

TR-360

類型タスク構造に基づく知識獲得

椿 和弘、滝 寛和、藤井裕一、山崎毅文

March, 1988

©1988, 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における知識獲得支援システムの一列: CTAS —
 < Classification Task Acquisition System >

ニキスパートシステムにおける重要なタスクの一つである分類タスクは、その初期段階においては若弱分類タスク、その末梢段階においては順位付けタスクからなっていると考えられる。本研究において、この分類型タスク構選を考慮した分類型問題向け知識獲得支援システムCTASをP S Iマシン上に構築した。

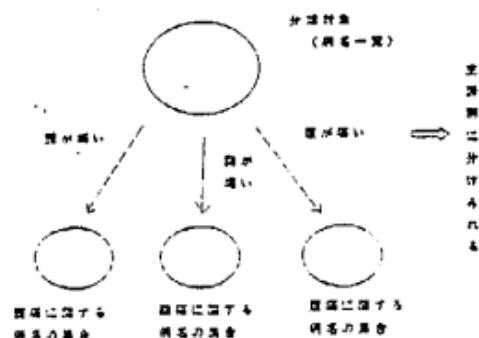
知識獲得支援システムは、知識ベースの構築の大きな問題である知識獲得ボトルネックの解消を目的として、研究開発が進められている。知識獲得の範疇としては、学習システム、知識ベース管理システムも入るが、本文では、インタビュースystemとして、開発した知識獲得支援システムCTASについて説明する。効率的な知識獲得には、獲得対象に合った枠組みが先行知識として必要である。これは、ある使用目的で知識ベースを構築しようとする、その使用目的とは何であるかを知識獲得支援システムに先行知識として与えておく必要がある。知識獲得支援システムは、先行知識をもとに知識獲得を行うことができる。そのような枠組みは、使用目的の知識表現や問題解決の構造をもったものでなくてはならない。その枠組みの候補として、ドメインモデル [Kahn 85] や類型タスク [Chandrasekaran 86] が考えられる。また、専門家が直接理解できる知識表現が望ましい。知識表現のレベルとしては、最下層に「インプリメンテーションレベルの知識表現」(例：プロダクションルール)があり、中間層に「タスクレベルの知識表現」(例：類型タスク)があり、そのうえに上位層として「アプリケーションレベルの知識表現」がある。アプリケーションレベルの知識表現はアプリケーション対応の表現であり、各専門家が持つ知識表現を意味する。しかし、知識獲得以前から色々な知識源について、そのレベルの知識表現を知識獲得支援システムが持つことは難しい。このレベルの知識表現は、獲得後に判明するものである。そうすると、タスクレベルの知識表現を知識獲得支援システムに持たせることが考えられる。本システムはそのような考えから分類タスクにおける類型タスクを考慮して開発した。

エキスパートシステムは、その質的構造により、診断問題等の分類型問題と設計問題等の合成型問題とに大別される。このうち、分類型問題の類型タスク構造は、ある程度明確になりつつある。そのタスク構造は、階層分類 (Hierarchical Classification) と特徴による順位付け (Ordering by Specification) とから成ると考えられる。例えば、MYCINのようなシステムは、病名の分類をCF値による順位付けタスクによって行っていると考えられるし、また機械の故障診断のような症状と故障原因とが一意に結び付くようなシステムは、診断木の探索によって解を探すタスクと考えられよう。また、一般的に、分類の初期段階では階層分類を、末期段階では順位付けタスクが多い。分類活動の初期段階では、分類すべき項目数が多いために、項目間には項目をグループ分割できるような特徴 (特徴1と呼ぶ) が多く、その特徴で分類対象をグループ分割す

分組をしない
討論会

この図はグループ分け
できない。
グループ内で意見を分
断する

分類の初期段階に見られる傾向で、医療診断の場合、診断の初期に行なう主訴別分類タスクにあたる。ある特徴で、集合要素をほぼ分けることができる。



— 1 —

(2) 順位付タスク向き

分類の末期段階に見られる傾向で、ある特徴で、集合要素を分けることができないタスク

要素AもBも特徴a, b, cを持っているときに、どういうやり方で、AとBを分類すればよいだろうか。「Aとa, b, cの関係の強さ」と「Bとa, b, cの関係の強さ」からどちらがよりa, b, cの関係を持つかによって、AとBを分類するしかない。

例えば、以下のようなRating Gridによる知識表現が与えられているとする。

A	B	C	要素/特徴
5	2	1	a (5...1)
2	5	4	b (5...1)
3	4	1	c (5...1)

図3 Rating Grid による知識表現の例

今、特徴a, b, cについて、観測された値が、4, 4, 4とすると、

Aに対するベクタ (5 2 3)

Bに対するベクタ (2 5 4)

Cに対するベクタ (1 4 1) 以上の各ベクタと、今観測された値からなるベクタ (4 4 4) とのノルムによって、A, B, Cどれが、今観測された状態に近いかを判断する。

医療診断の場合だと、要素(A, B, C)にくるのが病名で、特徴(a, b, c)にくるのが、症状と考えればよい。

3 分類型知識獲得支援システムCTAS

本システムは、ICOTで開発されたPSIマシン上に構築されている。PSIのOSであるSIMPOSは、ウィンドウやオブジェクト指向ロジックプログラミング言語ESP開発環境等の豊富な機能が利用でき、知識獲得を効率良く行うために必要なユーザインターフェースを構築することができた。使用言語は、ESPである。

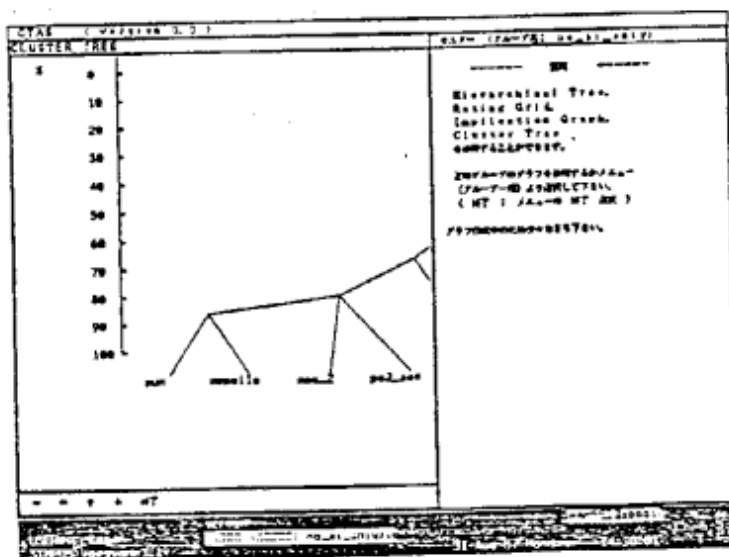
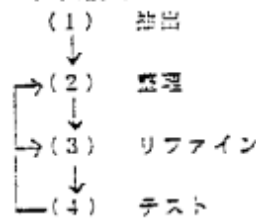


図4. CTASの全体画面

知識獲得プロセスは、次の4段階で行われる。



3-1 抽出プロセス

先の分類の過程を考えて、2通りの知識抽出法を、CTASでは提供している。

(1) ボトムアップ式抽出法

(1)まず、分類したい対象集合の要素項目(ELEMENT)を抽出し、

(2)その後、グループがBINARY TREEになるように、ELEMENTのグループ分割(階層分類)を行なう。

以上で、階層分類は終了し、次に順位付けタスクを行なう。

(3)各のグループ内のELEMENTの分類を行う。その分類に必要な特徴(TRAIT)と関係度の抽出を行なう。まず、特徴抽出を行なうが、その方法は、ETSで行われているPersonal Construct Theoryを用いた特徴抽出法を適用する。

(4)次に、ELEMENTとTRAIT間の関係度の抽出を行う。すると、グループ毎に、ELEMENTとTRAITの関係表(RATING GRID)ができる。

以上で、ELEMENTの階層分類とグループ毎のRATING GRIDが作成されたことになる。

E1	E2	E3	...	element / trait
5	4	1		c1/c1'
1	5	5		c2/c2'

図5. RATING GRIDの例

(B) トップダウン式抽出法

ELEMENTを階層分類した時の階層構造が、予め解っている時にこの方法が有効である。

(1)まず、最初に、階層構造の枠組みをBINARY TREEで抽出する。(A)でいうグループを先に抽出することに当たる。

(2)(1)によって、できた各グループに属するELEMENTを抽出する。

以下の操作(3),(4)については、先の(A)と同じである。

```

----- 実行抽出 -----

同じについての分類をするのですか。
分類名 >> machine

実行抽出からさらにしますか? (y/n) y

*** 実行抽出終了 ***
結果を1つずつ入力して下さい。
( 実行終了 : end)
実行抽出 >> symbolic

実行抽出 >> dec_20

実行抽出 >> vex_0000_series

実行抽出 >>
  
```

3-3 リファインプロセス

先の整理プロセスで表示した表・グラフを基に、層階木の修正、rating gridの値の修正、新たなtraitやelementの追加等を行って、知識をリファインする。

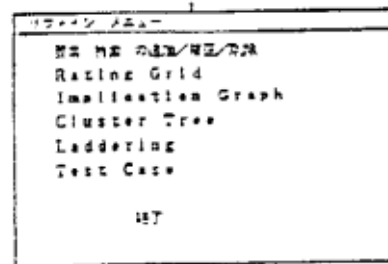


図8. リファインメニュー

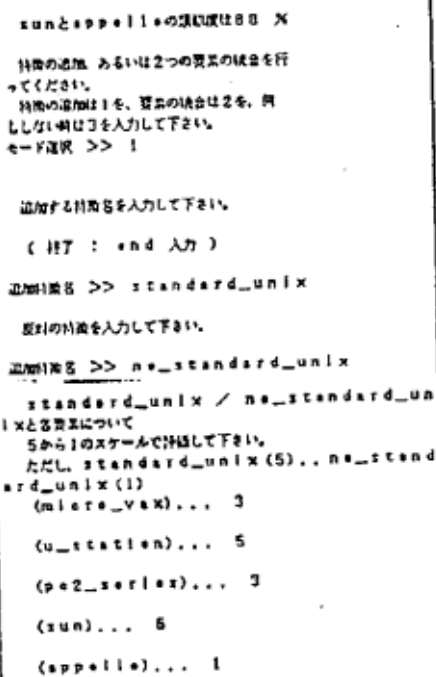
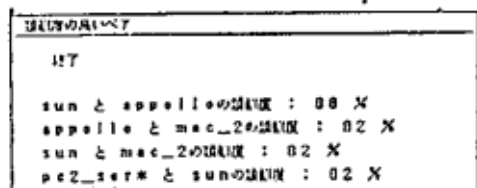
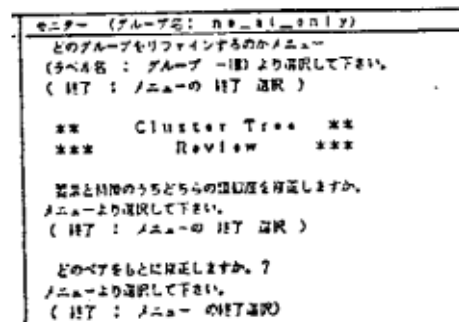


図9. cluster treeによるリファイン

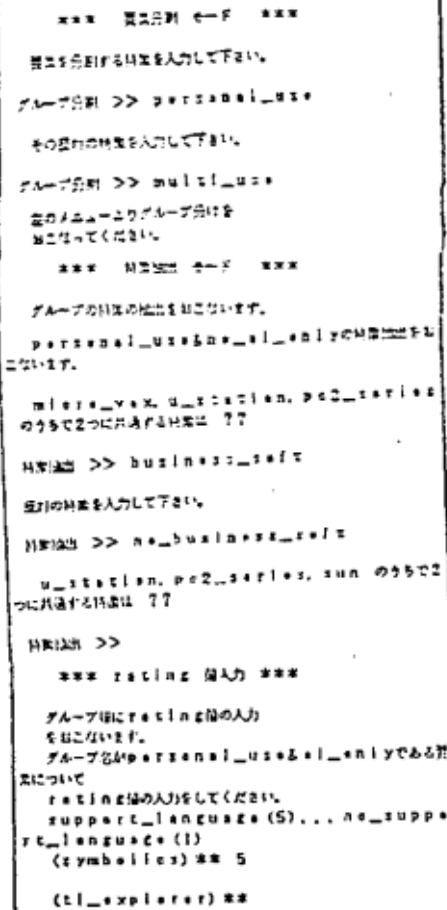


図6. ボトムアップ式抽出法

3-2 整理プロセス

グループ間の関係を表す層階木(HIERARCHICAL TREE)やグループ毎のRATING GRIDを作成する。また、rating gridを基に、implication graph(trait間の関係を示すグラフ)、similarity表示(trait間、及びelement間の類似度を示す)、cluster tree(trait間・element間の類似度を示すグラフ)を作成する。これらの表・グラフによって、ユーザは、知識の内容を視覚的に捉えることができ、知識ベースのリファインを行いやすくしている。

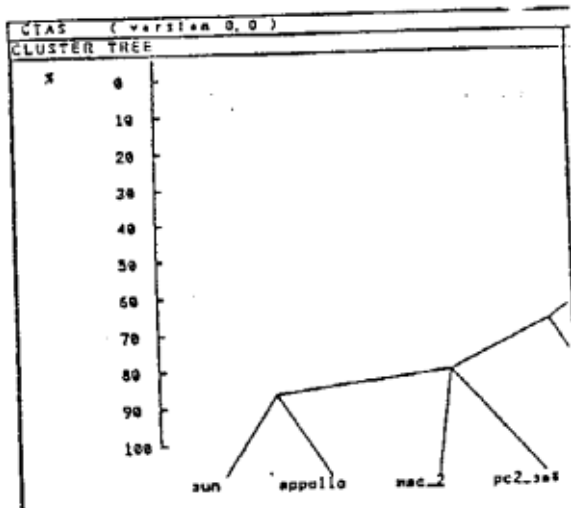


図7. cluster tree (element間)

3-4 テストプロセス

階層・rating gridを基にして、次の3種類のルールを作成する。

(A) HIERARCHICAL RULE (各ELEMENTが、どのグループに属するかを示すルール)

(B) CONCLUSION RULE (各特徴項目と結論項目間の関係を示すルールで、CF値付きルールとなる)

(C) INTERMEDIATE RULE (各特徴項目間の関係を示すルール、IMPLICATION GRAPHに表れる特徴間についてのみ、作成される。)

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。まず、Hierarchical ruleの実行が行われ、テストの対象になっている結論項目の属するグループが決定される。次にそのグループに属する結論項目と特徴に属するConclusion RuleとIntermediate Ruleが実行される。これは、該当グループに属する特徴の確信度に対する問いに対して、-1.0~1.0の値を入力することによって行われる。これらのルールの実行後、結論項目が、確信度の高い順にソートされて表示され、満足のかゆ結果が得られない時は、リファインモードに戻ってCTASの提供する種々の手法でリファインを行うことになる。

```
CTAS >> test.
Element group is test
Please input yes or no
yes - environment is personal-use (yes) or multi-use (no) ?
>> yes.
yes - is it only (yes) or both (no) ?
>> no.
This group name is personal-use&no-only
Group decision is over
Next elements are out in order of CF value
Please input CF value (-1.0~1.0)
yes - is low (1.0) / high (0.0) / unknown (0.0) ?
>> 1.0.
no - is standard (1.0) / no-standard (0.0) / unknown (0.0) ?
>> 0.0.
mode - is mode in Japan (1.0) / no mode in Japan (-1.0) / unknown (0.0) ?
>> 0.0.
yes - is business-type (1.0) / no business-type (-1.0) / unknown (0.0) ?
>> -1.0.
window - is multi-window (1.0) / no multi-window (-1.0) / unknown (0.0) ?
>> 1.0.
The result of test is below
machine - is mac-2 (0.94022481)
machine - is p-2 (0.875)
machine - is apple (0.5)
machine - is sun (0.1230489)
machine - is s-2 (0.9375)
machine - is m-2 (0.4298875)
machine - is w-2 (0.548875)
```

図10 テストの例

本システムでは、以上の抽出・整理・リファイン・テストを繰り返すことにより、知識ベースを構築していく。知識獲得は、システムとユーザとがインタラクティブに処理を進めて行くことにより行われる。

4 まとめ

分類タスクのタスク構造を考慮して構築した知識獲得支援システムCTASについて報告した。

CTASは、分類型問題に関する知識獲得について強力なツールであるが、次の2点について未解決な問題を残している。

1つは、ボトムアップ、トップダウンの階層分類は直感的には有効であると思われるが、理論的根拠に乏しいことである。Personal Construct Theoryがあったが、最初の階層分類における特徴抽出では、

決定的方法がない。もう一点は、階層分類と特徴順位付けとの項目をどこでつけるかという問題である。分類活動は、この2種類のタスクから成り立っているのは、明かであるが、その境界がはっきりしていないように思われる。これら2点の問題解決が、今後の検討課題である。

今後の研究、展開として、次の3点を挙げる事ができる。

(1) CTASの評価

CTASを使って知識ベースを作成していく過程で、どの技術が、知識獲得プロセスで重要なかが解明する。知識獲得のための必要技術として切り出せる部分があるかどうかを検討する。

(2) より良い知識モデルと論理的な枠組み

CTASは、Rating-Gridを知識表現として採用している。このほかMOREのドメインモデルなども知識獲得の知識表現として良い枠組みを提供している。ICOTは並列推論マシン・並列論理型言語を開発中であり、この枠組みからの知識表現として、「意識・無意識の認知モデル」[岡87]がある。この認知モデルも知識獲得の知識表現として利用できる可能性がある。

(3) 心理学的アプローチの研究

CTASは、特徴による順位付けにETSシステムのやり方を一部採用しているが、ETSにおいては、様々な心理学的手法が用いられている。その個々の手法についての有効性を調べたり、また計量心理学や専門家に対するインタビューテクニック等の他の心理学的手法を調べることは、有効であろう。

(4) ユーザインターフェースの研究

ユーザが使いやすいインターフェースは、どんなものか。ユーザにとって簡単に入力でき、また知識の内容も見やすく、リファインもしやすい知識表現形式は、何かを検討する。

[参考文献]

- [Boose 85] Boose, J.: Personal construct theory and the transfer of human expertise., Advances in Artificial Intelligence., North-Holland, 1985.
- [Chandrasekaran 86] Chandrasekaran, B.: Generic Tasks in Knowledge-Based Reasoning: High-Level Building Blocks for Expert System Design., IEEE Expert, Fall 1986.
- [Chikayama 84] Chikayama, T.: Unique Features of ESP. International Conference on Fifth Generation Computer Systems, November 1984.
- [ICOT 88] ICOT Technical Report, "Knowledge Bases for Expert Shell Evaluation", 1988 (to appear).
- [Kahn 85] Kahn, G., S. Nowlan, and J. McDermott.: Strategies for knowledge acquisition., IEEE transactions on Pattern Analysis and Machine Intelligence 7(5), 1985.
- [篠原 87] 篠原 靖志: 未整理な情報からの知識ベース構築. 情報処理学会第34回全国大会.
- [岡 87] 岡 夏樹: 意識処理/無意識処理の認知モデル. 日本ソフトウェア科学会第4回大会.
- [滝 87] 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.
- [椿 87] 椿, 滝: 解析型知識獲得メタシステムの構想. 日本ソフトウェア科学会第4回大会.