてのことである。

一般に、研究提案の一つ一つが個別に審査されるときには、小さな研究テーマの提案に比べて、大型プロジェクトはほとんどいつも有利な立場に立つ。たとえば、麻薬中毒者が家族に与える影響の解明、純粋数学の問題、考古学的発掘などの小さなテーマは、人間の全ゲノムの解明といった大型プロジェクトが約束できるような長期的な利益は、まず約束できない。

しかし、こうした小さなテーマの研究も、それなりの科学的価値があると思われるものは、やはりいつかは実行すべきものである。したがって、国のレベルでの研究政策に必要なことは、個々の研究の現在における長期的利益の見通しの大小によって、いまどの研究を支援するかを決めることではなく、どの研究を、時間的にどういう順序で実行すれば、研究の長期的利益が最大となるかといった、動的問題を解くことだ、というのがスタウト教授の指摘するところである。

そして、動的决定問題の理論の示すところによれば、結果の予測にかなりの不確定要素があるときには、現時点で最も有利と見られるテーマに全力を集中しないことが、長期展望において最大の投資効果をもたらすとされている。

これは次のように考えれば、直観的に納得できる。

第1に、長期的に見れば、科学の進歩は「科学資本」、つまり、教育、設備、試験研究、専門知識の集積などの総量が社会全体としてどれだけ恒常的に維持されているかに大いに関わっている。

したがって、いかに短期的成果や社会的利益につながるといっても、大型プロジェクトに予算を集中することが、このような科学資本の減少を伴ったのでは、長期的には最大利益に反することになってしまう。

だから、国家的規模における長期的科学政策の策定に当たっては、大型プロジェクトの短期的魅力に眩惑されることなく、科学的資本の充実に努力をしなければならないことを、スタウト教授は主張している。アメリカにおける巨大研究プロジェクトのはなばなしさに眩惑され、基礎研究の必要とする広い麓野の充実をとくなくおざりにしたまま、ともすると巨大プロジェクトの分野の振興に走りがちで、わが国一般の施策の責任者にとっても、こうした警告は大いに参考になるであろう。

20年ほど前、ジョンソン大統領の時代にNSFの長官代理を勤めていたウィルソン(John T. Wilson)は、「われわれの最良の大学は、いくらかは管理された、混沌状態にある。そしてそれが、知的興奮を引き起こすには最も良いところなのだ。」と述べている(1967年12月10日)。上に述べた動的系统や相転移の理論の指摘と考え合わせると、なかなか面白い発言である。

かつてはかくも自由を標榜していたNSFは、そしてアメリカの社会は、いま少しく変わりつつあるかに見える。こうした彼我の状況を検討してみると、国民の均質性に助けられ、数多くの、目的意識の強い開発プロジェクトに支えられて成功してきた日本が技術立国、工業立国として挙げた成果を見

て、アメリカの政府や研究関係者が強く刺激され、いま日本の後追いをしようとしているかにさえ感じさせる。その日本が、いま遂に、創造性の向上と基礎研究力の充実とに取り組まざるを得なくなり、その実現に向けて、個人のアイデアに基づく自由な研究の奨励を必要としているのは、歴史の皮肉であらうか。

ともあれ、創造性は決して安易に手にはいる資質ではない。そのためには、いままでわれわれが享受してきた均質的社会と、それに付随した、国民の高い平均的水準さえも、ある程度は代価として払うことを覚悟しておかなければなるまいというのが、本稿におけるわたしの考えの帰結である。

若い世代のあいだには、この新旧相克の兆しが、もうしばらく前から見えていていると感じるのは、わたしだけなのだろうか。

10. おわりに

本稿は、1989年9月14日に横浜で開催された、International Symposium on New Technologies : R & D Cooperation and the Development of New Technologies (新技術国際シンポジウム：研究開発の国際協調と新技術開発を考える)において、「日米の研究開発体制を考える」の題で講演したときに準備した参考資料に手を入れたものである。

この講演が実現したのは、横浜国立大学工学部の仁木克己教授のご甚力による。本センターの大野公男教授は原稿を精読され、数かずのご意見をたまわった。たびたび手を入れた草稿を、そのつどタイプして下さったのは堤妙子さんである。これらの方がたに厚く感謝したい。

(1990年3月)

参考文献

- ・朝日新聞、学歴・収入高い人ほど日本の取引に「不公正」、1989年4月12日。
- ・朝日新聞、外国と触れ合う「場」を：国際人育成を考えるシンポジウム、1989年5月10日。
- ・朝日新聞、制服の衣替えはなぜ一斉に6月1日なのか、1989年5月18日。
- ・朝日新聞、広がる日米ハイテク摩擦、強まる日本脅威論、1989年5月19日。
- ・朝日新聞、学校の指定業者はなぜ必要なのか、1989年6月22日。
- ・朝日新聞、「鉄鋼業界、商習慣は差別的」——米、日本に撤廃要求文書、1989年8月11日。
- ・朝日新聞、鉄鋼業界、強く反発、1989年8月12日(a)。
- ・朝日新聞、援助のプロ育成、通産も、1989年8月12日(b)。
- ・朝日新聞、来春の国公立大入試、総定員の42%「分離分割」に、新テスト課す教科減らす傾向も、1989年8月13日。
- ・朝日新聞、共通一次、配点変えても同じ人が合格？——電算機で模擬実験、夕刊、1989年8月16日。
- ・朝日新聞、「日本の受賞者少ないのは足の引っ張り合いのせい——ノーベル賞で科技庁長官が一説」、夕刊、1989年9月12日。

- ・ Benedict, R., *The Chrysanthemum and the Sword—Pattern of Japanese Culture*, Boston: Houghton Mifflin, Co., 1946.
- ・ Dutton, J. A., and L. Crowe, *Setting priorities among scientific initiatives*, *American Scientist*, 76: 599-603, 1988.
- ・ Fallows, J., *Containing Jappan*, *The Atlantic Monthly*, pp. 40-54, May 1989.
- ・ 学術会議、第4常置委員会、シンポジウム「大学院の充実と今後の展望」、東京、1989年5月8日。
- ・ 石田剛、際立つ日本の政治力の貧困、朝日新聞論壇、1989年8月9日。
- ・ 木村広男、もっと緩やかな検定制度を、朝日新聞論壇、1989年8月1日。
- ・ 国際理解教育を考える会、「国際人育成を考えるシンポジウム」、東京、1989年5月6日。
- ・ Levine, J. B., *Is the U.S. selling its high-tech soul to Japan?*, *Business Week*, June 26, pp. 97-98, 1989.
- ・ 松井茂、兵器輸入で「安い防衛」を：飛び抜けて高い国産品の調達価格、朝日新聞論壇、1989年8月10日。
- ・ Mervis, J., *Erich Bloch's campain to transform NSF*, *American Scientist*, 76: 557-561, 1988.
- ・ 斎藤進六、日本工学アカデミー第3回通常総会、材料専門部会報告、東京、1989年5月15日。
- ・ Stout, R. L., and G. R. Brankenridge, *Evaluating scientific initiative*, *American Scientist*, 77: 211-213, 1989.
- ・ 田中洋一、アイヌを独自の民族と認め伝統文化の継承に保障を——日本民族学会が五項目見解、朝日新聞、夕刊、1989年6月15日。
- ・ 東大シンポジウム、生命システムの時間的发展——生物的複雑性と情報、東京、1989年8月23-26日。
- ・ Ward, F., *Images for the computer age*, *National Geographic*, 175: 718-751, 1989.
- ・ 手塚晃、言語・思考の枠組としての文字システムの評価、日本語学、6(8): 88-103、1987。
- ・ 山田尚勇、文字体系と思考形態、日本語学、6(8): 43-64、1987。
- ・ 山田尚勇、日米の研究開発体制を考える、学術情報センター紀要、No. 3、pp. 139-160、1990年9月。
- ・ 読売新聞、「髪赤い」と黒染め強制——岡山の女子高“生まれつき”生徒10人に、1989年5月10日。
- ・ 読売新聞、生徒の個人情報パソコン管理!?!——導入めぐり是非論争、都民版、1989年5月11日。