

国立国語研究所学術情報リポジトリ

文献検索システム IRON

メタデータ	言語: Japanese 出版者: 公開日: 2017-06-13 キーワード (Ja): キーワード (En): 作成者: 中野, 洋, NAKANO, Hiroshi メールアドレス: 所属:
URL	https://doi.org/10.15084/00001306

文献検索システム IRON

中 野 洋

1. はじめに

文献検索は、電子計算機を使った仕事として、国内国外を問わず古くから研究され、また、実用されているものの一つである。国内では、古くは神戸大学計算情報センターがあり、大掛りなものは、日本科学技術情報センター（J I C S T）の機械検索サービス、オンライン情報検索サービスがある。また、多くの人に使われているものは、各地の大型計算機センターに登録されている文献検索プログラム*がある。文献データも磁気テープわたしでアメリカと日本、日本国内での交換利用が行なわれている**。

このような状況で、なお新しいシステムを作る必要がどこにあるか、それは以下の通りである。

1) 現在の文献検索の問題点は、いかにねらった内容の文献を見つけるかということにある。つまり、タイトルにもそれに付けられたキーワードにも、また抄録文にも直接には語としてあらわれてないある内容を、つまり手掛りがない状態で、どのようにして検索しえるかという問題である。

このことについては、今なお各方面で研究中である。たしかに言えそうなのは、文献の中の文章をいかに解析し、その意味内容を処理するかという問題が解決されない限り、本当の意味での文献検索はできないだろうということである。

この問題を考えるためには、現在可能なことは何か、問題は何かを明らかにするために、実験システムを作る必要がある。第一の目的は、この実験システムの作成にある。

* 東京大学大型計算機センター（TOOL-IR）その他が提供されている。文献10参照。

** JICST では、国内国外の文献、情報約585万件を磁気テープに蓄積し、提供サービスをおこなっている。文献11参照。

- 2) 現在わが言語計量研究部で作られた文献は付録に示したように二百編を越える。計算機を使用しはじめてから十四年を数え、この間メンバーも変わっている。同じような目的を持った文献が数編になるものもある（付録、件名索引参照）。オンラインキーワード検索を作り実用に供して欲しいという要望もある。これが第二の目的である。
- 3) 毎年業績報告なるものを書かされているがこれを計算機にやらせようというのが第3の目的である。
- 4) 外国人に計量研究部の業績を紹介するために人間の手をわずらわせることなく欧文目録、件名索引を作成する。これが第4の目的である。

作成した文献検索プログラム IRON (Information Retrieval Onami & Nakano) の機能は以下の通りである。

文献目録の作成	和文目録，欧文目録
件名索引の作成	和文件名索引，欧文件名索引
オンライン文献検索	和文キーワード または 著者による検索 欧文キーワード または 著者による検索 指定仮名文字連続による検索

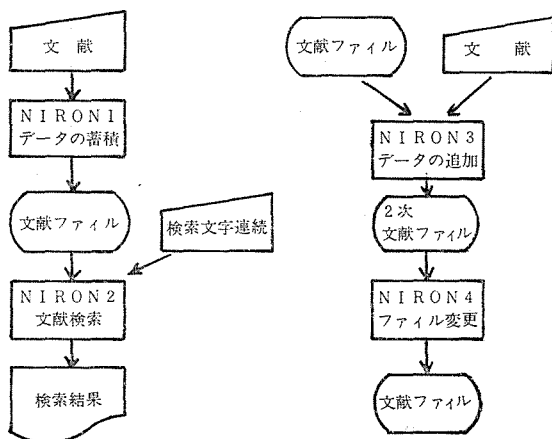
これらの目的は一応達成した。附録の文献目録はこのシステムで作成したものである。

また、他のシステムにはないと思われる新しい機能も付けることができた。任意語による検索機能とローマ字を五十音順に配列するプログラムである。

2. 方法

2-1 システムチャート

システムチャート 1 PDP-11使用



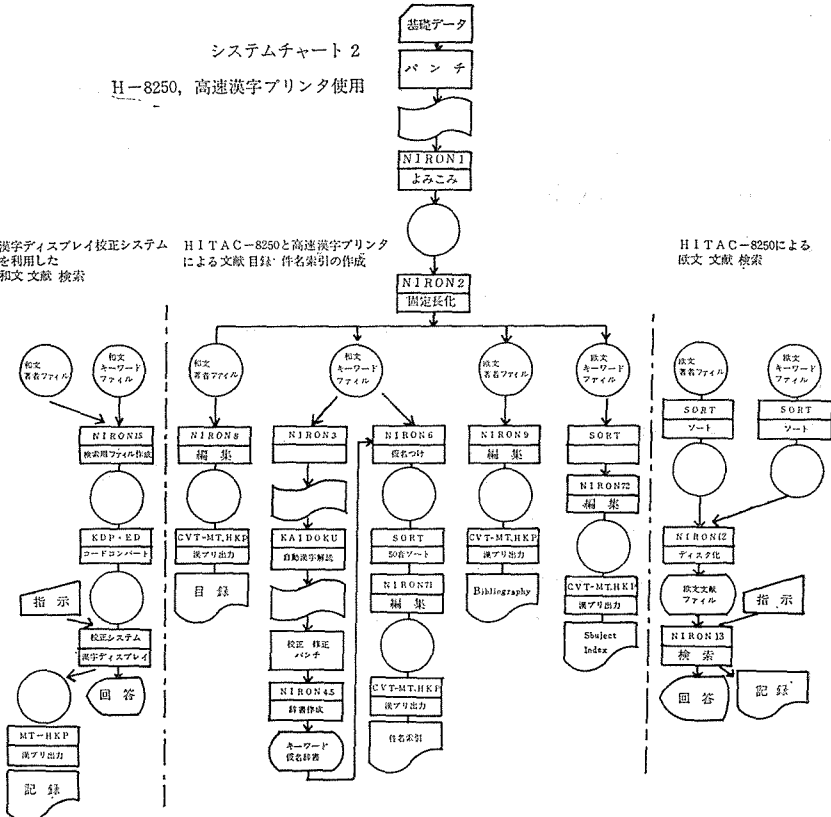
システムチャート 2

H-8250, 高速漢字プリンタ使用

漢字ディスプレイ校正システム
を利用した
和文 文脈 検索

HITAC-8250と高速漢字プリンタ
による 文献 目録・件名索引の作成

HITAC-8250による
欧文 文献 検索



2. 使用機器

文献目録の作成は、主な作業を HITAC-8250 でおこない、高速漢字プリンタで出力した。

文献検索は、欧文による検索を HITAC-8250 で、和文による検索を漢字プリンタとディスプレイで、カナ文字連続による検索を PDP-11 でおこなった。

2-3 データ・フォーマット

PDP-11 によるカナ文字連続の検索では、決ったデータ・フォーマットはない。Basic 言語がうけつける長さを最大長とするだけである。

H-8250, 漢字プリンタによる目録作成・検索のデータ・フォーマットについて以下に説明する。

2-3-1 基礎データ

次のページに示したものが、このシステムの基礎データである。普通の文献カード（これはマルゼンのアナリストカード 403）に必要な事項を書き込んだものだが、カードは何でもよい。あるいは、このカードがなくてもよい。筆者は人間の手による検索・分類・整理に利用するために、これを用いた。

品詞認定の自動化			
中野 洋			
電子計算機	言語研究	03	98 120 1971 03
Automation of Classification of Parts of Speech			
NAKANO Hiroshi			
品詞 (Parts of Speech) ・ 計算機 (computer)			
語結合 (word combination)			

パンチ原稿がわりに使っている。

2-3-2 入力データ

入力データは、漢テレによってさん孔された紙テープである。

一つの文献、つまり一枚の文献カードで一ブロック（最後E/I）を構成し、各項目（下記）はC/Rで区切られて一レコードを成す。項目の順序は自由である。

項目記号	スペース	データ	C/R
(一字)	(一字)	(可変)	(一字)

項目は下記のとおりである。

(項目)	(項目記号)	(備考)
標題	題	
著者・編者	著	共著の場合は・(ナカグロ)で区切る
資料・雑誌名	資	
ページ	ペ	××号××—×× ×:数字
発行年	年	××××/××/×× ×:数字
欧文題目	英	
欧文名	N	共著の場合は・(ナカグロ)で区切る
キーワード	K	先に和文,()内に欧文。・(ナカグロ)で区切る。
件名	件	

下に、例として入力データを漢字プリンタに出力したものを示す。

題 新聞における漢語のまぜ書き表記について

著 土屋 信一

資 電子計算機による国語研究

ペ 7号9-35

年 1975/3

英 A Method to write Chinese Word In Chinese Character and Kana in Newspapers

N TSUCHIYA Shin-ichi

K まぜ書き(a method write chinese word in Chinese character and Kana)・漢語(chinese words)・用字法(use of characters)・現代新聞(modern newspaper)

2-3-3 出力データ・フォーマット

次頁以降に出力例を示す。(1)~(4)は附録参照。

(1)和文著者ファイルから作られる主要記事目録

(2)和文キーワードファイルから作られる件名索引

(3)欧文著者ファイルから作られる Bibliography

(4)欧文キーワードファイルから作られる Subject Index

(5)欧文文献ファイルから、著者を指定して検索された結果

(6)同じく、キーワード・和集合指定の検索結果

積集合指定の検索結果

(7)和文文献ファイルから、著者またはキーワードを指定して検索された結果

PDP-11によるカナ文字連続指定の検索結果である。

(5) 著者指定検索の結果

(6) キーワード指定検索の結果

*****	*****
* NAKANO HIROSHI *	* MIZUTANI SIZUO.TANAKA SA *
* A PLAN FOR RECOGNITION O *	* ON COORDINATE CONNECTORS *
* F MEANING IN LANGUAGE DA *	* OF WORDS OR WORD GROUPS *
*****	*****
* NAKANO HIROSHI *	* KUSANAGI YUTAKA *
* CONTRASTIVE STUDY ON VOC *	* PATTERN MATCHING OF SNOB *
* ABULARY IN SIX KINDS OF *	* OL4 AND LINGUISTIC OPERT *
*****	*****
* NAKANO HIROSHI *	* MIZUTANI SIZUO *
* ON THE FREQUENCY-DISTRIB *	* INTRODUCTORY NOTE ON SNO *
* UTION PATTERN L OF VOCAB *	* BOL3 *
*****	*****
* NAKANO HIROSHI *	* MIZUTANI SIZUO *
* A TRIAL OF AUTOMATIC SYN *	* INFLUENCE OF THE DIFFERE *
* TACTIC ANALYSIS *	* NCE OF SPELLING SYSTEMS *
*****	*****

■中野 洋

71号17-24

構文自動解析の試み

意味辞書 学習機能 文法辞書 構文解析

(8) カナ文字連続指定による検索結果——線部, 応答

⇒ ケンサク シタイ コトオ タイフ とヨ
コトハ カナデ ヲシ シメイハ エイモシテ
?コト"モ
{ 「オトナノ コトハ" ノ コト"モ ノ コトハ" 」 GNG-008
「カ"クレイキ ノ コト"モ ノ コトハ" 」 GNG-018
「コト"モ ノ コトハ" 」 GNG-042
「コト"モ ノ フソヒ" コトハ" 」 GNG-160
「コト"モ ノ シ 」 GNG-191
「コト"モ ノ コトハ" ノ セカイ 」 GNG-283

READY

2-4 プログラムの概要

(1)よみこみ (NIRDN 1)

原データ (紙テープ) を読み, 磁気テープに出力する。

(2)固定長化 (NIRDR 2)

和文著者ファイル, 和文キーワードファイル, 欧文著者ファイル, 欧文キーワードファイルを作成する。

このシステムはMTベースだが, 将来はディスクベースにするべきである。データベースシステムを採用すれば, (1)(2)が一本化し, データのメンテナンス, 使用も効率化するはずである。

次頁に, MTフォーマットを記す。

(3)編集 (NIRON 8, NIRON 71, NIRON 9, NIRON 72)

出力データフォーマット(1), (2), (3), (4)を印字するためのファイルを作成

和文著者ファイル

資料名	件名	共著者	著者数	題目	頁					年月日					著者(単)	著者よみ	番号	入力年月日	番号作成
					号数	「号」	始頁	「」	終頁	年	—	月	—	日					
30	4	60	2	100	4	2	4	2	4	8	2	4	2	4	60	20	6	6	4

和文キーワードファイル

同 上															キーワード	キーワードよみ	同 上		
															40	40			

欧文著者ファイル

資料名	件名	共著者	著者数	題目	頁	年月日	著者(単)	番号	入力年月日	入力番号
					同 上	同 上				
30	4	100	2	200			40	6	6	4

欧文キーワードファイル

同 上															キーワード	同 上	
															100		

する。このファイルは漢字プリンタにかけられ出力印字される。

(4)かなつけルーチン

和文キーワードには、もともとよみがなが付いていない。しかし、件名索引を五十音順に配列するためにはキーワードによみがなが必要である。このルーチンは、自動漢字解読と人手による修正によって、よみがなをつけようとするルーチンである。

(4—1) キーワード紙テープ出力 (NIRON 3)

漢字表記のキーワードを紙テープに出力する。

(4—2) 自動漢字解読 (KAIDOK)

4—1の出力紙テープに読み仮名をつける。田中章夫氏作成のプログラ

ムである。(田中章夫「漢字かなまじり文を全文カナ書き・ローマ字書きに変換するシステムについて」電子計算機による国語研究Ⅱ 参照)

(4-3) 人手による修正

KAIDOKU の出力は紙テープである。これに印字・校正・修正をおこなう。

(4-4) 入力 (NIRON 4)

入力紙テープは次のフォーマットである。漢字は一字ずつによみがながつけられる。

漢 字	[よ み が な]	~	,
2	2	可 変		2	2	

MTに次のフォーマットで出力する。

見 出 し	よ み が な
30	30

(4-5) 辞書の作成 (NIRON 45, NIRON 5)

(4-4) 出力のMTによってディスク上によみがな辞書を作成する。同形語はそのうちの一つを取る (NIRON 45)。

(4-6) かなつけ (NIRON 6)

和文キーワードファイルに対し辞書によってよみがなをつける。よみがながつかないデータは、ディスプレイに表示され(紙テープにも出力)、これを見てコンソールから漢テレ3010コードによって入力することができる。この方法を用いないのなら4-3に戻る。

3. オンライン文献検索

機械によるオンラインの文献検索について考えよう。

オンライン文献検索において大切なことは次のことである。

必要とする文献を出来るだけ早く検索し、わかりやすく表示すること。

ここで、「必要とする」の意味は、必要とする文献が必ず入っていて、かつ必要としない文献は入っていないことである。

文献検索のスタート時では、「必要とする文献が必ず入っていること」だけを満足させて、よしとしがちである。しかし、これでは、関連のありそうな文献をすべて表示すればよく、極端な場合、表示された文献の中から本当に必要とする文献を探し出すまでにまた時間がかかることになる。

では、真に必要な文献だけをぬき出し、必要としない文献を排除するにはどうすればよいか。

そのひとつに適切なキーワードをつけることがあげられよう。しかし、万能のキーワードは存在しない。たとえば「文献検索の方法」というタイトルの文献に対し「文献検索」というキーワードは、それ自身による検索には能力を発揮するが、同意語「ドキュメント リトリユーバル」、類義語「文献の探索」、上位語「文献」「書物」「資料」、下位語「機械検索」「計算機利用」「図書館」などなどにはそれだけでは無力である。

これに対し、シソーラスを用意し、それによって関連をとろうとする方法が行われている。しかしながら、これにもふたつの問題点がある。その一つは、キーワードの語形・表記上の統一であり、他の一つは用意されたシソーラス以外のキーワードは使えないということである。つまり、実際上はキーワードつけは専門家によらねばならないということになる。

これはオンライン機械検索もまた、専門家がオペレートしなければならないことを意味する。というのは、検索を行おうとする人はシソーラスを知らないと考えなければならないが、シソーラスを持たないで専門語でかつシソーラスにある語を指定することは、シソーラスや検索したい文献の内容を知らない（これはあたりまえ）オペレータには到底無理だからである。つまり、オンライン文献検索がその効果を発揮しないということを意味する。この点の解決策は

IRON システムにおいて実現されている。あとで詳述する。

シソーラスを利用することは理想的な文献検索からみると蛇道であり、また効果的でないことを述べた。では、どうすればよいか。

この問題は、現在、各界で研究中である。文献の内容を自動的に抽出すること、その上で、効率的な検索方法を考えることが必要になる。その第一歩は自動単位切りであり、構文解析であり、文章処理である。わたしは、処理速度は遅くても言語学的にみて、オーソドックスなアプローチでこの問題について考えてみたい。IRON システムはその実験のためのシステムである。

次に、現状における文献検索システムの実用化を考えてみよう。

内容を自動的に抽出すること、範囲を限って自動的にキーワードをつけること（自動インデキシング）も現在は実用段階ではない。

このような現状では、適当なキーワードをつけてキーワードの管理を行う方法とキーワードは全くつけなくて題目や抄録文中の語を利用する方法の二つが考えられる。

どちらの方法でも肝要なことは、簡単にデータを入力・管理できること、データを手元において、いつでも好きなときにオペレートできることである。

後者の方法による、ミニコンピュータ、マイクロコンピュータにおいて実用に供することのできるシステムを作った。データはほとんど加工をくわえず、そのままで打ち込む。検索は指定した文字列によって行う。したがって、長い語でも（この場合、該当データは多くなる）、いくつかの語の組み合わせでも（この場合、より狙ったデータが選ばれる）、名前でも、雑誌名でも検索できる。しかし、検索時間は、データを頭から調べていくので、長くかかる。個人用のファイルならそれ程多くないだろうから実用に耐えるだろうという予測である。このシステムはミニコンピュータ PDP-11 で実現している。

前者の方法による検索システムは HITAC-8250、高速漢字プリンタ上で実現している。繰り返すが、この方法ではキーワードの管理が重要である。異形同語、類義語、および広範囲の概念を指す上位概念語（実は、これではキーワードの役割を果さないのだが）が頻出する。

以下、プログラムの概略を示す。

(1) 欧文文献ファイルの検索

著者による検索とキーワードによる検索がある。

著者による検索では、著者を指定すると出力データフォーマット(5)が表示される。各文献ごとにMT出力を選択できる。

キーワードによる検索では、和集合検索と積集合検索を選択することができる。

和集合検索は、指定したキーワード(複数)のどれか一つを持っていればその文献を表示する。積集合検索は、指定したキーワード(複数)のすべてを持っている文献を出力データフォーマット(6)によって表示する。

キーワードは最大5個まで指定できる。

日常語・専門語翻訳ルーチン

「連糸」(string)という語がある。言語学や計算機科学の世界で、「文字や語の連なり」などの意味で用いられる。計量国語学69号に水谷静夫著の「連糸概念の基礎」という論文がある。この論文の内容を示すキーワード「引用, 名前, 結合律……」の中に「連糸」が含まれている。

「連糸」についての論述では水谷論文は最も重要なものだから、その内容と関連がある分野の研究者は当然、その用語も論文も知っているはずである。それ以外の分野の研究者にとって、「連糸」という語は一般的ではないと思われる。その分野やそれに類する内容の文献を探す人にとって、「連糸」という語を用いるのはほぼ不可能といえる。では、文献検索用のキーワードとして、「連糸」は働かないのか。この論文において「連糸」は最もよくその内容を示すキーワードであって、「糸」「連なり」「線」では困る。たとえば、この論文は「言語」や「計算機」について述べているからといって、「言語」や「計算機」をキーワードとするとしよう。たしかに、これらの語によって、この論文は検索されるが、それと同時に、無数の「言語」や「計算機」に関する論文が提示される。計量国語学、国語学、言語学のほとんどの論文は「言語」について述べているからである。これではキーワードとしての働きをなしていないと言えよう。

ここに専門語とそれと同様・類義の一般語との翻訳ルーチンの必要が生ず

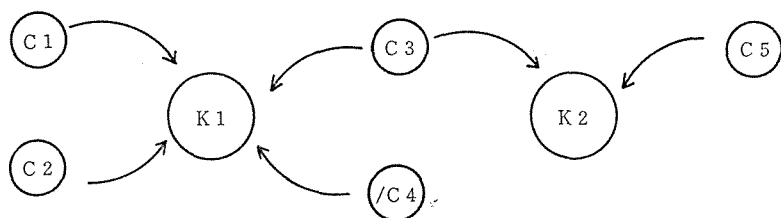
る。

本システムでは以下に述べる方法によってその機能を実現している。

ハードウェアの制約によって、英語キーワードの翻訳をおこなった。

辞書の作成

専門語を一般語に言いかえる辞書を作り、これを利用する際には逆に一般語から専門語を引こうというのである。この場合、専門語を的確に言い換えるたった一つの日常語を用意するより、当らずとも遠からず的な語をいくつか用意する方がより良い。検索の手掛りになる手足を多く持っていた方が、キーワードにたどりつきやすいからである。当らずとも遠からず的な語はシソーラスという上位概念を表わす語や下位概念を表わす語であってもよいが、それではなくてはならないとは考えない。なぜなら、語の意味関係を完全なトリー構造だけで表わすのには無理があると思われるからである。また、トリー構造の体系は新しい概念が入ってきた時に書き直さなければならないという理由もある。メンテナンスが大変なのである。トリー構造では、語を用いる時、常に全体構造を頭に置いていなければならない。語の意味関係はそれほど体系だったものとは思われない。以上の理由により、文献検索で用いられているシソーラスは本システムでは使用しない。それにかわるものとして、ネットワーク構造を用いる。図示すれば次のようになる。



専門語 K 1 は日常語 C 1 ～ C 4 で言いかえられる。専門語 K 2 は日常語 C 3 C 5 で言いかえられる。日常語は C 3 のように K 1, K 2 の複数の専門語を言いかえるのに用いる場合もあるし、ただ一つの専門語を言いかえるのに用いる場合もある。それぞれの日常語と専門語との関係は、たとえば C 4 の / が反対語を示すように示される。関係の種類は現在 / の一つしかない。関係の種類が

多くないということは、これを利用した形態が多くないことを示している。たとえば、「もっと詳しく」とか「上位概念は？」とかのオペレートができないことを示している。しかし、そのようなオペレートは何らかの立場に立つものであるから、その関係が入っている辞書は他の立場から使えないことを示す。本システムは利用の立場を限定しない。日常語 C 3 を介して K 1 と K 2 の関係を得ることができる。

この構造においては、新しい語を登録するのに何のさしさわりもない。ある語の構造をおかすということはない。専門語とそれを示す日常語を入れれば全く自動的に、この構造を作ることができる。

以上の理由により、既存のシソーラスでも、学術用辞典の解説でも何でもデータとして用いることができる。実際には、学術用語辞典は手元にないため、またロジェーのシソーラスは登録語が多すぎ実験システムとして適当でないために採用せず、同義語・反対語辞典 (“The Natural dictionary of English Synonyms and Antonyms” edited by G. Elgie Christ, London) を用いた。例を下に示す。

String, Cord, file, line, thread, twine, row, concatenation, series.

Behaviour, Conduct, deportment, comportment, bearing, demeanour, mien, carriage, air, port, action. *Misconduct*

Discourse, *v. n.*, Talk, converse, speak, expatiate, confer, confabulate, parley.

n., Talk, conversation, speech, disquisition, sermon, dissertation, homily, D. S. A.

先頭の語が専門語に当る語であり、それ以降がそれと同義の日常語である。最後の斜体で示された語は反対語であって、／の記号とともに入力される。小文字筆記体は品詞を示す。

プログラムは、オペレータが入力して来た語がキーワードであればよし、なければ用意された日常語の中を探し、みつければそれと関連のある専門語（キーワード）をすべて示し、選択を要求する。日常語の中にもなければ、ないことを表示する。

図に応答を示す。

SYUUGOO-KEITAI = (* / +)		
+		和集合指定
KEY-WORD IN NAKEREB A END		
ACT	:	} キーワード 指定
TALK	:	
WORD	:	
(TALK)		} そのうち、TALK はキーワードとして 登録さ れていない。
WA NAI		
TUGI NO GO KARA ERABE		} TALK と同義または類義のキーワード と送状させる。
SPEECH		
LANGUAGE		
DISCOURSE		
DISCOURSE		} TALK のかわりに SPEECH と指定
SPEECH		
*****		} 回答
* MIZUTANI SIZUO *	*	
* NO.24 01-19 *	*	
* 1963/03/25 *	*	
* STRATEGIES OF VERBAL ACT *	*	
* S CONCERNING TWO ALTERNA *	*	
* TIVE CONTENTS *	*	
* *	*	

(2) 和文文献ファイルの検索

著者による検索とキーワードによる検索とがある。

どちらもMT入力、漢字ディスプレイでの応答である。シーケンシャルなファイルなので検索時間がかかり実用には程遠い。

プログラムは、漢字プリンタ側のメモリーが少なくてアセンブルができな
いこと、漢字ディスプレイ用の入出力コマンドが難解なことによって、新た
に作成することはあきらめ、すでに用意されている校正・修正システムを利用
した。

(3) PDP-11 による文献検索

このプログラムは Basic で書かれている。

(3-1) データ蓄積 (NIRON 1)

フォーマットはいつでもよい、分かち書きにされていてもいなくてもよ
い。そのまま、ディスクに蓄積される。

(3-2) データの追加 (NIRON 3)

NIRON 1 のファイルにデータを追加する。

(3-3) ファイル名の変更 (NIRON 4)

データの追加が一回のランで終わらなかった場合、文献検索のランをおこなうとき、NIRON 3 出力ファイル名を NIRON 1 出力ファイル名に変更する。

(3-4) 文献検索

入力した文字連続（カナ、英文字、記号何でもよい）がファイル中にあれば、それを全て出力する。出力フォーマットは入力フォーマットと同じである。

プログラムの全てを下に示す。

```
NIRON1.
10 OPEN 'IRON.NAK' FOR OUTPUT AS FILE #1
20 PRINT 'デ-タ タイプセヨ'
30 INPUT A#
40 PRINT #1:A#
50 PRINT 'デ-タハ オワリカ、YES=FINAL、NO=C'
60 INPUT B#
70 IF B#='FINAL' GO TO 90
80 GO TO 30
90 CLOSE #1
100 END

NIRON2.
10 OPEN 'IRON.NAK' FOR INPUT AS FILE #1
20 PRINT 'カンククシクイ コオ タイプセヨ'
30 PRINT 'コハ カナデ、シムルイハ エイモシテ'
40 INPUT A#
50 IF END #1 THEN 110
60 INPUT #1:B#
70 A=POS(B#,A#,1)
80 IF A=0 GO TO 50
90 PRINT B#
100 GO TO 50
110 CLOSE #1
120 END

NIRON3.
10 OPEN 'IRON.NAK' FOR INPUT AS FILE #1
20 OPEN 'IRON1.NAK' FOR OUTPUT AS FILE #2
30 IF END #1 THEN 60
35 INPUT #1:A#
40 PRINT #2:A#
50 GO TO 30
60 PRINT 'デ-タハ ハンクセヨ'
70 INPUT A#
80 PRINT #2:A#
90 PRINT 'オワリカ、YES=FINAL、NO=C'
100 INPUT B#
110 IF B#='FINAL' GO TO 130
120 GO TO 70
130 CLOSE #1
140 CLOSE #2
150 END

NIRON4.
10 OPEN 'IRON.NAK' FOR INPUT AS FILE #1
20 OPEN 'IRON.NAK' FOR OUTPUT AS FILE #2
30 IF END #1 THEN 60
35 INPUT #1:A#
40 PRINT #2:A#
50 GO TO 30
60 CLOSE #1
70 CLOSE #2
80 END
```

4. おわりに

このシステムは狭い範囲での実用と文献検索自体の実験を目的として作られた。付録の文献目録、著者名索引、件名索引を作成できたという点では実用目的は達成されたと言ってよからう。オンライン文献検索は実験であったが、新しい試みができたという点で今回の目標は達成した。

次に、反省をのべる。

(1) 使用したハードウェアシステムの制限があった。漢字ディスプレイの自由な利用ができないため、和文文献検索の実験が十分できなかった。また、

データの修正が容易でなかった。英数字の2/3角文字が印字できなかった。

(2) 本来、このようなシステムはTSSによって動かすべきだと考えるが、ハード、ソフト上の制限により実現できなかった。文献検索やデータの修正はオペレータの思考作業時間が必要であり、かつ、複雑なプログラムが動かなければならぬ。

(3) これはバッチ処理むきではない。かなりの量のデータが各種に用いられ、かつ、随時データの追加がおこなわれる。したがって用いられた結果、エラーが発見されることが多い。このような場合データベース構造をとるべきだと考えるが、今回は実現できなかった。

(4) 文献につけられたキーワードに問題があった。語形・表記のゆれがあることと、キーワードらしくない語がつけられたことである。

(5) このようなマン＝マシンシステムでは、スピードや成功率よりいかに人間工学的な配慮がはからわれているかが重要である。使いにくいシステムは結局使われなくなる。この点の配慮は払われていない。

(6) 当然のことながら、自動インデキシング等に向かわなければならない。将来、一貫処理システム(文献12)とのドッキングを考えているが、今回は行わなかった。

このシステムは、昭和49年国立国語研究所図書館司書の大浪由起夫氏(現在日本学術会議)と筆者との共同研究で始めた。システム名IRONは、Information Retrieval, Onami & Nakanoの頭文字に由来する。入力項目の選定やシステムの概要を検討した段階で、種々の理由により中断したが、昭和51年になって筆者がこの研究・開発を再開し今日にいたった。したがって、大浪氏の協力・助言がなければ、この研究は始められなかったし、完成も見なかったであろう。深く感謝の意を表するものである。

また、プログラム・データの作成にあたっては、研究補助員長田厚子・アルバイタ藤生典子の両嬢の協力が大きかった。記して感謝の意を表する。

参考文献

1. 森下四郎：AIDOR インライン教育文献検索システム ——AIDOR システムとAIDOR 教育シソーラスについて——，情報管理 Vol. 20 No. 8~9, 1977
2. 計量国語学会：第1号—第80号索引，計量国語学11巻附録，1977
3. 中野洋：国語研試作版文献検索システムの紹介，計量国語学会第二十一回大会研究発表要旨，計量国語学11巻3号，1977
4. 高松忍・藤田米春・西田富士夫：係り受け関係に基づく文献の検索，情報処理 Vol. 19 No. 12, 1978
5. 遠藤勉・田町常夫：文献の主題概念を記述するための言語構造の分析，電気通信学会技術研究報告A L77—46, 1977
6. 堂下修司・西田豊明：アブストラクトによる文献検索，電子通信学会技術研究報告A L78—52, 1978
7. 飯田三郎・吉田雄二：名古屋大学大型計算機センターにおける文献検索サービスシステムについて，電子通信学会技術研究報告A L78—52, 1978
8. 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，昭和44年
9. 東京大学大型計算機センターTOOL—IR プロジェクト利用者マニュアル（上，下）第2版，昭和51年
10. 国内のオンライン情報サービス一覧。情報処理 21-2, 1980.
11. 中井浩：オンライン情報サービスの現状と今後。情報処理 21-2, 1980.
12. 中野洋：言語処理における一貫処理の研究。電子計算機による国語研究IX. 1978.