

専門用語研究

Journal of the Japan Terminology Association

No.13 1997. 03

「第10回 専門用語シンポジウム：脳と用語」から

能動的視覚情報処理の神経回路モデル	福島 邦彦	1
知識獲得とオントロジー構築	西田 豊明	10
用語間の意味関係の自己組織化	藤原 譲	16
情報流通と分類表の役割およびその体系	石川 徹也	21
「分類語彙表」の構築	中野 洋	27
TKE '96 にいたる 最近のターミノロジー研究	仲本秀四郎	35
スイス連邦官房ターミノロジー課訪問記	春山 暁美	46
JIS 用語規格制定の動き 1996		48
投稿規定		50

第10回専門用語シンポジウム：脳と用語

標記シンポジウムが1996年12月12日、日本学術会議講堂にて、約70名の参加者を得て開催された。今回は、日本学術会議（学術文献情報研究連絡専門委員会）、情報処理学会（情報学基礎研究会）、情報知識学会との共催である。発表は、下記のように、招待講演3件と一般講演6件の合計9件であった。

招待講演	脳の生化学および情報処理研究	電子技術総合研究所	松本 元
〃	能動的視覚情報処理の神経回路モデル*	大阪大学基礎工学部	福島 邦彦
〃	言葉のネットワーク	京都大学工学研究科	長尾 真
一般講演	オントロジーと知識構造	大阪大学産業科学研究所	溝口理一郎
〃	知識獲得とオントロジー構築*	奈良先端科学技術大学院大学	西田 豊明
〃	用語間の意味関係の自己組織化*	筑波大学電子情報工学系	藤原 譲
〃	情報流通と分類表の役割およびその体系*	図書館情報大学	石川 徹也
〃	【分類語彙表】の構築*	国立国語研究所	中野 洋
〃	TKE '96 にいたる最近のターミノロジー研究*		仲本秀四郎

本号では、各発表者に修正・加筆をいただいた6件（*印）を掲載する。

～「第10回専門用語シンポジウム：脳と用語」から～

能動的視覚情報処理の神経回路モデル

福島 邦彦* FUKUSHIMA, Kunihiko

1. はじめに

最近のニューロサイエンスの進歩によって、脳のメカニズムも次第に明らかになりつつある。しかし脳の高次中枢における情報処理機構に関しては、生理学や心理学の実験によって断片的な知見は急速に増えつつあるが、一つのまとまったシステムとしての機構説明にはまだほど遠いのが現状である。このような状況にあって、神経回路モデルを仲介とする脳研究の重要性がますます高まってきた。

神経回路モデルの研究では、まず、脳の持つ特定の機能に注目して、脳と同じ動きを示す神経回路モデルを作ることから研究を始める。このとき、生理学や心理学で解明されている事実モデルの中にもできるだけ忠実に取り入れる。しかし、まだ解明されていない部分に関しては、“その機能が実現されるためには神経回路はどのような構成になっていなければならないか”を考へて、大胆な仮説をたて、その仮説に従ってモデルを作るのである。モデルが出来ると、種々の刺激に対してそのモデルがどのような反応を示すかをコンピュータシミュレーションや数学的な解析によって調べる。このようにして求めたモデルの反応を、実際の生物の脳に同じ刺激を与えたときの反応と比較してみる。もし両者に違いがあれば、モデルのもとになった仮説に誤りがあったと考えて、仮説を修正していく。このような手順を何回も繰り返しながらモデルを改善して、実際の神経系に近付けていく。もっとも、モデルが脳と同じ反応をしたからといって、直ちに、生物の脳がこのようにして作られたモデルと全く同じ原理で働いていると速断することはできない。もう一度生理実験に立

ち戻って確認する必要がある。しかし、両者が同じ反応を示すのであるから、同じ原理で働いている可能性は高いと言える。このように、神経回路モデルの研究と従来の生理学・心理学研究との関係は、ちょうど理論物理学と実験物理学との関係と同じである。神経回路モデルは、脳研究の新しい手法の一つとして、特に生理学実験の困難な高次中枢の研究に対して威力を発揮すると期待される。

ところで視覚系の神経回路モデルの研究は、これまでは入力パターンを受動的に処理するものが大部分であった。しかし、われわれ人間は、網膜に映った外界のイメージを受動的にすべて受け入れているのではない。もっと能動的に自分が必要とする情報だけを選び取っている。関心のある箇所に眼球を動かしたり注意を集中したりして、その部分からの情報を選択的に取り込んで処理しているのである。何か重要な情報がありそうなところに注意を向けて、そこだけを詳しく観察したり、ある場所に存在するパターンをその前後の情報を用いて予測してみても、その予測が正しいかどうかを確認しながら認識操作を進めていくようなことも行なっている。

本稿では、筆者の行なっている研究の中から、能動的視覚情報処理機構のモデルをいくつか紹介する。パターンの意味や概念が脳の中でどのように表現され扱われているのかについても何らかの示唆を与えようことを期待する。

2. 選択的注意のモデル

脳の中での視覚情報処理には、末梢から中枢に向かうフォワードの信号の流れだけでなく、中枢から末梢に向かうバックワードの信号の流れも重要な役割を果たしている。筆者が提唱した選択的

* 大阪大学基礎工学部

注意モデル^{[1],[2]}は、階層型多層回路にバックワードの結合を追加した構造を持っていて、刺激パターンが与えられると、自分の見たいところだけに選択的に注意を向けながら入力情報を取り込んで処理していく。パターン認識能力を持つだけでなく、かすれや汚れのある不完全なパターンから完全なパターンを“連想”したり（パターンの修復）、複数のパターンの中から特定のパターンを切り出したり（セグメンテーション）する能力なども持っている。

モデルの回路構成を、フォワードの信号を伝える細胞の層とバックワードの信号を伝える細胞の層とに分けて描くと図1のようにになる。バックワードの系は、フォワードの系を鏡に映したような構造をもっている。両者の違いは信号の流れの方向だけである。フォワードの系では下から上へ信号が送られていくのに対して、バックワードの系では上から下に戻ってくる。フォワードの系の最下位にあるのが外部からの刺激入力を受け取る入力層であり、最上位にあるのがパターン認識の結果を出力する認識出力層である。バックワードの系の最下位には連想出力層がある。この回路の中で、フォワード信号はパターン認識の機能を受け持ち、バックワード信号は選択的注意および連想の機能を受け持つ。

フォワードの回路は、筆者が以前提唱したネオコグニトロンとほぼ同じ構造をもっていて、パターン認識の処理を行なっている。フォワードの系の

一番下にあるのが入力層で、ここに刺激パターンが与えられる。一番上にあるのが認識細胞層である。

ここでネオコグニトロン^{[3],[4]}について簡単に説明しておこう。ネコやサルの大脳の視覚野には、特定の傾きの線や輪郭に反応する細胞など、網膜に写ったパターンの部分的な特徴に反応する細胞がある。視覚野よりさらに上位の中枢には、丸、三角、四角といった図形や、ヒトの顔のような複雑なパターンに選択的に反応する細胞もあるといわれている。つまり視覚神経系は、簡単なものから複雑なものへ情報が処理されていくという階層構造を持つものと考えられる。ネオコグニトロンは、このような生理学的知見をヒントにして作られた神経回路モデルで、階層構造を持つ多層回路からなる。入力層の後ろには、S細胞と名付けた特徴抽出細胞の層（ U_s ）と、C細胞と名付けた細胞の層（ U_c ）とが交互に並んでいる。最上位層のC細胞が、ネオコグニトロンの認識結果を示す認識細胞の働きをする。回路は自己組織化の能力を持っていて、パターン認識能力を学習によって獲得していく。学習終了後は、入力パターンの位置のずれや形のゆがみの影響をほとんど受けることなく、入力パターンを正しく認識する。しかも汎化能力があるので、変形パターンをいちいち教えてやる必要はない。未学習のパターンでも、それが過去に学習したパターンと似ていさえすれば正しく認識するので、手書き文字認識など、広

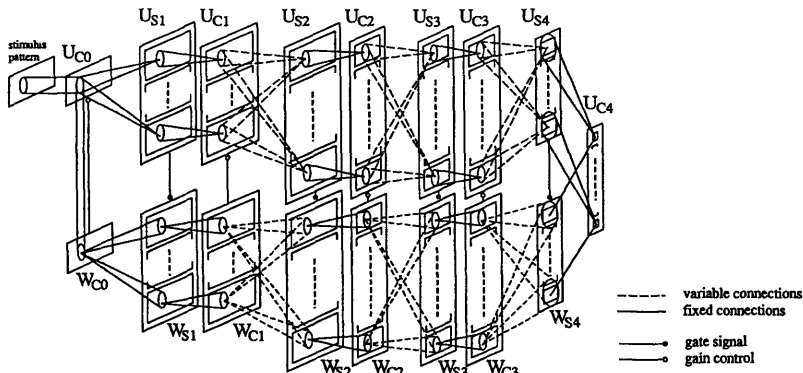


図1 選択的注意のモデルの構造。

い応用範囲を持っている¹⁰⁾。

学習時には、S細胞の入力結合の強度を変えることによって特徴抽出細胞を作り上げていく。学習には競合学習の原理を用いる。S細胞は、学習パターンが与えられるごとに自分の近傍の細胞と競合し、競合に勝った細胞だけが入力結合の強化を受ける。その強化量は、その結合に信号を送り込んでいる前段の細胞の出力に比例する。その結果、最大出力細胞の入力結合は、そのときの入力刺激に一致したテンプレートを形成することになる。C細胞は、S細胞が抽出した特徴の、位置ずれの影響を吸収する働きをしている。各C細胞は、わずかに場所が違ったところから同じ特徴を抽出してくるような一群のS細胞から興奮性固定結合を受け取っていて、特徴の呈示位置が入力層上で多少ずれても、同じC細胞が出力を出し続けるようになっているのである。入力パターンの情報は、多層回路内の各段でS細胞による特徴の抽出とC細胞による特徴の位置ずれの許容化を繰り返しながら上位段に送られていく。この過程において、下位段で抽出された局所的な特徴は、次第に大局的な特徴に組み上げられていく。このとき、入力パターンの変形にともなう局所的特徴の相対的な位置ずれの影響は、C細胞の働きによって少しずつ吸収されるので、最終的には、入力パターンの変形に対しても影響されない出力を得ることができる。

さて、選択的注意のモデルに話を戻そう。フォワードの系の最上位段の認識細胞層の細胞の出力は、今度はバックワードの系を通過して下に戻っていき、バックワードの系の最下位段にある連想出力層に達する。連想出力層の出力は再び入力層にフィードバックされる。このようにして回路内では、フォワードとバックワードの回路で形成されるフィードバックループの中を、信号がぐるぐる回りはじめる。しかし、信号は、フォワードの系とバックワードの系の中を、互いに独立に流れているわけではない。二つの系は、互いに相手側の信号の流れを強め合ったり弱め合ったりするような信号を授受し合っている。つまり、フォワードの系からバックワードの系に対しては、信号の流れの経路を制御する信号を送っている。このため、

バックワードの信号は、フォワードの信号と同じ経路を逆向きにたどって流れていくようになる。また逆に、バックワードの系からは、そのバックワード信号が流れている経路に対応したフォワードの系の経路に向けて、フォワードの信号を流しやすくするような制御信号を送っている。これは、入力刺激の中の特定の部分だけに注意を集中する働きをしているということが出来る。したがって、回路の中を信号がぐるぐる回っているうちに、回路の反応は時間とともに次第に変わっていく。

バックワードの系の最下段の連想出力層の働きについては、いろいろな解釈ができる。入力層に不完全なパターン、すなわち、かすれたパターンや汚れたパターンをみせても、回路の中を信号が回っているうちに、かすれや汚れが修復されたきれいなパターンが連想出力層に現われてくる。したがって、不完全なパターンから完全なパターンが連想されてくる層であるとも考えることもできる。あるいはこの連想出力層は、セグメンテーション(切り出し)の出力が現われる層であるとも考えることもできる。例えば入力層にAとBという二つのパターンが与えられたとすると、この回路はある瞬間にはAの方に注意を向けて、これはAであるという答えを認識出力層にだすと同時に、“Aはこういう形をしてこの場所にあった”という、Aの部分だけを切り出した結果をこの層に出すからである。

以上をまとめると、このモデルは図2に示すよ

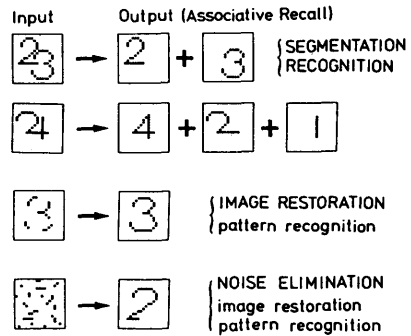


図2 選択的注意モデルの諸機能。これらの機能は、入力パターンが変形しても、位置がずれても正しく働く。

うに、二つ以上のパターンが同時に与えられても、個々のパターンを順番に認識し、切り分けていくことができる。不完全なパターンが与えられても、それを修復した完全なパターンを連想することができる。入力に含まれている汚れ(ノイズ)を除去する働きもある。しかもこのような機能が、刺激パターンが学習パターンとは異なった位置に与えられたり、あるいは形が歪んでいる場合にも正しく働く。

この選択的注意のモデルの考え方を用いると、英文筆記体の続け文字を読みとるシステム^{[6],[7]}を作ることでもできる。また、漢字を偏や旁に分けて認識したり^[8]、顔のパターンから目や鼻などを分離して切り出す目的にも使えそうである^[9]。

ところでこの選択的注意のモデルでは、バックワード信号が最上位の段の認識細胞だけから送り出されるような1重のフィードバックループだけしかもっていない。しかし実際の脳は、単一文字の認識に際しても、その文字の形の情報だけでなく、その前後関係や意味の情報なども利用して判断していると考えられる。したがって今後はモデルをさらに発展させて、多重のフィードバックループを導入して、文字認識細胞にも、もっと上位の、単語や文章の意味のレベルからのフィードバック信号を伝えるような構造にすることが重要な研究課題である。

3. 動きと形の並列処理

哺乳動物の視覚系では、後頭葉から側頭葉に至る経路で、視野内の物体の形や色など、主に物体視に関する情報を分析し(形のチャンネル)、後頭葉から頭頂葉に至る経路で、物体の動きや空間的位置など、主に空間視に関する情報処理を行なっている(動きのチャンネル)。両チャンネルとも階層構造を持ち、細胞の受容野は階層が高くなるほど大きくなり、その反応は網膜上の位置には影響を受けなくなっていく。そこで視野の中に複数の物体があったとき、各チャンネルの高次領野に生じる反応のうち、同一物体に起因するものどうしの対応をいかにしてとっているのか、ということが問題となる。このような問題は結びつけ問題(binding problem)と呼ばれ、その機構を説明

する仮説がいくつか提唱されている^{[10],[11]}。

一方解剖学的知見によると、形のチャンネルにも動きのチャンネルにも、フォワード結合だけでなくバックワード結合があり、各チャンネル内には一種の正帰還ループが形成されている。また、フォワード結合は、形のチャンネルと動きのチャンネルとはつきりと分離されているが、バックワード結合は両チャンネルにまたがって広がっており、下位の層にフィードバックされてくる信号は自分のチャンネルだけでなく相手側のチャンネルにも伝えられるようになってきているという^[12]。

このような知見をもとに筆者らは“脳は結びつけ問題を選択的注意機構によって解いている”と考え、神経回路モデルを提唱した^{[13],[14]}。モデルは図3のように、形の処理系と動きの処理系の二つのチャンネルからなる。形の処理系の最上位段はIT野を想定した認識層であり、動きの処理系の最上位段はMST野を想定した運動方向の知覚層である。それぞれの処理系はフォワードとバックワードの双方向の結合を持ち、バックワードの信号は自分のチャンネルだけでなく相手側のチャンネルにも送られる。

形のチャンネルは上述の選択的注意モデルとほぼ同じ構造を持っており、視野内の一つの物体に注意を向けて認識し、その物体を切り出してくる。動きのチャンネルも、フォワード信号とバックワード信号の相互作用によって、特定の一つの物体だけに注意を向け、その物体の動きの方向を検出する。この二つのチャンネルの注意選択機構はそれぞれ独立に働くのではなく、相手側のチャンネルから送られてくるバックワード信号の影響を受けて、常に同じ物体に注意を向けるようになる。その結果ある時刻には、形のチャンネルの最上位段では、視野内の一つの物体の形を認識し、動きのチャンネルの最上位段ではその物体の動きの方向を抽出する。一つの物体に対する注意の集中がしばらく続いて回路の反応が定常状態に近づくと、このことを図3の右端のモニタ回路が検出し、他の物体に注意を切り替える指令を送り出す。

モデルのシミュレーションを行なった。形のチャンネルには、あらかじめ三角形、長方形、菱形などを学習パターンとして呈示して教師なし学習によ

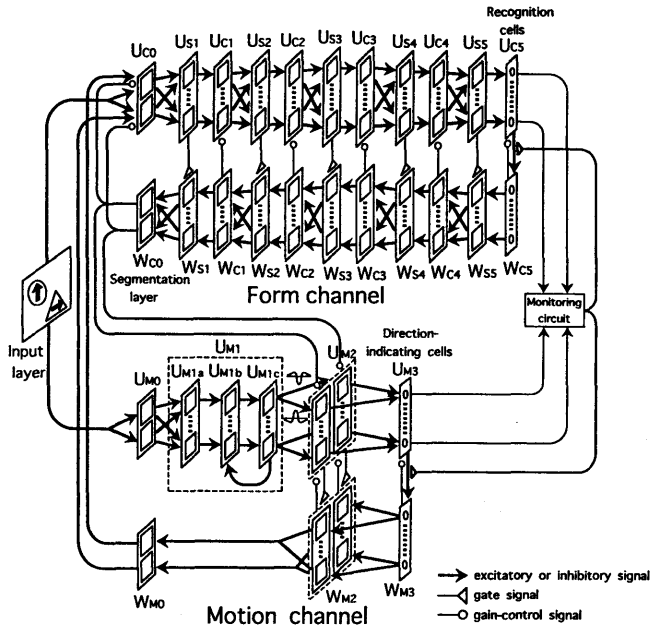


図3 形の処理系と動きの処理系をもつ視覚系の神経回路モデル。

て自己組織化させた。図4は、ランダムドットで形成されたテストパターンを与えたときのモデルの反応の一例である。テストパターンは、物体も背景も一様なランダムドットからなるが、物体の内部は特定方向に動き、背景は静止している。物体の輪郭は静止している。テストパターンの1フレームだけを見ても意味のある形は見えないが、連続したフレームを見れば動きがあるので我々は物体の形を明確に知覚できる。テストパターンは、図の左上部に示すように、右に動く菱形、上に動く三角形、および左に動く正方形を含んでいる。図には上から順に、連想出力層（形のチャンネルのバックワード系の最下位段）、認識細胞層（形のチャンネルの最上位段）、動き方向指示細胞層（動きのチャンネルの最上位段）の反応の時間経過を示した。モデルは最初に、右に動く菱形に注意を向け、その形と動きを検出した。少し時間が経って回路の反応が定常状態に落ち着くと（ $t=5$ ）、注意を自分で切り替えて、左に動く正方形を検出した。さらに $t=9$ には2回目の注意切り替えをして、こんどは上に動く三角形を検出した。モデル

はこのように、注意の切り替えによって視野内の複数の物体を順番に認識しその動きの方向を検出していく。

なおこのモデルは、視野内に同じ形の物体や同じ方向に動く物体がある場合でも正しく働く。

4. 不均一な受容野と眼球運動

ヒトの網膜は中心窩の部分が高い解像力を持っているが、少しでも中心窩をはずれると視力は急激に減少する。われわれがものを細かく識別できるのは網膜内でも中心窩付近に限られている。このために眼球を絶えず上下左右に動かして、見ようとする物体の像を中心窩付近にもってくる必要がある。

オン中心型やオフ中心型の受容野をもつ網膜神経節細胞の受容野の直径が、中心窩から離れるにつれて大きくなるように作られた網膜のモデルは、これまでもいくつか提唱されていた。しかし受容野のサイズの不均一性は、網膜だけでなく大脳の視覚野でも見られる現象である。だが大脳の視覚野の細胞の受容野の不均一性や網膜皮質間の投

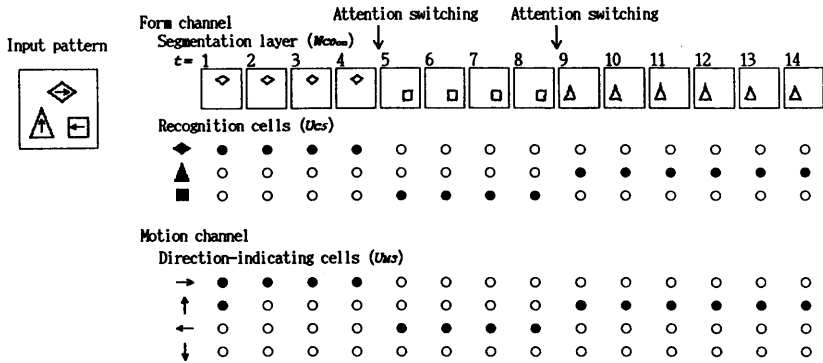


図4 動きと形の両処理系をもつモデルの反応例。ランダムドットで構成された3個の図形がそれぞれ異なる方向に動いている場合の反応の時間経過。

射の不均一性までも考慮した眼球運動のモデルは、まだあまり見られない。そこで筆者らは、図5に示すようなモデルを提唱した^{[5],[6]}。図6はコンピュータシミュレーションで求めたモデルの反応例である。

外界の光学像は網膜の視細胞層に投影され、網膜神経節細胞層でコントラスト成分が抽出される。神経節細胞の受容野は、中心窩付近では小さく網膜の周辺部では大きくなるように定めてある。次の中継細胞層は外側膝状体（あるいはその軸索終末の皮質視覚野の第4層での分布）を想定している。網膜視細胞層から中継細胞層へは不均一な投射があるので、中継細胞層では図6に示すように、注視点（刺激パターンの左目付近の*印の場所）の付近が拡大され、網膜の周辺部は圧縮されている。視覚野ではこの情報から、種々の傾きのエッジ情報（単純型細胞）や、その曲がりの情報（超複雑型細胞）を抽出している。

サッケードによる次の注視点の移動位置は、上丘を想定した細胞層（注視点決定機構）で決定する。ここには大脳皮質で抽出したエッジやその曲がりの情報が送られ、その情報密度の最も高い箇所が側抑制による競合によって選ばれ、そこを次の注視点とする。ただし、注視点が毎回同じ場所に繰り返して選ばれるのを防ぐために、過去の眼球位置を記憶する記憶ユニットを考え、注視点決定機構はその記憶ユニットからの抑制信号も受け

るようにした。

脳は、このような眼球運動によって取り込んできた断片的な情報をつなぎ合わせて一つのまとまった外界のイメージを作り上げていく。その生理学的なメカニズムはよくわかっていないが、モデルでは以下のような視覚情報統合機構を想定した。最初の注視点で取り込んだパターン（エッジ抽出細胞層などから送られてくる視覚情報）をひとまず一時的な記憶パターンとしてバッファに蓄える。その後注視点を移動させ、新しい注視点で取り込んだ情報を用いて、この記憶パターンを部分的に更新しながら解像度を上げていく。このとき、すでに蓄えられている記憶パターンと、新しく取り込まれたパターンとの位置合わせが必要になる。ところが心理学や解剖学の知見によると、視覚系には眼球位置の情報は必ずしも正確には伝えられていない。随伴発射などによって送られてくる眼球移動の情報は、大まかな位置あわせには用い得ても、細かい位置あわせに用い得るほど精度は高くないと考えられる。そこでモデルでは細かい位置合わせのために、新しく取り込んだパターンを、バッファ内の記憶パターンと比較して、類似度が最も大きくなる位置にシフトした後に、バッファを更新する機構を想定した。

5. 空間記憶のモデル

われわれは外界の空間情報を、脳内にどのよう

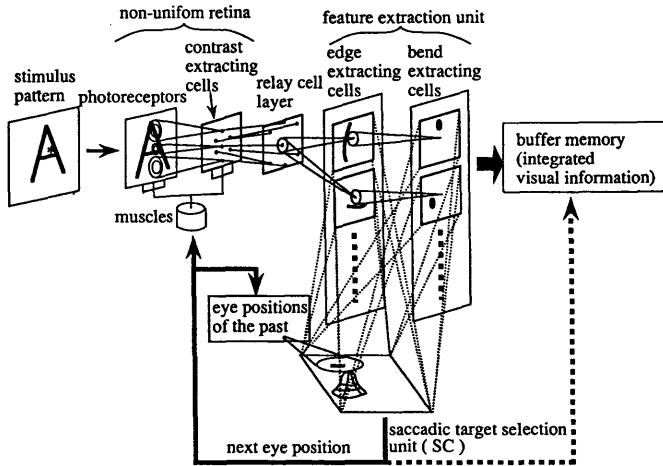


図5 不均一な網膜皮質間投射を持つ眼球運動のモデル。

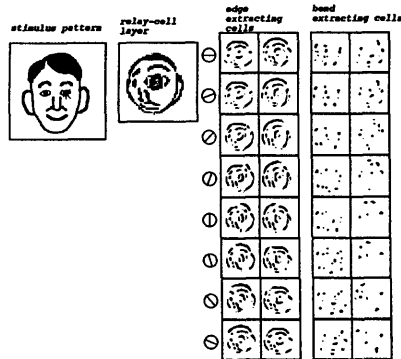


図6 眼球運動モデルの特徴抽出層の反応例。注視点の近く (*印で示す左目の付近)が拡大され、網膜周辺部は圧縮されている。

な形で表現し記憶し想起しているのであろうか。脳損傷患者による半視野無視などの症例から考えると、地図のような空間パターンは、脳の中でもそのトポロジーを保ったまま、やはり空間パターンとして表現されていると考えられる。

以前行ったことのある場所をドライブしている状況を考えてみよう。われわれは次にどのような地形や風景が現われるかを思い出すことができる。また最初に思い出した地形がもとになって、さらにその先に何があるかを思い浮かべることができる。われわれはこのように、広い範囲の地形や風

景を連鎖的に読み出してることができる。

もし外界の空間情報がトポロジーを保ったまま地図のような形で脳内に表現されているとするならば(脳内の細胞数が有限であるなどを考えると)、一定の解像度で一時に思い浮かべることのできる視覚イメージの大きさは、有限の範囲に限られているはずである。また視野の大きさは有限であるので、脳の中に一時に取り込むことのできる空間情報も限られているはずである。しかしわれわれは、隣接した場所で別々に取り込んだ情景を組み合わせ、一つのまとまったイメージを

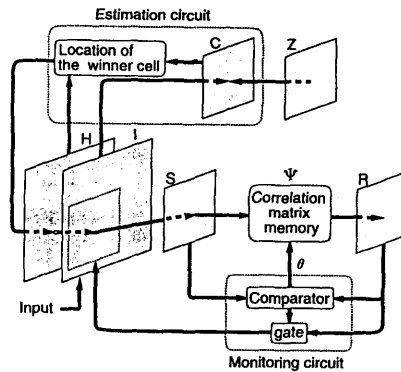


図7 空間記憶のモデル。

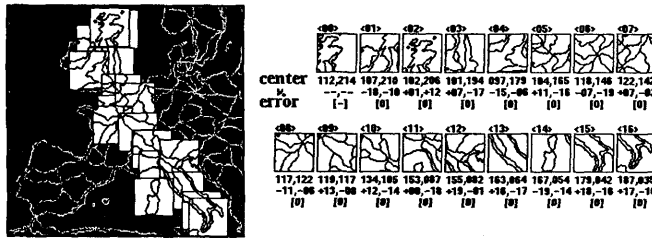


図8 連鎖的な地図の想起。

脳内に作り上げることもできる。

また行き先を告げられずにある場所につれてこられても、ひとたび見慣れた風景に出会えば、それが元になって想起の連鎖反応が始まる。このことは脳が、断片的な地図を、その位置を示す座標のような情報をつけずに記憶していることを示唆する。

そこで筆者らは、地図のような空間パターンを、断片的な形で取り込み、その記憶をもとに広範囲の空間のイメージを次々と連続して想起できるような機構の神経回路モデルを構築した^[1]。図7にモデルの構造を示す。脳の中ではどのような記憶回路が使われているのかまだよく分かっていないので、このモデルでは自己想起型の相関マトリクスを用いた。しかし相関マトリクスは刺激パターンの位置ずれを許容しないので、モデルでは、“積み重ねパターン”（記憶パターンを画素ごとに加算して作成したパターン）を補助的に用いて、

相関マトリクスに与える刺激パターンの位置合わせを行なうような機構を暫定的に導入している。

コンピュータシミュレーションでは、鉄道線路の記入されたヨーロッパの地図を、重なり合った多数の断片的な地図に分割し、これをモデルに記憶させた。その後、スコットランドからイタリアまで汽車で旅行する状況を想定し、スコットランドの地図をイメージ層（想起された空間パターンが現われる細胞層）に入力した。すると、この初期パターンが引き金になって、パリを經由しイタリア南部に至る経路の地図が、図8に示すように順次、とぎれることなく想起されてきた。

文献

[1] K. Fukushima: “Neural network model for selective attention in visual pattern recognition and associative recall”, Applied Optics, 26[23], pp. 4985-4992 (Dec. 1987).

- [2] K. Fukushima: "A neural network for visual pattern recognition", *IEEE Computer*, **21**[3], pp. 65-75 (March 1988).
- [3] K. Fukushima: "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", *Biological Cybernetics*, **36**[4], pp. 193-202 (1980).
- [4] K. Fukushima: "Neocognitron: A hierarchical neural network capable of visual pattern recognition", *Neural Networks*, **1**[2], pp. 119-130 (1988).
- [5] K. Fukushima: "Training neocognitron to recognize handwritten digits in the real world", *WCNN'96 - San Diego*, pp. 21-24 (Sept. 1996).
- [6] K. Fukushima, T. Imagawa: "Recognition and segmentation of connected characters with selective attention", *Neural Networks*, **6**[1], pp. 33-41 (1993).
- [7] 庄野, 福島: "折れ点処理回路を用いた選択的注意機構による英字筆記体連結文字列認識", *電子情報通信学会論文誌 D-II*, **J77-D-II**[5], pp. 940-950 (1994).
- [8] 芦田, 福島: "選択的注意機構のモデルによる漢字認識", *電子情報通信学会ニューロコンピューティング研究会*, No. **NC90-138** (1991).
- [9] 橋本, 福島: "選択的注意機構による顔の部分パターンの切り出し", *日本神経回路学会第6回全国大会*, pp. 274-275 (1995).
- [10] A. Treisman: "Features and Objects: The fourteenth Bartlett memorial lecture", *Quarterly J. of Experimental Psychology*, **40A**[2], pp. 201-237 (1988).
- [11] C. M. Gray, P. König, A. K. Engel, W. Singer: "Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties", *Nature*, **338**[6213], pp. 334-337 (March 1989).
- [12] S. Zeki, S. Shipp: "The functional logic of cortical connections", *Nature*, **335**[6188], pp. 311-317 (Sept. 1988).
- [13] K. Fukushima, M. Kikuchi: "Binding of form and motion by selective attention: A neural network model", *Brain Processes, Theories and Models: An International Conference in Honor of W. S. McCulloch 25 Years after His Death*, ed.: R. Moreno-Diaz, J. Mira-Mira, pp. 436-445, Cambridge MA • London: MIT Press (1996).
- [14] M. Kikuchi, K. Fukushima: "Neural network model of the visual system binding form and motion", *Neural Networks*, **9**[8], pp. 1417-1427 (Nov. 1996).
- [15] K. Fukushima: "Neural networks for selective looking", *ICONIP'94 - Seoul*, eds.: Myung-Won Kim, Soo-Young Lee, pp. 1367-1372 (Oct. 1994).
- [16] 青西亨, 福島邦彦: "網膜と皮質の不均一性を考慮した注視点移動モデル", *電子情報通信学会論文誌 D-II*, **J78-D-II**[9], pp. 1363-1371 (1995.9).
- [17] K. Fukushima: "Neural network model of spatial memory", *ICONIP'96 - Hong Kong*, Vol. 1, pp. 15-20 (Sept. 1996).

～「第10回専門用語シンポジウム：脳と用語」から～

知識獲得とオントロジー構築

西田豊明

奈良先端科学技術大学院大学

要旨

情報社会を支える知識基盤の構築を促進する人間とコンピュータの共同作業環境として、弱い情報構造を中心とするボトムアップの知識体系化の方式を提案する。この方式は、情報キャプチャ、視点変換、異なる情報ベースからの情報の統合、情報構造の整形、強構造の示唆からなる。提案のかなりの部分についてプロトタイプを作成し、予備的な評価を行った。

1. 知識基盤構築へのアプローチ

高度情報化社会では膨大な情報と格闘し、迅速な意思決定をする必要がある。迅速な意思決定を支える知識基盤として大規模知識ベースの重要性が指摘されてきた。これまで CYC[Lenat 89]など、知識基盤を直接構築する試みは行われてきたが、そこでは常識的で比較的变化しない知識を手作業を中心にコーディングするという手法が取られた。

本稿では、CYCのアプローチでは手をつけられなかった、具体性が高く、個別的・雑多で、刻々と変化していく知識に焦点をあてる。多数の不均質な情報源で生成される流動性の高い情報をゆるやかに関連付けて体系化する過程を支援する人間とコンピュータの協調的な作業環境(図1)の枠組みを提案する。

以下の議論を明確にするために、情報をタイプ

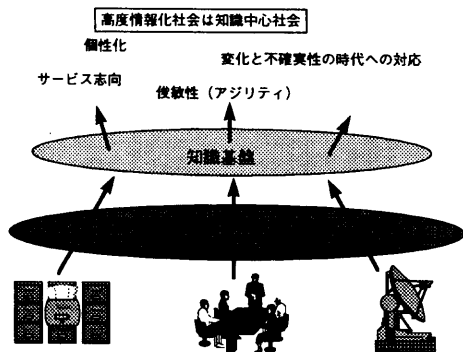


図1 知識基盤形成を促進する共同作業環

とレベルの点から図2のように整理してみた。横軸は動態による情報の分類である。フロー型情報は、電子メールや会議の議事録のように絶えず変化する状況を扱う。一方、ストック型情報は、シソーラスや書物のように固定され、体系化された視点から事態を捉えたものである。フロー型情報はメッセージとして移動するため前後の文脈に依存性が高いが、ストック型情報は逆に、文脈からの独立性が高い。

一方、横軸は情報表現の構文と意味に関する人工的な統制の程度による分類である。システムにおける情報の表現の構文と意味の統制が最も弱いものがメディアレベルである。形式レベルでは、言語の構文と意味が規定され、さらにモデルレベルでは言語の背後にあるモデルまでが明示的に規定されている。

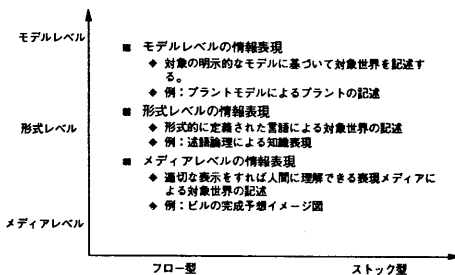


図2 情報表現のタイプとレベル

野中ら[野中 95]は知識を形式知と暗黙知に分類

しているが、本稿で提案するシステムが直接的に操作するのはすべて形式知に相当するものである。暗黙知はこうしたデータとして観測可能な情報の背後にあるものとして位置付けている。

メディアレベルの知識メディアは、暗黙知に近い部分の伝達に適している。人間に直観的でわかりやすいが、共有や流通に関してコンピュータによる支援が困難である。意図を伝達するためには、適切な文脈を与えなければならない。形式レベルの知識メディアはその対極にある。情報の論理構造の伝達には適しているが、暗黙知に近い部分の情報が欠落してしまう。エキスパートシステムの知識獲得ボトルネックに代表されるように、情報の形式化に関して少しでもユーザに負担を負わせると、ユーザの協力を得ることは非常に難しくなり、システムの扱う知識の質と量が極めて貧しいものになってしまう。両者の適度なバランスをとることが望ましい。

我々は知識の創出・体系化・流通・共有の枠組みとして知識コミュニティを提唱してきた。現在焦点を当てているのは、メディアレベルと形式レベルのストック型情報である。技術的な柱として、人間向きの情報表現とコンピュータ向き情報表現を関連付けた取り扱いを可能にするための知識メディア、自律的に行動し、情報の相互作用を引き起こすエージェント、情報ベースで参照される概念体系の明示的記述としてのオントロジーの3つを掲げている(図3)。

本稿では、メディアレベルの情報資源から情報を収集し、整理・体系化を行って形式レベルの知識体系をつくる、ボトムアップの知識・オントロジー構築過程の支援について報告する。ここで中心的な役割を果たすのは、メディアレベルの情報どうし、または形式レベルの情報との間のゆるやかな関連付けを表現する手法として連想構造である。

以下では、連想構造を用いた情報ベース CM2、

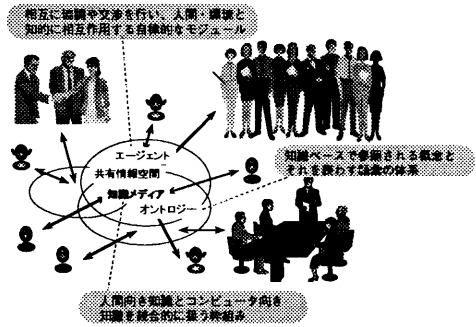


図3 知識コミュニティのアプローチ

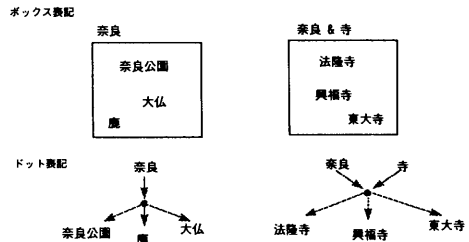


図4 連想構造

CM2 による情報の整理と体系化の支援について述べる。

2. 弱い構造を用いた情報ベースシステム CM2

CM2 (Contextual Media version 2)の中心になるのは連想構造である。連想構造は図4のように「何から何が連想されるか」を表わす関係である。例えば、図4左の連想構造は、「奈良から奈良公園と大仏と鹿が連想される」ことを表わし、図4右の連想構造は「奈良と寺から法隆寺、興福寺、東大寺が連想される」ことを表わす。「奈良」や「寺」を連想のキーと呼び、「奈良公園」「法隆寺」などを連想の値と呼ぶ。

CM2では、連想構造がどのような関係であるかをあえて明示的に定義しないようにしている。我々は、ドメインにおける意味の分節化がある程度進んだ段階で、何らかの関わりはあると認められるがその意味を明確に規定できないようなデー

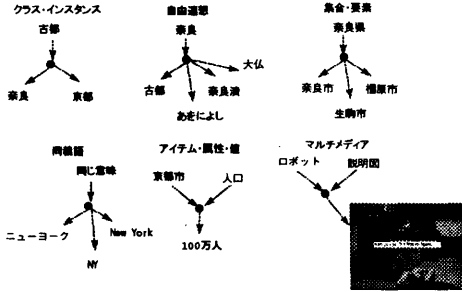


図5 連想構造による情報表現

タの存在を積極的に認め、体系化の手がかりや思考の支援として活用したい。連想構造は、そこに関係がありそうなことを書き留めておくために使用する。

従って、同じキーに対する連想の値は個人によって異なる主観的なものであり、個人々人においても状況とともに変化し得る。あくまでも連想構造は情報が構造化される中間段階である。まず、「連想」によって情報の間の関連を明示することによって関連のある情報を収集し、次にそれを分析し、体系化しようという帰納的なアプローチをとっているのである。

現実に連想構造として表記される関係のパラエティは図5のように、クラス・インスタンス、自由連想、集合・要素、同義語、アイテム・マルチメディアと概念の関連付けなどがある。

CM2の情報ベースは連想構造の集まりである。ユーザは情報ベースの一部を切り出したワークスペースを介して情報ベースを参照したり、操作したりする(図6)。CM2ではワークスペースやコンピュータの情報資源一般を連想構造のキーや値として参照できる。

CM2における情報ベースへの基本アクセスは、連想構造のパスを順にたどることである。例えば、図7では、「大仏」と「寺」が与えられたとき、両者を結ぶ連想構造の連鎖として、「大仏から東大寺が連想される」、「奈良と寺から東大寺が連

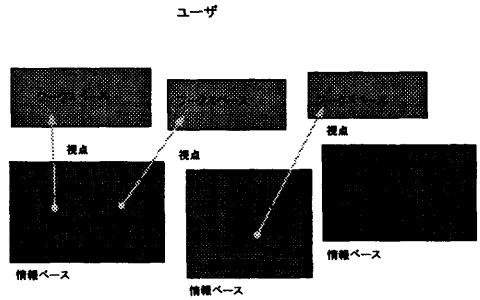


図6 CM2の情報ベース

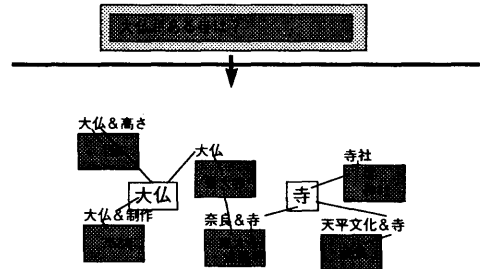


図7 CM2における連想検索

想される」をたどり、ユーザに提示する。

このように CM2 は基本機能レベルでは情報内容の理解に基づく動作をしているわけではない。CM2 が提示した情報を解釈して情報を獲得するのは人間である。一般に CM2 のアウトプットのもつ意味は受け手の背景知識によって異なる。人間に理解するための背景知識が欠如していると CM2 のアウトプットは理解できない。換言すると、CM2 自体は共通のコンセンサスの得られた意味論に基づく、規定の行動をするわけではない。

このような選択は CM2 を情報システムと考えると妥当なものではなかろうが、CM2 はそもそも意味論の定まる前の情報を扱うと考えると、発送支援システム的な意味は持ち得ると考えている。

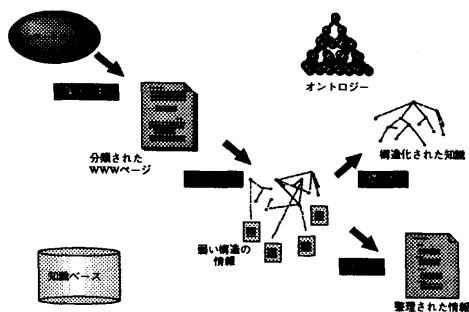


図 8 ボトムアップのオントロジー構築の支援

3. CM2 によるボトムアップな知識・オントロジー構築の支援

CM2 を用いた、インターネットなどの多様な情報源からのボトムアップな知識・オントロジー構築の支援システム (図 8) のプロトタイプのかなりの部分を構築した。CM2 は種々の情報源からの関連する情報をまとめて、構造化する過程を支援する。

CM2 による支援は

- (1) 情報キャプチャ: HTML 文書や自然言語テキストなどからの理想構造生成
- (2) 視点変換: 理想の視点を交換し、キーと値を適当に入れ替える。
- (3) 異なる情報ベースからの理想構造の統合。
- (4) 理想構造の整形: 観点の一定しないと考えられる理想構造を修正する。
- (5) 強構造の示唆。

から成る。CM2 は情報の意味は理解しないが、構造化を支援するプログラムは与えられた理想構造における情報の意味をヒューリスティックを用いて抽出することを試みる。

3.1 情報キャプチャ

完全な自然言語処理は期待できないので、形態素解析によってソースからキーワードを切り出し、HTML 文書の構造などを手がかりに理想構造を生成し、ヒューリスティックなどを用いて整形する (図 9)。

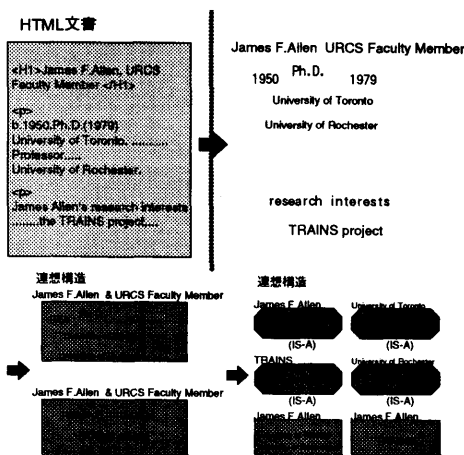


図 9 情報キャプチャ

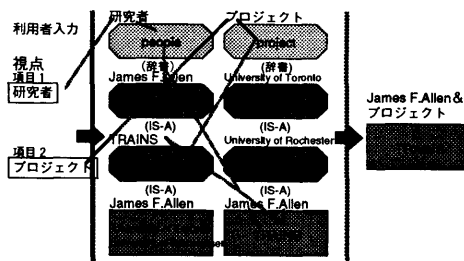


図 10 視点変換

3.2 視点変換

与えられた視点に基づいて理想構造を再構造化する。理想検索を用いて実現する。例えば、図 10 中央の理想構造は「James F. Allen」と「プロジェクト」の直接的な理想関係を明示したものではないが、この二つのキーワードによる理想検索を行うことによって、構造の中に暗黙的に含まれている理想関係を抽出し、明示している。

3.3 異なる情報ベースからの理想構造の統合

複数の情報ベースに含まれる理想構造を統合することによって、一様な視点からの包括的な理想構造の集まりを生成する。例えば、図 11 上の 3 つの理想構造を統合すると図 11 下のような理想構造を生成する。視点変換をしてから、理想構造のマージを行っている。

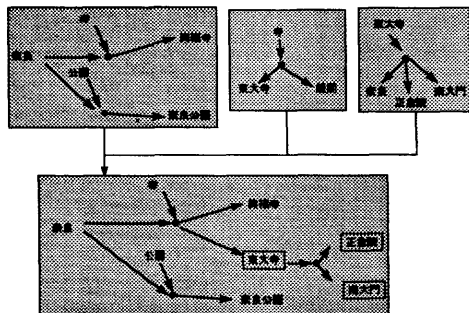


図11 連想構造の統合

連想構造の間で語彙の使い方に一貫性がない場合はあらかじめ語彙の整合を行う必要があるが、本研究ではこの問題に対する本格的な取り組みはまだ行っていない。

3.4 連想構造の整形

このフェーズでは、与えられた連想構造の集まりの観点のばらつきをヒューリスティックを用いて補正する。詳細度が異なる情報は詳細な方にあわせる方向に補正する。例えば、図12は直交分解と類推の詳細化とよぶヒューリスティックの適用例を示す。第1ステップで「生駒山上遊園地」と

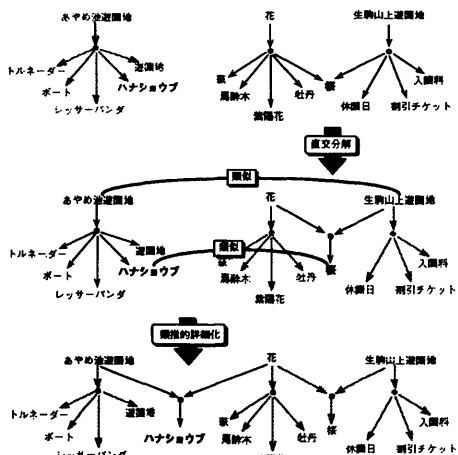


図12 ヒューリスティックを用いた連想構造の整形

「花」の両方から連想されている「桜」を認識し、それが「生駒山上遊園地」と「桜」の2つをキーとする連想構造として明示している。第2ステップで、与えられた情報ベースにおける「あやめ池遊園地」と「生駒山上遊園地」の挙動の類似性と、「ハナショウブ」と「桜」の類似性に基づいて、「ハナショウブ」を「あやめ池遊園地」と「花」をキーとする連想構造の値として推定している。

はじめは、「あやめ池遊園地」と「ハナショウブ」の間に単なる連想関係があることしか示されていないが、連想構造の整形によって両者の間の関係に関する情報が詳細化されている。

3.5 強構造の生成

以上のように収集・整理された連想構造に意味付けを行って強構造を生成する。強構造の生成は、強構造の抽出と一般化から成る。図13では、「奈良」を主要概念、「寺」や「公園」を属性としている。今のところどれを主要概念とし、どれを属性概念とするかの決定や、暗黙的な属性（例えば、「部分」）の挿入は人間が行うものとしている。

4. 現状と展望

これまでのところ、以上に述べた機能のかなりのところを実現し、予備的な評価を行っている。そのなかには CM2、CM2 のための情報キャプチャ

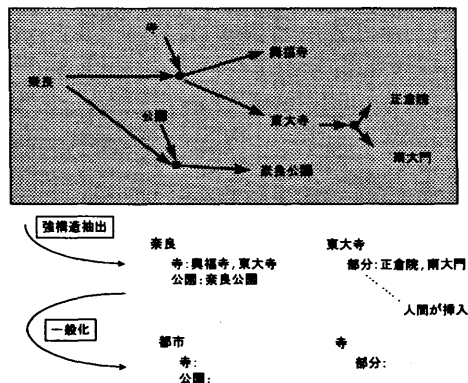


図13 強構造の生成

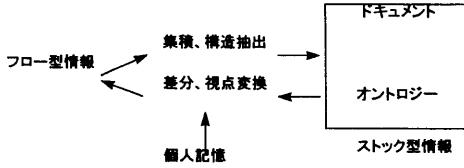


図14 フロー型情報とストック型情報の相互変換

ゃ、CM2を用いたオントロジー構築支援システムの試作を終えている。予備的な評価として、約2000個の概念を含んだ情報ベースの試作[前田96]、WWWからのオントロジー獲得支援システムの試作[糺谷95]などを行った。

意味を予め明確に規定しておかないことは、知識獲得のコストを下げる。データの質の不均質さによるデメリットがあっても、生々しい知識が集められ、明示されることのメリットの方が大きいという感触を得ている。今後本格的なシステム開発と評価に取り組んでいきたい。

今後CM2の研究を進展させていくとき、フロー型情報を取り扱えるようにすることがかなり重要度の高い拡張である(図14)。フロー型情報からストック型情報への変換は、例えば何回かに及ぶ共同作業の議事録からよく定義された共同作業のアクションプランを作成することに相当する。変更情報によって情報を書き換えることや、最後に残った情報片を構造化することなどが含まれる。

一方、ストック型情報からフロー型情報への変換は重要である。ストック型情報はえてして巨大なものになり手に負えなくなる傾向にあるので、それを要約して、情報の受け手の文脈に応じた差分だけを取り出す工夫が必要になる。ストック型情報からフロー型情報への変換は表層的にはプレゼンテーションの生成の技術として位置づけることもできる。現在図15のように連想構造を直接表示しないインタフェースや図16のようなグループコミュニケーションの支援などについて検討している。

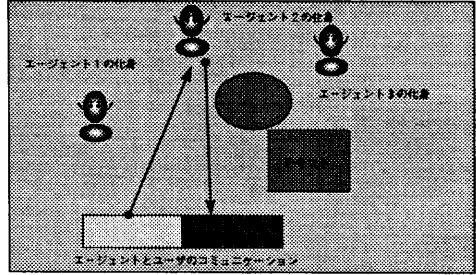


図15 CM2を利用したプレゼンテーションシステム

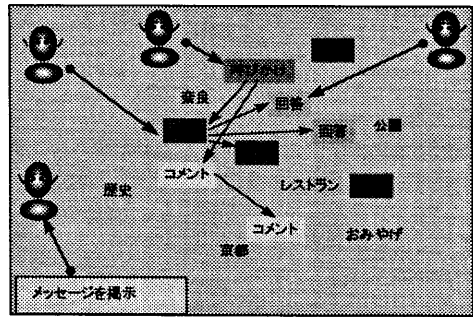


図16 CM2によるグループコミュニケーションの支援

文献

[糺谷95] 糺谷, 前田, 西田. 文脈メディアによる弱構造情報の提示と構造化. 第11回ヒューマンインタフェースシンポジウム論文集, 589-594, 1995.
 [Lenat 89] Lenat, Guha. Building Large Knowledge-based Systems, Addison-Wesley, 1989.
 [野中95] 野中, 竹内. 知識創造企業. 東洋経済新報社, 1995.
 [前田96] 前田, 糺谷, 西田. 情報ベースのユーザフレンドリなインタフェースのための連想構造の提案, Progress in Human Interface, Vol. 5, pp. 49-56, 1996.

～「第10回専門用語シンポジウム：脳と用語」から～

用語間の意味関係の自己組織化

藤原 譲* Yuzuru Fujiwara

Tel:0298-53-5532, Fax:0298-53-5206

E.mail:fujiiwara@is.tsukuba.ac.jp

要旨

情報の記録、通信の媒体として最も便利なものは言語であり、さらに専門家が思考活動において主として対象とするのはそれぞれの分野における専門的概念である。その専門概念を表現するのが専門用語である。従って知識、情報の課題はほとんど専門用語の課題ともいえる。これらの未解決課題の多くは用語の特性を解析することにより、本質が明確になるので専門用語のみならず広くマルチメディアを含め情報の高度活用の展望が得られることを示した。用語解析の応用の例として多言語概念構造自己組織化システムによる意味理解、情報生成などを述べた。

1. 序

用語は言語学や文学の問題でもあるが、計算機や通信技術の発展とともに、情報科学や通信工学の対象としても取り上げられるようになった。ここでは情報科学とくに情報を資源化し、それを適切に管理、処理し、高度利用に供するために情報の特性を解析し、情報の意味関係、論理関係、情報空間の構造などの理論を体系的に扱う分野が情報学基礎であり、ここではその立場からみた用語について述べる。

専門用語は計算機の普及、通信の高速化、多様化、およびエージェント、インターフェース技術などの進歩により、情報の利用が広がり、また高度になるに従ってその重要性が認められるようになった。また多種多量の情報の高度処理への要求が高まっているが、それに応えるには用語を通して情報の内容すなわち意味処理が必須である。

情報の記録、通信の媒体として最も便利なものは言語であり、さらに専門家が思考活動において主として対象とするのはそれぞれの分野における専門的概念である。その専門概念を表現するのが専門用語である。本研究はこれを専門用語の定義とする立場で進められている。従って知識、情報の課題はほとんど専門用語の課題ともいえる。

情報はそれ自身で実在することは少なく通常な

んらかの媒体上に記述、表現されて記録、管理、伝達、利用されるので、必然的に記述および表現の形式が媒体に依存することになる。例えば阪神大地震の被害を表現するのに媒体として文章のみまたは写真のみを用いるか文章と写真を併用するかを比較してみれば明らかである。

媒体として見ると文字に比し画像や音声は抽象化の水準が低いが、情報量が多く理解も容易である。これが最近マルチメディアへの期待が高まっている理由である。一方自然言語は簡潔に表現、伝達できるだけでなく、種々のレベルでの抽象化が行われ、人間が行う思考対象としても便利であるが、計算機にとっても形式的な取り扱いが簡単である。したがって用語は概念や事象の表現単位として重要であり、マルチメディア情報の処理に関しても抽象化は必要であり、結局用語の重要性は変わらない。

2. 用語解析と均質化 2部グラフ型モデル

用語の解析としては表現形式、意味関係、および構造、の3つの面を主たる対象とする。まず表現形式の面から専門用語の特徴をまとめると1) 定義が明確、2) 複合語が多い、3) 造語規則が単純ということになる。すなわち専門用語は明確に定義された内容を持つことが多く、その内容の規定のために複数の構成要素を有することになる。専門用語の長さは2～3語が最も多いが長いもの

* 筑波大学電子情報工学系

は10ケを超えることも珍しくはない。またこのことは一般用語に比して多義性が少ないことにもつながる。しかし当然一語からなる専門用語も多い。

次に意味に関連する用語の特性を挙げると、上で述べた媒体依存性の他にも種々のものがあり、専門用語も自然語から構成されるのが普通であるから多義であることは本質的なものである。記述、表現の多様性と多義性、論理の様相性、量的な非可算性、階層関係の多重性、概念の部分共有、入れ子構造、関係の相対性、2項関係(binary relation)のみでなく多項関係などがあり、以下にその簡単な説明を行う(1, 2, 3)。

a)多項関係 (n-ary relation)

概念間の関係はグラフで扱える2つの概念に対する2項関係だけでなく2つ以下および以上の各種の場合がありそれぞれ一項関係 (unary relation), n項関係(n-ary relation)と呼ばれる。

b)記述, 表現, 分類の多様性 (Diversity of Description, Representation and Classification) 情報の媒体が多様であるので記述, 表現の多様性があるのは、避けられないことであるが、同じ媒体であっても想像以上に様々な形態をとり得る。典型的な例は言葉でいえば同意語である。一般的に情報の記述, 表現の多様性の説明のため、単純な場合で包含関係だけとして次に説明する。

n個の属性で記述されるべき対象があったとして、記述項目の組合せは2のn乗通りとなり、表現法はn!通りとなる。なお分類の多様性も同じように説明できて分類概念がm個であるとm!通りとなることが示される。したがって属性が一つ増えると表現法も分類法もそれぞれ(n+1)倍、(m+1)倍となる。なおこれらは構造として束(lattice)になる。

c)様相性(Modality)

一般的に使われる情報は2値論理とは限らなくて、多値論理つまり「そうである」か「そうでないか」のどちらかに割り切れる場合だけでなく、「そうかもしれない」し「そうでないかもしれない」というような場合も含めた論理である。

d)非加算性 (Uncountability)

意味の関わる問題の一つは個別実体(Distinct Entities)の集合を通常仮定することである。順序関

係の成立する外延(extension)として概念を取り扱うことは対象を著しく制限することになる。

e)階層性 (Hierarchy)

情報, 概念の間には抽象化や総称表現に基づく包含関係などのため階層関係があり再帰構造も含むことになる。また、技術の進歩や生活様式の変化による新しい概念の生まれることが多く、入れ子型の構造にもなる。なお再帰構造は入れ子型とも考えられるがここでは基底集合の再帰構成のみでなく基底集合そのものに進化性に対応する内部構造を許すという意味の入れ子構造とを区別している。さらに抽象化の逆は上位概念から下位概念への属性の継承に対応し、応用面で重要である。

f)相対性, 双対性(Relativity, Duality)

実体と実体の間にある関係はそれぞれが固定されているのではなく、関係自体を実体としても扱いたいときまたはその逆に実体を関係として扱いたいときがあり、また実体と属性、階層関係における上下関係なども状況に応じて変化するので相対的である。これも従来型のシステムでは扱えない問題である。

g) 部分重複, 差分記述 (Partial Overlap, Differential Description)

h) 動態 (Dynamism)

概念の内容, 表現が時間, 場所や視点に依頼して用語の内容したがって構造が変化する。

概念の内容の少しづつの変化に対し、別々の用語があり、一般に最小の変化分毎に用語が対応するので、用語間の区別は最小変化分の記述つまり差別記述で済むことになる。

これらは概念の持つ内容を記述するためには従来用いられている外延型または個別的な方式はもとより、分類すなわち木構造およびそれを拡張した網型グラフ構造などでは不十分であることを示している。つまり用語の高度処理に関する基本課題も多いがそのことは情報の処理方式や論理体系の根本は見直しを迫っているとも云える。

結局思考機能に対応な柔軟で意味関係の充分記述可能なモデルは概念構造向きの均質化2部グラフ型モデル (CSHB: Conceptual Structure Oriented Homogenized Bipartite Model) であり (1, 2) ハイパーグラフ (4, 5) を拡張したもの

でもある。

これは脳における記憶構造に対応する構造であるが、導出したのは生理学的知見に基づくのではなく、上述のように用語の意味関係の解析の結果である。

3. 概念記憶構造のための意味記述表現

情報を利用するためには、まず関連情報を記憶装置に収録しなければならない。これが情報の資源化の第一階段である。しかも情報の内容を十分に記述・表現しておかなければならない。ところが概念には階層関係すなわち包含関係や、部分共有関係があり従来のデータベースや知識ベースの基本である外延型の記述形式では膨大な重複を生じ、また原情報が外延型でないため実際の意味内容の獲得そのものに大きな障害がある。資源化は情報の管理、利用に直結するので、そのために単に入力するだけでは充分ではない。例えば特許の文献を検索することはオンラインシステムでできるが、特許の内容検索の面から考えると、総称表現の問題がでてくる。即ち特許では出願する発明について権利の範囲を最大限拡張したいのと内容の具体的記述の洩れを防ぐため、どうしても個別的表现ではなく、できるだけ抽象的総称的な表現を用いることになり、抽象レベルの多様性のため符号照合では対応できないことになる。

さらに用語の解析で得られた他の意味関係の特性を考えると新しい均質化2部グラフモデルに基づく記憶構造で収録することになる。

4. 意味関係の構造化

実体や概念の間の様々な関係は主として用語の間の関係としてあつかうことができる。目的に応じて関係用語を網羅的に収集し、用語間の関係を、例えば同意語、多義語、階層関係、部分全体関係などをC-TRAN (Constrained Transitive Closure) (10) およびSS-KWIC (Semantically Structured Key Word Element in Terminological Context) (10, 13, 14)などを用いて抽出して用語の間の関係を扱えるようにして概念構造を自動的に作ることができる (10, 13, 14)。

情報構造の実現方法を簡単に述べると、例えば

日本語と英語の対訳用語集には英語にたいして日本語が対応関係が示してある。基本的には用語の訳は同値関係になるが、実際には人間の考えとしては同値関係の場合に上下関係も入れることが多い。それを全てが同値関係だけだとすれば、推移則が成立するので推移閉包をとり、単に全部の同値な用語を結んで同意語集合が得られる。これがC-TRANの基本である。例えばこれはJISの用語集に「耐熱性」という用語がある。「熱」の部分に対して「炎」と「火」や「燃」もあり、「耐」には「難」や「抗」があって、性質の程度を表すのに「性」と「度」がある。このような組み合わせのほとんど全てが実際に使われている。JISは勿論標準化の為に作るので用語も標準化されているが、それは専門分野別に行われるので全体としては標準化にはほど遠いということであり、多様性の典型的な例である。またSS-KWICは専門用語の大部分が複合語であり、造語規則の基本は修飾語が前で修飾される基本概念の語はその後に続くことを利用している。

これは学術用語でも同じであり、学術分野毎に用語も標準化されているが、標準化されたものが全分野に共通になっているのではなく、広く使われる概念であればあるほど多様な表現が使われている。

いろいろな用語について各種の意味関係抽出の仕方があるが、先ほどの上下関係や入れ子構造になる再帰関係がある場合には多義性によるノイズが拡大されるので、上位概念を抽出して推移閉包を求め、その結果を上位概念に結合することにより同意語集合の精度を上げること、抽出された上位概念はそれを利用して階層関係も構造化できるということ、割切簡単な方式で概念構造ができる。SS-KWICでも修飾語が複数で順序が可換の場合や接頭語、接尾語の取り扱いなど自然言語的諸要素を考慮した処理でなければならない。

同じような積み上げ方式によって論理関係とくに因果関係も自動的に収集構造化することができる。これにはSS-SANS (Semantically Specified Syntactic Analysis of Sentences) およびSANS (Semantic Analysis of Sentences)を用いる(13)。因果関係にも各種のものがあるが、自然科学で重

要なのは直接結果に結びつく原因結果関係と、いくつかの要因があって結果に結びつく要因結果の関係及び必然性が充分ではないけれども何らかの理由で結果につながる理由結果などの種類がある。これらは構造化すれば演繹推論は単なるナビゲーションとして実現でき、概念構造と併用して類推(14, 15)も実現できる。

これらの関係情報を抽出すると、概念構造として概念間の構造が組織化されるので、それには先ほどの各種の関係が含まれるわけであるが、例えば類似関係というようなことが直接扱えるようになり、情報の利用に関して非常に重要になる。また論理関係は論理構造として組織化される。更に元の情報が持っている書誌的な情報やインデックス、つまり物理的な所在情報などはシステムの扱い易い基礎的構造である。

つまり情報が持ついろいろな意味を構造化することによって、今までに述べた範囲内ではあるけれども計算機で意味が扱えるということである。

5. 概念記憶構造に基づく意味処理

以上の課題を解決すれば、利用機能の展開が可能となる。現在の計算機では四則演算や符号照合の処理、即ち数値解析、検索、演繹推論などは高速かつ高精度で処理される。より高度な予測や推定になると、完全ではないが種々の手法があり、実際に使われている。

更に高度な機能になると、類推、機能推論、仮説推論、発想、連想などと、それらを複合して問題解決、設計、意志決定、評価などをすることが要求されている。

このような高度な機能実現のためには意味とくに類似性の処理が重要であるが、情報が媒体経由の間接表現のため困難な問題である。しかし意味の関係を概念間の関係として構造の形で組織化ができれば、意味処理に道が開けることになる(1, 2, 14)。大量の情報の構造化は人手で行うことは極めて困難なことであるから、システムの、即ち自己組織的に行わなければならないし、そのような試みがなされているので以下に一つの例を示す。自己組織型情報ベースの例：

上で述べたような情報の構造化を行って実際

の研究開発に役に立つような応用システムの構築の例を示す(12, 13, 14, 15)。そのシステムはInformation-Base Systems with Self Organizing Receptor Interconnections, IBS:SORITESと名前付けられている。

要点のみを述べると、情報の持つ階層性、多項関係、相対性および部分重複などの基本特性は従来のグラフ構造型のモデルでは扱えないので、新しいモデルを構築し(1, 2, 13)、それに基づいてシステム開発を行っている。

IBSのモデルは上述の概念記憶構造に対応する均質化2部グラフ型モデルである。それに基づき検索や演繹推論のみでなく類推や帰納推論が使えるシステムが実現できる。

全体構成としては、まず網羅的に一次情報を収集、整理し、入力する。次に一次情報から概念関係、論理関係などを抽出し、それを用いて一次情報を構造化して意味処理に使うという方式である。このシステムは種々の研究用マルチメディア型情報に応用され高分子NMR、有機合成、半導体、超伝導、非線光材料(13, 14)、常温核融合等専門分野情報から電子博物館(13)、デジタルライブラリー(13)など広汎な情報までが対象となっている。なお今後これらのシステムの実用化のためには多言語自己組織化知識表現を資源化することが必要である。

6. むすび

専門用語は新しい情報化社会において一層その重要性が認められるようになった。しかし、用語の特性や意味関係を解析することにより、その構造のモデルを確立し、用語が表現する概念の内容を理解する方式が示される。したがって、専門用語の基礎を検討することは専門用語のみならず広くマルチメディアを含め情報の高度活用の展望が得られることを、先端科学技術への応用として多言語概念構造の自己組織化システムによる類推、帰納推論、仮説推論など思考支援機能を例として述べた。

参考文献

- [1] Yuzuru Fujiwara: "The Model for Self

- Structured Semantic Relationships of Information and Its Advanced Utilization". Proceedings of The 47th FID Conference and Congress, Oct. 1994.
- [2] Yuzuru Fujiwara and Hironobu Gotoda: "Representation Model for Relativity of Concepts". International Forum on Information and Documentation, vol. 20, 1, PP 22-30, January 1995.
- [3] Fugmann, Robert: "Theoretische Grundlagen der Indexierungspraxis". 1992.
- [4] C.Berge. "Hypergraphs". North-Holland, 1989
- [5] H.Boley. "Directed Recursive Labelnode Hypergraphs : A New Representation-Language". Artificial Intelligence, 9(1):49-85, 1997.
- [6] Z.Pawlak, "Rough Sets" Int. J. of Computer and Inf. Sti. Vol.11 pp341-54, 1982
- [7] Patric H. Winston, "Learning and Reasoning by Analogy" Communication of the ACM, 23(12), pp689-703, 1980
- [8] Y.Liu: "The Research and Test of Combination Schedule of Term Family with Classification in Thesaurus". Journal of Information Science, Vol.11, No2, PP 6069,1990. (in Chinese)
- [9] British Standard Institution. ROOT thesaurus, 3rd ed., MATCHING LIST. 1988.
- (Japanese version,1990).
- [10] J.Lai, H.Chen and Yuzuru Fujiwara: "Extraction of Semantic Relationships Among Term-SS-KWIC". Proceedings of The 47th FID Conference and Congress, Oct. 1994.
- [11] Peter Ingwersen: "Information Retrieval Interaction". 1992.
- [12] Hironobu Gotoda, Jianghong An, and Yuzuru Fujiwara: "Analogical Reasoning of Organic Reactions Based on the Structured Compound-Reaction Diagram". International Conference on Forum and Applied Practical Reasoning, FAPR'96, PP 276-290, Germany, June 1996.
- [13] Yuzuru Fujiwara, Ye Liu and Hanxiong Chen: "Semantic Understanding of Contents in Digital Library Based on Selforganized Knowledge". Proc. of International Seminar on Digital Library Seoul. Korea. p107-114. 1996.
- [14] Yuzuru Fujiwara and Jianghong An: "Semantic Structures of Chemical Data for Problem Solbing Systems." Proc. of 15 CODATA Tsukuba. Japan. Sept.30 - Oct.2. 1996 to be published.
- [15] Hironobu Gotoda and Yuzuru Fujiwara: "An Evolutional and Conceptual Structure Based Method of Analogical Reasoning".

Self Organization of Semantic Relationships among Terms

Yuzuru Fujiwara

Abstract

Natural languages are the most convenient as both recording media of information and communication media, and domain specialists mainly have concerns with scientific and technical concepts, which are represented by special terms. Therefore, most issues of special knowledge and information are regarded as those of special terms. Those issues to be solved are necessarily related to properties of terms and terminological analysis is used not only to understand special terms but also to develop advanced utilization of general information including multi media information. The multilingual self organization system of conceptual structures for semantic understanding and information generation, is shown as an example of direct application of terminological analysis.

～「第10回専門用語シンポジウム：脳と用語」から～

情報流通と分類表の役割およびその体系

石川徹也・ ISHIKAWA, Tetsuya

1. はじめに

現代における情報流通媒体を大別すると、① 図書、雑誌、CD-ROM等のスタンド・アロン型媒体と、② データベース、WebSite等のネットワーク型媒体に分けられる。

スタンド・アロン型媒体の主に排架整理に、例えば図書館間における排架分類の標準化を目的に、“媒体内容(情報)の主題と媒体形態”の組み合わせによって表現する分類標数が用いられている。この分類標数は、主に学問体系の十進細分化からなり、分類表に納められている。しかし、分類表は、図書館間の排架分類の標準化ツールと言え、国あるいは分野別に複数作成され、それぞれが利用されている(例：表1)。この中において、UDCが国際標準分類表となっているが、UDCは主にヨーロッパにおいて発展してきたため、例えばアメリカ、中国、日本等では主に図書館における排架分類として実際に利用されていない。

一方、ネットワーク型媒体のWebSiteそのものの存在はURLで表示されているが、インターネットの特性上無限に存在することから、それらを検索するために、現在では①Yahooに代表される情報検索サービス業^{註1)}および②検索ロボット(検索サービス^{註1)}・ソフト：例；Taitan, Altavista, Opentext等)の2つの“インターネッ

ト検索システム”が出現している。しかし、両者の各システムの検索利用において下記問題点が存在している。

i) 情報検索サービス業の提示しているURL分類リストは、情報検索サービス業において独自に設定しており(例：図1, 2)、スタンド・アロン型媒体の排架分類のように、標準分類になっていないことから、実際の検索利用において、URL分類リスト内の各分類範疇が明確でなく、実際の検索において不便をきたす。

ii) 一方、検索ロボットの利用においては、従来のデータベース検索システム利用における検索指示同様、検索語指示方式であることから、検索語の表記言語とWebSite内の情報記述言語が一致しなければ、検索の対象とならない問題点がある。この為、例えば日英同義語検索において、それぞれに対応する検索ロボットを立ち上げ利用するといった不便さがある。

そこで、本発表において、i)の問題解決に、従来から用いられてきている排架分類のための分類表の利用について提案し、ii)の問題解決に、検索ロボットに検索語翻訳システムの機能化について提案し、当問題解決に関連する筆者において現在進めている研究について紹介する。

2. 考察：国際間情報流通の要件

2. 1. 前提：問題点の認識

国際間情報流通の問題点について、事例として筆者がこの夏(1996年)得た2件の知見を下記に示す。

i) EU15カ国内において、現在45種の言語が使われていると言う。EC議会における公用語は通常、英語、仏語、独語の3ヶ国語であるが、公文書は45ヶ国語に翻訳しなければならない場

表1 分類表の例

国際標準分類表	国際十進分類法(UDC)
日本	日本十進分類法(NDC)
中国	中国図書館・図書分類法(CLC)
アメリカ	議会図書館分類法(LCC)
	デューイ十進分類法(DDC)

* 図書館情報大学 教授

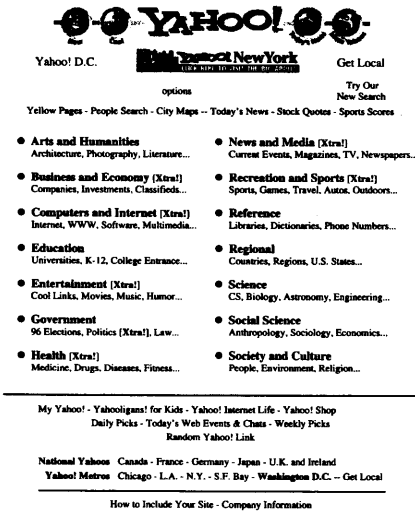


図1 インターネット情報検索サービス業のURL分類リスト例 (USA)

合があると言う。この結果、翻訳経費の削減、正確性、その処理の即時化の為に、現在、6ヶ国語対応機械翻訳システム：SYSTLANを利用しているが、現在尚、EC事務局に約1,100人の翻訳者がいると言う²⁾。

ii) WebSiteへの情報掲載に、国外から日本に対して、英語記述が強く求められているが、例えば、EUの各国に代表されるように、多言語の国においてはWebSiteへの情報掲載は、何ヶ国語で表現されているのだろうか。残念ながらWebSiteの記述言語種別を自動的に調べる方法は見当たらないが、聞くところによると、①多言語となると作成が面倒、②ファイル容量が大となる、③現時点におけるインターネット利用者の多くは英語を解する、ことからほとんどが英語記述のみとなっていると言う。この結果、インターネット利用情報流通によって、民族のアイデンティティの根幹をなす民族固有の言語が消滅する危険があり、さらにこのころから民族のアイデンティティ自身が崩壊する危険があるという議論が起こっていると言う³⁾。

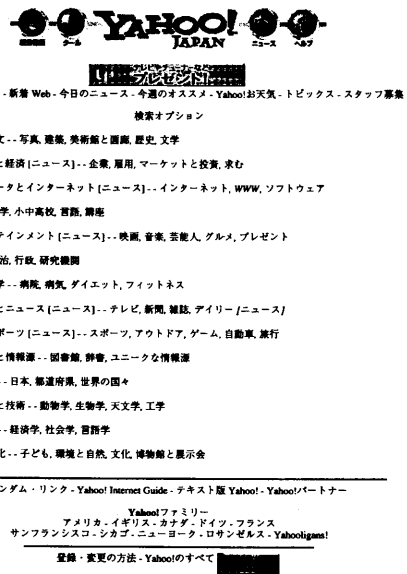


図2 インターネット情報検索サービス業のURL分類リスト例 (Japan)

2. 2. 考察：国際間情報流通の要件

以上の例から、国際間情報流通の要件として、下記のことが言えよう。

i) 情報記述は各国の公用語で発信することでよい(やむおう得ない)。

この結果、

ii) 多言語対応情報検索システム (Cross-Lingual IR System) が必要となる⁴⁾。

iii) 検索結果 (多言語情報) を解読するために多言語対応機械翻訳システムが必要になる。

3. インターネット情報検索システムに対する多言語対応化の提案および関連する筆者における研究

3. 1. 情報検索サービス業システムへのUDC利用提案および関連する筆者における研究

i) 情報検索サービス業システムへのUDC利用提案

表2は、現時点での代表的なインターネット情報検索サービス業Yahooの最初のURL分類リストに対し、筆者においてNDC、UDCおよびCLCの各分類標数を付けたものである。これを見る限

表2 インターネット情報検索サービス業のURL分類リスト例と分類標数付与例

ホームページ案内系		ロボット系	分類標数		
Yahoo			NDC ¹	UDC ²	CLC ³
Japan	U.S.A.				
芸術と人文	Art and Humanities	700,200	7+82+9	J	
ビジネスと経済	Business and Economy	330	33	F	
コンピュータとインターネット	Computers and Internet	007.6	681.3+654	TP311.13	
教育	Education	370	37	G4	
エンターテインメント	Entertainment	760,770	7	J6,J9	
政治	Government	310	32	D	
健康と医学	Health	490	61	R	
メディアとニュース	News and Media	699,070,550	070	G2	
趣味とスポーツ	Recreation and Sports	980,790	79	G8	
各種資料と情報源	Reference	010	01+02	Z	
地域情報	Regional	290	91	K9	
自然科学と技術	Science	400,500	5+62	T,Q,P	
社会科学	Social Science	300,801	3	C,F	
生活と文化	Society and Culture	360,069	06+64	G	

1. 日本十進分類法標数
2. 国際十進分類法標数
3. 中国図書館・図書分類法標数

り、例えばUDC分類標数においては、少なくとも5件の分類名に対して、異分野のしかも複数の分類標数が付いていることが分かる。このことはYahooのURL分類リストでは、1つの分類名内に異分野のしかも複数のWebSiteが割り振られていることになり、利用者にとって不便であることが分かる。

インターネット情報検索サービス業のURL分類リストは、ネットワーク型媒体に対して、スタンド・アロン型媒体の排架分類と同様の閲覧ツールの役割を果たしていることになり、この意味から分類体系は排架分類と同一であることが望ましいと考える。故に、ここに情報検索サービス業のURL分類リスト分類に、国際的標準分類であるUDCの利用を提案する。但し、URL分類名は、分類標数とその分類標数に対応する分類名辞を掲示する必要があると考える。

ii) 筆者における関連する研究

国および分野によって、固有の分類表を利用していることに対して、スタンド・アロン型媒体の排架分類を、UDCに切り替えることはほとんど不可能に近い。対して、情報検索サービス業のURL分類リストが、筆者の提案にてUDCに切り替わったとするならば、利用者は両分類表を利用することになり、不便をきたす。

現在、多くの図書館において蔵書データベース

が構築され、検索利用に供されている。この蔵書データベースを世界規模で検索可能にするには、検索指示語の多言語翻訳システムおよび各分類表の分類標数間の自動リンク・システムが必要になる。

当観点から、現在、筆者において、NDCとCLCの相互参照検索システムを構築している(図3、4)。しかし、この種のシステムの実現には、下記の様な問題点がある。

●分類標数対応名辞表現の問題点…図5に示すように(UDCの例)、従来の分類表は①主に図書館職員が利用することを前提に作成されてきたため、また、②印刷冊子として大部とならないようにするため、各分類標数の分類範囲を説明する分類名辞は、i) 連接省略表示、この結果、ii) 多くは句表記となっているため、分類名辞を直接検索処理するに困難さがある。この問題点解消に、分類名辞のソーラス表現(主に同義語列記)化が必要であり、例えば、UDCにおいては検討されだしている。

3. 2. 検索ロボットへの検索語多言語翻訳の機能化提案および関連する筆者における研究

i) 検索ロボットへの検索語多言語翻訳の機能化提案

前述のようにWebSite内の情報記述言語は、必

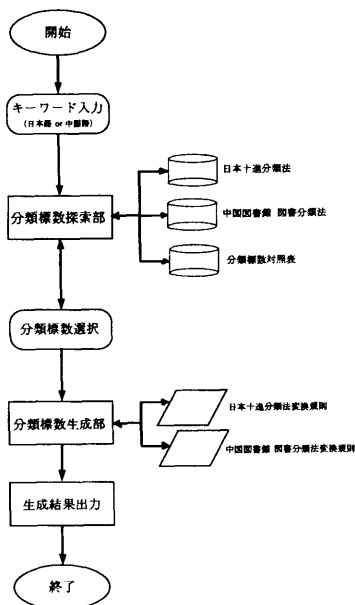


図3 NDC・CLC相互参照
検索システム機能概要

**** 日本十進分類法 ****

日本十進分類法:名辞-> 情報理論

```
[Total 2]
< 1 >
[NDC] 007.1 : 情報理論
[CLC] 6350-1 : 情報学理論

< 2 >
[NDC] 007.1 : 情報理論
[CLC] 6350-1 : 情報学理論

選択してください。 : □
```

**** 中国图书馆 图书分类法 ****

中国图书馆 图书分类法:名辞-> 情报学理论

```
{ 1 }
[CLC] 6350-1 : 情报学理论
[NDC] 007.1 : 情报理论
```

HIT RETURN KEY![]

図4 NDC・CLC相互検索結果例

```
+ 656.021.22
+ 656.021.24
! 656.11
```

Low density
High density. Traffic congestion
Highways, roads, bridges and their use. Organisation of road traffic
-361.811, 625.7/8

Example(s) of combination:

656.11:625.711.1 Organisation of traffic on trunk roads.
Arterial traffic

656.11.021.24 Road traffic congestion

656.11.021.24:159.942.6 Strong emotions caused by road
traffic congestion. Including: Anger.
'Road rage'

```
! 656.61
```

Transport by sea. Maritime navigation, shipping

Example(s) of combination:

656.61.052 Control of vessels. Navigation. Manoeuvres

656.61.052:561.326 Navigation etc through floating ice

656.61.052:681.786.3 Nautical sextants

656.61.091.7 Sea-letters (sea-peases)

図5 UDC分類名辞記述例 (BSI版)

ずしも英語によるとは限らない。基本的には各国言語で記述される。故に、フルテキスト検索型の検索ロボットにおいては、検索指示語を多言語に自動的に翻訳する機能化が必要になる。そこで、各種検索ロボットに検索語多言語翻訳システムの導入を提案する。

ii) 筆者における関連する研究

我々日本人からするならば、日本語で検索指示を行い、同義の検索指示語が少なくとも英語に自動的に翻訳され、日英記述のWebSiteが自動的

に検索できることが望まれる。

そこで、現在、検索語の対象となる複合語の日英/英日翻訳システムを構築している(図6)^{1,2)}。翻訳率は現在のところ、日英/英日翻訳とも約70%である。

しかし、この種のシステムの実現には、下記の様な問題点がある。

●専門語の正規表記問題——例えば、当『専門用語研究会』の「専門用語」は、英語の“Technical Term”の直訳と考える。しかし、現在、

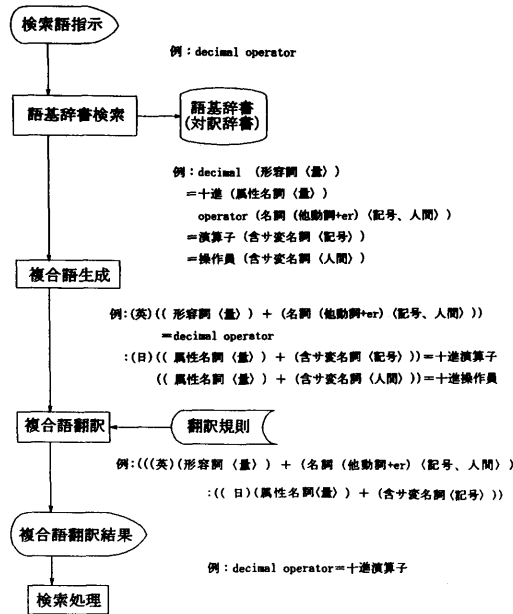


図6 複合語生成・番訳システム機能概要

日本で出版・利用されている中型国語辞典6点を見る限り“専門用語”の見出し語はなく、全て“専門語”なる語が見出し語として使用されている²⁾。このことから、“Technical Term”に対する日本語は“専門語”が正規表現であると理解できる。このように、特に専門語には異表記問題があり、検索指示語の異表記翻訳およびテキスト内の異表記に対する照合機能について、いかに対処したらよいかの問題点を残している。

4. おわりに

インターネットの普及に伴い情報の流通・検索・利用に改めて多言語対応問題を考える必要がある。当課題解決のために機械翻訳システムへの期待は大であるが、一方、情報検索システムのインターフェイス機能として、従来から図書館において利用されてきた分類表の利用活用を、また検索指示語の翻訳機能の実現について考える必要があるものとする。

「Terminology」の目的の1つに、特に専門語の標準表記の設定およびその普及があると考えられる。

本発表で取り上げた2件の情報検索システムのインターフェイス機能の課題解決に異表記辞書データの利用も考えられるが、一方において当「Terminology」の成果が待たれる。

注1) 両者呼称は、『日経産業新聞』1996年5月23日の記事内記述によるが、現時点において未だフィックスしていないと思う。

注2) TKE '96におけるEC事務局D. G. X III部のSoler, J. 氏の発表による。詳しくは、下記資料3を参照されたい。なお、言語の種類についてSoler氏に問い合わせたところ、下記資料4の呈示をいただいた。但し、当資料は1994年版であり更新版は未刊とのこと。pp.33-37の「The legal of the languages of Europe」の表には40種が示されている。

注3) TKE '96およびACM-SIGIR '96におけるコーヒ・ブレイク時に参加者複数人に聴取した結果。

注4) 当課題に対して、例えばACM-SIGIR '96において、'98と比較して多くの研究発表がなされ、

またワークショップが持たれた。

参考文献

- 1) 三日市紀子, 石川徹也: 電子図書館における専門用語の課題—検索インターフェイスとしての複合語生成・翻訳システム—, 専門用語研究, No.9, pp.3-9 (1995).
- 2) 石川徹也: 中型国語辞典の専門用語と専門用

語利用研究, 日本語学, Vol.15, pp.42-51 (1996年12月号)。

- 3) Commission of The European Communities: The Multilingual Information Society, COM(95)486 final, 34p.+i (8.11.1995)
- 4) Hearn, P.M. & Button, D.F., ed.: Language Industries Atlas, IOS Press, 406p. (1994).

～「第10回専門用語シンポジウム：脳と用語」から～

『分類語彙表』の構築

中野 洋* NAKANO, Hiroshi

0. はじめに

標題は、このシンポジウムの係りの方から頂戴したものである。しかし、この題で話ができるのは、『分類語彙表』編者の林大しかいないと思われる。発表者は、1981年以来その増補作業に参加した者にすぎない。そこで、本発表では『分類語彙表』の「まえがき」を読みながら、『分類語彙表』がどのようなものであるかを解説し、また現在、増補作業において種々の問題がどう処理されたかを概略報告することにする。

1. 『分類語彙表』概略

日本語のソソーラスとして、もっともよく知られているのは、国立国語研究所『分類語彙表』である。1964年の発表以来、言語研究のためにも広く使われており、言語情報処理でもよく利用されている。言語研究での利用状況については宮島・小沼(1992; 1994)を参照。宮島・小沼(1994)では、各種の研究136編が紹介されている。

『分類語彙表』のまえがきによれば、収録語数はおよそ3万2千6百である。これは国立国語研究所報告21『現代雑誌九十種の用語用字』第一分冊の語彙表に掲げる使用率の高い語や、阪本一郎氏の『教育基本語彙』などによって、日常生活で基本的な役割をはたしていると思われる語を選んだものである。この数は、一般の小型国語辞典の約半分で、実用にも、研究用にも、あるいはコンピュータによる日本語処理用にも、収録語数が少なすぎる。そこで原編者の林大氏を助ける形で、国立国語研究所所員を中心とする何人かの研究者たちが科学研究費を受けて増補・改定の作業をつづけてきた。この作業は、まだ進行中である。

* 国立国語研究所

平成元年以降は、以下の手順で進めてきた。現在は下記⑤～⑥の段階である。

- ①方針決定 体系の大きな変更は行わない。
- ②候補語の選択 複合語・新語・多義語・慣用語・オノマトペ・専門語等、全体に語を増やす。
- ③仮番号付け、段落および段落内の位置決定
- ④項目内の調整 語の追加削除・項目内の類義関係の修正・それにとりまなう新段落設定
- ⑤増補版を公開し、広く意見を聞く。
- ⑥項目間の調整 品詞分類間の調整ほか
- ⑦全体の見直し 機会あるごとに行う
- ⑧表記などの統一 ただし表記の基準を示すものとはしない。
- ⑨公刊

2. 『分類語彙表』の「まえがき」を読む

以下、「まえがき」を引用、編集し、点線で囲って増補作業での処置や解説を載せる。

2.1 分類語彙表の意味

同義類義の単語の目録には、次のような役割がある。

第一に、表現辞典、詞藻辞典としての役割

第二に、方言の分布や命名の変遷を知る手がかりとしての役割

第三に、表現上の特色を見る物指しとしての役割

第四に、基本語彙設定のため、常用の基本的な漢字の選定のための基礎データとしての役割である。

分類を妥当に組織だて、かつ分類項目に系統的に番号を与えることは、翻訳や情報処理たとえば文献要約などの機械化の一つの方法に、基礎的な作

業となるものである。

2.2 この分類語彙表の性質

(この分類語彙表の分類の仕方は), 一通り語彙の全分野を見渡す足場の役目は, 果たし得るものとする。

つまり, 「分類語彙表」は語彙を分類した表というより, ひとまずは「語彙の全分野を見渡す足場」の役割を, ついには意味地図の作成を目指したものだ。地図の中に配置されるべき語は無数にあらうが, いまの所「分類語彙表」の語はそれらの例である。

2.2.1 分類の大体を説明するならば, まず大分類として,

1. 名詞の仲間- 体の類
 2. 動詞の仲間- 用の類
 3. 形容詞の仲間- 相の類
形容詞, いわゆる形容動詞, 連体詞, ある種の副詞
 4. その他の仲間
接続詞の類, 感動詞の類, ある種の副詞の類等
- の4類に分ける。

「分類語彙表」が, 最初に品詞分類されていることは語彙研究者には幸いであった。品詞分布によっても, 意味分布によっても検索, 分析が可能だからである。

このようにして, 「分類語彙表」を語彙研究に用いた論文は数多く生まれたのである。

(1) 品詞論的分類での問題

このような, 品詞論的分類をまず第一に行なうことをしないで, 《美》と《美しい》, 《動き》と《動く》をそれぞれ一つの項目の中に合わせるやうかたも, 十分考えられることではあるが, 実際上は困難であった。

ただし, それぞれの仲間, 特に(1.)(2.)(3.)のそれぞれの細分にあたっては, なるべく平行するよ

うにし, 細分の番号をある程度一致させるように努めた。

小数点以下は品詞論的な4類の間で, 相互参照に便利するように内容上関係のある項目に, なるべく同じ番号を(少なくとも小数点以下第2位まで)与えるようにした。

名詞の仲間と動詞の仲間間で, 相互参照に便利するように下位までの調整を行っている。

(2) 名詞と形容動詞語幹

形容動詞と名詞との間では, その所属の決定に多少の動揺のあるものもある。ことに, 時に関することばでは, (1.)と(3.)とにわざと重出させた場合が多い。

増補作業では, 重出を避けることはしていない。増補版では, ある語がひとつの項目だけに配置されたのは57,842語である。複数の項目に配置された項目数と語数は, それぞれ2-10,214, 3-2,121, 4-468, 5-134, 6-47, 7-9, 8-11, 9-5, 10-2, 11-2, 14-1, 15-1, 18-1である。ちなみに18の項目に配置された語は「分」である。

(3) 名詞と複合サ変動詞語幹

今は作業上やむをえず, もとのまま, <変化>という名詞は(1.)の仲間, <変わる>という動詞は(2.)の仲間に入れてある。しかし, これらは, (1.1501)と(2.1501)というように細分の番号を共通にするから, 相互に検索することはたやすい。

動詞の仲間複合サ変動詞語幹を含めた。現在, 動詞の仲間は元版の4.53倍となっている。

(4) 形容詞・形容動詞の, 語尾<さ>を伴ったものは, 普通に名詞として扱われるが, ここで

は、〈高さ〉〈長さ〉など少数の慣用例を除いて、(1.)の名詞の仲間に掲げなかった。すなわち、〈快さ〉〈長たらしさ〉等は、(3.)の形容詞等の仲間に収める〈快い〉〈長たらしい〉等の一種の活用形として取扱ったことになる。

現在もこの方針にしたがっている。

2.2.2 各類の細分は次のようになる。(1.)の仲間は、

まず、人間活動に関するものと、そのあり方のわく組みに関するものとに大別される。

わく組みに関するもの(1.)は、《事実、存在、状態、力、変化、空間、時間、形、数量》

人間活動に関するものは、

活動の主体であるもの(2.)と

人間活動そのものの様相—精神と行為(3.)と、人間活動の相手として存在するもの、になるが、人間が直接に活動の結果として作り出した物および作り出すために利用する物(4.)

人間の主体的活動からは比較的自由に、外界として存在するもの(5.)となる。

(2.)の仲間、(3.)の仲間についても、(1.)(3.)(5.)に相当するものがある。

これを整理して示すと、次のようになる。

- 1.1 2.1 3.1 抽象的關係（人間や自然のあり方のわく組み）
- 1.2 — — 人間活動の主体
- 1.3 2.3 3.3 人間活動—精神および行為
- 1.4 — — 人間活動の生産物—結果および用具
- 1.5 2.5 3.5 自然—自然物および自然現象

(1) 各部門の間の区別は、常識的には比較的明

らかであると思われる。

(2) 語の譬喩的用法をもつもの

〈爆発、冷却〉：原義が支えになっているものとして、自然の部門

〈やく〉：嫉妬の場合と火の場合とで、別々に掲げておくほうが便利かのような問題は、この分類表で常に明白に処理されているとは言いがたい。

(3) 各部門が、さらにどのように細かく分けられているかは、分類項目の一覧によって概観されることを望む。

番号を与えた項目の数は、4類を通計して、798項になる。

(1.)の類545項、(2.)の類145項、(3.)の類85項、(4.)の類23項。そのうち(2.)(3.)の類の項目は、内容上ほぼ(1.)の類の項目に重なる。

(4) その分類と各項目の排列とについては、原理的に述べるのが至極困難である。各項目には細分の番号がつけてあるが、数多い項目をただ一つの順序にならべることは、全く検索の便宜によるものである。

ただ、互いに関連する項目は相接して排列されるが、その一般的総括的な内容を持つ項目は、部分的な内容を持つ項目よりも先にあげてある。たとえば(1.57)は身体に関する部分であるが、(1.570)は、(1.571~8)の身体各部位に関する総記の部分となす。

(5) なお、従前の研究所報告(4,13)(注2)における分類語彙表では、一つの語形は一つの項目にしか掲げなかった。すなわち、一つの語については一つの意味しか取りあげなかったのであるが、このたびの分類語彙表では、ある程度、多義を考慮して、一つの語でもいくつかの項目に掲げたものがある。

基本語の多くは多義的である。その意味のそれぞれには、当然別の分類番号が与えられなければならない。元版でも「月」は(1.1632 月)と(1.521 天体)とにある。「あがる」は、移動の意味では(2.1540 上がり下がり)に、終わる意味では(2.1502 開始・終了)に、食べる意味では(2.333 生活・衣食住)に分類されている。しかし、全体としては元版における多義語の処理は、不十分だった。今回の増補版ではこの点を改良して、同じ単語を意味に応じて何箇所にも出すようにした。体の部分

名称のうち、(1.5710 頭・目鼻・顔)と(1.5720 胸・腹・背)の例をあげよう。←の左側は元版にあるもの、右が付けたした分類番号である。

- 頭 (1.5710 頭・目鼻・顔) ← (1.1404 能力)
 顔 (1.5710 頭・目鼻・顔) ← (1.3030 表情), (1.3422 威厳・行儀・人望など)
 目 (1.5710 頭・目鼻・顔), (1.1850 模様・目), (1.1960 単位)
 ← (1.1303 調子), (1.3060 知・知識・意見), (1.3063 比較・区別・選択・参考)
 鼻 (1.5710 頭・目鼻・顔) ← (1.5770 血・乳・涙・汗・尿など)
 のど (1.5710 頭・目鼻・顔) ← (1.3031 声)
 首 (1.5710 頭・目鼻・顔) ← (1.3630 人事)
 背 (1.5720 胸・腹・背) ← (1.1760 前後・間・端), (1.1920 長短・広狭)
 胸 (1.5720 胸・腹・背) ← (1.3000 心)
 胴 (1.5720 胸・腹・背) ← (1.4230 下着・羽織・ズボン・コートなど)
 腹 (1.5720 胸・腹・背) ← (1.3000 心), (1.3045 意志)
 ウエスト (1.5720 胸・腹・背) ← (1.1920 長短・広狭)
 しり (1.5720 胸・腹・背) ← (1.1760 前後・間・端)
 しっぽ (1.5720 胸・腹・背) ← (1.1651 終始), (1.1760 前後・間・端)

複数の番号を付けたものの中には、多義というよりも、一つの意味を別々の観点から眺めたために違った番号が付けられたものもある。いずれにせよ、この処置もまた、『分類語彙表』の利用価値を高めるはずである。

3. 一つの項目中に収めた語の性質について

いわゆる待遇(敬意の方向と量)や、美醜、新古、詩性などの語感をぬきにして、同じ文脈の同じ地位(また転換された文脈の相当する地位)に現われることが許される限りのものである。

(1) 一つの項目に収めたのは同義類義の語の群であって、自由連想による語群ではない

(2) 日常生活でより基本的な役割をはたしている語は、大体において取りあげてある

所属項目の明白な語は、よほど重要と認めない限り、割愛したものがある。

いくつかの語が組になっているような場合に一

増補のための候補語の選択 今回の増補では、β単位のような短い語だけでなく複合語・慣用句など長い語も入れた。また、基本語彙だけでなく新語・専門語・オノマトペ等、全体に語を増やした。さらに基本語彙での多義も考慮し複数の項目に掲げた。

つか二つだけをあげて他を加えなかった。

比較的専門特殊に属すると認める語も多くを捨てた。

(3) 各項目に必要な語句を補充すること、それらの語句の必要さを他の語句と比べること、それが問題である。

語彙調査の結果から取り入れた語は、いわゆるβ単位(注1)であって、比較的短い語形である。

阪本語彙から取り入れた語や、主観的な増補による語は、必ずしもβ単位ではなく、β複合も含まれている。

この分類語彙表では、〈自転〉と〈車(シャ)〉を収めることはもちろんであるが、〈自転車〉もまた阪本語彙から補ってのせた。また比較的少数ではあるが、〈しりもちをつく〉というような慣用語句も補うことにした。

普通の表現単位をある程度採用しておくほうが、実用的な目的にかなうと考えた。

以下、←の左側が元版にある単語、右にあるのが今回増補した同義・類義の慣用語である。

ほれる ← 思いをかける 気がある

こがれる ← 胸をこがす

親しむ ← 気が合う 気を許す うまが合う 情が移る

なつかしむ なつかしがる ← 後ろ髪を引かれる

あまえる あまられたる ← 鼻をならす

きらう いたう うとんじる ← むしずが走る 嘔吐をもよおす

いやがる ← いやになる いや気がさす 気を悪くする 気(癪)にさわる

にくむ ← 牙をむく 修羅をもやす

ねたむ そねむ やく ← 角をはやす

尊ぶ ← 頭をさげる 一目おく

みくびる ← 尻目にかける 一笑に付する 眼中におかない 歯牙にもかけない

このことによって、表現辞典としての『分類語彙表』の利用価値が高まるものと期待される。

(4) 見出しは、努めてその項目の全体を示すような語を選んだが、いくつかの語を併記した場合も多く、また説明的な語句を用いた場合もある。

小数点以下第1位の分類には「抽象的關係」などの見出しがある。小数点以下第2位の分類にも見出しがほしいという声があるが、まだついていない。

(5) 項目の大きさ、すなわち収容語数は、必ずしも平均していない。しいて独立の項目に立てなくてもよさそうなものも、認められないわけではないが、分類表作成の過程での混乱を避けるため、

修正を最小限度に止めた。

たとえば用の類は約5倍に増補された。ひとつの項目の収容語数が多いと検索しにくくなる。また、類間の調整により語の移動もありえる。現在検討中である。

一つの項目の内部でも、さらにいくつかのグループが認められるものは、グループごとに行を変え、さらに番号をつけたものもある。そのグループの大きさ、またグループ分けのレベルについても、項目の場合と同様に、過不及がある。

分類語彙表は、人間用である。その分類は次のようになっている。たとえば、「におい(1.504)」の項目には、次の語例がある。

香(か・におい) かおり 芳香 香氣 臭み

臭気 異臭 悪臭 残り香 移り香

体臭 口臭 俗臭

この項目の中は、2つの段落に分かれている。「香(か・におい)」から始まる段落と、「体臭」に始まる段落である。前者は、においそのものを表す語であり、後者は何かのにおいを表す語である。この段落分けは形の上でも示してある。

ひとつの段落もいくつかに分かれる。前者は、「香(か・におい)」、「かおり」、「臭み」、「残り香」で始まる4つのグループに分かれそうである。後者は「体臭」と「俗臭」の2つのグループになるうか。これらのグループ分けの印はなにもない。しかし、利用者である我々には分かるし、したがっ

て検索も早くできる。

グループの中の語の並びにも意味がありそうである。たとえば、「臭み、臭気、異臭、悪臭」は、良くないにおいのグループだが、最初はより広い意味の語であり和語である。次に漢語で、さらに悪いにおいそのものを表す複合語で後ろ要素が「臭」となっている。これらの配列順序は必ずしも各項目に共通する規則にしたがって並んでいるのではない。しかし、それぞれを読むと分かる。もしこれを50音順に並べると「悪臭、異臭、臭み、臭気」となってしまう。これと比べると人間には先の配列の方が検索しやすい。

このような段落やグループ分け、またその並びに印がないから、あるいは明確な規則がないからといって、意味がないわけではない。さらに、機械処理に利用できないとも思わない。その分野での研究に期待したい。

「におい(1.504)」の項目は、増補版では次ぎのように増補、分類されている。斜線(/)は原文での改行を示す。

- 1 香(か) におい
- 2 薫り 薫香 芳香 香気 / 芳しさ ブーケ
- 3 臭さ 臭み 臭気 / 俗臭 異臭 悪臭 汚臭 死臭
- 4 残り香 移り香 余臭 刺激臭 / 腐臭 アンモニア臭 生臭み
/ 鼻持ち[3140-6]
- 5 体臭 口臭 俗臭 わきが
- A 防臭 消臭

専門用語と一般用語の分類にはおのずと違いがあろう。視点による分類だと思われる部分も「分類語彙表」に見られる。下表の右側に示した。もちろん、これは中野の解釈であり、厳密なものではない。

情報処理ではこの厳密さが問題になるのだろう。しかし、自然言語はもともと厳密ではないから、厳密ではないものを処理することこそが自然言語処理であるといえる。

1.4400 住居

- | | |
|------------------------------|----------------|
| 1 居住宅 お宅 家 うち | 家の一般的意味を表す語 |
| 2 本宅 別宅 私宅 | 主に住まいするか否か |
| 3 アパート マンション シャトー | 集合住宅 |
| 4 巣 ねぐら 巣穴 古巣 空き巣 | 人を含めた動物のすみか |
| 5 売り家 貸し家 借家 借り家 | 貸し借り |
| 6 屋敷 家屋敷 館 邸宅 邸 | 大きさ、規模、豪華さ |
| 7 山房 山荘 | 家がある場所、用途 |
| 8 宮 宮城 御所 皇居 王宮 天宮 | 住まう人：王など |
| 9 人家 民家 町家 山家 | 住まう人：庶民 |
| 10 城 城(じょう) 居城 城郭 古城 | 住まう人：戦士 |
| 11 本城 根城 出城 | 10の位置関係による言い分け |
| 12 墓 墳墓 奥津城 | 住まう人(?)：死んだ人 |
| 13 塚[4700-29] 古墳 前方後円墳 方墳 円墳 | 葬られた人の名も知れぬ古い墓 |

『分類語彙表』形式による語彙分類表(増補版)の語数

科研費を受けて作成した『分類語彙表』形式による語彙分類表(中野 1989, 下表白表紙)の掲載延べ語数は 52, 122 だった。次の科研費「言語研究におけるシソーラスの利用法」(平成元~2 年度)では 60,784 語を得た。国語研内の課題「分類語彙表の増補」を経て、一昨年で 82,828 語(下表 671 本)、現在 87,743 語(下表増補版)が増補の候補となっている。

分類	掲載延べ語数				増補率		
	元版(FD)	白表紙	671 本	増補版	白表紙	671 本	増補版
体の類	26,984	40,227	54,591	55,443	1.49	2.02	2.05
抽象的関係	6,780	9,026	12,663	12,558	1.33	1.87	1.85
人間活動主体	3,272	5,020	7,127	7,297	1.53	2.18	2.23
人間活動	9,920	14,708	19,247	19,724	1.48	1.94	1.99
生産物	3,277	5,656	7,960	8,145	1.73	2.43	2.49
自然	3,735	5,817	7,594	7,719	1.56	2.03	2.07
用の類	4,779	5,358	18,710	21,669	1.12	3.91	4.53
抽象的関係	2,153	2,380	8,145	8,322	1.11	3.78	3.87
人間活動	2,158	2,441	9,265	11,897	1.13	4.29	5.51
自然	468	537	1,300	1,450	1.15	2.78	3.10
相の類	4,653	6,147	8,928	9,890	1.32	1.92	2.13
抽象的関係	2,212	2,899	4,316	4,546	1.31	1.95	2.06
人間活動	1,788	2,515	3,571	4,228	1.41	2.00	2.36
自然	653	733	1,041	1,116	1.12	1.59	1.71
その他類	364	390	599	741	1.07	1.65	2.04
抽象的関係	99	105	127	187	1.06	1.28	1.89
人間活動	265	285	472	554	1.08	1.78	2.09
総計	36,780	52,122	82,828	87,743	1.42	2.25	2.39

この研究は、文部省科学研究費による研究「国際社会における日本語についての総合的研究(略称:新プロ「新プロ日本語」、研究代表者 水谷修)第3班の研究成果の一部を用いている。

参考文献

1. 国立国語研究所『分類語彙表』(国立国語研究所資料集6, 1964)
2. 中野洋『「分類語彙表」形式による語彙分類表』(科研費特定研究(1)「言語データの収集と処理の研究」, 1989)

3. 宮島達夫・小沼悦「言語研究におけるシソーラスの利用」(『国立国語研究所研究報告集』13, 1992)
4. 宮島達夫・小沼悦「言語研究におけるシソーラスの利用」(『語彙論研究』, むぎ書房, 1994)
5. 国立国語研究所『分類語彙表 [フロッピー版]』(国立国語研究所言語処理データ集5, 秀英出版, 1994)
6. 中野洋『「分類語彙表」形式による語彙分類表(増補版)』(科研費新プロ報告, 1996)

7. 宮島 達夫・中野 洋『分類語彙表』の増補について(『計量国語学』20巻6号, 1996)
 8. 国立国語研究所「現代雑誌九十種の用語用字 第一分冊」(秀英出版, 1962)
 9. 国立国語研究所報告4「現代語の語彙調査 婦人雑誌の用語」(秀英出版, 1953)
 10. 国立国語研究所報告12, 13「現代語の語彙調査 総合雑誌の語彙調査」(秀英出版, 1957, 1958) (前編、後編)
- 注1. β 単位は、語彙調査用の語の単位。国立国語研究所『雑誌九十種の用語用字(第1分冊)』(秀英出版, 1964)を参照。
- 注2. 国立国語研究所報告4『婦人雑誌の用語』(秀英出版, 1953)
国立国語研究所報告13『総合雑誌の用語後編』(秀英出版, 1958)

～「第10回専門用語シンポジウム：脳と用語」から～

TKE '96 にいたる 最近のターミノロジー研究

仲本 秀四郎* NAKAMOTO, Hideshiro

1. TKE '96

1996年8月、ウィーンにおいて、TKE '96が開催された。世界各地から多くの専門家が集り、日本からも20名を数える出席者を数えて、大方の関心の高さを示した。ここにいたる経緯ならびに3年毎に開かれる会議内容の変遷には、ターミノロジーに関わるものとして、いろいろ示唆に富むことも多いので、かいつまんで紹介することとしたい。

TKE '96: Fourth International Congress on Terminology and Knowledge Engineering 第4回 ターミノロジーと知識工学の国際会議

場所: Technical University of Vienna,
Vienna, Austria

日時: 1996年8月26日(月)より28日(水)まで

目的と内容: 用語学・情報学・計算機科学の学際的関心, 電算化技術を通じての情報と知識への普遍的な取得, 方法論の蓄積に基づく効率の応用への水平展開を目指し, 多言語情報処理, ニューメディア, ハイパーメディア, 情報ハイウェイの最近の開発を考慮に入れて, ターミノロジー研究, 知識工学, 言語工学, 電算機哲学, 分類理論, 情報とドキュメンテーション, 電算機支援研修/学習, 電算化ターミノロジー, 専門翻訳, 技術執筆の専門家の参加を求めた。

2. 経緯

第1回会議は1987年、ルクセンブルグ国境に近いドイツ・モーゼル河畔の古都トリアーで開かれた。ターミノロジーの国際団体 Infoterm と Association for Terminology and Knowledge

Transferの主催であったが、当時はユネスコの支援のもと、トリアー大学およびECの強い協力があった。その後、3年経て1990年の第2回会議もトリアーで開かれ、第3回は1990年にケルン、そして、このたび第4回をウィーンで開くにいった。ウィーンで開かれる経緯については、ターミノロジーの歴史と深い関係があるといつて差し支えない。

1936年、ISO(国際標準化機構)の前身ISA(国際標準協会)に、TC37(技術委員会37, ターミノロジー)がソ連の提唱でブダペストで発足してから1996年はちょうど60年にあたった。また Infoterm 創立25周年、ウィーン大学ターミノロジー研究所創立10年などが重なって、この年は記念すべき年であると、かねがね関係者の間でTKE会議のウィーン開催が取沙汰され、ようやく実現に漕ぎ着けたという経緯である。

この60年の進展については、いずれ他の機会に述べることもあろう。ただ、その経過のなかで、Eugen Wuesterを創始とし、Helmut Felber, Christian Galinskiに連なるウィーン学派の事績が、今回のウィーン開催に導いたといつて差し支えない。そのような事情を考慮すると、TKE '96の報告のみでは記念すべき意図にふさわしくないのではないかと考え、TKE '87前後からこの10年のターミノロジー研究を、TKEを中心に解説することによって、進展の模様とこの分野の概要を明瞭にしたい希望をもつようになった。

しかしながら、ターミノロジーの分野でなしとげられた成果について、すべてを尽くすことは困難である。特に、電算機の発達に伴って、ターミノロジーの活動に電算機が利用されてきたこと、さらには、機械翻訳・人工知能など、電算機利用にターミノロジーを必要とする趨勢が顕著になつ

* IRIS情報学研究所

て、ターミノロジーの関心は拡大し、概念の集成である「知識」工学の主題が加わってきた。ターミノロジーと知識工学が実質的に融合して、十分に熟した成果をあげるには、まだまだ多くの問題が残されているが、世紀末のこの時期は、注目に値する足跡を印していると考えている。

3. 最近10年の国際ターミノロジー会議

1987年前後からの最近のターミノロジー関係の国際会議を拾ってみると、次のとおりであり、このほかマレーシアその他の各地でも開催されたが、残念ながらすべてについて発表内容は把握されてない。以下の会議は、(4)を除いて日本からの出席者があり、内容もある程度わかっているし、発表を参照することも可能である。

- (1) International Congress on Terminology and Knowledge Engineering
29 Sept.- 1 Oct. 1987, Trier
- (2) Terminology Standardization and Harmonization in Theory and Practice
13 - 17, March, 1989, Tunis
- (3) Second International Congress on Terminology and Knowledge Engineering
2 - 4 Oct. 1990, Trier
- (4) Symposium on Standardizing Terminology for Better Communication: Practice, Applied Theory and Results
12 - 14 June 1991, Cleveland, Ohio
- (5) International Conference on Terminology, Standardization and Technology Transfer 2 - 6 July 1991, 北京
- (6) Third Infoterm Symposium "Terminology work in subject fields"
12 - 14 November 1991, Vienna
- (7) Third International Congress on Terminology and Knowledge Engineering
25 - 27 August 1993, Cologne
- (8) Fourth International Congress on Terminology and Knowledge Engineering
26 - 28 August 1996, Vienna

4. 国別発表数の経緯

それぞれの国で言語を異にする以上、ターミノロジー研究の進展を眺める際、国別の傾向を見る切り口もあるはずである。言語事情により、また政策など国の力の入れ方も異なり、歴史的経過もあって、その間の傾向はグローバルな解析に資するものと、国別の発表数を調べてみたところ、第1表のような結果が出た。チュニスと北京、A S

TMの各会議は国内会議の性格が多分にあり、統計からは除外した。また、(6)のInfoterm会議は主題別ターミノロジーをテーマとしており、統計には含めてない。TKEにしても当初はドイツ国内会議の色彩が強く、数の上でドイツの発表が多かったばかりでなく、ドイツ語による発表が第1回は1/3を占めた。

第1表 国別発表数の経緯

	1987	1990	1993	1996	計
ドイツ	18	20	11	7	56
ロシア	5	13	2	3	23
ウクライナ	-	-	1	2	3
ラトビア	-	-	1	1	2
日本	10	4	1	5	20
フランス	3	6	5	5	19
オーストリア	2	4	7	6	19
アメリカ	7	6	2	2	17
イギリス	2	1	3	6	12
オランダ	8	3	1	-	12
スペイン	2	-	2	4	8
カナダ	2	2	2	-	6
ルクセンブルグ	4	-	1	-	5
デンマーク	-	3	2	-	5
ベルギー	1	1	2	1	5
中国	1	2	2	-	5
ポーランド	1	-	2	-	3
スイス	-	1	1	1	3
イタリー	-	-	1	2	3
インド	-	3	-	-	3
旧ユーゴ	-	1	1	-	2
スウェーデン	-	1	-	1	2
ノルウェー	-	-	1	1	2
イスラエル	2	-	-	-	2
ギリシャ	1	1	-	-	2
アルゼンチン	-	1	-	-	1
イラン	1	-	-	-	1
南アフリカ	-	-	-	1	1
フィンランド	-	-	-	1	1
ブラジル	-	1	-	-	1
韓国	-	1	-	-	1
チリー	-	1	-	-	1
ナイジェリア	-	-	-	1	1
マレーシア	1	-	-	-	1
メキシコ	-	-	-	1	1
ハンガリー	1	-	-	-	1

結果は常識的な傾向を示した。常識的というのは、学術的な国際会議で今まで記されあるいは語られてきた国別傾向とほぼ同様であったということである。INFOTERMを抱える Austria はもちろん、地理的に近いフランスが多いのは当然に思える。ロシアの発表が主催国のドイツに次いでいるのが目を引くが、ほとんど出席していない。日本からの発表がフランス・オーストリアと肩を並べているのも、目立つ数字である。開発途上国の熱心さがもう一つの特徴で、C. Galinski も指摘したように、技術移転を目的に、方法とデータに関心が強いのは、かれらの置かれている状況からきている。多言語国家であるカナダやスイス、ま

た、中国のように政策としてターミノロジーの実際活動の盛んな国の発表がそれほどでの数でもないところを見ると、発表と実際活動には関連がないように見える。

5. 主題分類

発表は、基礎理論から実際の用語活動、標準化、教育にいたるまで多彩であるが、その傾向を概括すると、次表のとおりである。発表によっては二つのカテゴリに属するものもあり、また、分類に困難を感じるものもあって、必ずしも正確ということにはならないが、全体の傾向をつかむことは可能であろう。

第2表 主題の変遷

	TKE'87	TKE'90	TKE'93	TKE'96	ASTM	TSTT	Infoterm
基礎と原理	16	13	11	9	4	10	3
自然言語処理	3	3	0	2	0	4	0
オーダリング・分類	4	1	1	0	0	0	0
知識モデル	1	2	1	2	0	0	0
知識表現	4	9	4	2	1	0	0
知識情報の伝達と提供	2	5	4	1	3	1	2
用語知識データベース	4	3	8	5	1	3	4
用語電算機システム	7	4	3	4	6	4	1
シソーラス	1	5	3	2	0	2	0
辞書	6	2	1	1	0	1	1
電子化辞書	1	8	0	2	0	1	0
機械翻訳	6	5	0	1	0	3	0
知識工学	4	10	5	1	0	0	0
データ交換	0	0	5	1	1	0	0
翻訳指向ターミノロジー	0	5	4	3	1	0	1
多言語	0	0	2	0	0	3	1
非ヨーロッパ語の特殊性	14	6	1	7	2	3	5
実際応用	3	1	3	6	1	0	0
ネットワーク	0	0	0	5	0	0	0
マルチメディア	0	3	1	0	0	0	2
分野指向	5	3	1	6	2	0	9
教育	1	0	1	1	0	0	0
標準化	0	1	1	0	7	6	2
システム接続	1	0	0	0	0	1	0

6. 基礎, 原理

発表の40%を大学で占めているのが一つの特性で, Wuester 以来の理論的な報告が多く, 対象特性, 分類あるいは構造, 方法論, 意味ネットワーク, 概念分析, 知識取得, 記号論の研究など, より幅広く展開されている。科学史の側面を添え, 論理学, 言語学を超えて, 認識論, 認知, オントロジーなど, この10年は玉石混交ともいう状況であったが, 落ちつき方によっては, 期待のもてる感触がないでもなかった。

例えば1991年の infoterm の会議で, G. Rahmstorf は技術用語がその定義に知識を蔵していることから, 統語論的・意味論的解析を行い, 概念間のリンクは定義表現の統語構造の関数であり, 概念ネットワークは翻訳過程の結果で確立されるとして, 定義に基づく概念ネットワークの構造と分類による主題の構造とを比較した。また, 同じ会議で, V. Zhuravlev はオブジェクト知識と概念体系の標準的オントロジー構造, 概念のサブシステム, オントロジー構造の比較, 概念と用語の創造を論じた。

TKE'96において, Khurshid Ahmad は学術誌の索引語の変遷から, terminology dynamics が科学哲学の基礎になりえないかとの問題提起を行い, また, Gerhard Budin はターミノロジーの研究で科学史の再編が可能でないかと, conceptual dynamics を含む方法論を論じている。また, Heiner Benking は図書館学や情報検索におけるオーダーリングは多くの反対に直面しているとし, 概念のスーパー構造内で概念の接続点と比喻を論じ, 概念柱ないし意味空間の構築に言及した。にわかには理解するには努力の必要を感じたが, 意欲は評価できた。

TKE'96の圧巻は, 最終の発表でもあるし, 基調講演でもあった John F. Sowa のオントロジーの話であった。準備が間に合わなくて, 近刊予定の著書の摘出であるがと断った内容であったが, 最近のオントロジー活動の紹介から始め, 知識表現の基礎として論理, オントロジー, 計算化の3つを挙げ, 地理的, アリストテレス流カテゴリーツリーに言及して, オントロジー構造を英語・中国語の例で解説を試みた。著書の内容紹介の嫌い

はあるが, ターミノロジーとは密接な関係にあるオントロジーのアプローチは, AIばかりでなく, この分野にもいろいろの示唆を与えるものと思われた。

7. ターミノロジーデータバンク

この時期は, メインフレームからパソコン環境への変化が著しい時代特徴であり, 発表の内容もその変化に対応してきたといつてよいであろう。ネットワークの発表がTKE'96に現れたのもインターネットの展開と時期を同じくしている。

ターミノロジーデータベースは毎回, 着実な歩みを見せ, 用語集からデータベースへ印刷体からの脱皮を示している。その点は辞書から電子化辞書にいたる関心の変遷と軌を一にした。87年のECセッションの際, J. Goetschalcks は, 加盟12か国・9公式語をもつECの重要問題である多言語間の翻訳に必要な9か国語データバンクEURODICATION の開発を紹介し, 定義の困難さがあるけれども完成可能であると発表した。TKE'96会議後訪れたスイス政府官房ターミノロジー課でその実働を見るに及んで, 改めてこの10年の進展を実感した。

Infoterm Symposium は開催趣旨が各分野でのターミノロジー活動であったから, 分野指向が多いのは当然として, いずれの会議も自然科学が大部分で, 社会科学関係の少ないのが気になった。

Infoterm 会議において, ウクライナの F. Citkina が contrastive analysis によって, 語・意味・文法各レベルでターミノロジーの similarities と divergencies を求め, 英露翻訳に適用した経験を述べて, 多くの注目を惹いた。また同じ会議で, P. Losco らの多言語語彙管理システム GELEM のデータ構造に関する報告は, morphological dictionaries の構造として, インド・ヨーロッパ言語族, セム言語族, 表意文字を使う言語族用に, タイプの異なるものを用意しようとした意欲的な研究であった。

'96では Khursid Ahmad らが, 最近完成し EU が支援する TRANSTERM を紹介した。

8. 機械翻訳

TKE '87において、ECが開発中の機械翻訳について特別セッションが設けられ、Juergen VollmerのEUROTRA(機械翻訳システム)プロジェクトやI. PigottのSYSTRANの報告がC. Van der HorstのSYSLING(EC9公式語の翻訳サービス)についての紹介とともに、数百名に上る出席者の関心を引いた。PigottはSYSTRANについて、1976年から英語をsource languageにして、仏・独・露・日翻訳の研究をいくつかの相手先と契約して進めてきた歴史にふれ、ページフォーマットやスピード・経済性・正確さ・柔軟性などの点で、SYSTRANが優れていることを強調した。当時、機械翻訳の評価についての質問に発表者は明確な返事を避けたが、Van der Horstが返事のなかで、翻訳にはquality levelとquantity levelのあることを指摘したのは強く印象に残った。TKE '96後のヨーロッパ機械翻訳協会でのワークショップで、Dorothy Senezが「20年前、研究段階であった機械翻訳は今や実施への関心」を引く状況になったという発言が、この間の進展を如実に示している。翻訳家の協力が進んできたことが目立っている。

9. シソーラス

シソーラスの作成は用語編纂の一つであるが、目的が情報検索用であるため、体系・手法でターミノロジストとの間で議論があり、その効用への関心もあって、毎回の発表があった。国際協力による編纂やパソコンを用いた発表が目についた。

87年Ulrich Guentzerらは、概念間の関係を用いる方法によらず、実際の検索挙動から結論を自動的に引き出すTEGENという名のシステム開発を紹介した。TKE '90では、Ronald BuchanがNASAシソーラスの新登録語をデータベースに遡及してindexingしなおした経験を報告した。標題からmajor termを、抄録からminor termという選択が現実的である。Infoterm会議では国際原子力機関のC. TodeschiniがINISシソーラスとETDEシソーラスとの整合作業の経験を述べ、米国で使われない語、米国でしか使われない語、計画経済の国で使われない語などを挙げた。

L. de Lavieterは1990年に出版されたオランダ環境用語シソーラスに基づき、イタリア語及び英語への翻訳とソフトウェア開発を行って、約3,000語の3カ国語環境分野シソーラスを作成している発表があった。実際には、シソーラスというよりも分野別キーワードリストに近いが、この点は母体となったUNEPのシソーラスも同じである。TKE '93でErich SchweighoferとWerner Winiwarterは、ブール検索では、しばしば概念を特定できないことがあるとのべ、意味の知識表現利用や意味ネットワーク、ベクトル空間モデルなどの他の研究を参照して、知的検索システムであるKONTERMを構築、法律文の自動インデキシングのための選択的シソーラス作成を紹介した。

10. 知識工学

AIの関心に対応して、意味ネットやフレームに着目した知識モデルの発表から知識表現など、知識に関する基礎研究が常に発表されるのは、この会議のひとつの特徴を形成している。その応用ともいえる知識データベースの発表が、ターミノロジーデータベースに付随して各回会議での注目を招いた。

TKE '87でLa Raw Maranはプロダクション型のエキスパートシステムの農学知識応用システム開発の限界に関連する研究を発表した。また、抄録作成のAI化に関する手順モデル(Brigitte Endre-Niggemeyer)など、経験にもとづいた具体的な報告も、かなり発表された。

1991年、ウィーン大学のA. M. Tjoa and R. R. Wagnerはsemantic modelを要領よくレビューして、この分野においてさえ、共通の概念に異なる用語を用いていると指摘し、semantic data modelはdata modelへの応用環境でオブジェクトをマップするものと定義した。定義の可否はともかく、応用を意識していることと、意味とデータの領域を認識している点で同意できる内容と思われる。

TKE '93において、D. H. Fischerは、二言語ターミノロジーの語彙モデルとしてシソーラス・モデルを開発する際に適用できる、選択的な知識の表現法を扱い、実体と関係のデュアルモデル、

概念の命名と同形異義語の統制に関する原則、階層構造と結合関係の意味論、一種の構造保存型のマッピングとしての翻訳と多言語性に関する経験を報告した。

回かなりあったが、ラテン文字使用のターミロジーに携わる人々には異質感を与え、かつ歴史も浅くターミロジー分析の洗礼が十分でないからか、毎会議とも別扱いになっていた。しかし、その数は無視できない多さであり、参考のため第3表に集約してみた。

11. 非ヨーロッパ語の研究開発

また、ヨーロッパ語以外を対象とする発表が毎

第3表 非ヨーロッパ語対象の発表

[1987]

日本語	1.Masanobu Fujikawa, Tetsuya Ishikawa, 2.Masahiko Ishii, 3.Hiroshi Nakai, Tetsuya Ishikawa, Takaharu Osada, et al, 4.Akemi Haruyama, Kyoko Eriguchi, 5.Yasuhito Tanaka, Sho Yoshida, 6.Takeyoshi Hanada, Hajime Sasaki, 7.Ryuta Ito, Kunio Oota
中国語	1.Yongyuan Yu, 2.Feng Zhiwei
マレー語	Khalid M. Hussain, Zahrah Abd. Ghafur
東アジア	Yasuhiro Ota
ペルシャ語	Reza Mansouri
ヘブライ語	Ephraim Nissan
ギリシャ語	Spyros E. Diemessis

[1990]

韓国語	Deok-Bong Kim, Hyun-Jae Im, et al
中国語	Su Wubin, Jin Wanping
モンゴル語	Che Hosbayar, Hasbagan
非欧語	Christian Galinski
ギリシャ語	Diameisis
スペイン語(チリ)	Anna Maria Burdach, et al

[1993]

中国語	Feng Zhiwei
-----	-------------

[1996]

日本語	1.Kyo Kageura, Keita Tsuji, Atsuhiko Takasu, 2.Keita Tsuji, Kyo Kageura, 3.Akemi Haruyama, 4.Masumi Narita, Takashi Katooka and Yoshihisa Ooguro, 5.Werner Winiwarter, Osami Kagawa, Yahiko Kambayashi
ペルシャ語	Reza Mansouri
南ア英語	Milde Jordaen-Weiss

12. その他

第1回に比して、第2回で参加者が急激に減少したのは、実は会議の性格が明瞭になってきたことにある。機械翻訳に刺激されて第1回に集まってきた翻訳者は大部分がターミノロジーの利用者で、研究的色彩の強いこの会議には向かず、会議がこの規模になるのは当然といえた。出席者が減ったに対し発表は増えていることも裏付けている。あいかわらず、概念的な発表もみられたが、ターミノロジーデータバンクの要求、実績、経験など、実際の・具体的な発表が増加したのも特色である。ソ連がターミノロジーに熱心とはかねてからの評価で、発表は玉石混交であったが、経験にもとづいた実際のかつ優れた内容のものもあった。

Erich Ortner のデータモデリング企業での例、TRADOC (ヨーロッパ共同体の科学技術データベース)、また、品質管理の用語集が日本を含めて、既に多量の出版物になっていることも披露された (Joseph Sans)。

目的に応じて語を性格づけする試み (F. Sarret al) (Joseph Sans) や各種のターミノロジーシステムあるいはツール (Pierrette Bouillon and Lucia tovena) (Heinz-Dirk Luckhardt) の発表は毎回多く、色々の提案があり、それぞれの動機が理解できたが、いまひとつ汎用性と一般性が十分でない感を抱かせた。

ハイパーテキストに関する発表の多かったのも最近の傾向を示したものである。ハイパーテキストに関連して、語の関係表示に工夫がなされた (Rente Mayer and Karl-Heinz Hanne) 報告も目にした。

翻訳家の積極性が目立った。Infoterm 会議で D. G. Walker は翻訳者が当面している問題を、多言語問題を抱えるカナダや EC・国連機関での翻訳担当者によるターミノロジーの必要性、ターミノロジーデータバンクへの翻訳者の需要、社会的活躍の場としてのターミノロジー会議など、ターミノロジーへの翻訳家の参加が、トリヤー会議にも増して顕著になった印象を受けた。

S. E. Wright は ASTM, SAE, ASHRAE, APICS, US Military Specifications (MIL SPECS), NISO (MIL SPECS), ASME で作られ

た米国のターミノロジー規格について、一般的レイアウトの慣習、構造的な方法論、冗長性・多義性・一貫性、関係国際規格との調和を考察した旨報告した。また、I. Dahlberg は英独両言語の間の内容と形式の同等性を決めうるかどうかの研究を述べた。J. Datta は国連で計画中の6カ国語の会議文書を保存する Full-Text Data Bases システムを、P. Lewalle が WHO 加盟国間の技術移転に必要な用語の問題に関する活動を報告した。

TKE'93 では Gert Engel が基調講演において、1言語内での非曖昧性の解決を論じ、機械翻訳については「辞書を与え給え」と叫び、非曖昧性を可能にするマルチメディアコミュニケーションでは「言語筆記体系の独占は既に破れた」と宣言し、非曖昧性の追求がターミノロジーの目的であり、ホモファーバー (工人) としてのターミノロジストに課せられていることの重要性を強調した。

Uwe Reinke はターミノロジーデータ交換での困難を、Hans-Juergen Stellbrink はターミノロジーが教育に果たした貢献は計算できないし、標準化によるコストの軽減は莫大のはずであると、その中で、ターミノロジーの交換が無料なら翻訳のコスト削減は可能であるとした。C. Quist の報告は、概念分析から用語のプレゼンテーションに至るまでの仕事を支援する、標準的なパソコンで動作可能な応用プログラムの開発についてであった。ヨーロッパ議会の EUTERPE database について、TKE'93 では S. Ball が使用経験を、'96 には Cornelis van der Laan が www によるアクセスを報告した。

この点からいえば、TKE'96 の特徴はなんといっても、ネットワーク利用の発表が現れたことである。Erich Schweighofer と Scheithauer はインターネットで法律ターミノロジーシステム KONTERM を、Christine Jacquin and Maurice Lisouet はインターネットによる大量のテキストデータからのターミノロジーの抽出を発表した。また、Josef Wallmannsberger はネットワーク環境の急激な成長の中で、オンラインカタログにおける学生の実験とハイパーテキストブロージングを報告するはずで期待したが、結局現れなかった。

13. Prof. Czap のまとめ

この会議の創設に貢献のあったトリアー大学の Prof. Hans Czap が、TKE'87 の最終レビューでまとめた分類は '87 にかぎらず全般に有効である。すなわち、

1. Concept-oriented system
2. Term-oriented system
3. Logic-oriented system
4. Procedual-oriented system

とあって、ターミノロジーの全貌を概括する道標になる。そして、この分野の将来の傾向として

1. Automatic reasoning
2. Pattern- and structure-matching
3. Hybrid system
4. Picture processing system
5. Procedual concept system
6. Technology assessment

を挙げた。10年たった今なおこれらがまだ課題としてあるのは、今後待つ示唆であるとしても、やはり長いようで10年は短い年月であったといことができる。

14. むすび

TKE '87 および TKE'96 会議には、それぞれ 20 名を超える日本人参加者があり、諸外国に比して、ターミノロジーに対する日本の関心の高さを示した。また、発表も顕著な数を示した。このレビューで日本からの発表を意識的に除いているのは、一つは国内で内容が入手可能であり、一つは身内への無遠慮を避けた姿勢があるが、他方、日本の発表内容はその数の多さに比例しているという期待もあった。すなわち、異言語間の概念の移転について、日本は常に深い関心を示し、成果の蓄積もかなりの量である。その集成・組織化と解析・理論化に、積極的に取り組んできたし、またさらに進めねばならない状況にある。したがって、会議への多数参加が今回の意義ばかりでなく、将来への発展を期待して喜ぶべき事態と考えている次第である。

謝辞

この稿をまとめるにあたっては、春山暁美(元愛知淑徳大)、平井歩実(関西女学院短大)、門條司(三洋ソフトウェア)の諸氏の協力を得ているので、ここに謝意を表したい。

別表 1 シソーラスに関する発表

[1987]

Haruyama, A./ Eriguchi, K.: Construction and evaluation of an environmental microthesaurus

[1990]

Lutes-schaab, B.: An online thesaurus for fact retrieval

Rogozhnikova, T.: Individual information thesaurus: associative fields and a model of semantic development

Sarre, F./ Mittermeir, U./ Guentzer, U./ Guentzer, G. Learning behavior and user acceptance of TAGEN, A thesaurus - generating information retrieval system

Sunetti, R.: Thesaurus in law: some problems of construction

[1993]

Schweighofer, E./ Winiwarter, W.: Refining the selectivity of thesauri by means of statistical analysis

Muraskiewicz, M./ Rybinski, H./ Struk, W.: MULTHES-ISIS: A flexible software for multilingual thesaurus building

Roulin, C./ Cooper, C.: Bringing thesauri together: terminological problems and interests

[1996]

Anzaldi, C./ Bordoni, L./ Sano A.: Construction of a terminological interdisciplinary thesaurus

Haruyama, A./ Yamashita, Y./ Kubota, H.: Development of a multilingual indexing vocabulary based on a faceted thesaurus

Third Infoterm Symposium, 1991

L. de Lavier/ J.A.Deschamps/ B.felluga: A multilingual environmental thesaurus - past, present and future

Todeschini, Claudio / Thoeming, George: The thesaurus of the international nuclear information system: Experiences in an international environment

別表2 各会議にわたる発表者の経緯

10年にわたる各人の発表の経緯は、研究の動向、関心の変遷を示して、興味あるデータになろうかと、多発表者を拾ってみた。

Gerhard Budin, Universitaet Wien

1987 Terminology and Knowledge Data Processing
with Galinski, Nedobity and Thaller

1990 Scientific Knowledge Structure

1993 Knowledge Organization of Terminological Knowledge

1996 Evolution of Scientific Terminologies

Herbert A. Eisele, Institute Superieur d'Interpretation et de traduction

1987 Concept Assesment in Terminological Practice

1990 Les Frustrations du terminologue

1993 Speed training via Terminology

1996 The Ownership of Words and Knowledge

Ulrich Heid, University of Stuttgart

1987 Die Verbklassifikation von Steiner 1986 als Basis fuer die Generierung
Deutscher Saetze

1990 Terminological and Lexical Knowledge for Computer-aimed Translation and
Technical Writing (with Freibott, G.)

1993 On the Representation of collocational Phenomena in Sublanguage Lexicons

1996 Term extraction with Corpus Exploration - Experience from German
(with Hohmann, A.)

Khurshid Ahmad, University of Surry, U.K.

1987 Term Banks: A case Study in Knowledge Representation and Development

1993 Terminology and Knowledge Acquisition: A text-Based Approach

1996 A terminology Dynamic and the Growth of Knowledge: A Case Study in

- Nuclear Physics and in the Philosophy of Science
Engineering Terminology - A Case for a Linguistically-informed
Terminology Database
(with Ogonowski, A., Dauphin, E., Sta, J-D, & Arppe, A.)
Terminology Management and Lesser-used Living Languages: A critique of
the corpus-based Approach
(with Brekke, M. & Myking, J.)
Is Your Terminology in Safe Hands? Data Analysis, Data Modeling and Term
Banks (with Holms-Higgin P.)

- Faina Citkina, Uzhgorod State University
1990 Comparative Terminology Theory: Problems, Goals, Methods, Applications
1993 Methods and Procedures in Comparative Terminology Science
1996 Terminological relativity and translation

- Hans Czap, University of Trier
1987 Neue Anzaetze in Terminologie und Wissenstechnik zur Unterstuetzung von
Information und Kommunikation
1990 Construction and Representation of Concepts in Enterprises
1993 Guiding Principles for (Re-)Constructing Concepts

- Herbert Picht, Institut for Spansk, Handelshojskolen Kopenhagen
1987 Fachsprachliche Phraseologie - Die Terminogische funktion von Veben
1990 New Professional Profiles in Knowledge Engineering and Knowledge
Transfer (with Engel g.)
1996 Graphic and Other Semiotic Forms of Knowledge Representation in
Terminology Management

- Robert de Beaugrande,
1987 Systemic versus contextual Aspects of Terminology
1996 LSP and Terminology in a New Science of Text and Discourse

スイス連邦官房ターミノロジー課訪問記

春山 暁美 HARUYAMA, Akemi

1996年9月2日、TKE'96会議の帰途立ち寄ったベルンでスイス連邦官房中央言語部 (Central Linguistic Services) のターミノロジー課 (Terminology Section) を訪問する機会を得た。訪問者は、仲本、太田、稲見、吉野、長田、高橋、春山、平井、山下の9名である。月曜日の朝、ホテルまで迎えにきて下さった課長のRolf Moos氏の案内で、一同徒歩で町中のビルの中にあるこじんまりしたオフィスを訪ねた。

ターミノロジー課は、少人数 (筆者の記憶によると7名) のチームである。最初にMoos氏から全体的な話があり、つぎに用語データベースの管理者であるFrank Kauz氏が端末を使用して用語検索のデモンストレーションを行い、その後、かなり活発な質疑応答の後で、数人ずつに分かれて両氏のそれぞれの部屋に招かれ、見学及び討論を行った。

○内容

この訪問を受けるに当たって、Moos氏が提示した論点は、次の三つである。話もこれに沿って進められた。

- (1) 多言語ターミノロジー作業
- (2) ターミノロジーデータの処理
- (3) ターミノロジー協力

・多言語ターミノロジー作業

スイスは多言語国家であり、1994年刊行のスイス連邦統計によると、ドイツ語63.7%、フランス語19.2%、イタリア語7.6%、特定地域の人々が話すロマンス語 (Romansch) 0.6%、その他8.9%となっている (1990年全国調査による)。6言語の中から先に挙げた4言語が公用語となっており、法律は、コスト高となるロマンス語 (こ

れを話す人はドイツ語を解する) を除き3言語で書かれている。

約200人の翻訳者がターミノロジー・サービスのために協力している。10年前にはLinguistic Serviceは存在していなかった。15年前に用語データベースEURODICAUTOM (9ヶ国語、50万件以上、1973年開発) に基づく用語データベースTERMDATが開発された。(1988年)

標準的の用語リストは作らず、用語データベースを公開してフィードバックを受け入れた方がよいという考えである。対象ユーザーは、現在はcivil serviceのみであるが、将来的には開放される可能性がある。国内の大学は全部インターネット等を通じて用語データベースにアクセスできる。

・ターミノロジーデータの処理

用語データベースのソフトウェアは、ECが開発・提供したものである。コマンドの数40、対話型の検索システムと、用語管理のバッチシステムがある。用語レコード数約170,000件、略語も収録されている。用語レコードの構成要素としては、ソース言語、ターゲット言語、主題分野コード (46分類)、定義、参考文献 (出典情報)、用語の信頼性のレベルを表すコードなどがあり、複数のフィールドを組み合わせることで検索することができる。

データベースの更新は月2~3回ルクセンブルグ (EC本部) から送られるCD-ROMに基づきバッチで行う。年間約20,000件のデータが更新される。(以上、数字は筆者のメモによる)

文字セットとしては、ISO8859を使用している。

・ターミノロジー協力

3年くらい前から、協力体制が非常によくなっ

た。

○ “COTSOWES Terminology”

Moos氏から私達に上記表題の文献2部（バインダーにファイルしたもの）が贈られた。内容は、COTSOWES (Conference of Translation Services of West European States) の Working Party on Terminology and Documentationが編さんしたターミノロジー作業のための勧告指針 (Recommendations for terminology work, 1990年刊行) である。基本的な解説などのほかに、付録として、既存の用語データベースなどが簡単に紹介されている。TERMDATについては記載されているが、内容の詳細は記述されていない。



○感想

小会議室のテーブルを囲み、リラックスした雰囲気話し合いがもたれた。初対面にもかかわらず、同じような分野の仕事に携わる者同志の親近感のようなものが生まれ、実のある内容であった。

この訪問を企画し、実現のためにお骨折り下さった仲本氏と、多忙な時間を割いて私達のために密度の濃いプログラムを用意していただいたMoos、Kauz両氏に対し、心から感謝したい。



J I S 用語規格制定の動き

情報源：“標準化ジャーナル”，1996年1月号-1996年12月号

制定・改正される J I S

- 航空宇宙-流体系統に関する用語-第1部：圧力に関する一般用語および定義（制定）
Aerospace - Fluid systems - Vocabulary - Part 1: General terms and definitions relating to pressure
- 航空宇宙-流体系統に関する用語-第2部：流れに関する一般用語および定義（制定）
Aerospace - Fluid systems - Vocabulary - Part 2: General terms and definitions relating to flow
- 航空宇宙-流体系統に関する用語-第3部：温度に関する一般用語および定義（制定）
Aerospace - Fluid systems - Vocabulary - Part 1: General terms and definitions relating to temperature
- B0106 工作機械（部品・工作方法）用語（改正）
Glossary of terms relating to parts of machine tools and operation
- B0114 木材加工機械の名称に関する用語（改正）
Glossary of terms for wood working machinery
- B0152 クラッチおよびブレーキ用語（改正）
Glossary of terms relating to clutches and brakes
- B0171 ドリル用語（改正）
Glossary of terms for drills
- B3000 FA用語（改正）
Glossary of terms used in factory automation
- B8650 プラスチック加工機械用語（改正）
Glossary of terms used in plastics working machinery
- C0067 環境試験方法-電気・電子-耐火性試験用語（改正）
Terminology concerning fire tests
- K0212 分析化学用語（光学部門）（改正）
Technical terms for analytical chemistry (optical part)
- L0221 ジオシンセティック用語（改正）
Glossary of terms used in geosynthetic industry
- Z2300 非破壊試験用語（改正）
Glossary of terms used in nondestructive testing
- Z3001 溶接用語（改正）
Welding terms

官報告示・発行された J I S

- A8403-1-96 土工機械-油圧ショベル-第1部：用語および仕様項目（制定）
Earth-moving machinery - Hydraulic excavators - Part 1:

- Terminology and commercial specifications
- A8422-1-96 土工機械－ダンプトラック－第1部：用語および仕様項目（制定）
Earth-moving machinery - Dumpers - Part 1: Terminology and commercial specifications
- B0103-96 ばね用語（改正）
Springs vocabulary
- B0154-96 円すい用語（改正）
Vocabulary of cones
- B0176-96 ねじ加工工具用語（改正）
Threading tools - Vocabulary
- B3410-96 プロッタ用語（改正）
Technical drawings - Plotters - Vocabulary
- C5610-96 集積回路用語（改正）
Glossary of terms used in integrated circuits
- D0102-96 自動車用語－自動車の寸法、質量、荷重および性能（改正）
Road vehicles - Dimensions, masses, weights and performances -Vocabulary
- L0202-95 家庭用編機用語（改正）
Nomenclature of hand knitting machine
- Z0102-96 クラフト紙袋用語（改正）
Vocabulary for kraft paper sacks
- Z0103-96 防せい防食用語（改正）
Glossary of terms used in rust and corrosion preventive technology
- Z0108-96 包装用語（改正）
Glossary of terms for packaging
- Z6000-96 マイクログラフィックス用語（改正）
Micrographics - Vocabulary

「専門用語研究」投稿規定

1. 「専門用語研究」（以下会誌という）には、下記の内容に関する論文・記事を掲載する。

- ・ターミノロジーの理論と応用
- ・専門用語集の作成技術
- ・その他、専門用語に関するもの

2. 会員は、会誌に自由に投稿することができる。編集委員会からの依頼により執筆することもできる。

3. 原稿の書き方

3.1 原稿用紙

原稿は、通常のA4サイズ横書き原稿用紙（20字×20字）か、ワードプロセッサを使用する。ワードプロセッサを使用する場合は、A4用紙に1行20字、20行で作成し、印刷する。また、可能であれば、MS-DOSテキスト形式でフロッピーに保存し、印刷物と一緒に送付する。

3.2 原稿の長さ

全体で図表ほかを含めて、原稿用紙16枚から32枚とする。原稿用紙4枚で刷り上がり1ページとなる。執筆依頼時に別途指定ある場合はそれに従う。

3.3 原稿の仕様

原稿には、以下の内容を記入する。

- ・和文と英文の、表題、著者名、所属
- ・和文の、抄録(250字前後)とキーワード(5から10語)
(可能ならば、英文の抄録(150語前後)とキーワード(5から10語)も)
- ・本文(ページをつける)
- ・図表など(番号と表題をつけ、朱筆で文中に挿入位置を指定する)
- ・引用文献(本文中に肩付き数字¹⁾²⁾・・・を記入する)
- ・参考文献、参考図書(本文を読む上で参考になるものがあれば)

4. 原稿の受理、査読

投稿原稿は、当研究会事務局が受け付けた日を受付日とし、会誌編集委員会で査読を行なう。査読結果をもとに、会誌編集委員会で掲載の可否を決定する。委員会で内容・表現などについて修正が必要と認めた場合、執筆者に修正依頼する。

5. 校正依頼

執筆者に初校を依頼する。この際、大幅な修正・加筆は行なわないこと。なお、論旨に差し支えない範囲で、編集委員会が内容の変更を求めることがある。

6. 掲載原稿の扱い

会誌に掲載された原稿、フロッピーは返却しない。

7. 謝礼

執筆者には、掲載された会誌10冊を無料贈呈する。これ以上および抜刷を希望する場合は、有料となる。校正時に申し込むこと。

8. 著作権

本誌に掲載された論文、記事の著作権は、当研究会に帰属する。

9. 原稿提出先

専門用語研究会誌編集委員会

編集委員会委員

太田 泰弘 文教大学
四ノ宮明夫 大正製薬
戸塚 隆哉 KMK デジテックス (委員長)
中山 亮一 リョウ・プロダクション
牧野 正久 東京理科大学
山下 泰弘 電気通信大学
山本 昭 関東短期大学

専門用語研究 第13号

(1997年3月31日発行)

発行所 専門用語研究会

〒102 東京都千代田区一番町4-6

一番町中央ビル2F

日本総合技術研究所(JIST)内

Tel.03-3262-8956

Fax.03-3262-8960