学術情報センター電子図書館サービス

© 学術情報センター

ISSN 0913-5022

学術情報センター紀要

Research Bulletin of the National Center for Science Information Systems

第 9 号

1997年3月

文部省 学術情報センター

学術情報センター電子図書館サービス　　©学術情報センター

学術情報センター紀要第９号

目　　次

巻頭言　　近くて遠い国々との交流　　松村多美子(図書館情報大学　図書館情報学部教授)

10周年記念講演

21世紀の政治・経済システム
佐和　隆光 (京都大学経済研究所長) ····· 1

わが国のマルチメディアネットワークの将来
青木　利晴 (日本電信電話株式会社常務取締役・研究開発本部長) ····· 9

研究論文

学術情報分野

情報化社会の国際化と日本語
山田　尚勇 (中京大学情報科学部) ····· 33

CJK (中国語/日本語/韓国語) 環境下での国際標準件名標目システムの適用
相原　信也 (国立国会図書館)、内藤　衛亮 (学術情報センター) ····· 73

総合目録オンラインDBと情報検索システムの連携方式
大山　敬三 (学術情報センター)、髙須　淳宏 (学術情報センター)、鶴岡　弘 (学術情報センター) ····· 83

物理的な空間としての研究室やオフィスにおける情報組織と視覚的な手がかりによる探索
越塚　美加 (学術情報センター) ····· 91

多次元データに対する項目反応モデル
孫　媛 (学術情報センター) ····· 103

ヤクート語語彙研究 (2)：方向表現---サハ族英雄叙事詩オロンホを資料として---
藤代　節 (神戸市看護大学) ····· 113

システム分野

簡易幾何モデルを用いたコンピュータ合成動画像の圧縮
後藤田　洋伸 (学術情報センター)、小野　欽司 (学術情報センター) ····· 131

AODBMS: Phasmeとその形状検索処理機能を持つ画像検索プラグイン・モジュール
フレデリック　アンドレス (学術情報センター)、クリストファー　クローデ (東京大学)、
小野　欽司 (学術情報センター) ····· 137

単語間の係受け情報を用いた文献検索手法
池田　和幸 (東京大学大学院工学系研究科)、髙須　淳宏 (学術情報センター)、
安達　淳 (学術情報センター) ····· 143

認識誤りを含むテキストにおける検索手法
太田　学 (東京大学大学院工学系研究科)、髙須　淳宏 (学術情報センター)、安達　淳 (学術情報センター) 161

学術情報センター電子図書館サービス　　© 学術情報センター

ATM ネットワークにおける自動再送方式に関する研究
趙　偉平 (学術情報センター)、ルンキーラティクン　スワン (学術情報センター)、
淺野　正一郎 (学術情報センター) ····· 173

ATM 網におけるセル廃棄特性解析
ルンキーラティクン　スワン (学術情報センター)、　趙　偉平 (学術情報センター)、
淺野　正一郎 (学術情報センター) ····· 179

シェーピングによるトラヒック特性への影響
計　宇生 (学術情報センター) ····· 189

動画像情報サービスのためのオープンプラットフォーム技術について
相澤　彰子 (学術情報センター) ····· 195

DBMS におけるクエリーコスト評価のためのニューラルネットアプローチ
ジハド　ブロス (学術情報センター)、ヤン　ビエモン (ベルサイユ大学 (フランス))、
小野　欽司 (学術情報センター) ····· 203

インターネットにおける実時間通信の QOS 保証
森野　博章 (東京大学工学系研究科)、相澤　彰子 (学術情報センター) ····· 219

調査研究分野

Wellcome 財団の助成活動と研究評価を支援するデータベースの役割
山崎　茂明 (東京慈恵会医科大学医学情報センター) ····· 227

Medline データベースからみた臨床試験文献の分析：コクラン共同計画との連携
山崎　茂明 (東京慈恵会医科大学医学情報センター) ····· 233

学術情報センター電子図書館サービス © 学術情報センター

Research Bulletin
of
The National Center for Science Information Systems
Volume 9, March 1997

Table of Contents

T. Matsumura — Preface

Lecture Notes at NACSIS 10th Anniversary

T.Sawa 1 Political and Economic System in the Twenty-first Century

T.Aoki 9 The Future of Multimedia Network

Contributions

Scholarly and Science Information Research Area

H. Yamada 33 Globalization of Information Society and Japanese Language

N. Aihara, E. Naito 73 Adaptation of Universal Standardized Subject Heading System in the CJK (Chinese / Japanese / Korean) Environment

K. Oyama, A. Takasu, H. Tsuruoka 83 Configuration of an IR System linked to a Union Catalog On-line Database

M. Koshizuka 91 Information Seeking in Terms of the Visual Data in an Office

Y. Sun 103 A study on IRT models for multidimensional data

S. Fujishiro 113 A study on lexcal items in Yakut language (2) : Terms of direction in Saxa epic poem *olonxo*

System Research Area

H. Gotoda, K. Ono 131 Compression of Synthetic Movies Using Simplified Geometric Models

F. Andres, C. Claude, K. Ono 137 AODBMS: Phasme and its Image Retrieval Plug-ins Implementing shape-based retrieval processing

K. Ikeda, A. Takasu, J. Adachi 143 A Method of Document Retrieval using Dependency Relationships between Words

M. Ohta, A. Takasu, J. Adachi 161 Retrieval Methods for Text with Missrecognized Characters

W. Zhao, S. Runggeratigul, S. Asano 173 A Study on ARQ Mechanism in ATM Networks

学術情報センター電子図書館サービス
© 学術情報センター

S. Runggeratigul W. Zhao S. Asano	179	Analysis of Cell Loss Performance in ATM Networks
Y. Ji	189	The Characteristics of ATM Traffic Constrained by Different Shaping Mechanisms
A. Aizawa	195	Constructing Open Network Platform for Video Services
J. Boulos Y. Viemont K. Ono	203	Neural Networks Approach for Query Cost Evaluation of DBMS
H. Morino A. Aizawa	219	QOS guarantee for real-time applications in the Internet

Research Trend Research Area

S. Yamazaki	227	Activities of research funding and the role of research databases for supporting the evaluation of scientific performance in the Wellcome Trust
S. Yamazaki	233	Bibliometric analysis of the papers on clinical trials and randomized controlled trials in the Medline database, with reference to the Cochrane Collaboration

学術情報センター電子図書館サービス　　©学術情報センター

学術情報センター紀要 第9号 (1997)

巻頭言

近くて遠い国々との交流

図書館情報大学　図書館情報学部教授　松村多美子

近年わが国でもアジアに対する関心が高まりつつあるようだが、平成6年度から科学研究費で行っているユネスコとの共同研究でベトナム、バングラデシュ、モンゴル、タイ、ネパールなどを含むアジア諸国およびニュージーランド、フィジー、パプアニューギニアなど太平洋諸国の情報専門家と共同研究を行い、またこれらの国の現地調査を実施する機会があった。ユネスコはその総合情報計画(General Information Programme)という部門で図書館、文書・公文書館を含めて包括的な情報活動を行っているが、そのなかで特にアジア・太平洋地域については、ASTINFO (Regional Network for the Exchange of Information and Experiences in Science and Technology in Asia and the Pacific)と称する地域プログラムがあり、ドキュメント デリバリ サービスの実施、薬用植物データベースの共同構築、各種研修コース、セミナーの開催など地域レベルの様々な活動を行っている。しかしこのような活動をより強化し各国の情報基盤整備を促進するとともに、地域としての情報基盤を確立しGIIへの対応をはかる必要性はますます高まっている。

わが国はかねてからASTINFOの情報活動を支援してきたが、このような背景から平成6年～8年にわたりアジア・太平洋地域における情報基盤構造に関する調査研究を実施した。これには地域内11か国ーモンゴル、ベトナム、タイ、バングラデシュ、インド、フイリピン、インドネシア、フィジー、ニュージーランド、ネパール、パプアニューギニアが参加し、これにユネスコ地域専門家およびわが国の研究者を含めて国際研究チームを編成し、各国における現地調査の実施、national situational report の作成や3回のワークショップの開催、さらに最終的なRegional Assessment Report の作成を行った。

アジアと一口に言ってもそこに位置する国々は言語、宗教、社会制度から国民性に至るまで実に多様である。むろん情報基盤の現状も例外ではない。これはまた情報基盤の整備に対するアプローチの違いにも見られる。3年間の共同研究を通して各国の情報に関連する法制度、情報資源やシステム・サービス等に関する知見をある程度は得ることができ、また現地調査で社会・文化をほんの束の間ではあるが触れることが出来たが、交流というには程遠くやっと一歩を踏み出したにすぎない。

国際的ネットワークの急速な普及により時間空間に制限なく情報が流れ、その意味で世界が狭くなりつつあることは事実である。しかしながらその反面、近い国々との距離はむしろ遠くなっているのではないだろうか。

むろんこれまでにも経済援助の面で、そして学術においても研究者間や大学・研究機関の間で交流が行われてきている。しかし今後は情報の分野でこれらアジア・太平洋地域の国々との交流を、もっと活発に進めていく必要があると考える。

学術情報センター電子図書館サービス　　◎学術情報センター

学術情報センター紀要 第9号 (1997)

創立10周年記念講演

演題 「21世紀の政治・経済システム」

Political and Economic System in the Twenty-first Century

京都大学経済研究所長　佐和　隆光
Takamitsu SAWA
Director, Institute of Economic Research, Kyoto University

(この講演は、平成8年11月1日(金)、国立オリンピック記念青少年総合センターにて、行われたものである。)

平成不況という呼び名がほぼ定着したかと思いますが、この不況はいつ始まっていつ終わったのかというと、1991年5月に始まり1993年10月に終わったと、これは政府が勝手に決めていると言っていいかと思うのですが、そのように政府が決めているわけです。ということは、この平成不況は30か月続いたわけです。一つの不況が30か月も続くことは非常にめずらしいことです。通常の不況は大体12か月で底入れするのが普通なのです。

オイルショックのあとの不況が、やや長くて16か月でした。しかし、今回の不況はオイルショックの不況のちょうど倍で、30か月という異常に長い不況だと言って差し支えないかと思います。

この不況の原因はさておくとして、私はこの平成不況を戦後日本経済の第3の転換点であると位置づけております。

それではいったい第1の転換点はいつのことかというと、1957(昭和32)年7月から1958(同33)年6月にかけての鍋底不況です。これが第1の転換点です。

この転換点は何をどう転換させたのかというと、戦後復興期にピリオドを打って高度成長期の幕を切って落とした。そういう転換点だったわけです。ちなみに昭和33年というのはどういう年だったのかを思い起こしていただきたいのですが、一つはミッチーブームに沸いた年でありました。つまり正田美智子さんが皇太子妃に決まって、ミッチーブームに沸いた年です。それから2つ目に長嶋茂雄が巨人軍に入団したのが、ほかでもない昭和33年です。そして、1万円札が初めて発行されたのもこの年なのです。当時の大学卒の初任給は、1万2千百円余りでしたから、もし1万円札を使って給料を払うとなると、お札が3枚になるということで大騒ぎをしたりしました。それから、ボディビルが流行したのもこの年です。

今、4つのことを申し上げましたが、この4つのことのいずれもが高度成長期の幕開けを飾るのにふさわしい出来事なり、人物だったのではないでしょうか。

さて、第2の転換点はいつのことかと申しますと、1973(昭和48)年です。48年12月に始まったオイルショック不況です。そしてこの不況は16か月間続いて、昭和50(1975)年3月に底入れいたしました。16か月間続いたこの不況が第2の転換点です。

この不況は何をどう転換させたのかというと、高度成長期に終止符(ピリオド)を打った。そのあといったいどうなったのかというと、私はそのあとの時代のことを昭和50年以降の日本経済を減速経済期と呼ぶことにしております。もう少し具体的に言いますと、4%成長の時代に入ったわけです。

高度成長期というのは、昭和33年に始まって48年に終わったわけですが、ちょうど15年間です。この高度成長期の経済成長率は、平均年率で9.2%でした。そして、そのあとオイルショックを経て後は、再び9～10%成長の時代に戻ったのではなく、4%成長の時代に入ったというわけです。

その後4%成長の時代がずっと続いたのですが、私はこの4%成長の時代も、1990(平成2)年に終わったと見ております。したがって、この減速経済期というのも1975(昭和50)年に始まって、終わったのが平成2(1990)年ですから、高度成長期と同じく15年間続いたわけです。

この15年間の経済成長率は、平均年率で4.2%だったわけです。つまり高度成長期が9.2%だったのが、オイルショックを経て後の減速経済期には4.2%ということで、巡航速度が半分以下に下がったということです。

この15年間というのは、確かに成長率は半分以下に下がったわけですが、欧米諸国もみんな先進諸国はい

学術情報センター電子図書館サービス　　©学術情報センター

ずれもそのような減速経済期に入ったのです。相対的には、日本は高い成長率が続いたわけです。

そして、1987(昭和 62)年に、実は日本は一人当たりGNPというものさしで測って、とうとうアメリカを追い抜いたのです。戦後、考えてみますと50年近くにわたって、「追いつけ追い越せ」ということをモットーに掲げ、とにかく我々日本という国は、あるいは日本の企業は一目散に走り続けてきたわけです。その追いつけ追い越せの願いが、1987年にようやくかなえられたわけです。つまり42年間かかってみごとアメリカを追い抜いたというわけです。

では、この年日本は一人当たりGNPが世界一になったのかというと、実はそうではなくて1位はスイスでした。そして2番目がルクセンブルク、3番がアイスランド、4番が日本だったわけです。

スイスの人口は六百数十万人ですが、ルクセンブルクは35万人ぐらいです。そしてアイスランドは24万人ぐらいですから、国といっても大変小さな国です。ですから、一億二千数百万の日本が一人当たりGNPが世界で第4位になったということは、これは大変なことだったわけです。

それはさておき、この減速経済の時代も90年(平成2年)に終わったのだとすると、そのあといったいどうなったのかと。先程申し上げましたとおり、1991年に入るとにわかに景気が悪化して平成不況に陥ったわけです。そして30か月、93年の10月というと今からちょうど3年前ですが、3年前にトンネルを出たと。それにしては相変わらずもう一つパッとしないのではないかと皆さん方は思われると思います。あるいは、まだ不況が続いているとてっきり思っていたと言われる方も少なからずおられると思います。「どうもトンネルは抜け出たのだけれど、元に戻ったわけでは決してない」と。どうも経済構造のようなものが変わったのだと理解した方がわかりやすいと思います。

いったい減速経済期が終わってどうなったのかというと、私は今現在の状況を、日本はとうとう成熟化段階に入った、あるいは成熟化経済の時代を迎えたと言うことにしています。

成熟ということばを使いますと、おそらくここにおられる皆様方の半数は、「えらく暗いことを言うな」と言われると思います。あまりイメージがよくないわけです。つまり、「成熟すれば、あとは衰退して死ぬしかないのではないか」と言われるかもしれませんが、私はここで成熟化と言うのは、次のような意味においてのことです。

戦後50年間、とにかく工業化社会の階段を必死になって、息せき切って駆け上ってきたわけです。そして、階段の踊り場に到達いたしました。その階段の踊り場が、ほかでもない成熟化段階ということなのです。つまり工業化社会としては成熟化を遂げたという意味なのです。

そして、階段の踊り場は何のためにあるのかというと、これはしばしの足休めのためにあるわけです。踊り場の向こうには、もう一つの階段があるわけです。そして、この踊り場を経て後、新しい階段を上りはじめるわけですが、新しい階段とは何なのかと。それは一言で申し上げればポスト工業化社会なのですが、「工業化社会が終わって、次にやってくる社会はポスト工業化社会です」というのでは、これは同義反復も甚だしいわけです。

「では、ポスト工業化社会をもう少し具体的に内容を言え」と言われますと、もちろん言いだせばきりがないのですが、一言で言えば「高度情報化社会」と言って、あながち見当はずれではないと思っています。

とにかく新しい社会がまもなく始まろうとしているわけですが、階段の踊り場にいる間に、単にしばしの足休めをしていたのでは、下手をすると21世紀を通じてずっと階段の踊り場にいつづけなければならないと私は思っております。

つまり一言で申しますと、日本型システムというか、この国の社会、あるいは企業経営、教育、政治、行政等々のあり方が一切合切が、工業化社会向きに最適に設計されていると思います。

工業化社会向きに最適である日本型システムというのが、はたしてポスト工業化社会向きなのかと問われますと、私自身は首をかしげざるをえません。ひょっとすると最不適かもしれないとさえ思っております。

さて、話が飛躍するようですが、アメリカという国は1980年代にはもうさんざん悪口を言われていたわけです。「もうアメリカの製造業は衰退した。パックスアメリカーナの時代は終わった」と、「パックスジャポニカの時代がやってくる」などということさえ言われたわけです。にもかかわらず、なぜか過去2~3年間アメリカの評判がいいのです。経済の調子もいい。経済成長率も3%前後の成長率をあげています。それからビック3と言われる自動車会社も何やら元気になったようですし、コンピュータやソフトウエアの分野では、圧倒的な強みを誇っています。

学術情報センター電子図書館サービス　　©学術情報センター

これはいったいどういうことなのか。私はその答えをたずねられたとすれば、次のように答えることにしています。

80年代のアメリカは、ポスト工業化社会に一番乗りするための産みの苦しみを味わっていたと。そして、90年代に入ってみごと一番乗りに成功したと見ております。

それでは、ポスト工業化社会というものをもう少しわかりやすく、具体的に内容を説明すると、私は二つの側面があると思います。

一つの側面は、製造業が高度情報化技術をふんだんに取り入れることによって、ありていに言えば生産性を向上させる、いいものを安く作れるようになること。それがポスト工業化社会の一つの側面です。

今から3年ほど前に、リエンジニアリングということばが大変はやりました。このリエンジニアリングは、まさしく生産プロセス、経営プロセスに情報化技術をふんだんに取り入れて、抜本的な革新を図るということだったのです。

アメリカの自動車会社、あるいはコンピュータ会社、あらゆる製造業、場合によっては第三次産業が、そういったリエンジニアリングを80年代末から90年代初頭にかけて徹底的にやりました。そして、見違えるようによみがえったわけです。これが高度情報化社会の一つの側面です。

ポスト工業化社会においては、製造業は消えてなくなると言われる方がいますが、これはまったくの誤解なのです。むしろ、高度情報化技術を取り入れることによって、生産プロセス、経営プロセスの抜本的な革新をやって、製造業をよみがえらせるというのが、高度情報化社会、ポスト工業化社会の一つの側面なのです。

もう一つの側面は、ソフトウエアあるいは情報関連の産業が経済の中枢部に躍り出るということです。これはまさに今のアメリカを見ていると、まさしくそのとおりだとうなずいていただけると思います。そういう社会にアメリカはすでに転換を遂げました。

では日本もそういう転換を遂げるためには、システムを徹底的に抜本的に改変する必要があるわけです。例えば教育システムを考えてみますと、今日も文部省の方がおられるかもしれませんが、日本の初等、中等教育のあり方、あの画一的な教育は、考えてみれば、あれはまさしく工業化社会向きなのです。最適な教育のあり方だったわけです。だから日本は成功したのです。

ところが、ポスト工業化社会向きかと問われると、おそらく「そのとおりだ」と言われる方は一人としていらっしゃらないと思います。どうもポスト工業化社会向きには、ひょっとすると最悪な教育システムかもしれないと言わざるをえないわけです。

そういったことも徹底的に組み換える、改変する必要があるということをとりあえず一言申し上げておきたいと思います。

さて本日のテーマは、21世紀に向けて、世界そして日本、あるいはアジアの経済や政治はどう変わるかということです。世界全体を考えてもいいのですが、ちょっと広すぎますので、とりあえず近隣の東アジア一帯がどうなるのかについてのお話から始めさせていただきます。

2010年ごろ(今から14～15年先)に、いったい東アジアはどうなっているのかということについて、私に予測せよと言われれば、私は次のように答えることにいたしております。

おそらく2010年ころには、東アジア一帯は中国人、華人経済圏になっているであろうと。華人とは、中国籍を持たない中国人のことです。東南アジア諸国、香港、台湾はむろん東南アジア諸国のいたるところに華人がいるわけです。そして、東アジア一帯に住む華人の数は、5000万と言われています。5000万人の中国人が東南アジア諸国、香港、台湾に住んでいるわけです。

彼らはどこの国に行っても大変経済的に成功しています。そして、今中国のメインランドにどんどん華人資本が流れ込んでいます。華人はいろいろな国にいるわけですが、そういう人たちの間でみごとにネットワークが張りめぐらされていて、そのネットワークが、中国のメインランドにも今広がりつつあるというのが目下の状況であります。

私は、中国人や華人というのは、市場経済のプレーヤーとして実にタフであると思っています。皆様方もご存じの中国人というのを頭に想像していただきたいのですが、今30～40代の中国人は生まれたときから社会主義です。ですから、彼らは資本主義なんて全然わかっていない、市場経済なんて知らないのだと思われると、それは大まちがいだと思います。

彼らは、本能的に市場経済とか資本主義というのは大好きなのです。どういうことかというと、競争、リスク、選択の自由などというものは大好きです。

ところが、それにひきかえ我々日本人はどうなのかというと、競争は極力回避しようといたしますし、リ

学術情報センター電子図書館サービス　　©学術情報センター

スクも回避したがるのです。選択の自由もあまり好きじゃないという人が多いわけです。

そういう意味では、日本は市場経済の国だと建前上そうなっていますが、実は競争やリスクが大嫌いという人が作っている国ですから、市場経済が本当にいきわたっているかどうかというと首をかしげざるをえません。

実際、最近規制緩和と言われていますが、この国には1万1000個の規制があると、だれが勘定したのかは知りませんが、とにかくそれだけあるそうです。そして、それだけの規制があるというのは、別にお上がそういう規制を押しつけたわけではなく、競争とリスクが大嫌いな日本人が自ら好き好んで作り上げたシステムだと言わざるをえないわけです。

日本は、そういう意味ではアジアの中でも非常に特殊な国であるということを、とりあえず強調しておきたいと思います。

さて、それだけの市場経済のプレーヤーとしてのタフネスを持っている中国人、華人が、おそらく2010年ごろにはこの東アジア一帯を自分たちの経済圏にしているであろうということを申し上げました。そのときに日本という国は、中国人、華人経済圏の中に組み込まれているのかと問われますと、私はどうもそうではないのではないかと思うのです。では、どうなっているのかというと、中国人、華人経済圏の外側にぽつんといるという感じだと思います。

そのことを私はこのレジュメの中で、「アジアの香港になっているのではないか」と申し上げました。もう少し別の言い方をすれば、ヨーロッパのスイスのような国になっているのではないかと思います。

そして、おそらく2010年ごろに中国人は次のようなことを言っていると思います。

「あの国(日本)は、今から20年前には経済大国と言われた国だ。そして、93～94年ごろには、円が非常に高くて一人当たりGNPが世界一だった。そして今も結構豊かな国です」。

おそらく私は2010年ごろになりますと、為替レートは円の値打ちは今の半分ぐらいになっていると思います。仮にそれが180円としますと、94年に日本の一人当たりGNPは世界一になったわけですが、その年の平均為替レートが約90円です。1ドル90円でドル換算した一人当たりGNPが世界一だったわけです。それが、もし同じ今のGNPのままで、もちろん平均的に成長していくわけですが、90円が180円になったとすれば、一人当たりのGNPは半分になるわけです。今現在、一人当たりGNPが日本は3万8000ドルですが、これが1万9000ドルになったとすれば、世界で何番目ぐらいかというと、十何位です。ニュージーランド、スペインと肩を並べるぐらいです。しかし、考えてみれば日本の豊かさというのはそんなものかという感じもしないではないわけです。

いずれにせよ、今申し上げているのは、2010年ごろに中国人は次のようなことを言っているだろうということなのです。

「今も結構豊かな国らしいです。きれいで安全で、旅行者に大変親切ないい国です。しかし、国際政治や国際経済に対する影響力はまったくない国ですね」と言われている可能性はきわめて高いわけです。

ですから、この国のしくみを相当思い切って変えないと、今申し上げたようなことになりかねないのです。

私が今申し上げたのは、中国の今後の経済発展というのは、今の調子が今後当分続くということを前提にして申し上げたわけです。逆に大変悲観的なことを言う人もいるわけです。

今から3年ほど前に、世界銀行が「エイシアン・ミラクルズ」、アジアのミラクル(奇跡)というレポートを書きました。そして、アジアの経済発展をほめそやしたわけです。それに対して、ポール・クルーグマンというスタンフォード大学の教授で経済学者ですが、この人が「ザ・ミス・オブ・エイシアン・ミラクル」、アジアの奇跡は幻であるというタイトルの論文を書いて大変話題になったのです。今から2年ほど前のことです。

この中で彼は何を言ったのかというと、東アジアの経済発展は技術進歩なき経済発展であると決めつけたわけです。これは別に中国に対してのみ言っているのではなく、韓国、香港、台湾、シンガポールといういわゆるアジアNIESと呼ばれる国ないし地域、そして日本もその中に入っているわけです。

つまり、生産要素と呼ばれる労働力や資本の投入をどんどん増やすことによって経済は成長してきたのだと。それだけのことであって、技術進歩がないから、どこかでそれは行き止まると。日本についてはややことばを濁しているのですが、その論文の中で、例えば日本にしても、とうとう成長が止まったではないかと書いてあるわけです。

ですから、結局技術進歩なき成長というのは必ず止まると、あるところまでいくと止まってしまうという

学術情報センター電子図書館サービス　　©学術情報センター

わけです。

そういうことを言ったわけですが、このことについてはクルーグマン先生の話の進め方にはいろいろ問題があるわけです。私はそうではなくて、とりあえず日本は別にして、韓国から中国にいたるまでのアジアの経済発展は、技術と資本の移転による、あるいは導入による経済発展であるとみなした方が的確であると思っています。そして、技術と資本をどんどん取り入れることによって経済発展すると。

そして考えてみますと、欧米諸国あるいは日本の技術の水準が高いのであれば、あらためて町工場を造って技術開発をやらなくても、当然よそからちゃんと教科書に書かれた技術をどんどん取り入れればいいに決まっているわけです。ですから、今は技術導入によってどんどん経済が成長していると言うことができると思います。

ここで一言付け加えて申し上げますと、戦後の日本経済は短期間の間に大変な経済発展をなし遂げたわけです。その日本は、はたして今の中国にとってのモデル、模範になりうるのかどうかというと、私は必ずしもそうは思いません。

なぜそうなのかというと、戦後つまり1945～1950年ごろに日本が発展を開始したというのは、非常にタイミングがよかったわけです。あの当時の技術のレベルというのは、今に比べると圧倒的に低かったわけです。つまりキャッチアップ可能な水準にあったわけです。ですから、ソニーであれ、ホンダであれ、あるいは京セラであれ、昭和20年代から30年代前半にかけて、小さな町工場を数人で造って、そして10年後にはまさにハイテク企業に衣替えするという成功例がいくつもあるわけです。

それは日本人が勤勉であったとか、優秀であったとかいうことだけではなく、それ以上に技術の発展のステージが大変キャッチアップしやすいステージにあったということではないかと思います。

今の中国に日本を見習えということで、「さあ皆さん、町工場を造って頑張りましょう」と言っても、ここまで技術が高いと、とてもキャッチアップはできません。ですから、技術をどんどん導入して発展するという戦略は、決して誤っていないと私は思っています。

さて、あと4年と2か月で閉じようとしている20世紀というこの100年間は、どういう世紀だったのかと問われますと、もちろんそれに対する答え方はいろいろな答えがありうるわけですが、イノベーションの世紀であったというのが、一つの的確な答えではないかと思っています。この100年間の間に実に数限りないプロダクトイノベーション、そしてプロセスイノベーションがどんどん行われ、経済がどんどん成長ないしは発展して今日に至っているということになろうかと思います。イノベーションの世紀であるということは、言い換えれば、経済成長の世紀でもあったわけです。

20世紀の初め、つまり2001年の世界のＧＮＰはこれだけでした、今現在それが200倍になりましたと言ってもあまり意味のない比較なのです。それよりも、世界の人口がどれだけ増えたのか、これが一つのものさしになるわけです。

ロバート・マルサスという名前を皆さん方ご存じかと思いますが、マルサスは1798年に人口の原理(俗称、人口論)という著作を著しました。その中で彼は何を言ったのかというと、今後人口は幾何級数的に一定率で成長するというわけです。他方、食糧生産、あるいは食糧供給は算術級数的にしか伸びないから、やがては食糧不足に陥って、そこで食糧の奪い合いから人はお互いに殺し合い、戦争をやって、そこで人口の増加にブレーキがかかるということを言ったわけです。

マルサスが人口論を書いた18世紀末ないし19世紀初めの世界の人口はどのくらいだったのかというと、9億人と推計されています。それでは、19世紀末ないしは20世紀初めの世界の人口はいくらだったのかというと、16億人です。ということは、19世紀の100年間で9億が16億ですから、100年かかって2倍にもならなかったわけです。ところが今現在、世界の人口は58億人です。最近は1年間に9000万人ずつ増えていますから、20世紀末つまり2000年末には、世界の人口は62億ぐらいになっているということはまちがいないと思います。ということは、16億が62億ですから4倍近くに膨れあがったというわけです。

しかし、マルサスが心配したような食糧不足は起きませんでした。なぜ起きなかったのかというと、これもイノベーションのおかげなのです。

それから経済がどんどん成長することによって、地球の人口扶養能力が高まったわけです。マルサスのころには、世界で9億人の人口しか地球上に養うことはできませんでした。19世紀末には16億人、今やそれが62億人になったということは、それだけ扶養する能力が高まったということなのです。何でそんなに扶養する能力が高まったのかというと、先程来繰り返し申し上げていますように、まさにイノベーションのおか

学術情報センター電子図書館サービス　　©学術情報センター

げなのです。

ところで、もう一つ重要な点は次の点なのです。

1980年ごろまでは、世界の経済というのはいったいどうなっていたのかというと、欧米先進諸国そして日本、つまりOECD24か国、いわゆる工業国というのはその24か国と、ソ連とアジアNIES(韓国、台湾、香港、シンガポール)が、工業化を開始したわけです。この3つの地域で工業製品は作られていたわけです。そして、ものというのは放っておくと必ず作りすぎになります。作りすぎたものというのは、どんどん輸出されたわけです。

つまり、今申し上げた3つの地域の人口は、OECD、ソ連、アジアNIESの3つの地域を合わせて14億人です。そうしますと、発展途上国に住む人の数、つまり発展途上地域の人口はいくらかというと、40億ということになるわけです。どんどん工業製品を作り、その域内の需要を全部満たしてしまうと当然作りすぎになるわけです。作りすぎたものは、いくらでも発展途上諸国がアブソーブしてくれたわけです。どのようにしてアブソーブしてくれたのかというと、農林水産物、あるいは地下資源と交換にアブソーブしてくれたのです。だから世界経済は成り立っていたわけです。

ところが、最近は単に東アジアだけではなくてインドなどでもどんどん工業化が始まりました。それから中南米は一足早く工業化を始めました。

ということで、90～95年の時点に立ちますと、アジアと中南米が新たに工業化諸国の仲間入りをしてきたわけです。そうすると残る地域というと、中近東とアフリカぐらいになりますから、何と申しましてもアブソービングパワー、つまりアブソーブする力が明らかに弱まってきました。

このままいって21世紀に入ってどんどん中国でものが作られるようになると、あるいはインドでものがどんどん作られるようになると、明らかに生産過剰になるわけです。当然、資源、環境問題がそれに伴って引き起こされるわけです。

そうなると、どうも21世紀はいささかならず物騒な時代になるのではないかと。つまり、どんどんものを作り、このまま放っておけば何が起こるかというと、2010年ごろにはニューヨークに行こうが、パリ、ロンドンに行こうが、どこへ行っても中国製の自動車が走り回っているという時代に入ると思います。

そうなると、当然ヨーロッパの国々は、ヨーロッパ連合(EU)の垣根を高くします。つまり中国製の自動車が入ってくるのを防ごうとするのです。そういうことで、いわゆる地域のブロック化がより一層強固なものとなる可能性がきわめて高いのです。そして、残されたマーケット(市場)をめぐって、大変な争いが生じるということです。世界はいささかならず物騒な時代になると思います。

それからもう一つ申し上げたいのは、今日ここにおられる皆様方と大変関連の深いことかと思うのですが、私の思うところ、今現在プロダクトイノベーションが最も華やかな分野というのは何なのかというと、それは申すまでもなく情報通信の分野です。しかし私はいかなる技術であれ、いくところまでいけば必ず成熟化するというか、飽和状態に達すると思うのです。

情報通信の技術だって、21世紀を通じてどんどん新製品を登場させるということはないと思います。過去のいろいろな技術の例を参照しますと、大体2015年ないし2020年ごろに、情報通信の分野の技術開発が飽和状態に達すると言わざるをえないのです。では、その次の分野は何なのかというと、なかなかそれが思い当たらないわけです。

情報通信が、もしかしたらプロダクトイノベーションの最後の泉かもしれないと、私は思っています。

もちろんバイオのような分野もありますが、バイオの技術が大量生産、大量消費につながるような製品を作り出して、それが日本経済あるいは世界経済の牽引車になるということは期待できないとすると、情報通信の次の分野はいったい何なのか。

もし次の分野がないとすれば、つまり仮に情報通信の分野がプロダクトイノベーションの最後の泉だったとすると、2015年か2020年にそれが飽和するとすれば、21世紀は19世紀と同じく停滞の世紀になるのではないかと言わざるをえないわけです。

そうならないためにはどうすればいいのかということで、レジュメの[7]というところに処方箋が書いてあります。

まず一つは、やはり需要が足りないわけです。先程、生産過剰になるということを申し上げましたが、それを防ぐためには地球規模でのケインズ政策を行う必要があります。

例えば、この国に1兆円の公共投資をしても、それによって誘発される内需、消費は本当に微々たるものなのです。実際に公共事業の現場で働く人は、徹夜で働いて収入が2倍になったとしても、もしその労働者が外国人労働者だったとすれば、1円も消費を増やし

学術情報センター電子図書館サービス　　©学術情報センター

ません。増えた分は全部貯金しておくのです。当然です。もし、その労働者が日本人であったとしても、家に帰ればテレビもビデオも自動車もエアコンもあるということになると、収入が増えたからといって慌てて電気屋さんの前に飛んでいくなどということはめったにありません。そういうことで、国内でいくら公共事業をやっても、それによってワッと消費が誘発されることは望めないわけです。

ところが、同じ1兆円を中国に投資するとすれば、中国の内需はおそらく1兆円の3倍分ぐらいは喚起されることはまちがいないと思います。

今申し上げたようなことを乗数効果というわけですが、日本に比べてまだ中国の乗数は非常に高いのです。ですから、途上諸国へ公共投資を行うことによって内需を喚起して、オーバーサプライあるいは生産過剰を抑止しようではないかということです。

2番目が、先進国と途上諸国の分業による棲み分けを図ることです。例えば、製造業の生産拠点を中国、タイ、マレーシア等に移転して、日本は情報ソフト分野に産業構造の重点を移行させるのです。どういう棲み分けが望ましいかはさておくことにして、何らかのかたちを模索する必要があります。

3番目が、イノベーションの火種を絶やさないよう国際的な戦略を練ることです。つまり情報通信が、これがプロダクトイノベーションの最後の泉であっては困るわけですから、その次のイノベーション分野というものを探索していく必要があるということです。

4番目がメタボリズム、循環代謝型文明社会の構築です。これはまさに資源環境問題に対する一つの対応です。大量生産、大量消費、大量廃棄ということでやってきたわけですが、そういう文明を21世紀もずっと持ち越すならば何が起こるかというと、資源の枯渇、地球環境の汚染が免れえないわけです。

そういうことを避けるために、循環代謝型の文明を構築しようではないかということです。そして循環代謝型文明を我々日本人の手で作り出して、構想して、それを率先垂範するということが、おそらく何にもましてすばらしい国際貢献になるのではないかと私は思っています。

第5番目は、化石燃料の消費を抑制するための国際的な枠組みを構築しようではないかということです。これからの国際協力が一つの主たるテーマ、メインイシューは地球環境の保全、あるいは化石燃料の消費抑制ということになります。そのために日本は、国際的な枠組みを構築するためにもっと努力すべきだということです。

最後6番目は、究極の省エネルギー社会へ向けて、画期的なプロダクトイノベーションとプロセスイノベーションに挑戦することです。ひょっとするとエネルギーや環境といった分野が、この情報通信の次にやってくるプロダクトイノベーションの分野かもしれません。そういうことを期待して、この分野でのイノベーションに挑戦しようではありませんか。

以上、ここで21世紀のケインズ問題とマルサス問題への処方箋を6つ順不同で述べさせていただきました。

何かこういった思い切った対応なり何なりがなされないかぎり、21世紀はこのままいくと停滞の世紀とならざるをえないのではないでしょうか。その結果、工業製品の市場の奪い合い、天然資源の奪い合いといった悲惨な事態が招来されかねないという、警告のようなことを申し上げて私の話を終わりとさせていただきます。どうもご静聴ありがとうございました。

レジュメ「21世紀の政治・経済システム」

[1]　平成不況は戦後日本経済「第三の転換点」

「第一の転換点」は、1957年7月から58年6月にかけての「鍋底不況」：戦後復興期に終止符を打ち、高度成長期の幕を切って落とした。「第二の転換点」は、73年12月から75年3月にかけての「オイルショック不況」：高度成長期に終止符を打ち、減速経済期(4%成長の時代)へと移行。そして「第三の転換点」は91年5月から93年10月にかけての「平成不況」：減速経済期に終止符を打ち、成熟化段階(工業化社会のラスト・ステージ)へと移行。成熟化段階は階段の「踊り場」に例えられる。工業化社会の階段を息せき切って駆け上ってきた私たちは、とうとう階段の踊り場に到達したのだ。踊り場の向こうにあるもう一つの階段はポスト工業化社会にほかならない。成熟化階段を、ポスト工業化社会への「進化」に備えて、日本型制度・慣行の諸々を改編(トランスフォーム)するための時機と理解すべきである。成熟化段階における経済成長率は平均年率2%台前半、80年代のアメリカはポスト工業化社会(高度情報化社会)への一番乗りをするための「産みの苦しみ」を味わっていたのではなかったか？これからの日本は、80年代のアメリカと同じ苦しみを味わわねばなるまい。

[2]　20世紀の最後の10年に入って以降の予期せぬ

学術情報センター電子図書館サービス　　©学術情報センター

「激変」の頻発

①自民党一党支配体制の崩壊、②平成不況、③日本型制度・慣行への評価の逆転、④日米摩擦の激化、⑤円高、⑥価格破壊、⑦雇用問題の深刻化、⑧阪神大震災、⑨地下鉄サリン事件とオウム真理教、⑩究極の政治的無関心と不安定な政治、⑪しのびよる金融不安。

[3] 楽観論と悲観論とが交錯する東アジアの経済発展に対する見方

2010年頃には、アジアが中国人・華人経済圏となり、日本は「アジアの香港」となるとの説がある。他方、アメリカの経済学者ポール・クルーグマンは東アジア諸国・地域の経済発展を「技術進歩なき経済発展」と決めつけた。東アジアの当面の経済発展を「技術と資本の移転による経済発展」とみなすほうがより的確ではないか。また、そうした発展戦略は決して誤った方策ではなく、むしろ賢明な方策と評価すべきではないか。第二次大戦後の日本と現在の東アジア諸国の違い：技術の発展のステージの差異が、両者の発展戦略の違いを説明する。

[4] 20世紀はイノベーションの世紀

20世紀は「イノベーションの世紀」であり、それゆえに「成長の世紀」であった。実際、20世紀中に、無数のプロダクト・イノベーション(新製品)が絶え間なく立ち現れた。自動車、航空機、家庭用電化製品、コンピュータ、OA機器、等々。大量生産・大量消費につながるプロダクト・イノベーションと、労働生産性の向上をもたらすプロセス・イノベーションが経済を成長させた。と同時に、エネルギー多消費型経済社会への移行を促した。マルサスの『人口の原理』(1798年)が発刊された当時の世界人口は9億、1900年の世界人口は16億、1950年の世界人口は25億、2000年には61億。人口の増加は経済発展の結果である。国連等の推計によると2050年の世界人口は100億。19世紀は相対的には「停滞の世紀」であった。19世紀のイノベーションといえば蒸気機関と鉄道ぐらいしかなかった。

[5] 情報・通信分野がプロダクト・イノベーションの「最後の泉」か？

いま現在、プロダクト・イノベーションの宝庫は何かと問われれば、誰もが口をそろえて情報・通信分野と答えるであろう。しかし、いかなる技術革新であれ、やがては飽和状態に達する、すなわち新製品は出尽くし、それらの普及率も閾値に達してしまう。携帯電話、マルチメディア装置、その他情報・通信関連の技術革新が飽和点に達するのは2015年頃と見る。その後を引き受けるのはなんなのか？もし仮に後を引き受けるものがなければ、21世紀は「停滞の世紀」とならざるを得まい。

[6] 21世紀のマルサス問題・ケインズ問題

20世紀の80年代までは、高度なモノ作りは、OECD諸国(8億6500万人)、旧ソ連・東欧(4億1800万人)、アジアNIES(1億2400万人)で行っていた。つまり、14億人が工業化社会にすんでおり、残りの40億人は農林水産物と資源と引き替えに工業製品を入手していた。しかし、27億人のアジア人(日本とNIESをのぞく)と4億人の中南米人が高度なモノ作りにエントリーしてきた。過剰生産の危機：21世紀のケインズ問題、エネルギー・環境・食糧問題の深刻化：21世紀のマルサス問題。

[7] 21世紀のケインズ問題とマルサス問題への処方箋

(1) 地球規模でのケインズ政策の必要、先進諸国の乗数は小さく、公共投資等による内需喚起には限界あり。途上諸国の乗数は相対的に高いから、途上諸国への公共投資等による内需喚起は「費用対効果」という点からして望ましい。

(2) 先進諸国と途上諸国の「分業」による「棲み分け」を図る。たとえば製造業の生産拠点を中国、タイ、マレーシア等に移転し、日本は情報、ソフトウェアに産業構造の重点を移行させる。

(3) イノベーションの火種を絶やさないよう、国際的な戦略を練る。

(4) メタボリズム(循環代謝型)文明社会の構築、日本の省エネルギー技術の途上諸国への移転の積極化、設備更新のための資金援助と人口抑制のための教育援助。

(5) 化石燃料の消費抑制のための国際的枠組みの構築。

(6) 究極の省エネルギー社会へむけて、画期的なプロダクト・イノベーションとプロセス・イノベーションに挑戦する。

[8] 21世紀は「停滞の世紀」か？

以上のような対応が進まなければ、21世紀は「停滞の世紀」とならざるを得まい。工業製品の市場の奪い合い、天然資源の奪い合いという悲惨な事態が招来されかねない。

学術情報センター電子図書館サービス　　◎学術情報センター

創立10周年記念講演

演題 「わが国のマルチメディアネットワークの将来」

The Future of Multimedia Network

日本電信電話株式会社常務取締役・研究開発本部長　青木　利晴
Toshiharu AOKI
Executive Vice President, Senior Executive Manager,
Research and Development Headquarters, Nippon Telegraph and Telephone Corporation

(この講演は、平成8年11月1日(金)、国立オリンピック記念青少年総合センターにて、行われたものである。)

学術情報センターの創立10周年、まことにおめでとうございます。先進的なネットワークの運営をずっとされておられること、また学術研究のパイロットとして歩いてこられたことに心から敬意を表するとともに、ますますのご発展をお祈りしたいと存じます。

さて、本日はからずも講演をさせていただく機会をいただきましたが、私どもネットワーク提供業者として、猪瀬先生をはじめスタッフの先生方に常にご指導を受けておりまして、別のことばで言えば、いつもおしりをたたかれて新しいネットワーク作りを進めております。その一環として、例えば、コンピュータ通信ネットワーク建設、運営やそのサービスの改善に努めてまいりました。

現在、電話からマルチメディアへと大きく流れが変わりつつある中で、学術情報センターの先見性と申しますか、新しい流れを予測して日本のネットワークを的確な方向にお導きくださったことに心からお礼申し上げたいと思います。

ただいま、佐和先生からずいぶん広いお立場のお話がありました。本日は私の演題でありますネットワークの発展についてお話を申し上げたいと思います。何分、私はネットワーク業者でございますので、手前味噌もございますが、順繰りにお話し申し上げたいと思います。

(以下スライド併用)

〔マルチメディアネットワークの将来〕(図1)

〔ネットワークサービスの発展〕(図2)

〔電話加入者数〕(図3)

現在電話がどれくらいあるかというと、ご覧のように約6100万ほどの電話がありますが、100年以上前から電話を構築してまいりました。ここ30年間の間に、そのうちの5000万が作られました。それまでは加入数が少なく平らのカーブだったのですが、増加が見られたのはほんの30年間のことです。

この30年間はいろいろ技術進歩があり、もちろん社会のニーズが高まったということもありますが、アナログからデジタルになった、あるいはコンピュータ技術が進展したとか、あるいは光ファイバーの技術が進んでコストが安くなったなど、いろいろな技術的なバックグラウンドもあってこのように成長したものと思っています。

最近は、これの伸びがだいぶ飽和してきて、昔ほどに伸びておりません。ただ、昔学校を出て会社に勤め独身寮に入ると、独身寮に1台だけ電話があって、それを皆さんで使い合うという時代でしたが、今は独立するとすぐに電話を持ちます。あるいは、一軒のお宅で子どもさんが別の電話を持つぐらいの時代になってきました。年間約100万ぐらいの増加があるわけです。

しかしながら、毎年の増加量で、最近は携帯電話が一般の加入電話を超えるようになりました。

それから後ほど申し上げますが、ISDN(デジタルで64キロビット／秒の情報をお客様にお送りするサービス)をお使いになるお客様の増加量が、このアナログの電話の増加量よりも増えてしまいました。今年はそれが逆転したという大変画期的な年です。今後は電話はむしろ増えるのではなくて、だんだん減ってくるのではないかと予測しています。さてそれでは、電話以外のものについて若干最近の傾向をご説明いたします。

これは、パーソナルコンピュータの普及です。1996年のところを見ますと、約2000万台です。これが調査によりますと、ネットワーク率でいうと30%のコンピュータが、ネットワークつまり電話網を使ってインターネットやパソコン通信などにつながっています。

学術情報センター電子図書館サービス　　©学術情報センター

すなわち、約600万台ぐらいが現在ネットワーク化されていると考えられます。

それに対して2000年にはどうなるかというと、このパソコン現在の伸びからは、日本は約4000万台のパソコンになる。しかもご覧になってわかりますように、ネットワーク率が60～70%になります。こうなると、4000万台の70%として、2800万台のパソコンがネットワークにつながります。

先程、電話は6100万台と申し上げましたが、それの半分ぐらいの数の新しいパソコン通信、コンピュータ通信がそのころは一般的になっているだろうと考えられます。

〔携帯電話とPHS〕(図4)

これは、ご承知の携帯電話あるいは昨年から売り出したPHSという、いわゆる移動の端末のサービスです。なぜ伸びたかというと、もちろん規制緩和もございます。しかしながら我々技術者としましては、最近のPHSにも見られますように、100グラムで100ccぐらい、本当に胸のポケットに入れても重さを感じないぐらいのものでできるように技術が進歩したことも、大きくこの伸びに寄与していると思います。

今年非常に顕著に伸びていますが、来年の3月にはこのPHSを使いまして、32キロビット／秒のデジタルのサービスをする予定です。

今までは、電話を声を伝えるものにだけ使っておりましたが、今後モーバイルコンピューティングと申しまして、パソコンや携帯のパソコンのようなものと無線のPHSとが組み合わさって、端末からデジタルで遠くに飛ばす(コンピュータにアクセスできる)という世界が、今後どんどん開けてくる予定です。

これが携帯とPHS合わせて、これも2000年ごろには2000万ぐらいの需要になるのではないかと考えています。

〔ISDNサービス〕(図5)

もう一つ見逃してならないのは、ナローバンドISDN(64キロビット／秒のサービス)の伸びです。これももう10年以上前からサービスを始めましたが、ごく最近急に伸びてきました。

今年の7月には、このISDNのお客様の増加数の方がアナログの加入電話の増加数を超えてました。

これも紀元2000年ごろには、加入者数で500万、64キロのチャンネルにすると約1000万ぐらいのチャンネル数になるだろうと予測しています。

〔マルチメディア時代のネットワークサービス〕(図6)

さてそれではこのマルチメディア時代のネットワークサービスはどうなるかということで、多少技術的に細かくて恐縮ですが、私どもの考え方を申し上げます。一番上がもともと今電話網を使っているコンピュータの通信、インターネットです。これはTCP/IPというプロトコルを使っていますが、このお客様がパソコンを使って2000万、3000万と増えてくるということを予測しています。そのため、もっと安くて品質のいいコンピュータ通信網を供給しようということで、オープン・コンピュータ・ネットワーク(OCN)ということを今年の10月から開始しました。つまりコンピュータネットワークを発展させるための安いネットワークを提供しましょうということです。

それから下に示されているものは先程のISDNですが、これもどんどん伸びてきて、スピードの速いものを皆さんがどんどんお使いになるようになるだろうと考えています。

一番下の流れは、ビジネス用の高級なお客様用に使うネットワークでして、データ通信として現在も盛んに使われているものです。

これらがだんだん発展して2000年ごろになると、高度OCNといいますか、新しいいろいろなプロトコルが出てきてコンピュータ通信が盛んになると思います。

ナローバンドISDNはさらにスピードが上がり、高速の画像等を送るような通信網に変化していきます。

これらが、世に言う光ネットワークや、ATMバックボーンというものに乗って、いろいろなかたちでネットワークサービスが行われるだろうと考えています。

さらにこのうえにミドルウェアというものがあります。例えば、セキュリティであるとか、ディレクトリであるとか、いろいろなネットワークを安心して、信頼して使っていただけるような付加価値を乗せて皆様にお使いいただくというように考えています。

〔OCN回線サービスのネットワーク構成〕(図7)

今のオープン・コンピュータ・ネットワークの構成をご紹介いたします。

これは、まったくある意味でインターネットと同じですが、このバックボーンにあるオープン・コンピュータ・ネットワークはATMのネットワークを使います。それで加入者系におきましても、それから他事業者とつながっているところにおきましても、実はいろいろな事業者の方々が自由にこういうものを切り取って使えるということです。

学術情報センター電子図書館サービス　　©学術情報センター

つまり、例えばこの左下ですと、ルーターと書いてあるものの下側、あるいはルーターの上に相当するところ、あるいはゲートウェイと書いてあるところ、このいかなる点にもいろいろな事業者の方々がつないで使っていただけるように考えてあります。

〔マルチメディアネットワーク構築〕(図 8)

さて、それではこういう将来のサービスの基盤となるネットワーク、あるいはアプリケーションをどうやって構築していくかということについて申し上げます。3つ大きな課題があると思います。

1つは、基盤になる光ファイバー網をどうやって張っていくか。すでに中継系と申しまして、電話局と電話局の間には全部光網がすでに導入されていますが、いわゆる電話局からお客様のところ、ここまでのところに光ファイバーはまだ入っておりません。これを今後どうやって入れていくか。

それともう1つ重要なことですが、こういうものを使ってアプリケーション、いろいろなサービス、需要、これをどうやって開拓していくか。

それから最後にそういうものに必要な技術をどうやって開発していくのか、この3点です。

〔計画的な構築〕(図 9)

まず光ファイバー網ですが、今電話網に使われている銅線、これは15年から20年ぐらいたつと、例えば水が入ったりいろいろな影響でさびたりしてきて、いわゆるメンテナンスにお金がかかるようになります。

したがって、毎年古くなったケーブルを大量に取り替えているわけですが、従来これは全部銅線で取り替えていました。もし光ファイバーのコストが銅線並みになれば、当然光ファイバーに置き換えたいわけです。これが今から申し上げます技術のターゲットになりました。

〔アクセス網の光化展開の加速〕(図 10)

電話局からお客様のところに電話を送るときに、それを銅線で送るかあるいは光を使って、しかも電気回路との変換などいろいろあるわけですが、それでも光の方が安い、あるいは銅線と同じにするようにするにはどうしたらいいかということを研究してまいりました。つい今年の7月ですが、私ども数年の研究成果が実って、銅線と同じコストで光を通じて電話が供給できるということを発表いたしました。

〔FTTH への展開〕(図 11)

これはどういうシステムになるかと申しますと、これは真ん中のところの左側のところにNTTという局社がありますが、ここに光ファイバー終端装置があって、電気光変換をするわけです。ここから光に信号を乗せまして、お客様のところのごく近くの電柱、あるいは上にあるアパートの中のところに持ってきます。ONU(オプティカル・ネットワーク・ユニット)が右上にあるような箱に入っていますが、ここに10加入ぐらい多重化して、ここまで光でもってきて、ここからばらして電気信号にして銅線で送るという方式です。これだけではありませんが、これが主な開発の成果の中でも一番のものです。これを含めて、局社からここまでのところでトータルで電話を送るときに、銅線と同じコストになるというシステムを開発いたしました。

この方式で、来年から精力的に銅線を光に取り替えていく予定です。左上の電柱に、見慣れない扇形の箱がありますが、今後ご近所で上をご覧になったとき、ああいう箱が付いておりましたら「ああ、家のところまで光が来たな」と思っていただきたいと思います。

〔光設備の経済化トレンド〕(図 12)

これが、先程のコストの低下傾向です。この水平の線が銅の加入者線のコストですが、だんだん研究が進むにしたがって光のコストが下がってまいりまして、本年度銅線と光ファイバーシステムのコストがクロスしたということで、来年から導入するわけです。

別のことばで言えば、「電話料金を上げずに光に換えます」と、「余計な投資をせずに換えます」と申し上げています。

〔ユーザと組んだ構築〕(図 13)

それではもう一つの課題でしたアプリケーション、あるいは需要をどうやって開拓するかということです。

〔マルチメディア共同利用実験〕(図 14)

私どもは2年前にマルチメディア共同利用実験を提案して、日本全国にATMのバックボーンネットワークを張って、それをいろいろなお客様にご利用いただいて、新しいサービスを開拓してみようというプロジェクトを始めました。

ご覧のように日本に10のノードを設けて、この間をATMのネットワーク2.4ギガないし10ギガの高速ネットワークで結び、このポイントからお客様のところまで光ファイバーをひいて、150 メガあるいはそれより低いスピードのデジタルのATMの情報を流します。

これには大きく2つのものがあり、「高速広帯域バックボーン・ネットワーク利用実験」というのが、今申し上げました高速のデータ通信や一般向けの通信です。

それから下のものが、やはりファイバーを使った

学術情報センター電子図書館サービス　　© 学術情報センター

CATV の映像実験ですが、今日はこれについては触れないで、上の2つについてご説明を申し上げます。

〔高速コンピュータ通信利用実験〕(図 15)

まず、共同利用実験のうち 150 メガ以上の非常に高速のデータを使うような実験ですが、約 11 ぐらいのプロジェクトがあり、この左に挙げるような名前のプロジェクトをそれぞれいろいろな機関と共同で実験させていただきました。

ご覧のように一番上の学術研究支援情報網というのは、学術情報センターとの共同研究です。

〔高速コンピュータ通信利用実験〕(図 16)

こういう実験もございます。

例えば、超高速光通信利用実験という一番上のものですが、これは郵政省とやっています。何百キロか離れた電波天文台が電波望遠鏡で受けた信号を、両方で高速で符号化して高速の回線で持ってきて相関をとるわけです。そういたしますと、もしその2地点のところで地殻変動があって数センチ動きますと、到着時間のずれがわかるということで、例えば地震予知に使えるのではないかと思います。

それから遺伝子情報ネットは、東京と京都にある2つの研究所の高速コンピュータを結び、両方のコンピュータ同士で複雑な遺伝子の構造解析をするという実験です。

このようにいろいろ新しい研究課題というか、ネットワークができたために新しい研究ができるというものもたくさんあるわけです。

〔学術研究支援情報網〕(図 17)

これは、先程最初にご紹介した学術研究支援情報網の例ですが、学術情報センターとつないで、いろいろなプロジェクトをこの上で研究しているようです。例えば気象衛星からの情報をいろいろな大学に配って、それを研究するというようなものにも使われています。

〔GIBN (Global Interoperability of Broadband Networks)〕(図 18)

このネットワークはさらに発展して、GIBN と言っていますが、学術情報センターを中心にして、私どもの研究所やアメリカの大学を結んで、ここにあるような実験を今後やっていこうということで、国際的なものにもなっています。

〔NACSIS / NTT / UW　GIBM 実験〕(図 19)

これがそのネットワークですが、NTT、KDD、AT&T をつないで実験をしているところです。

〔一般利用向けマルチメディアネットワーク利用実験〕(図 20)

次は、一般利用向けの実験です。マルチメディアネットワークなど高速の基盤ができますと、どんな社会になるか、どんなネットワークサービスができるか、どのように個人生活が変わるかと、いろいろなことを考えていかなければいけないわけです。当面私どもでネットワークを供給いたしまして、左にあるようないろいろなプロジェクトを、右にあるような実験参加者の方々が実験をしています。

〔一般利用向けマルチメディアネットワーク利用実験〕(図 21)

新聞、医療、教育、さらに趣味とか生活、実際にデザインとかの製品開発にお使いになっているところもあります。

〔サービスニーズとタイアップした開発〕(図 22)

次は、もう一つの重要なポイント、R&D 研究開発の目標やタイミングをどう設定するかという課題です。つまり、私ども研究開発をする立場として、従来の研究開発と違って目標設定というのが非常に難しいのです。

昔ですと、例えば光の伝送方式を研究すると、その結果、どんどん多重化したものができてきますと、電話料金を下げることができます。コストが安くなるわけで、お役に立てると同時に論文も書けるということで、非常に目標設定しやすかったわけです。しかしながら、今は我々がせっかく技術やサービスを技術者の方で作っても、本当に世の中の方でそれを評価していただかなければ何もならないという時代です。先程の共同実験もそうですが、直に研究開発者が世の中と接触することが必要だろうということで、総合実験というものをやっています。

つまりネットワークを使って、最先端の研究室レベルで出来上がったようなものを、世の中にプローブとして出して、いろいろな評価をしていただくということです。

〔VI&P 総合実験〕(図 23)

これは少々古い時代からそういうことをやってきたのですが、現在のナローバンド ISDN を伸ばすために、5 年ぐらい前からここにあるようないろいろな項目について実験をやってきました。ネットワークは、武蔵野と横須賀を結んだ高速のネットワークを使っています。

〔VI&P 総合実験〕(図 24)

そのあと 93～94 年には、光ファイバーあるいは ATM

学術情報センター電子図書館サービス　　©学術情報センター

という広帯域の通信を用いた新しいマルチメディアサービスを実験しました。

右下にあるスーパー・ハイディフィニション・テレビ(SHD)と申しますが、2000×3000 ぐらいの非常に高精細な、HDTV の4～5倍の高精細です。これを使うことによって、病理診断もできます。

それから超高速の回線を使いまして、左下は離れたところのスーパーコンピュータ2台を使って、共同の作業をさせるという実験をしているところです。

〔VI&P 総合実験〕(図 25)

これは、去年から始まったＰＨＳを使った実験です。モービル・コンピューティングという分野、あるいはパーソナル・マルチメディア・コミュニケーションという将来の分野がどう発展していくかということを、実際に研究と同時に評価をしていただくものです。武蔵野のエリアにたくさんのアンテナを立て、研究者や来た方々が使ってみているという実験をやっています。

〔社会基盤としてのマルチメディアネットワーク〕(図 26)

今後のマルチメディア社会とネットワークの関係について少々述べたいと思います。

ネットワークといっても広い意味を持つわけですが、ファイバーやATM交換機というような物理的な意味でのネットワーク、それから社会活動、個人生活を支えるための公共ネットワークとして学術、医療、教育、行政、福祉、防災などのネットワークが必要です。これは社会基盤として必要であろうと思います。そのうえに、いろいろな企業が活動したり、あるいは個人が楽しんだりするためのネットワークができてきます。

〔社会基盤としてのマルチメディアネットワーク〕(図 27)

これは、私どもマルチメディアネットワークを構築するものとして、社会基盤として活用されることを考え、提供する事業者としてそれを支える技術をどう研究開発していくかということです。

かつて私どもの会社は、電気通信事業者と言われていたわけですが、例えば最近ネットワークにつながったパソコンを使ってマウスをクリックをしたときに、ほしい情報が、いったいアメリカにあるのか隣のビルにあるのか、お客様は何もそんなことは考えずにクリックをされるわけです。

つまり情報がデジタル化され、データベース化されて、それがネットワークとつながりますと、結局我々は情報共有というプラットホームを提供することが必要ではないかと思います。皆様すべての人がすべての情報に自由にアクセスできるようなプラットホームを提供するのです。

つまり情報通信事業のしかけ、ここに書いてありますが、例えば安心してお使いになれるようなセキュリティ、ディレクトリ、符号化など、そういうものを乗せて皆様にアプリケーション、パソコン通信、インターネットを使っていただいていると考えられます。

〔社会基盤としてのマルチメディアネットワーク〕(図 28)

さらにこれが発展すると、将来はいわゆる情報のデマンドとサプライを結び付けるような情報流通事業のしかけを、我々ネットワーク事業者は提供するのではないかと思うのです。

つまり先程の情報共有プラットホームのうえに、情報流通プラットホームというのを書いてみましたが、例えば将来ショッピングモールというのがネットワーク上でできるようになります。このためには、例えば電子図書館、サイバースペース、こういうサイバースペースに皆さんが入っていって、そこで実際にインフォメーションプロバイダや商店を歩き回ってショッピングして歩くのです。これが現実に、お宅にいてもそういう空間を皆さんが共有してやるわけです。

そこでデマンドとサプライの出会いをうまく作るようなしかけ、エージェントなど、いろいろなものがありますが、そういうものを私どもが供給していくのではないかと考えます。

〔マルチメディア時代を先導する学術情報ネットワーク〕(図 29)

先程3枚ほど、前のスライドで申し上げましたように、学術情報ネットワークはいわゆるマルチメディア時代の社会基盤であるとともに、私どもネットワーク事業者の常に前を歩いてこられたわけです。

学術的にも、それをサポートするうえでも、学術ネットワークは最初にN1ネットワークというのがあります。もう20年以上前で、今のデータ通信のはしりで、パケットネットワークを最初にご提案され、そのあとここにあるようにインターネットバックボーンとして、あるいは最近はATMをバックボーンに高速の学術情報ネットワークを構築しておられます。

〔学術情報ネットワーク〕(図 30)

さらにネットワークの具体的な活動として、学術研究情報の流通やキャンパス情報ネットワークの総合接続、国際的な学術研究情報ネットワークとの協力等々

学術情報センター電子図書館サービス　　© 学術情報センター

があります。

これらを通じて、一方さらなる社会基盤の学術情報ネットワークを最近新しく構築されました。

〔学術情報 ATM バックボーンネットワークの構成〕(図 31)

これは今年の秋から ATM のノードを 29 ほど入れて、前のスライドにもあったような活動をこの大きなネットワークの上でやるわけですが、これがすでに動きはじめております。

これはまさしく質のうえでも量のうえでも、世界でトップレベルの学術情報ネットワークです。

〔学術情報センター〕(図 32)

このように学術情報センターの事業が成熟されるにつれ、新たな時代の動向の先取りということも必要になっていて、これについても学術情報センターは、また先生方は、卓越した先見性によりいろいろな先端的な学術研究支援、あるいは国際的な協調等を含め産学共同研究等をやっておられるわけです。

さらに、今後はわが国の学術研究を革新するセンター・オブ・エクセレンスになられることが期待されているわけです。過去 10 年間いろいろとご発展されまして、たくさんの革新的な研究をされておられましたが、今後もますますご発展されて、私どもをご指導されることをお願いしたいと思っています。

〔マルチメディア社会のスムーズな発展のために〕(図 33)

最後にマルチメディアの社会というか、これは私どもから一方的ですが、情報化社会のスムースな発展のための課題ということを 2～3 申し上げまして、私の講演を終わりたいと思います。

まずこの発展のためにグローバルなルールづくりが今後必要でしょう。つまりこういうネットワークができたということを逆に考えますと、あるいは先程申しました情報共有基盤ができたと考えますと、すべての人が発信、受信できます。これを、権利をどう保障し、まただれにでも送れるという公平性を保障するか、義務づけるかということが課題になってきます。

それからもう一つは情報ハイウェイを、自動車のハイウェイに例えてみますと、これを利用するためのルールづくり、まさしく交通ルールのようなものが必要です。何もサインを出さずに車線を急に変える、あるいは真ん中の車線を非常にゆっくり走るとか、これは言語道断です。他人に迷惑をかけない、あるいは一般的に品質をどう考えるのかというようなことについて、やはりルールづくりが必要です。

もう一つ、ネットワークのうえでいろいろなコンテンツがたくさん流れるわけですが、この知的所有権、倫理なども話題になっています。もし知的所有権を厳しく保護しますと、実際に使う方々にとっては不便になりますし、逆にこれを保護しないと新しいコンテンツを作るインセンティブがわかなくなるということになるわけです。

さらにこういうマルチメディアの社会では、従来の How to ではなくて What、つまり新しいこと、目標を創造するということが、従来以上に必要になってくるかと思います。

そして、今後マルチメディアはたぶんマルチカルチャーというか、各国の文化を生かして、情報においても貿易と同じように一方的にどこかの国からどこかの国に輸出する、輸入するということではなく、バランスのとれた情報の流れが将来できるようなかたちに作っていかなければならないのではないかと思います。

以上、私の講演を終わらせていただきます。ありがとうございました。

学術情報センター電子図書館サービス　　© 学術情報センター

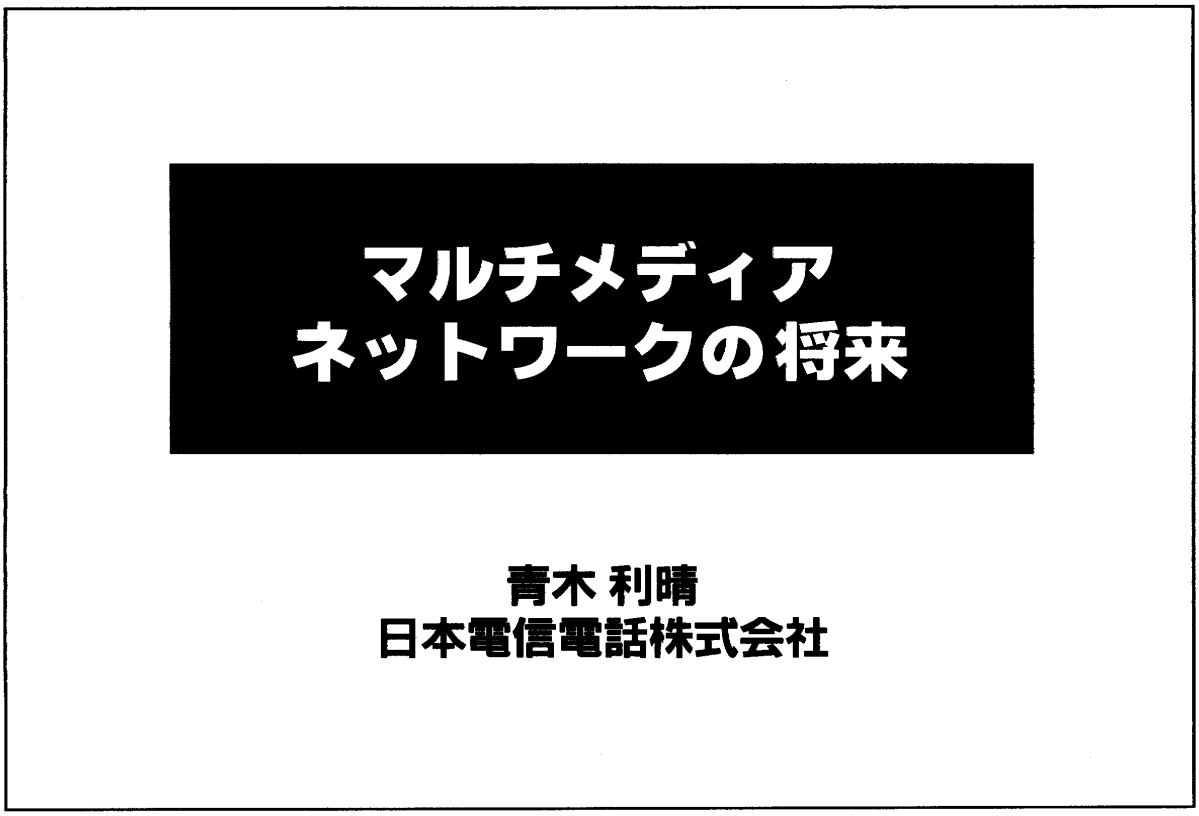


図1　マルチメディアネットワークの将来



図2　ネットワークサービスの発展

演題 「わが国のマルチメディアネットワークの将来」

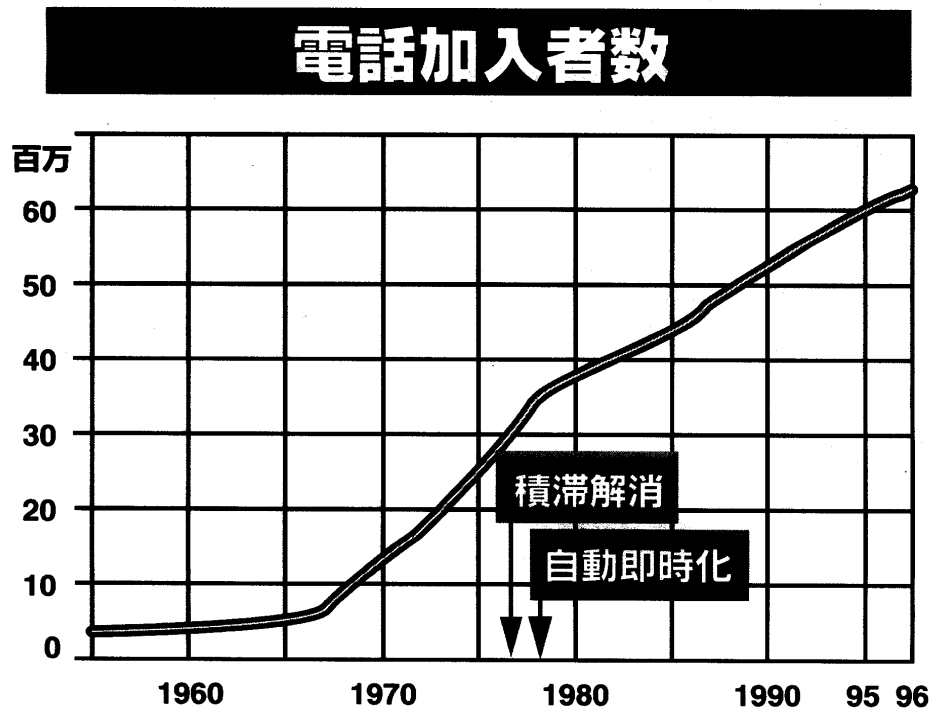


図3　電話加入者数

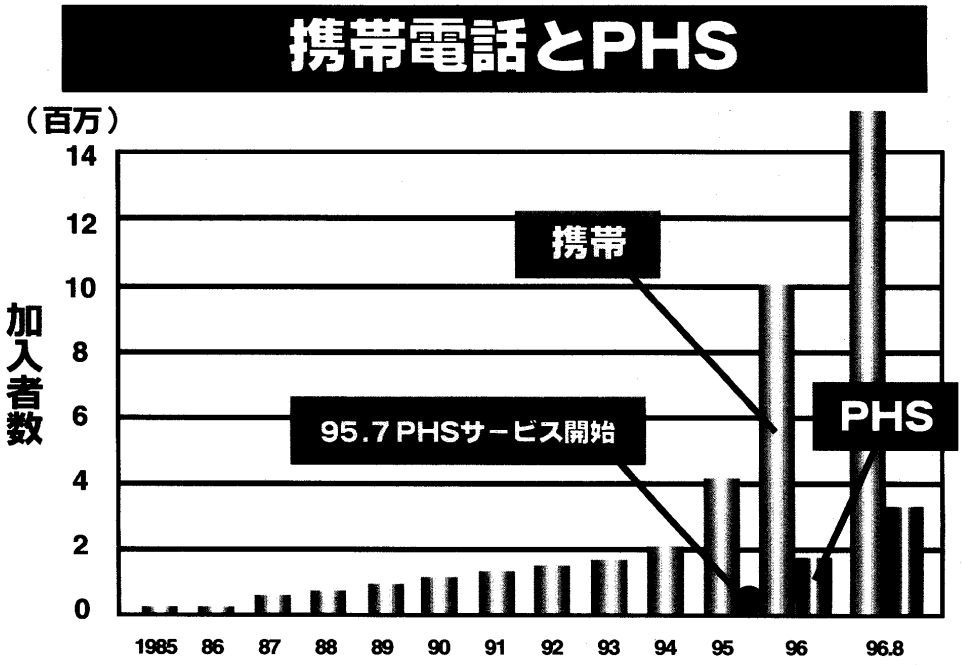


図4　携帯電話と PHS

© 学術情報センター

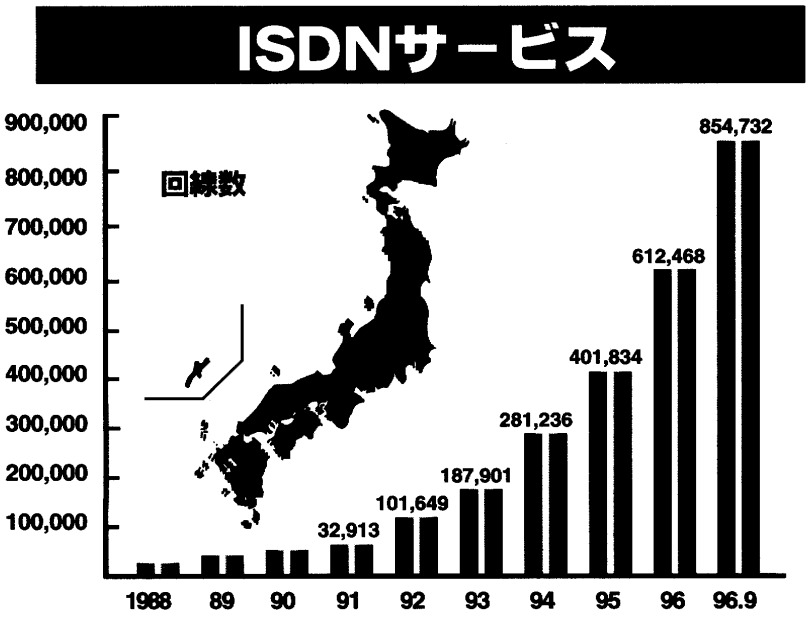


図5　ISDNサービス

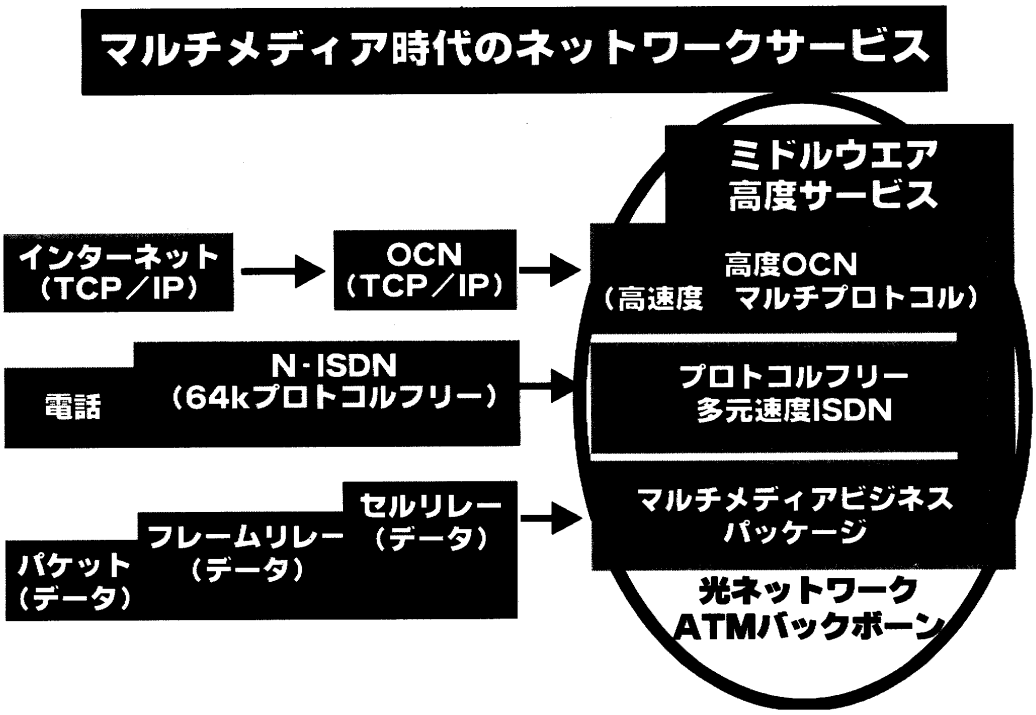


図6　マルチメディア時代のネットワークサービス

演題　「わが国のマルチメディアネットワークの将来」

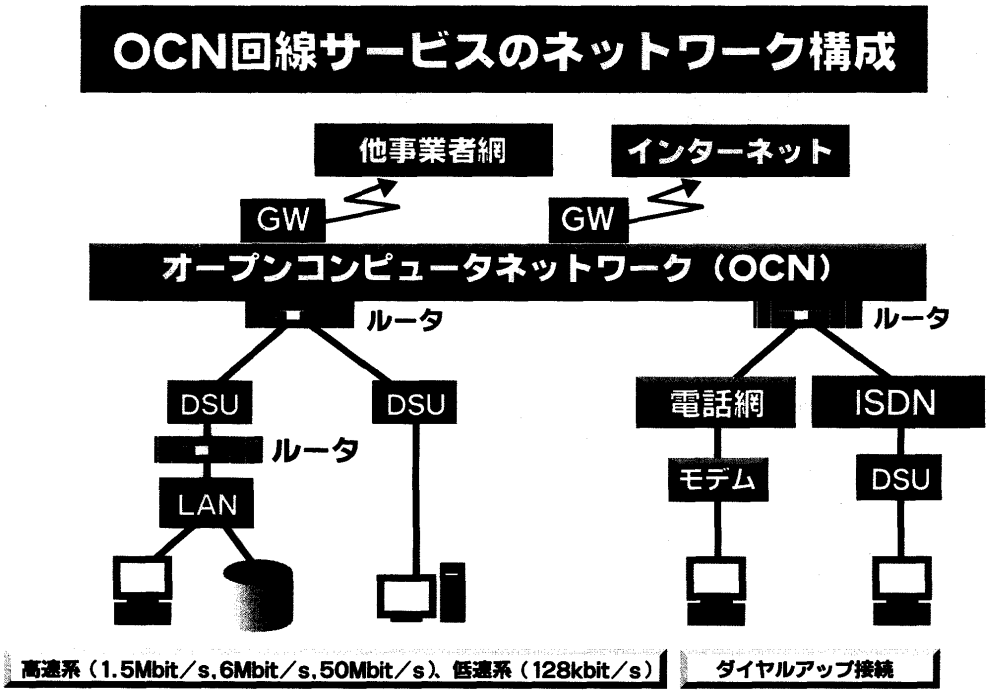


図7　OCN回線サービスのネットワーク構成

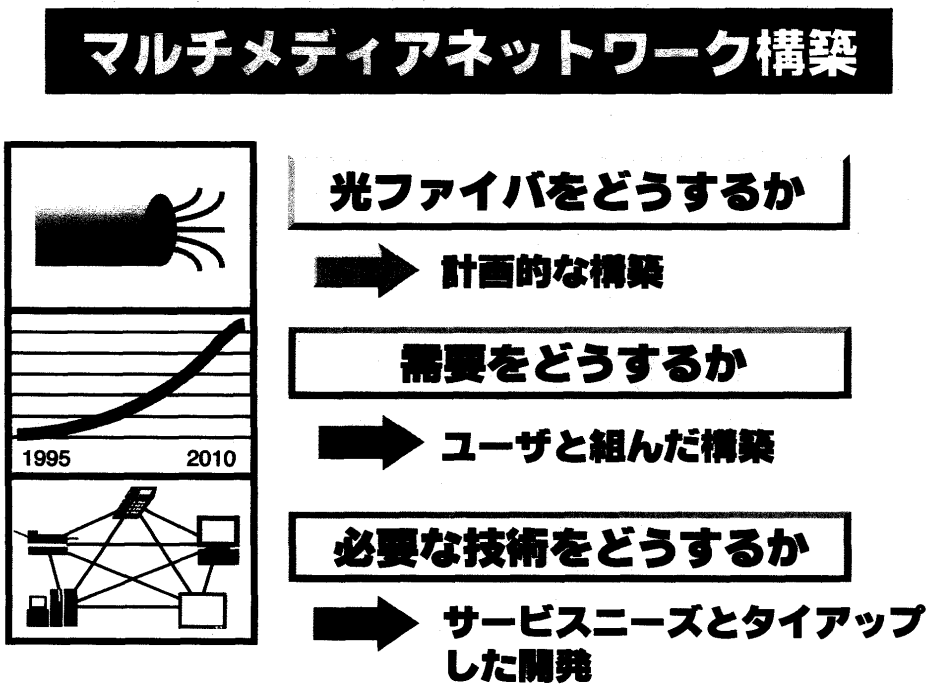


図8　マルチメディアネットワーク構築

学術情報センター電子図書館サービス　　© 学術情報センター

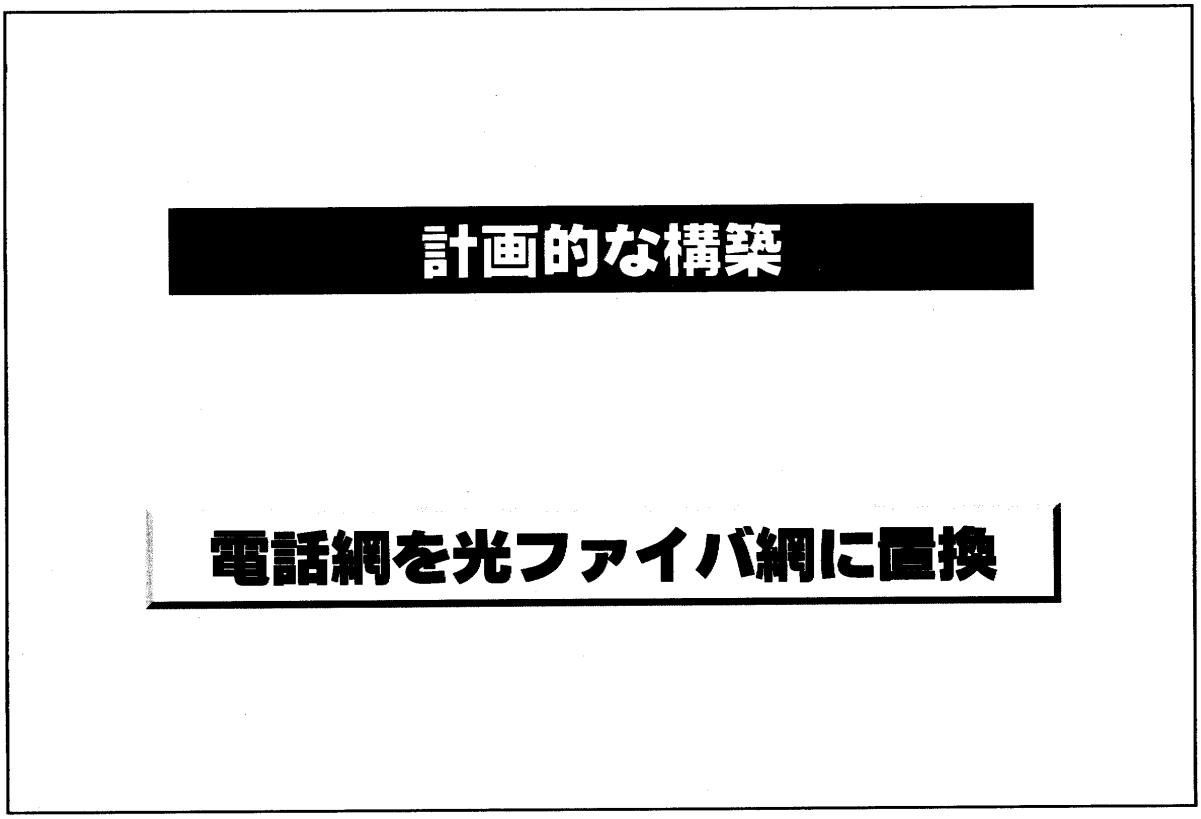


図9　計画的な構築

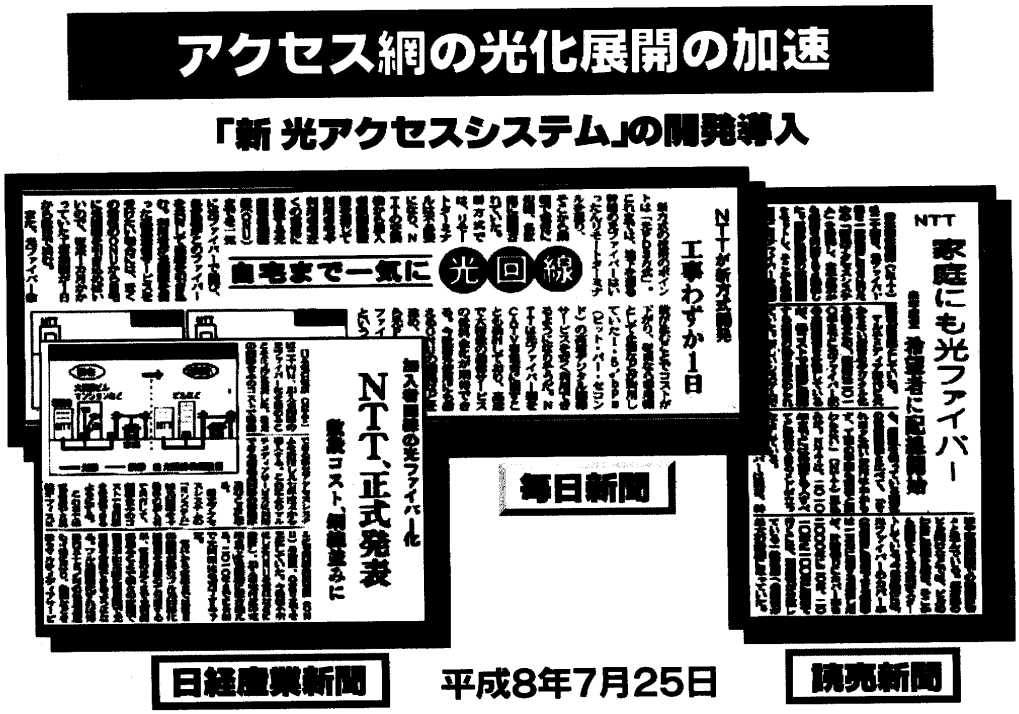


図10　アクセス網の光化展開の加速

演題　「わが国のマルチメディアネットワークの将来」

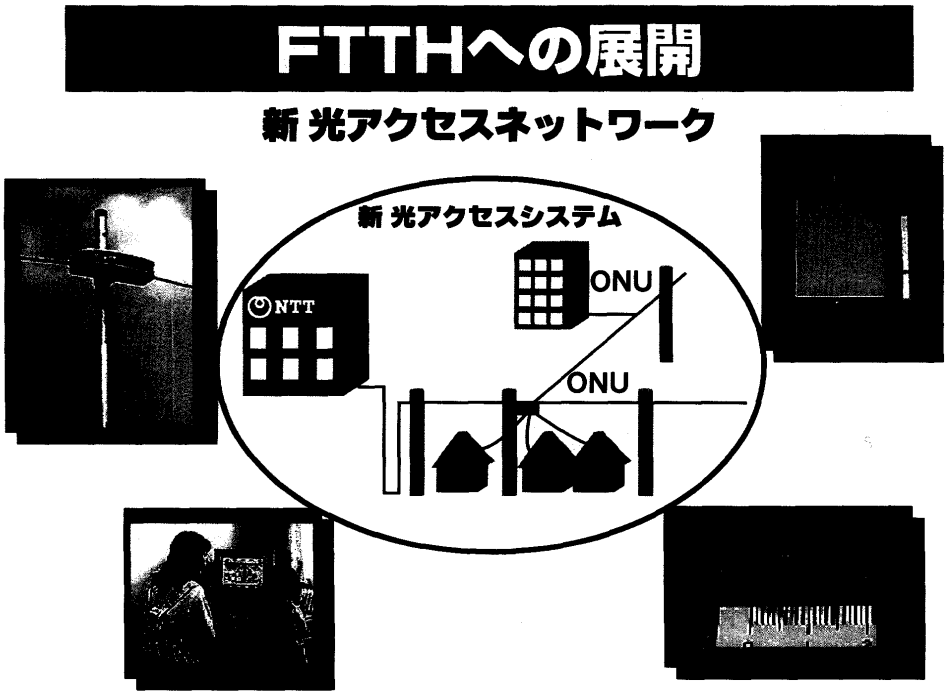


図 11　FTTH への展開

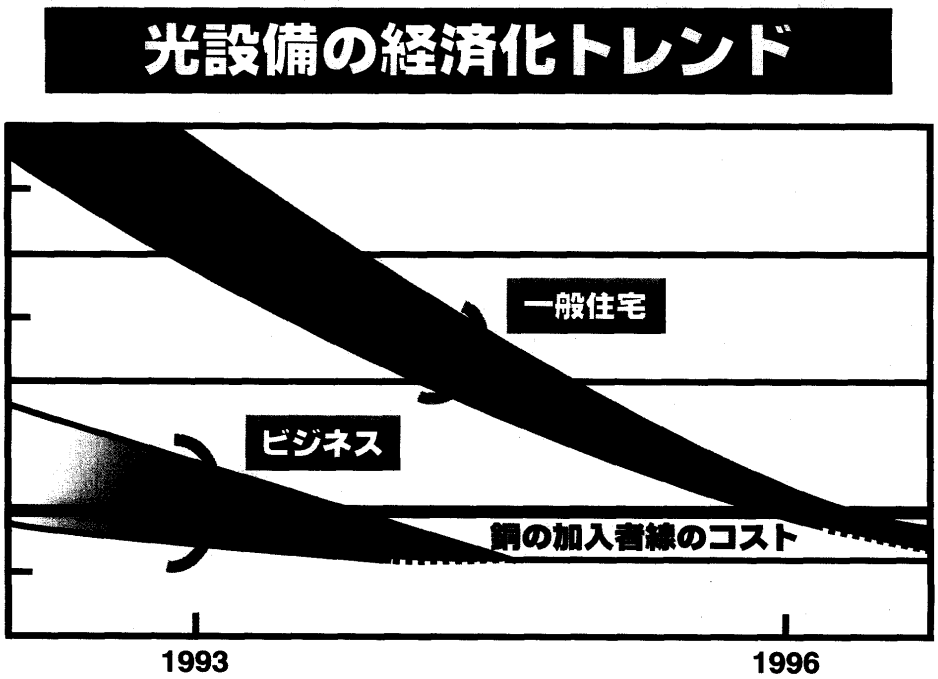


図 12　光設備の経済化トレンド

学術情報センター電子図書館サービス　　© 学術情報センター

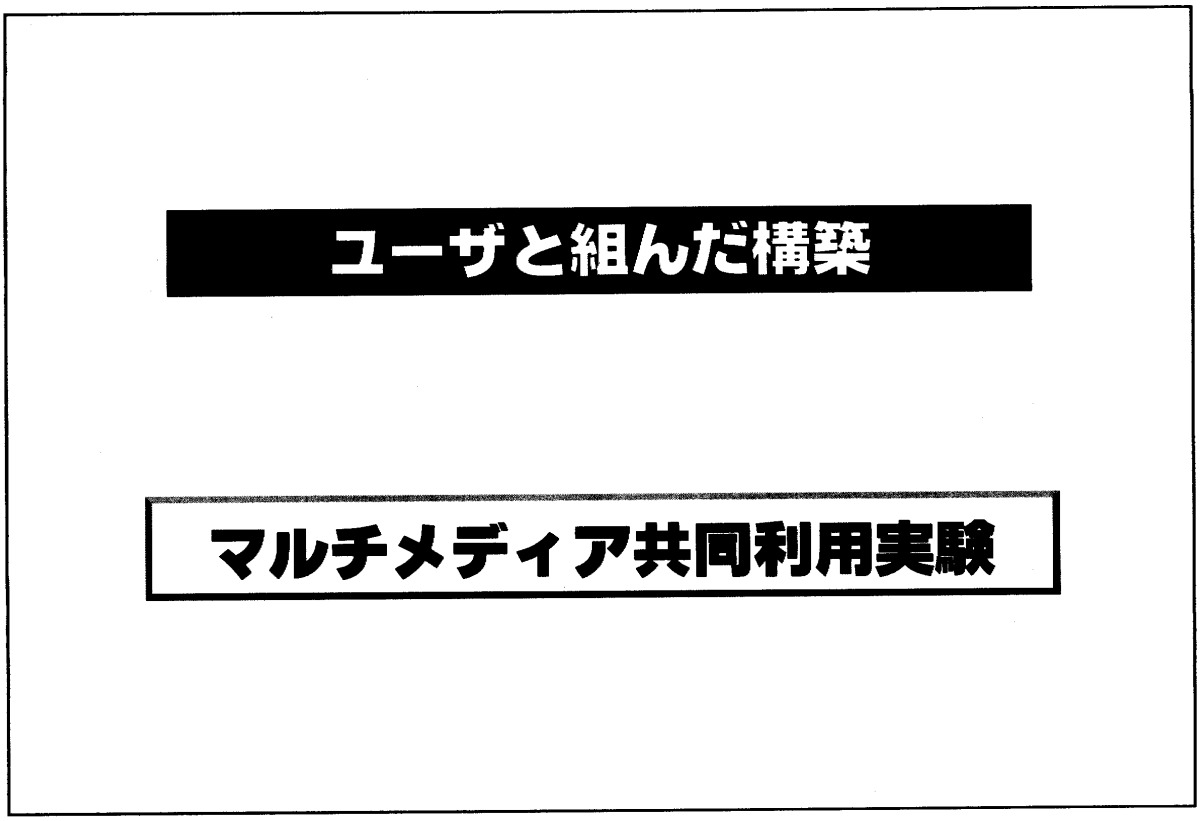


図13　ユーザと組んだ構築



図14　マルチメディア共同利用実験

演題 「わが国のマルチメディアネットワークの将来」

高速コンピュータ通信利用実験

アプリケーション	共同利用実験参加者
学術研究支援情報網	学術情報センター、東京大学 早稲田大学、九州大学、NTT等
集中的医用画像情報処理システム	東北大学、放射線医学総合研究所 理経、聖稜クリニック
ATM広帯域コンピュータ ネットワーク実験	大阪大学、奈良先端科学技術 大学院大学　等
超高精細画像マルチメディアと 超高速コンピュータネットワーク	慶應義塾大学 NTT情報通信研究所　等
高速WANによる 大学間共同システム	信州大学、名古屋工業大学
次世代映像メディア	東京工業大学、慶應義塾大学 東京大学、早稲田大学　等

図 15　高速コンピュータ通信利用実験

高速コンピュータ通信利用実験

アプリケーション	共同利用実験参加者
超高速光通信利用実験	郵政省通信総合研究所 NTT通信網研究所
遺伝子情報ネット	東京大学、京都大学 NTTソフトウエア研究所
超高速マルチメディア分散処理	東北大学、名古屋大学、九州大学 筑波大学、東京大学
HEPnet-J高速通信実験	高エネルギー物理学研究所、 学術情報センター、京都大学　等
医療診断実験	東京女子医科大学

図 16　高速コンピュータ通信利用実験

学術情報センター電子図書館サービス　　© 学術情報センター

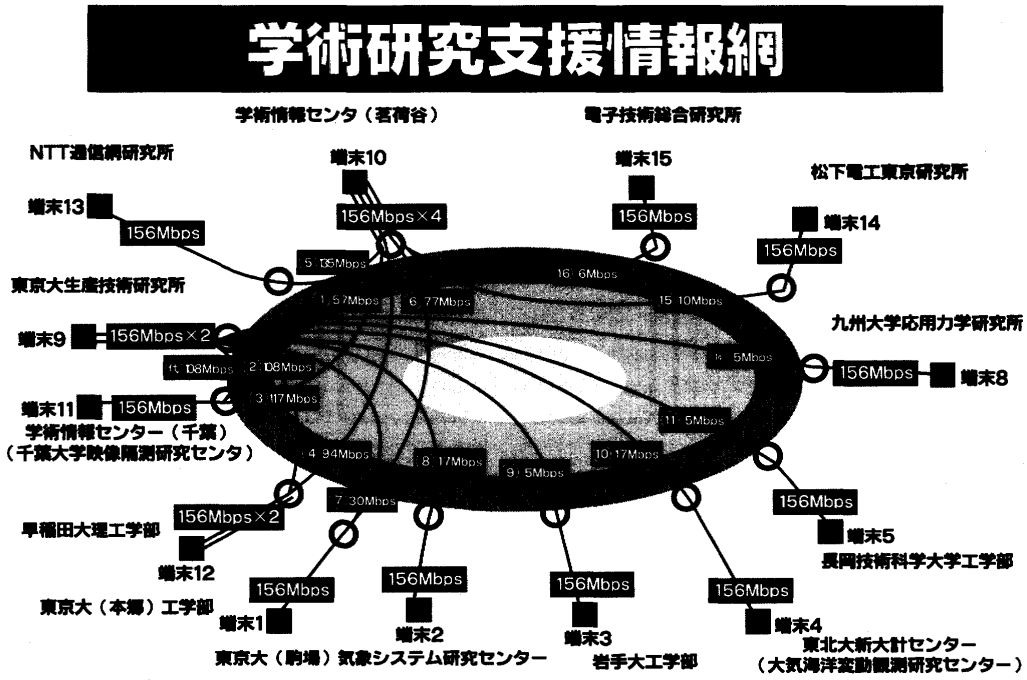


図 17　学術研究支援情報網

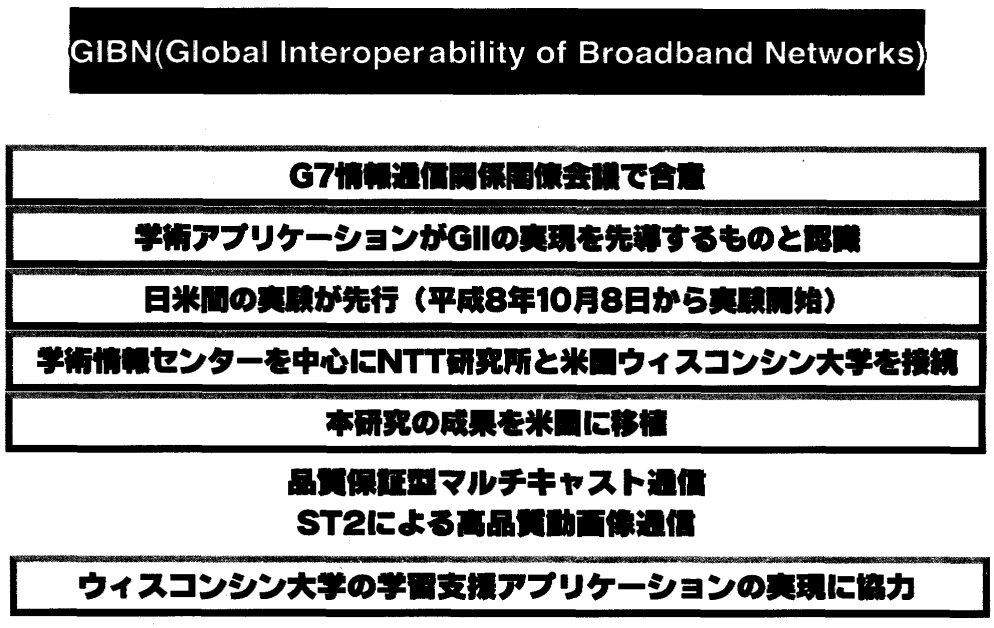


図 18　GIBN (Global Interoperability of Broadband Networks)

演題　「わが国のマルチメディアネットワークの将来」

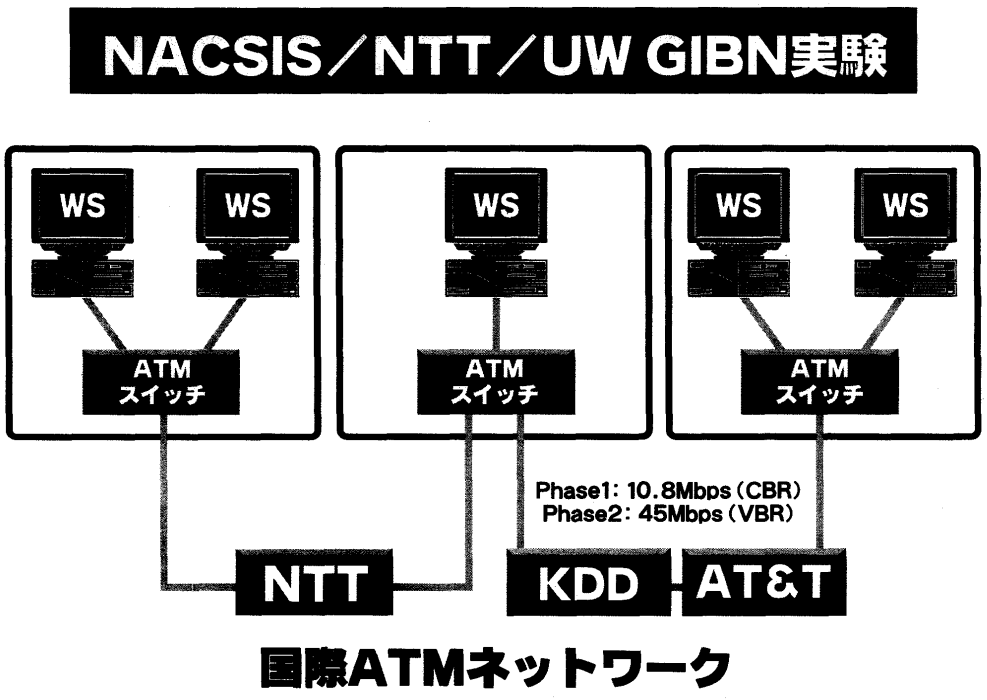


図 19　NACSIS / NTT / UW GIBN 実験

一般利用向けマルチメディアネットワーク利用実験

適用分野	アプリケーション	共同利用実験参加者
エンタテイメント	多人数参加型ネットワークゲーム	スクウェア、セガ
	アニメ・ビデオ・オン・デマンド	旭通信社
新聞出版放送	ニュース・オン・デマンド	朝日、毎日、日経、富士通、JCTV
	TV-CF・オン・デマンド	ビデオリサーチ、東京企画、花王
	電子出版	アスキー、大日本印刷　等
医療	遠隔医療	沖電気、コステム、厚木テレコムパーク
福祉	マルチメディア介護相談	東京海上火災
	身障者への映像制作支援	熊本ソフトウェア研修センタ
教育	遠隔教育	放送教育開発センタ、タカラベルモント、横河ヒューマンクリエイト日本電気　等

図 20　一般利用向けマルチメディアネットワーク利用実験

学術情報センター電子図書館サービス　　©学術情報センター

一般利用向けマルチメディアネットワーク利用実験

適用分野	アプリケーション	共同利用実験参加者
趣味教養	遠隔教養講座	尚美学園、花とみどり社
	ネットワークアート	武蔵野美術大学
	オン・ライン・ショッピング	丸紅、トーハン セイノー情報サービス 等
生活消費	遠隔コンサルティング	第一勧業銀行、住友銀行、東海銀行 第一生命保険、日興システムセンタ等
	マルチメディア・ショールーム	松下電工、東京ガス
	遠隔セキュリティシステム	総合警備保障
	地域生活情報	大分県、情報処理振興事業協会
開発生産	デザイン・設計協調作業	大成清水、鹿島、竹中、日立製作所 等
	製品開発協調作業	三菱電機、キャノン、スターネット

図21　一般利用向けマルチメディアネットワーク利用実験

サービスニーズとタイアップした開発

VI&P総合実験

図22　サービスニーズとタイアップした開発

演題 「わが国のマルチメディアネットワークの将来」

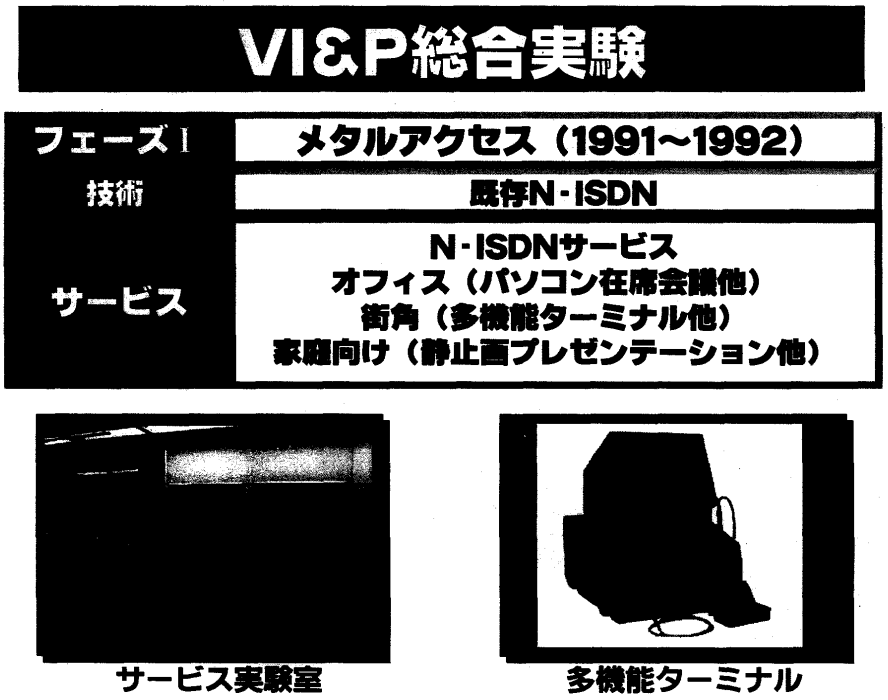


図 23　VI&P 総合実験

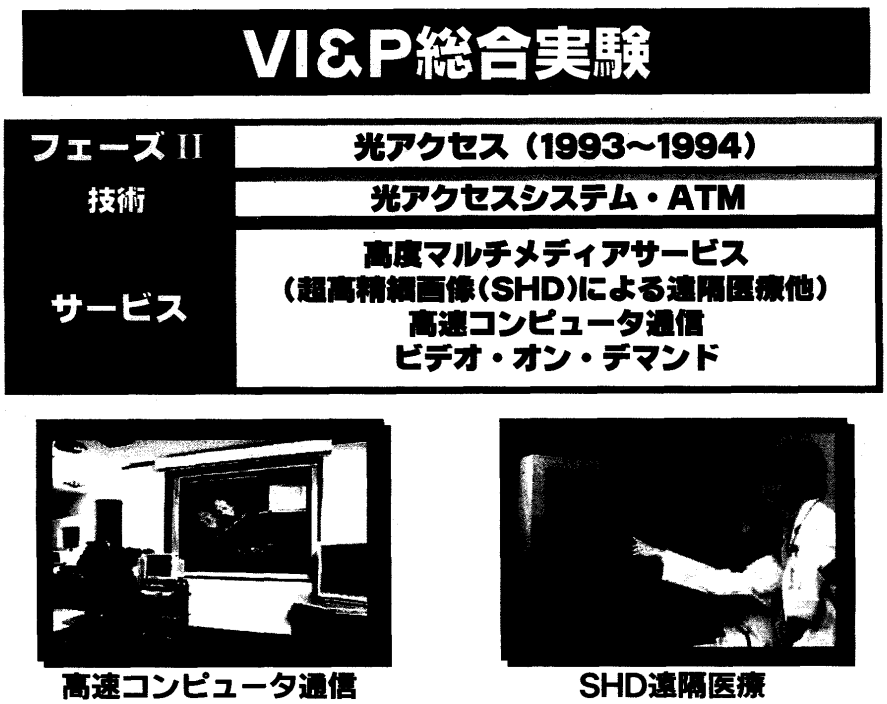


図 24　VI&P 総合実験

学術情報センター電子図書館サービス　　©学術情報センター

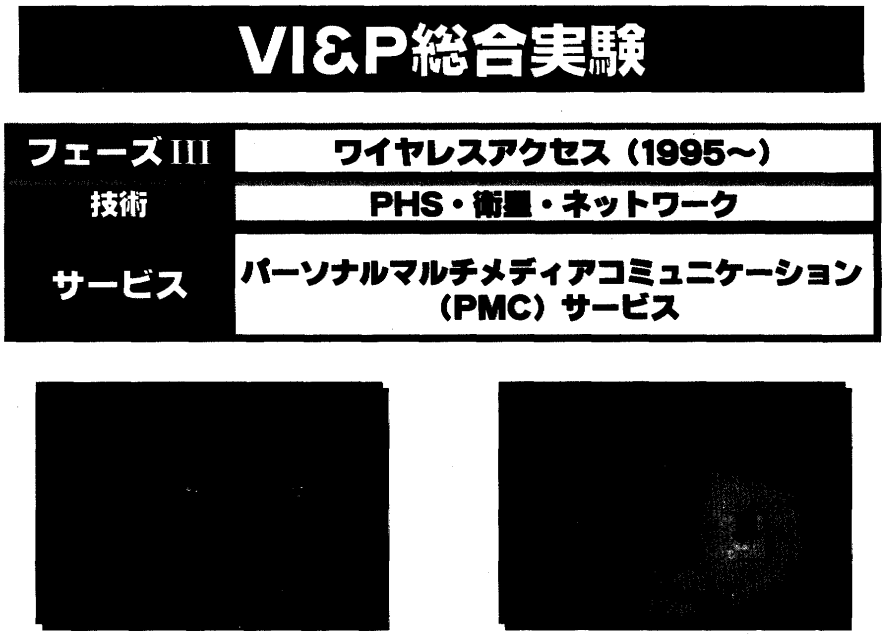


図 25　VI&P 総合実験

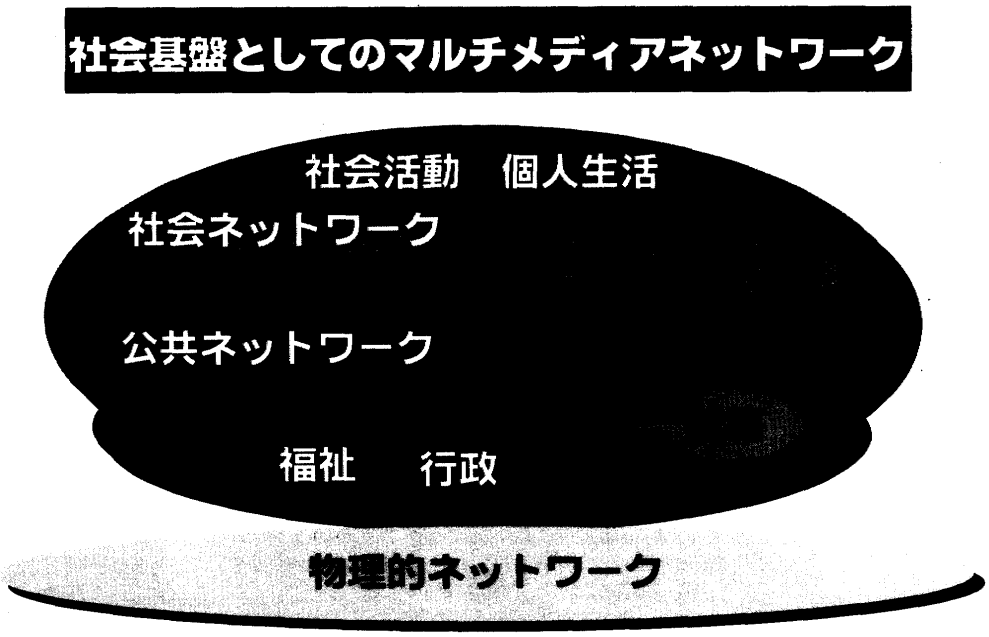


図 26　社会基盤としてのマルチメディアネットワーク

演題 「わが国のマルチメディアネットワークの将来」

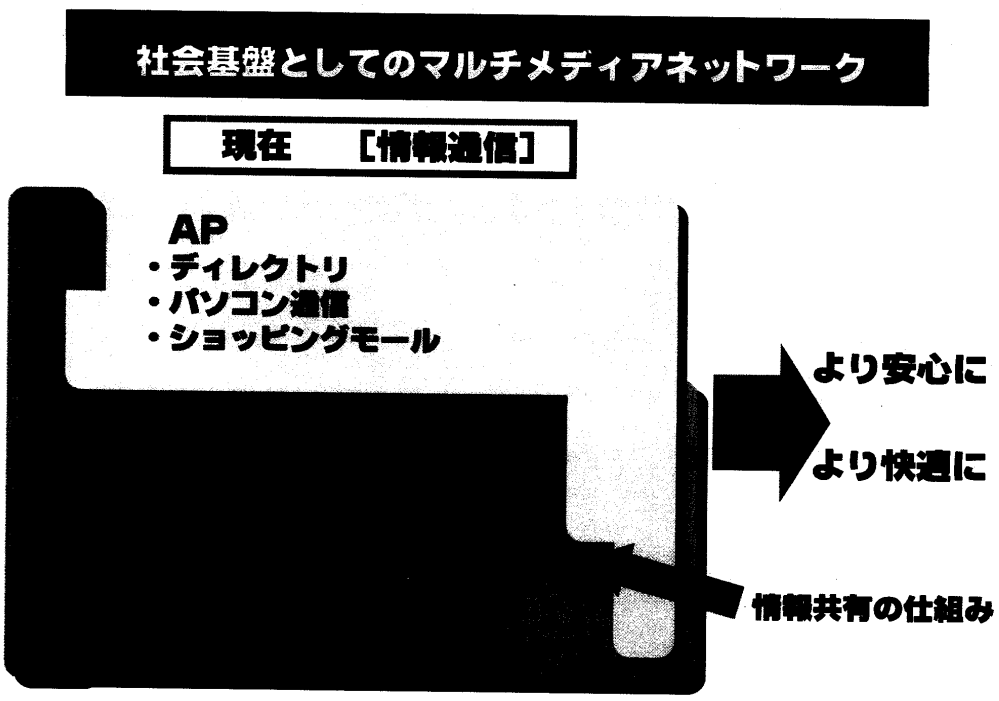


図 27　社会基盤としてのマルチメディアネットワーク

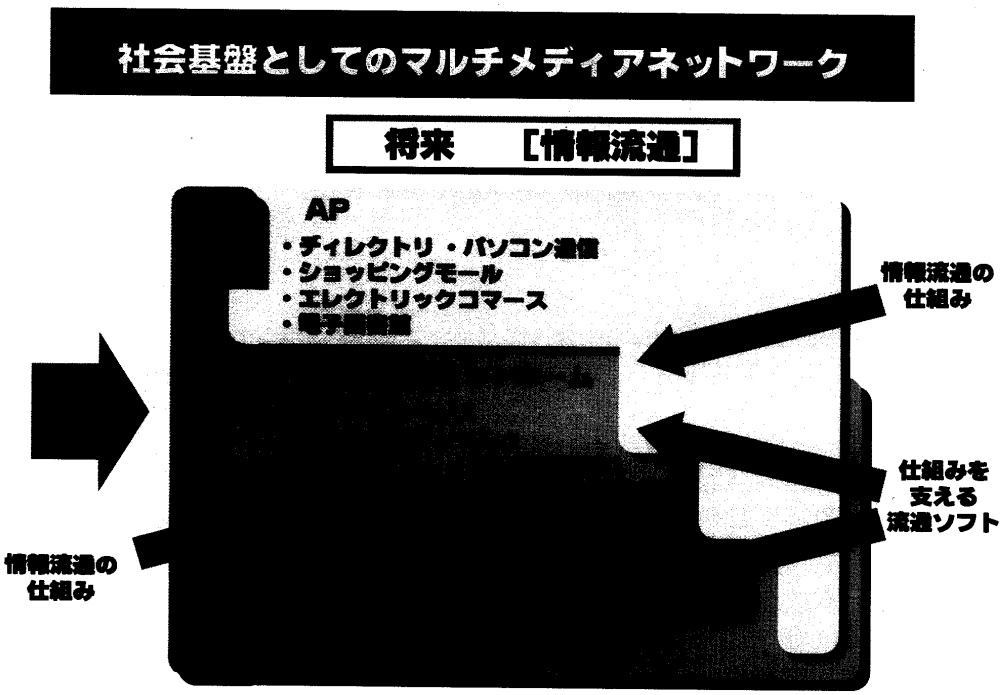


図 28　社会基盤としてのマルチメディアネットワーク

学術情報センター電子図書館サービス　　© 学術情報センター

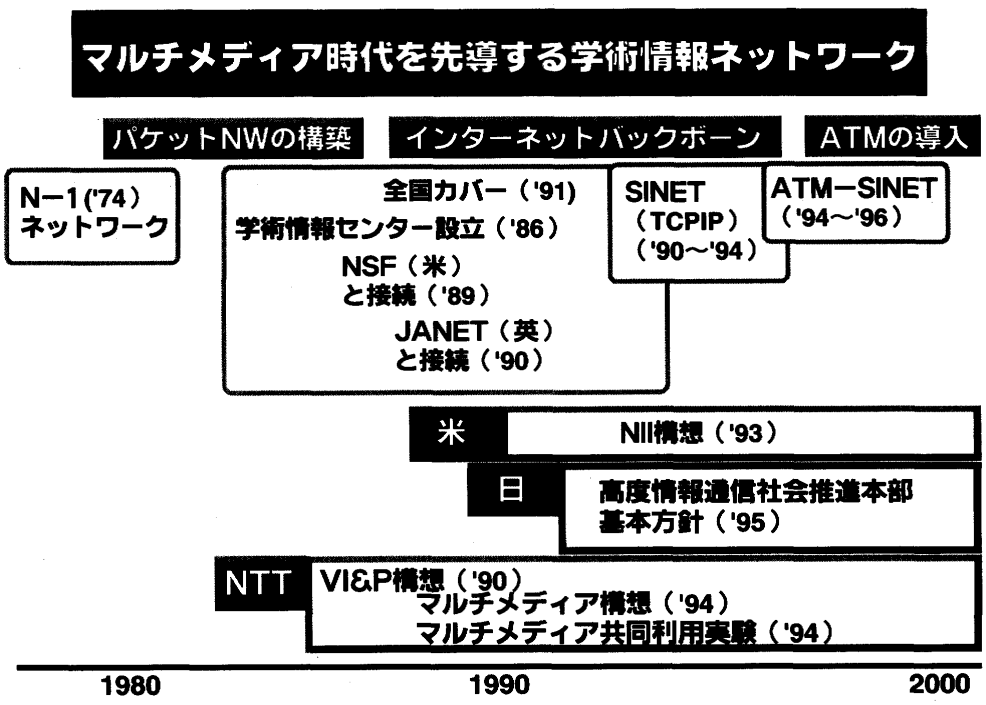


図29　マルチメディア時代を先導する学術情報ネットワーク

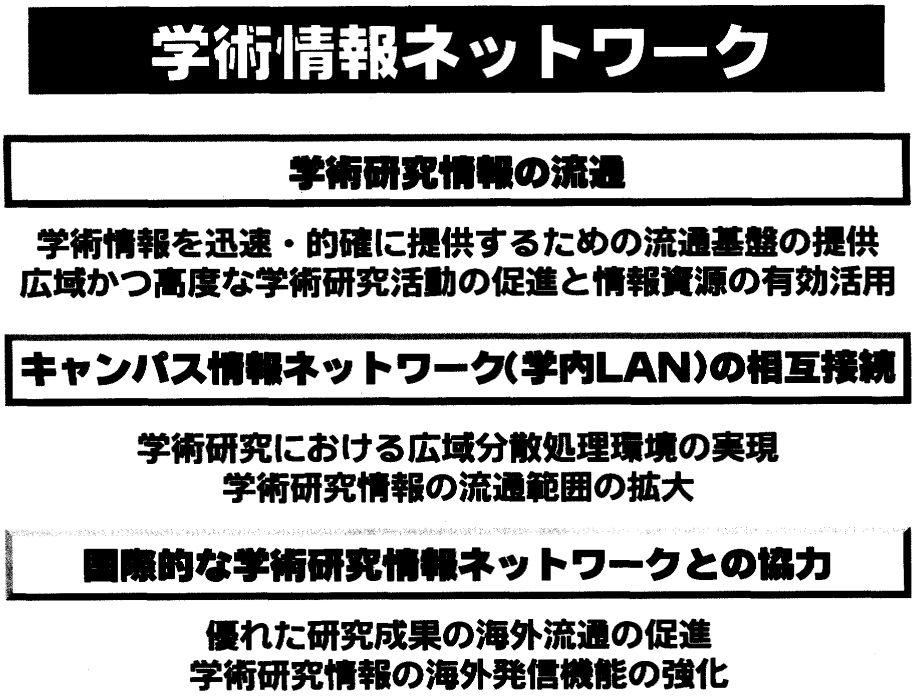


図30　学術情報ネットワーク

演題　「わが国のマルチメディアネットワークの将来」

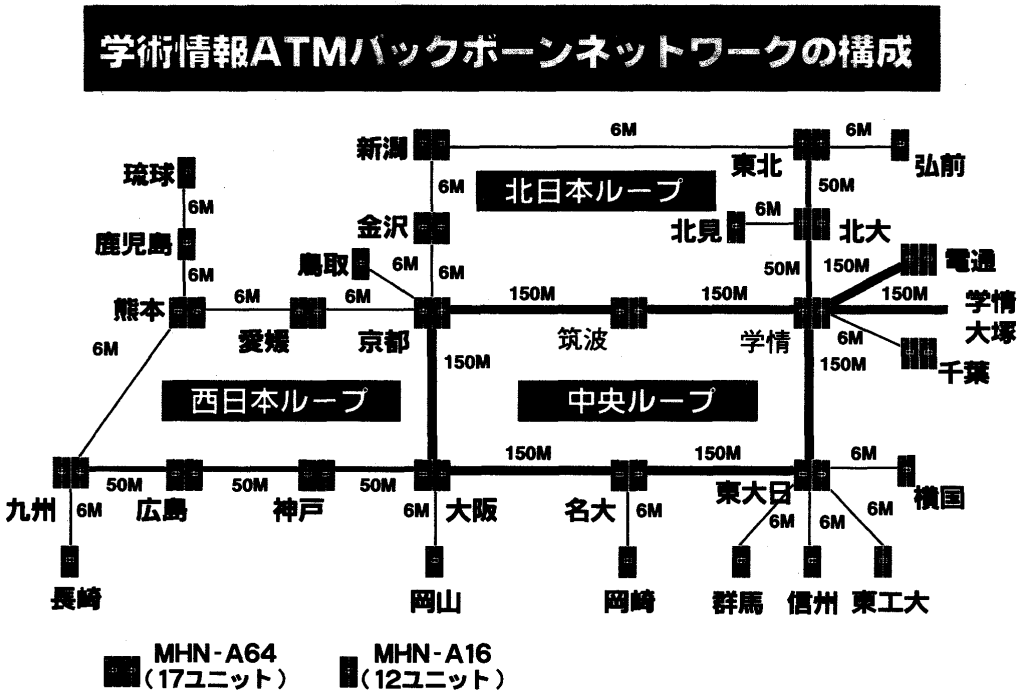


図 31　学術情報 ATM バックボーンネットワークの構成

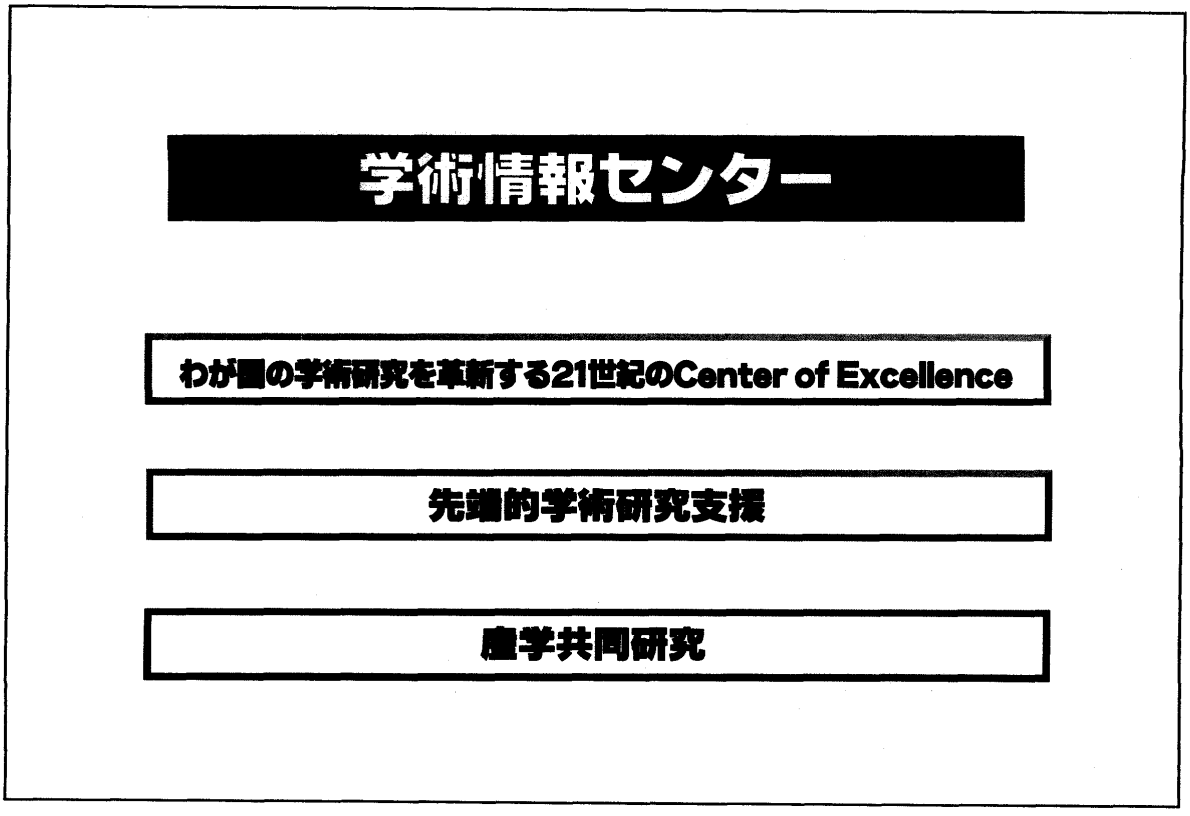


図 32　学術情報センター

学術情報センター電子図書館サービス　　◎学術情報センター

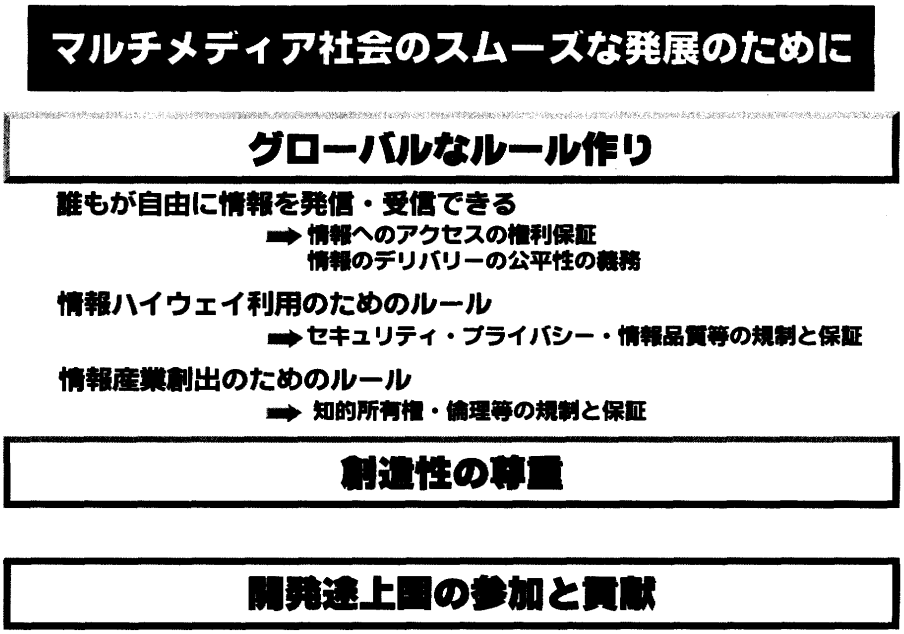


図33　マルチメディア社会のスムーズな発展のために

学術情報センター電子図書館サービス　　◎学術情報センター

学術情報センター紀要 第9号 (1997)

展望

情報化社会の国際化と日本語

Globalization of Information Society and Japanese Language

中京大学情報科学部　山田　尚勇
Hisao YAMADA
School of Computer and Cognitive Sciences, Chukyo(Tyûkyô)University, Toyota, Japan

要旨

現在世界的な規模で進行しつつある、社会の情報化と国際化に伴い、日本語とその表記法とは、いま新たな問題に直面している。すなわち、言語学的には世界でもっともやさしい部類に属する日本語が、世界でもっとも複雑な表記法を用いているが故に、外にはなかなか国際的に受け入れられず、内には情報化の出発点となる機械可読化、すなわち入力が複雑で非能率、かつストレスの多い作業となることである。

本稿では、長期的な課題として、いかにして日本語の表記を国際化するかについて、また短期的な課題として、いかにして現在の表記法のままの文章を能率よく、かつ楽に入力するかについて、考察する。

ABSTRACT

Because of the arrival of global information society, the Japanese language and its writing system now face new challenges. In spite of the fact that the Japanese language belongs to the structurally simplest category in the world, it has the most complex writing system, and, therefore, it is not readily embraced in the world theater, while it is also cumbersome, inefficient and stressful to render into machine readable form for processing.

In this note, we shall examine some possible means to make the writing system more globally acceptable as a long-term prospect, and also some possibilities to streamline its input process as a short-term task.

[キーワード]　外国語、キーボード、国際化、教育、情報化社会、日本語、入力法、表記法、平易化、文字
[Keywords]　foreign language, keyboard, globalization, education, information society, Japanese, input method, writing system, simplification, character

学術情報センター電子図書館サービス　　© 学術情報センター

目次

要旨 …… 33
1　情報化社会の到来と日本語 …… 35
2　英語に流される日本語 …… 36
3　ことばと表記の違い …… 37
4　ことばの民衆化 …… 38
5　復古思潮と国際化指向のはざまで …… 40
6　日本語のあり方を支える理念 …… 41
7　漢字制限をなくそうとする動き …… 42
8　表記法について何をなすべきか …… 43
9　漢字優越の先入主 …… 44
10　漢字の不具合 …… 46
11　和語の圧迫 …… 47
12　異体字などの問題 …… 49
13　ローマ字表記の問題 …… 51
14　文字と思考形態 …… 53
15　誤字はなぜ起こるか …… 55
16　意味の拡張性の欠除 …… 56
17　思考の歪曲 …… 57
18　日本文の入力処理 …… 59
19　入力法について何をなすべきか …… 61
20　キーボードの問題 …… 63
21　入力に関わる学校教育問題 …… 64
22　おわりに …… 66
謝辞 …… 68
参考文献 …… 68

学術情報センター電子図書館サービス　　© 学術情報センター

情報化社会の国際化と日本語

A man who has a shrine has transcended reason.
— Rex Stout

1 情報化社会の到来と日本語

コンピュータ通信網が予想をうわまわる勢いで普及し、個人を世界的な規模で結ぶインターネットの利用者は、1996年の夏には1億人にも達したという推定が出たが、その後もとどまるところを知らない増加が続いている。

そもそもその基盤となっている、コンピュータと通信の技術の生まれたのがアメリカ合衆国だったこともあって、いまインターネット上で飛び交っている通信のほとんどは、だいぶまえから事実上の国際語と言われてきた英語によってなされている。

第2次大戦後の廃虚の中から再出発し、土地も資源もない日本が、科学技術力と独特の経営技術によって営々と築きあげ、やっと世界の最先端に肩を並べた日本の経済力が、円通貨を安定させ、その十分な供給と使いやすさによって国際性をもたらしたのは1980年代になってであったろう。

しかしそのころから、世界の経済力を左右する実体が、工業的に生産されたハードウェアである「物」に加えるに、ソフトウェアである「情報」へと大きく片向き(傾き)始めた。それとともに、いまや経済における一国の競争力は、投入できる資本のほかに、国際的な通用性を持った数多くの人材の動員力によって左右される時代となってきた。すなわち、他の国ぐにの異文化と付き合える抱擁力と、自由に意思の疎通を果たせる言語力とを発揮できる民衆の層の厚さが、一国の経済力を決める重要な要素となってきた(山田 1984a[63])。

そうした国際情勢の中にあって、残念ながらわれわれの言語日本語は、互いに関連している二つの理由で、われわれを不利な立場に追いこんでいる。すなわち第一に、情報を蓄積し交換する第一歩としての、文章の機械可読化、つまり入力がめんどうなことと、第二に、日本語が外国人にとって甚だむずかしい言語とされていることである。

まず入力問題であるが、よく言われているように、アルファベットで国語を表記している国ぐにでは、基本になっているのが30たらずの文字で、大文字小文字、それに数字と句読点などを合わせて100内外の異なり字数ですんでしまう。

しかるに日本文では、片かな平がなのほかに、常用漢字だけに限ってみても、すぐに2000字ほどになってしまい、各文字を一対一のキーボードで入力することはきわめて繁雑になり、それを避けようとすると、たとえばかな、もしくはアルファベットで文章を表音化して入力し、それをコンピュータ内のきわめて大きく複雑なプログラムで変換処理をして漢字に戻さなければならない。それでも同音異義、つまり同音異字の語が多すぎて、いまのプログラムではまだかなりの誤変換が出るので、入力時には出力文章の表示スクリーンから目を離せず、また誤変換の訂正には時間をとられ、煩わしい。

したがって、たとえば英語の国では1分間に70語を入力する人はかなりいるが、漢字かな交じり文となると、ほぼ同じ情報量となる、1分間175字を変換法で入力することは、まず不可能である。

そんなぐあいだから、日本のオフィスでの文書の相対的な生産性は悪く、中には極端に、欧米の10分の1以下とさえ言われたこともあったが(岩村 1966[10])、いまでもアメリカの場合の60パーセントぐらいだとされている。

しかも、せっかくそうして漢字かな交じりの日本文を作ってみても、それを読みこなせる外国人の数は限られてしまうから、外国への発信はどうしても英文によることになる。アメリカの大学などに留学する際に求められる、英語能力の試験TOEFLの成績データによると、日本人の英語力は諸外国にくらべてかなり低いようだし、そのうえ「日本製品は優秀だが、付いてくるマニュアルは冗談だ」、「紙屑だ」などと言われたりするように、文章の作成能力も悪いようだから、国際社会における日本人一般のコミュニケーション能力はかなり低い(山田 1986b [67]参照)。

そんなこともあって、日本のあるビジネス誌が、1995年の夏に、ネットワーク時代の競争力について特集号を組んだときには、「問題は日本語だ」と表紙に大書したほどに、自からの日本語をさげすむようなことになってしまう(日経ビジネス 1995[45])。

もちろん日本の国際化をむずかしくしているのは言語の問題だけではないし、また経済における能率化を阻んでいるのは入力の問題だけではない(たとえばDeutsch 1983 [84] 参照)。しかし、日本語が世界的に広まり、かつ入力が容易になることによってコミュニケーションが活発になれば、そうしたほかの問題の原因となっている日本文化の特殊性も、国際的にかなり理解されてくることも、また事実である。

学術情報センター電子図書館サービス　　©学術情報センター

2 英語に流される日本語

いずれにしろ、国語についてのわれわれの無意識の劣等感には、一般に抜き難いものがあるとみえて、むかしから使われてきて、すっかり定着していることばがある場合でさえ、その多くをわざわざ英語らしからぬ和製英語で置き換える現象が急速に進んでいる。

たとえば「パネラー」などはその最たるものの一つで、むかしの「討論会」がいつのまにか「パネルディスカッション」になったときに、「討論者」が「パネリスト」になり、その後それを勝手に縮めて「パネラー」にしてしまったもののようである。英語では「paneler」は滅多に使われないことばで、たとえば審議会などの「審査員」の意味があることはあるが、日常の用法ではまず「羽目板職人」と受けとられる。

また「ベビーカー」や「ベビーサークル」などもその例である。「うば(乳母)」のいなくなったいま、「うば車」ではおかしいというのなら、「赤ちゃん車」など、もう少し別の言い方がありそうなものだと思うが、やはり英語もどきでなければ嬉しくないのだろうか。近ごろ広告などでよく見かけるようになった、住まいや洋服の「リフォーム」も、英語としてはおかしな用法で、「reform」は形のない制度や行ないなどを改めることを言い、家などの改造には「remodel」など、別のことばを使う。

そのほか、日本で発音を勝手に変えてしまったものもたくさんある。たとえば「pouch」は「パウチ」であるのに「ポーチ」(=porch、車寄せ)にされてしまった。また、ある地方テレビ局において、最近 (1997 年 3 月 15 日) の深夜に放映された映画の題名が「ペンダラム」と珍しいものになっていたので、何のことかと好奇心から覗いてみたら、なんと原題は「pendulum (ペンデュラム)、= 振り子」であった。どうして辞書を引いてみることぐらいしなかったのだろうか。しかし場合によっては「news」のように、誤った「ニュース」から、「ニューズレター(newsletter)」のように、場合によっては正しい「ニューズ」に戻されつつあるものもある。

こうした和製英語の例は数限りなくあるが、ここで言いたいことは、こうした造語は日本でしか通用しないのだから、それが本ものだと思って国際舞台で使うことは、たとえば漢字の共用が漢字圏の国際交流に資するものだという短絡的な考えなどに似て、実はかえってわれわれの国際化を阻害しているということである(山田 1994a[72])。

それでも、これなどはまだいいほうかもしれない。あるFM ラジオ放送局などでは、日本語と英語の発話をチャンポンに混ぜる司会者が、四六時中いかにも得意気にまくしたてているようで、いったいどんな人たちが聴いているのだろうかと思ってしまう。多民族国家のアメリカでは、当然放送にも数多くの言語が使われているが、それぞれは独立して用いられているのであって、このようにまぜこぜのものはない。国際化とは全く無縁の行為である。(ちなみに「チャンポン」は中国語の「chānhùn」からか。「chān」の漢字は JIS 規格に入っていないが、「讒言」の「讒」の言偏を手偏に置きかえた字で、「まぜる」の意。「hùn」は「混」。なおこれはマライ語の「campur」、かつての英国統治時代のつづりでは「champor」、意味は「まぜこぜ」、ではないかとも言われているが、このマライ語は中国語から出たとも考えられる。なお沖縄では「チャンプルー」がこれに相当することばである。)

1943 年(昭和 18 年)に従来のものを改編し、その後 1945 年(昭和 20 年)まで使われた、ある国文法の教科書は、冒頭が「国語は我等の言葉である。祖先伝来の言葉である。我等の祖先は、この言葉によって、神の声を聞き、わが国がらをたたえ、わが国土をほめ、君の栄え・民の誉まれを歌った。(中略)したがってこの言葉は、心なく用いられてよいものではなかった。」と始まり、全冊を「国語に輝きあらしめよ。美しく強く、明らかにゆたかに、まさしくまことのあらわれたらしめよ。純正な国語への努力こそ、我等の忠君愛国の最も深い基礎であることを、我等は寸時も忘れることはできないのである。」と結んでいる(現代表記化した) (教育総監部 1943 [25])。時代が時代であったので、当然表現には時代思潮の反映が見られるのはやむを得ないと思われる。しかし本筋となっている、国語に対する熱い思い入れは、今日でも国語を尊ぶ者の共感を呼ぶであろう。それはまた、日本以外の活力ある諸国でも、それぞれ当然のこととして支持されている考え方であるし、正直のところ、筆者もこの文章に触発されて本稿をまとめる気になったほどである。

ことばは生きものであるから、時代の要求するところに応えて変化をするのは当然である。しかしその変化には理念と秩序がなくてはならない。さもなくば、ことばは弱体化し、遂には消え去ってしまうことになるであろう。それがまた、しばしば文化と民族の消滅につながっていることは、歴史が数多くの例を示している。いまの日本には、そうした自覚が欠けていないだ

学術情報センター電子図書館サービス　　©学術情報センター

ろうか。

最近の日本における極端な英語礼賛の風潮を憂えて書かれた津田(1997[41])の現状分析といくつかの提案は傾聴に価するであろう。その津田が国際コミュニケーション論が専門の教授であることを考え合わせると、さらにその感を深くする。それに対するソーントン(1997[36])の反論は比較文学、米文学が専門の教授によるものであるにもかかわらず、残念ながら日本語の在り方という根本問題には、ほとんど答えていないと思う。

民俗学者柳田国男は、かつて「ほかの民族に支配される最もたまらない苦しみは、自分のことばをたいせつにして、これを発達させて行かれないこと」(原文はローマ字書き)だと書いたが(柳田 1925[59])、いまやわれわれはそうした民族としての誇りを自から投げ捨てようとしているのではないだろうか。

彼はまた、もしわれわれが日本語とその表記法をもっとやさしいものに改めていかなければ、遂には日本人は和語を捨て、むしろ欧米語の便利性を認めるようになるかもしれないことをも心配していた。最近の英語もどきの盛んな風潮は、その恐れを現実のものにしつつあるのではないかと思える。

日本語のそうした現状は、1928年に国際的な要請を受けてから、日本の政府及び専門家が、紆余曲折を経て、60年の歳月をかけて国際規格として採用させた、日本語のローマ字つづりの標準(Tatuoka & Yamada 1996 [94] 参照)を、内閣訓令に背いてまで退け、いまだに英語もどきのヘボン式に固執している外務省の態度の中にも見られるであろう。

そのような自国語、ひいては自からの文化の軽視があるからこそ、外務省は、ごく最近のことだけとっても、橋本竜太郎総理がマレーシア訪問まえの1997年1月7日に、英字紙だけに全面広告を出し、当然ながらずっと発行部数の多いマレー語紙は無視することにより、マレーシア人の心を傷つけることになるのに気づかなかったり(朝日新聞 1997[3])、また1996年暮れには、米軍機が那覇市沖に爆弾を投棄した事件で、米軍から通報を受けた外務省が、安全性が確かめられるまえに国民の保護問題を棚に上げて、海上保安庁に「公表を見合わせて欲しい」と、米国寄りの態度で要請したとのことである(朝日新聞 1996b[2])。まさに植民地において下僕が宗主国に対して持っていた精神構造なみである。われわれは、外国に対して主張すべきところは主張すべきであって(日下 1993 [26])、国語問題もその一つである。

本稿では以下に、これからいよいよ情報化し国際化する社会活動全般において、日本語を活力ある言語として駆使し、かつ国際舞台の場で世界にも認めてもらうための方略として、われわれ日本人がともすると苦手とする、ものごとの原点に立ち返り、長期的には表記法の問題、短期的には現在の表記法による文章の入力問題を、どのように考えていけばよいかについて、私見の一端を述べてみることにする。

3　ことばと表記の違い

それに先立って、まずはっきりさせておくことが必要なのは、ことばとその表記とは、互いに深く関わりあってはいるものの、本来は別ものであるということである。このことは、近代にいたるまで、いずれの国においても、充分に文章の読み書きができたのは国民のごく少数派であったが、それでも日常生活においては、誰もがことばを自由に使っていたことからも明らかである。

しかるに、日本語の問題を論ずるとき、相当の知識人のあいだにさえ、この二者を混同している者は多い。そうした、ことばと表記法との混同には抜き難いものがある。たとえば1996年9月26日の朝日新聞夕刊の「きょう」欄には、ある著名な作家はワープロも使うが、死んでも文章は日本語で書きたいとローマ字入力を拒否し、目隠しをして文字盤のカナ配列を覚えたそうだと紹介している。したがってこれを書いた記者(川村二郎)も、やはりこの二者を混同していることになる。

もっともこの作家は1986年5月に日本で行なわれた国際シンポジウム「漢字文化の歴史と将来」における討論会では、朝日新聞で伝えられたように、ワープロなど死んでも使いたくない旨の発言をしていたのだから、その点では大幅な進歩をとげているのだが。

同じようなことは技術者のあいだにもある。たとえば1995年4月11日の日本経済新聞に出された、「ローマ字変換入力をやめて、よりやさしいかな入力変換入力を可能にする」という、ある新しい文字配列のキーボードと、その接続のためのインタフェース・ソフトの広告では、「ラベルとソフトでキーボードを改良、ローマ字帝国への挑戦。日本人はローマ字で考えているわけではありません」と謳っている。

そのほか、たとえばよく聞かれる「日本語はむずかしいことばである」といった表現にも、この混同が無意識に出ている。

学術情報センター電子図書館サービス　　© 学術情報センター

たしかに、日本語の学習について、そうしたデータを示した報告はある(加藤 1979 [22] 参照)。一例として、アメリカ合衆国において、外交関係の人びとの教育を行なっている Foreign Service Institute では、教育の困難さの度合に応じて、外国語を 4 種類に分類している。最もやさしいのはフランス語、ドイツ語、スペイン語などの第1類で、最もむずかしいのはアラビア語、中国語、韓国語、日本語から成る第 4 類である。各言語とも熟達度を最低 0 から最高 5 の 6 段階に分けてあって、その第 2 段階は「何とかできる(barely managing)」能力を示している。この第 2 段階に到達するのに、第 1 類の言語では平均 480 時間の授業(contact hours)で済むが、日本語では平均 1320 時間の授業が必要である。しかしその大部分は読み書きの能力の習得に費やされているのだという(AAS 1995 [79])。つまりこれは、日本語学習の相対的なむずかしさが、主として漢字かな交じり文という表記法と、それに付随して起こってくる、あまりにも多い同音異義語の存在にあることを言っているのであって、日本語自体の言語としてのむずかしさを言っているのではない。

言語学者の指摘するところによると、世界に何千とある言語の中で、日本語の文法構造は、多数派を占める、むしろ単純なものに属している。また音韻構造も単純であるし、それに語囲(語彙)や運用論などは、どの言語にしても、習得に本質的な差はない。ただ、うえの第 1 類の印欧語のように、似た語囲を持った、同じ仲間の言語がほかにないことは、少しは習得の難易に影響するであろうが。

具体的な例としては、近ごろはわれわれのまわりにも、日本語をかなり流暢に話すのに、読み書きは全くだめという外国人がかなり増えている。それは印欧語系言語の話し手に限られていない。たとえば、第 2 次大戦中にビルマなどで日本語を教えさせられた詩人、じっこく・おさむ(1987. 十国修、本名草薙正典[34])は、かつてラングーンで日本語を教え、その後かれが重爆撃機で日本に留学させた教え子の一人について述べている。その後の教育で彼は、英語、フランス語、日本語を自由に操れる外交官となり、大使としても活躍したそうであるが、その彼が、話して一番やさしいのが日本語、書いてひどくむずかしいのも日本語だ、といっているとのことである。

そのほか、科学系のある専門分野において、研究論文が世界的にどのくらい読まれているかを示す指標の一つとして、Science Citation Index(SCI)というものがある。これはある論文が、発表後のある期間に他の論文によって引用された回数である。そして、ある学術雑誌が掲載した論文全ての SCI を 2 年間にわたって平均したものを、その雑誌のインパクト・ファクタと呼ぶ。ところがこれによると、日本の学会誌は英文で発行されているものでさえほとんど引用されていない(山田 1989[69])。

たとえば、1987 年に世界で出た最高のインパクト・ファクタは約 35.1 であったが、インパクト・ファクタが 0.001 以上の 4091 誌の中で、日本の雑誌の第 1 位はやっと 557 番目に 1.851 で登場する英文誌である。そのほかインパクト・ファクタがちょうど1になる第1250番目までに、日本誌は数誌あるが、ことごとく英文誌である。

ところが研究水準が日本以下の小国で、英語以外の言語で書かれた雑誌が、ここまでのリストにかなり入っている。筆者の経験によると、こうした彼我のあいだに格差の出る大きな原因は、どうも、日本の論文自体はいくら英文で書かれてあっても、それと一体として参照されるべく引用されてある文献の多くが、非アルファベット表記の日本語のものなので、それらの文献は内容の見当がつけられないから、かなりの人たちが、論文自体を読んでも無駄だと考えてしまうことにあるようである。

つまり、いま日本語の国際化を妨げているのは、ことばとしての日本語ではなく、その表記法にあることを、まずはっきりと認識すべきである。同時に、外国人にとってむずかしいものは、同じ人間である日本人にとってもむずかしいものであると悟ることは、日本語の国際化について考えることの第一歩であろう。

そして、世界的にやさしい日本語が、どうして世界で最もむずかしい表記法をとるようになったかを知るためには、その歴史的な流れを眺めることから始めるのがよいであろう。

4　ことばの民衆化

すでに述べたように、ことばが生きものであり、時代とともに、また地域や職業とともに変わってくることはよく知られている。あとに出てくるいろいろな記述を統一的にみる枠組みの準備として、ここでちょっと述べておくと、そうした言語の変化を、時代に沿った変化の体系としての「通時態」と、ある時期における、地域あるいは社会層などによる異なりの体系としての「共時態」という名のもとに論じたのは、しばし

学術情報センター電子図書館サービス　　©学術情報センター

ば近代言語学の祖ともいわれる、スイスの言語学者ソシュール(Ferdinand de Saussure, 1857-1913)であった。

そうした通時態の枠組みの中でみると、ヨーロッパの諸国においては、中世紀までの社会でラテン語が上位のことばであり、民衆の用いていた地域語は一段と低いものとされていた。しかし14世紀以降、ルネサンスに始まり産業革命を経て、地域語の共有が民衆に国家の意識を形成させ、地域語が国語として確立された。それは19世紀初頭までにほぼ安定化してしまい、もはや時代と共にたいした変化がみられなくなっている。

たとえば17世紀末に活躍したドイツの哲学者・数学者・政治家ライプニッツ(Gottfried Wilhelm Leibniz, 1646-1716)はラテン語で著作を行なったが、18世紀初頭、彼の思想を整理し体系化してライプニッツ・ウォルフ学派を開いたウォルフ(Christian von Wolff,1679-1754)は、ドイツ語によって著作を始めた思想家ということでも知られている。

西洋におけるそうした国語中心主義化の過程は、いまからほぼ百年まえまでに収束した。ことば、特に語囲の細かい変化はその後にも起こってはいるものの、表記法がかなり安定してしまっているから、百年まえに書かれたものをいま読んでみても、そんなに苦労させられることはない。

たとえば、いまイギリスのオックスフォード大学ボドリアン図書館に保存されている、1821年に行なわれた、国王ジョージ4世の戴冠式の儀式の次第書きの巻頭言などを読んでみると、平易で、昨日書かれたものと言われても、違和感はないであろう。

たしかに古いものを読むと、たとえば「four-in-hand(四頭立て馬車)」のような、いまでは使われなくなった乗りものの呼び名が出てくるが、これが分からなかったとしても、それはもの自体が身のまわりからなくなったからであって、ことばそのもののむずかしさや、あるいは表記法のせいではない。

そのように、かつて地域語から国語へのそうした収束が進んでいた時代には、ことば間の国際性を保つなどということはそれほど問題にされなかったから、ヨーロッパの言語のほとんどが属している、同じ印欧語族内の各国語を、いま共時態の枠で眺めてみると、類似の音素が全く異なるつづりで表わされていたり、また同じつづりの発音が全く異なっていたりしている。しかしいまとなっては、これらの違いを解消して、異国語間でつづりの統一をとる動きが出ることなどは、まず考えられないであろう。

日本では大和時代初頭、つまり4世紀半ば以後に、それまでの土着文化に比べて圧倒的に優勢な中国文化がもたらされたとき、その担い手としての漢字もともに受け入れたので、その後は、欧米におけるラテン語に対比されるべき漢語・漢文が上位のものとして貴ばれることになった。ずっと下がって18世紀、江戸時代の中期になって、それまでの中国朱子学から離れて日本朱子学へと向かい、さらに国学の台頭とともに培われていったナショナリズムが、やまとことばへの回帰の志向を示し出した。しかしかつてのヨーロッパにおけるように、自分たちのことばでものを書くという言文一致の運動が盛り上がるのには、19世紀末、西洋を模範として国の産業化が始まった、明治の中期まで待たなければならなかった。

しかしながら、ものごとの細部にこだわることには熱心だが、全体を大きくとらえる理念を明らかにすることには苦手な、つまり哲学性の弱い文化を持った日本(たとえば 渡辺 1949[78])は、明治中期には、政策として中国文明から離れ、西洋文明に近づく過程をたどっていたにもかかわらず、その結果として欧米から入って来る、あまりにも多くの新しい概念を表現するためのことばを作り出す必要に迫られるや、旧態を頼むという知的怠惰に落ちいり、漢字による新語をやたらに造ることに走った結果として、せっかく動き出していた、民衆のことばを正統化する流れには歯止めがかけられてしまった。

やまとことばは造語力に乏しくて、新しい概念に対応した新しい表現を作る力に欠けていたのだから、しかたがなかったのだとよく言われる。しかしそれは、かつてヨーロッパにおいて、ラテン語に対して民衆語を使おうという運動が起こったときに、時の体制派がしばしば口にした反対理由の一つだったということに似ているようである。

おそらくルネサンス以後のヨーロッパにおける言語改革のときにもそうだったのだろうと思うが、一般に何か革新のさなかにいる人たちの多くは、それに明確な理念を見いだせず、ただとまどうばかりであることが多い。それが今日の日本語やその表記法の改善をも大幅に阻害しているのであろう。

忘れてならないのは、こうした国語化運動は、実は、主観に基づいて築き上げられた権威に頼る他力本願の思想から、ルネサンス以後のヒューマニズムに基づく、現実の世界での実証に立脚して個体性を生かすという思想への転換の気運に刺激されて、16世紀から17世紀

学術情報センター電子図書館サービス　　◎学術情報センター

にかけて発達した、F.ベーコン、コペルニクス、ガリレオなどからニュートンなどに至る人たちの科学観や宇宙観と無縁ではなかった。それがいまの世界では、エレクトロニクスの発展の成果として情報化された社会の出現が、社会の通念をはじめとする、既成の文化の在り方や価値観に対して、逆に情報本位制を基盤とした大幅の再構成を迫りつつある。すなわち、もろもろの既存の聖域が、果たして理性を超越して築き上げられた錯覚ではなかったのかということの再検討が迫られていることになる。

5　復古思潮と国際化指向のはざまで

現代言語学の教えるところによると、どんな言語でも新しい要求にこたえて新しい表現を生み出す柔軟性と強靭性とを持っているということだから、わが国において似たようなやまとことばへの回帰運動が19世紀末に活発にならなかったのは、やはり上からの国語政策が大きく影響したと言えよう。それはよく指摘されているように、明治の新政府がとった、上からの早急な財閥育成政策によって、江戸時代から少しずつ成長してきた草の根的な資本主義生産が押しつぶされ、同時に、本来は資本主義とともに起こるべき、人権、個性、自由などの尊重という意識革命が置き去りにされてしまって今日に到ったことと、無縁ではないだろう(日下 1993 [26] 参照)。

言語問題に限ってみても、たとえば何千年ものあいだ、世界の桧舞台ではほとんど忘れられていたヘブライ語が、イスラエル国の再建とともに、古代ヘブライ語を基とした造語運動によって活性化され、一流の科学論文が書けるだけの活力をとりもどし、国際的学術活動の中心にまで躍り出るようになったという対比が、そのことをかなり裏書きしているであろう。

漢字・漢語中心の国語政策の結果、むずかしい漢字を多く使った文章がその後も依然として幅をきかせ、また音韻的変種に乏しい漢字によるやたらな造語が同音異義語の氾濫をもたらし、口頭によるコミュニケーションを複雑なものにしている。したがって、主権在民の民主主義社会におけるコミュニケーションの道具としてふさわしいことばや、その表記への模索は現在も続いており、いままだ日本語とその表記は、いまの欧米諸国ではみられないほどの速さで変わりつつある。

さらに、航空機をはじめとする交通手段や、エレクトロニクスの進歩に伴うコンピュータや通信技術の発展は、地球を相対的に小さなものとし、全ての国ぐにの人びとの活動を国際的に相互に依存するものとしたために、社会活動の全体が世界的に情報化、国際化へと急速に推移しつつある。その結果、かつてはほとんど日本国内に限られていた日本語の活用地域も、いまでは全世界に向けて広がってきている。

それに伴って、日本語の国際性の問題が取りあげられ、たとえばいま「国際社会における日本語についての総合的研究」というプロジェクトが、国費による創成的基礎研究の一つとして、文部省傘下の専門家たちによって真剣に取り組まれている(国立国語研究所 1996 [29] 参照)。

そうした研究の中心となる課題は、これからの全世界的な舞台において日本語の占める地位、そしてそれが果たす機能を明らかにすることにあると思われる。そのことに関して、ここで注目しておきたいことがいくつかある。

第一に、国際社会において日本語はむずかしいとよく言われるが、すでに述べたように、世界に何千とある諸言語の中で、文法や統語法の型からすれば、日本語は最多数派に属しているばかりか、むしろ易しい部類に属しているという、言語学者の指摘である。もっとも、日本人にはわざわざ日本語はむずかしいものと思いたがる母語観があって、外国人が日本語を自由に話すことに対する拒絶反応がしばしば見られる。しかしそうしたことを含めて、日本語がしばしばむずかしいと言われるのは、日本語ということば自体とその表記法とのあいだを混同しているからであって、いわゆる日本語のむずかしさというものは、実は漢字を使った表記法と、それから派生する同音異義語などが発話において引き起こす問題のことであろう。外国人が日本語を自由にあやつることに対する反発は、表記法の習得についての自からの苦労を軽くとられているような気がすることに対する無意識の反応であろうか。

その二として、自国語の国際性の良さといったテーマが大がかりな研究対象として取りあげられるということは、すでに民衆化したことばや表記法が安定してしまっている欧米の諸国では、ちょっと考えられないことであって、たとえばフランスなどでは、いま逆に、フランス語にまぎれ込んでくる、主として英語系のことばをいかにして排除し、純フランス語系のことばで置き換えるかに腐心しているほどである。

したがって、日本語の国際性といったことが研究テーマとして取り上げられ得ること自体が、実はまだ日本語やその表記が真に民衆化、安定化していないことを

学術情報センター電子図書館サービス　　©学術情報センター

暗黙のうちに肯定したうえでのことであり、またそれだからこそ、こうしたテーマが実行可能、かつ意義のある研究テーマとして扱われ得るのであろう。同時にそれは、ことばは生きものだと言いつつも、その通時態内での変化がレッセ・フェール・レッセ・パッセ(放任放往主義)だけにまかせられるべきものではなく、ことばの変化に方向づけをする影響力を、ある程度働かせ得ることをほのめかしていると思える。

6 日本語のあり方を支える理念

しかしながら、そうした積極的な、日本語の言語力の国際性増強が暗に意図されているにもかかわらず、第三番目として、そのために考えられている、各種の異なる努力の背景にあるべき、一貫した理念がいまひとつはっきりしていないと言いたい。

その片鱗として、たとえば使用頻度の低い漢字・漢語をよく知らない若い人たちについて、国語力が年々低下しているといった嘆きが年輩層からたびたび聞かされている(加藤・山田 1972 [21]、成瀬 1996 [43])。しかしその原因の一端となっているのは、1946年に政府が訓令・告示した当用漢字、さらに1981年に告示した常用漢字、また1990年の人名用漢字などの普及や、現在の日本語の発音に沿った新かな使いなどの導入によって、日常の文章が平易化したことにより、これら以外のむずかしい漢字には、さほどお目にかかれなくなったことである。すなわち、こうした苦情が出てくるのは、実はヨーロッパでのラテン語に相当する、漢文やその亜流が上位言語であったそれまでの状態からの脱却運動が、いまもまだ続いているからだと言える。

いまの日本において19世紀末、明治初期の文章がすらすらと読めないのは、たとえばいまのイギリスにおいて14世紀のチョーサー(Geoffrey Chaucer,1340?-1400)の作品がすらすらと読めないことぐらいに相当するのだから、そう気にすることもないであろう。

ことばやその表記の第一の、そして最も重要な目的は、それによって表現された意味内容を正しく理解し、かつ自からの考えをそれによってはっきりと表現することにある。すなわち、国民の国語力とは、全国民におけるそうした意味内容の授受の能力の総和であり、簡明になった現在のことばや表記は、日常において新聞・雑誌や書物を読む人口やその読む量、また日常において各種の文章を書く量およびその人口を格段に増加させた。その意味において、新しい表記法は国民の国語力を、日本の歴史始まって以来、いま最高の点にまで高めていると思える。

むかしに比べて現在の国民の国語力は格段に進歩しているが、その一方で大衆はおろか、相当の知識人でさえ、たとえば義務教育において学んだはずの物理や数学などの科学的知識の活用能力のほうは、まだかなり劣っていると言っても言い過ぎではあるまい。詳細については稿を改めて述べることにするが。

第2次大戦中の軍隊については、「帝国陸海軍の名誉のために」という思いやその他の理由によって、不都合だったことが今にいたるまでしばしば隠されたままになっている。にもかかわらず、たとえば以下のようなことがすでに明かるみにでている。

すなわち、当時は相対的に質が高いと思われていた航空隊の兵隊の中にも、字の読めない者がかなりいたという。そのため、航空機のエンジンの整備がゆきとどかず、せっかくの機体の稼働率が低くなってしまうようなことは珍らしくなかったようである(たとえば保科 1949 [52]、碇 1977 [7])。しかも新機種が配備されるときには、アメリカ空軍では前線の整備隊に対して、機体にマニュアルを付けて送ることで済ませられたのに(たとえば Caidin1966 [81])、文書作成能力の低かった日本では、充分なマニュアルが準備できず、また整備兵のマニュアルの読解力も充分でなかったので、工場で新しい整備兵を教育しては、機体と一緒に前線に送り出してやるという非能率的なことをしていたという。ちなみに第一線のイギリスの整備兵には、部品に分解して送られてきた新しい航空機を、マニュアルをもとにして組み立てることができる者たちさえいたようである。

そのほか中国戦線では、伝令のもたらした、兵器の破損部品の補充要求文書が正しく読めなかったために、間違った部品が送り届けられ、兵器が使えず、部隊が全滅させられたという話もある(保科 1949 [52]、塩田 1980 [32])。

当時、兵器の部分名称だけでも4000以上に達し、日常語や和語を用いればなんでもないものが、たとえば「チューラ(＝ボルト)」「ホーテツキ(＝ミシン)」「トボクキ(＝くつブラシ)」「キューガクレーキャクユソクトウ(＝ピストン冷しの油ポンプ)」といった漢語が用いられていた。いまなら大学卒業生とても、こうした漢語を読んだり書いたりできる者は少ない。したがって客観的に見れば、上に述べたような事態が起こったのは兵士の責任ではなく、コミュニケーションの道具としてのことば、さらにはその表記法が不適格であった

学術情報センター電子図書館サービス　　©学術情報センター

と言うべきである。

そうした事実は大戦直後の国民について行なわれた、読み書き能力調査委員会(1951) [77]の報告書の詳細な分析によっても明らかにされている。

したがって、第2次大戦中のそうした経験の末には、当時の軍隊からさえ、使用漢字数の制限と発音的かな使いへの改訂の声があがり、兵器名称や用語のための漢字を 1235 字に制限する通牒が出され、さらにゆくゆくは漢字を 600 字内外に切りつめるべく研究が進められていた。しかし、自身では戦場におもむかなかった軍国主義者や、当時の敵国である中国の文字に固執していた超国家主義者の強い反対に手を焼かされ、思うにまかせられなかったのである(保科 1949 [52] 参照)。天皇家の先祖が中国から渡来したのではないことは、ほぼ明らかにされていたにもかかわらず、やまとことばを退け、漢字に固執した超国家主義者の主張の意図は明らかでない。それは今日についても言えることである。

こうした歴史的事実があるにもかかわらず、国語表記をやさしくしようとした戦後の国語政策などが、占領軍の押しつけであったする、もと文化庁長官の理解(安島 1995 [58])はどんなものであろうか(Unger 1996 [98])。

7　漢字制限をなくそうとする動き

もっとも、詳細な学問的研究に裏付けされた Unger や Gottlieb(1996 [86])たちの指摘をまつまでもなく、日本における表記法の問題が純粋に国民の国語力の向上や文化的継承だけの立ち場から論じられたことはまれであり、その背後には必ず政治的な思想の対立がつきまとってきたことは明らかである。

表記法の簡易化を推す人たちが、自身では漢学に長けていた前島密や福沢諭吉などをはじめとする、民主主義的な思想の持ち主であったことはうなずけるが、それに反対する人たちは、しばしばむずかしい表記の維持によって、実は意識的あるいは無意識的に民衆を隔離しようとする、保守的なエリート思想の持ち主であった。

にもかかわらず、民主主義時代が到来するに及んで後者は、言語は民衆に属するから、政治が表記を制限するのは民主主義の精神に反すると主張している。(意図的なものかどうか、ここには言語と表記法との混同がある。)しかしそのかれらは、たとえばそれ以上に強権的と思われる、学習指導要領の存在や教科書検定制度に対しては、なんらの反対表明もしていない。

以下においては、こうした政治的な問題はいっさい抜きにして話しを進めることにする。

最近の朝日新聞の夕刊(1996 年 11 月 16 日)の「きょう」欄(川村二郎担当)には、漢字の大家某の、「漢字は覚えなくても、体を通過させておけば役に立つのですから、吸収力の旺盛な子どものうちに、たくさん食べさせておくことが大切なんです。漢字の知識は日本語の骨、カルシウムですからね」という言や、教える漢字を制限するのは誤りであるという意見を引いて、日本の現状の一因が漢字を粗末にしてきたことにあるかもしれないといった、片寄った考えを述べている。

しかし、義務教育時代にそんな時間的ゆとりがあったなら、もっと理科や数学をこそみっちりと身につけさせるべきであろう。義務教育で習った理科や数学などは、社会に出てから、ちっとも役に立っていないという声がよく聞かれるが、それは漢字は書けても、それを用いて実のある内容が書けるだけの、理科や数学の駆使力を持たないもの書きが多いからなのではないだろうか。

そうした中にあってむしろ問題になるのは、日常にますます多くの文字を使おうとする最近の動きである。すなわちコンピュータ用の漢字として、上記の漢字表とは独立した字体を含む、日本工業規格(JIS)による文字種および字体の標準が 1978 年に制定され、現在その 3 度めの改定が進行中である。現在の常用漢字は 1945 字、人名用漢字は 284 字であるが、1990 年制定の JIS/X0208 規格には 6355 の字種が含まれている。それでも特殊の分野や人たちからは、まだ文字種が少ないと、いっそうの拡張の要求が絶えない。

しかし、どうやら国語審議会にさえ、迷いの心が濃く、日本語や表記法の在り方について、一貫した理念を掲げることもなく、結果的には何十年もかけて、主としてただ各種の漢字表をいじって来ただけだったと言えないだろうか。そしてそのあいだに、国語政策の根底に大きく関わっている、漢字の JIS 規格がさっさと制定され、世に流通しているのを横目で見ているだけのようである。

現在では、これらの JIS 漢字だけがコンピュータで使用可能な漢字である。したがって、文書の作成、加工、蓄積などがコンピュータで行なわれ、電子メールなど、ネットワークによる通信が日常化しつつあるいま、これらの JIS 規格は国の国語政策に対して、いわば下剋上の相をなしていることになる。

それればかりか、この JIS 漢字表の制定についてさえ、字

学術情報センター電子図書館サービス　　©学術情報センター

数や字体の選択において、その理念が明らかでないという批判があるのに、コンピュータ発祥の地ながら、日常において漢字を用いていないアメリカのコンピュータ産業界の、アジアを目標とした、主として営業関係者たちのかなり強引な主導によって、その後、中国、韓国、日本などの漢字圏における漢字を含む、全世界の文字を16ビットで統一的に規格化する、国際標準化機構(ISO)の10646標準、正確には国際符号化文字集合(Universal Multiple-Octet Coded Character Set, UCS)、いわゆるユニコード(Unicode)が1993年に制定され、いちはやく1995年には、それがJIS/X0221規格としても制定されている。しかも、これには漢字が2万字ほど入っているのに、それでもまだ不充分だとする声が、日本をはじめ、漢字圏の国ぐにの一部の人たちから上がっているのである。

しかし漢語の表現の単純化としては、すでに「叡智」が「英知」、「窮極」が「究極」、「厖大」が「膨大」、「決蹶」が「決起」、「扮装」が「粉装」などと書かれるようになって久しいが、その伝でいくと、「語彙」なども「語囲」で充分と思われ、また一時はそう書いていた言語学者もあったが、学界には受け入れられなかったのは惜しいことである。同様に、「乖離」なども「解離」と書いて差しつかえないであろう。

そうは言っても、漢字の中には「央」のように、ごく限られた特定の文脈の中にしか出てこないにもかかわらず、しばしば使われ、しかもその一字だけではほとんどの人たちが正確な意味をよく知らないというものもある。こうした字はいまさら整理することは不可能であるし、ただ覚えて使う以外に手はないものとなっている。そんなことも、漢字の使用が日本語の表記を複雑にしている。

世の中には「面倒な」とか「草臥(くたび)れる」とかいった当て字や借字は平気で使うのに、上に示したような、同音で似た意味の代用字を使うのをひどく嫌う人たちがある。しかし文字の発達の過程では、表音化は自然の発展法則であり、それ自体は何も日本だけに限られた行為ではない。一例を示すと、すでに論語の述而第七には、詩経に由来する「暴虎馮河」の用字があり、「素手で虎を打ち、船なくして河を渡る」の意味とされるが、この4文字のうち、「暴」と「馮」は実は意味の異なる、同音の別字の代用字に過ぎない(山田 1984b [64])。この種の代用はむかしから中国でも数限りなく行なわれて今日に至っていて、われわれもそれらをそのまま受け入れてしまっているのである。

8 表記法について何をなすべきか

こうした国語の民衆化の要求やそれに逆行する動きとの取り組みにあたっては、少なくとも二つの方針が可能であろう。その一つは、ことばは生きものであり、時代の要求に従って変化するものだから、一歩退いて、そうした現象を客観的に記述し分析し、そこから上がってくる要求にそのまま応えていくことに終始するといった、観察の第三者として、Sein(存在)の立場をとることである。もう一つは、急激な技術革新がもたらしつつある、世界的な情報化・国際化のなかで、日本語が主要な国際語の一つとなれるためには、これからの日本語はどう発展させて行くべきかというSollen(当為)のために、実践の当事者となって、積極的に努力をすることである。

もちろん、前者の努力の積みあげがある程度はなければ、後者の作業の方向づけができないことは言うまでもない。それを承知の上でも、いまの日本語には後者が強く望まれていると思うし、またいまがその好機でもあると思える。

現在の日本語とその表記法はかなり民衆化してきたとは言っても、ことばの情報内容の伝達がますます重要になる、これからの社会におけるコミュニケーションの道具としては、まだまだ複雑でめんどうな点が多い。したがって、上に述べた意味において、国民の国語力がいっそう増大して行く過程では、これからも表記法の単純化、またそれにともなって起こる、ことば改めの道がたどられることになると思われる。繰り返しになるが、国際社会において、もし日本語、特にその表記が外国人にとってむずかしいものとされるのなら、それは同じ人間である日本人にとってもやはりむずかしいものであることは間違いない。したがって既成観念にとらわれず、それらを日本人にとって使い勝手のよい表記にするということは、そのまま日本語の国際性が高まることにつながっていくであろう。

そうした考えに基づき、ことばの従来の表記を変えようとすると、それがいかなる変更であるにせよ、伝統の破壊だとして、強い反対を唱える保守的な意見が出てくることは避けられない。しかし、冷静になって考えてみれば、日本語の表記は、古くは万葉がなに始まり、文字使い、かな使いはいままでに多くの変革を経て来ている。したがって、現在の表記法を変えていこうということに反対する心情の裏には、自分が幼いとき、若いときに習って身につけた表記法を捨てて、改めて新しいものを習得することを嫌うという精神的な

学術情報センター電子図書館サービス　　◎学術情報センター

惰性が、かなり大きく働いているであろう。しかもそれは大脳生理学的にも無理のない態度なのである(山田 1991 [71])。

しかし、もし人類がいままで全ての進歩を拒み、過去のしきたりに固執してきたとしたら、われわれはいまだもって裸ではだしで森の中を駆けずりまわっており、弓矢はおろか、棍棒や投石とかで獲物を狩り歩くということさえも始まらなかったであろう。人類の発展の過程における主要な発明のうち、おそらく文字の発明は、発火技術、石斧や弓矢などの道具類という、二つのハードウェア的な技術のあと、農業に次ぐ第二番目のソフトウェア的な技術の革新であった。したがって、古い文字は貴重な文化的遺産であり、主として学者によって記録され、受け継がれて行くべきなのは、古いハードウェア技術が博物館によって保存されていることと対をなす。同時に、文字は技術的発明の一つとして、時代の要請に応じ、絶えず改良され、更新されたものを使うべきであろう。

幸い、まだ日本語は、欧米ではすでに済んでしまった簡明化・民衆化の過程のただ中にある。この機会に後発者としての有利性を生かし、国際化に資する要素をも取り入れつつ日本語の民衆化を図れば、もともと言語学的にはやさしい部類に属する構造の日本語を、結果として世界でもっとも国際的なことばの一つとして育てていくことも夢ではあるまい。日本語の国際化と民衆化とは、まさに車の両輪であろう。

しかしながら、かりに将来に向かってはそれでよいとしても、過去に書かれたものが読めなくなり、文化や伝統の連続性がとだえるという、通時態的な問題のほうはどうであろうか。

それについても、そう心配することはないと思える。ことばは生きものであり、時代の要求に合わせて常に細かく変わっているものだから、いつの時代のものでも、時がたてば遂にはそのままで理解できなくなるのは仕方のないことである。しかし、ことばや文章の第一の目的は、書かれた内容情報の伝達にあるのだから、真に必要なものは、時代に合わせて書き改めればよい。

事実、われわれが外国で書かれたものを読むときには、ごく少数の人びとを別にして、外国語という「超共時態的」な隔壁を越えるために、ほとんどは翻訳で読んで充分間に合わせているのだから、同じ日本語のものを通時態の枠の中で読むのには、そうして書き改めることによって、目的は充分に達せられる。そうした書き改めの手間を省くために、より時代の要求に合った表記法の導入を拒むのは理にかなっていないし、また、そうした書き改めの手間に値しないものは、専門家や一部の好事家以外の人びとに忘れられてしまったとしても、大局的にはさしたる損失ではあるまい。なにしろ時間とともに、新しい書物はとても読みきれない量で生産されているのだから。

科学者や技術者などが、現在の繁雑な表記法による不便や困難にもめげず、苦労して生みだし、生産している工業製品など、文明の最先端の利器の恩恵を充分に活用している人たちの中に、ことが言語という、情報伝達の手段のことになると、情報化社会にもっとも適合した表記法の利便性と必要性を認めず、その採用を拒み、不便で非能率な旧套に固執して、それを全社会に押しつけ続けようとする、主として人文・社会系の人たちがいまだに少なからずあるのは、いささか身勝手な振る舞いだと言えるのではなかろうか。

9　漢字優越の先入主

すでに述べて来たように、情報化し国際化する未来の社会に日本語が適合し、国際的に活力を発揮できるためには、実はその表記法を平易にし、民衆化することが大事であろう。その具体策としては、すでに多くの人びとがいろいろな提案をしているし、また筆者も過去においていくつかの考えを述べてある。ここではその詳細について述べるゆとりがないが、その一部をかいつまんで記しておく(山田 1986b [67], 1991 [71], 1994a [72])。

いちばん基本になることは、漢字・漢語は貴く、やまとことばは俗なものであるという、長いあいだにわたって培われてきた無意識の先入主を見直すことである。この先入主はあまりにも根深く、有識者の多くが、現在ではもうそんなことはあまりないとさえ思っているようである。

しかし一例を挙げると、たとえば一般読者を対象とした新聞記事にさえ、いまでも「上梓」ということばが使われたりする(最近では朝日新聞 1996 年 11 月 10 日書評欄)。言うまでもなく、これは「出版」のことである。むかし書籍はあずさの板で版木を彫り上げて刷ったからこう書いたのだが、いまでは全く使われない技術であるし、それにほとんどの人たちはあずさの木を見ても名前が分からないだろう。にもかかわらず、こうした判じ物のようなことばがいつまでも使われているのも、それが漢語だからであって、もしこれが「あずさにのせる」のような和語であったとしたら、そん

学術情報センター電子図書館サービス　　◎学術情報センター

な言いまわしはおそらく使われ続けてこなかったであろう。

「てがみ」のことを「書簡」というのもその伝である。「簡(かん)」とは木や竹の札のことで、これは紙の発明されるまえに、木簡や竹簡に記録をとった時代の名残りにしかすぎない。いま「簡」に「てがる」「えらぶ」「はぶく」などの意味のあるのは、貴重な木簡や竹簡に文字を書くとき、全体をできるだけ短くする手順を踏まざるを得なかったことから出て来た二次的な意味である。また日本語の「かみ(紙)」ということばも、やはりこの「カン」から出たそうで、古代のアウストロネシア(マライ・ポリネシア)語族系であった、当時の土着日本語には「ン」の音がなかったので、それに近い日本語の音の「ミ」に変わったものである。

したがって、日本語から全ての外来語を排除しようなどというのは不可能なことである。しかし、少なくともすぐには意味のとれない、同音異義語の多い漢語は、できるだけ減らすように努力すべきであろう。

上に類した例はいくらでもある。たとえば「姑息」などがそれである。この語の字づらからすると、もっとも直接的な意味は「しゅうとめの息子」、すなわち(縮退の場合としての「夫」を含む)「義理の兄弟」のことを言っているように見え、なぜそれが「その場しのぎ」の意味になるのかが分かる人は少ないにもかかわらず、漢語だというだけでよく使われているようである。もっとも若い人たちのあいだには、「姑」の字が読めると、何となく「しうとめ」と結びつけて、「とりつきにくい」とか受けとったり、あるいは「姑」の字の読めない者は、字の中の「古」に引かれて、「古くさい」といった連想で、感覚的に「姑息」への意味付けをしている者もあるようだが。

もっと身近なことばでも、汽船の「出帆」などは漢語だから好んで使われているのであって、帆などない今日の汽船に対しては、やはりおかしな用語であろう。和語からすなおに作られた「船出(ふなで)」のほうが、よほど意味の透明性があるのに、それが和語だということだけの理由で、いまではあまり使われなくなってしまっているようである。

また、「生きる」「生まれる」「生ま」「生える」や「携える」「携わる」、「抱える」「抱く」など、日本語では全く異なることばを、それらに該当する適当な漢字がないために、同じ漢字で書いている。その一方で、「固い」「硬い」「堅い」や「使う」「遣う」、「熱い」「暑い」など、日本語では同じことばなのに、たまたま古代中国語に別字があったからというだけで、それらで書き分けることがなされている。このようなことも、漢字は貴く、和語は程度の低いことばであるという先入主があってこそ、いつまでも続けられているのであろう。将来こうしたことは、できるだけ避けるべきであろう。

上に挙げた例のように、日本語の同じことばを異なる漢字で書き分けることについては、かなりの人が、日本語はあいまいで、そうした書き分けによってはじめて意味がはっきりするというようなことを主張している(たとえば 橋本・他 1987 [50] 参照)。しかし、われわれが日常会話をするときに、そうしたやまとことばを用いても、意味が分からないということはないし、また場合によって別の字で表わされていることばだなどという意識も持っていない。

それに、英語の場合についてみても、やはり日本語に似て、ことばには概念を大きくとらえるところがあり、たとえば「hot」一つをとってみても「熱い」「暑い」「(からだが)ほてる」「おこりっぽい」「熱烈な」「激しい」「(味が)辛い」「(においなど)強烈な」「できたての」「(情報など)ほやほやの」「情熱的な」「熱心な」「追いつきつつある」「高圧電源の」「放射性を持った」「高速・高馬力の」などの意味があり、さらに俗語としては「ハイカラな」「即興的な」「すぐれた」「やばい」という使い方もある。以上は形容詞としての用法であるが、そのほか副詞、動詞としても使われ、さらに慣用句、熟語としては、「自慢話し(hot air)」をはじめ、「(悪事などの)温床(hot bed)」「(野球の)3 塁(hot corner)」など、まだまだ多くの意味を持つことは言うまでもない。

しかし聞いたり読んだりするとき、こうした異なった意味は文脈によって分かることであって、英文ではそうした意味の違いをわざわざつづりを変えて書き分けるようなことはしていないのだから、日本語の場合にも、わざわざ中国の文字で書き分ける必要はないであろう。

現在とんでもない和製英語が街に氾濫しているように、かつて漢字が輸入されたときにも、生半可な知識をもとにして、いい加減な漢字の用法が生まれたことは想像に難くない。それは現在でも起こっていることである。

たとえば上に挙げた「ことば遣い」のときのように、もともと中国における意味とは外れた誤用であるものに対し、わざわざ「使い」と区別した漢字を使い続ける

学術情報センター電子図書館サービス　　© 学術情報センター

必要があるのだろうか。

ほかに似たような例としては「宛て名」の「宛」がある。この「宛」の字には本来「曲がった」とか「屈曲した」とかの意味しかなく、日本でこれが「あてる」の意味に使われるようになったいきさつはよく分からない。まさか東京などの街並みがあまりにも曲がりくねっていて、なかなか目的地に到着できないことが多いということへの皮肉ではあるまい。とにかく、字義にうるさい人の中には、わざわざ「あて名」と書いている人もあるようである。しかし、日本では「宛」の字が本来の意味に使われることはないようだから、いまさら「当て名」と書き改めることもないかもしれない。

熟語でも、たとえばひと昔まえまでは「独擅場」であったものが、「擅(セン)」の字が「壇」に似ているために、「どくせんじょう」が「どくだんじょう」と読み誤まられ、ついに「独壇場」と書かれるようになってしまった。「擅」は「専」と同じく、「ほしいままにする」の意義の字であるから、早くから「独専場」あるいは「独占場」とでも書き改めておけば、ことばそのものが変わってしまうことはなかっただろうが、いまとなっては、もうもとに戻すことはむずかしいであろう。これなどは、先に述べた「姑息」の場合のように、漢字がなんとなく与えるイメージがわざわいして、ことばを歪めてしまった典型的な例であろう。

10　漢字の不具合

漢字とかなの交ぜ書きは日本文を読みやすくしていると、よく言われるが(たとえば 加藤 1979 [22]、安島 1995 [58]、阿辻 1996 [4])、そんなことは、膨大な量の漢字を覚えるよりもずっと簡単に覚えられる分かち書きの規則(たとえば 堀内 1959 [53])を採用し、実行すればすむことである。それに漢字の使用がそのまま語の分割を与えないことも少なからずある。

たとえば「見かけ上下がる」では「上下」に引かれてしまうし、最近話題になった縄文時代の文化についての報道では、「縄文」よりも「文学」のほうが使用頻度の高い単語なので、初期には「縄文学」がどうしても「縄・文学」と目に飛び込み、その後「縄文・学」と悟るのを経験した人は少なくなかったはずである。

さらに、以下でだんだんと見ていくように、日本語とは言語学的に本質的に異なる中国語の表記体系である漢字を使っていることにより、漢字かな交じりの日本文では漢字とかなが水と油のごとく相性の悪いことが多く、とてもエマルジョン化しているとは言い難い。すなわち、活用によって語尾が変わる日本語に、語に変化のない中国語の漢字を使うのだから、木に竹を継ぐように、しっくりしない場合が多く、今世紀になってさえ何度も送りがなの規則を変えてきたにもかかわらず、いまだにその場その場あたりの例外的な書き方が数知れず存在している。だから辞典にさえ「史的に見れば、送りがなは、訓みの確認のため漢字の傍らに随時小書きしたもので、一貫した理法など由来、存しない」と言い切ってある(山田 1981 [62], p.6)。

したがって、いまさらここに詳しいことを書くつもりはないが、一、二の例を挙げると、一般に動詞には「話す」のようにかなを送るが、これが名詞化すると「話」のように送りがなを省く。しかし、たとえば「組合」の場合では、「くみあい」なのか「くみあわせ」なのか分からない。したがって、後者は「組合せ」あるいは「組み合わせ」などと書かれる。もっとひどいのが「だきかかえ」であろう。「だく」にも「かかえ」にも「抱」の字が使われているから、「抱き抱え」でもあまり気持ちが良くないのに「抱抱え」や「抱抱」ではまず読めないであろう。「ひとさしゆび」も似たような例で、「人さし指」「人指し指」ならまだいいが、「人指指」とはとても書けない。ときには同字の重なりを嫌って、例外的に「人差し指」「人差指」と書いたりもする。このとき「差」は表音文字として機能している。

この送りがなの使い方について、以前は変化をする部分はかなを送るという書き方がかなり普及していたが、その後の改訂で、送りがなは読みの分かる限り少なくすることが決められた。言うまでもなく、むかしの書き方とて、例外はあった。たとえば「すこし」「すくない」から分かるように、変化する部分を全て送るとすると「少こし」「少くない」となるが、年配の人は「少くない」は「すくない」ではなく、「すくなくない」と読むので、結局「少し」「少ない」となっていた。しかし現在では「いった」「おこなった」がどちらも「行った」になったり、また上に述べたような例外がいろいろと出てきて、むかしよりむずかしいものになってしまった。

このように、おそらくわずかばかりの文字数を減らすことだけが目的の、明らかに改悪と思われる新表記法であり、それは常用漢字表などと同じく、目安であるべきだと思われるのに、官庁ではこれを強制しているところが多い。たとえば(銀行)「振込」を「振り込み」

学術情報センター電子図書館サービス　　© 学術情報センター

と書くと、書類を受け取らず、書き直しを要求するところがある。そのくせ、内閣訓令であり、またISOを通して国際標準として認めさせた、日本語の訓令式ローマ字つづりは、外務省(パスポート)、建設省(道路標識)をはじめとして、ほとんどの省庁がこれを無視し、依然として英語もどきのヘボン式の使用を強いているのである。

最近電車の吊り広告で「秋を色採る山河」という文言を見た(1996年10月、写真弘社による東京地下鉄内での広告)。「いろどる」はもともと「色取る」だから、「彩る」と書くよりも「色どる」あるいは「いろどる」のほうがよいのかと思うが、広告が、ある写真関係会社のものであったので、わざと「採る」の字を入れたのかもしれない。もっとも、写真なら慣用字は「撮る」ではあるが。

このように、日本語では2語の複合としてできたことばに、相当する中国語の漢字をただ一つ当てている表記はかなりある。「陥る」は「落ち入る」、「快い」は「心良い」、「掌」は「手の平」であり、「傾く」は「片向く」である。「たらい(盥)」は「手洗い」だし、「実る」あるいは「稔る」のもとは「実成る」あるいは「実乗る」だそうだし、「つまびらか」は「詳か」あるいは「審か」などと書かれるが、もともとは「つまびら(抓開)く」から来たものであろう。こうした表記法なども、やはり和語に対する漢字優先の心理が見え隠れしている例である。したがって日本語の起源から見たときには、「みのる」の送りがなは「実のる」が正しく、現行の「実る」はおかしなものということになる。そうは言っても、「苔(こけ)」が「木毛」から来ているところまでさかのぼったとすると、もう、ちょっと分からなくなるであろう。

そのほか、よく知られているように、漢字を使用することによって、意味は全く異なるのに、ことばとして全く同じ発音になる、いわゆる同音異字、同音異義のことばの問題が起こってくる。この問題はよく知られているので、ここでは詳しいことは省略するが、たとえば和語の「落とし穴」で済むものを、それでは威厳がないと考え、わざわざ「陥穽」(同音異義語に完成、官制、乾性、歓声、喚声、閑静、感性、管制、監製などがある)を使ったり、あるいは安易に漢字を並べて、たとえば「規整する(regulate)」「規制する(control)」「規正する(correct)」のように、漢字を見るか、英語で言い換えなければ分からないような同音異語を新しくつくることは避けるべきである。いまやたらに片かな語が増えているが、それは口語によるコミュニケーションがますます増える情報化社会において、人びとが同音異義語を避けようとすることも動機の一つになっていると思える。

中には、同音異義語を区別する必要上、漢字の使用は不可欠だなどという人もあるが、それは本末転倒であって、音韻構造に強い制約のある漢字をただ並べただけの安易な造語が同音異義語の問題をもたらしているのである。漢字の使いすぎを改めれば同音異義語が減ることは、すでにベトナムや南北朝鮮など、いくつかの国によって実証されている。それは人びとが古いことばや方言など、その国古来の言いまわしを掘り起こして、巧みなことば直しを活用することによって実現されている。そうした変化は、同時に言文一致を促進し、人びとの口頭によるコミュニケーションの能力を向上させてもいるようである。

しかし、漢字貴しの先入主の固い日本では、ホームドラマや純文学の作家のように、比較的飛躍の少ない物語りを書く者の中にはそうした試みをなすものは見当たらず、わずかに時代小説作家の時代考証や、奇想天外な着想が命である科学小説作家の古代ものなどの中で、例外的に試みられることがあるのが見られるぐらいのようである(たとえば 豊田 1987 [42])。

さらに漢字の存在が方言の研究そのものの進歩を妨げていることさえ、すでに1920年代に柳田(1925 [59])によって指摘されているのだから、こうした努力をするのは、そう簡単なことではないだろう。

11 和語の圧迫

日本語のことばに直接対応する漢字がないために、なんとか漢字を当てはめた、いわゆる当て字(あるいは宛字)には2種類ある。一つは「面白い」や「目出度い」「矢張り」「皮肉」のように音で当てたものであり、もう一つは「紅葉(もみじ)」や「文身(いれずみ)」、「経緯(いきさつ)」「等閑(なおざり)」のように、同じ意味の中国語の単語の漢字をそのまま当てたものである。

おもしろくても顔が白くなるわけはないし、めでたくて目が飛び出す度(＝法則)もない。漢字は見てすぐぱっと分かるなどと言われるときには、こうした困った例は全く無視されているが、その数はばかにならない。

それでも前者の類は和語そのものは保存しているが、問題なのは後者である。たとえばひと昔まえまでは、秋に木の葉の色が変わるのは「もみじ」と言って

学術情報センター電子図書館サービス　　©学術情報センター

いた。ところが「もみじ」に相当する漢字はなく、中国語の「紅葉(hóngyè)」を当て字として使っているうちに、いつのまにか「こうよう」と読まれるようになってしまい、いまでは「もみじ」のことを「こうよう」としか言わなくなってしまった。そして「もみじ」のほうは、代表的な紅葉樹である「かえで(楓)」を意味するようになり、それとともに「もみじ狩り」や「もみじ葉」「柿もみじ」といった古来からのことばも、いまでは忘れ去られようとしている。これは「花見」では、和語に対応する漢字が日本語の「はな・を・みる」の語順で書かれているために、依然として「はなみ」で留まっていて、「かけん」に変わらなかったのといい対照である。

同様に、いま「紅玉」と書き、「こうぎょく」と呼ばれているリンゴの品種も、もとは「べにだま」と呼ばれていたものらしい。いまでもリンゴ園では「べにだま」と呼んでいるところがある。この種の語も多く、「ゆずりわたし」、「やまづみ」などが、「譲渡」「山積」と、送りがな抜きで書かれているうちに、ついに「ジョウト」「サンセキ」と漢語化してしまったのは、そう古いことではない。「としつき(年月)」が「ネンゲツ」になり、「いきき(往来)する」が「オウライする」になったのもこれに類している。

これらの例のほかにも、漢字の使用が影響して滅びかけている和語は数知れずあることを忘れてはなるまい。たとえば、「たとえ(例え)」と「たとい(縦令)」とは本来別の日本語である。しかし読んだり書いたりするのがむずかしい漢語「縦令」が「たとい」に当てられていたために、むしろかなで「たとい」と書かれているうちに、漢字が使われている「例え」に引かれて混同され、いまではほとんど「たとえ」になってしまった。

また適当な漢字がないために低俗に思われ、あるいは漢字で書くと直截に読めないために嫌われて、文章に使われなくなった和語の多いのも残念なことである。「こなす(頭ごなし)」「しこり」などが前者の例であり、「あらが(抗)う」「はば(阻)む」「いじ(弄)る」「くたび(草臥)れる」などが後者の例である。テレビやラジオでニュースなどを聞いていると、そのことがよく分かる。

専門用語にしても、「perturbation」を物理学者は和語を用いて「ゆらぎ」と訳しているのに、工学者は同じものを「攝動」と訳している。また「feeder(給電線)」もかつては「饋電線」と呼ばれていた。ダーウィンの進化論においては、英語としてふつうのことばの「(natural)selection」に対して、漢学の流れを汲む日本の生物学者によって「(自然)淘汰」という、中国人でもよく知らなかったむずかしい古典語が当てられてしまった。まさに漢字貴しの思いの末のことであったろう。ちなみに「淘汰」にあたる日本語は「よなげ」であり、動詞は「よなげる」、古くは「よなぐ」であるが、漢字で書くと「沙げる」「汰げる」「淘げる」「簸げる」などで、どれもすぐには読めないので、使われなくなってしまったことばの一つである。こんなところにも、日本人が良いことだと思って、漢字に音読みと訓読みとを導入したことが、実はしばしば自からのことばのほうを失なわせることにつながってしまったことが表われている。

テレビやラジオのように、一般に口頭で情報を伝えるときにさえ、このごろでは不必要に漢語が増え、和語が使われなくなってきている。たとえば「こわれた」「落ちた」「解かれた」などと言えばよく分かるところを、「倒壊した」「落下した」「解除された」などと言っていた(1996年12月6日朝7時すぎ、NHK名古屋放送局でのテレビニュースの例)。そのほか「肥大化(＝ふとりすぎ)」、「捻出(＝ひねりだし)」、「克服(＝うち勝ち)」、「懐柔(＝まるめこみ)」、「変遷(＝うつりかわり)」など、毎日のニュースは漢語だらけである(松岡 1992 [54] 参照)。

NHKをはじめ、新聞などでの用語は、個人個人が勝手に決めているのではなく、用語委員会のような組織を作って研究し、決めているものであることは言うまでもない。しかし、かつてNHKの放送文化研究所の主任研究員であった、上の松岡の著作などを参照対比して見ると、現在のマスコミ一般では、もはやことばの平易化、民衆化への志向はあまり見られないものになっていると思える。

そのことは、高校で国語教育にあたっている小畑忠夫が用語問題についてかつてNHKの会長にあてた一連の公開書簡(たとえば おばた 1992 [17], 1993a, b [18][19])に対する、NHK放送文化研究所の用語研究班の回答(NHK 1993 [11])が、国語問題の核心に対して筋道を立てて応えていないことからもうかがえる。

言うまでもなく、できるだけ日本古来の和語を使い、漢語を避けるとは言っても、すでに完全に日本語化してしまったことば、たとえばすでに述べた「かみ(紙)」や、また「うま(馬)」や「うめ(梅)」などがその例であるが、そのほか「もちろん(勿論 wùlùn、論ずるなかれの古語)」

学術情報センター電子図書館サービス　　©学術情報センター

のような漢語を、いまさら全て和語で置き換える必要はないであろう。それにしても「もちろん」よりは「言うまでもなく」のほうが、意味することを日本語として直接的に表現しているではあろうが。

そのほか、おもしろい例としては、たとえば「けみする」がある。「月日をけみする(＝経る)」、「子細をけみする(＝調べる)」などと使われる「けみする」は、もともと中国語の「検(ケム)」が和語化したものであったが、日本語における意味が少し変わったあと、それに合わした漢字で「閲する」と書かれるようになり、その後「閲」の字がよく読めないことがあってか、いまではほとんど見かけなくなった。これなどはことばとして、「ふ(古・旧)る」から来た「へ(経)る」よりも日本語と相性がいい場合も多いように思えるから、今後も使われ続けてもよいように思われるのに、漢字のおかげですたれてしまったことばだと言えそうである。

12　異体字などの問題

もっと細かい話しとしては、異体字やフォントの問題がある。日本のみならず、世界の多くの地では、古くから文字が呪詛的な魔力を持つという考えが一部にあるが、それが日本ではまだかなり信じられているらしく、字の形のわずかな差異にこだわる人たちがあとを断たない。

現在まだ一般では漢字は表意文字だと言われているが、文字学においては漢字は意味を直接表わすのではなく、むしろ主として語を表わす表語文字(word script, logograph)と考えられている。しかし最近の学者、特に言語学者、文字学者やシナ学者(Sinologists)たちの中には、漢字は、極端に数が多く、またいまでは音声的に不正確になってしまってはいるが、視覚的かつ多分に表語的性格を残した「音節文字」であるという考え方に傾いた者が増えてきているようである(たとえば DeFrancis 1984 [82], 1989 [83]、Unger 1996 [98])。にもかかわらず、漢字はその発祥において物を形取った(＝象った)り示唆したりした表意文字から表語文字へと発達したものであるから、同じことばを表わすのに、いろいろの違った絵画的表現、つまり形があったことは否めない。

たとえば「峰」に対する「峯」や「島」に対する「嶋」のように、偏やツクリの位置を取り替えたり、あるいは一部分を置き換えたりした、(ときには動用字とも呼ばれる)別字はその一例である。これは数がかなり多く、たとえば「松」「海」「群」などにもそうした別字がよく使われている。

したがって、いまだに異体字や、細かな形の差異があるということは、漢字が画数の多い文字であるゆえに、標準形へ向けての収束が遅く、その意味ではまだ発達途上にあり、安定化していない文字体系であることの反映にすぎない。

文字使用の本質が、それによって表現された意味内容の伝達にあることを考えれば、同じことばを表わす異なった文字は、数学的に言えば同値類を形成しているのであって、一つの同値類は、その元の中から、万人の同意を得た任意の一つを選んで代表元として用いればよいことである。そのような、コミュニケーションの道具としての文字の性格を理解しようとせず、いつまでも異体字の細かい差異にこだわり続けることは論語読みの論語知らずで、ひょっとすると一種の機能的文盲であるとまで極言できそうである。

幸い大勢としては、現在そうした収束の方向に進んでいるのは明らかである。たとえば中国の書聖といわれる王義之や顔真卿などの残した書をみれば、常用漢字表にある多くの文字が、常用漢字の字体とはかなり異なる形で書かれている。いまになって、こうした異体字をいちいち書き分けるような必要はないし、また現在使われている異字体は、人名漢字も含めて、これからもどんどん整理してしまってもよいであろう。国語の能力に限ってみても、文章法など、われわれがもっと力を入れなければならない課題はほかに山積みしているのだから(山田 1986b [67])。

ここでちょっと現在の中国の簡体字に触れておく。むかしといえども、漢字は意味的な属性を示唆する形符(しばしば偏や冠)と、発音を示唆する声符(しばしばツクリ)からなっている、いわゆる形声文字がほとんどであった。したがって、中国ではむかしから使われてきた簡体字を参考にして、漢字を簡単化し、1956年に中国国務院が正式な簡化漢字表として発表したものも、現在の中国の発音に沿って簡略化された字体が多い。たとえば、「遷」は同じ発音の「千」にシンニュウをかけたもの、「達」「礎」「積」なども同じ原理で、それぞれ「大」にシンニュウをかけたもの、石偏に「出」をツクリとしたもの、ノギ偏に「只」をつくりにしたもの、などとなっている。

したがって、日本と中国とで同じ形にした簡略字を使えばよいといった、ナイーブな提案が繰りかえしなされているものの、同じ漢字の発音が日本と中国とでかなり違ってしまったいま、そうした提案はどちらの

学術情報センター電子図書館サービス　　◎学術情報センター

国にとっても実用的ではない。

さらに、もっと基本的な象形や会意文字についても、たとえば「骨」の字は、日本では頭部の四角形部分の中の4分の1の小四角形の2画が右下についているが、おそらく書くときの1画を減らすためにか、中国の簡体字ではこれが1画で左下につけられている。しかし、情報伝達の目的からすれば、こうした細かい形の違いはどうでもよいことであろう。

そう達観してのことかどうかは定かではないが、たとえば倉石(1965 [27])の中国語辞典では、初版から20年たって、筆者の使っている1985年版でも、中国語の部分の「骨」の字は、上に述べた中国の字体ではなく、日本の字体をそのまま使っているし、倉石・折敷瀬(1983 [28])の日中辞典では、正しい簡体字の字体を使ってはいるものの、両者のあいだでは、形が違うだけでなく、筆画数も1画分異なっているのに、簡体字と日本の字との差異を示した、巻末の表からはこの対を外してある。

朝日新聞によると、現在は中国政治の元老となっているテン・シァオピン(鄧小平)は、1974年6月6日、副首相として日本の西園寺公一と会見したときに、「中国は二つの面で、日本に歴史的な損害を与えている。一つは漢字であり、もう一つは孔孟の道である。・・・私はそれについて日本にあやまるべきであると思う」といった内容の発言をしたそうである。かれのような政治家が他国の文化に対して、独断で軽がるしくこうした発言をするとは思えない。したがって、(孔孟の道のことは政治的な色彩が強いから、ここでは立ち入らないとして)、日本における漢字の使用については、これが、「われらが文化の宗主」と思いこんでいる中国の識者のあいだに共通な、醒めた理解と考えてよかろう。

正書法とは同じ語をいつも誰もが同じ表記にするということなどの取り決めである。以上に述べてきたことからも分かるように、日本では多くの人たちにまだそれが理解されていない。したがってこの20世紀末にもなって、秩序立ち、明確な正書法が国語に確立できていないのは、世界の主要な国のことばの中で、おそらく日本語だけではなかろうか。そればかりか、むしろ同じ語がいろいろな表記で書けるということがいいことであり、日本語表記の特徴だとする考えがある。しかしそれは伝達される情報についての吟味ではなく、情報の入れものについての遊びだと言える。したがって日本語を国際的に開けた言語にするためには、そうした考え方の見直しを含め、思いきった国語政策の確立に向けて、早急に努力をする必要があるだろう。そしてその第一歩は、やはりやまとことばを中心に据えた、明確な表記法を確立することにあろう。

もちろん文字を、意味内容の伝達を主目的とせず、主として造形的芸術の対象とするような書道などは、こうした考慮の対象外であることは言うまでもない。また、学者が文字そのものの発展や、古典の書誌情報などについて研究するときなども、全く別の枠組みを必要とすることは明らかである。

しかるに、この後者の場合については、専門家の人たちの中から、現在のJIS規格には、古典や専門分野で用いられている漢字が入っていないので、何とかして欲しいという声がよく聞かれる。しかし、民衆一般のコミュニケーションの道具としての文字と、専門家のみが必要とする特殊文字とをいっしょにして扱って欲しいという要望は、たとえば通信販売の商品カタログに、家電製品と並べて考古学時代の石器や土器なども掲載しておいて欲しいといった要求に似ている。ことが物品については、日曜大工の道具や電気炊飯器と並べて、石斧や縄文土器を同一扱いにするのはおかしいと誰でも思うのに、ことが文字になると、そうした要求を不思議に思わないのは、われわれが、文字の主要目的はコミュニケーションの道具として使うことであるという、確固たる認識を欠いているからであろう。

しかし筆者は、専門家のための文字フォントの整備に冷淡なわけではない。かつて筆者は、研究者に対してデータベースから学術情報を提供するサービス業務を行なっている、文部省所轄の大学共用利用機関の一つ、学術情報センターに在籍していた当時から、研究者用の漢字データベースの構築に関して、以下のような積極的支持の考えを述べてきた(西原・他 1996 [44]参照)。

すなわち、日本文化形成の基盤となった漢字はわれわれの貴重な文化的財産なのであるから、日本や中国などの歴史や古典などの研究に携わる者にとっては必要不可欠でありながら、いまではまれにしか使われない漢字についての情報は、国家的規模のデータベースとしてサービスを提供すべきである。そのためには、どこか一か所に、データベースの半永久的な管理を司るべき施設を設け、そこでは専門家の英知を集めて、文字に関する必要項目の共通仕様を定めて漢字データベースを構築し、関係研究者が、自分の研究において必要とする漢字に関するフォントパタンや意味、用例、出典などのデータが、もしすでにそのデータベー

学術情報センター電子図書館サービス　　©学術情報センター

ス内に存在すれば、通信ネットワークなどを通してそれを取り出して利用し、もし存在しないときには、新たにそれを自分で構築して使うと同時に、このデータベースに入力し、その後の保守管理と他の研究者へのサービスとをまかせる、といったことができるような機関にするべきである。

そのようなデータベースは、研究者たちによって広く使用されることにより時間とともに完成度が高く、利用価値の大きなものとなり、異なる研究者が同じ文字に関するデータを必要とする都度、自分で作成するといった労力の重複を軽減してくれる。そうしたデータベースの構築は、内容として、廃字や地方字、それに俗字、日本の国字、ベトナムの漢字(チューノム)などの異体字を全て数え上げると、総計12万字にもなるという推定もあるほどの大事業となるが、それでも物理学における粒子加速機、天文学における超大型電波天文台などと肩を並べるほどの大型施設にはとてもならないのだから、学問的研究のためには、国として正面きって取り組むべき、貴重なプロジェクトであろう。

13　ローマ字表記の問題

以上では、現在使われている漢字かな交じり文の表記法の枠組みの中で、日本語を民衆化することにより、同時に国際化の程度を高めようと考えたときの、息の長い、長期的な改善方針について考えたものである。その実行でさえ、文化の中でも特に保守的性格の強い文字の問題であるから、なかなか簡単には進まないであろう。

ましてや、千年以上もかけて進化し、使われてきた漢字かな交じりの文の中に、たとえば新しくローマ字語を加えて、ほかの語と対等に扱うというようなことなどは、おいそれとは実現しないであろう。さらに一歩進めて、別にローマ字だけで日本文を書くことをも認めることなどは、ほとんど実現不可能だと思われている。

にもかかわらず、古代における文字の発展過程を見ても、また近代から現在に到る世界の動向を見ても、文字の表音化、特にアルファベット表記への収束の傾向は明らかである。世界最大の国中国は、その大きさのゆえもあってか、文化的変化のゆるやかさも際立っているようであるが、その中国でさえ、技術的な便宜さと情報処理面の要求との圧力のもとで、言語学者、工学者、コンピュータ・プログラマが協力し、いま懸命になって漢字にかかわる問題の解決に努力している(たとえば Mair & Liu 1991 [88])。それにもかかわらず、遠からず中国は、漢字文と、ピンインによるローマ字正書法の文との2表記システムの国になるだろうという国際的評価が生まれつつある(たとえば Mair 1996 [89])。

日本語とローマ字との関係については、1868年の明治維新の前後以来、かな文字表記の可能性と並んで、すでに多くの著作があり(橘田 1992 [24] 参照)、中でもたとえば田丸卓郎の名著「ローマ字国字論」[39]は、1914年(大正3年)の初版以来、80余年を経て、その記述における文字使いがいまでは古くなってしまったにもかかわらず、1930年(昭和5年)の第3版が、依然として今日でも版元から供給されている。

したがってここではローマ字の問題を、いままで本稿に述べた問題と関連することだけに絞って、少し述べてみる。

その一つは、今日科学や工学をはじめとし、日常生活にいたるまでのほとんどあらゆる分野で、外国語を表音的に書き改めた、いわゆるカタカナことばが氾濫していることである。一見これは日本語の国際化の流れに沿っているかに見える。しかし、もともと日本語はせいぜい一つの子音のあとに必ず単母音がついた開音節の拍(mora)がほとんどという構造を持っているから、この拍音を文字にしたカナ書きでは、重子音が多くあり、また子音や重子音で終ることの多い外国語の単語を表記するには無理がある。したがって、たとえば外国から日本への留学生にとって、これらカナことばを理解することがなかなか困難であるということはよく知られている。しかもそれは単語が純粋に外国語の場合のことであって、さきに述べた「和製外国語」についてのことではない。

この場合、もし日本文の中にローマ字つづりの単語を混ぜることを正当化し、第2次大戦直後、一部の専門書で試みられたように、原語の単語はアルファベットつづりのままで文章の中に取り込むといったことを許すなら、そうした困難はかなり解消できるであろう(岩村 1955 [8], 1956 [9], 1966 [10] 参照。ちなみに、この岩村清一は、かつて日本海軍の要職にあり、日本語表記法の不合理性を身に滲みて体験した者の一人である)。そればかりか、日本語と外国語の音韻構造の違いのゆえに、たとえば「interface」のように、同じことばが「インターフェース」「インタフェース」「インターフェイス」「インタフェイス」などと、いくつもの異なる表記になるといったわずらわしさはなくなる。

具体例として、たとえば日本化学会・編の「実験化

学術情報センター電子図書館サービス　　©学術情報センター

学講座」において、化合物名を全て英名とし、アルファベットつづりのまま文中で使ったところ、講読者から、使いやすい、読みやすい、よくやったとの褒めことばが寄せられたとのことである(日本化学会 1966 [48])。そのほか「薬学雑誌」が同様の文体を採用しているし、また、日本のこの試みの影響があってのことか、韓国の「薬学雑誌」もハングル文の中に英文の化合物名をそのまま用いている。なお、いまの韓国の学術雑誌では、漢字はほとんど使われていない(矢野 1997 [61])。

ちなみに、1920 年代から 30 年代にかけて日本でもローマ字書きの学術専門書が何冊か出版されたが、その中で、たとえば荒川文六著の「電気工学の実験」などは、ローマ字書き文の中に多数の英語の用語をそのままの形で使っている(荒川 1928 [5])。

こうした文体を一般的に用いた場合、専門家以外の一般の人たちにそれが読めるか、またどう発音されるかといったことが問題になる。小学校におけるアルファベットの教育は、戦後かなり詳しかったのに、文部省の学習指導要領の改訂のたびにだんだんと省略され、いまではほとんど形式だけになってしまった。しかし国際化の進む今日、再びローマ字教育は強化されるべきだろう。

発音に関しては、ある言語の変化は語囲、構文、発音の順にむずかしくなることがよく知られている。したがって原語つづりの単語をどう読むかは、時間とともに落ちつくにまかせるべき問題である。かつて日本に漢字が導入されたときにも、その読みはそうした過程を経て、中国におけるものとは異なるもの、すなわち現在の読みが定着したのである。

そうした日本語の音韻構造の制約を無視して、いまわれわれは英語の発音をいかにカナで表現するかに腐心しすぎていないだろうか。たとえば一部の放送局で「チーム」を「ティーム」と発音しようとしたり、「テレメディシン(遠隔医療)」を「テレメディスン」と書いたりしている。しかし、別のところで詳しく述べておいたように(Tatuoka & Yamada 1996 [94])、日本語の「ティ」(歯音的、dental)と英語の「ti」(歯茎音的、alveolar)とは別の音であり、同様に日本語の「シ」と英語の「sh」も、また別の音である。したがって日本人が「machine」や「function」を発音すると、英語の話者にはかえって反対に押しやられた音のように感じられ、「maseen」や「funcsion」とさえ聞こえるのである。だから日本語として、いまさらこれらをわざわざ書き分ける必要はあまりない。事実、「multimedia」は歴史的な理由で「マルチメディア」と、いま混合型で書かれており、「マルティメディア」でも「マルチメジア」でもない。とくに「b」と「v」や「l」と「r」が書き分けられず(たとえばバーチャル・リアリティ)、また「s」と「th」(魚の bass と bath)なども書き分けられていない中にあって、そうした書き分けはほとんど無意味であろう。

これに関係して、日本語のローマ字つづりの標準には、音声・音素学、つまり音韻学的に、そして形態論(morphology)的にいまのところもっとも妥当な訓令式つづりがあり、国の標準であると同時に ISO によっても国際標準としても認めさせてあるのに、皮相的に英語の音声に似せた英語式のヘボン式が依然としてまかり通っている(Tatuoka & Yamada 1996 [94])。これなどもいかに日本人が時代の強者におもねり、日本文化や日本語を低く評価してきたかの反映の一端であろう。商品名などはいざしらず、少なくとも万人の用いる地名などは、この ISO 標準に従って表示すべきであろう。

ちなみに、中国において長い伝統を持つ「青島ビール」のラベルでは、商品名のほうは、少なくとも中国では、1912 年以来の歴史的なつづりである Wade-Giles 式よりも古い、おそらく郵便事業に発するつづり方で、依然として「Tsingtao Beer」と記してあるが、それでも製造所の所在地のほうは「Qingdao」と、中国が ISO の標準とさせたピンイン式で記してある。

アルファベットつづりの単語を混ぜることは、日本文にもう一つの影響をもたらすことになるかもしれない。それは分かち書きの導入である。

すでに述べたような方針によって、日本文における漢字の使用量を減らそうとすると、いままでは漢字の存在によって二次的に果たされていた、語と語の区切りが定かでなくなるところが増え、読みにくくなる。アルファベット文字の文では、歴史的にいろいろの試行を経て、いまでは分かち書きが当然のものとなっている。したがって、かな書きの部分でも、同様に語間にスペースを入れて分かち書きをすれば、読みにくさの問題はなくなる。そうした試みをすることもなく、漢字の果たす分かち書き的効用のみをあげつらうのは意味のないことである。(分かち書きの規則については、かなりの数の出版物があるが、たとえば 堀内(1959 [53]) 参照)。なお、1922 年創刊の理化学研究所の彙報では、1929 年から 1940 年までのあいだ、ふつうの漢字かな交じり文が分かち書きがされていた。しかも多くの英単語がカナ書きされず、ローマ字書きのまま文中で使われていた。

学術情報センター電子図書館サービス　　©学術情報センター

とにかく、いまもし全ての文をローマ字書きにしたとすれば、これまで日本人を悩ませてきた送りがなの問題もまた、一挙に解決されることを忘れてはなるまい。これは大きな前進となる。

ちなみに、カナモジ会の会誌「カナノヒカリ」では、カナ文字だけの文が多く、当然ながら分かち書きがなされているが、そのほかにも一般の文章において、むずかしい漢字の漢語だけをカタカナで書くという文も載せられている。どうせ漢語は外来語あるいは外来語もどきなのであるから、それをカナ書きにすれば、ある程度は分かち書きと似た機能を果たしてくれる。たとえば「語彙」を「語い(＝かたらい)」と混ぜ書きするのをやめて、「語イ」と書くことなどである。

日本語はいわゆる粘着語(膠着語)型であるから、分かち書きがむずかしいということがよく言われる。(この粘着語型という分類は、あまり妥当でない点があって、いまでは言語学者は使わないようにしているが、ここでは便宜上この呼び名を使っておく。)しかし粘着語の性格をもっているウラル・アルタイ語族の数多い言語の中で、分かち書きをしていないのは日本語だけであることを考え合わせれば、日本語の分かち書きが、特にむずかしいものでないことは容易に理解できるであろう。

また、ローマ字書き文では同音異義語の区別ができないという、しばしばローマ字に向けられる不満(たとえば 加藤 1979 [22]、安島 1995 [58])は、実は本末転倒の考えであって、漢字を使うこと自体が同音異義語の問題を生んでいるのだということはすでに述べた。そのほか、ローマ字書き単語には意味の透明性がないとか(橋本・他 1987 [50] 参照)、速く読めないとか、いろいろの苦情が出されている。しかし、そうしたこといっさいは、実はわれわれが幼いときにどういった文字教育を受けたかによって決まってくる認知能力の発達の問題であって、文字そのものの本質ではない。詳しいことは長くなるし(Paradis et al. 1985 [91] 参照)、過去において筆者自身も何度か詳述してあることだから(山田 1990[70], 1991[71])、ここではこれ以上立ち入らないことにしよう。

ただ早合点してはならないのは、これからの日本文にローマ字書きを増加させていくといっても、それは従来の表記法の中に少しずつ加えていくことによってもできるし、また従来の表記のものとローマ字表記のものとの二本立てをとりつつ、少しずつローマ字表記文を普及させて行くこともできる。ケマル・アタチュルク将軍による1923年の革命のあとにトルコ共和国がとった、それまで使われていた表音文字のアラビア文字を、同じ表音文字であるローマ字で置き換える政策は、ほとんど一夜にして成し遂げられたかのように見えるが、実はその背後には、少なくとも50年にわたる慎重な準備があったということである。

それとても、革命当時のトルコは文盲者が人口の90パーセント以上もあり、また国としても、アラビア文字を用いていたイスラム文化圏から脱出する気運が盛んだった上に、ケマル将軍自身がローマ字化運動に熱心であったので、アラビア文字使用禁止の違反者を処罰する法律まで作って、強権的にローマ字化を推進したといった事情があって、やっとこの文字改革は成功したとのことである。ましてわれわれが表語文字である漢字からの改革を、しかも説得だけで成し遂げようということは至難の業である。

15歳を過ぎて、大脳神経の接続構造が固まってしまった者には、どれにしても、新しい文字体系の駆使力はなかなか身につけられない。このことは大脳生理学的にも実験心理学的にも示せることだし、また大人が苦労してもなかなか新しく使いこなせるようにならないパソコンを、頭の柔軟な小・中学生が、マニュアルを見ることもなく、試行錯誤でさっさと使いこなせるようになる事実からしても類推できる。

したがって表記法の改革は現在の国民のためのものではなく、次世代以後の国民の利益のために、順を追った長期計画のもとになされるべきものである(柳田 1936 [60] 参照)。

もう一つ、ローマ字は、従来の漢字かな交じり文の入力のときにも大きく役立っている。これについてはあとで述べる。

14　文字と思考形態

上に述べたのは、日本語の民衆化、そして望むべくは国際化へ向けた立場からみた表記法の考察であった。しかしながら、その背後には、もっと広く、われわれ日本人の思考形態に関わる問題が隠されている可能性がある。まず、ものごとが漢字によって記述されると、漢字の数の膨大さのゆえに、記述されたものごとの体系化が困難になり、それが中国における科学の誕生を抑圧したのだという指摘がある(Logan 1986 [87])。しかし、これは分類や索引作りなどにおける、漢字の処理の複雑さに由来した、思考にとってはいわば外的な事象である。

学術情報センター電子図書館サービス　　◎学術情報センター

たとえば、書物などにおいて引用文献リストや索引を作るときに、しばしば人名の漢字が読めなかったり、漢字の並び順に定まった規則がなかったりするので、大変な苦労をさせられることを考えてみただけでも、これは明らかである。したがって、相当に専門的な書物でも、引用文献リストや索引がないといった欠陥商品がまだ少なからずあるが、アメリカなどではそうした本はふつう専門書とはみなされず、大学などでも研究のためにはまず読まなくてよいとまで指導されている。

直接に人間の認知機構に関わる問題としては、「漢字は表意文字だから、見てすぐぱっと分かり、速く読める」と、かなりの人びとによってしばしば繰りかえし主張されているということがある(たとえば 加藤1979 [22]、安藤 1986 [6])。しかし、それらの内容を詳しく検討してみると、それがまだ科学的な根拠を与えられていない、漢字貴しの先入主にすぎないことがわかる。

第一に、漢字のなかで象形や会意によって意味を示唆しているものはほんの一部にしかすぎなく、すでに述べたように、大部分は不完全ながら音節文字的色彩の強い表語文字である。それが速く読めたり、すぐに意味が分かったりすると思うのも、実は大脳の成長期に、長いあいだをかけて培った慣れを基にして形成された機能についての、主観的な印象論にすぎない(山田 1986b [67], 1990 [70], 1991 [71])。

しかし、漢字は見ただけでぱっと意味が分かると言われること自体は、全くの誤りであるとは言えない。すなわち、いまのように同音異義語の多いことばを表音表記すれば、たとい文脈から意味の選択ができるとしても、それには時間がかかる。それに対して、漢字で書き分けてあれば、その時間が短かくてすむ。だがこうした場合には、そもそも漢字に頼りすぎて同音異義語を作りすぎたこと自体が誤りなのであって、文字の本質的な性質によるものではない。

またそのほかに、たとえば逆に、あることばの表記の構成要素である漢字一つ一つの意味は知っているが、しかし単語としてつづられた文字が、実はことばとして読めないという場合がある。いまでいうと中学レベルの、むかしの教科書で使われていた表記の中には「賞め」という和語があった(教育総監部 1943 [25])。いまでは大学生でもこれを読むのに苦労するであろうが、「賞」の字を知っていれば、おおよその意味の見当はつけられる。したがって、それがあたかも漢字は見ただけでぱっと意味がわかるという印象を与えるのであろうが、もしはじめからこれを「ほめ」と書いてあれば、漢字でなくてもその意味は明らかなのである。

とにかくそうしたことがあるから、文字学的には表語文字と考えるのが妥当といわれる漢字が(山田 1986a [66])、しばしば表意文字だと言われ続けているのであろう。

第二に、構成文字の意味の組み合わせを見ることによって、漢語には意味の透明性がある(加藤 1979 [22])というのも、やはり都合のよい場合だけを集めてきての議論であって、構成漢字の示唆する意味とは異なった意味を持つ漢語はいくらでもある。それだからこそ、漢和辞典の内容のほとんどは、漢字ではなく、熟語の意味の説明が占めることになってくる。

たとえば、「物理」「心理」「法理」とくれば、「料理」は「材料学」と思うのは当然であろうが、料理にはそんな意味はない。ちなみに中国語でも「料理 (lìaolǐ)」は一般に仕事を取り仕切るとか整備するとかいう意味であって、「料理家務」といって、はじめて「りょうり」が家仕事の一部として取り仕切られることを意味する。

つぎに「暖簾(のれん)」をみよう。「暖」が「の」と読まれるのは唐音の「ノン」がつまったものであるが、意味は「あたたかい」であることに変わりはない。「簾」は竹を編んだ「すだれ」であるから、「暖簾」の字からは日本における「のれん」の意味は出てこない。中国での「暖簾」とは部屋の入り口一杯に垂らして寒気を遮断する、しばしば綿入れ(キルト)でできた垂れ幕のことであったが、現在でもこれは中国の寒い地方で使われている。むかしの日本でそれが禅宗の寺に取り入れられたのであるが、比較的暖い日本では、いつのまにか本来の意味が失われて、商家の「のれん」に化けたのである。ちなみに現在の中国では「簾」は、同音「lían」で「酒家の印の幟」を意味する、「穴」の下に「巾」と書いた字といっしょにされている。

そのほか、すでに述べた「上梓」「姑息」をはじめ、字づらだけから見れば、「作家」は「家の建築」だし、「大家」は「大きな家」だし、「大兵」は「軍の大集団」だし、「若鷺」が「鷺」に成長するわけではないし、「田作り」がなぜ「干しいわし」なのだろうか。また「素敵」の「もとからの敵」は何だかよく分からないし、「地球儀」は「国際的な儀式」だろうし、あまり適切な例とは言えないかもしれないが、「臍繰り」にいたってはせいぜい「へそのゴマの掃除」だろうか。さらにこれは「綜麻繰り」と書かれることもあるそうであるが、そうなるともう何のことか推測の手がかりさえない。こ

学術情報センター電子図書館サービス　　©学術情報センター

うしたことばの意味がすぐ分かるのは、長年をかけて漢語と意味との連想が作り上げられてあるからであって、それはアルファベット国における単語の果たす役割りと全く変わらない。加藤(1979 [22])が指摘している、アメリカ人の多くが一見むずかしそうなつづりの単語の意味を知らないということは、アルファベット表記そのものの欠陥というよりは、むしろアメリカでは日本ほど初等・中等教育や、特にわれわれが漢字に対して行なっているほどに、語囲についての教育に力を入れていないことが主な理由である。

15　誤字はなぜ起こるか

認知科学的な課題として第三には、誤字の問題がある。

ここで最近採点した、ある大学の学生の試験答案用紙に見かけた誤記の例を少しお目にかけよう。まず、約150人中、3人もが「結果」を「決果」と書いていた。そのほか、「定義」が「定規」、「足りない」が「絶りない」などとあり、そのほかの誤記や誤字も少なからずあった。「最近は国語力がおとろえた」と感じている人たちが見れば、「それ言わんことじゃないか」と言うところであろう。事実、現在の表記法が認められている社会で、しかも大学生の表記能力としては大いに問題であろう。

しかし、こうした事例などからも、意味深い教訓は読みとれるのである。

その第一として、もしわれわれが漢字かな交じり表記ではなく、表音表記を使っていたならば、これらの誤りは起こらなかったであろうということである。すなわち、かれらは「けっか」「ていぎ」「たりない」ということばは確実に知っており、かれらの誤りは読み方を知っている別の字をそれらに当てはめたということなのだから、たとえば初めからかなのような表音表記文字を使っていたなら、こうした誤りは起こらなかった。つまり、ことばと表記とは別ものであり、ことばを知っている人間がそれを自由に表記しにくい表記法がいま使われているという、二者間の格差が誤りを生み出しているということである。

第二に、よく言われているように、漢字は見てなんとなく意味が分かるということは、同時になんとなく分からないということでもあることを、これらの例は示している。

「結果」が中立的な印象を与えるのに対して、「決果」のほうは何となく主観的な意志が働いているようにも感じられ、もしいま新しく表記を考えるのなら、実験などの報告についてそんなことがあっては困るが、場合によっては「決果」でもいい場合があるような気がしないでもない。別の大学の学生の話しによると、講義のときに、ときたま「決果」と黒板に書く先生がいるとのことだから、上の3人の学生がこう書いたということは、あるいはこちらでも誰かの影響があったのかもしれない。

「定義」に対する「定規」のほうは、ふつうには発音が違うが、しかし辞典をみると「定規」は「ジョウギ」のほかに、意味することによっては「テイキ」とも読まれるから、特に関東では発音が濁って「テイギ」になってもそうおかしくはない。また意味的にも、「いみ」の「義」に対して「きめ」の「規」を使ったのも、何となく意味が通っている。特に最近の数学などにおけるように、定義そのものが抽象化し、それ自体からは必ずしも直観的な意味が見えず、むしろ恣意的とさえ思える形式的なものが多いときには、「定義」よりも、かえって「定規」のほうが、漢字一つ一つの意味との親和性が高いとさえ言える。

ちなみに「定」の字の古い形は家を示すウ冠のしたに「正」だったそうであり、その「正」が、「足」の意味の「疋」と同類の、「足」の略形になっているのは、誰かがどこかで書き誤ったのが、いつのまにか定着したにすぎないのだろう。

最後に「絶りない」であるが、書き手には、「足りない」に比べて絶対的に足りなく、ほとんど「絶えて」しまったとの思いがあったとも考えられる。

これらや、すでに述べた「独壇場」の例が示すように、漢字は見て何となく分かる連想があるということを認めるのなら、同時にそれが災いして、何となくいい加減な用法がなされ、同時に情報の伝達が不明確になる恐れがあることも認めなくてはなるまい。

ついでに書いておくと、「ケッカ」にはふつうの同音異義語として、ほかに「欠課」「決河」などがあるが、あまり使われないし、また「定義」には「提議」、「定規(テイキ)」には「定期」、「定規＝定木(ジョウギ)」には「情宣」と、漢語にしては同音異義語の数が少なく、また「定期」以外は使用頻度も少ないものばかりであることも、こうした書き違いをさそったのかもしれない。いずれにしろ、漢字を二つ並べて安易に新語を作ることにより、同音異義語をふやしてしまうことがなければ、日本語とて表音文字表記で充分なことが多いことは、よく考えて見るべきであろう。

学術情報センター電子図書館サービス　　◎学術情報センター

さらに、フロイト(Sigmund Freud、1856-1936)に発する精神分析学によると、人間のなす錯誤は偶然ではなく、その背後には無意識の思考が隠されている場合いが多いということであるから(Freud 1904 [85])、表音文字表記と漢字表記との認知過程の差については、いままでのように、漢字かな交じり文で育った人たちを主とした被験者を用いた片寄った実験と、単なる印象論的主観とによる、読みやすさなどの議論を越え、今後はもっと科学的、客観的に、実験心理学などのパラダイムを駆使して、たとえば上のような錯誤を犯す過程の、定性的、定量的な議論にまで研究を進めるべきである。

16　意味の拡張性の欠除

認知科学的な課題として第四に、漢字では一つ一つの持つ意味が狭く限定されてしまっているので、漢字語ではかえって困るということがある。

元来ことばには、必要に応じて、その意味が拡張されていくという性格があるが、漢字語ではそれがむずかしいことが多い。たとえば「anemometer」はもともとギリシャ語で「呼吸、風」などを意味したことばと「計る」とから作られ、風速計を意味し、日本では「風速計」「風力計」などと訳されて使われてきた。しかし英語でも、またロシア語などでも、いまではその意味が拡張され、液体の流量を測定する「流量計」の意味でも用いられている。これなど、日本語では別の訳語にしなくてはわけが分からなくなるであろう。

また言語学では「phoneme」は「音(phone)」から作られたことばで、「音素」と訳されている。しかし、聴覚障害者の用いている手話(sign language)が、1950年代の言語学的研究によって、自己完結性をもつ、りっぱに独立した言語であることが認められた。それで手話の基本構成要素である、手の形、位置、運動などは、やはりphonemeと呼ばれるようになった。これなども「音素」では、やはりしっくりしないであろう。

このほか、かつては言語に関して用いられていた「context(文脈)」も、いまでは文章(text)を離れ、マルチメディアや作業過程をはじめ、考察の対象が系列や集合体の一要素であるときにも使われるようになり、ついには、まわりの「situation(状況)」と同じように使われるようにさえなってきた。同様に「ethnography(民族誌学、民族学)」は、直接ある人種についての考究を意味していたものだが、今ではもっと広く、異なる人種に対応できるような状況を考慮した技術開発の観点などをも意味するようになった。アメリカで近ごろはやりの、日本料理を含む「ethnic cuisine」は、そうした用法の流れに沿っているが、これも「ethnic」の直訳「人種的」を使って「人種的料理」としたのではぴんと来ない。せいぜい「民族料理」にしなくてはなるまい。

同じような拡張は、近ごろのはやりである「ecology(生態学)」についてもみられ、たとえばコンピュータ機器を使うときなどの環境条件を示すときに「ecological」といった形容詞が使われている。これも、「economy(経済学)」と同じく、もともとはギリシャ語の「oikos(家)」から出ていることばだが、漢字を使っているために、どちらも訳語の意味が狭くなってしまい、新たな意味の拡張にもついて行けず、また「生態学」と「経済学」との関連も明きらかでなくなったことばの例である。

このように、ものごとの概念を大きく掴む性質をもつ英語ややまとことばの単語にくらべて、漢語は意味の細分化の傾向が強く、それが思考の自由なはばたきに枠をはめることが多いから、それに伴って独創的な思考の飛躍ができにくくなっていると考えられる。

そうした例は枚挙にいとまがないが、たとえば「vocabulary(語彙)」がある。この語が「ダンスは彼のvocabulary of expressions(表現手段)の一つだ」といったぐあいに、抽象的な比喩に使われたりすることのあるのは言語の常として珍らしくない。だが具体的にも、最近ますます一般化しつつあるスポーツの一つとして、飛行機から集団で飛び出し、パラシュートを閉じたままの落下中に、個人が単独でいろいろの造形運動をしたり、また集団としての演技を競い合うといったスカイダイビングでは、そうしたときの体の運動の基本型を一括して、「action vocabulary」と呼んでいるようである(1996年10月10日、ワシントンD.C.で視た、スカイ・サーフィンのコンペティションに関する、あるテレビ局の番組で)。この場合にも、「vocabulary」という、もともとは言語学上の用語の直訳を使って、これを「動作語彙」などと呼んだのでは、あまり気持ちがよくないであろう。

なお「vocabulary」に関する同様の使い方は、米国ナショナル・フットボール・リーグ(NFL)の1996年シーズンの決勝戦で、フィラデルフィア・イーグルズ対サンフランシスコ・フォーティナイナーズのゲームのラジオ中継放送でも聞かれたから(1996年12月30日)、もはや一般化していることのようである。

ところが漢字と異なり、カタカナことばとなると、われわれもことばの意味の似たような拡張を平気で行なっ

学術情報センター電子図書館サービス　　©学術情報センター

ている。

ある老人が、車のうしろの窓ガラスのことを「うしろのフロント」と呼んでいた例については前に書いたことがあるが(山田 1991 [71])、われわれが「microwave oven」を「電子レンジ」と呼ぶのは、本当は意味をなさない。

もともとアメリカでガス・コンロ(焜炉)を四つほど束ねて台所に固定したとき、これを「ガス・レンジ(range、並び)」と呼んだものを、日本では意味も考えずに「microwave(極超短波)」を「電子」にし、かつ「range」を「oven」の代わりに使ったものである。(コンロは和語。英語は gas ring または cooking (gas) stove、中国語は煤気竈(méiqìzào)で、煤は石炭のこと。なお竈の簡体字は火偏に土を書く。)しかし日本語としてここまで定着した用語をいまさら誤り呼ばわりをする必要もないだろう。ただ、漢字を離れることによって、いかにことばの意味の拡張が容易になるかの一例としてとらえておけばよい。

ちなみに、ここにおける「電子」の用法も、あまり意味をなさない。「電子(electron)」は物質の構成要素の一つである粒子のことであり、これから「電子技術(electronics)」ということばが生まれ、英語ではそのまま使われているものを、日本では縮めて「電子」にしてしまい、複合語としては意味的に異和感のあるべき、「電子立国」とか「電子文書」とか言うようになったものである。

漢字のこうしたいいかげんな使い方が、われわれの思考方法をもだらしないものにし、日常生活における「建てまえ」と「本音」の二重構造の存在に対しての感覚を鈍らせ、ついには汚職のような道徳的腐敗にまで至らせる影響力となっていないかどうか、もし認知心理学的に調べることが可能になれば、非常におもしろい研究になるであろう。漢字がわれわれの文化を形づくっていると、よく言われることが直感として正しいことを、あるいはこうした現象は示しているのかもしれない。しかし、その深層における正確な内容については、日ごろわれわれは直視することを避けているのではないだろうか。

17　思考の歪曲

以上のように、アルファベットに比べて漢字の使用は、認知心理学的にみてかなり異なった効果をわれわれの思考過程に与えていることがわかる。認知科学的にみた第五の課題として、漢字の使用というものがわれわれの思考の内容そのものを歪めているかもしれないということがあって、これにはすでに心理物理学的な実験研究の報告がかなりある。

ここでは詳しいことに立ち入れないが、まず1920年代から30年代にかけて、「絵画の優越効果(pictorial superiority effect)」といって、われわれがある種の判断をするのに、絵画に頼るほうが言語に頼るよりも速いという現象が心理学者によって明らかにされた。その後1970年代半ばから、パイビオ(Paivio 1975 [90])にはじまり、ベスナーとコルザート(Besner & Colthert 1979 [80])などの心理物理学者によって、絵画による判断は、アルファベット単語による判断よりも速いには速いが、そのとき、たとえば表示図の大小によって、表示された内容の大小に関する判断時間が左右されるということが、実験的に証明された。

これらを受けて、中国系アメリカ人の研究者、ツェングとワング(Tzeng & Wang 1983 [95])たちのグループは、絵画の代わりに、多分に絵画的とされる漢字単語を用いて、やはり同様の実験結果を出した。すなわち、漢字語を通しての判断時には、その見かけによって、表示された内容そのものについての判断に対して、アルファベットつづりの語には出ない、一種のひずみが与えられている可能性があるということである(山田 1986b [67], 1987[68], 1991[71])。

これについて、直観的に分かりやすい具体例を一、二引いておくと、いま職業上の名が某「埜々」というモデルがいる。「埜」は「野」の異体字であり、「やや」と読ませるもののようだが、試みに二、三の人に「野々」と書いたものと比べてもらって印象を聞いたところ、「野々」のほうには粗野とか野性的な感じがあるのに、「埜々」だと何となく可憐さを感じるということであった。2者とも同じ意味の字から成っているのに、こうした印象の違いの出るのは、どうも後者の形が「楚々」を連想させるからのようである。これなどは、文字による遊び、いたずら、ふざけなどと考えられるが、しかし、それが可能だということは、やはり漢字の字づらが表記された意味内容をゆがめて知覚させるということがあってのことであろう。

大学で心理学を専攻し、それに加えて、論理的推論をも柱とする推理小説の大家である佐野洋は、「エンターティメント」が「娯楽」と訳されると「とたんに、言葉のニュアンスは変ってくるように思う。どことなく官能的な色彩が強まってしまうのだ。それは多分、「娯」という字が、女偏を持っているためであろう」と書い

学術情報センター電子図書館サービス　　©学術情報センター

ている(佐野 1974 [30])。これは漢字の視覚的な形態が、それによって記述された内容の判断に歪曲を与えることを言い得て妙であるまいか。

また、前節に述べた「音素」および「動作語彙」の用法例についての、ある研究者の意見だと、「音素」の使用のほうには抵抗があるが、「語彙」のほうは気にならないとし、その理由として、音素に比べて語彙のほうが抽象性が高いように感じられるからではないか、とのことであった。

しかし、語の意味内容からすると、語彙よりも音素のほうが抽象性が高く、一般には語彙の概念よりも音素の概念のほうが理解させるのにむずかしく、またそれだけ時間もかかるのがふつうである。したがって、この研究者の意見は、「音素」および「語彙」という漢語は、それらに含まれている「素」と「彙」の字づらの感じとして、相対的に前者がよく使われる字として具体性を感じさせるのに、後者はそれほど使われなくてよく意味を知らず、漠然としか意味を感じさせないということがあり、そのために表記された内容自体の抽象性に関する判断が影響されていることを示していることになる。すなわち、漢字を用いていると、字づらによって意味内容そのものの理解が影響される、つまり歪曲を受けている、ということのもう一つの例になる。

また前々節ですでに触れた、アルファベット表記と漢字表記とのあいだに出てくる、思考の錯誤の起こりかたの違いも、やはりこの認知科学的な歪曲の反映の一つである。

ちなみに、同じようなことは音声についても言える。一つの文章を異なる人が朗読すると、読み手の発声の個人的な特徴が付加されて、それぞれが異なる印象を与える。また、同じ脚本をもとにしても、(いま演出家の個性までには立ち入らないことにして)、それを異なる俳優が演じると、やはり異なる印象を与える。これらの場合、原典は文章として与えられたものであって、朗読や演技はその「解釈」である。したがって、原作者の思想にじかに触れるには、あくまでも原典によるべきであって、その解釈に頼るべきではない。文字が思考を表現するときにも同じことが言えて、文字は思考内容にできるだけ少ないひずみを与えるものが望ましい(田口 1936 [37] 参照)。

もちろん筆者は朗読者や俳優などの至芸を否定するものでもないし、文字の歪曲効果を有効に活用する文学の創造性を軽くみるつもりもない。しかし、情報のコミュニケーションの道具としての文字は、やはりそのような歪曲効果のできるだけ少ないものが良いと考える。事実、文学作品などでは、こうした漢字の持っている一種の心理的な歪曲や誇張効果が、しばしば巧みに活用されていることはよく知られている。しかし哲学や科学など、完全に純粋な思索が要求される分野において、非常にデリケートな思考を行なっている最中に、使用文字の外観によって、思索内容そのものが影響されるというようなことが、もし確実に起こっていることがよく管理された実験によって定量的に証明されたとすれば、それは文字の性質としては望ましくないことであると言える。

こうした、思索の補助として使用する文字自体が大脳において別の連想を引き起こすとともに、音読み訓読みや同音異義語の判別など余分な処理を行なう必要があることによって、主体であるべき思索そのものが影響されるといった現象に似たこととして、これから次節で述べる、漢字の機械可読化、すなわち入力過程において、使用する入力方式によっては、そこに関与する連想が作業の邪魔をすることが認知科学的に確かめられているという事実がある。それについては、簡単にではあるが、やはり次節で触れることにして、ここでは日本における漢字の使用が思考過程に与える干渉の影響の可能性を示唆する事象について、もう少し考察しておく。

それは国際的に比較すると、一般に日本人は相対的に話しが下手であり、特に外国語を話すことにおいては国際的にその悪名が高いという事実である。それで長年の知己である、ある中国系の研究者が、ある国際会議で筆者に「日本人は民族として失語症の遺伝子を持っているということはないのか」と問いかけたことがあった。

たしかに日本人の外国語、たとえば英語の能力は、一般に世界の水準よりはるかに低いと考えられている。それについては、日本人には外国語に関して精神分裂病的な心情があって、一方においてそれをうやまいつつ、他方においてそれを嫌悪する、心の中の葛藤から来ているという考えがある(岸田 1997[23])。しかしアメリカと日本の政治家などがそれぞれの自国語で演説をしているさまなどを見ていると、彼の疑問ももっともに思えるほど言語能力の差がみられる。したがって、ことは英語だけに限られていないことがわかる。

といって、それは遺伝子などによるわけではなく、可能性としては、われわれが改まって話しをするとき、幼

学術情報センター電子図書館サービス　　©学術情報センター

いときから一つ一つの漢字について、音読みと訓読み、しかも借字読み、当て字読みなどを思い浮かべ、しかもその上に、限りなく多い同音異義語の弁別のためにもいちいち漢語表記を思い浮かべるといった繁雑な操作を経て発話をするということをし続けてきたので、思考の概念から直接発話へと持って行く能力が充分に発達していないということがあるのではなかろうか。

これは人間の思考と発話とのあいだにある過程の認知科学的研究のテーマとして、たいそう複雑で取り組みのむずかしいものとなると同時に、奥が深く、はなはだおもしろいものとなるであろう。

同じ漢字を使ってはいても、中国ではほとんどの字は読みが一つしかなく、また同音異義語の数も日本よりは少ないことに注意すべきである。これは日本人と中国人の外国語に関する言語能力を比較するときに考慮に入れなければならない要素の一つであろう。

ちなみにアルファベット語としても、形から意味内容が印象的に示唆されるといった効果は皆無ではない。たとえば「eye」から顔が、そして「locomotive」から蒸気機関車がなんとなく連想されるというようなことは、むかしから言われて来た。しかし漢字に比べてその度合いはずっと少ない。したがって純粋に情報伝達を託す道具としては、いらざる連想を引き起こして思考を歪曲する分だけ漢字はアルファベットに劣っているということになる。

ツェングとワングらの報告には、そうした重大な意味が含まれているにもかかわらず、漢字は見てぱっと意味が分かると主張する人たちが、その理解過程が思考のひずみを伴っているものかどうかの研究のほうには、さっぱり関心を示して来なかったのは、漢字貴しの先入主に由来する思考の片寄りがあって、表の効果は強調しつつ裏の効果は無視するという、自己欺瞞ではなかろうか。

さらに前節でみたように、漢語の使用はものごとの概念を大きく掴んで一般化することには、あまり向いていない。ところが独創性は、しばしば一つの概念の飛躍的拡張によってもたらされるようだから、こうした制約のある漢語を使っていることが、われわれの独創性の発揮に影響してくるかどうか、はなはだむずかしい問題ではあるにしろ、独創性がなく、ものまねが上手なだけだと、国際的にしばしば評価されている日本人にとっては、心理言語学や認知科学の研究課題として、大いに関心のあるところであろう。

18　日本文の入力処理

以上では、日本語が世界でもっと認められ、かつ広く使われて行くためには、その前にもっと民衆化する必要があること、そして具体的には、その根底にある漢字の問題とどう取り組んでいくべきかを述べてきた。しかし過去の文化との連続性を保つ必要があるから、この問題は一朝一夕に解決できる見込みのあるものではなく、日本語の表記のあり方の理念として、今後長い時間をかけて取り組んでいくべき性格のものであろう。

一方、社会の情報化・国際化はいま急激に進んでおり、短期的には、現在使われている表記法で書かれた日本語を情報処理に必要な機械可読の形にすること、すなわち入力作業を高速化し、諸般のオフィス作業を高能率化する手段の開発と普及とが必要である。

漢字かな交じり文の存在が日本のオフィスなどの生産性を低いものにしていることは、コンピュータの発明とともに始まったことではない。英文タイプライタが1874年にアメリカで発売されたあと、すでに今世紀の初めにアメリカに赴いた多くの日本人が、タイプライタがオフィス作業の組織化や生産性の向上に大きく貢献していることを見て驚き、また感に入っていろいろと報告している(Yamada 1980 [99] 参照)。

その後、数多くの有識者の努力の結果、まずカナ文字タイプライタが生まれ、つぎに漢字3千字ほどを納めた邦文タイプライタが生まれたが、前者では日常に使うふつうの文章が打てず、後者での打字能率は英文タイプライタの数分の1にしかなり得なかった。さらに、1950年代にコンピュータが実用化し、70年代にパソコンが普及するに及んで、文書処理の能力における彼我の格差はいっそう大きくなっていった。

日本でも、日本語をかなで入力し、コンピュータのプログラムによってそれを漢字かな交じり文に変換する方式などの研究は、遅くとも1960年代には始まっていたが、初めてそれが商品化されたのは1978年になってからであった。

同音異義語の存在が通常のことであり、それらは漢字によってやっと区別し得ているのであるから、日本文をかなで入力するときには意味情報の大幅な喪失を引きおこす。そのうえ分かち書きの習慣がないために、かながべた打ちに入力されるのだから、それから漢字かな交じり文を復元するには、コンピュータによるかなり複雑な作業を必要とする。多大の努力の結果、それを成し遂げたこと自体は称賛に価するものの、同音異

学術情報センター電子図書館サービス　　©学術情報センター

義語の誤変換は完全に排除できないから、作業時には出力表示スクリーンの注視や、わずらわしい訂正作業はまぬがれ得ない。かてて加えて、プログラムによる変換作業は、どうしても文字使いなどを規制化してしまい、その点でも不満が残る(山田 1995 [74], 1996 [75])。

そのほか入力速度が本質的に英文なみには上がらないから、日常使われている入力作業がほとんど変換方式になったいまでも、入力サービス業者など高能率が欠かせないところでは、依然としてほかの入力方式も使われている(情報処理学会 1989 [35])。しかしながら、生産活動のコストダウンや、ホワイトカラーの生産性の向上を目指す団体などの調査報告などでも、不思議と文書の入力や処理の問題に触れていないものが多い(日本能率協会 1975[49]、日経連広報部 1993[46])。また学会などでは、近年、学会誌の編集発行が財政上の大きな負担となっているのは周知の事実であるが、ここでも欧米に比べて入力コストの高いことが重い課題となっているのに、その根本に立ち返って、入力方法の研究と基礎的に取り組むことは、大勢として、もはや過去のものとなったと考えられているようである。

いままでに本稿では、漢字かな交じり文の入力は英文入力に比べてかなり遅いということを、何回か述べた。しかし、実のところそれは全くは正しくない。「2ストローク入力」「直接入力」「コード入力」「無連想入力」「反射入力」などと、いろいろな名前で呼ばれている入力方式を用いることにより、はじめの技能習得にみっちりとした正しい訓練が 200 時間ほど要るが、そのあとでは、ふつうの文章、特に自分の専門分野の文だと、1 分間あたり、漢字かな交じり文を 200 文字以上も入力できるまでに腕があがった人はかなりいる。情報量で比べると、これは英文で 1 分間あたり 80 語から 100 語に相当する。中には直接入力方式が不得意とする、不特定多数の人たちの書いたものを入力する入力業者のもとで働いているオペレータにも、そういった人たちがかなりいる。

直接入力がうまく機能する理由の説明はかなり長くなるので、関心を持たれるかたは、ほかの解説文献をお読みいただきたい(たとえば Yamada 1983 [100]、山田 1985 [65], 1995 [74], 1996 [75])。いまこの入力法の原理だけを書けば、ふつうのキーボード上の 40 ほどのキーの二打鍵の組み、約 40×40=1600 をとり、それらに異なる基本文字を割り付け、言語的連想を避けてそれを丸暗記したうえで、その二打鍵の組みを「コード」として使って入力する、というものである。ふつう個人が文書を作成するときには、専門家であっても漢字 800 字、多くても 1000 字ぐらいしか使っていないことが分かっているから(情報処理学会 1989 [35])、個人用としてこれは充分な数である。

それでも、(1)そんなコードが全て暗記できるか、(2)かりに暗記できたとしても英文タイプほどの入力速度が出るか、という二つの疑問が出る。正直のところ、筆者もこの方式のことを初めて聞いたときには、常識的にそんなことは不可能だと思った。しかしあとになって分かったのだが、世の中にはそうした常識を無視し、独立して実験を行ない、実用化した人たちが、筆者の知る範囲でも 3 人はあったのである。直接入力法のそうした反常識性が、アメリカ在住が長く、高能率な日本文入力法の必要性をひしひしと身に感じていた筆者の関心を引き、そのさらなる洗練と、認知科学的理論付けとの研究に向かわせたのであった。

そうした認知科学的な研究の結果、いまでは直接入力法の人間科学的基礎がかなりはっきりとし、単に速いばかりではなく、実は変換入力法などよりもずっと人間の本性との整合性がよい、自然で楽な日本文の入力法であると言えるようになった。

ここで一つたいせつなことは、この二打鍵コードを覚えるのに、言語的連想を用いてはいけないということで、もしそうした連想を用いると、文字の視覚的入力に応える指の打鍵運動が直接的でなくなり、余計な思考の介在によって、入力速度が下がるばかりでなく、精神的ストレスが増える(詳しくは Yamada 1983 [100]参照)。筆者の考えでは、前節で触れた、日本人の話し下手、特に外国語下手には、これと同じような理由が関わっている可能性があるのではないかと思われる。したがってこの 2 者は、認知科学的研究のパラダイムとして取り組むに値するものではないかと考える。

日本の社会は、多くの組織で職場間にある、定期的な人事異動ということでも分かるように、専門家の集団よりも、むしろ「万能選手」たちに仕事をまかすことを尊ぶ「アマチュア主義」の色彩の強い社会であるから(山田 1994b [73])、組織の一部の人たちにまず特別な訓練を施してから、入力業務に専念してもらうというようなことはあまりしない。また一般の人たちも、実は成人するまでにおそらく 1 万時間にも及んでいる、漢字をうまく使いこなすという能力の獲得過程で息切れしてしまってのことなのか、その上さらに数百時間をかけて直接入力法を身につけるということに対しては、ほとんどがかなり強い拒絶反応を示す。

学術情報センター電子図書館サービス　　© 学術情報センター

そうしたわけで、自分から進んで直接入力法を身につけて使っている人たちはすでに少なからずいるのだが、どちらかというとかれらは好事家の集団となってしまい(E メール・アドレス・リスト tcode-ml@is.s.u-tokyo.ac.jp)、その使用を一般に向けて広める努力をしようとは考えていないよう見える現状は、筆者にとっては残念に思えてならない(山田 1996 [75] 参照)。

しかしながら、第2次大戦中の日本では兵器として、たとえば三菱重工業(株)などのようなところで、少数の天才が、少数の天才しか乗りこなせない零式艦上戦闘機のような逸品を作っていたのだが、アメリカでもドイツでも、少数の天才が多数の凡人の乗りこなせるものを作っていたところに、日本の敗因の一つがあったようである(奥宮 1979 [15])。

また、血の滲むような猛訓練を経た少数の精兵と、世界最上の光学兵器および射撃盤とを載せ、30km の距離から初弾を命中させる確率が50パーセントと、自信満々であった日本海軍の(黛 1972 [55], 1977 [56]、生出 1988 [13])、世界最強の戦艦「大和」も、その唯一参加した1944 年 10 月 25 日のフィリピン沖の対艦砲撃戦における、アメリカの空母「ガンビア・ベイ」に対する砲撃では、世界最大の 46 センチ主砲を 100 発も射って、命中弾が 1 発もなかったという(小倉 1997 [16])。

このように、技術の粋とも言うべき光学装置や、また別に部品総点数が数万にも及ぶ精巧無比の機械的計算機だった高射算定具なども、所詮はエレクトロニクス技術によるレーダーやソーナー、それに近接信管などの敵ではなかったのである(NHK 1994[12])。

その後の日本は柔軟性の高いエレクトロニクスに力を入れることにより、精密機械工業だった時代にはどうしてもかなわなかった、ドイツのカメラ産業を圧倒してしまい、さらにはテレビやビデオ・デッキ産業などでも世界を制覇してしまった。

こうしたことを考え合わせると、筆者としては、変換入力法がかなり使えるものになった今日、いまさら利用者の初期訓練が欠かせない直接入力法でもないのかなとも思う。

同時に、上記の機械技術とエレクトロニクス技術との性格の違いは、いみじくもまた漢字とアルファベットの能力の差をも暗示しているかのようである。すなわち、表記法は天才やエリートだけのものであってはならず、民衆が気楽に使いこなせるものに変化させて行くことこそ、民主主義化した情報化社会のとるべき道であろう。

にもかかわらず、処理能率が世界一悪いとされるいまの表記法を使っているうちは、やはり一部の専門職が直接入力法のようなものを使い、相当量の初期訓練を行なった上で実務における入力能率を上げ、国際的な情報化社会で漢字国日本が立たされている不利を補ってやる必要があるとも思えるのである。

なお、慶応義塾大学環境情報学部の大岩元教授の研究室にも、一般人と入力専門家双方を対象にして開発された、直接入力法のための TUT コードについて、タッチタイピング教育を含め、一般的な情報交換および討論のためのメーリング・リストが、1997 年 1 月から開設されている(参加用アドレス：tut-code-request@crew.sfc.kio.ac.jp)。

19　入力法について何をなすべきか

以上において述べてきたのは、社会の情報化にこたえて日本語の処理を高効率化し、また国際化するための第一課題として、過去と未来とをはっきり分けて考え、未来に向けて表記法そのものを民衆化すること、および文書処理能力の向上の具対策についての私見の一端であった。だがそうした思いきった方策には、過去の文化との連続性という観点からする強い反対もあって、実行はかなりむずかしいであろう。

しかし、もしいまそうした壁を越えることを諦めたとしたら、来たる 21 世紀の半ばには、日本が世界においてほとんど影響力のない小国に転落してしまっているであろうことは、ほぼ確実である。乏しい筆者の体験に照らしてみても、すでにその兆しは、アジア諸国においてあれこれと見られるようになってきた。

すなわち技術力においては、すでにしばらくまえから、情報化社会の基盤であるマイクロプロセッサ・ユニット(MPU)については、世界中がアメリカのインテル社の軍門にくだったに等しく、1995 年には基本ソフトにおいてマイクロソフト社の「ウィンドウズ 95」の前に手をつかねて追従した。そうした動きの結果、コンピュータのハードウェアおよびソフトウェアに事実上の国際標準規格ができあがるにつれて、東アジアを中心とする新興国もまた工業生産の技術力を身につけ、いまや日本で製造されるパソコンの各種部品は総額の 70 から 99 パーセントが海外からの調達品となっているという。

また、たとえば最新式の CDMA(code divided multiple access、符号分割多重接続)方式の技術によって、21 世紀には全世界のどことでも接続できるようになる携帯

学術情報センター電子図書館サービス　　© 学術情報センター

電話システムは、いま地球上の 7 か所にハブ局が建設中であるが、アジア地区での第一番目の局の建設は日本でではなく、隣の韓国で進められている。韓国ではすでに 2 社が CDMA 電話のサービスを開始しているが、いまのところ日本でのサービスは早くて 1997 年のこととなっている。

そのほか、情報化社会の活動一般において、ますます膨大化してきた各種の文書の管理のために、アルファベット国では、アメリカはもちろん先進国以外の国ぐにでさえ SGML(Standard Generalized Markup Language) の使用が成熟化してきているようなのに比べると、日本では文書の入力や漢字の処理がめんどうなことがあって、掛け声だけは高いものの、実施は大幅に遅れている。

また土地や公共サービスについての情報や国際的な貿易輸送情報などの、政府による管理が、アルファベットで行なわれているシンガポールなどの現状に比べて、日本での実施ははるかに遅れていると言える。

また経済面においても、アメリカの株価の上昇つづきにもかかわらず、日本の株価は低迷を続け、いまや世界の資本は日本を見限り始めたかの感がある。海外企業の在日拠点は東アジアの他国へと移動を続け、たとえばシンガポールやホンコンは、もはや日本を経済政策の重点対象とは考えていないという。

1984 年、筆者が短期の客員研究員としてアメリカのスタンフォード大学に滞在していたとき、依頼を受けて日本の産業について講演をしたことがある。[その内容については、同様の趣旨のものが Yamada(1991 [101]) にまとめてある。]当時のアメリカでは、日本の産業の目覚ましい発展ぶりに対して、日本は自からの独創力を発揮せず、欧米のアイディアを利用して稼ぎまくっているという、いわゆるアイディア只乗り論が、多分に非難の意味をこめて語られていた。

したがって講演のあとでは、聴衆の大半を占めていた大学生から、この只乗り論が出され、日本はもっと自からの創造力で世界に貢献すべきではないかと迫られた。そのとき筆者は、多分に皮肉を込めて「それでなくてさえ日本の産業は、いまアメリカの産業にとって大きな圧力となっているのではないか。その上もし日本が、創造性においてまでアメリカを凌いだとしたら、アメリカの産業は日本に完全に屈服しなくてはならなくならないか」と切り返したものである。聴衆は一瞬静まりかえり、しばらくは発言する者がなかった。

そのころからアメリカの全体でも、この問題について深く考えはじめ、日本に対して徹底した知的所有権の主張をするようになっていったのである。

その日本が現在では電子技術において台湾に、半導体製品で韓国に、組み立て産業で中国に圧迫されつつある上に、肝心の知的財産を生み出す研究能力においてさえ、国際的に大きく門戸を開いているシンガポールによって急速に追い上げられつつある。

現在のように、国際的な標準化の結果、国際的な技術移転がふつうになってきた状況の中にあって、われわれが工業生産力において国際的な競争力を維持して行くには、どうしても幅広く、技術以外の分野でも能率をあげて行かなければならない。オフィスの生産性はその重要な一角である。

とは言っても、文化に深く根を下ろしている表記法の抜本的な改善には、息の長い取り組みが必要である。したがって、いまの表記法から第二義的に派生してくることであるとはいえ、短期的に重要な課題として、日本文の入力の効率化の問題とは積極的に取り組む必要がある。しかしすでに述べたように、それに最適と思われる直接入力方式が、社会で広く受け入れられる可能性は、将来といえどもきわめて小さいであろう。したがって、以下では次善の策として、現行の変換入力方式の活用という制約の中で、その改善について、使用する側に立っていくつかの対策を考えてみる (Tatuoka & Yamada 1995 [93]、Yamada 1980 [99], 1983 [100]、山田 1985 [65], 1995 [74], 1996 [75])。

大きく分けて、そこには互いに深く関わり合っている課題が四つほどある。すなわち、(1)入力に用いる表音文字をかなにするかローマ字にするか、(2)キーボード上の文字配列をどうするか、(3)キーボードの形状や大きさをどうするか、そして(4)技能習得のための教育をどうするか、などである。

まず、日本語の入力なのだから、かな入力でと考えるのは自然である。しかしキーボードを見ないで打つタッチタイプを楽に行なうためには、キーの数が 30 ぐらいが理想であるのに、ローマ字だと問題ないが、かなの数はそれを大きくうわまわっている。さらに、これからのいよいよの国際化とともに、オフィスでも、また個人生活でも、たとえば英文のように、アルファベット書きの文書を作成する必要はますます増えてくることを考えると、かな入力のほかにローマ字入力も欠かせなくなるから、一つで両者が処理できるように、ローマ字を使った変換入力を採用したほうがよいと思われる。

学術情報センター電子図書館サービス　　© 学術情報センター

つぎに、タッチタイプの技法を能率よく、悪い癖を身につけることなく習得できる、よい教育法を開発する必要がある。この問題は、実は次節で取り上げる、キーボード上の文字配列や、キーボードの形状や大きさと深く関わっているが(Yamada 1980 [99], 1983 [100] 参照)、ここではそこまでは立ち入らない。

それでなくても、どういったテキストをどういった順に練習するかについてでさえ、まだ最適な方法が確立されているわけではない。たとえば、一番初めに指の練習を始めるのに、無作為的なつづり、いわゆるナンセンスつづりから始めている教師もあれば、それでは誤ったつづりを打つ癖をつけているようなものだから、意味のあるつづりだけを使って練習するのがよいという、いまではだんだんと受け入れられてきている主張もある。その場合でさえ、左右の手の運用の交互性の円滑化を習得し、かつ母音字同士間の混同を起こしにくくするようなテキストの構成法の必要性などについては、いまだに定説がないに等しい。そしておそらく、これらは習得者の素質にも依存している可能性があることも調べてみる必要がある。

20　キーボードの問題

変換入力方式として、かな入力とローマ字入力のどちらを選んだとしても、まだキーボード上の文字配列をどのようにするかという問題がある。現在広く使われている、上段(下から第3段目)左端の配列にちなんでqwertyと呼ばれるローマ字配列が、英文タイプにとってさえ、あまり良くないものであることは知られているし(Yamada 1980 [99])、日本文のローマ字入力用としても最適ではない(Yamada 1983 [100])。一方、1930年代にワシントン大学のドボラク(August Dvorak)教授によって開発され、アメリカの国家規格協会(American National Standards Institute, ANSI)がqwertyにつぐ代替標準としているドボラク配列は、英文はもちろん、日本文の入力にとっても、ずっと良いものであることが分かっているが、すでにqwertyがあまりにも普及していることなどもあって、商品として入手できなくはないものの、ほとんど使われてこなかった。しかし最近ではqwertyが、このところ急に増えた、手の腱鞘炎などの原因となっていることが分かってきたので、手に対する負担の軽いドボラク配列のキーボードが注目を集めているようである(山田 1995 [74] 参照)。

日本でいまいちばん多く使われているかな配列は、古くから使われていたものがそのまま規格化され(JIS / B9509)、その後いくらか改訂されたものであるが、日本語入力にはあまり向いていない。それで新しく配列を合理化した新 JIS(X6004)規格のキーボードが 1986年に制定されたが、その前後の取り組み方が中途半端であったために、ほとんど普及していない。

このキー文字配列の問題は、実はキーボードの大きさや形状の問題と深く関わっている。現在のキーボードの大きさは、はじめてタイプライタが生産されたとき、大の男がキー上の文字を目で探しつつ、2本指で打つことを前提にして決められた。それが広く普及したあとで、手の小さな女性が職場に進出し、さらに10指を使い、キーを見ないで能率よく打つタッチタイプが普及した。したがって使用文字数の少ない英文を速く打つのにも、現在のキーボードは一般には大きすぎるのだが、すでに大量に出回っていることを口実に、かなりお役所仕事的な作業のもとで、それがそのまま国際規格にまでなってしまった。

またその形状にしても、もともとが目視による索表的な操作を前提としたものを、19世紀の未熟な機械工業の生産技術に制約されて作ったから、キー配列がいまのように、上端に向かって不規則に左にずれた平行四辺形型に納められたのである。

しかし製造技術が発達し、充分な経済性を持って少量多品種生産が行なえ、しかもエレクトロニクスの技術が形状設計の機械的制約を大幅に緩和してしまったいま、同じ形、同じ大きさ、同じ文字配列の規格品キーボードを万人に使わせるということは、体格に個人差が大きく、異なる言語を持ち、また使用目的も多様化している世界中の人間を、むりやり一つの機械に合わせるという愚行である。それは当然個人ごとに合わせるべき洋服や靴を、大きさもファッションもただ1種類に統一するのと軌を一にしている。統一され、規格化されなければならないのは、こうした人間とのインタフェースのところではなく、個人の体格や使用目的に合わせて自由に設計され、かつ簡単に携帯できるようにした個人用キーボードと、それで行なう入力情報を処理する本体装置一般の結合部との、機械的かつソフトウェア的なインタフェースのほうである。

一時期広く世に受け入れられた、あるワープロ商品用の「親指シフト」キーボードの配列は、手にあまる大きさの規格品キーボードを、物理的にほとんどそのまま用いて、かな入力をする困難を緩和する目的で考え出されたものである。ほんとうなら、キーボードの大きさや形を手に合わせることこそ、まずとるべき問

学術情報センター電子図書館サービス　　©学術情報センター

題解決の手段であったろう。

そのような場合に、結合部インタフェースの標準化が必要となることは、研究者によってたびたび指摘されてきた。しかしメーカー側は一度ついたユーザーを自社製品に繋ぎ留めておく手段の一端として、いまだにインタフェースの標準化には甚だ消極的である。

ちなみに、コンピュータなどの操作にあたっては、近年キーボードによる命令の入力を離れ、表示スクリーン上においてメニューとポインタとをマウスで操ることが多くなり、1995年に発売された「ウィンドウズ95」においては、ことここに極まったの感がある。しかしマウスの使いすぎはキーボードの使いすぎよりも手の労働障害を起こしやすいらしいことが、最近アメリカなどでぽつぽつ報告されているという(Unger 1995 [97])。したがって操作に熟達した人たちに対しては、マウス一辺倒ではなく、キーボードからの操作も可能なソフトウェアが欲しいところであるが、まだそのように配慮された商品はない。

また老人にとっては、キーボード操作の習得は若い人たちよりもむずかしいとされているが、60歳を越えた人たちでも、実用上充分な技能の習得が可能であることも報告されている。しかし、歳をとると手の運動の制御精度が下がり、現行のソフトウェアではマウスによるポインタの細かい操作が困難になるということが報告されている(吉田 1996 [76]、pp.64-66 参照)。これなどは、マウスによるポインタの移動量を、ユーザーが相対的に大きくできるようなソフトウェアで解決できる問題であるが、いまのところそうした商品は販売されていないようである。

こうした入力法や入力機器にかかわる問題は、かつて今日の情報化社会の到来の必然性を予期し、それがいまのような奔流となるまえに、組織的に手を打っておくべきことと考えて、筆者なども警鐘を鳴らすべく努力をした一人であったが(たとえば Yamada 1980 [99], 1983[100])、いまにいたるまでその影響がほとんど見られないで来たのは残念である。

したがって、そうした問題を含めて、入力機器に関する開発の多様化やそれらについての啓発運動においては、これから実行して欲しいことがまだまだある。

21　入力に関わる学校教育問題

最後に日本文入力に関わる、小・中学校での教育問題について少しばかり考えてみる。

まず入力におけるつづり方であるが、日本語の漢字かな交じり表記には簡単に呑みこめるような正書法なるものはないし、また目安にしてもあいまいさが多い。それに比べてローマ字表記にはかなりきちっと正書法が定められてある(Tatuoka & Yamada 1996 [94])。

また漢字かな交じり文のつづり方のほうは歴史的な尾を引いていて、現在の発音とかけ離れている点があるが、ローマ字では構造言語学的にきちっと定められてあり、まだほとんど守られていないとはいえ、それが訓令式として国の省庁における標準であり、また国際標準化機構(ISO)にもそれを国際標準として認めさせてある(ISO 3602-1989)。それとともに、当然のことながら、日本工業標準調査会でも、JISにおける英文規格の場合、日本語の表記は訓令式のみによることが定められ(JIS 1990 [33]、p.17)、即日実行に移されている。

中国においては、長らく使われてきた、英語式のWade-Giles式を排してピンイン式のつづりを1982年にISO 7098 として世界に承認させ、中国人は一般にそれを誇りに思っているようである(たとえば「語文建設」1990年第6号、12月12日刊、p.42、王均・李楽毅の記事参照)。しかし日本ではISO 3602 に対して、マスコミが全くの沈黙を守り続けている(田中・他 1997 [38] 参照)。

日本語は英語とは全く異なる言語である。したがって、アルファベットで日本語をつづるときには、日本語の音韻と形態素との構造に合わせたつづり方を用いるべきであり、皮相的な英語もどきのつづりを用いるべきではない。その2者をはっきりと区別した教育をおろそかにしていると、いろいろな点で英語と日本語の区別もあやしくなってしまい、たとえば英語を日本語式に読んだカタカナことばが蔓延し、国際的に通用しない和製英単語が日本語にますます増えることになってしまう。

すでにそうした例はいくらもあるが、一、二を引くと、たとえば「award (授賞)」がある。これは英語では「アウォード」であるのに、英語と日本語のつづりを混同して、いまでは「アワード」として使われている。また、ある外資系のコンピュータ・メーカーの基本ソフトに「ワープ」というのがあるが、そのもとは「warp (ウォープ)」(=反り、引き綱)であるのに、日本市場におもねてか、わざわざ「ワープ」と呼んでいる。

そんな調子であるから、いま日本文をローマ字で入力しようとすると、第一に訓令式つづりのままで簡単に入力できるキーボードがない。また変換入力ソフトの商品もまだ一つもなく、実際にはかなつづりを1字

学術情報センター電子図書館サービス　　© 学術情報センター

ずつローマ字に直して入力しなければならないものだけである。しかもそのつづり方の方便も千差万別である。ローマ字文は分かち書きされるから、筆者たちの示したように(Tatuoka & Yamada 1995 [93])、訓令式つづりのローマ字文入力から漢字かな交じり文への変換は、現在使われているかな漢字変換方式の場合よりも、変換ソフトがずっと楽に作れる。したがって国の言語政策として、標準つづりで入力できる変換ソフトの供給は、当然奨励されるべきものであろう(村野 1979 [57])。

しかもそれが実現したあかつきには、小学校でローマ字を習うとすぐ、情報化社会で不可欠になるキーボード入力の技能の習得とともに、ワープロやパソコンを使い、キーの数が少なくてすむローマ字文入力からの変換によって、文章作成の技術の習得が格段に加速できることになる。小学生の国語の能力では、まだ自分たちの話しことばのほとんどが漢字では書けないのだから、そうした音声言語的な入力から文字言語が早くから出力できるということは、国語力や文章作成能力を豊かにする教育に、大いに役立つことと思う。

そのことは、出版社の編集者などで作っている「現代子ども国語表現力研究会」が実施した調査によって明らかになった、小学生は低学年ほど、ほかの科目に比べて国語、中でも漢字学習を苦手としており、高学年では作文、つまり表現力の発揮が嫌いになっているという事実からも推測されるであろう(朝日新聞 1996a [1])。

さらに、日本語が話せるのに日本文が書けない外国人にとっても、そうした変換ソフトは大いに役に立つものになり、日本語の国際化にとっても、技術的に有力な援助手段となるであろう(大坪 1992 [14])。事実、そうした目的で作成されたワープロソフトを外国人留学生の日本語教育で実験的に使用している学校では、明らかに成果があがっているという報告がある。

情報化社会において重要となる、キーボード技能の習得に関しては、組織的な教育の提供の問題がある。入力に際して、キーボードを目で見ながら、左右両手の人さし指の2本だけを使って打つ目視法より、キーボードを見ないで10指を使って打つタッチ法のほうが、初期の訓練こそ欠かせないものの、熟達すれば格段に速く、しかもそれ以上に、目や肩の疲れが少なくてすむことはよく知られている。したがってこれからの社会では、万人によるそうしたタッチタイプ技能の習得は望ましいことであり、すでに大学や専門学校などにおける情報処理教育では、どんどん取り入れられている。

しかし今後小・中学校レベルからの情報処理教育やそれに伴う入力技能習得が奨励されるとなると、問題が一つ起こってくる。それは小・中学生がタッチタイプを習うのには、現在市販されているキーボードは、明らかに大きすぎることである。

短期大学や専門学校において情報処理教育を担当している何人かの先生たちから筆者が相談を受けたときの話しによると、小さいうちからワープロやパソコンになじんできた学生ほど、正しい入力法の技能習得に拒絶反応を示すので、よい技能を身につけさせるのに苦労するということである。これは組織立ったタッチタイプの教育法がまだ確立されてなく、また訓練に必要な、手の大きさに合ったキーボードが手に入らないままに、小さい子が我流で規格品キーボードを使い始めることによる。

むかしからタイピストの養成やワープロ技能の検定などに力を入れてきた日本商工会議所(日商)は、日本国民のタッチタイプ技能の増強に深い関心を示し、早くからキーボード操作技能認定の制度化などと積極的に取り組んで来た。さらに高度情報化社会の到来に向けての基礎教育や環境整備の方針策定のため、1993年度には「キーボード操作技能研究委員会」を設置し、その答申を踏まえて、基礎教育の重要課題に関する簡潔な提言を公にしている(日商 1994 [47])。

その提言には、関係官庁をはじめ、業界団体、企業、教育界などに、早期実現のための役割りを期待することとして、(1)タッチタイプ技能の教育・普及、(2)キーボード本体の改善研究、(3)情報機器操作と健康との関わりの研究、(4)情報機器関連の地球環境への配慮、の4項目を掲げている。当然ながらその第二項には、個人の体格と感性に合わせたキーボードの供給と、それらが全機種に接続できるための、インタフェースの標準化の推進とを謳ってある。

さらに、小・中学校で情報処理教育にあたる先生たちの、当座の供給不足の問題に関して、側面から文部省を支援する可能性の一つとして、日商では、すでにかなりよく組織されているソロバン塾とのつながりを活用し、塾の先生たちの教育や、生徒の手に合ったキーボードの組織的な供給などのアイディアの実施の可能性を検討していると聞く。

もしうまく成功すれば、それは一石二鳥の快挙となり得るであろう。しかしそれに先立ち、まず入力にお

学術情報センター電子図書館サービス　　◎学術情報センター

けるつづり方の決定とその変換ソフトの開発、次いで人間工学的なキーボードの設計、そして最適な教育法およびカリキュラムの開発など、先立つ理念の確立や技術の開発などと早急に取り組む必要があることを、いまこそ理解しなければならない。

パソコンの基本ソフト、「ウィンドウズ95」の開発によって1995年以降に引き起こされた大騒ぎに流されて、日本文化の根底にある日本語の問題を根本から考えることなどは、なおざりにされているとしか思えない今日このごろである。応用ソフトの互換性の向上の立ち場からすれば、基本ソフトにいくつかの事実上の国際標準ができること自体は、大いに望ましいことである。それにしても、それらの上に載せる、日本の文化と密接に結びついている日本語処理のための応用ソフトに関しては、上に述べてきたように、それと組み合わせて使うべきキーボードとともに、われわれ独自の立場から早急に最良のものを開発していくべきであろう。

22　おわりに

以上では、社会がますます情報化する中で、日本語が世界的にもっと普及するためには、まず文字というものの理念を明確にし、それに従って、やさしい正書法が決まるような表記法を確立する必要があること、それもわれわれ自身の世代のためではなく、これからの世代のために教育のプログラムを準備し実行しなければならない性格のものであること、そしてその第一歩としては、われわれは漢字かな交じり表記を順次改めて行かざるを得ないであろうことなどについて考察してきた。

しかし、そうした目標を達成するには、一つには一般大衆を対象として、そのような理念に沿った文筆活動が大幅に展開される必要があるが、いままでのところそうした努力をした作家の数はごく限られている。もっとも、漢字貴しの先入主に凝り固まっている大衆を相手に、そうした活動で成功するのは、並たいていの才能の持ち主ではおぼつかないことであろうが。

にもかかわらず、文字というものは、あくまでも情報内容を伝えるべき、いわば入れものである。そのことを忘れて、文字の表現する外的な効果のみに心を奪われ、伝達内容についての精進をおろそかにしているような文芸の創造活動をみると、あたかも幼な児に箱入りのおもちゃを与えたとき、箱に目を奪われて中身を忘れ、夢中になって箱であそんでいることがあるのが思い浮かばれる。

またコンピュータの活用に関しても、日本文化をまっ正面に見据えての発想がたいせつなのは、実は日本文の入力処理に関してだけに留まらない。ここでは詳しいことに立ち入らないが、日本語自体の構文は各種のコンピュータ演算と大きな論理的整合性がある。

すなわち、いま簡単な例として、「あの・机の・上の・リンゴを・とって・居間に・いる・私に・持ってきて・下さい」という命令を実行することを考えてみる。その作業の時間的な系列は、まず「あの」によって机のある方角を決め、つぎに「机の」によって机を特定し、「上の」によって机に対する位置ぎめをなし、「リンゴを」によって行為の対象物を特定し、・・・といったぐあいに、この文の語順のとおりの運びとなり、最後に「下さい」で手渡すことで作業が終わることになる。

もちろん日本語は語順についてかなりの自由度があるから、必ずしも上の順に書かなくても、同じ意味をなす文章はいくつか書けるが、少なくともこの例のように、日本語では複合作業を指定するにあたって、必要となる単位作業の順に、はじめから単語を並べて命令文を作ることが容易である。

ところが、コンピュータがアメリカで一番さきに作られたこともあって、ほとんどのプログラム言語は印欧語の語順である「命令・実行対象1・実行対象2」、たとえば「割る・Aを・Bで」の、いわゆる命令前置形(ポーランド記法)をとるか、「A・割る・B」のように命令中央置形をとり、ほんのわずかの特殊なプログラム言語で命令後置形(逆ポーランド記法)の「Aを・Bで・割る」をとるようになっているから、プログラムの実行にあたってコンピュータは、まずそれらを命令後置形に改める必要がある。しかし、上に出したリンゴを扱う作業のようにこみ入った例では、それなりの手数がかかってしまう。

ちなみに、前置形が「ポーランド記法」と呼ばれているのは、言語学的にポーランド語と関係があるからではなく、たまたまルカシェーヴィッチ(Jan Łukasiewicz)というような、発音するのも書くのもむずかしい名前を持った、ポーランドの論理学者が初めて採用した記法だからである。

したがって、たとえば人工知能の開発の場などでこうした作業の命令プログラムを書くときには、日本人にとってプログラムが楽に書けるプログラム言語が初めからあれば、プログラム書きが自然になり、しかも

学術情報センター電子図書館サービス　　◎学術情報センター

コンピュータにとっても、そのほうが処理が自然で簡単になる。事実、相対的にはまだ簡単な仕様のものと言えるかもしれないが、そうしたプログラム言語とその処理ソフトはすでに存在し、市販もされている(片桐 1988 [20])。すでにちょっと触れたように、実は世界の言語の大多数は印欧語系ではなく、かなりの言語が日本語のような構造を持っているのだから、よいものさえ作れば、そうしたプログラム言語と処理系は、国際舞台においても大いに歓迎される可能性があると思われる。

現在、印欧語が世界の先進国の言語の主力をなしているが、なにもそれは言語構造が本質的に優れているからではなく、いま確からしい学説によると、太古にたまたま印欧語系の言語を持った種族が農業を発明し、それが食糧の確保を可能にした結果、人口を爆発的に増加させつつ地球上に広がっていったからに過ぎないらしい(Renfrew 1989 [92] 参照)。したがって、コンピュータ言語が印欧語系の構文で縛られなければならない必然性はない。

外国によって開発されたソフトウェアにいつまでも日本が振りまわされているのは残念であると、いままでにも創造的ソフトウェアの開発を支援するプロジェクトなどが政府によっていくつか組織されて来たが、それらの努力の主力は依然として欧米産のプログラムの亜流的なものの改善などにあてられている。今後は、真に日本文入力に向いた入力法やワープロソフトはもちろん、上記のような、日本語の構造に向いたプログラム言語やその処理システムなどの開発に積極的に目を向けた、組織的な研究支援体制が確立されるべきであろう。

以上本稿では、好むと好まざるとにかかわらず、いま日本にも国際化を迫っている、世界的な情報化社会の到来にあたって、日本文化が直面している問題のうち、長期的には日本語の表記法の問題、そして短期的には日本文入力の問題に主として焦点を絞り、筆者の考えていることの一端を述べてみた。

文字というものは、われわれの外部に蓄積された膨大な知識とわれわれとのあいだのインタフェースであり、また入力システムはわれわれとコンピュータ、さらにその先にある知識の集積などへのインタフェースである。したがって、これらについての問題の多くは、近年盛んになったインタフェース学の研究分野のものである。同時に、それらの研究に用いられる手法は人間の頭脳の活動と密接に関わっており、やはり近年台頭した認知科学に属している。

こうした分野の研究で、理論的にはすでにいろいろなことが分かり、それを踏まえて、改善策についてもかすかずの提案がなされているにもかかわらず、それらに基づいた具体的な開発や試行がなかなか実行されないのは、実はわれわれの素質や実行能力の欠除によるものではなく、日本文化、特に教育制度に根を下ろしている画一主義、平等主義によって、一般に動的活力が抑制されてしまっていることにあることは、むかしから多くの識者によって繰り返し指摘されてきたところである(たとえば、千早 1982 [40]、生出 1988 [13]、産経新聞社会部 1992 [31]、日下 1993 [26])。

幕末の開国後、工業国として日本が欧米先進国に追いつくには、勤勉で質の均一な労働力の確保が大きな頼みとなったことは明らかである。同時にそれは、他とは異なる個人の活動を抑圧することになり、独創性の発露や社会のすばやい変革を抑えることにもなった(山田 1994b [73])。さらに、一人の人間が個人として獲得できる知識の量には限りがあるから、たとえば教育行政における学習指導要領などによって、国民の知識や考え方が均一化されているということは、その一方で、集団として盲点だらけの民衆がつくり出されていることでもある。いくらたくさん集まっていても、井戸の中の蛙には井戸の外のことは分からない。したがって、個人が多様化すれば、個人の得意とする分野はそれぞれ分散してしまっているから、限られた数の同じ問題によるテストでは、個人の得点、したがってその平均値が見かけの上では下がるのは明らかかもしれないが、それでも教育をいっそう自由化して、社会全体としての弾力性、包括性を増強することこそ、国際社会における日本の活力を真に増大させる道となるであろう。

さらにこの情報化および国際化の時代における国語教育の内容一般についても、漢字重点の教育をはじめ、早急に改められなければならない問題点がいくつか存在する。従来の心情主義・感化主義的な鑑賞を中心とした、受動型の国語教育から脱却し、論理的な思考力と効果的なコミュニケーションを重視した能動型の教育に切り換えることの重要さは、日本言語技術教育学会から中央教育審議会にあてて出された提案(波多野 1996 [51])によっても明らかであろう(山田 1986b [67] 参照)。

またそうしたコミュニケーション重視の教育を推進することによって、よく言われているように、われわ

学術情報センター電子図書館サービス　　◎学術情報センター

れが持つ、民主主義の根幹である交渉能力に乏しく、また国際的な舞台においても、「親善でなくば孤立」といった、両極端の立場しかとれない、日本的な思考形態(日下 1993 [26])からも脱却できそうである。

謝辞

本稿をまとめるにあたって、もと国立国語研究所所長の林大氏、慶応義塾大学環境情報学部の大岩元教授、今は隠退しておられる化学者の矢野雄司氏、市立前橋高等学校教諭の木村一郎氏にいろいろとご教示をいただいた。また同上の林大氏、東京大学名誉教授の柴田武博士(言語学)、オハイオ州立大学教授の James Marshall Unger 博士(言語学)、東京理科大学教授の清水忠雄博士(物理学)は本稿を査読下さり、かずかずの貴重なご意見をいただくことができた。しかし本稿にいかなる不備な点があっても、それは筆者の責任である。

最後に、しばしば書き改めた草稿をそのつど浄書タイプして下さったのは木稲宏実さんである。

これらの方がたに対し、ここで厚く感謝の意を表わしたい。

なお、本稿の内容の約 3 分の 1 ほどは、岩波書店、現代日本文化論(河合隼雄・責任編集)、第 7 巻、『体験としての異文化』(養老孟司・編集)、1997 年に、「情報化社会における日本語表記と入力」と題して書いたものを、こまかく書き改めたものから成っていることをお断りしておく。

参考文献

[1] 朝日新聞, 1996a, 国語—小学生ら対象に民間研究会調査, 11 月 10 日.

[2] 朝日新聞, 1996b, 米軍機爆弾投棄—外務省, 非公表を当初要請, 12 月 13 日夕刊.

[3] 朝日新聞, 1997, 日本政府, 英字紙だけに広告—マレー語紙読者が反発, 1 月 14 日夕刊.

[4] 阿辻哲次, 1996, 現代日本と漢字, 中京大学文学部国文学科学術講演, 中京大学名古屋キャンパス, 1996 年 6 月 22 日.

[5] 荒川文六, 1928, 電気工学の実験(ローマ字書き), 丸善.

[6] 安藤貞雄, 1986, 英語の論理・日本語の論理, 大修館書店.

[7] 碇義朗, 1977, 戦闘機飛燕, 広済堂.

[8] 岩村精一, 1955, 日本生産性の基盤, 河出書房.

[9] 岩村精一, 1956, 第二次産業革命と国字問題, 一橋書房.

[10] 岩村精一, 1966, 生産性とこれからの実務用字, 日本生産性本部.

[11] NHK, 1993,「NHK の日本語について」のおこたえ, 放送文化研究所, 用語研究班, 8 月 31 日, 6pp.

[12] NHK, 1994, エレクトロニクスが戦いを制す, ドキュメント太平洋戦争, 3, 角川書店.

[13] 生出寿(おいで・ひさし), 1988, 砲術艦長・黛治夫, 光人社. [光人社 NF 文庫 1996 年]

[14] 大坪一夫, 1992, 日本語教育でのコンピュータ利用の過去, 現在と未来, 日本語教育, No.78, pp.9-19.

[15] 奥宮正武, 1979, 日本海軍航空隊戦史—さらば海軍航空隊, 朝日ソノラマ.

[16] 小倉磐夫, 1997, 古老の思い込み(2), 朝日カメラ, 第 82 巻, 第 1 号, pp.140-141.

[17] おばた・ただお(小畑忠夫), 1992, NHK のにほんごについて—もっとまともなにほんごを—, 公開書簡, 8 月 3 日, 2pp.

[18] おばた・ただお, 1993a, NHK のにほんごについて—また—, 公開書簡, 7 月 31 日, 1p.

[19] おばた・ただお, 1993b, NHK のにほんごについて—またまた—, 公開書簡, 9 月 23 日, 2pp.

[20] 片桐明, 1988, 日本語プログラミング言語 Mind, 翔泳社. [1992 年の改訂版が, 現在は(株)リギーコーポレーションからマニュアルとして入手できる]

[21] 加藤常賢, 山田勝美, 1972, 字源辞典, 角川書店.

[22] 加藤ひろき, 1979, 日本語表記法の有利性, 言語, 第 8 巻, 第 1 号, pp.65-72.

[23] 岸田秀, 1997, 英会話という病気, 官僚病の起源, 新書館, pp.143-163.

[24] 橘田広国, 1992, 日本のローマ字運動, 日本ローマ字教育研究会.

[25] 教育総監部, 1943, 国文法教程, 林大(改編).

[26] 日下公人, 1993, 名誉ある孤立の研究, PHP 研究所.

[27] 倉石武四郎, 1963, 中国語辞典, 岩波書店.

[28] 倉石武四郎, 折敷瀬興・編, 1983, 日中辞典, 岩波書店.

[29] 国立国語研究所, 1996, 国際社会の日本語, 新プロ「日本語」国際シンポジウム予稿集, 東京, 9 月 21 日, 49pp. [6 論文, パネル討論内容は収録なし]

学術情報センター電子図書館サービス　　©学術情報センター

[30] 佐野洋, 1974, 推理小説とエンターティンメント, 小説推理, 第14巻, 第10号. [佐野洋, 推理日記—佐野洋ミステリー論集, 潮出版社, 1976年, pp.246-251に収録]

[31] 産経新聞社会部・編, 1992, 大学を問う—荒廃する現場からの報告—, 新潮社.

[32] 塩田紀和, 1980, 保科孝一の回想, 私信.

[33] JIS, 1990, 規格票の様式, JIS/Z/8301-1990, 日本規格協会.

[34] じっこく・おさむ(十国修), 1987, 桃青の詩的生涯—試論詩:芭蕉伝, 勁草社. [あとがきの追伸から]

[35] 情報処理学会, 1989, 日本文入力, 情報処理ハンドブック, オーム社, pp.1135-1145.

[36] ソーントン不破直子, 1997, 日本の「英語病」は杞憂, 朝日新聞・論壇, 1月16日.

[37] 田口泖三郎(りゅうざぶろう), 1936, Mozi wo kaku kimari, Rômazi Sekai, 第26巻, pp.4-6, (1月号).

[38] 田中克彦・他(計18名), 1997, 「国際化」とローマ字, 国文学·解釈と鑑賞, 第62巻, 第1号, 至文堂. [18論文の特集号]

[39] 田丸卓郎, 1930, ローマ字国字論(第3版), 岩波書店. [現在も在庫管理]

[40] 千早正隆, 1982, 日本海軍の戦略発想, プレジデント社.

[41] 津田幸男, 1996, 目にあまる日本の「英語病」, 朝日新聞・論壇, 12月27日.

[42] 豊田有恒, 1987, 崇峻天皇暗殺事件, 講談社.

[43] 成瀬正勝, 1996, 貧弱な国語の力, 東大は主張する, 六甲出版, pp.100-101. [東京大学新聞, 573号, 1964年4月22日掲載文を収録]

[44] 西原一幸・他, 1996, コンピュータ国文学の展望：情報交換と日本の漢字, 第1回シンポジウム・コンピュータ国文学・講演集, 国文学研究資料館, pp.93-115.

[45] 日経ビジネス, 1995, 問題は日本語だ：ネットワーク時代の競争力, 5月1・8日合併号, pp.20-30.

[46] 日経連広報部・編, 1993, ホワイトカラーの生産性向上事例集, 日本経営者団体広報部.

[47] 日商, 1994, 高度情報化社会に向けての基礎教育等の重要課題に関する提言, 日本商工会議所, 4月21日, 3pp.

[48] 日本化学会・編, 1966, 実験化学講座(続), 月報, No.8, 丸善, 5月.

[49] 日本能率協会・編, 1975, コストダウンのための改善ポイント, 日本能率協会.

[50] 橋本万太郎, 鈴木孝夫, 山田尚勇・編著, 1987, 漢字民族の決断：漢字の未来に向けて, 大修館書店.

[51] 波多野里望, 1996, 「国語科」教育の抜本的改革—次期学習要領改訂へ向けての緊急提言, 日本語技術教育学会・編, 言語技術教育, 第5巻, 明治図書, pp.108-109.

[52] 保科孝一, 1949, 国語問題五十年, 三養書房.

[53] 堀内庸村, 1959, カナ・ローマ字共通, やさしい分かち書き法, ダイヤモンド社.

[54] 松岡新児, 1992, ニュースよ日本語で語ってほしい, 兼六館出版.

[55] 黛治夫, 1972, 海軍砲戦史談, 原書房.

[56] 黛治夫, 1977, 艦砲射撃の歴史, 原書房.

[57] 村野辰雄, 1979, コンピュータとローマ字表記法の問題, Rômazi no Nippon, No.321, pp.1-2, 12月.

[58] 安島弥(ひさし), 1995, 日本語ローマ字化論への疑問, 朝日新聞, 論壇, 1月28日.

[59] 柳田国男, 1925, Simabito no kotoba, Rômazi Sekai, 第15巻, pp.2-5, (1月号).

[60] 柳田国男, 1936, Kokugo no seityô ni tuite, 1 & 2, Rômazi Sekai, 第26巻, pp.24-27, 135-137, (1月, 2月号).

[61] 矢野雄司, 1997, 私信, 3pp.

[62] 山田忠雄・主幹, 1981, 新明解国語辞典(第3版), 三省堂.

[63] 山田尚勇, 1984a, 情報化時代とことば, 月刊国語教育, 第3巻, 第10号, pp.46-57.

[64] 山田尚勇, 1984b, ことばと表記のあれこれ, 愛媛ローマ字国語, 第16号, pp.1-5. [カナノヒカリ, 第747号, pp.12-16(1984年10月)に転載]

[65] 山田尚勇, 1985, 専任タイピスト向きタイプ入力法の研究経過, コンピュータソフトウェア, 第2巻, pp.55-64.

[66] 山田尚勇, 1986a, 技術革新と日本語のゆくえ—ヒューマン・インタフェースの視点から—, 基調講演, 国際シンポジウム・漢字文化の歴史と将来, 東京, 5月19—20日, [東京大学理学部情報科学科テクニカル・レポート, No.86-10, 1986年9月, 114pp.]

[67] 山田尚勇, 1986b, コミュニケーションツールとしてのドキュメントの重要性—アメリカでの体

学術情報センター電子図書館サービス　　◎学術情報センター

験を通して，第一回ドキュメンテーション・エンジニアリング・セミナー，東京，8月20日．[東京大学理学部情報科学科テクニカル・レポート，No.86-12，1986年10月，45pp.]

[68] 山田尚勇，1987，文字体系と思考形態，日本語学，第6巻，第8号，pp.43-64.

[69] 山田尚勇，1989，学術情報サービスから見た専門用語，日本語学，第8巻，第4号，pp.18-26.

[70] 山田尚勇，1990，文字の型と読みの速さ，日本語学，第9巻，第11号，pp.4-29.

[71] 山田尚勇，1991，文字論の科学的検討，学術情報センター紀要，第4号，pp.261-318.

[72] 山田尚勇，1994a，多すぎる漢字・漢語にどう応えるか，学術情報センター紀要，第6号，pp.249-289.[増訂私家版あり]

[73] 山田尚勇，1994b，創造性の発露について，学術情報センター，vi+143pp.[論文集]

[74] 山田尚勇，1995，常用者のための日本文入力法の基礎的研究について，学術情報センター紀要，第7号，pp.73-112.

[75] 山田尚勇，1996，直接入力法の人間科学的考察，ワークショップ・日本文直接入力法，学術情報センター，1月31日，43pp.

[76] 吉田敦也，1996，塾年パソコン上達のツボ，講談社.

[77] 読み書き能力調査委員会，1951，日本人の読み書き能力，東京大学出版部.

[78] 渡辺慧(さとし)，1949，未来をゆびさすもの，中央公論社.

[79] AAS (Association for Asian Studies), 1995, The Ideographic Myth and Its Impact on Asian Studies, Sessions 97 and 117, organized by Mary Erbaugh and J.Marshall Unger , Washington , D.C., April 6-9. [これらの部会で発表された論文は編集を終え，現在出版が意図されている．なお，同様の情報は Hector Hammerly の "Synthesis"にも掲載されている由であるが，いまのところ筆者は入手できないでいる]

[80] Besner, D.; and M. Colthert, 1979, Ideographic and alphabetic processing in skilled reading of English, Neuropsychologia, Vol.17, pp.467-472.

[81] Caidin, Martin, 1966, The Rugged Rugged Warriors, E.P.Dutton and Co., Inc.

[82] DeFrancis, John, 1984, The Chinese Language: Fact and Fantasy, University of Hawaii Press.

[83] DeFrancis, John, 1989, Visible Speech—The Diverse Oneness of Writing Systems, University of Hawaii Press.

[84] Deutsch, Mitchell F., 1983, Doing Business with the Japanese, New American Library.

[85] Freud, Sigmund, 1904, The Psychopathology of Evryday Life, in German, [A.A.Brill (trans. and ed.), The Basic Writings of Sigmund Freud, The Modern Library, 1938, pp.33-178]

[86] Gottlieb, Nanette, 1996, Kanji Politics—Language policy and Japanese script, London, Kegan Paul.

[87] Logan, Robert. K., 1986, The Alphabet Effect: the impact of the phonetic alphabet on the development of western civilization, William Morrow and Co.

[88] Mair, V.; and Y. Liu (eds.), 1991, Characters and Computers, Amsterdam, IOS Press.

[89] Mair, Victor H., 1996, Modern Chinese writing, P.D.Daniels, and W.Bright (eds.), The World's Writing Systems, Oxford University Press, pp.200-208.

[90] Paivio, A., 1975, Perceptual comparisons through the mind's eye, Memory and Cognition, Vol.3, pp.635-647.

[91] Paradis, M.; H. Hagiwara; and N. Hildebrandt, 1985, Neurolinguistic Aspects of Japanese Writing System, Academic Press.

[92] Renfrew, Colin, 1989, The origins of Indo-European langages, Scientific American, Vol.261, No.4, pp.82-90.

[93] Tatuoka, H.; and H. Yamada, 1995, Toward the Japanese input through Roman orthography, Proceedings of the 1995 International Conference on Computer Processing of Oriental Languages, Honolulu, November 23-25, pp.107-114.

[94] Tatuoka, H.; and H. Yamada, 1996, On the theory and practice of the romanized writing system for Japanese, 学術情報センター紀要，第8号，pp.27-74 and errata.

[95] Tzeng, O.J.L.; and W.S.Wang, 1983, The first two R's, American Scientist, Vol.71, pp.238- 243.

[96] Unger, J. Marshall, 1987, The Fifth Generation Fallacy—Why Japan is betting its future on artificial intelligence, Oxford University Press. (奥村睦世・

学術情報センター電子図書館サービス
© 学術情報センター

訳，コンピュータ社会と漢字，サイマル出版会，1992年)

[97] Unger, J. Marshall, 1995, Private communication, at AAS, April 6-9.

[98] Unger, J. Marshall, 1996, Literacy and Script Reform in Occupation Japan—Reading between the lines, Oxford University Press.

[99] Yamada, Hisao, 1980, A historical study of typewriters and typing methods—from the position of planning Japanese parallels, Journal of Information Processing (情報処理学会欧文誌). Vol.2, pp.175-202.(小笹和彦・訳，タイプライタとその入力方法の歴史的考察—日本語タイプライタの開発動向への視点，bit, 1981年6～11月，共立出版)

[100] Yamada, Hisao, 1983, Certain problems associated with the design of input keyboards for Japanese writing, W. Cooper (ed.), Cognitive Aspects of Skilled Typing, Springer-Verlag, pp.305-407.

[101] Yamada, Hisao, 1991, An analysis of Japan's present industrial success, 学術情報センター紀要，第4号，pp.319-330. [山田 1994b に収録]

(筆者の書いたものの多くは、ご要望があれば、別刷りを差しあげられると思います。)

学術情報センター電子図書館サービス　　◎学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

CJK(中国語/日本語/韓国語)環境下での国際標準件名標目システムの適用

Adaptation of Universal Standardized Subject Heading System in the CJK (Chinese / Japanese / Korean) Environment

国立国会図書館　相原　信也
Nobuya AIHARA
National Diet Library
学術情報センター　内藤　衛亮
Eisuke NAITO
National Center for Science Information Systems

要旨

本稿は国際図書館連盟(IFLA)第52総会において、IFLA多文化図書館サービスに関する常任委員会が開催し、IFLA目録常任委員会；分類・索引常任委員会が後援したOne Day Workshopにおいて発表したものの日本語版である。日本語版として原典に注及び参考文献を付して拡大した。

中国語、韓国語、日本語は、異なるスクリプト、すなわち表記形とその読み・ローマ字による翻字の組からデータが成り立っている。本論文では、このような異なるスクリプトを扱わなければならない言語における国際標準件名標目の適用可能性を論じる。件名標目の国際的な標準化の前提となるいくつかの観点に即して、日本の件名標目表を紹介する。MARCと件名標目の本質的な違いは、前者が「入れもの」で、後者は内容であるというところにある。人的・技術的要素、言語、管理・運用の観点から、国際件名標目システムを実現するために解決しなければならない要件を論じる[1]。

ABSTRACT

The paper is a Japanese version of the presentation made at the One Day Workshop held on Thursday, August 29, 1996 in Beijing, China by the IFLA Standing Committee: Library Services to Multicultural Populations and sponsored by IFLA Standing Committees: Cataloguing; Classification / Indexing during the 52nd IFLA General Conference. This Japanese version is enhanced by adding notes and references.

The paper explores possibilities of adaptation of a Universal Standardized Subject Heading System to multi-script language environment such as Chinese, Korean and Japanese language, of which data string is a set of description, reading and Romanized character string in many cases. Subject heading systems in Japan are reviewed with some quantitative aspects to set a background to the universal standardization of subject heading system. Essential difference between MARC and subject heading data is that the former is the container and the latter is the content. Factors to be solved for achieving the Universal Subject Heading Systems are discussed in terms of human resource factors, technology factors, language factors, and administrative factors.

[キーワード]　多文化社会図書館サービス、国際標準件名標目システム、中国語/日本語/韓国語環境、国立国会図書館件名標目表、基本件名標目表、Japan MARC、NACSIS-CAT、TRC MARC、運営課題、国際協力

[Keywords]　Library Services to Multicultural Populations, Universal Standardized Subject Heading System, CJK Environment, NDLSH, National Diet Libraly List of Subject Headings, BSH, Basic Subject Headings, Japan MARC, NACSIS-CAT, TRC MARC, administration, international collaboration

学術情報センター電子図書館サービス　　©学術情報センター

はじめに—日本語版のための前説

本節では、本稿成立の特殊状況から、読者の理解のために本稿の成り立ちおよび概要を解説する。

本稿は、国際図書館連盟(IFLA)第52総会(IFLA北京大会)において、IFLA多文化図書館サービスに関する常任委員会が開催し、IFLA目録常任委員会および分類・索引常任委員会が後援したOne Day Workshop(1996年8月29日)における口頭発表に用いた原稿の日本語訳である。日本語版として原典に注及び参考文献を付して拡大したものである。

第一章では、主催者であるIFLA多文化図書館サービスに関する常任委員会による開催主旨、議題とされ提案されている国際標準件名標目システム(USSH)の目的、性格について要約し、それに対して、汎用システムであるべきUSSHに搭載される場合の漢字を使用するデータの特殊性や要件について、日本語の書誌データの特徴を例にあげて検討した。すなわち、記述文字列が多くの機能を果たす英語の書誌データとは対照的に、日本語の書誌データでは記述データ列、読み、そしてローマ字に変換した文字列の三つで一組のデータ要素を構成している点が最大の特色であり、この要件は多くのアジア諸言語に共通する。また、読みを国際的に利用可能とするためには、外国語の名称を自国語読みするという習慣、すなわち、中国人名を日本では日本の漢字の発音で読むし、香港では広東語の発音で、さらにローマ字化するなど、多様性がある。これらの解決すべき課題はかならずしも十分には把握されていない。

第二章では、日本の主要な件名標目表の存在と状況を述べた。国立国会件名標目表(NDLSH: National Diet Libraly List of Subject Headings)および基本件名標目表(BSH: Basic Subject Headings)の概要に加えて、機械可読形態での件名標目としてJapan MARC、NACSIS-CAT、BSHおよびTRC MARCなどのおおむね共通する仕様について要約し、現在進行中のBSHの改訂作業について紹介した。

第三章では件名標目の適用と管理・運用の問題について論じた。Japan MARCにおける件名標目やNACSIS-CATにおける件名標目の適用状況に関わる先行調査事例で指摘されている問題点を要約して、日本の国内的課題が少なからぬことを示して、アジア諸国のそれぞれの事情を暗示しようとした。

以上のような国内的な調整課題が多く存在するなかで、国際件名標目システムを実現するための要件について、第四章で論じた。将来的には、あるいは必要なものであろう国際件名標目システムを、直ちに共同構築を開始することは容易ではないことを指摘するために、問題点を列記した。

書誌データベースのいわば単独の構築から、マルチメディア・データベースの構成要素としての書誌データのデジタル化の趨勢と、書誌情報の国際的アクセスの実現という図書館界の使命からすれば、国際件名標目システムの実現可能性を国際的に検討するべきであるし、そのための国内的な実現可能性調査が望まれる。そのためには最大限の支持を与えられた国内組織、国際組織の設置が必要である。

ワークショップでは、米国(議会図書館)、米国大学図書館における多言語件名、カナダ、中国、韓国、タイ、マレーシアなどの件名標目の状況などが報告された。主題アクセスが今後の第一の課題であるという共通認識によって成立したワークショップと見ることができるが、各国とも国内調整課題の解決が当面の課題と思われた。さらに、発表者の中には、書誌ユーティリティや書誌データベース作成者(この場合は例えばNACSIS-CAT入力館も含む)が、個別の解決を図りつつあるという状況を踏まえ、かつ、情報通信基盤や文字コードの整備が不十分であることへの認識もあり、将来的に必要となる可能性を否定はしないが、ただちに共同で着手しようとは提案しにくいという立場であろうものもいた。本稿もその一つである。

1 序

1.1 国際標準件名標目表実現に向けて(主催者の開催主旨)

ワークショップの主宰者は、CJK(中国語/日本語/韓国語)環境下における国際標準件名標目表の適用可能性について発表するように我々に要請した。この要請は、MARCのコンセプトを世界的に広め、図書館や情報サービスにおけるコンピュータの適用の基礎を築いてきたMARCの発展過程を跡づけるということである。そのうち特に重要なのは、世界の多くの図書館がMARCシステム(特に、LCのシステムと英米目録規則)に従い記述の標準化を図ったように、件名標目についても、以下の現状に基づき、標準化を図る可能性はあるのかどうかという点である。

1) LCSHは最も知られており、それゆえに、欧米語圏でもっとも利用されている件名標目表である。
2) LCSHは連邦の資金的裏付けがあるので、なくなったり廃れたりすることはありそうもない。いずれにしても、

学術情報センター電子図書館サービス　　©学術情報センター

永続性がある。

3) LCSH はすでに完全にデジタル化されている[2]ので、他の国々は LCSH に該当する自分の国の言語の件名標目を加えさえすればよい。

新しい技術によって、有力な件名標目システム(例えば、LCSH)に合わせて 1 つのものとしてそれぞれの件名標目をデジタル化することもできよう。そうすることで、CJK ユーザは、統一件名標目からそれぞれの言語で書かれている資料にアクセスすることができる。このためは、発音を転記する必要はない。それぞれの件名標目の下に、それぞれの国ごとに、それぞれの言語の記入を編成してもよい。これを検討することは、言うのは簡単だが、実際に行うのは難しい。なぜなら、時間と努力、そしてなにより資金が必要だからである。今はまさに、その実現可能性を論議する場を設けるべき時である。すなわち、国際標準件名標目表がどのようなことに益するか。例えば、

a) 統一件名からさまざまな言語の資料にアクセスする
b) 一地域の、国内の、そして国際的な情報ネットワーク化を促進する
c) 経済的、文化的な効果がある

[以上が開催者が用意したワークショップ開催主旨および国際標準件名標目表の目的である。]

この論文では、できるかぎり与えられたテーマを扱うが、日本語とその文字に限定し、それに即して議論を進めていく。

1.2 日本語の書誌データの特徴

1.2.1 データ要素の組

表示したり印刷したりするための日本語データ(記述データ列)は、データを操作するための文字列、例えば、ソートのために表音文字で表記された読み(発音)や検索のために表音文字をローマ字化した文字列を伴っている。記述データ列、読み、およびローマ字は、データの組を形づくる。このことは、コンピュータ上の件名標目も例外ではない。記述データ列への自動読み付与や分かち書きなど、さまざまなデータ処理手法が開発されているが、その場合でも、結局はこれら 3 つの要素の組み合わせである。記述データ列とそれに対する読み(発音)が必要であるという点は、中国語、韓国語、日本語いずれにも共通であると思われる。

1.2.2 読み

よく知られているように、漢字は字形、字義、音(読みあるいは発音)などの属性を持っている。字形、字義、音は、1 つの言語あるいは方言のなかでもさまざまな組み合わせがある。その組合せは、国(言語)と方言(ローカルな言語)ごとに別々なので、[CJK として扱うとすれば]3 通り以上になる。したがって、一つの表記された(印刷された)日本語(の名称)には 1 つ以上の読み方がある。

多くの日本語システムでは、データベースの検索キーは読みであるが、漢字(記述)も検索キーとするシステムもある。異なる読みによる検索は、検索の際の利用者の判断に委ねられるか、検索システム側で検索を可能にする手段を用意する[3]か、いずれかで行われる。

日本では、中国語、韓国語の適切な知識がないと、中国、韓国の名称(人名や地名など)は、もともとの言語の読みとは異なる日本語の慣用に従って読まれる。同様に、中国人、韓国人は、日本の名称を自分らの流儀で読む。これは、外国の名称については母国語による正規の読みにはかかわりなく、異なる読みの検索キーを用意しなければならないことを意味している。

2 日本における件名標目の現状

2.1 日本の主要な件名標目表

日本で使用されている件名標目表の種類は多くない。それらは 2 つの主要な標目表である国立国会図書館件名標目表(NDLSH)[4]、基本件名標目表(BSH)[5]と、医学件名標目表[6]と学校図書館件名標目表[7][8]などである。その他のものはほとんど使用されていないと思われる。

2.1.1 NDLSH：国立国会件名標目表

国立国会図書館件名標目表(NDLSH)[9][10][11]は、国立国会図書館図書部図書整理課が編集している。初版は1964 年に刊行された。国立国会図書館(NDL、1946 年に創立)は、1949 年に和書の件名目録の編成を開始し、その後、いくつかの件名標目表を使用するなど曲折を経た後、1964 年に NDLSH を刊行した。NDLSH の改訂版は、1973 年(第 2 版)、1980 年(第 3 版)、1986 年(第 4 版)にそれぞれ刊行されており、現在の版である第 5 版は 1991 年にでている。

NDLSH は 2 巻からなり、参照を除く標目 17,133、参照 5,391、合わせて 22,524 の標目が収められている。いくつかの例外を除いて、ほとんどの参照は「を見よ」参照である。NDLSH は NDL の蔵書の件名目録作成のための件名標目表であり、それ以外での利用については考慮されていない。NDLSH 第 5 版は、1949 年から

学術情報センター電子図書館サービス　　◎学術情報センター

1991 年までの間に NDL が整理した和書に付与した件名標目と、1964 年から 1991 年までの間の洋書に付与した件名標目が収録されている。件名標目は主として納本された資料に基づいている。

NDLSH は冊子体を作成するために機械可読形態で維持されている。今のところ、機械可読版は館内で事務用として利用されているだけである。受け入れられた資料に付与された全ての件名標目は、件名典拠ファイルとしてカード形式で維持されている。これらの標目の中で、一般件名標目(主としてトピックに関する件名標目)のみが件名典拠データベースにも入力され、機械可読化されるが、人名、団体名はこのデータベースには入力されず、カード形式でのみ維持されている。

NDL の人名と団体名の件名標目は、著者標目とは異なる形をとっている。NDL の目録は著者、書名、件名が別々になっており、それぞれに関わる目録作業は別々の作業として行われている。記述目録作業と主題目録作業それぞれに関わる規則の改訂の時期が異なった。そのため、欧米の図書館では一般的である著者名、件名共通の名称典拠ファイルが NDL にはない。NDL では、NDLSH を和書だけでなく、洋書にも適用している。

2.1.2 BSH：基本件名標目表

基本件名標目表(BSH)は、日本図書館協会(JLA)件名標目委員会によって編集され、JLA から発行されている。初版は、1956 年に刊行された。改訂版は 1971 年と 1983 年に刊行され、後者が最新版である。標目 4,270 と参照 2,540 からなる。2 巻から構成される冊子体が入手可能である。第 4 版は現在準備中である(付録　BSH 機械可読フォーマットフィールド一覧(案、1995 年発表)を参照)。

BSH は、公共図書館、高等学校図書館、大学の教養課程の図書館を対象としている。そのため、特別な蔵書や大規模参考図書館で必要となる標目を収録していない。地名、人名、団体名の標目は、収録対象から除外されている。参照標目、特に「をも見よ」参照の充実が BSH の特徴として挙げられる。TRC(図書館流通センター)MARC が BSH を採用していることは、最も重要な特徴であるかもしれない。TRC は書籍商であり、図書が出版されてから間を置くことなく MARC を頒布している。TRC MARC で使用されている件名標目数は、1996 年 5 月現在で 47,000 に達している。

BSH の記入は、標目の読み、標目、分類記号、参照あり注記などから成る。BSH のデータ要素は、次のとおりである：

通常の標目：標目の読み(カナ表記)、標目、NDC 分類記号、[「を見よ」/「をも見よ」あり注記]

参照標目：参照標目の読み、参照標目、参照先(「を見よ」/「をも見よ」)

標目と標目の読みの組合せが不可欠な点が重要である。標目の読みは、a)ソートキー、また、b)標目検索のためにキーになる。

2.2 機械可読形態での件名標目

2.2.1 Japan MARC [12]

1980 年以後、Japan MARC が入手可能であり、現在広く利用されている。Japan MARC は、以下の 2 つの件名標目フィールドを持っている。

650 個人名件名

658 一般件名(上記以外の件名)

各フィールドはサブフィールド A、X、B から構成され、この一組のサブフィールドは繰り返し可能である。

サブフィールド識別子	サブフィールド名
A	カナ
X	ローマ字
B	漢字

これらすべてのデータの内容は、日本の国家標準規格が定める 2 バイトの文字コードで記録されている。このコードに加え、NDL が独自に拡張したコードも若干存在する。サブフィールド B は印刷と表示のためにあり、2 バイトの文字コードである。サブフィールド A と X は、ソートや検索といったフィールド操作のためにある。これもまた、2 バイトの文字コードである。

$Aオダ，ノブナガ(１５３４～１５８２)
$XOda，Nobunaga(１５３４～１５８２)
$B織田//信長(１５３４～１５８２)#

$Aセイショウナゴン
$XSeisyônagon
$B清少納言#

フィールド 650(個人名件名)の$A と$X ではコンマ、サブフィールド$B ではダブルスラッシュが姓と名を区切るために使われている。生没年も判明するならば記録されている。人名では読みは分かち書きされている。

ローマンアルファベットの外国人名標目には、$B はく、$A と$X のみ記録されている。この場合、$X の中のデータは、<<と>>に囲まれている。

$APlato.
$X<<Plato. >>#

学術情報センター電子図書館サービス　　◎学術情報センター

$AMill,John Stuart(1860～1873)
$X<<Mill, John Stuart(1860～1873)>>#

フィールド658の一般件名は、人名以外の全ての件名標目が含まれる。事項、地名、団体名、著作のタイトル、家名などがこのフィールドに入力されている。

$Aハンドウタイ
$XHandôtai
$B半導体#

$Aフカウラマチ(アオモリケン)
$XHukauramati(Aomoriken)
$B深浦町(青森県)#

$Aワセダダイガク
$XWasedadaigaku
$B早稲田大学#

$Aヘイケモノガタリ
$XHeikemonogatari
$B平家物語#

$Aヨコタ(ケ)
$XYokota(ke)
$B横田(家)#

複合語や句の読みであっても分かち書きされていない。$Bと$Aが同じものであるとき、$Bは省略される。地名を含む件名標目の場合、$Xでは原綴が使われる。

$Aウォーターゲートジケン(1972)
$X<<Watergate>>ziken(1972)
$Bウォーターゲート事件(1972)#

細目は、各サブフィールドに入力され、ダブルスラッシュによって区切られている。

$Aニホン//レキシ//メイジジダイ//ショモク
$XNihon//Rekisi//Meizizidai//Syomoku
$B日本//歴史//明治時代//書目#

2.2.2 NACSIS-CAT

NACSIS-CATは、1986年に運用を開始し、450以上の教育・研究機関が参加しているオンライン共同分担目録システムである。NACSIS-CATで目録情報源となっているMARCは、洋書の場合はUS MARCとUK MARC、和書の場合はJapan MARCとTRC MARCである。そのため、件名データはこれらのMARCから取り込まれる。和書の場合、Japan MARCからNDLSHが、同じようにTRC MARCからBSHが取り込まれる。NACSIS-CATの件名フィールドの構造は次のようになっている[13]。

SH：件名標目表の種類：件名標目--細目II件名標目の読み--細目の読み//標目のタイプ

NACSIS-CATは、このような方法で、参加図書館が使用しているさまざまな件名標目表に対応できる柔軟性を確保している。NACSIS-CATでは自らの総合目録データベースだけでなく、目録情報源である各種のMARCも同様に検索可能である。

さまざまな検索キーは、ソースMARCからもオリジナルで入力した目録レコードからも同様に生成される。検索キーが生成されるデータ項目は次のとおりである[14]。

レコードID、出版年1・2、出版国、タイトルの言語コード、テキストの言語コード、本文の言語コード、ISBN、全国書誌番号(NBN)、NDLカード番号(NDLCN)、タイトルと責任表示に関する事項の表記および読み、出版地、出版者、異なるタイトルの表記および読み、リンクされた書誌レコードの表記および読み、著者の表記および読み、統一タイトルの表記および読み、分類番号、件名標目など。

2.2.3 BSHの現状と今後

BSHは日本図書館協会が維持・管理しており、TRC MARCに適用されている。現在、第4版に向けた改訂作業が、JLA件名標目委員会によって進められている。改訂方針は、1995年に委員会により公にされている[15]。

1) MARC作成機関に役立つツールにすることを目標にする。
2) 細目形式を維持する。また、使用できる細目を増やす。
3) 地理区分を、どの標目でも適用できるようにする。
4) 文献的根拠があるならば、例示件名ではなく、表中に実際に標目を掲げる。
5) JLA選定図書およびTRC MARCの追加採用標目をもとに、新主題に対応するために標目の増強を図る。
6) BSHと異なる形をもつNDLSHの標目は、参照として採用する。
7) 参照にシソーラス構造を取り入れ、「をも見よ」参照を充実する。
8) 冊子体だけではなく、機械可読版も頒布する。

このうち重要な方針の変更は6)、7)、8)である。6)により、BSHとNDLSHの間をスイッチングする機能を実現できる[16]。いくつかの大規模公共図書館とNACSIS

学術情報センター電子図書館サービス　　◎学術情報センター

は、目録情報源となるMARCとして、Japan MARCとTRC MARCの両方を利用している。そのため、Japan MARCのNDLSHとTRC MARCのBSHの間のスイッチングを行うためのツールが求められている。ただし、BSHの制約のため、表現が類似しているもの、概念に違いをもつものなど、取り入れられなかったNDLSH標目も多い。BSHは、これらNDLSHの標目を参照として取り込んでいる。

BSHは参照が充実している。参照にシソーラス構造を取り入れることで、参照関係の整合性と一貫性を図ることができ、また同時に、欠落していた参照を補うこともできる[17]。BSH4版に向けた改訂作業では、コンピュータを使い、このシソーラス構造化の作業を行っている。

機械可読版の提供により、図書館のコンピュータ化に対応することができるだろう。機械可読版BSHはUNIMARC典拠フォーマットに準拠しており、オンライン閲覧目録(OPAC)での検索や目録作業において利用されることを想定している[18]。件名標目委員会は、UNIMARC典拠フォーマットに件名典拠として必要な修正を加えた。また、将来的にJapan MARC典拠フォーマットと統合されることを想定し、互換性を確保するための検討を行った。

2.2.4　日本語の件名標目の特徴

以上、3つの例、すなわちNDLSH、BSHとNACSIS-CATは、いずれもJapan MARCが採用した基本的な構成要素を持っている。件名標目あるいは細目は、以下の3つの要素($A、$X、$B)から構成される。

A　カナ形
X　ローマ字
B　漢字

3　件名標目の適用と管理・運用

日本において件名目録あるいは主題アクセスは、図書館の機械化活動がピークを迎えた1990年以前のカード目録の時代にはそれほど一般的ではなかった。大部分の図書館では、主題検索の手段は分類目録あるいは書架目録しかなかった。NDLと少数の先進的な図書館、例えば神戸市立中央図書館が全蔵書の網羅的な件名目録を編成していた。また、郷土資料などのような特別資料に限定して件名目録を維持していた図書館もあった。1990年以後、OPACの開発が日本でも進み、図書館の利用者にとって主題検索が身近なものとなり、それにともない主題検索への要求も大きくなってきている。

3.1　Japan MARCにおける件名標目

中井は、日本全国書誌1702号(1989年6月30日発行)の書誌データについて、件名の付与状況を分析している[19]。この分析では、Japan MARCにおける件名の部分的な様相しか分からないが、とりあえず、1書誌レコード当たりいくつの件名が付与されているかが概観できる(表1)。約53%の書誌レコード(1,255件中の664件)は1つの件名標目だけ、そして、わずか5%のレコードに3つ以上の件名標目が付与されていた。これらの数字から、主題目録作業にいかに多くの手間がかかるかを見て取ることもできよう。21.3%のレコード(262件)は件名なしである。これら件名がないレコードの半分は、作業基準に従って、もともと件名を付与することにはなっていない文学作品である(表1参照)。

表1　Japan MARCにおける1書誌レコード当たりの件名標目数

標目の数	タイトルの数	%
1	664	52.9 %
2	270	21.5 %
3	56	4.5 %
4	3	0.2 %
なし	262	21.3 %
合計	1,255	100 %

3.2　NACSIS-CATの件名標目データ

芳賀と松井は、1995年にNACSIS-CATデータベース(JBCATデータベース)のうち和図書の件名データの分析を行った[20]。1985年から1994年の間の出版年を持つ書誌データ410,516件について、件名と分類の付与状況を調査したものである。これから、前記の10年分のレコードでは、件名付与率は64～73%、分類付与率は81～88%であることがわかった。1994年と10年分全体についてのみ調査結果を示すと、以下のとおりである(表2参照)。

また、いずれの件名標目表を使用しているかは、NACSIS-CATデータベース(JBCATデータベース)の「件名標目表の種類」コードから分かる(表3参照)。

1985年から1994年までの10年間では、NACSIS-CAT中の和書の書誌レコードでの件名付与率が64～72%の間にあるとしている。さらに、この研究では、各分類表ごとの件名付与の状況や、件名の表記形とタイトル

学術情報センター電子図書館サービス　　© 学術情報センター

表2　NACSIS-CAT データベース(和図書)における件名と分類の付与状況

年	1994		10年分の合計	
件名あり	28,171	68.5 %	278,517	67.8 %
分類番号あり	34,046	82.7 %	342,742	83.4 %
両方なし	6,671	16.2 %	63,685	15.5 %
重複を除いた全レコード数	41,152	100.0 %	410,516	100.0 %

表3　NACSIS-CAT データベース(和図書)における件名標目表の種類

年	1994		10年分の合計	
NDLSH	12,193	43.3 %	194,813	69.9 %
BSH	19,316	68.6 %	116,590	41.9 %
MESH	1,199	4.3 %	6,812	2.4 %
LCSH	356	1.3 %	2,366	0.8 %
フリーワード(非統制語)	304	1.1 %	4,160	1.4 %
重複を除いた全レコード数	28,171	100.0 %	278,517	100.0 %

中の単語の比較などの分析を行っている。この研究に基づき、芳賀と松井は以下のような提言をしている。

1) 主題検索を可能にするために、件名付与を奨励するべきである。
2) 目録情報源としてMARCを有効利用するべきである。
3) 件名の一貫性はエンド・ユーザのために不可欠である。したがって、件名典拠コントロールを行うことが望ましい。
4) オンライン総合目録において、件名典拠コントロールを行うことは困難ではあるが、やるべきである。NDLSHとBSHの互換性の問題を解決すべきである。
5) NACSISが運用のための規則を定め、オンラインシステム上の件名典拠コントロールのためのツールを開発するとともに、目録担当者に件名付与を奨励するべきである。

4　国際件名標目システムを実現するための要件

MARCシステムが世界的に推進され、成功を収めた理由の1つは、MARCフォーマットが書誌データの「入れ物」の仕様であるからである。これに関しては、英米目録規則が書誌記述を標準化するために以前から寄与してきたところである。しかし、言語や文化的な慣習の相違を越えた国際的な記述の標準化を成し遂げるためには、数多くのことをやらねばならない。

MARCは標準の入れ物である。国際件名標目システムはデータ内容を階層的に相互に結びつけることを指向している。しかし、件名標目は、言語と文化、そしてデータ内容それ自体に依存している。実現できれば、世界中の件名をリンクすることによって、すばらしい結果がもたらされることはまちがいない。この目標を実現するために、以下の要因に十分な手当が必要である。

4.1　人的要素

国際件名標目システムの意義とそれがもたらすと考えられうる恩恵は、以下の観点から決まる。

- エンド・ユーザ
- 目録作成担当者
- MARC作成者／プロバイダー
- データベース作成者
- 国内標準の件名標目表と各機関の件名標目表との対応表を維持する者
- 国際的な件名標目表と国内の件名標目表との対応表を維持する者
- 国際的な調整者／調整機関

4.2　技術要素

1) NDLSHとBSHに階層構造を持たせる必要がある。
2) NDLSHとBSHを、オンライン、機械可読形式のいずれでも利用できるようにする必要がある。
3) NDLSHとBSHの間でのリンクづけが必要である。
4) 国際件名標目システムと日本の標準件名標目表とのリンクづけはできるであろうが、日本の標準件名標目表と個々の図書館・機関の件名標目とのリンクづけがさらに必要になる(これにより国際件名標目システムとのリンクづけができる)。これは、国内における件名標目の管理・運営上および財政上の問題を引き起こす。

4.3　言語上の要素

1) 一つの漢字は、1言語内だけでなく方言においても、字形、読み、意味に多くのバリエーションを持つ。エンド・ユーザは自分の母国語に頼るので、件名はこれらのさまざまな組み合わせのうちの1つからアクセスできなければならない。
2) 外国人名の読みはそれぞれの国で異なっていることを考慮し、各国ごとに調整しなければならない。そうしないと、エンド・ユーザは近隣の国の人名にアクセスすることができない。
3) たとえ2つの件名標目表をリンクづけすることが可能

学術情報センター電子図書館サービス　　© 学術情報センター

であっても、同じ文字を使っている言語間で語彙の意味分析が必要になるだろう。

4.4 管理・運営上の要因

1) LCSH の専門家は、日本にはあまりいない。
2) 件名標目表の専門家が必要である。
3) 英語の知識を持った件名標目表の専門家が必要である。
4) NDLSH(Japan MARC)と BSH(TRC MARC)あるいは他の件名標目表との間で組織をまたがる調整が必要である。
5) 出張旅費は、日本の図書館コミュニティの中で以前から問題になっている。
6) 日本の標準件名標目表と他の件名標目表と同様、国際的な件名標目表と日本の標準件名標目表とをリンクづけする仕事のため財源を必要とする。

謝辞

BSH4 版の改訂に関する助言をいただいた JLA 件名標目委員会委員・北克一氏(大阪市立大学助教授)、NACSIS-CAT の設計者でありその件名データについて情報をいただいた宮沢彰氏(学術情報センター教授)、このテーマに関係したデータを提供いただいた松井幸子氏(図書館情報大学教授)、芳賀奈央子氏(当時、図書館情報大学)の名を記して感謝する。

付録 BSH 機械可読ファイルフォーマット (案、1995 年発表)のフィールド

出典：日本図書館協会件名標目委員会,「基本件名標目表第 4 版　機械可読典拠フォーマット(案)」, 現代の図書館, Vol.33, No.1, pp.55-67, 1995

0-- 識別ブロック
001 レコード識別子
005 バージョン識別子
015 国際標準典拠データ番号

1-- コード化情報ブロック
100 一般的処理データ
150 名前に関するコード化データフィールド
152 目録規則
154 統一タイトルに関するコード化データフィールド
160 地域コード

2-- 標目ブロック
200 個人名
210 団体名
215 地名
220 家族名
230 統一タイトル
235 集合的統一タイトル
240 名前＋タイトル
245 名前＋集合的統一タイトル
250 普通件名

3-- 注記ブロック
300 注記
301 日付に関する注記
305 文章による「をも見よ」参照注記
310 文章による「を見よ」参照注記
320 一般的説明参照注記
330 一般的スコープノート

4-- 「を見よ」参照トレーシングブロック
400 個人名
410 団体名
415 地名
420 家族名
430 統一タイトル
440 名前＋タイトル
445 名前＋集合的統一タイトル
450 普通件名

5-- 「をも見よ」参照トレーシングブロック
500 個人名
510 団体名
515 地名
520 家族名
530 統一タイトル
540 名前＋タイトル
545 名前＋集合的統一タイトル
550 普通件名

6-- 分類番号ブロック
686 日本十進分類法

7-- リンク標目ブロック
700 個人名
710 団体名
715 地名
720 家族名
730 統一タイトル
740 名前＋タイトル
745 名前＋集合的統一タイトル
750 普通件名

8-- 出典情報ブロック
801 レコード作成機関
810 出典データあり
815 出典データなし
820 用法または範囲に関する注記
825 例示注記
830 目録作成者一般注記
831 文字に関する注記
835 標目修正に関する注記
890 レコード維持・管理アドレス

9-- 標準手続きブロック
990 標準手続き

学術情報センター電子図書館サービス　　©学術情報センター

注・参考文献

[1] この論文に述べられていることは個人の見解であり、所属する機関および言及した機関・組織の見解を述べたものではない。

[2] US MARC Format for Authority Data, Washington, Cataloging Distribution Service, Library of Congress, 1987-.

[3] 書誌レコードや典拠レコードに異なる読みを収めるフィールドを設け、それからも検索できるようにする等が考えられる。

[4] 「国立国会図書館件名標目表」, 第5版, 国立国会図書館, 1991.

[5] 日本図書館協会件名標目委員会, 「基本件名標目表」, 第3版, 日本図書館協会, 1983.

[6] 「医学件名標目表」, 日本医学図書館協会, 1995.

[7] 「小学校件名標目表」, 全国学校図書館協議会, 1985.

[8] 「中学・高校件名標目表」, 第2版, 全国学校図書館協議会, 1985.

[9] 横山幸雄, 「A Memory of NDLSH」, 現代の図書館, Vol.33, No.1, pp.15-22, 1995.

[10] 中井万知子, 「国立国会図書館件名標目表」, 情報の科学と技術, Vol.39, No.12, pp.575-581, 1989.

[11] 千賀正之他, 「件名作業の現状と今後」, 図書館研究シリーズ,No.27, pp.75-196.

[12] 「JAPAN MARC マニュアル 図書編」, 国立国会図書館, 1992.

[13] 「目録システム利用マニュアル データベース編 目録情報の基準」, 改訂版, 学術情報センター, 1990.

[14] 「目録システム利用マニュアル 検索編」, 改訂版, 学術情報センター, 1989.

[15] 日本図書館協会件名標目委員会, 「件名標目表改訂の基本方針」, 図書館雑誌, Vol.89, No.3, pp.186-189, 1995.

[16] 北克一, 芝勝徳, 「我が国における主題索引ツール統合の試み」, TP&D フォーラムシリーズ, 3, pp.6-31, 1994.

[17] 北克一, 芝勝徳, 「基本件名標目表第3版(BSH3)のシソーラス構造化の試み 外形式を中心として」, 整理技術研究, 31, pp.53-75, 1992.

[18] 日本図書館協会件名標目委員会, 「基本件名標目表第4版 機械可読典拠フォーマット(案)」, 現代の図書館, Vol.33, No.1, pp.55-67, 1995.

[19] 中井万知子, 「国立国会図書館件名標目表」, 情報の科学と技術, Vol.39, No.12, pp.575-581, 1989.

[20] 芳賀奈央子, 松井幸子, 「NACSIS-CAT の和図書ファイルにおける主題情報の現状調査」, 越塚美加編：「研究成果流通システムの研究開発 平成7年度報告」, (科学研究費総合研究(A)課題番号 06302076 研究成果報告書), pp.101-115, 1996年6月.

学術情報センター電子図書館サービス　　©学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

総合目録オンラインDBと情報検索システムの連携方式

Configuration of an IR System linked to a Union Catalog On-line Database

学術情報センター　大山　敬三
Keizo OYAMA
National Center for Science Information Systems
学術情報センター　髙須　淳宏
Atsuhiro TAKASU
National Center for Science Information Systems
学術情報センター　鶴岡　弘
Hiroshi TSURUOKA
National Center for Science Information Systems

要旨

本論文では、学術情報センターの目録所在情報システム(NACSIS-CAT)と情報検索システム(NACSIS-IR)の連携方式について述べている。オンラインで更新される書誌・所蔵データベース中のレコードに対して、自動的に情報検索用のインデクスを付与してエンドユーザへ検索サービスを提供するシステムの構成の検討を行っている。データベースサーバへの負荷、検索応答性能、ファイル容量などの面から、いくつかの方式を評価して実際のシステムの基本設計を行った結果として、データベースの構造と性質に応じた最適な構成が示されている。

ABSTRACT

This paper describes a method to link NACSIS-IR, an information retrieval system, with NACSIS-CAT, a catalog and location information system. The authors studied the configuration method of a system which automatically builds index of bibliographic and holding records stored in databases updated on-line, and which provides information retrieval services to the end users. Several methods were evaluated from viewpoints of load to the database server, response time of retrieval, file capacity and so on in order of the basic design of an actual system. As the result, the optimal configurations corresponding to structure and characteristics of databases are obtained.

[キーワード]　情報検索システム、オンラインデータベース、データベースサーバ、インデクス、オンライン更新、総合目録、性能評価

[Keywords]　information retrieval system, on-line database, database server, index, on-line update, union catalog, performance evaluation

1　はじめに

学術情報センターでは、図書館向けのサービスとしてオンライン目録所在情報サービス (NACSIS-CAT、以下 CAT) を提供する一方、大学研究者向けのサービスとして情報検索サービス (NACSIS-IR、以下 IR) を提供している。CATは大学等の図書館が所蔵する図書・雑誌の書誌と所蔵情報をオンラインで検索・更新するための総合目録データベース (DB) システムである。IRは学術情報センターが構築したり外部から導入した学術情報 DB のオンライン情報検索システムである。CATは図書館向けのサービスであるが、ここで構築された DB は研究者に取っても有用なものであり、従来からIRでも検索サービスを提供してきた。

センターでは1996年1月にこれらサービス用のメインシステムの更新を行ったが、このシステムでは従来の互換性と将来のサービスの展開を考えて、いわゆる

学術情報センター電子図書館サービス　　© 学術情報センター

メインフレーム(互換系)と UNIX システム(オープン系)の複合構成をとっている。CAT ではデータベース機能の向上とインターネットへの柔軟な対応を可能とするため、データベース機能をアプリケーションから切り離し、1997 年 1 月からオープン系システム上での運用を開始した。一方、IR でもより高度な検索機能や、インターネットを通じたユーザフレンドリな検索インタフェースの提供、あるいは全文 DB への対応などを可能とするため、オープン系システム上に新しい情報検索システム(新 IR システム)の開発を進めている[1]。

本稿では、これらの開発の中で実現しようとしている、オンライン総合目録 DB システムと情報検索システムの連携方式について述べる。本方式は目録所在情報に限らず、データベースに蓄積されてゆく情報を自動的かつ実時間に近い早さで検索可能とするシステムの構成方式としての汎用性を備えている。インターネットをはじめとするデータ通信網の整備により、作成者と提供者の連携が緊密化し、オンライン出版や電子図書館によるネットワークを通じた情報流通が主流となる環境において、情報提供の円滑化をはかる上で有効である。

2　NACSIS-CAT システムの構成

2.1　データベース

CAT の DB は目録所在情報 DB(総合目録 DB)と参照目録 DB から構成されている。これらの DB の種類と規模を表 1 に示す。

総合目録 DB は和文・欧文、図書・雑誌の書誌・所蔵情報と著者名などの典拠情報を集積したものであり、参加図書館が日常の整理業務としてオンラインで検索・更新を行っている。参照目録 DB は書誌データの作成の効率化を図る目的で、外部の目録作成機関から提供を受けた MARC (Machine Readable Catalog) をフォーマット変換して提供しており、定期的にバッチ更新を行っている。

総合目録 DB のレコード構成を図 1 に示す。書誌、所蔵、典拠はそれぞれ別テーブルに格納され、書誌レコード中には典拠レコード ID、所蔵レコード中には書誌レコード ID を持つことによりリンクを形成している。書誌ごとに対応して所蔵が作成されるので、レコード間の対応は 1:*n* になる。一方、典拠は書誌とは独立なのでレコード間の対応は *m*:*n* になる。

表 1　NACSIS-CAT のデータベースの種類と規模

テーブル名称			収納件数 (1996.12.20.)	週間更新 (1996.12.16～)
総合目録 DB	書誌	和図書	1,275,082	9,261
		洋図書	2,523,331	12,288
		和雑誌	84,120	535
		洋雑誌	122,617	376
		合計	4,005,150	22,460
	所蔵	和図書	17,136,178	119,942
		洋図書	8,173,756	46,258
		和雑誌	1,872,827	7,093
		洋雑誌	1,164,812	3,808
		合計	28,347,573	177,101
	典拠	著者名	907,314	2,584
		統一書名	11,531	58
		変遷マップ	23,267	115
		合計	942,112	2,757
参照目録 DB	書誌	和文書(JP)	1,847,816	–
		和図書(misc)	523,406	–
		和雑誌(JP)	100,582	–
		洋図書(LC)	4,992,674	–
		洋図書(misc)	1,993,347	–
		洋雑誌(LC)	727,381	–
		非文字(LC)	268,147	–
		合計	10,453,353	–
	典拠	著者名(LC)	2,947,639	–
		著者名(JP)	327,561	–
		統一書名(LC)	175,450	–
		合計	3,450,650	–

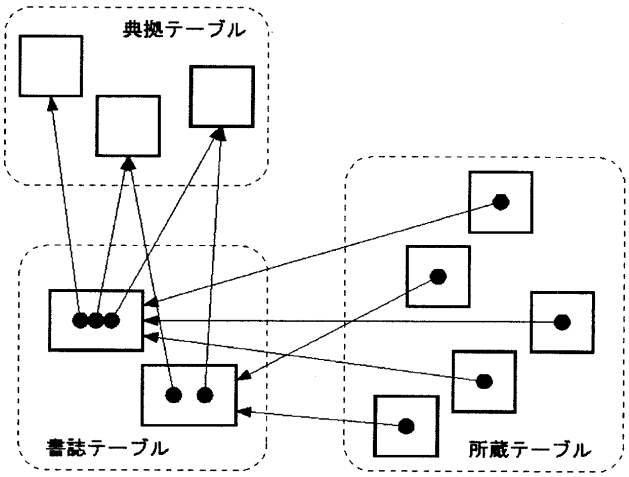


図 1　総合目録 DB のレコード構成

2.2 システム

CATシステムの構成を図2に示す。システムは全体として、DBサーバ、アプリケーションサーバ、クライアントの3つから構成されている。DBサーバはオープン系上に実現され、DBACと呼ぶアクセスインタフェースを通じてDBMS機能を提供する。アプリケーションは現在は互換系上の画面型 (SOA)のみであり、図書館の目録および文献複写業務に対応した機能を提供している。UIP はさまざまなプラットフォーム上に実装されており、参加図書館のローカルシステム上で稼働し、ローカルなDBとの連携を行っているものも多い。SOAと UIP の間は独自の仮想画面転送プロトコル[2][3]により通信を行っている。

現在、オープン系の上に新しい目録所在情報アプリケーションサーバ(新CATシステム)を開発中である。レコードを直接的に操作するためのプロトコルを有し、クライアントを介して図書館のローカル DB との柔軟な連携を可能とするものである。また、WWW を通じてOPACサービスを行うためのシステムも準備中である。これらのアプリケーションはすべて上記のDBサーバを共用し、DBACを APIとして用いる。

3 目録系IRシステムの構成

3.1 背景

NACSIS-IR では、学術文献を中心とする各種の索引や抄録などの二次情報や、学術論文などの全文の一次情報などの検索サービスを行っている。これらの DB の多くは外部から導入したりセンターが一括で作成したりしており、更新の頻度は1週間に1度程度から数年に1度程度までさまざまである。これらはいずれもIR専用のDBを構築して運用している。

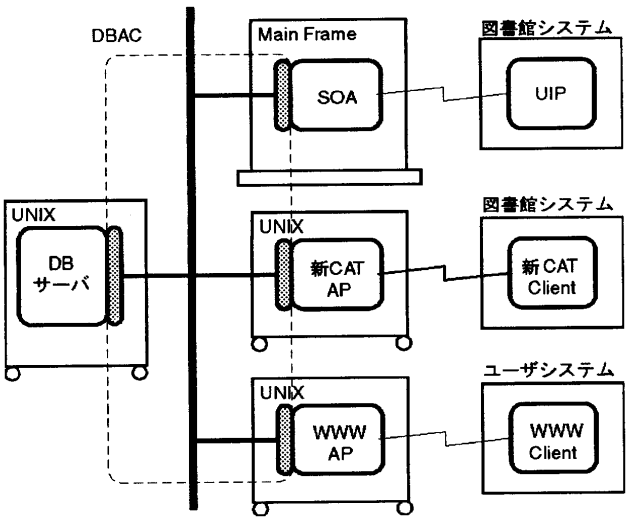


図2　NACSIS-CATのシステム構成

目録系DBについても、現在のIRシステムでは、CATのデータを抽出して専用のDBを作って運用している。更新も、CATの更新データをためておいて定期的にバッチ処理するという形態で運用している。しかし、総合目録 DB は参加図書館からのオンライン入力により構築されるため、IRとしてもできるだけ最新の情報を提供することが望ましい。また、参照目録 DB についても、CAT用とIR用にDBを重複して保守するのは容量的にも時間的にも無駄が多い。更新頻度を多くすると、更新処理中のサービス停止も問題となる。

目録系DB以外でも、学術雑誌目次速報DBは協力図書館から電子メールなどでデータを随時送ってもらい、情報をDBMSで管理している。IRサービス用には全レコードをダンプして専用の DB を作成しているが、これも管理用のDBをそのまま使えることが望ましい。

そこで、新IRシステムでは、これらのDBをできるだけ統一的な方式で、しかもデータの重複を減らして自動的な運用ができるようにすることを目標にしてシステムの設計を行った。

3.2 設計上の要求事項

CATシステムは図書館の目録業務や文献複写業務に利用することを前提に設計されており、検索もコード類や書名、著者名、出版社名などを主な対象としている。このため、書名中や注記中の自由語検索などへの対応は不十分であり、また、所蔵図書館名の部分検索などもできない。そこで、エンドユーザの利用性を高めるために、インデクスはCATとは別にIR専用のものを構築することを前提とした。従って、目録系IRシステムの設計に当たっては、主にデータ本体(テキスト部分)の連携について検討を行った。検討に当たっては以下のような事項を考慮した。

(1) エンドユーザ用のデータビューの提供

総合目録DBは書誌、所蔵などのレコードが独立した構造となっているが、IRでは書誌・所蔵を統合して検索できるようにする必要があり、IR用のデータビューを提供する必要がある。IR用のデータビューの概念を図3に示す。所蔵は書誌に包含する形で統合され、任意の組合せで検索可能である。典拠は書誌とは独立しており、検索も独立に行われる。書誌と典拠を組み合わせる場合は、まず典拠を検索して ID を取得し、これと書誌情報を組み合わせて検索を行うことになる。

学術情報センター電子図書館サービス　　©学術情報センター

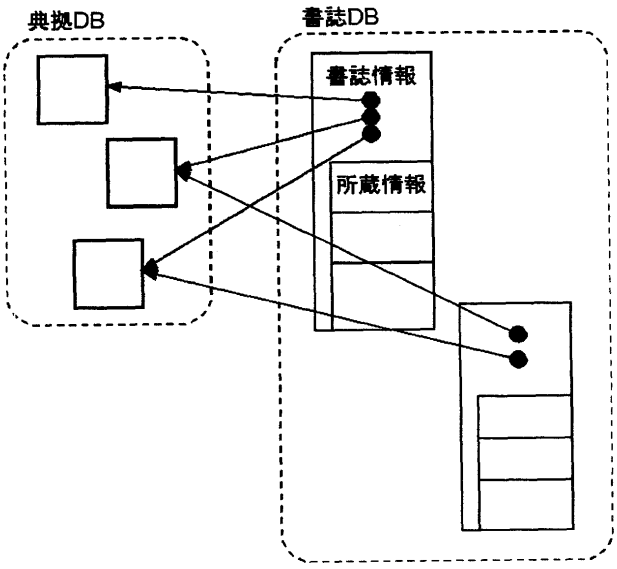


図 3　総合目録 DB の IR 用データビュー

(2)　適正な応答時間の確保
総合目録 DB、参照目録 DB ともかなりの大きさの DB であり、ユーザが満足できる応答時間を確保できる構成にする必要がある。平均的な負荷での通常の検索応答時間が 2 秒以内となることを目標としている。

(3)　DB サーバへの負荷抑制
CAT への性能上の影響を避けるため、DB サーバへの負荷が過大にならないよう、特に数秒以上にわたってバースト的な負荷が発生しないようにする。

(4)　更新処理の効率化
オンラインでの更新をできるだけ速やかに IR に反映してユーザに最新の情報を提供する。また、データベースの更新時間を短縮することによりシステム運用の負担を減らす。

(5)　データベースの重複の抑制
DB を重複して持つことにより発生する記憶容量の増加と保守のコストを抑制する。

特に、(1)と(2)はユーザの使い勝手に直接関係する部分であるので必須の条件である。また、(3)は図書館業務の効率に影響するため重要な条件である。(4)はユーザなどへの直接的な影響はあまりないので(1)、(2)、(3)ほどでないが、運用面から見るとやはり重要な条件である。(5)はディスク資源に関する限りは、他の条件と比べると追加投資で解決可能であるので、あまり強い条件ではない。

3.3　性能と容量の検討

3.3.1　検討対象システム構成

新 IR システムでは、自由語検索や階層を持つデータの検索への対応や、全文 DB 検索システムの実現のために、カナダ OpenText 社製全文検索システムである OpenText6 をエンジンとして使用する。そこで、ここでは OpenText6 のメタフィルター機能とオンライン更新機能を用いてシステムを構築することを前提とする。

メタフィルター機能では、ファイルやリレーショナル DB に格納されたテキストデータに対する検索エンジンからの仮想的なアクセスを提供するため、物理的なデータの記憶システムに依存することなく全文のインデクスを構築し検索を行うことができる。この中で、データ形式の変換も可能であるため、CAT のレコード形式を IR に適した形式(新 IR では SGML 風のタグ形式)に動的に変換したり、書誌と所蔵のレコードを統合してユーザに適したデータビューを作ったりすることができる。

また、オンライン更新機能では、検索エンジンを停止することなく、ファイルや DB のレコードなど記憶システムにおける論理的な格納単位(オブジェクト)ごとに、インデクスを更新することが可能である。更新トリガを与えるとそのオブジェクト中のテキストに対するインデクスを作成し、検索と並行して動的にインデクスファイルの更新を行う。

これらの機能を用いることにより、DB の共用を図りつつ、IR を通じて実時間に近い情報の提供を実現するためのいくつかのシステム構成が可能となる。

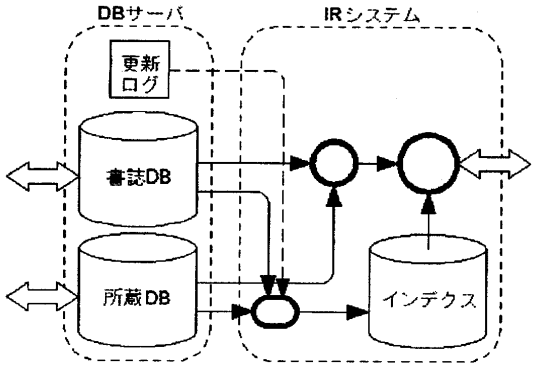
今回検討を行ったデータベース連携方式は以下の 3 種類である。これらの概念を図 4 に示す。

(1)　完全共有型
IR のサービスはすべてオンラインの DB を直接使って行う。必要なデータビューの変換はアクセス時に動的に行う。

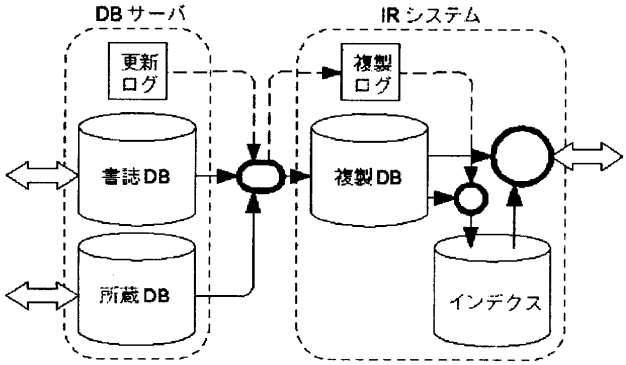
(2)　複製レコード型
オンライン DB からレコード単位の複製を作り、IR サービスはこの複製を使って行う。必要なデータビューの変換は複製時に行う。つまり、各書誌とそれに付随する所蔵は統合して 1 レコードとして格納する。

(3)　複製ファイル型
オンライン DB のダンプファイルを作成して IR 専用のテキストファイルを作成する。必要なデータビューの変換は複製時に行う。すべての書誌と

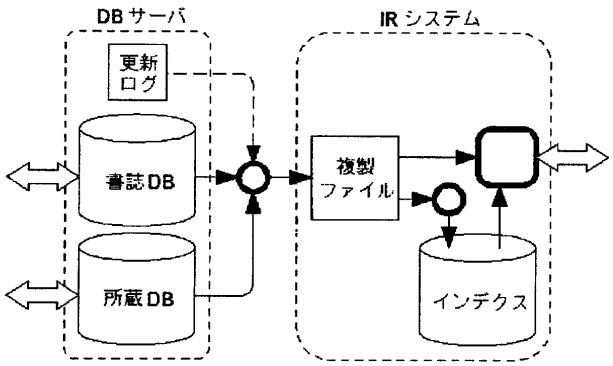
学術情報センター電子図書館サービス　　© 学術情報センター



(1) 完全共用型



(2) 複製レコード型



(3) 複製ファイル型

図4　DB連携方式の種類

所蔵レコードを統合して1ファイルとする。

結論としては、総合目録系DBでは(2)を、参照目録系DBでは(1)を採用することとした。なお、(2)では複製を格納する記憶システムとしてDBMSとファイルシステムがありうる。当面は実現の容易さを考慮してファイルシステムを使うことにしているが、将来は運用を容易にするために変更する可能性もある。

以下ではこの結論にいたる議論として、これらの各方式について、DBサーバへのアクセス回数、更新処理時間、応答時間、ディスク容量に関して行った検討について述べる。

3.3.2 DBサーバアクセス回数

DBサーバへアクセスする局面としては、DB初期作成、インデクス再作成、レコード更新、レコード検索がある。

(1) DB初期作成

書誌、所蔵、典拠の全レコード数をそれぞれ N_B、N_H、N_A とすると、完全共有型、複製レコード型、複製ファイル型のいずれの構成でも全レコードに1回ずつ、すなわち、$(N_B+N_H+N_A)$回のアクセスが発生する。現在の総合目録DBでは N_B=4.01×10^6、N_H=28.35×10^6、N_A=0.94×10^6 であるので、これは3,330万回になる。参照目録DBでは N_B=10.45×10^6、N_H=0、N_A=3.45×10^6 であるので1,390万回になる。

(2) インデクス再作成

完全共有型は初期作成と同じで(NB+NH+NA)回、複製レコード型、複製ファイル型は複製があるので0回のアクセスとなる。総合目録DBでは完全共有型は3,330万回になり、現実的ではない。参照目録DBでは完全共有型は1,390万回になり、日常的に行うのは不可能であるが、事前に十分計画しておけば不可能な規模ではない。現実的には、DBの定義やインデクスの正規化ルールの変更などを行うときに必要となる程度で、頻繁に行うことではない。

(3) レコード更新

総合目録DBの1日あたりの書誌、所蔵、典拠の更新レコード数の最大値をそれぞれ U_B、U_H、U_A とすると、完全共有型では書誌、所蔵では1レコードの更新につき、関連する書誌と所蔵のレコードにアクセスし、典拠では更新のあったレコードだけにアクセスする。1書誌に関連する所蔵数の平均は書誌と所蔵の全レコード数の比に等しいので N_H/N_B である。従って、この場合は $((1+N_H/N_B)(U_B+U_H)+U_A)$回のアクセスが発生し、過去1年間の統計では U_B=6,000、U_H=56,000、U_A=1,000程度であるので、最大50万回程度になる。複製レコード型、複製ファイル型では複製のうち、更新のあったレコードの部分のみを入れ換えればよいので$(U_B+U_H+U_A)$回になり、現状では最大63,000回程度になる。参照目録DBでは定期的にバッチ更新を行っており、最も更新レコード数の多いLC-MARCで1回あたり最大で31,000レコード程度である。所蔵データはないので、完全共有型、複製レコード型、複製ファイル型いずれでも

学術情報センター電子図書館サービス　　©学術情報センター

更新には最大で31,000回程度のアクセスが必要になる。

(4) レコード検索

通常の検索対象となるIRのレコードには書誌と所蔵が含まれているため、1レコード表示あたり完全共有型では$(1+N_H/N_B)$回のアクセスが発生する。複製レコード型、複製ファイル型では複製があるのでDBサーバへのアクセスは0回である。ピーク時のレコード表示頻度を10レコード/sと仮定すると、総合目録DBでは完全共有型は77アクセス/sとなり、CATサービスでのピーク時のアクセス頻度が200～225回/s程度であるのに比べて、無視できない大きさとなる。参照目録DBでは完全共有型は10アクセス/sとなり、ほぼ無視できる。

3.3.3 更新処理遅延時間

オンラインDBにレコード更新があったとき、完全共有型、複製レコード型ではOpenText6のオンライン更新機能を使って逐次更新が可能である。更新時間は主にデータアクセス時間とインデクス作成時間に分けられる。典拠レコードの更新頻度は小さいので無視すると、1更新レコードあたりのデータアクセス時間は、レコードアクセス時間をTとして、完全共有型では$(1+N_H/N_B)T$、複製レコード型ではTである。インデクス作成時間は更新レコード数に依存する部分と固定部分があるが、総合目録DBでは更新間隔が比較的小さいので更新レコード数も少なく、固定部分の方が大きな割合を占める。原稿執筆時点のOpenText6のオンライン更新性能の評価からは、高々10分以内には収まると見積もられる。更新間隔を10分と仮定すると、更新レコード数はピークで2,000個程度となり、T=100ms程度であることから、データアクセス時間は完全共有型では1,500秒程度、複製レコード型では200秒程度となり、完全共有型では多重処理をしない限り処理不可能であるが複製レコード型では逐次処理で十分処理可能である。

一方、複製ファイル型ではオンライン更新機能は利用できず、インデクスの再作成となり、最大で20時間程度かかると見込まれる。

3.3.4 検索応答時間

IRでのサーチ自体はインデクス上で完結するため通常のDBと同等であり、問題とはならない。データの表示のときに完全共有型、複製レコード型ではレコードへのアクセスが発生するため応答時間が問題となるが、複製ファイル型は通常のDBと同等であり十分高速である。完全共有型では1件あたりの表示に$(1+N_H/N_B)$回のDBサーバへのアクセスが発生し、総合目録DBでは800ms程度かかる。通常は10件程度をまとめて表示させるので8秒程度かかることになり、許容できない。DBACを改修して複数の所蔵レコードをまとめてフェッチできるようにすれば、1件あたりの時間を300ms程度まで短縮することは可能であると考えられるが、それでも3秒程度かかるので十分とは言いがたい。参照目録DBでは1件あたり100ms程度であり、10件まとめて表示しても1秒であり許容範囲である。複製レコード型では1件あたり1回のアクセスであるので100ms程度となり、10件まとめて表示しても1秒であり許容範囲である。

3.3.5 ディスク容量

インデクスは完全共有型、複製レコード型、複製ファイル型いずれでも同じであるので、本体データの複製に関する容量増加が検討の対象となる。完全共有型は複製を作らないので容量増加はない。

複製レコード型はレコード単位で複製を作成するため、DBMSやファイルシステムのブロックサイズに満たない部分が無効となる。複製レコード型で無効となる容量は、UNIXファイルシステムの論理的なフラグメンテーションのサイズを1KBとすると1レコードあたり平均で0.5KB程度となる。実際にサンプルデータで測定してもほぼ同様の結果となっている。レコードは書誌と典拠単位に作成するため、合計で総合目録DBでは2.2GB程度、参照目録DBでは6.5GB程度の無効容量が生ずる。これに本体のデータ容量を加えると、総合目録DBでは9.0GB程度、参照目録DBでは14.5GB程度の容量増加となる。

複製ファイル型ではブロック化による無効容量は発生しないので、レコードの正味の容量となり、総合目録DBでは6.8GB程度、参照目録DBでは8GB程度となる。しかし、サービスと並行して更新を行うためには新旧のデータとインデクスを二重に持つ必要があり、実際にはこの2倍程度の容量がさらに必要になるため、合計では総合目録DBでは20GB程度、参照目録DBでは24GB程度となる。

複製レコード型、複製ファイル型のいずれにしても、容量の増加自体はハードウェア資源の追加で対処でき

学術情報センター電子図書館サービス　　© 学術情報センター

るのであまり重大な問題ではない。むしろ、バックアップなどの保守のコストの方が運用上、大きな問題になりそうである。

3.4 その他の検討

前節ではおもに性能や記憶容量に関する検討を行ったが、この他にもいくつか検討を要する点がある。

第一に、CAT からのオンラインでのデータ更新と IR のインデクス更新の時間差による不整合の問題がある。何らかの理由でインデクス更新が中断すると、完全共有型ではその間の更新データとインデクスの間に大きな不整合が発生する。複製レコード型ではインデクス更新処理による遅延時間分の更新データに不整合が発生するだけで、これは高々10分程度であるので、比率は許容できる程度に小さい。複製ファイル型ではデータとインデクスの不整合は発生しないが、切り替えのタイミングに関する問題が発生する。これらのことから複製レコード型が比較的問題の少ない方式といえる。ただし、参照目録 DB 本体は定期更新であることから、IR のインデクス更新とタイミングを合わせることが比較的容易であり完全共有型でも不都合はあまりない。

第二に、CAT の保守のために DB サーバを停止する場合の IR サービスへの影響の問題がある。このような場合、複製レコード型、複製ファイル型では複製を用いて IR サービスを行っているので問題ないが、完全共有型ではデータ本体を DB サーバに依存しているので IR サービスも停止せざるを得ない。DB サーバは原則として無停止で運転することになっているが、実際の運用は始まったばかりであるためどの程度安定して動作するか不明である。もし仮にシステムダウンが頻発するようであれば参照目録 DB でも完全共有型の採用をあきらめざるを得なくなるであろう。

3.5 IR システムの構成

前節の検討をふまえて決定した目録系 DB の IR システムの構成を図5に示す。

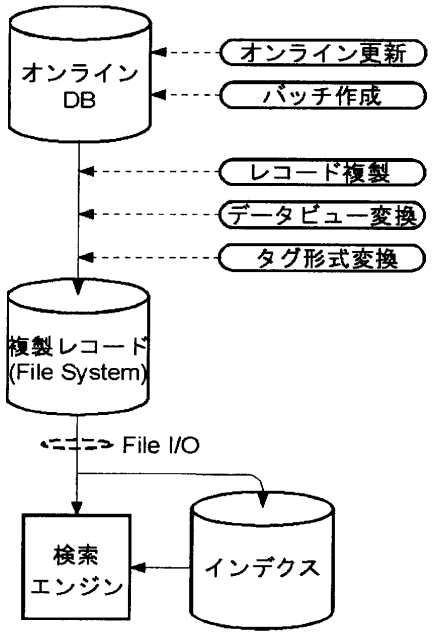
総合目録 DB では複製レコード型の構成とし、必要な変換は全てレコード複製時に行うことにより、OpenText6 のメタフィルタ機能を使用せずに標準的なインタフェースにより実現している。

参照目録 DB では完全共有型の構成とし、メタフィルタにより DBAC 経由で DB サーバにアクセスし、動的にタグ形式の変換を行っている。OpenText6 のメタフィルタとオンライン更新の動作は確認済みであり、DB サーバへのアクセス機能についてもほぼ開発を完了している。

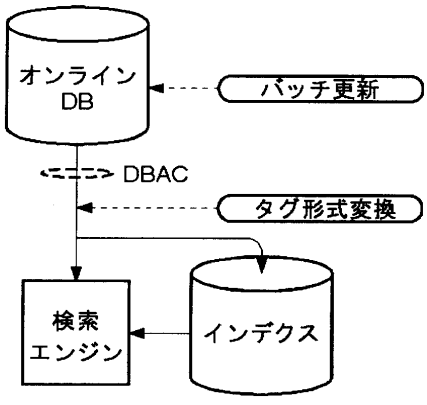
4 おわりに

本稿ではオープン系上で開発を進めている目録系 DB の IR システムの構成について、オンライン DB システムである CAT との運用上の関連、DB サーバへの性能上の影響、IR サービスの利用性などを中心に行った検討について述べた。現在、上に示した構成に基づき開発を進めており、開発を完了した部分から評価を行ってゆく予定である。

オンライン DB との連携を中心に検討を行ったため、オンライン更新に伴うレコードや検索結果集合のバー



(a) 総合目録DB



(b) 参照目録DB

図5　目録系 DB の IR システム構成

学術情報センター電子図書館サービス
© 学術情報センター

ジョン管理についての論理的取り扱いについては本稿では触れなかった。しかし、これもオンライン更新型の情報検索システムを実現する際に一般的に発生する問題であり、本システムの開発における重要な課題であった。この検討については別の機会に報告する予定である。

参考文献

[1] 神門典子, 木村優, 志津田嘉康, 大山敬三, 越塚美加, 小山照夫,「NACSIS-IR の検索機能の高度化」, 情報処理学会研究報告(情報学基礎研究会)95-FI-39, Vol.95, No.87, pp.57-64, 1995.

[2] 安達淳,「画面志向アプリケーション向けネットワークプロトコルの開発」, 東京大学文献情報センター紀要, Vol.1, pp.86-129, 1985.

[3] 安達淳, 橋爪宏達, 大山敬三,「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, Vol.1, pp.73-89, 1987.

学術情報センター電子図書館サービス　　©学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

物理的な空間としての研究室やオフィスにおける情報組織と視覚的な手がかりによる探索

Information Seeking in Terms of the Visual Data in an Office

学術情報センター　越塚　美加
Mika KOSHIZUKA
National Center for Science Information Systems

要旨

個人の情報利用行動を明らかにする上で、研究室における情報利用を無視することはできない。本稿では、図書館・情報学及びオフィス・オートメーション・システムの研究分野で発表された文献をもとに、研究室や会社のオフィス内での情報利用行動、パーソナル・ファイル、すなわちそこに蓄積される資料類の組織法と検索法について検討した結果、研究室内の情報探索は空間的な位置、資料の色や形等の視覚的な情報及びその資料を扱った時間に関する記憶によって行なわれることが明らかになった。

ABSTRACT

There are a few research work on information seeking and/or information organization in offices. This article discusses information seeking behavior and personal file organizations in offices. Spatial information, colors, forms of materials and usage time provide useful index to recall their contents and find a certain intended materials.

[キーワード]　情報探索、情報組織、研究室、空間、視覚情報
[Keywords]　information seeking behavior, information organization, offices, spatial information, visual information

1　個人レベルの情報利用としての研究室やオフィスにおける情報組織研究の意義

近年、情報要求や情報利用行動を「ある専門職集団」、「ある分野の研究者」などの「ある集団」に一般的な特徴として捉えるのではなく、「特定個人」がどのような情報をどのような状況下で要求したり、利用しているかを、利用者に視点をおいて個別に検討し、理解するという立場からの研究の重要性が指摘され、同時に、実際に調査や研究を行なう研究者が少しずつ増加してきた[1]。そのような立場をとる研究者は、情報要求や情報利用行動を個別の状況と結びつけて検討しながら、その上で個人が置かれた状況を越えて共通している点や共通していない点を検討することによって、情報要求や情報利用行動をモデル化し、説明しようとしている。このように個人に焦点を当ててその情報利用行動を明らかにしようとする際に無視することができない「場所」は、研究者にとってはその仕事の多くの部分が行われる研究室であり、会社に勤務している人の場合には勤務先のオフィスと言える。特に、研究者は、周辺によい図書館があり、その図書館でよい書誌や参考図書を利用できても、自分の研究室を始めとする、すぐに手の届くところにある資料のみを用いることがいくつかの既存文献で指摘されてきた。このことからも、研究室は情報源の蓄積場所としても重要な位置を占めると考えられる。

図書館・情報学分野では、数は少ないが、研究者が自分の研究室やその他の仕事場で、図書や文献、様々な書類やメモ等をどのように組織しているかの研究が行われてきた。これらの研究は、より使いやすい情報検索のインタフェース設計の基礎データとするために、研究で用いる様々な文献類をどのように組織し、また何を手がかりにその中から必要な情報を探しているかを研究するものと、研究過程に沿った情報利用の実態を明らかにし、個人レベルの情報利用をモデル化しようとするものとに大別できる。

例えば、前者の例として、情報検索システムの利用者が自分たちの研究室の中で非電子的なファイルをどのように組織するか、どうしてそのように組織するの

学術情報センター電子図書館サービス　　©学術情報センター

かという要因、どの様に検索を行なうか、どうして検索できるのかを明らかにしようとしたCaseの研究がある[2]。Caseは、現行のコンピュータのファイルを検索する方式は、手にとって内容を確認することのできない、ファイル名やそのファイルを保存した日時を頼りにした検索方式が中心であり、この方式の問題は、ファイル名を思い出せないと検索できないことと、ファイル名を見ただけではファイルの内容が分からないことにあると主張した。そこで、適切な検索方式を考案するために、人々が、実際に自分の蔵書や書類、メモ等、研究室にあるものをどのように組織化し、検索を行なっているのかを知る必要があるとしている。

後者の研究については、直接研究室内の情報利用を明らかにすることを目的としているわけではないが、人文・社会科学分野の研究者を中心とした研究者の情報利用に関する研究がある[3]。越塚は、ブラウジングは、問題状況を明確に定義できなかったり、単に手当たり次第にぱらぱらと雑誌をめくるだけではなく、研究において重要な役割を持っていることを示し、特に、一度ブラウジングされた文献類が研究室の書架上に並べられ、再度ブラウジングされる過程がアイディアをまとめる段階で本質的な役割を持っていると述べた。

個人の情報利用を対象とした研究は、オフィス・オートメーション・システムの設計の分野でも注目された。コンピュータシステムの設計者は、コンピュータの操作に習熟している。彼らがコンピュータに精通していない多くの一般の人々にとって利用しやすいシステムを設計するためには、一般利用者が自分たちの研究室やオフィスの中で何をしており、また室内に配置されている様々なものをどの様にして認知し、記憶し、記憶の中からいかにして思い起こすのかという方法に合わせて設計しなければならないと考えた[4], [5], [6], [7]。しかし、そのような考えに基づいてコンピュータシステムを設計した場合であっても、実際に利用者が机をどの様に使っているか、あるいは個人が自分のオフィスをどの様に組織するかを、きちんと理解し、体系立てた上で設計しているわけではないことがMaloneによって指摘されている[8]。そしてこれが、Cole[9]に端を発する、研究室やオフィスの中で情報をどのように組織化し、それらをどのように検索しているのかを明らかにしようとする研究の動機の一つとなっている。彼らは、個人にとってより効果的なオフィス・オートメーション・システムを構築するための基礎研究として、オフィスとは何か、オフィス内の資料類を日常的にどのように蓄積し、検索しているかを研究している。本稿では、こうした研究から研究室やオフィスの中で、どのように情報が組織され、また検索されているかを検討する。

2 研究室やオフィス、パーソナル・ファイルを捉える視点

2.1 既存の研究にみられる研究室やオフィスを対象とした研究の分類

Hirschheimは、オフィスを対象として行なわれた記述的な研究は二種類に大別できるとしている[10]。まず第一に、オフィスそのものを様々なタイプにカテゴリー化する研究で、たとえば、Panko[11]の研究を挙げることができる。

Pankoは、1981年に中規模の銀行の上級幹部(senior executive)を対象として、経営学を専攻した学生に何を期待するかというインタビューを行なった。しかしそのインタビューの中で、約30分の時間をそれら上級幹部が自分のオフィスに持たせている機能についての質問と、その機能をきちんと果たすようにオフィスを整備した過程についての質問に費やした。このインタビューの結果、Pankoはオフィスを二つのタイプに分けている。Pankoがタイプ I としたオフィスは、会社の日課となっている情報処理を行なう部署で、たとえば、経理課、給与課、ワープロセンターなどがこれに相当する。これらの部署の仕事の手順は固定的であるため、オフィスにコンピュータを導入し、作業をコンピュータに処理させることが可能であるとし、OA化によって業務の効率が上がることは明らかであるとした。これに対して、タイプ II にカテゴリー化されたオフィスは、遂行する仕事がルーティン化していないような部署で、たとえば、経営部門の管理職のオフィスや法律に関連した仕事を行なう部署、経営企画室などがこれに相当するとしている。そして、これらのオフィスの特徴として、仕事を行なう定まった手順がほとんど見られず、電子メールや意思決定支援システム、部署で作成したり会社で作成したデータベースなどのツールを取り入れることで業務の効率を挙げることができるとした。

Hirschheimがもう1つのタイプとして分類した研究は、オフィスにおいて目に見える様々な行為を分類する研究である[12] [13]。このカテゴリーに入る研究は、各研究者がそれぞれが設定した体系に従ってオフィスでの仕事の分類を行なっているが、基本的には各研究

学術情報センター電子図書館サービス　　©学術情報センター

表1　Hirschheimによって示された研究室やオフィスを分類する2つの見方

	分析的な見方	解釈学的な見方
① 研究室やオフィスの捉え方	組織をうまく運営するために、人々がその中で様々な機能を果たす環境としてオフィスを捉える。この際の「機能」とは、形式的で、構造化された行為によって示される。	ほとんどが構成化されておらず、インフォーマルな行為によって成り立っている。
② 研究室やオフィスの機能	決定論的で、合理的、かつ外部から見ても明らかなものとして示される。	非決定論的で政治的な要素があり、外部から見て明らかではない。
③ 組織	情報や仕事の流れを表現した線によって結ばれるロール・ボックス(role-box)を並べたようなイメージとしての「構造」と捉える。	行為者、あるいは文化として捉える。ある組織に所属している人にとって、組織は、社会的な目的を達成するための道具であり、その中で共有された意味や記号、慣例、認知的な枠組みを作り出す、小規模の制約された社会を実現するための道具であると考えられる。
④ 研究室やオフィスでの行為	明確な行動として現れるので、観察可能である。	行為者の間でのみ共有されると考える。
⑤ 研究で採用した方法論	経験的な方法を用いて形式的なモデルを構築し、そのモデルを通じて量的な手法によって研究できる。	現象学的な研究によって明らかにする。
⑥ 研究の焦点	オフィスとは何かという問題を、オフィスでの仕事や機能を構成要素に分割して分析することに焦点をあてている。	現象を理解することにある。すなわち、ある特定のオフィスで実際に起こっていることやオフィスでの行為を、そのオフィスで働いている人はどのような意味があるものと捉えているのか、また、どのような意図を持って行動しているのかを、理解しようとすることに焦点がある。

者が提示した分類は、共通性を備えていることが指摘されている[14]。

Hirschheimは、上述のように二種類に大別できる先行研究でのオフィスの分析方法では不十分であることを示した。なぜなら、前者のようにオフィスそのものを分類しても、後者のように目に見える行為によって分類しても、それはオフィスが備えている多面性を反映していないためである。「オフィス」という概念は言葉では定義しにくく、またオフィスの機能の多面性を考慮すると一つのオフィスを単一のカテゴリーのみに分類することは不可能である。Hirschheimによれば、一つのオフィスは次のような様々な側面によって分類可能である。

地理的な側面　(物理的な場所)
時間的な側面　(そこで仕事をする時間)
行為の側面　(行なわれる仕事)
構造的な側面　(そこで働く労働者によって報告される関係)
空間的な側面　(人々が同僚とある関係を保つ場所)
経済的な側面　(組織を運営し、それによって労働者が評価される経済的な規準)
社会的な側面　(オフィスで仕事をする人々を動機づける社会的、心理的要因)

このように様々な側面がある以上、一つのオフィスを単一のカテゴリーにのみ分類することは、本来オフィスが備えている機能を隠してしまう危険性があることが分かる。そこで、Hirschheimは、既存の研究でのオフィスの見方を、表1に示したように、二つに大別した。すなわち、分析的な見方と解釈学的な見方である。

学術情報センター電子図書館サービス　　© 学術情報センター

両者を比較すると、分析的なみかたは、最初から一般化ができることを前提としており、オフィスはどのような調査者が見ても、観察可能で明らかな機能から成り立っているという立場を取っている。一方、解釈学的なみかたは、調査対象者自身にとってオフィスやオフィスの中での行動がどのような意味を持つかを理解しようとする立場を取るという点で大きく異なっている。この点を捉えて、Hirschheim は、Burrell と Morgan の言葉[15]を借りて、分析的なみかたは「客観主義」、解釈学的なみかたは「主観主義」に対応すると述べた。

2.2 Hirschheim の分類に基づいた本稿で検討する研究の分類

本稿で検討する Cole[9]や Malone[8]、Kwasnik[16]の各論文は、研究対象としている研究室やオフィスとは何か、あるいはパーソナル・ファイルと呼ばれる、個人の蔵書や書類、メモ類はどのような役割を果たしているのかという点を特に論じることなく、研究室やオフィス、パーソナル・ファイルの中にはどのような情報があり、その情報はどのように組織され、またどのような方法で検索されるかを見ている。Cole らにとってオフィスやパーソナル・ファイルとは、「暗黙のうちに了解されている何か」であり、少数の特定のオフィスやパーソナル・ファイルに対する考察から一般化を行なおうとするという点で分析的なみかたをしている。

その一方で、Cole らは、オフィスとは何かを明らかにしようとするためには、実際にそのオフィスやファイルを使っている人自身からみて、その時の状況を考慮にいれた場合にどのように理解できるかという観点から調査を行なっている。この意味で、彼らは解釈学的なみかたをしているといえる。さらに、調査手法としては、文化人類学で用いられているフィールドワークの手法を取り入れたり[8]、非構造的なインタビューを用いる[9]など解釈的な手法をとっている点からも、研究全体を貫く基本的なみかたは、解釈学的なみかたであるといえる。ここで言う非構造的な面接とは、非指示的面接とも呼ばれ、調査票を用いるなどあらかじめ決められた枠組みに沿って決まった項目を尋ねる構造的な面接とは異なり、質問事項は、調査者が被調査者とのインタビューの過程で変えながら行なう面接法である。「対象者の生活状況や意識に一定の構造があることを主眼とする面接法である」と言われているように[17]、対象者自身の構造をインタビューによって引き出すことを主要なねらいとしており、質的な方法を採用する場合の主要な調査手法の一つである。

Cole や Malone、Kwasnik らの研究は、Hirschheim の分類では「分析的なみかた」と「解釈学的なみかた」の両方の性格を兼ね備えている。しかし、Hirschheim が述べたような、オフィスに関する多面的な検討は行なっていない。たとえば、人は状況に応じてさまざまな役割をこなす。一人の研究者は、場合に応じて、学部長であり、教員であり、外部の諮問委員会の委員であったりする。しかし、研究室やオフィスの持ち主のその時の役割との関連づけを行なった上で、オフィスの機能を明らかにしたり、オフィス内にあるさまざまな資料の組織法がどうしてそのようになっているかを明らかにするといった検討は行なっていない。また、後者の解釈学的なみかたに関しても、安易に「オフィスとは」という一般的な話に結びつけてしまっている点で問題がある。

しかし、そのような立場の不十分さはあっても、これらの研究が取る視点は意義深いと思われる。すなわち、オフィスやパーソナル・ファイルがそれらを組織している人にとってどのような意味を持ち、どのように組織しているかという点にとどまらず、なぜそのように組織しているのか、という点まで踏み込まなければ理解できないという立場に立ったことは、システムやインタフェースの開発にとどまらず、情報利用行動を明らかにする上でも意義深く、評価されるべきであろう。

3 研究室やオフィス、パーソナル・ファイルとは何か

第 2 章で、Cole[9]や Malone[8]、Kwasnik[16]は、研究室やオフィス、パーソナル・ファイルはどのようなものであるかは、直接問題にしていないことを指摘した。そこで次に、オフィスやパーソナル・ファイルが具体的にどのようなものであり、またどのように組織され、検索されているのかを検討する。

3.1 オフィスに置かれているものとその組織法

Case[18]は、1986 年に発表した論文で、60 人の大学教授を対象にインタビューをし、彼らのオフィスの中にある資料のスタックの数、本や雑誌、ノートが並べられている書棚の数、ファイル・キャビネットの引き出しの数、その他の書類の蓄積場所の数を実際に数えた。ここで言う「スタック」とは、後述する Malone が「パイル」と呼ぶ書類の山とほぼ同義である。その結

学術情報センター電子図書館サービス　　©学術情報センター

果、書類を蓄積するための家具類が不足しているわけではないのにもかかわらず、彼らは平均して19のスタックをオフィスの中に形成していることが分かった。

また、1991年に発表されたCaseの論文[2]では、歴史学者20人を対象に、単に資料類の量を数えるだけではなく、彼らが自分のオフィスで資料をどのように組織し、またそれらの資料をどのようにカテゴリー化しているか、どのようなラベルを付与しているか、すなわち記憶や検索に資料を表現するためにどのような隠喩が用いられているかを調査した。その中で、研究室は全体として乱雑であるが、そこに置かれている備品は以下の4つに大別した。

① 家具: 机、本箱、ファイル、キャビネット、椅子など
② 機器: 電話、コンピュータ、マイクロフィルムのリーダーなど
③ 個人の私物:衣類、食器、食料、ラジオ、美術品、おもちゃなど
④ 情報関連の人工品(information artifacts): 本、雑誌、新聞、論文、ノート、メモ、磁気テープ、AV機器

そして、研究室は特に、印刷物などの情報関連の人工品によって占められていることを明らかにした。被調査者の中には自宅で仕事をする人もおり、その人にとっては、研究室や教育や学内のマネジメントに関する資料の所蔵場所となっていることが示された。

Malone[8]は、専門家と事務職従事者がオフィスの中においてあるものを対象とした調査を行なった。その時、机、テーブル、ファイル・キャビネットなど、研究室内で情報を蓄積しておく場所を総称して「机(desk)」と呼び、それに床なども検討対象として加え、これらの組織法、検索法を調査している。

また、オフィス内に置かれている書類の束を二種類に区別した。そのうち、「パイル」と呼ばれるものは、タイトルがまったく付与されておらず、単に積み重ねられた書類の束で、研究室の比較的表面にある。もう一つをファイルと呼び、明確なタイトルは付与されていないが、論理的に整理されている書類の束を指している。ファイルは、通常キャビネットの中などに入れられており、外側からは見えない位置にある。

標準的な書式の書類を扱い、書類の流れがルーティン化しているような庶務の仕事を行なう人のオフィスの場合は、図1として示された通り、書類の流れに合わせて書類が配置されており、きちんと片付いたオフィスとなっている。そして、このような場合、役立つ資料はみなテーブルの上に出してあり、流れる書類は常に引き出しなどの入れ物の外にあること、未決ファイルの中にある書類は決して引き出しにはいれないといった資料の蓄積場所に関してきちんとした説明ができることを示した。実際に、このオフィスでは作業の流れと、オフィス内のパイルやキャビネットなどに整理され、綴じ込まれているファイルが密接に結びついていた。書類の山であるパイルも単に重ねられているだけではなく、その中に規則性がある精密に組織されたものであることが示されている。すなわち、新しくそのパイルに入れられる書類をパイルの一番下に置くことで、次に処理しなければならない書類が一番表面にくるようになっている。またパイルの数も常に5から10の間で、比較的少なかった。

一方、日常的な書類の流れがほとんどない研究者の場合、図2として示されているとおり、研究室全体は乱雑で書類の束も大部分がパイルであり、そのパイル類も非常に大ざっぱにしか分類されていないことが指摘されている。ただし、机(この場合は家具としての机)の上は、綴じ込まれたファイルよりも新しい書類が置いてある点では規則的だった。

このようにオフィス内には、多くの資料があり、それらは持ち主の仕事の種類に依存した形で、組織されていることが分かった。

3.2 情報組織を行なう際の制約と組織法

従来、図書館は媒体や分野が多様な大量の資料を組織し、その検索手段としての目録や二次資料類を提供してきた。図書館における資料の組織は、主題による分類に注目が集まりがちではあるが、配架の都合で図書、雑誌、マイクロフィルムといった媒体の形状によってまず分けられており、その中を内容の主題分類によって組織していることが多い。Maloneによれば、研究室やオフィスの中で、この図書館方式と似たような方法で、すなわち、資料の内容の分類や書誌事項をカードに書き込んで目録を作成し、相互参照をつけるなどして検索手段を維持しようとした研究者もいた。しかし、インタビューを行なったときには、既に床の上に多くの文献が散らばっていたことなどからも分かるように、実際に個人の研究室やオフィスの中で図書館方式を維持することは難しい[8]。

Cole[9]は、オフィス内の情報は、情報の使用頻度が組

学術情報センター電子図書館サービス　　◎学術情報センター

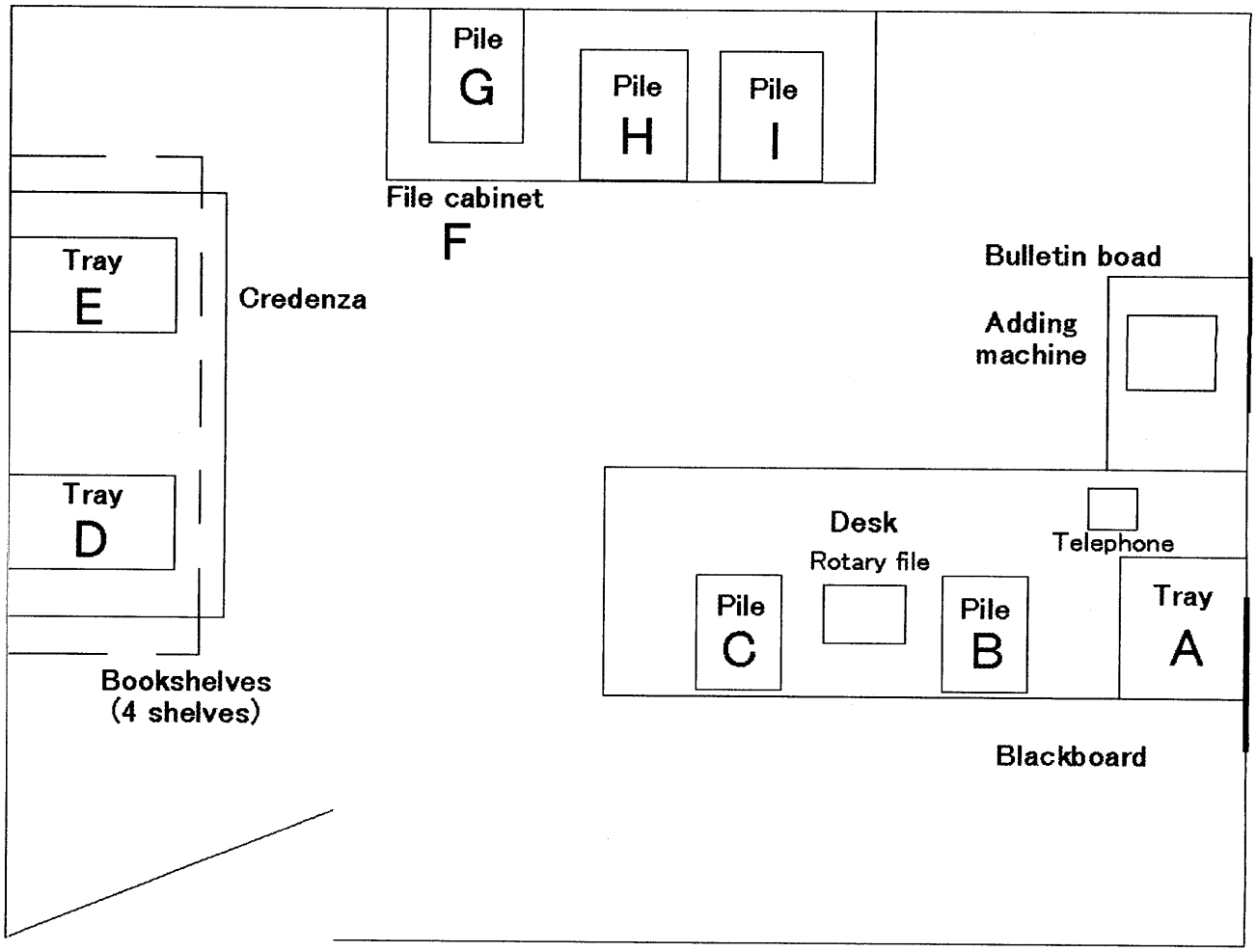


図1　整理されているオフィスの例

出典: Malone, Thomas W. How do people organize their desks?: implications for the design of office information systems. ACM Transactions on Office Information Systems, Vol.1, No.1, p99-112 (1983) より再トレースした。

織化の程度に関係があるとした。そして、使用頻度が高いほど組織化の程度は低く、逆に使用頻度が低いほど厳密に組織されることを示した。Cole は、情報を使用頻度が高い順に三段階に分けた。そして、この三つの段階を「action information」、「personal work files」、「archive storage」と呼んでいる。各段階は、次のような特徴を備えている。

まず、action information は、近い将来処理される、あるいは処理されるだろう情報のことを指しており、文献だけではなく、日記やするべきことのリストといったスケジュールに関するメモも含まれている。置かれている場所としては、未決書類入れ、机の上、椅子の上や床の上など、表面に近いところや作業領域である机のそばに置かれており、すぐに利用できる形になっている。ただし action information の量が増えすぎると検索が困難になるため、各情報とそれらの位置を記憶中に覚えていられる程度の量を保っているときにのみ効果的に利用される。そして、形式的に組織されている度合いがもっとも低い一方で、もっとも頻繁に用いられる情報でもある。

時間が経過するにつれ、action information として維持されなくなっても残される情報を、personal work files と呼んでいる。personal work files は、オフィスにじかに接している場所にある、ファイルされた情報である。ここで言うオフィスにじかに接している場所とは、ファイルの中や戸棚の中、ファイルキャビネットの中、利用者の机の中などを指すが、それらは、オフィスの中にある場合もあれば、たとえば廊下などのオフィスに隣接している場所も含んでいる。personal work files にある情報は、利用者の現在進行中の仕事に関連している。たとえば、進行中のプロジェクト等がこれにあてはまる。

学術情報センター電子図書館サービス　　© 学術情報センター

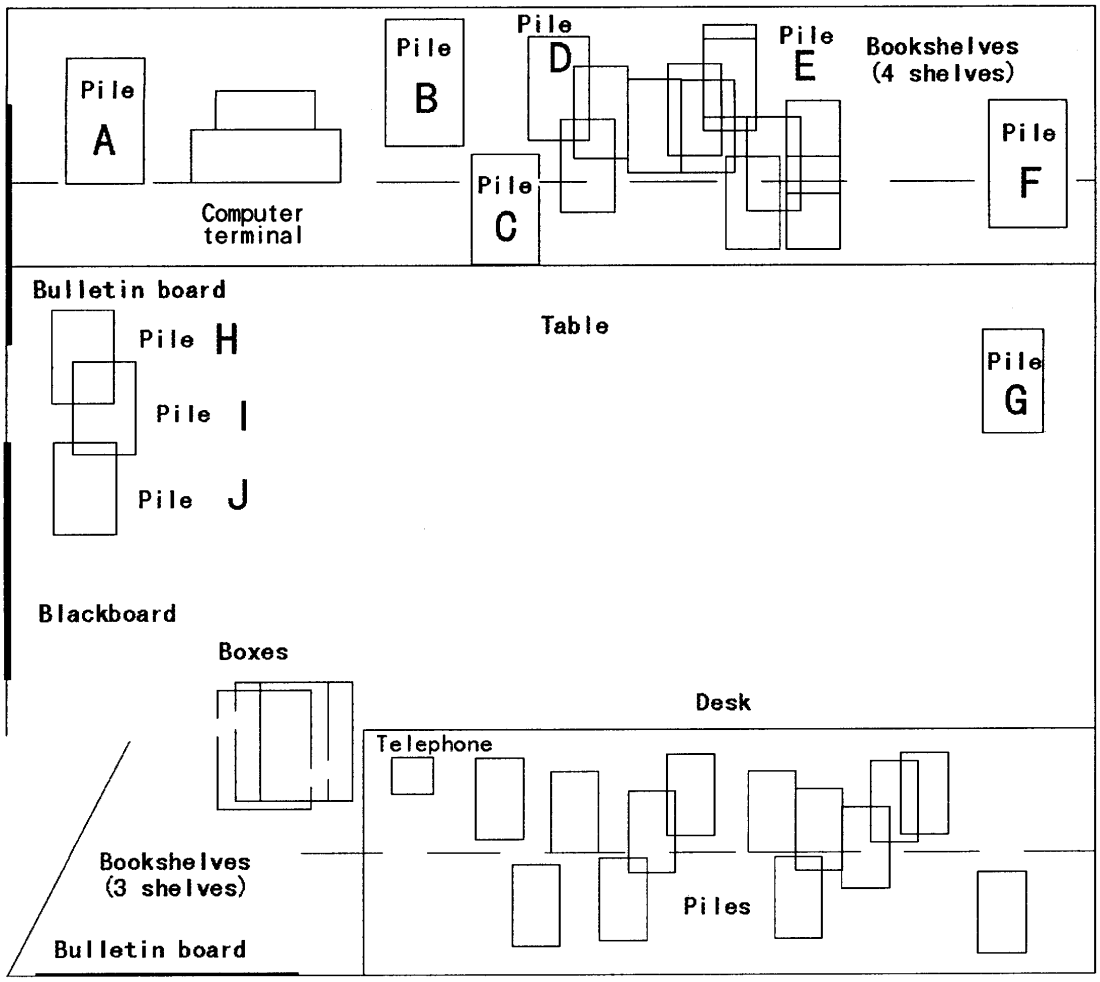


図2　乱雑なオフィスの例

出典: Malone, Thomas W. How do people organize their desks?: implications for the design of office information systems. ACM Transactions on Office Information Systems, Vol.1, No.1, p99-112 (1983) より再トレースした。

さらに、利用者のこれからの仕事に直接関係のない情報は、archive storage へ送られる。archive storage には、長期的に蓄積される書類が含まれ、蓄積システムが構築される場所は、オフィスからは離れている場合が多い。また、組織化の程度はもっとも厳密であることを指摘した。Cole が調査した範囲の被調査者は、基本的に action information と personal work files という、組織の程度が比較的緩やかなものを好むという考察を示している。さらに、新たに情報の蓄積システムを設計しようとする場合には、内容の分類に基づく厳密な組織法を考案するが、この方式は実際には長続きしないことが分かった。そして、実際に行なわれている組織法は、単純で緩やかな分類に基づいていることを示した。そして、離れた場所に蓄積されるほど、利用頻度が低くなるが、組織化の程度は高いことを示した。また、archive storage は、現行のコンピュータ・システムが採用している方式に近いが、実際にオフィス内で行なわれている組織法とは全く異なっている。

Case[2]は、オフィスや研究室内で行なわれている組織化が次の四つのレベルで行なわれると述べている。すなわち、① 空間的な制約による組織化、② 資料の形態による組織化、③ 主題による組織化、④ 扱い、目的、質による組織化である。そして、これらのうちのどれか一つに依存して、実際の組織化が行なわれるわけではなく、いくつかを組み合わせて組織化が行なわれているとした。

Case は、研究室やオフィスにおける組織化は、第一に「似たものを一緒に」という原則にしたがって行なわれているとした。「似たもの」には様々な種類があり、「よく使うもの」を意味したり、書架の棚の高さに合わせて資料の高さが似ているものを意味することもあれば、同じ曜日の授業に関連するテキストなど、

学術情報センター電子図書館サービス　　©学術情報センター

資料が果たす機能が似ていることを意味する場合もある。しかし、この原則には、いくつかの制約があり、そのひとつが空間的な制約であると述べている。この空間的な制約によって、自分が持っているスペース、すなわち、研究室やオフィス、その他、蓄積場所として利用できる場所の広さに依存して、蓄積される文献類の量が自ずと決められる。また、蓄積しようとする資料の物質的な制約によっても組織の方法が異なってくる。たとえば、図書を置こうとする書棚の高さに合わせて置く場所が制限される。また、生活空間としてのオフィスや研究室を考えれば、見た目による組織化も必要となる。これらのことから、オフィスや研究室で資料を組織しようとするときには、まず第一に空間的な側面に制約されると言える。

オフィスや研究室で資料を組織するときに、受ける第2の制約は、資料の形態によるものである。たとえば、雑誌は一般に図書とは別の場所に蓄積されており、図書はそれを所蔵している人が考えた順序で配架されているが、基本的に雑誌は、タイトル順に配架されている。調査対象の中には、ある図書と雑誌論文の相関をつけるために索引を作成している人もあるが、概ねそれらは簡略で、多くの研究者は論文執筆時に関連文献をまとめてリストアップする手法をとっている。すなわち、雑誌論文に関しては、あまり内容面での分類は行なっていない。しかし、研究の各プロジェクトごとに使用した文献について作成したカードをまとめておく歴史学者もいた。歴史学者は、文献カードを維持している割合が多く、約60％と過半数を占めている。しかし、そのカードは、現在の研究に関する文献のみ絞って作られている場合が多く、所蔵資料全体の索引を作成しようとする試みは例外なく失敗している。この結果は、実際に情報を蓄積する場合にはできるだけ単純で緩やかな分類を行なうことが重要であり厳密な分類は失敗するという、Cole[9]と同様の結果を示している。

空間的、形態的に組織された文献は、次に、トピックによって組織される。歴史学の場合、主題の他に、時間、地域による組織化がよく行なわれている。Case[2]は、トピックに関しては、研究分野の特性によって異なるだろうとしている。

文献を組織するときの一番下位にある制約は、その資料の扱いや資料を用いる目的、資料の質によるものである。たとえば、伝記などのジャンルやその伝記が読む価値があるものかどうかという質、授業のテキストとして用いるといった質による組織化が行なわれていることを示した。

3.3 資料類を組織する際のカテゴリー化の方法

Cole[9]は、一般に、計画段階では複雑なファイリング・システム、すなわち、資料の内容によってある色を割り当て、その色別に作成したファイルを維持しようとしたり、階層を作って厳密に分類するなど、複雑なファイリング・システムを構築し始めようとするが、実際に長期間実行され、維持されるのは単純な分類が多いと述べた。

資料類のファイリングの方法に対しては、Coleと同様に、Malone[8]も比較的単純な分類が好まれているとした。Maloneは、さらに同じ文献に対しても、蓄積を行なうときに割り当てたカテゴリーと検索しようとするときに割り当てるカテゴリーが一致しないことも指摘している。Maloneが指摘したように、同一の文献に対してであっても、それが用いられる時間的な違いや、その文献を用いようとする目的によって、同一人物が異なるカテゴリー化を行なうことはしばしばある。このように資料類に対して割り当てられたカテゴリーは永久に固定的に用いられるわけではない。また、用意されるカテゴリー自体も変動的であり、古いカテゴリーが統合されたり、分化して新しいカテゴリーとなったり、新しい研究プロジェクトに携わることで、まったく新しいカテゴリーが出現することもある。

Kwasnik[16]は、大学の教員が文献をどのように組織しているかを調べるために、彼らが毎日届く様々なメールをどのように仕分けしているか、すなわち、どのようなメールをどのようにカテゴリー化しており、既存のカテゴリーとどのように整合性を維持するか、新しいカテゴリーを作成するのはどのような時かを調べた。Kwasnikがここで「メール」と呼んでいるのは、大学のメールボックスに入れられる様々な資料類である。これには字義通りの郵便物をはじめとして、学内の会議の案内などの通知類など多様なものが含まれている。

Kwasnikは、まず、被調査者に研究室やその他の資料保管場所を案内してもらって、そこには何を置いているか、また、被調査者が個々の事物にどんな呼称を与えているか、すなわち、資料類を中心とする研究室内の事物をどのような名前の下にカテゴリー化しており、どのようなものが同一のカテゴリーに入れられるかを調べた。すなわち、資料類を組織する上での既存のカテゴリーの洗い出しを行なった。次に、各被調査者の

学術情報センター電子図書館サービス　　©学術情報センター

メールボックスに入れられているメールを、どのような状況でどのようにカテゴリー化するかを実際に被験者自身に作業を行なってもらうことによって観察した。その際に、頭の中で考えていることをすべて口に出していってもらい、観察結果とあわせて分析の対象としている。そして、Kwasnik はただ単に被調査者である教員が自分たちのメールを仕分けするところを観察し、報告するのではなく、彼らが行なっているカテゴリー化の作業を Kwasnik 自身が正確に理解できているかどうかを調べるために、自分自身でも同じ作業を行ない、被調査者自身が仕分けするのと同じように仕分けできているかどうかを確認してもらった。このように複雑な手順をとった理由は、カテゴリー化という処理が、その結果は目に見えるが、過程は目に見えない処理であることによる。そのため、Kwasnik が得たカテゴリー化についての結果の妥当性を確保する方法としてこのような方法をとっている。

その結果、次の図3に示すように、研究室をはじめとした最初の案内とメールをカテゴリー化する際に被調査者が用いていた側面は33種類であり、それは7つのカテゴリーに大別できることがわかった。そして、一番多く用いられたのは、カテゴリーとしては、「状況に関する属性」のカテゴリーで全体の33.3%を占め、次いで「文献に関する属性」(29.4%)、「処理」(12.6%)のカテゴリーだった。側面としては、「形式」(14.7%)、「使用」(13.1%)、「トピック」(8.5%)、「場所」(8.1%)、「環境」(7.6%)、「時間」(8.5%)の諸側面からのカテゴリー化が多いことがわかった。これらのカテゴリーは、図書館で行なっているような内容による分類と一致するものはトピックのみであり、コンピュータで通常行なわれているファイルの名称や検索のてがかりとも異なっている。

また、Kwasnik 自身がメール類を分類する実験の結果から、被調査者の分類方法と異なった分類をしてしまった要因を分析している。それによれば、メールの主題を抽出し間違えたことと、被調査者が行なった実験のときと Kwasnik が行なった実験のときとでは環境が変わってしまっていたことが最大の要因としてあげられている。後者の結果から、カテゴリー化の方法は、環境によって異なることが示されており、Malone の研究結果とも同じ結果を得ている。

Case は、自分の経験を整理する上で、各経験にラベルをはってカテゴリーに分類することは、人間の基本的な特性であるとし、その上で、カテゴリー化と比喩表現は関係が深く、用いられる比喩表現も日常的に使い慣れている道具に関する表現が、同じ機能を果たす他の事物を示すときに用いられることを見出した[2]。Case が調査対象としたのは、歴史学者であり、歴史学者は研究に図書館の目録に用いられるような「カード」を用いることが多いことから、「カード」に関する表現を比喩として用いていることを示した。

3.4 検索の手がかり

図書館では、所蔵されている図書を探すてがかりとして、通常、「著者名」や「書名」の他に「主題」や「分類」といった内容を示す、「言語による」てがかりを用意している。しかし、パーソナル・ファイルの場合、主として用いられるのは内容に関係ないてがかりであることが、今回検討したすべての文献で示されている。この要因としては、図書館とパーソナル・ファイルの全体量の違い、また、パーソナル・ファイルは受入れ作業や整理作業、配架などすべて特定の人が利用者でもある自分のために行なうこと、さらに、オフィスや研究室では、資料類のある場所という空間的、視覚的なてがかりを利用できることによる。

Cole[9]は、パーソナル・ファイルを維持する上で索引を維持しない代わりに、部屋という空間が検索に重要な役割を果たしていることを示した。すなわち、ある空間的な位置と特定の資料をそこに置いたという記憶を組み合わせて検索を行なっているとした。三次元のオフィスは、蓄積された情報を直接指示するような多くのポイント、たとえば、色や書架のどのあたり、といった多くのポイントを持っているとしている。しかし、これはパーソナル・ファイルとして持っているすべての資料類について言えないことも述べている。これを2節で述べた「action information」、「personal work file」、「archive storage」と関連させて説明すると次のようになる。

まず、「action information」の場合、形式的に組織されている度合いが最も低く、明確なカテゴリー化による分類や索引を作って相互に関連づけるといった作業は通常行なわれていない。しかし、最も頻繁に利用されている情報であるため、必要とされる情報があるとそれがどの位置にあるかということで検索することが可能である。つまり、頻繁に利用するため、それが置かれている場所を常に把握していられる。このことから、利用者にとっての認知モデルは、空間的な要素に支配されていると言える。しかし、逆に、ある情報がどこにあるかを空間的に識別できなくなることを避け

学術情報センター電子図書館サービス ◎学術情報センター

るために、相互作用を頻繁に行なわなければならないことを指摘している。

「personal work file」の場合は、カテゴリカルな関係と情報の空間的位置に対する認識の両方に基づいて検索が行なわれている。personal work file に入っている情報は、action information と比較して、利用される頻度が低くなる。それは、ファイルが作られてから、どれくらいの時間が経過したかということとも関係があり、時間がたつにつれて利用される頻度が低くなるとされている。そして利用されなくなってくれば、空間的に識別できる可能性が低くなってくるため、より形式的なカテゴリーの構造が出現し、空間的な認識の代わりに情報へアクセスするために用いられるようになる。「archive storage」は、最も形式的に組織されたレベルの相互作用が行なわれる。ここに含まれる情報に対しては、もはや空間的な位置に対する知識は乏しくなっており、長い時間にわたってあまり利用されることがなくなるか、仲介者や仲介するツールを通じてアクセスされるようになる。たとえば、情報管理のために雇っている人や、コンピュータを利用したデータベースシステム、索引などがそれにあたる。このように形式化されているので、逆から見れば、archive storage の組織に用いられているカテゴリーの構造を理解しなければ、適切なカテゴリーに組み込んでいくことで実行する「蓄積」も、探したい資料が含まれているであろう適切なカテゴリーを探し出して、その中にあるかどうかを確認することで実行する「検索」も行なえなくなってしまうことになる。

Case[2]もやはり、研究室内のパーソナル・ファイルを検索する上で、視覚的なてがかりが重要な役割を果たしていることを指摘している。Case が研究対象とした歴史学者は、検索は記憶に頼ることが多く、しかも自分の記憶力に自信を持っていることを見出した。たとえば、それらの被験者のうち、「自分のものがどこにあるか、全部把握している」と答えている人もいれば、「自分で作ったノートの内容は全部覚えている。何か月かに一回は全部通して読むからね」と答える人もいた。20人の歴史学者のほとんどが自分の記憶に頼って検索し、しかもどこにあるかの記憶について自信があることを述べている。例外的に、研究室内で必要なものををうまく探せなくて困っていると述べた人でも、ある文献に探している内容がのっているという記憶が間違っており、文献そのものは探すことができると述べた。

Case は、歴史学者の特性として、細かい物事をきちんと覚えていられる能力が要求されるため、記憶に頼った検索がうまくいくことは驚くに当たらないとしている。そして、この検索には、学習の際に行なったことと、その時の環境が記憶に大きな影響を与えるとしている。すなわち、ある文献を読んだときに、それはどのような仕事との関連で読んだのか、その時に同時に進行していた仕事はどのような仕事でその仕事に関連する文献はどのようなものだったか、といった環境が大きな影響を与えるとしている。その意味で、部屋が乱雑であっても、その乱雑な状態がある文献がどこにあるか、何と一緒にあるかといった記憶の助けになるとしている。

以上のような研究結果から、パーソナル・ファイルのうち、特に現在進行している仕事に関する文献の検索には、空間的、視覚的な記憶、および探そうとする文献を最後に使用した時にその他に何をしていたかという記憶が重要な役割を果たしていることが分かった。この節の最初でも述べたが、こうした検索方法は、主題をもとに行なう従来の図書館での検索や商用のオンライン・データベースの検索方法とは異なっている。

3.5 パーソナル・ファイルの特殊機能：「思い出させる」機能

Malone[8]は、オフィスにおかれている多くのパイルは、情報組織上のカテゴリが厳密であれ、ゆるやかであれ、ただ単に、「ファイリング」する機能だけではなく、しなければいけない仕事やスケジュールなどを「思いださせる(reminding)」機能を備えていることを指摘した。すなわち、机の上に読まなければいけない資料を表面に近いところにおいたり、出席しなければならない会議の案内を机の前に張っておいたり、作成しなければならない報告書に関係する資料や書類をひとまとめにして普段作業する場所に近いところへ置くことで、それらを思い出させるようにする機能もパイルには備わっている。

4 まとめ

研究室やオフィス、そこに置かれているパーソナル・ファイルについての研究を検討した結果、次のようなことが分かった。

- 各ファイルに付与した索引や主題に関する情報を手がかりにして検索を行なうだけではなく、空間的なてがかり、視覚的なてがかりを用いて

学術情報センター電子図書館サービス　　© 学術情報センター

状況に関する属性
- アクセス
- 環境
- ニード／要求
- 文献の所有者
- 私に関係がある
- 部屋／空間
- 情報源
- 利用／目的

文献に関する属性
- 著者
- 形式
- 主題
- タイトル
- 物理的な属性

処理
- 変更
- 廃棄
- 保存
- 場所
- 延期

状況／組織
- 蓄積
- グループ
- 分離
- 未整理

時間

価値
- 重要
- 面白い
- ぜひとも必要
- 価値がない
- 秘密／機密
- 価値があるか分からない
- 私にとって役に立つ

認知の状態
- わからない
- 覚えておきたい
- 知っている

図3　メール類の分類を決定するときに用いられたカテゴリー

出典: Kwasnik, Barbara H. The importance of factors that are not document attributes in the organization of personal documents. Journal of Documentation, Vol.47, No.4 (1991), p.392 より

検索を行なっている。
- 通常、厳密なカテゴリー化を行なわずにファイルを管理をしている。
- 各ファイルの機能には「ファイリング機能」と「検索機能」に限らず、「思い出させる機能」がある。

一度目を通した図書や文献、書類を研究室という空間に配置し、そこから必要な資料を検索する時にはそれらの資料の視覚的な手がかりを利用するという振る舞いは、言い替えれば、各人の頭の中に記憶される、その人自身が読み取った内容を研究室という物理的な空間の中に展開し、視覚的な手がかりを索引として検索に供するという過程であると言えよう。例えば、越塚は、ドイツ文学研究者の研究過程における情報利用の段階の一つとして、研究者が既にブラウジングした文献を研究室で自分がすわる位置から見える書架上に並べ、それらを常に視野に捉えることで、物理的な文献を索引代わりにして書かれた内容を頭の中で反復的に思い起こし、アイディアを熟成させていると述べた。すなわち、「他の本との物理的な位置関係等を用いて可視化し、索引とすることによって、具体的な操作も可能なようにしている」とし、さらに、「ブラウジングした結果によって、配架された本の位置を物理的に変え、その本で象徴している内容同士の関係を異なるものに組み替える」ことを指摘し、この過程が研究の

学術情報センター電子図書館サービス　　© 学術情報センター

重要な部分を占めていることを示した[3]。特に文献類を研究の対象としている分野では、研究室が実験室の役割を果たしていると言える。

参考文献

[1] Fidel, Raya., "Qualitative Methods in Information Retrieval Research", *Library and Information Science Research*, Vol.15, No.3, pp.219-247, 1993.

[2] Case, Donald Owen., "Conceptual organization and retrieval of text by historians: the role of memory and metaphor", *Journal of the American Society for Information Science*, Vol.42, No.9, pp.657-668, 1991.

[3] 越塚美加,「文献のブラウジングが研究過程に与える影響」, 学術情報センター紀要, No.8, pp.131-142, 1996.

[4] Goldstein, I.P., "PIE: a network-based personal information environment", *Proceedings of the Office Semantics Workshop*, Chatham, Mass., June, 1980.

[5] Goldstein, I.P.; Bobrow, D., "Descriptions for a programming environment", *Proceedings of the 1st Annual National Conference on Artificial Intelligence*, Stanford, Calif., August, 1980.

[6] Negroponte, N., "Books without pages", International Conference on Communications, Boston, Mass.: *IEEE156*, Vol.56, No.1, pp.1-8, 1979.

[7] Smith, D.C. et al., "Designing the Star user interface", *Byte*, Vol.7, No.4, pp.242-282, 1982.

[8] Malone, Thomas W., "How do people organize their desks?: Implications for the design of office information systems", *ACM Transactions on Office Information Systems*, Vol.1, No.1, pp.99-112, 1983.

[9] Cole, I., "Human aspects of office filing: implications for the electronic office", *Proceedings of the Human Factors Society*, 26th Annual Meeting, pp.59-63, 1982.

[10] Hirschheim, R.A., "Understanding the office: a social-analytic perspective", *ACM Transactions on Office Information Systems*, Vol.4, No.4, pp.331-344, 1986.

[11] Panko, Raymond R., "38 offices: analyzing needs in individual offices", *ACM Transactions on Office Information Systems*, Vol.2, No.3, pp.226-234, 1984.

[12] Conrath, D. et al., "The electronic office and organizational behaviour: measuring office activities", *Computer Networks*, Vol.5, No.6, pp.401-410, 1981.

[13] Kurke, L.; Aldrich, H., "Mintzberg was right! : a replication and extension of the nature of managerial work", *Management Science*, Vol.29, No.8, pp.975-984, 1983.

[14] Higgins, C.; Safayeni, F., "A critical appraisal of task taxonomies as a tool for studying office activities", *ACM Transactions on Office Information Systems*, Vol.2, No.4, pp.331-339,1984.

[15] Burrell, G.;Morgan, G., "Sociological Paradigms and Organizational Analysis", Heineman, London, 1979.

[16] Kwasnik, Barbara H., "The importance of factors that are not document attributes in the organization of personal documents", *Journal of Documentation*, Vol.47, No.4, pp.389-398, 1991.

[17] 森岡清美, 塩原勉, 本間康平,「新社会学辞典」, 東京, 有斐閣, 1993, 1726p.

[18] Case, D.O., "Collection and organization of written information by social scientists and humanists: a review and exploratory study", *Journal of Information Science*, Vol.12, No.1, pp.97-104, 1986.

学術情報センター電子図書館サービス　　© 学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

多次元データに対する項目反応モデル

A study on IRT models for multidimensional data

学術情報センター　孫　媛
Yuan SUN
National Center for Science Information Systems

要旨

テスト理論の分野で近年重要性が増している項目反応理論(IRT)は、テストデータの一次元性を仮定している。しかし、現実のテストでは、項目に正答するのに複数の能力が関与していると考えざるを得ない場合も多い。そのような多次元データに対して、被験者を多次元潜在空間における能力ベクトルとして位置づける多次元項目反応理論(MIRT)が提案されている。MIRTモデルは、特に認知診断テストへの適用が有望視されているが、いくつかの問題も残されている。一方、個人の能力特性値の変動を導入することによって、一次元IRTを拡張したものとして、一般化項目反応理論(GIRT)がある。GIRTモデルでは、攪乱能力次元の影響を、能力特性値の変動の個人差によってある程度説明できるので、多次元データへの適合が改善される。本研究では、MIRTモデルとGIRTモデルの解析的観点からの比較、両モデルの特徴・適用場面の検討が行われる。

ABSTRACT

Determining a correct response to many test items frequently requires more than one ability. Multidimensional IRT models which define *m*-dimensional latent space and locate persons as ability vectors are very useful to deal with multidimensional response data, especially in the case of cognitive diagostic testing. However, sometimes it is hard or even impossible to define and explain some dimensions as characteristics of items because dimensionality is an interaction of persons with items. GIRT model is a modified unidimensional IRT model which takes out one target ability dimension and consider the rest nuisance ability dimensions as person's ability distribution. In this paper a compensatory MIRT model and GIRT model are discussed, the mapping of the statistics in the two models is derived from an analytical aspect.

[キーワード]　多次元潜在空間、補償型多次元IRTモデル、一般化IRTモデル、項目識別力、項目困難度
[Keywords]　multidimensional latent space, compensatory MIRT, GIRT model, item discrimination, item difficulty

1　はじめに

ここ10数年の間、心理・教育測定学の領域において、項目反応理論(item response theory:IRT)に関する研究が中心を占めている。それは、この分野におけるさまざまな問題を解決する上で、項目反応理論が非常に役立つからである(Lord,1980 [13])。しかし、IRTの数理モデルはいくつかの非常に強い仮説に基づいている。なかでも、一つの能力または特性を想定するだけで、被験者のテストの出来具合やテスト項目間の関係を説明できるという一次元性(unidimensionality)の仮定は、重要である。現実のテストにおいて、厳密に一次元性が満たされるのは難しいとしても、測定の目的となる特性の影響が支配的であるならば、一次元モデルは頑健であると考えられる。たとえば、語彙理解力テストや計算力テストにおいては、様々な種類の単語や計算問題に共通な一次元を測定することに成功していると思われる。しかし、正答するために複数の能力が関与すると考えざるを得ない場合もある。たとえば、計算能力テストの中に文章題が含まれており、項目に正答するのに、計算能力だけでなく文章理解能力も必要とされる場合には、一つの特性だけで反応パタンを説明することはできないだろう。また、被験者の認知的構造を調べることを目的とする認知診断テストは、複数の認知構成要素(cognitive components)が解答に影響することを前提としており、通常の一次元モデルを適用することは不適切である。IRTモデルがデータに適合しない

学術情報センター電子図書館サービス　　© 学術情報センター

場合には、一次元性の仮定が満たされないことが主な原因であると指摘する研究も多い(Loyd & Hoover, 1980; Hutten, 1980; etc)。

2 多次元 IRT モデル

IRT の長所を生かし、なおかつ一次元性の満たされないデータにも適合するモデルを提供する理論として、多次元項目反応理論(multidimensional item response theory:MIRT)の研究が始められている。まず、従来の一次元 IRT の概念的枠組みを使い、以下のような MIRT の仮定を指摘することができる。

(1)局所独立性(local independence)

局所独立性とは、被験者の能力レベルを固定するとテスト項目に対する被験者の反応が統計的に独立であるという仮定である。局所独立性の仮定は、θ の条件付共分散がすべての項目間で 0、すなわち $Cov(u_j, u_{j'} | \Theta = \vec{\theta}) = 0$ であることを要求する。ただし、u_j は項目 j に対する反応を表す確率変数である。それに対して、Stout(1987,1990)は、よりゆるやかな局所独立の定義として、本質的局所独立(essential independence)の概念を提案している。これは、項目数 n が∞に近づくとき、項目間のすべての組み合わせの共分散 $Cov(u_j, u_{j'} | \vec{\theta})$ の平均が 0 になるとするものである。

$$\frac{\sum_{1 \le j, j' \le n} Cov(u_j, u_{j'} | \vec{\theta})}{\binom{n}{2}} \to 0 \quad as \quad n \to \infty$$

(2)次元性(dimensionality)

本質的局所独立の仮定を満たすのに必要な潜在能力の次元として、テストの本質的次元性(essential dimensionality)を定義することができる。テストの出来が m 個の能力(特性)によって決まるとき、この m 個の特性が m 次元潜在空間(m-dimensional latent space)を定義する。これは、m 個それぞれの能力次元において、被験者母集団の分散が 0 でないと表現することもできる(全被験者がすべての能力次元上で異なることを仮定するものではない)。m 次元を Cartesian 座標軸で表すと、被験者 i は、能力得点のベクトル $\vec{\theta}_i = [\theta_{i1}, \theta_{i2}, \cdots, \theta_{im}]$ として、能力空間に位置づけられる。同じ能力ベクトルをもつ被験者が項目に正答する確率は等しい。

(3)単調増加性(monotonicity)

どれか一つの能力あるいはいくつかの能力の合成得点が増加すれば、項目に正答する確率も単調に増加する、というのが単調増加性の仮定である。このことから、項目特性関数が m 次元の滑らかな項目反応曲面(item response surface:IRS)になることが示唆される。ある項目に対して同じ正答確率を持つ被験者どうしは、等確率等高線(equiprobability contour)上に位置づけられる。

多次元モデルの中でもっともよく使われるのが 2 パラメータ・ロジスティック補償型モデル(two-parameter logistic compensatory model: M2PL-C model)である(Hattie, 1981 ; McKinley & Reckase, 1983 ; Reckase, 1985 ; Reckase & McKinley, 1991)。このモデルにおける項目反応関数(item response function: IRF)は、以下のように定義される。

$$P(u_{ij} = 1 | \theta_{ik}, a_{jk}, d_j) = \frac{1}{1 + \exp(-1.7(\sum_{k=1}^{m} a_{jk}\theta_{ik} + d_j))}$$

ただし、

u_{ij} は、被験者 i の項目 j への応答。
　　正答ならば u_{ij}=1、誤答ならば u_{ij}=0
θ_{ik} は、被験者 i の次元 k 上での能力特性値
a_{jk} は、項目 j の次元 k 上での識別力パラメータ
$d_j = -\sum_{k=1}^{m} a_{jk} b_{jk}$ は、項目 j の困難度に関連する
　　スカラーパラメータ

である。

合成得点に対する θ_k $(k = 1, 2, \cdots, m)$ の寄与の程度は、a_{jk} によって変わってくる。たとえば、次元数が 2(m=2)のときで考えてみると、a_{j1}=a_{j2} ならば、2 つの次元が同じ程度よく測定されており、a_{j1}=0 ならば、項目 j の識別力が θ_1 次元では働いていないことになる。したがって、項目の各次元上においては、識別力 a_{jk} によって被験者の能力 θ_{ik} の重みづけられ方が違ってくる。すべての項目について a_{jk} が等しいならば、そのテストは完全に一次元となる。

Reckase ら(Reckase & McKinley,1991[21])は、多次元能力空間においては項目jの識別力ベクトルのノルムまたは長さを使って、最大識別力値(maximum amount of discrimination: $MDISC_j$)を定義した。

$$MDISC_j = \sqrt{\sum_{k=1}^{m} a_{jk}^2}$$

$MDISC$ が一次元 IRT モデルの識別力に対応することは明らかである。次元数 m=1 の場合、$MDISC_j$=a_j となる。θ_k 軸に対する項目 j の最大識別力値の点の方向余弦は以下のように定義される。

$$cos\alpha_{jk} = \frac{a_{jk}}{\sqrt{\sum_{k=1}^{m} a_{jk}^2}}$$

学術情報センター電子図書館サービス　　© 学術情報センター

原点から項目反応曲面(IRS)上の最大識別力値の点までの符号付き距離が一次元IRTモデルの困難度b_jに対応し、以下のように表される。

$$D_j = \frac{-d_j}{\sqrt{\sum_{k=1}^{m} a_{jk}^2}}$$

このことから、多次元項目困難度(multidimensional item difficulty: MID)の指標として、原点から勾配の最も急な点までの距離D_jとその点の方向余弦$cos\alpha_{jk}$または角度α_{jk}などの統計量を用いることが提案されている(Reckase,1985)。

つぎに、2次元(m=2)でa_{j1}=a_{j2}=1, d_j=0という特別な場合を想定すると、項目jのIRFは

$$P(u_{ij}=1|\theta_{i1},\theta_{i2}) = \frac{1}{1+\exp(-1.7(\theta_{i1}+\theta_{i2}))}$$

となる。この式より、$\theta_1=\theta_2=0$のとき、$P(u_{ij}=1|\theta_1, \theta_2)=0.5$となることがわかる。また、$\theta_1$がごく低い能力値、たとえば$\theta_1=-2.0$で、$\theta_2$がごく高い能力値+2.0であるときにも、同じ正答確率0.5が得られることも示される。つまり、補償型多次元IRTモデルでは、ある次元上での高い能力値が他の次元上での低い能力値を代償する。極端に言えば、非常に高い能力値の次元がひとつでもあれば、他の次元の能力値が低くても、項目への正答確率は高くなりうる。認知診断に項目反応理論を適用する場合、一次元のモデルよりも多次元モデルの方がふさわしいと考えられるが、複数の認知構成要素が同時に働いて正答が導かれるとするならば、補償型モデルには不適切な点が残る。

多次元IRTでは、以下のような2パラメータ・ロジスティック非補償型モデル(two-parameter logistic non-compensatory model)も提案されている(Sympson, 1978 [28])。

$$P(u_{ij}=1|\theta_{ik},a_{jk},b_{jk}) = \frac{1}{\prod_{k=1}^{m}(1+\exp(-1.7a_{jk}(\theta_{ik}-b_{jk})))}$$

b_{jk}は、項目jの次元k上での困難度を表す。非補償型モデルでは、項目への正答確率がそれぞれの次元のIRFの積となっている。したがって、項目に正答するのに必要な認知的構成要素のどれか一つの特性値が低ければ、正答確率は低くなる。つまり、項目への反応に対する認知的構成要素の関係が補償的ではない。その点では、補償型モデルよりも認知診断テストに符合すると考えられるが、パラメータの推定問題はずっと複雑になる。

3 個人パラメータの変動を考慮に入れた一次元IRTモデル

多次元IRTモデルでは、能力特性値は被験者固有の定数とみなされており、この点では、通常の一次元IRTモデルと同じである。これに対し、被験者の能力特性値の確率的な分布を導入し、項目の難しさと被験者能力の特性値の双方が心理連続体上で正規分布に従うと仮定したTorgerson(1958[29])の一般正規累積モデルがある。項目jと被験者iの分布がそれぞれ$N(b_j,a_j^{-2})$と$N(\beta_i,\alpha_i^{-2})$であるとすると、被験者の項目に対する正答確率は、

$$P_{ij} = \Phi(h_{ij})$$

と表わされる。ただし、$h_{ij} = \frac{a_j\alpha_i}{\sqrt{a_j^2+\alpha_i^2}}(\beta_i - b_j)$で、$\Phi$は標準正規分布関数である。Torgersonのモデルでは正規分布関数を用いているが、これより取り扱いの容易なロジスティック関数を用いたのが、一般項目反応モデル(generalized item response model:GIRT)である(孫,1989[25];孫・芝,1990[26])。GIRTモデルの式は次の通りである。

$$P_{ij}(\alpha,\beta,a,b) = \frac{1}{1+\exp(-1.7\frac{\alpha_i a_j}{\sqrt{\alpha_i^2+a_j^2}}(\beta_i-b_j))}$$

αが大きい被験者については、能力分布の標準偏差が小さくなり、その反応パタンは従来の項目反応モデルにおける反応パタンと近くなると考えられる。$\alpha_i \to \infty$という極端なケースでは、被験者の能力特性値の分散は0になる。このとき、

$$\lim_{\alpha_i\to\infty,\beta_i\to\theta_i} P_{ij} = \frac{1}{1+\exp(-1.7a_j(\theta_i-b_j))}$$

となり、一般項目反応モデルは従来の項目反応モデルと一致することが示される。

3.1 個人反応曲線

GIRTモデルにおいて、$a_j \to \infty$の場合も想定することができる。このとき、項目反応曲線ICCは階段状

学術情報センター電子図書館サービス　　© 学術情報センター

$$\lim_{a_j \to \infty} P_{ij} = \frac{1}{1+\exp(-1.7\alpha_i(\beta_i - b_j))}$$

になるが、正答確率 P を b_j の関数と見なすと、これはいわゆる個人反応曲線(Person Response Curve : PRC)(Lumsden, 1977 ; Trabin & Weiss, 1983)と一致する。個人反応曲線は、被験者の反応に関する信頼性および安定性を調べるために提案されたものであるが、GIRT モデルの特殊な場合として扱うことができる。

3.2 GIRT モデルと多次元性

通常の一次元 IRT モデルのもとでは起こりにくい反応パタンを示す被験者がいれば、その能力分布の分散が大きくなり、α は小さな値になるであろう。孫([25])、孫・芝([26])は、IRT に基づく日本語語彙理解力尺度を利用して、外国人留学生の語彙理解力を測定し、従来の一次元IRTモデルの当てはまりがよくないこと、GIRT を適用することでデータとモデルの適合が大幅に改善されることを示している。つまり、母語として日本語を習得した人が被験者であるときには一次元性の仮定を満たす語彙理解力尺度が、外国人にとっては多次元尺度になっていると考えられるのである(すでに母語を獲得した後に、外国語として日本語語彙を習得するときには、学習開始以前に形成されている関心、個人的経験、被験者の母語や自国の文化などの影響によって、母語として日本語を習得する場合と比べて、日本語語彙の習得過程が異なるのであろう)。GIRT モデルは、測定の目標(target ability)はあくまでも一次元にとどめ、そのほかの攪乱能力(nuisance ability)次元を、個人パラメータ値 (β,α) として表わしている、変形された一次元 IRT モデルであると解釈することができる。$\frac{1}{\alpha^2} \to 0$、すなわち $\alpha \to \infty$ のとき、被験者の反応データは一次元となる。一方、$\frac{1}{\alpha^2} \to \infty$ すなわち $\alpha \to 0$ のとき、データが多次元となる。ただし、次元数をパラメータ α から示すことはできない。

4 多次元データに対するモデルの比較

4.1 MIRT モデルと GIRT モデル

4.1.1 GIRT モデルの変形

GIRT モデルでは、尺度そのものは一次元であり、解答パフォーマンスに関与するその他の要因の影響は、ランダムであることが前提とされる。その意味で、GIRT モデルは変形された一次元 IRT モデルであると考えられる。日本語語彙理解力尺度を例にとろう。テストは日本語を母語として育った日本人にとっては、語彙理解力を測る一次元性の高い尺度を構成している。しかし、他の言語を母語として育った外国人がテストを受けるときには、被験者それぞれの文化的・個人的経験などの要因がテストのパフォーマンスに影響する。漢字圏の人にとっては漢語項目が、留学生にとっては自分の専門分野に関する単語が、容易であったりする。しかし、項目と被験者の間の相互作用は系統的なものではないと解釈され、パラメータ α は被験者に固有な性質ととらえられる。

次に、能力尺度(テスト)が被験者との相互作用で多次元性を持つ場合を考えてみよう。被験者が正答するために、ある項目に対しては多くの次元が影響し、別の項目に対しては一次元だけで足りるかもしれない。このような場合には、項目によって、被験者の能力分布の散らばりが違うといえる。このとき、GIRT モデル中の個人パラメータ α_i を、個人 i と項目 j との相互作用によるものとして、α_{ij} に置き換えることにする。これによって、GIRT モデルは以下のように変形され、MIRT モデルとの比較が可能になる。

$$P_{ij}(\alpha,\beta,a,b) = \frac{1}{1+\exp(-1.7\frac{\alpha_{i,j2\cdots jm}a_{j1}}{\sqrt{\alpha^2_{i,j2\cdots jm}+a^2_{j1}}}(\beta_i - b_j))}$$

4.1.2 MIRT-C と GIRT 間の解析的分析

MIRT では 2 次元目、3 次元目、m 次元目というように項目の特性としてすべての変動を取り入れるのに対して、多次元上の GIRT では一次元以上の変動をすべて項目と被験者と相互作用の変動で説明する。MIRT における多次元が互いに独立して、正規分布に従うならば、分散の可加法性から以下のように書くことができる。

$$\frac{1}{\alpha^2_{i,j2\cdots jm}} = \frac{1}{a^2_{j2}} + \frac{1}{a^2_{j3}} + \cdots + \frac{1}{a^2_{jm}}$$

ここで、ある特定の項目について考えると、上式は以下のように整理できる。すなわち、

学術情報センター電子図書館サービス　　© 学術情報センター

$$\frac{1}{\alpha_{i,2\cdots m}^2}=\frac{1}{a_2^2}+\frac{1}{a_3^2}+\cdots+\frac{1}{a_m^2}$$
$$=\frac{\sum_{l=2}^{m}\prod_{k=2}^{m}\frac{a_k^2}{a_l^2}}{\prod_{l=2}^{m}a_l^2}$$
$$=\frac{\sum_{l=2}^{m}\prod_{k=2}^{\prime m}a_k^2}{\prod_{l=2}^{m}a_l^2}$$

ただし、$\prod'$は、第l番目の要素を除いて乗ずることを意味する。

GIRTモデルの項目反応関数IRFの一部を以下のように変形する。

$$exp(\frac{a_1\alpha_{i,2\cdots m}}{\sqrt{a_1^2+\alpha_{i,2\cdots m}^2}}\beta_i)=exp(\frac{a_1\sqrt{\frac{\prod_{l=2}^{m}a_l^2}{\sum_{l=2}^{m}\prod_{k=2}^{\prime m}a_k^2}}}{\sqrt{a_1^2+\frac{\prod_{l=2}^{m}a_l^2}{\sum_{l=2}^{m}\prod_{k=2}^{\prime m}a_k^2}}}\beta_i)$$
$$=exp(\frac{\prod_{l=1}^{m}a_l}{\sqrt{a_1^2\sum_{l=2}^{m}\prod_{k=2}^{\prime m}a_k^2+\prod_{l=2}^{m}a_l^2}}\beta_i)$$
$$=exp(\frac{\prod_{l=1}^{m}a_l}{\sqrt{\sum_{l=1}^{m}\prod_{k=1}^{\prime m}a_k^2}}\beta_i)$$

m=2の場合、$\frac{1}{\alpha_{i,2}^2}=\frac{1}{a_2^2}$となり、上式は

$$exp(\frac{a_1\alpha_{i,2}}{\sqrt{a_1^2+\alpha_{i,2}^2}}\beta_i)=exp(\frac{a_1a_2}{\sqrt{a_1^2+a_2^2}}\beta_i)$$

となる。

次に、補償型多次元モデルの対応した部分は以下のようになる。

$$exp(\sum_{l=1}^{m}a_1\theta_l)=exp(\frac{\prod_{l=1}^{m}a_l}{\sqrt{\sum_{l=1}^{m}\prod_{k=1}^{\prime m}a_k^2}}(\frac{\sqrt{\sum_{l=1}^{m}\prod_{k=1}^{\prime m}a_k^2}}{\prod_{l=1,l\neq1}^{m}a_l}\theta_1+$$
$$\cdots+\frac{\sqrt{\sum_{l=1}^{m}\prod_{k=1}^{\prime m}a_k^2}}{\prod_{l=1,l\neq m}^{m}a_1}\theta_m))$$

ここで、θ_1の係数は以下のとおり。

$$\sqrt{\frac{\sum_{l=1}^{m}\prod_{k=1}^{\prime m}a_k^2}{\prod_{k=1,k\neq1}^{m}a_k^2}}=\sqrt{1+\frac{a_1^2}{a_2^2}+\cdots+\frac{a_1^2}{a_m^2}}$$
$$=a_1\sqrt{\frac{1}{a_1^2}+\frac{1}{a_2^2}+\cdots+\frac{1}{a_m^2}}$$
$$=a_1\sqrt{\sum_{l=1}^{m}\frac{1}{a_l^2}}=\frac{a_1}{w}$$

ただし、$\frac{1}{w^2}=\sum_{l=1}^{m}\frac{1}{a_l^2}=\frac{1}{a_1^2}+\frac{1}{\alpha^2}$は総変動(総分散の逆数)を表す。

同様に、θ_mの係数は以下のとおり。

$$\sqrt{\frac{\sum_{l=1}^{m}\prod_{k=1}^{\prime m}a_k^2}{\prod_{k=1,k\neq m}^{m}a_k^2}}=\sqrt{1+\frac{a_m^2}{a_2^2}+\cdots+\frac{a_m^2}{a_m^2}}$$
$$=a_m\sqrt{\frac{1}{a_1^2}+\frac{1}{a_2^2}+\cdots+\frac{1}{a_m^2}}$$
$$=a_m\sqrt{\sum_{l=1}^{m}\frac{1}{a_l^2}}=\frac{a_m}{w}$$

したがって、

$$\beta_i=\frac{a_1}{w}\theta_1+\cdots+\frac{a_m}{w}\theta_m=\frac{1}{w}\sum_{k=1}^{m}a_k\theta_k\equiv\sum_{k=1}^{m}\rho_k\theta_k$$

となり、βは、m次元能力$\vec{\theta}$空間のベクトルとみなすこともできる。

m=2の場合、

$$\beta=\frac{a_1}{\sqrt{\frac{a_1^2+a_2^2}{a_1^2a_2^2}}}\theta_1+\frac{a_2}{\sqrt{\frac{a_1^2+a_2^2}{a_1^2a_2^2}}}\theta_2$$
$$=\frac{a_1^2a_2}{\sqrt{a_1^2+a_2^2}}\theta_1+\frac{a_1a_2^2}{\sqrt{a_1^2+a_2^2}}\theta_2$$
$$=a_1a_2[cos\alpha_1\theta_1+cos\alpha_2\theta_2]$$

となる。

4.1.3 MIRTモデルとGIRTモデルの特徴および応用場面

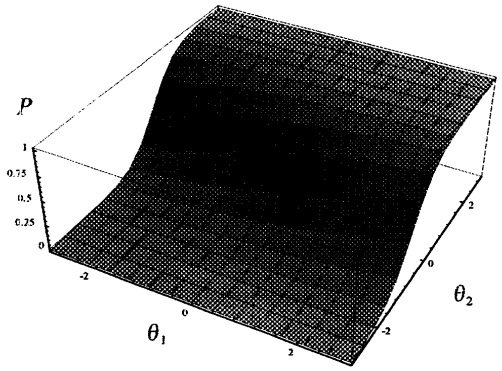
2次元の場合　MIRTとGIRTの項目特性関数と項目特性曲線は以下のように簡略化することができる。すなわち、

学術情報センター電子図書館サービス　　© 学術情報センター

$$MIRT\text{-}C\text{: } P_{ij} = \frac{1}{1+\exp(-1.7(a_{j1}(\theta_{i1}-b_{j1})+a_{j2}(\theta_{i2}-b_{j2})))}$$

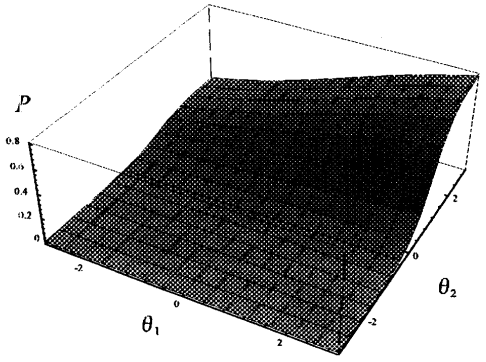
$$MIRT\text{-}N\text{: } P_{ij} = \frac{1}{(1+\exp(-1.7a_{j1}(\theta_{i1}-b_{j1})))} \times \frac{1}{(1+\exp(-1.7a_{j2}(\theta_{i2}-b_{j2})))}$$

$$GIRT\text{: } P_{ij} = \frac{1}{1+\exp(-1.7\frac{\alpha_{i,j2}a_{j1}}{\sqrt{\alpha_{i,j2}^2+a_{j1}^2}}(\beta_i-b_j))} = \frac{1}{1+\exp(-1.7\frac{a_{j1}a_{j2}}{\sqrt{a_{j1}^2+a_{j2}^2}}(\beta_i-b_j))}$$



a1 = 0.3
a2 = 1.4
d = 0.2

図 1　MIRT-C モデルの ICC



a1 = 0.3
a2 = 1.4

b1 = - 0.2
b2 = 0.9

図 2　MIRT-N モデルの ICC

図 1 と図 2 を比較すると、MIRT-C モデルにおいてはある次元上での高い能力値が他の次元上での低い能力値を代償できること、MIRT-N モデルにおいては項目への正答確率がそれぞれの次元の IRF の積となっているため、項目に正答するために必要な認知的構成要素のどれか一つの特性値が低ければ、正答確率は低くなることがわかる。ここで、両極端のケースを考えてみることにする。

(1)次元 2 における項目の識別力 $a_{j2}\to 0$ の場合

MIRT-C モデルは以下のようになり、2 次元目の働きが完全になくなる。

$$\lim_{a_{j2}\to 0} P_{ij} = \frac{1}{1+\exp(-1.7a_{j1}(\theta_{i1}-b_{j1}))}$$

それに対して、MIRT-N モデルでは 2 次元目の IRF が 0.5 に近づき、全体の IRF が

$$\lim_{a_{j2}\to 0} P_{ij} = \frac{1}{1+\exp(-1.7a_{j1}(\theta_{i1}-b_{j1}))} \times 0.5$$

となる。

また、GIRT モデルにおいては、$\frac{1}{\alpha_{i,j2}^2} = \frac{1}{a_{j2}^2}$ により、$\lim_{\alpha\to 0} P_{ij} = 0.5$ となる。つまり、2 次元目においては項目の分散が極端に大きいときには、GIRT モデルではその項目に正答する確率が 0.5 に近くなること、MIRT モデルは次元間が独立であるため、1 次元目の IRF が a_{j2} に影響されないことが示される。

(2)次元 2 における項目の識別力 $a_{j2}\to\infty$ の場合

MIRT-C モデルでは、項目正答確率が第 2 次元の能力に左右されることになり、モデルは以下のようになる。

$$\lim_{a_{j2}\to\infty} P_{ij} = \begin{cases} 1, \theta_{i2} > b_{j2} \text{のとき} \\ 0, \theta_{i2} < b_{j2} \text{のとき} \end{cases}$$

MIRT-N モデルは、以下のようになる。すなわち、

$$\lim_{a_{j2}\to\infty} P_{ij} = \begin{cases} \frac{1}{1+\exp(-1.7a_{j1}(\theta_{i1}-b_{j1}))}, \theta_{i2} > b_{j2} \text{のとき} \\ 0, \theta_{i2} < b_{j2} \text{のとき} \end{cases}$$

つまり、第 2 次元目においては、被験者の能力特性値が項目の困難度よりも高いときには、MIRT-N モデルは通常の一次元 IRT と一致し、能力特性値が困難度よりは低いときには、その項目に正答する確率が 0 に近くなる。

それらに対して、GIRT モデルでは、$\frac{1}{\alpha_{i,j2}^2} = \frac{1}{a_{j2}^2}$ により、IRF が以下のようになる。すなわち、

$$\lim_{\alpha\to\infty} P_{ij} = \frac{1}{1+\exp(-1.7a_j(\theta_i-b_j))}$$

$\alpha\to\infty$、つまり被験者の能力特性値の分散が 0 になるときには、GIRT モデルが従来の一次元 IRT モデルと

学術情報センター電子図書館サービス　　◎学術情報センター

一致することが示される。

項目間および項目と被験者との相互作用　次元がはっきり分けられないときには、MIRT モデルは適切ではないと思われる。MIRT-C モデル、MIRT-N モデルのいずれにおいても、多次元上のパラメータは互いに独立であり、次元間の相互作用やテストと被験者の相互作用などがモデルに反映されない。つまり、被験者に依存しない、互いに独立な項目パラメータによって能力空間が形成される。被験者はその能力空間に能力推定値として位置づけられるにすぎない。しかし、同じ次元に関して、被験者にとって項目パラメータが同じにならない場合もあるだろう。被験者によって項目またはテストの次元数さえも違ってくる可能性がある。変形した GIRT モデルでは、そのような相互作用がパラメータとして取り込まれている。第 1 次元が目標能力を測るもので、その次元における項目パラメータだけが項目固有の特性と考えられ、それ以外の次元においては項目間と項目・被験者の相互作用がパラメータ $\alpha_{i,j2\cdots jm}$ に含まれる。

モデルの応用　前に述べたように MIRT-C モデルには不適切な点があるにも関わらず、今の段階では、MIRT-C モデルについての研究と応用が主である。MIRT のパラメータ推定用のプログラムもいくつか開発されている。そのうち、TESTFACT と NOHARM がもっともよく使われている。TESTFACT は Bock-Aitkin(Bock & Aitkin,1981 [6])の EM アルゴリズム法を用いて、周辺最尤推定値(MMLEs)を求めるに対し、NOHARM は harmonic analysis に基づいて最小二乗解を求めるものである。これらのプログラムは、すべて MIRT-C モデルをベースにしたものである。

MIRT モデルは、差異項目機能(Differential Item Functioning: DIF)研究にも用いられている。たとえば、Ackerman などは MIRT の文脈から DIF の起きる原因を説明している(Ackerman, 1988[1]; Ackerman, 1992[4]; Douglas Roussos & Stout,1995[8], etc)

また、MIRT を適用すれば、(理論的には)テストの次元数を決めることができ、それぞれの次元ごとの項目パラメータと能力特性値を測定できる。そのために、多次元を想定した項目プール(item pool, item bank)を蓄えることが可能になり、テスト項目の選択、テストの作成、分析および実施などにあたって有用である。特に、認知診断テストへの適用に役立つと考えられる。

テストと被験者の相互作用が強く、多次元を明確にわけられないときには、GIRT モデルを用いると効果的だろう。たとえば、学力テストでは、多次元能力が項目反応に関与していても、テストで測りたい目標能力は一次元であることが多い。GIRT モデルは、その目標能力の次元(b, a) を取り出し、それ以外の攪乱能力次元をつぶして各被験者の能力分布(β,α)として扱う。MIRT と GIRT の解析的分析で示したように、GIRT の能力分布の平均値βは、MIRT の多次元能力の線形結合になる。GIRT モデルにおいては、各次元での詳しい情報を提供することはできないが、第 2 次元、第 3 次元などを分解しないことが長所として働くこともあると考えられる。当て推量(guessing)やケアレス・ミスの扱いなど、通常の一次元 IRT モデルでは処理しにくかった変動要因は、個人的変動のパラメータに吸収される。また、GIRT の項目反応関数においては、項目と被験者の乗法項があり、両者の相互作用が含まれることを示している。

通常の IRT モデルにおけるθの推定の精度(たとえば、θの標準誤差や事後標準誤差)には、個人ごとの反応パタンのモデルへの適合度が反映されないのに対して、GIRT モデルにおいては、個人パラメータαを導入することによって、間接的にそれが可能になる。パラメータαを被験者の反応パタンの特異性や反応の安定性の指標と解釈することによって、被験者の認知パタン、知識構造の相違、学力検査におけるケアレスミスや当て推量の度合い、あるいは性格検査における反応の一貫性の程度を示す指標として利用することもできる。

4.2　一次元 IRT の適用

多次元項目反応モデルを適用するに際していくつかの問題点がある。そのうち、パラメータの推定が最大の問題である。MIRT の推定すべきパラメータが多いために、精度の高い推定をすることが難しい。項目数が 100 より少ないときには、MIRT モデルによる被験者間の弁別は一次元モデルに及ばないことを示唆する研究もある(eg., Davey & Hirsch, 1990)。次元が増えるにつれて、構造パラメータと付随パラメータの推定が非常に難しくなっていくと言われている(eg., Holland, 1990; McDonald, 1982)。

そこで、多次元データに対して、もう一つ非常に実用的で、また妥当性のある方法は一次元 IRT モデルの適用である。こうした方向では、多次元データに対して一次元 IRT モデルを適用した場合のいろいろな可能性を検討する研究が行なわれている(Drasgow & Parsons, 1983; Ansley & Forsyth, 1985; Luecht & Miller, 1992; Ack-

学術情報センター電子図書館サービス　　©学術情報センター

erman, 1989; Ackerman, 1991, etc)。ここではWangの重みづけ平均法について言及する。

重みづけ平均法　Wang(1986,1988)は、多次元モデルのもとで生成されたデータについて、一次元モデルで能力特性値を推定し、その推定値と多次元の能力特性値との関係を線形代数的方法で分析した。Wangによれば、支配的な次元が存在するかどうかに関係なく、一次元モデルでの能力推定値が多次元の能力特性値の重みづけ平均(reference compositeとよばれる)になるという。具体的には、その重みづけ平均は、以下のように近似できるという。

$$\gamma_1 = \overrightarrow{w_1'\theta}$$

ただし、Aは項目数n×次元数mの識別力行列で、$\vec{w}_1$は行列$A'A$の最も大きい固有値と対応する固有ベクトルである。

5　今後の課題

古典的テストモデルと比較して、項目反応モデルが理論的に数多くの長所を持つのはまちがいない。それにもかかわらず、IRTモデルの利用範囲が限られているのには、いくつかの理由が考えられる。まず、IRTモデルが要求する強い仮定のために、データとモデルの適合、モデルの頑健性の程度が大きな問題になる。実際のテストにおいては、項目に正答するために複数の能力が関与していて、従来のIRTモデルの適用が制限される場合も多い。そこで、一次元性の仮定を満たさないデータにも適合する多次元モデルが、IRT理論の発展的応用の有望な方向として浮かび上がってきている。ここ数年、多数の研究者が精力的に多次元IRTモデルを研究し、いくつかの成果が上がっている。

多次元IRTの適用を妨げる大きな要因の一つは、モデルのパラメータ推定の問題である。パラメータ推定は、一次元のIRTにおいても重要な研究課題であるが、パラメータ数が増えるMIRTでは、精度の高い推定を行うことがいっそう難しい。事前情報を導入し、ベイズ推定法を用いることによって、パラメータの推定値が改善され、項目数・被験者数の少ないデータによるパラメータ推定が可能になるという指摘はあるが、適切な事前分布の選び方、ベイズ的アプローチを採ることの可否など、今後さらに検討を重ねる必要があるだろう。

GIRTモデルでは測定対象となる能力は一次元であるが、能力に確率分布を導入することによって、多次元データへの適合が改善される。モデルのパラメータ数はMIRTモデルよりも少ないが、モデルの適用を検討する上での最大の難関はやはりパラメータの推定問題である(孫・芝, 1990; 孫, 1991)。柴山・繁桝(1994)、前川・中村(1993)は、ベイズ推定法を用いて項目および被験者パラメータの同時推定を試みている。

GIRTは個人内の能力変動をモデル化しているが、文化や地域など集団的な違い・変動が一次元モデルの不適合の原因となることもあるかもしれない。同じテストであっても、集団によって、測定対象となる次元がずれていたり、ある集団にとって一次元であるテストが別の集団にとっては多次元になっている、といった事態も起こりうるかもしれない。そのような場合には、集団差を表すパラメータを導入すれば、通常のIRTより適合度に優れ、しかもGIRTよりパラメータ数が少なく、推定精度も改善されたモデルが可能になる。

参考文献

[1] Ackerman, T. A. (1988). An explanation of differential item functioning from a multidimensional perspective. *A paper presented at the annual meeting of the American Educational Research Association.* New Orleans, LA.

[2] Ackerman, T. A. (1989). Unidimensional IRT calibration of compensatory and noncompensatory multidimensional items. *Applied Psychological Measurement*, 13, 113-127.

[3] Ackerman, T. A. (1991). The use of unidimensional parameter estimates of multidimensional items in adaptive testing. *Applied Psychological Measurement*, 15, 13-24.

[4] Ackerman, T. A. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement*, 29, 67-91.

[5] Ansley, T. N. & Forsyth, R. A. (1985). An examination of the characteristics of unidimensional IRT parameter estimates derived from two-dimensional data. *Applied Psychological Measurement*, 9, 37-48.

[6] Bock, R. F., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 46, 443-459.

[7] Davey, T., & Hirsch, T. M. (1990). Examinee dis-

学術情報センター電子図書館サービス　　◎学術情報センター

crimination as a measure of test data dimensionality. Paper presented at the annual meeting of the Psychometric Society, Princeton NJ.

[8] Douglas, J., Roussos, L. & Stout, W. (1995). Refinement of the Shealy-Stout multidimensional IRT Model of DIF with practical implications.

[9] Drasgow, F. & Parsons, C.K.(1983). Application of Unidimensional Item Response Theory Models to Multidimensional Data. *Applied Psychological Measurement*, 7, 189-199.

[10] Hattie, J. (1981). Decision criteria for determining unidimensionality. Unpublished doctoral dissertation, University of Toronto, Canada.

[11] Holland, P. W. (1990). The Dutch identity: A newtool for the study of item response models. *Psychometrika*, 55, 5-18.

[12] Hutten, L. R. (1980). Some empirical evidence for latent trait model selection. Paper Presented at the annual meeting of the American Educational Research Association, Boston.

[13] Lord, F. M. (1980) Application of item response theory to practical testing problems. New York: Lawrence Erlbaum Associates.

[14] Loyd, B. H. (1980). Vertical equating using the Rasch model. *Journal of Educational Measurement*, 17, 178-193.

[15] Lumsden, J. (1977). Person reliability. *Applied Psychological Measurement*, 4, 477-482.

[16] Luecht, R. M. & Miller, T. R. (1992). Unidimensional Calibrations and Interpretation of Composite Traits for Multidimensional Tests. *Applied Psychological Measurement*, 16, 279-293.

[17] 前川眞一・中村知靖,「一般項目反応モデルにおける個人パラメタの周辺最尤推定法」, 教育心理学研究, Vol.41, No.1, pp.22-30, 1993.

[18] McDonald,R.P.(1982). Linear versus nonlinear models in item response theory. *Applied Psychological Measurement* ,6,379-396.

[19] McKinley, R. L. & Rechase, M. D. (1983). An extension of the two-parameter logistic model to the multidimensional latent space (Research Report ONR83-2). Iowa City, IA: The American College Testing Program.

[20] Rechase, M. D. (1985). The difficulty of test items that measure more than one dimension. *Applied Psychological Measurement*, 9, 417-412.

[21] Reckase, M. D., & McKinley, R. L. (1991). The discriminating power of items that measure more than one dimension. *Applied Psychological Measurement*, 15, 361-373.

[22] 柴山直・繁桝算男,「一般項目反応モデルにおけるパラメタの推定方法」, 行動計量学, Vol.21, No.1, pp. 57-65,1994

[23] Stout, W. F. (1987). A nonparametric approach for assessing latent trait unidimensionality. *Psychometrika*, 52, 589-617.

[24] Stout, W. F. (1990). A new item response theory modeling approach with applications to unidimensionality assessment and ability estimation. *Psychometrika*, 55, 293-325.

[25] 孫媛,「特異な反応パタンを示す被験者の能力推定」, 東京大学大学院教育学研究科昭和 63 年度修士論文（未公刊）, 1989.

[26] 孫媛・芝祐順,「特異な反応パタンを示す被験者の能力推定-一般項目反応理論の適用」,教育心理学研究, Vol.38, No.4, pp.360-368, 1990.

[27] 孫媛,「GIRT の個人パラメタ推定に関するシミュレーション研究」,日本教育心理学会第 33 回総会発表論文集, p.536, 1991.

[28] Sympson, J. B. (1978). A model for testing with multidimensional items. In D.J.Weiss(Ed.) Proceedings of the 1977 Computerized Adaptive Testing Conference(pp.82-98). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program.

[29] Torgerson, W. S. (1958) Theory and Methods of Scaling.

[30] Trabin, T. E. & Weiss, D. J. (1983). The Person response curve: fit of individuals to item response theory models. In Eeiss, D. J. (ed.) New Horizons in Testing. Academic Press.

[31] Wang, M. M. (1986). Fitting a unidimensional model to multidimensional item response data(ONR Rep.042286). Iowa City IA: University of Iowa.

[32] Wang, M. M. (1988). Measurement bias in the application of a unidimensional model to multidimensional item-response data. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

学術情報センター電子図書館サービス　　©学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

ヤクート語語彙研究(2)：方向表現
---サハ族英雄叙事詩オロンホを資料として---

A study on lexcal items in Yakut language (2): Terms of direction in Saxa epic poem *olonxo*.

神戸市看護大学　藤代　節
Setsu FUJISHIRO
Kobe City College of Nursing

要旨

東・中央シベリアを中心として、広い地域に分布するチュルク系言語ヤクート語の方向に関連する語彙についての研究である。ヤクート語にはその主要な話者であるサハ(or ヤクート)族が今日の分布域に至るまでに辿ってきた過程が、特に語彙において反映されており、そのことがチュルク諸語の中にあって、ヤクート語が特異な言語とされる所以でもある。この言語の特異性を明らかにするため語彙の中でもセットとして扱えるものを取り上げる。既に色彩名称に関する研究を発表したが、本稿ではこれに続いて同じくセットと出来る語彙として方向表現[1]を選び、研究の対象とした。サハ族の英雄叙事詩オロンホを主な資料とし、ヤクート語の方向表現の整理分析をこころみた。

ABSTRACT

This paper is concerned with lexical items in Yakut language, which belongs to the Turkic language family. The main part of the Yakut language community is consited with Yakut or Saxa people. Their language traces well the process of their immigration to Siberia. In this paper, the author presents the data of the set of lexical items "directions" from the olonxo "Kïïs Däbilijä", one of the most famous epics of the Saxa people and suggests that such approach to the study of lexicon is effective to uncover the complicated situation of languages of Siberia.

This paper is the second serial of the study on the lexicon of Yakut language, which we published in the preceeding number of the bulletin.

[キーワード]　ヤクート語、オロンホ、シベリア、方向表現、チュルク系諸言語
[Keywords]　Yakut language, olonxo, Siberia, Terms of directions, Turkic Languages

0　はじめに
1　資料について
2　ヤクート語における方向表現
3　方向表現分析
4　おわりに
5　データ
注及び参考文献[2]

0　はじめに

シベリアの東部を中心に広範囲に分布するヤクート語は36万人程度の話者人口をもつチュルク系の言語である。ヤクート語はチュルク系言語としての典型的な特徴を持ちながらも、他のチュルク諸語と少なからず異なりがみられ、他の言語とグループを成さず特異な位置にある。特に語彙において相違がみられるが、これは現在の分布域に至るまでに移動を繰り返してきたという点に要因がもとめられる。

ヤクート語の主要な分布域である旧ヤクート自治共

学術情報センター電子図書館サービス　　©学術情報センター

和国はソ連邦が崩壊して後、ロシア共和国になってからも、その版図にある。現在はサハ(ヤクート)共和国 rus. Respublika Saxa (Jakutija) と称され、サハ族(ヤクート族)[3]の民族性を強調した地域確立への志向が強い。そういった社会的風潮の中で、民族のアイデンティティとして言語を尊重する気運がこれまでにも増してサハ族の中では高まってきている。これをうけて他の旧ソ連邦内の「民族言語」と同様にその資料的出版物が増加の傾向にある。とはいえ、ヤクート語の言語資料としては、ごく断片的なものを含めても17世紀末に記録されたものより遡る資料は少なくとも現在のところ知られていない。そのような状況の中で、サハ族の成立を重要なモチーフの一つとする口承文芸であるオロンホ yak. oloŋxo と呼ばれる英雄叙事詩の言語資料としての価値は高い。

本稿ではサハ族の英雄叙事詩オロンホ《クース・デビリィエ》yak. Кыыс Дэбилийэ を主な資料としてヤクート語の語彙研究の端緒をつけることを目的とする。本稿で対象とした語彙は方向表現である。この英雄叙事詩において見られる方向・方角を表す語彙をグループにまとめ、データを提出する。現代標準ヤクート語でのこれら語彙の形式と意味について考慮しながら、ヤクート語の周囲に位置するシベリアの諸民族言語との類似点にも触れる。

シベリアの諸言語における方向表現が重要である点について少しふれておきたい。現在、シベリアに分布する言語にはロシア語の他にこの地域に先住の人々の言語がある。これらの言語は法律用語や行政用語等としての社会生活上の使用においてはおおむねその機能をはたしているとは言い難い状況におかれている。しかし、ロシア語以外の先住の人々の諸言語はシベリアの広域に分布し、民族の接触の状況を如実に映し出していると言える。方向表現はセットとして取扱い易いという側面の他に、シベリア地域の主要産業である放牧や狩猟など、広域にわたって移動することが必須である生活形態を考慮すると、この語彙セットは諸民族にとって重要で、興味深いものであることは疑いない。その反面、例えば、地勢が少し異なれば表現も異なるなど、同一言語の中においても方向名称と見なせるものが一貫して認められるのかというような、扱いにくさはある。とはいえ、伝統的に当該言語において方向表現として使われる語彙、もしくは句の構成が現代のシベリアの諸言語にどのように反映されているかは、言語学のみならず、文化人類学、民族学に対しても少なからぬ情報を提供することになろう。

本稿では主な資料としてオロンホを扱ったが、オロンホの研究は旧ソ連邦時代から現在に至るまで民族学及びフォークロールを扱う分野においては、比較的注意を向けられることが多かった。しかし、テキストの言語学的分析についてはまとまった研究がなされているとはいい難い。本稿ではヤクート語の語彙研究の端緒となるものを著すとともに今後のオロンホの言語学的視点を中心に据えた総合的研究の方向性を探ることをも目標としている。

1　資料について

オロンホはサハ族に伝わる英雄叙事詩である。一編毎の分量も様々である。口承文芸であり20世紀に入ってから文字や蝋管などの音声記録機材により収録されたものが多く残っているが、刊行されているものは少ない。それには、ソビエト政権時代、文芸作品としての価値をあまり公にされず、また、その継承も奨励されず、どちらかといえば、否定的に考えられていたという背景がある。近年、保存の気運が高まりつつあるとはいえ、その継承に亀裂が生じていることは否めない。詳しくは拙稿「ヤクート語語彙研究(1)」[4]を参照。

オロンホは民間の口承文芸であり、サハ族の始祖や創世についてなどサハ族の世界観などを反映している。又、口承文芸である故にそのテキスト自体はオロンホが語られる時代により変容するものであろうが、一方でサハ族が現在の地に居住するに至るまでに接触した言語の痕跡が残っている可能性も大きい。

なお、言語資料としてのオロンホはフォークロールの資料としてロシア科学アカデミーにより収集された録音資料がいくつかは残っているがその整理と分析はこれからの公開、公刊にかかっている。また、年々少なくなっている「オロンホの語り手」の演ずるオロンホの収集もさらに進められなければならないだろう。

1.1　サハ族英雄叙事詩オロンホ『クース・デベリィエ』

本稿で扱うオロンホはyak. Kïïs Däbilijä(乙女　燃えさかる(こと)-所有接辞)『炎の乙女』[5]をタイトルとし、ノボシビルスクで出版されたものである(以下『オロンホ』と略す)。収録は1941年6月から7月にかけてヤクーチヤ中央部のウスチアルダン地区で、ニコライ・ペトロビッチ・ブルナシェフ氏 Николай Петрович Бурнашев (1846-1950)から

学術情報センター電子図書館サービス　　©学術情報センター

同地区出身のジャーナリスト、ステパン・コンスタンチノビッチ・ヂヤコノフ氏 Степан Константинович Дьяконов (1913-1966)がおこなった。テキストの分量は4969行である。筋書き等、詳細は拙稿「ヤクート語語彙研究(1)」を参照されたい。

1.2 その他の資料

本稿では、上記の他に、補足的に以下も資料として用いた。

1) ヤクート語訳福音書[6]

かつて帝政ロシア時代のシベリア進出を背景にギリシア正教の布教活動が盛んに行われた。それにともない宣教師等による福音書の民族言語への翻訳活動が行われた。近年、ソ連邦時代は冷遇されていた宗教活動復活への兆しがある中で、その成果である出版物が復刻版の形で再版されることがある。1898年にカザンで出版された福音書のヤクート語訳の復刻版が1994年に刊行された。新約聖書である。『オロンホ』とこの福音書において使われたヤクート語は口承文芸と散文の翻訳本であるという点、ジャンルが全く異なるという点等でその性格を異にする。しかしながら、両者が現代ヤクート語の基盤となったヤクート語を反映している限られた資料であることは疑いない。

2) 現代ヤクート語の言語資料

上に掲げた文献による資料の他に1991年及び1996年にシベリアでのフィールド調査により収集した資料も参考にする[7]。ここでの資料はヤクート語(ドルガン語を含む)を母語とする人々に対して、どのようにして方向を表すかを問い、得られた資料である。地域により、異なる回答を得たが、これらとヤクート語の標準語と比較する。

2 ヤクート語における方向表現

2.1 ヤクート語語彙

ヤクート語は現在に至るまでにその主たる話者であるサハ族の移動によりもともとのチュルク語の語彙層にいくつもの語彙層が重なった。6割以上が借用語彙である可能性が指摘されている。なかでも量的に大きいのはモンゴル語からの借用語彙で、おそらくはバイカル湖沿岸域にあった頃に接触したブリヤート族からの借用語彙である。これらモンゴル系言語からの借用語彙はヤクート語の語彙の25%にのぼるとされる[8；p.548]。また、モンゴル語からの借用の他にも、レナ川に沿ってシベリア東部域へと移動する過程で、またシベリアにおいて有力な民族として先住のツングース系諸民族を圧迫、またはこれらと融合同化していく過程で多くの語彙を取り込んだ。ヤクート語の中には他の先住の諸民族、例えば、ユカギール族などとの接触により借用されたと考えられる語彙も存在し、またさらには起源の不明な語彙もみられ、ヤクート族の移動の軌跡を示唆するものでもある。

さらに16世紀、より盛んには17世紀になってからのロシア人の東漸によりヤクート語は大量にロシア語との接触を余儀なくされ、行政、貨幣経済、宗教等の多岐にわたる文化、文明等々に関連する語彙が導入される。特に10月社会主義大革命以後、ソビエト政権樹立後は「文化基地」がシベリア各地に建設され、教育制度をはじめとする同化政策の導入によりそれまでとは比較にならないほどの勢いでヨーロッパロシアの文化や文明や政治用語などの語彙がロシア語借用によってヤクート語に取り入れられた。

ソ連邦崩壊の少し前から、それまでは少なくとも表面にでることのなかった、民族固有の言語を尊重する傾向が見られるようになり今日に至っている。もっとも、かつてのロシア語からの借用語彙をヤクート語に翻訳借用し直すなどは、あまり、頻繁とは言えないようである。大革命後はヤクート人は若い年齢層を中心にヤクート語とロシア語との2言語併用状態にあり、現在に至ってはロシア語を第1言語とするヤクート人人口が加速度的に増え続けている。

ヤクート語語彙にはこのように現在の分布域に至るまでに形成された語彙層とその地に至って後に形成された語彙層とがある。これらの歴史的形成を経たヤクート語の語彙の方言差もまたその史的形成という点で非常に興味深い。ヤクート語は広範囲に分布する割には方言差の小さい言語であるとされることがある。しかし、地域によっての語彙の差異は必ずしも小さいものであるとは言い切れない。言語接触を考慮して分析していけばヤクート語が他のチュルク系言語と特に異なるとされる点について、また、ヤクート語の方言差に異系統の言語が及ぼした影響についてもこれまでより明らかになるであろう。それによりシベリアという地域での系統の異なる諸言語の語彙に共通した特徴が提出できるかも知れない。

学術情報センター電子図書館サービス　　©学術情報センター

2.2　ヤクート語における方向表現

2.2.1　方角名称の形式

現代ヤクート語の方角表現として使用されている語彙を主として以下に掲げる。まずⅠ表に現代標準ヤクート語を掲げる[9]。さらに、Ⅱ、Ⅲ表は1991年及び1996年にヤクート・サハ共和国、及び極北シベリアのタイムル自治管区で行ったフィールド調査により得た回答をまとめたものである[10]。これにより、現代ヤクート語においては方角を表現する語彙としてどのような語彙が使用されているかを示したい。[11]

現代標準ヤクート語

備考欄には各語彙が本来的にもつ意味を記した。詳しくは3章を参照のこと。

Ⅰ表

日本語	ロシア語	ヤクート語	備考
「東」	vostok	ilin	「前」
「西」	zapad	arγaa	「後ろ」
「南」	jug	soγuruu	「上へ」
「北」	sever	xotu	「下へ」

現代ヤクート語(方言形式との比較)

Ⅱ表

日本語	現代 yak.標準語	yak.北東(モムスク)方言
「東」	ilin	arγaa
「西」	arγaa	ilin
「南」	soγuruu	soγuruu
「北」	xotu	xotu

ドルガン語(ヤクート語タイムル方言)

Ⅲ表 [12:p.271]

日本語	ﾎﾟﾋﾟｶﾞｲ村	ｶﾃｨﾘｲｸ村	ﾍﾀ(ﾑﾁﾉ)村
「東」	tään	川の流れに	allara
「西」	taas	沿って表す	haapaat
「南」	kotu	mas diäk(i)	mas d'äk, üühä
「北」	muorattan	muora diäk	muora d'äk

täänは「平地」taasは「石」muoraは「ツンドラ、海」masは「木」diäki, d'äkは「～の方へ」üühäは「上へ」allaraは「下へ」haapaatは「西」(rus. zapad [za:pat]からの借用)

2.3　『オロンホ』に現れる方向表現関連語彙

『オロンホ』のテキストに現れる方向表現に関連した語彙を以下に掲げる。詳細は5章の単語の形式別のデータを参照のこと。なお、ここに掲げた語彙の分類は『オロンホ』におけるロシア語の対訳を参考にして行った[13]。

方向表現として『オロンホ』に見られるものを以下に一覧する。方向の出現の頻度については『オロンホ』の内容に当然、関わることであるので一概に特徴とすることはできないが、「東」「西」「南」「北」等が主要な方向表現として安定して見られる。また、方向表現に関連するものとして2.1.～2.2.にあげたように川の流れによっての「上(流)」「下(流)」がある。これらは必ずしも直接に方角を表すものとして一括して捉えることはできないであろうが、方向表現として扱い得る。同様に場所を表す語彙でありながらも方向表現に関わるものとして「(真)中」「方」「側」等を方向表現及びそれに関連する語彙のセットとして広義に考える。

2.3.1　「東」「西」「南」「北」

以上でみてきたように方角表現としてこれら4つの語彙はほぼ、安定して『オロンホ』にあらわれる。

東	ilin
西	arγaa
南	soγuruu
北	xotu

2.3.2　「上(方)」「下(方)」

これらの語彙は「東西南北」が本来的な方角名称であるのに対し、川の流れを表すことにより方角を示すのに使われる表現である。ヤクート語が現在のシベリア地域において使われるようになってから実際の立地にあわせて発展していった形式であろう。Ⅰ表に示したように方角名称の「南北」をあらわす単語はもともと意味上、「上下」の対語となっていた。しかしながら、方角名称として固定したため、「上下」の意を失った。一方、『オロンホ』やヤクート語の方言にみられる方向を表す語彙形式は本来の「上下」の意味を各々保持したまま方向表現に組み込まれたものである。従って、『オロンホ』に使われている方向表現としての語彙形式には、方角表現に比べて、バリエーションが見られる。「上下」が本来的に意味する垂直方向にのびる空間を示すことも多い。さらに、『オロンホ』のテキスト中の「上」「下」は現実世界である「中」を加えて、非現実世界の「上界」「下界」を表して熟語的に使用されることもある。

学術情報センター電子図書館サービス　　©学術情報センター

上下(高低)
上(高)　taŋnarï
　　üöhää
　　örö
下(低)　alïn, anna
　　allaraa

2.3.3 「側」「方」「地方」

これらはいずれも方角を表す語彙形式に付属して現れるが、単独で現れることはなく、語彙形式としては「拘束形式」といえるであろう。『オロンホ』においては「東西南北」と組合わさる例は少ない。

方　anaraa öttügär
　　alïn öttügär
四方　tüört ärkinin
万国　üs dojdu

2.3.4 「中」

『オロンホ』において「中」は「上界」及び「下界」とあわせて、「中界」として文字通り、中心的な世界、つまりは現実世界を表す表現としてこちらも熟語的に用いられることが多い。この表現はサハ族の世界観にも関係するものであるが、本稿ではその点には言及しない。

中心(真中)
　　orto
　　kiin

ヤクート語においては orto が方向表現を形成することがある[14:1875]。

kün orto 「正午」　tün orto「真夜中」
kün ortoto「正午、南中、南」 tün ortoto「真夜中、北」

『オロンホ』には orto を単独で用いた方向表現はないが、ヤクート語にこのように太陽の動きを組み込んだ表現があることは方向を表すのに実際的に有効である「川のながれ」や「地勢」等を基準とする表現に加えて、太陽を基準とした一連の方向表現がセットとして存在することを可能にする。語彙形式「前後」により方位を分割するについてはヤクート語のみに見られるのではなく、古代チュルク語にもその用例がある。従って、「前」＝「東」、「後ろ」＝「西」の組み合わせが古代チュルク語からヤクート語に至るまで継承されている、と見ることもできる。

2.3.5 「右」「左」「前」「後」

この項目については後掲の分析の章を参照のこと。

前　innin(ilin)
後　kännin(kälin)
右　uŋa
左　xaŋgas

2.3.6 方向表現よる「熟語」

『オロンホ』においては方向表現が特定の名詞と結びついて繰り返し生起している場合がいくつか見られた。それらについて詳細及び実際の形式は、5章のデータを参照されたい。固定した表現により「熟語」となっているものについても言及する。

例　「上界」　üähää dojdu
　　「中界」　orto dojdu
　　「下界」　allaraa dojdu
　　「西のシベリア」 Aryaa Sibiir

2.4 ヤクート語聖書に見られる方向表現

ここで特に述べておきたいのはヤクート語による福音書における方向表現は「東西南北」を明確に区別しているという点である。

例えば、ルカによる福音書 13 章 29 節は次のように訳されている：Ilintän, aryaattan, xotuttan, soyuruuttan da kälännär, Taŋara caarstvatïgar öttügästään sïtïaxtara.「それから人々が、東から西から、また南から北からきて、神の国で宴会の席につくであろう。」。(ヤクート語下線部は全て奪格形)

この他の箇所でも福音書において方向が標示されているところではいずれも上例にある方角名称が例外なく用いられている。その他の方向表現は使われない。福音書のヤクート語は翻訳がなされた当時(19 世紀末頃)のヤクート語に近いというよりむしろ、忠実に翻訳を行うため、ヤクート語話者の理解の範囲を越えない程度に「超」規範化を行っている傾向があると考えられる。この点に関連しては、拙稿「ヤクート語語彙研究(1)」(p.169)において「金」「銀」をあらわすヤクート語形式が一貫して特に区別されていた点に言及しているので参照されたい。

3 方向表現分析

前節にあげた方向表現はいずれも移動の方向を示したり、修飾語として使われたりする。ヤクート語において一般には移動の方向を表すには与格形、および後

学術情報センター電子図書館サービス　　©学術情報センター

置詞 diäk(i) が用いられる：Türgän soγustuk d'iätigär källä.「かなり速く彼は家に(d'iä-tigär「家-(彼の)へ」)着いた。」Küöl diäki barda.「彼は湖のほうへ(küöl diäki「湖+方へ」)行った。」[15]。本節ではこれらの一般に方向を表す方法を除いて、方角名称「東西南北」に近い意味をもち、方向を表すために使われる表現について分析する。

3.1 方角名称

ヤクート語の方角名称としては、I表にあげた語彙形式を用いるのが一般的である。これらは「東西」、「南北」の各々の組について方位以外の反意語のペアを本来なしていた。

「東」ilin ＜古代チュ.[16] alin 「前」「額、顔」
　～＜古代チュ. ilgärü 「前へ」「東へ」
「西」arγaa ＜古代チュ. arγaa「後」「項、背中」
「南」soγuruu ＜古代チュ. joqaru 「上へ」
「北」xotu ＜古代チュ. qodï 「下へ」

この中で、ヤクート語「東」ilin に対応する ilgärü と対になる語彙として qurïγaru 「西へ」「後ろへ」が古代チュルク語において対語となっている。

3.2 方向表現

方角名称以外の方向表現がヤクート語(ヤクート語の1方言とみなすことも可能な極北のドルガン語を含む)においてどのように表されているか、その意味するところから、以下のグループにわけることが出来る。

① 地勢的なもの
② 身体名称に起源をもつもの
③ 「左右」等

① このグループにはいるのは実際の地勢に沿ってあらわされるものである。

方角名称の他に有力な方向表現として方向を迂言的に表すものをあげることが出来る。例えば、上に掲げたⅢ表の muora diäki 「ツンドラの方へ」＝「北」、taas di「石」＝「西」についていえば、「ツンドラ」、「石、つまり山」は各々、この表現を収録した地点において地勢の上からみて「北」、「西」にあたる。

このグループに入る表現としてドルガン語を含むヤクート語において川の流れによって方向を表す例がよくみられたが、これも実際の地勢によっているといえる。ヤクート人にとってシベリアでの移動に大きな役割をはたしていたのは河川であった。そのことは現在に至るまでヤクート人の多くが大なり小なり河川の沿岸域に居住していることからも推察し得る。その移動の方向を表現するにおいて「川の流れ」を用いたことは自然である。このグループの方向表現はその形式の各々が実際に方角名称のいずれに相当するかはその使用地点により異なる。一般的には、「上流(へ)」「下流(へ)」として、各々、「南」、「北」を表わすが、これはヤクート人に限らず、諸民族の移動において重要であった大河川が北氷洋へと北へ流れるためであろう。

その他の方向表現として、主に野外調査から得られた表現をいくつか挙げることが出来る。方角を表す表現にはヤクート語のインフォーマントによれば、まず第一にあげられるのが「川の流れ」である。さらに「星の位置」、「風の吹く方向」、「地理上の目印」即ち、山や川等の地勢により示すなどがある。

上掲のⅢ表に見られるように方向を表すために必ずしも東西南北のような方角名称が使用されるわけではなく、むしろ地勢に応じた形式が見てとれる。Ⅲ表の「北」を表す語彙に使われる muora「ツンドラ」は各々の形式において muorattan「ツンドラから(単数・奪格形)」、muora diäki「ツンドラの方へ」として使われている。一方 mas diäki「木の方へ」は「南」を表わす。地勢からいえば、南方のタイガ地帯に対して極圏のツンドラ地帯が北方に位置している。

② 現代ヤクート語において方向表現の機能を果たしているが、同時に主として人間の身体部位を意味する、あるいは起源を身体名称にもつものがある。

例えば、arγaa は「西」を表すが同時にこの単語は「項」「背中」「後ろ」を意味する。「東」を意味する ilin は同時に「前」をも表し、もともと古代チュルク語の「顔」「額」を表す alïn にさかのぼると見ることもできる。[14:924]

方向名称というよりむしろ位置を表す語彙であるが、is 「内」は同時に「内臓」を、kiin「真ん中」は「臍」を、「側」「傍ら」を表す ojoγos は「脇」を意味する。

このように方向あるいは方位表現を身体名称になぞらえることは、必ずしも、特定の地域や民族に限ったことではないが、寄り広いジャンルの語彙について調査をすれば、パターン化できる可能性があると考える。

③ 広義に方向名称を捉えた場合、「左右」「真中」「中」等を考慮に入れ得る。このうち、「左右」は「南北」の意を持つことがあるのでやや詳しくここで触れておきたい。

「左右」については各々 xaŋas「左」、uŋa「右」がヤクート語にはあるが、このうち uŋa は古代チュルク語の uŋa、oŋa に起源を持ち、チュルク語に本来的な

学術情報センター電子図書館サービス　　©学術情報センター

語彙である。このuŋaはヤクート語においては「南」をあらわすことがある。『古代チュルク語辞典』[16]にはヤクート語の他にキルギス語においても同様の例があげられており、ラドロフ[17]によれば、アルタイ語にも「南」の使用例がある。その他のチュルク語には「南」の意味での使用は見られないようである。一方で「左」を表すxaŋas～xagasはペカルスキー[14:3311-2]によればモンゴル語のxaγas「半分」に対応する。ヤクート語にはxaŋasに「半分」の意味はないが、この2つの語彙が同源であるとすれば、ヤクート語にモンゴル語から借用されて後、この単語は独自に「左」を表すようになり、チュルク語に本来的な「右」とともに「左右」のペアとなったと考えられる。「左」を表すチュルク語本来の形としてはsolがあるが[17]、この形がヤクート語に継承されていれば、語頭のsが脱落してolとなるが、この形式はヤクート語において指示詞「あの」「あれ」を表すolと同じになる。そこで、同音衝突を避けて、「左」を表す語彙としてモンゴル語から借用したxaŋasをあてたのであろう。意味の変容の順を確定できないが、「半分」の意を持つモンゴル語の単語から発展したのであるとすれば、まず、同音衝突がきっかけとなり、xaŋasが「左」の意を獲得し、「右」と対になり、やがて「北」の意味を表す場合も生じたと考えられる。なお、ラドロフ[17]によれば、キルギス語においてはsolが「左」の他に「北」を表すという記述があり、『キルギス露語辞典』[18]にも古風な意味として「北」の意味が挙げられている。キルギス語においては「南北」を表して、各々、太陽の出る方角に向かって右が「南」、「左」が「北」、「到着、出現」の意のchïgïshが「東」、「(天体につき)入り」を表すbatïshが西をあらわす。このことはサハ族がかつて方向表現のための基準として太陽の出る方角を基準に用いていた可能性を示唆する。

4　おわりに

現代ヤクート語使用域においては方向表現として規範的な「東西南北」の他に実際の地勢に沿った表現方法をとることもある。

方向表現には規範的な方角名称「東西南北」を含め、方位の分割の基準として「前後」「上下」さらに「左右」等がある。また、その定め方として「太陽の軌道(日の出の方角)」、あるいは「川のながれ」等の立地点の要素が入る。ヤクート語においてはチュルク語におそらくは本来的であった「前後」及び「上下」を表す語彙の組み合わせによる方向表現が方角名称「東西南北」として定着し、本来の意味を失い、標準ヤクート語に保持された。『オロンホ』やヤクート語方言等にみられるその他の方向表現はヤクート語内部で独自に発展したものと考えられる。いずれにしても方向表現はその言語の実際に位置している地点により、微妙に異なりを見せるものであるので、特定の語彙形式をもって固定的に方向表現と考えるのは早計である。しかし、一方では方向表現が規範的な方角名称になっていく過程で現実世界をどのように分割するかについては例えば、「前後」と「東西」が対応する、あるいは身体の名称が方向の名称と対応するなど、いくつかパターン化して捉えることが可能である。

本稿で扱ったこの語彙セットはヤクート語の語彙研究において他にセットとして扱う語彙と同様、広くチュルク語全般からヤクート語の特異性を把握するために重要である。今後は、さらにヤクート語の接触した周辺の、つまり、シベリア全域の先住の人々の言語についても比較の対象とせねばならない。この点については稿を改め、より詳しく検討したい。本稿で提出した『オロンホ』のデータは、そのための基礎的資料となるであろう。

学術情報センター電子図書館サービス　　◎学術情報センター

5　データ

＜データ1＞

( )内の数字は『オロンホ』Kïïs Däbälijä のテキストに付されている行の番号である。原則として、『オロンホ』から、方向表現やそれに類した表現を抜き出してある。対訳文として付されているロシア語もそれに応じて示した。複数の箇所での出現はまとめて呈示した。

『東』 ilin

iäjiäxsittääx ilin xallaannaax,(39)(477)(4610)
с благодатно-добрым восточным небом,
ilin xallaan (214:abl.)(4138)
из-под восточного неба
ilin xallaan is äŋij ätittän (2828)(4145)
из-под внутреннего ската восточного неба
ilin öttünän ilbistääx xajaäŋärdääx (3180)
с востока кровожадной горой окаймленное,
iliŋŋi iäjiäxsittääx ïllïgïn(3983)
по восточной тропе с Иэйиэхсит
bu alaas sir ilin tardïïtïgar (3996)
на восточной окраине этого алааса
ilin diäki ärgitä battaat,(4280)
на восток круто повернув,
ilin-arγaa, (4784)
на восток, на запад,

『西』 arγaa

anïsxannaax arγaa xallaannaax,(38)
с прохладно-ветреным западным небом,
arγaa xallaan(107)(206:abl.)(278)(1231)(1784)(4089)(4211)
западного неба,
arγaa Sibiir(110)(113)(1234)(1786)(2507)(2892)(3944)(4091)(4119:dat.)(4134:dat.)(4212)(4335)(4623:dat.)(4835)
Западный Сибиир
arγaa tardïïtïgar (1035)
на западной окраине его
alaas sir arγaa bahïgar (1053)
на западной оконечности алаас-страны
ajïïlaax-andagardaax arγaa xallaan(1729)(2504)(2889)(3941)
[туда, где] под нижней окрайной
tus arγaa diäki (1733)(4637:abl. diäkittän)
прямо на запад,
ald'arxajdaax majdaan arγaa bahïnan (3111)
по задней стороне гибельной площадки,
arγaa öttünän adaar xaja d'ajï ŋnaax, (3178)
с запада остроконечными горамим окаймленное,
arγaa öttünääγi (3261)
на западной окраине [находящегося],
arγaa abaahï ajmaγa arγastar (3793)
по проторенному племенем абаасы западному [пути],
ilin-arγaa, (4784)
на восток, на запад,

学術情報センター電子図書館サービス　　© 学術情報センター

「南」 soγuruu

sollo ŋnoox soγuruu xallaannaax, (40)
с ненасытно-жадным южным небом,
soγuruu xallaantan (210)
из-под южного неба
xaxajdaax xald'aajïlarïn xa ŋas öttünän(407)
южный косогор со львами слева обойдя,
soγuruu chanchïgar,(499)(2134)(4965)
на южной стороне
soγuruu suruktaax xallaan (1645)
южного расписного неба,
soγuruu d'ajï ŋar, (1837)
на южной стороне
soγuruu öttünän (3177)
[на место,] с юга,
soγuruu öttünääγi (3257)
на южной окраине,
suruktaax soγuruu xallaan uola (3668)
сын расписного южного неба
xotu-soγuruu (4785)
на север, на юг,

「北」 xotu

xoloruktaax xotu xallaan,(41:xallaannaax)(2433)
с буйно-вихревым северным небом,
xotu xallaan (202:abl.)(1832)(2123:xotugu)
из-под северного неба к этой стране
xotu diäki salaja battaat, (2072)
повернув круто на север,
xotu d'ajï ŋar (2136)(4967)
у северной окраины
än samantan tus xotu (2686)
отсюда прямо на север,
xoptoox-d'iptääx xotugu xallaan (2715)
северного неба
xotu öttünän (3179)
с севера
xotugu öttünääγi (3265)
на северной окраине находящегося,
xotu-soγuruu (4785)
на север, на юг,

「前」 innin

innin diäki kimiris, (338)
вдруг вперед подалась:
innigär ürbälänän,(366)
впереди себя погонишь,
xotunnarïn tus innigär (588)
прямо перед своей госпожой
salbax bahïn salajan, (2115)
большую голову вперед устремив,

学術情報センター電子図書館サービス　　© 学術情報センター

ilin öttügär (2117)
спереди сопровождаемый
innigär-kännigär süürär (2817)
взад-вперед бегающми!
sordoox xallaan bätäräätä, (3081)(3099:dat.)
к переднему краю злокозненного неба,
dibdigirächchi töttörü-taarï xaamïtalaan baran, (3198)
гулко стуча, взад-вперед прошлась,
suol aanïgar suru suullaran, (4737)
на передней стороне серую зарезав,

「後ろ」 kännin
kännin diäki kääd'äs gïnan baran: (339)
потом назад откинулась:
kännigär sätiilänän taxsargïn (367)
за собой на веревке кого поведешь,
känninän täjiäkälään ti[ä]jän baran, (1439)
назад отлетев, попятился
kälin öttügär (2121)
сзади провожаемый
sobolonnoox xallaan ulaγata,(3080)
к заднему краю алчного неба,
kätäγiriin diäki käräni oxtoron, (4734)
на заднем дворе белую [лощадь] забив,

「上」①taŋnarï
taŋnarïtardïllan, (15)
сверху ниспадая,
taŋnarï iäγillän tühän,(27)
сверху вниз выгнувшись,
allaraa dojduga taŋnaarï sann'ïlïjan tühär. (524)
мигом в Нижний мир скользнула.
allara taŋnarï sann'ïlïjar, (943)
вниз усколъзать будет,
alta-sättä taŋnarï baγïïrdaax (1456)
шесть или семь книзу загнутых
taŋnarï üünän tüspüttär, (1476)
вниз опустившись, выросли;
taŋnarïta tamnaattaabït kurduk, (2003)
вверх дном перевернутые,
taŋnarï tuttaxxa chaskïjan ïlar, (2011)
если вниз острием опустить - визгом визжащий,
taŋnarï tamalaha xaallïlar, (3090)
расползаясь, вниз проваливались,
taŋnarï saxsïllan tühän istilär. (3537)
пока они вниз так летели,
taŋnarï baγïïrdaax tumustaax, (3627)
с загнутым вниз крючковатым клювом,
taŋnarï archïlaatïlar. (4157)
сверху вниз заметались.

学術情報センター電子図書館サービス　　© 学術情報センター

「上」② üöhää

üöhää taxsaachchï kïïs (513)
девка, поднимающаяся в Верхний мир,
üöhää taxsïbït kïïsätääri: (544)
девка, в Верхний мир поднимавшаяся:
üöhää dojdu uola (641)(656)(2230)(2655)(2673)(3309)(3789)(3871:acc.)(3883)
вдруг у парня Верхнего мира
üöhää dojdu oγoto, (696)(1245:oγoloro)
тебя, сына Верхнего мира:
üöhää uoha (734)
верхняя губа
üöhää diäki (1094)(3618)(3928)
поверх долины
üöhättän suruktaax buolbatätin duo? (1270)
предначертание сверху имеющий?
üöhättän tüspütä küdänä köstübätä, (1430)
упал ли сверху - воздух не шелохнулся,
ol üöhää öttün örö körön tahaardaxxa, (1445)(1450)
выше глаза поднял и присмотрелся:
ol üöhää öttügär ---(1463:pl.) (1469)(1474)
а выше этого -
üöhättän tölkölööx (1661)
предначертание свыше имеющий,
üöhättän törüttääx (2644)
сверху предопределение имеющий
üöhää böγöttön (3012)(3453:dat.)
Если от чудища Верхнего мира
üöhää buollar(3656)
и наверху, будь в Верхнем мире -
üöhää sügün sïld'ïbakkïn, (3657)
наверху тебе неймется,
üöhää ürdük xallaaŋŋar ütäjäbin, (3908)
на высокое небо твое повелеваю тебе подняться,
üöhää chäbdik xallaan (4352)
на верхнем чистом небе живущих,
üöhää dojduttan (4562:abl.)
со стороны Верхнего мира

「上」③ örö

örö körön turan, (21)
вверх вперив,
örö kinärijän, (410)
кверху поднявшись,
örö d'ulurujan, (412)
вверх взойдя,
örö d'abïdïjan, (414)
наверх пробравшись,
örö küöräjän, (418)
вверх выскочили;
örö suodujan taxsan, (426)
проворно вверх поднявшись,

学術情報センター電子図書館サービス ◎学術情報センター

örö suburus gïnan xaalar, (516)
вверх взметнулась;
soγotox xaraγïn örö (683)
единственным оком вверх уставившись,
örö kötör künnärä källäxtärinä, (946)
ввысь уноситься будет - дни такие придут,
örö kïstaan öröhölöön tahaarbïttar äbit. (1211)
высокую кучу из них сложили, оказывается.
örö d'aarïstaan tahaarbïttar. (1291)
[волки] ярусами вверх взгромоздили.
örö tuttaxxa üögülüü tühär, (2012)
если вверх поднять - криком кричащий,
örö kötön (3718)(4048)(4283)(4614)
вверх полетела.
örö xalïhïjan taxsan, (3797)
вверх устремилась,
örö üünän turarïn (3809)
ветви вверх взметнув, выросла,

「下」① alïn
alïn öttügär (30)(494)
под тем ～,
alïn tardïïtïgar, (108)(1004)(1232:dat.)(1785:dat.)(2505:dat.)(2890:dat.)(3942:dat.'западного неба')(4211)
под нижним слоем
anna nälägär, ürdä ümürük (415)
снизу широкому, сверху узкому,
alïn kïrïïtïgar, (480)
под нижним краем
sarïal kün annïgar (481)
под восходящим солнцем
annïgïn aγïs gïna xajïtïaxpït, (776)
нижнюю часть на восемь частей разрубим,
ajïïhïttaax xallaan (818)
под нижним краем
alïn ini öttügär, (819)
под внутренним краем
alïn tiräxpin (1252)(1678:tiräγä)(2516)(3555:tiräγiŋ)
нижние опоры
soduomnaax xallaan alïn kïrïïtïgar, (1646)
у нижнего края беспокойно-буйного неба,
annïnan xarïs saγa (2722)
к дому - снизу даже с вершок
soduomnaax xallaan anna (3082)(3101:annïgar)
к нижнему краю беспокойного неба -

「下」② allaraa
araj, Allaraa (233)
разве что Нижнего мира
allaraa ald'arxajlaax-arallaannaax (492)
у старика Ёлюю Ётюмэхтэя,
allaraa dojduga(521)(524)(3018:dajdïttan)(3953)(4564:dajdïttan)

学術情報センター電子図書館サービス　　© 学術情報センター

в Нижний мир
allaraa dajdï oγoto (707)
тебе, сыну Нижнего мира:
allaraa sïmïhaγa allara sïllarïchchï tarta. (735)
нижняя губа его вниз отвисла.
allara taŋnarïsann'ïlïjar, (943)
вниз ускользать будет,
allaraa adaγalaax-arbaγastaax (1235)
старика Нижнего мира Арсан Дуолая,
allarattan taxsïbïta bïrkïta billibätä, (1429)
снизу ли выскочил - признаков нет,
allara böγöγö argïar buolan argïjan, (3454)
Нижнюю твердь (мир) студеным ветром обдавая,
allaraa appaŋnïïr aat ajaγïn (3496)
На самом нижнем краю
allaraa buollar (3658)
и внизу, будь в Нижнем мире -
allaraa batan xaampakkïn! (3659)
тебе угомона нет!

「右」 uŋa
uŋa timir tühägäj n'ilbägär (593)
на правое железное колено
uŋa ämiijinültü sïstaran, (596)
правую грудь свою размягчила
Chugdaan Buxatïïrï uŋa bülgününän siksirijän kääspitä, (1437)
Чугдаан Бухатыыра правым плечом так двинул,
onton uŋa säbärgänätin xaptaγajïttan (2033)
затем с правой балки-перекладины [взял]
uŋa sototun tahïnan (2051)
и в правое голенище свое
kïmn'ïïlïax öttüttän (2078)
с правого бока
uŋa xonnoγum annïgar ugan kuluŋ! --- diän (3288)
под правую подмышку суньте мне, умоляю!
uŋa sostuojugar (3507)
с правой стороны
uŋa iliitinän (3874)(4178:dat.)
правой рукой
uŋa diäki oloror kihi (4431)(4735)
на лавке с правой стороны сидящий человек
uŋa oroŋŋo oloror (4941)
мимо правой лавки проходя,

「左」 xaŋas
xaŋas xara tühägäj n'ilbägär olordon, (599)
на левое черное колено посадив,
xaŋas ämiijin ültü sïstaran,(600)
левую грудь размягчила
xaŋas chanchïgar (957)(4502)
на левой стороне

学術情報センター電子図書館サービス ©学術情報センター

ikki uolu xaŋas kuoraj dalaanïtïhïnan (1435)
д в у х п а р н е й л е в о й з а г р е б и с т о й л а д о н ь ю
xaŋas oroŋŋo (1752)
н а л а в к е с л е в о й с т о р о н ы
xaŋas ïtïhïn (3858)(3873:dat.)
л е в а я л а д о н ь
xaŋas n'ilbägär tuora uurbut, (4175)
н а л е в о е к о л е н о р е б р о м п о с т а в и в ,
xaŋas diäki (4435)(4736)(4934)
н а л е в о й с т о р о н е
kätäγiriin oroŋŋo kälän, (4938)
н а л е в о й л а в к е п о д о й д я ,

『(真)中』 orto
orto turuu dojdu (224:dojdum)(3234)
с р е д и н н о й п р о ч н о й з е м л и м о е й ,
orto dojdu(313:dat.)(517:dat.)(548:dat.)(781:dat.)(1431:abl.)(2153:dajd ï ttan)(3025:abl.)(3455:dat.)(4333:abl.)(4560:abl.)(4570:dat.)
С р е д н и й м и р
orto aan dojduga (363)
н а С р е д и н н о й и з н а ч а л ь н о й з е м л е
ortotunan d'ölökköj (416)
а в с е р е д и н е п о л о м у ,
uraann'ïktaax ortotunan (424)
в с а м ы й ц е н т р ,
turuu baraan toloon ortotugar (427)
в с е р е д и н е у ю т н о й п о л я н ы
ortoluu ahaan ihän, (642)
в с е р е д и н е е д ы - т р а п е з ы
orto dojdu oγoto, (703)
т в о е , с ы н С р е д н е г о м и р а ,
orto dojdu uola,(722)(920)(1401)(1684)(3786)(3935)(4533)
с ы н С р е д н е г о м и р а
ortoγun otut gïna olomoottuoxput! (777)
с е р е д и н у н а т р и д ц а т ь к у с к о в р а с к р о м с а е м !
d'iä ottotun diäki d'irägäld'ihän tiijännär, (856)
н а с е р е д и н е ж и л и щ а в м и г с о б р а л и с ь
tüöhün tïlïttan ötürü üünän taxsïbït (862)
и з с е р е д и н ы г р у д и в ы р о с ш е й ,
xabïllar xaba ortolorugar (1156)
в с а м о й с е р е д и н е [с т а д - т а б у н о в]
ortoloox üöhää dojdu oγoloro (1245)
С р е д н е г о и В е р х н е г о м и р о в д е т и ,
orto duguj buorga oloxsujbut (2423)
с р е д н е й з е м л е о б о с н о в а в ш и й с я ,
ortoku oroŋŋo baran (2446)
к с р е д н е й л а в к е п о д о й д я ,
ortotunan tutum saγa (2724)
с в е р х у д а ж е в п я д ь
oholloox sostuoj ortotugar kiirän, (3199)
н а с е р е д и н у г и б е л ь н о й с к а л ы п о д н я л а с ь ,

学術情報センター電子図書館サービス　　© 学術情報センター

ortotunan kutald'ïjan, (3464)
в середине заколыхалась,
ortotunan tobulu ïstanan, (3524)
вмиг посредине разверзлась:
oholloox sostuoj ortotugar (3770)
на середине гибельной скалы
otto duguj buorga (4087)
уединенной Средней земле,
orto buor tolugar, (4112)
чтобы Средний мир защитить,
ortotunan d'ölördöj (4179)
с отверстием в середине,
d'iä ottotugar kälän (4526)(4795)
на середину дома выйдя,

「真ん中」 kiin 他
oloruoγun anaraa öttügär,(88)
еще не воссел,
kiinigär(90)
на самой середине
d'ularïjar orojugar, (1100)
в самом центре ее,
köxsüm xaraγïn d'ölö kïŋaajaŋïj, (3220)
в середину спины моей не направляй,
üllär büörünürdügär (3807)
в самой его середине
toγus atïïrïnüörügär baar (4474)
потому среди девяти табунов
buharan budulutan baran, (4505)
на середину желтого стола,

「内」 is
is äŋijätä, (478:dat.)(1002)
внутренний край
is äŋijätin diäki (4139)(4282)(4611)
в сторону внутреннего ската

「側」「地方」 örüt
anaraa öttüγar,(3)(98)
по ту сторону
ulaγa öttügär, (6)
за невидимой гранью
ürütγ öttügär (467)
обитель имеющего
ahïï oloron, antax xajïsta. (724)
во время еды в сторону стал смотреть
ikki öttügär (1361)(1441:abl.)(4511)(4701)
ojoγoluu sïld'an chuoxajdaata, (1596)
со стороны наблюдая, ликовал,
üsö ttülärinän (1597)
с трех сторон

学術情報センター電子図書館サービス
© 学術情報センター

ulaγa öttügär, (1645)
по ту сторону
bätärää öttügär (1647)(1839)
по эту сторону
kün diäki öttügär baar (1979)
с того, на солнечной стороне сеней находящегося,
tügäxöttügär, (3499)
на самой нижней кромке
künöttün (3876)
в стороту солнца,
kïrïïtïttan körör (4022)
со стороны посмотреть -
bättäx körön (4236)(4245)
в эту сторону погляди

「わき」 ojoγos
kïrïïtïnan (644)(3316)
по бокам
ojoγos öttügär (2119)
сбоку сопутствуемый
ojoγohunan süürälään, (2351)
по бокам разлилась,

「四方」
tüört ärkinin büdüülääbätäx, (23)
четыре его стороны разглядеть не могут,
tögürüjär tüört tulumnaax,(67)
с четырьмя сходящимися сторонами,
ürällä tüspüttärin (511)
в разные стороны вдруг рассыпались.
ürüö-taraa (1567)(4787)
в разные стороны
xallaan tüört örtüttän (4153)
со всех четырех сторон небес

学術情報センター電子図書館サービス　　© 学術情報センター

注及び参考文献

[1] 本稿では「東」「西」「南」「北」のように専ら方角を表すのに用いられるものを「方角名称」と示し、これらを含め、広義に方向を表すものを「方向表現」とする。

[2] 略号：rus.＝ロシア語、yak.＝ヤクート語、dol.＝ドルガン語(ヤクート語タイムル方言)、古代チュ.＝古代チュルク語
転字:キリール文字の転字に際しては、ь(軟音符号)は'(アポストロフィー)で表わした。

[3] ヤクート族という名称は近年になってから民族自称であるサハ族に替えられつつある。本稿では言語名としては慣用に従い、ヤクート語とする。

[4] 学術情報センター紀要第8巻掲載「ヤクート語語彙研究(1)：色彩名称」(藤代節：1996：pp.155-189)，以後「ヤクート語語彙研究(1)」と略す。

[5] *Якутский героический эпос Кыыс Дэбилийэ*, Наука., Новосибирск, 1993, 以後『オロンホ』と略す。

[6] *Святое евангелие (на якутском языке)*, Типо-Литография В.М.Ключникова с. д., 1898, Казань, (Издание Монастыря Параклита, Оропос Аттикис, Греция, 1994).

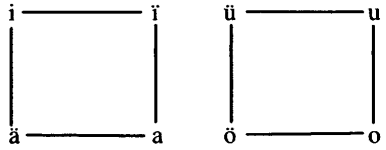
[7] 1991年に収集した資料については注12の拙論参照。ドルガン族の言語はドルガン語としてヤクート語から分岐した独立した言語とみなされることもあるが、ヤクート語のタイムル方言とされることも多い。1996年の野外調査においてはタイムル半島のドルガン語、サハ共和国北東部インディギルカ川沿岸のヤクート語モムスク方言の資料を収集した。

[8] 庄垣内正弘(1992)「ヤクート語」，亀井孝・他(編)『言語学大辞典』第4巻，亀井孝・他(編)，三省堂，東京，pp.544-550.

[9] Слепцов П.А.(ред.), Якутско-русский словарь, АН ССР, Советская энциклопедия, 1972, Москва.

[10] インフォーマントになって下さったのはサハ共和国北東部モムスク区のスレプツォフ・Ｉ・Ｉ氏、タイムル自治管区のキルギゾフ・Ｉ・Ａ氏等である。ここに記して感謝の意を表したい。

[11] ヤクート語の母音と子音については以下に掲げる表を参照。庄垣内 (1992) [8 : p.546-7] より。



短母音、さらにこれらと同種の長母音がある。
二重母音には、i ä, ï a, ü ö, uo がある。
子音には、次の20子音がみとめられる。

閉鎖音：p, b, t, d, k, g；
破擦音：č, ǰ；
摩擦音：s, x, y, h, ỹ；
鼻　音：m, n, ň, ŋ；
流　音：l, r；
半母音：y

[12] 藤代節「タイムィル半島(シベリア)のドルガン族の言語資料ノート」『言語学研究』11号，1992，京都，pp.261-284
Ⅲ表については拙稿「極北のドルガン語をたずねて」月刊『言語』23巻3号，pp.92-97でやや詳しく述べている。

[13] 『オロンホ』(1994)の対訳にあたった エメリヤノフ氏 Н. В. Емельянов、エフレモフ氏 П. Е. Ефремов、オイウンスカヤ氏 С. П. Ойунская はいずれもオロンホの研究者である。方向表現を広義にとらえたため、これらのヤクート語話者でもある研究者による対訳文は大変参考になった。

[14] Э.К.Пекарский, *Словарь якутского языка*, (АН ССР), 1907-1930, С.-Петербург, Петроград, Ленинград

[15] *Грамматика современного якутского литературного языка (Синтаксис)*, Академия Наук Республики Саха (Якутия), Институт языка, литературы и истории, Новосибирск, 1995, p.31

[16] *Древне тюркский словар*

学術情報センター電子図書館サービス ◎学術情報センター

ь, АН СССР, В. М. Наделяев, Д. М. Насилов, Э. Р. Тенишев, А. М. Щербак(ред), Наука, 1969, Ленинград.

[17] В. В. Радлов, *Опыт словаря тюркских наречий*, 1893, Санкт-Петербург.

[18] К. К. Юдахин, *Киргизко-русский словарь*, Советская энциклопедия, 1965, Москва.

学術情報センター電子図書館サービス ◎学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

Compression of Synthetic Movies Using Simplified Geometric Models

簡易幾何モデルを用いたコンピュータ合成動画像の圧縮

Hironobu GOTODA
National Center for Science Information Systems
学術情報センター 後藤田 洋伸
Kinji ONO
National Center for Science Information Systems
学術情報センター 小野 欽司

ABSTRACT

The MPEG standard is not necessarily suitable for compressing computer-generated synthetic movies. When transferring such a movie over a network, decomposing it into an outline and details, and compressing them separately, can lead to an improved compression ratio than compressing the whole movie using MPEG. Here, the outline is the portion of the movie that can be reproduced using simplified models, while the details are the sequence of difference images between the outline and the complete movie. We will elaborate this idea further by introducing mechanisms to exploit the interframe coherence. Our method adopts simplified geometric models both to reproduce the outline and to approximate the details. More specifically, an approximation of the optical flow is computed using the geometric models, which is then applied for estimating the coherence in the difference images. This leads to further reduction of the amount of details, and therefore less traffic over the network.

要旨

MPEG規格はコンピュータで合成された動画像の圧縮には必ずしも適していない。合成動画像をネットワークを通して転送する場合、全体を「概略部」と「詳細部」とに分け各々を個別に圧縮する方が、全体をまとめてMPEGを用いて圧縮するよりも、よりよい圧縮率が得られる。ここで概略部とは、動画像のうちの一部で、簡略化されたモデルを用いて受け手側で容易に再現できるものであり、一方詳細部とは、全体の動画像と概略部との差分動画像であって送り手側で計算される。本稿では、このアイディアを押し進め、さらにフレーム間相関をも利用することのできる方法について述べる。この方法では、簡略化モデルとして簡易幾何モデルを採用し、概略部を再現すると同時に詳細部の近似も行なう。具体的には、簡易幾何モデルを用いてオプティカル・フローを近似計算し、それを用いて詳細部のフレーム間相関を推定する。これにより、詳細部をさらに縮小し、ネットワーク上の転送量を軽減することができる。

[Keywords] MPEG-1, optical flow, motion compensation, distributed rendering, rendering hardware
[キーワード] MPEG-1、オプティカル・フロー、動き補償、分散レンダリング、レンダリング専用ハードウェア

1 Introduction

Recent rapid down-pricing of rendering hardware has enabled handy creation of 3D scenes and movies on desktop workstations using computer graphics. Many cheap, PC adaptable graphics boards are also on the market. However, to generate very realistic movies in real time, it is still necessary to employ high-performance graphics workstations equipped with expensive rendering engines. Since such workstations are so expensive, it is often desirable to develop some methods for sharing their processing power among low-end workstations via computer networks.

Without appropriate compression schemes, real-time transmission of high quality movies is almost impossible under the today's networking technology. Fortunately, the announcement of JPEG [12] and MPEG [9] compression standards motivated the development of image and movie com-

学術情報センター電子図書館サービス ◎学術情報センター

pression and decompression hardware, making it feasible for a high-performance graphics engine to operate as a rendering server. In the simplest setting (method A), the server compresses frames of images using either JPEG or MPEG-1, and sends the compressed images to a client, i.e., a low-end workstation, which decompresses and displays the images on a screen. In an improved setting (method B), the server generates high-quality and low-quality renderings for each frame, and sends only the difference images to the client in a compressed form. The client generates the corresponding low-quality renderings, to which the decompressed difference images are added. Method B requires more computation on both server and client than method A, but less traffic on the network.

We elaborate the latter approach (method B) to further reduce the required network traffic. Toward this end, we first review and analyze why method B leads to a better compression ratio than method A. Based on the analysis, we then present an improvement over method B designed specifically to deal with synthetic movies. The newly developed method exploits the interframe coherence using simplified geometric models. More specifically, the models serve to generate the low-quality rendering and to approximate the optical flow between consecutive frames. We use this approximate flow information to predict the motion in the sequence of difference images between high-quality and low-quality renderings. If both the server and client are to predict the motion in the same way, it becomes enough to transmit only the difference of the actual motion from the predicted one. This significantly reduces the required network traffic.

2 Related Work

Movie or video compression has long been a subject of intensive research in image processing and transmission. MPEG [8] and Quicktime [11] are two major industry standards for lossy video compression. Methods based on wavelet transforms have also been proposed [10]. These were originally developed for compressing natural scene movies, but are not necessarily suitable for compressing synthetic image sequences such as computer animation frames.

Researchers in computer graphics have attempted to compress specifically the rendered images of dynamic polygonal scenes. In contrast to image-based compression approaches, these approaches employed geometric information to improve the compression ratio. For instance, Guenter et al. developed a lossless compression scheme based on the geometric information derived from animation scripts [4]. They used the transformation matrices included in the scripts to enhance the motion prediction. Wallach et al., on the other hand, presented a scheme to accelerate MPEG compression itself using optical flow information derived from polygon movement [13]. Their method fitted well to the matrix-vector multiply and scan-conversion hardware installed on many graphics workstations.

The polygon-assisted compression proposed by Levoy stood in the midway between the image-based and geometric compression approaches [9]. He divided the compression task into two parts; one suitable for geometric coding, and the other suitable for image-based coding. The method succeeded to recover sharp edges of objects and smoothly shaded areas, which was difficult if the image-based coding alone was applied.

The Levoy's polygon-assisted compression is a broad scheme, which allows many variations. Levoy presented and discussed three variations, two of which worked well but one did not. A certain analysis is necessary to explain why and under what conditions the combination of geometric and image-based compression works better than using the image-based compression alone. In addition, the Levoy's scheme did not consider interframe coherence. Only the image-based coding using MPEG-1 could partially benefit from the interframe coherence. However, more efficient coding is possible if we combine the motion information from both parts of the geometry-based and image-based coding. We will elaborate on these issues in detail.

3 Basic Compression Scheme

We will briefly review and analyze the polygon-assisted compression proposed by Levoy [9]. The analysis will serve as the basis of the revised compression scheme that we present in the next section.

3.1 Scheme Description

Perhaps it is best to explain the scheme using the particular client-server relationship where the client submits a rendering task to the server and receives the rendered image. Fig.1 illustrates the elements of this relationship. The basic idea underlying the polygon-assisted compression is to split

学術情報センター電子図書館サービス ◎学術情報センター

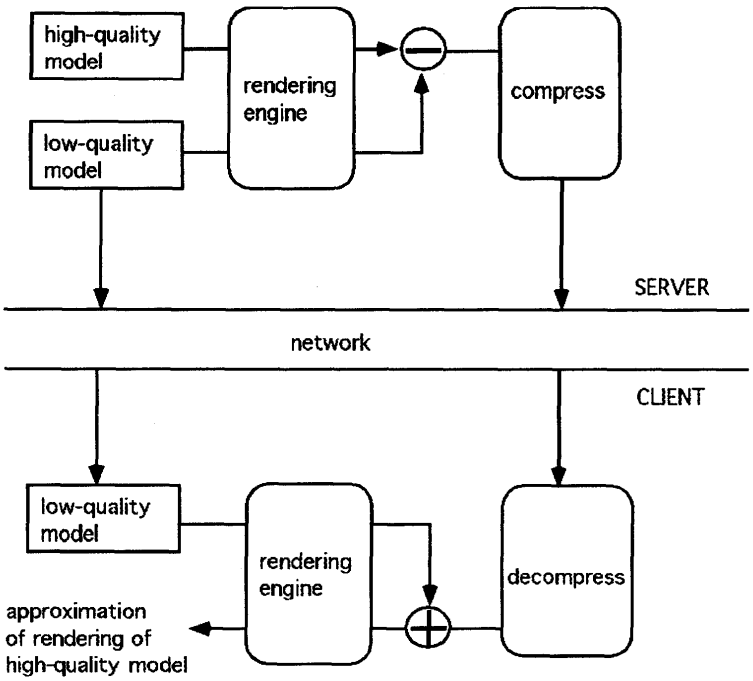


Fig.1 Data flow diagram of the polygon-assisted compression scheme

the rendering task into two parts, i.e., low-quality and high-quality renderings. More specifically, the server renders both low-quality and high-quality models, and subtracts the resulting rendered images, and sends the difference image to the client. The client, on the other hand, renders the low-quality model only, adds the received difference image to the rendered one, and displays the outcome. When sending the difference image, the server applies a lossy compression scheme to reduce its size.

There are many ways to derive the low-quality model from the high-quality model. Three particular methods were introduced and discussed in [9]: (1) untextured surfaces from textured ones, (2) coarse tessellation of surfaces from fine tessellation, and (3) nonantialiased Gouraud-shaded surfaces from antialised Phong-shaded ones. According to [9], strategies (1) and (2) worked satisfactory, while (3) did not.

3.2 Analysis

It is important to analyze why the polygon-assisted compression scheme results in a better compression ratio than the image-based schemes such as JPEG and MPEG-1. As remarked in [9], it is not possible to conduct objective, quantitative comparison of these schemes, since they are lossy compression schemes, and preserve different qualitative aspects. Therefore, any comparisons should be inherently subjective.

We view the effect of polygon-assistance from the following perspective: How much information does the client succeed to generate that simplifies the restoration of the rendered images of high-quality models. The client uses low-quality geometric models to generate approximate images. There is a tradeoff between the complexity of the used model and the fidelity of the approximation. Broadly speaking, we can measure the complexity by the size of the model's geometric description, and the fidelity by the entropy of the generated approximate image. The differences among the three partitioning strategies seem to suggest that geometric models that consist of large, untextured polygons will give a large fidelity/complexity ratio. We will call such models as *simplified geometric models* whose fidelity/complexity ratio is larger than a predetermined complexity limit.

Simplified geometric models can be derived through coarse tessellation or mesh optimization from high-quality geometric models [3][5]. The model derivation should be based on the available computational resources on the client computer, e.g., how many polygons can be processed for each second. As a general principle, the more information the client successfully generates, the less data the server needs to send over the network. We will show that it is possible to increase the amount of client-generated information at a reasonable additional computational cost. In particular, by predicting the motion in the sequence of difference

学術情報センター電子図書館サービス ◎学術情報センター

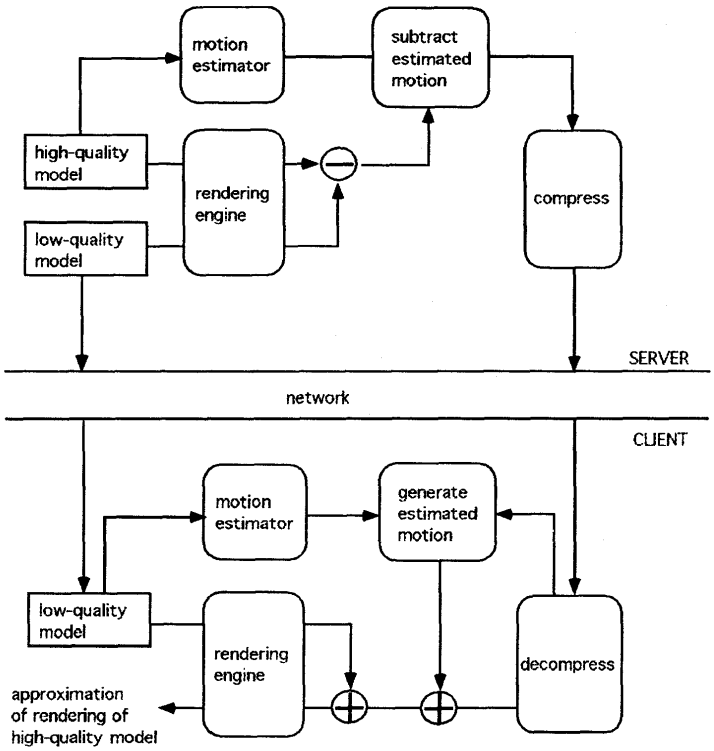


Fig.2 Data flow diagram of the revised compression scheme

images, we can significantly reduce the amount of necessary network traffic.

4 Revised Compression Scheme

We can improve the basic compression scheme in two independent ways. First, the low-quality geometric model can be chosen not to exceed the complexity limit described in section 3.2. Secondly, the interframe coherence can be used to further compress the sequence of difference images. In subsequent sections, we will discuss only the latter type of improvement.

4.1 Estimating Interframe Coherence

Optical flow is the 2-dimensional vector field that represents the motion of pixels in a sequence of images. It describes how a pair of consecutive image frames are related to each other. Optical flow has been extensively used in computer vision to recognize moving objects [7].

Computing the optical flow generally requires expensive search algorithms to match corresponding pixels. However, when dealing with rendered images of dynamic polygonal scenes, it is possible to accelerate the optical flow computation with hardware assistance. Wallach et al. described such a method that employed the matrix-vector multiply and scan conversion hardware installed on many today's graphics workstations [13]. The optical flow was then used to accelerate the motion compensation component of MPEG-1.

Although the complete optical flow can be computed only by evaluating the high-quality renderings, an approximate, partial flow can be computed by evaluating the low-quality geometric models using the Wallach's method. In the client-server diagram described in Section 3, this implies that both the server and client can obtain the same approximate flow information with some additional computation. Therefore, the server needs to send only the differences between the approximate and complete flows.

Fig.2 illustrates a revision of Fig.1 based on the above observation. On the server, the difference image D_1 between high-quality rendering H_1 and low-quality rendering L_1 is first computed for frame 1. The approximate optical flows $F_2,...,F_N$ are then computed using the low-quality geometric model for frames 2,...,N from frame 1. Combining the difference image D_1 and flows $F2,...,F_N$ will yield an image sequence $D'_2,...,D'_N$, which approximates $D_2,...,D_N$. Finally, the error components $E_i=D_i-D'_i$ ($2 \le i \le N$) are computed and transmitted to the client in a compressed form using MPEG-1. The client receives sequence of images $D_1,E_2,...,E_N$, computes $F_2,...,F_N$ using the low-quality geometric model, and

学術情報センター電子図書館サービス　　©学術情報センター

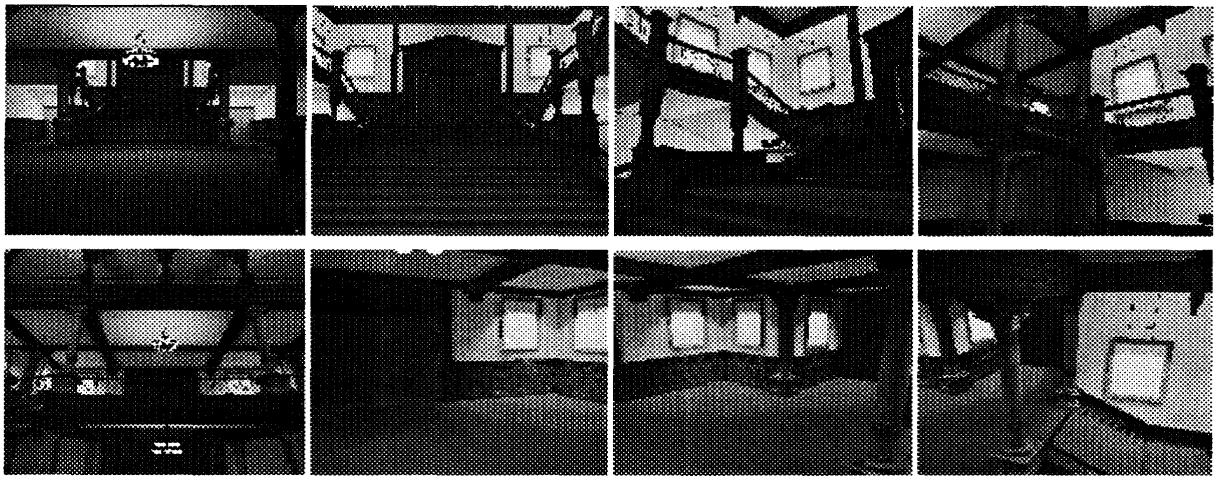


Fig.3 A sample synthetic movie for which the revised scheme is evaluated

obtains $D_1,D_2,...,D_N$ through reverse computation. It then combines the difference images $D_1,...,D_N$ with low-quality renderings $L_1,...,L_N$ to yield high-quality renderings $H_1,...,H_N$.

We have evaluated the effect of the proposed revision by experimenting on several sample synthetic movies. Fig.3 shows a series of thumbnail images of a sample movie. All the sample movies are also available at http://www.rd.nacsis.ac.jp/~gotoda in MPEG-1 format. For these movies, the revised scheme resulted in about 30% improvement over the basic scheme in the compression ratio. It should be mentioned, however, that more extensive study is necessary to draw a general conclusion on the degree of improvement introduced by our scheme.

4.2 Bidirectional Motion Compensation

The method described in Section 4.1 was unidirectional in that it computed $D'_2,...,D'_N$ using the estimated motion only from the preceding frame 1. It is easy to extend it to a bidirectional scheme, which computes $D'_2,...,D'_N$ based on the estimated motions from frames 1 and N+1.

This bidirectional compensation scheme is similar to the B-frame concept in MPEG-1 [8]. The differences are: (1) in MPEG-1, motion compensation is done on macro blocks, each consisting of 16x16 pixels, while our method approximates motion of each pixel; and (2) blending between forward and backward compensation can be performed more sophisticatedly in our method based on the feasibility information, e.g., motion information derived from polygons nearly vertical to the screen is less feasible than that from parallel ones. Such feasibility information is obtainable through evaluating the simplified geometric models.

5 Applications

We can apply the revised compression scheme to various domains, e.g., to reduce storage requirements, and to reduce the usage of network bandwidth. In the following subsections, we will outline two examples of network-oriented applications where the proposed scheme is considered useful.

5.1 Graphics Model Server

Sharing expensive graphics workstations among cheap personal computers via low-speed networks was what Wallach et al. and Levoy intended in their pioneering papers [9,13]. Since such workstations are typically equipped with high-performance multi-CPU system, it is natural to extend this framework to sharing their CPUs as well. For instance, high-performance graphics workstations will perform all operations (transformations, deformations, etc.) on graphics models, and low-end PCs will receive the results.

Being designed on the assumption that the low-quality geometric models have to be transmitted over the network, our scheme will enable such usage of graphics workstations more easily than previously proposed ones. The graphics servers can also be connected with other types of servers through high-speed networks to form a workstation cluster, which as a whole responds to various types of requests submitted by the clients.

学術情報センター電子図書館サービス ◎学術情報センター

5.2 Progressive VRML

The rapidly growing popularity of WWW has led to several problems in wide-area networking. For instance, many international lines are suffering from ever increasing traffic congestion. On the other hand, recent network-oriented applications, such as video-on-demand (VOD) systems or virtual reality (VR) systems [1], will require to transfer large data streams in real time, which is becoming more difficult to achieve. One possible solution to alleviate this problem is to transfer such data progressively while controlling their quality.

Our compression scheme can be extended to support the progressive mode, where levels of details are gradually increased. The quality control over the streams of difference images can be accomplished using MPEG-2, while geometric models are also transmitted progressively using the progressive mesh method [6]. The progressive VRML realized in this way can reduce the access latency, which provide more comfortable interface to the users operating at low-end workstations.

6 Concluding Remarks

We have proposed a scheme for compressing synthetic movies based on the combination of geometric and image-based coding, and on the approximate motion prediction using simplified geometric models. Compared to other compression schemes, the proposed one better utilized inter-frame coherence through motion compensation using low-quality geometric models. The compensation required small additional computational cost, but considerably reduced the storage requirement. The computation can be accelerated using the 3D rendering hardware present in many today's graphics workstations.

Many directions for future research were suggested in [9], most of which still seem to remain unaddressed. Here, we will only point out two other possibilities. One is to apply the proposed scheme to progressive or scalable transmission of multiresolution synthetic movies. Whether natural scene movies can be compressed more compactly by our scheme than MPEG may also be an interesting issue to be addressed again in the future.

References

[1] Chen, S., "QuickTime VR－An Image-Based Approach to Virtual Environment Navigation", *Proc. SIGGRAPH '95*, ACM SIGGRAPH, pp. 29-38, 1995.

[2] Deering, M., "Geometry Compression", *Proc. SIGGRAPH '95*, ACM SIGGRAPH, pp. 13-20, 1995.

[3] Eck, M.; DeRose, T.; Duchamp, T.; Hoppe, H.; Lounsbery, M.; Stuetzle, W., "Multiresolution Analysis of Arbitrary Meshes", *Proc. SIGGRAPH '95*, ACM SIGGRAPH, pp. 173-182, 1995.

[4] Guenter, B.; Yun, H.; Mersereau, R., "Motion Compensated Compression of Computer Animation Frames", *Proc. SIGGRAPH '93*, ACM SIGGRAPH, pp. 297-304, 1993.

[5] Heckbert, P.; Garland, M., "Multiresolution Modeling for Fast Rendering", *Proc. Graphics Interface '94*, Canadian Information Processing Society, pp. 43-50, 1994.

[6] Hoppe, H., "Progressive Meshes", *Proc. SIGGRAPH '96*, ACM SIGGRAPH, pp. 99-108, 1996.

[7] Horn, B., "Robot Vision", MIT Press, 1986.

[8] Legall, D., "MPEG－A Video Compression Standard for Multimedia Applications", *CACM*, Vol. 34, No. 4, pp. 46-58, 1991.

[9] Levoy, M., "Polygon-Assisted JPEG and MPEG Compression of Synthetic Images", *Proc. SIGGRAPH '95*, ACM SIGGRAPH, pp. 21-28, 1995.

[10] Lewis, A.; Knowles, G., "Video compression using 3D wavelet transforms", *Electronics Letters*, Vol. 26, No. 6, pp. 396-398, 1990.

[11] Radecki, S., "Multimedia With Quicktime", *Academic Press*, 1993.

[12] Wallace, G., "The JPEG Still Picture Compression Standard", *CACM*, Vol. 34, No. 4, pp. 30-44, 1991.

[13] Wallach, D.; Kunapalli, S.; Cohen, M., "Accelerated MPEG Compression of Dynamic Polygonal Scenes", *Proc. SIGGRAPH '94*, ACM SIGGRAPH, pp. 193-196, 1994.

学術情報センター電子図書館サービス　　© 学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

AODBMS: Phasme and its Image Retrieval Plug-ins Implementing shape-based retrieval processing

AODBMS: Phasme とその形状検索処理機能を持つ画像検索プラグイン・モジュール

Frederic ANDRES
National Center for Science Information Systems
学術情報センター　フレデリック　アンドレス
Christopher CLAUDE
Tokyo University
東京大学　クリストファー　クローデ
Kinji ONO
National Center for Science Information Systems
学術情報センター　小野　欽司

ABSTRACT

We describe Phasme, an Application-Oriented database system, designed both to provide high performance and to meet the needs and the requirements of today's multimedia system as an opened and customisable DBMS.

Phasme is an opened database system using the Extended Binary Graph(EBG) data structure and employing both inter operation and intra-operation parallelism. Phasme is fully customisable from the data type to the execution model. It provides a dynamic plug-ins mechanism and a set of many-sorted algebra to efficiently support different types of data structures.

We discuss how these features impacted the design and the implementation of shape-based image retrieval applications. We show how Phasme features provide both challenges and opportunities in building a high-performance image retrieval system on top of it. In particular, we show how a novative shape-based image retrieval method can be mapped into Phasme's Extended Binary Graph. Capitalizing from previous works as archetype transformation, we propose a solution to support similarity image retrieval processing.

要旨

本稿では、筆者らが開発を進めている「Phasme」と呼ばれるアプリケーション指向のデータベースシステム(AODBMS)について述べる。Phasme は高いパフォーマンスを提供し、オープンかつ柔軟な DBMS に必要とされる今日のマルチメディアシステムの要求条件を満たしている。

Phasme は、データ構造として Extended Binary Graph(EBG)を使用し、インターオペレーション(inter-operation)とイントラオペレーション(intra-operation)の並行性を利用したデータベースであり、データタイプから実行モデルまで、完全にカスタマイズ可能である。

また、異なったデータ構造を効率的にサポートするために、動的プラグイン(Plug-ins)機構や多ソート代数を提供している。さらに、これらの機能を kernel に持たせているので、商用の汎用 DBMS に比べて、カスタマイズしやすいのが特徴である。

本稿では、これらの特徴が画像検索アプリケーションの設計や実装にどれだけインパクトを与えるかを論じ、Phasme 上の高性能なイメージ検索システムの構築に際して有効なことを示す。

本文では、革新的な形状の画像検索方式が Phasme の Extended Binary Graph(EGB)にどのようにマッピングされるかを具体的に示す。またアーキタイプ変換を利用して、類似イメージ検索処理に対する解決策を示す。

学術情報センター電子図書館サービス ◎学術情報センター

AODBMS: Phasme and its Image Retrieval Plug-ins Implementing shape-based retrieval processing

[Keywords] AODBMS, Image retrieval, Image database
[キーワード] アプリケーション指向 DBMS、画像検索、画像データベース

1 Introduction

In recent years, consensus has been reached in the Image Retrieval Community about the importance of the content-based processing when pattern similarity is needed. In the field of museums, shopping catalog or photo banks, information of the image content items are captured by query by content-based search techniques which act as filtering mechanisms. Queries containing pattern similarity search can be formulated in SQL and be optimized. Several methods are available even in commercial product such as QBIC[15].

Storage and querying of images shapes in an extensible database system still pose severe performance challenges to current database technology. An image database used in a multimedia retrieval system is typically large of size (image banks, digital librairies or hypermedia information) and archival nature where deletions and updates are rare.

The large image volumes involved require the DBMS to use very efficient resource management strategies. Important aspects are I/O, multi-level cache performance, memory management and parallel processing. The optimization of image retrieval processing requires an efficient use of image access method to reduce the number of the image accesses.

In this article, we present Phasme previously named Chameleon[1], an Application-oriented database server, intended to satisfy user requirements and information system needs in various application domains. It has already achieved successes as an object manager[2], a Delivery Information System[3] and for supporting video on demands[4,5]. We discuss how the features of Phasme provide opportunities to support shape-based image retrieval (SIR). A major challenge of SIR technology is to determine a minimal set of features that describes the shape of the image which admits similarity measure. Many features are available from the fields of data analysis - for example archetype describing the shape. These features are hyperplan feature vectors.

The remainder of this paper is organized as follows. Section 2 provides an overview of the architecture of Phasme. Section 3 reviews major related works in shape-based image retrieval. In Section 4, we describe the shape-based image retrieval processing inside Phasme. Section 5 describes the implementation to support shape-based retrieval processing. Section 6 presents the conclusion.

2 Architecture of Phasme

Phasme is a novative database kernel under development since 1993. Its development is based on both our experience gained in building DBS3[6], a parallel ERDBMS and on the current market trends in database server technologies.

Given this background and the requirements of the users, we applied the following ideas in the design of Phasme:

- Extended Binary Graph data structure. Phasme combines three strong concepts: DBgraph[16], Decomposition Storage Model[17], and the physical feature of the Graph Data Model[12]. The Extended Binary Graph structure (see Figure 1), is a graph between oid and values or between oid and oid(OID1-OID relationship in Fig.1). The DBgraph avoids value duplication for complex data type. The Decomposition Storage Model introduces clustering and vertical partitions. It provides efficient caching management when it is combined to horizontal partitions. The Graph Data Model facilitates hierarchy relationship between objects.

  Figure 1 shows how data are stored in EBGs. The left column is referred to as *source*, the right column as *destination*. An EBG is a set of no-oriented arcs between items which are either oids or values. EBGs contain either fixed-size item values or variable-size item values.

  Persistency is managed in an orthogonal way from the data structure. So persistent index or persistent data structures are stored directly inside EBGs.
- Customisable and architecture. Phasme has been build as an opened database manager to reach user and information system requirements. This system provides full customisability features from data type definition to execution model. This vertical customisability is also available for data operations,

学術情報センター電子図書館サービス　　©学術情報センター

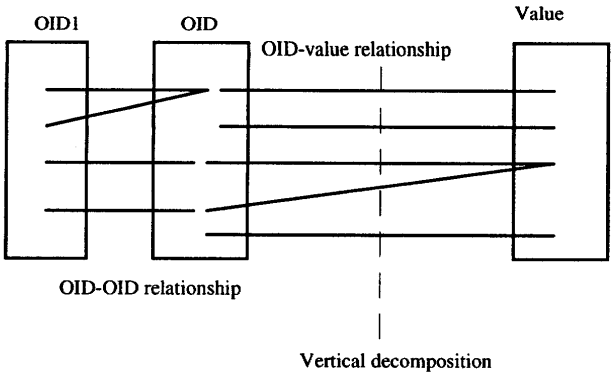


Fig.1 Phasme's Extended Binary Graph(EBG)

physical data structure, and query optimization. All the architecture of Phasme work on this assumption. Phasme lets the application specify new data structure and data management strategies according to its own needs or requirements. The EBG data structure pushes the semantic data model out of the kernel. Due to this feature, Phasme defines a novative class of DBMS called Application-Oriented Database System (AOD-BMS). Each application which accesses Phasme provides its own semantic to the data. It could be either traditional data semantics (e.g. relational model, object model) or exotic data semantics.

- Many sorted algebra. Phasme architecture is based on a set of languages based on the concept of many sorted algebra[11]. The set of many sorted algebras provides Phasme extensibility. Their nature allows for custom-isability requirements the addition of new item types, operations and data processing optimization on thoses types. Also SQL or OQL query calculus are supported to process many-sorted algebra transfor-mation as a framework where dynamic query optimization and parallel query processing are efficiently managed. Phasme's Interface Language (PIL) is based on an EBG many-sorted algebra. PIL as a Plug-ins Language allows the user to plug-ins modules at runtime. Those plug-ins are used to implement applications' requirements and their needs inside the kernel. They can contain new item types, their associated data structure with specific indexes, optimization rules or execution models.
- Inter and intra-operation parallelism. Parallelism feature is vertically incorporated inside Phasme from the PIL level to the execution model level. Parallel query processing integrates both inter and intra-operation parallelism. The parallelism mechanism uses the multi-threading functions of the operating system.

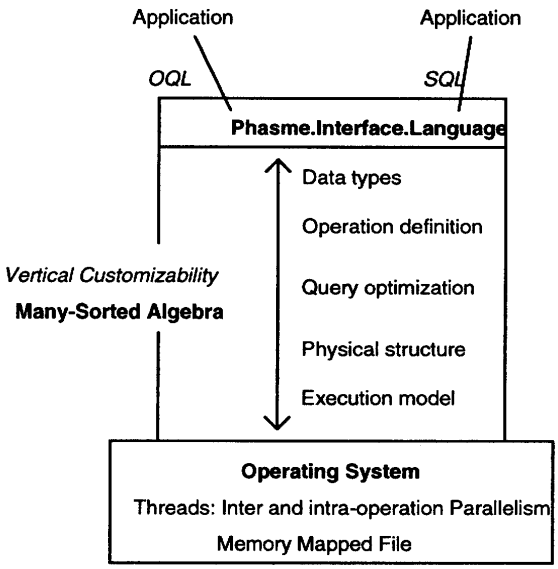


Fig. 2 Phasme server

- Uniform data format. Phasme uses a uniform data format on disk and in memory. The 0 memory mapped file technology is used to provide high speed data retrieval.

3 Shape-based Image Retrieval Processing

The problem of shape recognition in the context of query by content of image is both very important and very difficult. One usual wants to design a scheme where the user specifies a shape (by drawing it or extracting it from an example image) and subsequently asks for images containing "similar" shapes [14].

To cope with this problem one usually tries to associate a metric to the space of shapes. This can be done either directly, by defining a distance in the space of shape, or indirectly by extracting first the set of numbers (called features) defining a multidimensional vector in a presumably Euclidean space. With this approach, two shapes are considered "similar" if they are close with respect to the associated distance [15][13].

However, as pointed out by D.Mumford in [13], this approach does *not* fit with the psychological properties of similarity. Very interesting psychological experiments [14][19] have shown that, from a human point of view:

1) The concept of similarity is *asymmetric*, that is the

学術情報センター電子図書館サービス ◎学術情報センター

statement "A is similar to B" is not equivalent to "B is similar to A", more precisely the degree of similarity is not the same in both cases.

2) The similarity of two objects is strongly *context* dependent, that is A may be considered similar to B if both objects are to be paired (taking into account 1) among a given set; but in another set, A will be find more similar to a different object.

To take 1) into account in the problem of shape matching Mumford ([12]) proposes the following interpretation: if a given shape B, to be matched, has "outstanding" features which makes it *prototypical* of a class of shapes then the mind emphasizes on those features. Now if the input shape A has not these particular features then B will rapidly be recognized as not being close to it. But the reverse is not necessarily true as A has no prominent features. It is also clear that the outstanding feature for a particular shape depends on the set to which it belongs. A rounded form in a set of various squared form is outstanding but a rounded form is not outstanding in a set of other rounded form. This is consistent with the point 2.

This point of view suggests to use a feature extraction procedure which describes a *set* of shape by a given number of *prototypical shapes* and each particular shape as a combination of those prototypes. The solution we propose hereafter is a step towards that goal.

The mathematics underlying this idea is called archetypal analysis and was developed for a different purpose by Cutler and Breiman [17][7]. We will now briefly describe the key ideas of this approach. Taking the same formalism as in [16], we consider the following set of n elements.

$$\{x_i, i = 1,..,n\},$$

where each element is a m dimensional vector (T stands for Transpose)

$$x_i = (x_{1i},..x_{mi})^T$$

we then seek a set of p vectors $Z_1 \cdot \cdot Z_p$ (of dimension m), called archetypes, satisfying the following constraints

$$\hat{x}_i = \sum_{k=1}^{p} \alpha_{ik} z_k , \quad \alpha_{ik} \geq 0 , \quad \sum_k \alpha_{ik} = 1$$

$$z_k = \sum_{j=1}^{n} \beta_{kj} x_j , \quad \beta_{kj} \geq 0 , \quad \sum_i \beta_{ki} = 1$$

where $\hat{x}$ is an approximation of x, and minimizing the error

$$\varepsilon_p = \left\langle \left\| x_i - \hat{x}_i \right\| \right\rangle$$

In our problem, if we take x to be the set of shapes then the mathematical properties of archetypes, quite unlike other decomposition schemes like Karhunen-Loeve for instance, makes them suitable to be considered prototypical shapes:

1) Most importantly, the experience proves that archetypes closely *look like the data* (x). This can be partially understood from eq {} which shows that each archetype is a linear combination, with positive components, of shapes belonging to the set. Moreover, the sum condition on the betas insures an homogeneity of scale between archetypes and data.

2) It is obviously a *data dependent* decomposition scheme. Here this property is however not intended to reduce the error in the representation, rather it is a way to take the context into account.

3) They are not an orthonormal basis, and as such their number is not limited by the dimension m of the vectorial space in which we express the shapes. Their number can be chosen only on the basis of the variety of forms found in the set.

4) There is *no natural ordering* for the archetypes. Therefore the alphas can be considered components of a feature vector and those components play an equal important role in the description of the shape. This provides us an euclidean space which allows the use of powerful indexing techniques inside the database kernel [10].

It should be noted that those properties have a natural, and therefore unavoidable, computational drawback: if we already have a set of n shapes described by p archetypes then incrementing n or p leads to a complete recalculation of the new set of archetypes.

We suggest the following scenario for shape matching: Lets take a set of shapes, then calculate the archetypal decomposition for all but one shape, the "stimuli" . Then we project this stimuli on the archetypes (this is not unique however and one has to choose a way to do it) and use the Euclidean distance to decide whether or not this shape is close to any given shape of the set. In that context, calculating the archetype means extracting the outstanding

学術情報センター電子図書館サービス ©学術情報センター

features from the set. The subsequent matching helps to decide whether or not the stimuli has those features.

We can do so for another shape of the set. The archetypes will be different and therefore the Euclidean distance as well. If the number of shape is small this could introduce the asymmetry mentioned by [13]. However, if the number of shapes is large, the asymmetry is likely to be weaker, unless one shape is clearly different from the rest.

4 Shape-based Image Retrieval Processing inside Phasme

Using the Phasme database kernel to support shape-based image retrieval applications foreseen poses the following challenges on its design and implementation:

- efficiently dealing with huge image data, which keeping high quality on data delivery.
- multi-dimensional indexing that works for large, virtual memory-based databases.
- response time should be fast based on dynamic feature indexing.

Our approach is to use Phasme as a flexible and an application-oriented DBMS where the data format on disk is the same as the data format in memory. The semantic of the data is given by the application.

In order to implement the previously described scenario for shape retrieval, the archetype algorithm is a loop which iterates successively on the alpha and the beta until a prescribed error is obtained. Then we just run one iteration of the algorithm to determine alpha for our new shape. Obviously to do so, we must first load all the apha and beta for each shapes of the set, but this overhead is done only once during a session and is valid for all subsequent match. Once alpha is determined, we use the usual Euclidean distance to decide whether or not this shape is close to any given shape of the set.

5 Phasme implementation of the Image Shape-based Retrieval Processing

5.1 Extension Plug-ins

In our implementation, we have chosen to represent a shape by two vectors, one for each axis (in cartesian coordinates). The origin is taken at the center of gravity and the first point is the one with the smaller vertical coordinates and zero as horizontal one. Then the result is normalized in scale.

5.2 Test Configuration

In order to test the archetype decomposition we have chosen a set of car images because all entries are obviously related but at the same time there are clearly several type of cars to be decribed with archetypes. This is a typical situation in which archetype decomposition is expected to be efficient [7]. Compared to the type of cars, a larger number of archetypes is required to have a fine description of cars among a sub categorie, however, as all items are cars that we expect general features to appear here. Thereore we should end up with less archetypes than the number of type of cars times the number of sub features. The correct number is to be determined experimentally, by trial and error method.

6 Conclusion and the future study

In this article, we described how Phasme, an operational prototype, can be used to support large image data sets efficiently. We discussed how its Extended Binary Graph Data structure, and customizability provide a mechanism for shape-based similarity retrieval processing.

Phasme's memory management based on the memory-mapped file concept and the Extended Binary Graph (EBG) data structure provide the opportunity to manage efficiently contour-based image retrieval. The memory mapped file concept avoids the overhead for image retrieval introduced by traditional DBMSs. It gives an uniform data format for memory and disk data storage and manipulations. We produced an image shaped retrieval system by integrating the archetype transformation and shaped-based data structure inside Phasme kernel. A performance evaluation is in process in order to show the effectiveness of our approach.

Acknowledgments

We would like to thank Professor A Cutler (Utah State University) for her help and for her support about the archetype algorithms.

References

[1] Andres, F. ; Boulos, J. ,"Chameleon: A Parallel Application - Oriented Database System", *IFATEC report*, No. 941201, 1994.

学術情報センター電子図書館サービス ©学術情報センター

[2] Andres, F. ; Ono, K. ,"A High Efficient Multimedia DBMS in Object Management", Research Bulletin of the National Center for Science Information Systems, No8, March 1996, pp 301-315.

[3] Andres, F. ; Boulos, J. ,"DDS The Data Delivery System", *in IFIP 1996 proceeding, Advanced Intellegent Tool*, Canberra, Australia, 1996.

[4] Andres, F. ; Ihara, K. ; Boulos, J. ; Ono, K. ; Yasuda, Y. ,"The OLVP (OnLine Video Processing) System", *in Proc. Int. Workshop DMS96*, Hong kong, June 1996.

[5] Andres, F. ; Boulos, J. ; Ihara, K. ; Ono, K. ; Yasuda, Y. ,"Performance Evaluation of the OLVP System", *in Proc. Int. COMPSAC*, Seoul, Korean, 1996.

[6] Bergsten, B. ; Couprie, M. ; Valduriez, P. , "Prototyping DBS3, A Shared Memory Parallel Database System", *in Proc. PDIS 1991*.

[7] Cutler, A. ; Breiman, L. , "Archetypal Analysis", *Technometrics 36*, p.338, 1994.

[8] Gray R. S. ,"Content-based Image retrieval: Color and Edges", *technical report*, Dartmouth College, 1996.

[9] Guting R. H. ,"Second Order Signature: A tool for specifying data models, query processing and optimization", *in Proc. ACM SIGMOD*, May 1993.

[10] Jagadish, H. V. ,"A Retrieval Technique for Similar Shapes", *in International Conference on Management of Data, SIGMOD 91*, Denver, CO, ACM, May 1991, pp.208-217.

[11] Kunii, H. S. , "Graph Data Model and its Data Language", Springer-Verlag, Tokyo, 1990.

[12] Mumford, D. , "The Problem of Robust Shape Descriptors", In proc. of the IEEE first international conference on computer vision", London, England, June 8-11, 1987, pp602-606.

[13] Mumford, D. ,"Geometric Method in Computer Vision", Vol. 1570, SPIE, pp2-10, 1991.

[14] Mumford, D. ; Herrnstein, R. ; Kosslyn, S. ; Vaughan, W. ,"Analysis and Synthesis of Human and Avian Categorization of 15 Simple Polygons", preprint Harvard Univ. Dept. of Psych., 1989.

[15] Faloutos, C. ; Equitz, W. ; Flickner, M. ; Niblack, M. ; Petkovic, D. ; Barber, R., "Efficient and Effective Querying by Image Content", *IBM report*, 1995.

[16] Smith, J, R. ; Chang, S, F. , "Local Color and Texture Extraction and Spatial Query", *in IEEE International Conference on Image Processing*, 1996.

[17] Stone, E. ; Cutler, A., "Archetypal Analysis of Spatio-temporal Dynamics", Physica D , 90, 1996 ,pp.209-224.

[18] Thevenin, J.M., "Architecture d'un Systeme de Gestion de Bases de Donnees Grande Memoire", Ph.D. Thesis of Paris VI University, 1989.

[19] Tversky, A. ,"Features of Similarity", Psychological Review, 84, 1977, pp.327-352.

[20] Valduriez, P. ; Khoshafian, S. ; Copeland, G. , "Implementations techniques of Complex Objects", *in Proc. of the International Conference of VLDB*, Kyoto, Japan, pp.101-110, 1986.

Appendix

Image items:

Name	Data type
Id	text
Abstract	text
Title	text
Comments	text
Image data	image
Date	date
Representability rate	float
Normalized shape	set of float
Archetype coefficient	float

Archetype items:

Name	Data type
Id	text
Confidence rate	float
Point Cardinality	integer

学術情報センター電子図書館サービス　　© 学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

単語間の係受け情報を用いた文献検索手法

A Method of Document Retrieval using Dependency Relationships between Words

東京大学大学院工学系研究科　池田　和幸
Kazuyuki IKEDA
Graduate School of Engineering, The University of Tokyo
学術情報センター　髙須　淳宏
Atsuhiro TAKASU
National Center for Science Information Systems
学術情報センター　安達　淳
Jun ADACHI
National Center for Science Information Systems

要旨

キーワード検索は現在文献検索手法の主流を占めているが、出現単語とその統計的特徴量のみで検索要求に適合した文献を選別せねばならず、その限界は明らかである。そこで我々は、文献の表題やキーワードの検索においてそれらの構成単語間に成立する係受け関係に着目した検索を行うことで、言語の持つ情報を活用し検索精度を高めることを提案する。

本手法では、事例分析に基づいた文の係受け解析及び単語 2-gram を用いた複合名詞の係受け解析を通じて文献情報を二分木の形で構造化し検索インデックスとしている。問合せは文献表題と同様の構造を持った擬似的な自然言語で発行するため、検索者は検索言語を意識することなく検索要求を問合せに反映することができる。本稿ではこれらの手法を説明するとともに、プロトタイプシステムを用いた検索実験を通じて本手法の有効性を検証する。

ABSTRACT

Although keyword-based retrieval occupies the major position in document retrieval methods, it has obviously limitation in retrieval effectiveness because it is difficult to select documents that coincide with user's requests using only query words and its statistical characteristics. Thus we propose a method to enhance retrieval effectiveness using language information such as dependency relationships between words in retrieving documents' title and keywords.

In our method , we make structured index represented with binary tree through modification analysis by using case rules and compound nouns analysis which employes word bigram. Queries are issued by psuedo natural languages such as expressions found in documents' title, so users can reflect their requests on query without being concerned about query languages. In this paper, we describe these methods and inspect an effectiveness of our methods through retrieval experiments on our prototype system.

[キーワード]　情報検索、係受け、自然言語、構造化インデックス、CGI、複合名詞解析、単語 2-gram
[Keywords]　Information Retrieval, Dependency Relationship , Natural Language, Structured Index, CGI(Common Gateway Interface), Compound Nouns Analysis, Word Bigram

学術情報センター電子図書館サービス　　©学術情報センター

1　はじめに

文書の電子化が進むにつれて、文献データベースの果たす役割は日増しに大きくなってきている。また、安価で高性能な計算機やインターネットの普及によって誰でも気軽にデータベースサービスにアクセスできるようになることが予測されるため、

- 検索式を意識せず、頭の中にある検索意図を自然な形で表現するだけで
- 大量の情報の中から欲しい情報だけを的確に抽出できる

ことが今後重要になってくると思われる[1]。現在の文献データベースで主流となっているキーワードを用いた論理検索では、残念ながらこれらの条件が満たされているとは言い難い。問合せを発行する際には、検索者は検索意図からキーワードを選び出して論理式の形でシステムに提示しなければならない。また、検索の結果得られた出力文献の間に順位付けがなされていないため、最悪の場合には検索者が大量の出力結果の中から自分の目を頼りに所望の検索結果を探さねばならない。近年ではベクトル空間モデルや TF-IDF(Term Frequency Inverted Document Frequency)法などを用いた得点づけ[2]をする検索エンジンもある。代表的な例としては分散情報システムとして有名な freeWAIS-sf + SFgate[3]を挙げることができる。しかしながら、この場合でもあくまで得点は出現単語と単語の出現頻度等統計的特徴量のみに依存するため、その正確性には限界があると言わざるを得ない。

一方で、これらの限界を自然言語処理技術を用いて打破してしまおうという研究が近年盛んである[4]。しかしながら実用に耐え得る自然言語処理技術は形態素解析のみであると言える。構文解析では言語の自由度の高さによる曖昧性の処理が課題であるし、意味解析、発話解析となるとまだ研究の域を出ているとは到底言い難い。しかしながら、そのような自然言語処理技術も適用の場を限定すれば高い精度を出すことが可能である。このようなアプローチの研究としては、格関係に着目した、名詞と動詞の係受けを利用した検索[5]や、医療分野に特化した概念検索エンジン[6]などを挙げることができよう。

このような状況を踏まえて我々は、文献情報の中でも比較的文の構造の限定された表題やキーワードを対象として、それらの構成単語間の係受け関係を検索に活用することを提案する。また、検索文として表題に見られるような擬似的な自然言語を用いることで、検索時の検索者の負担を軽減することも考える。本論文では、2 章で現在の文献検索システムにおける諸問題について触れ、3 章でそれらを解決するために我々の取るアプローチについて説明する。4 章で係受けを付与したインデックスの作成法について述べ、5 章でそのインデックスの検索手法を説明する。6 章、7 章で本システムの採用した得点付け手法とそれによる出力結果を紹介し、最後に今後の課題と結論について述べる。

2　現在の文献検索システムにおける諸問題

検索者にとって使い勝手が良くかつ検索効果の高い検索システムを実現するためには、現在以下のような問題点があると考えられる。

2.1　問合せとその意味的曖昧性

キーワードベースの検索では、検索者は検索意図からキーワードを選び出して論理式で問合せを発行せねばならず、さらに検索結果を見ながら新たな問合せを発行して検索結果同士の論理積を取るなどの複雑な作業を必要とする場合が多い。これは慣れない検索者にとっては大変なことであり、結果として検索意図が十分に検索システムに伝わらないこともある。また、検索意図を論理式に変換した時点でキーワード同士の意味関係が失われるので、問合せ自体にも意味的な曖昧性が残る。

2.2　文献内容空間の縮小、差異

現在の文献データベースでは、文献から表題やキーワード・アブストラクトなどの主題要素を抽出し、それらを用いて検索インデックスを作成している。従って文献データベース自体が検索対象文献の主題を十分に保持しているとは言い難く、実際にそれによる検索漏れが問題になっている。この解決としては、全文データベースや文書要約[7]などが試みられている。また、個々の検索者によって検索意図の表現に違いがあることに起因して生じる検索者の文献空間と検索システムの提供する文献空間の間の差異も検索においては問題となる。

2.3　文字列検索の不正確性

論理検索ではキーワード文字列の一致/不一致がマッチングの条件となる。そのため、例えば「新聞の記事」「新聞に対する記事」という 2 つの異なる意味を持つ文は「新聞」「記事」という 2 つのキーワードを持っ

学術情報センター電子図書館サービス　　© 学術情報センター

た検索文によって検索されてしまい、それらの文の持つ意味の違いは検索結果に全く反映されない。また、日本語のように単語と単語の間が明確に区切られていない言語の場合は、複数の単語に跨った文字列が検索文字列と一致したり、複合名詞が部分一致で検索されたりすることによる検索精度の劣化の問題もある。

2.4 類義語・類義文に対する検索漏れ

現在のキーワードベースの検索では、同じ意味を表す別の語や文を検索できない。類義語の検索については類義語辞書やシソーラスを用いた解決が指向されており、キーワードをシソーラス展開して検索を行うシステムも実際に存在する。しかしながら辞書やシソーラスの保守や維持は問題点として残されたままである。類義文の検索については、本格的な解決は自然言語処理技術のうちの意味解析以降の成熟を待たねばならない。

3 提案する文献検索システム

以上の情報検索における諸問題のうち、我々は検索者の発行する問合せと検索精度の問題に焦点を当て、その解決を試みる。検索精度の向上のために、我々は単語間の係受け関係を検索に利用することを提案する。単語間の意味関係を用いることで、キーワードベースの検索において存在していた意味的な曖昧性が排除され、検索効果が向上することが期待できる。また、問合せとしては、文献表題に見られるような、構造の限定された擬似的な自然言語を用いることを提案する。これは検索者にとっては日常用いている言語であるため、検索意図をスムーズにシステムに伝えることができると思われる。このようにして、我々は1章で述べたような

- 検索式を意識せず、頭の中にある検索意図を自然な形で表現するだけで
- 大量の情報の中から欲しい情報だけを的確に抽出できる

検索システムを実現することを志向している。提案する検索システムは、図1のように示すことができる。

検索システムは、文献の表題とキーワードを検索対象文として用いる。これらの文は

- 自然文に比べて構造が単純である。具体的には名詞と付属語が繰り返し出現する構造を有する。
- 文献の主題を端的に表しており、かつ抽出が容易である。

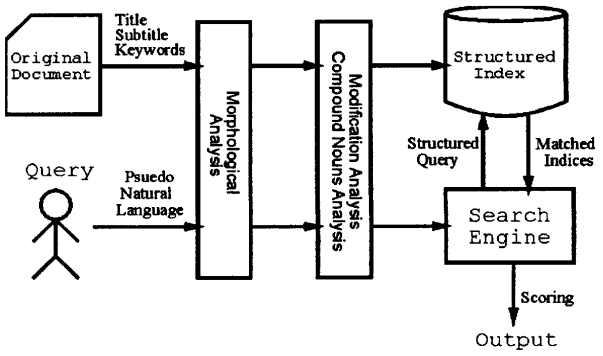


図1　提案する検索システムの全体像

という特徴を有している。構造の単純さゆえに係受けの付与が自然文と比べると容易であり、また文献の主題を端的に表すため文献の持つ情報を代表させるのにも相応しいと考えられる。本研究では、学術情報センターの提供するデータベース NACSIS-IR からデータベース分野の論文1490件を抽出し検索対象文献としている。

これらの文は、4章で説明する形態素解析、係受け解析、複合名詞解析を通じて文の構成単語間に係受けを付与され、最終的には係受けの二項関係を表す二分木に構造化されデータベースに保存される。このように表現された検索インデックスを構造化インデックスと呼ぶことにする。構造化インデックスは図2に示すように、二分木の葉の部分に文を構成している単語、内部ノードに単語同士を結ぶ付属語及び付属語の属するカテゴリを配するデータ構造を有している。単純に全単語間の係受けを二項関係で持つ場合と比べ、構造化インデックスは係受けの階層性を表現できる点で記述力が高い。図2を例にとると、「日本語データベースにおける情報検索→研究」という名詞句同士の係受けの下に「日本語データベース→情報検索」という複合名詞同士の係受けがあり、さらにその下に「日本語→データベース」「情報→検索」という単語同士の係受けがあることを表現できる。係受けの階層性は検索された文献の得点付けにおいて利用している。

一方、検索者は擬似的な自然文で問合せを発行する。検索者は日常用いている言語を検索言語として用いることができるため、検索意図を的確に反映した問合せを発行することができる。また、キーワード検索と異なり問合せ中に単語間の意味関係が保存されているため、それらを検索に活用することができる。

検索エンジンは SUN Ultra2 Workstation 上に C 言語で実装されている。また、CGI(Common Gateway Interface) を用いて書かれているため汎用の WWW Browser を検

学術情報センター電子図書館サービス　　©学術情報センター

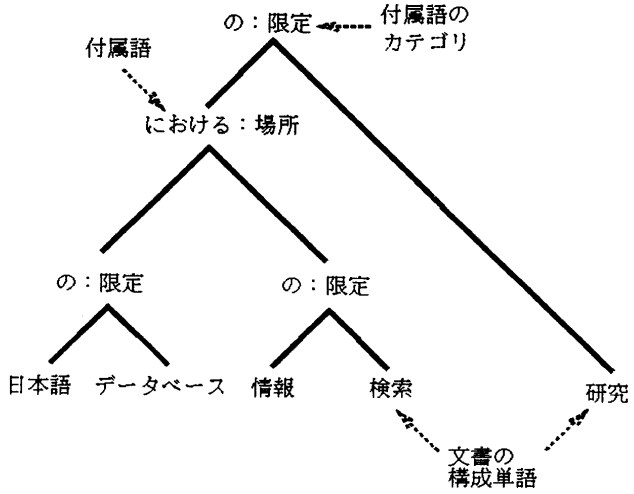


図2　構造化インデックス

索インタフェースとして用いることが可能になっている。検索時には、まず検索対象文と同様の手法で問合せを二分木に構造化し、インデックス・問合せ双方の二分木のマッチングを行いそれに応じた得点を文献に付与する。得点順に文献を整列することでシステムは検索意図との適合度の高い順に検索者に文献を提示することができる。詳細については5章及び6章で説明する。

4　構造化インデックスの作成手法

4.1　形態素解析

文の持つ係受けを判定するために、まず形態素解析を行う。形態素解析の役割は、文を名詞と付属語のトークンに分け、さらに名詞についてはその構成単語に分割することである。本研究では、市販のソフトウエアを形態素解析に用いている。但し、辞書としては付属語や固有名詞、専門語を格納した辞書を人手で作成し、解析に用いている。

付属語の中には「に関する」「についての」のようにほぼ同じ意味を有するにも関わらず表層文字列の異なるものが多数存在する。これらの付属語を表層文字列のみで検索することは検索漏れの原因となる。そこで、付属語をその意味に応じてカテゴリに分類し、カテゴリを用いて付属語の検索を行うことを考える。これは5.3節で詳説する。表1にカテゴリとそこに分類される付属語の一例を示す。カテゴリは18作成され、約98%の付属語がいずれかのカテゴリに分類される。どのカテゴリにも意味的に分類されない付属語は、「その他」というカテゴリを設けてそこに分類している。

4.2　名詞間の係受け判定

名詞数が1の場合、名詞間に係受けは存在しない。また、名詞数2の場合は両名詞を結ぶ係受けが唯一の係受けであるため、係受け判定に曖昧性は生じない。しかしながら3名詞以上の文では複数通りの係受けが可能性として存在するため、その曖昧性を排除せねばならない。

4.2.1　タイトルパターンによる係受け規則の作成

3名詞以上から構成される文の係受けの曖昧性を排除するために、「タイトルパターン」の概念を導入する。タイトルパターンとは文中の名詞と付属語の組合せのことであり、例えば本論文の表題「単語間の係受け情報を用いた文献検索手法」であれば「A間のBを用いたC[1]」というタイトルパターンに属することになる。ここで、「同じタイトルパターンに属する文の係受けは全て等しい」という仮説を立て、実際に検証を行った。科学技術系論文の表題3666件を事例としてタイトルパターンの抽出を行い、そのうち頻度3以上のもの103種について、そこに含まれる表題に対して人手で係受けを付与した結果、殆んどの表題に唯一の係受けが付与される確定的な場合と、表題に応じて複数種の係受けパターンが付与される曖昧な場合があることが分かった。後者の例を図3に示す。

係受けが曖昧な場合について更に調べると、表題の末尾に位置している語が

- 「研究」「提案」「評価」など表題全体の特徴を表し
- 表題中の特定の語との係受けにあまり関与せず、前全体を受ける係受けを有する

語であるか否かで係受けを高確率で同定できることが分かった。これらの語を文末一般語と呼ぶことにする。文末一般語として、上記の語の他「実現」「考察」等53語を定義している。

以上のことから、タイトルパターンと文末一般語の有無によって係受け規則が定まる。この係受け規則を用いて事例分析に用いた文献表題を検定したところ、3名詞の表題の83.8%(1147/1369)、4名詞の表題の48.7%(319/655)に係受けを付与することができ、そのう

[1] A、B、Cは名詞

学術情報センター電子図書館サービス　　©学術情報センター

表1　付属語とそのカテゴリ分け

カテゴリ	主な付属語	カテゴリ	主な付属語
限定	の、な、としての	目的	のための、を目指した、を指向した
場所	における、上の、での	所有	を有する、を持つ、を備えた
手段	による、を用いた、に基づく	内容	に関する、について
and	と、および、ならびに	所有	を有する、を持つ、を備えた
起点	からの、から	考慮	を考慮した、に着手した、を重視した

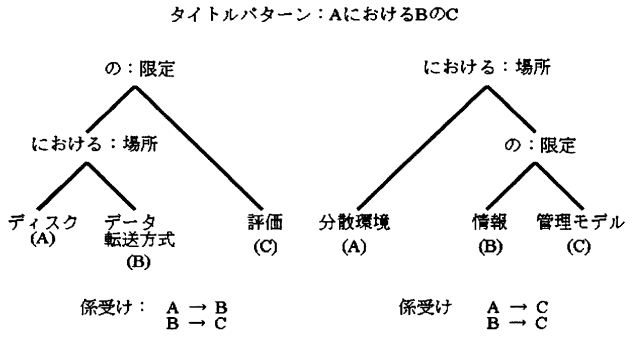


図3　係受けパターンが曖昧な例

ち、3 名詞では 92.0%(1055/1147)、4 名詞では 84.7%(272/319)に正しい係受けが付与された。

4.2.2　拡張タイトルパターン

4.2.1 節の検定から分かるように、4 名詞の場合には半分以上の文献表題に対してタイトルパターンによる係受け規則を適用した係受け判定を行うことができない。そこで、意味の似た付属語は文中で似たような使われ方をするため結果的には同じような係受け関係を有するという経験則を用いて、タイトルパターンの概念を付属語のカテゴリに適用したものを考える。これを拡張タイトルパターンと名付ける。拡張タイトルパターンは 3 名詞の文で 40、4 名詞の文で 33 抽出されており、タイトルパターンで係受けを付与できなかった場合に参照する。拡張タイトルパターンを用いても係受けが判定できなかった場合には、3 名詞の場合には文の末尾の語が文末一般語であるか否かで、4 名詞以上の場合は4.2.3節で説明する経験則を用いて係受けを付与する。

4.2.3　経験則に基づく係受け付与

(拡張)タイトルパターンに基づく係受け規則を用いても係受けが付与できなかった場合には、経験則から係受けを定める。5 名詞以上の文や一部の 4 名詞の文がこれに該当する。経験則としては、以下の 3 種を用いている。

- 語間関係の強い名詞の仮想的な結合
  「な」「された」等「限定」のカテゴリに分類される付属語などは、その付属語の前後の名詞を強く結ぶ性質がある。そこで、付属語と前後の名詞を連結して仮想的な一つの名詞と考えて見かけの名詞数を減らし、よし少ない名詞数のタイトルパターンを適用して係受けを付与する。
- 特定の付属語による文の分割
  「における」等「場所」のカテゴリに分類される付属語、「と」等「and」のカテゴリに分類される付属語は、その付属語の前後の句の間に係受けを成立させることが多い。そこで、それらの付属語を境に文を切断し、切断した各文に対して係受けを定め、最後に切断した句同士を結ぶ係受けを付与する。
- 文末一般語による文の分割
  文末一般語は、4.2.1 節に述べたように直前までの文全体を受ける性質がある。そこで文末一般語の直前の名詞で文を区切り、区切られた文に対して係受けをまず付与する。最後に区切られた文を文末一般語で受けるように係受けを定める。

いずれも文の名詞数を少なくし、3 名詞や 4 名詞のタイトルパターンを用いて係受けを定めるというアプローチである。この手法を用いると、文全体としては完全に正しい係受けが付与されなかった場合にも局所的には正しい係受けが付与される可能性が高く、係受け付与のミスによる検索精度の劣化が抑えられる。経験則を用いた係受け付与の例を図 4 に示す。この例の場合は、経験則を用いたにも関わらず文全体として正しい係受けが付与されている。

4.3　複合名詞解析

現在供用になっている多くの検索システムでは、複合名詞は部分一致などで検索されており、その構成単

学術情報センター電子図書館サービス　　◎学術情報センター

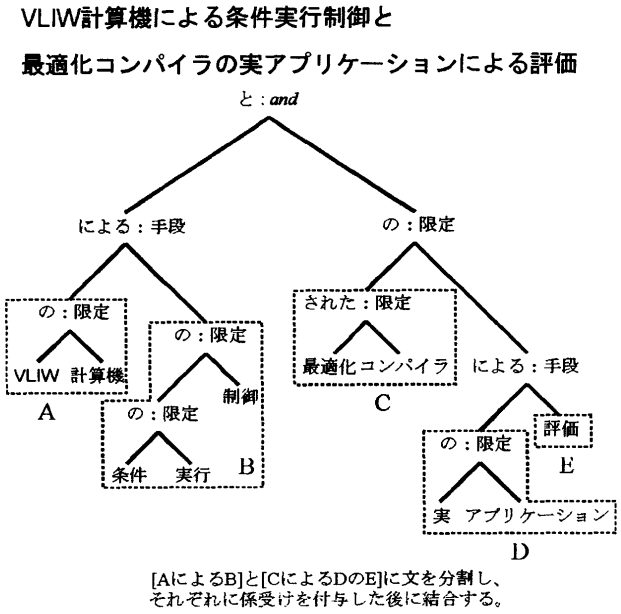


図4　経験則を用いた係受け付与の例

語の意味関係まで考えた検索は行われていない。本研究では、複合名詞を構成する単語を先頭から、適当な付属語を補いながら読み換えることで通常の文に変換できる[8]ことに着目して複合名詞を分割し、構成単語間の係受けを定め構造化インデックスの枠組へ組み込むことを提案する。この手法により複合名詞を通常の文同様に扱うことができ、2.3節で述べたような複合名詞検索の問題点が解決できると考えられる。

複合名詞の分割、係受け付与においては、

- どの付属語で単語間を補うか
- 単語同士が意味的にどのように結合しているか

が重要となる。

4.3.1　単語間を補う付属語

まず、複合名詞に対してどのような付属語を補うのが適当かを考える。複合名詞の構成単語の末尾が「～化」「～的」のような接辞で終っている場合には、それぞれに特徴的な付属語を補って読み換えることができる。この例を表2に示す。

それ以外の場合については、原則的に「の」という付属語で名詞間を補う。付属語と名詞の共起関係を付属語が先行する場合、名詞が先行する場合双方について抽出した結果、「への→応用」「による→実装」などの共起関係が得られた。しかしその頻度と同数以上に「の」と名詞との共起関係が得られることが多いため、統計的に考えれば「の」で補うのが正しいと考えられる。宮崎の論文[8]や石崎の論文[9]では、名詞を意味に応じてカテゴリに分け、そのカテゴリの結び付き方で名詞の意味関係や付属語の補い方を定めているが、その辞書の保守が容易でなくかつ人手を介さなければならないため、本研究では表層情報を主体にした複合名詞の係受け解析を考える。

4.3.2　固有名詞

固有名詞は検索対象文献から予め人手で切り出し、辞書に登録している。固有名詞は単独あるいは複合名詞の末尾の語として出現する場合が多い。そこで、複合名詞の末尾の単語で固有名詞辞書を引き、もし該当すれば直前の付属語を「としての(同格)」に変える。また、固有名詞は前の複合名詞全体を受ける性質があるので、次節で述べる単語2-gramの頻度には最低値を割り当てる。

4.3.3　単語2-gram統計の複合名詞解析への利用

複合名詞解析においては、その構成単語間に強い意味的結合関係のある語同士に優先的に係受けを付与しなければならない。その単語間の意味的結合関係の強さの指標として、単語2-gram統計を用いることを本研究では提案する。単語2-gramとは、複合名詞中の2個の連続した単語の組合せがどのような頻度で生じるか統計を得たものである。意味的結合関係の強い単語同士はしばしば同一名詞中で隣合って出現するため、十分な標本数を取れば単語2-gramがそのまま単語間の意味的結合関係の強さを表すことができると考えられる。本研究では、検索対象文献1490件の表題から計10347の単語2-gramを抽出し、そのうち頻度数が4以上のもの504種類を単語2-gramデータとして格納し、複合名詞解析に用いている。

以下に、単語2-gramの複合名詞解析の適用について説明する。複合名詞中の構成単語間における単語2-gramの状態としては、ある単語2-gram頻度に対して

1. 頻度 b_{th} 以上の単語2-gramが連続しない場合
2. 頻度 b_{th} 以上の単語2-gramが連続する箇所がある場合

の2種類が考えられる。

1の場合、単純に高頻度の2-gramが付与されている単語間ほど強い係受け関係を持つ。よって単語2-gramの高い順に単語間に係受けを成立させていく。複合語を構成している単語を $w_1,w_2,...,w_k$、w_j と w_{j+1} の間の単語2-gram頻度を $b_{j,j+1}$ とする。まず、頻度最大の単語2-gram $b_{i,i+1}$ を探し、w_i と w_{i+1} の間に係受けを定める。こ

学術情報センター電子図書館サービス　　◎学術情報センター

表2　接辞と補うべき付属語

接辞	補う付属語	複合名詞	読み換えられた文
間	間の	車々間通信	車々／間の／通信
上	上の	基盤上配線	基盤／上の／配線
的	な	段階的構造化	段階的／な／構造化
用	用の	検索用言語	検索／用の／言語
可能	可能な	拡張可能データベース	拡張／可能な／データベース
化	された	構造化文書	構造化／された／文書
付き	付きの	子機付き電話	子機／付きの／電話
向け	向けの	利用者向け言語	利用者／向けの／言語

れを順次2番目に高い頻度の単語2-gram以降にも適用する。但し、ある単語2-gram$b_{x,x+1}$について、w_xが既に$b_{x-1,x}$によって係受けに使われていた場合には、w_xに付与された係受けの根$R(w_x)$に対して係受けを付与する。

一方、2の場合には頻度b_{th}以上の単語2-gramが連続する箇所がある1つの意味的なまとまりである可能性が高い。よって、隣接2単語間の単語2-gramのみでは単語間の結合関係が判断しにくい。この場合には、連続3単語におけるそれぞれの単語間の単語2-gramを用いて、連続3単語間における意味的結合関係を判定する。複合語を構成する単語のうち、頻度b_{th}の単語2-gramが連続する部分を$w_x,w_{x+1},...,w_y$、その単語2-gramを$b_{x,x+1},b_{x+1,x+2},...,b_{y-1,y}$とおく。まず、先頭の3単語のみを見る。単語$w_x,w_{x+1},w_{x+2}$の間に成立する単語2-gramのうち、$b_{x,x+1}$、$b_{x,x+2}$を見ることで、$w_x$が$w_{x+1}$、$w_{x+2}$のどちらとより強い結合関係にあるかを推定する。もし$b_{x,x+1} > b_{x,x+2}$ならばw_{x+1}と、$b_{x,x+1} < b_{x,x+2}$ならばw_{x+2}とw_xは意味的により強く結合していることになる。本手法では単語2-gramの頻度で結合関係の強さを反映するため、上記の大小関係に基づいて、必要に応じて$b_{x,x+1}$、$b_{x+1,x+2}$の頻度を入れ換える。すなわち

- もし$b_{x,x+1} > b_{x+1,x+2}$で、かつ$b_{x,x+1} < b_{x,x+2}$ならば$b_{x,x+1}$と$b_{x+1,x+2}$を入れ換える。
- もし$b_{x,x+1} < b_{x+1,x+2}$で、かつ$b_{x,x+1} > b_{x,x+2}$ならば$b_{x,x+1}$と$b_{x+1,x+2}$を入れ換える。

ことになる。これを$w_{x+1},w_{x+2},w_{x+3},...,w_{y-2},w_{y-1},w_y$まで逐一行って単語2-gram頻度を入れ換えた後、1と同様の操作で複合名詞の構成単語全体に係受けを付与する。

例として「関係データベース管理システム」という複合語の係受け付与を考える。まず、複合名詞を付属語を補って読み換える。この例の場合は付属語がないので「関係のデータベースの管理のシステム」となる。

構成単語はw_1＝「関係」、w_2＝「データベース」、w_3＝「管理」、w_4＝「システム」である。また、単語2-gram頻度はそれぞれ$b_{1,2}$＝101、$b_{2,3}$＝125、$b_{3,4}$＝149である。単語2-gram頻度がみな$b_{th} = 30^2$以上であるため、この語は1つの意味的なまとまりで構成されている語である。

まず、先頭から3単語を取り出し、それぞれの単語に対する単語2-gramを比較する。$b_{1,3}$＝1であり、$b_{1,2} > b_{1,3}$が成立する。一方で$b_{1,2} < b_{2,3}$である。よって$b_{1,2}$と$b_{2,3}$の頻度を取り換える。次にw_2から3単語を見る。$b_{2,4}$＝143であり、$b_{2,3} < b_{3,4}$である。また、$b_{2,3} < b_{2,4}$であるため、頻度の入れ換えは必要ない。

頻度入れ換え後の単語2-gram頻度は$b_{1,2}$＝125、$b_{2,3}$＝101、$b_{3,4}$＝149であるから、w_3とw_4の間、w_1とw_2の間、w_2とw_3の間に順に係受けを定める。前2者は係受けに関与する単語が他の単語2-gramによって使われていないが、後者に関しては使われているので、$R(w_2)$と$R(w_3)$の間に係受けを付与する。最終的な解析結果を図5に示す。

4.4　インデックス作成上の工夫点

本節では、構造化インデックスの品質向上のための工夫について述べる。

4.4.1　表記揺れの吸収

「DB」と「データベース」のような表記揺れは検索精度の低下につながる。そこで、検索対象文献から表記揺れを抜き出して表記揺れ辞書を作成した。この辞書には「元の表記」と「統一表記」が登録されており、様々な表記はある一語に吸収される。また、複合名詞

[2] 総単語2-gram数の0.3%を閾値としている

学術情報センター電子図書館サービス　　©学術情報センター

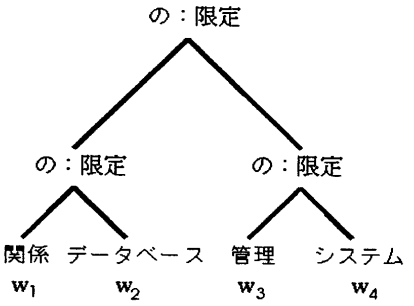


図5　複合名詞「関係データベース管理システム」の解析結果

中の構成単語間の係受けを本システムではサポートしているため、統一表記には係受けを持つ語を優先している。表記揺れ辞書の一例を表3に示す。

4.4.2　文内照応解析

文献表題には、しばしば「その」「それ」等の指示代名詞が出現する。この内容を正しく補うことは検索精度の観点から大切である。本研究で対象としている文の場合には、構造が定型的でありかつ前後の文脈のない単独の文であることから、補うべき文(これを照応文という)は経験的に

- 先頭から指示代名詞の直前の名詞まで
- 先頭から指示代名詞の2つ前の名詞まで(直前が一般語の場合)

のいずれかである。照応文の二分木を指示代名詞の位置に代入することで指示代名詞の内容を補った検索を行うことができる。また、二分木において「その」の置に単純に照応文を代入してしまうと二分木中に重複する部分が生じてしまう。これはインデックスのオーバヘッドとして好ましいことではない。そこで、可能であれば重複部分を省略してインデックスの大きさを小さくする。この例を図6に示す。

5　構造化インデックスのマッチング

本システムでは、問合せも文献表題と同様の構造を持った擬似的な自然言語で発行する方式を採っている。よって、検索者は検索言語の存在を意識せずに自分の意図を明確にシステムに伝えることができる。一方、検索システム側から見ると、問合せに関しても検索対象文同様4章に示す処理で係受けを付与して二分木に構造化可能であり、検索は両者の木構造同士のマッチングを取ればいいということを意味している。図1はそれを端的に示している。

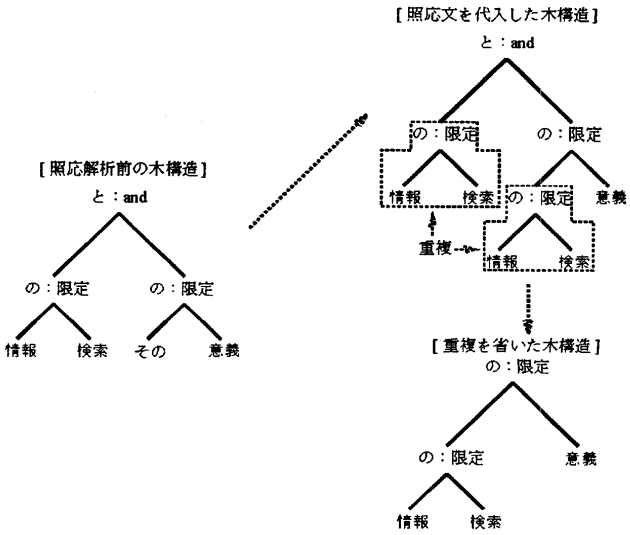


図6　文内照応解析におけるインデックスのオーバヘッドの縮小

5.1　マッチング条件とその問題点

係受け関係を用いた検索では、問合せにおける係受け関係、すなわち「係る語」「係られる語」の組関係の等しい検索対象文が適合文書候補となるため、同じ単語を用いているものの係受け関係の異なる文を検索において排除することができる。また、「係り関係」などを用いて係受けの成立状態によって文献間に得点差をつけることができるため、より検索要求に適合した柔軟な検索を実現することができる。

ここで、係受け関係の満たされた検索対象文を抽出するためには、問合せの二分木を包含する検索対象文を抽出すれば良い。しかしながらこの場合、以下のような問題点が生じる。

- 問合せに含まれる情報のうちの1つでも情報の欠落した文献は検索されない
- 問合せに含まれる情報に修飾語を付加した文献は検索されないことがある
- 問合せに多くの情報を含ませるほどマッチしにくくなる

そこで日本語の係受け関係に則った木構造の縮退動作を行うことでより柔軟な検索を行うことを考える。

5.2　二分木の縮退

5.2.1　縮退の操作

ここでいう二分木の縮退とは、文中の係受けのうちのいくつかを消去して残りを保存することで、文意を極力変化させずに二分木を縮小する動作を指す。図7の左側の二分木について考えると、二分木の縮退のためにはノードqとその子孫について「ノードpを1ま

学術情報センター電子図書館サービス　　© 学術情報センター

表3　表記揺れの例

元の表記	統一表記
DBMS	データベース管理システム
World Wide Web、ウエブ、ワールドワイドウエッブ	WWW
インデクス、索引	インデックス

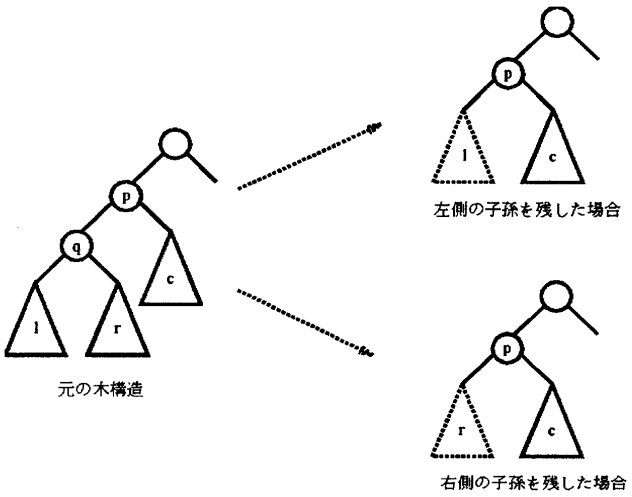


図7　インデックスの縮退のモデル

たはrの新たな親とし、ノードq及びその子孫l、rのうち選ばれなかった方を削除する」という操作が必要になる。この操作に従って縮退を行った二分木が図7の右側の二分木である。

5.2.2　縮退の規則

縮退後に文意をあまり変化させないためには、縮退は日本語の係受け規則に則って行なわれる必要がある。日本語の係受け規則は以下の4点に集約される。

- 前方の単語から後方の単語に係る
- 単語の係先は一つである
- 複数の単語を一つの単語が受けてよい
- 係受けは交差しない

木構造にこれを当てはめると、ある語が直接係る語は、木構造を1つ右上に登りそこから右下に降りた末端の語である。これは名詞句や複合名詞についても同じことがいえる。

よって、左の子孫を残した場合には係先が変化する。これはを見れば明らかであり、縮退前には部分木lはrに係っていたが、縮退後はcに係っている。この場合、文意を大きく変えてしまう可能性が高い。一方、右側の子孫を残した場合にはl→rの係受けが消滅するだけで他の係受けは変化しないため、縮退後の文意もそれほど変化しないと考えられる。以上の考察から、縮退においては右側の子孫を残すことを基本とする。具体的な縮退の例を図8に示す。

縮退において、例外的に左側の子孫を残した縮退木を作成する場合がある。それは左の子孫にも検索上重

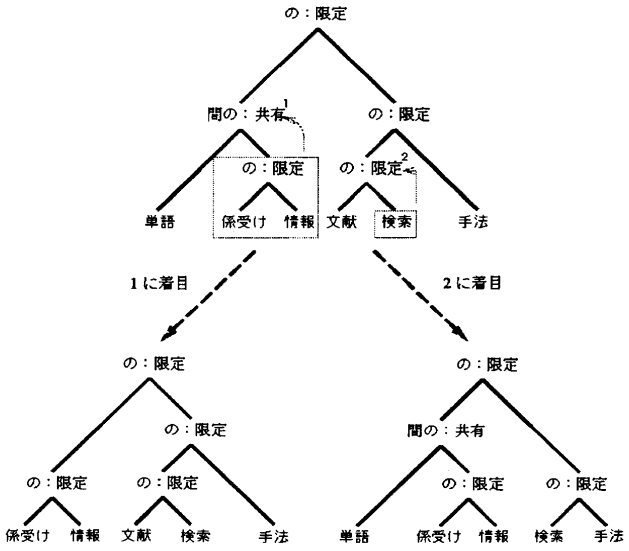


図8　二分木構造の縮退の実例

要な要素が残されている場合や左の子孫と右の子孫が並列の関係にある場合であり、具体的には以下の3つである。

- 二分木の根に着目して縮退を行う場合
- 着目しているノードが「and」「or」カテゴリの場合
- 着目しているノードの右の子孫が文末一般語相当の語の場合

5.3　縮退を踏まえたマッチング処理

5.3.1　マッチングアルゴリズム

以上に紹介した縮退処理を踏まえて、本研究では以下の方法で問合せと文献情報との間でマッチングをとっている。本手法を用いることで、問合せ、検索インデックス双方の任意の係受け・単語の間でその一致判定を行うことができる。

1. 問合せを4章に説明した手法で二分木に構造化する。

学術情報センター電子図書館サービス　　©学術情報センター

2. 問合せを縮退する。元の問合せも含めた全縮退木を問合せ集合と見なし、集合中の全ての問合せに対し順次3-6の処理を行う。
3. 問合せ集合から、問合せを 1 つ選ぶ。これを q_i とする。
4. 問合せq_iの持つ全ての単語を含む検索インデックスをシグネチャファイル法[10]で抽出する。もしq_iが1単語の問合せであれば、ここで抽出された検索インデックスは無条件でマッチするため、6の処理を行う。
5. 抽出した検索インデックスに問合せが包含されるか確認する。包含関係は、検索インデックスを縦型探索して部分木を生成し、それとq_iの二分木が一致するか否かで判定する。包含されるならば6へ。包含されないならば、検索インデックスを縮退して再度包含するか確認する。これを縮退木が生成しなくなるまで行う。縮退木が生成されなくなっても包含されない場合はマッチしなかったと判断し、3に戻る。
6. マッチした検索インデックスに対し、元の問合せと一致した部分とマッチング情報を付与する。全問合せ縮退木についてマッチング判定が完了していなければ、3に戻り処理を続ける。

マッチした文献情報には、マッチング部分にマッチング情報を付与し、6 章で説明する得点付けに活用する。

5.3.2 マッチング情報

マッチング情報としては、以下に述べる 3 種を用いている。

- 係り関係の一致度
- 係受け中における一般的概念の語の有無
- マッチした問合せにおける係受けの頻度情報

がある。係り関係の一致度については、以下に述べる3つの段階を定めている。

Exact Match 係られる語句、係る語句、係り関係を表す付属語が問合せと検索インデックスで等しかった係受け。その係受けは意味的にも問合せに等しいと考えられる。

Category Match 係られる語句、係る語句が等しく、係り関係を表す付属語は問合せと等しくないもののそのカテゴリは等しかった係受け。その係受けは問合せと若干の意味の違いがある可能性があるが、ほぼ等しい。

Wild Match 係られる語句、係る語句は問合せと等しいものの、係り関係を表す付属語もそのカテゴリも等しくない係受け。その係受けは問合せと意味的に等しい保証はない。

なお、マッチした単語には Word Match というマッチングの段階を付与する。係り関係に段階を設けることで、同じ「係られる語句」「係る語句」を持つ問合せに対して「係り関係」、つまり係受けの意味的類似度に応じた得点付けが可能となる。

「係る語句」「係られる語句」の少なくとも一方が一般的概念を表す単語の場合、その多くはたとえ係受け関係が一致していても検索ノイズになる場合が多い。よって、一般的概念を表す単語で占められた「係る語句」「係られる語句」の数を記録し、その数に応じて係受けの得点を下げる。

係受けの頻度とは、ある係受けが全文献中に何度出現したかを表す値である。係受けの頻度の低いものに高い得点を与えることで、検索対象文の特徴を表す係受けに高い得点を与えることができる。ここで係受けの頻度を正しく求めることは、木構造を展開して考え得る全ての係受けを求め、それを重複なく数えなければならないため計算のコストが高い。そこで近似的に、相対的な頻度の大小関係が正しくなるように係受け頻度を定めることを提案する。

係受けの頻度の特徴として、頻度の少ない係る語と頻度の高い係られる語から構成される係受けの頻度は低いということが挙げられる。そこで本研究では、子の頻度の相乗平均を係受け頻度と定義している。あるノード番号 k の子孫ノード(葉)の番号を l、r とし、それぞれの頻度を f_k、f_l、f_r とすると、f_k は式(1)で計算される。

$$f_k = \sqrt{f_l \times f_r} \tag{1}$$

初期値として二分木の葉に単語の頻度を与え、ボトムアップ的に根に向けて式(1)を計算することで、二分木の全ノードに係受けの頻度を付与できる。式(1)によって係受けの頻度を計算した例を図9に示す。

検索インデックスと問合せとの一致部とそのマッチング情報は、各インデックスごとに縮退前の問合せと同一の木構造(これをマッチング二分木と呼ぶ)を準備しておき、一致部とマッチング二分木が対応する部分格納される。

学術情報センター電子図書館サービス　　© 学術情報センター

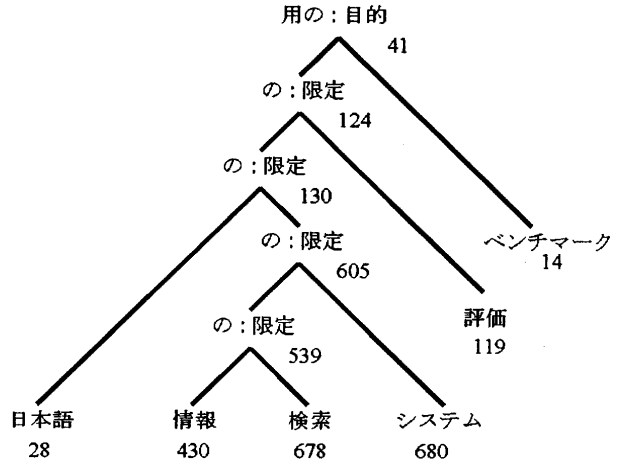


図9　相乗平均を用いた係受け頻度付与の例

6　文献に対する得点付け

本章では、検索された文献に対する得点付けについて説明する。文献の得点は、表題に対する得点とキーワードに対する得点の総和で計算する。また、表題、キーワードの得点は、それに含まれる係受け及び単語の一致度によって求める。

6.1　文献主題要素の統合

本研究では、文献表題・副題とキーワードで文献の持つ主題を代表させている。問合せと文献の主題との一致度が得点として出力されるべきだから、性質の異なる主題要素をどのように統合するかは重要である。

一番単純な方法は、検索対象文毎に問合せとマッチングをとって得点付けをし、それぞれの得点の線形結合で最終的な文献の得点を決める方法である。しかしながらこの場合には、同じ係受けが複数の検索対象文に含まれていた場合の重複検索の問題、検索対象文毎の重み付けの難しさの問題があり、主題と問合せの適合度を反映した得点付けは難しい。

そこで、本研究では図10に示す手法で主題要素を統合し、得点付けに活かすことを考える。

文献表題・副題、キーワードの特徴として、

- 文献表題は文献を代表する見出しであるため、比較的多数の単語を用いて記述される場合が多く、文の中で係受けが閉じている。また副題は表題を補足するものであるが、特徴的には表題と類似である。
- キーワードは1名詞、多くて2名詞から構成されており、多くの文献には複数個付与されている。また、キーワード付与者は文献において重要と思う概念を細切れにしてキーワードという形に分割する。よって係受けが単独のキーワードの中では閉じているとは限らない。

ということが挙げられる。

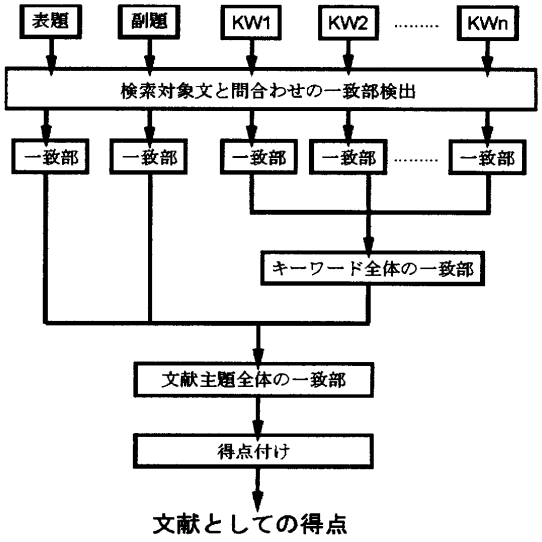


図10　文献主題の統合とそれを用いた得点付け

そこで、キーワードについてその和集合を取ったものを考える。これはキーワード付与者が重要と思う概念の集合と等価であり、係受けはこの集合の中で閉じていると考えられる。さらに表題・副題とキーワード和集合の和を取ると、それは文献についての主題と一致すると考えられる。

以上の考察を踏まえて本研究で用いている得点付けの手法を説明する。まず、文献表題・副題、キーワードそれぞれについて、5.3節で述べた手法で問合せとのマッチングを取る。ここで、各文献情報の問合せとの一致部は、各々の文献情報のマッチング二分木に格納されている。次に、キーワードのマッチング二分木の和を取り、キーワード集合全体のマッチング二分木をを作成する。キーワードは単独では係受けが閉じておらず、キーワード集合となって初めて係受けが閉じることを考慮して、キーワード集合全体のマッチング二分木において「係る語句」「係られる語句」があるもののマッチング段階が付与されていないノードにはWild Matchのマッチング段階を与える。次に、キーワードのマッチング二分木、表題・副題のマッチング二分木の和を取り、文献全体の一致部を作成する。文献全体の一致部に対し、次節で説明する得点付け手法を適用することにより文献としての得点が付与される。

6.2　係受けに対する得点

係受け点を決める要素としては、5.3節で説明したマッ

学術情報センター電子図書館サービス　　©学術情報センター

チング情報の他に、係受けの階層を考える。構造化デックスでは名詞句、複合名詞、単語の係受けを階層的に表現することができる。文中に上記の 3 つの係受けが存在した場合、名詞句、複合名詞、単語の順に係受け関係の文章全体に与える影響力は大きくなるため、二分木の根に近い係受けに高い得点を与える。

これらを踏まえて、本研究ではある文献 b の係受けに対して式(2)で得点を与えている。

$$S_b(D)=\frac{\sum_{j=1}^{N}\frac{w_j t_j l_j}{k^{dj}}}{\sum_{j=1}^{N}\frac{t_j l_j}{k^{dj}}} \qquad (2)$$

但し、l_jは各係受けのマッチング時の頻度による重みであり、

$$l_j=(\log_2\frac{S_{all}}{f_j})^2 \qquad (3)$$

である。なお、式(2)、式(3)中の変数はそれぞれ

$S_b(D)$ ： 文献 b における係受けの点数
j ： マッチング二分木におけるリンクの番号
N ： ノードの総数
w_j ： マッチング段階に応じた重み(≦1.0)
t_j ： 一般語の重み(≦1.0)
k ： 階層の深さにおける重み(≧1.0)
d_j ： 階層の深さ(根が 0)
S_{all} ： 全文献情報数
(=表題、副題、キーワードの総数)
f_j ： ノード j の頻度

式(2)は、取り得る最高点で正規化されており、1 点満点となる。

6.3 単語に対する得点

単語点とは、問合せに一致した文献情報中の単語に対して与えられる得点である。単語の得点の決定要素としては、その単語の文献集合中における頻度が考えられる。多くの文献に出現する単語では文献を絞り込むことはできないが、希少な単語であれば一単語だけで文献を絞り込むことができる可能性が高くなるため、自ずと希少な単語ほど重要度が高くなる。この条件を満たす得点付け手法に、TF-IDF 法がある。本研究で扱っている文献情報は表題やキーワードという高々名詞数個の文であり、TF はほぼ 1 であるから、TF の項は無視している。すなわち、式(4)を単語点の算出式として用いている。

$$S_b(W)=\frac{\sum_{i=1}^{n}\log_2\frac{S_{all}}{S_i}}{\sum_{i=1}^{n}\log_2\frac{S_{all}}{S_i}}\cdot\delta_i \qquad (4)$$

但し、

$S_b(W)$ ： 文献 b における単語点
i ： マッチング二分木中の単語の番号
n ： 単語の総数
δ_j ： $\begin{cases}1 & \text{単語} i \in \text{検索インデックス}\\ 0 & \text{otherwise}\end{cases}$
S_{all} ： 検索対象文の総数
S_i ： 単語 i が出現する検索対象文数

式(4)は、取り得る最高点で正規化されており、1 点満点となる。

6.4 係受け点と単語点の結合

係受け点と単語点にそれぞれある重みをつけて線形結合し 100 点満点換算したものを本手法における文献の得点とする。これは

$$S_b=\frac{100\times x_D}{x_D+x_W}S_b(D)+\frac{100\times x_W}{x_D+x_W}S_b(W) \qquad (5)$$

で表される。ここで、x_D及びx_Wはそれぞれ係受け点、単語点の重みであり、式(5)は 100 点満点で正規化されている。本手法において最適な重みは、7 章で検索実験を通じて検証する。

7 評価と検討

7.1 検索出力の例

本システムの検索の特徴を表すために、同じ単語から構成されるものの係受けの異なる以下の 2 件の問合せを実際にシステムに発行した。結果を表 5、表 6 に示す。なお、各パラメータは表 4 のように設定した。

Query 1 並列関係データベースにおける結合演算処理
Query 2 関係データベースにおける並列結合演算処理

Query1 では「並列(関係)データベース」、Query2 では「並列結合演算(処理)」がそれぞれの問合せにおいて特徴的な係受けを含む部分である。表 5 と表 6 を見比

学術情報センター電子図書館サービス　　© 学術情報センター

表4　本検索実験におけるパラメータ

パラメータ	設定値
w_j	Exact Match 1.0 Link Match 0.9 Wild Match 0.5
k	1.4
t_j	片方が一般語 0.2 双方一般語 0.1
x_D	75
x_W	25

べると、それぞれ特徴的な係受けを含む文献が上位に出力されている。

7.2　適合率による有効性評価

本節では、出力を順位順に区切った上での再現率・適合率を測定することで、キーワード検索と比べて問合せの意味内容に適合した文献が真に上位に検索されることを示す。キーワード検索では出力文献は同列に扱われるが、比較対象として論理検索の結果をTF-IDF法で重み付けしたものを用いる。TF-IDF法はキーワード検索においては精度の高い得点付け手法としてしばしば用いられている手法である。まず、正解集合について定義をしなければならない。正解としては以下に示す2種類の正解ランクを考え、その和集合を正解集合と定義する。

ランクA正解集合　問合せの内容を多く含み、意味的に殆んど問合せの内容が削られることなくマッチするもの。

ランクB正解集合　問合せの内容の一部を含むだけで、意味的には問合せの持つ内容の多くが削られてマッチするもの。

「日本語テキストDBからのキーワード抽出による検索インデックスの作成」という問合せを例にとると、「日本語インデックスからのインデックスの作成」のように問合せの内容の多くを含む文献情報はランクA、「テキストDB」「日本語DB」のように問合せの内容の一部を含むものはランクB正解集合となる。また、単に「データベース」「テキスト」のように問合せの内容をある程度十分に記述しているとはいえない単語、「テキストDBの作成」のような問合せの内容とは異なる係受けを持つ名詞句等は不正解とする。

本手法の評価指標としては、適合率を用いる。その定義は

$$\text{適合率} = \frac{\text{検索文書集合中の正解文書数}}{\text{検索文書数}} \qquad (6)$$

である。すなわち、適合率は検索ノイズの少なさを表す値である。また、本研究では出力文献の数ではなく出力文献の順位付けを問題としているため、適合率は出力文献全体に対して測定するのではなく、正解文書数のある割合までの出力文献に対して測定する。

評価においては、12件の問合せを発行してその検索出力を見る。評価に用いる適合率の計算手法を以下に示す。

- 検索システムのプロトタイプに対して問合せを発行。
- 出力された検索結果に対して、人手で正解ランクを付与。
- 正解文献の10%ごとに出力文献を区切り、そこまでの正解数を求める。10%から150%までこの処理を行う。但しぴったり割り切れなかった時には、最も近い値を採用する。
- 求められた正解数から、式(6)に従って適合率を求める。
- 正解文書数がぴったり割り切れなかった場合には、その前後の適合率を用いて線形補間を行う。ある割合sにおける適合率P(s)の値を、sを挟む2点s_-,s_+における適合率$P(s_-)$,$P(s_+)$から線形補間する式は式(7)となる。

$$P(s) = P(s_-) + \frac{P(s_+) - P(s_-)}{s_+ - s_-} \qquad (7)$$

- 各問合せにおいて求められた適合率の平均値を計算する。

評価においては、式(5)の最適なx_D、x_Wの重みも求める必要がある。そこで、表7に示す5種類の重み付けで得点をそれぞれ計算し、適合率を求める。ここで、方式Eの場合は単語の得点のみで文献の得点を計算したことになりTF-IDF法と等価になるため、これを本手法との比較対象とし係受け関係を検索に用いることの有効性を検証する。

まず、全正解集合に対して評価を行った。適合率を図11、再現率－適合率曲線を図12に示す。ここで、再現率－適合率曲線とは、再現率、すなわち出力文献に含まれる正解数を増やした時に適合率がどのように変化するかを示す曲線である。

学術情報センター電子図書館サービス　　©学術情報センター

表5　「並列関係データベースにおける結合演算処理」に対する検索結果

順位	文献番号	得点	文献(上段：表題、下段：キーワード)
1	1325	79.04	主記憶データベースに関する研究 主記憶データベース、性能評価、並列処理、結合演算処理、関係データベース
2	157	67.79	スーパーデータベースコンピュータ SDC-II における 並列結合演算処理に関する性能評価 データベースマシン、並列アーキテクチャ、関係データベース処理、性能評価
3	1426	65.88	SymmetryS81 における結合演算の並列処理に関する考察 関係データベース、結合演算、並列処理
4	520	64.85	並列関係データベースシステムにおけるバッチ問合わせ処理最適化技法の検討 並列関係データベースシステム、多重結合演算、バッチ処理、最適化
5	1185	59.62	トランスピュータを用いた並列データベースマシンにおける 結合演算の性能評価 関係データベース、結合演算

表6　「関係データベースにおける並列結合演算処理」に対する検索結果

順位	文献番号	得点	文献(上段：表題、下段：キーワード)
1	157	78.50	スーパーデータベースコンピュータ SDC-II における 並列結合演算処理に関する性能評価 データベースマシン、並列アーキテクチャ、関係データベース処理、性能評価
2	1426	70.91	SymmetryS81 における結合演算の並列処理に関する考察 関係データベース、結合演算、並列処理
3	519	70.91	並列 SQL サーバ SDC-II におけるトランスポーズ型ファイル 編成適用の検討 関係データベース、トランスポーズファイル、多重結合演算、並列処理
4	1325	66.57	主記憶データベースに関する研究 主記憶データベース、性能評価、並列処理、結合演算処理、関係データベース
5	662	60.29	並列計算機 AP1000DDV における多重結合演算の実装とその評価 並列多重結合演算、並列データベースシステム、分散メモリ型並列計算機

グラフから、方式E、すなわちTF-IDF法を用いて得点付けをした場合に最も検索効率が低く、係受けの重みを高めていくに従って検索効率が高まることが分かる。また、係受けの得点の重みを増す程検索効率が高まったものの、方式Bと方式Aでは有意な差が見られなかった。これは、両方式の以下のような利点が均衡したためだと考えられる。

- 手法Aの重み付けでは係受けの一致が最も強調される。正解文書の多くは係受けの一致を伴うため、その部分を強調した得点付けならば、それらの文書を上位にランクすることができる。
- 手法Bの重み付けでは係受けの一致を強調しながらも、係受けの付与のミスや一単語で問合せの主題を述べている場合のように係受けの得点と適合度がマッチしない場合の検索において、検索漏れを単語の得点で防ぐことができる。

係受け点と単語点に対する重み付けが有効であったことを立証するために、ランクA正解文書、ランクB正解文書、不正解文書についてそれぞれの単語点及び係受け点を、単語点と係受け点を2つの座標軸とする2次元平面にプロットした。これを図13に示す。

学術情報センター電子図書館サービス　　© 学術情報センター

表7　評価を行う係受け・単語の重み

	方式	係り受けの重み x_D	単語の重み x_W
係受けを利用	A	100	0
	B	75	25
	C	50	50
	D	25	75
単語のみを利用	E	0	100

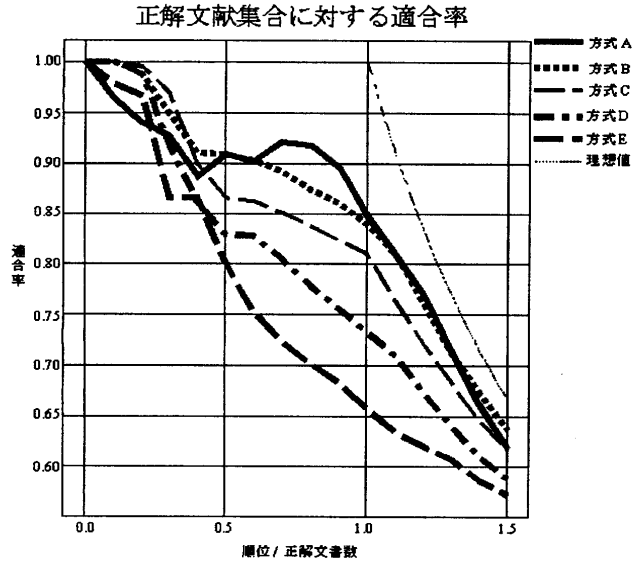


図11　正解集合に対する適合率曲線

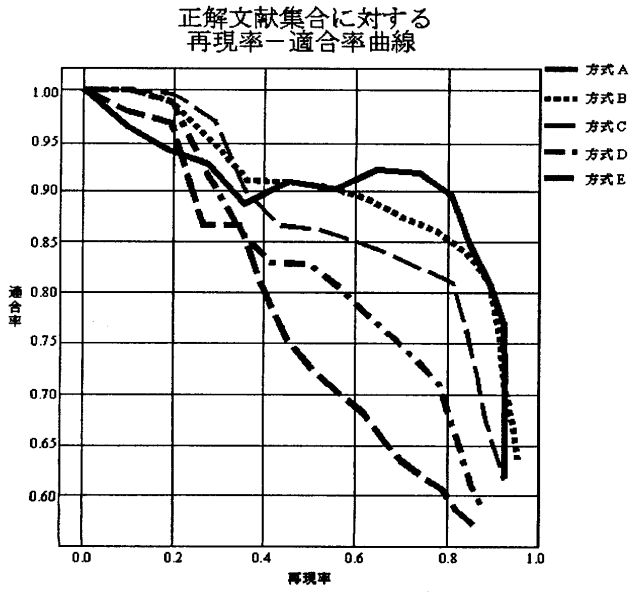


図12　正解集合に対する再現率--適合率曲線

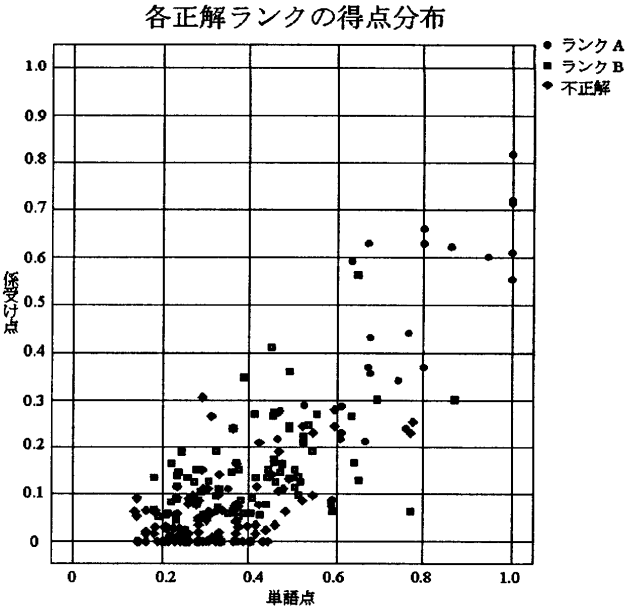


図13　全体の得点分布

図13から、ランクA正解文書はその多くが係受け点、単語点の高い位置にあることが分かる。これらの文書については、正解文書であることを識別することは容易である。よって正解文書を正しく判断するためには、ランクB正解集合と不正解文書集合をいかに正確に判別するかを考えなければならない。判別において、両者の集合を明確に分離できる曲線を図13上に引けることが望ましいが、図13を見る限りはそれは難しい。そこで、できる限りの分離を考えると、平面においてある傾きを持った直線を引くことでおおよその分離ができることが分かる。逆に単語点に垂直な軸で両者の集合を分離することは困難である。これは単語の得点の大小だけでは正解文書と不正解文書を判断できないことを示している。

このことから、単語点のみを用いるのではなく、係受け点と単語点にある重みを与えたものの和をとることの有効性が分かる。

図11、図12のグラフでは、係受け関係を検索に用いた場合と単語のみを検索に用いた場合の検索効率の差が現れている。しかしながら、本来は両手法の間に本質的な差がないにも関わらず偶然に測定値に大きな差を持つ結果を用いたために両手法の間に差が生まれたとも考えられる。その偶然性を排除するために、統計的検定を用いて両手法の差を検定し、本手法の優位性を裏付ける。

検定の対象としては、最高の検索効率を出した方式Bと、TF-IDF法を用いた方式Eについて、正解文書数と

学術情報センター電子図書館サービス　　© 学術情報センター

等しい順位における適合率を扱う。検定手法として、2つの母集団の平均値の差を検定する2標本量t検定[11]を用いる。ここで方式B、方式Eの適合率平均はそれぞれ0.840、0.657であり、両者の標本分散値は6.67×10^{-3}、4.14×10^{-2}である。よってt値は2.88となり、$t_{\alpha/2}(16)$の値から信頼区間を求めると、98%以下では両者の平均値が等しいという仮説が棄却される。ゆえに、98%の信頼度で両手法の間に有意な差があることが実証できる。

以上の実験結果及び考察から、文献に対して係受け点と単語点を定めて、それらの線形和で文献の得点を定めることで、キーワード検索と比べて精度の高い検索を行うことができることが立証されると考えられる。

8 今後の課題と展望

本研究では、文献情報のうち比較的構造の単純な表題やキーワードについて、その構成単語間の係受けを考慮した検索を行うことで検索要求に適合した検索出力を得ることができることを立証した。また、本システムでは検索者は疑似的な自然言語で問合せを発行するため、検索者にとっては問合せの発行や検索結果の論理演算などの手間が必要なくなり、使い勝手の優れたシステムとなっている。

今後の課題としては、以下の事柄を挙げることができる。

8.1 検索におけるオーバヘッドの削減

検索におけるオーバヘッドとして、検索時間と必要記憶容量を挙げることができる。従来のキーワード検索よりも複雑な検索を行っているためオーバヘッドの増加はやむを得ない面があるものの、なるべく削減することが望ましい。特に検索時間については検索者の使い勝手に直接関係してくるため、アルゴリズム的な改良によって高速化を図る必要がある。

8.2 主題要素の拡充

現在は文献の主題として、比較的抽出が容易な表題・副題・キーワードを用いている。しかしながらそれだけでは文献の主題を表現するのに十分ではない。論文の梗概や全文データから自動的に検索インデックスを抽出することで文献の主題要素を充実させ、より主題を漏れなく検索できるシステムにする必要がある。

8.3 類義語・関連語の処理

本手法では付属語については類義検索をサポートしているが、類義語・関連語については検索を行うことができない。本研究では単語間の係受け関係という語間関係を検索に用いているため、そこにシソーラスという別の語間関係を持ち込むことは得策ではない。よって、類義語や関連語を検索者に提示したり、専門用語を知らなくても検索が行えるような、発見的な検索との融合を図ることを考えて類義語・関連語を検索できるようにする必要がある。

9 おわりに

本稿では、文献データベースにおいて格納されている文献表題やキーワードを精度良く検索する手法として、係受け関係を用いて検索インデックスを構造化する方法を提案し、検索実験を通じて本手法の有効性を検証した。本手法は単語間の係受けという文献表題やキーワードでは比較的簡単に抽出できる情報を用いているため動作も頑健であり、問合せを擬似的な自然言語で行うことも含めて検索者の意図を反映した検索結果を得るのに適したシステムであると考えられる。

今後の課題としては、検索オーバヘッドの削減、特に実用的な速度での検索に向けてのアルゴリズムの工夫、全文情報や内容梗概を用いた文献主題の充実、類義語・関連語の本手法への導入を挙げることができる。

参考文献

[1] 拝原 正人：「日本語文献データベースへの知的アクセス」，電子情報通信学会誌 Vol.72, No.7, pp.797-806, 1989.

[2] Donna Harman : "Ranking Algorithms" in *Information Retrieval* , Chapter 14. Prentice Hall, 1992.

[3] Ulrich Pfeifer, Norbert Fuhr, Tung Huynh : "Searching structured documents with the enhanced retrieval functionality of freeWAIS-sf and SFgate", *Computer Networks and ISDN Systems*, Vol.27, No.6, pp.1027-1036, 1995.

[4] Alan F. Smeaton : "Progress in the application of natural language processing to information retrieval tasks", *The Computer Journal*, Vol.35, No.3, 1992.

[5] 那須川 哲哉：「文や句による日本語テキストの検索」，第53回情報処理学会全国大会，5T-1, 1996.

[6] C.Berrut : "Indexing Medical Reports : the RIME Approach", *Information Processing and Management*, Vol.26, No.1, pp.93-110, 1990.

学術情報センター電子図書館サービス
©学術情報センター

[7] "Introduction to Oracle Context", Oracle Corporation, 1993.

[8] 宮崎 正弘 :「係受け解析を用いた複合語の自動分割法」, 情報処理学会論文誌, Vol.25, No.6, pp.970-979, 1984.

[9] 石崎 雅人 :「複合名詞における語と語の関係について」, 第37回情報処理学会全国大会, 3C-2, 1988.

[10] C.Faloutsos and S.Christodoulakis : "Signature Files : An Access Method for Document and its Analytical Performance Evaluation", *ACM Transactions on Offive Information Systems*, Vol.2, No.4, pp.267-288, 1984.

[11] 「統計学入門」, 東京大学教養学部統計学教室編, 東京大学出版会, 1991.

学術情報センター電子図書館サービス　　© 学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

認識誤りを含むテキストにおける検索手法

Retrieval Methods for Text with Missrecognized Characters

東京大学大学院工学系研究科　太田　学
Manabu OHTA
Graduate School of Engineering, University of Tokyo
学術情報センター　髙須　淳宏
Atsuhiro TAKASU
National Center for Science Information Systems
学術情報センター　安達　淳
Jun ADACHI
National Center for Science Information Systems

要旨

本稿では、OCRによる認識誤りを含む和文のテキストに対する3つの確率的な全文検索手法を提案する。提案手法では、認識誤りの存在を考慮に入れた上で検索を行なうため、OCRによる文書認識後に必要だった手作業による修正編集を行なう必要がない。検索時に認識誤りを吸収するために、誤る可能性のある文字とその確率を保持した類似文字テーブル及び文字の2-gram統計に基づいた文字の接続確率を保持した2-gramテーブルを用いる。検索の手順は、1つの入力検索語に対して類似文字テーブルを参照することで複数の検索文字列を生成し、それぞれの検索文字列を用いて全文検索を行なう。検索された文字列の適否は、類似文字テーブルによるOCRの誤り易さに基づいた確率と2-gramテーブルによる文字の接続確率によって判断する。具体的にはこれら2つの確率に基づいた得点を各検索結果に与え、その得点が閾値を越えていれば検索条件を満足するとし、さもなくば棄却する。本手法を用いて検索効率(再現率と適合率)に関する実験を行なった結果、認識誤りを考慮しない場合と比較してはるかに検索効率が改善されることを示せた。

ABSTRACT

This paper presents three probabilistic text retrieval methods designed to carry out a full-text search of Japanese documents containing OCR errors. By searching for any query term on the premise that there are errors in the recognized text, the methods presented can tolerate such errors, and therefore manual post-editing is not required after OCR recognition. In the applied approach, confusion matrices are used to store characters which are likely to be interchanged when a particular character is missrecognized, and the respective probability of each occurrence. Moreover, a 2-gram matrix is used to store probabilities of character connection based on 2-gram statistics, i.e., which letter is likely to come after another. Multiple search terms are generated for an input query term by making reference to confusion matrices, after which a full-text search is run for each search term. The validity of retrieved terms is determined based on error-occurrence and character-connection probabilities. Those with a validity value greater than a given threshold are judged to satisfy the input query. In addition, the performance of these methods is experimentally evaluated by determining retrieval effectiveness, i.e., by calculating recall and precision rates. Results indicate marked improvement in comparison with exact matching.

[キーワード]　光学的文字読取装置(OCR)、認識誤り、類似文字テーブル、2-gram、全文検索
[Keywords]　OCR, recognition error, confusion matrix, 2-gram, full-text search

学術情報センター電子図書館サービス　　◎学術情報センター

1　はじめに

これまでに印刷文書の形で蓄積された大量の情報をデータベース化(以下、DB 化)する場合に、逐一手作業でテキストに直すのでは多くの労力を伴う。そこで本のページを 1 枚の画像データとして扱うことのできる文書画像 DB が電子図書館などで用いられている。また OCR(光学的文字読取装置)を用いると文書画像をテキストに比較的簡単に変換できることから、大量の印刷文書を画像で入力し、OCR を使って全文 DB を構築する試みが最近行なわれるようになってきた[1][2]。これは、紙の世界と電子的な世界との架け橋となる文字認識(OCR)と、全文検索を効果的に融合する技術といえるが、その場合 OCR による文字認識の結果には認識誤りが含まれるため、この誤りへの対処が必要不可欠である。本稿ではこの認識誤りに対処するために、文字誤りを含むテキストにおける検索手法の提案を行なう。つまり本研究の目的は、OCR を用いて得たテキストをなるべくそのまま人手を介さずに蓄積して検索に利用するために、このテキストに含まれる OCR の認識誤りを訂正するのではなく検索段階で吸収するための手法を実現して、その検索能力を評価することである。ここで提案する 3 つの検索手法を簡単に説明する。

1)　Confusion Matrix Retrieval Method (CMR 法)

この手法は、検索時に OCR の認識誤りを含むテキストにおいて検索洩れを起こさないように、類似文字テーブル(Confusion Matrix)を用いることで 1 つの入力検索語から複数の検索文字列を生成するものである。この類似文字テーブルは、OCR の出力した文字誤りを含むテキストとそれに対応する誤りのないテキストを比較することで予め作成しておき、OCR が認識を誤る可能性のある文字(置換誤り)が全てその誤り易さに基づく確率とともに格納されている。全文検索を行なう際には、適切な検索文字列を生成すべくこの類似文字テーブルが参照される。複数の検索文字列は単語部分照合にも用いられ、検索終了後各検索文字列には OCR の誤る確率を元に計算した得点を与える。その得点が閾値を越えていれば、その部分を検索結果として出力しさもなくば棄却する。

2)　Expanded Confusion Matrix Retrieval Method (ECMR 法)

この手法は CMR 法とほぼ同じアルゴリズムを採用しているが、検索時に置換誤りのみを扱う類似文字テーブルだけでなく、欠落・挿入・結合・分解誤りを扱う拡張類似文字テーブル(Expanded Confusion Matrix)をも考慮する点が異なる。これによって、文字誤りを含むテキストと誤りのない元のテキストとが 1 対 1 に対応しない文字切り出しの誤りにも対応できる。

3)　2-gram Matrix Retrieval Method (2MR 法)

この手法も CMR 法とほぼ同じアルゴリズムを採用しているが、検索結果の文字列の得点を計算する際に、OCR の誤る確率を保持した類似文字テーブルだけでなく文字の 2-gram 統計に基づいた接続確率を保持した 2-gram テーブル(2-gram Matrix)をも参照する点が異なる。

本稿では、まず第 2 節で研究の意義と、本研究と密接な関係をもつ OCR をとりまく文字認識技術の現状について述べた後、文書画像を OCR で処理したテキスト(認識誤りを含むテキスト)に対する検索手法の研究を紹介する。第 3 節では、実際に日本語活字 OCR の認識誤りを分類してその分類に基づいて類似文字テーブル及び拡張類似文字テーブルを構築する方法について述べる。第 4 節では、2MR 法で用いる 2-gram テーブルと、そこに蓄えられた統計的に得られる文字の接続確率をどのように OCR の誤り易さに基づいた確率と結び付けて得点計算に利用するかを説明する。第 5 節では、第 3、4 節で説明したテーブルを用いた検索アルゴリズム、特に得点付けのアルゴリズムについて述べる。第 6 節では、提案した 3 つの検索手法を用いて実際に検索実験を行ない、検索精度の評価を行なう。第 7 節をもってまとめ、今後の課題について述べる。

2　研究の背景

2.1　研究の意義

文書 DB を構築する上で画像との関係を考えると次のような 2 つの立場がある。

- 文書を画像として蓄積する。
- 文書の構成要素として画像が含まれる。

前者の立場には、学術情報センターで提供している電子図書館システム NACSIS-ELS[3]の例がある。後者は、文書をテキスト、図、表、写真などの構成要素に分解して管理するマルチメディア DB 指向の捉え方である。本稿では前者の立場から構築された文書画像 DB を念頭におきその検索手法について提案を行なう。

利点	欠点
○ DBへの入力が容易(過去の資料の経済的な遡及入力が可能) ○ 印刷文書のもつ画像情報の活用 ○ 複数の言語・外字への対応が容易	● 検索機能が貧弱 ● 大容量の記憶装置が必要

図 1　文書画像 DB の性質

学術情報センター電子図書館サービス　　©学術情報センター

文書画像DBは一般的に図1に示すような性質をもっており、メディアの違い(文書画像とテキストコード)からいくつかの利害得失が挙げられるが、これをユーザに意識させるのは好ましくない。そこでこの両者を統合的に扱うために本稿では、図2のような全文検索システムを考える。このシステムでは、ユーザのブラウジング要求には文書画像を用いることにしてユーザにメディアの違いを意識させない。一方ユーザの様々な検索要求に柔軟に対応するために、文書認識によって得た全文データを利用することで検索機能を強化している。この全文データをもつことによって、全文検索(full-text search)の実行は勿論、これを文書画像のハイパテキスト化などに利用することもできる。

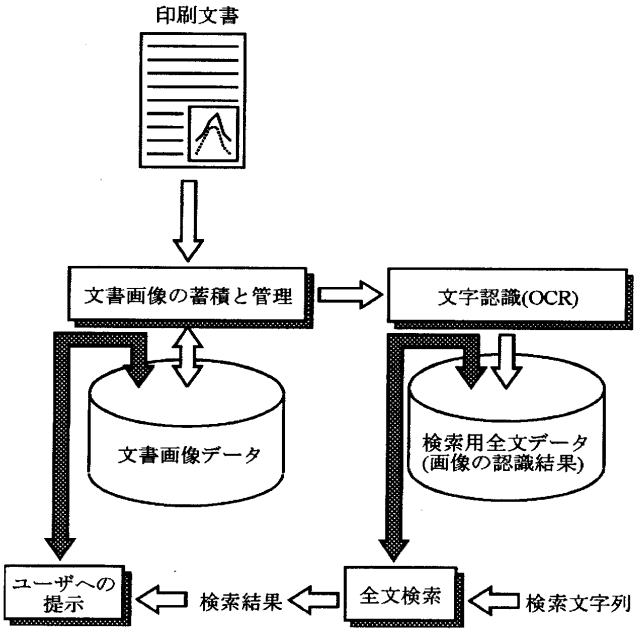


図2　文書認識を利用した全文検索システム

全文データを文書画像から得るには大量の印刷文書の電子形態への変換作業を伴うためOCRを用いるのが普通である。しかしOCRを用いると、得られたテキストデータに認識誤りが含まれるためそれらの誤りに対処する機構が必要になる。現在までのアプローチとしては、自然言語処理などOCRによる文字認識の後処理を行なうことで少しでもOCRの認識率を上げようとする試みが多い。すなわちOCRの認識誤りを少なくするというのが主流である。しかし認識率を100%にすることは難しく、実用的にはOCRによって得られた結果を効率良く修正するエディタを組み合わせる必要があり、その作業量は無視できない。そこで本稿では、DBへの入力作業をできるだけ省力化するという観点から、OCRで認識したテキストをそのまま蓄積・検索できるように、その認識誤りを訂正するのではなく検索段階で吸収するための手法を提案する。

2.2　OCRによる文字認識

2.2.1　OCR

印刷文書をスキャナ等で画像データに変換してDBに格納する場合、その画像データを認識してテキストコードにも変換することで、入力処理の自動化や省力化、検索機能の強化を図ることができる。ここでは、印刷文書にとって重要なOCRの認識技術を説明する[4]。

OCR(光学的文字読取装置)は、文字を含む画像データを計算機上で扱い易いようにテキストコードへ変換するための文字認識技術であり、その認識過程は一般に図3のようになる。まず入力された文書画像のテキスト領域に対して文字の切り出しを行ない、個々の文字に対して文字の認識を行なう。文字認識処理の結果としては複数の候補文字を類似度(その候補文字の辞書パターンと文書画像中に切り出された当該文字の間の類似性)とともに与える。そして不使用文字の除去や類似度による候補文字間の順位や各候補文字間の類似度の差や比などを利用して候補文字の絞り込みを行ない、さらに言語的制約や背景知識を利用して候補文字を決定する後処理を行なう。

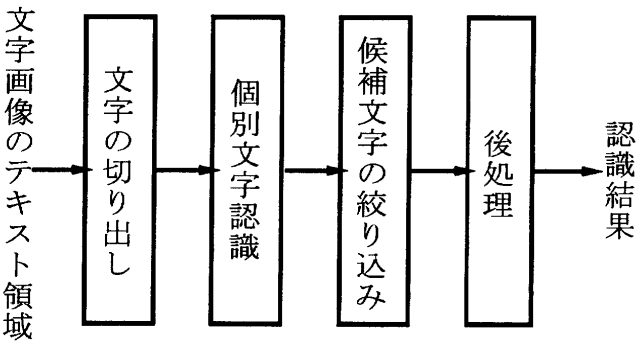


図3　OCRによる文字認識過程

2.2.2　OCRによる日本語文字認識の問題点

OCRの認識誤りの中には、日本語の一般的な文章の後処理を行なっても誤読文字を検出・修正できないものや、また時には、隣接する文字を誤訂正してしまって認識率の低下をもたらしてしまうものなどがある。その主な理由としては以下のものを挙げることができる。

1. たまたま単語辞書に登録されている単語と照合し、文法的エラーも検出されない(特に1文字単語)

学術情報センター電子図書館サービス　　©学術情報センター

2. 入力文中の辞書未登録語
3. 候補文字集合の中に正解文字が存在しない
4. 文字の切り出しの誤り

2.3 文字誤りを含むテキストにおける検索(類似研究)

文字認識率の高いOCRにおいても前節で述べたように認識誤りを起こすため、OCRの出力したテキストをそのまま DB に蓄積して検索を行なえるようにするには、認識誤りを許容する検索手法が不可欠である。本節では現在までに実際どのような手法が用いられてきたかを紹介する。

2.3.1 類似文字の共通コード化

これは OCR が混同しやすい文字(類似文字)を予め何らかの方法で共通のテキストコードに変換することによって、検索時に検索文字列と照合するようにする方法である[5]。この手法では、検索文字とそれに照合すべき文字誤りを含むテキストにおける文字が、うまく同じグループに分類されて共通のコードに変換された場合検索洩れを減らすことができる。この方法は、いかにうまく類似文字の分類を行なえるかが鍵となる。

2.3.2 複数認識候補型検索手法

この手法では、OCR が出力する複数の認識候補文字(第1位候補文字以外)中に正解文字が含まれている率が高いことを考慮して、これらの第 2 位以下の認識候補文字を検索に利用する[1][2]。検索手順を簡単に示すと以下のようになる。

1. まず不要な検索ノイズを低減するために、類似度を用いて各候補文字の絞り込みを行なう。
2. 各文字について精選した候補文字のみを認識結果のテキストにもたせて出力する。
3. この検索対象となるテキストを有限オートマトンで表現し、オートマトンと検索文字列との照合を行なう。

この手法のポイントは候補文字の絞り込み方法で、実際には類似度(距離情報)が用いられている。これは、OCRへの入力文字パターンと候補文字パターンの類似性を 0 から 256 の距離情報で表したもので実験では類似度の閾値が $\frac{10}{256}$ 以内のものをOCRの出力候補文字として選んだ場合の検索精度が良かったと報告されている。またこの場合、複数候補文字を保持することでデータ量は第1位候補文字だけを出力する場合の 1.28 倍になっている。

2.3.3 従来の検索手法の問題点

類似文字の共通コード化によって検索洩れは少なくなるが、新たに検索ノイズの問題が浮上する。この両者は一般的にトレードオフの関係にあるが、類似文字を完全に同一のものとみなすため、一定量の検索ノイズがのるのは避けられない。この原因の 1 つに、実際のOCRによる認識誤りには方向性があることが挙げられる。つまり文字'x'を'y'と誤る確率と'y'を'x'と誤る確率は一般に異なるのに、この両者を同一のコードで扱うために起きる弊害である。

複数認識候補型検索手法においては、OCR の性能をフルに活用するという点が評価できる反面、検索精度がOCRの性能によって決定されてしまう危険がある。最近の日本語活字OCRの認識率は良いもので99%を越えるが、それでも欠落、挿入、結合、分解誤りといった文字の切り出しに関わる誤りが存在し、この場合複数の認識候補中にも正解文字が含まれないことが多い。また、検索対象テキストに検索洩れを防ぐために冗長性をもたせるため、複数の認識候補を持たせなかった場合のデータ量に比例した一定量多いテキストデータを蓄える必要がある。

3 類似文字テーブルの構築

3.1 OCRの認識誤り

まず日本語活字OCRの認識誤りを具体的に把握するために予備実験を行なった。具体的には学術論文誌から得たテキスト 41854 文字について、OCR に認識させるための元のテキストと認識結果のテキストの比較を行なって認識誤りを以下のように分類した。この実験で使用したのは(株)リコーのIMAZONE日本語活字OCRで認識率は98.3%であったが、以下の分類は一般のOCRについても成り立つ。

- 置換(誤字)
  1 文字がOCRによって別の1文字に置換される誤り。
  例:な→は　ボ→ポ　問→間
- 欠落(脱字)
  1 文字以上がOCRによって抜け落ちる誤り。
  例:E-R→ER
- 挿入
  1 文字以上がOCRによって挿入される誤り。
  例:グラフィックス→グラ、フィックス
- 結合
  2 文字以上がOCRの文字の切り出し誤りによって

学術情報センター電子図書館サービス　　© 学術情報センター

1文字として認識される誤り。
例:し、→い　(こ→に　R→Pr

- 分解
  1文字がOCRの文字の切り出し誤りによって2文字以上として認識される誤り。
  例:利→夫U　指→ま旨　ル→Jし

またその出現頻度を調べた結果を表1に示す。ここで“その他”は、上記の方法で分類できなかった誤りを表している。なお認識誤りの原因の約4割は英数字(半角文字)の認識に絡むものであったが、これは本研究で使用したOCRが日本語活字対応のものであったことによる。

表1から、OCRの認識誤りのうち文字列の長さが変わらない置換誤りが約81%と最も多く、つづいて文字数が増加する挿入、分解誤りが約9%、文字数が減少する欠落、結合誤りが約8%とほぼ同程度の割合で起こることが分かる。また、調べたテキストにおいては、連続した2文字以上の欠落、挿入、3文字以上への分解はなく、3文字以上の結合は4件だけ(しかも英数字)であった。つまりこの実験結果から、置換誤りと1文字の欠落、挿入誤り、2文字の結合、分解誤りを考慮すれば、OCRの認識誤りの約98%をカバーできることが分かる。

表1　日本語活字OCRの認識誤り†

誤りの種類	出現頻度	出現頻度(%)
置換 (誤字)	586	80.7
欠落 (脱字)	10	1.4
挿入	60	8.3
結合	49	6.8
分解	8	1.1
その他	13	1.8

†41854文字に対する予備実験の結果

3.2　文字の確信度

文字の確信度とは、正しいテキストにおける文字がiである事象をA_i、それに対応するOCRのテキストにおける文字がjである事象をB_jとするとき、OCRのテキストにおける文字jが正しいテキストにおいて文字iである(と確信できる)確率 $P(A_i|B_j)$のことである。この$P(A_i|B_j)$は、Bayesの定理から式(1)で計算される。

$$P(A_i|B_j)=\frac{P(A_i)P(B_j|A_i)}{\sum_{k=1}^{n}P(A_k)P(B_j|A_k)}. \qquad (1)$$

この文字の確信度は、得点付けの際に全ての提案手法が利用する。

3.3　類似文字テーブル

全ての手法で用いている類似文字テーブルは、OCRの出力した文字誤りテキストとそれに対応する誤りのないテキストを比較することで得られるもので、日本語活字OCRにおいて最も頻度の高い置換誤りに対処するために構築する。このテーブルには、置換誤りの可能性のある文字が全てその確信度とともに格納されている(図4参照)。またCMR法は予備実験の誤り分類(表1)で約8割を占めた置換誤りに対処するための手法であるため、使用するのはこの類似文字テーブルのみである。

B_j=認識結果の文字j

A_i=元の文字i	B_1	B_2	B_3	.....
A_1	$P(A_1 B_1)$	$P(A_1 \mid B_2)$	0	
A_2	0	$P(A_2 \mid B_2)$	0	
A_3	0	$P(A_3 \mid B_2)$	$P(A_3 \mid B_3)$	
.....				

図4　類似文字テーブル

3.4　拡張類似文字テーブル

欠落、挿入、結合、分解誤りは、結果としてテキストの文字数を変化させるため、CMR法の単語部分照合によっても検索できない場合が多い。そこでECMR法では、図4に示す類似文字テーブルと同様な構成をもつ欠落文字テーブル、挿入文字テーブル、結合文字テーブル、分解文字テーブルを作成してこれらの誤りに対処する(図5、6参照)。またこれら4つの文字テーブルを総称して拡張類似文字テーブル(Expanded Confusion Matrix)と呼ぶことにする。

図5(a)、(b)はそれぞれ、欠落文字テーブルと挿入文字テーブルを表しており、ここで仮想文字V_mとV_iはそれぞれ認識結果の文字と元の文字を表す。このような仮想文字を考える事で、これらのテーブルを類似文字テーブルと同様に扱うことができ、ECMR法でもCMR法と同じ得点付けのアルゴリズムを用いる事が可能になる。また図6(a)、(b)はそれぞれ結合文字テーブルと分解文字テーブルを表し、同様の理由からここでも

学術情報センター電子図書館サービス　　© 学術情報センター

$a_i=(A^0A^1)_i$, $b_j=(B^0B^1)_j$という元のテキストと認識結果のテキストにおける 2 つの連続する文字を定義する。さらに認識結果の文字 B_j が元のテキストにおいて $a_i=(A^0A^1)_i$ という文字列とみなせる確信度(確率)$P(a_i|B_j)$が結合文字テーブルには蓄えられており、同様に認識結果の文字列 $b_j=(B^0B^1)_j$ が元のテキストにおいて A_i という文字とみなせる確信度 $P(A_i|b_j)$が分解文字テーブルに蓄えられている。

4　2MR 法

4.1　2-gram 統計に基づいた文字の接続確率

2MR 法では、類似文字テーブルに保持する OCR の誤り確率とは別の観点からの確信度を導くために、2-gram 統計から得られる文字の連接情報(接続確率)を利用している。日本語の文字種は約 3000 種類と英語に比べて桁違いに多いため、充分な学習量のテキストデータに対して n-gram 統計 ($n \geq 3$) を求めるのは困難である[1]。そこで文字の連接情報として 2MR 法では、比較的頻度統計がとり易く、自然言語のモデル化能力が高い 2-gram[7] を用いている。

本稿でいう文字の接続確率とは、元のテキスト中の文字 A^i が A^{i-1} の次に続く確率で式(2)で求められる。

$$P(A^i|A^{i-1}) = \frac{f(A^{i-1}A^i)}{f(A^{i-1})}. \tag{2}$$

式(2)において、$f(A^{i-1}A^i)$ と $f(A^{i-1})$ はそれぞれ文字列 A^{i-1} A^i (2-gram)の出現頻度と文字 A^{i-1} (1-gram)の出現頻度を表す。そしてこのようにして求めた文字の接続確率を図 7 に示す 2-gram テーブルに保持し検索結果の得点付けの際に参照することにする。

4.2　2MR 法における文字の確信度

認識結果の文字 B_j^1 が元のテキストにおいて A_k^1 とみなせる確率は、類似文字テーブル(第 3.3 節)によると $P(A_k^1|B_j^1)$ であった。しかし 2MR 法では認識結果のテキストにおける B_j^1 の直前の文字 B^0 に着目して式(3)の確率を計算する。

(a) 欠落文字テーブル

	A_1	A_2	A_3	.....
V_m	$P(A_1\|V_m)$	$P(A_2\|V_m)$	$P(A_3\|V_m)$	

V_m = 欠落文字A_iに対応する仮想的な文字

(b) 挿入文字テーブル

	B_1	B_2	B_3	.....
V_i	$P(V_i\|B_1)$	$P(V_i\|B_2)$	$P(V_i\|B_3)$	

V_i = 挿入文字B_jに対応する仮想的な文字

図 5　欠落文字テーブルと挿入文字テーブル

(a)結合文字テーブル

	B_1	B_2	B_3	.....
$a_1=(A^0A^1)_1$	$P(a_1\|B_1)$	$P(a_1\|B_2)$	$P(a_1\|B_3)$	.....
$a_2=(A^0A^1)_2$	$P(a_2\|B_1)$	$P(a_2\|B_2)$	$P(a_2\|B_3)$	.....
$a_3=(A^0A^1)_3$	$P(a_3\|B_1)$	$P(a_3\|B_2)$	$P(a_3\|B_3)$	.....
⋮	⋮	⋮	⋮	

(b)分解文字テーブル

	$b_1=(B^0B^1)_1$	$b_2=(B^0B^1)_2$	$b_3=(B^0B^1)_3$	.....
A_1	$P(A_1\|b_1)$	$P(A_1\|b_2)$	$P(A_1\|b_3)$	.....
A_2	$P(A_2\|b_1)$	$P(A_2\|b_2)$	$P(A_2\|b_3)$	.....
A_3	$P(A_3\|b_1)$	$P(A_3\|b_2)$	$P(A_3\|b_3)$	.....
⋮	⋮	⋮	⋮	

a_i = 元の文字列 A^0A^1 の i 番めの組合せ
b_j = 認識結果の文字列 B^0B^1 の j 番めの組合せ

図 6　結合文字テーブルと分解文字テーブル

$$P(A_k^1|B^0) = \sum_i P(A_k^1|A_i^0)P(A_i^0|B^0). \tag{3}$$

つまりこれは認識結果のテキストにおいてその前の

[1] 最近では大容量テキストから n-gram 統計 ($n \geq 3$) をとることも可能になりつつある[6]。

学術情報センター電子図書館サービス　　© 学術情報センター

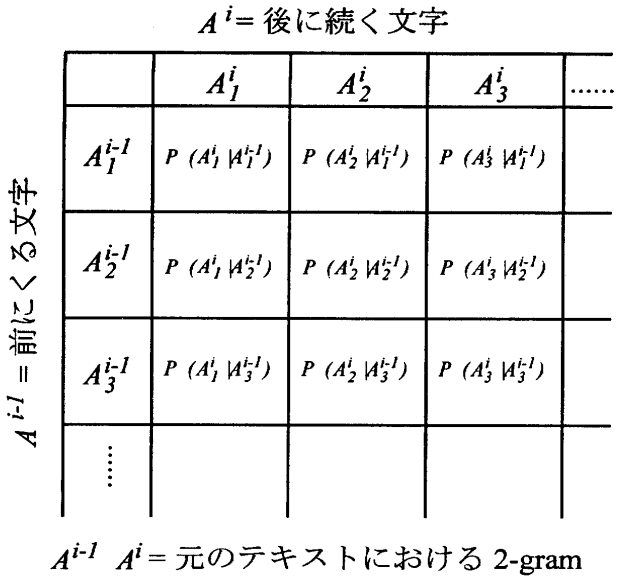


図7　2-gram テーブル

文字が B^0 であるとき、次の文字が元のテキストにおいて A_k^1 とみなせる確率である。式(3)の右辺の $P(A_k^1|A_i^0)$ が2-gram テーブル(図7)を参照することで得られる文字の接続確率で、一方 $P(A_i^0|B^0)$ が式(1)で定義した、認識結果のテキストにおいて B^0 である文字が元のテキストにおいて A_i^0 とみなせる確率、つまり確信度である。このように類似文字テーブルと2-gram テーブルを参照することで、認識結果のテキストにおける直前の文字 B^0 から、正しいテキストにおける次の文字が A_k^1 である確率を求めることができる。その結果、類似文字テーブルのみから得られる確信度 $P(A_k^1|B_j^1)$ とは別に、認識結果のテキストにおけるその直前の文字から、$P(A_k^1|B^0)$ という異なる観点からの確率(確信度)が得られる。

さてここで類似文字テーブルのみによって定まる $P(A_k^1|B_j^1)$ を $m_1(A_k^1)$、式(3)で求めた $P(A_k^1|B^0)$ を $m_2(A_k^1)$ とおく。するとこの2つの確率(確信度)を統合した新たな確率 $m(A_k^1)$ は、Dempster の結合規則[8][9]を用いると式(4)のように計算される。

$$m(A_k^1)=\frac{m_1(A_k^1)m_2(A_k^1)}{1-\sum_i m_1(A_i^1)(1-m_2(A_i^1))}. \qquad (4)$$

2MR 法では、式(1)で定義した元の確信度 $m_1(A_k^1)=P(A_k^1|B_j^1)$ の代わりに、文字の接続確率も考慮したこの $m(A_k^1)$ を新たな文字の確信度として使用する。

5　検索アルゴリズム

5.1　検索の手順

本稿でいう検索は全文検索を念頭においており、テキストは原則として1つの長大な和文の文字列として記憶されている。認識結果のテキストに対してユーザの入力した検索語で検索を行なう前に、類似文字テーブルを参照して以下の手順で複数の検索文字列を生成する(図8参照)。拡張類似文字テーブルを用いた場合も同様であるのでここでは類似文字テーブルによる検索文字列の生成のみを示す。

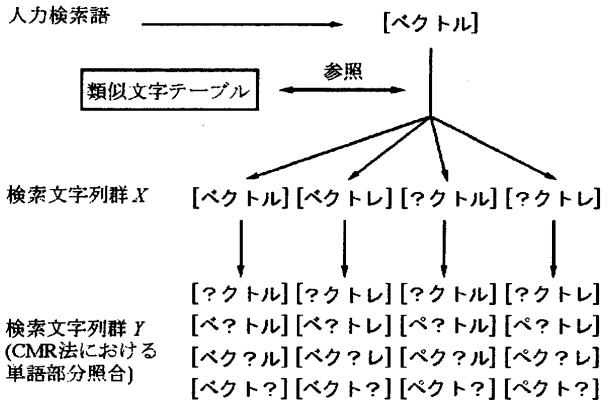


図8　検索文字列群の作成

1. ユーザから検索語が入力されると、類似文字テーブルを参照してその入力文字列中の各文字について置換誤りの可能性のある全ての文字を置き換えて新たに検索文字列群 X を作成する。図8では入力検索語は“ベクトル”で、類似文字テーブルを参照した結果、“ベ→ペ”、“ル→レ”という置換を行なっている。
2. CMR 法においては、検索文字列群 X の各検索文字列中の1文字を前から順にワイルドカード(長さ1)に置き換えた単語部分照合用の検索文字列群 Y を作成する。これは CMR 法が置換誤りにしか対処できないためで、単語部分照合を用いることでそれ以外の誤りにもある程度対処できるようになる。
3. CMR 法では検索文字列群 X と Y の、ECMR 法及び2MR 法では検索文字列群 X の全検索文字列を用いて全文検索を行なう。各検索結果には第5.2節で詳説する得点が付けられており、閾値を用いて各検索結果の適否を判断する。

5.2　文字列の確信度と部分確信度

検索文字列の得点として、CMR 法は文字列の確信度

学術情報センター電子図書館サービス　　©学術情報センター

と部分確信度の2つを用い、ECMR法と2MR法は文字列の確信度のみを用いる。

5.2.1 文字列の確信度

CMR法では、認識結果のテキストを検索して得られる文字列が正しい確率を、文字列中の各文字が正しい確率の積と仮定する。つまり、認識結果の文字列 $B^{012...n}$ が元の文字列 $A^{012...n}$ に対応する確率を、

$$P(A^{012...n}|B^{012...n}) = P(A^0|B^0)P(A^1|B^1)...P(A^n|B^n). \quad (5)$$

で表す。ここで $P(A^{012...n}|B^{012...n})$ のことを**文字列の確信度**と定義する。

2MR法でも文字列の確信度は式(5)で表されるが、式(4)で定義した $m(A^i)$ を式(1)で定義した $P(A^i|B^i)$ の代わりに用いて計算しなければならない。

しかしECMR法では、文字列の確信度の計算方法が若干異なる。これは拡張類似文字テーブルが入力検索語から複数の検索文字列を生成すると、入力文字列と生成文字列との間に1対1の文字の対応関係が成り立たなくなるためで、よってECMR法では文字列の確信度は以下のようにして算出する(図9参照)。

- 欠落誤りを含む場合
  欠落文字に対応する仮想的な文字 V_m を認識結果のテキストにおいて考え、欠落文字の確信度 P_m を次式で与える。

$$P_m = P(A_i|V_m) \times P_{m_0}. \quad (6)$$

  式(6)において $P(A_i|V_m)$ は図5に示した欠落文字テーブルより得られる。また P_{m_0} は、欠落の起こる確率で、表1から式(7)で求められる。

$$P_{m_0} = \frac{10}{41854} \fallingdotseq 0.00024. \quad (7)$$

  実際には認識結果のテキストにおける文字 V_m は仮想的なもので観測できないため、P_{m_0} は必ず考慮しなければならない。文字列の確信度の算出は、この欠落文字の確信度を考慮した上で式(5)と同様に算出する。
- 挿入誤りを含む場合
  認識結果のテキストに挿入された文字に対応する仮想的な文字 V_i を元のテキストにおいて考え、図5に示した挿入文字テーブルから挿入文字の確信度を得る。式(5)による文字列の確信度の算出の際にその挿入文字の確信度も余分に掛け合わせる。
- 結合、分解誤りを含む場合
  それぞれ図6に示した結合文字テーブルと分解文字テーブルを参照して結合文字の確信度、分解文字の確信度を得る。文字列の確信度は、結合文字あるいは分解文字の確信度を考慮した上で式(5)と同様に算出する。

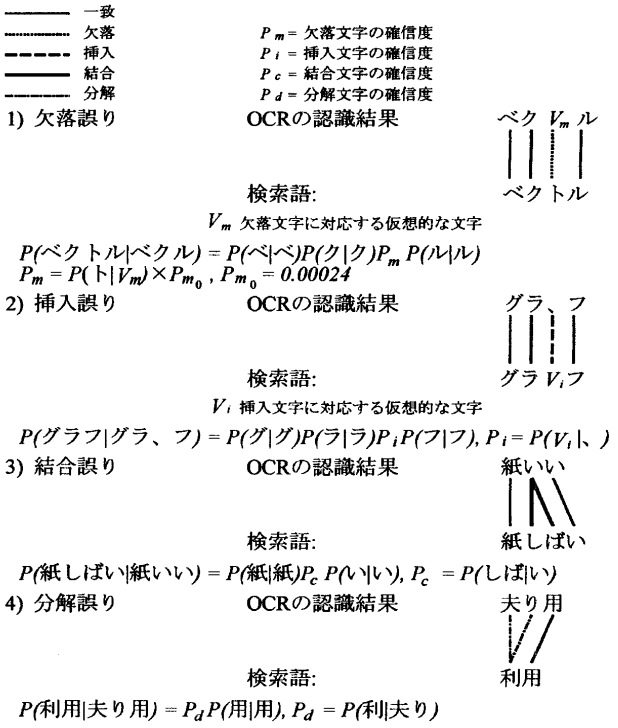


図9　ECMR法における文字列の確信度の計算

5.2.2 文字列の部分確信度

CMR法で用いられる単語部分照合は、n 文字の検索文字列のうち任意の(n-1)文字が完全一致すれば成り立つというものである。つまり、CMR法が図8の検索文字列群 X と Y の文字列を生成して検索する時、検索対象となる認識結果のテキストにおいて検索文字列群 X の文字列には照合しないがワイルドカードのおかげで検索文字列群 Y の文字列には照合する場合に成り立つ。しかしこの場合式(5)に基づくとこの文字列の確信度 $P(A^{012...n}|B^{012...n})$ は0となるため、**文字列の部分確信度** $PCD(A^{012...n}|B^{012...n})$ を式(8)のように定義する。換言すれば部分確信度は、文字列中の各文字の確信度を比較してその中で最低のものを除いて掛け合わせた文字列の確信度であるといえる。故に式(5)の確信度が0となる検索結果の文字列の適否の判断に利用できる。

学術情報センター電子図書館サービス　　© 学術情報センター

$$P(A^k|B^k) = \min\{P(A^0|B^0), P(A^1|B^1), \ldots, P(A^n|B^n)\}\text{のとき、}$$
$$PCD(A^{012\ldots n}|B^{012\ldots n}) = P(A^0|B^0)\ldots P(A^{k-1}|B^{k-1})P(A^{k+1}|B^{k+1})\ldots P(A^n|B^n). \quad (8)$$

$$\text{再現率} = \frac{\text{認識結果のテキストを検索して得た文字列のうち正しいものの数}}{\text{元のテキストを検索して得た文字列の数}}. \quad (9)$$

$$\text{適合率} = \frac{\text{認識結果のテキストを検索して得た文字列のうち正しいものの数}}{\text{認識結果のテキストを検索して得たすべての文字列の数}}. \quad (10)$$

6 実験

6.1 実験条件

提案した 3 つの検索手法の検索効率について評価実験を行なった。その時の条件を以下に示す。

1. 使用した OCR は、(株)リコーの IMAZONE 日本語活字 OCR で認識率は 98.3%であった。
2. 全文検索の対象としたテキストは、類似文字テーブルや 2-gram テーブルの作成に用いた学習用テキストに含まれる closed data(約 80KB)とそれに含まれない open data(約 55KB)である。
3. 文字の 2-gram 統計は、情報分野の学術論文雑誌から得たテキストデータ約 500 万文字(約 8MB)についてとり、文字の接続確率が 0 になったものは 10^{-5} に底上げした。
4. 検索語としては、検索用テキストに含まれる文字列で、名詞のものを無作為に 50 単語を抽出して用いた。検索効率の値は、それぞれその 50 単語で検索した結果の平均値である。
5. 検索効率とは検索結果における再現率・適合率のことでそれぞれ検索洩れの少なさ及び検索ノイズの少なさを表し、式(9)及び式(10)を用いて計算される。
6. CMR 法によって検索された文字列には全て、確信度 P 及び部分確信度 PCD が付けられており、ECMR 法及び 2MR 法によって検索された文字列にはそれぞれ確信度 P_E 及び P_2 が付けられている。
7. CMR 法では $P \geq$ 閾値、$PCD \geq$ 閾値及び $(P>0) \cap (PCD \geq$ 閾値) という 3 つの検索条件を用いる。一方 ECMR 法と 2MR 法ではそれぞれ、$P_E \geq$ 閾値、$P_2 \geq$ 閾値という 1 つの検索条件を用いる。

6.2 実験結果

6.2.1 確信度 P、P_E、P_2 の閾値と検索効率

まず確信度 P(CMR 法)の閾値と検索効率の関係を調べるために、閾値を変えながら closed data に対して検索実験を行なった。図 10 にその結果を示す。図 10 から、閾値が≈0.1 付近において再現率・適合率ともに高くなることが分かる。また P_E(ECMR 法)、P_2(2MR 法)の閾値と検索効率の関係を調べる実験でもほぼ同様な結果が得られた。

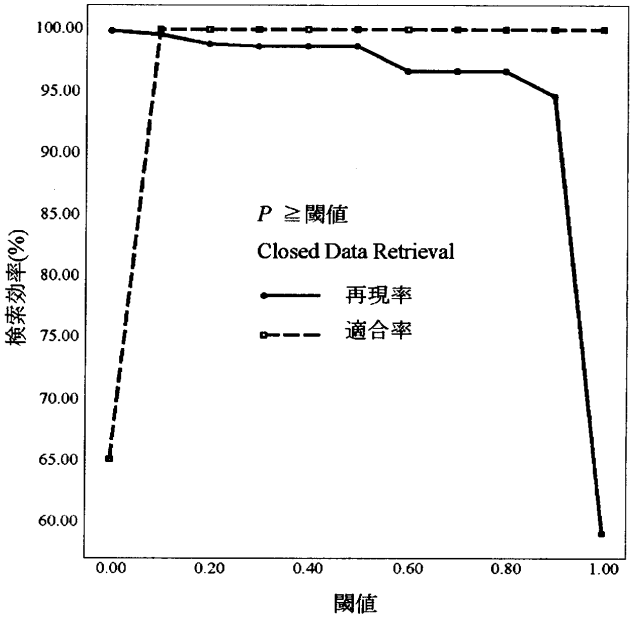


図 10　P の閾値と検索効率

6.2.2 部分確信度 PCD の閾値と検索効率

同様の実験を PCD の閾値と検索効率の関係についても行なって図 11 に示す結果を得た。PCD による評価ではワイルドカードによる単語部分照合を許容するため再現率は高いが適合率は全般的に低い。

6.2.3 確信度 P と部分確信度 PCD を組み合わせた評価

再現率・適合率ともに高い値を保つために、$(P>0) \cap (PCD \geq$ 閾値) という条件を用いて評価した結果を図 12 に示す。これは、検索文字列中の各文字が類

学術情報センター電子図書館サービス　　◎学術情報センター

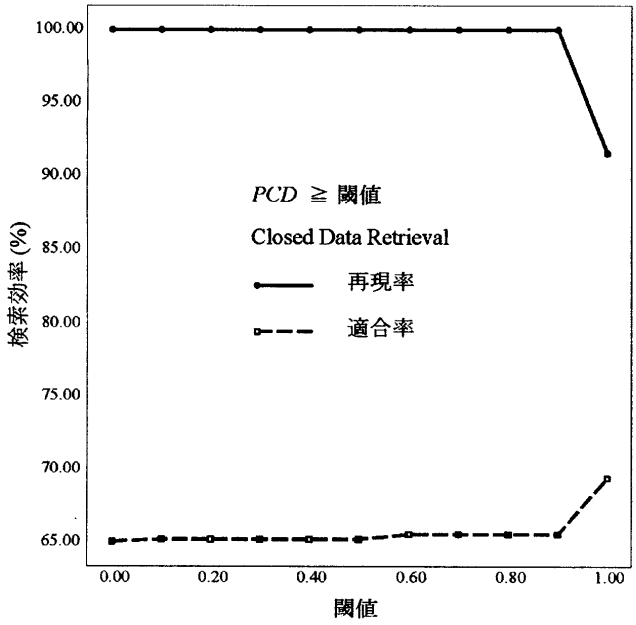


図 11　*PCD* の閾値と検索効率

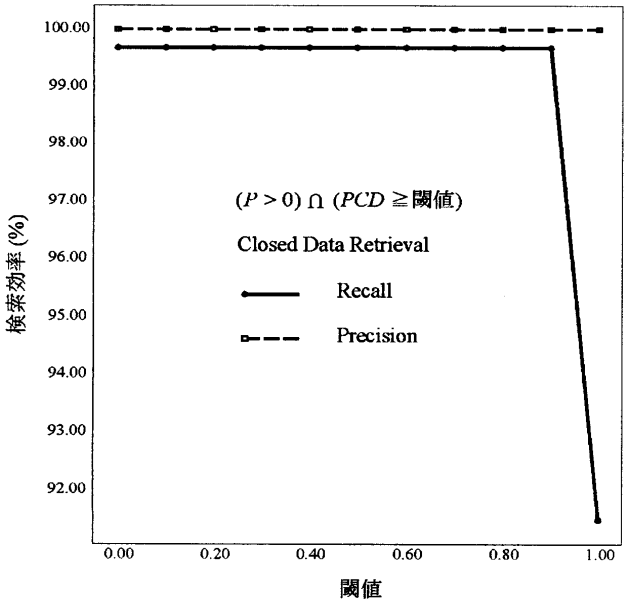


図 12　$(P>0) \cap (PCD \geq$ 閾値$)$ の検索効率

似文字テーブルに存在し$(P>0)$、かつ *PCD* が閾値を越えるもの(適合率を高い値で維持するため)を検索結果としたものである。

6.2.4　最適閾値における検索効率

closed data と open data それぞれの最適閾値における検索効率を表 2 及び表 3 に示す。検索対象テキストである closed data 及び open data を完全照合で検索した場合、再現率はそれぞれ 96.62%(表 2)、96.01%(表 3)であった。これに対して CMR 法$(P \geq$ 閾値など$)$を用いると、closed data では 99.69%の再現率(表 2)、open data では 99.26%の再現率(表 3)を実現している。結局、open data においては若干適合率を下げる(99.60%から 99.28%)ものの、どちらのテキストにおいても完全照合と比べて再現率を約 3%改善できることが分かる。

ECMR 法は closed data に対しては適合率を下げることなく再現率を完全に回復することができた(表 2)。これは、拡張類似文字テーブル(ECMR 法)によって類似文字テーブルだけ(CMR 法)ではカバーできなかった認識誤りにも対処できたことを示している。一方 open data においては CMR 法と比較して検索効率の差がみられないが、これは拡張類似文字テーブルにも含まれない認識誤りが原因であった。具体的には次のような認識誤りが存在したためである。

1. 比較→上ヒ薮
   “比→上ヒ”という認識誤りが拡張類似文字テーブル(分解文字テーブル)に存在せず、“較→薮”という誤りも類似文字テーブルに存在しなかったので全く検索することができなかった。
2. 非線形→ヲ E 線形
   “非→ヲ E”という誤りが分解文字テーブルに存在しなかったため確信度 P_E=0 となってしまい、検索結果として不適と判断された。

これらはどちらもテーブルを構築する際の学習用テキストに存在しなかった誤りであり、このことからテーブルの構築には充分な量の学習用テキストが必要であることがいえる。

2MR 法は CMR 法と比較して、適合率には殆んど差がみられなかったが再現率は全般的に高く保たれた(図 13)。これは、検索結果として適した文字列の確信度が文字の接続確率を考慮することで引き上げられたことを示している。つまり CMR 法の与える *P* よりも 2MR 法の与える P_2 の方がより得点として好ましいといえる。統計という性質上、常に 2MR 法が CMR 法よりもすぐれた得点付けを行なっている訳ではなく、中には文字の接続確率を考慮することでかえって確信度が改悪されることもあったが、図 13 にも示される通り全般的には 2MR 法の方が CMR 法よりも優れているといえる。

7　おわりに

本稿では、OCR の出力した認識誤りを含むテキストを有効に活用するという観点から、認識誤りを含むテキストにおける検索手法を 3 つ提案し、検索精度の評価を行なった。

学術情報センター電子図書館サービス　　© 学術情報センター

表2　closed data における検索効率

検索条件	再現率(%)	適合率(%)
完全照合	96.62	100.0
$P \geq 0.01$ (CMR 法)	99.69	100.0
$PCD \geq 0.9$ (CMR 法)	99.95	65.64
$(P > 0) \cap (PCD \geq 0.9)$ (CMR 法)	99.69	100.0
$P_E \geq 0.01$ (ECMR 法)	100.0	100.0
$P_2 \geq 0.01$ (2MR 法)	99.69	100.0

表3　open data における検索効率

検索条件	再現率(%)	適合率(%)
完全照合	96.01	99.60
$P \geq 0.00001$ (CMR 法)	99.26	99.28
$PCD \geq 0.001$ (CMR 法)	99.91	68.84
$(P > 0) \cap (PCD \geq 0.001)$ (CMR 法)	99.26	99.28
$P_E \geq 0.00001$ (ECMR 法)	99.26	99.28
$P_2 \geq 0.01$ (2MR 法)	99.26	99.28

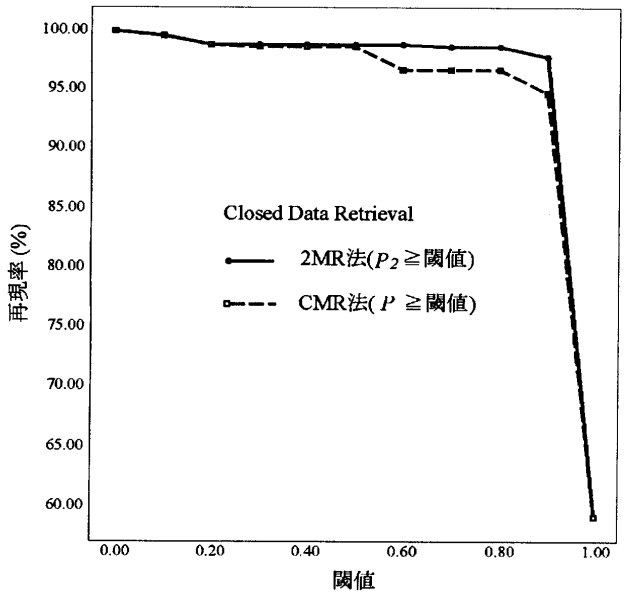


図13　closed data における CMR 法と 2MR 法の再現率の比較

これまでの文字認識に関する研究・開発は文字の認識率向上に主眼が置かれてきたが、最近では単に文字認識精度の向上だけでなく、文字認識をどのように利用するかという後処理を含めたシステムとして考える動きがある。その一例として、大量の印刷文書を画像で入力し、OCR を使って全文 DB を構築する試みが最近行なわれるようになってきており、これは紙の世界と電子的な世界との架け橋となる文字認識(OCR)と、全文検索を効果的に融合する技術として注目に値する。そこで本稿ではこの場合問題になる OCR の認識誤りに対処するため、この誤りを訂正するのではなく検索段階で吸収するための手法を提案した。

提案手法の具体的なアルゴリズムは第3、4節に述べた通りであるが、いずれの手法も OCR の認識誤りを学習したテーブルを検索時に参照することで検索洩れを減らすとともに、各検索結果に確率に基づく得点を与え、その得点を閾値によって評価することで検索ノイズの低減を図るものであった。またその得点付けでは、OCR の文字認識においてその誤り易さを考慮した確率、および統計的に得られる文字の接続確率の 2 つを用いた。

第 6 節で示した実験の結果、どの手法も検索効率の改善に有効であり、殆どの誤りを提案した検索手法によって吸収できることが分かった。しかしこの実験から、提案手法は類似文字テーブル構築の際に学習用テキストに存在しなかった誤りには対処できないので、充分な量のテキストデータを学習に用いなければならないことも示唆された。また本稿では検索速度の評価は行なっていないが、実験で用いた検索語の各文字は類似文字テーブルを参照することで平均 1.26 文字に拡

学術情報センター電子図書館サービス ©学術情報センター

張されており、これは類似文字テーブルしか用いないCMR法でも、長さnの1つの検索語から平均1.26^nの検索文字列が生成されることを意味している。よって提案手法の実現にあたっては、検索速度についての考慮も必要不可欠である。

参考文献

[1] H. Fujisawa and K. Marukawa. Full-text search and document recognition of Japanese text. In *Proc. of 4th Symp. DA & IR*, pp. 55-80, Las Vegas, Nevada, 1995.

[2] 丸川 勝美, 藤澤 浩道, 嶋 好博. 文書認識と全文検索の融合技術に関する実験的検討. 情報処理学会研究報告 *95-FI-39*, 95(87):65-72, September 1995.

[3] 安達 淳, 橋爪 宏達, 片山 紀生. 学術情報センターの電子図書館システムの概要と試行実験. 情報処理学会研究報告 *95-FI-37*, 95(45):23-30, May 1995.

[4] 飯沢 篤志. 文書画像データベースシステム. 情報処理, 33(5):497-504, 1992.

[5] A. Myka and U. Güntzer. Fuzzy full-text searches in OCR databases. In *Advances in Digital Libraries (Preliminary Version)*, chapter 7, pp. 87-100. SPRINGER-VERLAG, New York, 1995.

[6] 長尾 眞, 森 信介. 大規模日本語テキストのnグラム統計の作り方と語句の自動抽出. 情報処理学会研究報告 *93-NL-96*, 93(96):1-8, July 1993.

[7] Michèle Jardino. Multilingual stochastic n-gram class language models. In *Proceedings of ICASSP'96*, Vol. 1, pp. 161-163, 1996.

[8] 石塚 満. DempsterとShaferの確率理論. 電子通信学会誌, 66(9):900-903, 1983.

[9] 小林 邦勝, 鈴木 伸明, 根元 義章, 佐藤 利三郎. DempsterとShaferの確率理論に基づく情報量に関する一考察. 電子通信学会論文誌, J68-A(8):741-747, 1985.

学術情報センター電子図書館サービス　　©学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

A Study on ARQ Mechanism in ATM Networks

ATM ネットワークにおける自動再送方式に関する研究

Weiping ZHAO
National Center for Science Information Systems
学術情報センター　趙　偉平
Suwan RUNGGERATIGUL
National Center for Science Information Systems
学術情報センター　ルンキーラティクン　スワン
Shoichiro ASANO
National Center for Science Information Systems
学術情報センター　淺野　正一郎

ABSTRACT

ATM networks are designed to support a wide variety of services with different requirements of QoS. Some applications requiring very high QoS, and the progresses in IC technology and processing capacity make it possible to introduce ARQ mechanism in high-speed ATM networks economically and efficiently. In this paper, an ARQ scheme based on selective repeat strategy is proposed. Its performance in terms of throughput, traffic increment and required buffer size is evaluated by both theoretical analysis and simulation. The analytical and simulation results show that the proposed scheme has a very good error detection capability while maintaining high throughput with an acceptable requirement of buffer size.

要旨

ATM ネットワークは、異なるサービス品質(QoS)を要求するさまざまなサービスを提供している。電子バンクや電子キャッシュなど信頼性の高いアプリケーションを実現するために自動再送方式(ARQ)が有力な手段として期待されている。本論文では、まず、ATM ネットワークに ARQ を導入する際の要素を観察し、次に、選択再送(SR)を基づいた ARQ 方式を提案する。本方式では受信側で誤りセルを検出した際、それ以降に受信した正しいセルは保持しておき、誤りセルのみ再送要求する。異なる転送距離と異なる転送誤り率について性能評価を誤り訂正率、スループット及びバッファーサイズを性能指標として理論解析とシミュレーションにより行った。その結果、本提案方式は誤り訂正性能を大幅に改善し、高いスループットを維持することが分かる。

[Keywords]　ATM network, ARQ, QoS, throughput, error detection
[キーワード]　ATM、自動再送、サービス品質、スループット、誤り検出

1 Introduction

Broadband networks (B-ISDN) based on the Asynchronous Transfer Mode (ATM) technology are designed to support a vast variety of services with different bit rates and quality of service (QoS) requirements. Researches on QoS assurance have been recently studied widely and many control methods have been proposed [1] [2] [3] [4]. Most of proposed methods were traffic controls including call admission control (CAC) and source policing (UPC). For video multicast communications, Sebayashi [5] proposed a method which can guarantee the receiver-initiated QoS using "media scaling" on the dynamic resource environment of the receivers.

For conventional computer networks, automatic repeat request (ARQ) protocol is often used as a dedicated method enabling high reliable data communication [6] . But, ARQ protocol is currently not used in ATM network because of two rea-

学術情報センター電子図書館サービス ◎学術情報センター

sons. One is that the transmission speed of ATM is high and the size of ATM packets (called cells) is fixed and small, which makes the window size of ARQ large. Another is that the bit error rate of optical fiber transmission line of ATM network is small (normally 10^{-7}), which enables ATM network to be error free for most of applications. However, some applications such as electronic bank and electronic cash that do require very high QoS, and the progresses in IC technology and processing capacity make it possible to introduce ARQ mechanism in high-speed ATM network economically and efficiently. If ARQ is only introduced and applied in the services which need very high QoS, the window size measured in buffer size for ARQ and the processing complexity of ATM layer can be reduced.

There are several ARQ strategies: stop-and-wait [7] , go-back-N [8] and selective-repeat (SR) [9] . In terms of undetected error probability and throughput, the SR strategy is markedly superior to the other two strategies [10] .

In this paper, an ARQ scheme based on the SR concept for ATM network is proposed and analyzed. In practice, transmission distance of most applications is short, like in ATM-LAN. In our scheme, ARQ is only applied in transmitter and receiver, i.e. end-to-end mode. Analysis and simulation results have shown that the proposed scheme has good performance when it is used in an ATM environment.

The remainer of the paper is organized as follows. In the next section, some considerations of ARQ are discussed when it is applied to ATM network. The proposed ARQ scheme is presented in detail in Section 3. In section 4 we gives its performance analysis in detail. Some simulation results are described in Section 5. Finally, Section 6 makes some concluding remarks.

2 Considerations of ARQ in ATM network

ARQ protocol can only be applied to non real time applications in ATM environments, because the retransmission of erroneous packets needs the acknowledgment of packet between the transmitter and the receiver which takes around propagation time. On the other hand, services supported by ATM network can be classified based on three parameter: timing relationship between source and destination, bit rate and connection mode. Each parameter can take two values. Among the service classes defined by the International Telecommunication Union Telecommunication Standardization Sector (ITU-T), Class C and Class D correspond to variable bit rate (VBR) services with no timing relationship between source and destination. Class C is connection-oriented while Class D is connectionless. One of currently defined four ATM adaptation layer (AAL) protocols is AAL3/4 which is used for services of those two classes [11] [12] [13] . Therefore, ARQ protocol can be used for the services provided by AAL3/4.

The ITU-T recommended ATM cell structure is shown in Fig.1. The 4-bit sequence number (SN) field and the 10-bit multiplexing identification (MID) field are specified for assembly and multiplexing use.

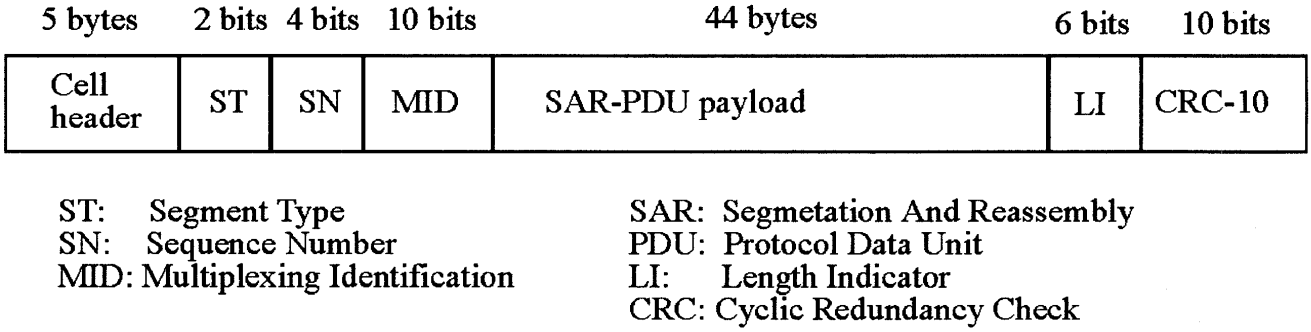


Fig. 1 ITU-T recommended AAL 3/4 cell structure

3 Design of ARQ

In order to realize ARQ technique in ATM network, the information for cell sequence number and error detection must be included in the cell. In our scheme, cell number (CN) field of two bytes is introduced in the cell whose structure is shown in Fig.2. Cyclic redundancy check (CRC) field is 10 bits as the same as the recommended one.

3.1 Cell Number

In a system where applies ARQ protocol, the relationship

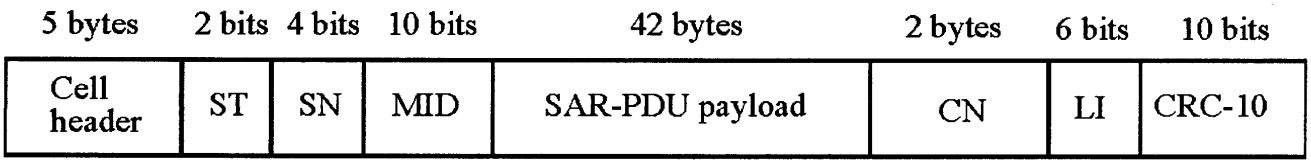
学術情報センター電子図書館サービス　　© 学術情報センター

between the window size (W) and the transmission distance (D) approximately obeys

$$W = \frac{2SD}{180km/ms}, \quad (1)$$

where S is the maximum transmission speed.

The two bytes cell number field is designed for the maximum window size up to 65536 cells. The half of numbers ranging from 0 to 32767 called TCN are used to number the transmitted cells while the rest half of numbers called RCN are used for the retransmitted cells, which will be described in detail in Section 3.4 and 3.5. In this scheme, the effective maximum window size is 32768 that corresponds to about 19,000km transmission distance when the transmission speed is 155*Mb/s*.



ST: Segment Type
SN: Sequence Number
MID: Multiplexing Identification
CN: Cell Number
SAR: Segmetation And Reassembly
PDU: Protocol Data Unit
LI: Length Indicator
CRC: Cyclic Redundancy Check

Fig. 2 Proposed ARQ cell structure

3.2 Cyclic Redundancy Check

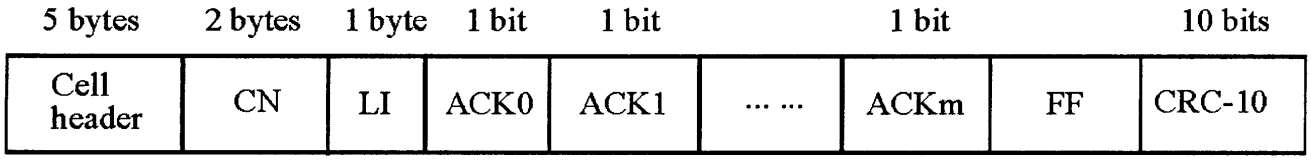
It is recommended by ITU-T to be used in operation, administration and maintenance (OAM) cell and ALL3/4 ATM cell for error detection. Its generator polynomial is defined as

$$G(x) = x^{10} + x^9 + x^5 + x^4 + x + 1$$

When this CRC-10 code is used for error detection, it can detect any odd number of errors, any double-bit errors, any burst errors with the length below 10, and any most burst errors with the length above 10 [6] .

3.3 Acknowledgment

Acknowledgment is one of indispensable factors in ARQ protocol. Since all the transmitted packets are fixed length cells in ATM network, the acknowledgment sent by the receiver to the sender must have the same ATM cell structure. In our scheme, it follows the structure shown in Fig.3.



CN: Cell Number
LI: Length Indicator
ACKi: Acknowledgment for the ith cell counted from the cell identified by CN
FF: Fill Field

Fig. 3 Acknowledgment cell structure

An acknowledgment can acknowledges up to 255 cells, the LI field tells the number of cells acknowledged. The CN field indicates the cell number CN of the first cell. ACKi is 1 bit long and is used to acknowledge the th cell counted from the cell numbered CN.

3.4 Transmitter Operation

There are two states in transmitter operation. The transmit-

学術情報センター電子図書館サービス ©学術情報センター

ter continuously sends and buffers cells numbered in TCNs when it is in the normal state, and releases cells after it receives their acknowledgments. When NAKm (negative acknowledgment) is received or when ACKm is not received after time out, it will enter the retransmission state. In this state, it checks whether the retransmitted cell m is requested for the first time. If yes, it renumbers the requested cell by m + 32768 and resends it, then retuns to the normal state. Note that the resent cell is buffered in new CN which is RCN. If the requested cell m had been already retransmitted, the transmitter simply resends it again and goes back to the normal state.

3.5 Receiver Operation

Receiver operation also has two states. In the normal state, after receiving a cell, the receiver checks whether it is a retransmitted cell. If its CN is TCN which means it is not a retransmitted cell, the receiver checks errors and sends ACK if no error is found, otherwise sends NAK. In the case of retransmitted cell, if there is not any error in it the receiver recovers its CN and inserts it in the proper sequence, then enters the normal state. If there exists any error, it sends NAK again.

4 Performance Analysis

The analysis is made under the following assumptions:

- ARQ window size is large enough to make the transmission continuously;
- The sender always has information to send;
- ATM network can handle the traffic load increment caused by the retransmission.

4.1 Error Detecting Capability

Error detecting capability of a block code can be approximated by using its weight distribution [15] and its syndromes [16] . In this paper, the analysis method based on the syndromes of code is used, since it is difficult to obtain the weight distribution of a code.

For a n-bit codeword, the total number of error patterns E_i can be calculated by

$$E_i = C_i^n = \frac{n!}{i!(n-i)!}, \qquad i = 1, 2, \ldots, n. \tag{2}$$

Then, the undetected probability for i erroneous bits $p_u(i)$ can be obtained by

$$p_u(i) = \frac{S_i}{E_i}, \qquad i = 1, 2, \ldots, n, \tag{3}$$

where S_i is the number of error patterns whose syndromes are zero for i erroneous bits.

The undetected probability for all errors p_u can be computed by

$$p_u = \sum_{i=1}^{n} p_u(i) C_i^n p^i (1-p)^{n-i}, \tag{4}$$

where p is the bit error rate.

Since the CRC-10 code used in our scheme can detect any odd number and any double errors, $E_i = 0$ for i = 1,2,3,5,7,.... What we need to calculate are E_4, E_6, ... To evaluate the performance, only E4 is needed because $p << 1$ and $p_u(i)$s are very small and do not vary much for different i. S4 is 17,989, which is obtained by a program that calculates the syndromes of all the error patterns of 4 erroneous bits by decoding them and to obtain the number of error patterns whose syndromes are not zero. Therefore, eq.4 can be approximated by

$$p_u \approx p_u(4) C_4^n p^4 = S_4 p^4 = 17{,}989 p^4. \tag{5}$$

Table 1 shows the QoS performance improvement with the CRC-10. For example, when the bit error rate of transmission line is 10^{-7}, the bit error rate of received data is about 1.8×10^{-24}. Therefore, when the ARQ is used, the number of bit errors in received data is much less than that without the ARQ.

Table 1 Error performance with ARQ

p	p_u
10^{-4}	1.8×10^{-12}
10^{-5}	1.8×10^{-16}
10^{-6}	1.8×10^{-20}
10^{-7}	1.8×10^{-24}
10^{-8}	1.8×10^{-28}
10^{-9}	1.8×10^{-32}

4.2 Throughput

The probability that a cell has error(s) p_c is

$$p_c = 1 - (1-p)^{384}. \tag{6}$$

学術情報センター電子図書館サービス　　© 学術情報センター

Then, the retransmission probability of a cell p_r can be obtained by

$$p_r = p_c + p_{cl}, \tag{7}$$

where p_{cl} is the cell loss probability.

Define the cell retransmission rate R_c as

$$R_c = \frac{Retransmitted\ cell\ number}{Total\ transmitted\ cell\ number},$$

then the throughput T_p can be obtained by

$$Tp = \frac{1}{1 + R_c}.$$

For ARQ, because it retransmits only the erroneous cells, the throughput becomes

$$Tp = \frac{1}{1 + \sum_{i=1}^{\infty} p_r^i} = 1 - p_r,$$

where p_r is defined by eq.7.

Taking the additional two bytes CN field into account, the overall throughput of ARQ is

$$Tp = \frac{42(1 - p_r)}{44}. \tag{8}$$

Note that the overall traffic load will be slightly increased when ARQ is used. Let R_s be the original source bit rate and assume that R_s is constant, the traffic increment T_{inc} caused by the retransmission can be obtained by

$$T_{inc} = R_s R_c = R_s[\frac{44}{42(1 - p_r)} - 1] = R_s \frac{1 + 21p_r}{21(1 - p_r)}. \tag{9}$$

For example, when the bit error rate of transmission is 10^{-7} the traffic will be increased by 4.77%.

4.3 Buffer Size

The sender must buffer the transmitted cells till they are positively acknowledged by the receiver. Since buffering time of cells t_c is equal to the sum of the emitting, queuing, and processing times (t_e, t_q, t_p), and the round propagation time t_r, and plus some safety margin t_m, i.e.

$$t_c = t_r + 2t_e + 2t_q + 2t_p + t_m,$$

where t_q and t_p are assumed to be the same in both ways. Assume that t_q and t_p are equal to the transmission time of two cells, t_m is $0.2t_r$, then,

$$t_c = 1.2t_r + 10t_e.$$

The required minimum buffer size BS (number of cells) for our ARQ scheme is

$$BS = t_c S_t \ / \ 424bits. \tag{10}$$

Fig.4 shows the required buffer size of ARQ for different distances and service rates. Pick up a sample, the required buffer size is about 4884 cells or about 260 Kbytes when the source bit rate is *155Mb/s* and the distance is *1000 km*, which is acceptable.

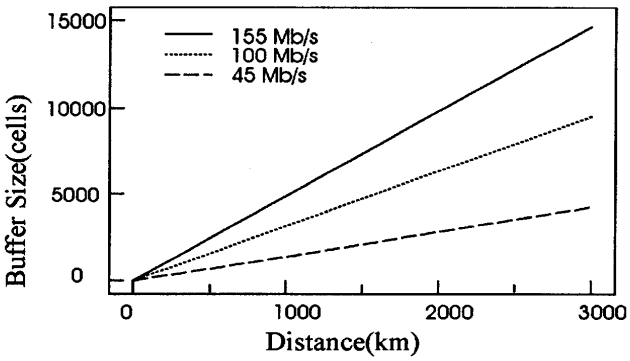


Fig. 4 Required buffer size for ARQ

5 Simulation Results

The performance of the proposed ARQ scheme is also evaluated by the simulation. The bit error rate varies from 10^{-4} to 10^{-9}.

Table 2 shows the simulation results for throughput. From the results it can be found that the throughput is close to the upper limit 42 / 44 = 0.9545 which is caused by the additional tailor of two bytes even when the bit error rate is as high as 10^{-4}. Note that there is no relationship between the throughput and the transmission distance.

The simulation results of traffic increment is shown in Table 3. It has the same conclusion as the throughput since the traffic increment is related to the throughput. The results show that the traffic increment of the proposed ARQ is very small, which implies it is an efficient protocol when it is used in ATM networks.

6 Conclusions

We have presented a selective repeat ARQ scheme for ATM networks. The scheme retransmits cells when they are lost during transmission or checked to be in error by the receiver. The scheme has been analyzed and simulated. Its per-

学術情報センター電子図書館サービス ◎学術情報センター

Table 2 Simulation results of throughput

p	T_p(Calcul)	T_p(Simul)
10^{-4}	0.9186	0.8203
10^{-5}	0.9509	0.9437
10^{-6}	0.9542	0.9531
10^{-7}	0.9545	0.9542
10^{-8}	0.9545	0.9545
10^{-9}	0.9545	0.9545

Table 3 Simulation results of traffic increment

p	T_{inc}(Calcul)	T_{inc}(Simul)
10^{-4}	8.86%	21.19%
10^{-5}	5.16%	5.97%
10^{-6}	4.80%	4.92%
10^{-7}	4.77%	4.80%
10^{-8}	4.77%	4.77%
10^{-9}	4.77%	4.77%

formance in terms of throughput, traffic increment and window size has been examined. Numerical examples showed that the proposed ARQ scheme has very high throughput and has an excellent error detecting performance. If it is used normal optical fiber networks, whose bit error rate is usually 10^{-7}, the undetecting probability is about 10^{-24} and cell loss can be avoided. The required buffer size of the scheme is proportional to the transmission distance and the source bit rate. When the source bit rate is *155Mb/s* and the transmission distance is *1000 km*, the buffer size is about 260 Kbytes, which can be easily implemented.

As the future research, we will improve the proposed ARQ protocol by introducing error correction code into acknowledgment. We will also explore the possibility to implement it.

References

[1] R, Jain.,"Congestion control and traffic management in ATM networks: Recent advances and a survey", *Computer Networks and ISDN Systems*, Vol.28, No.13, pp.1723-1738, 1996.

[2] Y, J, Kim.; S, C, Chang.; C, K, Un.; B, C, Shin., "UPC/NPC algorithm for guaranteed QoS in ATM network", *Computer Communications*, Vol.19, No.3, pp.216-225, 1996.

[3] I,Hsu.;J,Walrand., "Admission control for multi-class ATM traffic with overflow constraints", *Computer Networks and ISDN Systems*, Vol.28, No.13, pp.1739-1751, 1996.

[4] K,R, Kidambi.; R, Sankar.; A, Ottensmeyer., "QoS mapping scheme and model for MAN-ATM internetworking", *Computer Communications*, Vol.19, No.3, pp.276-286, 1996.

[5] K, Sebayashi. ; H, Uose.,"ATM multicast communications method with receiver-initiated QoS guarantee", Global Information Infrastructure (GII) Evolution, IOS Press, 1996, pp.151-160.

[6] S, William., "Data and computer communications", Macmillan Publishing Company, 1988.

[7] W,R,Cowell.; H,O, Burton.,"Computer simulation of the use of a group code with retransmission on a Gilbert burst channel", *AIEE Trans. Commun. Electron.*, Vol.81, pp.577-580, 1962.

[8] R,J,Benice.; A, H, Frey., "An analysis of retransmission systems", *IEEE Trans. Commun. Syst.*, Vol.CS-12, pp.135-144, 1964.

[9] R,D, Stuart., "An insert system for use with feedback communication links", *IEEE Trans. Commun. Syst.*, Vol.CS-11, pp.142-143, 1963.

[10] E.J. Weldon, "An improved selective repeat ARQ strategy", *IEEE Trans. Commun.*, Vol.30, No.3, pp.480-486, 1982.

[11] ITU-T Recommendation I362, "BISDN ATM adaptation layer (AAL) functional description", March 1995.

[12] ITU-T Recommendation I363, "BISDN ATM adaptation layer (AAL) specification", March 1995.

[13] R, Onvural., "Asynchronous transfer mode networks", Artech House Publishers, 1995.

[14] A,S, Tanenbraum., "Computer networks", Third edition, Prentice Hall, 1996.

[15] L, Shu., "Error control coding, fundamentals and applications", Prentice Hall, 1983.

[16] L, Shaw-min., "The syndrome distribution of m-bit error patterns and the analysis of ATM forward error correction codes", *IEEE Trans. Commun.*, Vol.43, No.2/3/4, pp.1392-1399, 1995.

学術情報センター電子図書館サービス　　◎学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

Analysis of Cell Loss Performance in ATM Networks

ATM網におけるセル廃棄特性解析

Suwan RUNGGERATIGUL
National Center for Science Information Systems
学術情報センター　ルンキーラティクン　スワン
Weiping ZHAO
National Center for Science Information Systems
学術情報センター　趙　偉平
Shoichiro ASANO
National Center for Science Information Systems
学術情報センター　淺野　正一郎

ABSTRACT

Asynchronous Transfer Mode (ATM) is known as a target solution for implementing broadband high-speed communication networks. To study the performance of an ATM network, it is necessary to consider an analysis model that can take general traffic model into account since the network can provide service to various types of sources.In this paper, cell loss performance in ATM networks for general traffic model is analyzed by studying the upper bounds of cell loss probability in *GI/D/*1 queue.It can be shown that generating function of traffic arrival probability is an important factor for the bound analysis but derivation of this kind of function is usually complex, especially for bursty traffic models. We then propose an approximation derivation method for the arrival probability generating function of bursty traffic source models. The idea of the proposed method is based on the fact that bursty traffic model often consists of multiple states each of which is a simple stochastic process including Poisson process, deterministic process. By some numerical results, it is shown that the proposed method gives very accurate results comparing to those obtained from exact calculation. Moreover, the computation amount of the method is very small. These features lead to the conclusion that the proposed derivation method is very effective to the control and design problem of ATM networks.

要旨

ATM は超高速通信ネットワークの実現方式として検討されているが、バースト性の通信トラヒックを含めて様々なトラヒックに対応することができるため、ATM 網の性能評価には一般のトラヒックモデルを考慮する解析方法が必要とされている。本論文では、ATM網を 待ち行列にモデル化し、セル廃棄特性の解析について廃棄率の上限に着目する。実際、廃棄率の上限値を求めるとき、トラヒックモデルの到着確率の母関数が重要な要素となるが、一般の到着過程、特にバースト性をもつものについては、このような関数の計算が非常に複雑であることが知られている。そこで、本論文では、バースト的なトラヒックモデルの到着確率の母関数に対する近似計算法を提案する。本計算法はバースト性をもつトラヒックモデルの性質、すなわち、バースト的な確率過程は、複数の基本的な確率過程から構成されていることに基づいている。厳密計算法との結果比較では、提案した近似計算法が高い精度で結果を与えており、かつその計算量が非常に小さいため、ATM 網の性能評価法として有効に適用できることがわかる。

[Keywords]　Communication networks and services, ATM, cell loss performance, queueing analysis, arrival probability, generating function

[キーワード]　通信網、ATM、セル廃棄特性、待ち行列解析、到着確率、母関数

学術情報センター電子図書館サービス ©学術情報センター

1 Introduction

Asynchronous Transfer Mode (ATM) is an advanced multiplexing and switching technique which was proposed and studied widely since the last decade. It is now agreed to be considered as the target solution for implementing Broadband Integrated Services Digital Network (B-ISDN). Since ATM networks provide service to various types of communication sources including non-bursty and bursty sources, it is necessary to consider network model that can take general traffic source into account when studying network performance. Traffic source model in ATM networks is studied widely in the literature. However, several of the studies deal with specific model for specific type of traffic source, or unified model for general traffic source.However, studies with specific traffic model lack the generality of analysis model, and there is no guarantee that any unified model will be suitable for general sources. Consequently, efficient analysis model and result cannot be obtained in these studies.

In this paper, we study cell loss performance in ATM networks for general traffic model, especially for bursty source. Besides the cell loss performance, cell delay and jitter performances can also be considered as quality of service in ATM networks. However, we study only cell loss performance in this paper since the above delay-related performances can be managed and controlled by assigning switching buffer capacity in the network [14]. As a consequence, we need an analysis method for cell loss probability in *GI/D/*1 queueing model. To obtain such kind of analysis method, we focus on the upper bound analysis of the cell loss probability (e.g., [10], [4], etc.). It is found that the generating function of cell arrival probability is an important factor in deriving the bound of cell loss probability. Unfortunately, this kind of function especially in the case of bursty model is difficult to determine and numerical computation amount for exact function is quite large and complex.

In this paper, to solve the above complexity problem, we propose an approximation method for the derivation of cell arrival probability. The method also gives a way to the computation of the generating function. The idea of the proposed method is based on the fact that bursty traffic source model often consists of multiple states each of which is a simple stochastic process including Poisson process, deterministic process. This makes the proposed method be applicable to many bursty traffic models studied in the literature, such as Interrupted Poisson Process (IPP) [8] , Markov Modulated Poisson Process (MMPP) [5], [2], etc.

The rest of this paper is organized as follows. First, analysis of the upper bound of cell loss probability in *GI/D/*1 queue is described in Sect.2 and bursty source models studied in the existing works are reviewed in Sect.3. Then, an approximation method for computing the generating function of cell loss probability for bursty stochastic models is proposed in Sect.4. Some numerical results are given in Sect.5 and the paper is concluded in Sect.6.

2 Upper Bound Analysis of Cell Loss Probability

In this section, the analysis of *GI/D/*1 queueing model is discussed with a purpose to obtain the upper bound of cell loss probability.It is also shown below that the generating function of arrival probability is an important factor in the analysis.

Since we need an analysis method for an ATM network that is capable with general traffic model, we consider the network (e.g., ATM switch, transmission link, virtual path) as a discrete time *GI/D/*1 queueing system with infinite buffer capacity. Arriving cell at the system is assumed to be served upon the first-come-first-served basis. The time-scale is divided into slots where the interval of one time slot is equal to the service time of one cell (we call this time interval as one unit time t).

Next, let's introduce the following notations.

- a_k — the number of cells arrives at the system at the kth time slot,
- Q_k — the queue length at the end of the kth time slot,
- Q — the stationary queue length, i.e., Q_∞ ,
- q_i — probability that $Q=i$,
- p_j — probability that j cells arrive at the system during one time slot.

The queue length Q_k can be computed by the following well-known recursion formula.

$$Q_k = \max(0, Q_{k-1} - 1) + a_k \qquad (1)$$

In the same manner, by considering state of the queue between two adjacent time slots at stationary state, we have the following equations.

学術情報センター電子図書館サービス　　© 学術情報センター

$$q_0=p_0q_1+p_0q_0,$$
$$q_1=p_0q_2+p_1q_1+p_1q_0,$$
$$q_2=p_0q_3+p_1q_2+p_2q_1+p_2q_0,$$
$$\vdots$$

The above equations can be summarized as

$$q_i=\sum_{j=0}^{i}p_jq_{i-j+1}+p_iq_0,\quad i\geq 0. \tag{2}$$

After the z-transform and some mathematical manipulations of Eq.(2), we obtain

$$\Theta(z)=\frac{(1-\rho)(1-z)\Psi(z)}{\Psi(z)-z}, \tag{3}$$

where $\Psi(z)=\sum_{j=0}^{\infty}p_j$, $\Theta(z)=\sum_{i=0}^{\infty}q_i$, and ρ is the service utilization.

It is easy to show that there is only one non-trivial pole z_0 of Eq.(3). Consequently, we can approximate q_i as

$$q_i\approx-\frac{c}{z_0^{i+1}}, \tag{4}$$

where $c=\lim_{z\to z_0}(z-z_0)\Theta(z)$.

It is well-known that cell loss probability L of the actual system with finite buffer capacity has the following upper bound.

$$L\leq P[Q\geq K]=\sum_{i\geq K}^{\infty}q_i, \tag{5}$$

where K is the buffer capacity (cells).

As a consequence, we obtain the bound of L as L_1.

$$L_1=\frac{c}{z_0^{K}(1-z_0)}. \tag{6}$$

There are also many studies regarding bound analysis in the literature. For example, the method that applies Chernoff bound technique to the tail probability obtained from the recursion in Eq.(1) [10], [12], the method that uses some mathematical properties (e.g., Kingman-Ross inequality, Wald's fundamental identity) to the recursion equation [4].

Result proposed in [10] can be briefly shown as

$$P[Q>q]\leq\sum_{i\geq 1}z^{*-(i+q)}_i\Psi_i(z^*_i), \tag{7}$$

where $\Psi_i(z)$ is the probability generating function of the random variable N_i, $i\geq 1$, and z^*_i is the minimum of z in $z^{-r}\Psi_i(z)$, where r is an integer and $z^*_i\geq 1$. Here N_i denotes the number of arriving cells during i time slots.

Also result given in [4] is shown as follows.

$$P[Q>q]\leq\hat{z}^{-q}, \tag{8}$$

where $\hat{z}$ is the non-trivial characteristic root of $z^{-i}\Psi(z)=1$, i.e., $\hat{z}\neq 1$. The value of integer i is assumed to be chosen appropriately when determining $\hat{z}$.

To summarize this section, we point out that cell arrival probability p_j as well as its generating function $\Psi(z)$ are necessary in determining the bound of cell loss probability in *GI/D/1* system.

3 Bursty Traffic Model

We discuss in this section about bursty traffic source models for ATM networks. It is realized that finding a model to fit a bursty source is a very difficult task since a same set of traffic parameters describing burstiness of cell arrival can cause several completely different traffic models. There are many proposed sets of traffic descriptor and some of them are a combination of the following traffic parameters: average bit rate, average cell arrival rate, variance of cell arrival rate, peak bit rate, peak to average bit rate ratio, burst length (average duration time of peak bit rate) [6].

Thus, to find a suitable common model for general bursty traffic is nearly impossible. The following are examples of bursty source models studied in conventional research.

- **IPP model:** Interrupted Poisson Process was proposed to approximate the overflow traffic to reduce the cost of computer simulation by analytic method [8]. In IPP, there are two states called ON and OFF state. The time duration of the both states are both exponentially distributed. Cell arrival follows Poisson process during the ON state and there is no cell arrival in the OFF state. For ATM traffic source model, IPP is adopted to be a model of single video (moving picture) source when peak and average bit rate of a video source are given [1].
- **MMPP model:** Markov Modulated Poisson Process is a more general form of IPP. In N-state MMPP, there are N independent states and duration time (sojourn time) of each state in the process is exponentially distributed, and arrival process in each state follows Poisson process.

学術情報センター電子図書館サービス ◎学術情報センター

The applications of MMPP are: traffic model for video source [7], [13], approximated model for superposed stream (e.g., multiplexed stream from homogeneous ON-OFF voice sources [5]).

- ON-OFF model: In this model, there are two states called ON state and OFF state. In the ON state, cell arrival is at a fixed rate while there is no arrival in the OFF state. The time duration of the two states are both exponentially distributed.

  ON-OFF model is studied widely for modeling bursty traffic in various manners. In [5], [7], [9], [11], etc., it is used to model a single voice source with speech activity detector (SAD) which generates cells only in a talkspurt period and no cell is generated in a silence period. Consequently, talkspurt and silence period are the ON state and OFF state of the model, respectively. ON-OFF model is also used to fit general bursty sources. For example, general source is modeled by using ON-OFF model when a set of traffic descriptor is given [3], [16].
- MMDP model: Markov Modulated Deterministic Process is a more general form of the above ON-OFF model. In *N*-state MMDP, there are N independent states and duration time (sojourn time) of each state in the process is exponentially distributed, and arrival process in each state follows deterministic process (fixed cell arrival rate). MMDP is used as an approximated model for multiplexed stream of homogeneous ON-OFF source model [15].
- D+P model: There are two states in this model where cell arrival in one state follows deterministic process (D state), and arrival in the other state (P state) is Poissonian. It should be noted that this model is named by the authors of this paper because this kind of source model, although has been studied in the past, has not been given any name. Clearly, the name given above is very straightforward that D stands for deterministic process and P is for Poisson process.

  There are many manners to set the distributions of state duration in the process, e.g., both are exponential, one is exponential and one is fixed, one is exponential and the sum of the two times is fixed, etc. Studies on D+P model are: application to video source modeling with simulation [11], usage on general bursty source modeling with traffic descriptor [17].

By the above discussion, we can see that the bursty traffic models have the following characteristic: the process is a multi-state stochastic process where each state i is an individual process B_i with cell arrival rate λ_i, and state duration r_i. This characteristic can be depicted as in Fig.1.

Furthermore, the individual process is one of the following fundamental stochastic processes.

P	This is the well-known Poisson process where cell arrival in this process is random and cell interarrival time is exponentially distributed.
D	This is a deterministic process where cell arrival is at a fixed rate.
Z	This is a process that none of cell arrival occurs, i.e., cell arrival rate is equal to zero.

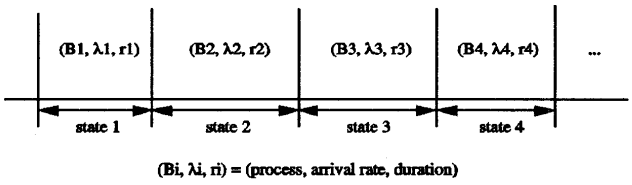


Fig.1 Bursty traffic model

Thus, the bursty traffic models discussed above can be decomposed into a combination of several fundamental processes as

IPP	P and Z,
MMPP	combination of P's,
ON-OFF	D and Z,
MMDP	combination of D's,
D+P	D and P.

It will be seen in the next section that the above fact is very important and useful in analyzing bursty traffic.

4 Approximation Method for Bursty Models

We apply the characteristic of bursty traffic models obtained from Sect.3 to approximate cell arrival probability and its generating function in this section.

First, cell arrival probability $p_k(t)$ and generating function $\Psi(z,t)$ of the three fundamental processes are given below. Here, $p_k(t)$ is the probability that there are k cells arrive at the system during time period t. Now, $\Psi(z,t)$ is considered as the function of both z and t.

- P process: Arrival in Poisson process is random and interarrival time is exponentially distributed. From literature, it is well-known that $p_k(t)$ can be determined

学術情報センター電子図書館サービス　　© 学術情報センター

by

$$p_k(t) = p_k^p(t) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}, \quad k \geq 0, \tag{9}$$

where λ is the mean cell arrival rate of the process. Generating function of P process is then

$$\Psi_P(z,t) = e^{\lambda t(z-1)}. \tag{10}$$

- D process: Arrival in this process is at a fixed rate. The value of $p_k(t)$ of this process can be determined as

$$p_k(t) = p_k^d(t) = \begin{cases} 1 - \frac{t}{T} + n, & k = n, \\ \frac{t}{T} - n, & k = n+1, \\ 0, & \text{otherwise}, \end{cases} \tag{11}$$

where T is the fixed interarrival time, $n = \left\lfloor \frac{t}{T} \right\rfloor$, and $\lfloor x \rfloor$ is the maximum integer that is equal or smaller than x. Thus, the generating function of D process can be obtained as

$$\Psi_D(z,t) = z^n(1 + (\frac{t}{T} - n)(z-1)). \tag{12}$$

- Z process: Since there is no cell arrival in this process, it is clear that $p_k(t)$ in this process is

$$p_k(t) = p_k^z(t) = \begin{cases} 1, & k = 0, \\ 0, & \text{otherwise}. \end{cases} \tag{13}$$

The generating function is then $\Psi_Z(z,t)=1$.

In the case of bursty traffic models, the determination of $p_k(t)$ as well as $\Psi(z,t)$ is much more complex than the above fundamental processes. We will consider IPP as an example here. Since IPP is a special case of MMPP, we can use the result of MMPP to determine $\Psi(z,t)$ of IPP that can be given as below [5].

$$\Psi_{IPP}(z,t) = \prod \exp((R + (z-1)\Lambda)t)e, \tag{14}$$

where

$$\Pi = \frac{1}{1/\alpha + 1/\beta}[1/\beta \quad 1/\alpha],$$

$$e = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$R = \begin{bmatrix} -1/\alpha & 1/\alpha \\ 1/\beta & -1/\beta \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix},$$

λ is the mean arrival rate during ON state, α, β are the mean duration time of ON and OFF state, respectively.

As can be seen from above, the computation complexity of $\Psi(z,t)$ for complex bursty model is very large so some approximation technique is needed. We propose such kind of method in this section and again take IPP as an example. The computation method for $p_k(t)$ is described first and then by this result the generating function $\Psi(z,t)$ is discussed.

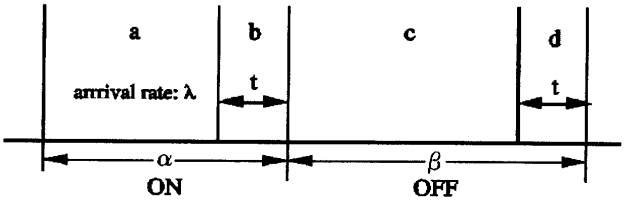


Fig.2 Approximation method for p_k of IPP

Remind that there are two states in IPP called ON state which is a Poisson process (part a and b in Fig.2) and OFF state which has no cell arrival (part c and d in Fig.2). We then define a time of interval t and move it along the ON state and OFF state as in Fig.2. Note that t is the time of which we need to calculate $p_k(t)$. We get the following.

- If the starting point of interval t is in part a, $p_k(t)$ in this part is equal to $p_k^p(t)$ which is the $p_k(t)$ of Poisson process and can be determined by Eq.(9).The length of time of part a is clearly $\alpha - t$.
- If the starting point of interval t is in part b, $p_k(t)$ in this part is the $p_k(t)$ of the IPP itself. The length of part b is t.
- If the starting point of interval t is in part c, $p_k(t)$ in this part is equal to $p_k^z(t)$ which is the $p_k(t)$ of Z process and can be calculated by Eq.(13). The length of part c is $\beta - t$.
- If the starting point of interval t is in part d, $p_k(t)$ in this part is the $p_k(t)$ of the IPP itself as in the case of part b. The length of part d is also t as in part b.

By summarizing all the above four cases and taking the average value of $p_k(t)$, the following holds.

学術情報センター電子図書館サービス ◎学術情報センター

$$p_k(t) = \frac{(\alpha - t).p_k^p(t) + t.p_k(t) + (\beta - t).p_k^z(t) + t.p_k(t)}{\alpha + \beta}. \quad (15)$$

After some mathematical manipulating, the above becomes

$$p_k(t) = \frac{(\alpha - t).p_k^p(t) + (\beta - t).p_k^z(t)}{\alpha + \beta - 2t}. \quad (16)$$

In general cases, t is much smaller than the time duration of both ON and OFF states. Hence, Eq.(16) can be approximated by

$$p_k(t) \approx \frac{\alpha.p_k^p(t) + \beta.p_k^z(t)}{\alpha + \beta}. \quad (17)$$

It is obvious that the above expression can also be obtained directly by ignoring part b and part d in Fig.2 and taking average of probability in part a and c.

From the above result, $\Psi_{IPP}(z,t)$ is approximated by

$$\Psi_{IPP}(z,t) \approx \frac{\alpha.\Psi_p(z,t) + \beta.\Psi_z(z,t)}{\alpha + \beta}. \quad (18)$$

It is clear that $\Psi_{IPP}(z,t)$ from the above approximation is much simple than the exact one given in Eq.(14). To deal with general bursty traffic processes, the same approach can be adopted as follows.

If the bursty process consists of N independent states where each state i is a fundamental process with mean duration time r_i, $p_k(t)$ of this bursty process can be approximated by

$$p_k(t) = \frac{\sum_{i=1}^{N} r_i \cdot p_{(i,k)}(t)}{\sum_{i=1}^{N} r_i}, \quad (19)$$

where $p_{(i,k)}(t)$ is the $p_k(t)$ in state i of the process which can be determined by Eqs.(9), (11), and (13) according to the type of fundamental process that the state is.

The above result leads to the generating function of the bursty process as

$$\Psi(z,t) = \frac{\sum_{i=1}^{N} r_i.\Psi_{(i,k)}(z,t)}{\sum_{i=1}^{N} r_i}, \quad (20)$$

where $\Psi_{(i,k)}(z,t)$ is the generating function of arrival process during state i.

When the number of states in the bursty process gets larger, there is no complexity problem in computing the generating function by the proposed method.Moreover, clearly from Sect.2, the time of computing $\Psi(z,t)$ to determine the upper bound of cell loss probability is quite large, so the proposed approximation method is very useful and applicable to this computation task.

5 Numerical Results

In this section, we apply the approximation method proposed in the last section to IPP and MMPP to confirm the accuracy of the method. IPP and MMPP are described here as follows.

For IPP: IPP(λ, α, β) where λ, α, β are cell arrival rate during ON state, mean duration time of ON state, and mean duration time of OFF state, respectively.

For 2-state MMPP: MMPP (λ_1, λ_2, r_1, r_2) where λ_1, λ_2, r_1, r_2 are cell arrival rate during state 1 and 2,mean duration time of state 1 and 2, respectively.

Without loss of generality, the unit time (time for transmitting one cell) is equal to 1 in all examples. In each example, dash-lines are the result from the proposed method.

Fig.3 and Fig.4 show the results of IPP (0.2,1000,1000) and IPP (0.4,1000,1000) when the number of multiplexing (VC) varies.

Fig.5 and Fig.6 show the results of multiplexed IPP sources for several ranges of parameters when the utilization is 0.9.

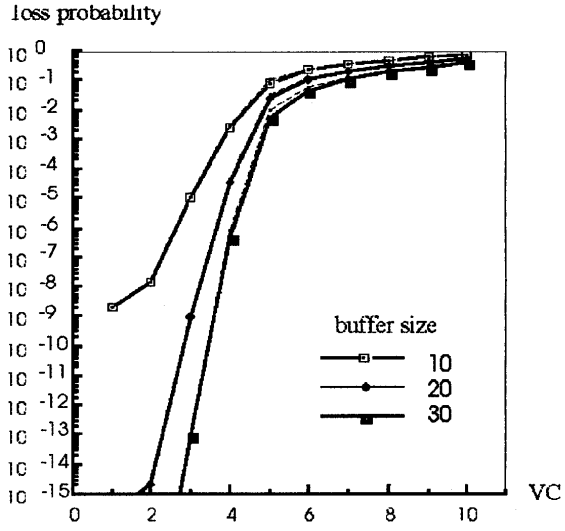


Fig.3 Results comparison: IPP (0.2, 1000, 1000)

学術情報センター電子図書館サービス　　© 学術情報センター

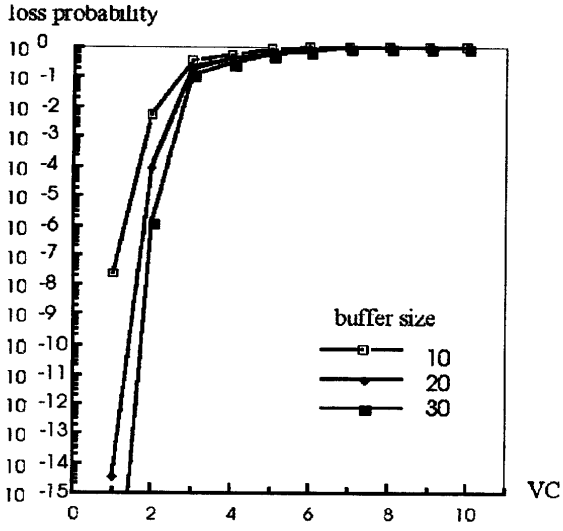


Fig.4 Results comparison: IPP (0.4, 1000, 1000)

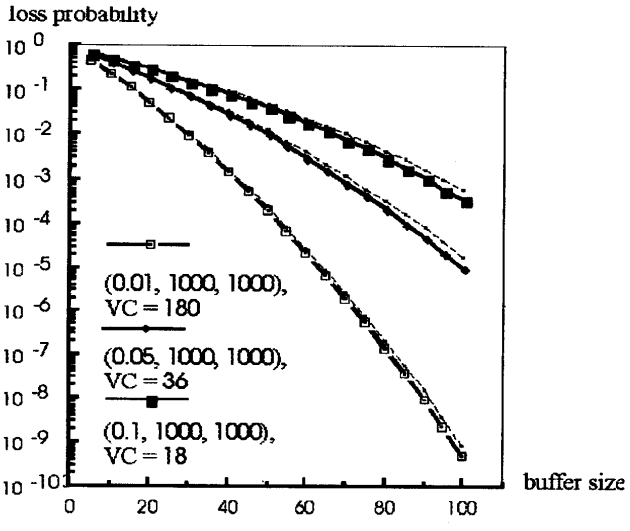


Fig.5 Results comparison: multiplexed IPP #1

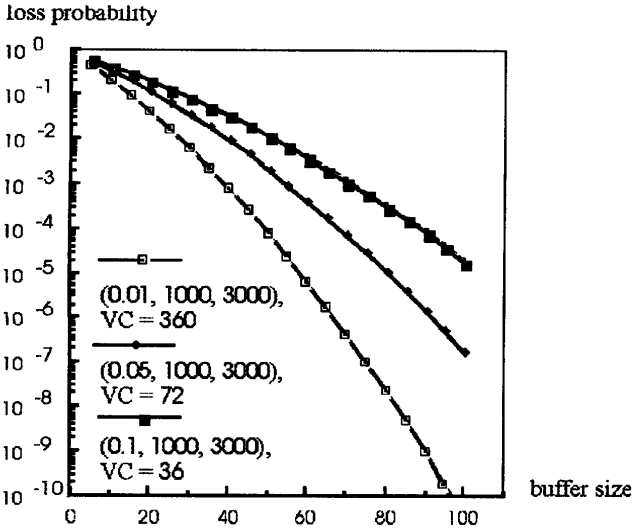


Fig.6 Results comparison: multiplexed IPP #2

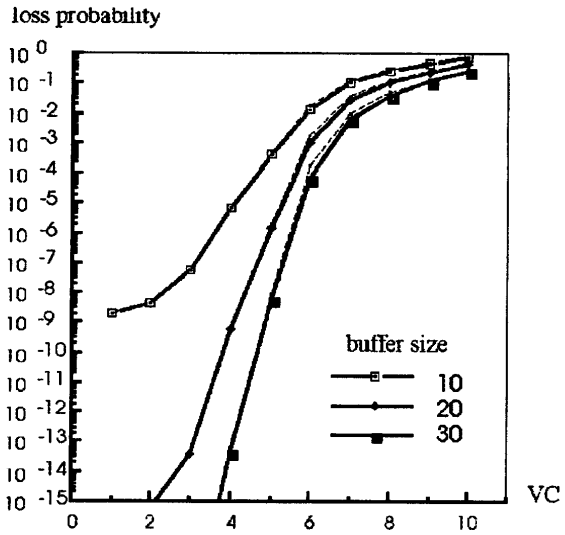


Fig.7 Results comparison: MMPP (0.05, 0.15, 1000, 1000)

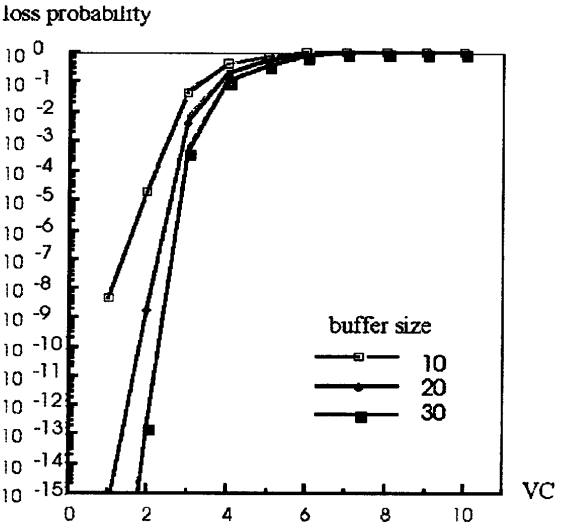


Fig.8 Results comparison: MMPP (0.1, 0.3, 1000, 1000)

Fig.7 and Fig.8 show the results of MMPP (0.05,0.15,1000,1000) and MMPP (0.1,0.3,1000,1000) when the number of multiplexing (VC) varies.

Fig.9 and Fig.10 show the results of multiplexed MMPP sources for several ranges of parameters when the utilization is 0.9.

Fig.11 and Fig.12 show the effect of state duration in IPP and MMPP (the length of duration is equal for each state in the process).

From the above results, it is clear that the proposed method gives very good and accurate result to a wide range of parameters of the IPP and MMPP. Thus, the method is very powerful in using as a tool for ATM network design and control (e.g., call admission control, etc.). Although

学術情報センター電子図書館サービス ◎学術情報センター

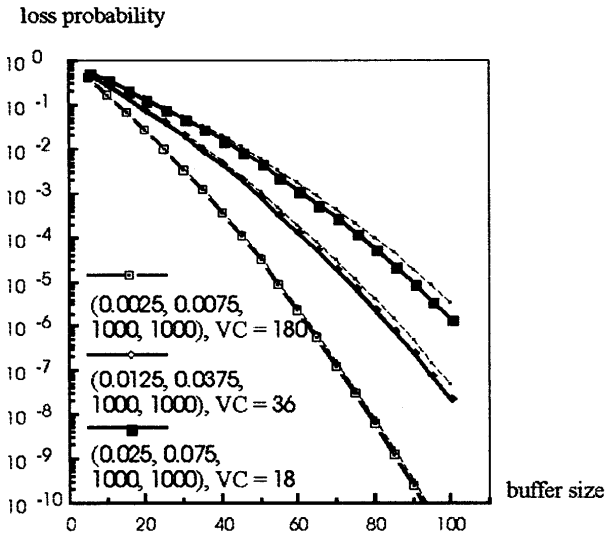


Fig.9 Results comparison: multiplexed MMPP #1

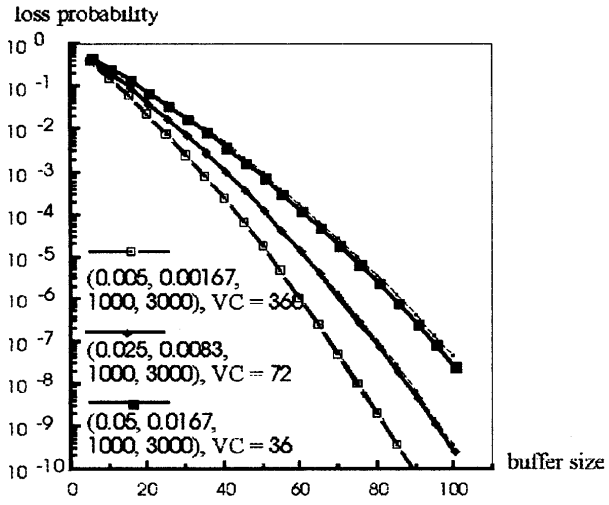


Fig.10 Results comparison: multiplexed MMPP #2

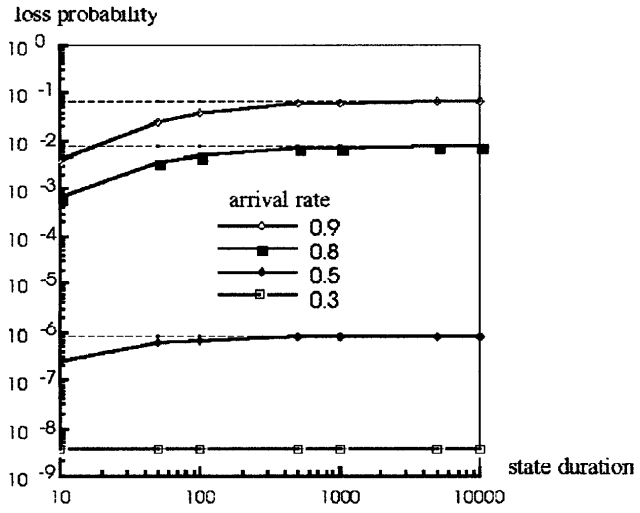


Fig.11 Effects of state duration: IPP

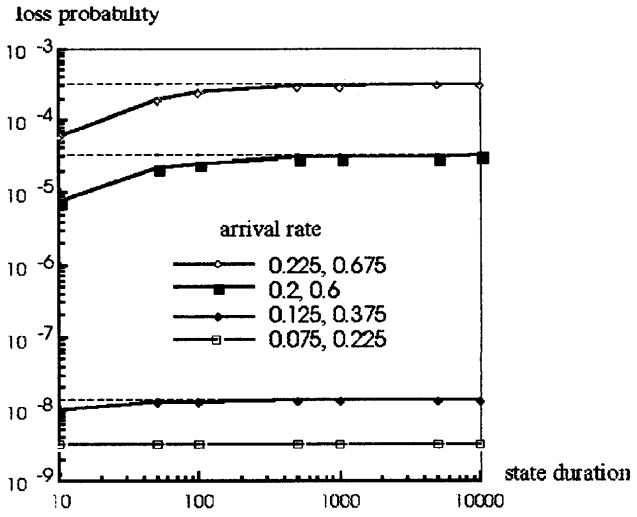


Fig.12 Effects of state duration: MMPP

there is problem when state duration is small, it is shown that the proposed method still gives an upper bound for the result. In fact, for typical bursty source in ATM networks,state duration is quite large comparing to cell transmission time, and it is already shown that the proposed method gives accurate result for this case.

6 Conclusions

In this paper, we study the analysis of cell loss performance in ATM networks by focusing on general traffic source. The method of analysis is conducted for *GI/D/*1 queueing model to get the upper bound of cell loss probability. Clearly, to determine such bound the generating function of cell arrival probability is very important, and this makes it necessary to find an approximation method for the function due to the complexity of exact method.

We then propose an approximation method for computing arriving generating function by using an important characteristic of the bursty process. It is shown by several numerical examples with wide range of parameters that the proposed method gives very high accuracy of results comparing to those obtained from exact computation. Thus, it is concluded that the proposed method is very effective for applying to ATM network control and design problem.

Acknowledgements

The work in this paper is supported by the Japan Society for the Promotion of Science (JSPS) under Postdoctoral Fellowship of the first author (S. RUNGGERATIGUL).

学術情報センター電子図書館サービス ©学術情報センター

References

[1] Arai, H.; Kawarasaki, M.; Nogami, S., "An Analysis of Call Admission Control in the ATM Network", *IECEJ*, Vol.J72-B-I, No.11, pp.1000-1007, Sept. 1989, (in Japanese).

[2] Baiocchi, A.; Melazzi, N.B., "Steady-State Analysis of the *MMPP/G/1/K* Queue", *IEEE Trans. Caommun.*, Vol.41, No.4, pp.531-534, Apr. 1993.

[3] Gallassi, G.; Rigolio, G.; Fratta, L., "ATM: Bandwidth Assignment and Bandwidth Enforcement Policies", *Proc. GLOBECOM'89*, pp.49.6.1-6, 1989.

[4] Geihs, K.; Kobayashi, H., "Bounds on Buffer Overflow Probabilities in Communication Systems", *Perf. Eval.*, Vol.2, pp.149-160, 1982.

[5] Heffes, H.; Lucantoni, D.M., "A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance", *IEEE J. Select. Areas Commun.*, Vol.SAC-4, No.6, pp.856-868, Sept. 1986.

[6] Kawashima, K.; Saito, H.,"Teletraffic Issues in ATM Networks", *Computer Networks and ISDN Systems*, 20, pp.369-375, 1990.

[7] Komine, T.; Yamada, H., "An Analysis on Statistical Multiplexing Characteristics of ATM Cells with Voice and Video Inputs", *IECEJ*, Vol.J75-B-I, No.8, pp.509-516, Aug. 1992, (in Japanese).

[8] Kuczura, A., "The Interrupted Poisson Process as an Overflow Process", *Bell Syst. Tech. J.*, Vol.52, No.3, pp.437-448, Mar. 1973.

[9] Nagarajan, R.; Kurose, J.F.; Towsley, D., "Approximation Techniques for Computing Packet Loss in Finite-Buffered Voice Multiplexers", *IEEE J. Select. Areas Commun.*, Vol.9, No.3, pp.368-377, Apr. 1991.

[10] Nakagawa, K., "Loss and Waiting Time Probability Approximation for General Queueing", *IEICE Trans. Commun.*, Vol.E76-B, No.11, pp.1381-1388, Nov. 1993.

[11] Nogami, S., "Quality Control at the Cell Level and Its Characteristics", *IECEJ*, Vol.J73-B-I, No.9, pp.671-680, Sept. 1990, (in Japanese).

[12] Park, H.S.; Kwak, D.Y.; Rhee, W.S.; Jeon, M.Y.; Kim, J.K., "Global Traffic Control in ATM Networks", *IEICE Trans. Commun.*, Vol.E78-B, No.4, pp.476-484, Apr. 1995.

[13] Saito, H.; Kawarasaki, M.; Yamada, H., "An Analysis of Statistical Multiplexing in an ATM Transport Network", *IEEE J. Select. Areas Commun.*, Vol.9, No.3, pp.359-367, Apr. 1991.

[14] Saito, H.; Shiomoto, K., "Dynamic Call Admission Control in ATM Network", *IEEE J. Select. Areas Commun.*, Vol.9, No.7, pp.982-989, Sept. 1991.

[15] Yang, T.; Tsang, D.H.K., "A Novel Approach to Estimate the Cell Loss Probability in an ATM Multiplexer Loaded with Homogeneous On-Off Sources", *IEEE Trans. Commun.*, Vol.43, No.1, pp.117-126, Jan. 1995.

[16] Yoshida, M.; Okazaki, H., "A Study on ATM Network Planning Based on Evaluation of Design Items", *IEICE Trans. Commun.*, Vol.E76-B, No.11, Nov. 1993.

[17] Zhao, W.P., "Researches on Control of Qualities of Service in ATM Networks", Doctor Dissertation, The University of Tokyo, Dec. 1992.

学術情報センター電子図書館サービス　　©学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

シェーピングによるトラヒック特性への影響

The Characteristics of ATM Traffic Constrained by Different Shaping Mechanisms

学術情報センター　計　宇生
Yusheng JI
National Center for Science Information Systems

要旨

ATM 網におけるユーセージパラメータ制御(UPC)のための機構として GCRA が標準化されている。それは Leaky Bucket として知られている方式と同等である。その他に、Jumping Window や Moving Window などの方式もそのために提案されている。一方、トラヒックシェーピングでは UPC と同様のアルゴリズムが使われることが多いが、異なるシェーピング機構によるトラヒックの特性への影響についてはまだ明らかでないことが多い。本稿では、シェーパー及び交換ノードの両方における遅延及び損失に着目し、Leaky Bucket と Moving Window 方式をシェーピングに用いた場合の比較を行う。それによって、Leaky Bucket が他の方式よりもすぐれた特性を有していることを示す。

ABSTRACT

GCRA has been made as the standard Usage Parameter Control Algorithm for ATM networks. It is equivalent with the so called Leaky Bucket algorithm. Besides, other algorithms, such as Jumping Window or Moving Window, have also been proposed. On the other hand, algorithms used in traffic shaping are usually as same as those used in UPC. Nevertheless, performance of these algorithms used in traffic shaping is not quite clear as them used in UPC. In this paper, we focus on cell delay and loss both at the shaper and the network switching nodes, and give a comparison of the characteristics of traffic shaped by Leaky Bucket and Moving Window. It is shown that the Leaky Bucket algorithm has a more disirable feature than the other one.

[キーワード]　トラヒックシェーピング、リーキー・バケツ、ウィンドウ制御、移動ウィンドウ
[Keywords]　traffic shaping, Leaky Bucket, Jumping Window, Moving Window

1　はじめに

GCRA(Generic Cell Rate Algorithm)は ATM 網における UPC(Usage Parameter Control)のための機構として、ITU-T などで標準化されている [1] 。その基本は Leaky Bucket として知られている方式と同等である。その他に、Jumping Window や Moving Window およびそれらの変形 [2] もそのために提案されている。これらをポリシング機構として用いた場合のトラヒックの特性については文献 [3] で検討されている。一方、トラヒックシェーピングでは、UPC と同様のアルゴリズムが使われることが多いが、シェーピング機構として使う場合のトラヒック特性への影響については、まだあまり知られていない。

文献[4]では、Shift Register Traffic Shaper (SRTS) と呼ばれるシェーピング機構が提案されている。SRTS は Moving Window を拡張したものであると言えよう。文献 [4] によれば、SRTS を使えば、期間の短いバーストに対して、従来の Leaky Bucket より柔軟に対応することができ、パケットの損失を減らすことに有効であるという。しかし、同文献で行なわれている比較は、ネットワークの入口におけるシェーピングのための遅延および損失であり、バーストを制限することによって、ネットワークの交換ノードにおける輻輳回避および遅延削減の効果についてはふれていない。一般的には、ネットワークの入口でのシェーピングによる制限が緩いほど、ネットワーク内における遅延が大きくなる可能性が高い。このような観点にたって、本論文では、ネットワークの入口及び交換ノードにおける遅延の両方に

学術情報センター電子図書館サービス　　© 学術情報センター

着目し、Leaky Bucket および Moving Window 両方式の有効性の比較を行う[5]。

2 レート制御アルゴリズム

ATM トラヒックの特性は、ピークセルレート、平均セルレート、および最大バーストサイズによって表すことができる。トラヒックのピークレート、平均レートおよびバースト性に対して制限を加えることにより、ネットワークの負担を軽減し、ネットワーク内に入ったトラヒックの品質を保証すると同時に、ネットワークの帯域をより有効に使うことがシェーピングの目的である。ここでは、平均セルレートおよび最大バーストサイズを制限するために使われるシェーピング機構(GCRA の第 2 段目に相当する)に注目し、ピークセルレートについてはすでに整形ずみであると仮定する。このようなシェーピング機構は一般に、ポリシングで使われる機構と同様に、レート制御機構を使うことができ、以下で述べるような各種のアルゴリズムが提案されている。また、シェーパーを直ちに通過できないセル(ポリシングにおける違反セルに相当するもの)については通常、マークしたり、直ちに廃棄したりするようなポリシングの場合と違って、バッファリングして、シェーパーを通過できる時点まで待ってから送出する。この場合、バッファの長さを有限にすることによって、シェーパー内での最大遅延を制限することができる。

2.1 Jumping Window および Triggered Jumping Window

文献[6]で提案されている方式によって代表される Jumping Window と呼ばれる機構では、ある決まった時間間隔 I (時間ウィンドウという)内に通過するセル数をカウントし、それを $\bar{\lambda}I$ の値以下に制限するものである。前のウィンドウの直後に始まる新しいウィンドウ区間では、セル数のカウンタはリセットされ、再び 0 からカウントされ始める。

明らかに、Jumping Window を通過できるトラヒックの平均セルレートの最大値は $\bar{\lambda}$ である。また、二つの隣り合う時間ウィンドウの境界点の前後を合わせると、ピークセルレート $\bar{\lambda}$ で連続的に通過できるトラヒックパターンは図 1 のようになる。このときの最大バーストサイズは $2\bar{\lambda}I$ である。

このような Jumping Window の実装は簡単である反面、各時間ウィンドウの開始および終了時刻はソーストラヒックの挙動と同期を取っていないことが欠点である。この欠点を補うために提案されたのが Triggered Jumping Window である。この時、時間ウィンドウは一連に続くのではなく、最初のセルの到着時刻で開始される。これによって図 1 のようなパターンが避けられる。

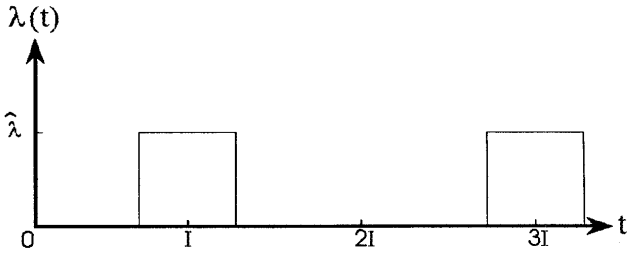


図 1　最大バーストを持つトラヒックパターン：Jumping Window の場合

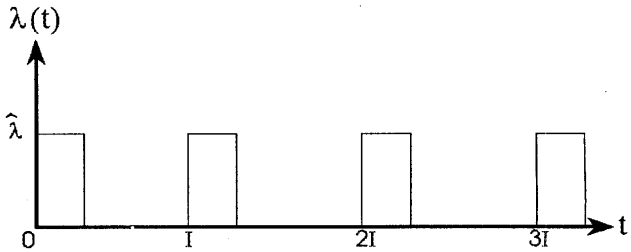


図 2　最大バーストを持つトラヒックパターン：Moving Window の場合

2.2 Moving Window および SRTS

一方、時間ウィンドウの間の固定な境界点をなくし、時間軸上で自由に移動するようなウィンドウを Moving Window と呼ぶ。これはすなわち、フロー制御のための Sliding Window のアルゴリズムにおけるパケットウィンドウの概念を時間に適用したものである。文献[7]では、$\lambda(t)$ を時刻 t に通過するトラヒックの瞬時セルレートとして、Moving Window による制限を次の式で表している。

$$\forall t,\ \int_t^{t+I} \lambda(\tau)d\tau \le \bar{\lambda}I \qquad (1)$$

Jumping Window と同様に、Moving Window を通過するトラヒックの平均レートは $\bar{\lambda}$ 以下に制限される。また、平均レートが同じでも、その平均化区間である I の値の大きさによって通過するトラヒックのバースト性が左右される。すなわち、I の値が大きいほど、バースト性が大きく、平均レートについて求めるのに必要な時間間隔が長い。

学術情報センター電子図書館サービス　　© 学術情報センター

ここで、ピークセルレートは $\hat{\lambda}$ に制限されていると仮定し、また、一般性を失わないように、時刻 t(0 におけるシェーパー内の待ち行列の長さが 0 であるとする。もし時刻t=0 から、ソースが可能な限りセルをネットワークに送り込もうとすると、Moving Window をピークセルレートで連続的に通過できるトラヒックパターンは図 2 のようになり、このときの最大バーストサイズ B_{mw} は

$$B_{mw} = \bar{\lambda} I \tag{2}$$

である。

また、式(1)のような制限を複数個同時に設けられるように Moving Window を拡張したのが Shift Register Traffic Shaper である。これらの機構の実装は、Jumping Window に代表されるものよりむずかしいが、より厳密にトラヒックの制限ができる点で注目されている。

2.3 Leaky Bucket

Leaky Bucket [8] - [10] として知られている方式では、それを通過するセル流は次の式を満足する。

$$\forall t, x \geq 0, \quad \int_t^{t+x} \lambda(\tau) d\tau \leq \bar{\lambda} x + \sigma \tag{3}$$

同様に、$\lambda(t)$ は時刻 t に Leaky Bucket を通過する瞬時セルレートであり、σ はセル数を単位とした最大トークン数(バケツの大きさ)であり、また、$\bar{\lambda}$ はトークンの生成率である。

ここでは議論をしやすくするために、時刻t=0 におけるトークン数が最大値の σ であると仮定する。このような仮定のもとでは、一般性が失われることがない。なぜなら、入力セルレートの長時間平均がトークンの生成率 $\bar{\lambda}$ より小さいことがシステムが安定するための条件であるが、それはトークン数が最大値になる時が必ずあることを意味し、このような時点を t=0 とすることができるからである。時刻 $t = 0$ において、Leaky Bucket をピークセルレートで連続的に通過できるセルの最大数(最大バーストサイズ)B_{lb} は、最大トークン数 σ と B_{lb} をピークセルレートで転送している間に再生されるトークン数の和、すなわち図 3 における X 点の縦座標である。したがって、

$$B_{lb} = \frac{\hat{\lambda}\sigma}{\hat{\lambda} - \bar{\lambda}} \tag{4}$$

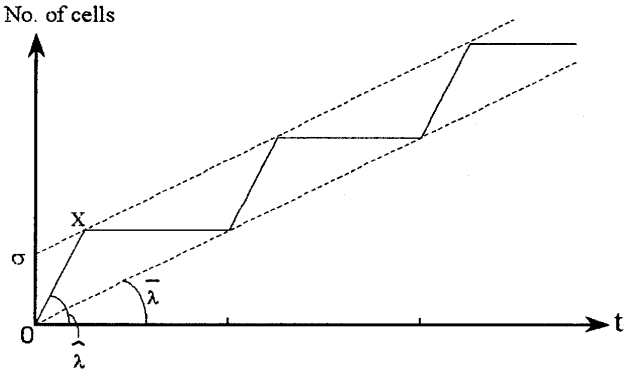


図3　最大バーストを持つトラヒックの累計セル数：Leaky Bucket の場合

もしソースが時刻 $t = 0$ において最大バーストサイズを転送し始め、トークンが0になった時点で転送を中止し、トークンが最大になるまでためてから、再び転送を開始するといったようなことを繰り返せば、図2と同じようなトラヒックパターンが得られる。

Moving Window および Leaky Bucket のそれぞれを通過できる最大バーストサイズを同じにすることで、式(4)と(2) より最大トークン数 σ と時間ウィンドウ I の関係を求めると、次式のようになる。

$$\sigma = \bar{\lambda} I \left(1 - \frac{\bar{\lambda}}{\hat{\lambda}}\right) \tag{5}$$

3 シミュレーションによる評価

3.1 シミュレーションモデル

シェーピングによるトラヒックの特性への影響を評価するため、図 4 に示すネットワークモデルを用いてシミュレーションを行なった。各々のコネクションは複数個のソースホストから一本の中継リンクの両端につながる交換ノードを通って、目的ホストまで張られている。トラヒックによるリンク帯域の競合が中継リンクにおいてのみ起こり得るものとする。中継リンクの帯域が 1 cell/tick となるよう、時間の最小単位 tick を決める。交換ノードにおけるサービス規律 [11]～[12] はもっとも単純な先着順サービスとする。各交換ノード自身は十分な処理能力をもっており、入線側から入るセル流に対して、停滞なく交換処理を行えることとする。

シェーピングはソース側のホスト上で行なわれ、シェーピングパラメータとして、式(1)および(3)における平均レート(トークンの生成率) $\bar{\lambda}$ を 0.1 cell/tick とする。ホストと交換ノードをつなぐリンクは、トラ

学術情報センター電子図書館サービス　　© 学術情報センター

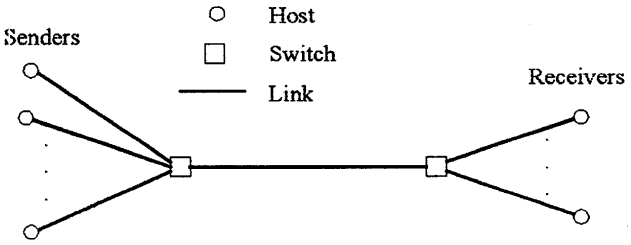


図4　シミュレーションに用いたネットワークモデル

ヒックがピークセルレートで通過するのに十分な帯域を有する。セル遅延はセルがシェーパーまたは交換ノードの待ち行列における待ち時間としてカウントされ、待ち行列の長さに制限がある場合、長さの制限に達した待ち行列に到着したセルは直ちに廃棄される。

それぞれのセル流は ON-OFF モデルにしたがって生成され、ONおよびOFF期間の長さはそれぞれ平均が 20 ticks と 130 ticks である幾何分布に従うものとする。ON 区間におけるセルレート(ピークセルレート)は中継リンク容量の半分の 0.5 cell/tick とする。したがって、それぞれのソースの平均レートは約 0.067 cell/tick である。シェーパーを通過したトラヒックにも一定のバースト性を許容するため、シェーピングパラメータの平均レート $\bar{\lambda}$ とソーストラヒックの平均セルレートとの比(オーバーディメンジョニング・ファクター)を 1.5 とした。

3.2　シェーパー内の遅延と損失

図5では、 Leaky Bucket および Moving Window のアルゴリズムを採用したそれぞれのシェーパーにおけるセルの遅延の平均および標準偏差と、バースト性を制限するシェーピングパラメータのウィンドウサイズ(式(1)における $\bar{\lambda}$ I)との関係を示す。Leaky Bucket の場合では、式(5)にしたがってウィンドウサイズを最大トークン数 σ に換算している。

同図では、ウィンドウサイズが大きく、すなわち、バーストに対する制限が緩やかになるにしたがって、遅延およびその分散は小さくなることを示しているが、その変化の度合は Leaky Bucket のほうがより大きい。このことは、文献 [3] でこれらのアルゴリズムをポリシングに用いたとき、Leaky Bucket の場合の違反確率(violation probability)のほうがウィンドウサイズが大きくなるにつれて大きな改善が見られることと矛盾していない。

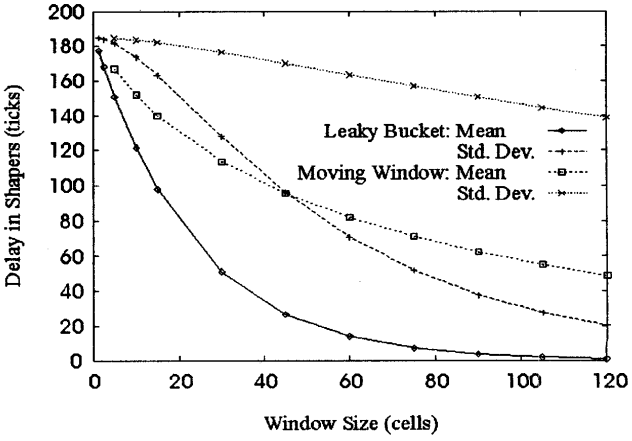


図5　シェーピングによる遅延と標準偏差

また、同じウィンドウサイズに対して、Moving Window シェーピングによる遅延および分散が Leaky Bucket シェーピングよりも大きいことを示している。これは、同様なウィンドウサイズに対し、Moving Window によるシェーピングはトラヒックのバースト性に対して、より厳しい制限を加えていることを意味する。

以上のすべての場合、シェーパー内の待ち行列の長さには制限がないものとした。一方、シェーパー内のバッファの大きさを有限、したがって最大遅延を有限とした場合のバッファの大きさと損失率の関係を示したのが図6である。ここでは、ウィンドウサイズが 15 cells と 60 cells のそれぞれのときについて調べた。バッファが大きくなるにつれ、損失が少なくなるが、同じバッファサイズでは Leaky Bucket のほうの損失率が小さくなっており、Leaky Bucket のウィンドウサイズが 15 cells のときの損失率は Moving Window のウィンドウサイズが 60 cells のときとほぼ同じになっている。

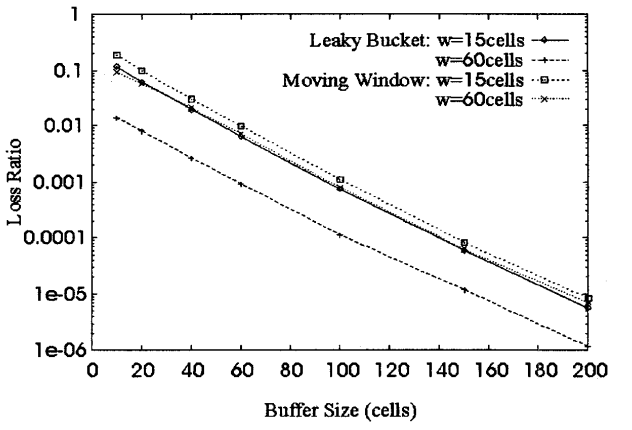


図6　シェーパー内における損失

学術情報センター電子図書館サービス　　© 学術情報センター

3.3 交換ノード内の遅延および遅延の総和

シェーピングの効果を確認するために、シェーピングされたトラヒックがネットワークの交換ノードにおける遅延について調べた。図7に中継リンクの手前の交換ノードにおけるセルの遅延の平均および標準偏差と、シェーピングパラメータの一つであるウィンドウサイズとの関係を示す。同じ交換ノードに8個のホストが接続され、それによってできる中継リンクのトラヒック密度は約 53.3 % である。図7では、ウィンドウサイズが大きくなるにつれ、シェーピングされたトラヒックの遅延はシェーピングしないときの遅延に近づいていくが、シェーパー内での遅延と同じように、ウィンドウサイズの大きさの変化に対しては Leaky Bucket のほうがより敏感に反応している。また、同図では、Leaky Bucket と Moving Window は大きな差がないものの、ウィンドウサイズ 10 cells から約 40 cells までの区間では Leaky Bucket、それ以外の区間では Moving Window によってシェーピングされたセル流のほうの遅延の平均が大きくなっている。

一方、図8では、ホスト数を12、したがって中継リンクのトラヒック密度を 80 % とした場合の遅延を示している。同図では、Leaky Bucket のほうの遅延が大きくなっている。しかし、図5のシェーパー内の遅延と合わせて比べると、次のようなことがわかる。例えば、Leaky Bucket のウィンドウサイズが 45 cells のときと Moving Window のウィンドウサイズが 90 cells のときとで両者の交換ノードにおける遅延の平均がほぼ同じになるが、これらの場合のシェーパーにおける遅延の平均は Moving Window のほうが Leaky Bucket の2倍以上になっている。すなわち、Leaky Bucket シェーパーのほうがバーストに対する制限が緩やかにもか変わらず、ネットワークにとってはそれほど大きな負担を掛けることにはなっていない。言い変えれば、よりバースト性に厳しい制限を課している Moving Window による網内での遅延の軽減の効果はあまり期待できない。

また、バーストに対する制限を緩やかにするためには、Moving Window のほうではより大きなウィンドウザイズが必要であり、実装上においても不利である。

図9では、Leaky Bucket と Moving Window のそれぞれの場合のシェーパー内の遅延と、交換ノード内の遅延の総和を示している。シェーピングによって起こるシェーパー内での遅延と交換ノード内での遅延の総和はシェーピングしない時の交換ノードのみでの遅延よりは

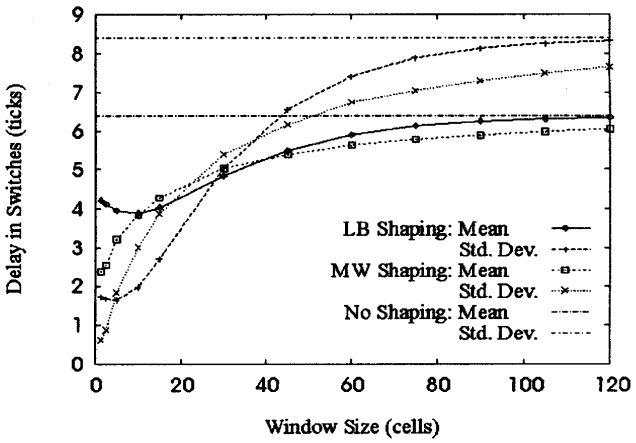


図7　交換ノード内の遅延と標準偏差：多重度8の場合

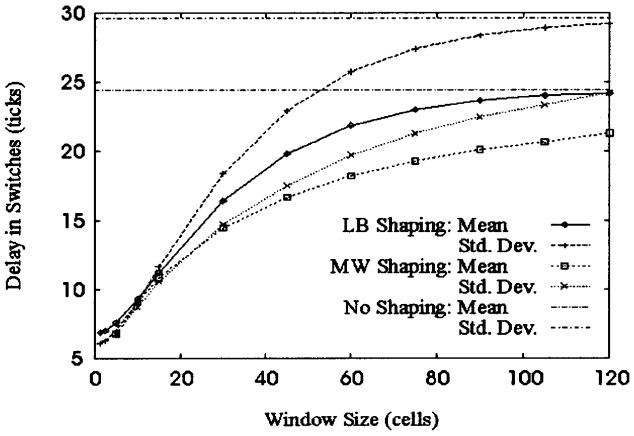


図8　交換ノード内の遅延と標準偏差：多重度12の場合

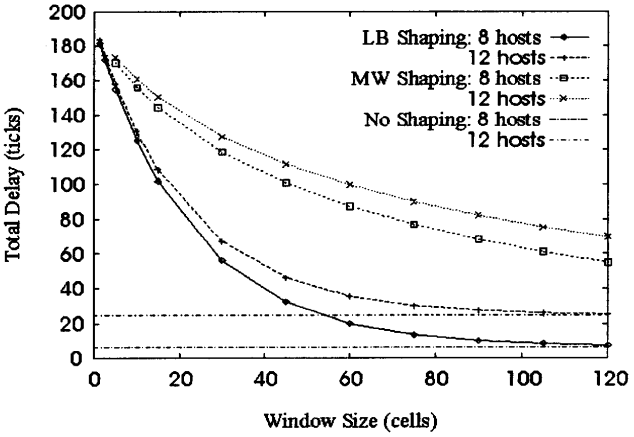


図9　遅延の総和の比較

大きくなっていることはいうまでもないが、トータルな遅延特性として、ウィンドウサイズが 15 cells および 60 cells の2つの場合とも、Leaky Bucket のほうがより優れていることがわかる。

学術情報センター電子図書館サービス　　©学術情報センター

4 おわりに

トラヒックシェーピング機構の代表として、Moving Window と Leaky Bucket について取り上げ、それぞれのアルゴリズムがトラヒックの性能に与える影響について調べた。遅延については、ネットワークの入口に設定するシェーパー及びネットワーク内の交換ノードの両方に着目し、またシェーパー内での損失についても、二つの方式の比較を行なった。それによって、Leaky Bucket が他方の方式よりもすぐれた特性を有していることを示した。

この結果は SRTS シェーパーのほうが Leaky Bucket よりも短かい期間のバーストに対する許容度が高く、シェーパー内の遅延と損失が小さいという文献 [4] の結果と相反している。それは、同文献で使われているウィンドウサイズと Leaky Bucket の最大トークン数の関係と本稿の式 (5) に示した関係が異なることから起因していると思われる。文献 [4] では、シェーパー内での遅延と損失についてだけ比較を行なっているが、二つのシェーパーを通過したセル流のネットワーク内での挙動については明らかでない。本稿で用いた簡単なネットワークモデルによって、それらのセル流のネットワーク内での遅延特性も明らかになったと言えよう。

参考文献

[1] "Traffic Control and Congestion Control in B-ISDN", Draft Revised ITU-T Recommendation I.371 , Perth, 1995.

[2] Leland, W., "Window-Based Congestion Management in Broadband ATM Networks: The Performance of Three Access-Control Policies", *Proc. IEEE GLOBECOM'89* , pp. 1794–1800, 1989.

[3] Rathgeb, E., "Modeling and Performance Comparison of Policing Mechanisms for ATM Networks", *J. on Selected Areas in Comm.* , Vol. 9, No.3, pp. 325–334, 1991.

[4] Radhakrishnan, S.; Raghavan, S. ; Agrawala, A., "A Flexible Traffice Shaper for High Speed Networks: Design and Comparative Study with Leaky Bucket", *Computer Networks and ISDN Systems* , Vol.28, pp.453–469, 1996.

[5] 計宇生,「トラヒックシェーピング機構の性能についての検討」, 信学技報 SSE96-115, Vol. 96, No. 387, pp25–30, 1996.

[6] Ramamurthy, G. ; Dighe, R., "Distributed Source Control: A Network Access Control for Intergrated Broadband Packet Networks", *IEEE J. on Selected Areas in Comm.* , Vol. 9, No. 7, pp. 990–1002, 1991.

[7] Faber, T.; Landweber, L. ; Mukherjee, A., "Dynamic Time Windows: Packet Admission Control with Feedback", *Proc. ACM SIGCOMM'92* , 1992.

[8] Turner, J., "New Directions in Communications (or Which Way to the Information Age?)", *IEEE Commun. Magazine* , Vol. 24, No.10, pp. 8–15, 1986.

[9] Sidi, M.; Liu, W.-Z.; Cidon, I. ; Gopal, I., "Congestion Control Through Input Rate Regulation", *Proc. IEEE GLOBECOM'89* , pp. 1764–1768, 1989.

[10] Bala, K.; Cidon, I. ; Sohraby, K., "Congestion Control for High Speed Packet Switched Networks", *Proc. IEEE INFOCOM'90* , pp. 520–526, 1990.

[11] 計宇生,「高速通信網の交換ノードにおけるサービス規律」, 学術情報センター紀要, 第7号, pp.129–140, 1995.

[12] Ji, Y. ; Asano, S., "Virtual Rate-Based Queueing: A Generalized Queueing Discipline for Switches in High-Speed Networks", *IEICE Trans. Commun.* , Vol. E77–B, No. 12, pp. 1537–1545, 1994.

学術情報センター電子図書館サービス　　© 学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

動画像情報サービスのためのオープンプラットフォーム技術について

Constructing Open Network Platform for Video Services

学術情報センター　相澤　彰子
Akiko AIZAWA
National Center for Science Information Systems

要旨

ネットワークを利用した動画像情報サービスの現状を概観し、広域バックボーン上でサービスを実現するために今後必要となる中継／変換技術について考察する。

ABSTRACT

An overview of current technologies for video services over wide-area communication networks is given. The importance of application level video gateway for heterogeneous internetworking is also discussed.

[キーワード]　ビデオサービス、動画像通信プロトコル、メディアスケーリング、応用層プロトコル中継サーバ、異種ネットワーク間接続

[Keywords]　video service, video protocol, media scaling, application protocol gateway, heterogeneous internetworking

1　はじめに

学術情報としての動画像情報(Video Data)には、講義ビデオ、学会発表ビデオ、研究素材データビデオ、実験結果プレゼンテーションビデオ、研究所紹介ビデオなどがあり、近年のネットワーク基盤の整備を背景に、その流通への期待が高まっている。

そこで本稿では、広域バックボーンを介した動画像情報サービスの実現に必要なネットワーク技術について考察する。まず動画像情報サービスを、通信網の利用形態および動画像情報の受け渡し方法に基づき4つに分類する。そして各サービスについて現在の技術動向を概観し、サービス実現に向けて必要となる機能を整理する。次に、ネットワーク接続点における上位層での中継／変換サービスの重要性を指摘し、動画像通信中継サーバの具体的な構成を検討する。

2　動画像情報サービスの分類

動画像情報のサービス形態は一般に、「ビデオオンデマンド型」と「ブロードキャスト型」の2つに大別される。前者ではサーバとクライアントが1対1で通信を行うのに対して、後者ではサーバが送出する1つの動画像情報を、多数のクライアント端末が同時に受信することが特徴である。特に後者の場合には、通信資源の節約の観点から、分岐点を持つ通信路(いわゆるマルチキャストツリー)を最適に構成する技術が盛んに研究されている。

ここで、ブロードキャスト型のサービスの場合に分岐点(中継ノード)の果たす役割に注目すると、中継ノードは上流ノードより受信した動画像情報を複製して下流ノードに送出する(図1)。中継ノードは、仮想的にサーバおよびクライアントに対して1対1の通信を提供することから、ブロードキャスト型のサービスは、マルチキャストツリーの動的な構成管理と中継ノードにおける複製処理という付加機能を備えた多数の片方向ビデオオンデマンド型サービスの集合として捉えることができる。

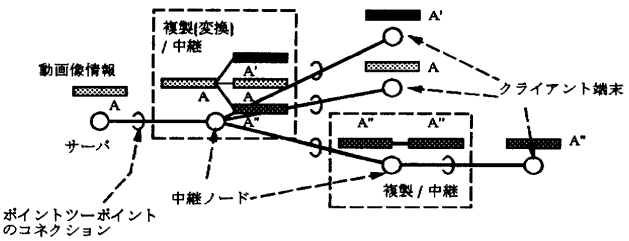


図1　ブロードキャスト型サービスにおける中継ノードの役割

学術情報センター電子図書館サービス　　© 学術情報センター

本稿では以下、サーバとクライアントによる 1 対 1 の動画像通信を中心に議論を進めることにし、動画像通信サービスを次に述べる 2 つの視点に基づき分類する。

まず、動画像通信サービスがネットワークに要求する機能に注目したものとして、(1)「QoS(Quality of Service)保証型」および (2)「ベストエフォート型」の 2 者への分類がある。前者では、ある基準を満足する動画像情報をサービスするために、ネットワークに対して遅延やパケット損失など一定の通信品質の保証を要求するが、後者では、特にサービスの品質に基準を設けることはなく、ネットワークが提供可能な通信資源を最大限に利用することを目標とする。今日のインタネットの基本となる IP(Internet Protocol)が後者の型のサービスを提供するプロトコルであることはよく知られており、今後いかに QoS 保証型のサービスを実現するかが話題の中心であるといえる。

次に、サーバとクライアント間での動画像情報の受渡し方法に注目したものとして、(a)「リアルタイム型」および(b)「バッチ型」の 2 者への分類がある。前者では、受信したデータをローカル記憶装置に蓄えることなく順次表示するが、後者では受信データを一端バッファやファイルに蓄積してすべてのデータの受信を確認した後に端末上で表示を行う。

この 2 者の組合わせで決まる動画像データのサービス形態は表 1 に示す 4 通りである。以下では各々について現在の技術動向を概観し、サービス実現に向けての課題を整理する。

表 1　動画像情報サービスの分類

	(a)リアルタイム型	(b)バッチ型
(1)QoS 保証型	一定の品質の動画像をリアルタイムで端末に表示する	一定時間以内に転送を終了して端末に動画像を表示する
(2)ベストエフォート型	受取った動画像をとにかく(端末の能力が許す限り)リアルタイムで表示する	動画像を通常のファイルとして受取った後表示する

2.1　「QoS 保証／リアルタイム型」動画像情報サービス

現在、動画像通信という場合に前提とされるのは、「QoS 保証型」かつ「リアルタイム型」のサービスである場合が多い。いわゆる連続的メディア(Continuous Media)としての扱いである。このようなサービスでは、一定の通信品質での継続的なデータ転送が必要であるが、現在コンピュータネットワークの中核をなすインタネットは、前述のように残念ながらこのための仕組を持たない。サービス実現のためには、従来のインタネットの枠組を越えた次世代のネットワーク基盤が必要である。

このための具体的なアプローチとして以下の 2 つが考えられる。1 つは ATM を中心にしたまったく新しいネットワーク基盤を構築することであり、いま 1 つは既存のインタネット上に新たなサービスを定義して動画像通信に必要な機能を実現することである。前者にかかわる標準化として、ATM-Forum による MPEG over ATM [1]、後者にかかわるものとして、IETF(Internet Engineering Task Force)による RSVP(Resource ReSerVation Protocol) [2] や ST2 (Internet Stream Protocol Version 2) [3]、より包括的な機能標準として DAVIC (Digital Audio-Video Council)による DAVIC 1.0 [4] などがある。しかし、いずれのアプローチとも、広域ネットワーク上でこれを実現するためには、通信にたずさわるすべての機器の更新や通信回線の大幅な増強などを必要とし、現在のインタネットに匹敵する規模での情報流通の実現からはほど遠い段階にあるのが現状である[5]。

一方、ローカルエリアネットワーク(LAN)環境においてはバックボーンの高速化や高機能化が比較的容易であることから、サーバとエンドユーザのクライアント端末間での動画像通信が一部ではすでに現実のものになっている。特に、動画像情報の検索とリアルタイム転送を基本とするビデオオンデマンドシステムについては、各社から多くの製品が発表されており、先進的なサービスを目指した開発も盛んである。これらは均一の高速ネットワークでむすばれた小規模なシステムを前提としている場合が多く、ネットワーク上に 1 つあるいは複数個の動画像サーバを立ち上げ、GUI に工夫を凝らした専用のクライアント端末上で表示を行うのが標準的な使い方である。図 2 に典型的なビデオオンデマンドシステムの構成例を示す。著者の所属する学術情報センターにおいても、すでに数社のビデオオンデマンド製品を試行的に導入しているが、問題点として、これらの製品が動画像の蓄積や通信に独自の方式を用いており、容易に相互接続できないことがあげられる。

また、上記に関連する話題として、近年実用化され話題を呼んでいる IP スイッチ[6] がある。ATM 交換器の高速なスイッチングハードウェアを用いて IP パ

学術情報センター電子図書館サービス　　© 学術情報センター

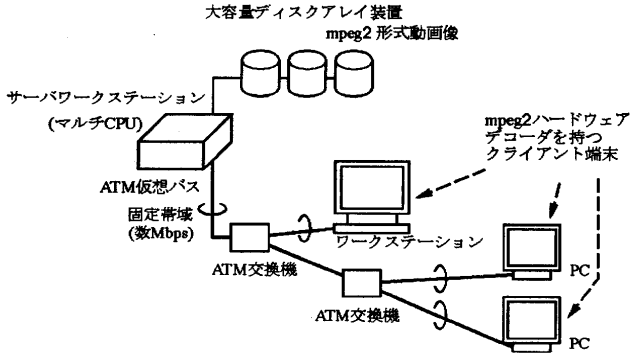


図2　高速 LAN を前提とした
ビデオオンデマンドシステム構成例

ケットのルーチングを高速に実行するもので、連続データ転送に対して動的に ATM の仮想パスを設定できる。仮想パスの設定はパケットの統計的性質に基づき交換器と端末の NIC (Network Interface Card)の間で自動的に行われる。通常、アプリケーション側では従来のインタネットと区別せずに通信を行うが、必要に応じてこの ATM 仮想パスを帯域保証の目的で用いることも可能である。また文献[5] で参照しているメディア別交換機なども資源割当ての概念を含んでおり、このタイプのサービスの延長として分類が可能である。

2.2　「QoS 保証型／バッチ型」サービス

現在のインタネットは IP プロトコル、すなわちデータを一定長さ以下の独立のパケットに分割して送出するコネクションレス型通信を基本にしているため、バッチ型転送の場合でも QoS(この場合には主に最大遅延時間)を保証したサービスの提供はむずかしい。一方で、リアルタイム型サービスと異なり、バッチ型に関しては研究例もそれほど多くはなく、意外にこのタイプのサービスが話題にのぼることは少ない。

しかし現実問題として、通信回線が極めて混雑した状況では、このタイプのサービスの有用性は高いと考えられる。具体的には、たとえば、トラヒックのピーク時にファイル転送がタイムアウトしてしまう海外回線を用いて、ある時刻までに発表用の動画像データを相手先ホストまでファイル転送したい、あるいは一定量の観測データを最大遅れ 24 時間で転送したい、という形態の利用である。

このようなサービスの実現方法としては、リアルタイム OS におけるタスクスケジューリングの技術が適用可能である。すなわち、利用可能な通信容量を CPU と同様に資源とみなして、その両端で資源要求の受付けおよび資源割当てのスケジューリングを行う。回線を 1 本のインタネット専用線ではなく、あらかじめ帯域を確保した仮想的な複数本の通信路とみなす必要があることから、物理的に回線が 1 本である場合には、フレームリレーや ATM などの設備が必要である。

また上記の方式において、転送終了の期限を端末の表示速度にあわせて設定してタスクスケジューリングを行えば、バッファ機能つきのリアルタイム型サービスを実現できる[7]。このような技術は、統計多重した可変レートの動画像情報サービスなどにおいて、今後必要になると考えられる。

通信の経路中に複数の中継ノード(ルータ)が存在する場合には、全体として資源受付けおよび割当ての制御が複雑になるが、たとえば国際回線などのように、ボトルネックになる回線が明らかである場合には、問題となる回線の両端で制御を行えばよいので実現は比較的容易である。

2.3　「ベストエフォート型／リアルタイム型」サービス

このタイプのサービスは、現在のインタネットが提供するベストエフォート型のネットワーク機能をそのままの形で利用して、その上でリアルタイム型の動画像サービスを実現するものである。

ここでサービス実現のポイントとなるのは、時々刻々と変化するネットワークの状況に対応して動画像データの転送レートを調整する適応的トラヒック制御技術である。ネットワーク自体はこのような機能を持たないので、サーバとクライアント上のアプリケーション間で適切に転送データ量を制御して、過剰なトラヒックによるネットワークの輻輳を回避しなければならない。

たとえば、現在のインタネットのテレビ会議ソフトウェアでは一般に、画面サイズと単位時間あたりのフレーム数を GUI を介して会議の参加者が予め指定することにより、転送データ量の設定を行うようになっている。利用者自身がネットワークの混雑状況を見極めて転送量を調節するわけである。

より高度な制御を行う例として、イリノイ大学にある NCSA の研究グループによる VDP (Video Datagram Protocol)[8] やカリフォルニア大学バークレー校の研究グループによる CMPlayer (Continuous Media Player)[9] がある。図 3 に、これらの動画像通信プロトコルの概略を示す。まず、サーバとクライアント間では制御用と動画像用の 2 本のコネクションを設定する。ここで、

学術情報センター電子図書館サービス　　©学術情報センター

前者には誤り回復機能を持つコネクション型プロトコル(TCP)、後者には誤り回復機能はないがオーバヘッドが少ないコネクションレス型プロトコル(UDP)を用いるのが一般的である。動画像データの転送が開始されると、クライアント側では制御用のコネクションを用いて受信状況を定期的にサーバに通知し、一方、サーバ側では受取った情報に基づき動画像の転送レートを調整する。

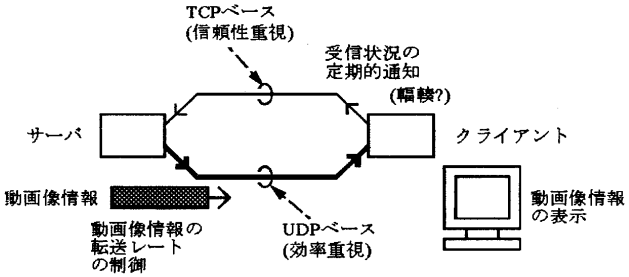


図3　インタネット動画像通信プロトコルの実現方法

VDP では、受信状況を表す指標として損失率、すなわちサーバの送信データ量に対するクライアントの表示データ量の割合を用い、損失率がある値以上(以下)に達した場合に転送レートを増す(減らす)。CMPlayer では、受信状況の指標として「ペナルティ」、すなわち連続したフレーム欠落に対して重みを加算したフレーム受信率を用い、ペナルティ値に応じて常時転送レートの修正を行う。両者とも動画像データは mpeg 符号化形式で蓄積されており、転送レートの調整は B または P フレーム(欠落した場合でも比較的影響が少ない mpeg のフレーム情報)を間引いて転送フレーム数を減らすことで実現している。

その他、より洗練された転送レート制御アルゴリズムに関する研究例として 文献 [10][11][12]、また、すでに製品化されているビデオオンデマンドシステムとして Xing Technology Corp.の StreamWorks や VDOnet Corp.の VDOLive などが存在するが、いずれも既存のインタネットを用いながら、図 3 の方式で適応的に転送レートを調節するものである。ここでクライアントからサーバに送られる受信状況情報には時間遅れが存在するため、アプリケーションレベルでの転送レート制御は本質的にネットワークの輻輳を回避するものではないことには注意が必要である。

最後にインタネット上では、アプリケーションごとに独自の動画像通信プロトコルが使われており、異なる方式によるサーバとクライアントどうしでは動画像情報サービスが成立しないのが一般的である。これに関しては、Netscape の PlugIn 機能などを利用することによってサービスのオープン性を高めることができる。また、JAVA のような仮想的な実行可能コードをダウンロードすることによって、独自の動画像通信プロトコルを任意の端末で実現することも将来の可能性として検討されている。

2.4 「ベストエフォート型／バッチ型」サービス

現在のインタネットの典型的なファイルアクセスは、このタイプのサービスである。たとえば通常の WWW(World Wide Web)ブラウザを用いて mpeg 形式の標準的な動画像情報にアクセスすると、他のファイルアクセスの場合と同様に HTTP (Hyper Text Transfer Protocol)による転送が行われ、クライアント側ではすべてのデータを受信後に端末のビューアを起動することになる。

大容量の動画像情報の場合、たとえ LAN のような高速ネットワーク環境であっても、ローカルディスクの容量や表示開始までの遅延時間などが問題になる。たとえば標準テレビ品質で 24 分程度の学術情報センター紹介ビデオ「すべての研究者とそれを支える人達へ」は mpeg2 形式で約 1GB あり、LAN 環境でも表示可能端末に制約が生じてしまうのが現実である。また広域ネットワークの場合は、動画像情報への過剰なアクセスはいたずらなトラヒック増加を招き、結果としてより優先度の高い他の通信を妨害してしまうので好ましくない。ネットワーク運用の現場においては、通信回線を効率的に運用するために、頻繁に参照されるファイルを中継ノードに一時的に蓄積するキャッシュサーバの導入が一般的であるが、必ずしも動画像情報を効率的に転送するためものではない。

3 動画像通信中継／変換サービスの検討

3.1 ネットワーク間接続と動画像情報サービス

前節で概観したように、現状のインタネットによる広域バックボーンにおいてもっとも容易に実現できるサービスは、「ベストエフォート型／バッチ型」の組合わせであり、その他のサービス形態については、当面、通信環境の優れた個々のネットワークを中心に広がって行くと考えらる。

ここで、個々のサービスに最適な通信方式は、利用可能なネットワーク資源や提供する動画像情報の種類によって異なる。したがって動画像通信に関しては、インタネットにおける IP のように、すべてのネット

学術情報センター電子図書館サービス　　© 学術情報センター

ワークが共通の方式を採用することは考えにくい。

たとえば現在、インタネット上での帯域保証が盛んに検討されているが、動画像情報サービスであってもネットワークによる QoS保証は必ずしも不可欠な要素ではない。ネットワーク資源が潤沢で課金の必要もない LAN 環境においては、実現が容易で負荷が少ないベストエフォート型がむしろ適しているといえる。圧縮された動画像情報は本質的に可変レートである場合が多く、一定帯域の確保などはネットワークの視点から見た QoS の定義でしかないことに注意が必要である。

さらに、クライアント端末側で表示用の専用ハードウェアを備えている場合には、表示可能な符号化方式が特定されるし、独自のサービス方式によって、動画の品質を高めたり、早送りや検索などの高度な機能を実現する場合もある。

そこで重要になるのが、多様な動画像情報サービスを相互に接続し、少しでもオープンな環境を実現するための、ネットワーク接続点における中継／変換技術である(図4)。特に、広域ネットワークを利用して広く動画像情報サービスを提供するためには、異なる情報通信基盤が接続するネットワーク境界点において、これらの通信の差異を吸収することが重要である。

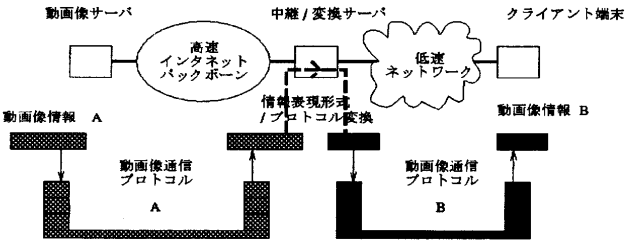


図4　ネットワーク境界点における中継／変換機能の提供

ここで、QoS 保証型のサービスは、ネットワークの接続点において転送レートを調整することによって、ベストエフォート型に変換できる。リアルタイム型はキャッシュによってバッチ型に変換できる。逆に、転送元のネットワークにおいて、遅延などが問題にならないほど資源が豊富である場合には、ベストエフォート型を QoS 保証型、バッチ型をリアルタイム型に変換することも等価的に実現可能である。

また同種のサービス形態の場合でも、中継／変換サーバによって、符号化形式の変換やパラメタ値の変更による圧縮率の変更、プロトコルヘッダの書き換えを行うことによって、ネットワーク間のサービスの差異を吸収し、情報の流通性を高めることができる。

3.2　動画像通信中継／変換サーバに必要な機能の実現

本稿では、バックボーンネットワークと LAN などのネットワーク境界点において動画像情報の変換および中継を行うサーバに必要な機能として、(1)動画像通信プロトコルの高速中継および変換、(2)提供される動画像情報と利用者に関する属性情報の管理データベース、(3)変換前後の動画像情報のキャッシング、の3つをあげる。以下、各々の実現について簡単に述べる。

(1)　動画像通信プロトコルの高速中継および変換

プロトコルそのものや符号化形式の違いを吸収するため、上位層における高速の中継／変換機能の提供が必要である。動画像の通信においては一般に、ネットワークの能力に応じて転送データ量を動的に制御する必要があるため、情報の通信方式(プロトコル)と表現形式(符号化方式)を同時に定める必要がある。これに対応して中継／変換サーバ提供する機能としても、上位層パケットプロトコルヘッダの高速解析と書換え、および、専用のハードウェアなどによる高速の動画像情報変換の2つがある。

特に後者は QoS フィルタリング(QoS Filtering)あるいはメディアスケーリング(Media Scaling)と呼ばれ、このための動画像の符号化方式として mpeg2 をはじめとする種々の階層的符号化が提案されている[13]。一方で、現実に存在する圧倒的多数の画像データは、JPEG や MPEG1 などの形式で蓄積されていることから、これらに対応したスケーリング技術も重要である[14]。このようなメディアスケーリング技術の分類として [15][16] では、(1)符号化方式の変換、(2)フレーム落とし、(3)DCT 係数の変換、(4)量子化ステップの変更、(5)複数画像の合成、の 5 つをあげている。その他、画素数および画素値の変換などもあり、それぞれ必要となる復号／符号化レベルやデータ圧縮効果が異なるため、目的に応じた使い分けが必要である。

(2)　提供される動画像情報と利用者に関する属性情報の管理データベース

変換／中継サーバで提供するべき機能は、実際にコネクションを設定するサーバとクライアントの組合わせで決まる。適切な変換処理を起動する

学術情報センター電子図書館サービス　　©学術情報センター

ために、サーバおよびクライアントの属性情報を管理するデータベースが必要である。

このような管理データベースでは動的に変化する大量の情報を扱うため、エントリの自動生成や問い合わせに対する高速マッチング方式などを工夫する必要がある。具体的には、HTTP 要求／応答メッセージのヘッダに含まれる属性情報の利用、データベースの並列化などが考えられる。また、クライアント側の端末属性にあわせてデータ変換を行うとともに、サーバ側でのデータの表示形式や利用者範囲の指定を可能にして、提供データのオリジナリティを確保することも重要である。

(3) 変換前後の動画像情報のキャッシング

ネットワーク接続点におけるキャッシングは、回線の有効利用だけではなく、中継／変換サーバの負荷の軽減のために有効である。中継／変換サーバ上では、同一の動画像情報に対して品質の異なる表現が複数生成されるため、いずれの情報をキャッシングするのが有効か、アクセス履歴などに基づく適切な判断が必要である。

4 おわりに

ネットワーク情報の流通が急増、多様化する中で、通信回線の効率的運用や利用者への高度なサービス実現のため、上位層での情報変換機能の提供が強く要求されている。本稿で述べた動画像の中継／変換技術は、たとえば大規模科学技術シミュレーションや衛星画像など、多次元観測データの提供にも適用可能である。これらの動画像は、標準テレビ画像と比較すると特殊で複雑な扱いを必要とする。表示端末の属性、利用目的、通信路の状態などに応じた様々な変換サービスを行うことによって、もとになる情報に付加価値を与え、提供者と利用者双方の利便をはかることができる。

また動画像情報の変換以外にも、文字コードの変換、アプリケーションにかかわる管理情報の収集と交換、セキュリティのための高度パケットフィルタリング、アプリケーションセションの暗号化、登録話題に関連する情報を選択的に配送する情報フィルタリングなどのサービスが考えられ、今後の急速な展開が予想される。

参考文献

[1] http://www.atmforum.com.

[2] R, Braden. ; L, Zhang. ; S, Berson. ; S, Herzog. ; S, Jamin. , "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification" , IETF draf-ietf-rsvp-spec-14.txt, 1996.

[3] Delgrossi and Berger eds. , "Internet Stream Protocol Version 2 (ST2) Protocol Specification - Version ST2+" , IETF Request for Comments 1819, 1995.

[4] http://www.davic.org

[5] 森野,「インターネットにおける実時間通信のQoS保証」, 学術情報センター紀要,第9号,1997.

[6] Ipsilon Networks, "IP Switching: The Intelligence of Routing, the Performance of Switching" , IPSILON Product Information, 1996 (http://www.ipsilon.com /productinfo/techwp1.html).

[7] H, Woo. ; J, Parc. ; C, K, Kim. , "Multicast Scheduling for VOD Services" , Multimedia Tools and Applications, Vol.2, No.2, pp. 157-171, 1996.

[8] Z, Chen. ; S, M, Tan. ; R, H, Campell. ; Y, Li. ,"Real Time Video and Audio in the World Wide Web" , http:/choices.cs.uiuc.edu/Papers/New/vosaic/vosaic. html.

[9] L, A, Rowe. ; B, C, Smith. , "A Continuous Media Player" , Proceedings of 3rd International Workshop on Netwrk and OS Support for DigitalAudio and Video, 1992.

[10] S, Chakrabarti. ; R, Wang. , "Adaptive Control for Packet Video" , Proceedings of the International Conference on Multimedia Computingand Systems, IEEE, pp. 56-62, 1994.

[11] S,Cen. ; C, Pu. ; R, Staehli. ; C, Cowan. ; J, Walpole. , "Demonstrating the effect of software feedback on a distributed realtime MPEG video audio player" , Proceedings of the ACM International Multimedia Conference & Exhibition,1995.

[12] P,Goyal. ; H, M, Vin. ; C, Shen. ; P, J, Shenoy. , "A Reliable, Adaptive Network Protocol for Video Transport" , Proceedings of IEEE INFOCOM, Vol.3, pp. 1080-1090, 1996.

[13] 藤原,大久保,「インターネット時代の画像圧縮技術 -MPEG からケーブル・モデムまで」,アスキー出版局, 1996.

[14] E, Amir. ; S, McCanne. ; H, Zhang. , "An Application Level Video Gateway" , Proceedings of the ACM International Multimedia Conference & Exhibition,1995.

学術情報センター電子図書館サービス
© 学術情報センター

[15] N, Yeadon. ; A, Mauthe. ; F, Garcia. ; D, Hutchison. , "QoS Filters: Addressing the Heterogeneity Gap" , Proceedings of European Workshop op Interactive DistributedMultimedia Systems and Services, 1996.

[16] N, Yeadon. ; F, Garcia. ; D, Hutchison. ; D, Shepherd. , "Continuous media filters for heterogeneous internetworking" , Proceedings of SPIE - Multimedia Computing and Networking, 1996.

学術情報センター電子図書館サービス　　© 学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

Neural Networks Approach for Query Cost Evaluation of DBMS

DBMSにおけるクエリーコスト評価のためのニューラルネットアプローチ

Jihad BOULOS
National Center for Science Information Systems
学術情報センター　ジハド　ブロス
Yann VIÉMONT
Laboratoire PRiSM Univ. de Versailles-St-Quentin France
ベルサイユ大学(フランス)　ヤン　ビエモン
Kinji ONO
National Center for Science Information Systems
学術情報センター　小野　欽司

ABSTRACT

This paper presents a new approach for cost evaluation of query execution plans that may help or replaces the known analytical approach. Our proposed approach is based on neural networks and the connectionist concept. A neural network is trained to learn the execution cost of the implementation algorithm for a logical algebra operation (e.g., a merge-join implementation of the join operation) under some database conditions and in a certain environment; after that, this network is used to estimate the cost of this operation with other entries. The approach is based on a curve fitting like since neural networks have been proven to be "universal approximators." The main advantage of this approach is its applicability to user defined methods where the user does not need to estimate the cost of its method since the system may apply this method several times, collects measurements, and capture its behavior with its curve fitting capacity. Another advantage of this approach is its implicit covering of environment depending parameters.

要旨

本論文はデータベース管理システムのクエリー実行計画のコスト評価における既存の解析モデルに代る新しいアプローチを提案している。本アプローチはニューラルネットワークと conectionist concept に基づいている。

ニューラルネットワークはデータベースの一定条件環境下での論理的な代数操作やユーザ定義のメソッドの実行コストを算出することを学習するように訓練される。そしてこのネットワークが他のケースでの同様の操作のコストを評価するために使用される。このアプローチはニューラルネットワークが普遍的に適切であることが既に証明されているため、本アプローチの一番の利点は、ユーザがコスト評価しなくてもユーザ定義メソッドへの適用性が良いことである。他の利点は環境依存のパラメータを暗黙的にカバーすることである。

[Keywords]　Databases, Optimization, Performance Prediction, Neural Networks, Cost Evaluation
[キーワード]　データベース、最適化、性能評価、ニューラルネット、コスト評価

学術情報センター電子図書館サービス ©学術情報センター

1 Introduction

In recent years database systems had to introduce several new functionality in their kernels to respond to the ever increasing demands of users. This evolution introduced several complexities in database optimizers and cost models from old ones [27, 28, 19] to new ones [2, 12, 10, 17, 21, 25, 14, 11] in order to cope with it. However, in most cases the cost models used by these optimizers remain primitive regarding the real execution complexity. This simplification may mislead the optimizer to choose a Query Evaluation Plan (QEP) that was estimated cheaper than others in a certain environment (platform, DBMS, data), but is more expensive in the real one since most of these cost models are very simple, do not evolve with the environment, and make some assumptions on several factors (cache hit ratio, system charge, size of sorting area, available number of processors, data placement) influencing the execution cost at execution time.

In this paper we study the feasibility and effectiveness of replacing an analytical cost model for database performance evaluation by a set of neural networks. The approach builds upon replacing each analytical cost formula for estimating the execution cost of a physical algebra operation by a neural network. This network is trained to capture the operation execution cost in a certain environment by a set of inputs (operation cost influencing parameters) and an output (the execution time). This implies that a phase of input-output measurement followed by a training phase must proceed the exploitation phase.

This paper is organized as follows: Section 2 exposes some weaknesses and new requirements in current database cost evaluation techniques and Section 3 gives an idea about the theoretical foundations of neural networks and their utilities in our problem. Section 4 discusses ways of constructing and training neural networks for queries and operations cost prediction and Section 5 presents issues in using neural networks in evaluating some of the different execution algorithms of database operations. In Section 6 we present the result of experiments we made to test the efficiency of neural networks in database performance evaluation and in Section 7 we extend the experiments to cost learning and prediction of user defined methods. Section 8 discusses some related works. A conclusion and some future directions are stated in Section 9.

2 Limitations and Future Needs

Several limitations exist in analytical cost models for query processing cost evaluation; we stress here these limitations and state some of the most emerging requirements in database technology.

2.1 Analytical Model Limitations

We have already said that analytical models remained simple and inaccurate in most optimizers and performance evaluation tools; we give here evidence to this claim.

As a first example of a cost formula often over simplified in most known query optimizers is the $(n \times \log_2 n)$ number of I/O for a sort operation; the base 2 is usually used for the logarithm while in reality the correct base must be the number of available clusters (merging units) [13]. This last parameter may radically change the sort behavior from an I/O-bound operation to a CPU-bound one, when there is a sufficient number of buffers. The estimated cost for a QEP that includes a sort operation may get completely out of a reasonable error rate. A second example is given in [17] where the authors showed that, whenever sufficient disks are available, a sequential scan is first CPU-bound when the tuple sizes are small and only becomes I/O-bound when these sizes get large.

Another simplification in most optimizers is the non-consideration of the buffer cache hit ratio; System R [27] and Postgres [30] compute the cost of a sequential scan as follows:

$$Cost = \#page_io + w \times \#tuples_processed;$$

w is an adjustable weighting factor between I/O and CPU. The "#page_io" is taken as is without any consideration to the cache or at best with a uniform hit ratio assumption. There are many situations where this formula is not correct. [17] identified two situations in the execution of single operations that are sensible to the available buffer size: 1) choosing between a sequential scan and an index-scan with an unclustered index, and 2) choosing between a nested-loops with an index-scan over the inner relation and a hash-join.

An additional influencing limitation of cost models in most commercial and research DBMSs is the relative rather than absolute cost estimate approach. With relative estimates,

学術情報センター電子図書館サービス ©学術情報センター

expected numbers of I/O, CPU instructions, and messages are predicted without any consideration to the underlying platform, its disk throughput, CPU power, and communication bandwidth. Even with a fundamental modification to the underlying platform, the cost model does not change and its error rate may grow out of a reasonable range.

2.2 New Database Requirements

The next generation of database systems may converge to an industry standard combining the extended relational model with the object oriented model. This combination will lead to a further complication of cost evaluation for query processing.

An OQL [5] query may need to apply several methods during its execution. Some of these methods may be written and integrated in a database system by normal users. What is the selectivity of each method? How much instructions are executed during the processing of each method? In which order should the system apply the methods? Until nowadays, and as far as we know, there is no implemented algorithm to predict these valuable estimates.

[18] stated that when a user adds a function to Postgres, there is little hope for the system to correctly estimate the cost and selectivity of predicates containing this function in an initial state. To add a function to Postgres, the user writes the function in C and compiles it. He must then issue a Postquel's *define* function command with a number of special flags to capture the expense information of the function; these flags are summarized in Table 1 (reproduced from [18]). The cost of a predicate is then computed by adding up the costs of each expensive function in the expression. The cost per tuple of a predicate $p(a_1, \ldots, a_n)$ is recursively defined as:

$$e_p = \begin{cases} \sum_{i=1}^{n} e_{a_i} + percall_cpu(p) \\ \quad + perbyte_cpu(p) \times (byte_pct(p)/100) \\ \quad \times \sum_{i=1}^{n} bytes(a_i) + access_\cos t \\ \qquad\qquad \text{if } p \text{ is a function} \\ 0 \qquad \text{if } p \text{ is a constant of tuple variable} \end{cases}$$

The problem in this process is that the user must be an expert who can give the exact values of the flags. A solution that the paper suggests is "*after repeated applications of a function, one could collect performance statistics and use curve-fitting techniques to make estimates about the function's behavior.*" We believe that our proposed technique may produce correct estimates: the curve fitting solution is applied using neural networks.

[15] studied expensive predicate placement in QEPs. He proved that the common relational heuristic of "selection pushdown" is no longer advantageous in all situations. For several selection predicates in a query, and in addition to its selectivity, the cost of applying a predicate must be taken into account: the cheaper the predicate, the earlier we wish to apply it, since its benefits may be repeated at a low cost. Selection predicates over a given relation are hence ordered in ascending order of the metric

$$rank = \frac{selectivity - 1}{\cos t}$$

In the join case, the used cost model needed to allow the selectivity of a join node to be different for each input stream to the join, something that known global cost models cannot capture. The adopted approach states that given two relations R and S, and a join predicate J of selectivity s over them, the selectivity of J over R is represented as $s.\{S\}$, where $\{S\}$ is the number of tuples that are passed into the join from S. Similarly the selectivity of J over S is represented as $s.\{R\}$.

To cope with the mentioned advances and requirements of new technologies in database systems, we propose here a technique based on neural networks for query processing performance evaluation. To identify the concepts that interest us in neural networks, we give in the following section some background knowledge about this connectionist approach.

3 Neural Network Backgrounds

Neural networks have proven their utility in several domains such as telecommunications, pattern classification, prediction, ... [32]. We are mostly interested here with the use of neural networks for the approximation of functions. This section begins with a brief review of basic neural network definitions. Then, recent theoretical results of neural-networks for the approximation of any measurable function are summarized.

学術情報センター電子図書館サービス ©学術情報センター

Table 1 Function Expense Parameters in Postgres

Flag name	*description*
percall_cpu	*execution time per invocation, regardless of the size of the arguments*
perbyte_cpu	*execution time per byte of arguments*
byte_pct	*percentage of argument bytes that the function will need to access*

3.1 Neural Network Definitions

Neural networks started as a model for a set of neurons put connections, a bias value, an activation value computed from all its inputs, a threshold function, an output value, and a set of output connections. Each connection has an associated weight which modulates the effect of the output value of its source unit on the activation value of its target unit. Each of these values is represented by a real number.

With simple networks, the input values of a unit are summed to compute the activation value. Hence, the activation value η for unit i at time t is given by

$$\eta_i(t) = \sum_j \omega_{ij} x_j(t) + \beta_i$$

where ω_{ij} is the weight of the connection from $unit_j$ to $unit_i$, b_i is the unit bias value, and x_j is the output value of unit j. The output value of a unit is normally given by a non-linear function (the activation function) of its activation value. A typical choice of such a non-linear function –also called a squashing function– is the "*sigmoid.*" The reader is referred to [26, 20, 23] for an introduction.

Specific units of a neural network are used as global inputs and as global outputs. Input units receive their input values from the outside world with no weight associated and output units deliver their output values to the outside world. As a whole, a neural network may be seen as an adaptive function of its outside inputs to its outside outputs. This result is achieved through learning. The problem of learning in neural networks consists in finding a set of connection weights and unit biases which allow the network to carry out the desired computation. A special class of neural networks is the feed-forward network (Fig.1) where units are organized in layers from the input unit layer to the output unit layer and where connections are only allowed between two adjacent layers (in that direction). The back-propagation learning algorithm is the most popular method to train feed-forward networks.

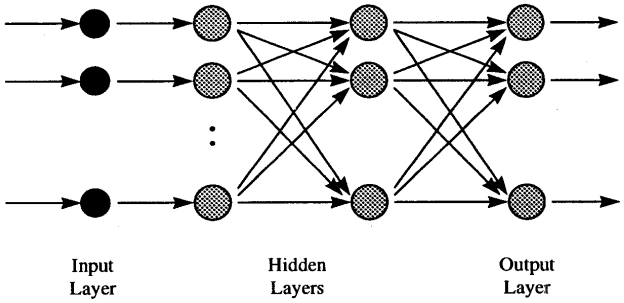


Fig.1 A Feed-forward Neural Network with Two Hidden Layers

For a neural network to learn how to model a given system, it must be trained with a set of input/output pattern pairs. Each input pattern is itself a set of all outside input values for the system at a given time and an output pattern is the set of corresponding (desired) output values. The training patterns are generated from the real system to model. After the training patterns have been given to the neural network in the learning phase, and after usually many training cycles on these data, the network may "generalize" and give the desired output for a new input.

The main interest of neural networks for our work is their capability to approximate to a desired degree any measurable function. In our case, the desired outputs are mainly the costs of basic database operations. The objective is hence to obtain a set of neural networks that by interpolation can approximate the input/output set of data points that capture the execution cost of physical database operations. This problem is classical but tricky in statistical regression: the shape of the approximation function must be established in advance, meaning trying several shapes before reaching a good approximation. Neural networks are more versatile. While it remains necessary to try several network structures to find the one that fits the best, the choice is only driven by the complexity of the algorithm implementing a physical database operation for cost estimation with no other assumption on the shape of the curve.

学術情報センター電子図書館サービス ◎学術情報センター

3.2 Function Approximation

Hornik, Stinchcombe, and White were the first to prove in the following theorem the universal approximation property of feed-forward neural networks [16].

HSW Theorem: *standard multi-layer feed-forward neural networks with as few as a single hidden layer can approximate any continuous function uniformly on any compact set and any measurable function arbitrarily well in the corresponding metric, regardless of the squashing function (continuous or not), regardless of the dimension of the input space, and regardless of the (finite) measure used.*

Hence, feed-forward networks are capable of approximating almost any multidimensional function to any desired degree of accuracy, provided sufficiently many hidden units are available. As a consequence, HSW theorem establishes that "*failures in applications can be attributed to inadequate learning, inadequate numbers of hidden units, or the presence of a stochastic rather than a deterministic relation between input and target.*"

The HSW theorem above formally proves the power of neural networks for function approximation. In [22], the authors prove that for a neural network to approximate a given function, its activation function must be non-polynomial. [7] generates a robust back-propagation learning algorithm for function approximation that can tolerate some imprecise measurements or undesirable measurements peaks as those found in real world environments. Finally, the problem of generating a good network structure for function approximation is addressed in [24]. The algorithm generates several network structures based on generate-and-test procedure and the one with the smallest error is kept. However, it is unclear if a minimal structure is obtained with this algorithm.

4 Neural Network Evaluation Approaches

From the beginning of our investigation on the possibilities of using neural networks in query processing cost estimation, we faced two main issues to resolve. The first issue was to decide whether is it better to model and evaluate whole queries or basic operations in their QEPs? The second issue was the choice of the statistical parameters that capture in the most accurate way the specifications of the database and the queries and that are relevant as entries in a neural network. We discuss in this section the nature and possible approaches of query processing cost estimation by neural networks. The next section addresses practical issues of the approach and identifies some of the most influencing and relevant parameters in our problem.

The first question that we asked to ourselves was whether is it better to have a single large network structure for each type of queries or to have a small network for each atomic operation in a QEP and to connect these small networks in a large one according to each query? The advantage of the first choice is its bounded global error rate while its inconvenience is the large structure of the networks and the null inputs that these networks may have when a simple query is injected into one of them. On the other hand, the advantage of the second choice is its adaptability to the QEP structure of any query while its inconvenience is the propagation of the error rate of the different networks.

In a context where cost evaluation of queries is the main subject, a single neural network may perform better than a network of neural networks while in an optimization context where different QEPs are to be evaluated, the second choice may be more convenient. The choice may hence be taken according to the usage needs, i.e. a single network for performance evaluation, and a network per operator in optimization. However, if the user has access to the QEP generated by the underlying DBMS in performance evaluation, it may be more convenient to use the second solution.

4.1 Modeling Queries

In a first approach of query processing cost evaluation by neural networks, whole queries may be measured, learned, and evaluated. In this context a classification of query types must be processed to identify the fittest network to a query, and then the network is used to estimate the cost of the query. A finite set of query types must be identified and modeled.

The types of queries may by classified according to several criteria: the number of joins in the query, the number of predicates in the query, the complexity of the functions that must be applied in the query. Other criteria may also be considered. Since neural networks have a powerful capability of classification, they may classify in a first phase an evaluated query in a certain type (Fig.2 a), and then evaluate the cost ofthe query according to the neural network that captures the execution model of the type (Fig.2 b).

This approach is relatively simple but the identification of query types and the classification of queries complicate the evaluation process. More research on these two aspects in this approach must be done.

学術情報センター電子図書館サービス ©学術情報センター

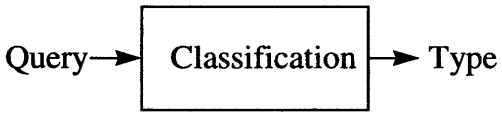


Fig.2 a: A First Phase of Query Classification

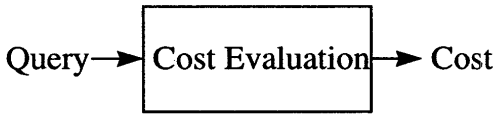


Fig.2 b: A Second Phase of Query Cost Evaluation

4.2 Modeling Operators

In a different approach of query processing cost evaluation by neural networks, basic operations in QEPs may be measured, modeled, and evaluated. The constraint in applying this approach is that the evaluator must have access to the QEPs generated by the DBMS. In this case, a query execution tree is generated by the DBMS without a real execution of the query; this query tree is then passed to the evaluator and the cost of each node in the tree is evaluated by a separated neural network according to the node type. The whole cost of the query is the sum of the query tree node costs.

A set of neural network in this approach of query processing cost estimation must be generated and trained. The number of networks in this set is finite since a network captures the cost of a query node type, which is an element in a finite set of physical operations. Statistics on a node in the query tree, and eventually the result size of a preceding node, are the main inputs to the neural network modeling its execution cost. An illustration of the mapping of a query tree to a tree of neural networks is given in Fig.3.

This approach of query processing cost evaluation by neural networks is more complicated than the first one, but its applicability and integration in an optimizer or a performance evaluation tool seem to be more realistic.

5 Neural Network Issues

During our work on performance evaluation by neural networks, we faced several problems and choices to take. The kind and structure of the networks that mostly adapt to our needs had to be identified and constructed. Influencing and irrelevant environment parameters for a platform, a DBMS, and an application had also to be identified and separated, and measurements had to be done to obtain a significant amount of data for training and validating the networks.

A neural network cannot be used in evaluating the cost of operations in a first stage of an evaluation process. A simple solution to this problem is to proceed in two steps: with the installation of a database application, a performance evaluation tool may use in a first step a simple analytical model to predict the cost of some pre-planned queries, and the execution times of the algebraic operations of these queries are measured; in a second step, a neural network is trained with these measurements, and the performance tool may switch to the neural network estimations.

An applied rule for testing a neural network generalization is to train the network on a part of the I/O data and to test its performance on a second part. We applied this rule in all the experiments we report in this paper: we used about 20% of the measurements as training data, and we tested the network on the whole data.

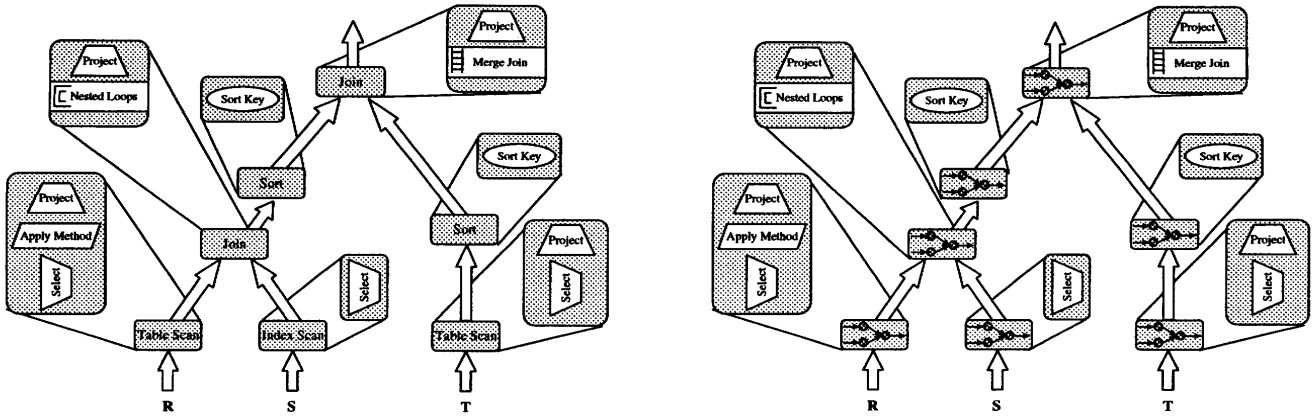
An additional observation we noticed on the input data is that it is more convenient for a neural network to take inputs that have small values; large values of some input parameters may be "cooked" in order to bring them to more convenient ones. An example of such parameters is the size of a relation. In this case, it is better to inject in the neural network the logarithmic value of the size (e.g., $\log_{10}(\text{size})$). In addition, and since neural networks do not accept symbolic input data but only numerical data, all symbolic inputs must be mapped into numerical values.

5.1 Network Structure

The theorem of HSW (stated above) guarantees the existence of (at least) a neural network that approximates a specific function. However, it did not give a method to generate the network; a trail-and-error process for the network structure must be undertaken. During the learning phase, and in the cases where a network reaches a stable state that does not satisfy the user (i.e., stuck in a local minimum), the network is then reinitialized with other random values and the learning phase is repeated. If after several trials, the network did not converge to an acceptable rate, a different structure of the network must be regenerated and the process of learning is again reinitialized.

The most fitting neural networks for function approximation are Feed-forward Neural Networks (F-NN). The possible learning algorithms for these networks are the Back-

学術情報センター電子図書館サービス　　© 学術情報センター



(a): A QEP with its Statistics　　(b): The Equivalent Network of Neural Networks

Fig.3 The Mapping Between a QEP and a Network of Neural Networks

Propagation and the Radial Basis Function (RBF) [20].

We have used the Back-Propagation algorithm in all our experiments in this paper. The reasons behind this choice are that the Back-Propagation is the most known and used algorithm for learning in neural networks, and because we have developed a certain knowledge in using this algorithm.

A question that we also asked to ourselves was when shall we stop the training phase? It is in the basic foundations of the neural network approach to learn from examples. The learning phase makes as much cycles on these learning examples as the user is unsatisfied with the error rate showed by the network. When the user decides that the error rate is sufficiently small, he stops the learning phase and tests the network with other input data. Hence, we decided to stop the learning phase each time the network converged to a Mean Squared Error (MSE) that went bellow a fixed threshold.

5.2 Input Parameters

The choice of the input statistical parameters to a neural network must characterize a database application and at the same time be relevant in the network. In addition, there is a part of these parameters that is dependent on whether we adopt the general (modeling queries) or the building block (modeling operators) approach and on the data model and its optimization strategy. However, some parameters are the same for both choices.

We state here some of the common inputs that, we believe, are the most representative for most queries. These may be: the value of a predicate constant, the size of the input, the size of the output (selectivity), indexes for each relation, clustered indexes or not, average size of a tuple, number of attributes in a tuple, minimum and maximum values for each attribute, kind of the predicate(s) (=, <, ≤, >, ≥), symbolic or numerical predicate, CPU power, disk access time, network bandwidth.

5.3 Data Sets for Training and Validation

While in traditional methods of performance evaluation assumptions are usually made (e.g. uniformity of data distribution, attribute independence) [8], in our approach of neuronal performance evaluation the system makes its evaluations independently from any assumption on the data, on a query complexity or on the environment. An influencing parameter in an analytical formula may be irrelevant for a neural network. Hence, careful attention must be given to identify important parameters for neural networks.

We present here the most influencing parameters and the different experiments we made on our working DBMS to extract measurements for constructing and comparing the neural network approach in cost evaluation against analytical models. We used several SQL queries on two benchmarks: the AS^{3}AP [31] and the one defined in [9] (we have implemented this benchmark and used it in several experiments of this paper). Several database sizes were used to identify influencing parameters. We then calibrated a cost model and constructed several neural networks capturing the execution costs of most interesting database execution operations. The different database operations that we tried to model were: table scan, index scan, sort, nested-loops join, and merge-join. We could not apply the calibration and training process on the hash-join algorithm since this last was not vailable in both Oracle7 versions that were available to us

学術情報センター電子図書館サービス ©学術情報センター

(version 7.1.0.16 and version 7.2.3). We also exclude here hardware dependent parameters such as I/O access time and CPU power since we only made measurement on one single platform; however, these are not difficult to add to the measurement process.

SQL queries were executed with different shapes, predicates, predicate constants, and database sizes. For example, the SQL query that must be executed to extract the cost of a table scan is:

Select Count(*) **From** *R*;

The "Count" is used to avoid the cost of result transfer between Oracle7 working space and the client working space and it was used for the same reason in all the other queries. Some of the inputs that we used in different experiments during our work included:

Table Scan:
- Scan size (number of rows)
- Row size (in bytes)

Index Scan:
- Number of selected rows
- Index number of entries (number of rows)
- Index page size (in bytes)
- Clustering factor (e.g., 0 for non clustered indexes and 1 for clustered ones)

Sort:
- Number of sorted rows
- Memory size

Nested-Loops Join:
- Number of participating rows from R (i.e., after predicate application)
- Selectivity of the selection predicate
- Selectivity of the join

Merge Join:
- Number of rows in R
- Number of rows in S
- Result size

6 Experimental Results

In this section we present some results we obtained from experiments we made on Oracle7 Release 7.2.3 on a Sun-Sparc20 machine. Selection queries from the AS^{3}AP [31] and [9] benchmarks were executed several times on several sizes of the databases and measurements were collected. The tests were made in two ways: measurements and estimations were made on global relational queries, and then measurements and estimations were made on basic algebraic operations. For all the experiments reported in this section, 20 measurement points were taken for each experiment where 20% were used to train the neural network and the whole 20 points were used to test it. We report the performance of some neural networks in query processing cost estimation for one global joining query and several other basic algebraic operations.

We have also calibrated general analytical cost formulas for some basic operations to compare their output performance against that of neural networks. The accuracy of the results of each method is given for some experiments with both average absolute and average relative errors; this is because relative errors become meaningless with small values while absolute errors tend to be dominated by errors on the largest values. The absolute error is computed as $|O - C|$, where O is the estimation output and C is the real cost, and the relative error becomes $|O - C| / C$.

Finally, the tests on neural networks were done with a freeware generator developed at the University of Stuttgart, the Stuttgart Neural Network System (SNNS) [29]. This comprehensive tool offers a graphical interface greatly reducing the task of specifying, training, and testing a appropriate structure for most of the operations.

6.1 Global Approach

In the case where cost predictions are applied to queries, a classification of query type and/or size must precede the estimation process. We present here an example of cost estimation for a global query.

In this test we tried to estimate the execution cost of a query joining two relations. We took two relations of 100,000 tuples each and joined them with the sort-merge-join algorithm. A predicate was applied on one of the two relations in order to change the result size of this relation and hence the merge size. The query has the following shape:

Select Count(*)
From *R1, R2*
Where *R1.A1 = R2.A2*
and *R1.A1 = cst*

The optimizer choose the following QEP for this query:

Select (Count)	1
Merge-join	Result-size
Sort(join)	Result-size
Table-Access-Pred(R1)	100000
Sort(join)	100000
Table-Access-Full(R2)	100000

学術情報センター電子図書館サービス　　© 学術情報センター

Table 2　A Comparison of Error rates on a Sort-Merge Join Query

	Avg. Absolute Error	Avg. Relative Error
NN Estimates	0.21	0.99%
Analytical Estimates	0.34	1.5%

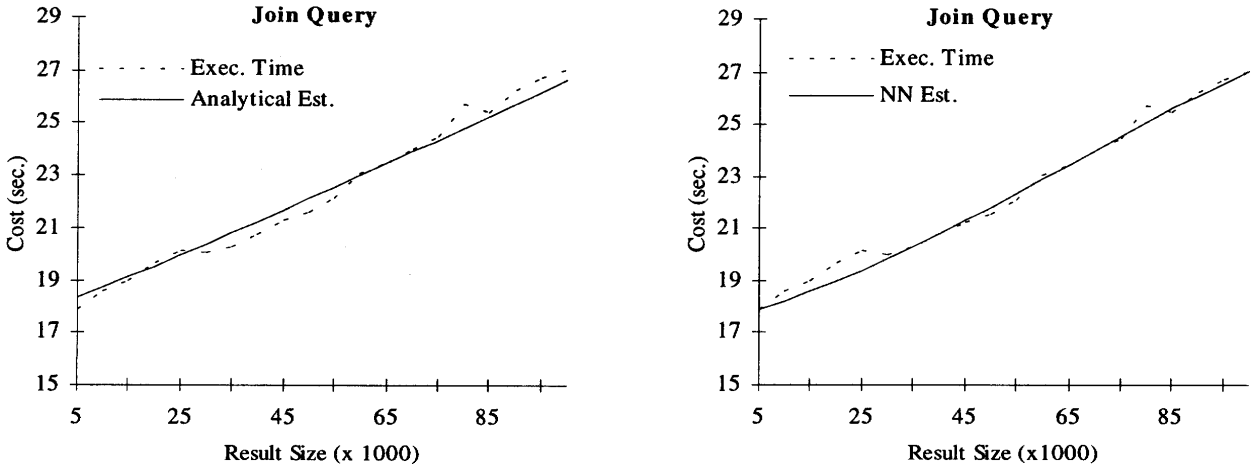


Fig.4　Performance of a Neural Network against an Analytical Formula for the Sort-Merge Query

The first relation is full scanned and the 100,000 tuples are sorted, while the second relation is full scanned and the predicate is applied and then the remaining tuples are sorted. Since we choose a uniform distribution for both attributes in the two relations, the result size from merging the two relations is the same as the size of the second relation after applying the predicate. The analytical cost formula that we calibrated for this query was computed as follows:

$$\begin{aligned}\text{Cost(Sort-Merge-Join)} = {} & (2 \times 100000 \times \textit{Access_Tuple}) \\ & + (100000 \times \log_2(100000) \\ & \quad / \textit{machine_dependent_constant}) \\ & + (\text{Result_Size} \times \log_2(\text{Result_Size}) \\ & \quad / \textit{machine_dependent_constant}) \\ & + \text{Result_Size} \times \textit{Merge_Join}.\end{aligned}$$

The only input parameter to the network was the predicate constant value -since all other parameters were unchanged- and the output was the elapsed cost time. At the same time we "cooked" the I/O measurements in order to bring the input to the range of [0..1] -i.e., we divided all the values of the constant by its maximum- and the output to the range of [3.0..11.0] -i.e., we divided the cost by 10. The most appropriate structure of the neural network was 1-3-2-1; it converged after 2,000 cycles in about 5 seconds.

Table 2 compares the absolute and the relative error rates for both the neural network and the analytical formula. In Figure 4 we illustrate the curves for the whole query real execution time, the neural network estimates, and the analytical formula estimates. From both the comparison table and the figures, it is chown that the neural network estimates are at least as efficient as those of the analytical formula.

6.2　Building Blocks

With this approach of cost estimation where each physical operation is isolated and estimated alone, we tested the efficiency of neural networks with different input parameters for several operations. We report here simple experiments for the scan, the sort, and the merge-join operations.

Scan

The first test consisted of scanning a relation with several sizes. The size of the relation was varied between 5,000 and 100,000 tuples, and the same query was repeated on all sizes. A neural network was generated (with a structure of 1-3-2-1) and trained on a subset (20%) of the measurement data, and as expected- converged after few cycles (1000 cycles that

学術情報センター電子図書館サービス ©学術情報センター

Table 3 A Comparison of Error rates on an Elapsed Cost of a Scan

	Avg. Absolute Error	Avg. Relative Error
NN Estimates	0.26	11.7%
Analytical Estimates	0.77	113.4%

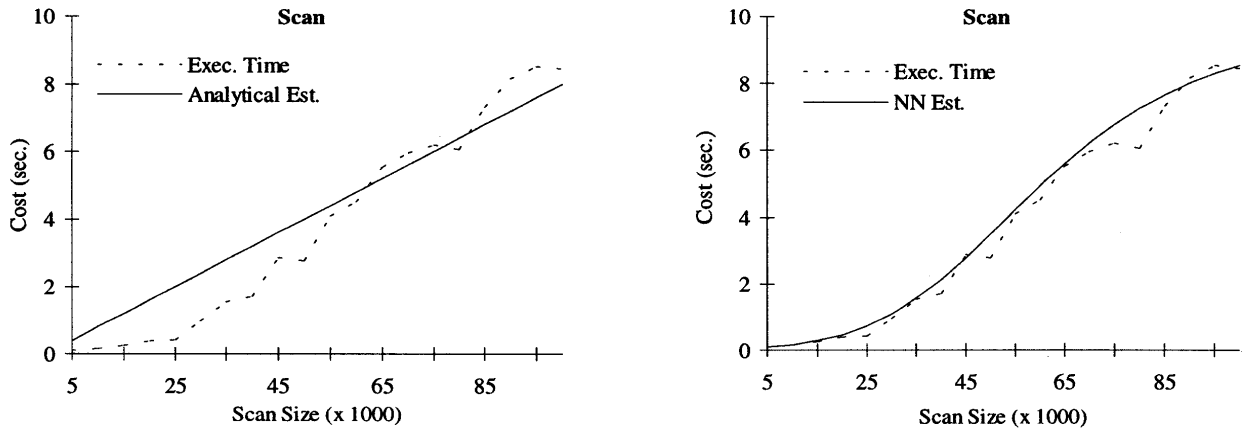


Fig.5 Performance of a Neural Network against an Analytical Formula for the Scan Operation

took less than 5 seconds). The measured cost of this test was the elapse time of the scan, i.e., the CPU and the I/O cost. Table 3 compares the absolute and the relative error rates of the neural network against those of the following calibrated analytical formula:

$$\text{Cost(scan)} = \text{Scanned_tuples} \times \textit{Access_tuple}$$

Figure 5 illustrates the estimates of both the neural network and the analytical formula against the real execution time of the scan. The neural network outperformed the analytical formula by a considerable margin because of the cache effect on the small scan sizes and where the analytical formula could not capture it.

Sort

A second test consisted of sorting a relation with several sizes. The size of the relation was also varied between 5,000and 100,000 tuples, and the same query was repeated on all the sizes. A neural network was generated (with a structure of 1-3-2-1) and trained with a subset of the measurement data. The measured cost of this test was the elapse time of the sort, i.e., the CPU and the I/O cost. Table 4 and Figure 6 compare the performance of the output of the neural network against that of the following calibrated cost formula:

$$\text{Cost(sort)} = \text{Sort_Size} \times \log_2(\text{Sort_Size}) \;/\; (\textit{machine_dependent_constant})$$

Here, the neural network outperformed the analytical formula by a small margin.

Merge-Join

The third test consisted of merging two relations with several result sizes. The size of the result was varied between 5,000 and 100,000 tuples. A neural network was generated (with a structure of 1-3-2-1) and trained with a subset of the measurement data. The measured cost of this test was the elapse time of the merge. Table 5 and Figure 7 compare the performance of the neural network output against that of the following calibrated cost formula:

$$\text{Cost(merge-join)} = \text{Result_Size} \times \textit{Merge_Join}$$

One may argue that the error rates of the analytical formulas for the sort and the merge-join are small enough and do not need to be reduced any more. We must, however, remember that these operations are usually operations in a larger QEP: several such error rates in the QEP increase the error rate of the whole query cost estimates.

学術情報センター電子図書館サービス　　© 学術情報センター

Table 4　A Comparison of Error rates on a Sort

	Avg. Absolute Error	Avg. Relative Error
NN Estimates	0.57	15.8%
Analytical Estimates	0.56	16.11%

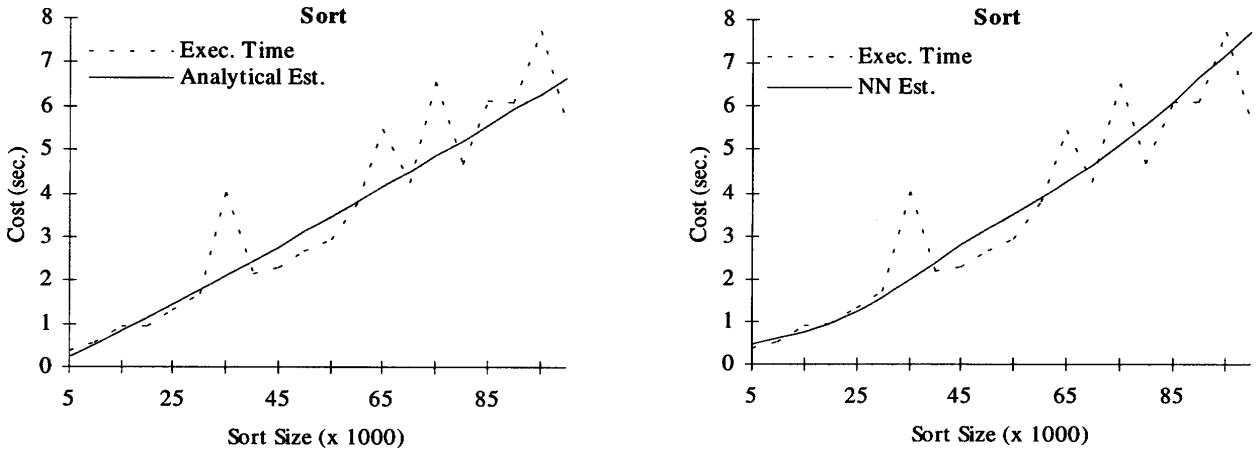


Fig.6　Performance of a Neural Network against an Analytical Formula for the Sort Operation

The neural network approach for estimating the cost of simple query processing operations proved its effectiveness to us at this point of our investigations, we extended our work to more complex operations.

7　Estimating Method Costs

In this second set of experiments, we tried to use neural networks for estimating the costs of user defined methods. As we have already said, the feasibility and usefulness of this approach is based on the neural network capabilities of approximating measurable functions. We defined and tested several methods with different execution costs and influencing input parameters. We report here on two methods with one and two inputs. Increasing the number of input variables is straightforward.

7.1　A Simple Method

In a first experiment, we implemented a PL/SQL function with one input variable and where this function makes merely CPU processing depending on the input variable length. The input variable has an exponential influencing cost on the method. The method was executed with 20 values for this input variable and the network was trained with 20% of the sample points and tested with all points. The experimented network had a structure of 1-3-2-1. About 2000 cycles over the learning set were sufficient for the network to converge, which took about 5 seconds. The relative and absolute errors (Table 6) of the estimates were higher than expected because of the high execution cost in seconds. We can see from Figure 8 the shapes of both the real execution cost and the neural network estimated cost of this method.

This method was relatively simple and the neural network efficiently captured its execution cost, we tried to push the complexity of the method processing to a higher degree.

7.2　A Complex Method

In this second experiment a method with two input variables was used. The first input variable had an exponential influencing cost on the method while the second input variable had a logarithmic influencing cost. The method was executed with 10 values for each variable, which makes 100 executions. As above, we trained a network with 20% of the sample points and tested it with all points. The experimented network had a structure of 2-4-3-1. About 3000 cycles over the learning set was sufficient for the network to converge, which took about 10 seconds. The performance of the net

学術情報センター電子図書館サービス ©学術情報センター

Table 5 A Comparison of Error rates on the Merge-Join

	Avg. Absolute Error	Avg. Relative Error
NN Estimates	0.33	7.44%
Analytical Estimates	0.38	8.44%

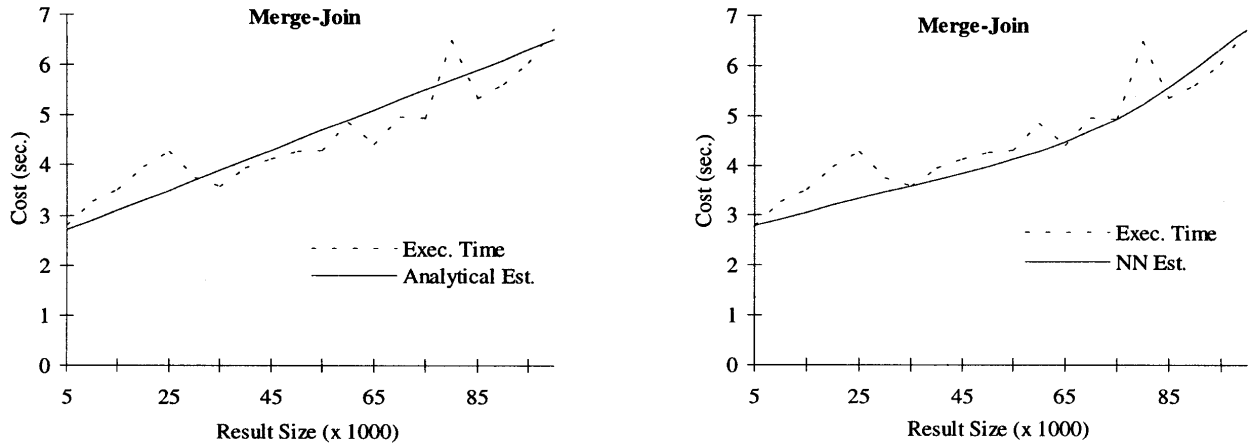


Fig.7 Performance of a Neural Network against an Analytical Formula for the Merge-Join Operation

Table 6 Performance of a NN in Estimating the Cost of a Simple User Defined Method

	Avg. Absolute Error	Avg. Relative Error
NN Estimates	28.8	33.8%

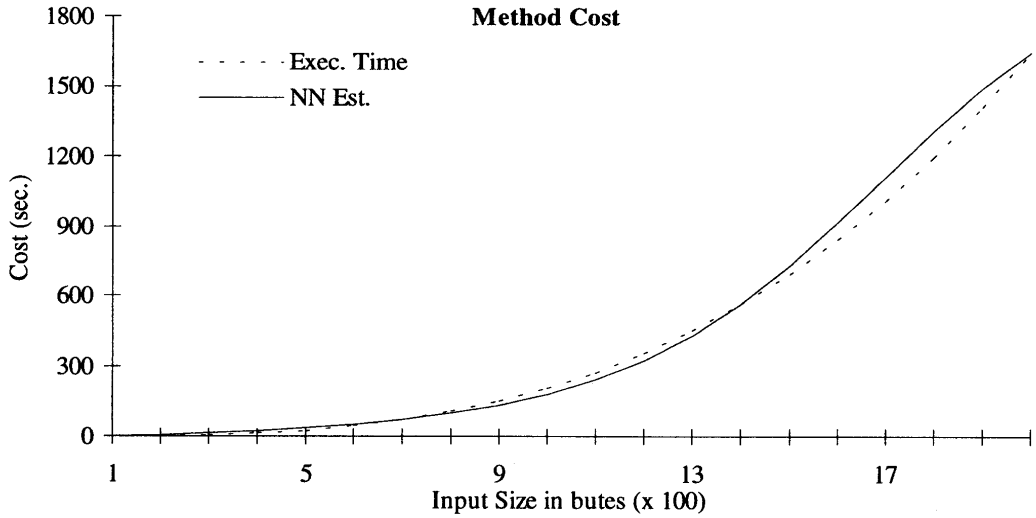


Fig.8 Comparison between the Real Execution Cost and the NN Estimates for a Simple User Defined Method

学術情報センター電子図書館サービス　　© 学術情報センター

Table 7　Performance of a Neural Network in Estimating the Cost of a Complex User Defined Method

	Avg. Absolute Error	Avg. Relative Error
NN Estimates	0.58	0.39%

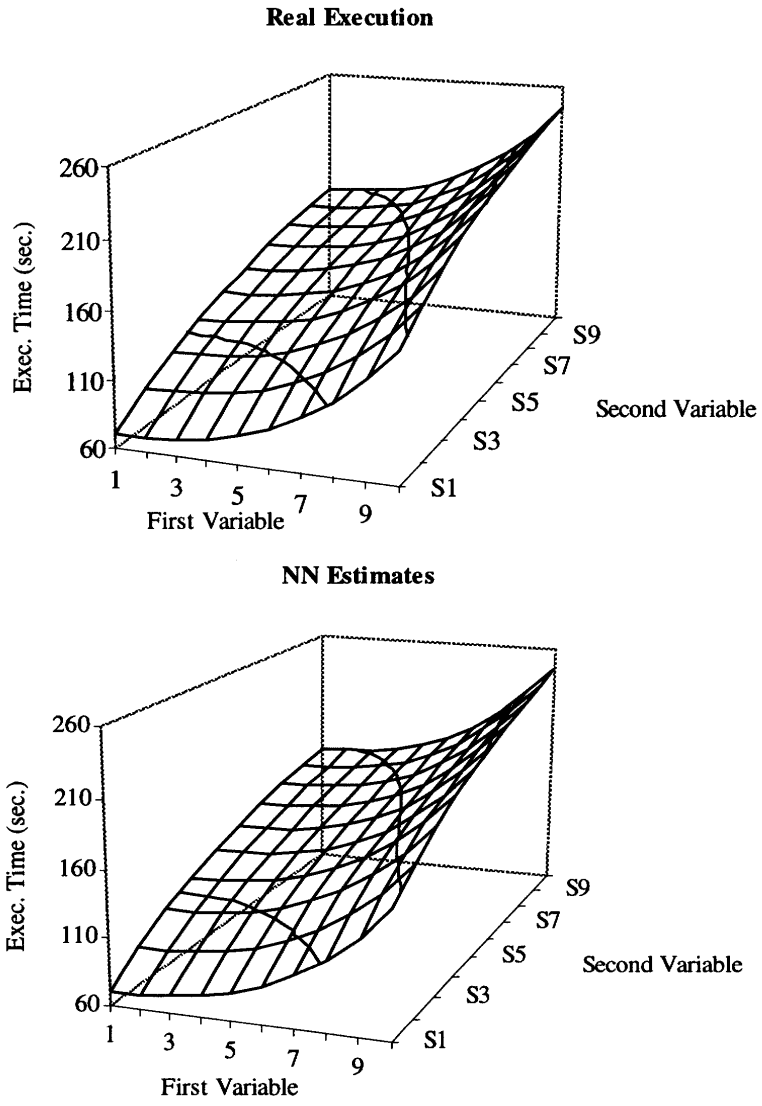


Fig.9　Comparison between the Real Execution Cost and the NN Estimated Cost of a Complex Method

work went beyond our hopes and the average relative error between the estimated cost and the real execution cost for the 100 execution points was bellow 1% (Table7.) We can see from Figure 9 the shapes of both the real execution cost and the neural network estimated cost of this method.

8 Related Work

The most closed works to ours is those of [9] and [33] for cost estimations in heterogeneous DBMSs. In both papers, the authors used a calibration process for some predefined analytical model. [1] uses almost the same approach as [9] for analytical model calibration. In all cases, and to the best of our knowledge, no one has until now tried to estimate method costs with a curve fitting like approach. Existing DBMSs that accept user defined methods require the user to give the cost of its method.

[9] and [1] established a calibration method to extract for a predefined analytical cost model its parameter values that depend on the platform and the DBMS. A benchmark is executed on a system to measure its performance, and an equa-

学術情報センター電子図書館サービス ◎学術情報センター

tion system is then constructed and resolved. This process produces a calibrated cost model for a specific platform. However, this process always needs an expert in database performance evaluation to establish the model and the availability of the calibrated environment for a long time (hours and maybe days). We have used this technique in generating and calibrating a cost model [3] for the GeoSabrina geographical database system and in SMART2 [4].

[33] extended a previous work [6] for selectivity estimation by curve fitting. In this paper ([33]), the authors approximate the execution cost of a query in a heterogeneous environment by executing some predefined queries on the target DBMS with different result sizes and then approximating the queries costs by polynomials using a regression method well known in linear algebra. Contrary to [9] where basic operations were modeled, in [33] global queries are modeled. We have in our approach of cost estimation by neural network investigated cost modeling of global queries and an example of a query joining two relations was report in this paper.

9 Conclusion

We have proposed in this paper a new approach for query processing cost estimation. The approach is based on the capabilities of neural networks to approximate any measurable function in a curve fitting like manner. A set of neural networks is trained with measurements from a working database environment in a first phase, and it produces cost estimates for evaluated queries in a second phase.

Limitations of analytical cost formulas and new requirements of emerging database technology motivated our work in searching for a new method to predict the performance of query processing algorithms. The neural network approach to overcome these limitations and meet the new requirements was decomposed, analyzed, and successfully experimented.

A discussion on global or building block ways of evaluating the cost of a query by neural networks was exposed and analyzed. We have the feeling that the building block approach where a tree of neural networks matches the tree of nodes in a QEP may outperform the global one in its effectiveness in a real working environment. We have, however, reported experimental results for both approaches -i.e., the building block and for global queries- and comparisons against calibrated analytical cost formulas proved the effectiveness of both of them. In addition, the effectiveness of neural network estimations of user defined method costs was also reported with two examples.

More research work on feeding additional input parameters of a database environment into larger neural networks must be done. We did not until now integrate any platform dependent parameter in our tests; we intend to investigate this influencing aspect on performance by varying the CPU power, the disk throughput, and the network bandwidth in a distributed environment.

References

[1] Andres, F.; Kwakkel, F.; Kersten, M., "Calibration of a DBMS Cost Model with the Software Testpilot", *Lecture Notes in Computer Science, Springer*, 1006; Proc. 6th Int. Conf. CISMOD, Bombay, Nov. 1995.

[2] Batory, D., "Extensible Cost Models and Query Optimization in GENESIS", *IEEE Database Engineering*, Vol. 9, No. 4, Dec. 1986.

[3] Boulos, J.; Andres, F.; Viemont, Y., "Generating Cost Models for Geographical OODBMSs", BULL S.A., Les Clayes-sous-Bois, Report No. BULL/RAD /DB/93017, Sept. 1993.

[4] Boulos, J.; Boudigue, D., "An Applications of SMART2: A Tool for Performance Evaluation of Relational Database Systems", *Proc. of Performance Tool'95, 8th Int. Conf. on Modeling Techniques and Tools for Computer Performance Evaluation*, Heidelberg, Germany, Sept. 1995. Also printed in "Quantitative Evaluation of Computing and Communication Systems", *Lecture Notes in Computer Science*, Springer, No. 977, Sept. 1995.

[5] Cattel, G., "Object Databases: The ODMG Standard", Ed., Morgan & Kaufman, 1993.

[6] Chen, C.; Roussopoulos, N., "Adaptive Selectivity Estimation Using Query Feedback", *Proc. of the 1994 ACM SIGMOD Int. Conf. on Management of Data*, Minneapolis, Minnesota, May 1994.

[7] Chen, D.; Jain, R., "A Robust Back Propagation Learning Algorithm for Function Approximation", *IEEE Transactions on Neural Networks*, Vol. 5, No. 3, May 1994.

[8] Christodoulakis, S., "Implications of Certain As-

学術情報センター電子図書館サービス ©学術情報センター

sumptions in Database Performance Evaluation", *ACM Trans. on Database Systems*, Vol. 9, No. 2, June 1984.

[9] Du, W.; Krishnamurthy, R.; Shan M.C., "Query Optimization in Heterogeneous DBMS", *Proc. of the 18th Very Large Data Bases Conference*, Vancouver, British Columbia, Canada, 1992.

[10] Freytag, J. C., "A Rule-Based View of Query Optimization", *Proc. of the 1987 ACM-SIGMOD Annual Conf.*, San Francisco, May 1987.

[11] Gardarin, G.; Gruser, J.; Tang, Z., "A Cost Model for Clustered Object-Oriented Databases", *Proc. of the 21st Int. Conf. on Very Large Data Bases*, Zurich, Switzerland, Sept. 1995.

[12] Graefe, G., "The EXODUS Optimizer Generator", *Proc. of the 1987 ACM-SIGMOD Annual Conf.*, San Francisco, May 1987.

[13] Graefe, G., "Query Evaluation Techniques for Large Databases", *ACM Computing Surveys*, Vol. 25, No. 2, June 1993.

[14] Graefe, G.; McKenna, W., "The Volcano Optimizer Generator: Extensibility and Efficient Search", *Proc. of the 9th Int. Conf. on Data Engineering*, Vienna, Austria, April 1993.

[15] Hellerstein, J., "Practical Predicate Placement", *Proc. of the 1994 ACM SIGMOD, Int. Conf. on Management of Data*, Minneapolis, Minnesota, May 1994.

[16] Hornik, K.; Stinchcombe, M.; White, H., "Multilayer Feedforward Networks are Universal Approximators", *Neural Networks*, Vol. 2, pp. 359-366, 1989.

[17] Hong, W.; Stonebraker, M., "Optimization of Parallel Query Execution Plans in XPRS", *Proc. of the 1st Int. Conf. on Parallel and Distributed Information Systems*, Miami Beach, Florida, Dec. 1991.

[18] Hellerstein, J.; Stonebraker, M., "Predicate Migration: Optimizing Queries with Expensive Predicates", *Proc. of the 1993 ACM SIGMOD, Int. Conf. on Management of Data*, Washington, DC., May 1993.

[19] Jarke, M.; Koch, J., "Query Optimization in Database Systems", *Computing Serveys*, Vol. 16, No. 2, June 1984.

[20] Jain, A.; Mao, J.; Mohiuddin, K., "Artificial Neural Networks: A tutorial", Computer, March 1996.

[21] Lanzelotte, R.; Valduriez, P.; Zaït, M., "Optimization of Object-Oriented Recursive Queries using Cost-Controlled Strategies", *Proc. of the 1992 ACM-SIGMOD, Int. Conference on Management of Data*, San Diego, California, June 1992.

[22] Leshno, M.; Lin, V.; Pinkus, A.; Schoken, S., "Multilayer Feedforward Networks With a Nonpolynomial Activation Function Can Approximate Any Function", *Neural Networks*, Vol. 6, pp. 861-867, 1993.

[23] Looney, C., "Advances in Feedforward Neural Networks: Demystifying Knowledge Acquiring Black Boxes", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 8, No. 2, April 1996.

[24] Nabhan, T.; Zomaya, A., "Toward Generating Neural Network Structures for Function Approximation", *Neural Networks*, Vol. 7, No. 1, pp. 89-99, 1994.

[25] Pirahesh, H.; Hellerstein, J.; Hasan, W., "Extensible/ Rule Based Query Rewrite Optimization in Starburst", *Proc. of the 1992 ACM-SIGMOD, Int. Conference on Management of Data*, San Diego, California, June 1992.

[26] Rumelhart, D.; Widrow, B.; Lehr, M., "The Basic Ideas in Neural Networks", *Communications of the ACM*, Vol. 37, No. 3, March 1994.

[27] Selinger, P.; Astrahan, M.; Chamberlin, D.; Lorie, R.; Price, T., "Access Path Selection in a Relational Database Management System", *Proc. ACM-SIGMOD Conf.*, Boston, May 1979.

[28] Sevcik, K., "Data Base System Performance Prediction Using An Analytical Model", *7th Int. Conference on Very Large Data Bases*, Cannes, France, Sept. 1981.

[29] Stuttgart Neural Network Simulator, User Manual, V.3.2, Report No. 3/94.

[30] Stonebraker, M.; Rowe, L.; Hirohama, M., "The Implementation of Postgres", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 2, No. 1, March 1990.

[31] Turbyfill, C.; Orji, C.; Bitton, D., "AS3AP - An ANSI Sequel Standard Scalable and Portable benchmark for Relational Database Systems", in the Preformance Handbook for Database and Transaction Processing Systems, San Mateo, CA: Morgan Kaufmann, 1991.

[32] Widrow, B.; Rumelhart, D.; Lehr, M., "Neural Net-

学術情報センター電子図書館サービス
© 学術情報センター

works: Applications in Industry, Business and Science", *Communications of the ACM*, Vol. 37, No. 3, March 1994.

[33] Yao, Z.; Chen, C.; Roussopoulos, N., "Adaptive Cost Estimation for Client-Server based Heterogeneous Database Systems", University of Meryland Report, CS-TR-3648, May 1996.

学術情報センター電子図書館サービス　　©学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

インターネットにおける実時間通信のQOS保証

QOS guarantee for real-time applications in the Internet

東京大学工学系研究科　森野　博章
Hiroaki MORINO
School of Engineering, University of Tokyo
学術情報センター　相澤　彰子
Akiko AIZAWA
National Center for Science Information Systems

要旨

インターネットの需要はデータ通信から動画・音声通信へと広がっているが、TCP、UDPといった既存のトランスポートプロトコルは動画・音声に必要な実時間伝送の品質を十分に保証することができず、新たなプロトコルが求められている。

本稿ではインターネットで実時間伝送に対する品質保証を可能にするプロトコルの研究を紹介する。プロトコルの中心的機能はネットワークの負荷に応じた動的レート制御とメディアスケーリングであり、提案された各プロトコルについてレート制御のアルゴリズムと、画像品質への影響を論じる。

最後に各方式の性能比較を行うとともに、ネットワークインフラストラクチャの分野で品質保証を実現しようとする研究例を報告する。

ABSTRACT

Application in the Internet is currently being expanded form simple data transmission to video and audio communications. As the existing protocols such as TCP or UDP are not intended to guarantee QOS(Quality Of Service) for real-time applications like video or audio transmission, new protocols are required. In this article, we present some protocols which can guarantee the QOS for the real-time applications in the Internet. Main functions of these protocols are dynamic rate control corresponding to network load and media scaling. We discuss the rate control algorithm for each protocol and its effectiveness in improving picture quality. Finally, we show a brief result about performance comparison and make a report on the research which will guarantee the QOS in the area of future network infrastructure.

[キーワード]　インターネット、動画、音声、QOS、レート制御
[Keywords]　Internet, video, audio, QOS, rate control

1　はじめに

動画・音声は、長い間、テレビ放送、レンタルビデオといった、一方向性のメディアとして用いられることが主であったが、インターネットの普及と計算機の能力向上、記憶装置の低廉化によって、個人が自らを表現する手段として動画・音声を取り扱うことが可能となった。WWW(World Wide Web)における自己紹介、あるいは研究報告のホームページに用いられている例から分かるように、テキストや静止画に比較してより正確に情報を伝えることができるのは明らかであり、ネットワークの用途として今後ますます増大していくものと思われる。

動画・音声といった実時間情報を伝送するにあたっては、ネットワークの側でデータの遅延時間、パケット損失率等のQOS(Quality Of Service:サービス品質)を保証する必要があるが、現在のインタネットを構成するネットワークの主な通信方式のEthernet、FDDI、そしてその上位の通信プロトコルであるTCP、UDP/IPは電子メール、ファイル転送といった従来型のデータ通信に適しており、

学術情報センター電子図書館サービス　　©学術情報センター

実時間通信を行う場合、負荷が高い際の品質に問題が生じることが知られている。このような背景から、データ通信と実時間通信に対して合わせてQOSを保証するあらたなプロトコル体系が求められている。

通信方式に関しては、QOS保証を実現する方式として以前からATM(Asynchronous Transfer Mode)がITU-T(International Telecommunication Union-Telecommunication)、ATM Forumを中心に活発に研究され、基幹ネットワークおよびLANの一部に実用化が始まっている。ただし、ATM交換機が一般のネットワークに広く浸透するまでには相当の時間がかかるうえ、従来との互換性を重視する用途も多く、現在の通信方式も私設網を中心に残り続けるであろう。したがって、異種網を接続するインターネットのプロトコルにおいてQOSを保証する機能が重要となってくる。本稿では、特に動画像通信に焦点を当て、インターネット上でQOS保証を実現する通信プロトコルの研究例を報告した後、ネットワークインフラストラクチャの将来像について概観する。

2 研究の背景

2.1 QOSとは

動画の伝送方式は、ビデオ・オン・デマンドのようにあらかじめ送信側に格納されたデータを受信側の要求に応じて送信する蓄積型と、テレビ会議に代表される、送信側で実時間で情報を取得して送信を行なうライブ型の大きく2つに分けられる。現在のネットワークの伝送速度が低いため、ライブ型はもちろん、蓄積型であっても、画像フレームを実時間で伝送して受信側でそれを逐次再生する方法が多く用いられる。

ネットワークにおいて実時間伝送を実現するために必要なQOSは、主に次の4つである。[1]

- スループット
- 平均転送遅延時間
- 遅延分散
- パケットの誤り率

また、ライブ型ではマルチキャスト(同報通信)が多く用いられることから、マルチキャストの環境下での品質保証も必要となる。

2.2 インターネットにおけるQOS保証の問題点

現在のインターネットのプロトコルは表1のような4階層を持っている。[2]

Ethernet、FDDI等のさまざまなネットワーク間を相互接続し、端末間でのパケット転送を行う役目を持つのはインターネット層のIP(Internet Protocol)である。IPはコネクションレス型のプロトコルであり、制御が簡単である反面、ネットワークの負荷の制御ができないために、経路上でパケットの損失が起こることは避けられない。そこでその上位のトランスポート層のTCP(Transmissi-on Control Protocol)では、失われたパケットの再送制御、端末間でのフロー制御を行っている。さらにパケットの損失から輻輳(過負荷の状態)を検出して、送信レートを制御する輻輳制御の機能も持つ。

表1　TCP/IPプロトコルスタック

Process/Application layer	Telnet, FTP, e-mail
Transport layer	TCP, UDP
Internet Layer	IP
Network access layer	Ethernet, FDDI, ATM etc.

TCPはデータ通信に対しては十分な機能を提供しているが、パケットの転送遅延時間の保証は考えられていないために、実時間伝送では不都合が生じる。例えばTCPを用いて動画の送信を行うと、パケットの損失が生じた場合、再送制御によって、受信側では一時的に画像が停止してしまう。つまり再送制御は必要でない。

一方もう一つの代表的なトランスポート層プロトコルとしてUDP(User Datagram Protocol)がある。こちらはTCPと異なり、フロー制御、再送制御および輻輳制御を行わないため、パケットの転送遅延時間が予測でき、実時間伝送に適している。しかしながら負荷が高くなるとパケット損失率の保証が出来ないことが欠点である。

2.3 QOS保証方式の分類

前節の問題点を踏まえ、インタネット層とトランスポート層のプロトコルで実時間伝送のQOSを保証することを考えた場合、そのアプローチは、適応制御型と資源予約型の大きく2つに分けられる。

適応制御型では、データ転送中に、ルータまたは受信側の端末がパケット損失等からネットワークの負荷に関する情報を検出して送信側に定期的に通知し、送信端末は、その情報をもとに送信レートを変化させることにより、パケット損失を最小限に防ぐ。この方法は、TCPで行われている制御から再送制御を除いたものとほぼ同等である。さらに、データの転送遅延時間を一定にするために送信側で、負荷の変動に対応して

学術情報センター電子図書館サービス　　© 学術情報センター

送信すべきデータの情報量そのものを削減するメディアスケーリングの処理を行う。動画の場合には、毎秒の送信フレーム数を削減する方法と、フレーム数は変えずに画素数の削減等によって個々のフレームの情報量を削減する方法の主に2つが用いられている。適応制御型はコネクションレス型通信しか提供しないネットワークの上でも、上述した制御を送信側、受信側の各々の端末に実装するだけで品質保証が実現できる。

一方、資源予約型ではデータ転送の開始にあたりあらかじめ呼を設定する。この際、送信側がルータに対して必要な帯域を要求し、経路上の各ルータが伝送路の空き帯域幅を調べ、要求された帯域が利用可能である場合に呼が受け付けられる。一旦呼が設定されれば、送信側が申告した帯域にしたがって送信レートを調節することにより、通信の終了時までスループット、転送遅延時間、パケット損失率の各品質がほぼ完全に保証される。この方法では、ルータが各々の呼に対して伝送路帯域を割り当てる必要があるため、IPでは実現できないが、IPに代わり帯域確保の機能を持つインターネット層プロトコルとして、RSVP[3]、ST-II[4]等が提案されており、経路上のすべてのルータにこれらを実装することにより、制御が可能になる。

3 提案されているプロトコル

3.1 適応制御型

3.1.1 VDP

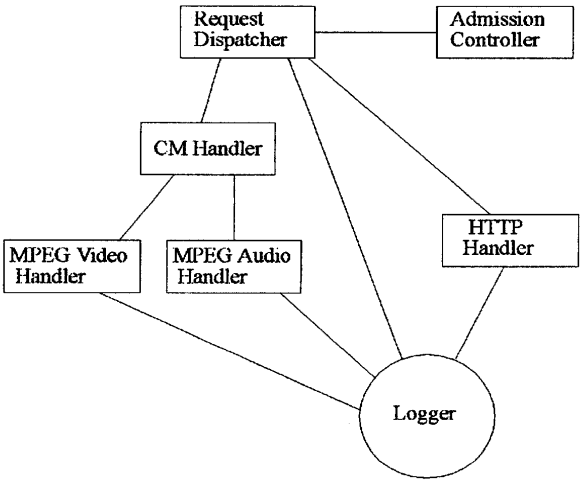


図1　vosaicのアーキテクチャ[5]

VDP(Video Datagram Protocol)[5]はイリノイ大学で開発されたWWWブラウザであるvosaicに組み込まれたプロトコルであり、蓄積型の動画像および音声を実時間で送信することを目的としている。vosaic全体のアーキテクチャは図1に示される。受信端末からサーバーへ接続要求が発せられると、送信すべきデータの種類によって、制御データの場合はHTTP Handler、動画・音声の場合はMPEG Video Handler、MPEG Audio Handlerにそれぞれ要求が引き渡されてデータの送信が開始される。VDPはこのうち動画・音声のデータ転送のプロトコルとして用いられている。

データ転送が開始されると、受信端末は、30フレームを受信するごとにネットワークの輻輳あるいは受信側の能力不足によって損失したフレーム数の割合を記録しておき、サーバーに通知する。サーバーはフレームの損失率が15%を越えた場合レートを下げ、5%を下回るとレートを上昇させる。レートの調節は、フレーム数を削減することによって行なっており、MPEG1画像フレームのI、P、Bの3つのピクチャのうち、重要でないBピクチャを削減することで画質の劣化を防いでいる。

輻輳によってフレームが失われた場合、受信端末が符号化モードを判別して、Iピクチャあるいは、Pピクチャの場合にのみフレームの再送を要求し、復号画像の品質を保証する。

表2　評価用画像の諸元

画像	レート (fps)	解像度	再生時間
model.mpg	9	160x120	14(sec)
startrek.mpg	5	208x156	128(sec)
puffer.mpg	5	320x240	35(sec)
smallogo	5	320x240	324(sec)

表3　端末名

端末名	役割
indy1.cs.uiuc.edu	受信端末(イリノイ州)
showtime.ncsa.uiuc.edu	送信端末(州内)

表4　実験結果

画像	フレーム欠落率	遅延分散
model.mpg	0	46.3ms
startrek.mpg	0	57.1ms
puffer.mpg	0	34.3ms
smallogo	0.2	50.0ms

表2に示された評価用画像を用いて、図3の端末間で伝送実験を行なった結果が表4に示されている。レー

学術情報センター電子図書館サービス　　© 学術情報センター

ト制御の効果によって、フレームの損失率はほぼ 0 に抑えられ、かつ遅延時間も人間の視覚に影響を与える閾値(100ms)を下回っている。

3.1.2 CMPlayer[6]

CMPlayer(Continuous Media Player) は米バークレー大学の Plateau Multimedia 研究グループの開発した分散ビデオ・オン・デマンドシステムに組み込まれた、動画・音声用の再生ソフトウェアである。

アーキテクチャは図 2 に示す通りであり、ファイルサーバーの CM Source プロセスと受信端末の CM Server プロセスの間で動画・音声データの通信が行なわれる。通信プロトコルとしては、受信側から送信側に送られる再生、停止などの制御データに対しては TCP が、動画・音声データに対しては UDP を用いており、さらに UDP の上には独自のレート制御が施されている。

UDP の上で行なわれるレート制御のアルゴリズムは次の通りである。まず、送信を開始する前に送信レートの最大値と最小値を決定しておく。データの転送が開始されると、受信側の CM Server はフレームの受信率を元にペナルティと呼ぶ値を計算する。ペナルティは、フレームが再生時刻までに端末に到着せず、フレームの欠落が生じた場合に 10 点、2 つの連続したフレームが欠落した場合にはさらに 10 点が加算される。したがって、たとえば 2 つの連続したフレームが欠落すると合計で 30 点のペナルティとなる。さらにネットワーク上でパケットの損失が生じるとやはり 10 点のペナルティが加算される。ペナルティは最高で 100 点まで累積され、300ms ごとに送信側に通知される。送信側は次の式で送信レートを計算する。

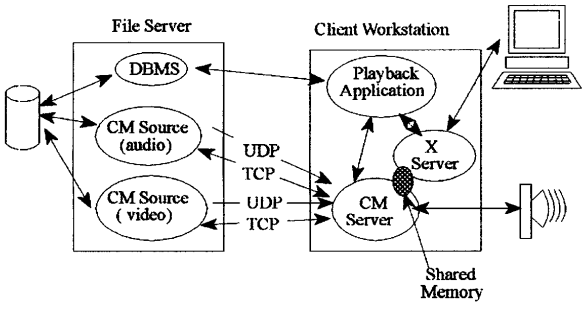


図 2　CM Player のアーキテクチャ[6]

$$currentRate = currentRate \times (1 - \frac{penalty}{100}) + minimumRate \times (\frac{penalty}{100})$$

すなわち、ペナルティが 0 ならばレートの変更は行なわれず、ペナルティが 100 ならば最小レートに設定される。一方、送信側は定期的に、送信レートを加算的に増加させる。レートの調整は VDP と同様に、送信フレーム数を削減することによって行なわれる。一般にネットワークでは、パケットの連続的な損失が輻輳の兆候であることが知られており、ここではフレームの連続的な欠落に対してペナルティを多く与え、レートの低減幅をより大きくすることによって輻輳の発生を防いでいる。

3.1.3 VCR[7]

VCR は、前節で述べた CMPlayer の改良版であり、アーキテクチャは図 3 に示す通りである。動画を送信する VS(Video Server)の送信プロトコルにおいては UDP の上に図 4 に示すフィードバックレート制御が施されている。

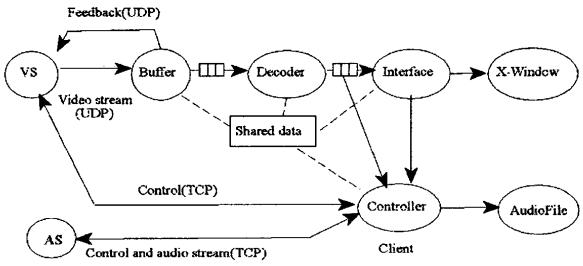


図 3　VCR のアーキテクチャ[7]

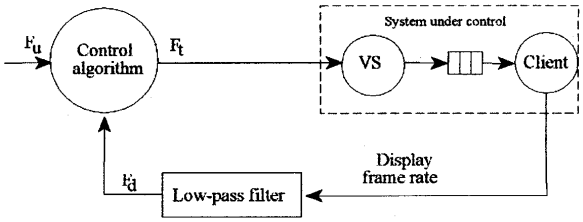


図 4　レート制御機構[7]

送信レートの目標値はユーザが指定することが可能でこれを F_u、ネットワークの負荷によって変動する送信レートの目標値を F_t、受信側での再生レートを F_d とすると、F_t は F_d の値により、次の式にしたがって変更される。

$F_d < F_t - T_h$ の場合　$F_t = F_t - \Delta$
$F_d < F_t - T_t$ かつ
$F_d < F_u$ の場合　$F_t = Min(F_t + \Delta, F_d)$
$(T_h, T_t, \Delta$はパラメータ$)$

Low-pass filter は、フィードバック入力値 Fd を平滑化

学術情報センター電子図書館サービス　　© 学術情報センター

する役目を果たしており、入力 input[i]と出力 lowpass[i] (i≧0)は、

$$lowpass[0] = input[0]$$

$$lowpass[i+1] = (1.0-R)*lowpass[i] + R*input[i] (1 \geq R \geq 0)$$

の関係にある。

VCR についても他の方式と同様に、インタネット上での性能評価が行われている。図5は図4に対応して、横軸に設定された F_u が、"sent"、"display" がそれぞれ Display frame rate と F_t の値を示している。この実験では回線の帯域幅の制限によって受信側の再生速度は15フレーム/秒が上限であるが、レート制御機構がないと F_t が F_u に追随するため、パケット損失率は最大で83%に達するのに対して、レート制御機構を有効にすることにより F_t がこの上限値に抑えられ、パケット損失率は30%まで減少する。

動画像通信の品質評価を行うに当たっては、パケット損失率よりも、再生の滑らかさを基準とする方が適切であることから、本方式では次のようにして再生の滑らかさを定義している。すなわち、送信すべき画像ストリームの各フレームを(f_0、f_1、..f_n)、受信端末で再生されたフレームを(f_{i0}、f_{i1}、..f_{in})とした時、フレームの本来の再生時刻 i と、実際の再生時刻 i_k の間には通常誤差があるが、この誤差 $e_i=i-i_k$ を用いて表示の滑らかさ S は次の式で表される。

$$S = \sqrt{\frac{\sum_{i=0}^{n} e_i^2}{n}}$$

S が小さい程再生が滑らかであることを示し、パケット損失および再生時刻の揺らぎがまったくない場合には S=0 となる。図6は図5の各の画像に対して S を算出したものであって、レート制御を行うことによりパケットの連続的損失が抑えられ、表示の滑らかさが向上することが示されている。

3.2 資源予約型

3.2.1 HeiTP

HeiTP(Heiderberg Transport Protocol)[8] はインターネット層プロトコルとして帯域予約が可能な ST-II を想定し、その上で動作するトランスポート層プロトコルである。

3.2.2 呼の設定

HeiTP では1対1呼の他にマルチキャスト呼を設定することが出来る。さらに、呼の品質保証方式として

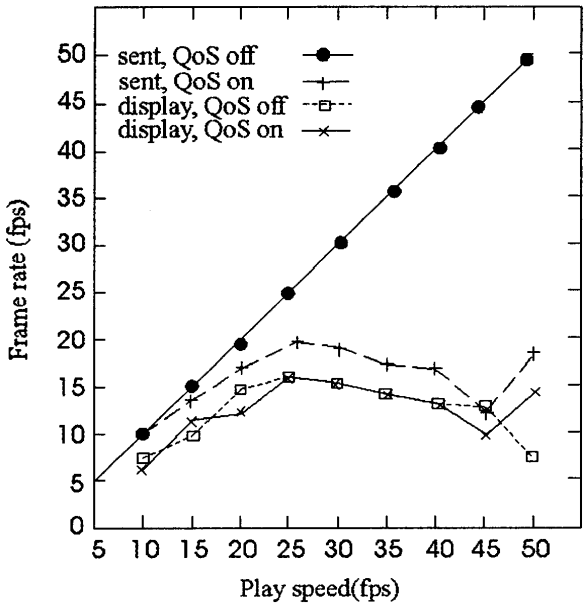


図5　VCR の性能評価(1)[7]

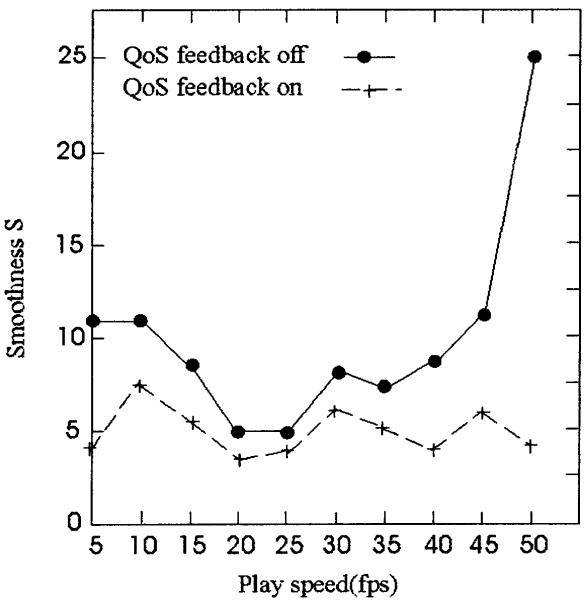


図6　VCR の性能評価(2)[7]

は、guaranteed connection と best-effort connection の2種類があり、それぞれ資源予約型、適応制御型の制御に対応している。guaranteed connection では、送信端末はあらかじめ、次に示すパラメータにより品質を申告することができる。

- 最大データ長
- 送信レート
- 遅延時間および遅延分散

HeiTP ではこれらのパラメータを ST-II に申告して、ST-II が経路上の帯域の確保を行なう。一方 best-effort connection は、トラヒック品質の保証は行な

学術情報センター電子図書館サービス　　©学術情報センター

表5　適応制御型プロトコルと資源予約型プロトコルの性能比較

	適応制御型	資源予約型
ルータの機能の簡略化	○	△
回線利用効率	○	△
品質保証能力	△	○
適切なアプリケーション	・ファイル転送 ・可変速度画像・音声通信	・固定速度動画・音声通信
適切な通信形態	コネクション型 またはコネクションレス型	コネクション型

われない。

HeiTP で特徴的であるのは、マルチキャスト呼を設定した際のレート制御である。ここでは、送信すべき画像は画像の品質に応じてサブストリームと呼ばれる画像に分割されている。図7 に示す例では、保証すべき画質レベルのサブストリームを資源予約型呼で、それ以上の画質をもつサブストリームを適応制御型呼で転送することにより、画質の保証と、帯域の有効利用が両立される。またマルチキャストの場合は受信側の受信能力と経路の利用可能帯域幅は端末によって異なることが普通であるが、サブストリームの数を調節して、端末の能力に応じた画像を再生することが可能になる。

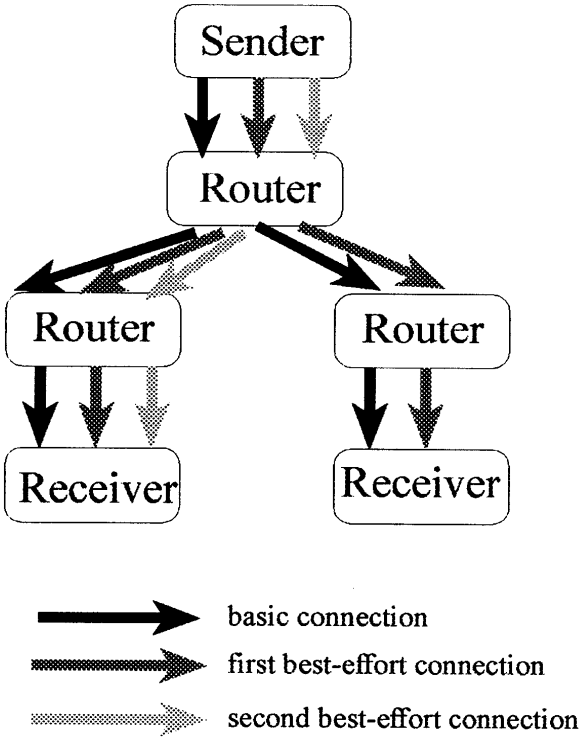


図7　HeiTP におけるマルチキャストのレート制御 [8]

4　まとめ

4.1　制御方式間の性能比較

本稿で取り上げた適応制御型プロトコルの中には、画像品質の評価が行われていないものがあるが、レート制御の帯域利用効率と変動する帯域への対応速度によって品質を定性的に評価することは可能である。

VDP、CMPlayer、VCR の3つの方式はいずれも単位時間あたりの転送フレーム数を調節してレート制御を行っている点では同じであるが、VDP と VCR ではレートの低減が一定値であるのに対して、CMPlayer はパケットの連続的損失に対してレートの低減幅をより大きく設定している。したがって、帯域の短期間変動に迅速に対応できる点で最も優れているのは CMPlayer であると考えられる。

また、適応制御型と資源予約型それぞれの制御方式の性能を大まかに比較すると表5 のようになる。適応制御にはルータの機能が簡単であることのほかに可変速度動画・音声通信のようにデータ量が時間によって変化する場合は、トラヒック統計多重効果によって、収容呼数の増加が見込める。通信形態としては、コネクションレス型、コネクション型のいずれもが対応できる。ただし、フィードバック制御であることから、遠距離では性能が劣化すると考えられる。資源予約型制御はコネクション型通信の上でしか動作しないが、品質保証能力に優れており、特にマルチキャスト通信および固定速度の動画・音声通信に適している。

適応制御型と資源予約型はそれぞれ利点があることから、HeiTP の例に見られるように、ネットワークでは双方のサービスをサポートして、最低限保証すべき品質レベルの情報を資源予約型で、それ以上の品質を持つ情報を適応制御型で転送する方法が最も有効な利用方法であろう。

学術情報センター電子図書館サービス　　© 学術情報センター

4.2 将来のネットワークインフラストラクチャ

本節では、これまでの議論をふまえて、マルチメディア通信を実現するための、特に広域網に求められるネットワークインフラストラクチャのアーキテクチャについて考える。IPの提供するコネクションレス型通信は本来、トランザクション処理、分散ファイルシステムのように少量のトラヒックが頻繁に発生するデータ通信で性能を発揮することが知られており、4.1で述べたように、さまざまなトラヒック特性をもつ動画・音声通信に対して総合的に品質保証を行うためには、帯域保証の機能を持つコネクション型通信が必要となる。

したがってネットワークとしては、端末とのインターフェースは単一ながら、通信形態としてはコネクションレス型通信、コネクション型通信の双方のサービスを提供し、ユーザの側で伝送するメディアに応じてどちらかを選択することができれば理想的である。

その一つの方法として、ATM網をベースとした、広域VLAN(Virtual Local Area Network)アーキテクチャが提案されている。[10]コネクション型通信はATMのVC(Virtual Circuit)で、コネクションレス型通信はネットワーク内部のノードのCLSF (Connection Less Server Function)により実現される。本方式の特徴は二つあるが、一つは、コネクションレス型通信の中でtelnet、ftpといった保留時間の長い通信に対してATMのSVC(Switched Virtual Circuit)をオン・デマンドに設定してコネクション型通信に切り替えることにより高速なパケット転送を可能にしている点である。同様のアプローチは、Ipsilon社のIpsilon Switchにも見られる。[10]もう一つは、網の構成を中継網と地域網に階層化するとともに、コネクションレス型通信のパケット転送にIPとは異なる独自プロトコルを用いて、ネットワークアドレスも階層化している点である。これにより、パケットの中継段数が減少するとともに、CLSFノードはアドレスの一部だけを参照することによりパケットの宛先を決定することができ、高速なルーティングが可能になっている。

一方、全く別の方法としては、Post-ATMアーキテクチャ[11]があげられる。ここでは、動画・音声・データ等の要求品質の異なるメディアごとに別々のネットワークが構築される。ユーザから発信されたパケットはmedia-multiplexerで時分割多重により単一の固定長フレームに統合され、media-demultiplexerでパケットのメディアの種類に応じたネットワークに振り分けられる。通信形態としては、動画・音声通信のネットワークではコネクション型通信が、データ通信のネットワークではコネクションレス型通信が用いられる。本方式は、入力されたパケットがほぼそのままの形で中継され、ATMで問題となるようなパケットとATMセルの相互変換のオーバヘッドが存在しないことが特徴である。研究課題としては各メディア別ネットワークの交換機、および性能に大きな影響を与えるmedia multiplexer、media demultiplexerの構成法等がある。

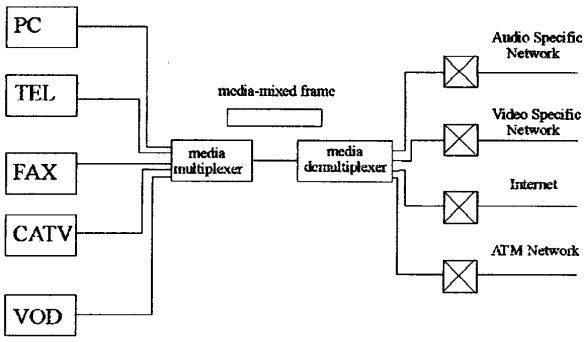


図8　PostATMネットワークアーキテクチャ

アーキテクチャの評価基準としては、伝送速度の高速化への対応能力、種々のLAN(Local Area Network)の通信方式に対する接続性、コストパフォーマンス等があげられるであろう。上に挙げたほかにもまだ多くのアイデアの出る余地があり、「使いやすく、高品質で、安価な」ネットワークサービスを目指して、今後インフラストラクチャに関する議論が深まってゆくことを期待したい。

参考文献

[1] 西田竹志 他,「マルチメディア・ネットワーク」, 日経エレクトロニクス, 日経BP社, No.583, 1993年6月.

[2] William, Stallings., "Data and Computer Communications", Macmillan Publishing Company, 1991.

[3] L, Zhang. ; S,Deering. ; D, Estrin. ; S, Shenker. ; D,Zappala., "RSVP: A new resource ReSerVation Protocol", *IEEE Network magazine*, Vol. 9, No. 5, 1993.

[4] Topolcic, C., "Internet Stream Protocol Version 2 (ST-II)", RFC 1190, October 1990.

[5] Zhang, Chen.; See-Mong, Tan. ; Roy, H, Campbell. ; Yongcheng, Li., "Real Time Video and Audio in the World Wide Web". **http://choices.cs.uiuc.edu/Papers/New/vosaic/vosaic.html.**

学術情報センター電子図書館サービス ©学術情報センター

[6] Lawrence, A, Rowe. ; Brian, C, Smith., "A Continuous Media Player", *Proceedings of 3rd International Workshop on Network and OS Support for Digital Audio and Video*, San Diego CA, November 1992.

[7] Shanwei, Cen.; Calton, Pu. ; Richard, Staehli. ; Cripin, Cowan. ; Jonathan, Walpole. , "A Distributed Real-Time MPEG Video Audio Player", Available via anonymous FTP from **http://hulk.bu.edu/nossdav95/papers/CEN_Shanwei.ps.**

[8] Luca, Delgrossi., "Design of Reservation Protocols for Multimedia Communication", KLUWER ACADEMIC PUBLISHERS, 1996.

[9] 牧野将哉, 中村陽一, 会田雅樹, 村山純一, 「広域VLANアーキテクチャ」,NTT R&D, Vol.45, No.10, pp.961-970, 1996.

[10] P, Newman. ; W, Edwards. ; R, Hinden. ; E, Hoffman. ; F,Ching, liaw. ; T, Lyon. ; G, Minshall., "Ipsilon's General Switch Management Protocol Specification Version 1.1" Network Working Group Request for Comments:1987, August 1996. **http://www.ipsilon.com/protocols/rfc1987.txt.**

[11] 青木輝勝, 相田仁, 齊藤忠夫, 「Post-ATM Networkアーキテクチャの展望とその技術的課題」, 1996年電子情報通信学会通信ソサイエティ大会, B-795.

学術情報センター電子図書館サービス　　©学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

Wellcome財団の助成活動と研究評価を支援するデータベースの役割

Activities of research funding and the role of research databases for supporting the evaluation of scientific performance in the Wellcome Trust

東京慈恵会医科大学医学情報センター　山崎　茂明
Shigeaki YAMAZAKI
Medical Information Center for Education and Research, Jikei University School of Medicine

要旨

Wellcome財団は、1936年に死去したSir Henry Wellcomeの意志で創設された。その目的は、生物医学と医学史の研究を支援することである。現在、世界最大の私立慈善助成財団となり、1994年の助成金支出は2.4億ポンドにのぼっている。その一部門であるPRISM (Unit for Policy Research in Science and Medicine)は、財団として取り組むべき重要なテーマの識別や申請の採択にあたり、その意思決定を支援している。さらに、研究資金によって助成された研究が、実際どのような成果を産みだしているかを分析するためにROD(Research Outputs Database)データベースを作成し、研究評価のための調査活動を行なっている。RODデータベースは、イギリスの生命科学研究から産み出された論文の書誌情報に助成機関の情報をリンクしたものであり、助成機関にとっては提供した資金の効果を判定することができる。また、Wellcome財団は、WISDOMシステムを開発し、人事募集、研究助成金、科学政策情報などの3つのユニークなデータベースを提供している。今後、わが国において効果的な研究助成方式を確立するためには、助成された研究活動の評価が重要になり、Wellcome財団の活動はそのモデルとなるだろう。

ABSTRACT

The Wellcome Trust was founded under the will of Sir Henry Wellcome, who died in 1936. The objectives of the Trust are to support research in the biomedical sciences and in the history of medicine. Now, the Wellcome Trust that is the most richest medical charity in the world, spent £242 million for the grant expenditure of 1994. PRISM, the Unit for Policy Research in Science and Medicine, has been developing ways of helping the Trust choose between areas of research and between different applicants. It has also been developing experimental approaches to priority setting and creating an exciting new tool for evaluation of outcomes, the Research Outputs Database. This unique database is develpped with the bibliographics detailes of UK biomedical research linked to their funding sources. Using the Research Outputs Database, UK funding organizations have been able to track the output of the research they have supported. The Wellcome Trust have also developed other useful databases. WISDOM consists of three databases on funding sources for biomedical research, science policy and the catalogue of the library. Evaluation of scientific performance would become more important in oredr to establish the effective research funding. The Wellcome Trust appears to present us a model for the system.

[キーワード]　ウエルカム財団、データベース、研究助成、研究評価
[Keywords]　Wellcome Trust, Database, Research Funding, Research Evaluation

1　はじめに

日本の研究活動の問題点として、欧米から、基礎研究ただのり論が、また大学と企業との連携の弱さが指摘されている。最近の自然科学領域での、日本の論文生産数の高さからは、活発な研究活動が示されている[1-3]。しかし、日本の研究費総額に占める政府資金の比率は、アメリカ、ドイツ、フランスなどと比較して、著しく低いのが現状である。基礎研究より企業を中心とした応用研究が主体になっていることは間違いないであろう。基礎研究へ政府資金の投入量を増大させることで、従来の経済や産業優先の研究体制を転換していくことがより一層求められる。同時に、大学と企業

学術情報センター電子図書館サービス　　©学術情報センター

との連携を強めることが求められている。日本の企業の一部には、海外の大学との連携をはかり、基礎研究を充実させようとしている例もある。このことは、日本国内にふさわしい提携先を見い出せないという現状を示している。この際、大学や大規模研究所などの強化が、基礎研究の充実のために必要である。そして、このような科学政策上の問題点を解決できるような方向で、データベースに関する政策を検討していくことが大切である。また、科学研究活動への予算増に対する一般の理解を得られるよう、人々が研究成果を直接利用できる方法を確立しておくことも忘れてはならない。

1995年11月に成立した議員立法である科学技術基本法により、日本は「科学技術創造立国」を目指すことを決定し、科学技術振興を推進していくことになった[4]。日本学術会議では、科学技術基本法の成立前に、伊藤会長が日本の学術体制を巡って多面的な視点から見解を表明していた[5]。各省庁では、この基本法の趣旨に沿って、さまざまな施策を準備している。文部省においても、情報化の推進と関連し、学術研究データベースの充実が課題としてあげられた。そこで、平成8年度科学研究費補助金「学術情報データベースの形成・利用の国際化と研究動向の国際比較に関する調査研究(研究代表者 根岸正光教授:学術情報センター)」により、欧米の先進国を中心に関連機関の訪問調査が行なわれた。本稿は、イギリスのWellcome財団に焦点をあて、その助成活動の概況と研究管理や研究評価などに果たすデータベースの役割を論じたものであり、Wellcome財団をモデルとして、日本のデータベース関連政策を検討することを目的としている。

2 Wellcome財団とその資金

Wellcome財団は、1936年に死去したHenry Wellcome卿の意志で創設された。その目的は、生物医学と医学史の研究を支援することである(図 1)。Wellcome 財団の収入は、その60パーセントが世界の300社以上の企業へ投資した有価証券の利子、そして残りの40パーセントが世界最大の製薬企業であるGlaxo-Wellcomeの4.7パーセントの株式などから形成されており、それらから1995年には、3.3億ポンド(£330 million)の収入があった。日本円で578億円(1ポンドを175円で換算: 1996年9月現在)に相当する。

1992年のScience誌によれば、同年のWellcomeの予算は£175millionであり、イギリス政府のMRC(Medical Research Council)の予算と匹敵する金額であった[6]。なお、Wellcome財団が特に焦点をあてている研究領域は、分子生物学、熱帯医学、生理学、薬学・薬理学、神経科学などの分野である。医学領域で重要な癌研究については、すでにその他の有力な支援財団が存在しており、あまり力をいれていない。

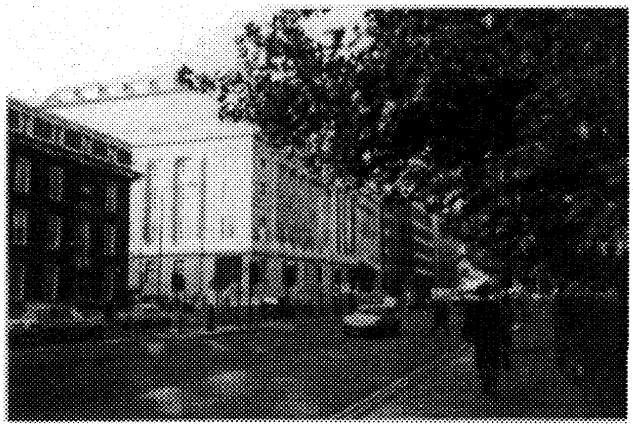


図 1　Wellcome財団本部ビル(London, UK)

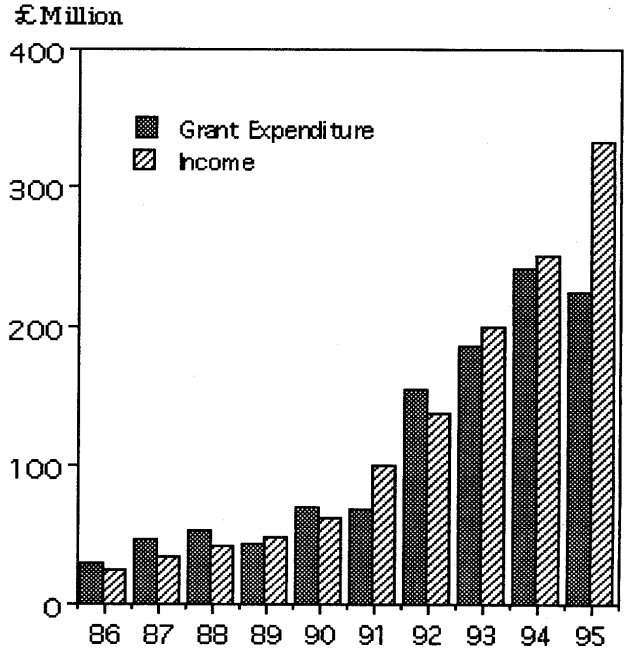


図 2　Wellcome財団の助成額と収入額の年次変化

実際の助成額についてみると、1995年では389億円(£222 million)にのぼる(図 2)。このうち18億2250万円(£7million)が医学史関連の助成額であり、助成資金の大部分(97 パーセント)は基礎医学研究から臨床研究といった最新の生命科学領域の研究に支出されている[7]。日本の文部省科学研究費の総額は1995年には1000億

学術情報センター電子図書館サービス　　©学術情報センター

円の大台に近づいたが、Wellcome財団の資金だけで日本の科研費の40パーセントに相当している。Wellcome財団は、文部省の科学研究費のように人文社会科学から自然科学のすべてをカバーしているのではなく、生命科学領域だけを対象としているので、この領域に果たす財団の資金力は巨大なものといえる。

3 Wellcome財団の組織と活動

Wellcome財団の初期の主要な活動は、医学史研究のための図書と資料の収集であった。このうち、図書は医学史図書館(Wellcome Institute for the History of Medicine Library)に保管され、世界的な医学史コレクションとしてまとめられている[8]。歴史資料はロンドンの科学博物館(Science Museum)に移管され、まとめて展示されている。現在の財団の活動は、分子生物学研究への助成に示されるように、医学史だけでなく、最先端の生命科学研究に助成の中心を置いており、英国の医学研究を支える最も重要な助成機関になっている。なお、Wellcome 財団は、設立の母体となった Glaxo-Wellcome 製薬からは独立した機関であり、研究助成金の支出にあたり、いかなる影響も受けていない。現在の財団の主要な部門は、助成事業部、医学史、助成金管理、財務、人事・サービス部、医学センターから形成されている(表1)。

表1　Wellcome財団の主要組織

1. Science Funding
2. History of Medicine
   1) The Wellcome Institute Library
   2) The Wellcome Unit for the History of Medicine
3. Grants Administration
4. Finance
5. Personnel and Services
6. Wellcome Centre for Medical Science
   1) Centre Administration
   2) Communication and Education
   3) Information Service
   4) PRISM(Policy Research in Science and Medicine)
   5) Medical Photographic Library
   6) Tropical Medicine Resource
   7) Audiovisual Resources Unit

4 Wellcome財団で制作されているデータベース

Wellcomeは、世界最大の生物医学領域における研究財団であり[9]、取り組むべき重要なテーマを識別して、研究資金を適切に配分し、そしてそれらの効果を評価する必要性がある。そのため財団内にPRISM (Unit for Policy Research in Science and Medicine)を組織し、ROD(Research Outputs Database)を作成している。また、Information Service 部門では、インターネットとJANET(Joint Academic Network)を通してアクセスできるデータベースサービスを提供している。このデータベースサービスは、WISDOM(Wellcome Information Service Databases on Medicine)と命名されており、つぎのような3つのユニークなデータベースから形成されている。

1) 人事交流を活発にするための人事募集データベース(Biomedical Research Assistant Vacancies Database)
2) 研究活動を活発にするための助成金データベース(Sources of Biomedical Research Funding:Research Grants Database)
3) 科学政策情報の速報的な提供サービス(SPIN:Science Policy Information News)

Wellcome財団のPRISMとInformation Serviceでは、研究助成事業の効果的な発展と運用のために、これら4種類のユニークなデータベースを制作している。

また、Information Serviceでは、図書館情報サービスを、一般の人々を含め、科学政策研究者、研究管理者、科学研究者、教員、学生などに広く提供している。科学・医学知識の普及と科学政策や研究評価といった専門領域のための情報サービスをあわせて展開している。

さらに、科学教育を支援するためのシステムが、Wellcome財団も参加して、イギリスで形成されている。Science-Netは、4000以上の学校を結び、科学に関するさまざまな質問を生徒や教員から電子メールで受け、その回答も電子メールでなされるという相談サービスである。回答は300名以上のそれぞれの専門研究者によってなされている。これらの質問と回答は毎月データベースに蓄積され、10の主題トピックスに分類され、キーワードから自由に検索できるようになっている。現在、毎週100件の質問が寄せられている。このサービスは、Wellcome財団、Royal Society、Office of Science and Technologyによって助成され、一般の人々への科学の普及と社会教育を目標にしており、科学をより身近なものにするシステムである。電子メールばかりでなく、電話による同様のサービスもScience-Lineとして提供されている。なお、電子メールによるScience-Netの優れている点として、データベース化により、蓄積された質疑応答を検索できることがあげられる。科学を専門家の世界に閉じ込めず、一般の人々への回路をつ

学術情報センター電子図書館サービス　　©学術情報センター

くる姿勢は、科学の透明性を高めるという今日的な方向とも適合しており、重要な視点である。

また、Wellcome 財団は、OMNI(Organising Medical Networked Information)プロジェクトを推進する7つの機関のひとつになり、インターネット環境下での生命科学領域における効果的なナビゲーションシステムの開発を行なっている。Science-Net と OMNI プロジェクトは、Wellcome財団独自のものではなく、他機関と協力して行なっている事業であるが、データベースの形成やサービスに対する財団の立場を考える際に役立つであろう。

以下、まず研究評価や研究動向の調査を行なっているPRISMとそこで作成されているRODデータベースについて、つぎにInformation Serviceで提供されているデータベースサービスであるWISDOMについて、そしてゲートウエイサービスであるOMNIプロジェクトについて解説する。

5 研究評価と研究動向を調査するPRISMとRODデータベース

Wellcome財団の最も主要な生命科学への寄与は、優れた研究者への資金提供にある。PRISMは、Wellcome Centre for Medical Scienceに所属しており、財団の助成すべき研究領域の選択や助成申請の採択を支援する役割を担っている。現在では、基礎研究と応用研究という枠組みだけでなく、将来どのような研究が広く社会に必要となるのかといったニーズにこたえる戦略的研究(strategic research)が求められている。PRISMは、そのための基礎データをまとめ調査研究を行なっている。さらに、研究資金によって助成された研究が、実際どのような成果を産みだしているかを分析するために、ROD (Research Outputs Database)データベースを作成している。このデータベースは、多くの助成機関にとって有用な情報を提供するものであり、研究評価のためのユニークなツールとみなすことができる。

研究助成機関は、財政状況や研究環境の変化のなかで、時代に適応した新しい助成方針を設定するため、かれらが支援した研究成果の評価をする必要がある。生命科学領域では、助成を得た数年後に、その研究成果が学術雑誌の掲載論文として発表される。研究者は、研究助成を受けていれば、論文の謝辞のなかで助成機関名や支給年度などを記載する。そこで、RODデータベースは論文の書誌情報、著者の所属機関名や住所にくわえ、助成機関名を記録している。

具体的には、ISI社(Philadelphia,米国)のScience Citation Index(SCI)誌から、著者の所属機関がイギリスである生命科学論文を選び、実際にそれらの論文の謝辞にあたり助成団体名の情報を採録している。1988年から毎年約2万5千論文の情報が蓄積されている。財団名は、3文字のコードで入力されており、例えばWellcomeは「WEL」、Medical Research Councilは「MRC」となっている。さらに、団体の所属国名と助成団体の性格を識別できるようにしている。著者の所属機関住所は、イギリスの郵便コードでも記録されており、これは地域的分析のデータにもなっている[10]。癌領域を例にすると、1988年から1992年の5年間に英国で11580の論文が発表されており、このうちの62パーセント(7147件)に助成機関への謝辞が存在した。その内訳は、私立慈善財団(charity)が66パーセント、政府機関(MRC:Medical Research Council, SERC:Scientific and Engineering Research Council, AFRC:Agriculture and Food Research Council など)が23パーセント、産業界が12パーセントであった。さらに、それらを個々の財団別に分析することができる[11]。

RODデータベースは、1995年7月現在、27のメンバーによるクラブ組織で維持されている。メンバーは、各々の助成金額に応じた予約金を支払い、自分の助成金による研究成果である論文のリストを受けとる。例えば、助成金額が年間1億ポンド以上の機関はクラスAに分類され、1万ポンドの年間予約料を支払っている。100万ポンドから1000万ポンドの機関はクラスBで3千ポンドの予約金であり、年間10万ポンド以下の助成機関はクラスEで予約金は100ポンドとなる。さらにクラブメンバーは、PRISMの行なった調査レポートや特定の問題に関わるより詳細な調査データの提供を受けることができる。

6 データベースサービスとしてのWISDOMシステム

Wellcomeは、研究成果を評価するためのデータベースであるRODの他に、研究活動を活発にするため、助成金データベース(Sources of Biomedical Research Funding)、人事交流を活発にするための人事募集データベース(Biomedical Research Assistant Vacancies Database)、科学政策関連ニュースの速報(SPIN:Science Policy Information News)などのデータベースを、WISDOMシステムにおいてJANETとインターネットを経由して広く世界に提供している(図3)。

学術情報センター電子図書館サービス　　©学術情報センター

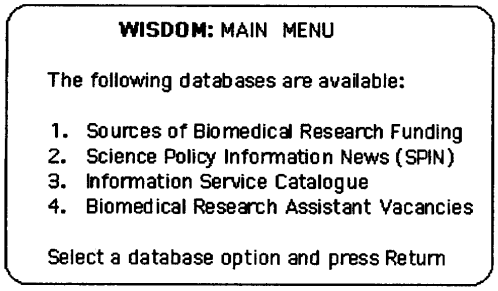


図3　WISDOMのメニュー画面

Sources of Biomedical Research Fundingは、医学研究のための約450以上の助成金募集情報を掲載している。レコードには、助成の目的や種類、助成額、応募方法などの情報が入っている。Biomedical Research Assistant Vacancies Databaseは、英国内の学術機関での大学院生、ポストドクなどの若手研究者が、適切な研究の場を探すためのもので、研究主題、地域、給与レベルなどのさまざまな視点から検索できる。科学政策に関するニュースの速報的な提供サービスがSPINであり、Nature、Science、New Scientist、Lancet、BMJなどの総合誌や科学政策の専門誌など150誌からの記事が収載されている。これは印刷体でも刊行されており、カレントアウアネスに有効である。

7　ゲートウエイとしてのOMNI(Organizing Medical Networked Information)

このプロジェクトの目標は、インターネット上に存在する医療と生物医学領域の質の高い情報資源へのゲートウエイを確立することである。インターネットは、生命科学領域の研究者や医療サービスの専門家にとって、今後ますます重要になる情報源であり、OMNIはネットワークに存在している適切な資源を発見するための目録を提供する。情報源への索引やキーワード付けを行ない、注記をくわえ、さらに選択を行ない、質の高い情報提供を目標にするものである。OMNIプロジェクトは、7機関によるコンソーシアムにより開発されている(表2)。膨大な量に成長しているインターネット資源への、ナビゲーションとゲートウエイを提供しようとしている。

8　Wellcome財団から学ぶもの

世界最大の私立助成財団であるイギリスのWellcome財団をモデルとして、日本のデータベース関連施策に参考になる事項をまとめると次ぎのようになるであろう。学術データベースは、基本的に商用データベースとして成立することが困難であり、政府資金の投入が

表2　OMNIプロジェクトに参加している7機関名

1. British Medical Association
2. Cambridge University
3. King's Fund
4. National Institute for Medical Research
5. Nottingham University
6. Royal Free Hospital School of Medicine
7. Wellcome Centre for Medical Science

成功の鍵になる。その際、資金量的な側面だけでなく、研究活動や研究体制の質的な転換に資するよう、資金を戦略的に配分することが大切である。この点で、研究評価を意図したRODデータベースの意義は重要である。PRISMでは、これを用いて現在の研究動向を分析し、今後主要なテーマとなる研究領域を予測し、それらを主要先進国との連携と競合関係のなかで分析している。さらに、研究資金の投入という意思決定場面において、research managementやresearch evaluationなど、研究活動自体を対象とする研究の役割が認識されている。「研究費の増額よりは、そのより効率的な運用によって目的を達しようとしている」[12]イギリスの学術政策の特色が、ここに示されているといえよう。これらの点は、日本における科学政策とデータベースの役割を検討する際にモデルとなるであろう。

Wellcomeの研究助成活動に果たす、データベースの役割から示唆されるように、研究助成金の獲得、人事募集情報、科学政策ニュース、研究評価といった、研究プロセスのさまざまな局面で有用なデータベースを制作することが大切である。これらのデータベースは、研究者や科学界の全体に役立つものであるが、個々の専門研究者が制作するとは考えられない。それだけに、公的機関が取り組むのにふさわしいデータベースである。研究活動を活発にし、人事の流動性を高め、研究評価を行ない、そして限られた資金を投入すべき重要分野を明確にするなど、これらのデータベースは研究動向を分析し科学政策を立案するうえで重要な役割を果たす。研究費の量的な増額をはかるだけでなく、効果的な運用(研究の質)を目標に置いた科学政策関連データベースの強化が必要である。

また、データベースの量的増大は、それ自身の情報洪水をもたらすものであるだけに、OMNIプロジェクトに示されたように、利用者のニーズに適合したデータベース資源へのゲートウエイとナビゲーションシステムの開発も要請される。利用者が、それぞれの問題解決にあたり、適切なデータベースを選択できるよう

学術情報センター電子図書館サービス　　© 学術情報センター

な支援機能が開発強化されるべきであろう。

謝辞

本論は平成 8 年度文部省科学研究費補助金「学術情報データベースの形成・利用の国際化と研究動向の国際比較に関する調査研究」による海外視察にもとづくものであり、研究代表者の根岸正光教授(文部省学術情報センター)のご指導に感謝いたします。また、Grant Lewison 博士(Wellcome 財団 PRISM)から受けた助言と有益な情報に対し感謝の意を表します。

参考文献

[1] 根岸正光,「学術論文数の国際比較調査－結果の概要と分析視点」, 情報管理,Vol.39, pp.245-257, 1996.

[2] 石井啓豊,「学術論文数の国際比較調査－先進 6 カ国の動向」, 情報管理, Vol.39, pp.326-338, 1996.

[3] 山崎茂明,「学術論文数の国際比較調査－医学領域の分析」, 情報管理, Vol.39, pp.391-407, 1996.

[4] 水間英城,「科学技術基本法について」, 情報管理, Vol.38, pp.1080-1093, 1996.

[5] 伊藤正男,「始動する英国の新学術政策」, 日本学術会議月報, Vol.35, pp.112-115, 1994.

[6] Aldhous, P., "Wellcome Trust: Britain's big biomedical spender", *Science*, Vol.256, pp.1132-1133, 1992.

[7] "The Wellcome Trust annual record 1995 CD-ROM", London, The Wellcome Trust, 1996.

[8] 酒井シヅ,「イギリスの医学史近況(1)」, 科学医学資料研究, No.187, pp.1-5, 1989.

[9] Moran, N.,"Wellcome reveals plans for research funding", *Nature Medicine*, Vol.1, p.394, 1995.

[10] O'Driscoll, M.; Chisholm, P.; Anderson, J., "Immunology research in the UK: a scientific audit", *Immunology Today*, Vol.16, pp.413-417, 1995.

[11] Lewison, G., 私信, 1996 年 6 月 28 日.

[12] 伊藤正男,「我が国の学術体制を巡って」, 日本学術会議月報, Vol.36, pp.56-65, 1995.

学術情報センター電子図書館サービス　　◎学術情報センター

学術情報センター紀要 第9号 (1997)

研究論文

Medlineデータベースからみた臨床試験文献の分析：コクラン共同計画との連携

Bibliometric analysis of the papers on clinical trials and randomized controlled trials in the Medline database, with reference to the Cochrane Collaboration

東京慈恵会医科大学医学情報センター　山崎　茂明
Shigeaki YAMAZAKI
Medical Information Center for Education and Research Jikei University School of Medicine

要旨

本稿の目的は、臨床試験と無作為化比較試験に関する文献を対象に、論文数の成長、言語と発行国別にみた分布などを明らかにすることである。1966年から1995年のMedlineデータベースにもとづき、臨床試験と無作為化比較試験に関する文献を、1991年に新しく導入された出版タイプを識別キーにして検索した。その結果によれば、日本は世界の科学文献の生産において、高い占有率をしめているが、臨床試験と無作為化比較試験に関する文献のシェアは、その他の主要先進国と比較して小さい。今日、臨床試験や無作為化比較試験は、医療の有効性を評価するために利用されるようになっている。コクラン共同計画では、米国国立医学図書館と協力して、臨床試験や無作為化比較試験によって出版された論文を識別している。これにより、臨床家や医療従事者は、Medlineデータベースに収録された臨床試験文献から科学的な証拠を見いだすことができる。わが国においても、医学領域におけるデータベースを改良するために、JMEDICINEや医学中央雑誌は、臨床試験や無作為化比較試験についての十分な情報を索引する必要があろう。

ABSTRACT

The purpose of this survey is to study the growth of number of papers and the distribution of language and country of publication in the field of clinical trials (CTs) and randomized controlled trials (RCTs). Based on the Medline database between 1966 and 1995, the number of papers on CTs and RCTs were obtained using the publication type that was introduced into the Medline in 1991. Though Japan has a large share of world publications in science, its share in the field of CTs and RCTs is smaller than that of the other developed countries. CTs and RCTs are now increasingly being used to evaluate the effectiveness of health care intervention. The Cochrane Collaboration is working with the National Library of Medicine to identify the published papers on CTs and RCTs. As the consequence, practicing physicians and allied health personnel are able to find the best evidence from clinical trials in the Medline database. In order to improve the use of Japanese databases in medicine, JMEDICINE and Igaku Chuo Zasshi (Japana Centra Revuo Medicina) should be suggested to index adequate information on CTs and RCTs.

[キーワード]　計量文献学、コクラン共同計画、データベース、臨床試験、無作為化比較試験
[Keywords]　bibliometrics, Cochrane Collaboration, databases, clinical trials, randomized controlled trials

1　はじめに

米国国立医学図書館(National Library of Medicine)で制作しているMedlineデータベースは、1879年に創刊された世界の医学文献索引誌であるIndex Medicusを母体に発展したものであり、医療の問題解決ツールとして、世界の医学研究と臨床、健康サービスに貢献してきた。体験や経験に依存するだけでなく、情報と知識をベースにした医学・医療の世界を支援する最大のリソースとみなすことができる。世界の医師によって最も信頼されている総合医学雑誌New England Journal of

学術情報センター電子図書館サービス　　©学術情報センター

Medicineの編集委員長であるKassirer[1]は、電子情報環境下の医師に求められる資質として、Medlineデータベースの効果的な活用を提言していた。臨床の場において、さまざまな問題に直面しなければならない時、これまでの膨大な医学知識が蓄積された文献データベースこそ、解決のための重要な情報源になるからである。同時に、このデータベースを、常に信頼できる最新の文献情報を適切に検索できるように改良していくことは、医療の質の向上に大きな影響を持つ。

このMedlineデータベースに、1991年から出版タイプ(publication type)のフィールドが新設された。これにより、無作為化比較試験(randomized controlled trial:RCT)などの臨床試験(clinical trial)論文がデータベースから検索できるようになったということは、医師だけでなく最終的には患者に役立つものであり、科学的な根拠にもとづいた医療(evidence based medicine)の基礎になるものである。

本稿の目的は、Medlineデータベースの出版タイプのフィールドをもとに、臨床試験論文の定量的な分析を行ない、世界の主要国のこの分野における貢献度と日本の現状を明らかにすることである。

2　問題解決ツールとしてのMedlineデータベース

Medlineデータベースから、適切な文献を検索していくためにMedical Subject Headings(MeSH)と呼ばれるキーワード体系がつくられている。実際の論文では、同一の疾患をさまざまな言葉で表現したり、不統一な略語の使用などがあり、効果的な検索のためには、用語のコントロールが欠かせないからである。MeSHの1996年における語数は約1万7千語であり、1文献にたいして平均12語程度が与えられている。また医学では、常に診断、治療からはじまり、病理、副作用、合併症、疫学など共通するアプローチが存在する。そこで、疾病や薬物などを表現するキーワードとともに、約80語のサブヘッディングが付与された。Medlineは、このような主題を示すであるキーワードであるMeSHと、アプローチを識別するサブヘッディングを用いて、利用者の多様な問題解決に対応できるよう発展してきた。

このような発展プロセスのなかで、1980年代に入り、医療界では医療の質や倫理への関心が増大してきた。そこで、Medlineデータベースを制作している米国国立医学図書館は、1992年になり「論文の撤回、注記、誤りなどにたいする方針」を発表した[2]。データベースに収載された論文に関する撤回、注記、誤植などについて、もとの文献とリンクさせることで、検索利用者に注意を喚起している。Medlineを検索した利用者は、検索された論文がその後に誤りのために撤回されていることを、ディスプレイやプリントアウトから識別できる。また、エラータが発表されていることも知ることができるようになった。具体的には、1991年に創設された出版タイプのフィールドが、これらの識別に当てられる。こうして撤回された情報がそのまま流通したり、誤りを含んだ情報が知らぬ間に使用されないようにし、出版倫理へ対応している。

3　コクラン共同計画、臨床試験とMedline

イギリス人医師Archie Cochraneの名を冠したコクラン共同計画(The Cochrane Collaboration)は、1992年からイギリスのNational Health Services(NHS)の一環として始まった医療技術評価プロジェクトである[3]。Cochraneの主張は、1972年に発表した単行書“Effectiveness and Efficiency: Random Reflections on Health Services”(London, Nuffield Provincial Hospitals Trust: Reprinted in 1989)で展開されており、人々へ提供される医療サービスは、無作為化比較試験によって導かれた明らかな証拠にもとづくべきであるというものである。彼のアイディアは明快であり、医療関係者だけでなく一般の人々にも広く認められるようになったが、その考えを実行に移すには20年近くの時間が必要であった[4]。コクラン共同計画の目的は、薬の正しい評価を行ない、医師が必要とする信頼できる情報を提供することにある。すなわち、何が標準的な治療であり、何がまだ評価の定まらない治療であるかを明快に示し、薬剤選択の意志決定を容易にするものである。

治療のために使用する薬剤は、多くの開発費と時間を費やしてできあがる。創薬のプロセスは、前臨床試験と呼ばれるドラッグデザインから始まり、化合物の合成、薬効スクリーニング試験、動物実験までのプロセスと、ヒトを対象に薬物動態、安全性、薬効を調査する臨床試験の段階とに大きくわけられる。この臨床試験の情報は、より信頼性の高い臨床サービスのために重要であり、特に証明力の強い無作為化比較試験による結果は価値がある。そのために、コクラン共同計画において、治療法に関するすべての無作為化比較試験を収集し、それらを吟味する機関と体制が組織された[5]。これは、科学的証明にもとづいた質の高い医療

学術情報センター電子図書館サービス　　©学術情報センター

サービスの基礎になり、Cochrane のアイディアを実現するものであった。

現在、コクランセンターは 9 ケ所あり、それらはアメリカ、イギリス、オーストラリア、カナダ、オランダ、イタリア、デンマークの 7 国に設立されている[6]。そして、コクラン共同計画では、臨床家が、無作為化比較試験を中心とした臨床試験論文を Medline データベースから検索できるように、臨床試験を示す出版タイプのタグ付けに協力している。それは、専門家の点検によると、これまでの Medline データベースでは、無作為化比較試験文献の 50 パーセントしか識別できないことが指摘されているからである[7]。具体的には、コクランセンターでは、Medline に収録されながら、臨床試験に関する出版タイプのタグが付けられていないものをチェックしている。例えば、英国コクランセンターでは、欧州連合(EU)の助成を得て、1985 年から 1993 年の間で 10 万件の Medline データベースの抄録をチェックした結果、約 1 万 9 千件の臨床試験の論文が発見され、再度データベースのレコードに臨床試験のタグが付与されている[8]。

なお、コクラン共同計画では、Medline データベースに収載されていない雑誌に発表された無作為化比較試験を収集し、さらに雑誌で発表されていないものの収集も意図しており、世界で行なわれているすべての無作為化比較試験の登録を呼びかけている。そして、特定の医学領域における臨床試験を実際に収集し、それらの評価、解析を行なう共同レビューグループが組織され、無作為化比較試験のデータベースである CDSR(The Cochrane Database of Systematic Reviews)を制作している[9]。さらに、これまでの無作為化比較試験による発表論文の記載には、方法論的に貧弱なものが多いという指摘がなされ[10]、臨床試験をどのように行なったかについての標準的な記載基準が、米英の代表的な医師会雑誌に提言されるようになった[11,12]。

4 調査対象と方法

本調査では、Medline データベースの 1966 年から 1996 年 11 月更新分までを調査対象とした。Medline は医学領域における代表的な文献データベースであり、1995 年には世界の 3700 誌におよぶ代表的な医学雑誌から 39 万件の文献情報を収載している。1966 年から 1995 年までに約 860 万件のレコードが蓄積されたことになる。

個々のレコードは、書誌事項だけでなく、キーワードである MeSH や抄録などを含んでおり、さらに既述のように 1991 年から出版タイプ(PT)が付与されるようになった。この出版タイプは、それまでの citation type を発展させたものであり、1996 年の用語数は 49 である。論文の種類や出版倫理に係わる用語だけでなく、研究・統計方法についての用語が含まれている。特に、臨床試験についての重要な出版タイプが識別されている。“clinical trial”, “clinical trial phase I”, “clinical trial phase II”, “clinical trial phase III”, “clinical trial phase IV”, “controlled clinical trial”, “meta-analysis”, “multicenter study”, “randomized controlled trial” などである。出版タイプは、論文記事の形態を識別するだけでなく、出版の倫理や、研究手法など、論文の質的な側面を明示する重要な検索ポイントになっている。

今回の調査では、臨床試験文献の全体をとらえるために「臨床試験:clinical trial」を使用し、さらにその中でも重要な「無作為化比較試験:randomized controlled trial(RCT)」に焦点をあて分析した。出版タイプの clinical trial は、clinical trial phase I, clinical trial phase II, clinical trial phase III, clinical trial phase IV に細分されており、さらに controlled clinical trial, multicenter study, randomized controlled trial などの出版タイプの用語が必要に応じて付与されている。clinical trial という用語によって、臨床試験に関係する文献を広く網羅的に検索することができる。

5 結果と考察

5.1 文献数の成長(1966-1995):clinical trial と randomized controlled trial

1966 年の Medline データベース全体の文献数は 17 万 5 千件で、1994 年には 38 万 6 千件となっている。すなわちこの 28 年間で、データベースの文献量は 2.2 倍に成長している。Medline 全体の文献数に占める臨床試験文献の比率をみると、1966 年は 0.7 パーセント(1225 件)であったが、1994 年には 5.0 パーセント(19140 件)を占めるまでに成長した(図 1)。1980 年代の初頭に停滞があるものの、1960 年代の 1000 件台が、1970 年代に 3000 件台になり、1980 年代の後半から、急激な上昇傾向を示している。1966 年から 1994 年で、臨床試験文献は 15.6 倍に成長したことになる。無作為化比較試験の文献数は、特に 1980 年代後半から急激な増加に転じ、1995 年で 9162 件と 1 万件に近づいている。

5.2 発表論文の言語別にみた比較(1992-1994)

無作為化比較試験に絞って、1992 年から 1994 年の

学術情報センター電子図書館サービス　　© 学術情報センター

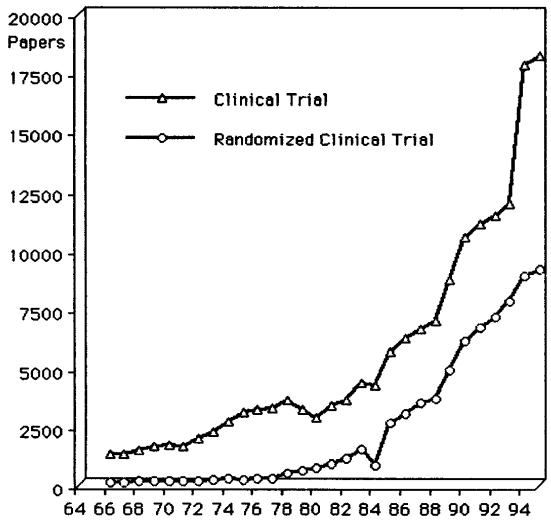


図 1　臨床試験と無作為化比較試験の文献の成長 (1966-1995)

表 1　発表言語による比較：全論文と RCT 論文 (1992-1994)

Language	All Papers	%(1)	RCT	%(2)	(2)-(1)
English	975352	85.2%	21752	91.4%	6.2%
German	28588	2.5%	625	2.6%	0.1%
French	28148	2.5%	322	1.4%	-1.1%
Japanese	26080	2.3%	150	0.6%	-1.7%
Russian	24387	2.1%	99	0.4%	-1.7%
Spanish	14632	1.3%	209	0.9%	-0.4%
Italian	9021	0.8%	223	0.9%	0.1%
Danish	3693	0.3%	96	0.4%	0.1%
Dutch	3327	0.3%	21	0.1%	-0.2%
Swedish	3102	0.3%	21	0.1%	-0.2%
Others	28840	2.5%	270	1.1%	-1.4%
Total	1145170	100%	23788	100%	0.0%

最近 3 年間を分析対象として、発表言語からみた分布をまとめた(表 1)。ここでは比較の意味で、Medline データベース全体の発表言語別の百分比をも示した。無作為化比較試験文献では、英語が 91.4 パーセントを占め、中心的な発表言語になっており、Medline データベース全体での比率(85.2 パーセント)よりも多くの文献が英語で発表されている。英語以外では、ドイツ語、イタリア語、デンマーク語の 3 言語による発表が、Medline 全体の発表言語よりも高い比率を示していたが、日本語、ロシア語、フランス語などは全体よりも低い値でしかない。日本語の文献は、Medline データベース全体では 2.3 パーセントを占めているが、無作為化比較試験文献では 0.6 パーセントにすぎない。Medline データベースへの日本誌の収載は約 120 誌でしかないが、日本医学会に属している学会誌や主要な医科大学雑誌が主に収録されていることを考えると、無作為化比較試験文献の少なさは問題であろう。

5.3　主要国における臨床試験文献の成長 (1986-1994)

次に、雑誌の発行国にもとづき、1986 年から 1994 年までの臨床試験の論文数変化を見てみよう。まず主要国であるアメリカ、イギリス、ドイツの 3 国と日本の変化を示した(図 2)。どの国の雑誌が、どれくらい臨床試験の文献を発行しているかを表わしている。

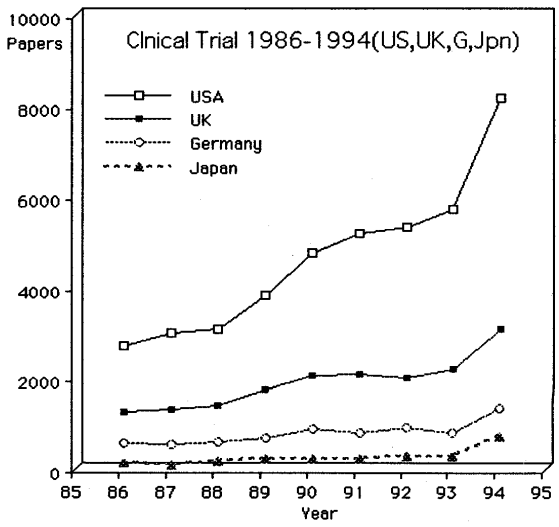


図 2　主要 4 ヶ国における臨床試験文献の成長 (1986-1994)

主要 3 国と日本についてみると、1989 年と 1990 年に臨床試験論文数が上昇に転じた。その後、1991 年から 1993 年頃まで、アメリカは上昇を持続させていたが、イギリスとドイツでは横ばい傾向となり、1994 年に再び急激な上昇を示した。日本は 1986 年に 101 件、1993 年に 255 件、そして 1994 年には 688 件と増加を示しているが、絶対数では全体的に低調のままである。日本は臨床試験論文の数と質の両面で、欧米に遅れていると批判されているが[13]、この点がこのデータにも示されている。

主要 3 国を除いた変化を、スイス、イタリア、デンマーク、オランダ、フランス、ロシアの 6 ケ国と日本について見ると、1994 年の日本の上昇が顕著である(図 3)。逆

学術情報センター電子図書館サービス　　© 学術情報センター

に、フランスは下降しており、スイスは、米国、イギリス、ドイツに次いで世界4位のランクを1986年から維持してきたが、1994年にはランクを落としている。日本、イタリア、オランダ、デンマークなどは、1994年に大幅な増加を示したが、スイスやフランスは減少に転じている。

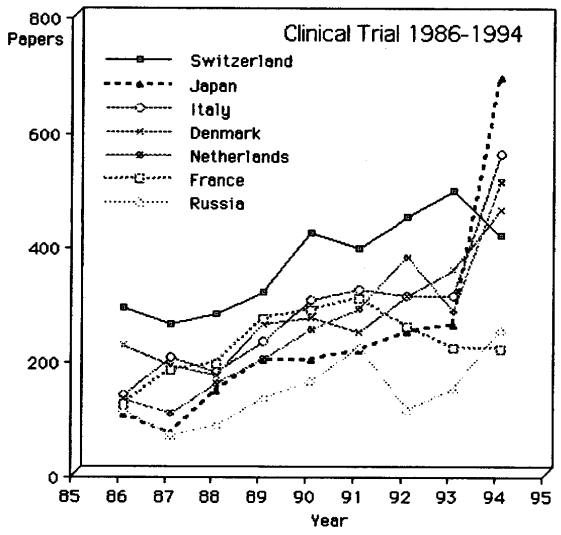


図3　7ヶ国における臨床試験文献の成長(1986-1994)

5.4 主要国における無作為化比較試験文献の成長(1986-1994)

まず、主要3国であるアメリカ、イギリス、ドイツと日本について図4にまとめた。アメリカの直線的な急上昇とイギリスの着実な増加が示されている。臨床試験全体の変化と比較すると、無作為化比較試験領域では、日本の増加がさらに目立たず、絶対数においても低迷している様子が顕著である。主要3国を除いた変化を見るために、図5にその他の6ケ国と日本の変化を示したが、ここにも、日本の低調さとフランスの下降が現われた。6ケ国のなかでは、デンマークの上昇傾向が明らかである。日本は、臨床試験論文では上昇傾向に転じていたが、無作為化比較試験文献では、日本の雑誌からの発表が増加していないばかりか、絶対数も低調なままであり、わが国がevidence based medicineへの取り組みにも遅れていることを表している。

5.5 全論文と、臨床試験、無作為化比較試験文献の出版国別百分比(1992-1994)

これにより各国の臨床試験と無作為化比較試験文献の生産への貢献度が見え、コクラン共同計画参加への活発さが示されるのではないだろうか。既述のとおり、

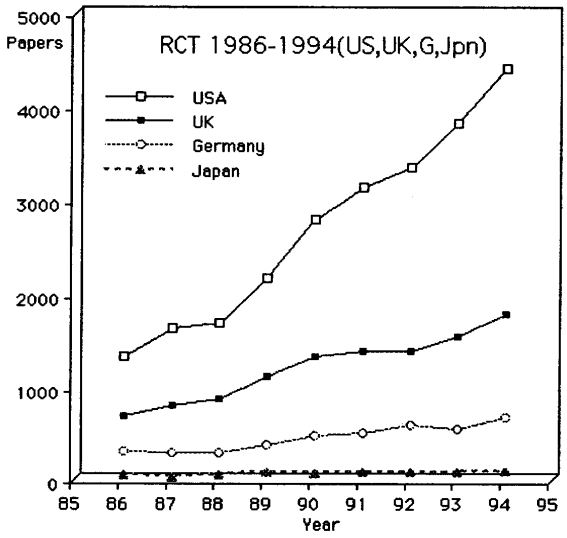


図4　主要4ヶ国における無作為化比較試験(RCT)文献の成長(1986-1994)

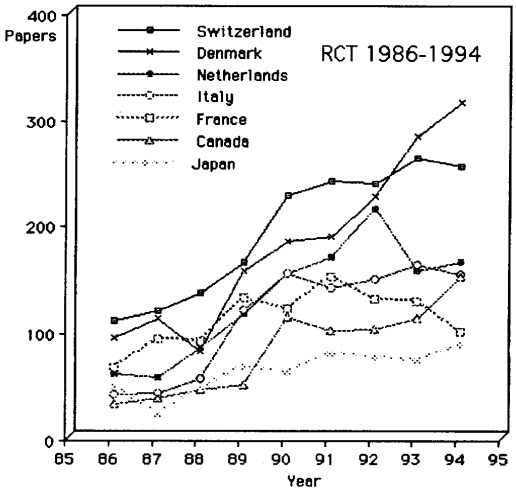


図5　7ヶ国における無作為化比較試験(RCT)文献の成長(1986-1994)

現在、コクランセンターは9ケ所あり、それらはアメリカ、イギリス、オーストラリア、カナダ、オランダ、イタリア、デンマークの7国に設立されている。すでに、発表言語による臨床試験文献の分布から、コクランセンターの所在する国の言語の貢献度が高いことがわかっている。そこで、Medlineデータベースに収載されている1992年から1994年の3年間のすべての記事を、主要な発行国にわけて占有率を計算し、これを、臨床試験(図6)と無作為化比較試験(図7)の文献の発行国別占有率と比較した。

学術情報センター電子図書館サービス　　©学術情報センター

アメリカ、イギリス、ドイツ、スイス、デンマーク、イタリアの5ケ国は、臨床試験と無作為化比較試験文

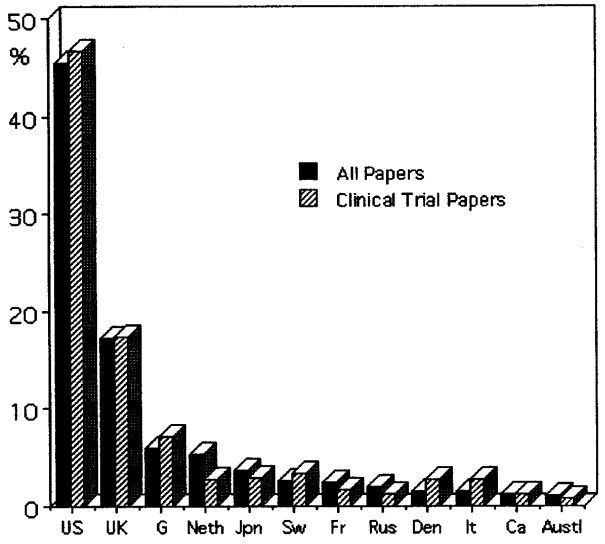


図6　臨床試験(CT)論文と全論文の出版国からみた百分比の比較(1992-1994 Medline DB)

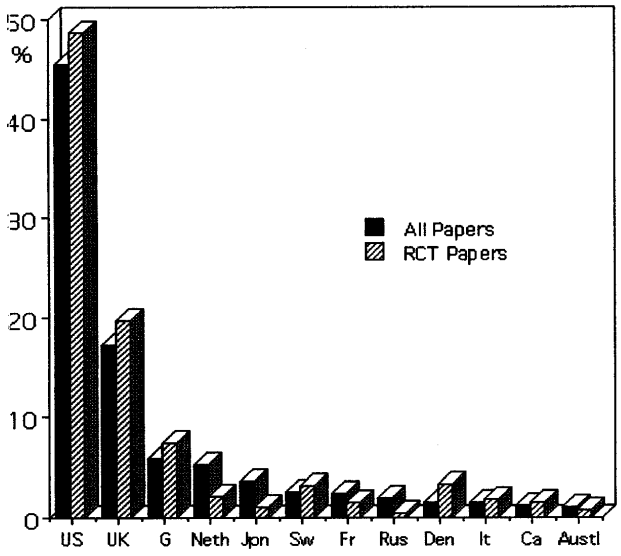


図7　無作為化比較試験(RCT)論文と全論文の出版国からみた百分比の比較
(1992-1994 Medline DB)

献の生産がともに活発で、全論文での占有率を上まわっており、薬学研究や製薬学における中心国であることが確認される。カナダは臨床試験文献では全論文と同じ比率であるが、無作為化比較試験文献では貢献している。この5ケ国のなかでは、デンマークの無作為化比較試験文献の高い比率が注目できるであろう。一方、オランダ、日本、フランス、ロシア、オーストラリアの国々は、臨床試験と無作為化比較試験文献の文献生産が低調で、既にみた発表文献の絶対数の動向だけでなく、データベースに占める比率からみても、主要国の活動状況との違いが示されている。

6　まとめ

Medlineデータベースへの臨床試験論文の収録は、データベース制作機関である米国国立医学図書館とコクランセンターとの協力により増大傾向にある。しかし、主要国の対応には違いがり、特に日本の貢献度は低い。日本は、近年、政府からの公的な科学研究費を増大させ、科学技術創造立国をめざしている。科学研究の成果としての論文生産も活発である[14]。しかし、臨床試験への取り組みには、まだ欧米に立ち遅れているといえよう。また、日本国内での医学領域のデータベースの編集制作状況をみると、Medlineデータベースで1991年に新設された倫理や臨床試験に関する出版タイプのようなフィールドが、十分には整備されていないのが現状である。印刷体とCD-ROM版で提供されている医学中央雑誌やオンラインデータベースであるJMEDICINEなどは、論文の質的な判断基準になる臨床試験、とくに無作為化比較試験の医療に占める重要性に配慮した編集・索引方式を検討すべきであろう。

日本では、1994年にJANCOC(Japanese Informal Networking for Cochrane Collaboration)が設立され、コクラン共同計画の普及とプロジェクトへの参加が検討されている[15]。しかし、わが国は、現在コクランセンターを設立していない先進国のひとつとなっている。コクラン共同計画は臨床の質を向上させることに役立つ国際協力によるデータベース事業でもあり、わが国においても、公的な資金の援助による試行と体制の整備をへて、相応のデータベース制作機関で事業化されるべきものである。こうして、医療に役立つ質の高い臨床試験の実施ともに、研究成果の円滑な流通と配付のシステムを確立していくことが重要であると考えられる。

本稿は、1996年12月1日に東京医科歯科大学で開催された、第1回コクラン・ハンドサーチ・ワークショップで発表した内容に加筆したものであり、発表の機会をいただいた津谷喜一郎助教授(東京医科歯科大学、JANCOC日本代表)に感謝いたします。また、本論の執筆にあたり、適切なご教示をいただきました根岸正光教授(文部省学術情報センター)に深く感謝いたします。

学術情報センター電子図書館サービス　　© 学術情報センター

参考文献

[1] Kassirer, J.P., "Learning medicine-too many books, too many journals", *New England Journal of Medicine*, Vol.326, pp.1427-1428, 1992.

[2] Colaianni, L.A., "Retraction, comment, and errata policies of the US National Library of Medicine", *Lancet*, Vol.340, pp.536-537, 1992.

[3] 津谷喜一郎,「コクラン共同計画とは」, 医学図書館, Vol.43, pp.119-120,1996.

[4] Chalmers, I., "The Cochrane Collaboration: prepairing, maintaining, and disseminating systematic reviews of the effects of health care", *Annals of the New York Academy of Sciences*, Vol.703, pp.156 -163, 1993.

[5] 別府宏圀,「情報の受け手としての医師」, 臨床と薬物治療,Vol.14, pp.583-587,1995.

[6] Bero, L.; Rennie, D., "The Cochrane Collaboration", *JAMA*, Vol.274, pp.1935-1938, 1995.

[7] Dickersin, K.; Scherer, R.; Lefebvre,C., "Identyfication of relevant studies for systematic reviews", *BMJ*, Vol.309, pp.1286-1289, 1994.

[8] McDonald, S.J.; Lefebvre, C.; Clarke,M.J., "Identyfying reports of controlled trials in the BMJ and the Lancet", *BMJ*, Vol.313, pp.1116-1117, 1996.

[9] 津谷喜一郎,「情報の質の確保」, 臨床と薬物治療, Vol.14, pp.593-597,1995.

[10] Fahey, Y.; Hyde, C.; Milne, R.; Thorogood, M., "The type and quality of randomized controlled trials (RCTs) published in UK public health journals", *Journal of Public Health Medicine*, Vol.17, pp.469-474, 1995.

[11] Altman, D. G., "Better reporting of randomised controlled trials:the CONSORT statement", *BMJ*, Vol.313, pp.570-571, 1996.

[12] Begg, C.; Cho, M.; Eastwood, S.; Horton, R.; Moher, D.; Olkin, I.; Pitkin, R. Rennie.; D. Schulz, K.F.; Simel, D.; Stroup, D.F., "Improving the quality of reporting of randomized controlled trials:the CONSORT statement", *JAMA*, Vol.276, pp.637-639, 1996.

[13] 西條長宏, 吉田茂昭,「臨床試験は社会に受け入れられているか?」, 癌と化学療法, Vol.23, pp.129-135, 1996.

[14] 根岸正光,「学術論文数の国際比較調査－結果の概要と分析視点」, 情報管理, Vol.39, pp.245-257, 1996.

[15] Tsutani, K., "The Cochrane Collaboration scene in Japan", *Cochrane News*, No.6, p.9, 1996.

学術情報センター電子図書館サービス
© 学術情報センター

NACSIS