

**この資料は日本Mテクノロジー学会員専用です。**

**この資料を学会員以外がコピーしたり、学会員以外に配布することを禁じます。**

Copy right : M Technology Association - Japan

日本Mテクノロジー学会事務局

〒259-1193 神奈川県伊勢原市下糟屋 143

東海大学医学部・基礎医学系

大櫛陽一

Tel: 0463-93-1121 ext.2140

Fax: 0463-96-4301

Email: [youichi@keyaki.cc.u-tokai.ac.jp](mailto:youichi@keyaki.cc.u-tokai.ac.jp)



*Technology  
Association  
Japan*

# Mumps

Vol.25,2010

Journal of MTA-Japan

目 次

頁

■巻頭言

木村 一元 1

■論文

M 言語上の確率論理プログラミングの取り組み . . . . . 内田 達弘 3

知覚連語の同値性と日本語文簡易化の方法

--- M 言語による日本語解析システム「ささゆり」の意味解析 --- 高橋 亘 9

オノマトペを含む日本語文の M 言語による代替表現機能

--- 聾者のための情報保障の技術 --- . . . . . 津村 雅稔 23

高橋 亘

M 言語による PIC シンボル・プロセッサの実現と

視覚記号の言語学 . . . . . 池田茉莉子 35

柳内 英二

高橋 亘

医療情報システムと連動した CTI システムの開発 . . . . . 土屋 喬義 51

田中千恵子

栢森 猛

木村 一元

CSP での医師国家試験問題検索システムの構築 . . . . . 木村 一元 59

五十嵐吉彦

■日本 MTA 事務局からのお知らせ

「日本 M テクノロジー学会」ご入会のご案内 . . . . . 67

「日本 M テクノロジー学会」規約 . . . . . 69

■資料

投稿規定 . . . . . 75

編集後記 . . . . . 79

表紙装丁 . . . . . 岡田 好一

## 卷頭言

日本 M テクノロジー学会

会長 木村 一元

2009年度から学会長を拝任することとなりました。よろしく、お願い致します。昨年に引き続き、学会大会論文集を学会大会抄録集と雑誌 Mumps に分け、大会発表内容の抄録としての大会抄録集の発行と大会にて発表して頂いた内容を充実して頂いて論文として雑誌 Mumps へ投稿して頂き、査読審査を経て、本年も6編の論文が掲載される事となりました。また、今回の雑誌から国際標準逐次刊行物番号 (ISSN) が符番され表紙に印刷される事となりました。これも、会員みなさんの努力のお陰と感謝しております。

M システムは 1960 年代後半に米国マサチューセッツ総合病院のコンピュータ科学研究所で開発され、Massachusetts General Hospital Utility Multi-Programming System から MUMPS と命名されたシステムで、現在 M システムと呼ばれています。大型機が全盛の時代に開発され、ミニコンピュータで動き、しかもタイムシェアリングで処理し、外部 Disk 領域が内部メモリーと同等に扱え、スパースアレイと言った機能を持ち当時では考えられないほど進んだデータベースシステムでした。しかし、精査されたコマンドや関数の数はあまりにも少なく、単にマニュアルを見た人は、この様なコマンドや関数では、まともなシステムは組めないと思った様です。またさらに、ソートに関するコマンドや関数が無いため、ソートも出来ないシステムとされていた様です。B-Tree 型のデータベースを持ち、斬新なアイデアが盛り沢山組み込まれていたシステムですが、マニュアルだけではその良さが十分に伝わらなかった様です。

時代は進み M システムは、ポスト・リレーショナル・データベースとして米国のみならずヨーロッパの各国で見直されています。また、Caché はオブジェクトの概念を取り込んでおり、今後、大きく種々の分野に発展する事が期待されます。個人利用のスモールスケールから大量のデータを扱う医療や金融のラージスケールのシステムにと幅広く利用できるこのシステムは多くの領域で使用されています。このため、どこにターゲットを置いて説明するかは大変神経を使います。この様な中で日本 M テクノロジー学会としてどの様な活動をして行くか、システムの柔軟さ、高速アクセスの多次元データベース、ユニークな言語使用等、どこに力点を置いて行くか非常に悩む所であります。斬新なアイデアの塊である M についてその良さを引き出し、会員それぞれの活動の中で生かして方策を AGPL に従ったフリーライセンスである GTM と市販のトップシェアである Caché について進めて行きたいと考えています。どうぞ、よろしくお願い致します。

2010 年 7 月吉日



# M 言語上の確率論理プログラミングの取り組み

## Probabilistic Logic Programming on M

内田達弘

Tatsuhiko Uchida

名城大学総合数理教育センター

Comprehensive Scientific Education Center, Meijo University

1-501 Shiogamakuchi, Tempaku-ku, Nagoya 468-8502 Japan

**要旨** GNOSIS は曖昧さを含まない閉世界仮説を採用している。曖昧さを含む現実問題に対処させるため、M 言語上の確率論理プログラムの装備を目指し、その理論と方法および実現の可能性について述べる。

**Abstract** GNOSIS didn't treat ambiguous events because of GNOSIS is based on the closed world assumption. I made one solution of this problem. I explain the theory and ability of the probabilistic logic programming on M toward the implementation.

キーワード: GNOSIS、NOUS、確率論理、確率論理プログラミング、知識データベース

Keyword: GNOSIS, NOUS, Probabilistic Logic, Probabilistic Logic Programming, Knowledge-database

### 1. はじめに

M 言語のデータベースのナレッジデータベースへの応用として論理プログラミングを装備した GNOSIS がある[1]。GNOSIS は一階述語論理に基づくため、曖昧さを含まない閉世界仮説を採用している。そのため導出されない命題は全て偽と結論される。現実的な問題、例えば統計的な意味を持つデータベースを扱うような場合に使うことができない。この問題を改善する方法として確率論理プログラミングがある。確率論理は不確定な問題を対象に出来る論理であり、Prolog を基本にした確率論理プログラミング PRISM も研究・開発されている[2]。M 言語上の確率論理プログラムの装備を目指し、その理論と方法および実現の可能性について述べる。

### 2. 目的

指定した条件に対し、完全ではないが、高い確率をもって一致する情報をデータベースから得たいことがある。例えば、症状リストから

```
>Prove [#^Diag(*diagname,symptomlist)]1)
>Write diagname, “: 確率=”, $Probability, !
糖尿病: 確率=0.9
>Prove
>Write diagname, “: 確率=”, $Probability, !
高血圧症: 確率=0.6
```

といった結果が得られるなら診断支援に役立つ。従来の GNOSIS ではこのような曖昧な検索ができなかった。確率論理はこの問題を解決する有効な理論であり、GNOSIS に応用することでより現実的な知識データベースを目指す。

1) Prove は推論を開始するコマンド、[#^Diag(...)]は推論の対象とする変数^Diag を表す述語リスト、\*diagname は推論により一致した値が代入されるローカル変数を意味する。

### 3. M言語上の知識データベース

長い年月の間に M データベースの資産が蓄えられてきた。また、1972年に論理プログラミング言語 Prolog が出現し、その推論機能は逆向き検索を可能にするため、柔軟性があり、知識データベースの処理言語として脚光を浴びた。筆者らはM言語に Prolog の推論機能を組み込んだ GNOSIS を提案し、1985年に開発した(図1)。

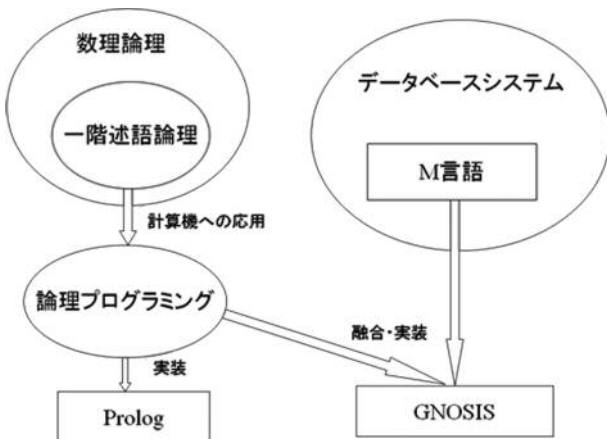


図1. GNOSIS とM言語の関係図

### 4. 確率論理と確率論理プログラミング

確率論は不確実性をもつ現象を数理的に取り扱うことを可能にした。確率論理は確率的な数理モデル上で推論を行う理論を提供した。そのコンピュータ上の実現システムとして確率論理プログラミングが提案されている(図2)。確率論理プログラミングの実装システムとしてPRISMがある。PRISMはB-Prologのアプリケーションとして実装されている[2]。

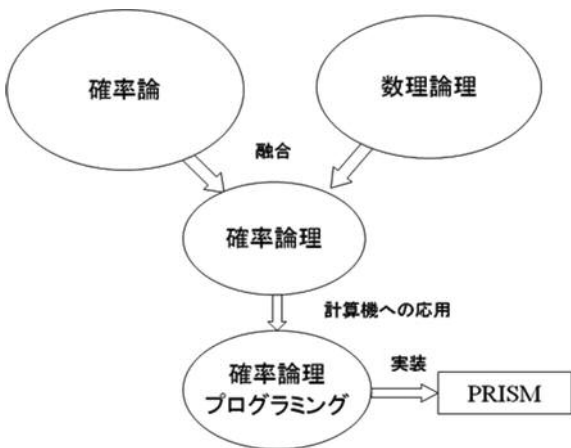


図2. 確率論理プログラミング

PRISMがProlog上のアプリケーションとして実現しているが、これはまたGNOSISも確率論理プログラミングの実装を可能にすることを示唆する(図3)。

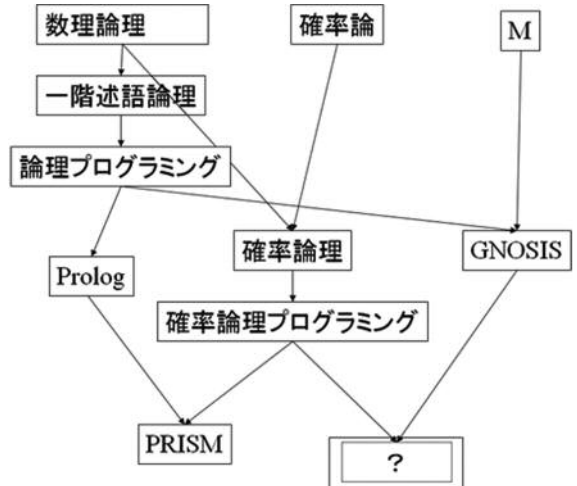


図3. M上の確率論理プログラミング

論理プログラミングは

プログラム=事実と規則の宣言の集り

事実は原始命題

規則は命題の論理積

実行=推論とユニフィケーションおよびバックトラック

により、構成されている。曖昧さを含まない形式的なロジックを提供する。

例：事実：棲息(鮎、水)

(意味) 鮎は水中に棲息する。

規則：魚(X)=呼吸(X,エラ) ∧ 棲息(X,水)

(意味) 魚はエラで呼吸しかつ水中に棲息する。

「Xが魚である」とは「Xはエラで呼吸する」かつ「Xは水中に棲息する」を表している。

GNOSISではこれを次のように実現した。

```
>SET ^LIVE(“鮎”, “水”) = “”
>SET ^BREATH(“鮎”, “エラ”) = “”
>SET ^FISH(*x) = [#^LIVE(*x, “水”),
                  #^BREATH(*x, “エラ”)]
>PROVE [#^FISH(*name)] WRITE name, !
```

鮎

#### 4.1 確率論理

確率論理は個々の命題にそれを真とする確率を与える。0を偽、1を真とすれば論理プロ

グラミングと同じである。すなわち、これまで論理プログラミングでは命題：リンゴの色は赤い「 $P(\text{apple}, \text{red})$ 」が真として定義されたのに対し、確率論理では命題に0から1の値、すなわち確率を定義する。命題「 $P(\text{apple}, \text{red})$ 」に確率0.9を与えるには「 $\text{Pr}(P(\text{apple}, \text{red}))=0.9$ 」のように記述して定義する。

論理プログラミングが、例えば、

$$P(\text{apple}, \text{red}).$$

$$Q(X) = Q_1(X) \wedge Q_2(X) \wedge \dots \wedge Q_n(X).$$

であったのに対し、確率論理は、

$$\text{Pr}(P(\text{apple}, \text{red})) = 0.8$$

$$\text{Pr}(Q(X)) = \text{Pr}(Q_1(X)) \cdot \text{Pr}(Q_2(X)) \cdot \dots \cdot \text{Pr}(Q_n(X))$$

のように命題に確率を与えて表される。文献[3]の確率論理の例を引用して示す(図4)。

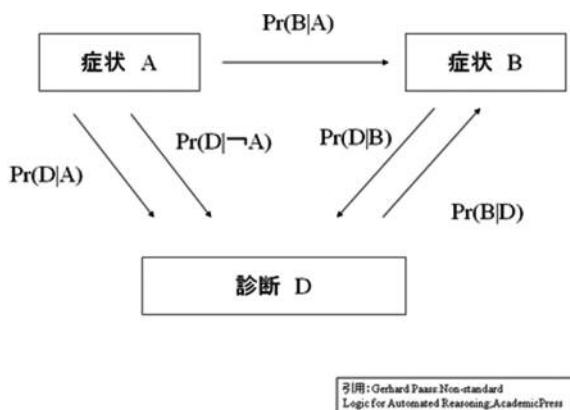


図4. 診断のネットワーク図

A, B を症状、D を病名とする。

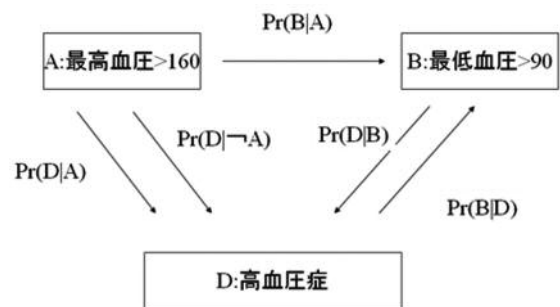
- 規則1 : A なら D である確率は p1
- 規則2 : A でないなら D である確率は p2
- 規則3 : B なら D である確率は p3
- 規則4 : D なら B である確率は p4
- 規則5 : A なら B である確率は p5
- 規則6 : A である確率は p6

これらを式で表す。ここで、「 $\text{Pr}(A)$ は事象Aの起きる確率とする」

- 規則1 :  $\text{Pr}(D|A)=p1$
- 規則2 :  $\text{Pr}(D|\neg A)=p2$
- 規則3 :  $\text{Pr}(D|B)=p3$
- 規則4 :  $\text{Pr}(B|D)=p4$
- 規則5 :  $\text{Pr}(B|A)=p5$
- 規則6 :  $\text{Pr}(A)=p6$

もう少し具体的に示すと(注: 数値は適当)

D:高血圧症、A:最高血圧>160、B:最低血圧>90 とすると (図5)、



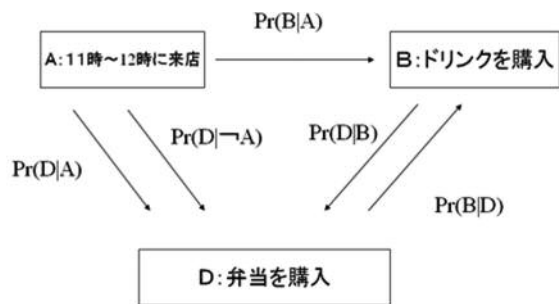
診断のネットワーク図 (高血圧症)

図5. 診断のネットワーク図 (高血圧症)

- 規則1 : 最高血圧>160 なら高血圧症である確率は 1
  - 規則2 : 最高血圧>160 でないなら高血圧症である確率は 0.9
  - 規則3 : 最低血圧>90 なら高血圧症である確率は 0.7
  - 規則4 : 高血圧症なら最低血圧>90 である確率は 0.5
  - 規則5 : 最高血圧>160 なら最低血圧>90 である確率は 0.8
  - 規則6 : 最高血圧>160 である確率は 0.1
- これらを式で表す。

- $\text{Pr}(\text{高血圧} | \text{最高血圧}>160)=1$
- $\text{Pr}(\text{高血圧} | \neg \text{最高血圧}>160)=0.9$
- $\text{Pr}(\text{高血圧} | \text{最低血圧}>90)=0.7$
- $\text{Pr}(\text{最低血圧}>90 | \text{高血圧})=0.5$
- $\text{Pr}(\text{最低血圧}>90 | \text{最高血圧}>160)=0.8$
- $\text{Pr}(\text{最高血圧}>160)=0.1$

この診断規則は他の分野でも適用できる。図6はコンビニのA: 11時から12時に来店した客のB: ドリンクを購入した客とD: 弁当を購入した客の関係図である。このように確率論理は命題に確率を与えることで、自然な診断の構造に近付いている。



診断のネットワーク図 (コンビニ)  
引用: Gerhard Pass: Non-standard Logic for Automated Reasoning, Academic Press

図 6. 診断のネットワーク図 (コンビニ)

### 4.2 確率論理プログラミング

確率論理のコンピュータへの応用は確率論理プログラミングとして研究されている。前述の PRISM がその代表例である。基本的な機能は論理プログラミングと似ている。述語命題に確率をどのように与えるかが重要な要素である。M 上の確率論理プログラミングを装備するために、次のシンタックスとアルゴリズムおよび確率空間の作成方法について提案する。なお名称について従来の GNOSIS と区別するために GNOSIS の国際名であった NOUS を M 上の確率論理プログラミングとして呼ぶことにする。

#### 1) シンタックス

GNOSIS のシンタックス;

述語リテラル、リストリテラル、PROVE コマンド、リスト関数

\$LAPPEND,\$EXTRACT,

\$LFIND,\$LENGTH にシンタックス

\$PROBABILITY 特殊変数

を追加する。

#### 2) 確率推論アルゴリズム

GNOSIS 推論アルゴリズムをつぎのように拡張する。

##### (1) 述語の推論値を

0 ~ 1 の数値にする。(述語の確からしさ) 0 と 1 のときはそれぞれ GNOSIS での偽と真と同じ意味になる。

##### (2) バックトラック

述語の推論値が 0 のときのみ行う。

##### (3) 述語の積の値は述語の推論値の積とす

る。

例:  $[\#^A P(*x), \#^A Q(*x)]$  の推論値は

$$\Pr(\#^A P(*x)) \cdot \Pr(\#^A Q(*x))$$

(4) PROVE コマンドで推論した述語の推論値は \$PROBABILITY 変数に記憶される。

##### (5) 事実 (FACT)

・値のない変数の推論値を 0 とする。

・リストリテラルの要素として 0 ~ 1 の数値を持つ変数をその述語の推論値とする。

例:  $S \wedge Pcolor("Apple", "red") = [0.8]$

$$S \wedge Pcolor("Apple", "green") = [0.1]$$

$$S \wedge Pcolor("Apple", "yellow") = [0.1]$$

・上記以外の変数の推論値を 1 とする。

##### (6) 規則 (RULE)

リストリテラルを持つ変数をその変数名をもつ述語の規則とする。

### 3) 確率空間の作成

確率論理プログラミングのプログラムの中心はデータベースから確率空間を作成することである。次の手順で作成する。

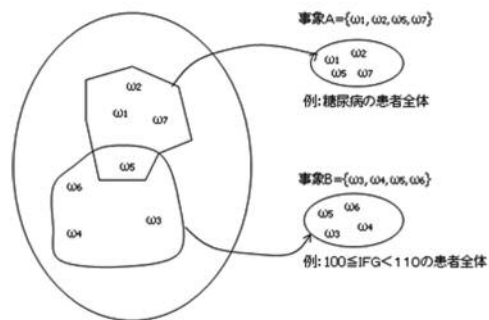
#### (1) 母集団の作成

例: 患者データの集合

#### (2) 事象の集合 (σ 加法族) の作成

例: 男性の集合、年齢 1 歳の集合、...

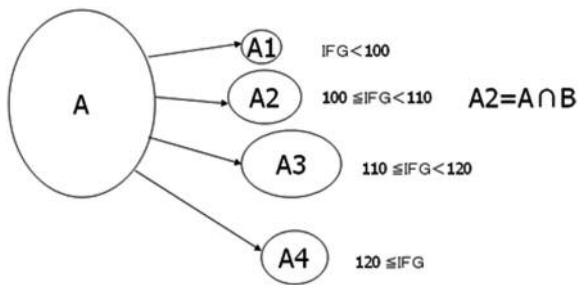
母集団  $\Omega = \{\omega: \text{原子事象}\}$



例: 地域検診情報...住民の健康診断データ

#### (3) 各事象に確率を与える

### 糖尿病の患者の集合を母集団とする確率空間



$$\Pr_A(A2) = \frac{|A2|}{|A|} = \frac{\Pr(A \cap B)}{\Pr(A)} = \Pr(B|A)$$

↑  
元の確率空間Ωから作ることができる

上記(1)～(3)の方法で確率空間を作成すれば、指定した事象の起きる確率が PROVE コマンドにより計算される。

例 患者の IFG 値から病名を推定するプログラム。

```
>S ^Paitient(2009001,"IFG")=110
>S ^Diag1("糖尿病",*IFG1)=
    [#$Eval(*IFG1<109),0.01]
>S ^Diag1("糖尿病",*IFG2)=
    [#$Eval(*IFG2<126),0.5]
>S ^Diag1("糖尿病",*IFG3)=1
>S ^Diag("糖尿病",*ID)=[#^Diag1("糖尿病",
    #$Eval(^Paitient(*ID,"IFG"))]
>Prove [#^Diag(*ID,*Diag)]
>W "患者 ID= ",ID," 診断名=",Diag," 可能性= ",$P
患者 ID=2009001 診断名=糖尿病 可能性=0.5
>
```

ここで、^Diag1(病名,IFG 値)は病名が IFG 値をとるという事象に確率を与えた規則である。

### 5. NOUS 確率論理版実現に向けて

ここまでの、理論面とプログラミングについて述べた。しかし、実際の計算量や計算時間やその他の実用上の問題点などは実装システムで検証しなければならない。そのため、NOUS を実現する必要があるが、次の2つの方法が考えられる。

方法 A GNOSIS の書き直し

メリット：推論速度の高速化がしやすい・修

正と追加ですむ（開発期間）・管理しやすい  
デメリット：既存のデータベースからデータの移植が必要、アプリケーションの開発環境が悪い

方法 B Cache 上のアプリケーションとして実装

メリット：既存のデータベースがそのまま利用できる・実用性が高い・開発環境が整っている

デメリット：実装までに多くの時間と労力を要する・推論速度の高速化が困難

現在、方法 A で開発中である。すでに M 言語の仕様に基づいた M を C 言語で作成済みであり、NOUS（確率論理版）に必要なシンタックスの構文解析および内部コードの生成などできている。GNOSIS Ver.1.02 の推論エンジンを移植中である。

方法 B の Cache 上に装備できないかは、検討中である。

### 6. 今後の課題

M言語上の確率論理プログラムについて提案した。今後、機能をより明確に定義し、理論的な面でも検証しなければならない。また、応用面でも有用性を示すため、検証のためのシステム作りを行う必要がある。

### 参考文献

1. 内田達弘,鈴木利明,ドナルド・スミス,若井一朗:SET GNOSIS=MUMPS+Prolog:米国 MUG 季刊誌 Vol.15,No.3,pp.14-19
2. 亀谷由隆,佐藤泰介,周能法,泉祐介,岩崎達也.:PRISM: 確率モデリングのための論理プログラミング処理系. コンピュータソフトウェア, Vol.16, No.5, pp.78-94, 1999
3. Philippe Smets,E.H. Mamdani,Didier Dubois, Henri Prade:Non-Standard Logic for Automated Reasoning:Academic Press,pp.213-251,1988



## 知覚連語の同値性と日本語文簡易化の方法

--- M 言語による日本語解析システム「ささゆり」の意味解析 ---

Two Equivalence Classes of the Perceptive Collocations  
and the Reduction Methods of Japanese Sentences

--- Semantic Analysis with the Japanese Analysis System SASAYURI  
on the M Language Scheme ---

高橋 亘

Wataru Takahasi

関西福祉科学大学社会福祉学部

〒 582-0026 大阪府柏原市旭ヶ丘 3-11-1

TEL 0729-78-0088, FAX 0729-78-0377

E-mail takahasi@fuksi-kagk-u.ac.jp

**要旨** 知覚連語間の意味的距離が導入される。指定された意味要素の組から意味的距離がある範囲内にある知覚連語の高速検索の技術が議論される。さらに、共通の単語を含む知覚連語の同値類と意味的に距離の近い知覚連語の同値類という二種の同値類の関係が述べられる。同時に、これらの同値類を適確に操作するインターフェイスが議論される。

新しい意味空間の理解に関連して、日本語解析システム「ささゆり」の日本語文簡易化の機能が議論される。

第一は、形式名詞が連体修飾をうける複文の単文化の機能である。複文を捌き、形式名詞の意味を推定する方法はこれまでの我々の研究で明らかである。この論文では、文脈によって限定された形式名詞の意味に最も近い日常語の検索技術が議論される。

第二は、難解語の言い換えに関する機能である。これは上述の二種の同値類の双方の特性を活用した言い換え表現を提供する機能である。特に意味的に距離の近い知覚連語の同値類の活用はキーワード検索では不可能な異表現かつ同義の知覚連語の検索を可能にさせる。

**キーワード** 日本語解析システム「ささゆり」、知覚連語、知覚連語の形成規則、意味解析、意味的距離、日本語文の簡易化、コミュニケーション支援、M 言語

## 1. はじめに

この論文は、先に“知覚連語の同値性と M 言語による日本語解析システム「ささゆり」における日本語文簡易化の方法”と題する題目で第 36 回日本 M テクノロジー学会において口頭発表された後、“日本語解析システム「ささゆり」における日本語文簡易化の方法と知覚連語間の意味的距離”と題する論文として「総合福祉科学研究」に投稿されたもの [1] に、特にインターフェイスの充実に重点を置いて、その後の発展を付加し、理論及び技術の整合性と有効性を強調するものである。

近年我々は、日本語解析システム「ささゆり」の基礎をあたえる知覚連語の言語学の構築を推し進めてきた [2 - 4]。この新しい言語学の基本原理をまとめると次のようになる。

知覚連語の言語学は、通常の言語において単語が明確な意味を保持せず、連語を形成するときに意味が純粋化されるという原理を基礎にとって、明確な意味を伝達すること、つまり明確な知覚を誘発する連語を形成することこそが言語習得の原動力となることに基礎を置いている。知覚連語とはこの明確な知覚を誘発する連語のことである。

知覚連語は二つの側面を有する。その一つは知覚連語の形成過程に由来するものである。知覚連語の言語学によれば、知覚連語はその形成規則にしたがって形成され、形成規則は知覚連語の構成要素の範疇列によって定義されるのであるから、知覚連語は最終的に語彙範疇の列によって構成される。もう一つは、十分な長さをもつ知覚連語は純粋な意味に対応するのであるから、より少ない意味要素によってラベルされることに由来するものである。知覚連語の言語学によれば、単語は通常多くの意味要素を保持しているが、単語が連語を形成し、連語が長くなるにつれて、連語の形成語間の相互規定のために連語が保持する意味要素はきわめて少なくなる。つまり、うまく構成された連語は純粋な意味を保持するようになる。この原理が語の結

合を単に連語と呼ばずに知覚連語と呼ぶ根拠である。

知覚連語の形成規則はある種の語の結合を意図的に禁止している。その一つは連体修飾をする動詞文と被修飾名詞の結合の禁止である。一般に動詞文による連体修飾の構造をもつ複文では、修飾する動詞文が長くも短くもなりうるので、短い動詞文と被修飾名詞の結合を連語名詞（連名詞）と把握してしまうと長い動詞文の修飾関係が把握できなくなるためである。我々の日本語解析システムでは動詞文による名詞の修飾を知覚連語の形成規則から排除することによって、期せずして複文の構文解析の機能を保持することになった。

このような複文の構文解析では修飾する動詞文（修飾子）と修飾される名詞（接合名詞）との対応関係と接合名詞を含んで後続する知覚連語（後続子）の構成関係の二つの関係が接合名詞の意味を限定する [5, 6]。この接合名詞の意味限定の原理は、形式化した名詞、つまり形式名詞についても同様である。一般に複文のような構造を持つ文は聾者にとって分かりにくいものとされている。複文の要としての接合名詞に形式化したものが入ると、さらに理解を困難にする。我々は聾者に対する情報保障のために、形式名詞を含む複文を単文化する問題に我々の意味限定の方法を適用することを試みた [7, 8]。このような目的を解決するためには、意味を限定された形式名詞を限定された意味に最も近い、意味の明確な、日常的名詞に置き換える必要がある。ここに語（知覚連語や単語）が保有する意味要素の集合の間の意味的距離を規定しなければならないという問題が浮上するのである。この論文の第一の主題はこのような意味要素の集合の間の意味的距離の測度の問題である。

知覚連語の形成規則のいま一つの知覚連語形成の禁止則は、機能性の高い機能語が端に来るような結合の禁止である。構文が長くなると語の結合は単に意味を規定し合うだけでなく、異質な意味の間の関係性を規定するようになるこ

とがある。このような関係性を把握するためには、機能性の高い機能語を知覚連語に内包させるべきではない。このような構文については、知覚連語間の接続関係によって、我々が高次の総合関係と称した指標付き直和ベクトル空間として把握すべきである。指標付き直和ベクトル空間の問題は、機会を改めて議論したい。

この論文で問題にする第二の主題は知覚連語間の同値類の問題である。先に述べた知覚連語の二つの側面は二種の同値類の存在を示唆する。その一つは知覚連語を構成する範疇の共通性をもたらす同値性であり、共通な範疇を保持する知覚連語が一つの同値類を構成する。このような同値類は、いわば共通のキーワードを保持する知覚連語の探索技術と直接的に関わるものである。このような知覚連語の同値類を、今後、共通語同値類と呼ぶことにする。いま一つは共通の意味要素を保持する知覚連語の同値類であり、意味的距離の近い知覚連語が一つの同値類を構成する。これは意味空間の構成が前提であり、知覚連語の構成によって始めて定義が可能になるものである。このような同値類は、共通のキーワードを保持しない知覚連語の探索技術を提起するものであり、情報検索の新しい技術である。このような知覚連語の同値類を、同義性同値類と呼ぶことにする。

二種の同値類、つまり共通語同値類と同義性同値類は、双方の利便性が相まって情報検索に寄与するものであるが、コミュニケーション支援の基本的方法を提示する。コミュニケーションは、人物 A が自分の記号化規則を用いて、自らの意図を言語記号によるメッセージを構成することに始まる。メッセージが他者に伝わると、メッセージを受け取った人物 B は、これをその人の記号化規則を逆に用いて復号し、メッセージ構成者の意図を推し量る。一般的に記号化規則は人ごとに異なっているから、復号された意図はメッセージ構成者の元のままではない。そこで、B に、B が理解した A の意図を、B の記号化規則にしたがってメッセージ化し

てもらおうことをしてみると B の構成したメッセージによって、A は自分の意図がどのように相手に伝わったのかを知ることが出来る。この時、多くの場合、B の理解した A の意図はしばしばいくらかのずれをともなって伝達されることが観察させる。コミュニケーションが正確に成り立つためには、少なくとも数回のメッセージのやり取りが必要な所以である。

やり取りされるメッセージの中に A と B のどちらかが、知らないか、もしくは分かりにくい言葉（記号）があると、このような対話は特異的になる。特異的な対話では、対話が進行するためには、一方が理解できない言葉を含むメッセージ（これは共通の言葉を含む知覚連語の集合、つまり共通語同値類を構成する）を、その言葉を理解出来る側から、同じ意味を持つ、他の言葉で表現された、別のメッセージ（これは意味的に近い知覚連語の集合、つまり同義性同値類を構成する）を作成する必要性が生じる。つまり、未知の語を既知の語を用いた言い換えの集合を構成することで、未知の語を含むメッセージの理解が促進されることが期待されるのである。

このような未知の言葉を含む文の言い換えの技術は、障害があるために、ある種の言葉に系統的に分かりにくい言葉が集中する時にも有効な方法を提供する。聾者に分かりにくい日本語文は (1) 語彙自体が分かりにくいもの、(2) 構文が難解なもの、(3) 日本語と日本手話で補助化する単語が異なるもの、などが挙げられる[2 - 4]。先に言及した複文は (2) の例であり、形式名詞は (1) と (2) にまたがっている。(3) の問題は、今は触れないにしても、(1) の語彙として分かりにくいものには機能語を筆頭に、補助化した内容語や抽象名詞、オノマトペ（擬音語・擬態語）などが系統的に分かりにくいとされている。こうした言葉には上述の言い換えが必要なのである。

**第 2 節**では、知覚連語間の意味的距離の測定方法および知覚連語・意味要素の相関関係の



捌き方が、第 3 節では、2008 年以來取り組んできた形式名詞を含む日本語文の簡易化の問題が取り扱われる [7, 8]. これには構文の簡易化と形式名詞の意味推定の技術、さらに限定された意味要素の集合に最も近い日常語を見つける技術などが総て投入される. この節で述べられることは、その前半は、宮地との共同研究として 2008 年の Mumps に公表された内容の要説、後半は、宮地の修士論文として発展させられた内容を筆者の観点から補遺するものである. 第 4 節では、知覚連語の二種の同値類、共通語同値類と同義性同値類の一般論が述べられる. また、この技術の適用例である、聾者の情報保障のためのオノマトペの言い換え技術が、津村との共同論文として本誌に公表される [9]. 第 5 節では、二種の同値関係を観察し、知覚連語とそれにもなう意味空間のデータベースを操作するインターフェイスの設計について述べられる.

## 2. 知覚連語間の意味的距離と知覚連語-意味要素の相関関数

一つの知覚連語にはその知覚連語が表現している意味要素の組が対応する. 二つの知覚連語があって、それぞれの知覚連語が意味要素の組  $A, B$  を持っていたとする. このとき、二つの知覚連語間の意味的距離をどのように定義すればよいのだろうか?  $A, B$  の要素数がそれぞれ  $n_A$  個,  $n_B$  個であったとし,  $A, B$  共通のものが  $n_{A \cap B}$  個あったとすれば, 共通性のない意味要素の個数は  $n_A + n_B - 2n_{A \cap B}$  であり, この要素数が二つの知覚連語の意味的隔たりの程度を表していることに異論はなからう. これに対して共通の意味要素の個数  $n_{A \cap B}$  個は二つの知覚連語の意味的共通性の程度を表していることもまた然りである. しかしこれらの二種の程度を表している量はどちらも個数という物理的には意味のよく分からない量を単位として測定するものであることも事実である. 数学や物理学では物事の特徴を示す数量を定義す

る際にしばしば単位系によらない数量として無次元数を定義する. これに習えば, 今の場合, 知覚連語間の意味的距離を次の式で定義するのが妥当であると考えられる.

$$d_{AB} = \frac{n_A + n_B - 2n_{A \cap B}}{n_{A \cap B}}.$$

この定義では,  $A$  と  $B$  が集合として等しい場合に意味的距離が 0 となり,  $A$  と  $B$  に共通する意味要素がないときに無限大となる. (もちろんコンピュータで無限大の数量を扱うことは出来ないので, 情報科学的には意味的距離の最大数をコンピュータで扱える最大数によって制限することになる) また, この定義は, 隔たりの程度と共通性の程度の比であることから, 隔たりの程度の大きいものに対して値が大きく, 共通性の程度が大きいものに対して値が小さくなるという特性も知覚連語間の意味的距離の性質をよく保持していると考えられる.

ここで, 我々は知覚連語間の意味的距離の測定を必要とするいくつかのアルゴリズムについて触れておきたい. 日本語解析システム「ささゆり」においては, 知覚連語が一つ存在するとそれが保持する意味要素の集合は, システムが保有する知覚連語-意味要素相関辞書に対応する大域変数  $\wedge\text{NCDIC}$  によって一意的に決定される. 大域変数  $\wedge\text{NCDIC}$  は,

$$\wedge\text{NCDIC}(\text{Col}, \text{PS}) = \text{SemSet}$$

の形を持っており, ここで大域変数をラベルする  $\text{Col}$  は知覚連語,  $\text{PS}$  は知覚連語の連語範疇である. 大域変数の値である  $\text{SemSet}$  は知覚連語が保有する総ての意味要素をセパレータ“/”を挟んで接合したものである. したがって, 知覚連語とその連語範疇が与えられるとその意味要素の集合が一意的に与えられるので, 二つの知覚連語とその連語範疇が与えられれば, これらの間の意味的距離を測定する方法は意味的距離の定義から自明である.

大域変数  $\wedge\text{NCDIC}$  を編集する知覚連語-意味要素相関辞書 (概念辞書) を図 1 に示す.

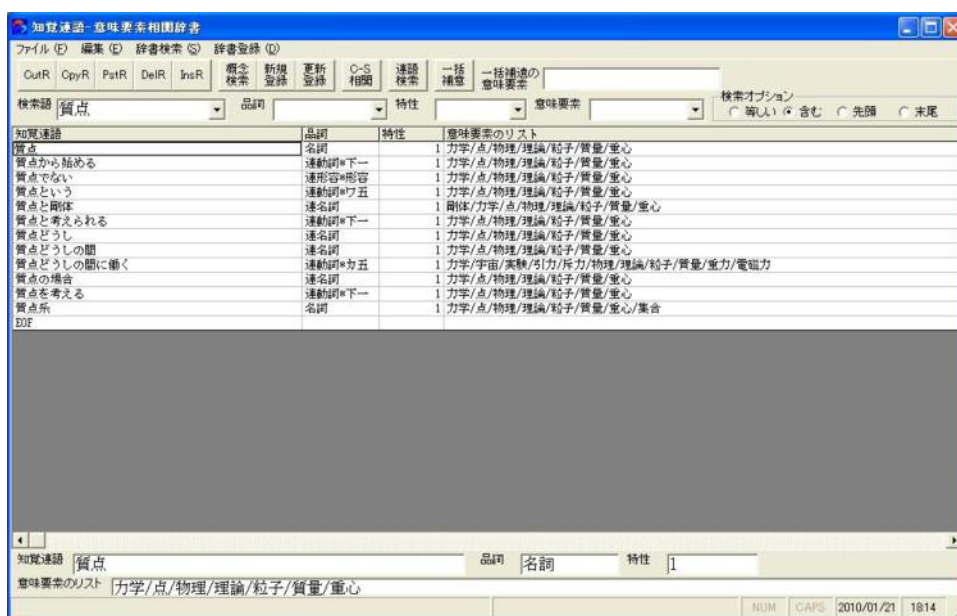


図 1 知覚連語と意味要素の相関を定義する概念辞書

検索のためにオプションを設定し、検索語他の情報を入力して、[概念検索] ボタンを押せば大域変数 ^NCDIC が記憶している意味要素のリストが表示される。表示内容は編集可能で、左上部 5 つのボタンは編集時にグリッドの行単位の切り取り (CutR)、コピー (CpyR)、貼り付け (PstR)、削除 (DelR)、挿入 (InsR) をするものである。[連語検索] ボタンを押せば、知覚連語辞書に登録されているものを表示し、それに対応する ^NCDIC の存在・非存在にしたがって意味要素がリストされる。検索語から推測して全検索内容に共通する意味要素が存在すれば一括して補遺することも可能である。編集結果を大域変数 ^NCDIC に反映させるのが [新規登録] と [更新登録] の 2 ボタンである。

ここで我々が問題にしたいのは、一つの知覚連語かあるいは一つの意味要素の集合があるときに、いずれの場合も意味要素の集合が一つ決まるわけであるが、与えられている知覚連語を除いて、この集合に最も近いか、もしくは意味的距離がある基準の範囲内にあるかの知覚連語を見つけ出すアルゴリズムである。現在、日本語解析システム「ささゆり」が知覚連語として学習完了しているものは単語も含めて 700,000 語を超えている。このような膨大なデ

ータを総てなめ尽くして検索するのに必要な時間は、4 GHz を超える処理速度を持つコンピュータであっても、数秒かかることになる。したがって、検索効率を上げるには、あらかじめ知覚連語と意味要素の相関、逆相関の関係を記憶しているデータを、M 言語の大域変数として、

$$\wedge\text{NWCSAMP}(\text{Col}, \text{Sem}),$$

$$\wedge\text{NWCSIAMP}(\text{Sem}, \text{Col})$$

のようなものを定義しておく方法が考えられる。<sup>1</sup> ここで Col は知覚連語、Sem は個々の意味要素である。このような大域変数を ^NCDIC から生成する方法は半ば自明であるが、この操作に必要な時間は数 10 秒と推定される。しかし、一度定義しておけば、ある知覚連語を指定して、その知覚連語が依存する意味要素を総てリストするときには大域変数 ^NWCSAMP を用い、逆に意味要素を指定して、これを含む知覚連語の総てをリストするときには大域変数 ^NWCSIAMP を用いて、それぞれ第二階層を手繰れば、どちらの検索もほとんど一瞬に完了する。これらの大域変数を登録するボタンは、図 1 では [C-S 相関] ボタ

<sup>1</sup> 2 つの大域変数は知覚連語と意味要素を結びつける広い意味の相関関数に対応している。

ンである。

意味要素と知覚連語の相関が定義されたところで、指定された意味要素の組  $S$  に最も近い知覚連語を探すアルゴリズムについて考察しておく。その第一段階は、意味要素の組  $S$  の個々の意味要素に相関を持つ知覚連語をリストしてこれらを候補とすることである。このとき、一般には意味要素の組  $S$  の複数の要素を保持する知覚連語も存在する可能性があるわけであるから、単純に候補をリストしていく手法では、重複して候補を挙げる可能性がある。このような知覚連語に対し一度候補に挙げた知覚連語を重複してリストしない仕掛けが必要である。

アルゴリズムの第二段階は、こうして、リストされた知覚連語の一つ一つと指定された意味要素の組  $S$  との意味的距離を測定し、その最短のものを見つけることである。先に二つの意味要素の組の意味的距離を求める方法は確立しているので、この段階のアルゴリズムはほぼ自明である。

### 3. 形式名詞の意味推定と形式名詞を含む日本語文の簡易化

一般的に連体修飾のある複文は聾者に分かりにくいとされているが、被修飾名詞が形式名詞になると、形式名詞の意味が単語のみでは推定し難いために、理解が一層困難になる。我々は、2008 年以来、形式名詞が連体修飾されている複文の機械的簡易化の技術について開発を続けてきた[7, 8]。

これまでの我々の方法を要約すると次のようになる。日本語解析システム「ささゆり」によれば、動詞文による連体修飾のある複文は、連体修飾する動詞文（修飾子）と修飾される名詞（接合名詞）、接合名詞が体言の役割を果たす骨格文に分解される[2 - 4]。被修飾名詞が接合名詞と呼ばれるのは、これが修飾子と骨格文の接合部分を形成するからである。修飾子は、それ自体が知覚連語であり、接合名詞は、それが体言として働く骨格文の、特に接合名詞とそれに

後続する部分の中で知覚連語（後続子）を形成する。

接合名詞の意味は一般に二つの関係によって限定される。二つの関係とは、

- (1) 修飾子と接合名詞の対応関係、
- (2) 接合名詞を含む知覚連語（後続子）の形成関係、

の二つである。接合名詞の意味は、前者の関係によって、修飾子の保持する意味要素の集合と接合名詞が元々保持していた意味要素の集合の積集合に限定され、後者によって、後続子の保持する意味要素の集合と限定前に接合名詞が保持していた意味要素の集合の積集合に限定される。したがって、二つの関係によって接合名詞の意味は、接合名詞がもともと保持していた意味要素の集合と修飾子の保持する意味要素の集合、後続子の保持する意味要素の集合の三つの集合の積集合に限定される[5, 6]。

形式名詞の意味推定の方法は、このような接合名詞の一般的な意味限定の方法を形式名詞に適用するものである。先の論文 [7, 8] で詳説したように、我々の立場では形式名詞とは、「連体修飾を受け、骨格文で体言として振る舞う」ということによって始めて明確な意味を持つ名詞であると言える。そして、意味の上から言えば、形式名詞は「大変多くの意味を保持してしまいが故に単独では意味を特定できないもの」と言うことができる。形式名詞は接合名詞で、かつ形式化したものであるから、一般の接合名詞の意味限定の方法によって意味が限定される。

以上がこれまで我々が開発に際して明らかに為し得た内容の要約であるが、ここで、この論文の主題の一つである、形式名詞を含む複文を単文化する最終段階の議論に移りたい。日本語解析システム「ささゆり」によれば、複文を修飾子と接合名詞の組と骨格文に分解する方法は既に確立しているのであるから、形式名詞を含む日本語文の簡易化は、上述の方法で意味を限定し、限定された意味要素の組に最も近い日常

的名詞を見つける方法が確立すれば、その大半のアルゴリズムが確立することになる。いま我々が必要とするのは指定された意味要素の組に最も近い日常的名詞を見つける方法であるが、この方法は第 2 節で既に確立している。指定された意味要素の集合に最も近い知覚連語を、連語範疇を名詞に限定して、第 2 節のアルゴリズムを適用するだけのことである。

以上で、限定された意味要素の集合に最も近い意味を持つ日常名詞を見つける方法のアルゴリズムが確立したことになるが、ここで、我々の方法で簡易化される日本語文の実例を挙げて、本当に分かりやすい日本語文が提供されるのかどうかを検証しておくことには意味がある。先の論文で引き合いに出したいくつかの例について、我々の方法を適用した結果を例示しておきたい。まずは形式名詞の特性を示す例文として、作成的に構成され、よく引き合いに出されるものを挙げてみよう。

【例 1】“彼がいうことは信用できない。”

このような例文に我々の方法を適用すると

(1) 彼がいう こと[事柄/内容/言葉; 話]

<こと>(1)は信用できない。

のような解析結果が出る。一行目が修飾子と接合名詞の対応関係、二行目が骨格文を表している。[ ] 内の “;” の前が限定された意味要素であり、後が形式名詞の代替されるべき日常的名詞ということになる。解析結果から、【例 1】の単文分解した簡易化文は、

“彼がいう。その話は信用できない。”

ということになる。つまり修飾子と接合名詞の対応関係で形式名詞を捨象して言い切る。そして、骨格文で、限定された形式名詞の意味を表現する日常的名詞を用い、これに指示的連体詞を付加して形式名詞と置き換えた文を後続させたものが、元の文を簡易化した言い換え文ということになる。

同様の趣向をもつ、短い例文を 4 例あげておきたい。

【例 2】“毎朝牛乳を飲むことにしている。”

(1) 毎朝牛乳を飲む こと[当為/習慣; 習慣]  
<こと>(1)にしている。

【例 3】“ホッチキスとは紙を留めるのに使うものです。”

(1) 紙を留める のに 使う もの[道具/学習; 学用品]

ホッチキス とは >> <もの>(1) です。

【例 4】“昔、あそこへよく行ったものでした。”

(1) あそこへよく行っ た もの[経験/習慣; 習慣]

昔、 >> <もの>(1) でした。

【例 5】“人間は死ぬものです。”

(1) 死ぬ もの[事態/必然/運命; 定め]

人間 は >> <もの>(1) です。

【例 4】に関しては、使用される名詞“習慣”にかかる連体詞は“その”ではなく、“そんな”の方が自然である。使用するべき自然な連体詞を決定する方法はこれからの問題であると考えられる。

上に挙げた例は、例文として作成的に構成されたものであるが、実際に文献に使用された例文についても我々の方法は有効であり、このような例を 2 例挙げておく。

【例 6】“太陽の引力とみあうだけの反対方向の力が惑星に対して働かねばならない。”

【P. C. W. デイヴィス著、松田卓也、二間瀬敏史訳(ブラックホールと宇宙の崩壊)】

(1) 太陽の引力とみあう だけ[平衡/程度/計量; 同程度]

<だけ>(1)の反対方向の力が惑星に対して働かねばならない。

【例 7】“このうち 3 羽は、去年巣立ったばかりの若い鳥だった。”【記者不詳(朝日新聞サイエンス動物)】

(1) 去年巣立っ た ばかり[時期/状態/状況/直後; 直後]

このうち 3 羽 は、 >> <ばかり>(1)の若い鳥 だった。

いずれも、機械的推定の妥当性が観察される。



#### 4. 共通の単語を含む知覚連語と意味的に距離の近い知覚連語

この節では、知覚連語の二種の同値性、構成要素として共通の単語を含むという意味での知覚連語の同値性（共通語同値類）と、意味的に距離が近いという意味での知覚連語の同値性（同義性同値類）とについて知覚連語の検索技術について述べ、その後で、これらの同値性を用いた難解語の言い換え技術について考察したい。

同義性同値類に関連する検索技術は、ある単語を含む特定の知覚連語に意味的に近い知覚連語をリストする技術である。これは二段階の操作で完了する。第一段階は、与えられた知覚連語の保有する意味要素の集合を取得することであるが、この集合は第 2 節で既に述べられている大域変数  $\wedge\text{NCDIC}$  の検索で直接的に与えられる。第二段階は、与えられた意味要素の集合に最も意味的に近い知覚連語の一群を検索することであるが、この第二段階は、第 2 節の最後で既に述べられている。

共通語同値類に関連する検索技術は、共通の単語を含む知覚連語の高速検索の技術である。日本語解析システム「ささゆり」によれば、知覚連語はその形成規則にしたがって機械学習される。知覚連語形成規則は範疇列の数語のパターンについて定義されているから、機械学習された知覚連語はその構成要素の列と構成要素の品詞の列が、ある系統性をもって配列されている。これらの配列は知覚連語の辞書に相当する大域変数  $\wedge\text{NWDIC}$  の値の一部として記憶されている。この大域変数は、

```
 $\wedge\text{NWDIC}(\text{Col}, \text{PS}, \text{No})$ 
= $\text{YOMI\_HT\_ADD\_HT\_STRING\_}$ 
 $\text{HT\_PARSING\_HT\_SGN}$ 
```

のような構造を持った大域変数である。ここで大域変数をラベルする Col は知覚連語、PS は知覚連語の連語範疇、No は作品の特性などを指定する指標（通常は 1 に固定）である。大域変数の値は HT (Horizontal Tabulation) で区

切られた個々の特性であり、STRING が単語列、PARSING が品詞列である。（単語列と品詞列はそれぞれ半角スペース“ ”で区切られている）第 2 節で議論した大域変数  $\wedge\text{NCDIC}$  は大域変数  $\wedge\text{NWDIC}$  と双対的に生成されるものであるから、大域変数  $\wedge\text{NWDIC}$  もまた総数 700,000 個を超えた数量を持っている。したがって、構成要素としてある単語を保持する知覚連語を総て見つけ出すという今の我々の課題は、ある意味要素を保持する知覚連語を総て見つけ出すという第 2 節の課題と酷似している。少し事情が異なる点は、STRING がスペース区切りの単語列であり、個々の単語は文中の形として活用形を保持しているから、これらの原型をデータベースに問い合わせ確認する必要があることである。このような確認作業を含めると、検索に要する時間は数 10 分に達する。したがって、検索効率を上昇させる必要性は先の例よりもはるかに大きい。我々はここでもまた、予め知覚連語 (Col) とその構成要素である共通語 (Word) の相関と逆相関の関係を記憶している M 言語の大域変数として、

```
 $\wedge\text{NWCWAMP}(\text{Col}, \text{Word}),$ 
 $\wedge\text{NWCWIAMP}(\text{Word}, \text{Col})$ 
```

のようなものを定義しておくことが得策である。<sup>2</sup> 議論を繰り返すことを避けて結論を述べれば、ある知覚連語を指定して、その知覚連語が依存する構成要素を総てリストするときには大域変数  $\wedge\text{NWCWAMP}$  を用い、逆に構成要素を指定して、これを含む知覚連語を網羅するときには大域変数  $\wedge\text{NWCWIAMP}$  を用いて、それぞれ第二階層を手繰れば、どちらの検索もほとんど一瞬に完了する。

以上の議論で既に明らかのように、特定の単語を構成要素として含む知覚連語をリストする問題は、大域変数  $\wedge\text{NWCWIAMP}(\text{Word}, \text{Col})$  の定義と検索の問題に集約される。

<sup>2</sup> 2 つの大域変数は知覚連語と共通語を結びつける広い意味の相関関数に対応している。

以上で二種の同値性に関連する検索技術の議論は完了するが、残された問題は、二種の同値性を用いた難解語の言い換え技術である。我々の置かれた状況は次のようなものである。ある人にとって難解な言葉が与えられたとすると、当該の言葉を含む知覚連語の集合  $A$  が存在する。ここで、集合  $A$  に属する知覚連語の一つ一つを  $X_i$  で表すことにする。一般に、一つの言葉はいくつも意味を持っているから、集合  $A$  は言葉の使われ方、つまり知覚連語の形成のされ方にしたがって、 $X_i$  の保持する意味が規定される。いま、第 2 節で考察した知覚連語どうしの意味的距離を考えると、一般的に、ある知覚連語に意味的に近い知覚連語の集合という同値類が考えられる。意味的に近いという同値性はある単語を含んでいるか、いないかという同値性とは独立したものであるから、この同値類に含まれる知覚連語には、指定された単語を含んでいるものと含んでいないものが存在する。問題となる一つの言葉を決めると、これを含む知覚連語の集合  $A$  の要素である個々の知覚連語  $X_i$  に対して、意味的に近い知覚連語の集合  $B_i$  が存在する。つまり、当該の言葉を含む知覚連語の集合は、その個々の要素に意味的に同値な知覚連語の集合によって類別される(図 2)。ここに知覚連語の集合  $B_i$  の一つ一つの要素には当該の言葉を含んでいるものといないものが存在する。つまりある言葉が難解であるとする人にこの言葉の意味を理解することを支援するためには、この言葉に対する言い換え表現が必要なわけであるが、いま述べた知覚連語の類別に登場する各々の意味的同値の集合における当該の言葉を含むものと含まないものとの照合体系こそ、我々が必要とする言い換え表現である。つまり我々に必要な言い換え技術は、当該の言葉が含まれる知覚連語の意味を同定し、これと意味的に近く、当該の言葉を含まない知覚連語を検索する技術である。この技術のアルゴリズムはこの節の前半の議論で既に明らかであろう。

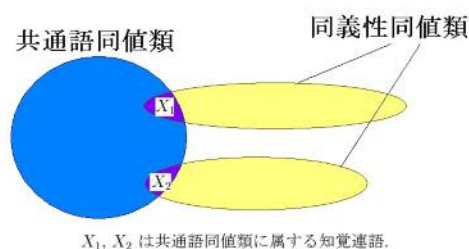


図 2 共通語同値類を類別する同義性同値類

我々の技術は、様々な障害によってコミュニケーションに困難が生じている個々の状況に応じて、有効な方法を提示するものと期待される。この技術の適用例として、聾者の情報保障のためのオノマトペの言い換え技術は、津村との共同論文として本誌に公表される[9]。

## 5. 知覚連語の二種の同値類に関連するデータベースを操作するインターフェイス

この節では、第 4 節で述べた知覚連語の二種の同値類を観察し、知覚連語とそれともなう意味空間のデータベースを操作するための実用的インターフェイスについて述べる。

最初に述べるインターフェイスは専門用語辞典である。いままで“未知の言葉を含む文の言い換えの技術”や“共通語同値類”について述べてきた中で、一般性を保持するために、“未知の言葉”、“難解語”、“難解な言葉”などと言ったり“共通の単語”と言ったりして、漠然と述べてきた“語”や“言葉”は、実用的には、これを分野別に種分けする必要がある。分野を分けて難解な言葉を設定し、それに解説を付加したものは専門用語辞典である。したがって二種の同値類を効率的に捌くために最初に必要なものは専門用語辞典である。第 4 節の最後でふれた“オノマトペの言い換え技術”について言えば“専門”は“オノマトペ”のようなものになるが、一般的に言えば、“物理学”、“医学”、“心理学”・・・、と言った様々なジャンルが考えられる。

我々が用意した専門用語辞典を図 3 に示す。分野と専門用語やその品詞を指定して専門用語の意味や想定される意味要素の組が検索できる

ようになっている。右上にある検索オプションは指定された用語に対して“等しい”，“含む”，“前方一致”，“後方一致”などの検索方法を指定するものである。図 3 では専門分野として“物理”，専門用語として“相互作用”がそれぞれ指定され，検索オプションは“含む”が選択されている。

辞書は編集可能で、左上部 5 つのボタンの働きは知覚連語-意味要素関連辞書 (図 1) にあるものと同様である。グリッド編集後に検索ボタン右の新規登録もしくは更新登録ボタンを押すことで辞書データベースに変更が反映される。以上はデータベース編集の通例の機能である。

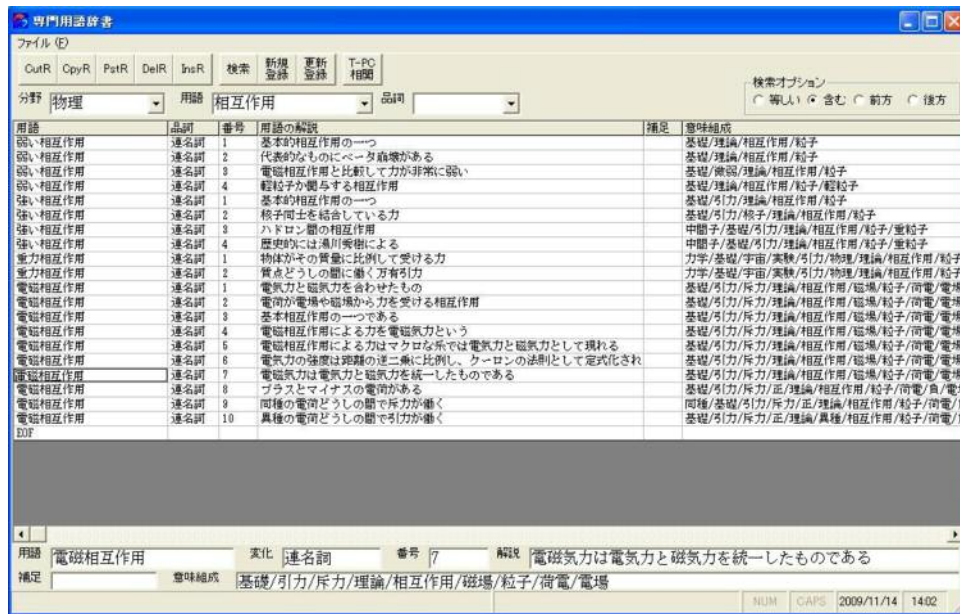


図 3 専門用語辞典

我々の専門用語辞典の特殊な役割は、これが専門用語の枠組みを決定する判定関数の役割を担うものであり、この辞書に登録されている専門用語に関して、第 4 節で述べられた、知覚連語と専門用語の相関が設定されるところにあると言える。[T-PC 関連] ボタンを押すことによって知覚連語と専門用語の相関が設定されるようになっていく。知覚連語と専門用語の相関について、今一つ補足しておかなければならないことは、図 3 の例中の“強い相互作用”，“重力相互作用”などからも明らかなように、多くの専門用語は形態素的には複合語になっているので、相関関係の登録には複数語列の包含関係を判定する機能を内包させる必要がある点である。このような包含関係の判定にはアルゴリズムを洗練した上でもさらに多くの時間が必要である。

次に述べるインターフェイスは共通語同値に関係するものである。図 4 は単語、複合語に限定

しないで、指定された専門用を含む知覚連語を検索し、概念辞書に各知覚連語に想定される意味要素の組を追加登録するためのインターフェイスである。このインターフェイスでは、検索用語を記入し [検索] ボタンを押すことによって、検索用語と相関関係のある知覚連語がリストされ、その時点で概念辞書に登録されている意味要素の組が表示される。検索用語以外に補助キーを書き込むと検索される知覚連語は検索用語と補助キーのどちらも含むものに絞り込まれる。こうして表示された知覚連語の集合の対比から想定される意味要素を、表示されている意味要素の組に追加して、書き込み、追加登録ボタンを押すと、概念辞書にこれが追加登録される。したがって、このインターフェイスは専門用語が含まれる知覚連語の類似した用法の観察から想定される意味要素の組を適正に判断して概念辞書に意味要素を追加登録するためのものであると言える。



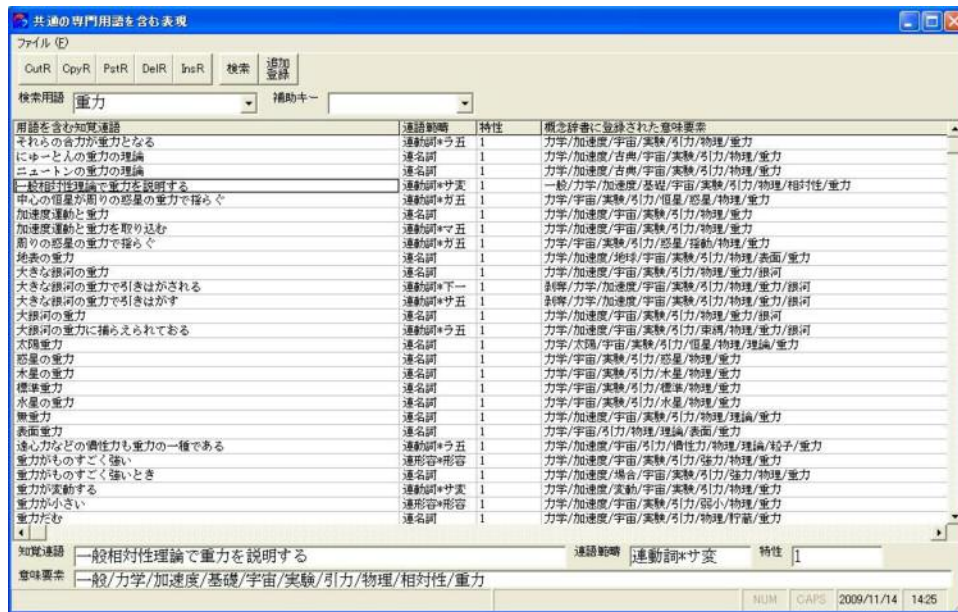


図 4 共通の専門用語を含む表現

この節の最後に述べるのは、難解な専門用語を含むテキストを同義同値の知覚連語を検索することによって、当該の用語の意味を理解するための言い換え表現を具体的に提示するものである。

図 5 は専門用語に関連する叙述をテキストとして入力することによって、テキストに含まれる専門用語を抽出し、想定される意味要素を推定し、

テキストの構文解析の結果や、意味的に距離の近い知覚連語を表示するものである。課題文を書き込み、分野を選択し、意味的距離の制限半径を書き入れ、同義検索ボタンを押すことで解析を遂行する。同義文として表示されたもののうち、専門用語を含まない表現が専門用語の意味内容を別の表現で叙述していることが容易に観察できる。



図 5 専門用語を含む文と意味的に近い知覚連語の検索 (専門用語を含む場合)

図 5 に示されたテキストは専門用語を含む叙述である。同じインターフェイスで専門用語含まないで同義の文をテキストとして課題文に入力す

るとどうなるだろうか? 図 6 はそのような場合の例である。テキストとして“物体がその質量に比例して受ける力”という叙述が入力されてい





を意味的に距離の近い知覚連語を一つの同値類(同義性同値類)として扱う技術に適用した。知覚連語はその形成規則によって学習されるものであるから今一つの知覚連語の同値類,つまり共通の単語を含むものという意味での同値類(共通語同値類)が存在する。二つの同値類は知覚連語について二種類の検索技術を提供する。共通語同値類の検索はいわばキーワードによる検索であるから,旧来からよく知られたものである。我々は共通語同値類の検索の効率を上げるために,知覚連語と共通語の相関関数と逆相関関数を記憶する M 言語の大域変数を定義しておく手法を導入した。同義性同値類に属する知覚連語の検索技術は新しい検索技術であり,同じ意味内容を異なる表現の仕方で表現している知覚連語の検索技術である。

コミュニケーション支援の立場からみれば,前述の二技術は日本語文の簡易化表現を与える技術として捉えられる。複文は聾者にとって分かりにくいものとされているが,複文の要となる接合名詞に形式名詞が来ると難解さはさらに増加する。こうした問題に対し我々の技術は,問題の日本語文を,意味の具体的な日常名詞を用いた単文に機械的に分解する手法を与える。

二種の同値類を組み合わせた検索技術は,一般的に,二者のコミュニケーションで一方が知らない言葉をもう一方が言い換えによって意趣を伝えるコミュニケーションの手法を与える。この技術は障害などによって系統的な言葉が分かりにくい場合の情報提供の方法としても重要である。たとえば,オノマトペは系統的に聾者に分かりにくいとされるが,我々の方法はオノマトペを文脈に従って別の言葉で表現する方法を与える。

**第 5 節**では,二種の同値類を有効に活用するためのインターフェイスの設計について述べた。専門用語辞典の編集と専門用語-知覚連語の相関関係の登録を行うもの,共通語同値類の検索と概念辞書の追加登録を行うもの,課題文を構文・概念解析し,解析結果と課題文に意味的に近い知覚

連語を抽出するものの 3 つであるが,これらと先の知覚連語-意味要素相関辞書を連携的に活用することによって,日本語解析システム「ささゆり」の,正確かつ広汎な専門用語理解の支援機能としての有効性が拡大していくと期待される。

## 引用文献

- [1] 高橋 亘, “日本語解析システム「ささゆり」における日本語文簡易化の方法と知覚連語間の意味的距離”『総合福祉科学研究』, Vol. 1, 91~100 (2010).
- [2] 高橋 亘, 『コミュニケーション支援の情報科学』, 現代図書 (相模原, 2007, 4 月).
- [3] 高橋 亘, “日本語解析システム「ささゆり」の言語学”, 『Proceedings 2007 M Technology Association of Japan』, 14 ~ 18 (2007).
- [4] 高橋 亘, “日本語解析システム「ささゆり」の基礎を与える言語学”, 『関西福祉科学大学紀要』, Vol. 11, 41~48 (2008).
- [5] 高橋 亘, “M 言語による日本語解析システム「ささゆり」の意味解析--- 連体修飾のある日本語文の意味解析 ---”, 『Mumps』, Vol. 24 (2008) 27~33.
- [6] 高橋 亘, “日本語解析システム「ささゆり」における連体修飾のある日本語文の意味解析”, 『関西福祉科学大学紀要』, Vol. 12, 21~30 (2009).
- [7] 宮地絵美, 高橋 亘, “M 言語による聾者のための日本語簡易化機能--- 連体修飾のある日本語文の単文化と形式名詞の意味推定 ---”, 『Mumps』, Vol. 24, 35~40 (2008).
- [8] 高橋 亘, 宮地絵美, “聾者のための日本語簡易化法 --- 連体修飾のある日本語文の単文化と形式名詞の意味推定 ---”, 『関西福祉科学大学紀要』, Vol. 12, 31~39 (2009).
- [9] 津村雅稔, 高橋 亘, “オノマトペを含む日本語文の M 言語による代替表現機能---聾者のための情報保障の技術---”, 『Mumps』, Vol. 25, 23~33 (2010).

# オノマトペを含む日本語文の M 言語による代替表現機能

## ---聾者のための情報保障の技術---

### Alternative Representation Function of Japanese Sentences with Onomatopoeia on the M Language Scheme

#### ---Information Presentation Techniques for the Deaf---

津村雅稔, 高橋 亘

Masatoshi Tsumura and Wataru Takahasi

関西福祉科学大学社会福祉学部

〒 582-0026 大阪府柏原市旭ヶ丘 3-11-1

TEL 0729-78-0088, FAX 0729-78-0377

E-mail takahasi@fuksi-kagk-u.ac.jp

**要旨** 日本語解析システム「ささゆり」による、聾者にわかりやすい日本語表現を提供する機能が考察される。特に、聾者に分かりにくいとされるオノマトペを含む日本語文の代替表現を提供する原理とインターフェイスが詳説される。

あらかじめオノマトペの範囲を決め、使われ方を類別するためのオノマトペ辞書がシステムに内包させられた。ある単語がオノマトペであるかどうかの判定はこの辞書への問い合わせによって決定される。

オノマトペの使用例を列挙する技術として、日本語解析システム「ささゆり」が対象とする知覚連語の中に共通のオノマトペを含むという意味での知覚連語の同値類が考察される。

次に、オノマトペを含む個々の知覚連語に意味的に近い知覚連語の集合という同値類が考察される。この同値類は、その意味によってオノマトペを含む知覚連語を類別する。同義性同値類に含まれる知覚連語には、指定されたオノマトペを含んでいるものと含んでいないものが存在するから、このオノマトペを含まない表現が代替表現の候補になる。

**キーワード** 日本語文の簡易化, コミュニケーション支援, M 言語, 日本語解析システム「ささゆり」, オノマトペ, 代替表現

## 1. はじめに

この論文は、先に同じ題目で第 36 回日本 M テクノロジー学会において口頭発表された後、“オノマトペを含む日本語文の代替表現機能---聾者のための情報保障の技術---”と題する論文として、論文誌「総合福祉科学研究」に投稿されたもの [1] に、特にユーザインターフェイスの操作性を中心に、その後の発展を付け加えて、理論及び技術の整合性と有効性を強調するものである。

我々の研究室では、近年、日本手話と日本語の構造比較から聾者への日本語の情報提供のあり方に関する研究を続けてきた [2-5]。テレビの字幕放送や要約筆記による情報保障など、聾者への情報保障を考えると、聾者に分かりやすい日本語、分かりにくい日本語という二分律が重要である。聾者に分かりやすい文章には、大きく分けて次の 3 つの配慮があると考えられる [6]。

- (1) 構文の適切な選択
- (2) 語彙の適切な選択
- (3) 日本手話と日本語の言語構造の相違に関する配慮

2008 年から 2009 年にかけて、我々は、日本語解析システム「ささゆり」を用いて、特に (1) と (2) の問題に関連があると考えられる形式名詞を含む日本語文の簡易化の技術を考察した。形式名詞は数多くの意味を保持しており、これを文や句が修飾することによってはじめて明確な意味が確定されるものである。したがって構文の困難さと語が形式化することによる多義性ということが問題であった。このような問題に対して、日本語解析システム「ささゆり」は構文解析の技術と意味解析の技術の両面からの特性を遺憾なく発揮したと言える [4-6]。

日本語解析システム「ささゆり」は、知覚連語を単位にして日本語文を切断する。通常の言語において単語自体は明確な意味を保持しない。語が結合して、ある程度の長さを持つ連語を形成するときに意味が純粹化されるのである。純粹化された意味は明確な知覚を誘発する。知覚

連語とはこの明確な知覚を誘発する連語のことである。このような知覚連語を単位として日本語文を切断する我々の日本語解析システムは、正確な意味解析の機能を保持している。

形式名詞を含む日本語文の簡易化を研究していく中で、意味が限定された形式名詞を日常的な名詞で置き換えるという問題が生じたとき、我々は、ある意味要素の組と日常的な名詞との意味的距離を明確にする必要に迫られたのである。知覚連語の間の意味的距離の問題は本誌の別の論文 [7] で、筆者の一人によって明らかにされる。

言葉と言葉の意味的距離を扱えるようになるとさらに別の視界が開けてくる。異なる表現の間で意味的に近い表現の集合は、一般的に、難解な表現を分かりやすい表現に言い換える手法を与える。この論文で我々が問題にしたいのは、聾者が系統的に語彙として苦手とする単語を聾者に分かりやすい表現で伝えようという新しい情報保障の技術である。この論文ではこうした技術の内、特にオノマトペに焦点をあてて議論したい。

オノマトペ (onomatopée) はフランス語を原語として、直訳すれば擬声語であるが、擬音語と擬態語の総称として使われている。日本語百科大辞典 [8] によれば、擬音語とは、「動物や人間など生物の発する声音、生物や物体が起こす物音、自然界で発せられるさまざまな音響を、言語音でできるだけ忠実に模倣して表現したことば」であり、擬態語とは、「生物の動作、容態、感覚、感情、心理状態、事象の状態・変化といった音響とは直接かかわりないものを言語音によって象徴的に表現することばである」とされている。

一方、聾者がコミュニケーション手段として用いている日本手話は、「日本手話は聾社会で発達した自然言語で、聾社会の中で使用されてきた」[9] と述べられるように、日本語とは独立した自然言語として発達している。したがって日本手話は日本語とは異なる言語構造を持つ言語であると言える。日本手話は視覚言語であり、



視覚における表象性の強いものである。一方、日本語のオノマトペ、特に擬音語は音声を模して表現される語であるから、聴者にとっては表象性の高いものであるが、この表象性は、音声を聞くことのない聾者にとっては何のリアリティーもないものである。二つの言語の表象性のずれによって、オノマトペを含んだ日本語文は聾者に意味が通じにくいことがしばしばある。このような事情から、オノマトペは一般に聾者にとってわかり難いものとして考えられている[6, 10]。さらにこの事情の二次的影響として、聾者が記した意味の分からないことばの特徴としてオノマトペが多いことが観察されている[11]。聾者にとって分かりにくいオノマトペを含む日本語文を、同じ意味を持つ別の表現に言い換えることによって、もとの文と言い換え文との対比から、オノマトペそのものの聾者による理解を深める技術を確立することがこの論文の主題である。

聾者がオノマトペを含んだ日本語文の理解に困難があることにより社会における人とのコミュニケーションや情報保障に支障が生じていることが考えられる。聾者への情報保障支援として母語の手話による手話通訳を介した支援もあるが、書字日本語を用いたテレビの字幕表示や講義・講演等における要約筆記による支援が併せて行われている場合も多い。また、手話を介さない場合の聴者と聾者を含めた聴覚障害者による主なコミュニケーション手段として口話と筆談、近年の携帯電話によるメッセージがあげられる。会話の内容や相手によって、口話と筆談を使い分けているが、相手の口の動きが読み取りにくいときや、話の内容が複雑なときなどの状況では、口話より確実性の高い筆談を用いてコミュニケーションを行っている[12]。聾者が母語である手話による言語理解だけでなく、日常生活において触れる機会が多い書記日本語を理解可能な言語とするために、聾者に分かりやすい日本語表現による情報保障支援が求められるのである。

我々は日本語解析システム「ささゆり」に新

たにオノマトペ辞書を内包させ、M 言語による、聾者のためのオノマトペの言い換え機能を構築することを試みた。

論文 [7] で述べられるように、日本語解析システム「ささゆり」によって機械学習される知覚連語は、二種の同値類を持っている。二種の同値類とは、共通の単語を含む知覚連語という意味での同値類と意味的に近い知覚連語の集合という意味での同値類である。我々は、M 言語の階層性を活用した検索技術を用いて、共通の単語としてオノマトペを含む知覚連語の同値類を、意味的に近い知覚連語の同値類でラベルし、部分集合に分類することで、聾者の感覚にそった言い換えを実現する方法を考察してみたい。

このような技術は、聾者のための分かりやすい文字情報を提供する技術として、テレビの字幕放送や要約筆記による情報保障の技術として活用されることが期待される。

## 2. オノマトペ辞書とオノマトペの使用例を操作するインターフェイス

オノマトペは一般に副詞に属するが副詞の総てがオノマトペではない。したがって、オノマトペの範囲を決め、使われ方を類別する辞書が必要である。我々は、このような目的を果たすものとして、オノマトペ辞書を日本語解析システムに内包させた。

オノマトペ辞書の内容を記憶している大域変数は ^NONOMAT でその構造は、

```
^NONOMAT(Onomatopoeia,No)
=Mean_ “^” _ADD
```

のような形をしている。ここで、大域変数をラベルする Onomatopoeia と No は、それぞれ、オノマトペとそれを類別する番号である。大域変数の値は、代表的なオノマトペの意味 (Mean) と副詞の使われ方にしたがって付加する変化語尾 (ADD) をセパレータ “^” を挟んで接合したものである。このような大域変数を我々は図 (図 1) のようなグリッド型のインターフェイスを用いて編集することが出来る。

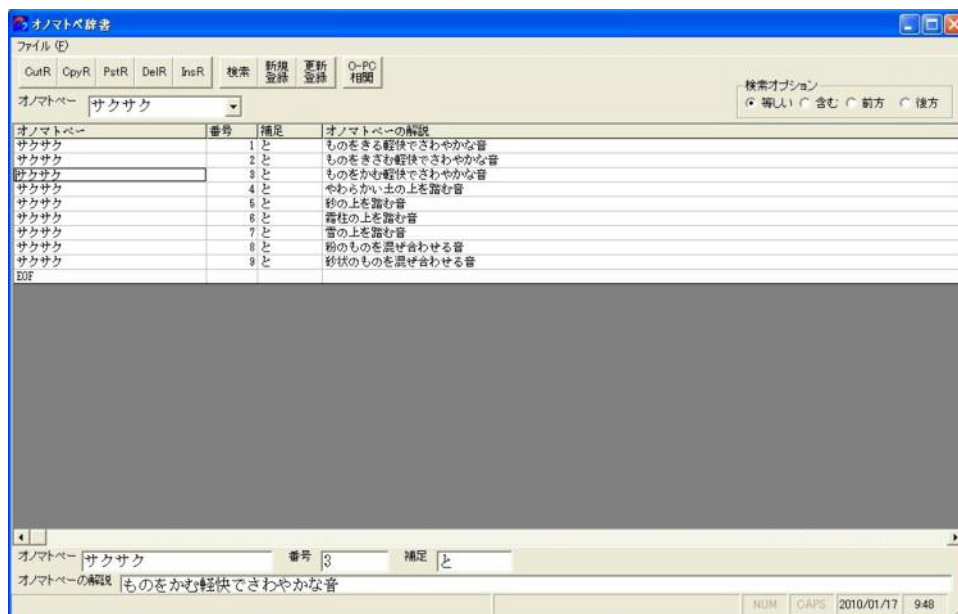


図 1 オノマトペ辞書 (オノマトペの定義と概観を操作)

検索のタイプを検索オプションで指定し、検索ボタンを押すことによって既に登録されているデータが表示される。左上に並ぶ、CutR, CpyR, PstR, DelR, InsR の 5 つのボタンは、セルを一行単位で、切り取り、コピー、貼り付け、削除、挿入を行うためのものである。グリッドの個々のセルを編集した後、新規登録ボタンを押せば、既に登録されているデータを削除することなく、新しく追加されたデータのみを追加登録し、更新登録ボタンを押せば、検索時グリッドに呼び出されたデータを一旦削除した後にグリッド編集後の新しい情報にしたがって、データを登録し直す。図では“サクサク”というオノマトペが検索され、9 種類の意味が既に登録されている。登録内容は、オノマトペに関する辞書 [13] や国語辞典、インターネット上に公表されている解説、知覚連語の使用例などを参考にして決定する。

オノマトペにはカタカナ表記のものがあり、カタカナ表記と平仮名表記では少しニュアンスが異なるという説もあるが、このような微妙なニュアンスの差は我々のシステムでは区別しない。したがって、このデータ登録システムには、カタカナ表記のオノマトペが登録されると、自動的に平仮名表記のものが同じ意味で登録される機能が付加されている。

我々の立場ではこの辞書に登録されたものがオノマトペなのである。ある単語がオノマトペであるかどうかの判定はこの辞書のデータが存在するかどうかで決定される。我々が最初に構築するのは、オノマトペに属する個々の単語に対して、日本語解析システム「ささゆり」が対象とする知覚連語の中に共通のオノマトペを含むという意味での知覚連語の同値類（共通語同値類）を生成する機能である。この機能はオノマトペの使用例を列挙する技術でもある。

日本語解析システム「ささゆり」では、総ての知覚連語が大域変数  $\wedge$ NWDIC の形で記憶されているが、この大域変数を検索して、特定の単語が含まれているかどうかを判断する相関関数を生成する一般論は論文 [7] で述べられている。ここでは、一般論にしたがって知覚連語とオノマトペの相関関数と逆相関関数を記憶している M 言語の大域変数、

$\wedge$ NWCWAMPO(Col,Onomatopoeia),

$\wedge$ NWCWIAMPO(Onomatopoeia,Col)

を生成する。ここで Col は知覚連語を、Onomatopoeia はオノマトペを表している。それぞれのデータ変数の値を如何に設定するかは自由度が残されているが、M 言語の \$Data 関数を用いさせれば大域変数が存在すれば 1 が、しな

ければ 0 が応答として返ってくるので、その値が何であれ、大域変数の存在、非存在によって目的を達成することが出来る。相関関数に対応する大域変数 ^NWCWAMPO は知覚連語がオノマトペを含んでいるかいないかを判定することに用いられ、逆相関関数に対応する大域変数 ^NWCWIAMPO は特定のオノマトペを含む知覚連語を列挙することに用いられる。図 1 における [O-PC 相関] ボタンがこれらの大域変数の登録に対応するものである。

これまで知覚連語とオノマトペとの相関関数と逆相関関数に対応する大域変数を生成する機能の大綱を述べてきたが、この機能を完成させるには、いま一つ注意すべきことがある。それは同じ形態を持つが品詞の異なるものが存在する構成要素の問題である。つまり同一形態の品詞的多重性の問題である。たとえば、“くすり”は擬音語であることも [薬剤] の意味を持つ名詞であることもある。また、“ふっ(ふっど)”は擬音語・擬態

語であることも“ふる”という動詞の連用形であることもある。このような名詞や動詞を誤ってオノマトペとして扱わないようにするためには判断すべき構成要素に引き続く構成要素を監視してそれが副詞に後続するのにふさわしくないものである場合はオノマトペとしてピックアップしないような機能が必要である。例として、後続するとオノマトペでないと判断してよい後続詞は、“が”、“の”、“に”、“を”のような格助詞や、“て”のような接続助詞のようなものが挙げられる。このような機能を内包させることによって、我々は、オノマトペの定義とその使用例である知覚連語の間に相関関係を生成する機能を構成することが出来る。

この段階で、オノマトペの使用例を具体的に表示し、意味空間の構築に重要な役割を果たすインターフェイスについてふれておきたい(図 2)。オノマトペの使用例はすなわちオノマトペに関する共通語同値類である。

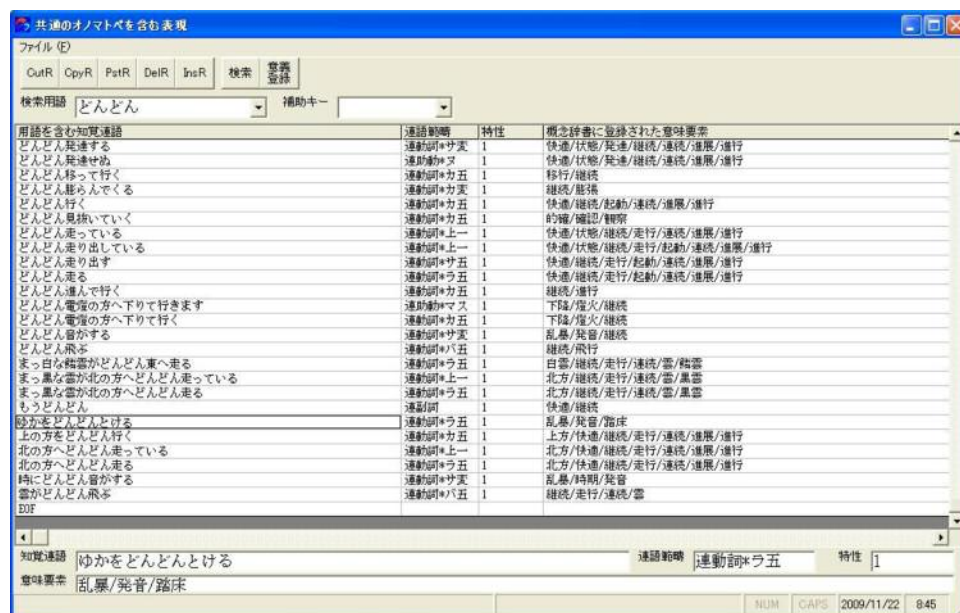


図 2 オノマトペの使用例を表示するインターフェイス

図 2 で示されるインターフェイスでは、検索ボタンを押すことで、大域変数 ^NWCWIAMPO の第二階層をたぐって、オノマトペを含む知覚連語、すなわちオノマトペの使用例、を表示する。そしてこの大域変数の値から連語範疇と特性を把握し、これを表示する。知覚連語、連語範疇、

特性の 3 つの情報が与えられれば概念辞書に対応する大域変数 ^NCDIC の値を見ることによって、知覚連語に振りあてられている意味要素の組を表示することができる。グリッドの最右列に表示されているのは、この意味要素の組である。このインターフェイスは概念辞書への意味要素

の追加登録機能も保持している。オノマトペの使用例の列を観察することによって不足している意味要素をグリッドに書き込み、意義登録ボタンを押すことで、追加された意味要素が概念辞書に反映される。この意味でインターフェイスは概念辞書の整合性を保つ上で、積極的な機能を保持しているのである。

### 3. オノマトペを含む知覚連語に意味的に距離の近い知覚連語の集合

二つの知覚連語の意味的距離の定義は筆者の一人によって既に定義が与えられている[7]。二つの知覚連語があつて、それぞれの知覚連語が意味要素の組  $A, B$  を持っていたとする。  $A, B$  の要素数がそれぞれ  $n_A$  個,  $n_B$  個であつたとし、これらの意味要素の中で共通のものが  $n_{A \cap B}$  個あつたとすれば、知覚連語の間の意味的距離は次の式で定義される。

$$d_{AB} = \frac{n_A + n_B - 2n_{A \cap B}}{n_{A \cap B}}.$$

この定義では、 $A$  と  $B$  が集合として等しい場合に意味的距離が 0 となり、 $A$  と  $B$  に共通する意味要素がないときに無限大となる。(もちろんコンピュータで無限大の数量を扱うことは出来ないで、情報科学的には意味的距離の最大数をコンピュータで扱いうる最大数によって制限することになる。) また、この定義は、隔たりの程度と共通性の程度の比であることから、隔たりの程度の大きいものに対して値が大きく、共通性の程度が大きいものに対して値が小さくなるという特性も知覚連語の間の意味的距離の性質をよく保持している。

我々は第 2 節で、一つのオノマトペを含む知覚連語の集合を如何に定義するかを考えた。いま、知覚連語どうしの意味的距離を考えると、一般的に、指定された単語を含んでいる各知覚連語には、これと意味的に近い知覚連語の集合という同値類(同義性同値類)が考えられる。意味的に近い

という同値性はある単語を含んでいるか、いないかという同値性とは独立したものであるから、この同値類に含まれる知覚連語には、指定された単語を含んでいるものと含んでいないものが存在する。一つのオノマトペを決めると、これを含む知覚連語の集合  $A$  に属する各知覚連語  $X_i$  について、意味的に近い知覚連語の集合  $B_i$  が存在する。つまり、一つのオノマトペを含む知覚連語の集合は、その個々の要素に意味的に同値な知覚連語の集合によって類別される。ここに知覚連語の集合  $B_i$  の一つ一つの要素にはオノマトペを含んでいるものとないものが存在する。あるオノマトペが難解だと考えられる人がいて、この人にこのオノマトペの理解を支援する言い換え表現は、このような知覚連語の類別に登場する各々の意味的同値の集合における当該のオノマトペを含むものと含まないものとの照合体系である。

意味要素の組が与えられたときに、この意味要素の組から意味的距離が指定された範囲内にある知覚連語を高速検索する技術はこの節の冒頭で引用した論文 [7] に詳説されている。

### 4. オノマトペを含む知覚連語の代替表現を判断する機能とその操作インターフェイス

この節では、いくつかの実例をあげ、前の二節で考察した一般論を用いて、実際にオノマトペを含む知覚連語の意味の決定法と、知覚連語の代替表現を見つける機能の動作について考察し、実際に聾者の実用に提供するインターフェイスについて述べたい。

まず、例として、オノマトペ“どんどん”について、現在我々のオノマトペ辞書に登録されている“どんどん”の意味は次の 8 個である。

- ① “大きくひびく連続音”
- ② “大砲を発射する音”
- ③ “花火を発射する音”
- ④ “戸を強くたたく音”
- ⑤ “床を荒々しく踏み鳴らす音”
- ⑥ “水がはげしくながれてぶつかる音”



⑦ “物事がつごうよくはかどるようす”

⑧ “あとからあとから続くようす”

これらの意味の記述はいずれも知覚連語の連動詞によって名詞が修飾されているものであるから、先の論文 [4, 5] で修飾子と接合名詞<sup>1</sup>の対応関係と呼ばれている関係を呈している。この対応関係によってオノマトペの意味を規定しているのである。これらの表現が代替表現の候補の一つを与えるが、この表現自体は複文であるから聾者に分かりやすい構文形式を持っていない。

そして日本語解析システム「ささゆり」の知覚連語辞書に登録されている“どんどん”というオノマトペと相関関係のある知覚連語をリストすると、大半は ⑦ もしくは ⑧ と同義で使用されているものであり、最初のいくつかを挙げると次のようなものである。

- ・ “お星さまたちの下をどんどんどんどんかけて行く”
- ・ “そこをまっ黒な雲が北の方へどんどん走っている”
- ・ “そこをまっ黒な雲が北の方へどんどん走る”
- ・ “そのうえどんどんちぢんでゆく”
- ・ “その下をまっ白な鱗雲がどんどん東へ走る”
- ・ “ただ雲がどんどん飛ぶ”
- ・ “どんどんあるいていく”
- ・ “どんどんあるく”
- ・ “どんどんかけて行く”

そして、特に擬音語と考えられるものは、次のもののみで、非常に少ないことが分かる。

- ・ “どんどんとける”
- ・ “ゆかをどんどんとける”
- ・ “どんどん音がする”
- ・ “時にどんどん音がする”

これらは、先の例を含めて、日本語解析システム

「ささゆり」が折にふれて機械学習したものである。

擬音語の中で“どんどんとける”と“ゆかをどんどんとける”については、知覚連語概念相関辞書で次のような意味要素が登録されている。

- ・ “どんどんとける”；[乱暴/発音/踏床]
- ・ “ゆかをどんどんとける”；[乱暴/発音/踏床]

ここで、直線的な推論からは少し外れるが、我々が置かれている状況を概観するために、知覚連語概念相関辞書に、これら二例のオノマトペを含む知覚連語と関連する、オノマトペを含まない知覚連語がどのように登録されているかを見ておきたい。二例のオノマトペを含む知覚連語とオノマトペ辞書からは、オノマトペ以外のキーワードとして、“床”、“踏み鳴らす”、“踏みならず”などのキーワードが想定されるので、これらをキーとして、やはり知覚連語概念相関辞書を検索すると、次のような意味要素が登録されていることが分かる。

- ・ “踏み鳴らす”；[発音/踏床]
- ・ “床板を踏み鳴らす”；[発音/踏床]
- ・ “荒々しく踏み鳴らす”；  
[乱暴/発音/踏床]
- ・ “床を荒々しく踏み鳴らす”；  
[乱暴/発音/踏床]
- ・ “一同が床板を踏み鳴らす”；  
[発音/踏床/集団]
- ・ “踏みならず”；[発音/踏床]
- ・ “床板を踏みならず”；[発音/踏床]
- ・ “荒々しく踏みならず”；  
[乱暴/発音/踏床]
- ・ “床を荒々しく踏みならず”；  
[乱暴/発音/踏床]

以上が知覚連語概念相関辞書の登録状況である。

このような状況下で、“ゆかをどんどんとける”というオノマトペを含む知覚連語が与えられると、

“ゆかをどんどんとける”

<sup>1</sup> 複文で修飾文から修飾を受け、骨格文で体言となる名詞。

---> これが知覚連語として保持する  
意味要素の組

---> [乱暴/発音/踏床]

---> これと意味的距離が 0.5 以下  
の知覚連語

とたどると、つぎのような知覚連語がリストされることになる。

- “どんどんとける”；連動詞\*ラ五
- “ゆかをどんどんとける”；連動詞\*ラ五
- “床を荒々しく踏みならす”；連動詞\*サ五
- “床を荒々しく踏み鳴らす”；連動詞\*サ五

• “荒々しく踏みならす”；連動詞\*サ五

• “荒々しく踏み鳴らす”；連動詞\*サ五

これらの内，“どんどん”を含む表現を，“どんどん”を含まないものを例示してみせることによって，“どんどん”のこのコンテキストでの意味が了解されると思われる。

ここで、以上のような機能を実際に聾者の実用に提供するインターフェイスについて述べる。

図 3 に示されるのがオノマトペを含む知覚連語と意味的に近い知覚連語を探索し表示するインターフェイスである。



図 3 オノマトペを含む文と近い意味を持つ知覚連語

この画面では、問題になる課題文と意味的距離の境界を与える数値（境界距離）を入力し、同義検索ボタンを押すことによってオノマトペの欄に課題文が含むオノマトペを、意味要素の欄に課題文を意味解析した結果想定される意味要素の組を表示する。そして同義文の欄にはこの意味要素との意味的距離が境界距離の範囲内にある知覚連語を表示する。各行の最も右側にはその知覚連語と意味要素との距離が表示されていることが分かる。

このインターフェイスはオノマトペを含む文

のみに有効なのではない。図 4 はオノマトペを含まない場合や複文などが入力されたときにどのように働くかを示している。インターフェイスはオノマトペを含むか含まないかにかかわらず課題文を意味解析する。また複文が入力されたときには複文を解析し、接合名詞が保持する意味を想定して、これに最も近い日常語を割り当てる機能を持っている。図では“物事がつごうよくはかどるようす”という複文を解析し、接合名詞である“ようす”に対し、“進行状態”という意味の明確な連語名詞を示唆している。



図 4 オノマトペを含まない文や構造を持った文

現時点では概念辞書の登録が十分ではないが、今後概念辞書が充実するにつれこうした機能は正確さを獲得していくことが期待される。そしてこのようなインターフェイスの正確な応答を目指して、概念辞書の意味要素の充足をしていく視点や道筋を与えるという意味でもインターフェイスの有効性が理解できる。

## 5. まとめと展望

日本語のオノマトペは擬音語と擬態語の総称として使われている。音声を模して表現される擬音語は、聴者にとっては表象性の高いものであるが、音声を聞くことのない聾者にとっては何のリアリティーもないものである。表象性のずれによって、オノマトペを含んだ日本語文は聾者に意味が通じにくいことがしばしばある。

日本語解析システム「ささゆり」の知覚連語間の意味的距離を測定する方法を用いて、オノマトペを含む日本語文を、オノマトペを含まない同義の表現に変換することによって、聾者に分かりやすい表現を提供する方法を考察した。

我々は、オノマトペの範囲を決め、使われ方を類別するためのオノマトペ辞書を日本語解析システムに内包した。ある単語がオノマトペである

かどうかの判定はこの辞書への問い合わせによって決定される。

オノマトペ辞書には指定されたオノマトペの複数の意味が収録されている。その個々の意味の記述は多くが知覚連語の連動詞によって名詞が修飾されている書式をもつものであり、修飾子と接合名詞の対応関係になっている。この対応関係によってオノマトペの意味が規定されるから、これらの表現が代替表現の候補の一つを与えるが、この表現自体は複文であるから聾者に分かりやすい構文形式を持っていない。

我々は、オノマトペに属する個々の単語に対して、日本語解析システム「ささゆり」が対象とする知覚連語の中に共通のオノマトペを含むという意味での知覚連語の同値類（共通語同値類）を生成する機能を構築した。この機能はオノマトペの使用例を列挙する技術でもある。我々は、与えられたオノマトペを含む知覚連語の検索効率を向上させるために、知覚連語とオノマトペの相関関数と逆相関関数を記憶している M 言語の大域変数を定義した。我々はこの逆相関関係を活用してオノマトペの使用例を高速に表示し、意味空間の構築に重要な役割を果たすインターフェイスを構築した。このインターフェイスでオノマト

ペの使用例を観察することによって追加された意味要素が即時に概念辞書に反映される。この意味でインターフェイスは概念辞書の整合性を保つ上で、積極的な機能を保持している。

我々の、オノマトペの理解を支援する言い換え表現を実現する論理は次のようにまとめられる。知覚連語どうしの意味的距離を考えると、指定されたオノマトペを含むある知覚連語に意味的に近い知覚連語の集合という同値類（同義性同値類）が考えられる。同義性同値類に含まれる知覚連語には、指定されたオノマトペを含んでいるものと含んでいないものが存在するから、このオノマトペを含まない表現が代替表現の候補になる。

一つのオノマトペを決めると、これを含む知覚連語の集合に属する個々の知覚連語について、同義性同値類が存在する。つまり、一つのオノマトペを含む知覚連語の集合は、その個々の要素に対応する同義性同値類によって類別される。類別された知覚連語の集合の一つ一つの要素にはオノマトペを含んでいるものといないものが存在する。あるオノマトペが難解だと考えられる人がいて、この人にこのオノマトペの理解を支援する代替表現は、このような知覚連語の類別に登場する意味的同値の集合におけるオノマトペを含むものと含まないものとの照合体系である。

我々のシステムにおいて、オノマトペを含む知覚連語の意味を決定する機能と、知覚連語の代替表現を見つける機能は次のように動作する。我々のシステムでは、現実的な文の中でオノマトペを含む知覚連語が与えられると、

知覚連語

---> これが知覚連語として保持する

意味要素の組

を知覚連語概念相関辞書で検索し、

知覚連語が保持する意味要素の組

---> これと意味的距離が 0.5 以下

の知覚連語

とたどって、代替表現の候補を与える知覚連語がリストされる。

我々は以上のような機能を実際に聾者の実用に提供するインターフェイスを具体的に構築した。このインターフェイスはオノマトペを含む知覚連語と意味的に近い知覚連語を探索し表示するだけでなく、オノマトペを含まない場合や複文などが入力された場合にも、課題文を意味解析する。それが複文であった場合には構文解析して、接合名詞が保持する意味を想定して、これに最も近い日常語を割り当てる機能を持っている。これは聾者に難解なもう一つの典型、つまり複文や形式名詞を含む文の理解に寄与するものである。

最後に、本文中に示した新しい検索機能を実用化するインターフェイスのいくつかについて言えば、こうしたインターフェイスは単独で活用するというよりは連動させて活用することによって、さらに概念辞書を充実させる効用を持っていると言える。聾者への具体的な支援に資するためにはインターフェイスの表示内容をもう少し簡略にしてインライン化を図るべきであるが、こうした問題については将来の問題としたい。

## 引用文献

- [1] 高橋 亘, 津村雅稔, “オノマトペを含む日本語文の代替表現機能---聾者のための情報保障の技術---”, 『総合福祉科学研究』, Vol. 1, 115 ~ 122 (2010).
- [2] 高橋 亘, 仲内直子, 宮地絵美, 村上裕加, “日本手話と日本語の構造比較と聾者にわかりやすい日本語の表現”, 『関西福祉科学大学紀要』, Vol. 10, 75~82 (2007).
- [3] 高橋 亘, 仲内直子, 宮地絵美, 村上裕加, “聾者の日本語使用データベースに観る日本手話的言語感覚”, 『Proceedings '07 M Technology Association of Japan』, 24~29 (2007).
- [4] 宮地絵美, 高橋 亘, “M 言語による聾者のための日本語簡易化機能 --- 連体修飾のある日本

語文の単文化と形式名詞の意味推定 ---” , 『Mumps』 , Vol. 24, 35~40 (2008).

[5] 高橋 亘, 宮地絵美, “聾者のための日本語簡易化法 --- 連体修飾のある日本語文の単文化と形式名詞の意味推定 ---” , 『関西福祉科学大学紀要』 , Vol. 12, 31~39 (2009).

[6] 高橋 亘, 『コミュニケーション支援の情報科学』 , 現代図書 (相模原, 2007, 4 月).

[7] 高橋 亘, “知覚連語の同値性と日本語文簡易化の方法 --- M 言語による日本語解析システム「ささゆり」の意味解析 ---” 『Mumps』 , Vol. 25, 9~21 (2010).

[8] 金田一春彦, 林大, 柴田武, 『日本語百科大辞典』 , 大修館書店 (1988).

[9] 福田友美子, 赤堀仁美, 乗富和子, 木村晴美, 津山美奈子, 鈴木和子, 市田泰弘, “聾者間の対話を対象にした日本手話の研究” , 『電子情報通信学会技術研究報告』 WIT99-1~22[福祉情報工学], 第二種研究会資料 Vol. 99, No. 1, 15~22 (1999).

[10] 岡田美里, 高橋 亘 “聾者の日本語使用データベースと聾者にわかりやすい文字情報” , 『関西福祉科学大学紀要』 Vol. 9, 185~192 (2006).

[11] 米川明彦, 『手話ということば』 , PHP 研究所 (2002).

[12] 水野映子, “聴覚障害者の職場におけるコミュニケーション--聴覚障害者・企業対象の調査にみる現状と課題” , 『Life Design Report (2007/11・12)』 (2007).

[13] 小野正弘, 『擬音語・擬態語 4500 日本語オノマトペ辞典』 , 小学館 (2007).



# M 言語による PIC シンボル・プロセッサの実現と 視覚記号の言語学

## Development of a PIC Symbol Processor on the M Language Scheme and the Linguistics of Visual Codes

池田茉莉子, 柳内英二, 高橋 亘

Mariko Ikeda, Eiji Yanagiuchi and Wataru Takahasi

関西福祉科学大学社会福祉学部

〒 582-0026 大阪府柏原市旭ヶ丘 3-11-1

TEL 0729-78-0088, FAX 0729-78-0377

E-mail takahasi@fuksi-kagk-u.ac.jp

**要旨** 二語文脈のメッセージを構成するシンボル・プロセッサの新機能が構築された。第一の新機能は、想定される二語文脈データを視覚的に編集するものである。二語文脈データはメッセージ編集のための候補となる PIC シンボルの表示を制御する働きを持っている。この機能は PIC シンボル (視覚記号) の自然な接続を観察するために重要である。

第二の新機能は一つのカテゴリーに属するシンボルの分類記号についての再記号化機能である。この機能は優先順位の高いシンボルを前に持ってくるのに重要である。

二つの新機能の活用は、視覚記号の表現と接続に関連する言語学的問題の観察を容易にする。「名詞 動詞」、「形容語 名詞」などの接続の性質、色彩を表す記号の表現法、副詞や固有名詞の表現法、機能語が少ない記号体系における時間発展の叙述の方法など、視覚記号の言語学の新しい知見が述べられる。

**キーワード** 知的障害児, 自閉症児, コミュニケーション支援, PIC シンボル, シンボル・プロセッサ, 表象性と含意性, M 言語, 二語表現

### 1. はじめに

この論文は、先に同じ題目で第 36 回日本 M テクノロジー学会において口頭発表された後、“PIC シンボル・プロセッサの実現と視覚記号の言語学”と題する論文として、論文誌「総合福祉科学研究」に投稿されたもの [1] に、その後の発展を付け加えて理論及び技術の整合

性と有効性を強調するものである。

知的障害児や自閉症児の中には日常的な言葉に興味をほとんど示さない児童がしばしば見かけられる。言葉はコミュニケーションの手段を与えるだけでなく、思考の手段を与え、認知の根幹をなす重要なものである。言葉に興味を示さない子供たちの中に、ものの形や色に興味を示す人がいることに着目して、我々は近年 PIC シンボルによって意思表示が可能なシンボル・

プロセッサーの開発に取り組んできた。

PIC (Pictogram Ideogram Communication) とは、ピクトグラム (Pictogram) と呼ばれる具象的なシンボルとイデオグラム (Ideogram) と呼ばれる抽象的なシンボルで構成され、拡大代替コミュニケーション AAC (Augmentative and Alternative Communication) の一つとして S. C. Maharaj (1980 ; Canada) によって開発されたものである。現在、約 1600~1700 のシンボルが世界的に普及している。PIC シンボルのような表象記号によるコミュニケーションが近年、知的発達障害者のコミュニケーション支援に特に有用であると考えられている[2-4]。

AAC ツールには、絵カード、コミュニケーションボード、コミュニケーションブックなどがあるが、これらの試みは、言語記号としてのシンボルの選択候補を数多く空間的に配置することに制御上の問題を抱えている。

こうした欠点を是正するためには情報機器の導入が避けられない。我々はコンピュータを活用して PIC シンボルを系統的に表示するシステムの開発をすることにした。日本語に興味を示さない児童に PIC シンボルで言語感覚を芽生えさせるために、我々は当初から二語文を構成できるシンボル・プロセッサーの開発を目指した。二語文の形成が児童に文法習得の一步となるという認識からである。児童に言語感覚を誕生させるには、発生した言語感覚が無理なく高次の言語感覚に結びつけられるものでなければならない。我々は M 言語による階層型データを活用して二語文脈を制御するシステムシステムを考案した[5-8]。

もちろん知的障害者の発達段階には個人差がある。ピアジェ (Piaget, J.) の発生的認識論を、知的障害者に適用したインヘルダー (Inhelder, B.) の発達理論によれば、知的障害者の発達段階は、次の通りである。重度・最重度レベルは感覚運動的知能段階、中度レベルは前操作段階、軽度レベルは具体的操作段階にあたる。中でも感覚運動的知能段階では表象段階に達している

か否か、前操作段階では、象徴機能が発達しているかどうかの区切りが重要な意味を持っていると考えられる[9]。各段階について発達段階に即応したツールの開発が非常に大切である。

開発が進むにつれて、現在日本で使用されている PIC シンボル自体に多くの言語学の問題がみつき、PIC シンボルの言語学的記号としての進化が必要であると考えられるようになった。日本語に興味を持たない児童にとっては、PIC シンボル自体が日本語に取って代わりうる言語にならなければならない。象徴機能が発達している児童にとってこのことが重要なばかりでなく、象徴機能が十分に発達していない幼児にとっても、自然な象徴機能の発達促進のために適切な表象言語が必要である。PIC シンボルは、日本手話がそうであるように、PIC シンボル自体が、日本語とは独立して一つの言語体系を形成する必要がある。このような視点から、我々は、PIC シンボル・プロセッサーを、障害のない児童が自然に二語文を生成できるプロセッサーとして完成させ、そこから象徴機能の発達段階に応じて、使用の簡易性を考慮し、使用できる文脈を児童に即して制限して行く方式を採用することにした。

採用するシンボルは、知的障害児にとって使用容易性が高く、日常生活を送る上で必要度の高い語彙を優先する必要がある。そして、使用頻度の高いシンボルをすぐに引き出すことができる配慮が必要である。このような視点から見れば、知的障害者に概念認知が容易な PIC シンボルが意外に少ないということがわかる。したがって、知的障害児の目線に合わせた、PIC シンボルの改良、あるいは新しく作成することが必要である。現在使用されている PIC シンボルの総数は先述の通りであるが、表現したい思念があっても表現に必要な語彙が用意されていないという問題がある。例えば施設でよく行う知的障害児の作業で“ポスターを貼る”と言いたくとも、[ポスター]と[貼る]の語彙が PIC シンボルにない。また名詞シンボルと動詞シン

ボルの接続の問題もある。童話に現れるメタファーが表現できないのである。「アリとキリギリス」のような童話を表現しようとして“キリギリスが歌う”を表現しようとする昆虫の〔キリギリス〕に人間の〔歌う〕は接続できない。このような言語学的ないくつかの課題を我々は先の論文で指摘している[7, 8]。

この論文で議論する新機能の第一は、二語文脈の自然な接続を観察できる二語文脈データの視覚的編集機能である。新機能では視覚的に記号としてのシンボルの接続を観察しながら文脈データを登録したり、削除したりすることを可能にしたい。二語文脈データの視覚的編集機能については第3節で述べられる。

第二は、PIC シンボルの使用頻度を考慮して、一つのカテゴリーに属するシンボルのソートを適切に変更する、再記号化の機能である。二語文脈は元来カテゴリーとシンボルが二組、つまり4階層の構造で決まるものであるが、一つのカテゴリーに属するシンボルの再記号化が4階層に即座に反映される機能については第4節で述べられる。

二つの新機能のもとに、我々はPIC シンボルの言語学についてさらにいくつかの新しい観察を得ることが出来た。これらの観察は多くの新しいPIC シンボルの創作を促すものである。第5節ではこれらの知見のいくつかを紹介される。

## 2. 二語文脈を編集するシンボル・プロセッサ

我々はこの数年、知的障害児や自閉症児の意思表示のAAC支援装置として、二語文脈メッセージを編集できるPIC シンボル・プロセッサの開発を行ってきた[5-8]。この節では、我々のメッセージ編集インターフェイスの基本的な設計と背後にあるシンボル繰り出し操作のメカニズムをレビューしておきたい。

二語文脈メッセージの編集インターフェイスは次のようなものである。つまり、限られた画面に多くのシンボルを表示するためには、メッ

ッセージを表現するためのシンボルのカテゴリー化が必要であり、一語に対してカテゴリー分類に対応するカテゴリー・シンボルと具体的にメッセージを構成するメッセージ・シンボルの対がコンピュータの画面半分に表示される必要がある。カテゴリー分類に関して、我々は幼児の二語文に現れるカテゴリーを中心に障害児の日常生活への必要性や概念形成上の必要性を考慮した19のカテゴリーを用意した。それらは、①人物、②植物、③動物、④道具、⑤乗物、⑥衣類、⑦食物、⑧飲物、⑨おやつ、⑩属性、⑪感情、⑫行為、⑬場所、⑭時間、⑮気象、⑯数量、⑰指示、⑱否定、⑲疑問の19カテゴリーである[5-8]。



図1 シンボルの連合関係

メッセージ編集画面では、これらの内の一つがクリックされることにより、そのカテゴリーに属するメッセージ・シンボルが、クリックにしたがって、20を単位にして周期的に繰り出されるようになっている。(図1) こうした一つのカテゴリーを指定して、そのカテゴリーに属するメッセージ・シンボルの選択を促すことは、ソシュールの言う連合関係を規定していることになる。連合関係を操る背後のメカニズムを仕切るものとしては、階層型データベース、M言



語の大域変数

^PICTREE(Categ,Obj)

=PicPath\_“^”\_PicName

が定義されている。ここで大域変数の二つの添え字 Categ、Obj はそれぞれカテゴリー・シンボルとメッセージ・シンボルの ID であり、カテゴリーの選択は第一添え字の指定によって決定される。メッセージ・シンボルの周期的繰り出しは第二添え字を手繰ることによって機械的に行われる。右辺の PicPath、PicName はそれぞれ PIC シンボルの画像ファイルへのパスと PIC シンボルを日本語に対応させる際の呼称であるがデータとしてはこれらが“^”を挟んで接合されている。画面に繰り出される各シンボルはこの大域変数の検索機能に従っているわけである。

メッセージ編集画面の最近の変更点として、背景色のデザインが挙げられる。クリックによってカテゴリー・シンボルやメッセージ・シンボルが選択された場合にシンボルを表示するピ

クチャー・ボックスの縁を、通常時平坦に設定されているものを立体表現に変えて選択状態を強調するように設計されているが、この差が、背景色が薄いグレーであると、識別されにくいという欠点を持っていた。そこで、長期使用の目の保護という観点も含めて、背景色を濃い緑を主張とする色彩に変更した。

第一カテゴリーのメッセージ・シンボルがクリックされると、選択されたメッセージ・シンボルが下のメッセージボックスに表示され、その下のテキストボックスに対応する日本語表現が表示される。

第一カテゴリーに属するメッセージ・シンボルが選択された段階で、二語文の後続部編集用の画面としての右半分の動作が開始される。想定される文脈に応じて、第二カテゴリーのカテゴリー・シンボルが表示され、第二カテゴリーのシンボルのクリックに応じて、第二カテゴリーに属するメッセージ・シンボルが、やはり 20 個を単位に周期的に表示される(図 2)。



図 2 二語文の連辞関係を編集するシンボル・プロセッサのインターフェイス

この段階で、カテゴリー・シンボルとメッセージ・シンボルの組は二セットがすべて表示さ

れることになる。第二メッセージ・シンボルが選択されると、二語文が決定される。第一メッ

セージ・シンボルと第二メッセージ・シンボルの組がソシュールの言う連辞関係をなす。

二語文の連辞関係の表示を背後で制御しているのは 4 階層を持つ、M 言語の大域変数

```
^ConnectAmp(Categ1,Obj1,Categ2,Obj2)
=Func
```

である。引数の前二つ `Categ1` と `Obj1` がそれぞれ二語文先行詞のカテゴリ・シンボルとメッセージ・シンボルの ID であり、引数の後二つ `Categ2` と `Obj2` は、それぞれ後続部のカテゴリ・シンボルとメッセージ・シンボルの ID である。右辺の大域変数の値は文脈に応じて想定される助詞を記憶している。この大域変数の存在、非存在にしたがって二語文後続詞のカテゴリ・シンボルとメッセージ・シンボルの表示が制御されるわけであるから、この変数を登録したり削除したりすることで想定される文脈を定義したり排除したりすることが出来るわけである。今後データ `^ConnectAmp` に言及するときには“二語文脈データ”と呼ぶ。

後続詞メッセージ・シンボルのクリックに応じて、選択されたメッセージ・シンボルが下のメッセージボックスの先述部分に後続するように付加され、その下のテキストボックスに、先述部分に加えて、想定される助詞を挟んだ、対応する日本語表現が追加される。

この節の最後として、二種の大域変数 `^PICTREE` と `^ConnectAmp` の編集方式について触れておく必要がある。今までのシステムでは、これらの変数の編集に対し、グリッド型のインターフェイスを持つ文字ベースの編集画面のみが用意されていた[2, 3]。しかし、我々の研究が進むにつれて、絵記号を接続させて、その絵記号の列が一つの言語としての役割を担わなければならないことがはっきりと自覚された段階で、絵記号の言語記号としての接続の問題を詳細に観察する必要性が生じたのである。**第 3 節**ではこのような目的に合致する二語文脈データの視覚的編集機能の設計について述べる。

### 3. 二語文脈データの視覚的編集機能

PIC シンボルが絵記号としてそのまま言語記号となるためには、統語規則の第一歩を与える二語文の自然な構文が絵記号によって構成されなければならない。通常の音声言語では、記号の無契性によって二つの記号の接続に対して何の制約もないが、絵記号を言語記号とするためには二つの記号の接続に強い制約が働く。それは、絵記号の本来的な意義として、記号が絵として強い表象性を保持する点にある。言葉に興味を示さない児童に対して絵記号によって意志を表現させようとする試みは、絵記号が表象性を保持してこそ、音声言語とは異なった立場を提供することが出来るのである。しかし、強く表象性を保持している二つの絵記号を構文として接続させるには、この表象性にいくらか制限を加える必要がある。「アリとキリギリス」の童話があるとして、“キリギリスが歌う”、“キリギリスが食べる”などの内容を絵記号で表現しようすると、“きりぎりす”に対応する PIC シンボルを作成することは難しいことではない。リアルなキリギリスの絵を一枚用意すればよいのである。しかし、“歌う”や“食べる”に対応する絵記号が人の絵を用いて表現されていれば、二記号の接続に不自然さが生じる。このような構文を可能にするためには動詞の表現が、人間によらない、表象性を少し押さえたものになる必要がある。このような事情を子細に観察しながら想定される文脈を設定するには、二語文脈データの視覚的編集機能が欠かせない。

我々は二語文脈データの編集を視覚的に行えるインターフェイスを導入することにした。これには、既に構成されているメッセージ編集用の画面が参考となる。二語文脈のシンボルを表示するためにメッセージ編集画面では次のステップを踏んでシンボルを繰り出している。

- (1) 第一カテゴリ・シンボルを表示。  
(`^PicTree` に従う)
- (2) 第一カテゴリ・シンボルの選択  
(`Categ1` の決定) に従って第一メッセー

ジ・シンボルを表示。( ^PicTree に従う)

(3) 第一メッセージ・シンボルの選択 (Obj1 の決定) に従って、メッセージボックスに第一メッセージ・シンボルを表示し、第二カテゴリー・シンボルを表示。(二語文脈データで制限)

(4) 第二カテゴリー・シンボルの選択 (Categ2 の決定) に従って第二メッセージ・シンボルを表示。(二語文脈データにより制限)

(5) 第二メッセージ・シンボルの選択 (Obj2 の決定) に従って、メッセージボックスに選択された第二メッセージ・シンボルを追加し、第一カテゴリー・シンボルを除く総てのシンボルを画面から消去。

以上がシンボルを繰り出すステップであるが、(5) で第一カテゴリー・シンボルを除く総てのシンボルを画面から消す動作を止めた画面が、想定される可能な文脈を編集する画面の原型を与える。この原型に対して (4) の段階で第二カテゴリー・シンボルの選択に応じて表示される

第二メッセージ・シンボルを総て立体表示にするようにすれば、この立体表示が文脈の存在を意味するようになる。(第二メッセージ・シンボルのクリックによって立体表示と平面表示を反転させることで選択状態を示すことが出来る)

以後、これを原型として二語文脈編集画面の設計を考察する。第二メッセージ・シンボルが選択された段階で、メッセージボックスに表示された二シンボルに対応する ^ConnectAmp の 4 つの添え字、Categ1、Obj1、Categ2、Obj2 の総てが確定しているわけであるから、メッセージボックスのシンボルの列を観察して、この文脈が必要の無いものであれば、対応する大域変数

^ConnectAmp(Categ1,Obj1,Categ2,Obj2) を削除すればよい。まず、画面に [文脈削除] のボタンを用意する。[文脈削除] ボタンを押すことで上述の大域変数が削除され、第二メッセージ・シンボルの総てを文脈データにしたがって再表示するようにする(図 3)。

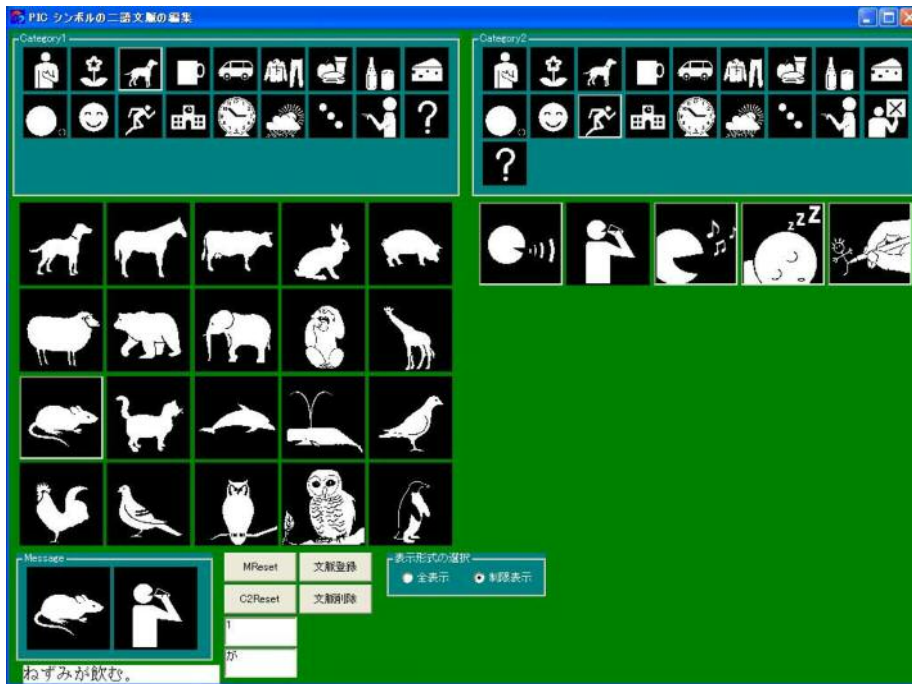


図 3 制限表示編集; 不要な文脈データを削除したり既存の文脈データの値を変更したりする

図 3 にはさらにいくつかの機能が付加されている。まず、次の 2 つのボタンを追加する。

(i) MReset ボタン; メッセージボックスのメッセージをすべてクリアし、対応する



第一メッセージ・シンボルの立体表示を解除し、画面右半分を解除。

(ii) C2Reset ボタン; メッセージボックスの第二シンボルをクリアし、第二メッセージ・シンボルを文脈データにしたがって再表示。

以上の基本的な装置の他に、文脈の存在 (1) 非存在 (0) と文脈に想定される日本語文の助詞を表示するテキストボックスを追加しておく。ここで、既に登録されている文脈データの値である助詞を変更するためのボタンとして [文脈登録] ボタンを追加する。このようにして既に登録されている制限された文脈のデータの修正・削除をする編集画面は、ひとまず図 3 の段階に到達する。(この段階で文脈データの追加登録も射程に置いて、表示形式の選択オプションを配備しておく)

文脈の中には今は制限されているが追加したい文脈もある。組合せとして存在する総ての文脈 (それが必要なものであるかどうかは別として) を表示し、追加したい文脈を選択させるには、上述のステップの (3)、(4) の段階を次のよ

うに変更する。

(3) 第一メッセージ・シンボルの選択 (Obj1 の決定) にしたがって第二カテゴリー・シンボルを表示。(PicTree に定義されているカテゴリー・シンボルを総て表示。; これが全表示であり、先の選択オプションで切り分ける)

(4) 第二カテゴリー・シンボルの選択 (Categ2 の決定) にしたがって第二メッセージ・シンボルを表示。(PicTree に定義されているカテゴリー・シンボルを総て表示し、二語文脈データにより想定される文脈を与える第二メッセージ・シンボルのみを立体表現にする)

この設定で文脈が制限されている第二メッセージ・シンボル (平坦表示) が選択された場合、これを立体表示に変え、メッセージボックスに文脈登録候補を表示する。候補の文脈データを、助詞表示テキストボックスに書かれた助詞を値として、登録する機能を [文脈登録] ボタンに追加して、二語文脈データベースの視覚的編集画面は、一応完成する(図 4)。

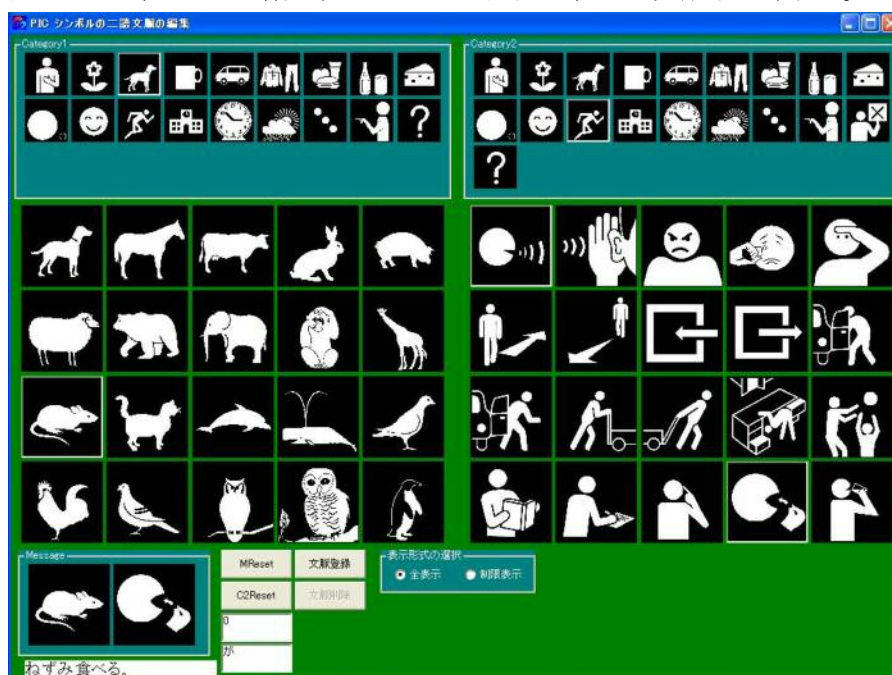


図 4 全表示編集; 存在し得る文脈を全表示して必要な二語文脈データを登録する

しかし、この文脈データの編集機能はまだ十分ではない。これまでの設計では、文脈データ

の削除・追加の操作が個別的で時間がかかりすぎる。第二メッセージ・シンボルをクリックす

る毎にクリックされたオブジェクト名 (Obj2) をリストボックスに蓄えておき、蓄えられたオブジェクト名をもとに文脈データを一括削除するような機能が欲しい。

我々は、表示形式オプションの右にリストボックスを追加した(図 5, 6)。

制限表示編集の場合(図 5)、クリックされた

第二メッセージ・シンボルのオブジェクト名が削除の候補としてリストに追加され、シンボルが平面表示になるようにする。平面表示のシンボルのクリックで立体表示にもどり、リストから取除かれるようにすれば、削除候補の取り消しもできる。[一括削除] ボタンで不要な文脈データの一括削除を実行する。

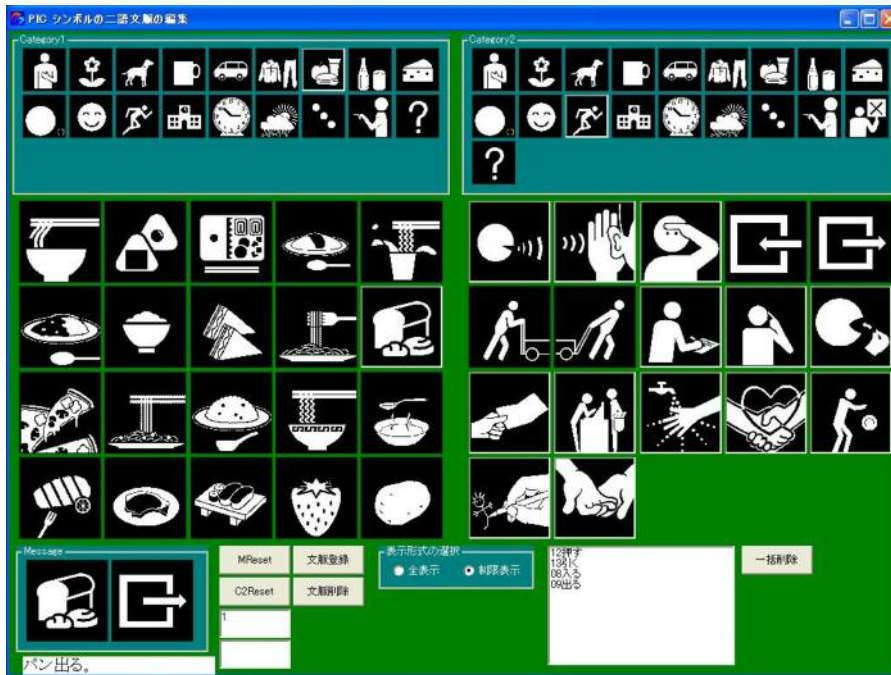


図 5 制限表示編集の修正; 一括削除を可能にする

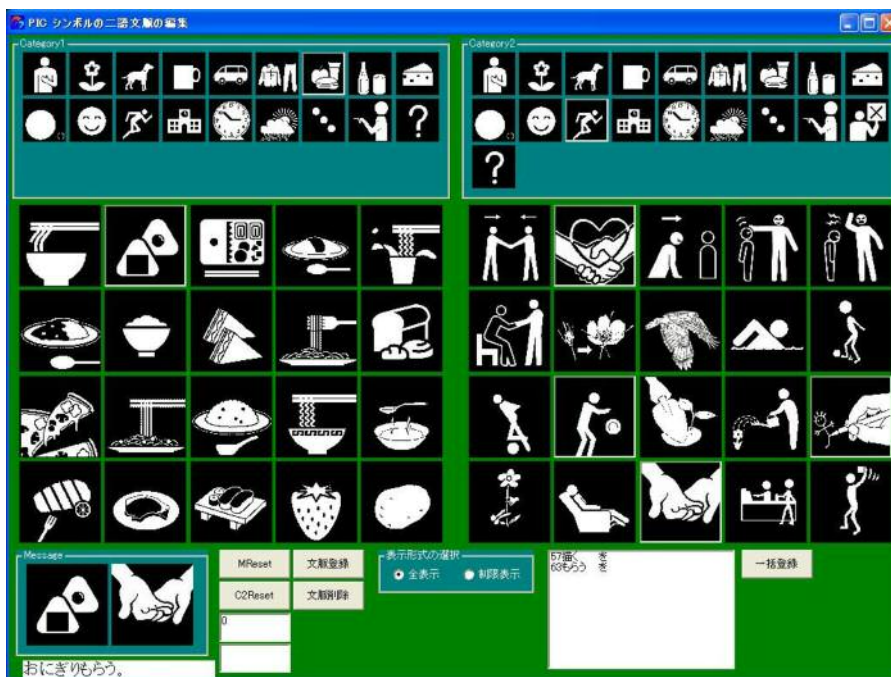


図 6 全表示編集の修正; 一括登録を可能にする



全表示編集についての場合(図 6)、[一括削除] ボタンは [一括登録] ボタンに変更され、リストボックスは登録候補をリストするようにする。第二メッセージ・シンボルについても、平面表示のものをクリックすることで、オブジェクト名のリストへ追加が実行され、立体表示に切替えられる。立体表示になったものを再びクリックすることでリストからはずすことも可能である。リストに追加されるテキストは第二メッセージ・シンボルのオブジェクト名と助詞表示用のテキストボックスのテキストを HT<sup>1</sup> を挟んで結合させた文字列である。一括削除と同様に一括登録ボタンでリストに対応する文脈データを一括登録することが出来る。

#### 4. 一つのカテゴリーに属するシンボルの分類と再記号化

一つのカテゴリーに属するメッセージ・シンボルの表示される順序は、よく使われるものほど前に来る方がよい。このことは筆者の一人によって、2008 年度の修士論文で主張された。我々のシステムを活用するにつれて妥当な表示順序が当初の想定と異なってくるのが想定される。例えば動物というカテゴリーを考えたときに我々の扱っているメッセージ・シンボルは生物学的な分類に従うと大きく分けて、① 哺乳類、② 鳥類、③ 爬虫類、④ 両生類、⑤ 節足動物、⑥ 軟体動物の六種類である。当初は必要に応じて暫定的に順序を決めていたものを、まずはこの順序に並べ替えてみたい。またさらに活用が進むにつれて利用者の使用頻度などに合わせて順序を変更してみようといった要望が想定される。

M 言語の大域変数の照合順序は添え字のアクセス番号、シフト JIS 番号などの順であるから、この問題はメッセージ・シンボルの ID の付け替えで解決する。例えば、生物学的分類に従わせるには、元々の ID の前に先の分類順

に 01、02、03、04、05、06 などの数値を付加することで目的が達せられる。

しかし問題は、メッセージ・シンボルの ID が大域変数 ^PicTree の第 2 添え字、大域変数 ^ConnectAmp の第 2 添え字、第 4 添え字に入っているのであるから、これらを総て矛盾のないように付け替える必要がある。M 言語のプログラミングとしては難しいものではないが、定義されているデータを消去してしまうことのないようにプログラムするには少し注意が必要である。古い ID (OCode) を新しい ID (NCode) に変換し、変換した大域変数 ^ConnectAmp の個数 (Z) を返す関数

```
Set Z=$$^ReCodeObj(OCode,NCode)
```

を Appendix として添付する。

新旧 ID の対応表をカンマもしくは HT 区切りのファイルで与えて、関数 ^ReCodeObj を用いて対応表にある総ての ID の書き換えを行うプログラムは簡単であるから、ここでは省略する。

#### 5. PIC シンボルの表現と接続の言語学

第 3 節で述べられた視覚的二語文脈編集インターフェイスを活用するにつれ、PIC シンボルの表現と接続に関連する多くの言語学的課題が浮上してきた。この節では、これらの言語学的課題についての我々の観察と対応、今後の見通しについて述べたい。

##### (1) 名詞と動詞の接続

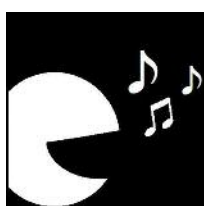
名詞に対応する PIC シンボルと動詞に対応する PIC シンボルとの接続関係は、先の論文で述べたように動詞の表象性と含意性との間には大変難しい背反性がある。童話のような多くの擬人化や隠喩が入るものをメッセージ化させようとするとその問題はさらに顕著になる。(第 3 節の冒頭で述べたように、文脈編集の視覚的インターフェイスが必要になる所以でもある。“ギリギリスが歌う”を表現するのに、人をイメージさせる[歌う]は不適切である(図 7, 8)。

<sup>1</sup> Horizontal Tabulation

図 7 [キリギリス]<sup>2</sup>

図 8 [歌う]

“歌う”という行為が人に固有の行為でないという認識をもって表現すると図(図 9)に示すような新作の PIC が出来る。この方が人にも他の動物にも接続がよくなる。

図 9 新作 [歌う]<sup>3</sup>

“食べる”を表現する動詞も、既成の PIC シンボルでは人が何かを口に入れようとしている様子で表現している。これは人が何かを食べる様子を表現するのであれば問題ないが、動物が餌を食べるときに使うには適切な絵記号とは言えない。障害児の認識を高めるのにシンボルの表象性は重要なものであるが、記号の接続が自然になるためには記号の表象性を少し抑えて、記号の含意性を高める必要がある。障害児が“ねずみがチーズを食べる”という情景を思い描いたり、実際食べている場面をみたりして、その様子を頭に思い浮かべて、メッセージで伝えようとするとき、二語文で表現するには、[ねずみ] [食べる]、[チーズ] [食べる] という具合になる。[キリギリス]、[ねずみ]、[私] などとの接続関係を考えると [食べる] は既成のもの(図 10)よりは新作の [食べる](図 11)の方が接続性が良い。

[ねずみ](図 12)との接続を考えるとその自然さが判断できる。既成の [食べる] の PIC

<sup>2</sup> 既成の JIS カテゴリー・シンボルから独立させたものである。

<sup>3</sup> 関西福祉科学大学 2008 年度卒業生、佐々木 亜里紗と仲地侑子による。

シンボルでは、“ねずみ”が食べている絵ではないので、障害児が作文するにも、理解するにも、“食べている”主体は“人間”であり、“ねずみが食べる”という意味に理解されるのは難しい。

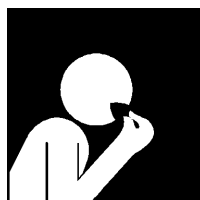


図 10 [食べる]

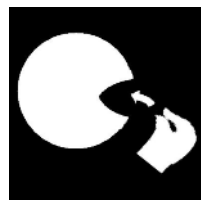
図 11 新作 [食べる]<sup>4</sup>

図 12 [ねずみ]

目的語の [チーズ](図 13)に対しても新作のもので十分に自然な表現が実現される。新作の PIC シンボルは [チーズ] を“食べる”のは [私] でも [ねずみ] でも接続がよい。



図 13 [チーズ]

ここで日本語の語順に関連した問題が浮上する。「名詞 動詞」の構文の格の問題である。上記の例では [ねずみ] と [チーズ] があがっているので主語になるのは [ねずみ] で目的語になるのは [チーズ] という自明性が主語と述語を決定している。文脈によっては [ねずみ] が目的語になる可能性がある。“猫がねずみを食べる”のような場合である。PIC シンボルの格はどのように決めればよいのか。日本語では助詞が格を決定する。つまり“猫が食べる”、“ねずみを食べる”と助詞を挿入することでこの問題は解決する。しかし、利用者の状況を考えると

<sup>4</sup> 脚注<sup>2</sup>に同じ

助詞の導入には大きな難度がある。助詞の導入が困難であれば格を語順が担う言語をモデルにすることはどうだろうか。英語なら〔ねずみ〕〔食べる〕は“ねずみが食べる”を意味し〔食べる〕〔ねずみ〕は“ねずみを食べる”を意味する。何れの格の方式が言語に興味のない知的障害者や自閉症児にとって難度が少ないのかは現場における試行が必要な所である。

PIC シンボルと同様に視覚的表象性の強い日本手話では格の問題をどのように考えているのだろうか。手話は動作で表現されるため、多くの手話単語は「主語 動詞」もしくは「目的語 動詞」、「手段 動詞」、「方法 動詞」のような組み合わせになっている。手話をモデルに作成された新作 PIC シンボルを次に挙げておく。

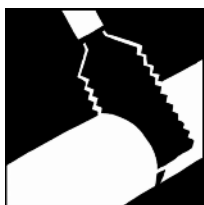


図 14 〔鋸で切る〕



図 15 〔鋏で切る〕



図 16 〔包丁で切る〕 図 17 〔カッターで切る〕  
いずれも表現された概念は一義的で理解されやすい。

既成の PIC シンボルでも、〔顔を洗う〕、〔手を洗う〕、〔食卓の用意をする〕、〔スイッチをつける〕などは名詞と動詞の組み合わせられた成句が一つの PIC シンボルとして表現されている。一般に、「名詞 動詞」の成句を一つの PIC シンボルにすると、一義的で分かりやすいという長所がある。しかし、この方式はもう一方で、一つの動詞をいろいろな名詞につけることがで

きないという意味で記号の含意性の低下を生じさせるので、必要な PIC シンボルの数が膨大になり、選択の困難さを増大させることになる。記号の表象性と含意性には一方を増大させるともう一方が減少するというある種の背反性があるのである。

## (2) 形容詞と名詞の接続

日本語の語順では、「名詞 形容詞」の場合は叙述的であり、「形容詞 名詞」の場合は修飾的である。日本手話では「名詞 形容詞」、「形容詞 名詞」はいずれも意味的差がなく捉えられが、重要なものを先に述べるということが一般化されているので、「名詞 形容詞」の形で表現されることが多い。

視覚記号の場合は、「名詞 形容詞」、「形容詞 名詞」のように逆に並べても意味伝達度にあまり差はないように思われる。知的障害者や自閉症児はどちらの順番で表出するのも変則的であり、「形容詞 名詞」という言い方をよくする。どちらの言葉を強く思うかで順番が決定されるようである。

PIC シンボルは、原則的には色を着けず白黒で表現することになっているが、知的障害児には色に強い関心を示すものが多い。意志を表示することの少ない障害児が、身近な物品を指さすことなどにより、色彩を訴えている光景は施設でもよく観察されるものである。その意味では色を表現する PIC シンボルが必要である。ここに新作 PIC シンボルを例示しておきたい。



図 18 〔青い〕

特定の形を与えず、色彩のみを提供するような PIC シンボルとして、図 18 のようなものが適切であろう。ここでは印刷の仕上がりが白黒であることを考慮して、一色のみを例示している

<sup>5</sup> これら 4 つは関西福祉科学大学 2008 年度卒業生、川上美幸による。

が、赤、緑、青、ピンク色、水色、黄色などの他、白、黒、灰色、茶色なども必要である。これらは中央の領域をそれぞれの色にすることで作成される。

日本語では名詞を修飾する形容語は「形容詞」と「形容動詞」の二種類があるが、これらは PIC シンボルでは同じ品詞と考えて良い。こうした理解の上に立った上でも、形容語を視覚的に表現することは割合と難しいものである。特に含意性の高い形容語の表現は動詞と同じく困難を極めると思われる。やはり形容する対象の表象性を用いて表現することが容易である。“きれいだ”という形容語は既成の PIC では [美しい女性] や [衣類の清潔さ] で表現されているので、自然に形容される「名詞」は限られてくる。

### (3) 副詞の表現

学部学生が卒業研究で絵本の LL 化<sup>6</sup>を考えたときに判明したことであるが、PIC シンボルで副詞を表すものがない。“雨がぼつぼつ降る”、“雨がざあざあふる”という日本語の文を PIC シンボルで表現しようとして、はたと困ったわけである。PIC シンボルには副詞の語彙が用意されていない。しかし、既成の [雨] (図 19) が絵記号としては“雨がざあざあふる”の意味で使用される可能性があることを考えると、相対的に表現すれば、[雨がぼつぼつ降る] を、含意性の低い、対象に依存した画像を構成することはそれほど困難なことではない(図 20)。



図 19 [ざあざあ降る] 図 20 [ぼつぼつ降る]

日本語における副詞の表現には、最初に導

入されたときの意味に対する表象性があり、その後使用されている間に違った意味で使用されるようなものが少なくない。[とても]、[非常に]、[ものすごく]、[きわめて]などは同じ表現であるし、[霧] [とても] は“霧が深い”の意味で使用される[10]。このような意味では PIC シンボルでも、何かの理由で導入されたものが、使用が重なるにつれ、別の意味で使用されるようになり、転用されるものがある。

### (4) 固有名詞の表現

固有名詞は、特に、障害児の身近な人物や施設、店舗などの日常生活に関連するものや絵本・物語の主人公などの趣向に関連するものが必要であると考えられるが、これに相当する絵記号を事前に想定することができないから、個別的にならざるを得ない。随時、必要とされる固有名詞は新作で対応する必要がある。

固有名詞を PIC シンボルで作成する方法は、身近な人や施設などの場合は写真を、童話など、物語の主人公の場合は原画をそのままスキャナで取り込んで使用すればよいだろう<sup>7</sup>。学部学生の卒業研究で対象とされた絵本「はっばのおうち」[11] では、主人公“さち”に対して図 (図 21) のような [さち] が使用された。



図 20 主人公 [さち]

### (5) 機能語が少ない記号表現における時間的発展の叙述

肯定や否定の意を表現するのに [うれしい]、[悲しい] の表現が効果的であるということは先の論文 [7, 8] で述べた。一般的に PIC シン

<sup>6</sup> LL はスウェーデン語の Lättläst の略で、「易しく読める」の意。英語の ER (Easy to Read) に相当。絵本などに PIC シンボルのメッセージを付与すること。

<sup>7</sup> 現時点では研究のための試作のみであるが、現場で使用するなどの実用化が進めば、このような複製使用は出版社及び著者と著作権に関する許諾が必要となると思われる。



ボルで機能語を表現することは困難である。時制の助動詞に至ってはなおのことである。漢文などの例にも見られるように視覚的言語の場合に時制の助動詞がきわめて少なくなることは避けられないのかも知れない。日本手話においても漢文においても時制の多くは時を表す名詞が明示的に入ることによって時間を表現している [10]。日本手話では、時間発展を表現するために叙述の途中に過去・完了を意味する [過去詞 (手おろし)] を挟んで叙述を中断して、次々に接続していく話法がとられる。多くの絵本では 1 頁に記入される日本語文の叙述が結構長いものであるが、LL 化の際に創作絵本にする場合には 1 頁に述べられる叙述を少なくして、場面展開のコマ数を増やすことによって時間発展を表現する手法が取られた。

## 6. まとめと展望

近年我々は、言葉に興味をほとんど示さない知的障害児や自閉症児の意思表示を支援するために、PIC シンボルによって意思表示を可能にするシンボル・プロセッサの開発に取り組んできた。我々は、言葉はコミュニケーションの手段を与えるだけでなく、思考の手段を与え、認知の根幹をなすという基本的な認識のもとに、障害児に言語感覚を芽生えさせるシステムの構築に取り組んでいる。

我々は、二語文の形成が児童に文法習得の一步となるという認識を持ち、当初から二語文を構成できるシンボル・プロセッサの開発を目指してきた。我々のシステムでは二語文の先行詞の候補を 19 のカテゴリーに分類されるシンボルを 2 階層の大域変数  $\wedge\text{PicTree}$  に従って繰り出し、後続詞の候補を二語文脈データ  $\wedge\text{ConnectAmp}$  に従って繰り出す。二語文脈は、カテゴリーとシンボルの 2 セットによって決まるから  $\wedge\text{ConnectAmp}$  は 4 階層の構造を持っている。候補となるメッセージ・シンボルの繰り出しは、系統的で無理がなく、児童に言語感覚を誕生させるために、発生した言語感覚

を無理なく高次の言語感覚に結びつけるようなものでなければならない。

この論文で議論された新機能の一つは、想定される二語文脈データを視覚的に編集する機能である。

我々はシンボル・プロセッサのメッセージ編集用画面をモデルにして、まず、二語文脈の後続詞の選択に対して、二語文脈データによって制限された候補のみを表示する画面を構成した。この画面には、編集作業として、不要な二語文脈データを削除する [文脈削除] ボタンと既に登録されている文脈に想定される助詞を更新するためのボタンが必要である。助詞の更新用のボタンは、後で未登録の文脈を追加する機能を追加することを想定して [文脈登録] ボタンとした。

新しい文脈を二語文脈データとして登録するには、組み合わせとして可能な総ての後続詞が画面に表示されなければならない。全表示 (組み合わせとして可能な後続詞の候補を総て表示する) と制限表示 (二語文脈データに応じて後続詞の候補を表示する) の双方はオプションボタンで切り替えられるようにインターフェイスを設計し、文脈データとして必要なものを登録する役割は前もって設置されている [文脈登録] ボタンに兼務させた。

以上の二語文脈編集機能は一々の文脈に対して作業が個別的であるから効率が悪い、我々はさらに削除候補や登録候補の ID を蓄えておくリストボックスを用意して、一括削除、一括登録を効率良く行うことが出来るようにした。

もう一つの新機能は、一つのカテゴリーに属するシンボルの分類記号 (ID) についての再記号化機能である。我々のシステムでは、画面に繰り出されるメッセージ・シンボルは M 言語の大域変数 (メッセージ・シンボルの ID を引数としている) によって制御されているので、優先順位の高いシンボルを前に持つるために、一つのカテゴリーに属するメッセージ・シンボルの ID をデータの照合順序を考慮して



付け替えればよい。シンボルの繰り出しに関わる大域変数は ^PicTree (第二階層がメッセージ・シンボルの ID) と ^ConnectAmp (第二階層と第四階層がメッセージ・シンボルの ID) の二つであるから、これらのデータの階層を逐次たどって更新に対応する総ての ID の書き換えを完了するプログラムを作成した。

二つの新機能、特に二語文脈のデータベースを視覚的に編集する機能は PIC シンボルの視覚的記号としての自然な接続を観察するための基本的な手段を与える。我々は視覚記号の表現と接続に関連する言語学的問題のいくつかを観察した。名詞と動詞、形容語と名詞などの接続について、動詞や形容詞の表象性と含意性との間に、先の論文でも指摘されている大変難しい背反性があることが追確認された。こうした問題に対し、同様の問題をもつ、日本手話などの視覚的言語の方法を参考に、可能な限りの記号の含意性の確保するデザインのあり方と個別的ではあっても使用頻度の高い表象的なシンボルのデザインのあり方を考察した。また、知的障害児の意思表示上重要であると考えられる固有名詞や色彩を表現する新作 PIC についての議論をした。

「名詞 動詞」の構文を持つ二語文を考えると、日本語の語順が持っている基本的課題にぶつかる。それは、この語順で主格と目的格をどのように考えるかという問題である。日本語では助詞がこの問題を解決するが、知的障害児に助詞の導入は難度の高い障壁を突きつける。多語文への拡張を考えると言語学的基本問題に直面するが、これは近い将来の課題としたい。

副詞や固有名詞の表現をどうするのかについて、やはり個別的ではあっても使用頻度の高い表象的なシンボルのデザインのあり方を述べた。

さらに、機能語が少ない記号体系における時間発展の叙述の方法について、場面展開のコマ数を増やすことによって時間発展を表現する手法などを考察した。

## 引用文献

- [1] 高橋 亘, 柳内英二, 池田茉莉子, “PIC シンボル・プロセッサの実現と視覚記号の言語学”『総合福祉科学研究』, Vol. 1, 101～114 (2010).
- [2] 清水寛之, 井上智義, 北上慎司, 高橋雅延, 西崎友規子, 林 文博, 藤澤和子, 『視覚シンボルの心理学』, ブレーン出版 (2003).
- [3] 藤澤和子編著, 『視覚シンボルでコミュニケーション: 日本版 PIC 活用編』, ブレーン出版 (2001).
- [4] 藤澤和子・井上智義・清水寛之・高橋雅延 『視覚シンボルによるコミュニケーション: 日本版 PIC』, ブレーン出版 (1995).
- [5] 高橋 亘, 柳内英二, “PIC シンボルによる知的障害者のコミュニケーション支援システムの M 言語による実現”, 『Proceedings 2007 M Technology Association of Japan』, 19 ～ 23 (2007).
- [6] 高橋 亘, 柳内英二, “PIC シンボルによる知的障害者のコミュニケーション支援システム”, 『関西福祉科学大学紀要』, Vol. 11, 49 ～ 54 (2008).
- [7] 柳内英二, 高橋 亘, “PIC シンボルによる知的障害者の意思表示システムの M 言語による実現”, 『Mumps』, Vol. 24 (2008) 41 ～ 47.
- [8] 高橋 亘, 柳内英二, “PIC シンボルによる知的障害者の意思表示システム”, 『関西福祉科学大学紀要』, Vol. 12, 41 ～ 48 (2009).
- [9] 中村義行, 大石史博編著, 『障害臨床学』, ナカニシヤ出版 (2005).
- [10] 高橋 亘, 『コミュニケーション支援の情報科学』, 現代図書 (相模原, 2007, 4 月).
- [11] 征矢清作, 林明子絵, 『はっばのおうち』, 『はっばのおうち』, 福音館書店 (1989).

## Appendix

この Appendix では、プログラミング言語 MUMPS による、関数、ReCodeObj(OCODE,NCODE) のプログラムを記載する。この関数は、二種の大域変数 ^PicTree、^ConnectAmp に含まれる総てのメッセージ・シボルの ID を古い ID (OCODE) から新しい ID (NCODE) に変換し、変換した大域変数 ^ConnectAmp の個数を返す。このプログラムで扱っている大域変数 ^ConnectAmp は本文で述べたものに加えて、複数の助詞が想定される場合の添え字として、第 5 の添え字 (No) を持っているが、プログラムの基本的アルゴリズムに大差はない。

```

ReCodeObj(OCODE,NCODE) ;
New (OCODE,NCODE)
; Programmed by W. Takahasi, E. Yanagiuchi and M. Ikeda
; 02:20 PM, 29 Jun 2009
; 用法: Set Z=$$`ReCodeObj(X,Y)
Set Z=0
Set S=""
For {
    Set S=$Order(`PICTREE(S))
    Quit:S=""
    Set CHK=$Data(`PICTREE(S,OCODE))
    Do:CHK=1 recode1
}
Quit Z
;
recode1 ;
Set Categ=S
Set TDATA=$Get(`PICTREE(Categ,OCODE))
; `PICTREE の再記号化
Kill `PICTREE(Categ,OCODE)
Set `PICTREE(Categ,NCODE)=TDATA
;
; `ConnectAmp の再記号化へ
; 前方固定; Categ1=Categ, Obj1=OCODE---->NCODE
Do rank3
; 後方固定; Categ2=Categ, Obj2=OCODE---->NCODE
Do rank1
Quit
;
rank3 ;
Set S3=""
For {
    Set S3=$Order(`ConnectAmp(Categ,OCODE,S3))
    Quit:S3=""
    Do rank4
}
Quit
;
rank4 ;
Set S4=""
For {
    Set S4=$Order(`ConnectAmp(Categ,OCODE,S3,S4))
    Quit:S4=""
    Do rank5H
}
Quit
;

```

```

rank5H ;
Set S5=""
For {
  Set S5=$Order (^ConnectAmp (Categ, OCode, S3, S4, S5))
  Quit:S5=""
  Do:CHK=1 recode2H
}
Quit
;
recode2H ;
Set Categ2=S3, Obj2=S4, No=S5
Set CADATA=$Get (^ConnectAmp (Categ, OCode, Categ2, Obj2, No))
; 前方固定 ^ConnectAmp の再記号化
Kill ^ConnectAmp (Categ, OCode, Categ2, Obj2, No)
Set ^ConnectAmp (Categ, NCode, Categ2, Obj2, No)=CADATA
Set Z=Z+1
Quit
;
rank1 ;
Set S1=""
For {
  Set S1=$Order (^ConnectAmp (S1))
  Quit:S1=""
  Do rank2
}
Quit
;
rank2 ;

Set S2=""
For {
  Set S2=$Order (^ConnectAmp (S1, S2))
  Quit:S2=""
  Do rank5L
}
Quit
;
rank5L ;
Set S5=""
For {
  Set S5=$Order (^ConnectAmp (S1, S2, Categ, OCode, S5))
  Quit:S5=""
  Set CHK=$Data (^ConnectAmp (S1, S2, Categ, OCode, S5))
  Do:CHK=1 recode2L
}
Quit
;
recode2L ;
Set Categ1=S1, Obj1=S2, No=S5
Set CADATA=$Get (^ConnectAmp (Categ1, Obj1, Categ, OCode, No))
; 後方固定 ^ConnectAmp の再記号化
Kill ^ConnectAmp (Categ1, Obj1, Categ, OCode, No)
Set ^ConnectAmp (Categ1, Obj1, Categ, NCode, No)=CADATA
Set Z=Z+1
Quit
;

```

## 医療情報システムと連動したCTIシステムの開発

### Development of CTI system linked to medical information system

土屋喬義<sup>1)2)</sup>、田中千恵子<sup>1)</sup>、栢森猛<sup>1)</sup>、木村一元<sup>3)</sup>  
Takayoshi Tsuchiya<sup>1)2)</sup>, Chieko Tanaka<sup>1)</sup>, Takeshi Kayamori<sup>1)</sup>,  
Kazumoto Kimura<sup>3)</sup>

1) 土屋小児病院、2) 獨協医科大学小児科

3) 獨協医科大学病院医療情報センター

1) Tsuchiya Children's Hospital

2) Department of Pediatrics, Dokkyo Medical University

3) Center for Medical Informatics, Dokkyo Medical University

埼玉県久喜市中央1-6-7

1-6-7 Chyuuou Kuki Saitama Japan

TEL:0480-21-0766

FAX:0480-21-2230

e-mail: takayoshi@tsuchiya.or.jp

**要旨** 救急病院の電話対応は非常に困難かつ重要な仕事である。土屋小児病院では主にこの電話対応を看護師が行い、受診希望者のトリアージに大きな成果を上げている。しかしながらかかってくる電話の中には、いたずら電話、強引なセールスの勧誘なども含まれ、電話の応答に恐怖感を訴える看護師もいる。電話に応答する前に相手を識別できるようになれば、不安感を払拭し、なおかつ緊急の電話に的確に対応できると考え、院内情報システムと連携した電話自動識別装置 (CTI) を開発する事とした。CTI 運用は既に1年が経過し、夜間の電話応答率は運用開始直後の91.1%より97.2%と向上した。CTI 装置と院内医療情報システムとの連携を行ったため、受診歴があれば診断名、投薬歴、などが瞬時に参照でき電話の対応時間の短縮に繋がった。また、あらかじめ相手が判る場合が多いため電話対応への恐怖感が減少した事と救急外来へ電話をかける患者保護者の理解の深まりが好結果に結びついた。

**Abstract** The telephone support for responding to the emergency calls is important and yet is a very difficult task to carry out. Nurses usually provide this support at our institute. However the quality of such emergency calls varies to such an extent that nurses often feel fear to even answer the calls during the night. The needs for new information system, which automatically identify the caller, and provide necessary information such as medical history, are realized. A new Computer Telephony Integration (CTI) system was developed, and the system is integrated in the hospital information system. CTI system could provide not only shorter time of telephone response time, but also necessary preparation prior to patient arrival at the hospital. Nurses fear to emergency telephone call has been significantly eased. Better communication between hospital and callers (patient's parents / guardians) was also achieved.

キーワード: 電話自動識別装置 (CTI)、病院情報システム (HIS)、電話対応、トリアージ

Key word : Computer Telephony Integration (CTI), Hospital Information System (HIS), Triage

### 1. はじめに

土屋小児病院は埼玉県の北東部久喜市にある25床の小児科専門病院である。小児医療は少子化を背景に縮小を余儀なくされており、また保険財政の逼迫と小児救急、高度な医療への要望、その結果として起こる小児科医の減少と小児科を希望する看護師の減少など困難な問題に直面している[1-3]。安定した病院運営のためには、いかに職員個人の安全を図り安心して継続的に就業できる環境を作ることが極めて重要である。当院では診療の専門化、救急を軸とする24時間対応、職員能力の向上ならびに合理的な職員の配置とネットワーク化したコンピュータシステムの使用により合理的で医療の提供者側も安心できる

体制を作るよう努めている[4]。

土屋小児病院では、1991年よりM言語であるU-MUMPS上で動作する医事システムを採用した[5]。2007年より医事システムはM言語をエンジンに持つCacheベースのものへ変更している[6]。また院内情報システムは同じくM言語であるMSMとCacheで独自に開発している。当院では医事システムより派生するデータを有効活用するため殆どのアプリケーションが医事システムのデータを参照している。医事システムと院内情報システムはM言語間で使われる通信仕様のひとつであるDDP接続で通信し、医事システムよりリアルタイムに得られる情報を活用している(図1)[7-9]。

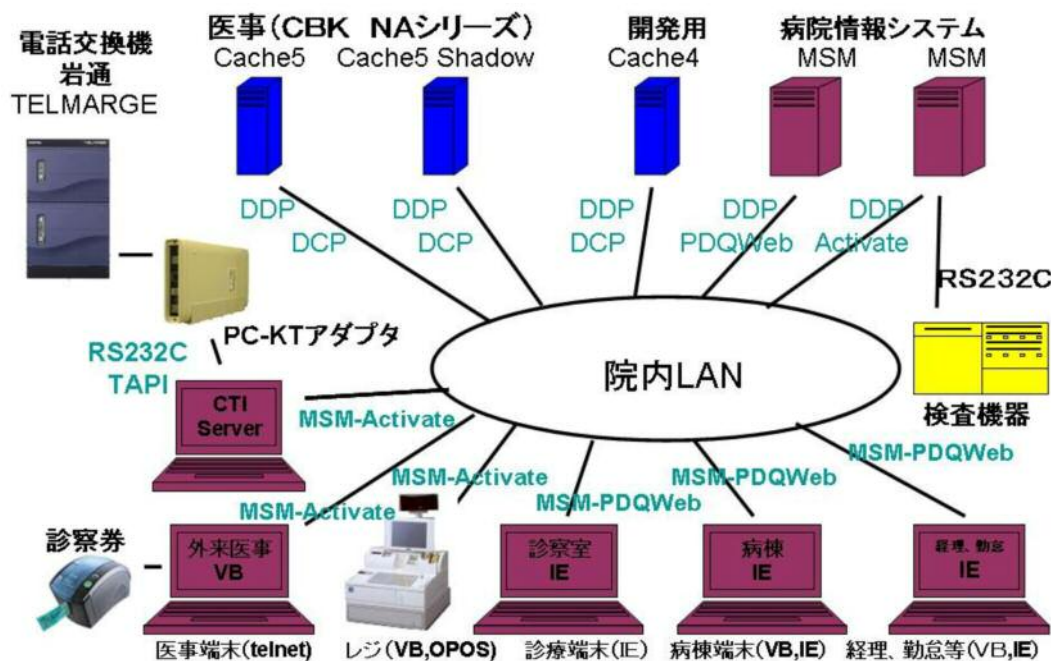


図1 院内システム構成

### 2. 対象と方法

CTI (Computer Telephony Integration) は、コンピュータと電話やファックスを連動させたシステムにより、企業のコールセンタや営業部門その他における顧客や取引先との対応で広く利用されている。顧客データベースと連携した電話情報システムにより、顧客のプロフィールや過去の対応履歴、購入履歴などを参照しながら

的確なサポートを提供することができるようになっている。こうしたCTIを含めた統合型の顧客対応システムはCRMと呼ばれることもある。このようなシステムを導入することにより宅配のピザ屋などで電話をただで名前、住所などが判り、簡単にすぐ注文ができる。医療の現場にもこのようなシステムを導入することにより電話をかけた患者に対して確実な対応がとれ、



瞬時に患者情報が参照できるため簡潔に電話対応ができると考えた。更に、かかってくる電話の中には、いたずら電話、強引なセールスの勧誘なども含まれ、電話の応答に恐怖感を訴える看護師もおり、診療の質のみでなく病院職員の安全性にも寄与すると期待した（図2）。

- 患者満足度向上(CS)か？組織防衛(ES)か？
  - 電話がかかってきた時相手の名前が判るとサービスが向上できる。  
ヒザ屋、「はい〇〇屋です。XX様、ご住所は△△ですね。」
  - 病院でも、昼間はもちろん便利
- 病院では夜間、看護師(女性)が電話に対応
  - いたずら電話、迷惑電話
  - 業務に全く関係ない業者の売込み(投機商品、不動産)
    - 当直の先生はいますか(XX商事です)。
    - 無言電話
- 電話に出るのが怖い、しかし重要な電話も
  - かかりつけ患者の急変
  - 飛込みの救急
  - 職員の連絡

図2 CTIシステムの必要性

病院では夜間の電話トリアージは看護師が行うため診察室、受付、処置室など電話を受け場所が定まらない。このため電話の近くの適当な端末より適宜患者情報を参照できる必要がある。この要求を実現するため各回線に入った電話の情報を院内全ての端末で参照できる必要があった。このため CTI システムはクライアント・サーバシステムとし、院内のどの場所からでも電話情報を見る事を可能とすることを考えた。同時に CTI 情報を多数の端末がアクセスしても負荷がかからず、ライセンスに問題を起こさない構成をとするため MSM-PDQWeb (Web Link と等価) を使用して動的な Web ページを作り院内端末の Web ブラウザより参照する形態を取った。

### CTI サーバ

当院は電話交換機に岩通 TELMARGE を採用している。外線は12回線で約150回線の内線が接続している。交換機自体はTAPIに対応したインタフェースを持たないため、電話番号情報を取り出すためには交換機の前にあるターミナルアダプタか若しくは交換機と、ビジネスフォンの間に CTI アダプタ (PC-KTアダプタ) を取付けて情報を取り出す方法が考えられた。PC-KTアダプ

タを使用することによりかかってきた外線を内線へ転送するなど機能の付加が可能となるため今回はPC-KTアダプタを使用することとした。岩通のPC-KTアダプタはMicrosoft Telephony API (TAPI Ver. 2.0) に準拠している。TAPI コマンドをコントロールするため TAPI コントロールミドルウェアとして free tool の Julmar Technology の ATAP.NET を採用した。開発言語は Microsoft .NET 2005 (VB) を使用し、MSM-Activate を介して M サーバのグローバルに電話イベントを書き込む事とした。

今回開発した CTI システムで、監視を行う回線は代表回線 4 回線、救急用回線 2 回線の計 6 回線とした。PC-KT アダプタより発生する TAPI イベントは電話が鳴動すると発生するため、全ての回線に対して PC・KT アダプタ取り付けこれを監視するのは不合理と考え、6 回線全ての電話が鳴動する電話を設定しこのイベントを監視することとした。どの回線が鳴動したかの確認は TAPI イベントでは判定できないため 150 ミリ秒毎に電話のボタンランプを観察し、鳴動した電話の回線を同定した（図3）。



図3 CTI 装置概観

通話イベントは M サーバの ^CTI グローバル (図4) に記録し、同時に ^CTI ACT グローバル (図5) の第 1 ノードに各回線の最新の状態を、第 2 ノードにその状態が起きた時刻を記録した (図6)。

### CTI サーバー ^CTI

```

^CTI("2009/08/08","11:17:35","11:17:58",1)="IN;11:17:43;Unknown"
^CTI("2009/08/08","11:29:34","11:35:54",1)="IN;11:29:42;080XX04XX24"
^CTI("2009/08/08","11:37:27","11:38:06",6)="OUT;00:00:00;Unknown"
^CTI("2009/08/08","11:45:22","11:47:02",1)="IN;11:45:31;090XX33XX73"
^CTI("2009/08/08","11:59:50","12:00:54",1)="IN;11:59:57;066XX43XX2"
^CTI("2009/08/08","12:01:44","12:06:29",1)="IN;12:01:50;090XX113XX29"
^CTI("2009/08/08","12:12:35","12:17:25",1)="IN;12:12:42;048XX78XX3"
^CTI("2009/08/08","12:24:10","12:25:28",5)="OUT;00:00:00;Unknown"
^CTI("2009/08/08","12:26:22","12:26:58",5)="OUT;00:00:00;Unknown"
^CTI("2009/08/08","12:27:55","12:30:58",5)="OUT;00:00:00;Unknown"
^CTI("2009/08/08","12:31:19","12:36:07",5)="OUT;00:00:00;Unknown"
    
```

図4 ^CTI グローバル

### CTI サーバー ^CTIACT

```

^CTIACT(1)="Connected,Unknown"
^CTIACT(1,1)="2009/08/08,15:56:28"
^CTIACT(2)="Disconnected,Unknown"
^CTIACT(2,1)="2009/08/08,10:28:41"
^CTIACT(3)="Disconnected;048XX37XX4"
^CTIACT(3,1)="2009/08/07,18:22:09"
^CTIACT(4)="Disconnected,Unknown"
^CTIACT(4,1)="2009/07/31,10:32:46"
^CTIACT(5)="Disconnected,Unknown"
^CTIACT(5,1)="2009/08/07,18:22:09"
^CTIACT(6)="Disconnected,Unknown"
^CTIACT(6,1)="2009/08/07,23:48:27"
^CTIACT(7)="Disconnected,Unknown"
^CTIACT(7,1)="2009/08/07,23:48:27"
^CTIACT(8)="Disconnected;090XX48XX05"
^CTIACT(8,1)="2009/08/07,23:48:27"
^CTIACT(9)="Disconnected,Unknown"
^CTIACT(9,1)="2009/05/30,21:06:41"
^CTIACT(10)="Disconnected,Unknown"
^CTIACT(10,1)="2009/05/30,21:06:41"
^CTIACT(11)="Disconnected,Unknown"
^CTIACT(11,1)="2009/05/30,21:06:41"
^CTIACT(12)="Disconnected,Unknown"
^CTIACT(12,1)="2009/08/08,15:56:28"
^CTIACT(13)="Disconnected,Unknown"
^CTIACT(13,1)="2009/08/08,15:56:28"
^CTIACT(14)="Disconnected,Unknown"
^CTIACT(14,1)="2009/08/08,15:56:28"
^CTIACT(15)="Disconnected,Unknown"
^CTIACT(15,1)="2009/08/08,15:56:28"
^CTIACT(16)="Disconnected,Unknown"
^CTIACT(16,1)="2009/08/08,15:56:28"
^CTIACT(17)="Disconnected,Unknown"
^CTIACT(17,1)="2009/08/08,15:56:28"
^CTIACT(18)="Disconnected,Unknown"
^CTIACT(18,1)="2009/08/08,15:56:28"
^CTIACT(19)="Disconnected,Unknown"
^CTIACT(19,1)="2009/08/08,15:56:28"
^CTIACT(20)="Disconnected,Unknown"
^CTIACT(20,1)="2009/08/05,20:46:31"
^CTIACT(21)="Disconnected,Unknown"
^CTIACT(21,1)="2009/08/06,09:26:14"
^CTIACT(22)="Disconnected,Unknown"
^CTIACT(22,1)="2009/08/08,15:05:44"
^CTIACT(23)="Disconnected,Unknown"
^CTIACT(23,1)="2009/08/08,15:05:44"
^CTIACT(24)="Disconnected,Unknown"
^CTIACT(24,1)="2009/08/08,15:05:44"
    
```

図5 ^CTIACT グローバル

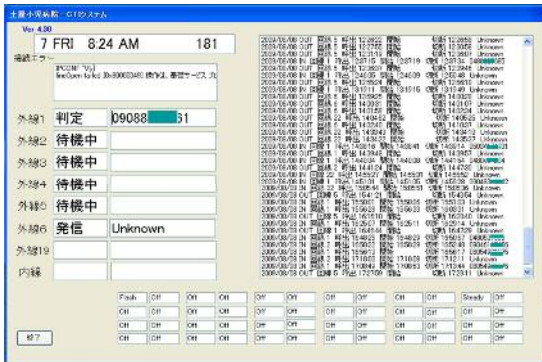


図6 CTIサーバ監視画面

### CTIクライアント

CTIクライアントはMSM-PDQWebを使用して院内端末のWebブラウザより^CTIACTグローバルを参照し、最新の電話回線状態を表示する。ブラウザは5秒毎に更新するように設定し常に最新の情報を表示するようにした(図7)。また、医事システムにはカルテ情報の中に電話番号が1つだけしか記録されないため、別にグローバルを用意し、同一患者に複数の電話番号の登録を可能にした。



図7 CTIクライアント動作画面

上段は過去10件の受診履歴、下段は現在コール中、もしくは接続中の電話情報。バックグラウンドが黒く(実際は赤)表示されているのは48時間以内の受診歴のある患者。

電話番号よりID番号が見つかった場合は医事システムに問い合わせを行い、かかって来た電話に属する家族の名前、それぞれの最終来院日と直近の受診時の診断名を表示する。新規の電話番号登録画面は、図7のCTIクライアント画面の登録ボタンを押す事により簡単に携帯電話などの医事システムに登録の無い電話番号をその場でID番号を入力する事により登録が行えるようにしている。また患者IDの無い場合は直接名前、もしくは組織名を記録している(図8)。



図8 患者番ID録画面

患者IDを入力すると下段に患者情報が表示され、確認後登録ボタンを押す。



夜間の電話対応で、どのような内容で相談されたか、どのような指示を与えたのかを把握している事は、診療の質と安全性を高めるのに重要な事である。このため CTI クライアントには電話での相談内容、指示内容を記録する機能を設けている。電話番号には複数の家族が登録されている場合があるため、電話相談の内容を記録するために該当患者名を選択して相談内容の登録を行っている。記録にはテンプレートを用意しこれを選ぶ事により簡単に作成できるようにしてある(図9)。

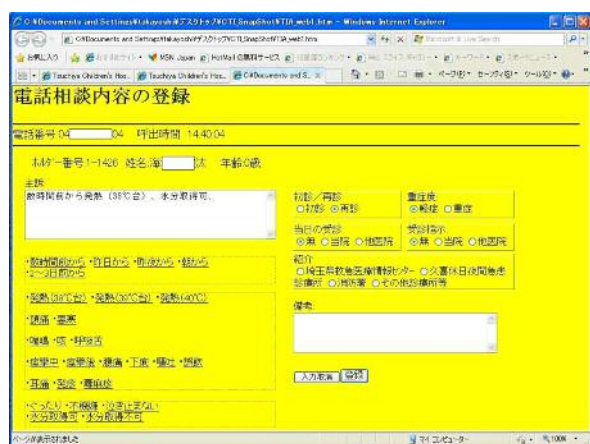


図9 電話相談内容登録画面

その他の機能

電話通話記録ログ

通話記録のログ参照画面を示す(図10)。ログには電話の鳴動開始時間、通話開始時間、切断時間、電話番号、患者ID、氏名を表示させている。通話終了後でも電話相談内容を後から記録するために通話内容の登録、修正ボタン備えた。また誤って登録された登録情報を修正するためのボタンも用意した。日付よりの表示だけでなく患者IDや電話番号をキーにしての表示も行えるようにした。

電話相談記録

過去の電話相談の内容を日、月単位で表示(図11)。電話相談した時間、名前、性別、年齢、相談内容、指示内容、緊急度などを表示させている。



図10 電話記録のログ



図11 電話相談内容の一覧

3. 結果

現在、MSM-PDQWebを介したWebクライアントでCTIクライアントを3台から5台で運用している。5秒毎の頻度で各クライアントがサーバにアクセスしているが、参照時間が極めて短時間で済む事によりMSM-PDQWebとMSMサーバ間の接続数はデフォルトの3で十分に機能している。

運用開始後既に1年が経過、2008年10月より電話の通知設定を呼びかけた結果、半数近くあった非通知電話は2009年7月には38%まで減少、さらに夜間の非通知電話に対しては自動で通知設定にしてから再度電話をかけ直すようにメッセージを流した結果、2009年12月には6.1%まで減少し、電話マナーも向上した。小児専門病院として職員の小児医療と救急医療に対する熱い情熱により当院の夜間電話応答率は運用開始直

後でも91.1%と非常に高いものであったが、2009年7月97.2%と向上している。

本システムによる2009年4月の通話時間の解析結果では、日勤時間帯（8時から17時台まで）2104本の平均通話時間110.6秒に比べ夜間（18時～7時台まで）667本の平均通話時間は180.7秒と1.63倍も長いことが判った。また、深夜2時台に時間当たりの電話件数は9件と最低を記録したが、平均通話時間は346秒と最大値を記録した(図12, 13)。

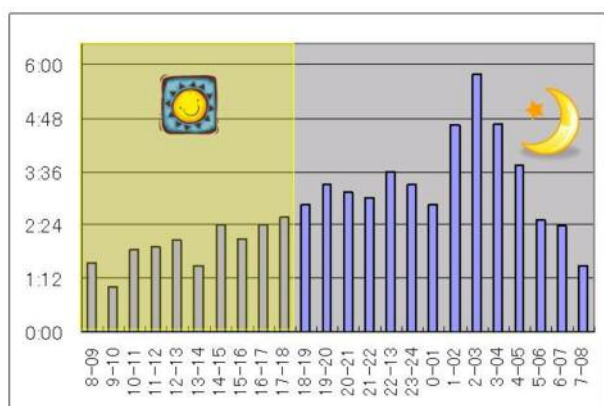


図12 1通話当たり時間帯別平均通話時間 (2009年4月)

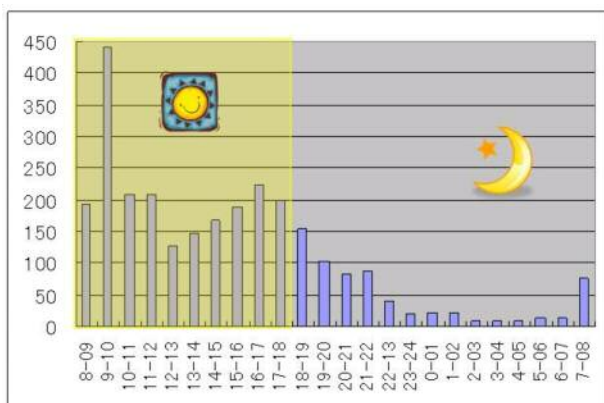


図13 時間帯別通話回数 (2009年4月)

#### 4. 考案

医療崩壊が叫ばれている中で、特に救急医療を含む小児二次医療のおかれている環境は非常に厳しい。二次医療を行う小児科運営の困難さは単に保険収入の少なさだけの問題ではなく、訴訟のリスク、患者の過大な期待など収入に見合わない過酷な勤務内容にある[1-4]。この様な

環境の中では小児科を志そうと考えこの世界に入った医師のみでなく、看護師、事務員などのパラメディカルですら小児医療にたいする情熱を失いこの過酷な現場から去って行く。中でも、夜間起こる医療従事者に対する暴力、暴言、言葉の齟齬による被害は深刻であり、当院の職員の中にはこれらを苦にして離職たり、救急や電話対応の現場を離れざる得ない者も出ていた。積極的に救急医療に携わる職員にとっても心理的ストレスは大きなものであり、小児二次医療を行う現場での大きな障害となっている。我々はこの様な状況を座視せず少しでも改善するため、努力している。

夜間当院に救急で受診する患者は電話で連絡をした上で受診している。このため電話は救急で受診する際の最も重要な手段である。しかしながらCTIシステム導入前は、病院の業務と無関係なセールス、いたずら電話、無言電話、性格障害者からの電話もあり、電話に出るまで内容を判別できない。このため事前に相手が判るCTIシステムの導入は、職員の離脱を防ぎ救急医療体制を維持するための組織防衛のツールとして必須のものとなると考えた。実際、CTIシステムを導入している事により、電話の匿名性がなくなり、患者保護者の電話マナーも向上してきていると感じている。また電話の相手に対して的確な対応が出来るためにより積極的な対応がとれるため（仕事に対する満足度、ESが向上）電話の応答率向上など、職員の救急に対する取り組みが積極的となっている。さらに、CTIと院内医療情報システムとの連携を行ったため、受診歴があれば診断名、投薬歴、などが瞬時に参照でき、スムーズな電話対応に繋がっていると考えている。

深夜2時から5時の間の救急外来は救急患者数の少ない比較的落ち着いた時間帯であるが、今回の電話情報解析の結果では最も電話対応の大変な時間帯であることが判り、今後、救急外来病院運営の検討課題と考えられた。

主に夜間の電話に回答に対する不安感を払拭



するために開発したシステムであったが、CTIシステムは電話がかかってきた時相手の名前、属性が判たためサービスと業務の効率化が期待できる。夜間の電話対応で、どの様な内容で相談されたか、どの様な指示を与えたのかを把握している事は、日常の診療の質と安全性を高めるのに重要な事である。このため現在当院で開発を進めている患者情報表示端末では患者診察時に電話相談内容を表示する機能を付加している。また予防接種や診療予約などもCTI画面より直接行えるようにしており患者満足度（CS）の向上に役立っている。

## 参考文献

- 1) 土屋喬義：小児科独自の診療報酬体系を：月刊 保険診療 vol.50 No.2 Ser.No.1274 1995.2
- 2) 読売新聞医療情報部：こどもの医療が危ない：中央公論新社，2002.5.25
- 3) 土屋喬義，田中千恵子：土屋小児病院のご紹介－小児医療の危機の中で医療法人立土屋小児病院の取り組み－：第33回Mテクノロジー学会大会論文集 17-22, 2006
- 4) 土屋喬義，加来裕康：小児科病院を継いだ院長－“不採算”を克服するために選んだ道－：日経ヘルスケア 1993.12

- 5) 土屋喬義，土屋恭子，木村一元：個人病院に於ける病院情報システムの活用－レセプト専
- 6) 土屋喬義，田中千恵子，木村一元：MSM,U-MUMPSよりCacheベースシステムへの移行経験，：第33回Mテクノロジー学会
- 7) 土屋喬義，木村一元：エンドユーザー用の言語としてのM当院での活用方法：第23回日本Mテクノロジー学会大会 大会論文集，48-51, 1996
- 8) 土屋喬義，田中千恵子，駒田智彦，木村一元：土屋小児病院の院内診療医療システムⅡ：第28回日本Mテクノロジー学会論文集，5-6, 2001
- 9) 土屋喬義，田中千恵子，木村一元：医事システムと連動したPOSレジの開発：MUMPS Vol.24, 21-25, 2008

## 参考URL

1. Microsoft Telephony Overview, [http://msdn.microsoft.com/en-us/library/ms733433\(VS.85\).aspx](http://msdn.microsoft.com/en-us/library/ms733433(VS.85).aspx), MSDN
2. Telephony Application Programming Interface Version 2.2, [http://msdn.microsoft.com/en-us/library/ms737220\(VS.85\).aspx](http://msdn.microsoft.com/en-us/library/ms737220(VS.85).aspx), MSDN
3. Welcome to JulMar Technology, <http://www.julmar.com>

# CSP での医師国家試験問題検索システムの構築

## Retrieval System of Questions on the National Examination for Medical Practitioners using CSP

木村一元<sup>1)</sup>、五十嵐吉彦<sup>2)</sup>

Kazumoto Kimura<sup>1)</sup>, Yoshihiko Igarashi<sup>2)</sup>

1) 獨協医科大学病院医療情報センター、2) 土屋小児病院

1) Center for Medical Informatics, Dokkyo Medical University Hospital

2) Tsuchiya Children's Hospital

**要旨** M の柔軟な文字処理機能と Caché の持つ CSP 機能を用いて医師国家試験問題の検索システムを構築した。医師国家試験問題の中には多くの画像があり、画像の表示はとても重要である。この画像の表示には、HTML が扱える CSP を使い、利用者からはブラウザの使用にて目的の問題の検索と画像を含む結果の表示が行える様にした。検索のキーワードは最大 4 つまで指定でき、検索された文字列中のキーワードは色分けを行い、どのキーワードが文中のどの位置で検索されたのかが判る様にした。キーワードを 4 つ指定した場合の検索・表示時間は、1 秒以下であり、また、ファイルサイズを 10 倍とした場合でも検索・表示時間は 4 秒で、利用に十分な検索・表示時間であった。

**Abstract** A novel search engine is constructed to search specific questions and information contained in the data of national examinations for medical practitioners using flexible function of processing characters by M system and CSP of Caché. There are a lot of image data in the national examinations. The display of image data is very important for recent examinations. Since image data were displayed by using 'img' tag of HTML in CSP, users can find questions and answers containing keywords and figures by user's browser. Key words of the retrieval were specified up to four words. Each key word was shown with different colors as they can be found easily in the sentences. When four key words were specified in a search, total time of information retrieval and display was within one second or less from the data containing eight years examinations. When file size was increased 10 times, retrieval and display time was 4 seconds. In conclusion, the system enables searcher to search or obtain desired information promptly from the national examinations data.

キーワード 検索システム、医師国家試験問題、Caché Server Pages (CSP)

Keyword Retrieval System, the national exam for medical practitioners question, Caché Server Pages (CSP)

### 1. はじめに

我々はこれまで、M の柔軟な文字処理機能とデータベース機能を用い種々の検索システムを開発して来た（細胞診・組織診検索システム、図書館所蔵雑誌検索システム[1]、業績集検索シ

ステム)。M の文字処理機能は、これまでの文献検索の規定された範囲の前方一致や後方一致機能と異なり、全文を対象としたものである。従来のコンピュータ化された医師国家試験問題システムは、年度別や分野別に問題が表示されるだけで、設問中の文字列の検索が出来なかった。そこで今回、医師国家試験問題を検索対象とし

た。しかし、医師国家試験問題は、これまで我々が開発して来たシステムとは異なり表示対象に文字列だけでなく画像が含まれておりその表示を考慮しなければならなかった。M そのままでは画像の表示が容易で無いので Caché の持つ CSP (Caché Server Pages) [2] 機能を用いて検索システムを実現する事とした。

## 2. 対象と方法

開発は、シングルバージョンの Caché (v5.1) を用い、検索の基本システムは M で作成し、検索キーの入力や結果の表示の部分は CSP を用い、利用者からはブラウザの使用にて目的の検索と結果の表示が行える様にした。

検索された文字列中のキーワードは色分けを行い、どのキーワードがどの位置で検索されたのかが判る様にした。また、JPEG 画像は、その縦横サイズを確認し、一定の表示サイズに納まる様にした。

検索対象となるデータは第 95 回から第 102 回までの国家試験問題で、そのファイルは、全体で 4,188 問題で文字列の大きさは強調文字、上付き、下付き文字、表組のための html タグを含め 2,154KB、JPEG 画像の大きさは 125.4MB である(図 1)。

- **検索対象: 第95回から第102回までの  
国家試験問題**
- **サイズ : 4,188問題、2,154KB  
約8万行**  
(強調文字, 上付き, 下付き文字, 表組の html タグを含む)
- **JPG画像: 1,133枚、120.6MB**
- **開発: シングルバージョンの Caché (v5.1)**

図 1 検索対象データ

なお、ルーチンはキーワード入力部分 (Keyword04.csp) と 検索 実行 表示 部分 (Kensaku04.csp) の 2 つの CSP から成る。

このシステムの学内への公開は、サーバマシン HPxw6200 (Xeon CPU3.4GHz, 3GBRAM) にて Caché (V5.1) と Apache (V2.0.55) を稼働させて

行った。

### ルーチンの機能

検索実行ルーチンは、4つまでのキーワードが指定でき、さらに検索されたキーワードは色分けができる。キーワードの指定きは、ワイルドカードでの指定が可能で、A%B では文字列 A と B の間に任意の 1 文字であるものが検索され、A%%B では文字列 A と B の間に任意の 2 文字であるものが検索される。また、文字列 A と B の間が何文字でも良い場合には、A\*B で指定できる。ただし、この場合、句読点があればその検索は中止される。なお、指定されたキーワードは、And 条件としてそれらの文字をすべて含む問題が検索される。

設問内の画像は、その大きさを判定し、表示する様にした。

### 試験問題のファイル構造

試験問題のファイル構造は、「問題番号」、「設問」、「解答群と正答」、「コメントと画像ファイル名」の 4 つの項目からなり、それぞれ、「nnAxxxx」、「Question:」、「Answer:」、「Point:」の識別子で区別される。また、これらの識別子は大きく「問題」の識別子で区別されている。

問題番号は nnAxxxx の形式で表現され、nn は医師国家試験の回数、A のアルファベットは A ~ I までの問題の種類を示し、xxx はその中での問題番号を示す。設問部分は、「Question:」から始まり、複数行の設問が記載されている。解答群と正答部分は、「Answer:」から始まり、これも複数行の記載があり、解答群の選択肢の前には、その選択肢が正解か否かの O X の正答が記載されている。コメントと画像ファイル名の部分は、「Point:」で始まり、コメントや画像表示があれば記載される。この部分も複数行となり、次の問題の識別子「問題」の前で終り、この「問題」の識別子の次の行から新しい問題が始まる。すべての問題の終わり、すなわち、ファイルの最後は、識別子「:E:」で終る(図 2)。

問題  
 nnAxxx 問題番号  
 Question:  
 設問 複数行  
 Answer:  
 解答群と正答 複数行  
 Point:  
 コメントと画像ファイル名  
 問題  
 :  
 :E:

図2 試験問題ファイルの構造

### 試験問題ファイル例

図3に実際の試験問題ファイルの記載例を示す。この例は101回の医師国家試験C問題 33番目のものである。「Question:」の次の行から設問が書かれている。この例では2行であるが、一般にこの箇所には複数行書かれる事が多い。また、検査データを表形式で表記する場合にはHTMLにて記載ができる。識別子「Answer:」の次の行には、設問に対する解答とその選択肢の正解、不正解が○×にて示されている。識別子「Point:」には設問に対するコメント（例えば医師国家試験出題基準[3]のどの基準に分類されているかなど）と設問に画像が示されているには画像ファイル名が記載される。

問題  
 101C033  
 Question:  
 頭部単純CT(別冊No.1)を別に示す。  
 診断はどれか。  
 Answer:  
 X= 脳出血  
 X= 脳梗塞  
 X= くも膜下出血  
 X= 急性硬膜外血腫  
 O= 急性硬膜下血腫  
 Point: 必-08- L-01,エックス線画像(単純、CT、造影),頭部単純CTによる画像診断  
 101C033A.JPG

図3 試験問題ファイルの記載例

### ルーチンの流れ

キーワード入力部と検索実行表示部のルーチンの流れは次のようになる(図4)。キーワードの入力が終わると試験問題ファイルの読み込みが始まり、各問題毎に入力キーワードとの一致が調べられ、入力されたキーワードを全て含む問題

があれば、次に画像ファイルの有無をチェックし、画像があれば画像サイズの確認を行い、問題と画像の表示を行う。この操作をファイルの最後まで繰り返す。

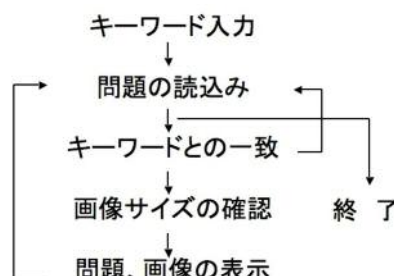


図4 ルーチンの流れ

### ルーチンの実際

検索用キーワード入力ルーチン(Keyword04.csp)のHTML入力フォームを図5に示す。HTMLのACTION機能を用い、キーワードが入力された後にOKボタンが押されると検索実行表示ルーチン呼び出す。

```

<HTML>
<HEAD><TITLE>Keyword04.csp</TITLE></HEAD>
<BODY>
<FORM ACTION=Kensaku04.csp>

Input Key word 1 <INPUT type="TEXT" name="key1" size="41"><BR>
Input Key word 2 <INPUT type="TEXT" name="key2" size="41"><BR>
Input Key word 3 <INPUT type="TEXT" name="key3" size="41"><BR>
Input Key word 4 <INPUT type="TEXT" name="key4" size="41"><BR>

<INPUT TYPE=SUBMIT VALUE="OK"></FORM>
</BODY>
</HTML>

```

図5 キーワード入力ルーチン

検索実行表示ルーチン(Kensaku04.csp)は、

```

<HTML>
<HEAD><TITLE>Kensaku04.csp</TITLE></HEAD>
<BODY>
<SCRIPT LANGUAGE="CACHE" RUNAT="SERVER">
set key(1)=%request.Get("key1")
:
set key(4)=%request.Get("key4")
:SearchDisp
open file
quit
do SearchDisp
:PhotoDisp
close file
:
set %session.EndSession=1
quit
goto END
:
:
:END
</SCRIPT>
</BODY>
</HTML>

```

図6 検索表示ルーチン



キーワードの受取り部分、国家試験問題ファイルの検索表示部分とセッションの終了を示す部分からなる。また、検索表示部分には、検索と画像データの読み込み、表示のルーチンが含まれる(図6)。

### 画像データの読み込みと表示

画像を表示するためには、その画像のサイズ(ピクセル数、分解能)を確認する必要がある。このために画像ファイルを open fn: ("RK¥RAW¥")にてバイナリモードで開き、データを用 use fn read \*datにて1バイト読みを行い、画像ファイル内にバイナリーにて書かれている解像度単位と横解像度、縦解像度を読み取る事とした。これらの情報は、JPEG File Interchange Format (JFIF) [4-6]にて規定され(図7)、この情報は Start of Image segment (SOI)に、また、縦横の画像サイズは Start Of Frame type n segment (SOFn)に書かれている。SOI はファイルのトップから始まっているが、SOFn の途中には DHT や DRI の情報が含まれており、バージョンにより異なるのでこれらの内容を読み飛ばした後に FFC2 から始まる箇所から実際の画像サイズを調べる必要がある。

- JPEG File Interchange Format (JFIF)
- SOI (Start Of Image segment)
- SOFn (Start Of Frame type n segment)
- FFC0, C2
- DHT (Define Huffman Table) FFC4
- DRI (Define Restart Interval) FFDD

図7 JPEG ファイル交換フォーマット

SOI には、7バイトから 10 バイト目に JFIF の識別子、12-13 バイトにバージョン番号、14 バイト目には解像度単位(1: dpi, 2: dpcm)、15-16 バイト目に横解像度、17-18 バイトに縦解像度が記されている(図8)。また、SOFn の FFC0 から4-5 バイト目に横画像サイズ、6-7 バイト目に縦画像サイズが記されている(図9)。

### SOI (Start Of Image segment)

SOI	App0	L	JFIF 識別子						U	Xd	Yd	サムネイル情報			
			字	号	1	号	号	号				Xt	Yt	ピクセル(2d)	
2	2	2	1	1	1	1	1	1	2	2	1	1	1	RGB	-

名称	サイズ	内容
La	2	セグメント長
Id	5	ASCII文字で"JFIF"と文字列編
Ver	2バイト	JFIFのバージョン
U	1バイト	解像度単位(単位なし、アスペクト比を表す 1: dpi 2: dpcm)
Xd	2バイト	横解像度
Yd	2バイト	縦解像度
Xt	2バイト	サムネイル横サイズ
Yt	2バイト	サムネイル縦サイズ
RGBn	2バイト	1ピクセルのRGBデータがXt x Yt 倍記録されます。サムネイルが記録されない場合、Xt、Yt共に0になります。

図8 SOI に記載の内容[5]

### SOFn (Start Of Frame type n segment)

SOPn	L	P	Y	X	Nt	色コンポーネント情報		
						Ct	抽	調
2	2	1	2	2	1	1	1	1

名称	サイズ	内容
Lf	2バイト	セグメント長
P	1バイト	成分数
Y	2バイト	画像縦サイズ
X	2バイト	画像横サイズ
Nf	1バイト	成分数
以下の値は成分番に依ります		
Cn	1バイト	成分ID
Hn	4ビット(上位)	水平サンプリング値
Vn	4ビット(下位)	垂直サンプリング値
Tan	1バイト	対応量子化テーブル番号

図9 SOFn に記載の内容[5]

## 3. 結果

キーワード入力画面と検索結果

図10に実際のキーワード入力画面を示す。この画面では、key word1に「肺」、key word2に「胃」、key word3に「腸」、key word4に「炎」の4つのキーワードが入力されている。

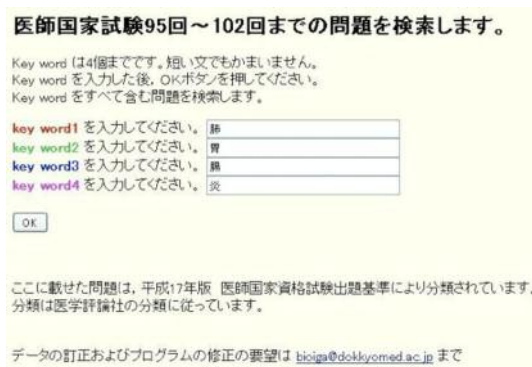


図10 キーワード入力画面

図11は上記条件での検索結果である。第95回から第102回までの試験問題でこの4つのキーワードを持つ問題は1件で第95回のE問題の23番目のものであった。key word1の「肺」



図 11 検索結果画面

の文字は赤で表示され、key word2 の「胃」は緑、key word3 の「腸」は青、key word4 の「炎」はピンクでそれぞれ色分けされて表示されている。

図 12 は、設問中に表組がある問題の表示例である。表は予め問題ファイル内にて HTML の table タグによって記載している。

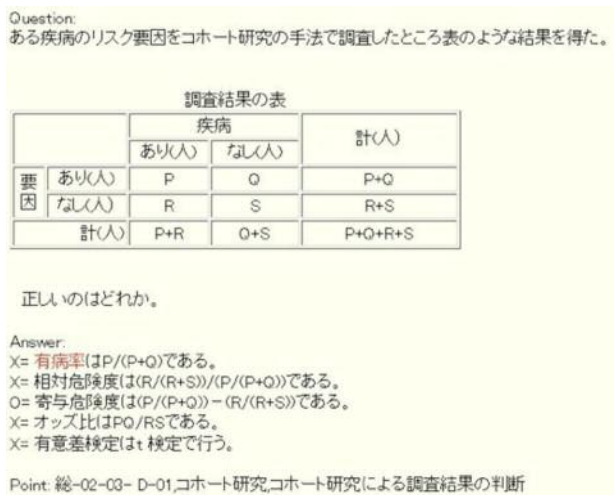


図 12 表組がある問題の表示例

図 13 は画像の表示例である。画像ファイル名は、試験問題の分類様式（国家試験回数、問題種別、問題番号）に画像の数(A, B, C, ...) で示される。画像の表示は、HTML の img タグを用い、得られた画像サイズ、分解能を基に width と height を定め行った。

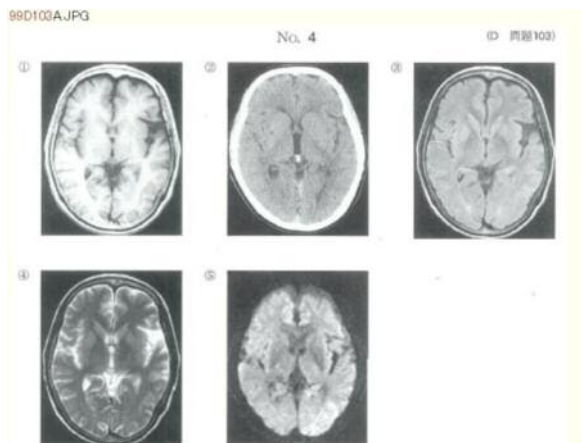


図 13 画像の表示例

#### 検索表示時間

キーワードを4つ指定した場合の検索・表示時間は、1秒以下であった。但し、表示される画像数により若干異なった。また、ファイルサイズを10倍とした場合での検索・表示時間は4秒であった。

#### 文字列の検索について

医師国家試験の場合、二文字の漢字は、間に1つのスペースを入れて入力され、表示されている。キーワードに二文字の漢字を指定した場合でも、文字間にスペースのある文字列の検索も可能であった。例) 気胸 酸素 等

また、指定された4つまでのキーワードは、それぞれ色分け表示され、設問の何処に検索キーワードがあるのかが示される。

#### 画像の検索について

この検索には、画像に関するキーワード、例えば、眼底、脳波やCTと言った語句と画像ファイルに付けられた拡張子JPGを入力するで行える。

#### 4. 考案

検索表示時間は、キーワードを4つ指定した場合で1秒以下と早く、検索時間として支障は無いと考える。また、ファイルサイズを10倍とした場合でも4秒であり、今後、問題が追加さ

れた場合としても利用に十分な時間であると考えられる。

指定されたキーワードの色分けは、設問の何処に検索キーワードがあるのかが一目瞭然で、過去出題の類似問題の把握が容易となり、検索システムの利用勝手が大幅に上がった。

試験問題にどのような画像が提示されているかを知る事は重要である。これには、画像に関連する語句と画像ファイルの拡張子 JPG をキーワードとして入力する事により行ったが、CT、MRI、眼底と言った語句の使用により、かなり絞られて形で同種の画像が過去問題として出題されている事が把握できた。

今回、試験問題のシーケンシャルファイルのファイル構造はXMLでのタグ形式(図14左)による記載では無く、ファイル内容の読み取りが容易かつ編集が行い易い形式を採用した(図14右)。これにより、試験問題の入力ミスへの対応がスムーズであった。

#### 各問題を XML にする:

<Ask>	問題
<Number: >	nnAxxx 問題番号
:	Question:
</Number: >	設問 複数行
<Question: >	Answer:
:	回答群と解答 複数行
</Question: >	Point:
<Answer: >	コメントと画像ファイル名
:	問題
</Answer: >	:
<Point: >	E:
:	
</Point: >	
</Ask>	

図 14 医師国家試験保存ファイルの構造

試験問題をシーケンシャルファイルに保存し、グローバルデータとして保存しなかった理由は、キーワードとして何が指定されるか不明で常に全文検索となるためグローバル化のメリットが無く、さらに、試験問題の修正時に一般のエディタで無く専用のツールが必要となり、データの修正が複雑となるためである。

今回の検索機能においてキーワード間に Or や Not 機能を持たせず、And 機能だけとしたのは、多数ある試験問題の中からの該当の問題の絞り込みを重要視したためである。

また、検索されたデータのページ毎での分割表示機能であるが、今回の場合、特に必要で無く、ブラウザのスクロール機能で充分に対応できた。

このシステムでは、MとWeb表示のためにCSPを用いた。今回のシステムを構築するに当たっては、ASP (Active Server Pages)としてCSP以外にJava、Perl、Rubyなどが候補に上がった。特にRuby [7]は、オブジェクト指向言語であり、文字列処理機能とWebアプリケーションとの親和性の良い事が知られており、Rubyだけでシステムの構築も考えられた。しかしながら、CSP内で透過的にMが利用でき、Mの習得の容易さと\$Pieceに代表される豊富な文字列処理機能が利用できる。また、Mとの組み合わせでWebLink [2, 8]の使用も考えられたが、プログラムの開発の行き安さでCSPを採用する事とした。

#### 課 題

検索システムは出来上がったが、今後、新しデータの追加入力(問題、画像、解答、出題基準)をどのように維持して行くのかが大きな課題である。

他の検索ツールでは、漢字、ひらかな、カタカナなど記載方法が異なっても検索対象とするものもあるが、これはグローバルに同義語の辞書を持たせ対応することが可能であろう(なお、医師国家試験では、使用される言葉は概念によって使い分けが行われている。例えば、病名の場合は癌、健診の場合はがんと記載される)。

医学用語では、肉月やヤマイダレの複雑な漢字が多く使われ、国家試験内でも記載されている。これらの漢字は、機種依存性のある新JISで追加された漢字が多く、どの機種どのブラウザをターゲットして良いか課題であったが、今後、データファイルの漢字をShift JISで無く、UTF-8として対応して行けば解決されると思う。また、同様にドイツ語、フランス語の特殊な文字の表示や単位のリットル(l)についても

課題であったが、dl を dL と記載せず、UTF-8 を用い dℓ とする事で解決できる。さらに○アイウの記号もこれまで(ア)(イ)(ウ)と表記して課題となっていたが、UTF-8 の利用にて解決の目処が立った(図 15)。

**新JIS で追加された漢字(機種依存)**  
**胆嚢、胆嚢 耳癬(ジセツ) 瘡痒(ソウヨウ)**  
**ドイツ語、フランス語の表示 ö é etc.**  
**単位 dl → dL or d/ dℓ UTF-8**  
**アイウ**

図 15 機種依存性のある文字

Mシステムでは、メモリーに置かれたデータよりもグローバルデータとして扱った方がアクセス速度が速くなる傾向がある。今回の場合、問題の入力ミスにすぐに修正できる様に試験問題をシーケンシャルファイルとしたが、今後、新しい試験問題が追加され、より速い検索速度が要求される場合には、試験問題と検索キーワードとその検索結果をグローバルデータ化する必要があろう。

## 6. まとめ

これまでのコンピュータ化された医師国家試験問題システムは、年度別や分野別に問題が表示されるだけで、設問中や解答群の文字列の検索が出来なかった。今回作成した検索システムにより、最大4つのキーワードにて過去にどのような問題が出題されたかが把握できるようになった。

キーワードを4つ指定した場合の検索・表示時間は過去8回分のデータに対し、1秒以下で検索速度は問題無かった(但し、表示される画像

数により若干異なる)。また、ファイルサイズを10倍とした場合での検索・表示時間は4秒であり、利用に十分な時間であった。

指定された4つまでのキーワードは、色分けされて各問ごとに表示され、設問の何処に検索キーワードがあるのかが一目瞭然となり検索システムの利用勝手が良くなり、過去出題の類似問題の把握が容易となった。

いくつかの表示上の課題点については UTF-8 の利用で解決できる目処が立った。

## 参考文献

- 1) 木村一元, 矢口裕子: 図書館所蔵雑誌検索システム, 第28回日本MTA大会論文集, 25-26, 2001, 北海道
- 2) 木村一元: WebLinkアプリケーションからCSPアプリケーションへ, 第33回日本MTA大会論文集, 55-60, 2006, 東京
- 3) 医事試験制度研究会: 平成17年版 医師国家資格試験出題基準, pp148, まほろば, 2004
- 4) [http://en.wikipedia.org/wiki/JPEG\\_File\\_Interchange\\_Format](http://en.wikipedia.org/wiki/JPEG_File_Interchange_Format)
- 5) [http://epi.fm.senshu-u.ac.jp/~ph150187/distribution/050722\\_jpeg\\_etc.doc](http://epi.fm.senshu-u.ac.jp/~ph150187/distribution/050722_jpeg_etc.doc)
- 6) <http://hp.vector.co.jp/authors/VA032610/contents.htm>
- 7) まつもと ゆきひろ, 石塚 圭樹: オブジェクト指向スクリプト言語 Ruby, pp576, アスキー出版, 1999
- 8) 木村一元: WebLink を使った医学部学生向け教育システム, 第27回日本MTA大会論文集, 41-44, 2000, 名古屋



## 「日本Mテクノロジー学会」ご入会のご案内

日本Mテクノロジー学会（日本MTA）は、M言語（MUMPS）の利用・改良・普及を目的とした団体で、個人や法人が加入して活発な活動を行っております。M言語はANSIにFORTRAN及びCOBOLに続いて3番目の標準コンピュータ言語として制定され、米国連邦情報処理標準言語にも採用されました。さらに1992年5月にはISO標準言語として制定され、また、1995年には日本工業規格にも採用されました。一方、近年のコンピュータのダウンサイジングの流れにあつて、ユーザーも着実に増えつつあります。

日本MTAは先に述べたような目的に向けて種々の活動を続けておりますが、貴方にも、是非とも日本MTAに参加し活動を盛り上げて頂きたくご案内申し上げます次第です。

### A. 日本MTAの活動

- 1) 年次学術大会、研究会や講習会の開催
- 2) M言語に関する技術情報の提供
  - PC通信Nifty-SERVE上にMUMPSフォーラムを設置
  - MTAニュースの発行
  - 各種資料の配布
- 3) 学術雑誌「Mumps」の出版
- 4) M言語改良仕様の検討・・・米国M Development Committeeと連携
- 5) 国際MTA、各国MTA（MUG）との交流
- 6) M言語のJIS化推進
- 7) ソフトウェアの公開流通

### B. 会員の特典

会員になることにより次のような特典が考えられ、充分満足頂けるものと考えられます。

#### \*個人会員の特典

- 1) 日本MTA年次大会、M言語関係学術集会、研究会、講習会のお知らせ
- 2) 日本MTA主催の学術集会、研究会、講習会などの参加費用の割引
- 3) M言語に関する各種資料の実費提供
- 4) 流通、ソフトウェア（MTAPAL）の低額頒布
- 5) 「MTAニュース」の無料配布
- 6) M言語ベンダーの折々のプロダクト紹介・パンフレット・カタログ類の頒布
- 7) 雑誌「Mumps」の無料配布

・上記の各種活動を通じて、M言語に関する全世界の最新の技術情報が得られます。

#### \*法人会員の特典

法人会員は「日本MTAの目的に賛同する法人で、日本MTAの目的を遂行するために積極的に事業を後援する事を表明した者とし、正副各1名の代表者を登録し、正副代表者とも個人会員と同等の資格を持つ」こととなります。尚、正副代表者には正会員と同様の日本MTAの役員としての道があります。

- 1) 日本MTA主催の集会には5名迄、会場費、講習会費などを会員割引

- 2) 日本MTA主催の医療人、企業人を対象とする講習会へ法人会員から優先的に出講
- 3) 日本MTA主催の集会への出品、展示に関する料金の割引
- 4) 日本MTA 学術大会論文集、MTAニュース等への広告費の割引
- 5) 法人会員のプロダクトのパンフレット、カタログ類の会員への頒布
- 6) ユーザー法人にはM言語ベンダーないしシステムエンジニアの紹介
- 7) 日本MTAの流通パッケージ（MTAPAL）を割引料金で利用
- 8) MTAニュースを単なる広告ではなく、新しいプロダクトの紹介等の質の高いPRのために利用可能

注意) 法人会員は、国際MTAが設けている施設会員と企業会員に相当するものですが、学校法人・国立施設など税法上非営利団体扱いの法人を非営利法人とし、国際慣例よりも40%低い基本会費を申し受けます。その他は企業法人ないしベンダー法人としての会費を申し受けます。ご入会の手続きは「法人会員入会申込書」によってお願い申し上げます。

- ・上記の各種活動を通じて、M言語に関する全世界の最新の技術情報が得られます。
- ・M言語ユーザ間、M言語を取り扱うベンダー・メーカー間とのコミュニケーションが充実します。

#### C. 会費

##### ア) 個人会員

入会費 ￥4,000 (学生会員: ￥1,000)

年会費 ￥6,000 (学生会員: ￥2,000)

##### イ) 法人会員

入会費 ￥10,000 (営利・非営利法人共通)

年会費 ￥50,000 (1口) ←営利法人

￥30,000 (1口) ←非営利法人

注意) 会計年度は、毎年4月1日から翌年3月31日までです。

#### D. ご入会手続き

- 1) 入会資料請求<電話・FAX・郵便>
- 2) 事務局から送付された「会員登録票」(法人会員の場合は正・副代表者の「会員登録票」及び「法人会員申込書」)に必要事項を記入の上、事務局までお送り下さい。
- 3) 郵便払込で入会金、年会費を事務局に納金して下さい。

#### E. 入会費・年会費お支払方法

日本MTA事務局より会費の請求がございましたら、以下口座にお支払下さい。但し、お手数料は振込人払いとさせていただきますことをご了承下さい。

- 郵便振替 口座番号: 01440-8-4520  
口座加入者名: 日本Mテクノロジー学会

以上

# 「日本Mテクノロジー学会」規約

## 第一章 総 則

第1条 本会は日本Mテクノロジー学会 (M Technology Association of Japan)という。

第2条 本会の事務所は幹事会の承認を経て、学会長が指定するところに置く。

## 第二章 目的および事業

第3条 本会は「M言語」並びにこれに関する情報システムの利用、応用、改良、並びに普及を行うことを目的とする。

第4条 本会は前条の目的を達成するため次の事業を行う。

- 1) 学会大会、フェア、研究会、講習会などの開催
- 2) 学会誌、ニュースなどの刊行物の発行
- 3) M言語の日本語装備の標準化
- 4) M言語の標準装備の監視
- 5) 海外のMTA (MUG) などとの連携活動
- 6) 内外の関連諸学会との連絡ならびに協力活動
- 7) M言語利用技術の相互交換の促進、本会に提供された資源の整備、管理ならびに会員への還元
- 8) 日本Mテクノロジー学会出版会に関する事業
- 9) その他目的達成のために必要な事業

## 第三章 会 員

第5条 本会員は個人会員と法人会員からなる。

- 1) 個人会員は本会の目的に賛同し、本会の対象とする領域、又はそれと関連する領域において活動する個人とする。
- 2) 法人会員は本会の目的に賛同する法人で、本会の目的を遂行する為に積極的に事業を後援する事を表明したものとする。法人会員においては正副各1名の代表者を登録するものとする。正副代表者は個人会員と同等の資格を有する。

第6条 本会に入会を希望する者は所定の申込書に入会金及び会費を添えて本会事務所に申し込まねばならない。

第7条 本会会員は、毎年所定の会費を前納しなければならない。

第8条 本会会員で住所変更のあったものは速やかに住所変更届を、また退会しようとするものは退会届を本会事務所に提出しなければならない。本会会員で、住所不明となるか催促にも拘らず2か年を越えて会費納入遅滞のあったものは退会の扱いを受ける。物故会員は退会の扱いを受ける。

第9条 本会の規約に背く行為のあった会員は、幹事会の議決を経てこれを除名することができる。

#### 第四章 役員その他

第10条 本会に次の役員を置く

1) 学会長	1名
2) 日本Mテクノロジー学会大会長(以下「大会長」という)	1名
3) 日本Mテクノロジーフェア実行委員長(以下「フェア実行委員長」という)	1名
4) 幹事 庶務財務担当	1名
国際担当	1名
流通担当	1名
広報担当	1名
雑誌担当	1名
ネットワーク担当	1名
5) 会計監事	1名
6) 評議員	若干名
7) 日本Mテクノロジー学会出版会理事長	1名
8) 日本Mテクノロジー学会出版会理事	若干名

第11条 各役員の選出または構成を次のように定める。

- 1) 評議員に欠員が生じた場合、学会長は評議員会の推薦者を総会に諮り、その承認を得て決定する。評議員の定数は学会長が定める。但し、各評議員の構成割合は会員の職域構成割合に近いものとする。
- 2) 学会長及び会計監事は、評議員会の推薦者を総会に諮り、その承認を経て決定する。



- 3) 幹事は学会長が推薦し、総会の承認を経て決定する。学会長と幹事は併任できない。
- 4) 大会長は学会長が幹事会の推薦者を総会に諮り、その承認を経て決定する。
- 5) フェア実行委員長は学会長が幹事会の推薦者を総会に諮り、その承認を経て決定する。
- 6) 出版会理事長並びに理事は学会長が推薦し、総会の承認を経て決定する。

第12条 各役員の任務は次のように定める。

- 1) 学会長は会を代表し、総会、幹事会、評議員会の議長となる。
- 2) 大会長は、年次日本Mテクノロジー学会大会を総括する。
- 3) フェア実行委員長は、年次日本Mテクノロジーフェアを総括する。
- 4) 庶務財務担当幹事は、本会に関する庶務及び全ての資金及び財産の管理を行う。また、最新の名簿の管理、総会その他の議事録の管理を行う。
- 5) 国際担当幹事は、海外のMTA (MUG) 組織との連携並びにM言語開発委員との協力を司り、その他の国際的協力を行う。
- 6) 流通担当幹事は、M言語応用プログラムのユーザー間相互交換の促進、MUGプロトタイプ・アプリケーション・ライブラリー (MUGPAL) など M言語資源の整備、管理、維持、会員に対する資料提供等のサービスを行う。
- 7) 広報担当幹事は、Mテクノロジーニュース等を通じ広報活動を行う。
- 8) 雑誌担当幹事は、学会誌「Mumps」の編集を兼ね、出版の進行を司る。
- 9) ネットワーク担当幹事は、ネットワークを活用した会員間のコミュニケーションの向上を図る。
- 10) 会計監事は、年次会計の監査を行い総会に報告する。

第13条 各役員の任期を次のように定める。

- 1) 学会長、幹事、会計監事の任期は、4月1日より翌々年3月31日までの2年間とし再任を妨げない。
- 2) 大会長の任期は、前学会終了時に始まり学会の残務処理の終了までの期間とする。
- 2) フェア実行委員長の任期は、前Mテクノロジーフェア終了時に始まりMテクノロジーフェアの残務処理の終了までの期間とする。
- 3) 評議員の任期は特に定めないが、4年間続けて評議員会に出席しなければ評議員資格を失う。

## 第五章 会議および委員会

第14条 (総会)

- 1) 総会は本会の最高の議決機関である。

- 2) 総会は学会長が毎年1回召集する。但し、幹事会の議決による場合または会員の5分の1以上から請求された場合、学会長は臨時総会を召集しなければならない。
- 3) 総会の議長は学会長とする。
- 4) 次の事項は総会に提出してその承認を受けなければならない。
  - a. 事業報告および収支決算
  - b. 事業計画および収支予算
  - c. その他幹事会が必要と認めた事項
- 5) 総会の成立に必要な出席者数は会員のうち50名または10%の少ない方を上回る数とする。
- 6) 総会の議決は本規約に別に定めるものの他、出席会員の過半数による。

#### 第15条 (幹事会)

- 1) 学会長が必要に応じて召集する。但し、幹事の過半数から請求があった時は、学会長は幹事会を召集しなければならない。
- 2) 幹事会の議長は学会長とする。
- 3) 幹事会は学会長、大会長、フェア実行委員長、幹事、会計監事により構成される。
- 4) 学会長は必要に応じて各種委員会の委員長を出席させることができる。
- 5) 幹事会の議決は構成員の過半数による。

#### 第16条 (評議員会)

- 1) 学会長が毎年1回召集する。但し、学会長は必要に応じて臨時評議委員会を召集する。
- 2) 評議員会は学会長の諮問に答え本会の重要案件を審議する。議長は学会長とする。
- 3) 評議員会は学会長、会計監事、Mumps 編集委員、新評議員を総会に推薦する。

#### 第17条 (学会誌 Mumps 編集委員会)

- 1) 雑誌担当幹事は必要に応じて学会誌 Mumps 編集委員会を召集する。
- 2) 学会誌 Mumps 編集委員会の議長は雑誌担当幹事とする。
- 3) 学会誌 Mumps 編集委員は編集委員会が任命する。任期は3年とし、再任を妨げない。

#### 第18条 (各種委員会)

- 1) 学会長は必要に応じて幹事会の議を経て各種委員会を設置、統合、分化、改廃することができる。

#### 第19条 (日本Mテクノロジー学会大会)

- 1) 本会は年1回以上の日本Mテクノロジー学会大会を開催する。

第20条 (日本Mテクノロジーフェア)

- 1) 本会は年1回以上の日本Mテクノロジーフェアを開催する。

第21条 (日本Mテクノロジー学会出版会)

- 1) 日本Mテクノロジー学会出版会の規約は別途定める。

## 第六章 資産および会計

第22条 本会の資産は次の通りとする。

- 1) 本会の設立当初からの財産
- 2) 入会金および会費
- 3) 事業に伴う収入
- 4) 資産から生ずる利子など
- 5) 寄付金品
- 6) 負担金
- 7) その他

第23条 本会の資産は、学会長及び庶務財務担当幹事が管理する。

第24条 本会の重要な財産（基本財産）に関しては、これを消費し、または担保にしてはならない。但し、本会の事業遂行上止むを得ない理由があるときは、幹事会の出席者の2/3以上の議決と総会の出席者の3/4以上の議決を経てその一部に限り処分し、または担保に供することができる。

第25条 本会の事業計画およびこれに伴う収支予算は、年度毎に学会長および庶務財務担当幹事が編集し、幹事会の議決を経て総会の承認を得なければならない。

第26条 本会の事業報告書および収支決算は、年度毎に学会長および庶務財務担当幹事が作成し、会計監事が監査し、幹事会の議決を経て総会の承認を得なければならない。

第27条 本会支援のため各種団体よりの負担金、寄付、研究費などの交付があった場合、幹事会の承

認により本会の資産として受け入れる。

#### 第七章 規約の変更ならびに解散

第28条 本規約の改正は幹事会および総会において各々出席会員の2／3以上の議決を経なければならない。

第29条 会を解散するには総会において出席会員の3／4以上の同意を必要とする。

第30条 会の解散に伴う残余財産は、法律による制限のあるものの他は世界保健機構（WHO）に寄付するものとする。

#### 第八章 付 則

第31条 本会の略称を日本MTA、英文略称をMTA-JPという。

第32条 本会の入会費、年会費は別に定めるものとする。

第33条 学会長は本会の発展に功績のあった特定個人に対し名誉会長、名誉会員の称号を与えることができる。

第34条

- 1) 本規約は1977年10月29日より発効するものとする。
- 2) 本規約は1979年 9月14日より改訂し発効するものとする。
- 3) 本規約は1987年 7月29日より改訂し発効するものとする。
- 4) 本規約は1991年10月31日より改訂し発効するものとする。
- 5) 本規約は1992年 8月 1日より改訂し発効するものとする。
- 6) 本規約は1992年10月29日より改訂し発効するものとする。
- 7) 本規約は1993年 4月 1日より改訂し発効するものとする。
- 8) 本規約は1994年 8月 6日より改訂し発効するものとする。
- 9) 本規約は1995年 9月30日より改訂し発効するものとする。
- 10) 本規約は1996年 9月15日より改訂し発効するものとする。
- 11) 本規約は2003年 9月 6日より改訂し発効するものとする。



## 「Mumps」投稿規定

(1991年7月10日制定)

(1994年12月1日改正)

(2008年11月24日改正)

本規定は日本 M テクノロジー学会誌「Mumps」に、会員が自発的に寄稿する論文（以下投稿論文という）に関する必要事項を定めたものです。学会誌「Mumps」には、編集委員会が依頼する原稿（依頼原稿）も掲載しますが、それについての必要事項はそのつど定めます。

### 1. 論文の主題

投稿を受ける論文の主題は、コンピュータシステム／言語である MUMPS に直接、間接に関係するものとします。

例えば、MUMPS の利用技術についての考案や開発、MUMPS の言語についての直接仕様や提言、MUMPS システム装備、MUMPS と他の世界とのインターフェース、MUMPS の教育など、MUMPS に関係するあるいは関係しそうなテーマについて広く受け入れます。ただし、他の雑誌に掲載された、あるいは投稿中の論文はお断りします。

### 2. 投稿論文の種類

投稿論文は次の 6 種類に限ります。

#### 1) 原著論文

未投稿で、論文の主要部分に独創性、独自性のある論文。既に発表した問題について別の視点からまとめた論文も未投稿原著論文であり得ます。また、応用開発、調査等であっても、その過程で創意工夫や独自性があれば原著論文の対象とします。

#### 2) 総説

ある主題について、過去の研究業績を詳細にまとめ文献を伴って記述し、その主題に関する現状と将来展望を明らかにした論文。

#### 3) 研究速報

新しい研究成果が原著になるほどにはまとまっていないが発表に価値があると考えられるもの。

#### 4) 技術ノート

作成したプログラムや新しいシステムの紹介など、MUMPS 技術に関する論文で、会員の相互の利益になると思われるもの。

#### 5) フォーラム

意見、提案、提言、感想、著書や学術集会の紹介など、上記以外で会員の利益になると思われるもの。

#### 6) Letter to the editor

原著論文に対する質問やコメント、日本 MTA の活動に関係のあるコメントなど。

### 3. 投稿論文の長さ

原則として下記の表の通りの長さとしします。A 4 用紙（横 21 字×縦 41 行×2 段組=1722 文字）で刷り上

がりページ1枚となります。ただし、これを越える場合でも、編集委員会が必要と認めた場合には別に定める超過料金を支払って掲載することができます。

論文の種類	論文のページ数（刷り上がり）
原著	10ページ（以内）
総説	30ページ
研究速報	6ページ
技術ノート	6ページ
フォーラム	4ページ
Letter to the Editor	1ページ

#### 4. 投稿者の条件

- 1) 筆頭著者は日本 M テクノロジー学会会員であること。
- 2) 共著者も原則として会員であることとします。

#### 5. 原稿の送付

原稿（2段組の印刷形式原稿でも可）を下記編集委員会宛てに送って下さい。原稿到着日を投稿の受け付け日としその日付を誌上に明記致します。

#### 原稿送付先・連絡先

〒321-0293

栃木県下都賀郡壬生町大字北小林 880 番地

獨協医科大学病院医療情報センター 木村 一元

TEL : 0282-87-2344 FAX : 0282-86-2606

e-mail : kimura@dokkyomed.ac.jp

#### 6. 掲載の採否

投稿された原稿は、編集委員会が依頼する2名の査読者が査読します。そしてその査読の意見を考慮して編集委員会はその原稿の採否を決定します。査読の結果によっては、原稿の内容や論文の種類を修正変更することを投稿者にもお願いすることもあります。

#### 7. 原稿作成要領

##### 1) 原稿の構成

投稿原稿はおよそ次の構成に従って作成して下さい。

- a) 論文の題名
- b) 著者名、所属、所在地
  - a) と b) は日本語と英語の両方を記入して下さい。
- c) キーワード・・・8語以内（日・英）

- d) 和文要旨・・・200字から400字
- e) 英文要旨・・・200words から300words
- f) 本文
- g) 謝辞・・・・・・・・・・必要に応じて
- h) 文献リスト

文献の引用は本文中の引用箇所に出現順に通し番号[1], [3-5]等を記し、本文の末尾に一括して引用番号順に並べて下さい。雑誌の文献は引用番号、著者名、論文題名、雑誌名、巻号、最初と最後の頁数、西暦年号の順です。

単行本の文献は引用番号、著者名、題名、版数、引用頁、発行社、発行地、西暦年号の順です。

(例)

1. 福井太郎：糖尿病患者管理システムの開発，医学情報学，10(2):30-35(1990).

- i) 図表・・・図や表はそれぞれを本文中に入れて下さい。図や表の大きさは基本的に著者の意向に沿いますが、大き過ぎると判断された場合は、縮小されることがあります。
- j) 特殊文字・・・特殊文字は原則として禁止しますが使用される場合は使用位置を別紙にて明示して下さい。

## 2) 投稿原稿 (FD, CD 等の記憶媒体または電子メールで提出)

原稿は標準的なワープロ（一太郎、MS-Word）で、A4用紙に横42文字×縦41行を1頁として作成して下さい。印刷原稿の形式でも受付けます。

また、原稿には表紙を付け、表紙にはつぎの事項を記入して下さい。

表紙・・・題名

連絡先（氏名・住所・電話・FAX）

原稿の種類

原稿の枚数（本文・図・表別に）

別冊希望部数（50部の倍数）

その他・・・特殊文字等を使用されている場合は明記して下さい。

## 3) 印刷原稿 (FD, CD 等の記憶媒体または電子メールで提出)

採用が決定した印刷原稿は、標準的なワープロにてA4用紙2段組のカメラレディの原稿にて提出して下さい。

(印刷原稿、カメラレディ原稿作成時の注意事項)

\*基本的に横21字×縦41行×2段組が1頁になって印刷されます。

\*原稿は題名（日本語）、題名（英語）、著者名（日本語）、著者名（英語）、著者所属・住所（日本語）、著者所属・住所（英語）、和文抄録、キーワード（日本語）、英文抄録、キーワード（英語）、本文の順で同一ファイル名に保存して下さい。

\*印刷は、モノクロで行われます。原稿にてカラーを使われる場合には、この点に留意して原稿を作成して下さい。

\*なお、編集側にてタイトル、著者名、所属、要旨の形式の統一を行います。また、タイトルページの左下に原稿受付の月日、査読後の受理月日を入れますので提出して戴いた原稿と異なることがあります。

#### 8. 著作権

掲載論文の著作権は日本 M テクノロジー学会に帰属するものとする。

#### 9. 別刷

別刷は30部まで無料とし、それ以上は実費とする。別刷の部数は投稿時または校正原稿提出時に申し出て戴ければ、30部単位で増刷する。



### 「Mumps」誌編集委員

編集委員長 木村 一元（獨協医科大学病院医療情報センター）  
編集委員 大櫛 陽一（東海大学医学部基礎医学系）  
山本 和子（日本 MAT 名誉会長 医療情報技師育成会）  
本多 正幸（長崎大学医学部附属病院医療情報部）

#### 編集後記

皆様の協力を得まして学会誌「Mumps」の第 25 巻を発行することが出来ました。今回は、投稿〆切を大会時（9月）では無く、11 月としました。昨年同様、大会での発表内容を論文として投稿頂き、査読者の協力を得て、何とか発行できる段取りとなりました。

お陰様で掲載論文は、医療関係 2 編、言語開発 1 編と言語処理関係 3 編の計 6 編と成りました。これらの論文は、昨年と同様、M を用いて作成したシステム、M で作成したデータベースからの知見、M で作成したシステムの改良、M の柔軟な文字処理機能を活用した日本語解析、聾者へのその解析機能の応用、また、M の樹状構造を活用した絵記号表示システムと多彩であった。

今後、種々の分野での M の柔軟な機能を生かした利用と知見が報告され、それぞれの分野での活動が活発になることを期待する。（木村）

発行者 日本Mテクノロジー学会  
会長 木村 一元  
〒321-0293 栃木県下都賀郡壬生町北小林 880  
獨協医科大学病院 医療情報センター  
Tel : 0282-87-2344  
Fax : 0282-86-2606

編集者 日本Mテクノロジー学会 編集委員会  
委員長 木村 一元  
〒321-0293 栃木県下都賀郡壬生町北小林 880  
獨協医科大学病院 医療情報センター  
Tel : 0282-87-2344  
Fax : 0282-86-2606

事務局 日本Mテクノロジー学会 事務局  
庶務財務担当幹事 大櫛陽一  
〒259-1193 神奈川県伊勢原市下糟屋 143  
東海大学医学部 基礎医学系 医学教育・情報学  
Tel : 0463-93-1121 Ex.2140, 2143  
Fax : 0463-93-5418

印刷 株式会社松井ピ・テ・オ・印刷  
〒321-0904 宇都宮市陽東 5-9-21  
Tel : 028-662-2511  
Fax : 028-662-4278

CONTENTS

■ Editorial

Kazumoto Kimura 1

■ Original Articles

Probabilistic Logic Programming on M . . . . . Tatsuhiko Uchida 3

Two Equivalence Classes of the Perceptive Collocations  
and the Reduction Methods of Japanese Sentences  
--- Semantic Analysis with the Japanese Analysis  
System SASAYURI on the M Language Scheme --- . . . . . Wataru Takahasi 9

Alternative Representation Function of the Japanese  
Sentences with Onomatopoeia on the M Language Scheme  
---Information Presentation Techniques for the Deaf--- . . . . . Masatoshi Tsumura 23  
Wataru Takahasi

Development of a PIC Symbol Processor on the M  
Language Scheme and the Linguistics of Visual Codes . . . . . Mariko Ikeda 35  
Eiji Yanagiuchi  
Wataru Takahasi

Development of CTI system linked to medical  
information system . . . . . Takayoshi Tsuchiya 51  
Chieko Tanaka  
Takeshi Kayamori  
Kazumoto Kimura

Retrieval System of Questions on the National  
Examination for Medical Practitioners using CSP . . . . . Kazumoto Kimura 59  
Yoshihiko Igarashi

■ Other Information

Information for MTA . . . . . 67

Information for Authors . . . . . 75

Editor's Postscripts . . . . . 79