

TM-0669

CTASの評価のまとめ

椿 和弘, 滝 寛和,
藤井裕

January, 1989

©1989, ICOT

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03) 456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

ICOT Technical Memorandum

CTASの評価のまとめ

ICOT 第5研究室 椿 和弘、滝 寛和、藤井 裕一

1 はじめに

I C O T 第五研究室で開発された分類型問題向きの知識獲得支援システム CTAS の評価を再委託先メーカーに依頼して行った。本 T M ではその評価レポートを基にまとめた、CTAS の評価結果について報告する。知識獲得とは専門家の知識を効率よく収集し、整理・体系化を行い、知識システムシェルの知識表現に変換する作業であり、エキスパートシステム開発のボトルネックとされている。知識獲得支援システムはその知識獲得作業を支援するシステムであり、ナレッジエンジニア (KE) 無しに専門家との対話によって知識ベースの構築支援を行う。知識獲得支援を行うためには高次推論が必要となるが、その実現のための要素技術はまだ研究段階にあり、支援システムによる実用的知識ベース構築支援は難しい。今回の CTAS の評価は実用的な知識ベース構築が目的ではなく、知識獲得支援システム、及び CTAS の問題点の洗い出しにある。また、CTAS の評価結果で CTAS についての誤解がある部分や補足が必要であると思われる部分については、評価レポートの内容を一部修正した。評価レポートの詳細については備え付けの資料を参照して欲しい。

2 CTAS 概要

CTAS は分類問題解決に必要な類型タスクである階層化分類と順位付けのタスクをもつ初期知識ベースを専門家との対話により構築する知識獲得支援システムである。CTAS では、分類対象 (項目) をトップダウンとボトムアップの 2 種類の方法で階層化し、階層化された分類項目に対して George Kelly の Personal Construct Theory に基づいて分類に必要な特徴と、項目と特徴間の関係度の抽出を行う。そして分類項目と特徴を種々の表とグラフを用いてこれらのリファイン (洗練化) を行い、CF 値付のプロダクションルールから成る知識ベースの構築を行う。CTAS の特長は階層化による分類型問題の知識獲得作業の軽減・獲得知識の精度の向上を図ったこと、分類対象を簡単に階層化する方法を提供していること、各種のグラフィック表現により知識ベースのリファインを効率的に行うことが挙げられ、診断・設計等の各ドメインに含まれる分類タスクへの適用が期待される。CTAS は PSI 上の ESP で開発されており、獲得された知識ベースは CTAS 用推論エンジンで実行されることになる。(なお、PSI 上への CTAS の実現は、山崎研修生 (NTT 横須賀通研) 及び、大崎氏 (日本情報処理開発協会) の協力のもとにおこなわれた。)

CTAS の知識獲得のためのコンサルテーションは、抽出、整理、リファイン、ルール変換、テストの 5 つからなる。抽出プロセスでは分類項目、特徴、関係度の抽出が行われ、整理プロセスでは抽出された内容を基に図・表が作成され獲得知識を視覚的に見ることができる。リファインプロセスでは分類項目、特徴、関係度のリファインが行われ、ルール変換プロセスではプロダクションルールの作成が、テストプロセスでは作成されたルールの評価が行われることになる。各プロセスは再帰的に実行される。

2 CTAS の評価内容

CTAS の評価方法としては再委託メーカが適当な問題を用意して、簡単な知識ベースを作ることにより評価をおこなった。そして例題の性質、獲得方法、獲

得結果等の評価項目を基に評価レポートを作成された。以下にシャープ、松下の評価協力メーカーごとに記す。

2. 1 シャープの評価

(1) 問題概要

例題は皿洗い器の故障診断であり、分類型問題である。診断知識は専門家によって作成されたチェックシートであり、KEを介して専門家の知識の一部であるチェックシートからどの程度の知識が抽出できるかを調べることになる。チェックシートの知識を説明すると、故障原因は電源不良や短絡等があり、全部で18項目ある。また、故障現象は洗浄中の水漏れ、ホースのひび割れ等があり、全部で50項目ある。皿洗い器の構造は下記である。

ア) 前面にドアがあり、箱型である。

イ) 後ろ面には給水部と排水部があり、それぞれ給水ホース・排水ホースで水道の蛇口・排水口につながれている。

ウ) ドアを開けると皿の洗浄部がある。底面にはヒーター、ノズル等があり、ノズルから水を噴射して皿を洗浄する。

エ) 正面には電源・セレクト・タイマーなどのスイッチ類がある。

(2) 知識獲得方法

知識獲得方法は専門家から対話によって抽出するのではなく、チェックシートをKEがみてからCTASを用いて知識ベースを構築することになる。知識抽出ではKEはCTASとの対話により故障原因を4つのグループに分類する。次に各グループごとにチェックシート中の50個の故障現象の内各グループの必要なものだけを選択して入力する。そして関係度(1、2、3、4、5の値をとる)の抽出であるが、現象がみられるものには5を、みられないものには1を、故障原因に直接関係のないものには3を割り当てる。2と4がないのはチェックシートからはそれらの値を読み取れないためである。リファインについては常識的な判断で抽出された知識の不備の解消が行なわれている。implication グラフに基づくリファインは行なわれていない。

(3) 評価結果

CTASの長所としては専門家が直接入力することが難しい知識(特徴間の類似度や包含関係などの知識)が要素間の関係度のような直感的な値として容易に入力可能であること、4種類の図・表が知識の抜け落ちの発見や特徴間の包含関係の理解に役立つこと、整理、リファイン、テストをシステムとの対話によって処理することができる点が挙げられている。

短所としては下記の2点が挙げられている。

- ア) 階層化分類を行う知識（ルール）にCF値が使えないこと
- イ) 順位付け知識の抽出時に階層化を行えないこと

前者については、「ブレーカーの落ちる頻度」といった性質の階層分類のための特徴についてはyes/noでデータを入力することは難しいからである。また一般エンドユーザは専門家が分かる特徴を理解できるとはかぎらないために、知識システムの使用時にyes/noのデータを入力することは難しいと考えられるからである。後者については専門家は完全に対象問題を理解していないので階層化分類と順位付けの境界がはっきりしているわけではない。そのため階層化分類が充分に行われていない場合には、順位付け知識の抽出時に階層構造の混入している場合がある。そこで順位付け知識の抽出時における階層化が必要になってくる訳である。

要望として上記の短所の改善が挙げられており、特にイ) については implication graphを用いることにより順位付け知識の抽出の際に、階層構造の抽出が行えるのではないかと言う指摘があった。

2. 2 松下の評価

(1) 問題概要

本の分類問題であり、ユーザが自分の思っている本をすぐに捜しだせることと、新たに購入した本を適切な場所に分類できることである。無作為に選択された30冊の本を分類するための知識獲得をCTASは行なうわけである。30冊の内容は自然言語関係、画像関係、計算機関係、認知科学関係である。但し、知識の獲得対象となるのは専門家ではない。

(2) 知識獲得方法

30冊の本の順位付けによる分類を行なおうとしたができなかった。これはグループの中から順に3つの要素を取りだして、それを2つに区別する特徴を聞いてくるという手法をCTASが使用しているため、予め分類に使用することを考えていた特徴がそのまま使えないことがあるからである。また30個の要素を分類する場合には56個の特徴の入力をCTASは求めてくることになり、すべて入力することが困難なためである。そこでNL(自然言語関係)、ZUKEI(画像関係)、COMPUTER(計算機関係)、NINTI(認知科学関係)、その他のグループで階層分類を二分木で行なった。次に、各グループごとに予め想定されている特徴の入力を行なう。そして関係度の値の入力を行い、リファインを行なう。

(3) 評価結果

長所として順位付けのための特徴抽出として、Personal Construct Theory手法を用いて、特徴発見を支援していることである。

短所としては下記が挙げられている。

- ア) 順位付けを行なうためには $2 * (N - 2)$ 個 (N はグループを構成する要

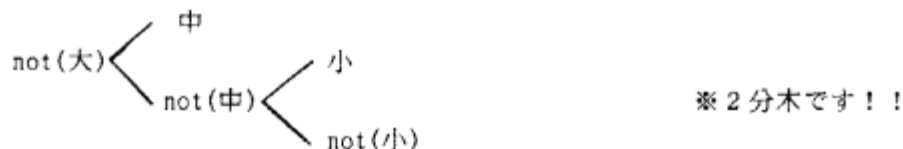
素の数)の特徴が必要であること、このため順位付けのための要素は5～6個以内であること

イ) 予めユーザが順位付けタスクのための特徴を予めもっている場合には
Personal Construct Theoryはいれずらいこと

ウ) 階層構造が2分木でなければならないこと

エ) 順位付けのための特徴を入力するときに同じ特徴を入力できないこと

ア) とイ) についてはPersonal Construct Theoryの欠点といえる。即ち、10個の要素があると、10個の要素を最初から3個ずつ表示して、16個の特徴の抽出を行なう。要素の数が少ない場合には要素を分類するための特徴を容易にいえるが、多くなるとなかなか思いつかなくなるためである。このため階層化が必要ないのに要素数が多いために無理にグループ分けしなければならない。また既に、人間が特徴について理解している場合には3つずつ要素を表示しなくても特徴の入力は可能なのに、機械的な質問を繰り返すからである。ウ) については階層化を2分木で行なうようになっているために、生じる問題である。階層化が“熱い／寒い”のような2値で行なえる場合には階層化の抽出が容易にできるが、“大／中／小”のような場合には、下記のように階層化する必要があるためこれが階層構造の抽出を困難なものとしている。



エ) についてはインプリメント上の問題であり、柔軟な知識抽出をやりずらくしている。

2. 2 CTASの改善案

今回評価した内容には CTASの知識獲得方法の問題からマンマシン・インタフェース、インプリメントレベル問題まで渡っている。よりパフォーマンスのよい知識ベースを構築するためには、階層化分類タスクの獲得問題、順位付けタスクの獲得問題が特に重要である。上記の問題点についての改善について下記に述べる。またCTASのデモンストレーションで指摘されているCF値の問題点についても述べる。

(1) 階層化分類タスクの獲得問題の改善

階層化分類タスクの獲得問題の改善には階層構造の抽出の柔軟性と階層化分類へのCF値の導入が必要である。そのためには順位付け知識の抽出の際に、gridやimplication graphの情報の利用による動的な階層構造の抽出支援が有効であると思われる。また階層構造を2値で行なうのではなく、複数個で階層化できるようにすることも必要である。階層化分類へのCF値の導入にあたって、階層化ルールへのCF値の割り振りをgridを用いて行なう。

(2) 順位付けタスクの獲得問題

順位付けのための特徴抽出を柔軟に行なうためには下記の3点が上げられる。

- ①入力すべき特徴がないときにはスキップが可能にすること
 - ②特徴抽出をトップダウンとボトムアップの2種類の方式を用意すること
 - ③特徴の入力時に同じ特徴の入力を可能にすること
- ①により特徴が思いつかなかった場合に、無理に要素をグループ化する必要はなくなる。②により予め入力すべき特徴が分かっている場合にはトップダウン方式により直接特徴の入力が可能になる。ボトムアップ方式とはPersonal Construct Theoryに基づく従来の手法による抽出方式のことである。

(3) CF値の問題

CTASによってプロダクションルールへ割り振られるCF値は特徴間の相対的な値であり、絶対的なものではない。このため実用レベルのCF値とするためにはデータによるCF値のリファインを行なう必要がある。そのためにはMOLE[Eshelman 87]で用いられているようなpreferencesを利用して専門家から間接的にCF値を獲得することが有効であると思われる。

3 おわりに

今回のCTASの評価は実際の専門家からの知識獲得という形式は行なえなかったもの、CTASの知識獲得における必要最低限度の問題点は明らかになったのではないと思う。今回明らかになった問題点を参考にしてCTASの機能拡張、及び問題解決知識の獲得システムEPSILON/OneとCTASの融合を行なっていく予定である。

[謝辞]

最後に、CTASの評価を行なった再委託先メーカーのシャープの研究員(勘座氏)、松下の研究員(池田氏、福重氏、小山氏)に感謝致します。

[参考文献]

- [Boose 84] Boose, J.: Personal construct theory and the transfer of human expertise., In Proceedings of the National Conference on Artificial Intelligence, Austin, Texas, 1984.
- [Boose 85] Boose, J.: Personal construct theory and the transfer of human expertise., Advances in Artificial Intelligence., North-Holland, 1985.
- [Eshelman 87] Larry Eshelman.
- MOLE: A Knowledge Acquisition Tool That Buries Certainty Factors, Proceedings of the 2ND KNOWLEDGE ACQUISITION FOR KNOWLEDGE-BASED SYSTEMS WORKSHOP, 6-0, 1987.
- [大崎 88] 大崎, 滝, 椿: 知識獲得支援システムEPSILON/One(2) - 専門家モデルの推論エンジンとPSI上でのインプリメント -, 第8回「知識工学シンポジウム」資料, 計測自動制御学会, 1988.
- [滝 87](1) 滝, 椿, 岩下 「知識獲得支援システム (EPSILON) における専門家モデル」

情報処理学会, 知識工学と人工知能研究会報告 52-4, 1987年5月

[滝 87](2) 滝, 椿, 藤井「知識獲得の為の専門家モデル」計測自動制御学会, 3 部会合同研究会「ヒューマンモデル」, 1987年7月

[滝 87](3) 滝, 椿, 藤井「EPSILON/EM: 専門家モデルを用いた知識獲得支援システム」情報処理学会 第35回全国大会, 1N-3, 1987年9月

[椿 88]椿, 滝, 大崎:知識獲得支援システムEPSILON/One(1) -異なる知識表現を用いたリファイン-, 第8回「知識工学シンポジウム」資料,

計測自動制御学会, 1988.

[山崎 87]山崎, 滝, 椿: 類型タスク構造に基づく知識獲得, 日本ソフトウェア科学会第4 回大会論文集, 1987.

[CTAS 評価レポート]

(1) シャープ株式会社の評価レポート

タイトル: CTAS の評価

(2) 松下電器株式会社の評価レポート

タイトル: 知識獲得支援システムCTAS の評価

5. CTASの評価

5.1. 対象問題

- 問題のタイプ
解析型（診断）
- 問題の概要
皿洗い器の故障診断

CTASを評価する問題を、皿洗い器の故障診断に設定した。この問題は次のような理由で選んだものである。

- 1) CTASは分類型問題を対象とした知識獲得支援システムである。
- 2) 故障診断は現実には、分類型のエキスパートシステムでよく扱われる問題である。
- 3) 手元に、皿洗い器の故障診断に関する資料が存在する（ここでは資料として、以前エキスパートシステムの講習会で用いたものを使用した）。

CTASは分類型問題向けのシステムであるが、扱う問題に向き、不向きがあると思われる。例えば、動物の分類問題のような、初期段階と末期段階がそれぞれ階層分類タスク、順位付けタスクに、ある境界をもって分別できる問題はCTASには適する。しかし、すべての分類型問題がこのような構造であるとは限らない。したがって、ここでは特にCTAS向きの問題を選択せず、現実によく使用されている診断型問題を設定した。

診断型問題を設定するにあたり、皿洗い器を選んだ理由は、手元に資料があり、規模もそれほど大きくないことがある。ただ、ここで問題となるのは、評価する筆者が皿洗い器の故障診断に関して専門家でないことである。存在する資料は、エキスパートシステムの講習会で使用した皿洗い器のチェックシート（付録参照）である。

今回の評価の立場は、CTASがKEの手作業を軽減する目的で作成されたシステムであるならば、専門家の知識の一部であるチェックシートから、KEに代わってどの程度の知識が抽出できるかを調べる、ということになる。もし、チェックシートだけからでは、抽出できない知識があれば、CTASはどのような知識を専門家からどのように引き出しているのかを知る手掛かりになる。このような立場から、今回の評価はチェックシートとCTASからエキスパートシステムの知識ベースを構築することを問題設定とした。

次に、ここで対象となる皿洗い器の構造を簡単に示す。

- 1) 前面にドアがあり、箱型である。
- 2) 後面には給水部と排水部があり、それぞれが給水ホース・排水ホースで水道の蛇口・排水口につながれている。
- 3) ドアを開けると皿の洗浄部がある。底面にはヒーター、ノズル等があり、ノズルから水を噴射して皿を洗浄する。
- 4) 正面には電源・セレクト・タイマーなどのスイッチ類がある。

以上のような皿洗い器に対して、故障診断を行うシステムを構築することが目的である。

5. 2. CTASの用途

5. 2. 1. 使用予定概要

資料では、皿洗い器の故障原因として下記のごとく18項目がリストアップされている。これらをCTASの要素項目として与え、グループ分け、要素の分類を行う。故障原因（18項目）を以下に示す。

故障原因

- | | | |
|---------|------------|-----------|
| ・電源不良 | ・電気容量不足 | ・短絡 |
| ・タイマー | ・セレクトスイッチ | ・ドアスイッチ |
| ・注水側バルブ | ・注水側バルブコルク | ・ホースと締め金 |
| ・ポンプシール | ・ドアガスケット | ・フロート |
| ・オーバサージ | ・バルブ開放 | ・ノズル |
| ・ろ過器 | ・排水バルブ | ・排水バルブコルク |

知識の抽出は、まずこれらの故障原因を数個のグループに分類する。そして、各グループごとに、チェックシートに基づいて要素項目の抽出を行う。要素項目には故障現象が対応し、故障現象はチェックシートに50項目書かれてある。実際診断に使用するのはこれらの一部である。

チェックシートの内容をもとに、抽出した要素項目の順位付けを行い、整理、リファインする。

5. 2. 2. 獲得予定内容

筆者がチェックシートの内容をもとに、18項目の故障原因をグループ分けした。この作業は専門家がCTASを使用する際は簡単に処理できる。それぞれのグループの特徴が最初に獲得される知識である。

次に、それぞれのグループで特徴の抽出を行う。チェックシート中の故障現象のうち、必要なものだけが選択される。これらが第2に抽出される知識である。

ここで抽出された特徴と要素間の関係度を与え、知識の整理、リファインを行う。第3の知識は、これらの特徴・要素間の implication graph, similarity, cluster tree 等で表現されている関係度の知識である。

5. 3. CTAS 評価

5. 3. 1. 獲得方法

まず、18項目の故障原因を以下のように、グループ分けした。

- (1) 電源投入不可
- (2) 注水の異常
- (3) 洗浄時の異常
- (4) 排水の異常

次に、それぞれのグループ内で順位付けタスクを行った。特徴及び関係度の抽出は資料の『皿洗い器メンテナンスシート』（付録参照）に基づいて決定した。このシートは故障に関しての現象とその有無を記したものである。その現象が見られるものには5、見られないものには1、その現象が故障原因に直接関係の無いものには3をそれぞれ関係度の値として与えた。また、知識の整理・リファインは、筆者の常識的な判断で行った。類似度の高いもの、満足のできない implication グラフについては、リファインで修正した。この際、要素をグループ内に制限すれば成立する implication グラフは修正しなかった。

5. 3. 2. 獲得内容

それぞれのグループで獲得された知識を表1に示す。

表1. 各グループの要素・特徴間の関係度

(1) 電源投入不可

	電不	短絡	容不
皿洗い器稼働時のブレーカー on/off	5	5	1
電源供給 正常/異常	1	3	1
複数の電気機器稼働 可/不可	3	5	5

電不(電源不良) 容不(電気容量不足)

(2) 注水の異常

	注コ	注バ	ドア	セレ	タイ
ブレーカーの落ちる頻度 多/少	1	1	5	5	5
タイマーの動作 正常/異常	3	3	5	5	1
セレクトスイッチ 正常/異常	3	3	5	1	3
注水側バルブのゴミ 有/無	1	5	3	3	3

注コ(注水側バルブコルク) 注バ(注水側バルブ) ドア(ドアスイッチ)

セレ(セレクトスイッチ) タイ(タイマー)

(3) 洗浄時の異常

	ろ過	ノズ	バル	オー	フロ	ドア	ボン	ホー
洗浄中の水漏れ 有/無	1	1	1	1	1	5	5	5
ホースのひび割れ 有/無	3	3	3	3	3	1	1	5
ポンプの水のしみだし 有/無	3	3	3	3	3	1	5	3
汚れ落ちの悪さ 有/無	1	1	5	5	5	3	3	3
フロート部分のゴミ 有/無	3	3	1	1	5	3	3	3
1回の洗浄量 正常/過多	3	3	5	1	3	3	3	3
ノズルの水のムラ 有/無	3	5	3	3	3	3	3	3

ろ過(ろ過器) ノズ(ノズル) バル(バルブ解放) オー(オーバサージ) フロ(フロート) ドア(ドアスイッチ) ボン(ポンプシール) ホー(ホースと締め金)

(4) 排水の異常

	排コ	排バ
排水側バルブのゴミ 有/無	1	5
排水側電磁弁の動作 正常/異常	1	3

排コ(排水バルブコルク) 排バ(排水バルブ)

5. 3. 3. CTASの長所

各グループ内の要素間の関係を rating grid 値を与えることで抽出することができる点が評価できる。専門家は特徴間の類似度や包含関係などを知識として把握しているが、これらを直接抽出することは難しい。それを、直感的な値を入力することで可能としたことはCTASの大きな長所と言える。

また、rating grid から類似度、包含関係等を計算し視覚的に表示する点が、知識を整理するうえで役に立った。全く異なる2つの要素について、ある特徴が欠落したために類似してしまうようなことを簡単に発見することができる。また、implication graph の表示により、特徴間の包含関係が分かるので、理解し易い。

さらに、整理・リファインとテストをインタラクティブに処理できることも利点である。これらを繰り返すことで、より良い知識ベースが構築されるであろう。

5. 3. 4. 問題点

第1の問題点として、階層分類の方法があげられる。専門家にとってグループ分けするための特徴は理解しやすいため、要素をグループ分けする作業は容易なことである。かし、エキスパートシステムを使用するユーザにとっては、専門家が分かる特徴を理解できるとは限らない。いきなり最初に yes とも no とも答えられないような質問をされることもある。このような時、ユーザがCF値を入力することができ、CF値によってシステムがグループを決定できるような機能が望まれる。

第2の問題点として、グループ内に階層構造を抽出できない点である。グループの中でも、明確に yes, no では分けることはできないが、階層構造が存在することもある。皿洗い器の故障診断の例では、注水異常のグループで特徴「ブレーカーの落ちる頻度」が多いか少ないかで、二つのグループに分けられている。ブレーカーの落ちる頻度が極めて少ない場合、タイマーの動作とセレクトスイッチには関係が無く、注水側バルブのゴミについてののみ、質問すれば質問項目が減ることになる。ただ、今回は、非専門家が評価を行ったという点で、階層分類の段階でグループ分けが十分行われなかったために、グループ内に階層が生じたのかもしれない。しかし、専門家でも完全に問題の階層構造を把握しているわけではなく、順位付けレベルに階層構造が入り込んでしまうことも考えられる。また、階層分類しようとしても「・・・の頻度」という特徴を yes, no の二者択一で判断するのは難しい。階層分類におけるCF値導入とともに、順位付け時の階層構造の抽出はできないものだろうか。

5. 3. 5. CTASへの要望

CTASの問題点として、階層分類における柔軟性と、順位付けにおける階層構造の取り扱い方を取り上げた。このことは、CTASの目指す階層分類と特徴分類の分離と相反するものかもしれない。しかし、故障診断の例から判断すれば、階層分類と特徴分類については、あいまいな情報と確定的な情報のどちらも扱えたほうが便利であると思われる。

CTASへの要望としては、

(1) 階層分類へCF値の取り入れ

(2) 順位付へ時の階層構造の抽出

の二点があげられる。(1)についてはCTAS 2で検討中ということなので是非実現して欲しい

(2)についてもう一度整理すると、専門家が階層分類時には気付かなかった階層構造をCTASが rating grid 値をもとに抽出できないかということである。階層分類と順位付けの境界ははっきりしないため、このようなことは、現実によく起こることではないかと感じた。階層構造が見付かれれば、それを階層分類に吸い上げることで、候補の絞り込み、システムの質問回数の減少につながり、効率良い処理が行える。

階層構造の抽出の方法は、implication graph を利用する。ある特徴に対して、そのグループ内の他のすべての特徴からの包含関係があれば、階層関係とみなす。CTASはそれをグラフで示し、ユーザに階層関係があることを提示する、ということができないのだろうか。

例として、皿洗い器の注水異常のグループの implication graph を図1に示す。タイマーの動作・セレクトスイッチ・注水側バルブのゴミのいずれの特徴からもブレーカーの落ちる頻度への包含関係が存在することが分かる。これはすなわち図2のような階層関係が成立することを意味する。同様に、洗浄時の異常のグループでは図3の implication graph から図4の階層関係が得られる。

5. 3. 6. まとめ

今回の評価では、CTASの階層分類と順位付けの境界を中心に考察し、獲得されない知識として、順位付け内の階層構造を取りあげた。階層構造の抽出の必要性やその方法については、議論の余地があると思われる。

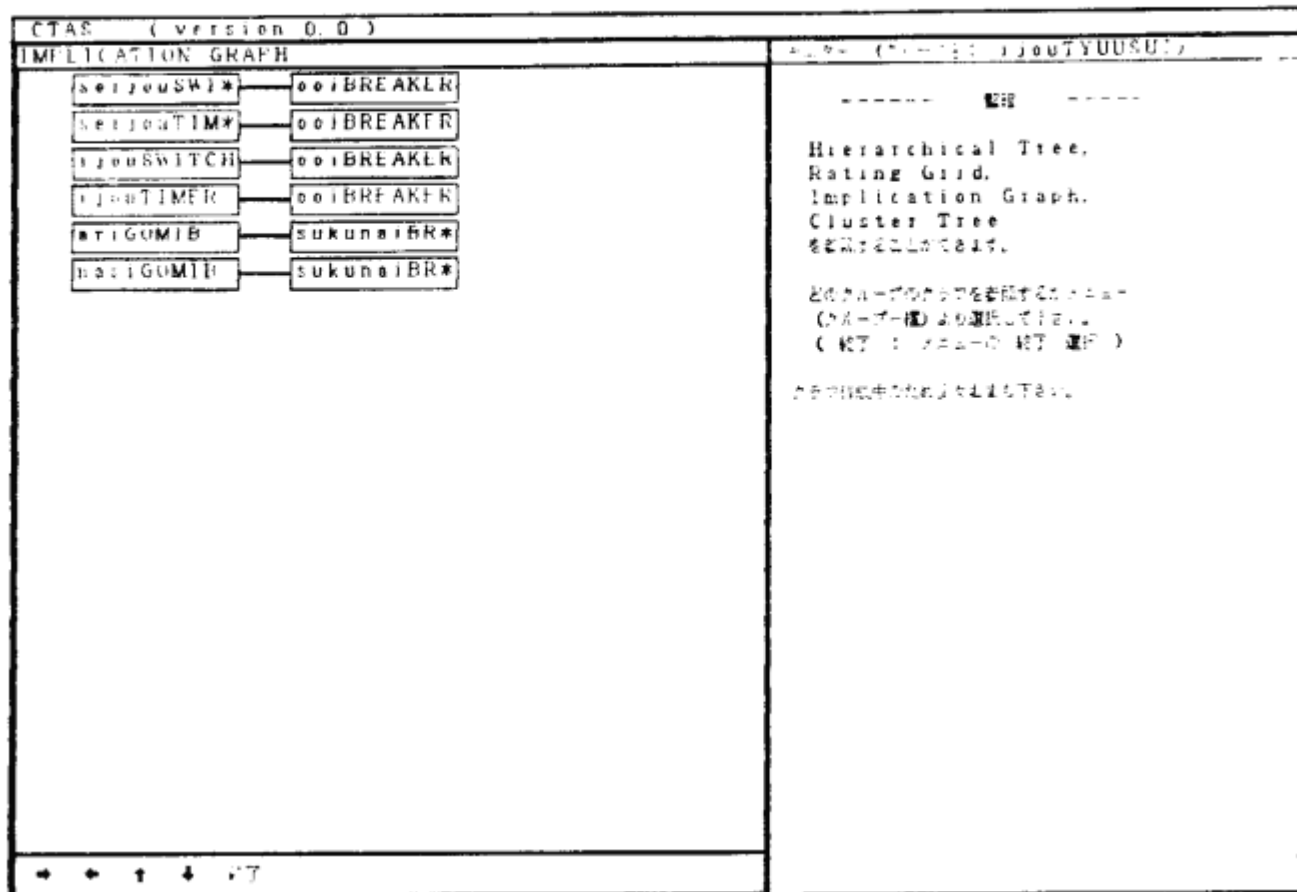


図1 注水異常グループの implication graph

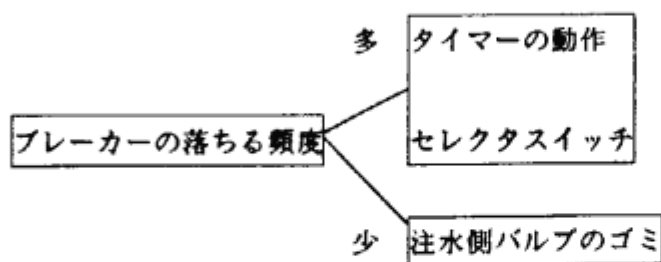


図2 注水異常のグループの階層図

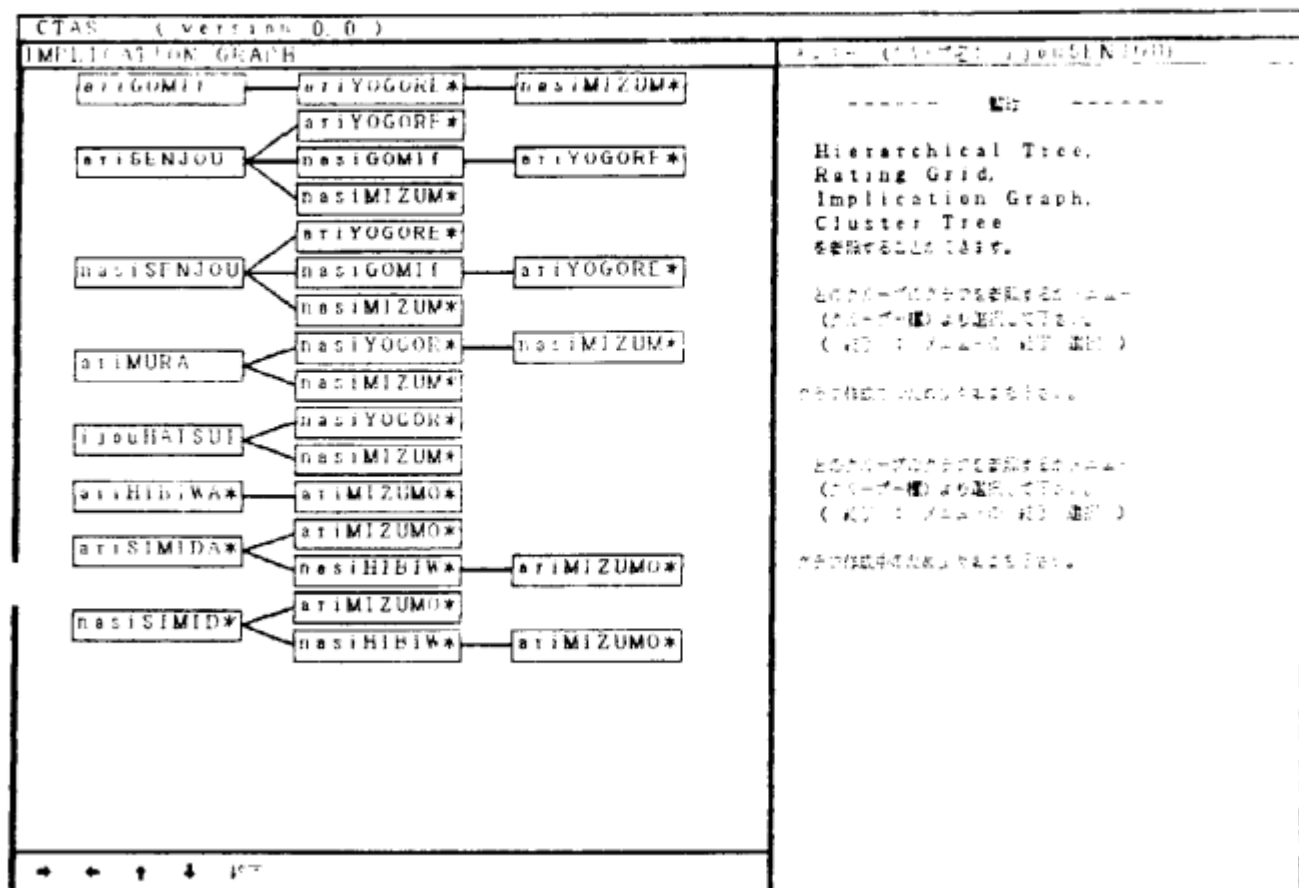


図3 洗浄部異常グループの implication graph

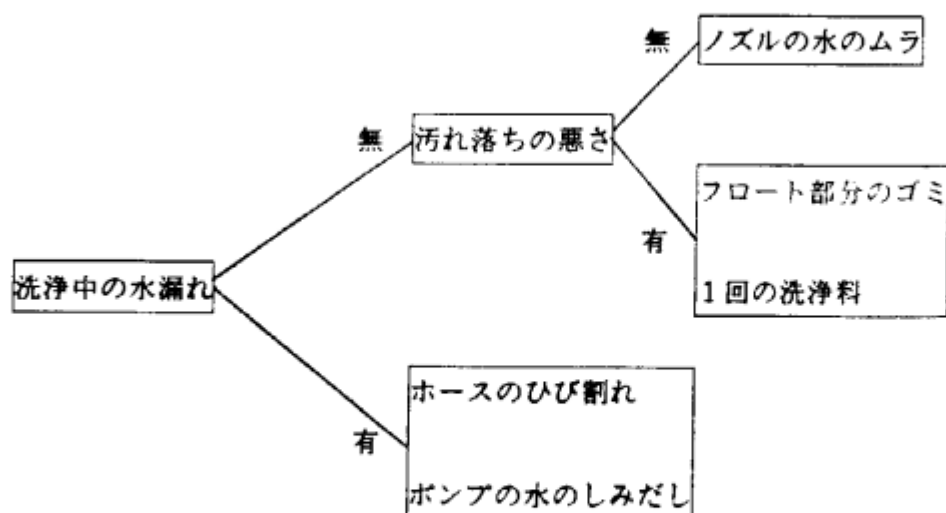


図4 洗浄時異常グループの階層図

皿洗い器メンテナンスシート

対 処 日	年 月 日 ()	客 先		
受 付 内 容				No.
CHECK LIST				
1. 皿洗い器の不自由 2. 皿洗い器の清掃 (1年以内) 3. 皿洗い器の電源投入 4. アレカの落ちる頻度 5. 皿洗い器稼働時のアレカ 6. 複数の電気機器稼働 7. 電源供給 8. 動作不可の電気機器 9. 電気機器の動作 10. テレビの画面 11. テスターによる導通テスト 12. タマの動作 13. タマのまわりの汚れ 14. 電源の電気容量は15 [A] 15. 皿洗い器レクタスイッチの汚れ 16. 皿洗い器以外の大電力消費 17. 皿洗い器の注水 18. 皿洗い器ホーススイッチの汚れ 19. 皿洗い器セクタスイッチの機能 20. 皿洗い器ホースの閉閉 21. 皿洗い器ホースの汚れ 22. 皿洗い器注水側バルブのゴミ 23. 皿洗い器注水時異常の時期 24. 皿洗い器注水側電磁弁動作 25. 皿洗い器洗浄中の水漏れ	有・無 有・無 可・不可 多・少 ON・OFF 可・不可 正常・異常 有・無 正常・異常 正常・異常 導通有・導通無 正常・異常 有・無 以上・未満 有・無 有・無 正常・異常 有・無 正常・異常 正常・異常 有・無 有・無 正常・異常 正常・異常 有・無 突然・徐々に 正常・異常 有・無	26. 皿洗い器ホースのひび割れ 27. 皿洗い器ホース締め金具の錆び 28. 皿洗い器ホース締め金具取替え 29. 皿洗い器ホースまわりの水のしみ出し 30. 皿洗い器ホースまわりの水のしみ出し 31. 皿洗い器ホースアサルトの弾性 32. 汚れ落ちの悪さ 33. 皿洗い器ボート部分のゴミ 34. 皿洗い器ボートの位置 35. 皿洗い器1回の洗浄量 36. 皿洗い器電源波形 (計測機器使用) 37. 皿洗い器と使用地域の電源周波数 38. 皿洗い器注入サイクル後の注水 39. 皿洗い器洗浄中の水位 40. 皿洗い器洗浄中のノズルからの噴射 41. 皿洗い器ノズルからの水の量 42. 皿洗い器ノズルのゴミ 43. 皿洗い器内部の清掃頻度 44. 皿洗い器ろ過フィルターのゴミ 45. 皿洗い器ろ過フィルターの取替え 46. 皿洗い器洗浄後の排水 47. 皿洗い器洗浄後の排水時間 48. 皿洗い器排水側バルブのゴミ 49. 皿洗い器排水バルブの折れ曲がり 50. 皿洗い器排水側電磁弁動作	有・無 有・無 最近・長期間経過 有・無 有・無 有・無 有・無 有・無 正常・異常 正常・異常 正常・異常 一致・不一致 有・無 定位・異常 正常・異常 有・無 有・無 多・少 有・無 最近・年数経過 正常・異常 正常・異常 有・無 有・無 正常・異常	
故障 原因	バルブ開放		備 考	

知識獲得支援システムCTASの評価

池田光生 福重貴雄 小山隆正
松下電器産業 技術本部
東京研究所 第一研究室

1. はじめに

ICOT(新世代コンピュータ技術開発機構)で開発された知識獲得支援システムCTASの評価を行った。CTAS(Classification Task Acquisition System)は分類問題に必要なタスクとして階層化分類と順位付けの2つのタスクを持ち、初期知識ベースを専門家との対話により構築する知識獲得支援システムである。

CTASはエキスパートシステムの知識情報処理のうち分類タスクについての知識獲得支援を考えている。分類タスクは、さらに階層分類タスクと順位付けタスクとに分けられる。階層分類とは、例えば動物を哺乳類とそれ以外に分けるように、ある特徴を用いてグループ分割を行うことである。順位付け分類とは、分割できなくなったグループに対して、グループ内の各要素と関係のある特徴を用いて順位付けすることである。

例えば、チータ、ヒョウ、トラ、ライオンを順位付けするための特徴として、足の速さ、体の模様、体の大きさなどがある。

今回、評価を行うために設定した対象問題は、当室が所有する図書の分類である。短期間で評価を行うために、設計型、あるいは診断型のエキスパートシステムを作成することはせず、分類のみとした。はじめに、当室が所有する図書の中から無作為に30冊を選び、これを、いくつかの特徴と本との関係度による順位付けによって分類することを考えた。しかし、30冊の本を、順位付けすることは、非常に困難であることがわかった。そこで、階層化することによっていくつかの小グループに分けて、各小グループ内での順位付けを行った。これについては、3章で詳しく述べる。

CTASの評価を行った結果、CTASはエキスパートシステム作成のための知識獲得支援システムとしては、いくつかの問題点があり実用的でないことがわかった。最も大きな問題は、順位付けを行うために必ず(グループを構成する要素の数-2)個の、特徴を必要とすることであった。また、階層構造が必ず二分木でなければならない点も大きな制約となった。

2章で、CTASの概要を述べる。3章で、CTASの評価に用いた対象問題、CTASの使用概要、CTASの評価について説明する。4章で、結論を述べる。

2. CTASの概要

CTASは、分類問題に必要な初期知識ベースを構築するための知識獲得支援システムである。分類問題のタスクは、階層分類と特徴による順位付けとからなる。一般的に、分類の初期段階では階層分類が行われ、末期段階では順位付けが行われる。

例えば、医療診断では、「頭が痛い」、「おなかが痛い」などの症状によって、「頭痛を伴う病気」「腹痛を伴う病気」のように病名を分類することができる。これは、病名からなる集合を症状という特徴を用いてグループ分割しているといえる。

分類活動の初期段階では、分類すべき要素の数が多いために要素間にグループ分割できるような特徴が多いと考えられる。ところが、分類の進んだ末期段階では性質の似ている要素を分類しようとしているので、それらをグループ分割できるような特徴が存在しなくなる。

このような要素を分類するために順位付けを行う。順位付けは、分類しようとする要素すべてと何らかの関係がある特徴を用いて行われる。各要素と特徴との関係度の強さは要素によって異なる。このことを利用して、順位付けを行うことができる。

例として、チータ、ヒョウ、トラの3つを、走る速さ、体の大きさという特徴で順位付けすることを考える。チータ、ヒョウ、トラが要素であり、走る速さ、体の大きさが特徴である。これらの要素と特徴が図2-1の関係を持っているとする。CTASでは、要素と特徴との関係度を5段階評価で表わす。チータ、ヒョウ、トラのそれぞれが持つ

ベクトル (5, 3), (4, 3), (3, 4) の計算によって8つの要素を順位付けすることができる。

チータ	ヒョウ	トラ	要素/特徴
5	4	3	走る速さ
3	3	4	体の大きさ

図2-1

CTASは、抽出、整理、リファイン、テストの4つのプロセスで知識を獲得する。

CTASの抽出プロセスには、ボトムアップ式とトップダウン式がある。ボトムアップ式ははじめに要素を列挙してからこれをグループに分割することによって階層分類し、その後に順位付けを行う。トップダウン式は、階層構造がわかっている場合に用いる。はじめに、階層構造の枠組みを作り出してから各グループに属する要素を抽出しそれを順位付けする。

CTASでは、順位付けに必要な特徴を抽出するためにPersonal Construct Theoryを用いている。これは、要素の中から3つを取り出し、その中の2つに共通する特徴を考えることによって順位付けに必要な特徴を挙げていくやり方である。

例として、a, b, c, d, eの5つの要素を順位付けすることを考える。CTASは、初めにa, b, cの3つの要素をユーザに提示して、そのうちの2つに共通する特徴を尋ねる。つぎにb, c, dの3つを提示して同様に特徴を尋ねる。こうして、CTASは順位付けする要素の数-2個の特徴をユーザに入力させる。

整理プロセスでは、グループ間の階層木やグループ毎の要素と特徴の関係表を作成する。さらに、関係表をもとにして特徴間の関係を表わすグラフ(Implication Graph)や特徴間、要素間の類似度を示すグラフ(Cluster Tree)を作成する。

リファインプロセスでは、整理プロセスで作成した表・グラフをもとにして、階層木の修正、関係表の値の修正、あらたな要素、特徴の追加などを行って知識をリファインする。

テストプロセスでは、階層木、関係表をもとにして、3種類のルールを作成する。作成されたルールはCTAS用推論エンジンで実行される。ここではテスト対象となっているグループに属する要素を一つ想定して、その要素について、特徴の確信度を-1.0~1.0の間で入力する。CTASは先に作成されたルールを実行して、想定された要素に最も近いと思われる順にグループの要素をソートして表示する。

CTASでは、以上の4つのプロセスを繰り返すことにより知識ベースを構築していく。

3. CTASの評価

3-1. CTAS評価のための対象問題

CTASの評価を行うために、当室の所有する図書の分類を行った。短期間で評価を行うために、特にエキスパートシステムの作成は行わず、図書の分類を行うだけとした。目的として次の二つを想定した。第一は、自分の思っている本をすぐに捜しだせることであり、第二は、新たに購入した本を適切な場所に分類できることである。

方法は以下の通りである。

はじめに、当室が所有する図書の中から30冊を無作為に抽出した。次に、それらを順位付けによって分類するために必要な特徴を設定した。特徴の数は、7種類になった。

ところが、実際にCTASを使用してみると、30冊の本を順位付けするためには28種類の特徴を与えなければならないことがわかった。これは、CTASが順位付けしようとするグループから3つの要素を取り出し、それらに対して1つの特徴を必要とするためである。この問題点については後述する。順位付けを行うためには、要素の数が5~6個以内であることが望ましい。そこで、CTASの階層化の機能を用いて、要

素の数が5～6個のグループに分け、それぞれについて順位付けを行った。グループは、NL（自然言語関係）、ZUKEI（画像関係）、COMPUTER（計算機関係）、NINTI（認知科学関係）、とその他の、計5つとした（図3-1）。

CTASの知識獲得の4つのプロセスを実行して、どのような知識ベースが構築されるかを調べることにした。

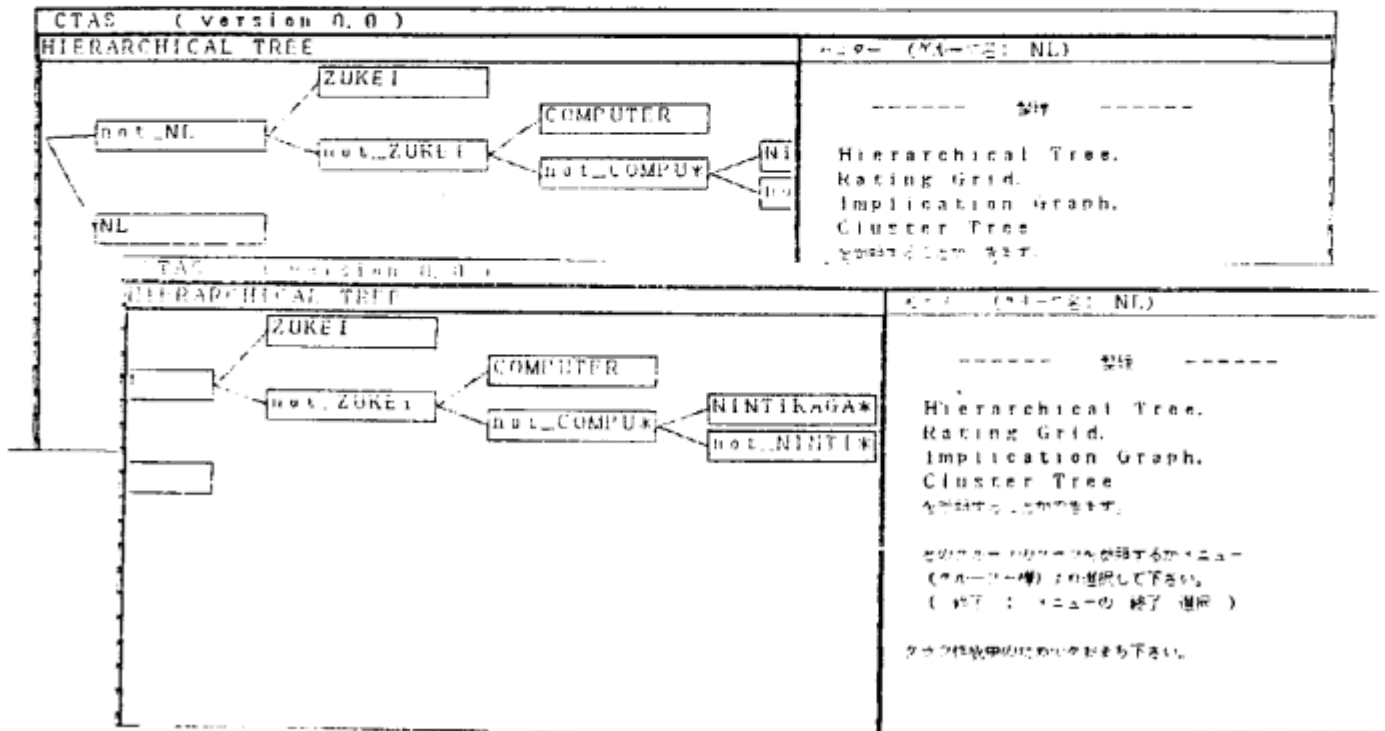


図3-1

3-2. CTASの評価

3-2-1. 知識獲得の方法と内容

まず、順位付けが可能なグループを作るために階層化を行った。階層構造は二分木でなければならぬため、はじめにグループNLとグループnotNLに分け、さらにnotNLをZUKEIとnotZUKEIに分けるというようにしていった。

次に、各グループの要素を5～6個ずつ決めていった。また、順位付けに必要な特徴もある程度決めておいて、順位付けのための特徴を抽出しながら適当に当てはめていった。当初は、あらかじめ特徴を決めて順位付けするつもりであったが、この方法ではむずかしいことがわかり、ある程度特徴の候補を上げておいてその場で適当に当てはめていくことにした。

これは、CTASではシステムがグループの中から順に3つの要素を取り出して、それを2つに区別する特徴を聞いてくるという方法を採用しているため、あらかじめ考えていた特徴がそのまま使えないことがあるからである。しかし、同じ特徴を二度続けて入れることはできないために、どの時点でどの特徴名を使うかということが重要であり、ある程度特徴名を考えておくことは必要である。

以上のようにして、各グループについて特徴とそれぞれの特徴に対する各要素の値を与えた。

例として、グループNLについて、説明する。NLの各要素と順位付けのための特徴との関係表を図3-2に示す。各要素には、番号が付けられている。NLの1番は「自然言語の処理」である。CTASは、はじめに要素1, 2, 3をユーザに提示してそのうちの2つに共通する特徴を尋ねる。ここでは、「自然言語の処理」、「A Grammar of

Anaphora」, 「The Philosophy of Language」のうちの2つに共通する特徴を尋ねてくる。ここで, gengogaku と答えると, その反対の特徴は何かと尋ねてくるので keisanki と答える。これは, 2つに共通する特徴が「言語学に近い」ことであり, その反対の特徴が「計算機科学に近い」ことであることを示している。

同様に, CTASは要素2, 3, 4をユーザに提示してそのうちの2つに共通する特徴を尋ねる。このようにして, ユーザは4つの特徴をCTASに入力することになる。

特徴名の入力が終わると, rating値の入力を行う。rating値は5段階評価で入力する。図3-2は, グループNLの要素「象は鼻が長い」が言語学, 文法, 日本語に近く, 意味論か構文論かに関してはどちらともいえないことを示している。

以上でCTASの抽出プロセスを終了する。同様のことを他のグループについても行う。

CTAS (version 0.0)		メニュー (グループ: NL)
RATING GRID		
1 2 3 4 5 6		----- 整理 -----
2 5 5 5 5 5	: gengogaku / keisanki	Hierarchical Tree.
3 5 5 1 5 1	: ippan / nihongomuke	Rating Grid.
2 3 5 2 5 3	: imiron / koubunron	Implication Graph.
3 4 2 5 1 5	: bunpou / not_bunpou	Cluster Tree
1 : sizengengo_shori		を参照することができます。
2 : grammar_of_anaphora		とのグループのグラフを参照するメニュー
3 : philosophy_of_language		(グループ=) を参照して下さい。
4 : nihonjinha_donna		(終了 : *メニューの終了 選択)
5 : nazasito_hituzensei		
6 : zouha_hanaga_nagai		クラスタリングのために参照して下さい。

図3-2

次に整理プロセスを行う。整理プロセスでは, 抽出プロセスで入力したデータを使って, 要素と特徴間の関係表, 特徴間の関係を表わすグラフ (Implication Graph) や, 特徴間, 要素間の類似度を示すグラフ (Cluster Tree) を作成する。

図3-3にNLのImplication Graphを示す。図3-3では, NLの要素で文法という特徴を持っているものは, 言語学という特徴も持っていること, 日本語向けという特徴をもっているものはやはり, 言語学という特徴も持っていることなどを示している。線の太さが太いほど結び付きが強い。

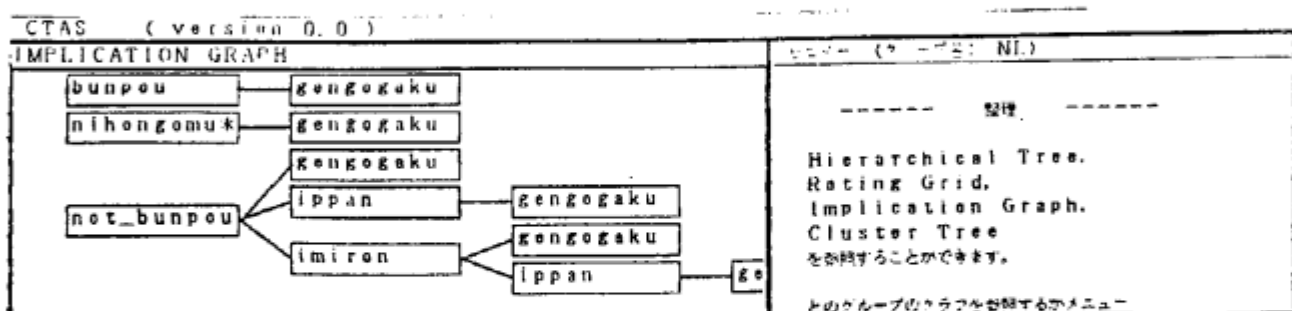


図3-3

図3-4に Cluster Tree の例を示す。図3-4は、「日本人はどんな文を使っているか」と「象は鼻が長い」が類似していること、「The Philosophy of Language」と「名指しと必然性」が類似していることなどを示している。

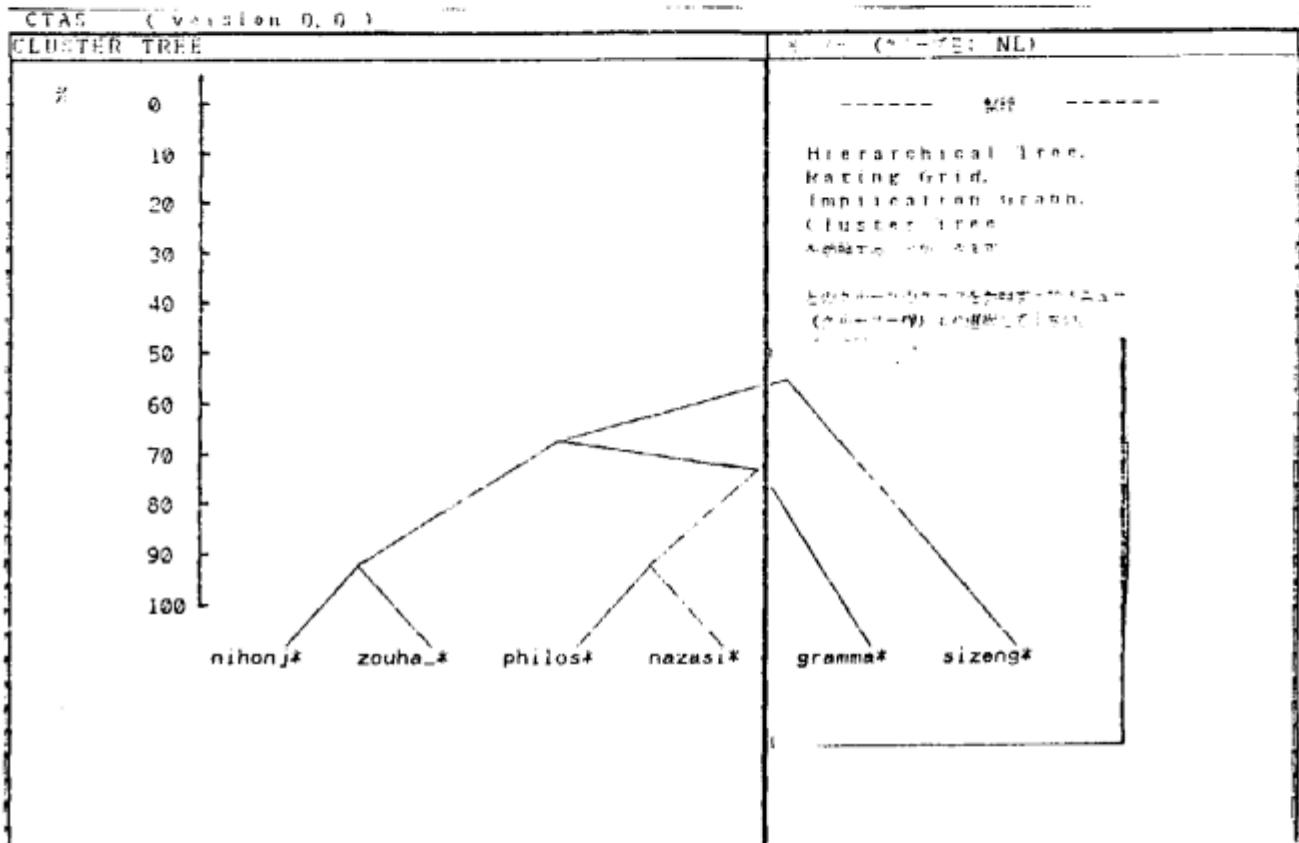


図3-4

次に、整理プロセスで得られたグラフをもとにして、階層木の修正、関係表の値の修正、あらたな要素、特徴の追加などを行って知識をリファインする。

次のテストプロセスでは、こうして出来上がった知識ベースを使ってテストをする。CTASは、まず、階層木と関係表をもとにして、3種類のルールを作成する。第一は、各要素がどのグループに属するかを示すルールである。第二は、各特徴と要素間の関係を示すルールで、CF値付きのルールである。第三は各特徴間の関係を示すルールである。作成されたルールはPSI-PROLOG上のCTAS用推論エンジンで実行される。テストプロセスの例を図3-5に示す。

図3-5では、ある要素を想定して、それが"robot(1.0)"に近いか"ningen(-1.0)"に近いかという問いに対して、きわめて人間に近いと答えている。文法的かどうかという問いに対しては-0.8くらい文法的であると答えている。さらに、日本語向けか一般の言語向けかという問いに対しては0.4くらい日本語向けであると答えている。

以上の結果、最も確信度の高い要素は「日本人はどんな文を使っているか」と「象は鼻が長い」であり、確信度は0.566であることが分かる。

prolog_debugger_45	Variable
<pre> Element group is set Please input yes or no nl_ka'is 'A(yes) or not_NL(no) ? >> yes. This group name is A Group decision is over Next, elements are set in order of CF value Please input CF value(-1.0~1.0) robot_human'is 'robot(1.0)/ningen(-1.0)/unknown(0.0) ? >> -1.0. grammar_ka'is 'not_bunpou(1.0)/bunpou(-1.0)/unknown(0.0) ? >> -0.8. which_lang'is 'nihongomuke(1.0)/ippan(-1.0)/unknown(0.0) ? >> 0.4. The result of test is below toshio'is 'nihonjinha_donna(0.566) toshio'is 'zouha_hanaga_nagai(0.566) toshio'is 'grammar_of_anaphora(0.294) toshio'is 'philosophy_of_language(0.186) toshio'is 'nazasito_hituzensai(0.122) toshio'is 'sizengengono_shori(-0.199) </pre>	

図3-5

3-2-2. CTASの問題点

最も大きな問題は、順位付けに必要な特徴の数が決められているということである。そのために、一つのグループの要素の数は5~6個に制限される。このことは、対象問題によっては、大きな制約となる。

図書の分類では当初30冊の本を順位付けしようとした。しかし、この方法はうまくいかないことが分かり、5~6冊のグループに分けた。30冊の本を順位付けするためには28個の特徴を要求されるからである。また、整理プロセスにおけるグラフの表示もうまくいかないようである。

どうしても数十個の要素を順位付けしなければならない分類問題も、存在するであろう。階層木によるグループ分けを行うことによって、必ず小さなグループに分割できるとは限らない。順位付けのための特徴の入力の仕方はもっと柔軟である必要がある。

CTASでは、順位付けのために3つの要素を提示してその3つの要素を2つに区別するための特徴をユーザに問い合わせる方法を採用している。

しかし、以前に入力した特徴名を再び使いたい場合や、特徴名を思いつかない場合がある。このようなときには、特徴名の重複のチェックをしたり、必ずしも特徴名を答える必要のないようにしたりすればよいであろう。

3つの要素を提示して特徴名を答えるというやり方は特徴名を思いつくためにはよい方法であるが、あらかじめ特徴名を考えている場合には、この方法はかえって邪魔になった。あくまで、ユーザが順位付けするための特徴を考えつくための手段としたらよいであろう。

順位付けのための特徴の数の最小値が決められているため、抽出プロセスの特徴名やrating値の入力が煩わしく感じる。CTASでは一通りの知識を獲得してからでないと整理プロセスやリファインプロセスに入れないようになっているが、不完全な知識を作ってから、少しずつ特徴名やrating値を追加していくようにしてはどうであろうか。

4. まとめ

図書の分類を対象問題として，CTASの評価を行った。その結果，次の点に問題があることがわかった。第一に，順位付けを行うことのできる要素の数が少ない。第二に順位付けのための特徴を入力するときに同じ特徴を入力できない。第三に，特徴の数を自分で決められない。

CTASの長所として，順位付けのための特徴発見の支援方法がある。3つの要素を提示してそれらを2つに区別するための特徴を考えることによって順位付けを行うという方法は，順位付けのための特徴を考え出す上で強力な方法である。