



## シソーラスを用いた研究室向き会話型文献検索システム THEODORES\*

平松敏祐\*\* 福村晃夫\*\*\*

### Abstract

Thesaurus oriented documents retrieval system (THEODORES) is developed. The THEODORES accepts different thesauri which represent various aspects of subjects, and the performance of retrieval depends significantly upon the quality of the thesaurus. Modifications of the thesaurus are possible through a conversation with the data base. Exploration of a model structure of the subject is also possible by the effective use of the thesaurus.

A logical connector and a pertinent data structure are designed to represent the Boolean operators and the thesaurus respectively. The THEODORES is possible to interpret an arbitrary query which is expressed by a combination of key words and the Boolean operators including NOT function. It is found that the efficiency of retrieval depends on the order of sequence of key words as well as on the length of key words of an input query.

### 1. はじめに

学術情報の激増にともなって、学術情報の組織化、蓄積、および効率のよい検索方式の開発が要求されるようになってきた。本研究は、

(1) 小型計算機を用いて、研究室規模の文献データベースを、シソーラスを活用しながら会話的に検索する。

(2) 検索システムの有効性と、将来の発展性を調査する、の2点に眼目をおいた、問題向き文献検索システムを開発することを目的としている。はじめに、研究の背景について述べる。

#### 1.1 システム開発の背景

環境学は総合科学または、学際的な研究の一分野で、多方面にわたる大量のデータを扱う必要がある。最近、環境学のための情報システム SMLESIS (SMLES In-

formation System) が島津<sup>1)</sup>により提案された。SMLESIS はモデル構造の設計と情報検索の両機能を備え、その基本的な考え方と特徴は次の通りである。

(1) 数値情報、文章、文献、研究者の住所、シミュレーション・プログラムが格納され検索できるデータベースである。SMLESIS にはシミュレーションのプログラムだけでなく、各シミュレーション・システムのモデル化の手法も含まれている。

(2) 研究者の哲学を反映する自作のシソーラスを検索の際に利用する。検索の適合度が低い場合にはシソーラスを変更して適合率を高める。

(3) データベースの拡張と修正が容易であり、検索システムが手軽に使用できる。

(4) 各プロジェクトごとに研究者の名前と住所が組織的に検索できる。

(5) 学際的な研究者の養成のための教育を柔軟なシソーラスを通して行う。

SMLESIS のプロトタイプは、浦部<sup>2)</sup>等によって、共同利用の大型計算機システムの上で開発された。このシステムは、FORTRAN で記述され、バッチで動く。上に述べた基本的な諸機能は実現されたが、質問文の表現に制限があったり、処理速度等が問題点とし

\* Thesaurus Oriented Interactive Documents Retrieval System (THEODORES) for Laboratory Data Base by Toshisuke HIRAMATSU (Computation Center, Nagoya University) and Teruo FUKUMURA (Department of Information Engineering, Nagoya University).

\*\* 名古屋大学大型計算機センター

\*\*\* 名古屋大学工学部情報工学科

て残った。

## 1.2 システムの設計方針

本研究で開発された検索システム THEODORES (Thesaurus Oriented Documents Retrieval System) は、前述の SMLESIS の基本概念に沿った、つぎの方針に沿って設計された。

(1) 使用計算機として小型機を用いる。その理由は

(a) 価格、性能面の条件が、本研究の目的に合うように向上してきた。とくに、文献検索では、ポイント操作、パターンマッチング、ファイルへのアクセス等が主な操作であるため、とくに大型機を必要としない。

(b) 研究室規模の文献を検索対象とするため、それほど大きなファイルを必要としない。すなわち、研究室固有の方針に沿って、文献を精選して格納するから、文献数は、千件から一万件程度ですむ。

(c) 会話による発見の問題解決の利用形態をとるには、大型機のオペレーティング・システムの制約から逃れることが望ましい。

(2) システムの機能上の特徴を次のように定める。

(a) 検索は会話形式とする。

(b) 検索処理速度を上げるため、アセンブリ言語を用いる。

(c) 特定の key word による検索ではなく、任意の key word を用いて質問文が構成できるようにする。質問文は、ブーリアン演算子と、key word との自由な組み合わせでつくられるようにする。

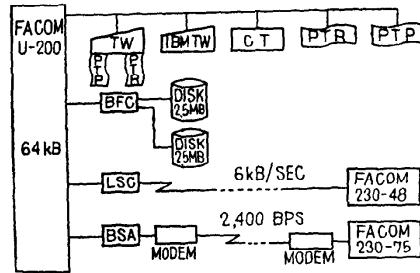
(d) (c)のため、インバーテッド・ファイルを用いることなく、全文献に対してしらみつぶし検索を行う。そのさい、key word の部分マッチングも可能にする。

(e) 処理効率の向上を考慮して、質問文の内部表現には論理結合子を、シソーラス・データの内部表現には、サーキュラー・リストと木構造を組み合わせた、特殊リスト構造を採用する。

(f) 質問文の、人による理解と修正を助けるため、Pretty-Printing を用意する。

(g) シソーラスの内容変更と拡張とを容易にするため、会話型文脈エディタ<sup>2)</sup>を利用する。

(h) 文献とシソーラスの入力データ形式は、SMLESIS と同じにする。また、性能比較のため、使用するデータも同じにする。



BFC: Basic File Channel  
BSA: Binary Synchronous Adapter  
LSG: Low Speed Interface Controller

Fig. 1 Hardware constitution.

## 2. 機器構成

THEODORES がインプリメントされた小型計算機システムの機器構成は、Fig. 1 に示されるとおりで、CPU、64kB のコアメモリ、タイプライタ 2 台、光電式の高速 PTR/PTP、カセット・テープ装置、およびディスク装置 (2.5MB\*2) からなる。また、大型システムとは、転送速度 2,400BPS の MODEM 経由で、中型システムとは、低速チャンネル間結合 (最大転送速度 6kB/秒) によって結合されている。

## 3. シソーラスと質問文のデータ構造

### 3.1 シソーラス

シソーラスの拡張と変更を容易に行うためには、入力するシソーラスのデータ形式をできるだけ簡単にする必要があり、Fig. 2 (次頁参照) に、ENVIRONMENT と POLLUTION に関するシソーラス・データの記述例を示す。この indentation を用いたデータ形式は、一般に木構造を表現するのに適しており、研究者がシソーラスの内容を読んだり、記述したり、さらに修正するのに便利である。各語のレベルは各々の語の前に置かれたブランクの数で識別し、同義語あるいは関連語は、その語の後に括弧でくくって列挙する。たとえば、Fig. 2 において SOX, SULFUR OXIDE, SO<sub>2</sub>等は、SULFUR OXIDE の同義語あるいは関連語を意味する。付加する同義語は、客観的な語の他に、研究者の主観的な語であってもよい。

THEODORES は、シソーラス・データが格納されているファイル名が指定されると、シソーラスをすみやかに検索するために、Fig. 2 に示したようなカード・イメージのシソーラス・データをディスクから読み、Fig. 3 (次頁参照) に示すようなリスト構造として主

ENVIRONMENT (ENVIRONMENTAL)  
 ENVIRONMENTAL QUALITY (QUALITY OF ENVIRONMENT)  
 ENVIRONMENTAL INDICATOR  
 INDEX BIOTA  
 INDEX OF HUMAN CONCERNS (INDEX OF HUMAN ENVIRONMENTS)  
 ENVIRONMENTAL IMPACT  
 ENVIRONMENTAL ASSESSMENT  
 RISK ASSESSMENT  
 ENVIRONMENTAL MONITORING  
 CONSERVATION (NATURAL CONSERVATION, NATURAL RESERVATION)  
 GENETIC MATERIAL  
 WILD LIFE (NATIONAL PARK)

POLLUTION  
 WATER POLLUTION (WATER QUALITY)  
 BOD  
 COD  
 SS

AIR POLLUTION  
 INVERSION LAYER  
 SULFUR OXIDE (SO<sub>2</sub>, SULFUR DIOXIDE, SO<sub>2</sub>)  
 NITROGEN OXIDE (NO<sub>x</sub>, NITROGEN DIOXIDE)  
 CARBON MONOXIDE  
 HYDROCARBON  
 PHOTOCHEMICALS (OXIDANT)  
 OZONE (STRATOSPHERE, SST)  
 AEROSOL  
 PARTICULATE (VISIBILITY)

SOIL POLLUTION  
 SOIL-PLANT

SOLID WASTE (WASTE DISPOSAL)  
 PLASTICS  
 GARBAGE  
 METAL  
 PAPER

THERMAL POLLUTION (THERMAL WASTE)  
 NOISE

Fig. 2 Examples of original indented thesaurus data concerning with the environment and the pollution.

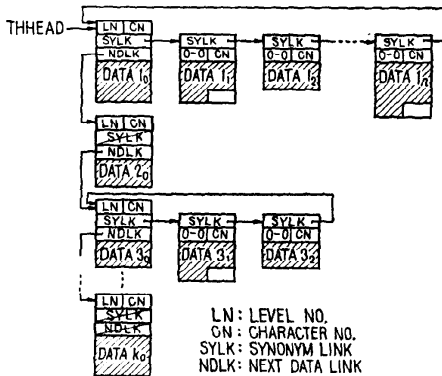


Fig. 3 List representation of thesaurus data.

記憶上に展開する。Fig. 3において THHEAD はシソーラス・データの先頭を示すポイントであり、SYLK, NDLK は、それぞれ同義語へのリンク、次のデータへのリンクを意味する。同義語や関連語は、シソーラス・データのリスト構造上でのサーキュラー・リストとして表現する。たとえば DATA 1<sub>0</sub>, DATA 1<sub>2</sub>, ..., DATA 1<sub>n</sub> などは、DATA 1<sub>0</sub> の同義語または関連語を意味する。LN と CN は、その語のレベルの深さと、文字数を示す。同義語または関連語を持たない語の SYLK 部は 0 である。シソーラス・データの最終ブロックは、NDLK 部が 0 か否かで識別する。

### 3.2 質問文

THEODORES は、ブーリアン演算子 (AND, OR, NOT) と key word と括弧とを用いて表現される任意の質問文を解釈する。ここでブーリアン演算子は、次の記号で表現する。

AND 演算子…\*, OR 演算子…+, NOT 演算子…-。内容は同じでも、質問文の記述の仕方が質問者によって異なる場合がある。たとえば「A⊕B」という質問文は次のように、いく通りにも記述される。

$$\begin{aligned} A \oplus B &= A * (-B) + (-A) * B \\ &= (A + B) * (-A * B) \\ &= (A + B) * (-A + (-B)) \\ &= -(A * B) * (-(-A * (-B))) \\ &= -(A * B + (-A) * (-B)) \end{aligned}$$

ここで「⊕」は排他的論理和を意味し、A と B は、それぞれ異なった key word を意味する。

THEODORES は、上記のいずれの表現も解釈することができる。

一つの key word の長さは 60 文字以内に限定しているが、内容については何等制限はなく、どのような単語や熟語でも検索の対象にできる。質問文中のブランクは、二つの key word の間の一つのブランク以外は、すべて無視する。たとえば、次の質問文

“WATER POLLUTION \* SIMULATION + BIO-CHEMISTRY”

は、次の質問文と等価である。

“WATER POLLUTION\*SIMULATION+BIO-CHEMISTRY”

同じマイナスの記号「-」が NOT 演算子とハイフンの両方の意味に使用されるが、それらは文脈によって識別する。たとえば、「-GEOCHEMISTRY」の「-」は、「NOT GEOCHEMISTRY」を意味し、上の例の「BIO-CHEMISTRY」の「-」は、ハイフンを意味する。

質問文は、Fig. 4 (次頁参照) に示すようなリスト構造に変換され、主記憶に格納される。質問文のリスト構造は、論理結合子と key word ブロックから構成される。Fig. 4 は、「A<sub>0</sub>\*(-B<sub>0</sub>+C<sub>0</sub>)」型の質問文の内部表現の例を示す。A<sub>1</sub>, ..., A<sub>x</sub> および C<sub>1</sub> は、それぞれ、指定されたシソーラスから発生された A<sub>0</sub> および C<sub>0</sub> に関する同義語、または関連語を意味する。論理結合子は 3 語から構成され、その詳細な構成を Fig. 5 (次頁参照) に示す。AND 結合子と OR 結合子は、論理結合子の ANDOR 部の値で識別する。

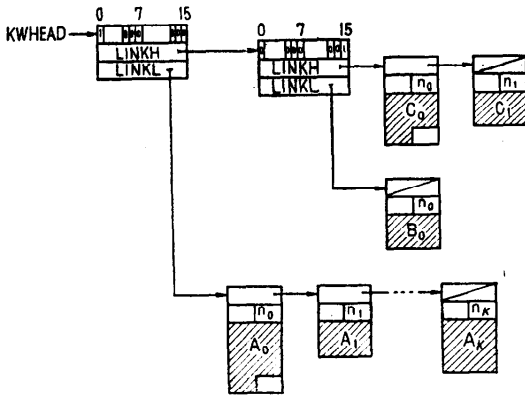


Fig. 4 An example of internal representation of "A\*(-B+C)" type query connected by logical connectors.

ANDOR=0...OR 結合子,  
ANDOR=1...AND 結合子.

LINKH と LINKL は、共に、リンク部で他の論理結合子、または key word ブロックの先頭のアドレスを示す。VALH と PASH, NEGH は、LINKH に対する制御ビットで、その値により次のような意味を持つ。

PASH=0...LINKH に対するマッチング・テストが済んでいない。

PASH=1...LINKH に対するマッチング・テストが済んでいる。

NEGH=0...LINKH に対する key word が肯定語である。

NEGH=1...LINKH に対する key word が否定語である。

VALH の値は NEGH の値によって意味が異なる。

VALH=1...NEGH=0 であり、LINKH に続く key word ブロックの少なくとも 1 つがマッチング・テストに成功するか、または NEGH=1 であり、LINKH に続く key word ブロックのすべてがマッチン

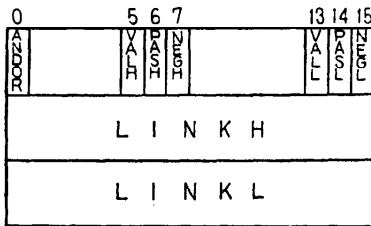


Fig. 5 Logical connector.

グ・テストに失敗するとき。

VALH=0...上記以外の条件のとき。

PASH と NEGL, VALL の意味は、LINKL に対して上に述べたのと同様である。論理結合子 AND の論理条件は、LINKH と LINKL の両方が真となる時のみ真となり、論理結合子 OR の場合には、LINKH または LINKL のどちらかが真であれば真となる。

### 4. 2分木の探索と Pretty-Printing

#### 4.1 2分木の探索

マッチング・テストまたは Pretty-Printing を行うために、Fig. 4 に示されているような 2 分木を探索するとき、論理結合子と key word ブロックとの識別が必要となる。論理結合子と key word ブロックは、そのアドレスで識別する。論理結合子は、特別なシステム領域に置き、key word ブロックと識別する。一度の質問文で使用できる論理結合子の最大数は 25 個に制限している。

#### 4.2 Pretty-Printing

Pretty-Printing は、複雑な質問文の論理関係を図示し、入力された質問文のユーザによる理解とデバッグを助ける。複雑な質問文とその Pretty-Printing の例を Fig. 6 に示す。ここでブーリアン演算子: \*\*AND\*\* と \*\*OR\*\* は、key word と視覚的に識別するために赤い文字で印字される。各々の key word の前に置かれた記号 - は、その語が否定語であることを意味する。

THEODORES に適用されている Pretty-Printing のルールは次の通りである。

```

**PLEASE TYPE IN KEYWORD(S)**
1>(-AAAA+BBBB*(-CCCC+(-DDDD))+EEEE)
2>+(1111+2222*3333)#
*INPUT KEYWORD(S) IS(ARE) ---
(-AAAA+BBBB*(-CCCC+(-DDDD))+EEEE)+(1111+2222*3333)#
*GENERATED KEYWORD(S) IS(ARE) ---
-AAAA
**OR**
BBBB
**AND**
-CCCC
**AND**
DDDD
**OR**
EEEE
**OR**
1111
**OR**
2222
**AND**
3333
    
```

Fig. 6 An example of long and complex query and its pretty-printing.

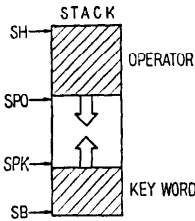


Fig. 7 Operator stack and key word stack.

(1) すべての行は、いくつかのスペースで始まり、そのスペースの数が、Fig. 4 に示されているような木構造の上での各々の key word の深さを示す。この深さは、木を探索するさいのスタックの深さを意味し、1つのスタックの深さに対して4つのスペースを出力する。

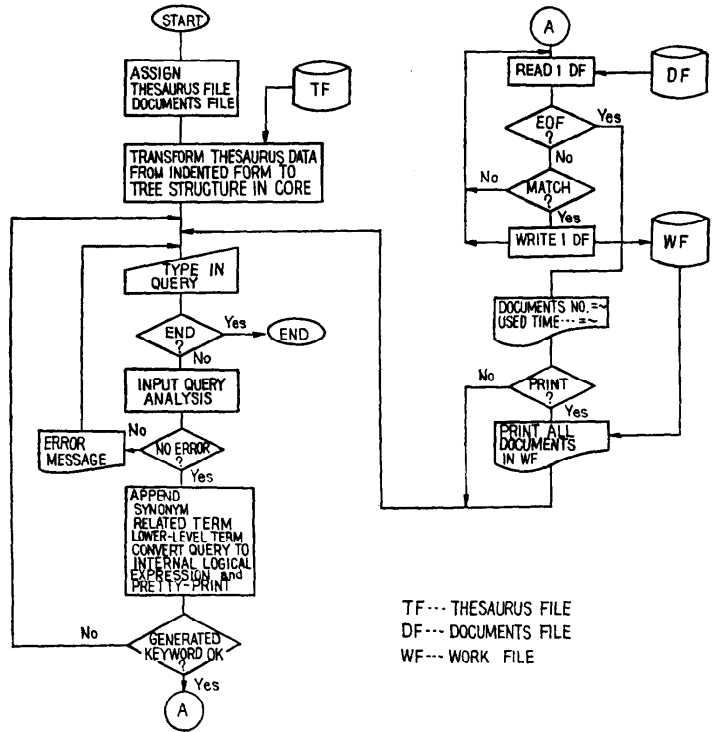
(2) 与えられた key word の順に印字し、もしシソーラスから同義語または下位レベルの語が発生されたならば、与えられた key word の下に並べて印字する。

4.3 スタック

演算子スタックと key word スタックの二種類のスタックが存在し、質問文の解析と Pretty-Printing のために使用される。演算子スタックは、ブーリアン演算子を格納し、key word スタックは key word の先頭アドレスを格納する。これらの二つのスタックは、共に同じ領域 STACK を両端から使用する。STACK の大きさは50語である。Fig. 7 に二つのスタックの構成図を示す。SH と SB はそれぞれ STACK の先頭と底の番地を示すポインタで、同時に演算子スタックと key word スタックの底を示すポインタでもある。SPO と SPK は、それぞれ演算子スタックと key word スタックの現在の位置を示すポインタである。

5. システム全体の流れ

Fig. 8 に THEODORES の概略の流れ図を示す。最初に、検索の対象にする文献ファイル (DF) と、使用するシソーラス・ファイル (TF) とをファイル名で指定する。もしシソーラスを必要としなければ、TF のファイル名として「NO」と答える。作業用ファイル (WF) は、システムで暗黙のうちに定義されている。シソーラスの検索を高速に行うために、THEO-



TF--- THESAURUS FILE  
DF--- DOCUMENTS FILE  
WF--- WORK FILE

Fig. 8 A general flow chart.

DORES は、指定されたファイルからシソーラス・データを読み、Fig. 3 に示すようなリスト構造に変換して、主記憶上に展開格納する。質問文が入力されるとシンタックス・チェックを行い、エラーが発見されると再び質問文の入力フェーズにもどる。エラーがなければ、シソーラスから同義語や関連語、下位概念の言葉を付加する。つぎに、このように生成された質問文に対して、Pretty-Printingを行う。質問者は Pretty-Printing の結果を見て、生成された質問文の内容が意図したものであるかどうかを判断する。もし生成された質問文が満足できなければ、質問を変更するか、さらに、必要ならばシソーラスを修正する。このようにして、満足できる質問文が生成されるまで変更または修正を繰り返す。ただし、シソーラスを修正するのは、研究者がシソーラスに記述した考え方に間違いを認めた場合に限られ、検索結果が好ましくないからといって、むやみにシソーラスを修正するのは本質的な解決策とはならない。

満足できる質問文が生成されると、THEODORES は、文献ファイルから1つの文献データを読み出し、生成された質問文との間に論理チェックを行う。論理

```
//TE02
**WELCOME TO THEODORES(V2-L2.1975)**

*THESAURUS FILE NAME ? =THESAU
*IR DATA FILE NAME ? =IRALL
*ANY CONTROL ? (Y/N)=Y
  C1. APPENDED WORD(S) PRINT ? (Y/N)=Y
  C2. PARTIAL MATCHING ? (Y/N)=N

**PLEASE TYPE IN KEYWORD(S)**
I>MODEL * TOXITY#

*INPUT KEYWORD(S) IS(ARE) ---
MODEL*TOXITY#

*GENERATED KEYWORD(S) IS(ARE) ---
MODEL
+
SIMULATION MODELLING
+
MODELLING
+
OPTIMIZATION
+
SIMULATION
+
SYSTEMS DYNAMICS
+
SD
**AND**
TOXITY
+
TOXICOLOGY
+
MINOR COMPONENT
+
TRACE ELEMENT
+
PESTICIDE
+
INSECTICIDE
+
HEAVY METAL
+
MERCURY
+
CADMIUM
+
LEAD
+
ARTIFICIAL CHEMICALS
+
PCB

*GENERATED KEYWORD(S) OK ? (Y/N)=Y

*DOCUMENT(S) NO..      4

USED TIME---  0 MIN. 39 SEC.  955 MSEC.

*DOCUMENT(S) PRINT ? (Y/N)=Y
```

条件が満足されれば、その文献を作業用ファイル(WF)に出力し、次の文献の論理チェックに進む。このようにして、すべての文献に対する検索が終了すると、論理条件を満足した文献数と検索に要した時間を出力し、文献の詳細を出力するか否かをたずねる。必要なら文献の内容を出力し、必要としなければ質問文の入力フェーズにもどる。

## 6. 実験および検討

THEODORES は、主プログラムと 52 個の副プログラムからなり、ステップ総数は約 4,600 で、およそ 12 kB(バイト) を占める。主記憶上にシソーラス用領域として 20 kB を確保すると、カードで 800 枚程度のシソーラスを展開でき、key word 用領域として 2 kB を確保すると、一度に約 150 個の key word を検索の対象にできる。THEODORES 自身の作成、および文献データ、シソーラス・データの編集に会話型文脈エディタ<sup>2)</sup>が使用され、その有効性が実証された。

システムの実用性のテストのために、SMLESIS で使用されている三種類のシソーラス (1. SCOPE/MAB プロジェクト: カードで 103 枚, 2. 環境学: カードで 615 枚, 3. モデル・アンケート: カードで 93 枚) と、二種類の文献データ (1. IRALL.....1,026 件の文献: カードで約 2,400 枚, 2. IRNDATA...1,652 件の文献: カードで約 4,000 枚) とを使用して検索のテストを行った。これらのデータ\*は、主に地球物理学や社会経済学、生態学、環境学等いくつかの学問分野にまたがる学際的な研究分野の文献を集めたもので、計算機間結合<sup>3)</sup>により中型機から小型機のディスクの中に転送格納された。

THEODORES の検索時間の評価のために、上記の二種類の文献データ・ファイル (IRALL と IRNDATA) に対して 8 種類 (内容的には 4 種類) の簡単な検

1. EBERHARDT, L.L., R.L. HEEKS AND T.J. PETERLE (1970) DDT IN FRESHWATER MARSH - A SIMULATION STUDY. BATTELLE MEMORIAL INST. PP.53. (SL199-BIOGEOCHEMICAL CYCLE, TOXICITY.)
2. IRIYAMA, J. AND M. YAMAMURA (1970) ENERGIES EMITTED IN RECONSTRUCTION OF VENUS, MARS, AND MERCURY. JOUR. SEISMOL. JAPAN 23, 152-153. (SMLES29-JAPANESE, THERMAL PLANET, MATHEMATICAL MODEL.)
3. LEISTRA, M. (1973) COMPUTATION MODELS FOR THE TRANSPORT OF PESTICIDES IN SOIL. RESIDUE REV. 49, 87-130. (PEST CONTROL MODEL.)
4. SASABA, T., K. KIRITANI AND T. URABE (1973) A PRELIMINARY MODEL TO SIMULATE THE EFFECT OF INSECTICIDES ON A SPIDER-LEAFHOPPER SYSTEM IN THE PADDY FIELD. RES. POPULATION ECOLOGY 15, 9-22. (SMLES60, SYSTEM ECOLOGY MODEL, PEST CONTROL.)

(continued)

\* これらのデータは、名古屋大学理学部の島津康男教授をリーダーとする SMLES (Seamless Earth Science) 研究グループに所属している。

**Table 1** Comparison of retrieval time (use time) for two document files.

	IRALL (1,026 documents)	IRNDATA (1,652 documents)
A*(-A)	10 sec. 189 msec.	17 sec. 451 msec.
-A*A	11 sec. 792 msec.	19 sec. 412 msec.
A+(-A)	31 sec. 903 msec.	55 sec. 359 msec.
-A+A	31 sec. 407 msec.	53 sec. 973 msec.
ABC*(-ABC)	12 sec. 986 msec.	22 sec. 544 msec.
-ABC*ABC	17 sec. 400 msec.	30 sec. 83 msec.
ABC+(-ABC)	32 sec. 732 msec.	58 sec. 27 msec.
-ABC+ABC	31 sec. 950 msec.	55 sec. 754 msec.

索を行った結果を **Table 1** に示す。さきに大型機で開発されたバッチ・システムとくらべ 2~3 倍の処理効率を持っている。Table 1 から明らかなように、質的には同一の質問文でも、与える key word の順序により検索時間に相違が出てくることが分る。AND 結合の場合が特に顕著であるが、一般的な結論として、AND 結合の場合には、狭義→広義の順に、OR 結合の場合には、その逆順に key word を与えることにより検索時間が短縮されるといえる。

THEODORES は、文献データとシソーラスの内容を変えれば、どのような研究分野にも適用できる。たとえば、現在、名古屋大学大型計算機センター図書室の洋書、マニュアル等の文献検索に利用されている。

実際の検索例の一つを **Fig. 9** に示す。ここに示す例の質問文は、「MODEL\*TOXITY」である。//TED 2 の TED は THEODORES の愛称であり、2 はシステムのバージョンが 2 であることを示す。この例の最初に指定した THEAU と IRALL は、それぞれ、シソーラスと検索の対象にする文献データとのファイ

```

**PLEASE TYPE IN KEYWORD(S)**
1>MODEL * TOXITY
2>*(-PLANET)#

*INPUT KEYWORD(S) IS(ARE) ---
MODEL*TOXITY*(-PLANET)#

*GENERATED KEYWORD(S) IS(ARE) ---

MODEL
+
SIMULATION MODELLING
+
MODELLING
+
OPTIMIZATION
+
SIMULATION
+
SYSTEMS DYNAMICS
+
SD
**AND**
TOXITY
+
TOXICOLOGY
+
MINOR COMPONENT
+
TRACE ELEMENT
+
PESTICIDE
+
INSECTICIDE
+
HEAVY METAL
+
MERCURY
+
CAESIUM
+
LEAD
+
ARTIFICIAL CHEMICALS
+
PCB
**AND**
-PLANET

*GENERATED KEYWORD(S) OK. ? (Y/N)-Y

*DOCUMENT(S) NO.- 3

USED TIME--- 0 MIN. 39 SEC. 732 MSEC.

*DOCUMENT(S) PRINT ? (Y/N)=Y

1. EBERHARDT,L.L.,R.L.MEEKS AND T.J.PETERLE (1970) DOT IN FRESHWATER MARSH - A SIMULATION STUDY.
BATTTEL MEMORIAL INST. PP.53.
(SL199.BIOGEOCHEMICAL CYCLE.TOXITY.)

2. LEISTRA,M. (1973) COMPUTATION MODELS FOR THE TRANSPORT OF PESTICIDES IN SOIL.
RESIDUE REV. 49.87-130.
(PEST CONTROL.MODEL.)

3. SASABA,T..K.KIRITANI AND T.URABE (1973) A PRELIMINARY MODEL TO SIMULATE THE EFFECT OF INSECTICIDES
ON A SPIDER-LEAFHOPPER SYSTEM IN THE PADDY FIELD.
RES. POPULATION ECOLOGY 15. 9-22.
(SMLES60.SYSTEM ECOLOGY MODEL.PEST CONTROL.)

**PLEASE TYPE IN KEYWORD(S)**
1>#
*INPUT KEYWORD(S) IS(ARE) ---
#
**END OF IR**
J STOP 0000

```

**Fig. 9** An example for the query: "MODEL\*TOXITY"

ル名を意味する。また検索の結果、文献を出力するさいには付加語をも一緒に出力し、検索のさいの部分マッチは行わないことを指定している。入力質問文の終りは「@」記号で指定する。

最初の質問文「MODEL\*TOXITY」に対する Pretty-Printing の結果において、「SIMULATION MODELLING」と「MODELLING」は与えた key word 「MODEL」の同義語であり、「OPTIMIZATION」や「SIMULATION」、「SYSTEMS DYNAMICS」、「SD」は、「MODEL」に対する下位レベルの言葉であって、ともに指定されたシソーラス THESAU から発生された key word である。

第一回目の質問文で、論理条件を満足する4件の文献の内容が付加語と共に出力されている。これらの4件の文献の内容をよく見ると、1件だけ質問の意図に反した文献が検索されているのが分る。その文献は、MODEL と MERCURY の二つの key word で論理条件を満足した二番目の文献である。これは、システム・エラーではなくて、いわばセマンティック・エラーである。この場合「MERCURY」には、二つの意味、すなわち、金属の MERCURY、つまり水銀と、惑星の MERCURY、つまり水星とがある。そこで第二回目の質問文で、「\*(PLANET)」という、意味を考えた上での否定の key word を追加した結果、質問の意図に適合した3件の文献が出力されている。THEODORES は、このような会話を繰り返して、必要な文献を効率よく検索する。

## 7. むすび

THEODORES は、Indentation を用いて記述した木構造のシソーラスを利用している。木構造のシソーラスは人間にとって直観的で理解しやすく、かつ計算機で処理しやすいが、より複雑なシステム構造を扱うのには不十分である。一方、人間の頭の中にあるシソーラスは、多次元の網構造として複雑にからみ合っているように思われる。したがって、それをより直接的に反映するネット・シソーラス、あるいは、さらに複雑なシソーラス構造を図形化して直観的に把握できるようなグラフィック・シソーラスの開発が将来の課題として残されている。

本研究は、名古屋大学理学部の島津康男教授をリーダーとする「環境学のための情報システム：SMLES-IS」の一環として行った。

## 参考文献

- 1) Y. Shimazu, Ed.: Simulation Modelling in Environment Problems-Activities in Japan, SMLES RESEARCH GROUP, Nagoya Univ., p. 166 (1976).
- 2) T. Hiramatsu: Interactive Context Editor by Buddy System, Memoirs of Research and Development Division, Computation Center Nagoya Univ., No. 1, pp. 57~83 (1975).
- 3) 平松敏祐: 小型計算機を介した中・大型計算機結合, 昭和50年度電気関係学会東海支部連合大会, p. 236 (1975).
- 4) 浦部達夫, 島津康男, 茂泉俊夫: 学際研究用データ・ベース・システム SMLESIS, 名古屋大学理学部地球科学教室, p. 33 (1976).

(昭和53年2月15日受付)