

文献目録の機械編纂の技法と問題点

— 社会科学分野での現状と実験例 —*

松井幸子

I 目録機械編纂の現状

文献検索のための電子計算機利用は、文献目録の編纂、SDIサービス、RSサービス等に大別される。その実用的効果はSDIやRS等の機械検索においてより、主としてデータの配列替や計数からなりたつ目録編纂においての方が大きいと考えられる。あるいは少なくとも、目録編纂の技法がSDIやRSの前提として、不可欠であるとは言えよう。

この計算機による目録編纂の利点は、短時間に大量のデータを処理できる高速性、判断処理の機械的逐行による一貫性、正確性などにある⁽¹⁾。その結果、手作業の場合よりはるかに容易に、多種類の索引を作成できる。しかもそれには、分類、著者、書名、刊年順など手作業でも作成可能なもののほかに、KWIC、KWOC、引用文献索引などのように、計算機によらなければ実用化が困難なものがある。そして、それらが1950年代に開発されたことが、文献データの計算機処理を急速に普及させたと言える⁽²⁾。

* 本稿は昭和47年度文部省科学研究費試験研究(1)「社会科学における標題・抄録によるKWIC索引、コーディネート索引等の検索効率の比較実験」による研究成果の一部であり、研究会では共同研究者や参加者各位に種々御教示を頂いた。また計算機処理には北海道大学大型計算機センターを利用した。

(1) この点はC.D. Batty [31]に簡潔にまとめられている。ただし、主題書誌編纂はほとんど自然科学・技術領域に限定されており、社会科学領域では実験の域を出ないというのは、本書の出版年(1971年)を考えると疑問である。(p. 83参照)たとえば表1を見よ。

(2) たとえば、KWIC索引と同様に標題中のキーワードを索引語に使用したのとして、イギリスのマンチェスター公共図書館で19世紀に作成した蔵書目録があるという。また、わが国でも1950年代初頭に手作業でKWOC索引が作成されている。(木田橋喜代慎 [9])ただし、これらは長時間を要するため継続的刊行には至らなかったわけである。

そこで、目録機械編纂の現状をみるために、社会科学分野で機械編纂された遡及的文献目録（逐刊形態の索引誌を除く）⁽³⁾を表1に掲げた。同表で索引の種類が最も多いのは Pollin である。これはウィリアム・ゴドウィンに関する文献3,379点を収録し、あらかじめ1点毎に文献記号を与えて、抄録も作成してあり、その文献記号や抄録中のキーワード（人名・書名）を索引項目に加えることにより、11種の索引や統計表を作成している。そこで、種々の文献タイプや人名・書名（文献の著者名・書名だけでなく、その中で言及された人物や書名を含む）、出版年、使用言語などさまざまな角度から検索することができる。また記入点数を手がかりとして、研究動向を数量的に把握するよう⁽⁴⁾なこともできる。ただし Pollin は人名・書名検索に重点をおいたためか、主題検索用索引は作っていない。この点では K W I C 索引に加えて、Subject Index（分類順索引）も作成している Aldous & Hill と対照的である。なお一般に機械編纂されている索引のうち、chain index や引用文献索引などは表1に含まれていない。

計算機利用の大きな利点は自動編纂にあると言われているが、実際に作成された目録類についてみると、その前提として主題分析がすまされていることが多い。たとえば自動編纂を大きな特徴とする K W I C 索引においてさえ、その検索効率を高めるためには、標題へのキーワードの追加が必要とされている（Janda, Aldous & Hill）。また上記のように Pollin では全文献に抄録を作成している。これらは主題専門家が関与しなければ難かしい。⁽⁶⁾

そこで表1の目録編纂組織をみると、すべて主題研究者が中心となっている。文献採録には各国研究者の協力を得ているものもある（Aldous & Hill,

(3) たとえば, *Jurindex* [8].

(4) A (Article on Godwin), R (Review of a book by Godwin), P (Passage on Godwin), C (Short comments, interspersed), M (Mention of Godwin), B (Book solely on Godwin), N (Necrology of Godwin), Q (Quoted material from Godwin), W (Work dealing in part with Godwin), RW (Review of a work dealing with Godwin), GQ (Godwin quoted), RB (Review of a book on Godwin). B. R. Pollin [3], p. xiii.

(5) 藤川正信他 [7], pp. 1105-06.

(6) K. Janda [1], p. xviii.

表1 機械編纂された目録一覧

作成者* (出版年)	[1] Janda (1964)	[2] Aldous & Hill (1967)	[3] Pollin (1967)	[4] Gotterer (1970)	[5] Matsuda (1973)
Master List Name Approach Author Title Periodical Subject Approach Evaluation (Review) Chronological Approach Language Approach 索引の種類数 主題範囲 収録期間 収録文献 (A=Article) (B=Book) 文献採録方法 非英語文献の扱い**** 収録録点数 入力データ・カード枚数 判形 (マス・リスト) ページ数 ①目録 ②解説 計算センター 計 算 機 プログラム言語	Author Alphabetized Bibliography Author Cross-Reference Author List Periodicals List KWIC Index** (I.e. KWIC Index)	Complete Reference List Author List Periodicals List KWIC Index** Subject Index (分類)	Synoptic Bibliography / Authors of Books and Articles / Persons Mentioned / Important Studies of Godwin / Books Mentioned Names of Periodicals Annotated / Reviews of Books on Godwin / Reviews of Godwin's Books / Chronological Listing of All Entries / Writings of W. Godwin Language of the Original Sources 11 (他に統計3あり)	Bibliography Author Index Permuted KWIC Index	Master List Author Index Title Index KWIC Index Reviewer Index Chronological Index
	3	5	3	3	6
	政治学 1906-1963	家族・婚姻 1900-1964	ウイリアム・ゴドウィン 1783-1966	コンピュータ・マネジメント 1955-1969	(1) 計 量 (2) 主 義 原 典 (3) サ ン ・ ソ ン ・ ソ ン 1947-1969 1817-1971 1831-1970
	A	A, B	A, B	A, B	B B B
	現物	二次資料***	1. 現物 2. 二次資料**	n.a.	二次資料 1. 二次資料 2. 現物 2. 現物
	なし	タイトル英訳, 原語省略	原語, 難解語英訳つき	なし	Pollin に同じ
	2,614	12,850	3,379	900以上	864 455 231
	n.a.	50,000以上	約20,000	n.a.	8,520
	B 5, 1 段組	A 4 大, 2 段組	B 5, 1 段組	B 5, 1 段組	A 4, 2 段組
	① 225, ② 12	① 496, ② 9	① 659, ② 46	① 152, ② 2	① 145 ① 92 ① 49 ② 32 ② 32 ② 32
	Northwestern Univ. Computing Center	Univ. of Minnesota Data Processing Center	New York Univ. Heights Academic Computing Facility	Johns Hopkins Univ	北海道大学 天理計算機センター
	IBM 709	IBM 1401	IBM 360-30	n.a.	FACOM 230-60
	MAP, FORTRAN	n.a.	PL/I, FORTRAN	n.a.	PL/I

* 書誌的事項は稿末「引用・参考文献一覧」の [1]-[5] 参照。
** タイトルへのキーワードの追加あり。
*** 文献採録には各国研究者の協力を得ている。
**** 非英語文献の扱いは [1]-[5] すべて、英字は大文字のみ、発音符省略、日本語、露語などは翻字。

Pollin)。プログラミングは大学付属計算センターの技術的援助を受けているものが多い。入力データ作成（転記・せん孔）にはパートタイム学生の参加が目立っている。ただし、いずれの場合でもライブラリアンはほとんど関与していない。ところで、最近の MARC テープの出現や、計算機処理を前提とした国際的書誌記述標準化の動きなどを考慮すると、ライブラリアンが専門プログラマと完全に交代することは無理としても、少なくとも両者が主題専門家に協力するという体制をとることが必要な時期になっているように思われる。上記のような事態の推移を考えると、従来のプログラマによる文献データ処理では、必ずしも満足できるとは言えないからである。⁽⁷⁾

そこで本稿のⅡでは、従来プログラマに一任されてきたためか、ライブラリアンにとって理解しやすい叙述が少なかったと思われる、計算機による文字列処理の基本的技法や問題点を、目録の機械処理を考えているライブラリアンを念頭において考察する。Ⅲでは具体例として筆者がプログラミングを担当し、文字列処理に特有な技法の開発にも努めた文献目録編纂プロジェクトについて検討する。最後にⅣで今後に残されている問題点を考察する。⁽⁸⁾

Ⅱ 目録機械編纂の概要

1 システム設計の基準

(1) ハードウェアとソフトウェアの選定

(a) ハードウェア

ハードウェアの選定では、計算機の容量の問題とならんで、入出力機器のキャラクタ・セットが入力文献に使用されている文字をどの程度含むか、および含まれない文字をどのように表現するかの問題がある。たとえば表1の目録に収録されている文献では、英語文献のみの Janda, Gotterer を除くと、書名を英訳し原語は省略するもの (Aldous & Hill) と、原則

(7) F. G. Kilgour [32], p. 142.

(8) なお、入力前の主題分析や、作成された索引の利用上の問題などについては別稿の予定である。

として原語のままに難解語にのみ英訳をそえるもの (Pollin, 松田) がある。後者ではロシア語, 日本語は翻字と英訳を用いる。⁽⁹⁾そして, すべての目録で英字は大文字のみが使用され, 発音符も無視されている。⁽¹⁰⁾句読符号は Aldous & Hill や Gotterer ではかなり限定されている。これら使用字種は最終的には出力ライン・プリンタの字種に制約されるわけである。⁽¹¹⁾

(b) ソフトウェア

ソフトウェアに関しては, どの言語を使用するかが問題である。通常, 文字列処理には, 機械語に近いアセンブラ言語か, 高水準言語の COBOL, PL/I 等が用いられる。特に細かい操作が必要な場合を除き, 高水準言語が使いやすく, そのうち, PL/I は機種により使用できない場合もあるが, COBOL より豊富な機能を持つと言える。⁽¹²⁾

(2) 入力文献の選定と出力リストの種類

-
- (9) 非英語文献を英訳しているのは, 表1では松田を除くとすべてアメリカ中心で作成されているからであろう。ただし, われわれ非英語国民にとっては, 原語省略, 英訳のみというのは明らかに望ましくない。なお, このような異言語混配は印欧語では語頭がかなり一致するので, 検索上便利でさえあるが, 日本語の翻字分は語系が全く異なるので意味がない。松田 [5-1], p. 5.
- (10) 現在では英字の大文字と小文字を打ち分け, 発音符も扱える機種もある。(ただし, 国産機では特注しなければならないということである。注(4)参照) これは印字速度がおちるが, 読みやすさはます。たとえば, *Widener Library Shelflist* も当初は英大文字だけであったが, 1966年6月以降, 小文字や発音符も使えるようになった。(De Gennaro [11], p. 319) なお, MARC II テープのキャラクタ・セットについては [47] 参照。
- (11) プリンタに小文字があれば, カードせん孔機にはないとしても, せん孔時に特殊文字を付加して, 大文字小文字を識別しておけばよい。なお, 現在使用しているプリンタに小文字が備わっていないとしても, せん孔時に上記の区別をしておけば, 将来, 小文字の打てるプリンタが導入された時, 過去のデータを作りかえなくてすむ。R. M. Hayes and J. Becker [27], p. 623.
- (12) PL/I 入門書のうち, 文字列処理に重点をおいているのは, T. H. Mott, et al. [13] である。本書は全体にわたりサンプル・プログラムが豊富であるが, 機械検索の技法に関するものに大きな比重がおかれている反面, 目録編纂の技法はそう重視されていない。なお, 文字列処理に重要な DEFINED 属性に全く言及していないのは疑問である。COBOL については, 坂本徹朗 [19] が KWIC 索引作成を例として, かなり詳細にプログラミング技法を解説している。沢山のプログラム例を挿入しているのはよいが, あとわずかの紙幅の増加で完全なプログラムを提示できたはずである。この点が惜まれる。もっとも, そうした場合, マスタ・ファイルの作成・更新フローについて, 現在より明確に言及しなければならぬだろう。

(a) 入力文献の選定

現在のところ光学式読取装置が未発達なため、入力文献をそのままでは計算機で使用できない。このため入力の前処理で人手に頼るので、多大の時間と経費を要し、しかもエラー・データの発生が不可避となっている。そこで近年、多くの機関で重複して機械可読入力データを作成するのではなく、ある機関で作成したものの磁気テープ・コピーを、その他の機関でマスタ・テープとして利用する方式も採用されるようになってきた。このテープが市販磁気テープと呼ばれている⁽¹³⁾。そこで入力文献の選定にあたっては、これらテープや既存の文献目録類の有無や品質を考慮して、それらと競合しない文献であることを確認すべきである⁽¹⁴⁾。

(b) 出力リストの種類

出力リストの種類は、原則として完全な書誌的事項を含めるマスタ・リストと、各種索引リストである。後者は原理的にはマスタ・リスト上のどの項目についても作成可能であるが、通常は、著者、書名、刊年、分類、件名、KWIC、KWOC、引用文献、chain indexなどが作成されている。

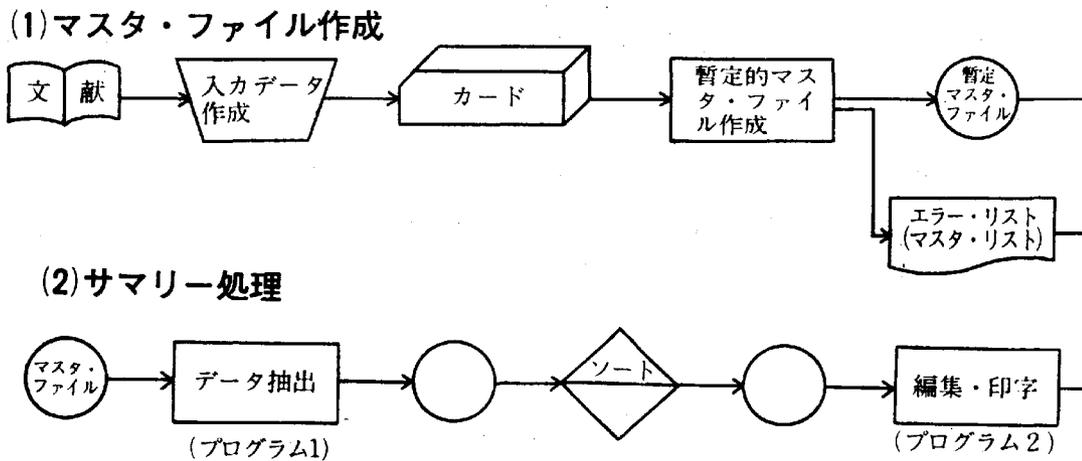
(3) プログラムの構成

計算機による文献データ処理は、一般に、マスタ・ファイルの作成とサマリー処理とに分かれる。前者は文献データをファイルに格納するまでの作業であり、後者はそれを用いて種々の索引を作成する作業である。このように分けるのは次の理由による。人手によって作成された機械可読データには種々のエラーが含まれており、それをカードのままで訂正することはごく少量

(13) 菊池敏典 [50] ほか参照。

(14) ただし、現在のところ遡及的に文献を収録しているテープはほとんどないと言える。たとえば MARC II テープの場合、過去にさかのぼって入力する RECON プロジェクトをもってしても、現在 1968 年以降の文献が入力されているにすぎない。(吉川藤一他 [48], p. 142) そこで遡及的書誌については、表 1 に示したような比較的せまい分野について、新たに作成されているのが現状である。

図1 目録機械編纂の



の場合を除いて不可能である⁽¹⁵⁾。そこで、まず、種々のエラーを含むであろうカードを入力して暫定的マスタ・ファイルを作り、その内容をチェックして、発見したエラーを訂正カードにより訂正し、漸次、正しいマスタ・ファイルとする。そして完成したマスタ・ファイルから、各種索引を編集するという手順をとるわけである。(図1参照)

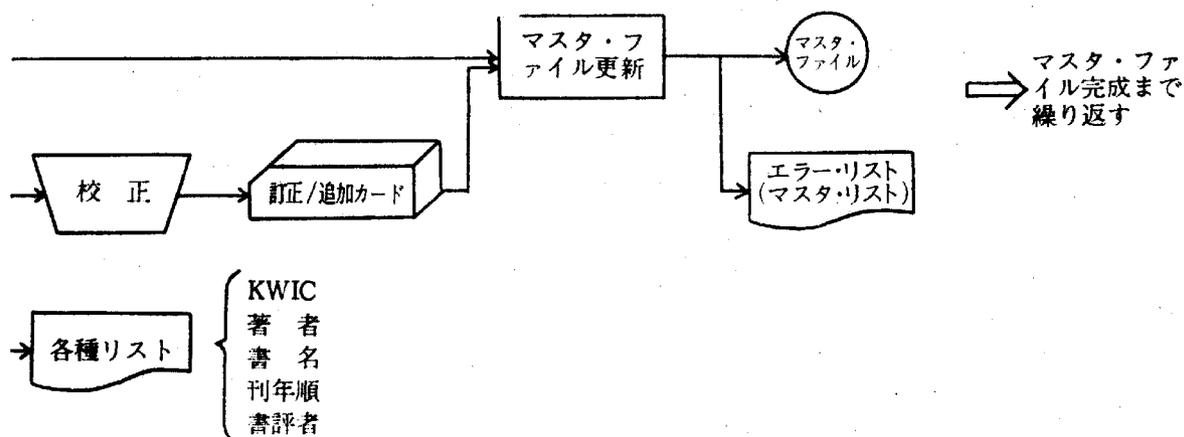
マスタ・ファイル作成作業における3つの計算機処理は、実際には1つのプログラムで行なうことが多い。これは、たとえばマスタ・ファイル更新における新文献の追加機能が暫定的マスタ・ファイル作成機能と同じであり、またファイル内容の出力は完成したマスタ・ファイルに対してだけでなく、エラー・チェック用にも必要だからである。次のサマリー処理では、データの抽出、配列替、印字の3本のプログラムとすることが多い。(詳細はⅡ3参照)

(4) 入出力フォーマットの設計

文献データ処理では、入出力フォーマットの設計が、科学技術計算の場合とは比較にならないほど重要である。設計にあたっては、出力リストに必要な情報、リストの読みやすさ、データせん孔作業の容易さ、プログラミング

(15) たとえば、たった1字のせん孔もれのために、80桁全部せん孔し直さなければならぬこともあり、それに伴うカードのさしかえも容易な作業ではない。

ジェネラル・フローチャート



技法などを考慮する必要がある。なお、文献情報機械処理の普及に伴い、国際的書誌記述標準化の動向に対する配慮も望まれている。⁽¹⁶⁾

(a) 入力カード・フォーマット

1枚のカードにせん孔できる字数は80字であり、これは文献データの長さとは無関係である。そこで入力カード・フォーマットは数項目を適当に組合せて1枚としたり、あるいは1項目を数枚のカードに続けたりする。前者では、無理に80欄全部を使用せず、適当に余白を残した方が、せん孔ミスの予防や発見、訂正に適している。⁽¹⁷⁾

(b) マスタ・ファイル・フォーマット

機械処理では1つのマスタ・ファイルからさまざまなリストを作成できる。そのフォーマットは最初から決定しておかなくてもよいが、マスタ・ファイルにはそれらを作成するのに必要十分な情報が含まれていなければならない。しかしながら、将来の利用を見込んで、余りに多くの情報を入れることは、不経済であるのみか、不正確な情報を作り出すことになり有効ではない。⁽¹⁸⁾ただし、マスタ・ファイル・フォーマットに、将来のための

(16) 「特集・ISBDをめぐって」[38]参照。

(17) 裏田和夫 [76], p. 154.

(18) なぜなら、現在すぐに必要としない項目は、往々適当なチェックしかされないため、実際に使おうとする時、誤りが多く使いものにならないことが多いからである。

予備スペースを残しておくことは否定されない。⁽¹⁹⁾ なお、テープ・フォーマットでは固定長と半固定長の2形式が一般に採用されている。1巻のテープにはかなり大量のデータを格納できるので、文献量がそう多くなければ、固定長形式の方がプログラムが簡単である。⁽²⁰⁾

(c) 出力リスト・フォーマット

計算機処理目録の利点の1つは、出力リストをそのまま印刷用原版として校正その他の手間を省けるところにある。従って出力リスト1ページのフォーマットは、縮尺して最も読みやすい大きさで、同時に経済性の面から1ページに収録する文献数ができるだけ多いことが必要である。たとえば表1のマスタ・リストでは、A4判2段組かB5判1段組が多く、収録点数では2段組の方がすぐれているが、プログラミング上、1段組よりはるかに複雑になる。(II 3(4)参照) また1文献の出力フォーマットは、マスタ・リストでは原則としてマスタ・ファイルの全内容を印字するが、その他の索引ではマスタ・リストへ案内する文献番号、参照コード(KWIC索引の場合)以外の項目を簡単化または省略することが多い。(表1では松田を除くと、ほとんど文献番号のみの記入である。) この場合、2回検索しないと完全な情報が得られないが、記入スペースは非常に節約になる。索引の使いやすさと経済性を勘案して決定すべき問題である。

2 マスタ・ファイルの作成

(1) 作業の概要

マスタ・ファイルは、II 1(3)で述べたように、人手による機械可読デー

(19) これは項目追加などのフォーマット変更時に、全プログラムに変更が及ぶのを避けるためである。半沢孝雄 [16], p. 170.

(20) たとえば、1,600 BPI, 2,400 フィートの磁気テープ1巻には50,000点以上記録できる。(丸山昭二郎 [46], p. 104) ただし、大量データでは可変長あるいは半固定長形式が採用される。たとえば、国立国会図書館編『欧文逐次刊行物所蔵目録』の場合、各項目の長さをそのうちの最長のものに合わせて固定長としたところ、16,000点の記録に556 BPI, 2,400 フィートの磁気テープで6巻に及んだ。これを半固定長形式に直したところ、1,600 BPI, 2,400 フィートのテープ1巻に十分収まったと言う。(小田泰正 [53], p. 4-6.) なお、可変長および半固定長形式については、中井浩他 [29], p. 81-84 に詳しい。

タの作成， 計算機による 暫定的マスタ・ファイル作成， 人手および 機械による内容チェック， それにもとづいて作成した訂正データによるファイルの更新の繰り返しで作成される。完成したマスタ・ファイル内容の出力をマスタ・リストと呼ぶ。

(2) 入力データ作成

はじめに入力文献を計算機処理用に編集する。たとえば表記法の統一や標題へのキーワードの追加などである。次にワーク・シートへの転記を行なう。一般に，原データを直接カードへせん孔することが困難だからである。この時，正確に転記しなければならないのは無論であるが，特にカードせん孔時に見間違いやすい文字については，あらかじめ，せん孔者側と打ち合わせを行ない，きちんと書き分ける必要がある。⁽²¹⁾なお，作業の統一をはかるため「ワーク・シート記入要項」を作成し，それを遵守することがきわめて重要である。この後，カードのせん孔，および検孔を行なう。なお，カード・イメージ・リストでの校正は後述のように得策ではないので，プログラミング上，どうしてもチェックが必要な項目に限定すべきである。⁽²²⁾

(3) 暫定的マスタ・ファイル作成フロー

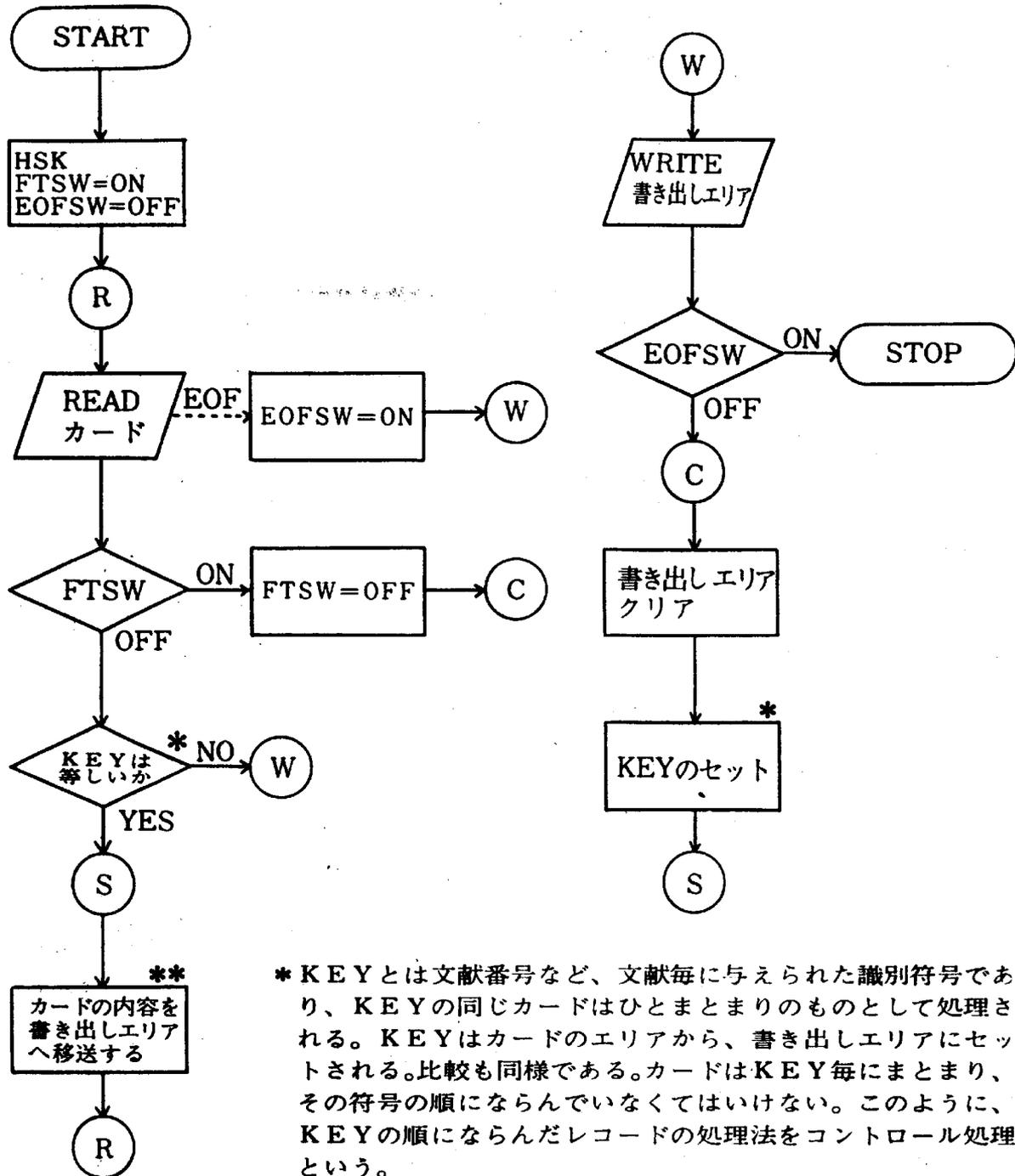
文献データでは，通常，入力カード・フォーマットが多種にのぼる。そこで，同一文献のカードを全部計算機へ読み込み，次の文献のカードがきた時，その前の文献のデータを一度にマスタ・ファイルへ書き出すというフローをとる。これは計算機に特有な処理フローの1つで，コントロール・フロー⁽²³⁾と呼ばれる重要なものである。(図2参照)

(21) 半沢孝雄 [16], 「見間違いやすい文字」 p. 122.

(22) 小松原康敏他 [43], pp. 316-17.

(23) フローチャートを完全に作り，入出力命令に習熟しておくことが必要である。なお，半沢孝雄 [16], 「第6章コントロール・ブレイク」 pp. 115-29 および「第10章2枚のカードにおたるレコードの処理」 pp. 213-21 参照。ただし，事務計算を想定しているので，そのままでは使えない部分もある。(例：入力カード枚数は固定されている)。

図2 コントロール・フロー



* KEYとは文献番号など、文献毎に与えられた識別符号であり、KEYの同じカードはひとまとまりのものとして処理される。KEYはカードのエリアから、書き出しエリアにセットされる。比較も同様である。カードはKEY毎にまとまり、その符号の順にならなくてはいけない。このように、KEYの順にならんだレコードの処理法をコントロール処理という。

** カードの内容を書き出しエリア（マスタ・ファイル・フォーマット）にセットするのは、1つのルーチンではできない。カードの区分によって、そのフォーマットが異なるからである。従って、このルーチンはカードにせん孔されている区分コードによって、各々の処理に分解するのが普通である。

(4) 内容チェック

前述のように、入力データ作成ではエラー・データの発生が不可避的である。そのためエラーをいかに効率的に発見し、訂正するかが、文献処理における最も重要な問題の1つとなる。以下にそのチェックについて述べる。

エラー・データは人手の介入するすべての時点で起こり得るが、その種類⁽²⁴⁾によって、計算機で容易に発見できるものとできないもの、人間にとって容易なもの⁽²⁵⁾とそうでないものがある。従って、プログラミングの難易度やデータ量を勘案して、人間と計算機がそれぞれ有利なチェックを分担すればよい。機械によって発見可能なエラーは全体の50—60%であるという報告もあるが、計算機で処理可能なチェックの種類としては次のものがあげられよう⁽²⁶⁾。

① 個別チェック (項目内のチェック)

スペル・チェック⁽²⁸⁾、コード・チェック、数値チェック、上下限チェック (例: $1832 \leq \text{出版年} \leq 1973$)

② 更新時チェック

更新すべきマスタ・レコードの有無チェック、「スペル訂正」(後述)での指定スペル有無チェック

(24) 小松原康敏他 [43] は日本での欧文文献処理時のエラー発生⁽²⁴⁾の報告である。アメリカのハーバードおよびスタンフォード大学図書館での例については、J. L. Dolby et al. [26], pp. 75-83 参照。

(25) たとえば、使用言語欄に定められたコード以外のものがあれば、計算機により容易に発見できるが、定められたコードの1つであれば、それが正しいかどうか判別することは簡単ではない。一方、人間ならば書名から使用言語を判別することは通常容易である。(難解語の判別については、溝口歌子 [42] 参照。) 逆にブランク1個のところを、あやまって2個にした場合など、人間には発見しにくい⁽²⁵⁾が、機械ならば容易である。

(26) J. L. Dolby et al. [26] p. 82.

(27) 事務計算を対象としたものではあるが、半沢孝雄 [16], 「第3章シーケンス・チェック」 pp. 85-95, 「第4章重複のチェック」 pp. 97-103, 「第14章エラー・レコードの処理」 pp. 265-80 参照。(事務計算の場合なので、コード・チェック、数値チェックが主で、スペル・チェックなどは問題外である)。

(28) 日笠衆司 [44] には、ディスクリプターのスペルの自動修正手法が紹介されている。ただし、著者名その他の固有名詞については自動修正が難しい。J. L. Dolby et al. [26], pp. 78-80.

③ 関連チェック

トータル・チェック，過不足チェック，シーケンス・チェック（磁気テープ），項目間の関連チェック（データ全体として妥当性を持つかどうか）

なお計算機によるチェック結果の出力，すなわちエラー・メッセージ・リストのフォーマットは，読みやすく，すぐに理由が分かり，できればリストを見ただけで訂正カードが作成できるとよい。人間による内容チェックのためには，特に校正用にフォーマットを工夫したリストを作成してもよいが，アルファベット順ソートのすんだ各種出力リストを利用することもできる。⁽²⁹⁾たとえば著者索引では著者名を，KWIC索引ではキーワードを集中的に校正する。この時，特に出現頻度の少ないものに注目すると，スペル・ミスが発見しやすいという利点があり，カード・イメージ・リストによる校正よりもずっと能率的である。なお，マスタ・リストにはファイル内容のほとんどすべてが印字されるので，上記の各種リストでチェックできない項目は，マスタ・リストで一挙にチェックするのがよい。

なお機械および人間による内容チェックは，時間と経費をかければかけるほど向上すると言えるが，無制限に作業を続けることはできない。どれほどの経費でどれほどのエラーを許容するかと言う考え方も必要であろう。⁽³⁰⁾

(5) 訂正データ・カード作成

チェックにより発見されたエラーに対して，訂正用入力カードを作成する。文献データでは長い文字列からなる項目が多いので，間違っただけせん孔する特別形式が望ましい。（次項の「スペル訂正」参照。）

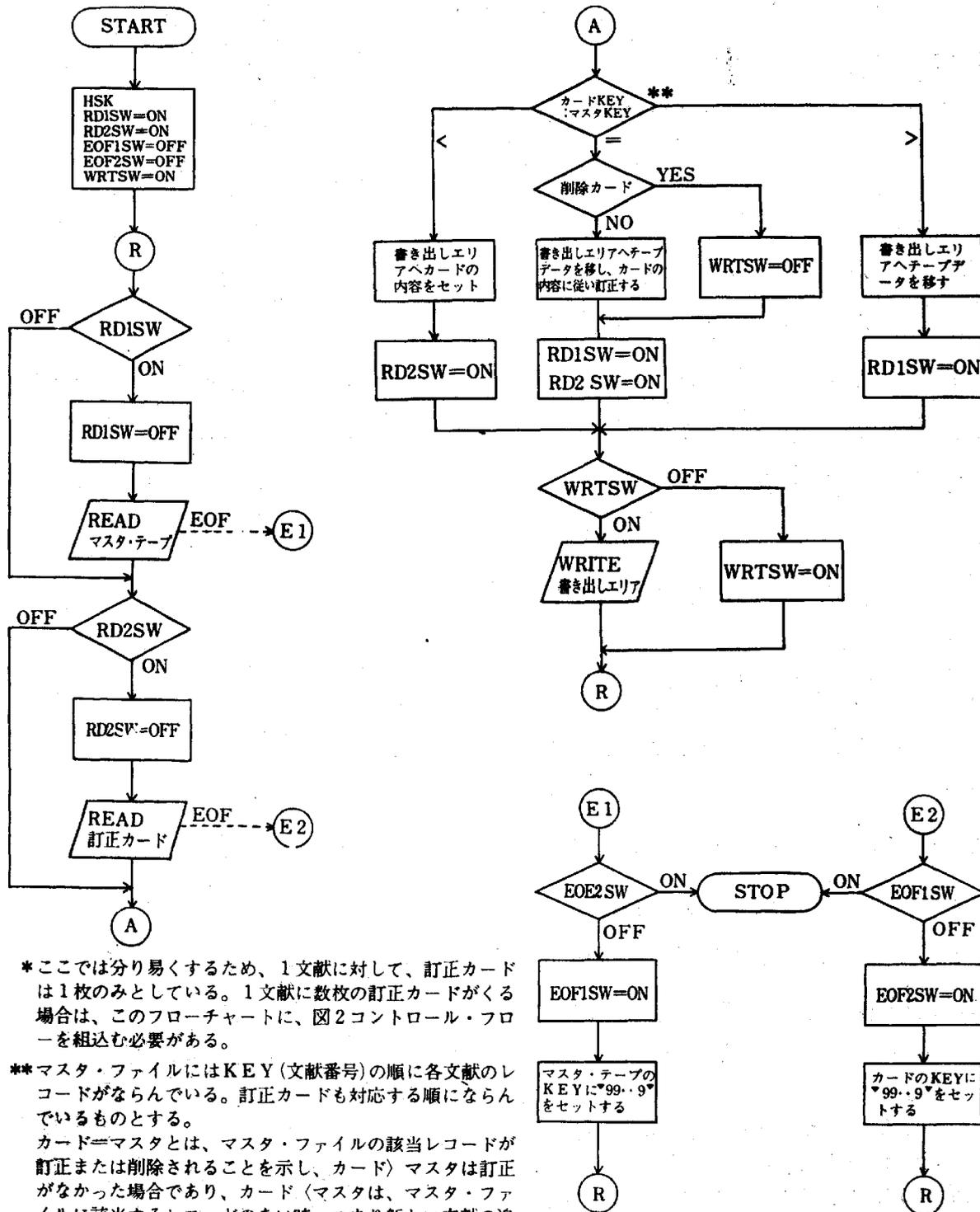
(29) たとえば，表1の Pollin では校正用に4種のフォーマットの異なるリストを作成したが，最終出力フォーマットのもの（各種索引リスト）も非常に校正に役立つという。B. R. Pollin [3], pp. xli-xlii.

(30) R. M. Hayes and J. Becker [27], p. 204.

(6) マスタ・ファイル更新フロー

新しい文献データの追加・挿入と、すでに入力してあるデータの置換え、削除、訂正作業をマスタ・ファイルの更新と言う。新文献の追加・挿入は、

図3 更新フロー*



*ここでは分かり易くするため、1文献に対して、訂正カードは1枚のみとしている。1文献に数枚の訂正カードがくる場合は、このフローチャートに、図2コントロール・フローを組込む必要がある。

**マスタ・ファイルにはKEY(文献番号)の順に各文献のレコードがならんでいる。訂正カードも対応する順にならんでいるものとする。
カード=マスタとは、マスタ・ファイルの該当レコードが訂正または削除されることを示し、カード<マスタは訂正がなかった場合であり、カード<マスタは、マスタ・ファイルに該当するレコードのない時、つまり新しい文献の追加であることを示す。

人手が足らず一度に入力データを作成できない時や、継続的にデータを追加していく時に有効なのである。文献データでは、一般に文献番号をキーとしてマスタ・レコードと更新データを関連づける。磁気テープ・ファイルでは、ファイルとカードに同じシーケンスを持たせマッチング処理をする。暫定マスタ作成後の計算機処理は、すべて更新作業の繰り返しとなるから、非常に重要なルーチンである。⁽³¹⁾ (図3参照)

次に、Ⅲで検討する目録編纂プロジェクトでの更新作業で最も多く用いた「スペル訂正」方式を、文字データ訂正の具体例として述べておく。

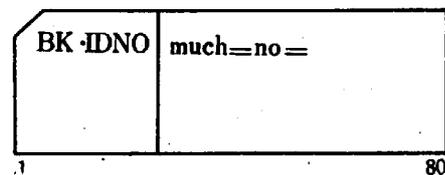
すなわち、文献データは一般に長い文字列からなっているから、転記やせん孔の際のスペル・ミスは避けることができないことは前述のとおりである。ところでその訂正において、カード1枚分の文字列全体をオリジナル・フォームで再作成するのは、作業量からも再度のエラーの危険性からも明らかに得策ではない。そこで誤まったスペルと正しいスペルのみを指示して訂正する「スペル訂正」方式を採用したわけである。⁽³²⁾

例：(誤) I HAVE MUCH MONEY.

↓
(正) I HAVE NO MONEY.

上記の例文を訂正する際のデータ・ **図4** スペル訂正カード・フォーマット

カード・フォーマットは図4のとおりであり、処理の流れは以下のようになる。①ワーク・エリアに、もとのデータ・エリアの訂正される文字列より前



の部分(▽I△HAVE△▽)を移送する。②それに続けてワーク・エリアに、訂正後の正しい文字列(▽NO▽)を移送する。③さらに続けてワーク・エリアに、もとのデータ・エリアの訂正される文字列の後の部分(▽△MO

(31) C. T. Meadow [28], 「第10章ファイルのメンテナンス」 pp. 335-47, および半沢孝雄 [16], 「第11章ファイル・メンテナンス」 pp. 223-36 参照.

(32) このようにスペルをキーとして訂正するのではなく、スペルの位置をキーとしてデータを置換える方法もある。塩田卓和他 [45] 参照.

NEY.▼) を移送する。④最後にもとのデータ・エリアに、ワーク・エリアの内容 (▼I△HAVE△NO△MONEY.▼) を戻す。「スペル訂正」方式は図4のように、訂正される文字列と訂正する文字列だけを入力すればよく、両者の文字数が異なることも許されるので、比較的長い項目（著者名や書名）のスペルの部分的誤りの訂正において、きわめて便利な方法⁽³³⁾と思う。

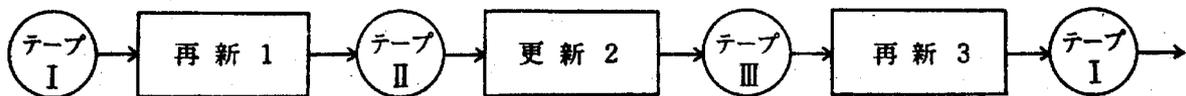
(7) マスタ・リスト作成

完成したマスタ・ファイルの内容はリストとして印字する。その技法はサマリー処理と同一になるので、Ⅱ 3 参照。（ただしファイル内容と同一のシーケンスで出力するのが普通なので、ソートは含まない。）

(8) マスタ・ファイルの保守

暫定的マスタ・ファイル作成後は更新作業の繰り返しとなる。そこで事故（例：I/O エラー，誤演算，紛失など）に備えて、3本のテープを繰り返して使用することが望ましい。（図5参照）たとえば、図5の更新3で入力テープⅢの途中で I/O エラーが発生し読めなくなった時は、すでにテープⅠは途中まで書き出されていて使えないので、テープⅡを入力して更新2を繰り返しテープⅢを再作成する。このためテープを3本使用することになる。データ・カードも1回前のものは必ず保存しておき、事故の起こった場合、再計算が可能ないようにしておく。

図5 マスタ・テープの保守



(33) プログラムは PL/I によれば、INDEX 組込み関数（文字列の中からある指定された文字の入っている位置を見出す）や SUBSTR 組込み関数（文字列の中から部分文字列を取り出す）などを用いることにより、わずか12ステップですんだ。この例は PL/I が文字列処理にきわめて便利なことを示している。

3 サマリー処理

(1) 作業の概要

マスタ・ファイルからの索引作成をサマリー処理と呼ぶ。原理的にはすべてのデータが索引対象となり得るが、よく用いられるのは著者名、書名、出版年、分類番号、件名、キーワード、引用文献などである。なお、該当データのすべてを対象とするか、あるいはさらに選択を加えるか（例：KWIC索引でのキーワードの選択）にも分かれる。また各種索引はその目的から言って、一定のシーケンスを持たせなければならないし、前述のように読みやすいフォーマットで印字しなければならない。これらの作業は、通常、抽出、ソート、編集・印字の3本のプログラムで処理される。

(2) データ抽出プログラム

データ抽出は必ずしもマスタ・ファイルの1文献について1つではない。また同一データ項目のすべてを抽出するのではなく、一定の選択を行なうこともある。各索引の目的に応じて、必要なデータを選択し、抽出するわけである。このためのプログラミング技法として、テーブル・サーチその他がある。⁽³⁴⁾

(3) ソート・プログラム

出力リストの配列順になるようデータをソートする。データ量がごく少ない時はコア・ソート⁽³⁵⁾が効率的でよいが、一般にはソート・ユティリティを用いる。なお、計算機によるソートでは文字の内部コード（例：EBCDIC、III 1 (2) (c) 参照）の順に照合されるため、空白も含めて特殊文字も無視され

(34) テーブル・サーチの方法としては、バイナリー・サーチ、シーケンス・サーチ等がある。テーブルの大きさは、原則として、一度にコアにおさまる方がよい。半沢孝雄 [16], 「第1章テーブル処理」 pp. 23-66 参照。なお、KWIC索引作成でのキーワード選択のためのテーブル・サーチの例として、拙稿 [18], pp. 19-20 参照。また、テーブル処理でなく、選択すべき文字列の前後をあらかじめ特殊文字で囲んでおき、それをタグとして抽出する方法もある。B. R. Pollin [3], p. xl.

(35) 半沢孝雄 [16], 「第2章コア・ソート」 pp. 67-84, R. P. Rich [17] 参照。

ず、厳密な字順配列となることに注意する必要がある。もちろん、少量のデータならばそのままでも検索上そう支障がないが、大量になる場合は慣用の配列（通常、英字の部分だけで配列順序を定める）に近づけるための手法を組込むことが望ましい。⁽³⁶⁾

(4) 編集・印字プログラム

マスタ・ファイルから抽出したデータは、各索引リストのフォーマットに合わせて編集する。この場合、1文献のフォーマットと目録1ページのフォーマットの2つの編集があり、前者については、上述のソートの前に行なうことが多い。これらの編集では、先に出力リスト・フォーマットの設計の項で述べた諸条件を満たすように配慮する。もちろん、そのための技法は、個々のリストのフォーマットにより大きく異なるが、ここでは、リストの読みやすさと経済性を高めるために重要な技法の例として、マスタ・リスト作成における2段組印字と行末処理について述べておく。

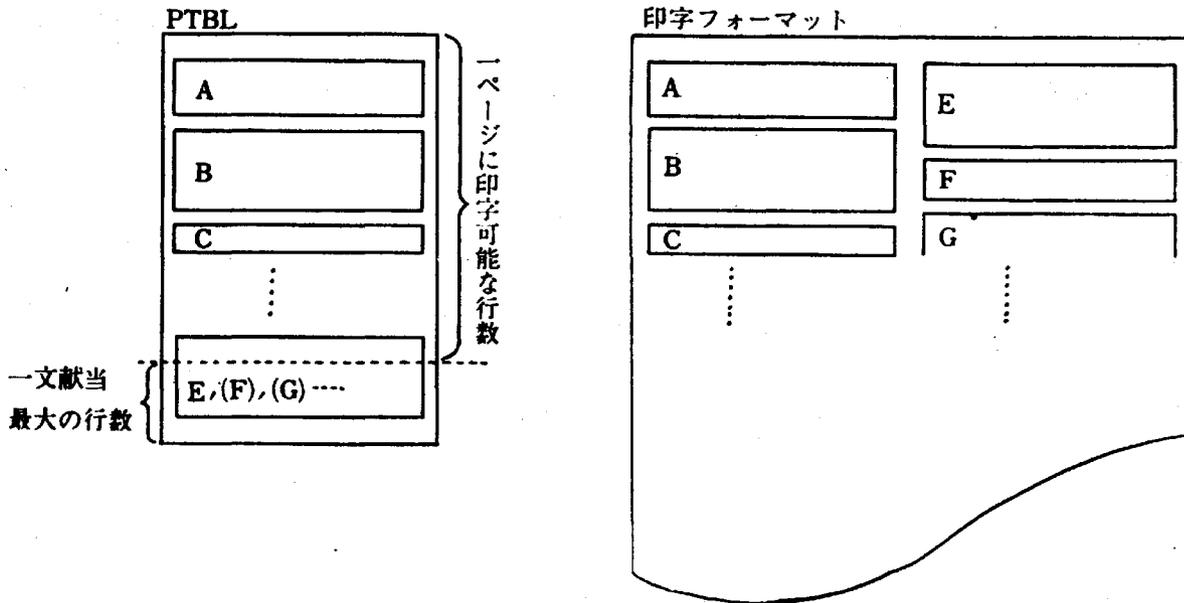
(a) 2段組印字ルーチン

これは、たとえば表1のAldous & Hill および松田のマスタ・リスト印字に用いられている方式で、1ページを左右2段に分けて使用するものである。読みやすくするために、段毎に文献単位で記入し、1文献の記入が2段にわたらないことが必要である。このため各段のプリント行数は一定数（例：A4判で100行、松田の場合）以下ではあるが、必ずしも段毎に一定とはならない。

処理の流れは、表1の松田では次のようにした。（図6参照）①各データ行を、順次、PTBLに貯える。（文献A, B, …）②そのうち、1ページの行数に入り切らないもの（文献E）が現われる。③この時、文献Eと、以前に貯えておいた文献A, B, …とを横にならべて書き出す。④文献Eの全行を書き出し終わったら、PTBLの文献Eを貯えていた場所

(36) 西村恕彦 [40] 参照。なお、図書館の配列規則と機械コードとの調整については、R. M. Hayes and J. Becker [27], pp. 610-16, および成井恵子 [61] 参照。

図6 二段組印字



に、次の文献Fを貯える。⑤そして③にもどる。⑥こうして、1ページ分いっぱいになるまで繰り返す。

上記の流れはプログラミングすると、かなり複雑なフローとなるので、場合によっては1段分だけ印字し、後ではり合わせてもよい。

(b) 行末処理

同一項目のデータが2行以上にわたる時(例:書名)、単語の途中での行替の必要が起きた場合、音節と無関係に行替をしてしまうと非常に読みにくい。これを避けるために、より読みやすい行末処理が必要となる。ここには、従来行なわれてきた行末処理の諸方法およびⅢで検討するプロジェクトで開発した「ヘア・スペース方式」について述べる。

① ⁽³⁷⁾ ブランク方式

最も一般的に使われている方法で、行末が単語の終わりにならない時、その単語全体を次行へ移し、残りの桁にはブランクをつめる。これは行末位置がそろわないので読みにくい。

(37) 表1の松田を除くすべてがこの方式である。

② 事前処理方式⁽³⁸⁾

データ・カード1枚を印字1行分にあて、分節をあらかじめ人手によってすませておく。これは分節に人手を要することと、カード1枚の桁数と印字1行の桁数が必ずしも一致せず、カード使用法が経済的でないという欠点がある。

③ 3字方式⁽³⁹⁾

行末の単語が3字以内の時は①と同じ処理。4字以上の時は、その行に印字できる字数-1字までを印字し、最後にハイフンを入れ、残りを次行へ移す。これは音節の切れ目と無関係になって読みにくくなることが多い。

④ 3・5・7字方式

3・5・7…字と奇数字目までを入れ、ハイフンをつけ、残りを次行へ移す。これは統計的に奇数字目が音節の切れ目になることが多いのに着眼した方法という。ただし、例外も多い。(例: clas・si・fi・ca・tion)

⑤ 辞書方式

計算機内に辞書を用意し、それを参照して分節する。ただし、完全な辞書を作ることは大変である。

⑥ ヘア・スペース方式

上記諸方法にはいずれも難点が認められたので、今回のプロジェクトでは新たに「ヘア・スペース方式」を開発した。これは、①ブランク方式で行末に残されたブランクを1つずつ、その行の単語の切れ目にヘア・スペースとして配分し、行末の印字位置をそろえる方法である。

処理手順としては、あらかじめ単語の切れ目のデリミターのタイプを分類しておき(表2参照)、①1行に入り切らない項目の文字列から1字ずつ取り出して、デリミターであるかどうか調べる。②デリミター

(38) たとえば、アジア経済研究所 [12].

(39) たとえば、吉川藤一他 [48] で報告されている MARC II テープから打出されるカードがこの方法で印字されている。読みにくい場合も多いが、カードでは1行の字数が限られるので、一つの便法と言えるかも知れない。

表2 デリミターの分類

区 分	デ リ ミ タ ー	備 考
左側の単語につくもの	(1) , . ; :) (2) - ? ! =	右側の単語との間にヘア・スペースをおく
右側の単語につくもの	(左側の単語との間にヘア・スペースをおく
どちらの単語にもつかないもの	空白	この位置にヘア・スペースをおく
デリミターとならないもの	英数字, ▼(引用符), ▼▼(複引用符)	引用符(左向きのみ)を両側に用いるためここに分類した。

(1) 「英字、数字」(例: No. 7) の場合、全体を1単語として扱い、ピリオドのあとにヘア・スペースを与えない。行分けもしない。

(2) 行分けのデリミターにはなるが、ヘア・スペースは与えない。

の時、そのタイプを分類して、単語識別テーブルを作る。③1行の区切りの位置を見つけ、ヘア・スペースの桁数を求める。④デリミターの分類に従って、ヘア・スペースを与える。

ヘア・スペース方式は、PL/Iでも105ステップを要し、かなり複雑なルーチンとなった。この方法は、今回のマスタ・リストが2段組のため、1行当61文字までと比較的短かかったこともあって、たとえばドイツ語の長い単語の語尾2~3字が入り切らずに次行へ移された場合など、ブランク(ヘア・スペース)がかなり多くなるという欠点を生じたが、全般的には読みやすい印字方式であったと思う。

現在の段階では、本章で概観した目録作成において、統一的技法が確定しているとは言いがたい。たとえば主題により便利なフォーマットも異なってくるはずだからである。

Ⅲ 目録機械編纂の実例

1 編纂プロジェクトの概要

(1) 組織と分担⁽⁴⁰⁾

(40) 詳細は松田[5]の各の解題参照。

これまでの一般的サーベイの上で、遡及的書誌編纂の一例として、われわれの共同作業である、表1に示した松田の3点の書誌作成において、筆者が主として担当した機械処理の手順と結果を検討する。書誌編纂の最大の問題である入力の前処理については、主題専門家として松田芳郎、佐藤茂行、広田明が参加し、書誌記述の統一等は松田、松井等が受持ったのであるが、ここではそれらの問題は一切省略し、機械処理の問題点を検討する。

(2) 使用ハードウェアとソフトウェア

(a) 使用機器

計 算 機	F A C O M 230-60 160 kw
カードせん孔機	(北大) H-1564 (IBM E Lコードに準じる) (小樽商大) I B M 26, 29 (Hコード)
磁気テープ	(800 B P I)
ライン・プリンタ	F-642 (数字, 英大字, カナ, 特殊文字)

(b) プログラム言語

F A C O M 230-60 M O N I T O R V システムのものと PL/I および S O R T ユティリティ使用。

(c) 文献データ処理の場合の使用機器による制約

① 文字エリアの長さ

F A C O M 230-60 の PL/I では、1つの文字エリアの長さは256文字以内という制限がある。⁽⁴¹⁾ PL/I では構造体や配列で処理できるが、プログラムは複雑となる。今回は書名の長さを256文字以下におさえてい⁽⁴²⁾る。

(41) これは、F A C O M 230 の機械語の命令で、長さを示す部分が8ビットで構成されていて、 $2^8=256$ が示しうる長さの最大であるためと考えられる。ただし、本来は機械語に制約されるべきでない高水準言語の PL/I が、このような制限を持たされていることは望ましくない。コンパイラの改良をメーカーに要望したい。ちなみに、I B M 360 では32,767字まで許されている。(『I B M システム/360 PL/I 解説書』[15], p. 34).

(42) 書名は時代がさかのぼるほど長くなる傾向にあるので、遡及的書誌編纂において、この制限は非常に不便なわけである。

② 使用可能文字

今回使用ライン・プリンタのキャラクタ・セットには、英小文字やラテン文字を使用する言語の発音符のついた文字（ウムラウト，アクセント等）が含まれず，カナが入っている。欧文文献処理⁽⁴³⁾には，当然のことながら，カナよりもこれら字種の備わった機種が必要である⁽⁴⁴⁾。なお特殊文字としては，PL/I 60 文字セットの特殊文字（21 種）に，60 文字セットで英字として扱われる ¥ @ # を加えたもの⁽⁴⁵⁾があり，[] が無いほかは，ほぼ十分であった。

③ 文字の内部コードの大小関係

FACOM 230-60 では内部コードに EBCDIC を用いているので，数字 > 英字 > カナ > 特殊文字の順になっている。特殊文字には句読符号，算術演算子，比較演算子などがあり，それらがランダムな大小関係を持っている⁽⁴⁵⁾。このため計算機による文字列処理は複雑となる。たとえば文字処理では基本的技法の 1 つである単語の識別には，特殊文字を単語の区切り記号として利用するが，それだけでは ▽（引用符，アポストロフィ）や -（ハイフン）等で単語が分離してしまう。同様に文字処理には欠かすことのできないソートにおいても，内部コードだけに頼ると，前述のように，慣用の配列とは大部異なったものになる。

④ カードせん孔機

IBM の Hコードと ELコードの機種を併用したが，変換プログラムがあり支障はなかった。コードの対応関係を図 7 に示す。同図で Hコードが空欄の場合は ELコードと同じ欄に重複せん孔すればよい。ただし IBM 26 せん孔機は，機構上重複せん孔に手間がかかるので，通常，特殊文字が多く出現する文献データには，不向きと思われた。

(43) たとえば，松田 [5] での使用言語は 12 か国語にのぼる。(3) 入力文献の項参照。

(44) 1972 年秋，全国 7 大学の大型計算機センターのすべてに間合わせたか，英字の大文字と小文字を打ち分けられる機種を見い出すことはできなかった。現在のところ，国産機では特注しなければならないと言うことである。

(45) 特殊文字を大→小の順に示す：▽▽=▽@#：?>-%./-;)*¥!&|+(<. 空白。

図7 カードせん孔機コード比較*

日立1564コード	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	/	S	T	U	V	W	X	Y	Z	1	2	3	4	5
IBM EL コード	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	/	S	T	U	V	W	X	Y	Z	1	2	3	4	5
IBM H コード	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	/	S	T	U	V	W	X	Y	Z	1	2	3	4	5
	Y																															
	X																															
	0																															
	1																															
	2																															
	3																															
	4																															
	5																															
	6																															
	7																															
	8																															
	9																															

日立1564コード	6	7	8	9	0	~	.	<	(+		!	¥	*)	:	-	082	,	%	_	>	?	:	#	@	▽	=	"	&	-	
IBM EL コード	6	7	8	9	0	ç	.	<	(+ <td> </td> <td>!</td> <td>\$</td> <td>*</td> <td>)</td> <td>:</td> <td>-</td> <td>082</td> <td>,</td> <td>%</td> <td>_</td> <td>></td> <td>?</td> <td>:</td> <td>#</td> <td>@</td> <td>▽</td> <td>=</td> <td>"</td> <td>&</td> <td>-</td>		!	\$	*)	:	-	082	,	%	_	>	?	:	#	@	▽	=	"	&	-	
IBM H コード	6	7	8	9	0	.)						\$	*		,	(=	▽		+	-			
	Y																															
	X																															
	0																															
	1																															
	2																															
	3																															
	4																															
	5																															
	6																															
	7																															
	8																															
	9																															

* 『北海道大学大型計算機センター・ニュース』1巻1号, p.38 より転記

(3) 入力文献

入力文献は表1のとおり, 計量経済学文献 864 点, サン・シモンおよびサン・シモン主義文献 686 点, 計 1,550 点である。使用言語は前者では英語, 後者ではフランス語が主であるが, そのほかに, イタリア語, スペイン語, ポルトガル語, ドイツ語, オランダ語, ロシア語, ポーランド語, デンマーク語, ハンガリー語などもかなりあり, 日本語文献も含まれている。

(4) 入出力フォーマット

(a) 入力データ・カード (12 種, 図8参照)

図8 入力データ・カード・フォーマット

1	4	5	6	7	8	17	19	28	80
BKNO	11	文献掲載場所				所蔵記号			
"	21	個人著者1 (25字)			個人著者2 (25字)			個人著者3 (23字)	
"	22	団体著者 (200字)							
"	23								
"	24								
"	31	書名 (252字)							
"	32								
"	33								
"	34								
"	41	出版地：出版者 (74字)							
"	42	出版年1	ページ数	書評者 (25字)					
"	43	出版年2	→ 使用言語記号						

→ 継続物の場合に最後の出版年を記入する。

図9 磁気テープ・フォーマット

BKNO	K/L	所蔵記号10字	使用言語記号	出版年1	出版年2	ページ	出版地・出版者	74字
1	5 6	16 17	21	25	29			102

E/S	文献掲載場所 10字	書評者 25字
103 104	114	138

個人著者1	25字	個人著者2	25字	個人著者3	25字	個人著者4	25字	団体著者	200字
139	164	189	214	239				438	

書名の長さ	書名 252字
439 442 443	694 695

予備エリア 25字
696 720

(b) 磁気テープ (固定長形式, 図9 参照)

(c) 出力リスト (図10 参照, なお目録1ページ当フォーマットについては, III 3 参照)

2 マスタ・ファイル作成

(1) 入力データ作成

作成手順は、① 原文献の収集と編集，② ワーク・シート記入，③ カードせん孔，検孔の順で，① は科研分担者が担当，② は学生筆耕者，③ は北大大型計算機センターと小樽商科大学計算センターに依頼した。作成カード枚数は8,520枚，1点当約5.5枚であった。使用字種は前述のライン・プリンタ字種の制約により，英大文字と数字，特殊文字であり，発音符はすべて省略し，ロシア語，日本語などはラテン文字へ翻字した。

作業としては，②ワーク・シート記入で，かなりの時間と経費を要し（1時間当平均12点転記），しかもかなり多くの転記ミス（スペル・ミスやせん孔時に見間違いやすい文字の書き分けのエラー，例：Iと1）を生じた。③カードせん孔では，両計算センターとも，これまで文献データせん孔経験がほとんどなかったらしく，多大の負担となったようである。もちろん，②③とも経験をつむことにより，ある程度は正確さや速度がますことは期待できるが，②は臨時の筆耕者であり，③は共同利用施設であるから，専用の仕事に習熟することは難かしいという事情もある。今回の実験でも，入力データ作成が文献データ機械処理の隘路であることが確認されたわけである。

(2) 暫定的マスタ・ファイル作成と内容チェック

データ・カード8,520枚を3回に分けて入力し，暫定的マスタ・ファイルとした。計算機による内容チェックでは，処理不能カード（例：更新すべきマスタ・レコードのないもの）を除外するにとどめ，残りはマスタ・ファイルに書き込んだ上，人間が各種出力リストを校正する方式をとった。今回は文献数もそう多くなかったので，計算機によるチェックを最小限にとどめた⁽⁴⁶⁾わけである。

人間による校正では，検孔および検孔後のカード・イメージ・リストによ

(46) ただし，Ⅱ 2(4)にあげたチェックの種類のうち，少なくともシーケンス・チェック（文献番号順にデータがきているか）だけは，計算機ですましておくべきであった。文献番号はコントロール・フローのキーとして用いられる重要なものであり，一方，計算機によるチェックのアルゴリズムも明白だからである。（J. L. Dolby et al. [26], p. 77).

る校正での訂正もれが発見されたわけであるが、それらの中には、スペル・ミス、⁽⁴⁷⁾桁ずれ、重複カード、せん孔もれ等と共に、かなりの表記不統一が見い出された。⁽⁴⁸⁾これは人間が読む場合にはそう支障がないとしても、計算機ソフトでは異なった文字列として処理されてしまう。言うまでもなく、計算機利用の場合は、従来の図書館での目録規則以上に厳密に表記法を定め、それを厳守する必要がある。今回の実験では、「記入要項」に従って転記を行なったが、その点十分でなかったと言える。

(3) マスタ・ファイル更新

更新の種類は、新文献の追加・挿入、すでに入力してあるデータの置換え、削除、訂正など

であり、カード・フォーマットの種類は表3のようにした。これらのうち、「スペル訂正」方式が最も多く用いられ、便利であった。(II 2 (6) 参照)

表3 更新カード・フォーマットの種類

更新のタイプ	カード・フォーマット
追加・挿入	オリジナル・フォーマット
置換え	(削除+挿入)
削除	削除フォーマット*
訂正	オリジナル・フォーマット スペル訂正フォーマット**

* BKNOのあとに、DELETEとせん孔したもの。

** 図4参照。

3 サマリー処理

作成した索引の種類は、著者、書名、刊年順、書評者、KWICの5種であり、前4種については、文献番号、著者、書名、出版年、書評者(書評者索引のみ)を無条件に抽出したので、プログラミングは非常に簡単であった。KWIC索引では、書名からの単語の抽出、キーワードの選定、KWIC列

(47) スペル・ミスの中では、著者名、書評者名が最も多く、校正においても原則として1字ずつ原文献とつきあわせたので長時間を要した。

(48) たとえば、著者名、書評者名の表記で、同一人名について、フル・スペル方式とイニシアル方式を混在させている。この問題については、日笠衆司[41]参照。

表 4 プログラムの構成****

作成リスト	プログラム名	入力カード数			LP枚数		プログラム数		CPU使用時間 (MS)			CORE占有時間 (MS)			CORE/CPU		北ソープ別*			
		プログラム・カード枚数	カード数	計	リスト作成	計	プログラム別	作成リスト別	コンパイル結合・編集	RUN SORT	計	コンパイル結合・編集	RUN SORT	計	コンパイル結合・編集	RUN SORT		計		
Master list	ACT 80	687	687	4,585	5,272	73	106	557	557	70,565	205,058	275,623	220,863	421,007	641,870	3.13	2.05	2.33	B	
	ACT 72	150	150	195	412	4	16	122	122	28,078	69,424	97,502	103,378	151,894	255,272	3.68	2.19	2.62		
	SORT	13	217	—	—	—	1	13	167	—	45,614	45,614	—	333,289	333,289	—	7.31	7.31		A B**
	ACT 73	54	—	—	—	98	105	32	15,985	96,721	112,706	103,628	127,845	291,473	10.24	1.32	2.39			
	ACT 83	76	—	—	—	—	1	11	54	17,964	24,916	42,880	140,343	115,301	255,644	7.81	4.63	5.96		
Author index	SORT	13	157	—	157	—	1	13	114	—	11,591	11,591	—	139,048	139,048	—	12.00	12.00	A	
	ACT 84	68	—	—	—	26	35	47	17,292	24,296	41,588	134,138	52,030	186,168	7.56	2.14	4.48			
	ACT 85	71	—	—	—	1	11	45	16,743	19,226	35,969	114,014	106,273	220,287	6.81	5.53	6.12			
	SORT	13	143	—	143	—	1	13	92	—	8,824	8,824	—	111,071	111,071	—	12.59	12.59		
	ACT 86	59	—	—	—	20	29	34	17,322	17,360	34,682	540,537	38,149	578,686	31.21	2.20	16.69			
Title index	ACT 87	71	—	—	—	1	11	47	17,105	19,281	36,386	97,102	102,630	199,732	5.68	5.32	5.49	A		
	SORT	13	148	—	148	—	1	13	105	—	8,582	8,582	—	92,634	92,634	—	10.79		10.79	
	ACT 88	64	—	—	—	20	29	45	16,948	17,977	34,925	112,056	34,814	146,870	6.61	1.94	4.21			
	ACT 89	66	—	—	—	1	10	47	16,877	18,954	35,831	87,568	99,565	187,133	5.19	5.25	5.22			
Reviewer index	SORT	13	149	—	149	—	1	13	112	—	8,972	8,972	—	109,533	109,533	—	12.21	12.21	A	
	ACT 90	70	—	—	—	20	29	52	17,654	24,980	42,634	124,527	58,052	182,579	7.05	2.32	4.28			
合計	合計	1,501	1,501	4,780	6,281	265	397	1,147	1,147	252,533	621,776	874,309	1,838,154	2,093,135	3,931,289	7.28	3.37	4.50	A	
	内訳	計	597	—	597	90	169	423	423	120,583	187,599	308,182	809,748	1,020,951	1,830,699	6.72	5.44	5.94		
	平均	74.6	74.6	74.6	11.3	21.1	52.9	52.9	17,226	26,780	44,026	115,678	145,850	261,528	4.26	2.48	2.86			
Bソープ (3件)	平均	904	904	4,780	5,684	175	228	724	724	114,628	416,817	531,445	487,869	1,034,035	1,521,904	4.26	2.48	2.86	A	
	平均	301.3	301.3	1,593.3	1,894.6	58.3	76	241.3	241.3	38,209	138,939	177,148	162,623	344,678	507,301	7.05	2.32	4.28		
	平均	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—		

* CPU使用時間3分, LP枚数100枚以内がAソープ, 30分, 500枚までがBソープ扱いである。
 ** ACT 72はAソープに相当するが, KWIC索引作成ソープとしての連続ランではBソープ扱いとなる。
 *** ACT 86のCORE占有時間のうち, コンパイル結合編集が異常に高いので, Aソープの計と平均では, ACT 86の使用時間等は除外した。
 **** カード枚数, LP枚数, 使用時間等は, 計量経済学文献 864 点の場合であり, ACT 80では第1回目のデータ・カード入力時, その他は最終原稿打出時のデータである。

の編集などの技法が必要であった。⁽⁴⁹⁾

ソートでは、KWIC索引編集プログラムにおいて、頻度順ストップ・ワード・リストの作成のために、コア・ソートを行なったほかは、SORTユーティリティを使用した。

作表のフォーマットは、オフセット印刷の原稿としてそのまま使用するために、最も読みやすい活字の大きさ等を検討の結果、ライン・プリンタ用紙2枚分(132行, 136字)を1単位として、縦は100行使用し、横は左右の余白1字分を残して全部印字して、それを約53%縮尺してA4判1ページとした。

マスタ・リストには、マスタ・ファイルのほぼ全内容を印字するためデータ量がかなり多くなるので、ページ数を節約するため2段組とした。その結果、1ページ当収録点数は約23点となった。その他各種索引では、ある程度の情報はマスタ・リストを参照しなくてもすむように、1文献当1行(134字)をあて、文献番号、著者、書名、出版年を可能な限り印字した。長すぎる場合は、書名の後半を省略した。(図8, 9, 10参照) 各種索引は1文献当1行なので問題はないが、マスタ・リストでは2段組とした関係もあり、ほとんどの書名が2行以上となった。この場合の行末処理として、今回「ヘア・スペース方式」を開発した。従来からの行末処理手法にはそれぞれ難点が認められたからである。(II 3(4)参照)

4 プログラムの分析

(1) プログラムの構成

今回の実験でのプログラム構成は表4の通りである。すなわち、プログラムACT 80でマスタ・ファイルを作成・更新し、マスタ・リストを印字した。そのマスタ・ファイルを使用して、KWIC索引など5種の索引を作成した。これら索引は、いずれもプログラム2本とソート1本で作成した。

プログラム・ステップ数の点で、全体の半分以上を占めているACT 80

(49) 拙稿 [18] 参照。

の構成を表5に示した。
すなわち、メイン処理の中
のセット・ルーチンと
ライト・ルーチンとで
ACT 80全ステップ数
の約8割を占めている。
ここでは、スペル訂正、
2段組印字、ヘア・スペ
ース方式など、文献デー
タ処理ルーチンとして今
回組込んだもののほか
に、第1回目の12種の
データ・カードのセット
にもかなりのステップを
要している。

メイン処理のうち、コ
ントロール・フローと

は、データの読み込み、処理、書き出しと言うプログラム全体のフローのコントロールである。通常、多種類のカードを読み込むために、この部分のプログラミングが難かしいのが、文献データ処理の特徴の1つと言える。同様に、実行命令に比して、エリアの定義が大きな比重を持っていることも、文献データの特徴と言える。なお、ACT 80プログラムの全体を通覧すると、ルーチン・コントロールに使われる判断命令と、その他の作業命令のうちの移送命令が多く、両者のうちでは前者の判断命令の占める割合が大きいと言える。

各種索引作成プログラムは、マスタ・ファイルからのデータ抽出、ソート、編集・印字である。KWIC索引を除くと、これらのプログラムは、1文献当1行印字としたため、1文献を1単位として処理でき、ライト・ルー

表5 ACT 80プログラムの構成

ジョブの分類	ステップ数
ハウス・キーピング処理	10
エリアの定義	23 (110行)
メイン処理	
コントロール・フロー	50
セット・ルーチン	174
内訳	
メイン	44
{ データ・カード・セット	41
{ スペル訂正カード・セット	73
{ 削除カード・セット	3
{ エラー・ルーチン	13
ライト・ルーチン	275
内訳	
メイン	115
{ 2段組印字	55
{ ヘア・スペース方式印字	105
エンド処理	25
合計	557

* PL/I では命令の行数とプログラム・ステップ数が異なる。特にエリアの定義では、1つの定義の終わりまでが1ステップになるので、()内に行数を示した。

チンでも2段組印字やヘア・スペース方式を採用する必要がなかったので、非常に簡単になっている。

(2) ジョブの性格

今回の編纂ジョブの入出力量，計算機使用時間などを，北大大型計算機センター・ジョブの平均と比較すると表6のとおりである。入出力量は筆者の予想に反し，センタ・ジョブ平均よりやや多い程度にとどまったが，計算機使用時間では大きな差が生じた。すなわち，CPU使用時間に対するコア占有時間の比で，AジョブBジョブともセンター平均の2倍以上になったのである。これは，前述のプログラム構成において，実行命令に比してエリアの定義が大きな比重を占めたことに対応するが，これらのジョブが計算センターにとって処理能率がよくないことが示されたわけである。

北大大型計算機センターは共同利用施設の性格上さまざまなジョブが持ち込まれるので，平均的ジョブが最も効率よく処理できるように運営されている。⁽⁵¹⁾そこで，たとえば，使用言語の点では，FORTRAN 88.5%，

表6 ジョブの比較

	CPU 使用時間(分, 秒)		CORE 占有時間(分, 秒)		CORE/CPU		入力カード枚数		LP枚数	
	A ジョブ	B ジョブ	A ジョブ	B ジョブ	A	B	A	B	A	B
文献目録編纂 ジョブ(平均)	0.44	2.57	4.22	8.27	5.94	2.86	571		21.1	76
北大センター・ ジョブ(平均)*	0.44	7.05	1.49	9.37	2.49	1.36	437.9		29.7	67.2
文献目録/センタ ー・ジョブ	1.0	0.4	2.4	0.9	2.4	2.1	1.3		0.7	1.1

* 文献目録編纂ジョブを実行した，1973年2月の平均（『北海道大学大型計算機センター・ニュース』5巻2号，p. 58）

(50) 中村鎮雄 [24] 参照.

(51) たとえば，テスト・ジョブによるFACOM 230-60の処理能力測定において，テスト・ジョブの分布がなるべく実ジョブの分布と等しくなるように配慮され，その結果，現在のシステムでは多重度3がよいという結論を得ている。江丸敏夫 [20] 参照。なお，国立大学大型計算機センターの運営一般については，国井利泰他 [22] 参照。

ALGOL 1.61%, FASP 1.63%, COBOL 1.62%, PL/I 1.02%の割合で使用されているので, FORTRANが最も効率よく処理されるようにファイルの割付けが行なわれている。⁽⁵²⁾このように, FORTRANに比して, COBOLやPL/Iの利用率が極度に低いことから, センター・ジョブのほとんどが数値計算で, 文献データ処理はまれであり, 従って, 現在のセンターでは文献データ処理の効率が低いことが理解される。

IV 考 察

1 プログラミング技法上の改善

今回の目録編纂実験について, 今後プログラミング技法上, 改善を予定している点を以下に記す。

① 書名の長さを256文字以上でも扱えるようにする。② 著者, 書評者名の表記を, 同一人についてはフル・スペル方式かイニシアル方式かに統一する。③ 書名索引では, ソート時に初語の冠詞を無視する。④ ドイツ語等の複合語で字数が多いものは, あらかじめ構成要素間に分綴記号を挿入しておき, ヘア・スペース方式やKWIC索引で利用する。すなわち, 前者では, ハイフンと同様に区切り記号に用いることにより, 1行におかれるヘア・スペースが多くなりすぎることを避ける。後者のKWIC索引では, 分綴記号の次の文字からもキーワード位置に印字して, 2番目以降の構成要素からも検索できるようにする。⑤ 入力データのせん孔ミスを可能な限り訂正する。

2 文献データ処理センターについて

今回使用した北大大型計算機センターでは, 文献データ処理はきわめてまれであり, 処理効率も低かったと言える。このことは他の大型計算機センターでも同じように思われた。⁽⁵³⁾ただし, 本稿の冒頭でも記したように, 文献デ

⁽⁵²⁾ 中村 孝 [23], p. 41.

⁽⁵³⁾ たとえば, 東大大型計算機センターでも, ほとんどのジョブがFORTRANを使用している。石田晴久他 [21], p. 715.

ータ処理への計算機利用は、今後ますます増加すると考えられるので、どこか1つのセンターだけでも、文字データ処理に向けた機種を、周辺機器も含めて導入すること、あるいは別に文献データ処理専用の共同利用施設を開設することが必要である。文献データの場合、大量のデータ処理を前提とするので大型機の導入が望ましく、それを1機関で専用することは不経済だからである。そのような文献データ処理センターのあり方を、これまで本稿で考察してきたことを土台として記すと次のようである。

① CPU使用時間に対するCORE占有時間の比を小さくして処理効率をあげるため、CPUの処理速度はそう高速でなくてもよいから、コアは大容量にして、演算の多重度をあげること。

② 文字データ処理に向けた入出力機器を導入すること。たとえば、ライン・プリンタは英字の大文字と小文字を打ち分けられるようにする。なお、今回使用のライン・プリンタでは印字位置が上下にかなりのばらつきを生じた。これも一直線上にならぶよう、ハードの改良が必要である。また、出力リストをそのまま印刷原版とすることが多いので、印字リボンを常に最良の状態に保つことが必要である。なお、新たに開発される入出力機器（例：光学式読取装置、日本語入出力機器など）を積極的に試用できる体制をとる。

同時にデータ・ライブラリ⁽⁵⁴⁾を開設し、たとえば、既存のデータ・ベース(MARC II, CAS, MEDLERS など)を導入し、利用できる体制をとる。

③ 文字列処理用プログラム・ライブラリを充実させる。たとえば、KWIC索引は多くの分野で作成されるが、その計算機処理はほぼ一定である。そこでKWIC索引を必要とするユーザがセンター側によって示される仕様によって入力データを作成し、かつエラー・データの修正を行ないさえすれば、望むリストが得られることが望ましい。

④ ユーザに対する文献データ処理技法の教育体制を確立する。

(1973. 6. 25)

(54) 国井利泰他 [22], p. 107.

引用・参考文献一覧

〔A〕 機械編纂された遡及的目録類

(表1に掲げたもの)

- [1] Janda, Kenneth, ed. *Cumulative Index to the American Political Science Review, Volumes 1-57: 1906-1963*. Evanston: Northwestern University Press, 1964. xxvi, 225 pp.
- [2] Aldous, Joan and Hill, Reuben. *International Bibliography of Research in Marriage and the Family, 1900-1964*. Minneapolis: University of Minnesota Press, 1967. 508 pp.
- [3] Pollin, Burton R. *Godwin Criticism, A Synoptic Bibliography*. University of Toronto Press, 1967. xlvi, 659 pp. Computer Programming Directed by Dr. George W. Logemann.
- [4] Gotterer, Malcolm H. *KWIC Index, A Bibliography of Computer Management*. Princeton: Brandon/Systems Press, 1970. viii, 152 pp.
- [5] Matsuda, Yoshiro, ed. *KWIC Index Series for Social Sciences, 1-3*.
 No. 1 *Bibliography: Works on Econometrics — reviewed in Econometrica 1951-1971—*. Otaru, 1973. viii, 24, 145 pp.
 No. 2 *Bibliography: Works by Saint-Simon and Saint-Simonians*. (to be published) viii, 24, 92 pp.
 No. 3 *Bibliography: Works on Saint-Simon and Saint-Simonians —1831-1970—*. Otaru, 1973. viii, 24, 49 pp.

(同上参考文献)

- [6] Janda, Kenneth. *Information Retrieval: Applications to Political Science*. Indianapolis: The Bobbs-Merrill Co., Inc., 1968. xxiii, 230 pp. ([1] の解説を含む)
- [7] 藤川正信, 白井 厚「書評 パートン・R・ポリン編『ゴドウィン論評書—総文献目録』『三田学会雑誌』61巻10号(1968. 10), 1102-10. ([3] の書評)

(その他)

(KWIC索引)

- [8] Institut de Recherche d'Informatique Juridique. *Jurindex: l'informatique appliquée à l'actualité bibliographique du droit français*. 1971- Paris: Masson & Cie Éditeurs.

(KWOC索引—手作業)

- [9] [木田橋喜代慎] 編 *Bibliothèque du Feu Prof. Juro Tedzuha, Section: Philosophie, Table Analytique*. [ca. 1951] 100 pp.

(蔵書目録)

- [10] Harvard University Library. *Widener Library Shelflist, 1 Crusades — 1965—* —. Cambridge, Mass.: The Harvard University Press. (Whole shelflists contain approximately 1.6 million titles.)
- [11] De Gennaro, Richard. "Harvard University's Widener Library Shelflist Conversion and Publication Program." *College & Research Libraries*, vol. 31, no. 5 (September 1970), 318-31.
- [12] アジア経済研究所『発展途上国の統計資料目録 (1969年版)』1970.
(筆者未見のものでは下記のようなものがある.)
- (i) Bwy, Douglas. *Social Conflict: A Keyword-In-Context Bibliography on the Literature of Developing Areas, with Supplementary References from Latin America*. (Prepared in connection with research supported by the National Science Foundation and the Northwestern University Council for Intersocietal Studies.) 1966.
- (ii) Farley, Earl, ed. *Kansas Slavic Index*. Lawrence: The University of Kansas Libraries, 1963.
- (iii) Grazia, Alfred de. (general ed.) *The Universal Reference System: Political Science, Government, & Public Policy: An Annotated and Intensively Indexed Compilation of Significant Books, Pamphlets, and Selected and Processed by The Universal Reference System — A Computerized Information Retrieval Service in the Social and Behavioral Science*. vol. 1 *International Affairs*. 1965, 2nd ed. 1969— vol. 10 *Comparative Government and Cultures, with Annual Supplements*. 1970-72.
- (iv) Janda, K. and Cultler, Neal E. "The Cumulative Index to the Midwest Journal of Political Science." *The Midwest Journal of Political Science*, vol. 11 (May 1967), 225-55.
- (v) Straus, Murray A. and Graham, Susanne C. *Cumulative Index to Marriage and Family Living, Volumes 1-24 (1939-1962)*.

[B] プログラム言語・技法・計算センター

(プログラム言語)

- [13] Mott, Thomas H. Jr., Artandi, Susan, and Strumingre, Leny. *Introduction to PL/I Programming for Library and Information Science*. New York: Academic Press, 1972. xvi, 231 pp.
- [14] 竹下 亨『PL/I 複合プログラミング言語』日本経営出版会, 1970. 412 pp.
- [15] 『IBMシステム/360 PL/I 解説書』日本アイ・ビー・エム, 1969. 364 pp.
(N: GC 28-8201-1)

(プログラミング技法)

- [16] 半沢孝雄『プログラミング技法集』竹内書店, 1970. 366 pp.
- [17] Rich, Robert P. *Internal Sorting Methods Illustrated with PL/I Programs*. Englewood Cliffs, N. J.: Prentice-Hall, 1972. 154 pp.
- [18] 松井幸子「プログラミングに関するノート (11)—PL/I による文字列の処理」『北海道大学大型計算機センター・ニュース』4巻3号 (1972. 7), 13-25.
- [19] 坂本徹朗「電子計算機の応用」(高橋達郎, 広田広三郎, 坂本徹朗『情報検索の手法と機器』南江堂, 1969), 209-301. (COBOLによるKWIC索引プログラミング)

(計算センター)

- [20] 江丸敏夫「FACOM 230-60 システムの処理能力」『北海道大学大型計算機センター・ニュース』3巻4号 (1971. 9), 36-41.
- [21] 石田晴久, 高橋延匡, 斎藤五郎「東京大学大型計算機センターにおけるジョブ統計」『情報処理』13巻10号 (1972. 10), 714-19.
- [22] 国井利泰 (座長) 他「計算センター運営論——国立大学大型計算機センターの場合——」『情報処理』13巻2号 (1972. 2), 104-12.
- [23] 中村 孝「最近のセンター事情」『北海道大学大型計算機センター・ニュース』4巻6号 (1972. 12), 32-41.
- [24] 中村鎮雄「CPU時間とコア占有時間の関係について」『北海道大学大型計算機センター・ニュース』3巻5号 (1971. 10), 59-61.

[C] 文字データ処理の手順と技法

(一般・図書)

- [25] Cox, Nigel S. M. and Grose, Michael W. *Organization and Handling of Bibliographic Records by Computer*. Newcastle upon Tyne: Oriel Press Limited, 1967. xiv, 187 pp.
- [26] Dolby, J. L., Forsyth, V. J., and Resnikoff, H. L. *Computerized Library Catalogs: Their Growth, Cost, and Utility*. Cambridge, Mass.: The M. I. T. Press, 1969. ix, 164 pp.
- [27] Hayes, Robert M. and Becker, Joseph. *Handbook of Data Processing for Libraries*. New York: Becker and Hayes, Inc., 1970. xv, 885 pp.
- [28] Meadow, Charles T. 『情報検索 検索言語・情報構成・ファイル処理』[*The Analysis of Information Systems: A Programmer's Introduction to Information Retrieval.*] 渡辺茂監訳, 高地高司, 笹森勝之助訳, 日本経営出版会, 1970. 357 pp.
- [29] 中井 浩, 笹森勝之助『情報検索システム』日本経営出版会, 1971. 221 pp.
- [30] 高橋達郎, 大場賢一, 林 省三, 福島芳直『情報検索』一橋出版, 1971.

325 pp. (コンピュータによる実務技法, ケース・スタディ・シリーズ, 3)

(一般・論文)

- [31] Batty, C. D. The Use of Computers in Bibliographical Compilation. (in Robinson, A. M. *Systematic Bibliography: A Practical Guide to the Work of Compilation*. 3rd ed. London: Clive Bingley, 1971). 78-86.
- [32] Kilgour, Frederick G. "Computer-Based Systems, A New Dimension to Library Cooperation." *College & Research Libraries*, vol. 34, no. 2 (March 1973), 137-43.
- [33] 桜井宣隆「冊子目録作成の機械化」『図書館学会年報』15巻2号(1969. 8), 43-59.
- [34] 「特集・医学図書館における電子計算機の利用」『医学図書館』15巻4号(1968. 12)
- [35] 「特集・図書館の機械化」『図書館雑誌』63巻8号(1969. 8)
- [36] 「特集・図書館の機械化 その実態と問題点」『現代の図書館』8巻1号(1970. 3)

(標準化)

- [37] 海宝 顕『データ処理の標準化』産業図書, 1971. 204 pp.
- [38] 「特集・ISBD (International Standard Bibliographic Description) をめぐって」『現代の図書館』10巻3号(1972. 9)

(文字列処理)

- [39] 安西郁夫「電子計算機による文字列の処理」*Library and Information Science*, no. 7 (1969), 109-16.
- [40] 西村恕彦「文字列の配列順序についての問題」『情報処理』10巻1号(1969. 1), 21-25.
- [41] 日笠衆司「著者索引と日本人名の省略」『ドクメンテーション研究』20巻9号(1970. 9), 271-78.
- [42] 溝口歌子「外国語雑誌の誌名と使用言語の判別法」『ドクメンテーション研究』23巻4号(1973. 4), 117-23.

(エラー・チェック)

- [43] 小松原康敏, 古谷 実, 日笠衆司「書誌情報のインプット時におけるエラーの発生」『情報科学技術研究集会発表論文集』第8回(1971), 311-17.
- [44] 日笠衆司「INIS ソースの構造とディスクリプター・ファイルの自動チェック」『ドクメンテーション研究』21巻5号(1971. 5), 145-51.
- [45] 塩田卓和, 小関和俊, 正井 樹「文献情報のチェックシステムについて」*TOSBAC REPORT*, no. 8 (1972. 3), 80-84.

(市販磁気テープ)

- [46] 丸山昭二郎「MARC II フォーマット」(丸山昭二郎, 井上哲也『洋書目録マニュアル』日本図書館協会, 1970), 100-27.
- [47] [The Library of Congress MARC Distribution Service] *Specifications for Magnetic Tapes Containing Monograph Catalog Records in the MARC II Format, II. Character Set.* pp. 4-15.
- [48] 吉川藤一, 中山和彦, 高田博行, 松村多美子「MARC II 磁気テープの利用による目録カード配布プロジェクト」『ドクメンテーション研究』23巻5号(1973. 5), 141-46.
- [49] 安田晶次, 岡田真一「トヨタ自動車工業㈱における市販磁気テープファイルの利用」『ドクメンテーション研究』22巻7号(1972. 7), 225-33.
- [50] 菊池敏典「二次資料の磁気テープファイルの利用」『情報管理』13巻10号(1971. 1), 657-62.
- [51] 栗田 洵, 菰田秀子「市販検索用磁気テープのファイル形式について」『情報管理』13巻12号(1971. 3), 749-63.
- [52] 「特集号 J I C S T 文献検索磁気テープ提供開始」『情報管理』14巻11号(1972. 2)

[D] 文字データ機械処理の具体例

(国立国会図書館)

- [53] 小田泰正「欧文逐次刊行物所蔵目録の機械編さん」『国立国会図書館月報』134号(1972. 5), 2-9.
- [54] 島村隆夫, 田村貴代子, 水野恵美子「国会会議録総索引の機械編さん」『国立国会図書館月報』136号(1972. 7), 2-13.
- [55] 高橋徳太郎「コンピュータに載せるといふこと——国会会議録索引業務での一例——」『国立国会図書館月報』106号(1970. 1), 2-10.

(日本科学技術情報センター)

- [56] 不破 博, 西尾元孝「文献速報自動作成システムについて」『情報管理』12巻6号(1969. 9), 293-308.
- [57] 亀山芳子, 田中 功, 増田勝治「逐次刊行物書誌ファイルの設計—JICST 資料管理システム」『情報科学技術研究集会発表論文集』第7回(1970), 1-10.

(アジア経済研究所)

- [58] 佐野敬夫『MARC-SDI システムの開発——アジア経済研究所における実験——』アジア経済研究所, 1972. 26 pp. (統計参考資料, No. 67)

(日本原子力研究所)

- [59] 古谷 実, 清水昭郎, 平川 隆「電子計算機による洋書冊子体目録の作成 I」『ドクメンテーション研究集会発表論文集』第3回(1966), 191-98.

[60] 古谷 実「電子計算機による索引および管理統計の作成」『ドキュメンテーション研究』19巻2号(1969.2), 37-43.

[61] 成井恵子「電子計算機によって作成する目録の記述の設計に関する一考察」『ドキュメンテーション研究集会発表論文集』第4回(1967), 107-14.

(東京大学)

[62] 遠藤哲郎, 高橋伸夫, 青木孝雄「雑誌受入業務の機械化計画—東大医学図書館における CURRENT JOURNAL 処理業務(第1期)を中心として—」『医学図書館』15巻4号(1968.12), 339-85.

[63] 田辺 広, 柿沼正男, 金子 豊, 湯浅富士夫「東京大学附属図書館における外国雑誌受入業務の機械化」『ドキュメンテーション研究』21巻3号(1971.3), 75-83.

[64] 牛島照夫「OLIMPAS—数学書の標題検索および自動分類をめざした作業の報告と検討—」(森口繁一『大学図書館と情報処理』科学研究費・特定研究Ⅰ「学術情報処理に関する基礎的研究」昭和47年度研究報告, 1973), 156-216.

(名古屋大学)

[65] 川原和子「電算機による経済統計資料の目録編集と検索—実験報告」『経済資料研究』6号(1973.2), 25-44.

[66] 名古屋大学附属図書館 特定科学研究費実験(経済学部班)『経済統計資料目録編集・検索(実験)』1972.1. 26 pp. (経済資料協議会, 昭和47年度総会研究報告配布資料)

[67] 島田静雄『土木工学図書館の情報検索(洋書編)』1973. 98 pp. (文部省科学研究費特定研究(1)「図書館業務機械化に関する基礎的研究」昭和47年度研究実績報告)

(神戸大学経営分析文献センター)

[68] 神戸大学経済経営研究所経営分析文献センター『文献検索プログラム操作説明書』1969, 100 pp.

[69] 塩田卓和「文献情報処理の現段階—目録の機械処理をめぐる—」『経済資料研究』6号(1973.2), 1-24.

(京都産業大学)

[70] 荒木雄豪, 平木茂子, 坪田 毅「京都産業大学における図書館業務の機械化」『ドキュメンテーション研究』22巻3号(1972.3), 75-85.

(大阪大学)

[71] 田保橋彬, 門田泰典「大阪大学附属図書館における業務機械化システム(LICS)の概要」『ドキュメンテーション研究』22巻11号(1972.11), 361-76.

[72] 浅野次郎, 門田泰典, 松浦 正「大阪大学附属図書館の電算化」『丸善ライ

ブラリーニュース』90号 (1973. 12), 865-69.

(群馬大学)

[73] 柿沼良介「コンピューターによる図書館業務処理システム—群馬大図書館の場合—」『丸善ライブラリーニュース』90号 (1973. 1-2), 870-72.

(慶応義塾大学)

[74] 安西郁夫「索引の自動化システム MISA (Mechanized Indexing System by Anzai)」*Library and Information Science*, no. 8 (1970), 101-21.

[75] 小川治之, 野添篤毅「逐刊物記録処理の機械化」*Library and Information Science*, no. 7 (1969), 131-48.

[76] 裏田和夫「『医学雑誌総合目録』編集の機械化システム」*Library and Information Science*, no. 10 (1972), 143-67.