

会話型ネット・シソーラス処理システム NESSY-I†

平松敏祐^{††} 福村晃夫^{†††}

会話型のネット・シソーラス処理システム NESSY-I を開発した。本システムは、対象とする研究分野において使用される用語や概念を、研究者自身の手により、ネット・シソーラスとして表現し、表現された概念構造の論理的な矛盾の発見や修正、新たな概念の追加などを計算機との会話を繰り返しながら行い、扱う研究対象を適確に把握したり、問題点の分析や抽出、さらには解決策の早期発見を短時間のうちに達成することを目的とする。

ネット・シソーラスは、二つのノード名と四種類の記号 (>と<, =, <>) の一つを用いて記述した二項関係の集合として定義する。ネット・シソーラス上の一つまたは二つのノード名を指定することにより、そのノード名に関する上位語、下位語、関連語、道筋の検索、ループの検索などを行う。ネット・シソーラスの生成や修正には、会話型コマンドを利用する。

網構造は、マトリックスではなくリレーション・テーブルを用いて内部表現する。ネット・シソーラスの構造解析を行うために、リンク・ブロックを考案した。用語を高速に処理するためには、ハッシング技法を用いる。

具体的な応用として、NESSY-I を「地震災害とその対策」に関するネット・シソーラスの解析に適用し、NESSY-I の有用性が実証された。

1. はじめに

1.1 本研究の目的

研究の成長期を過ぎ、使用される用語の内包する概念や意味が定着してきた特定の研究分野に対しては、各分野における長年の知的活動の所産であるほう大なシソーラスの蓄積がなされている。これらのシソーラスが、個々の研究分野の概念構造の把握や情報検索に果たす役割の重要性については言をまたない。

しかし最近のように、諸種の新研究分野が諸所に現われて急速な発展を見せたり、あるいは、いくつかの研究分野にまたがる学際的な研究領域が流動的に構成されるようになると、既存のシソーラスには記述されていない新用語が当然のように作られるし、また、すでに記述されている用語でも、従来とは異なった意味で用いられるようになる。しかも、その程度は時代とともに強まるものと考えられる。

したがってこのような状況のもとでは、問題の核心にふれる諸概念、諸用語の間の因果関係をつねに明確化していくことによって研究の対象とする問題点を適確に把握し、かつ、問題の分析から解決策の発見に至るまでの過程を短時間に終わらせるように努めること

が、きわめて重要なこととなってくる。

本論文は、このような要請に応じて、従来のシソーラスによる概念表現法を用いて問題向き小型シソーラスを短時間に、しかも柔軟性を持たせて構成する方法を提案し、かつ、それを実際の問題に適用することを目的としている。

1.2 従来の研究の動向

本研究の目的に合うシソーラスは網構造でなくてはならないが、この形のシソーラスについては、古くは、その表現法に関するいくつかの提案がなされており、例えば、オランダ陸軍技術ドキュメンテーション情報センターの用語を同心円の形式で表示した「円形式シソーラス」^{1),2)}や、Lauren Doyle による、用語間の関連度の強さを矢印グラフで記述した「意味の道路地図」^{3),4)}、アルファベット順表示と図式表示の両方を含んだ「EURATOM シソーラス」^{5),6)}などがある。

また後に、シソーラスの具体的なコンピュータ処理が考えられるようになってからは、Hersey と Hammond の水資源シソーラス⁷⁾、Martinez らの採油・製油シソーラス⁸⁾、Wolf-Terroine らの癌のシソーラス⁹⁾、Rolling によるシソーラスの編集法¹⁰⁾などがある。その他に、三浦によるシソーラスの数学的解析¹¹⁾ Gotliebらや Augustson らのグラフ理論の応用によるクラスター技法を用いたシソーラスの自動生成の研究^{12),13)}がある。

1.3 システムの特徴

本研究で用いる諸概念、諸手法は前節の諸例で用い

† Interactive Processing System for Net Thesaurus (NESSY-I) 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

られるものの範囲を越えるものではないが、構成されるシステムに与える目的の特徴を明記すると次のようである。

(1) 研究の対象とする分野で使用する用語や概念の意味論的な処理はすべて専門家が行い、それらの用語間の概念構造を従来の網構造の表現法を応用した因果関係として記述する、

(2) 記述された概念構造の論理的な矛盾の発見や修正、冗長な部分の削除、新たな概念の追加などをブレーン・ストーミングの最中に行う、

(3) 編集し終わった概念構造により、扱う研究対象を適確に把握したり、問題点の分析や抽出、さらには、解決策の早期発見に役立てる、
の3点からなる一連の研究過程を、計算機との会話を繰り返しながら短時間のうちに達成する。

1.4 研究開発の具体的項目

本研究は、

(1) 従来の網構造シソーラスによる概念表現の手法を応用しながら、より平易に、また柔軟に概念が記述できるネット（網構造）シソーラスを計算機で会話的に処理する、

(2) ネット・シソーラスの簡明で処理効率のよい内部および外部表現を調査・検討する、

(3) ネット・シソーラスの生成と修正、検索のソフトウェアを作成する、

(4) システムの効率と応用、将来の発展性を調査する、

の4点を、研究開発の具体的項目とする。

1.5 システムの設計方針

本研究で開発されたネット・シソーラス処理システム NESSY-I (Net Thesaurus Processing System-I) は、つきの方針に沿って機能設計を行った。

(1) ネット・シソーラスの生成と修正、検索はすべて会話形式とする。

(2) 他機種への移行を考慮して主に FORTRAN 言語で記述する。

(3) 網構造の表現には、将来の拡張と処理効率を考慮に入れ、マトリックスではなく二項関係を基本とするリレーション (relation) テーブル (関係表) を用いる。

(4) 検索の効率を上げるため、ハッシング技法を用いる。

(5) 網構造における二つのノード間の道筋およびループの検索には、リンク (LINK) ブロックによる特

殊リスト構造を利用する。

(6) 複雑な網構造の理解を助けるため、検索結果は indentation (字下がり) を用いた木構造に展開して出力する。

(7) システムおよびネット・シソーラス・データの作成と修正を容易に行うため、会話型文脈エディタ¹⁴⁾を利用する。

以上の基本原則に従って、プロトタイプ NESSY を 64 kB (バイト) のコアメモリと2台のディスク装置 (2.5 MB*2) を持つ小型機のもとで開発¹⁵⁾した。その後、小型機での経験をもとにプロトタイプ NESSY を大型機に移行し、システムの機能と処理能力を増強して NESSY-I^{16),17)} と名づけた。

2. システムの概要

NESSY-I は 79 個の FORTRAN プログラムと 4 個のアセンブラ・プログラムからなり、全体で約 2,500 ステップで 97 kW を占め、一度にノード数が 750 個、ノード間の二項関係が 2,500 個までのネット・シソーラスを解析することができる。

2.1 ネット・シソーラス・データの入力形式

複雑な網構造のネット・シソーラスを正確に記述するためには、入力データを人間にとって直観的で分かりやすく表現する必要がある。図1に国土開発に関するネット・シソーラスを図式化した例を示し、図2に、その等価なネット・シソーラスの記述例を示す。図2に示すように、ネット・シソーラスはノード名の上に4つの記号 (=, >, <, <>) の1つをはさんだ、二項関係の集合として表現する。

4種類の二項関係が示す意味は次の通りである。

1. $A = B$ …… A と B が同義語である、
2. $A > B$ …… A から B が生ずる、
3. $A < B$ …… B から A が生ずる、
4. $A <> B$ …… A と B が双方向関係を持つ。

二項関係は、一行の中に複数個並べて記述してもよい。また図2の最後の行の「()」は、入力データの終りを意味する。

2.2 網構造の内部表現

図2に示すようなネット・シソーラス・データは、前もってディスク・ファイルに格納しておく。ネット・シソーラス・データがディスクから読み出されると、ノード名はハッシュ関数 (付録) により数値化され、図3に示すようなハッシュ表と名前表に格納される。

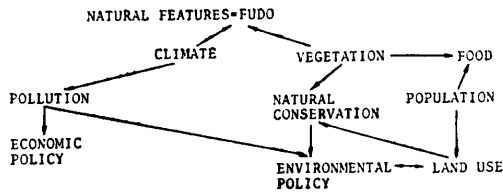


図 1 国土開発に関するネット・シソーラスの図形表現
Fig. 1 A graphical representation of net-thesaurus on a national planning.

(NATURAL FEATURES=FUDO<CLIMATE>POLLUTION)
(POLLUTION>ECONOMIC POLICY)
(POLLUTION>ENVIRONMENTAL POLICY)
(ENVIRONMENTAL POLICY<NATURAL CONSERVATION)
(NATURAL CONSERVATION<VEGETATION)
(ENVIRONMENTAL POLICY<>LAND USE)
(LAND USE>NATURAL CONSERVATION)
(FUDO<VEGETATION)
(VEGETATION>FOOD<POPULATION> LAND USE)
()

図 2 ネット・シソーラス・データの記述例。この記述は、図 1 に示したネット・シソーラスの図形表現と等価である

Fig. 2 An example of description of net-thesaurus data. This description is equivalent to the graphical net-thesaurus shown in Fig. 1.

ハッシュ表は、名前表へのポインタ、名前表には文字数とノード名を格納する。ノード名は、空白も含めて 40 文字以内で表現する。未登録のノード名でコリジョンが発生した場合には、ハッシュ値を +1 ずつ増加させて、ハッシュ表の空領域をサイクリックに探す。ハッシュ表と名前表の大きさは、それぞれ 4,096 語と 3,000 語で、NP は名前表のポインタを示す。

網構造は、図 4 に示す関係表で表現する。従来から網構造は、マトリックス（隣接行列、到達可能性行列など）や複雑なリスト構造によって内部表現されてきたが、これらの方法では多くの記憶容量を占有したり、網構造の操作が複雑になるなどの欠点があった。そこで全体の網構造を図 4 に示すような関係表による二項関係の集合で表現することによってメモリを節約し、同時に網構造の操作（生成と修正、検索）も容易に行うことを可能にした。図 4 の KWN1 と KWN2 領域には、ノード名のハッシュ値、RELA 領域には、二つのノード名間の関係子（「>」、「<」、「<>」、「=」）を数値化（1, 2, 3, 9）して格納する。一つの二項関係は次の 3 項組

(KWN 1 (I), RELA (I), KWN 2 (I)), $1 \leq I \leq 2500$ で表現し、全体で 2,500 個までの二項関係を一度に解析できる。

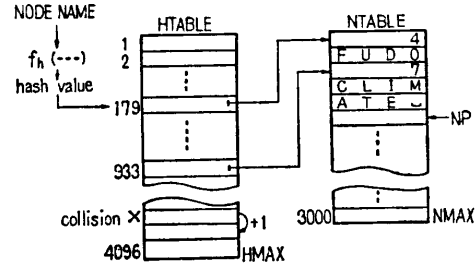


図 3 ハッシュ・テーブルと名前・テーブル。
Fig. 3 Hash Table and Name Table.

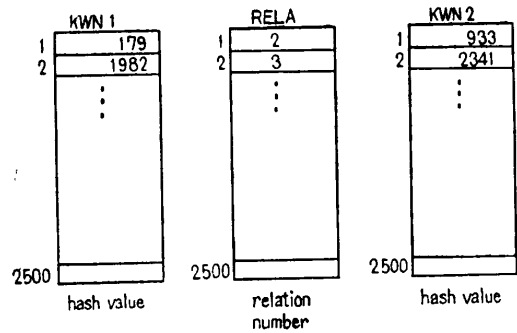


図 4 リレーション・テーブル。二つの関係：(FUDO<CLIMATE)と(ENVIRONMENTAL POLICY<>LAND USE)が格納されている。

Fig. 4 Relation table. Two relations: (FUDO>CLIMATE) and (ENVIRONMENTAL POLICY<>LAND USE) are stored.

2.3 網構造の操作

網構造の操作は、生成と修正、検索の三つからなる。これらのうち、網構造の生成と修正は、図 4 に示した関係表への二項関係の追加および削除により行う。網構造の検索は、次に述べる LINK ブロックと 7 個のスタックを用いて、網構造を木構造に展開して探索する。

2.3.1 LINK ブロック

LINK ブロックは、網構造を木構造に展開するのに使用する 7 語からなる結合子で、その構造を図 5 に示す。7 語の持つ意味は、次のとおりである。

- (1) Hash Value……ノード名のハッシュ値,
- (2) Flag……LINK ブロックの属性を示す,
 - { 0……普通の LINK ブロック,
 - { 7……以前に展開されたノードである,
 - { 9……ループを検出した,
- (3) Level No. ……ルート・ノードからの深さ (レベル) を示す.
- (4) Synonym Link……同義語へのポインタ,

Hash Value
Flag
Level No.
Synonym Link
Parent Link
Child Link
Brother Link

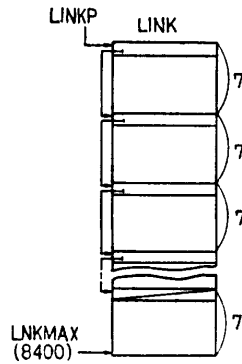


図5 リンク・ブロック 図6 リンク・ブロック・リスト
Fig. 5 LINK block. Fig. 6 LINK block list.

- (5) Parent Link……親ノードへのポインタ,
- (6) Child Link……子ノードへのポインタ,
- (7) Brother Link……兄弟ノードへのポインタ.

未使用のLINKブロックは、図6に示す一方向リストにより結合され、LNKGET と LNKPUT ルーチンが LINK ブロックの獲得と返却を管理している。LINKP と LNKMAX は、それぞれ未使用 LINK ブロックの先頭と末尾を示すポインタで、最終LINKブロックのポインタ部は0である。一度に最大1,200個までのLINKブロックを利用できる。

2.3.2 スタック

網構造の検索を行うために次の4種類のスタックを利用する。

- (1) NAME スタック……木構造に展開されたノード名を記憶しておくためのスタックである。
- (2) SEARCH スタック……展開された木構造の探索や印刷をするときに使用するスタックである。
- (3) HIT スタック……ネット・シソーラス上の任意の2つのノード間を結ぶ道筋や、任意のノードを通るループの検索を行うときに利用するスタックである。
- (4) WORK スタック……同義語を調べるための、または一般的作業用のスタックである。

NAME スタックと SEARCH スタック、HIT スタックには、それぞれ左と右の二つのスタックがあり、これらの各スタックの大きさは、すべて1,500語である。スタックする内容は、展開された各LINKブロックの先頭番地である。WORK スタックの大きさは、100語で、スタックする内容はノード名のハッシュ値である。

スタックの数は全体で7個あるが、検索の種類により使用するスタックが異なる。たとえば、一方向検索

の場合、「F(A>)……Aから到達できるノードを探せ」では、左 NAME スタックと左 SEARCH スタック、および WORK スタックを使用し、「F(A<)……Aのルーツを探せ」では、右 NAME スタックと右 SEARCH スタック、および WORK スタックを使用する。二方向検索の場合、たとえば「F(A>B)……AからBへの道筋を探せ」または「F(A>A)……Aを通るループを探せ」では、すべてのスタックを利用する。

2.4 コマンド

NESSY-I は会話型システムで、ネット・シソーラスの生成、修正、保存、検索などは、すべて会話型コマンドにより簡単に行うことができる。表1にNESSY-I の会話型コマンドの一覧表を示す。コマン

表1 コマンドと機能

Table 1 Commands and their functions.

コマンド	機能
1. A(B=C)	「BとCが同義語である」という関係を追加する。
A(B>C)	「BからCが生ずる」という関係を追加する。
A(B<C)	「CからBが生ずる」という関係を追加する。
A(B<>C)	「BとCが双方向関係にある」という関係を追加する。
A(B/C)	BとCの間の関係を切断する。
2. DN	すべてのノード名を出力する。
DR	すべての関係を出力する。
3. E	NESSY-I の実行を終了する。
4. F(B=)	Bの同義語を検索する。
F(B>)	Bからnステップ以内で到達できるノードを検索する。「n」の値は、L(Level)コマンドで設定する。
F(B<)	Bのnステップ以内のルーツを検索する。
F(B<>)	Bから双方向関係「<>」でnステップ以内に到達できるノードを検索する。
F(B>C)	Bからnステップ以内でCに到達できる道筋を検索する。
F(B<C)	Cからnステップ以内でBに到達できる道筋を検索する。
F(B>B)	Bからnステップ以内でBに到達できる道筋(Bを含むループ)を検索する。
5. I	NESSY-I を初期化する。
6. Ln	ネット・シソーラスの検索の制限ステップ数「n」を正の整数で指定する。nの初期値は5である。
7. QL	検索の制限ステップ数「n」の値を出力する。
QR	名前表、ハッシュ表、LINKブロック、スタックなどのリソースの使用状況を出力する。
8. Selement name	「element name」で指定された名前のネット・データを指定されたファイルから読む。
9. Velement name	ネット・データに「element name」という名前をつけて、指定されたファイルに格納する。

```

$SCPS ZN0907
. FACOM 230 M-6/7 CPS(U06-L01) START 78.02.17 18:11:59
. PASSUORD ?-#####
. YOUR CPS JOB NO.- 0907142175
. CANNOT READ MAIL FILE.
. CPS(U-06/L-01) INITIALIZATION NORMAL END
. READY FOR OUTPUT: NONE
$SUPRU ENNESSY
$ATTACH READ .T
$ATTACH SVSPAT .T
$ATTACH U01.NETDATA
$ATTACH U02.NETDATA
$CALL .P/NESSY
>SHAZARD (a)
>QR (b)

**USED RESOURCE LIST**
RELATION TABLE---( 280/2500)
NAME TABLE----- ( 875/3000)
HASH TABLE----- ( 185/4096)
LINK BLOCK----- ( 0/1200)
SEARCH STACK L--- ( 0/1500)
SEARCH STACK R--- ( 0/1500)
NAME STACK L----- ( 0/1500)
NAME STACK R----- ( 0/1500)
WORK STACK----- ( 0/100)
HIT STACK----- ( 0/1500)

>QL (c)
LEVEL NO. 5 (d)
>L3
>F(JISHIN) (e)
$PRINT (Y/N) ?-Y

JISHIN
DENWA FUTSU
SHINRI FUANTEI
DEMA HASSEI
CHIKAGAI SHUKKA
YUDOKU GAS
CHUDOKU
KEMURI
SHIBO
HINAN FUNO
SHINRI DOYO
SHOKAKI TSUKAEZU
SHOKI SHOKA SHIPPAI
JUTAKU SHUKKA
SHOKI SHOKA SHIPPAI
CHIIKI KASAI
YUDOKU GAS
KEMURI
KANENBUTSU
CHIKAGAI SHUKKA
YUDOKU GAS
KEMURI
BIL SHUKKA

TETSUDO FUTSU
KAZOKU RISAN
SHINRI FUANTEI
KITAKU SHODO
TETSUDO DASSEN
JIKO
SHIBO
FUSHO
CHIKATETSU FUTSU
HINAN FUNO
JIKO

>QR (f)

**USED RESOURCE LIST**
RELATION TABLE---( 280/2500)
NAME TABLE----- ( 875/3000)
HASH TABLE----- ( 185/4096)
LINK BLOCK----- ( 171/1200)
SEARCH STACK L--- ( 171/1500)
SEARCH STACK R--- ( 0/1500)
NAME STACK L----- ( 75/1500)
NAME STACK R----- ( 0/1500)
WORK STACK----- ( 1/100)
HIT STACK----- ( 0/1500)

>L12 (g)
>F(CHUIHO)PANIC (h)
$TOTAL PATH NO. 23
$PRINT (Y/N) ?-N (i)
>LB (j)
>F(CHUIHO)PANIC (j)
$TOTAL PATH NO. 9
$PRINT (Y/N) ?-Y

<PATH NO. 1>
CHUIHO(-1 YEAR)
SOKAI SAUAGI
KOTSU MAHI
SHOBO KATSUDO MAHI
SHINRI FUANTEI
DEMA HASSEI
PANIC

<PATH NO. 2>
CHUIHO(-1 YEAR)
SOKAI SAUAGI
KOTSU MAHI
SHOBO KATSUDO MAHI
CHIIKI KASAI
SHUDAN HINAN FUNO
PANIC

<PATH NO. 7>
CHUIHO(-1 YEAR)
SOKAI SAUAGI
KOTSU MAHI
RYUTSU MAHI
BUSSHI FUSOKU
SEIKATSU FUAN
SHINRI FUANTEI
DEMA HASSEI
PANIC

<PATH NO. 8>
CHUIHO(-1 YEAR)
KEIHO(-1 MONTH)
KINKYU TAISAKU
KOTSU KANSEI
KOTSU MAHI
SHOBO KATSUDO MAHI
CHIIKI KASAI
SHUDAN HINAN FUNO
PANIC

<PATH NO. 9>
CHUIHO(-1 YEAR)
TAISAKU HONBU
KINKYU TAISAKU
KOTSU KANSEI
KOTSU MAHI
SHOBO KATSUDO MAHI
CHIIKI KASAI
SHUDAN HINAN FUNO
PANIC

>QR (k)

**USED RESOURCE LIST**
RELATION TABLE---( 280/2500)
NAME TABLE----- ( 875/3000)
HASH TABLE----- ( 185/4096)
LINK BLOCK----- ( 177/1200)
SEARCH STACK L--- ( 117/1500)
SEARCH STACK R--- ( 60/1500)
NAME STACK L----- ( 82/1500)
NAME STACK R----- ( 48/1500)
WORK STACK----- ( 0/100)
HIT STACK----- ( 9/1500)

>E (l)

* END OF FORTRAN *
. PROGRAM NAME: NESSY ; RETURN CODE: 000.
$BYE
. CPS RETURN CODE : 000
    
```

図 7 NESSY-I の応用例

Fig. 7 A test running of the NESSY-I

ドは、1文字または2文字の英字で始まり、必要ならその後詳細情報を続けて記述する。コマンドが受け付けられるときには入力促進文字「>」がシステムから出力される。表1の「B」と「C」は、相異なるノード名を意味する。A(Add)コマンドでは、複数個の関係を一行の中に並べて書いてもよい。

2.5 実験および検討

NESSY-I の作成、およびネット・シソーラス・データの編集に会話型文脈エディタ¹⁴⁾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、「地震災害とへの対策」に関するネット・シソーラス* に対して検

索のテストを行った。このネット・シソーラス・データは、破壊的な地震によってひきおこされる都市型災害の発生過程と、その被害を最小限にいとめるための対策が、交通、生活、企業、行政など19ユニットにわたって記述されており、ノード数で185個、その間の二項関係は280個ある。

実際の検索例を図7に示す。ここに示す例で、NESSY-I のコマンドは次のように使用されている。

(a) ディスク・ファイル(ファイル名はNETDATA)からネット・シソーラス・データ(エレメント名はHAZARD)を読む。

(b) システム・リソースの使用状況を尋ねる。この例では、ノード数が185個、その間の二項関係が280個で、NAME TABLE が875語使用されている

* このデータは、名古屋大学理学部の島津康男教授により記述された。

のが分かる。

(c) 検索レベルを尋ね、システムで定められたデフォルト値5が得られる。

(d) 検索レベル値を3に変更する。

(e) JISHIN (地震) から検索レベル3で到達できる語を検索する。検索結果は、indentation を用いた木構造で出力される。重複して検索された語には後に「*」、ループを検出した語の後には「%」記号を付加する。~~~~は中略を意味する。

(f) 検索に使用されたシステム・リソースの状況を探る。今回の検索「>F (JISHIN>)」では、171個のLINKブロック、171個の左サーチ・スタック、75個の左名前スタック、それに1個のWORKスタックが使用されたのが分かる。

(g), (h) 検索レベルを12に変更し、注意報(CHUIHO)からパニック(PANIC)への道筋を検索する。

(i), (j) 検索レベルを8に変更して、同様の検索を行う。その結果9通りの道筋が得られる。

CHUIHO (注意報) の後の括弧の中の「-1 YEAR」は、CHUIHO の同義語を意味し、注意報は一年前に発令することを意味する。同様に、(PATH NO. =8) のKEIHO (警報) は、1カ月前に発令することを意味する。

(k) リソースの使用状況を探る。

(1) NESSY-I の実行を終了する。

この検索にはソニー・テクトロ 4014-1 型端末を用い、全体の計算時間は、CPU 時間で4.4秒、端末接続時間で516秒であった。NESSY-I は、このような会話を繰り返して、複雑な網構造で記述されたネット・シソーラスを効率よく検索する。さらに、検索ばかりでなく、ネット・シソーラスの生成や修正も NESSY-I のコマンド(A)により容易に行える。

3. む す び

NESSY-I の開発により、人間の頭の中にある複雑なシソーラスを直接的に反映するような、網構造で表現されるシソーラスの構造解析や生成、修正が容易に行えるようになり初期の目的を達成した。図7の実験結果からも NESSY-I の実用性が実証される。

ここに開発したネット・シソーラス処理機能を、「木構造のシソーラスを用いた研究室向き文献検索システム THEODORES」¹⁸⁾ に組み込むことが今後の課題である。さらに、複雑な網構造のシソーラスを直観的

に把握するために、図形入出力装置を利用した、グラフィック・ネット・シソーラス、あるいは、NESSY-I を拡張して、網構造上のノード間の推移時間や遷移確率などを解析する因果関係の数値的解析シミュレータを開発することが将来の課題として残されている。

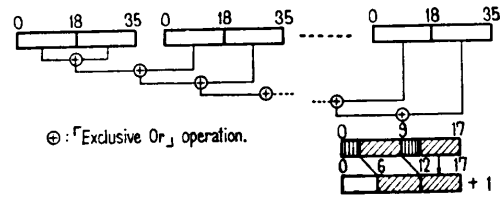
本研究は、名古屋大学理学部の島津康男教授をリーダーとする「環境学のための情報システム: SMLE SIS」の一環として行った。

参 考 文 献

- 1) Koekoek, M. and Schuller, J. A. : The TDCK-Compact System, *Journal of Documentation*, 18, pp. 176-182 (1962).
- 2) Netherlands Armed Forces Technical Documentation and Information Center : TDCK Circular Thesaurus System, The Hague(1963).
- 3) Doyle, L. B. : Semantic Road Maps for Literature Searchers, *J. ACM*, 8, pp. 553-578 (1961).
- 4) Doyel, L. B. : Indexing and Abstracting by Association, *Amer. Doc.* 13, pp. 378-390 (1962).
- 5) Rolling, L. N. : The Role of Graphic Display of Concept Relationships In Indexing and Retrieval Vocabularies, *European Atomic Energy Community, Brussels, EUR 2291.e.* (1965).
- 6) European Atomic Energy Community : EURATOM Thesaurus : Indexing Terms Used within EURATOM's Nuclear Documentation System, 2nd ed., 2 vols., EUR 500.e., Brussels (1966-67).
- 7) Hersey, D. F. and Hammond, W. : Computer Usage in the Development of a Water Resources Thesaurus, *Amer. Doc.* 18, pp. 209-215 (1967).
- 8) Martinez, S. J., et al. : Computer Processing of Thesaurus Data, *Proc. of the Amer. Doc. Institute*, 4, pp. 269-275 (1969).
- 9) Wolff-Terroine, M., et al. : Use of a Computer for Compiling and Holding a Medical Thesaurus, *Methods of Information in Medicine*, 8, pp. 34-40 (1969).
- 10) Rolling, L. N. : Compilation of Thesauri for Use in Computer Systems, *Inform. Stor. Retr.* Vol. 6, pp. 341-350 (1970).
- 11) 三浦宏文 : シソーラスの作成とその評価技術, *数理科学*, No. 41, pp. 29-35 (1966).
- 12) Gotlieb, C. C. and Kumar, S. : Semantic Clustering of Index Terms, *J. ACM*, Vol. 15, No. 4, pp. 493-513 (1968).
- 13) Augstson, J. G. and Minker, J. : An Analysis of Some Graph Theoretical Cluster Techniques, *J. ACM*, Vol. 17, No. 4, pp. 571-588 (1970).
- 14) Hiramatsu, T. : Interactive Context Editor by

- Buddy System, Memoirs of Research and Development Division, Computation Center, Nagoya University, No. 1, pp. 57-83 (1975).
- 15) 平松敏祐: THEODORES Tree Thesaurus から Net Thesaurus への拡張, 昭和 52 年度情報処理学会第 18 回全国大会, pp.293-294(1977).
 - 16) 平松敏祐: 網構造を持つ因果関係の解析システム NESSY-I/II, 昭和 53 年度情報処理学会第 19 回全国大会, pp. 971-972 (1978).
 - 17) Hiramatsu, T.: Interactive Net Thesaurus Processing System NESSY, Memoirs of Research and Development Division, Computation Center, Nagoya University, No. 4, pp. 21-39 (1978).
 - 18) 平松敏祐, 福村晃夫: 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, 情報処理, Vol. 19, No. 12, pp. 1137-1144(1978).

ッシュ値に変換するルーチン(HASH1)は,ビット処理が必要で,しかも高速性が要求されるためアセンブラで記述してある.例として図1に示したノード名に対して, HASH1 で求めたハッシュ値を付表1に示す.



付図 1 ハッシュ値の計算法

A. Fig. 1 Calculation method of a hash value.

付表 1 ノード名とハッシュ値

A. Table 1 Node names and their hash values.

NODE NAME	HASH VALUE
FUDO.....	179
CLIMATE	933
FOOD.....	1043
NATURAL FEATURES.....	1722
ENVIRONMENTAL POLICY.....	1982
LAND USE	2341
POPULATION	2480
VEGETATION	3358
ECONOMIC POLICY.....	3783
POLLUTION	3889
NATURAL CONSERVATION.....	4053

(昭和 53 年 10 月 23 日受付)

(昭和 54 年 1 月 18 日採録)

付 録 ハッシュ関数

空白を含む EBCDIC の文字列として表現されるノード名は,次に示す方法でハッシュ値に変換される. NESSY-I がインプリメントされた大型機 F 230-75 は, 1語が 36 ビットからなり 1語に 4 文字格納できる. 不定長の文字列をハッシュ値に変換するため,各語の上位 2 バイトと下位の 2 バイトを単位にとり,これらの間のビットごとの排他的論理和を付図 1 に示す要領でとる. つぎに,最終的に得られた 2 バイトの各バイト上位 3 ビットを削除して 12 ビット (1~4096) からなるハッシュ値を得る. EBCDIC の文字列をハ