PNC TJ9604 98-001



# 大規模実時間知識ベース構築における矛盾検知と学習に関する研究

（動力炉・核燃料開発事業団　研究委託内容報告書）

1997年3月



京　都　大　学

公　開　資　料
PNC TJ9604 98－001
1997年　3月

# 大規模実時間知識ベース構築における知識の矛盾検地と学習に関する研究

椹木 哲夫[1)]

## 要旨

現在までに原子力プラントで発生したトラブルの多くは、運転操作ミスなどの人的要因に関係したものである。そこで安全性、信頼性向上を目的として、従来人間の運転員や保守員の果たしてきた役割を人工知能技術を主体とした自動化システムに代替し、できるかぎり人的要因を排除していこうとする自律型プラントの開発研究が盛んに進められている。自律型プラントにおいては、プラントに関する膨大な量の知識を実施し、この知識を実時間で利用可能にしていくものであるが、そこで問題になるのが、これら膨大な知識の獲得と管理である。すなわち合理的な知識として実装できる知識の量には自ずと限界があり、このような不完全な知識を用いる限り、想定外事象の生起に対しては極めて脆弱なものにならざるを得ない。もちろん最終的には人間オペレータの判断に委ねることが考えられるが近年の人間オペレータへの監視業務一辺倒に偏りがちな自動化システムとの役割分担やモニターすべき情報の飛躍的な拡大は自律機能のみならず、オペレータに対して想定外事象が生じた差異の緊急時に冷静沈着な判断を支援することも併ねた知識化が必須となる。すなわち自律型プラントに求められるのは、このような状況で生起している事象を自身の知識の不備をも考慮した解釈を行うことで的確に現時点での矛盾を検知し、その解消策の提案をオペレータに対して提示したり、自らその解消を試みることで、自身の知識を常に更新していける機構の実現である。さらに知識ベースの実時間運用を考える際には、時間的に不変な知識ばかりを想定することは難しく、対象システムの時間的推移に合せて適用すべき知識もダイナミックに追従しながら更新していくことが求められるが、この意味においても、いつどの時点で現状の知識の認識し、これを改定していけばよいかについてのメンタルモデル判断機能が不可欠となる。

本報告ではこのような状況に際して、想定外の事象に対して現在の知識ベースの知識との矛盾を速やかに検知し、自律プラント自らが、あるいはオペレータがその矛盾を解消を的確に行うことを可能にするための方策について、人工知能における機械学習（machine learning）分野から、帰納学習、演繹学習を中心に概説する。そしてそれらの学習手法における矛盾検知とその解消による知識獲得がそのように実現されるかについて、1)既存知識の矛盾検知、2)合理的な矛盾解消、3)既存知識の修正による新規知識の学習獲得の観点から取りまとめる。

---

本報告書は京都大学が動力炉・核燃料開発事業団の委託により実施した研究内容成果である。
契約番号：
事業団担当部課室：本社　核燃料サイクル技術開発部
1)京都大学大学院工学研究精密工学専攻

PNC TJ9604 98－001
March, 1 9 9 7

# Studies on Learning by Detectiong Impasse and by Resuling It for Building Large Scale Knowledge Base for Autonomous Plant

Tetsuo Sawaragi[1)]

## Abstract

Recently, due to the tremendous improvement of information infrastructures such as networking facilities, the idea of a large scale knowledge base with realtime operations for technological plants has emerged. The major bottleneck for building a large scale knowledge base for an autonomous plant lies in its design phase. The acquisition of knowledge from human experts in an exhaustive way is extremely difficult, and even if it were possible, the maintenance of such a large knowledge base for realtime operation is not an easy task. The autonomous system having just incomplete knowledge would face with so many problems that contradicts with the system's current beliefs and/or are novel or unknown to the system. Experienced humans can manage to do with such novelty due to their generalizing ability and analogical inference based on the repertoire of precedents, even if they with new problems. Moreover, through experiencing such breakdowns and impasse, they can acquire some novel knowledge by their proactive attempts to interpret a provided problem as well as by updating their beliefs and contents and organization of their prior knowledge. We call such a style of learning as impasse-driven learning, meaning that learning dose occur being motivated by facing with contradiction and impasse. The related studies concerning with such a style of leaning have been studied within a field of machine learning of artificial intelligence so far as well as within a cognitive science field. In this paper, we at first summarize an outline of machine learning methodologies, and then, we detail about the impasse-driven learning. We discuss that from two different perspectives of learning, one is from deductive and analogical learning and the other one is from inductive conceptual learning(i.e., concept formation or generalization-based memory). The former mainly discuss about how the learning system updates its prior beiiefs and knowledge so that it can explain away the current contradiction using some meta-cognition heuristics. The latter attempts to assimilate a contradicting problem into its prior memory structure by dynamically reorganizing a collection of the precedents. We present those methodologies, and finally we introduce a case study of concept formation for plant anomalies and its usage for an intelligent decision support system.

---

1)dept. of Precision Engineering, Graduate School of Engineering, Kyoto University

# 目 次

# 表目次



# 図目次

# 第１章　緒言

近年，原子力発電プラントに代表される大規模・複雑システムの安全性，信頼性向上を目的として，従来人間の運転員や保守員の果たしてきた役割を人工知能技術を主体とした自動化システムに代替し，できるかぎり人的要因（ヒューマンファクタ）を排除していこうとする自律型プラントの開発研究が盛んに進められてきている．そこではプラントの設計・運転・保守といったプラントの全ライフサイクルをカバーすることが要請されることになるため，当然のことながらこのような自動化システムがもつべき知識は大規模にわたる．その一方で実際に実装できる知識の範囲には限界があり，限られた既存知識からいかにして想定外事象を含む新規な問題に対して解決を行うことができるかが課題となる．このような知識獲得ボトルネックは第一世代エキスパートシステムの提唱以来，すべからく認識されてきた問題であるが，これに対して近年の知識情報処理からのアプローチでは，従来の正攻法的な問題解決システム，すなわち一般普遍則としての知識を網羅的に準備しこれに基づく演繹操作のみで解を得るような第一原理に基づく問題解決の手法に代わり，知識のソースとなる具体的な事例群を大量に蓄積し，これを柔軟に活用することでその集積の中から一般則を導出したり，類推により解を得るスタイルの問題解決が主流となりつつある．

大規模実時間知識ベースを構築する上で問題になるのが，これら膨大な知識の獲得と管理である．すなわち合理的な知識として実装できる知識の量には自ずと限界があり，このような不完全な知識を用いる限り，想定外事象の生起に対しては極めて脆弱なものにならざるを得ない．もちろん最終的には人間オペレータの判断に委ねることが考えられるが，近年の人間オペレータへの監視業務一辺倒に偏りがちな自動化システムとの役割分担やモニターすべき情報の飛躍的な拡大は，自律機能のみならず，オペレータに対して想定外事象が生じた際の緊急時に冷静沈着な判断を支援することも併せた知能化が必須となる．すなわち自律型プラントに求められるのは，このような状況で生起している事象を自身の知識に照らし合わせてその状況を認識する際に，自らのカバーする知識の範囲を自省でき，自身の知識の不備をも考慮した解釈を行うことで的確に現時点での矛盾を検知し，その解消策の提案をオペレータに対して提示したり，自らその解消を試

みることで，自身の知識を常に更新していける機構の実現である．さらに知識ベースの実時間運用を考える際には，時間的に不変な知識ばかりを想定することは難しく，対象システムの時間的推移に合わせて適用すべき知識もダイナミックに追従しながら更新しくことが求められるが，この意味においても，いつどの時点で現状の知識の不備を認識し，これを改訂していけばよいかについてのメタレベル判断機能が不可欠となる．

本報告書では，まず上述のような技術課題を達成するための方法論として，人工知能の機械学習分野で提案されている各種の手法について概説する．つぎに想定外の事象に対して既存の知識ベースの知識との矛盾を速やかに検知し，自律プラントら自らが（あるいはオペレータが）その矛盾解消を的確に行うことを可能にするための方策についてまとめる．すなわち大規模知識ベースとしてあらかじめ完全な知識ベースを事前に準備することは極めて困難な技術的状況にあるが，このような矛盾検知と矛盾解消を通して新たな知識を学習獲得することができ，これを既存知識に付加していくことのできる成長型知識ベースの実現により知識獲得ボトルネックの解消が期待される．このことは今日，大規模実時間知識ベースシステムのみならず，よりオペレータと対等なレベルでのパートナーたるべき人工的なエージェントの設計に際しても必須の能力として認識されている．本報告書ではこのような知識の矛盾検知と学習において解決すべき技術課題の抽出，整理を行い，これに対して現在までに提供されている技術として，人工知能における機械学習（machine learning）分野から，帰納学習，演繹学習，類推学習を中心に概説する．そしてこれらの学習手法における矛盾検知とその解消による知識獲得がどのように実現されるかについて，１）既存知識の矛盾検知，２）合理的な矛盾解消，３）既存知識の修正による新規知識の学習獲得，の観点から取りまとめる．

以下本報告書では，続く２章ならびに３章において，演繹学習や類推学習の分野を中心に展開されている矛盾検知とその解消による知識学習の考え方についてまとめる．さらに４章ならびに５章においては，帰納学習の一つである概念学習の手法を中心に教師なし帰納学習において，過去に集積された事例集合と矛盾をきたすような新規事例が新たに入力された場合に，これを既存知識に同化するべく事例記憶をダイナミックに再組織化していくための手法を中心にまとめる．また続く５章ではこのような考え方に従って，プラントのトレンドデータからの概念学習のケーススタディを紹介し，意思決定支援システムとして構成した例について報告する．そして最終章では本報告をまとめる．

# 第２章　機械学習の概要

## 2.1　機械学習の分類

一般に学習（learning）とは，本来人々や計算機がその知識を増大させ，彼らの技能を習熟していく方法全般を表すものである．学習研究には２つの理由がある．一つは学習過程そのものを理解することであり，いま一つは，学習する能力をもつ計算機をつくりあげるという理由である．

機械学習に関する４つの視点は以下のように与えられる．

1. システムによるタスク実行の仕方を改良する過程
2. 明示的な知識の獲得　エキスパート・システムでは専門的知識（expertise）というものを，獲得され，組織化され，拡張される必要のある大量のルールの集合体として表現する．この知識の獲得ならびに修正，検証といった作業を計算機に行わせる．
3. 技能の獲得　人間の技能は長期にわたり実践を通して改良し続ける．あるタスクのやり方に関する後述の教えを容易に理解できるが，その後述の知識を精神的あるいは身体的な作用に変えていく技能獲得（skill learning）の過程．
4. 理論（仮説）の形成と帰納的推論　特定のデータ集合を説明する１個以上の妥当な仮説を発見するという活動．すなわち，特定の例題から一般的法則を推論する過程．

学習システムの議論を系統的に行うために図2.1に示す学習モデルを考える．このモデルにおいて円形部分は宣言的な情報の実体（例えば，述語論理で表現される事実であるとか専門家によって与えられる文）を表示している．また箱型部分は手続きを表す．環境は学習部に関連する情報を規定し，学習部はこの情報から知識ベースを明示的に改良するのに利用し，そして実行部はこの知識ベースを，タスク実行のために使用する．最後に，タスク実行の試行の間に得られた情報が，学習部へのフィードバックとして利用される．学習システムの設計に影響を与える最も重要な因子は，環境によってシステムに提供される情報の種類，とくにこの情報の抽象化の水準（level）である．

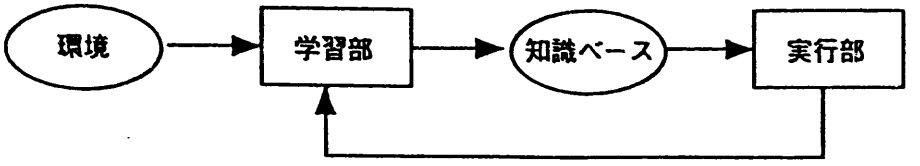



Fig. 2.1 学習の単純なモデル

情報の水準は，実行部の必要性に応じた情報の汎用性（適用可能分野）の程度を表している．高水準の情報とは，広いクラスの問題に適用可能な抽象的情報のことで，低水準の情報とは，単純な問題に対してのみ適用可能な詳細な情報のことである．学習部のタスクとは，その情報が環境によって提供される水準とその実行部がその機能を実行するのにその情報を用いることができる水準との間のギャップを橋渡しするタスクである．もし学習システムがその実行タスクについて抽象的すぎる（高水準の）情報を与えられているとすると，そのシステムは失われた詳細情報を推定しなければならない．もしシステムが特定の状況での実行の仕方に関する固有の（低水準の）情報を与えられているとすれば，学習部はこの情報を，重要でない詳細を無視して，状況の広範なクラスに実行部を導くのに使用することができる規則へと一般化しなければならない．しかしここでの詳細情報の推定の仕方や無視の仕方は前もって学習部は知らないので，その都度仮説を形成し，この仮説に基づいて実行してみてその評価をフィードバックしながら学習する内容を修正していくことが必要になる．従って環境によって提供される情報の水準により，以下の４つの学習状況が識別される．

1. 暗記学習　環境は実行タスクの水準で情報を正確に提供する．仮説は不要．
2. 助言による学習（演繹学習）　環境によって提供される情報があまりに抽象的もしくは一般的であるので，学習部は失われた詳細情報を仮説化しなければならない．
3. 例題からの学習（帰納学習）　環境によって提供される情報があまりに固有かつ詳細であるので，学習部はより一般的な規則を仮説形成しなければならない．
4. 類推による学習（類推学習）　環境によって提供される情報は類似の実行タスクに対しても適切である．学習システムは現在の実行タスクに対して，環境から提供された情報との間の類推を発見し，かつ類似の規則を仮説形成しなければならない．

以下本章ではこれらの各々の学習手法の概要について記す.

## 2.2 暗記学習

最も単純な学習状況は，実行部に直接利用できる形式の知識を，環境が提供する場合である．学習システムにおいては，環境によって提供される情報を理解あるいは解釈するために，何らかの処理をする必要はなく，後で使用するために入力情報を記憶しておくことだけである．暗記学習は実行部が解くべき問題をとりあげ，その問題と解答を記憶化することによってなされる．実行部は入力パターン $(X_1,\cdots,X_n)$ を受け取り，出力値 $(Y_1,\cdots,Y_p)$ を計算する関数 $f$ と考えることができ，$f$ に対する記憶部は，単に連想対 $[(X_1,\cdots,X_n),(Y_1,\cdots,Y_p)]$ をメモリ中に格納する．

暗記学習の問題点は以下の３つにまとめられる．

1. 記憶の組織化　暗記学習は，もし望みの項目を再計算するよりも，それを検索したほうが時間がかからない場合においてのみ有効である．高速に検索できるようにするためには記憶を適当に組織化しておかねばならない．
2. 環境の安定性　暗記学習に関する重要な仮定は，ある時刻に蓄積された情報はあとになってもなお妥当であるという仮説である．しかしもし情報がしばしば変化すればこの仮定はこわされる．
3. 計算対記憶のトレードオフ　記憶された情報を格納し，検索するコストが，それを再計算するコスト以上にかからないようにする必要がある．このためには，一つのアプローチとして，ある情報が最初に利用される時点でそれが後での再利用のために格納されるべきか否かを決定する選択的格納と，別のアプローチとして，進んで格納し後にそれを忘却するかどうかを決定する選択的忘却のやり方がある．

## 2.3 帰納学習

第一世代エキスパートシステムにおいては，知識ベース内の知識の獲得は，ナレッジ・エンジニアと呼ばれるシステム設計者が専門家に対し，質問シートやインタビュー等の手段により，一般則の形で抽出・獲得するのが常であった．このような場においては，

専門家の有する具体的な個別事例というものは，その具体性の故に適用範囲が狭すぎるとして興味の対象外におかれ，あくまで普遍的な一般則を，しかも言語的な形で記述することが要求されてきたが，このプロセスが困難さを極めた．これは，そもそも個別事例というものが，一度限りの変形にすぎず，その中には当該事例に固有な特徴と，一般則に普遍化できる本質的な特徴が混在しており，この中から後者のみを選択的に切り出して言語化することの困難さに拠るものであり，このような事例からの一般則の抽出作業自身，その問題領域に精通した真の専門家に限られることになる．

機械による学習（machine laerning）は，そもそもこのような高度知的作業をコンピュータに代替させるべく，既知の事例記述から一般則を獲得するための方法論に関する研究分野である．中でも例題からの学習（learning from examples）は，対象世界の個々の事例，例えば，システムの呈する断片的な現象から，その現象生起の背後に隠れた一般則を帰納的（inductive）に導くものである．個々の事例には，求めようとする目標概念の一般則（概念仮説）に直接は関係のない冗長性が含まれており，また一般則に該当する事例（正事例）のみならず，該当しない誤った事例（負事例）も含まれている．例題からの学習は，このような各事例に対する正・負の判別を外部教師が指示して与えたうえで，目標概念にあう例を説明できあわない概念を排除するような一般化記述を学習することが目的である．

このような帰納的手段でデータを処理する試みは，既に各種実験科学・社会科学の分野で急速な進展をみたクラスター分析などにも見られ，データの持つ潜在的な（概念の）区別をコンピュータで行なわせようとするものであった．ただし，これらの古典的手法においては，事例は特徴ベクトル・多次元空間内座標値ベクトル等の数値データ群であり，その目的も極めて多量のデータの集積の中から，その集団的特性を発見するためのもので，一般則を抽出できるか否かは，ひとえにこれらの手法を用いる解析者の能力に委ねられていた．これに対し，近年の知識情報処理分野における学習は，より限られた数の事例（極端には単一の事例）が個々に有する意味，すなわち，事物や事象の意味属性に注目した定性的記述でもって事例の持つ固有性に着目し，そこから上述の一般化概念を記述的な表現で陽に得ようとする点が異なる．

このような一般化の手法は大きく帰納学習と演繹学習に分けらる．後者については次節以下で述べるとして，本節では前者の帰納学習の基礎を与える Mitchell により提案

されたバージョン空間法について述べる．これは個々の事例記述を最も「特殊」な仮説，すべての事例を包含しうるような空の記述を最も「一般的」な仮説として，その間に一般化の半順序関係を介して構成される束構造の仮説空間（バージョン空間）を構成する[26]．

あらゆる表現言語において，文はその一般性（generality）の度合に応じて半順序関係に並べることができる．図2.2は *RED* と *BLACK* の述語を含む述語論理におけるいくつかの文の一般―特殊の順序関係を示している．規則空間の中で最も一般的な1点は

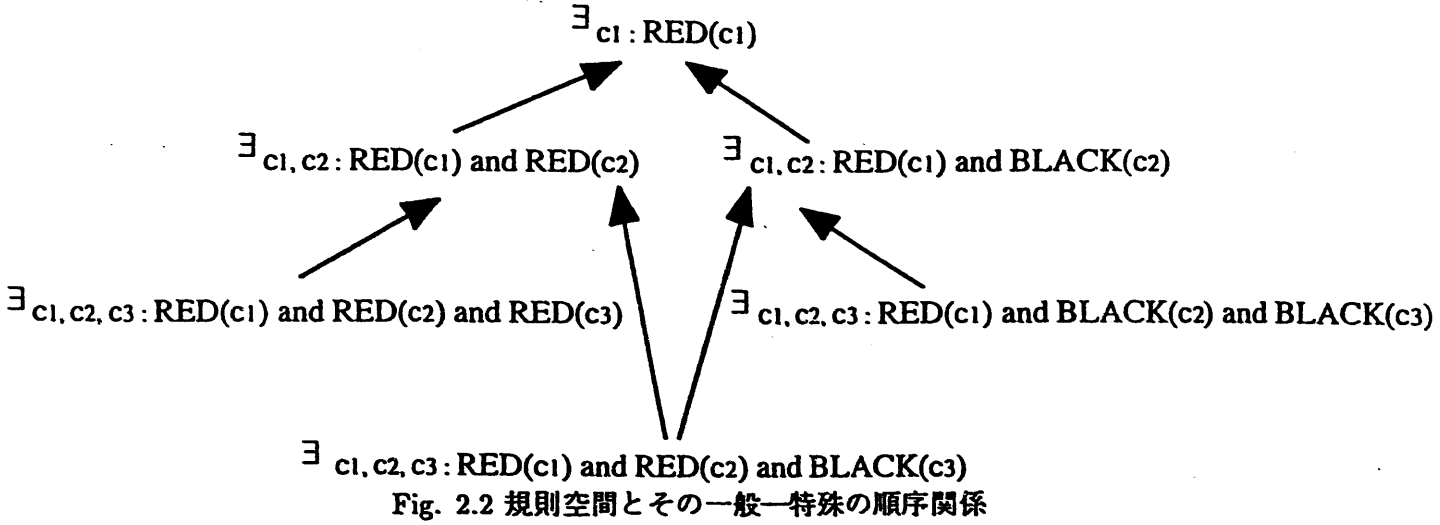



Fig. 2.2 規則空間とその一般―特殊の順序関係

普通，空の記述（すべての条件が落とされた記述）であり，訓練例に何の制約も与えずに，したがって「全て」の例を表わす．規則空間の最も特殊な点は，規則空間と同じ表現言語で表わされた訓練例そのものに対応する．半順序をなした集合の中の点の集合はその最も一般的な要素と最も特殊な要素により表わすことができる．つまり図2.4に示されるように，もっともらしい仮説の集合 $H$ は2つの部分集合―$H$ の中の最も一般的な要素の集合（$G$ 集合と呼ばれる）と $H$ の最も特殊な要素の集合（$S$集合と呼ばれる）―によって表わされる．このようにすべての可能な仮説の集合 $H$ をバージョン空間と呼ぶ．すなわちバージョン空間 $H$ はこれまで見たどの訓練例とも矛盾しないようなあらゆる概念記述の集合である．訓練例が与えられる前の段階の最初のバージョン空間は可能な概念の空間全体である．訓練例がプログラムに与えられるにつれて，候補概念がバージョン空間から消去されていく．そしてただ一つの候補概念が残ったとき，望んだ概念が見つかったことになる．正の実例は，プログラムに一般化することを強制する．すなわち非常に特殊な概念記述が $H$集合から除去される．逆に負の実例はプログラムに特殊

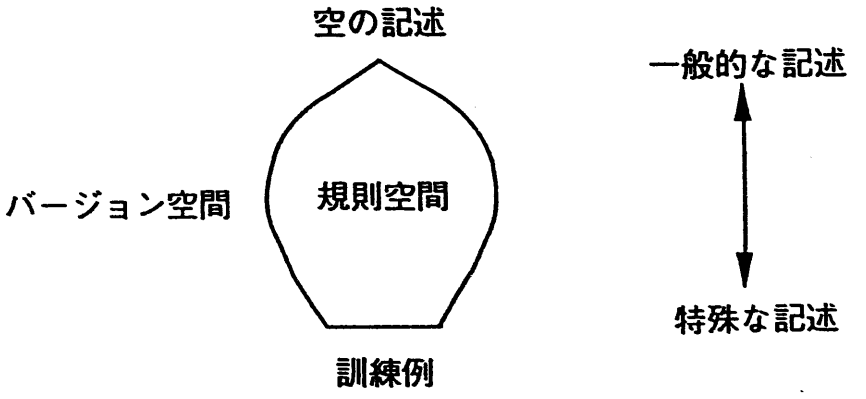



Fig. 2.3 規則空間の図式的表現

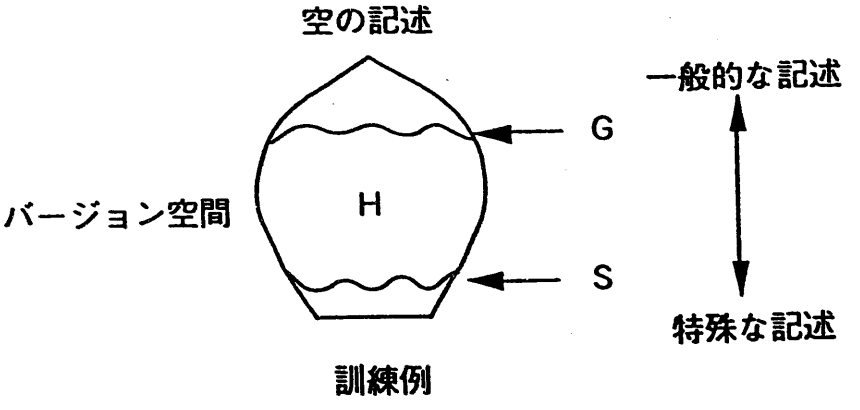



Fig. 2.4 規則空間中の部分空間を表わすための境界集合

化することを強制して，非常に一般的な概念記述が $H$ 集合から除去される．バージョン空間はこのようにして次第に小さくなり，ついには望ましい概念記述のみが残る．

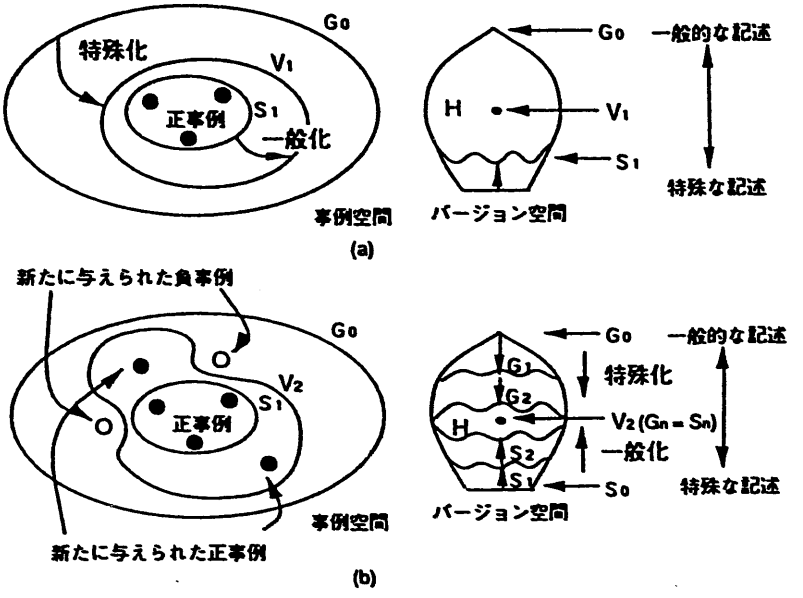



Fig. 2.5 バージョン空間の絞り込み

例えば，図 2.5(a) に示すように，●で表した３つの正事例をすべてカバーするような一般化記述は，この３つの事例以外の考えうるすべての事例をもカバーするような極めて抽象度の高い記述（$G_0$）から，３つの正事例記述のみをカバーような抽象度の低い一般化記述（$S_1$）の間に無限に存在する（$S_0$ は個々の事例記述に対応する）．しかし，これらの正事例に加え，新たに同図 (b) に示すような正負各々の事例が訓練例として与えられた場合には，$S_0$ の記述では正事例をカバーし切れず，また $G_0$ の記述では負事例をも含んでしまうことから，すべての正事例をカバーし負事例を一つもカバーしないような唯一の概念記述を発見しなければならない．そのためには，負の事例を含まないように $G_0$ から徐々に特殊化を施す一方で，新たな正事例もカバーするように現在までに生成された概念の一般化を行なうという操作を繰り返して，バージョン空間を次第に絞り込んでいくことになる．バージョン空間上で，蓋然的（plausible）な仮説，すなわち「今までの例により除外されていない仮説」の集合 $H$ は，$H$ の中で最も一般的な要素の集合 $G$ と，最も特殊な要素の集合 $S$ とによって挟まれた領域として表現される．従って，一般化概念の獲得は，このような空間を外部から与えられる訓練例に応じてつぎつぎに

更新していくことによって達成される（同図中，$S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow \cdots$ ならびに $G_0 \rightarrow G_1 \rightarrow G_2 \rightarrow \cdots$）．このような更新は原則として，正の例が与えられた場合には，その例をカバーしない仮説を $G$ から取り除き，次にその例と今までの $S$ との間で共通の一般化を施し，その中で最も特殊な仮説をすべて含むように $S$ を更新する．負の例が与えられた場合には，その例をカバーする仮説をすべて $S$ から取り除き，その例を含まないように $G$ を特殊化した上で，その中で最も一般的な仮説をすべて含むように $G$ を更新する．このような一般化と特殊化の操作を繰り返しながらバージョン空間を絞り込んでいき，最終的に $S_n = G_n$ となるような仮説を決定する．

以上のアルゴリズムをまとめると以下のようになる．

Step 1: $H$ を初期化して全空間とする．すなわち，$G$ 集合は空の記述のみからなり，$S$ 集合は空間中の最も特殊な概念の全てを含む．

Step 2: 新しい訓練例を受け取る．正の実例の場合には，まずその例をカバーしない概念をすべて $G$ から取り除く．次にその例といままでの $S$ の要素に，共通の一般化をほどこした概念のうち最も特殊な概念をすべて含むように $S$ を変更する．これは Update-S ルーチンと呼ばれる．負の例が与えられたときは，まずこの反例をカバーするすべての概念を $S$ から取り除く．つぎにその新しい実例といままでの $G$ の要素に共通の特殊化を施した概念のうちの最も一般的な概念を全て含むように $G$ を変更する．これは Update-G ルーチンと呼ばれる．

Step 3: $G = S$ かつこれが単一要素の集合となるまで Step 2 を繰り返す．繰り返しが終わったとき，$H$ はただ一つの概念しか含まないようにつぶれている．

Step 4: $H$（すなわち $G$ または $S$ を出力する）．

## 2.4 演繹学習

以上の帰納学習では，一般化すべき部分の抽出が複数の事例記述の共通部分を摘出することを基本としているのに対し，演繹学習の一手法として知られる説明に基づく一般化（explanation-based generalization, EBG）は，学習対象の領域固有知識を利用することによって単一の訓練例から正当な一般化記述を演繹的に得るための学習手法である [27]．すなわち，図 2.6 に示すように，与えられた事例が何ゆえに学習の目標概念を満

足するものかについての説明を生成することによって，事例記述の中から他の事例にも普遍的に当て嵌まるような本質に拘わる部分のみを選択的に切りだし，この説明木中の定数を変数に一般化することによって学習目標の十分条件の形で事例の一般化記述を得るための手法である．

例えばいま，「椅子」の一つの事例として「応接いす」を見せられたもとでの「椅子」の一般化概念の獲得を考える．このような場合，我々は与えられた事例の構造をそのまま詳細に写し取るような記述，すなわち「色は‥，大きさは‥，脚が4本あって，材質は‥」というような記述をつくりだすことは比較的容易い．しかしこのような記述をもって一般化概念とすることはない．「椅子」のプロトタイプ（原型）とはどのようなものであるか，すなわち，「安定していて，快適に体全体を支えることができる」というような機能的な定義を念頭におき，そのもとで当該事例の記述の中から，このような抽象的な機能定義が具象化されている部分を拾い上げ，例えば「長さの等しい4本の脚とほどよい位置で感触のいい水平状の座部，その端から垂直状に突き出した背もたれ部をもち柔らかな詰め物をした構造物」のように「椅子」の一般化概念を抽出してくるであろう．このとき椅子の「色」や「材質」といった当該事例に固有の属性は自動的に捨象される筈である．勿論このためには，どのような機能がどのような構造要素で実現されて得るかについての図2.6(a) に示すような背景知識（＝専門知識）が必要になる．

このようにEBGの手法は，上例の機能定義のように学習目標となる「粗い」概念と，このような概念が具体的に実現されている個別事例（訓練事例という）を与えることによって，予めシステム内に用意された一般的な背景知識（領域理論という）による演繹処理を行ない，当該事例の中から，他の「椅子」の事例にも普遍的に通用するような本質に関わる部分のみを選択的に抽出して一般化し（図2.6(b)），これらをコンパイル知識として獲得するための理論である（図2.6(c)）．従来の帰納的な学習が，多数の事例の中に見いだされる共通性に着目した一般化，従って多分に発見的で，大きな計算コストを要するものであったのに対し，EBGは背景知識に基づいて単一事例から有用な一般的知識の獲得を目指すものである．

このような説明に基づく学習のもとになった考え方が助言に基づく学習である．すなわち学習システムがその実行タスクについて抽象的すぎる（高水準の）情報を与えられたときに，そのシステムが失われた詳細情報を推定するのが助言に基づく学習である．

上例の機能定義のように学習目標となる「粗い」概念は，まさにこれに該当する事例が幾つも存在するという意味で「高水準」の情報である．一方，この情報はそのままでは「椅子」を認識するための「実行部」で利用されるべき情報としては抽象的過ぎて用いることができない．「実行部」にとって望ましいのは，より「低水準」の構造要素の関係として「椅子」の概念を記述した知識である．そこでこのギャップを埋めるのが，生成された説明構造である．そしてこの説明生成を通して，低水準の記述として知識を生成する．一般にこのような演繹による学習目標概念の具体化は唯一に決まるものではない．EBGでは，このプロセスで「訓練事例」といういま一つの情報を学習システムに与えることで，説明生成を当該事例に限定したものに抑え込む役割を果たしている．

## 2.5　事例に基づく問題解決（事例ベース推論）

従来，問題解決を行なう知識システムとしては，ルール型エキスパートシステムに代表されるように，問題解決に必要な知識を一般的な形で知識ベース内に用意し，これを解決の対象となる事例に対して演繹的に適用するというものであった．前節で述べたEBGも，本質的にはこのような知識ベースの逐次的拡張，問題解決効率の漸次的改善を意図したものであった．しかしながら，現実に問題解決の要請されている問題対象においては，このような一般則の獲得が必ずしも容易には行なえない場合，また例えできるとしても，膨大なコストを要する場合が少なくない．特に問題対象として，過去に人間エキスパートにより実際に問題解決のなされた事例の蓄積が豊富にあったり手に入れられることはあっても，その一般的な定式化が困難であったり，あるいは大規模な知識ベースの構築が可能になったとしても，実際それを使って推論を進める際には，推論の連鎖がとてつもなく長いものとなったり，一義的な解に絞り込める系統だった手法が存在しないということは十分に考えられることである．このような問題に対し，「類似性は因果関係を保存する．すなわち似た状況は似た結末を生じやすい．」という原則[40]に則って，単一の解の与えられた既知事例から一般化のプロセスを介さずにダイレクトに関係を写像することにより，未知状況中に特定の関係の存在を推論して解を見いだそうとするものが，類推（analogy）による問題解決（学習）である．そしてこれを現実的な問題に適用するための枠組みとして提唱されているのが事例ベース推論（case-based reasoning）

である [8],[16].

図 2.7に示すように，事例ベース推論は既知事例をシステム内に事例ベースとして格納し，入力された問題事例の解決の局面に応じて，原則として最も類似する既知事例を検索，これを当該事例に適合するように修正することによって，既知事例の結果を有効に活用しながら問題解決を行なうための枠組みである [40],[41]. ここで格納される事例情報には，単に解決の就けられた結果としての解のみならず，問題の設定や記述，解決に至った経緯を表す因果説明構造，事例の持つ特異性，等の情報が特徴リストとして記憶されている．図 2.8に示すように，これらの諸特徴と当該事例（target）にみられる特徴の間の部分的照合操作により，まず最類似の既知事例（base）が検索される．そしてこの検索された既知事例中の特定の関係が，当該事例にマッピングされ，これから未知項目を推論しながら問題解決を行なうものである．このマッピングに際しては，対象領域において常識的に認められているような意味構造や概念構造，あるいは因果モデルなどの領域固有な知識が用いられる．

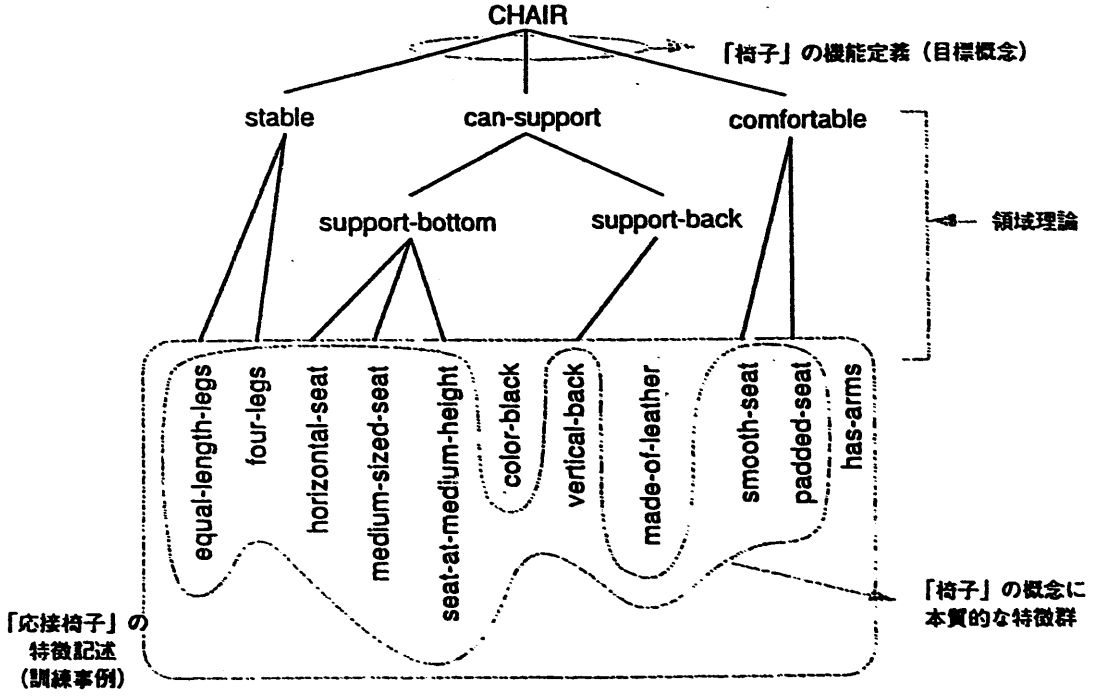
一般に上述のような既知情報からの類推を進める場合には，与えられた事例に対して既知の何れの事例からのマッピングを行うか，すなわち，類推の対象となる既知事例を，事例間の類似性に基づいて決めることが重要なプロセスとなる．このような事例間の類似性判断としては，従来統計処理の分野において，[項目－値] からなる特徴リストの重複関係から数値的な類似性尺度を定義し，メトリック空間上での近接性＝類似性と見なして行う手法が一般的であった．しかし，単純な数値尺度による類似性判断には，項目毎に priority の違いがある場合や，事例の本質に関連性をもたないような項目が混入する場合に問題が残された．これに対し CBR は，観察データの背後に潜在する因果モデルや，各々の特徴項目の差異に固有な類似性判断規則を用意することにより，問題の目的において個々の項目のもつ意味概念を慎重に吟味しながら補間を進めるための枠組みである．類推の研究で先駆的な役割を果たした Gentner は，当該事例に最も類似する既知事例の特定や写像により転移される情報の決定には，事例記述の背後に隠された高階の関係の存在が大きく寄与することを構造写像原理 [13],[14] として提唱しているが，CBR はまさにこのような原理を現実の問題に適用したものと言える．図 3.1に構造写像原理の図を示す．

上述のことからも明らかなように，事例ベース推論は，いわゆる解析的な手法でもっ

て問題の最適解を導出するためのものではなく，あくまで既知の事例の範囲内での満足解を得るための手法として理解するのが妥当である．また分断された一般則を問題の部分部分に繰り返し適用することによって結論を導く従来のエキスパートシステムとは根本的に異なり，実例のもつ全体像を残したままで，その全体的な類似性に着目する点は，領域エキスパートに固有な「直感的な理解」の側面を維持することができ，具体的な実例を介してのエキスパート―システム設計者の間の意思疎通が，より現実性と具体性を帯びたものになって円滑なコミュニケーションを可能にするという期待も寄せられている．

chair :- can-support, comfortable, stable.
can-support :- supports-bottom, supports-back.
supports-bottom :- horizontal-seat, seat-at-medium-height, medium-sized-seat.
supports-back :- vertical-back.
comfortable :- smooth-seat, padded-seat.
stable :- four-legs, eqaul-length-legs.

(a)「椅子」認識のための背景知識（領域理論）




(b)「応接椅子」の説明木

chair :-
equal-length-legs, four-legs, horizontal-seat, medium-sized-seat,
seat-at-medium-height, vertical-back, smooth-seat, padded-seat.

(c) 獲得された「椅子」の概念

Fig. 2.6 説明に基づく一般化

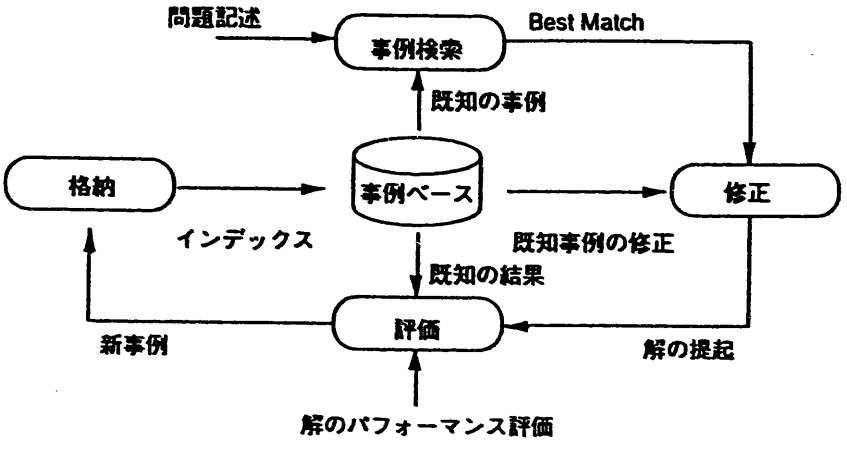



Fig. 2.7 事例ベース推論

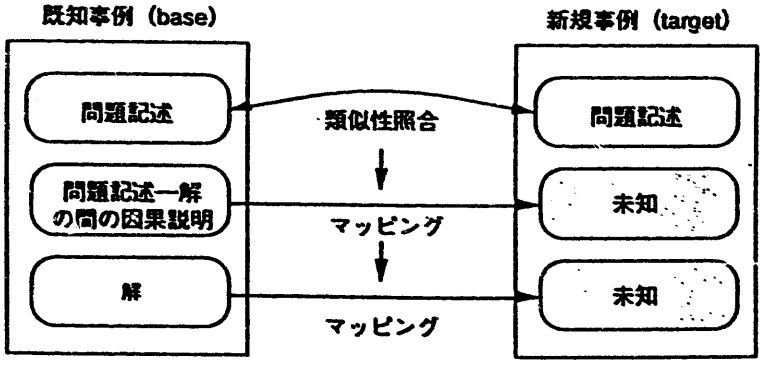



Fig. 2.8 事例ベース推論のマッピング

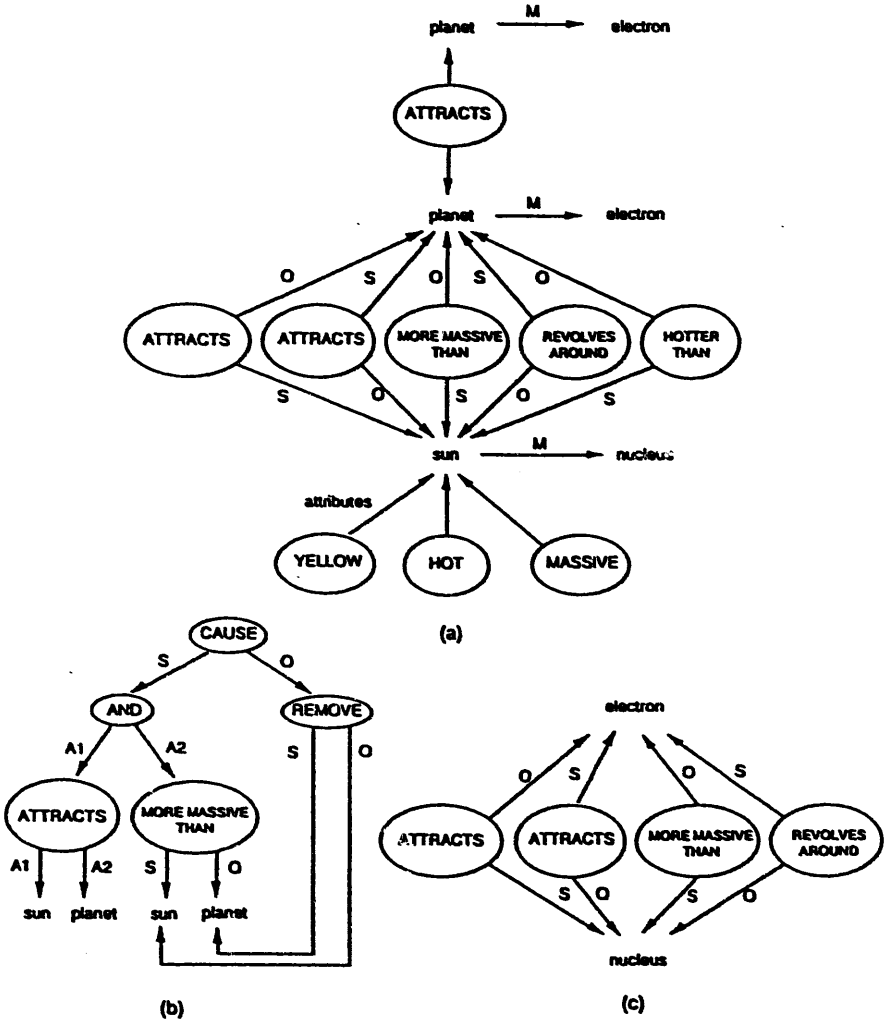



Fig. 2.9 構造写像原理

# 第 3 章　矛盾解消による知識獲得

## 3.1　矛盾の検知と学習

前章で述べた演繹学習では，事例から学習すべき目標概念を何等かの形で定義しなければ学習は進行しない．それでは一体，学習システムは何を新たな概念として獲得することになるのか，すなわち，「既知のことから既知の知識をつなぎ合わせてどんな新たな知識が生成されるのか」，これがEBGの提唱以来疑問視されてきた説明パラドクス (explanation-based paradox) である．また，EBGに対するいまひとつの批判として，説明構造生成に必要な領域知識が，現実問題として容易に備わるものかという疑問がある．

Schankは「説明」に対する分類として，人間の行う説明のパターンを，

1. Canned Explanation
2. Explaing Away Explanation
3. Additive Explanation

の三種類に分類している [34]．第一のCanned Explanationは，既知のことを対話の相手に伝えるだけの行為であるが，このようなタイプの説明は，説明が与えられる時点で既に説明者の内部に用意されていたものであり，説明を行うことによって説明者自身にはなんら新たな創造は望まれない．第二のタイプの説明は，説明対象が一事例と見なせるような，既知の仮説世界を見いだすことを意味するが，説明を生成することによって理解システム自身が知識構造の変容を加えられるというものではない．これに対し，第三のタイプの説明は，既有知識の中に，入力に該当するものが見いだせない場合に，ある一連の手段を講じてその理由付けを行い，より深い理解に到達しようとするもので，説明生成を通して以前には未知であったことが新たに付加される．そしてこのタイプの説明こそが「学習」と密接な関連をもつものであるといえる．

Schankはこのような説明の分類を，説明が要請されるニーズに対応させて，外部供与的なニーズ (2.) から，内部発生的ニーズ (3.) に亙る次元上で対応づけており，前者に比べ後者の説明ほど，説明生成のための手続きが，より複雑なものになり，生成の困

難さも増す一方，逆に学習システムにとってはより有意味な説明であるとしている．

以上の説明タイプの分類からも明らかなように，学習は，いくつかの事例間の類似性・差異点を認識するような受動的な情報処理では達成され得ないし，また既有の一般的な知識をやみくもに当てはめてみるだけの説明づけであっても，そこから新たな知識の生成を望めるものではない．学習における知識の役割は，その知識を使って対象を理解することの他に，逆に既有の知識の範囲内では理解できない，説明できないこと，すなわち「矛盾」（インパス（impasse））を検知するためにこそ存在価値があるのであり，このことが学習システムに創造的な学習を達成させるためのトリガーとなるものである．

このような矛盾の検知は事例からの学習を進める上で，２つの重要な役割を担うことになる．一つは事例の中のどの部分で矛盾が生起したかによって，学習すべき対象を特定できること．そして，矛盾検知によってそれまでの受動的な情報処理から，学習達成への能動的な情報処理への切り替えのタイミングが示唆されることになる．ここでいう「能動的な処理」とは，以下に示すように自らの有している知識に対する新たな仮説生成やメタ知識への言及，過去の事例からの類推や対象モデル構築に基づく実験の創案など，様々な形態が考えられるが，これらの情報処理を経て矛盾が解消されるとこの解消のためにとられた一連の手続が，新たな知識単位として学習獲得されることになる．

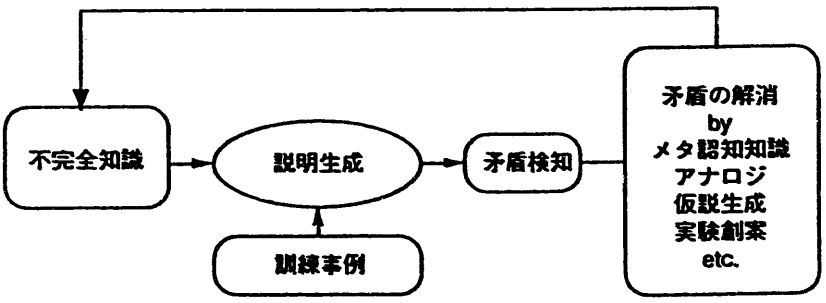



Fig. 3.1 矛盾解消による知識獲得

1. 仮説生成による矛盾解消：何か新しい情報が与えられた場合，我々は通常，自分のもつ知識に何らかの仮説を付け加えて，その情報を説明しようとする．そしてその説明に矛盾がなければ与えられた情報が説明できたとみなす．勿論ここでは，仮説を生成・選択する方法や，無矛盾の説明を行える複数の仮説から一つを選択する方

法，そしてやはり説明づけられないと解った時点で，どこまでバックトラックして仮説を棄却すればよいか，についての解決策が必要になる．

2. メタ認知知識による矛盾解消（Repair Theory）：従来から，認知心理学の分野においては，CAI システム構築のために，幼少児の学習プロセスの解明が様々な実験を通して進められてきた．そこでは，人間の犯すヒューマン・エラーには，系統的な systematic error と，ランダムに発生する slip の 2 つの形態が存在するが [28]，このうち前者のエラーが，ダイレクトに適用可能な知識が見いだせなかったり，不完全な知識しか持ち合わせていないが故に，一意に決めかねて思い悩むといった，いわゆる矛盾の検知であって，このような矛盾解消の手だてを学習者がパターン化していることから系統的なエラーが生じるというものである．すなわち矛盾が生じると同時に，学習者の視点はそれまでの主なる問題解決の視点から，その矛盾回避のための局所的な問題解決に切り替えられ，問題空間が全く別のものに切り替わると共に，そこでの新たな副目標が生成される．このとき，どのような矛盾のもとでは，どのような問題空間・副目標への切り替えを行うべきかを主問題の領域とは独立に一般的に規定したものがメタ認知知識である．そして，一旦この矛盾を解消するに十分な一連の操作が見いだされると，これを repair として，新たな知識単位として生成することになる．以上が，学習者のバグ生成を説明づけた repair theory である．この考え方に基づいた学習システムは，既に CMU で開発された SOAR の中にも見られる [20]．8 クイーン問題解法時に複数の番手の中から次手を決めかねる矛盾が生じると，8 クイーンの問題空間から，候補オペレータを評価するという副目標に，そしてオペレータの評価空間に切り替わり，局所的な問題解決を実行することによって，チャンク生成を行う手法を取り入れている（図 3.2）[1]．

3. 過去事例からの類推による矛盾解消：以上のように，発生した矛盾の解消について，知識として規定したり，演繹や解析による矛盾解消が不可能なような問題対象においては，過去の事例からの類推（analogy）に依存せざるを得ない．旧事例（base）を参照し，その中の特定の関係を写像することによって，新事例（target）で成立することを推論する類推の定式化においては，事例中のどの関係に焦点を

[1]ここでは「矛盾」と考えるよりは「行きづまり」（breakdown）の意味として捉えるほうが妥当である．

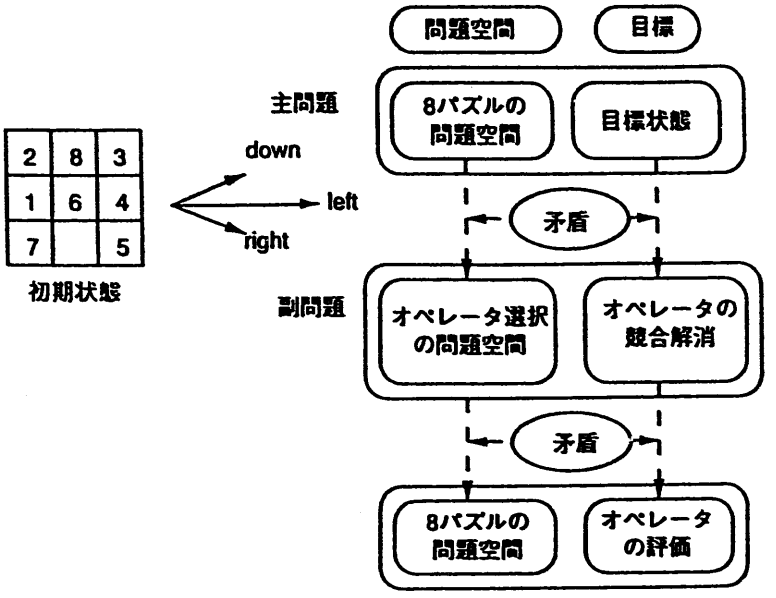



Fig. 3.2 SOAR における矛盾解消による知識獲得

絞るのか，そして，多数ある過去事例の中からどれを base として採用するのかが常に問題になるが，これは過去事例を，共通の「矛盾」の分類毎に，あるいは矛盾解消に取られた共通の手だて（矛盾に対する説明構造）毎に，組織化しておくことで，新事例処理に伴う矛盾発生時に，効率よく過去事例を検索してくることが出来る．また同一事例上での複数の矛盾に対しても，各々の視点からの参照事例を集めることによって，Winston の ANALOGY のシステム [40] で見られるようなパッチワーク的な問題解決を系統的に押し進めることが期待できる．このような矛盾に基づく事例記憶の提案は，既に Schank により失敗からの想起（failure-driven reminding）として記憶構造のモデルの提唱がなされている [35]．また学習システムへの応用としては，後述する Schank の SWALE のシステム [34] や Carbonell[2] らに見られる．

4. 実験の創案による矛盾解消：以上までに述べた矛盾解消が，既有の知識・記憶構造に対する学習システム内部への交渉によるものであったのに対し，システム外部の教師や教材に交渉を持つことにより矛盾を解消することが考えられる．例えば学習事例として与えられたものが，既有知識の範囲内では矛盾が発生して理解ができない場合に，自らの知識の一部を疑問視する仮説の生成を行う．そしてこの仮説の正

当性を実証するために，学習事例中のパラメータを修正したり新たな実験事例を創案し，この実験事例に対する外部の判断に基づいて，元の学習事例の理解を達成し，そこから新たな知識獲得を目指すというものである．ただし，このような実験事例を創案する際には，学習システム自身が実験結果をある程度予測しておく必要があり，そのためには学習対象に対する的確なモデル構築のための手法，例えば，定性物理・定性推論の導入が不可欠となる [32],[31]．このような考え方にたった事例研究を次節以下で後述する．

矛盾の検知は，与えられた学習事例が既有の知識の範囲から外れることを推論しなくてはならないが，この既有知識から既知とされる世界と未知の世界の境界をどこに定めるかについて，蓋然推論（default reasonig〉や状況推論（circumscription）などの各種非単調推論〈nonmonotonic inference）や常識推論（commonsense reasoning）など，人間の高次推論の定式化に負うところが大きい．

## 3.2 説明からの類推による矛盾解消

Schank らは，従来より自然言語処理を手掛け，人間の言語理解のモデルとして，CD（conceptual dependency）理論やスクリプト理論を発表し，それまでの自然言語研究の大半が入力文の構文解析や，主語・述梧，係り受けの分析に主眼がおかれていた自然言語研究に，意味的な曖昧性や文脈理解のための知識が必須であることを提案して，その後の自然言語研究のみならず，画像・信号の解釈・理解等の分野においても多大な影響を残した．その後，彼の研究はプラン（plan）・ゴール（goal）・MOP（memory organizing packet）へと進み，固定化されライブラリー化された当初のスクリプトの考え方から，より大局的かつ柔軟な情報処理を可能にするべく，高階の知識構造の導入に視点が移ってきた．以上のような考え方は，彼の弟子達により，学習のテーマへと展開されていくことになるが，そこでは Mitchell や Michalski に代表される論理的な機械学習のアプローチに対して，心理学のスキーマ理論 [1] に根拠をおきつつ，動的に変化する人間の記憶構造（知識構造），特に意味記憶・エピソード記憶の構造に根ざした独自のアプローチとして展開を見せている．この一連の研究の延長として，説明と理解そして学習の間の関連性に着目した学習システムを提案している（図 3.3）[34].

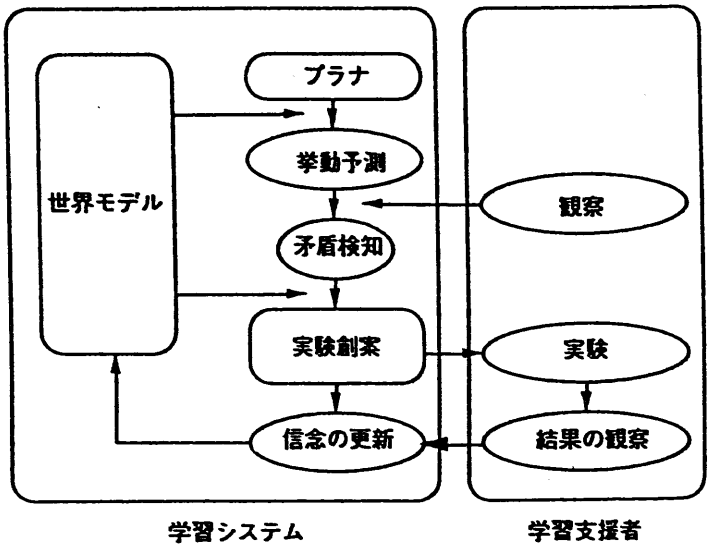



Fig. 3.3 SWALE における矛盾解消

入力された文章に対して，それが位置づけられるべき既有の記憶構造が見いだされたとき，理解されたことになる．換言すれば，それが適切なコンテクストのもとに置かれたとき，その意味が見いだされた（make sense）とみなせる．これに対し，入力された文章が予期された結果とくいちがう，すなわち適切なコンテキストが存在しない場合，あるいは決定できないときには，何故にそのような違いが生じたのかを理由付ける説明が必要になり，ここでの説明生成こそが，スクリプトタイプの理解に代わって真に深い理解へと導く，創造的な学習を達成するものである．ここで説明の果たす役割は，理解の失敗，即ち予測の失敗からの回復であり，表層的な特徴からだけでは結び付けようのないような過去の類似事例を想起するためのインデックスとしての機能である．SWALE は，このような視点のもとテキスト入力されたエピソードに対し，過去の説明を検索し，これを現状に合わせるように修正することによって，新たな説明を生成するプログラムである．

このプログラムの動きを SWALE の例題から説明する．まず，入力文「優駿 Swale は，ベルモントステークスに勝利した 1 週間後に馬小屋で死んでいるのが発見された」が与えられると，これを既有の知識構造の中に位置づけるべく処理を行う．これが難なく位置づけられたときには，新たな説明生成は行われないが，本例の場合，Swale の死が突然であったことから，temporal anomaly を認識する．ここでシステムはこの入力がダイ

レクトな過去の記憶への同化が困難なことを認識し，説明生成のメカニズムを起動することになる．

システムは過去の説明パターン（explanation patterns，以後XPと略す）の検索に取り掛かるが，この探し方には，説明されるべきイベントの一般的なカテゴリ毎に組織化されたXPを探す方法（rotine pattern search）により，本例の場合には「病死」とか「事故死」とかのタイプのもとでインデックスされたXPが検索される．各XPには，その説明構造が適用可能なanomalyの種類や，その説明構造が適用できるための前提条件が規定されている．システムはこれらのXPの前提条件が満足されているか否かを入力テキスト中に情報を調べながらチェックするが，満たされていない場合には，方略を変えて新たなXPの検索を行う．すなわち，この入力の異常さを呈している根拠は，どのような属性であるのかを調べ，この異常属性に着目したXPの検索を行う（unusual feature search）．

本例の場合，「体調が万全のもとでの突然の死」という異常属性のもとで，Jim Fixx XP＝（ジョガーがジョギングで体力を消耗するあまり心臓マヒで死にいたらしめられた）を検索し，このXPの前提条件があてはまるか否かを調べるが，入力の主人公Swaleはジョガーではないので，ダイレクトにこのXPを適用することはできない．このような矛盾が発生すると，システムは「主人公に食い違いが生じて，XPが使えないときには，主人公が関わる他のテーマに視点を転ずる（substitute alternate theme)」というメタ認知知識（SWALEではtweaking rulesと呼ばれている）に基づいて，「競馬」（horse-race）としての視点から捉えなおす．システムは，この「競馬」の概念を意味解折し，ジョギングとの融合点（共に「走る」行為を伴う）が見いだして，Jim Fixx XPの「ジョガー」をSwaleで置き換えた新たなXP：(Swaleは心臓が弱かったのでレースで体力を消耗し心臓マヒを起こした）を仮説として生成する．そして，テキスト入力された他の情報が，この仮説と矛盾をきたさないか，この仮説をcomfirmできるか否かについて処理を進める．矛盾が生じなければ，システムは元のXPおよび，現在修正されたXPの両者に共通する項目を調べ，これらを一般化して，両説明構造を包含するようなXPにとりまとめる．すなわち，本例の場合，「体力的に激しい運動をしているものの突然死」が引金となって起動されるような一般化説明構造を作り上げることになる．

ここでの説明は，従来のEBGのように，一般的な領域知識に基づいて演繹的に生成

されるわけではなく，パターン化された行為—状態の系列，いわゆるスクリプト的なものである．また矛盾検知の役割は，既有の説明が使えないということを見極め，類推の進め方を一般的に規定したメタ認知知識に制御を移すことにより，解決すべき副問題を設定する（上例では，「ジョガーと競争馬 Swale の連想関係を意味解析により見いだす」という副問題）．そしてこの結果から仮説として立てられた類推の正当性を検証し新たな説明パターンを生成すると共に，各事例はこの説明構造のもとでインデックスづけされることになる．

このように，SWALE では，与えられた入力エピソードに対して過去の事例エピソードを引っ張りだしてその上で類推を進めるのではなく，過去の事例に対して与えられた説明構造の上での類推を行うことによって，表層的な事例記述にとらわれず，また事例を分解して解析するということはせず，本質的かつ深層レベルで類似する事例を探り当て，事例の全体像を残しながら，創造的な知識の学習を実現している．

## 3.3 実験に基づく学習

Rajamoney らは，従来の説明学習が既有の知識を改善したり，範囲を拡大するものではなく，スキーマ構造へのチャンク化を通して，知識利用時の効率をあげるためのものにすぎない（schematic knowlede acquisition）として問題を提起し，不完全な既有の知識では解決できなかったことが学習により解決できるようになるような，真の意味での創造的な学習（conceptual knowledge acquisition）を達成するための手法の必要性を訴えている [29]．そのためには，学習システムが不完全なりの既有知識を用いて，世界モデルを構築し，そのモデルに基づく予測と実際に観測される対象との差異・矛盾を，自らのモデルにフィードバックして修正し仮説を立て直した上で，この仮説を実証するべく外部に情報を求める．すなわち，外部環境のもつ意味を認識するため，学習システム自身が initiative をとって能動的な外部対象との相互交渉を経ることにより初めて学習が達成されるとし，その考え方に立ったプロトタイプシステムを試作している．

彼のシステムの概念的処理フローを図 3.4に示す．以下 Rajamoney らの例題に基づいて説明する．まずシステムでは初期状態として不完全な世界モデルを信念（belief）の集合で表している．この中には「固体を境に隔てられている流体の間に流れは生じない」

という信念が含まれている．ここでシステムは，図3.4に示すような観察を行ったとする．すなわち，固体の壁に隔てられた2つの容器にそれぞれ濃度の異なる同分量の水溶液が，各々別の容器からそそぎ込まれ，放置されている世界を観察しているとする．システムは，この観察と並行して，観察の各ステップにおいて，自らの既有の知識に基づいて，次はどのようになるはずかについての予測を行っている（例えば，注ぎ込まれるプロセスでは，元の容器の水溶液の量は徐々に減少し，注ぎ込まれた容器内の量は徐々に増えて，やがて同分量になり，そのまま放置しておいても，分量は変わらない）．これに対し，世界での観察が，放置されている時間の進行にともなって，各容器の分量が同図に示すような変化をきたしたとする．システムは，「浸透圧現象」の知識を有していないので，自らの予測と食い違った「異常」(anomaly) を認識し，この異常を既有の知識の範囲内で何とか説明づけるべく，液体の量の変化を引き起こすような物理プロセスを，知識ベースから引出し，これによる説明を試みる．

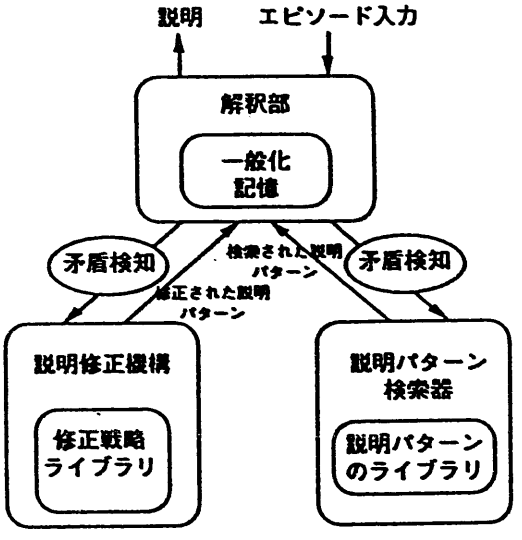



Fig. 3.4 実験創案による矛盾解消

本例の場合，蒸発 (evaporation)・凝結 (condensation)・吸収 (absorption)・放出 (release)・流れ (flow) のプロセスが想起されるが，いずれのプロセスもその前提条件が自らのもつ世界モデル上で満足されないので，説明不能に陥る．ここで矛盾を検知し，システムは現時点で世界モデルとして有している信念の再チュックに取り掛かる．すなわち，世界モデルの中から，上述のプロセスの前提条件と矛盾を引き起こしているもの

を特定しにかかる．すなわち，(1) 境界の固体が吸水性の材質ではないと分類したこと，(2) 流体が流れるにはクリアな，固体で隔てられていない流路が必要であること，(3) 蒸発・凝結には外気との接触が必要であること，の３つの信念である．観察が有効であるためには，これらの何れかの信念が誤りでなければならないが，そのような信念を特定するべく，どの物理プロセスが進行しているかを識別できるようにシステムは実験を創案する．

システム内の物理プロセスに関する知識には，そのプロセスに介在する各種物理量と，プロセスが進行する速さとの間の定性的関係が規定されている．この知識を用いてシステムは，特定の物理量のパラメータを変更した実験を創案する．すなわち対象の原因となっているプロセスを仮想的に一つ選定し，そのプロセスを速め，かつ競合するプロセス以外のプロセスの進行は逆に遅らせる方向にパラメータを変更する．その上で，実験の結果が当初観察された進行速度よりも格段に早められたものであるときには，この着目したプロセスが観察された現象の原因であることを示唆することになる．図 3.5(c) は，流れ（flow）のプロセスに着目して提案された新たな容器である．システム内には，同図 (b) に示すように「流れプロセス」が定義されているため，これへの参照から，両水溶液間の流路断面積を最大化し，流路長を最短化したものになっている．

実際にこのような実験を通し，２つの水溶液間流れの進行速度が速められたことを確認することによって，「流れプロセス」が原因となっていることを実証し，システムはその前提条件であった，「流体が流れるにはクリアな，固体で隔てられていない流路が必要である」という信念を世界モデルから抹消し，「流出プロセス」に類似した新たなプロセスを創り出して，モデルを改訂していくことになる．ただしここで創案されるプロセスは，浸透圧の原理を理解した上でのものではなく，実験に見られたように「特定の二種類の水溶液が特定の固体を境に接したときに流出が生じる」といった，一般的な流れプロセスを特殊化したものにすぎない．このような，特殊例から一般的な浸透圧現象の知識を獲得していくためには，上述のような特殊な実験結果，ならびに本質的な介在物理量を特定するための一連の実験の創案と結果の観察を繰り返しながら，より一般的な概念学習（すなわち，浸透圧現象のプロセススキーマの獲得）を進めていかねばならない．

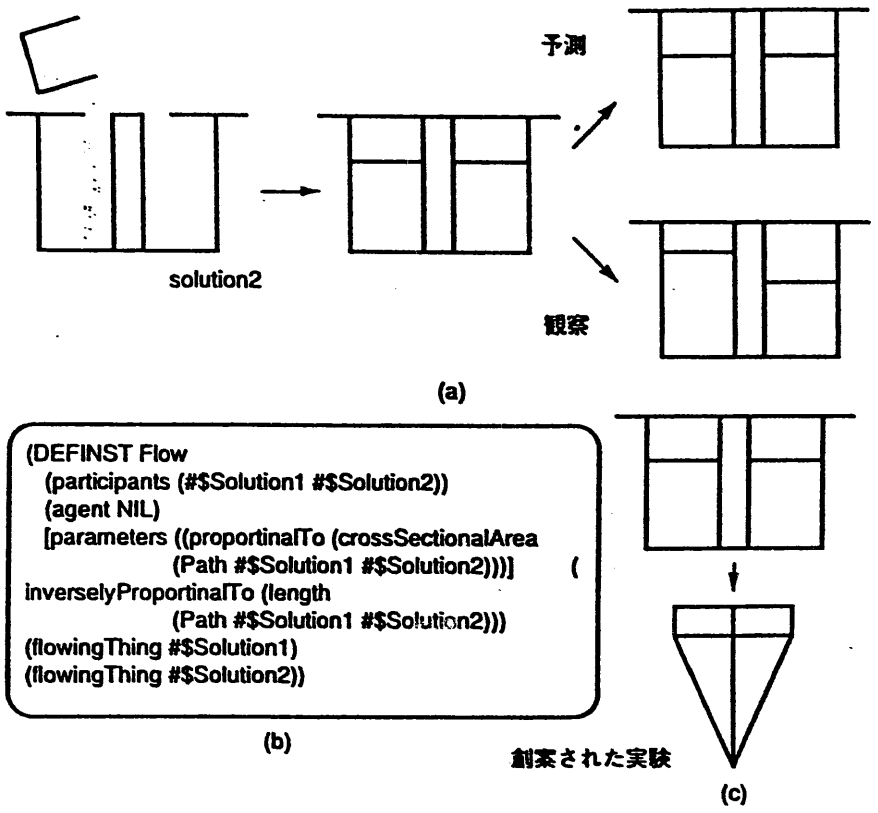



Fig. 3.5 実験創案による矛盾解消の具体例

## 3.4 事例ベース推論における矛盾検知とその解消

図3.7はMITのKotonにより開発された，事例ベース推論を用いた心臓疾患診断システムCASAY[18]からの例である．ここでの事例は，同図(a)に示すように各患者の兆候データ，テスト結果，病歴を含む約40項目の［項目－値］の特徴リストからなる問題記述と，その患者に実際に施した治療内容とその結果，さらにこれらの観察データが互いに患者のどのような内部状態を介して結び付いているかを表す因果モデル（同図(b)）を事例記述として有する．システムは，新規の患者（同図(c)）における問題記述から，これと最も類似する既知の事例を照合し，そこでとられた治療と因果モデルをマッピングすることによって，当該患者の疾患を説明づけるような因果モデルと治療を提示するというものである．

勿論このような場合，問題記述の上で完全な照合をみることは希であり，一方に観察され一方に観察されない項目，項目は一致してもその値に食い違いが生じていたり，一

| Principle Name | Prior Case Value | New Case Value | Repair |
|---|---|---|---|
| Rule Out | Present | Incompatible | Ignore |
| Other Evidence | Supports State S | Missing | Find Another Support for S |
| Unrelated Prior Case | Not Used | Missing | Ignore |
| Supports Existing State | Missing | Present | Find Support in Prior Case |
| Unrelated New Case | Missing | Not Used<br>Abnormal | Mark as Unexplained |
| Normal | Missing | Present | Ignore |
| No Info. in Prior Case | Missing | Absent | Assume Absent |
| No Info. in New Case | Absent | Missing | Assume Absent |
| Same Qualitative Region | V1 | V2 | Accept if V1 and V2 in same region |

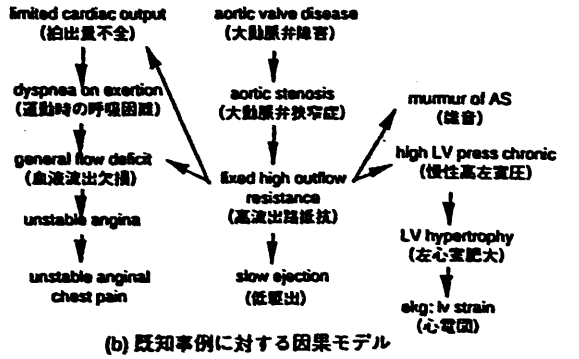
Fig. 3.6 CASAY での Eidence Principle

方では値が未知である場合などが通常である．CASAY ではこのような不一致項目を因果モデル上で評価するための知識として，問題領域に依存しない普遍的な経験則を図 3.6に示すような Evidence Principles として用意し，この規則に則った修正操作を適用しながら，既知事例—当該事例間の相違点が許容できるものであるか否かを調べる．

各々の Evidence Principles には，不一致が生じた場合の既知事例の因果モデルに対してどのように修正（adaptation）すればよいかについての方策（repair）が関係付られており，これに従って修正されたものが，当該患者の疾患を説明付ける因果モデル（同図 (d)）として決定され，治療策を提示するというものである．

| No. | Feature Name | Value for prior case |
|---|---|---|
| 1 | age（年齢） | 72 |
| 2 | pulse-rate（脈拍数） | 96 |
| 3 | temperature（体温） | 98.7 |
| 4 | orthostatic-change（立ち眩み） | absent |
| 5 | angina（アンギーナ） | unstable |
| 6 | mean-arterial-pressure（平均動脈血圧） | 107 |
| 7 | syncope（失神） | none |
| 8 | auscultation（聴診） | murmur of AS |
| 9 | pulse（脈） | normal |
| 10 | ekg（心電図） | normal sinus & lv strain |
| 11 | calcification（石灰化） | none |

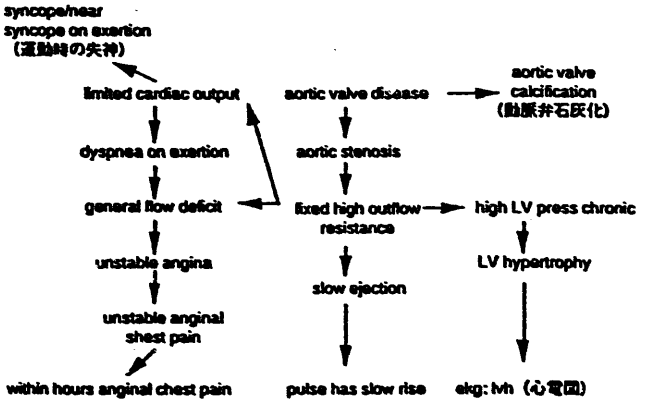
(a) 既知事例の問題記述




(b) 既知事例に対する因果モデル

| No. | Feature Name | value in new case |
|---|---|---|
| 1 | age | 65 |
| 2 | pulse-rate | 90 |
| 3 | temperature | 98.4 |
| 4 | orthostatic-change | unknown |
| 5 | angina | within-hours & unstable |
| 6 | mean-arterial-pressure | 99.3 |
| 7 | syncope | on exertion |
| 8 | auscultation | unknown |
| 9 | pulse | slow-rise |
| 10 | ekg | normal sinus & lvh |
| 11 | calification | mitral & aortic |

(c) 新規事例の問題記述




(d) 新規事例に対する修正された因果モデル（　は既知事例との共通部）

```
(substitute-evidence hr:90 hr:96)
(remove-evidence mean-arterial-pressure:107)
(add-evidence within-hours unstable-angina)
(add evidence syncope-on-exertion limited-cardiac-output)
(remove-evidence murmur-of-as)
(remove-evidence lv-strain)
(add-evidence lvh lv-hypertrophy)
(add-state aortic-valve-disease)
(add-evidence aortic-calcification aortic-valve-disease)
(add-state mitral-valve-disease)
(add-evidence mitral-calcification mitral-valve-disease)
```

(e)

Fig. 3.7 CASAY のシステム

# 第４章　概念学習手法の概要

## 4.1　概念学習について

機械学習は当初エキスパートシステム（ES）の核となる知識ベースの構築，すなわち専門家の有する知識を ES に移植するプロセスを自動化するものとして，さらに ES がその経験からその後の問題解決効率を向上させたり，解決可能な問題の範囲を拡大していけるような学習あるいは広い意味での適応の能力を具備させるための，いわば ES 構築のための付加的な方法論として研究が進められてきたが，近年では広くエージェント（＝知識を有する行動主体）が具備すべき知的能力の核なる能力としてその重要性が認識されている．サイモンは学習の定義として「あるタスクの遂行の効率を改善するように過去の経験を組織化すること」としている．本章ではとくに概念学習（concept learning）と呼ばれる分野の学習について詳述する．概念学習はニューラルネットのようなパラメトリック学習の手法に対して，知識ベース部分をシンボル（記号）により明示的に表現された概念として獲得する学習手法である．概念学習はシステムに与えられる学習材料がどの概念に属するかについての情報が与えられるか否かによって，すなわち教師の有無によって例題からの学習（learning from examples）と観察による学習（learning by observation）に分けられる．例題からの学習ではシステム外部の教師が，環境から入力される事例がどの概念に含まれ，あるいはまたどの概念に含まれないかについての情報を与え，これよりシステムは前者のみを含み後者を除外するような一般的知識を出力するものである．２章で述べたバージョン空間法もこれに該当する．一方，観察による学習ではシステムに与えられる学習対象がどの概念に属するか，またどれだけの種類の概念が存在するかについても一切与えられることなく，ただ入力される大量のデータを適切に分類する方法を，データ自体から獲得するものである．

観察からの学習は，別名教師なし学習（unsupervised learning）とも呼ばれる．これは学習システム外部の教師が，与える学習材料（対象）に対する類別や評価，報酬が与えられないもとで，これらの学習対象を組織化する範疇の形成を行なうものである．通常このような範疇の形成は前節で述べた心理学的な知見に倣い，階層的な多段の分類概

念の階層の導出を目的とするもので，このことから概念形成（concept formation）あるいは概念クラスタリング（conceptual clustering）とも呼ばれる[1]．

## 4.2 概念に関する心理学的知見

概念とは，普通我々が使う言葉を用いた文章に現れる言葉の一つ一つが，あるいはまたこれらの言葉が結び付いてより大きなまとまりとしての知識を表現する際のモジュールの一つ一つを表すこともある．どのような概念においても，それに属するものは一つとは限らないしその定義自身，明確に規定できるものではない．ある概念に属する要素の全体—具体的事例の集合—を，その概念の外延（extension）という．ある事例が，ある概念に属するか否かの判断を行なうには，その概念を定義する特徴，いわば概念集合の特性関数に相当するものが必要である．この概念を規定する特徴を内包（intension）という．概念は通常図4.1のように階層構造のもとに組織化されており，その下位概念は上位概念の性質（内包）を合わせ持つ，すなわち継承（inherit）している．このような概念構造において，下位概念の外延は上位概念の外延の部分集合であるのに対して，内包すなわちその概念で成り立つ特徴の集合（内包）は下位になるほど増大する．

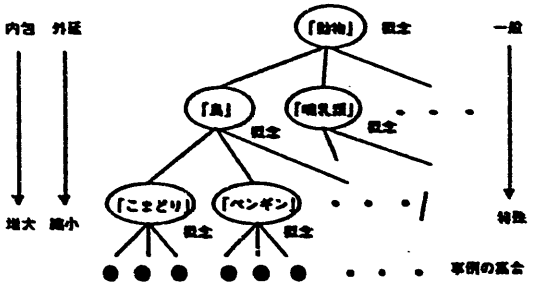



Fig. 4.1 概念の階層構造と内包・外延

心理学における概念形成にまつわる研究としては範疇形成（category formation）の研究として展開されてきた．すなわち被験者がある範疇，すなわち概念の事例（＝外延）を観察することによって新たな範疇を発達させるのはどのような処理によるものかを問

[1]研究者によっては，概念クラスタリングは原則として事例集合から一括式に平板な一レベルでの分類を生成する手法を，これに対して概念形成は多段の分類階層を逐次的に生成する手法として区別する場合もあるが，本章ではとくにこれらの間の区別をしない．

うものである．とくに「トリ」や「スポーツ」のような現実世界で起こる事物の分類は自然範疇（natural category）と呼ばれるが，このような範疇の構造に関する研究で重要なものとして Rosch の一連の研究の概要を以下にまとめる [30]．

1. 族類似（family resemblance）の概念：ある項目がある範疇の事例であるかどうかを何らかの固定された特徴群が定義することはありえない．範疇の成員（事例）間の類似は，家族の成員の間の類似と同じで，例えば，息子が母親に似ており，母親はその姉に似ているかも知れないが，息子とは違った特徴の上で似ているのであり，すべての家族が共通に有している特徴などは挙げることができない．にもかかわらず異なる家族の成員が混じりあった人達の間でこれを２つの家族に分類する際には，成員間で部分的に共有されている特徴の現れ方を観察するだけでも比較的正しい分類が可能となる．
2. 典型性効果（typicality effect）：上述の族類似の概念は，範疇を構成する個々の成員の間でその典型性の度合いに違いがあることを指摘している．このような典型性評価は実際には，知覚的再認実験，すなわちある項目を提示してこれがある範疇の事例になっているかどうかを尋ねた際の判断を出力するまでの応答時間を計測することによって評価でき，応答時間が短いものほどその範疇の事例として典型性が高く，長いものほど非典型，いわばその概念の中心に位置づけられるのではなく周辺的な成員ということになる．Rosch and Mervis らの実験では，典型的な成員であればあるほど，他の事例とより多くの特徴を共有する傾向があることを発見し，自然範疇が固定した境界を持たないことを主張した．そして被験者が自然範疇を学習しているのは，それに固有な特徴の集合（内包）を抽出しているわけではなく，またその範疇に属する成員の一つ一つ（外延集合の元）を記憶しているわけもなく，範疇の最も典型的な特徴の組み合わせとしての中心化傾向（central tendency）を記憶しているに過ぎないのであり，新たな成員がその範疇に属するかどうかを判断する際の手助けとしてこの中心化傾向を利用しているという点を強調している．
3. 自然範疇における基底レベル（basic level of natural categories）：

## 4.3 クラスター分析

教師なし学習と類似の目的を設定した従来研究の分野としては，各種実験科学，社会科学の分野で展開されてきたデータ解析手法のうち，外的規範をもたない手法，例えばクラスター分析（cluster analysis）がある．そこではデータが潜在的にもつ区別（範疇の別）を明らかにするべく，オブジェクト（事例）を多次元空間上の座標値ベクトルなど数値データ群で記述し，この空間上で定義される類似度あるいは距離尺度のもとに，類似したオブジェクト群を同一のカテゴリー（クラスター）にそして非類似なオブジェクト同士は互いに異なるカテゴリーに属するようにオブジェクトを配するものである．

クラスター分析での議論の中心は，まずカテゴリー生成における方略の違いである．すなわちすべてのオブジェクト同士の間の類似度を算出し，その中で最も類似した最近隣のもの同士を一つのカテゴリーに統合する処理を順次繰り返しながら最終的に全オブジェクトを包含するカテゴリーをボトムアップ式にカテゴリーを見いだしていくやり方(aggromerative startegy)，さらにトップダウン式方略として，まず最初に全てのオブジェクトを包含するカテゴリーから開始し，これをより細かなカテゴリーに順次分割していくやり方(divisive strategy)が考えられる．いずれの戦略をとるにせよ階層的なクラスター生成が可能になるが局所的な一対間での類似性の比較からのみでは，有用なクラスターが得られないとの指摘がある．一方上述のような階層的なカテゴリー形成を目指すのとは異なり，分割のレベルを固定した上で，その最適な分割を探索する手法としてのクラスター分析に関する研究も盛んに行なわれている．ここではオブジェクトの一対間の類似性尺度に基づくのではなく，クラスターそのものの質，すなわち同一クラスター内にどれほど類似したオブジェクトが集められているか，ならびに異なるクラスター間での非類似性をいかに維持できるか，の両基準をどのように満足させながらクラスターを生成するかによって生成されるクラスターの数が異なってくる．そこで通常は得るべきクラスターの数を予め指定し，そのもとでの繰り返し最適化問題としてオブジェクトのカテゴリーを決定するものでその代表的な手法として k-means clustering がある．このような繰り返し最適化を用いたクラスター分析の手法としては，複数のクラスターが互いに重複するような，すなわち一つのオブジェクトが同時に２つ以上のクラスターに帰属することを許容するようなファジィクラスタリングの手法も提案されている．

以上のようなクラスター分析の手法における限界としては，まず第一にその扱えるデータが基本的に数値データ群に限定される点である．しかし実際にはデータを記述するためには，数値の他，「大」「中」「小」などの順序数，「三角」「四角」，…，「多角形」のような階層構造や包含関係をもった分類量，「赤」「青」「白」などの名目量の組として与えられることが多い．そして事例の持つ特徴を精緻に表現しようとするならば，単純な属性の集合としてのみ記述できることはまれであって，オブジェクトがどのような要素から構成されており，また属性同士の間にどのような関係が成り立っているか，属性を埋める値が他の値との包含関係など，概念を構成する意味上の関係を明示することは必須となる．つぎにクラスター生成とその利用法，すなわち生成されたクラスターに基づいて新事例を分類する，という問題は基本的には切り離されている点である．従って通常のクラスター分析では，その出力結果を眺めながらデータ表現のスケーリング，関連属性の決定，重みの決定などを変更して試行錯誤的に，求める目的に合致した結果が得られるまで繰り返し実行されるが，その際の探索の制御は基本的にユーザに任せられる．さらに第三の限界として，生成されたクラスターはそれに属するオブジェクトを指示するのみであり，いわば前節で述べた概念の外延集合を特定するに過ぎない．一方の概念の内包記述，すなわち生成されたクラスターのもつ意味に対する解釈を生成する能力はなく，外部におかれた解析者の洞察力に委ねられる．機械学習の観察による学習では，このような従来のクラスター分析の限界を克服するものである．

## 4.4 観察による学習：概念クラスタリング

クラスター分析を機械学習で実現したものに概念クラスタリング[25]がある．これは前節末尾で述べたようなクラスター分析を利用する際の総合的な分析タスク，すなわちオブジェクト集合を分割してクラスターを生成するというタスク（clustering task）と各々のクラスターに対する概念的な意味づけを行った解釈を与えるというタスク（characterization task）を一括処理で行う点がクラスター分析との大きな違いである．後者のタスクはクラスターが包含する外延集合に対してこれらの各元において真なる内包記述を生成することである．このような概念クラスタリングとして最も著名なシステムに Michalski らによる CLUSTER/2 というシステムがある．このシステムでは拡張述語論理と呼ばれる

形で表現された事例に対して，k-means 法の繰り返し最適化に類した処理を実行しながらクラスターを生成するが，ここでのプロセスはクラスターに属する全ての（あるいは大部分の）事例において共有されている特徴群の連言記述が以下のような基準を満たす意味で適切なものとなるようにガイドされる．

1. 簡潔性（simplicity）：クラスタを特徴付ける記述は“単純”で対象のクラスへの割り当てやクラス間の区別が容易にできるものでなければならない．
2. 適合性（fit）：クラスの記述は実際のデータに“よく一致する”必要がある．

一般に非常に正確な“一致”を達成するためには記述は複雑にならざるを得なくなるのが通常であり，従って単純さと適合性の要求は互いに矛盾するものである．

ここでは例として，図 4.2(a) に示すような 10 個の対象（事象）がそれぞれ 4 つの属性変数によって記述されたものを考える．それぞれの変数の定義域は 3 値（3 種類の値）であり，$x_1, x_2$は線形変数，$x_3, x_4$は各々構造変数，名称変数とする．ここでは簡単のためこれら 10 個の対象を 2 つのクラスに分類することを考える．図 4.2(b) は同図 (a) の対象を拡張論理図（カルノーマップ）上にプロットしたものである．処理の手順は以下のとおりである．

1. まず 10 個の対象から求める分割数（$k = 2$）に等しい数の種（シーズ）を選ぶ．ここでは仮に $e_1, e_2$を選ぶことにする．
2. つぎにスターと呼ばれる記述を生成する．スター $G(e \mid E_0)$ とは，事象 $e$ をカバーするが $E_0$に含まれる事象はカバーしないような最も一般的な複合体すべての集合のことである．なお複合体とは，属性 $x_i$が値 $V_{ij}$をとることを $[x_i \# V_{ij}]$ のようなセレクタとして関係表現するとき，セレクタの論理積として概念を記述したものである．スター中の複合体 $G(e \mid E_0)$ は対象の最も一般的な記述であるので過剰な一般化を行っているかもしれない．そこで観測事象をカバーしつつ各複合体の稀薄度（sparseness），すなわち複合体が表わす外延集合の数に対する実際に観測されている事象の数の相対的比率，をできる限り減らした縮小スター $RG(e \mid E_0)$ を生成する．アルゴリズムでは各々の種に対して他の種をカバーしないようなスター

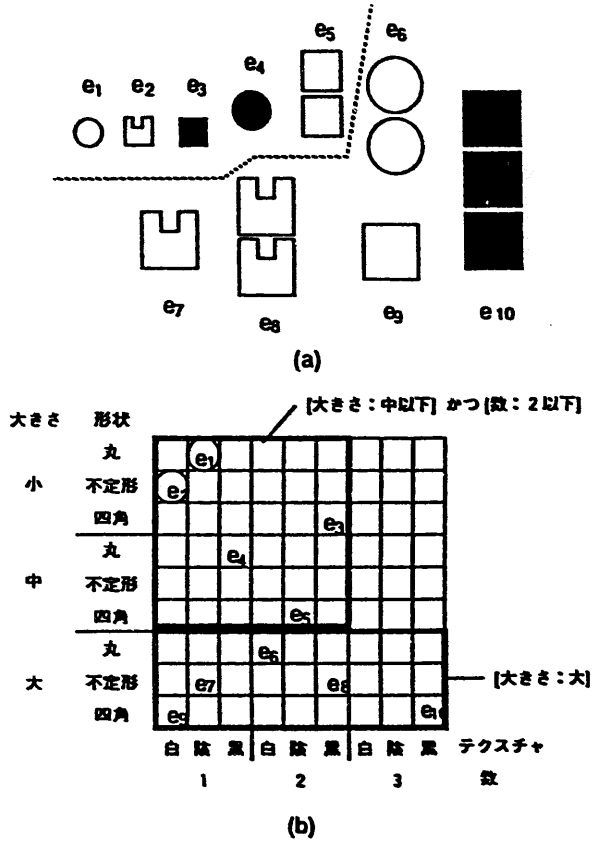



Fig. 4.2 CLUSTER/2 による概念クラスタリング

$G(e_1 \mid e_2), G(e_2 \mid e_1)$ を生成する．例題では，

$$\begin{aligned} RG(e_1 \mid e_2) &= \{[x_2 = a][x_3 = 0 \wedge 1], [x_4 = 1 \wedge 2]\} \\ RG(e_2 \mid e_1) &= \{[x_2 = b \wedge c], [x_4 = 0 \wedge 2]\} \end{aligned} \tag{4.1}$$

が得られる．これらのうち順序変数に対して区間連続化規則を，また構造変数に対して概念一般化規則を適用して，

$$\begin{aligned} RG(e_1 \mid e_2) &= \{[x_2 = a][x_3 \leq 1], [x_4 = 1 \wedge 2]\} \\ RG(e_2 \mid e_1) &= \{[x_2 = f], [x_4 = 0 \wedge 2]\} \end{aligned} \tag{4.2}$$

が得られる．ここで変数 $x_2$ に対する値 "$[x_2 = b \wedge c]$" はより一般的な値 $f$ で置き換えられている．

3. それぞれのスターを構成する複合体を組み合わせて $k=2$ 個の複合体が全事象集合 $E$を互いに重複することなくカバーするように変換する[2]．例題ではそれぞれのスターが2つの複合体で構成されているので $2\times 2=4$ 組の複合体が考慮の対象となるが，このうち素なクラスターとしては以下の複合体，

   $[x_2=a][x_3\leq 1]$

   $[x_2=f]$

   のみとなり，制約を満足しない複合体に対しては互いに素なクラスターになるように分離手続きがとられる．生成された複合体の組み合わせのうち，クラスターの質の評価基準である稀薄度と複雑度を算出し，最適あるいは準最適な組を残す．上記では前者の複合体が稀薄度 15 ，複雑度 2，後者が稀薄度 47，複雑度 1 となる[3]．
4. 終了基準，すなわち予め決められた適用回数に達しているか否かをチェックし，終了基準を満たせば終了，それ以外の場合は次ステップに進む．
5. 前々ステップで求められた複合体 $[x_2=a][x_3\leq 1]$ と $[x_2=f]$ によってカバーされている事象は各々 $\{e_1,e_4,e_6\},\{e_2,e_3,e_5,e_7,e_8,e_9,e_{10}\}$ であるので，これらの各々の集合の中の中心事象（形式距離により決まる）は $e_4$ と $e_8$ であり，これらが新しい種となる．そしてステップ2からの処理を繰り返す．

以上の手順により最終的に以下の最良のクラスターが得られる．

$[x_1\leq 1][x_3\leq 1]$ 稀薄度 31 ，複雑度 2

$[x_1=2]$ 稀薄度 22 ，複雑度 1

図 4.2(b) にこの解を図示したものを示す．以上の手順は $k=2$ 以外，すなわち任意の数のクラスターを求める場合も同様である．また $k$を順次変更しながら最適な分割数を見い出しクラスター階層を導出する手法も提案されている．

以上の CLUSTER/2 では概念の内包に相当する複合体の論理記述は，その外延集合にあたる全ての事例において真でなければならない．これに対して，全ての事例について真となるようなクラス記述を見いだすのではなく，そのクラスの成員間での属性値の

[2]このようなクラスターは互いに素なクラスターと呼ばれる．

[3]稀薄度は複合体の外延集合の数から観測事象の数を差し引いた数，複雑度は複合体を構成しているセレクタの数として算出される．

確率的な分布を考慮しながらクラス生成を試みる手法も存在する．このような概念クラスタリングの手法としては，WITT[17] や AUTOCLASS[7] のシステムがある．WITT は，k-means アルゴリズムにおけるクラスター生成戦略に類した手法でクラスターを生成し，前節で述べた Rosch and Mervis の族類似の知見に倣ってクラスター内での属性間の相関を高く，クラスタ間での属性の相関を小さくするように行われる．カテゴリー内ならびにカテゴリー間での構造の質のトレードオフを最適化するもので，予めクラスター数を指定するのではなく，この質を最大化するように数が決定される．また AUTOCLASS は，クラスター分割をクラスター内での属性値の分布ならびにクラスターそのものの生起分布に基づいて観察のサンプリングを行ったときに，ベイズの定理に従って未知の属性が最も確からしく予測できるように分割を行うものである．このためには各クラスター内での属性値の分布がなるべくバラツキが小さくなるように生成する必要があるがこれだけの基準では小さなサイズのクラスターが幾つも生成される結果となる．これに対し全オブジェクトにわたる属性値分布に適合するようなバイアスを付加することによって幾つものクラスターが生成されるのを抑え，これにより適切な数のクラスター生成を実現している．このような確率的な概念表現は CLUSTER/2 にはない柔軟な概念表現ということができる．

# 第５章　矛盾事例の同化のための概念の動的組織化手法

## 5.1　概念形成の手法

前章で述べた概念クラスタリングが階層構造をもたない平板な概念クラスの集まりを生成するのが目的であるのに対して，概念形成の手法は一般に以下のように特徴づけられる．

- - Given: 学習対象の事例記述が系列的にストリームとして与えられる
- - Find:
  - 上記の事例をカテゴリー化したクラスターを発見する
  - カテゴリーの帰属する事例群を要約する内包記述
  - カテゴリーの階層化構造

概念形成では前節で述べた概念クラスタリング同様，観察（事象や物）の多次元上での類似性から環境内に潜在している規則性に基づいてカテゴリーを発見することが目的である．教師なし学習であることからどの事例をどのクラスに帰属させるかのみならず，そこで必要になるクラスの数をも自律的に決められなければならない．一旦このようなカテゴリーが発見できれば，前節の自然カテゴリーの基底レベルに関する考察で述べたように，同一環境中の未知事例のもつ多くの特徴の予測が可能になり，過去の経験に基づいてより効率的な問題解決を可能にすることに繋がる．学習された概念の利用という観点からは，概念形成では，多くの属性にわたって平均的にその予測精度を上げるような概念階層を導出することを目指すものであり，従って概念階層の構築は，このような予測を少しでも多くの属性に関して正しく推論することができるように事例を組織化することに相当する．許容される事例記述としては，原則的には，属性—値の集合からなる記述であるが，これ以外にもより複雑な関係記述や構造を有した記述として与えることも許される．

概念形成においては学習のフェーズと学習された概念を推論に用いて新規事例の分類を実行するフェーズは不可分の関係にある．すなわち，既存の概念階層中に新たな入力事例を逐次的（incremental）に組み込む際には，まず既存の概念階層を用いて入力事例の分類が試みられ，最も類似した既存の事例や概念クラスが参照事例（クラス）として同定されると局所的にその構造に変更が加えられ，更新された後は基本的にその事例は忘れ去られる．概念形成の各種のアルゴリズムでは，このような構造への局所的な変更を与えるいくつかの操作子（オペレータ）が定義されており，どのような方略でこれらのオペレータを作用させ，最終的に少しでも多くの属性についてその値を正しく推論できるという意味において最適な概念階層を探索することができるかが問題となる．概念形成ではこのような探索は，基本的に可能な概念階層によって張られた空間内の漸次的山登り法（incremental hill-climbing method）として進行する．過渡的な学習概念自体，非常に複雑な構造を有するものであることから，横型探索やビーム探索で行なわれるような複数の学習概念の候補を維持しながら可能な組み合わせ考え入れた仮説空間を想定して，その中を必要に応じてバックトラックするような探索は困難であり，メモリが追い付かない．勿論このことは局所的な最適解に陥る危険性をはらむ．通常の山登り法では評価関数が固定されており，この評価関数の急勾配の方向へと解をシフトしていくやり方がとられる．しかし概念形成ではこの評価関数自身固定できるものではなく，学習対象としての事例が与えられる度に概念階層自身が変わるだけでなく，この概念階層によって張られる空間自体も変化してその評価関数が変更を受けることになる．従って最終的に得られる概念構造は，事例の与えられる順序に大きく影響を受ける．これを回避するために，概念形成では双方向（bidirectional）の概念構造の修正操作が必要になる．双方向とは，新たなカテゴリーを生成する操作に対して既存カテゴリーを削除していく操作や，カテゴリーを分化させるとともに統合する操作を考えて，過去の学習の効果を必要に応じて打ち消すための操作である．これらによって，例え記憶されているのがその都度変化を受ける一つの状態だけであるにもかかわらず，可能な状態を記憶している下でバックトラックすることと類似の効果が得られることになる．

## 5.2 概念形成に関する初期の研究

### 5.2.0.1 Feigenbaum の EPAM

初期の人間の記憶モデルとして提案されたものに Feigenbaum の EPAM と呼ばれるシステムがある [10]. EPAM では事例は原則として属性—値の連言記述で与えられ，認知心理学で弁別木あるいは識別木（discrimination tree）とよばれる構造を概念階層として逐次的に学習するシステムである.

- 概念階層の特徴　EPAM で構成される概念階層の非終端ノードでは概念記述が付与されているわけではなく，属性のテストが行なわれる分岐点を表す．ここから出ている枝により，テストの対象となる属性の特定の値ならびにその値以外（other）に分岐する．ここでの具体的な値の枝への割付は，具体的な事例が入力されて処理される度にその事例が有する属性値を枝に割り付けていくことによってなされる．終端ノードは，個々の事例が充てられる場合と，「イメージ（image)」と呼ばれる具体的な事例を伴わない概念記述（＝内包）が充てられる場合がある．後者の「イメージ」とはルートからの一連のテストを経て調べられた諸属性が経由してきた各枝に対応する値をとるとして記述された属性—値の連言記述であるが，いまだテストされていない属性については値は特定されないままとなっている．従って「イメージ」の表す外延集合としては終端ノードであるにもかかわらず２つ以上の事例の集合が対応する.
- 概念構造の更新操作　EPAM では入力された事例がそのまま概念階層中のいずれかのノードが表す概念クラスに明示的に帰属させられるのではなく，識別木としての概念構造を更新する役目しか果たさない．この際の更新の形態としては，

  1. 識別木の横方向への展開（テストによる値の候補値の拡張）
  2. 識別木の縦方向への展開（新たな属性テストの追加による識別木の深化）
  3. 終端ノードのイメージの内包記述の付加（イメージの具体化）

  の３種類を考えており，このための概念構造の操作子（オペレータ）として，

  1. familiarization: 入力事例の記述が，ルートノードから現在の識別木を辿って終端ノードのイメージに合致した場合に適用される．この場合，入力事例の

属性記述のうち，終端ノードのイメージの記述として未だ含まれていない属性—値が新たにイメージの記述として付加される（終端ノードのイメージの内包記述の付加）．

2. discrimination: 入力事例の記述が，ルートノードから現在の識別木を辿って終端ノードまで辿りついた際に，そのイメージ記述と矛盾する場合，既にテストされている属性の中で，入力事例—イメージ間で食い違っている属性のテストノードに対して，新たな属性の値（＝入力事例を記述している当該属性の値）の選択肢を生成する（識別木の横方向への展開）．もしも既にテストされている属性の中に上述のような属性が見いだせない場合には，終端のイメージと入力事例を識別するために必要な新たな属性を選択し，これをテストノードとする非終端ノードを生成してその下に以下の２つの終端ノードを生成する（識別木の縦方向への展開）．一つのノードにはここへ辿り着くまでに行なわれたテスト結果による属性—値の集合がイメージとして定義され，もう一方のノードのイメージには，修正前のイメージ記述に新たに追加したテスト属性の値を付加したものとなる．

#### 5.2.0.2 Lebowitz の UNIMEM

Lebowitz により開発された UNIMEM (UNIversak MEMory) はその名の通り，一般化記憶モデル（generalization-based memory）として本来自然言語処理におけるエピソード理解のための記憶モデルとして開発されたもので，複数の事例間で共有される特徴に基づいて組織化することにより，それら集合の背後に潜在するその世界での有意義な規則性（regularity）が抽出でき，この規則性から未観測の属性値に関する推論ができる [22],[23].

- 概念階層の特徴　EPAM との違いのみをまとめると以下の通りである．

  1. 終端ノードのみならず概念階層の各ノードに対して概念記述が与えられているとともに，その外延集合が明示的に格納されている[1]．

[1]ここでの概念記述は完全な内包に相当するものではなく上位ノードの記述に付加されるものについてのみ与えられ，それ以外の属性については上位概念クラスの記述内容が継承される．この意味で Epam のような識別木としてみることができる．

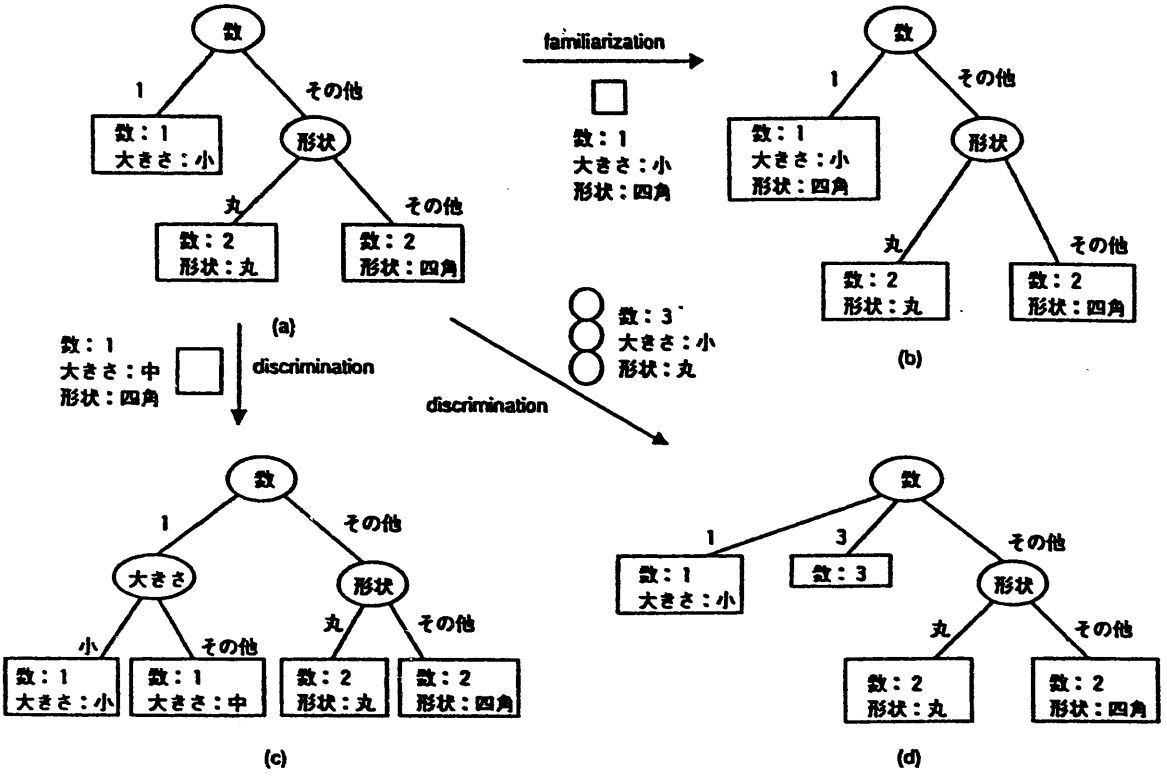



Fig. 5.1 Feigenbaum の EPAM

2. 概念木の各分岐が意味するのは，EPAM では単独の属性についてのテストであるのに対し，UNIMEM では複数の属性に関するテストが対応する．各枝の各々の属性に対しては，その値を知ることによって下位サブクラスをユニークに示唆できる度合い（predictiveness）が定量化されて付与されている．
3. 各ノードの概念記述は平板な属性—値の連言集合であるが，各属性—値の組には確信値（confidence value）の整数値が付与されており，ある事例がその概念クラスに帰属することを知らされた場合に各属性の値をユニークに推論できる度合い（predictability score）が定量化されている．
4. 各事例は複数のノードに帰属し得る．

- 概念構造の更新操作　UNIMEM における概念構造の操作子をまとめると以下のようになる．

1. 既存のノードクラスに基づく新事例の分類：UNIMEM では既構築の概念構造に対して新事例が与えられると，ルートノードから順次各ノードの記述に対

して類似性評価を行ない，十分な類似性が見いだせればそこから分岐している枝を通して該当するサブクラスへと処理を進める．

2. 複数の事例間で共有される属性—値の一般化による新たなノードの生成：上記において十分な類似性を見いだせない場合には，その親ノードに帰属されている全ての既存事例を参照し，新事例との間で属性—値を共有しているようなものが存在するか否かをチェックし，もしあればこの共有属性—値をノードの概念記述とする新たなノードを生成し，その下に事例を帰属させる．またそこで関与している外延を構成する事例群を参照して，predictability score と predictiveness の値を更新する[2]．
3. ノードの概念記述（内包）の固定化と削除：上記の手順による一般化は原則として2つの事例上で行なわれるため，時として過剰に特殊な値をもった一般化が生成される．これを回避するために，UNIMEM では，概念記述内の属性—値と整合する同一値が後のこのクラスに帰属させられる事例においても繰り返し観察されれば，この属性—値対に対する信頼度を上げ，整合しない値が観察されれば信頼度を下げる．そしてユーザの指定するしきい値以上の高い信頼度が得られた属性—値対はそこで固定され，以後の観察例の影響を受けないようにしている．逆にあるしきい値以下に下がると，この属性—値対を概念記述から削除する．
4. 過剰に一般的なノードの削除：上記の概念記述の削除により各ノードの概念記述として定義された属性—値対があまりに少数になりそこから何ら推論するものが得られなくことが生じる．この場合にはこのノード自身がそこから派生している枝と共に概念構造から削除される．
5. nonpredictive な属性—値対に対する子ノードへのインデックスの削除：ノードから派生している枝に付与されている predictiveness の値が高くなり過ぎると多くの子クラスをインデックスづけることになり，生成された概念を利用する際の推論の実行性能の観点からは，このような多数の照合を許容するでのは決して効率のいいものにはならない．そこでこのような枝を削除し選

[2]Unimem では単純に各属性—値対がいくつの事例で成り立っているかについて整数値のカウンタを用いている．HIERARCH と呼ばれるシステムではこの点を改善するものとして情報理論に基づく尺度で概念形成を誘導する手法を提案している (Nevin 1990)．

択的な属性チェックで事例の分類が行なわれるようにしている．

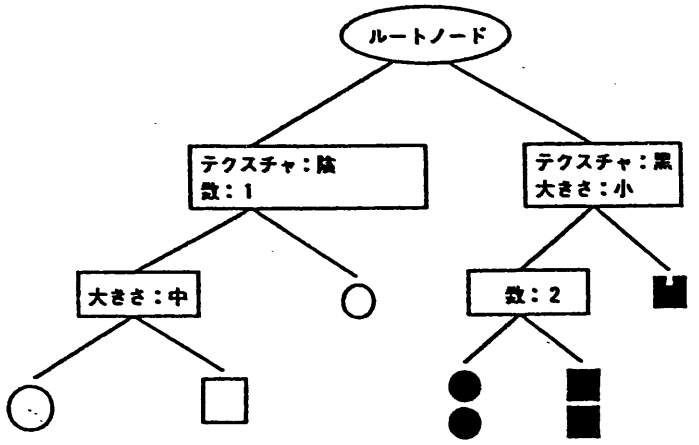



Fig. 5.2 Lebowitz の UNIMEM

#### 5.2.0.3 Fisher's の COBWEB

- 概念階層の特徴　COBWEB は事例や特徴の出現頻度により，事例の概念への包含性を確率的に表現したモデルである [11]．前２つのシステムと異なり構築される概念階層は識別木ではなく，木の枝は $is-a$ もしくは $subset-of$ の関係として定義され，終端ノードには必ず事例が対応する．また階層は相互に重複しない disjoint なクラスに分割する．図 5.3は COBWEB で生成される階層的なクラス概念を例示したものである．ここでは３つの属性（サイズ・色・形状）から記述される４つの積み木の事例群に対して形成された概念クラスを表し，最下位のリーフ部のクラスは単一事例の記述になる．各クラス概念は図に示すように，そのクラスに属する事例集合の中での各属性値の確率分布として定義される [37]．また各クラスにはその親ノードにあたるクラスのメンバ数に対する比率が保持されている．すなわち，クラス分けされた概念 $C_k(k=1,2,\cdots)$ があり，その特徴群 $A_i(i=1,2,\cdots)$ が値として $V_{ij}(j=1,2,\cdots)$ を持っているとする．このとき各クラスには以下の２つの指標が定義される．

  - クラス内類似性（intra-class similarity）$P_s = P(A_i = V_{ij}|C_k)$：あるクラス $C_k$ の事例が特徴 $A_i = V_{ij}$ をもつ条件付き確率であり，「特徴値 $V_{ij}$ がクラス

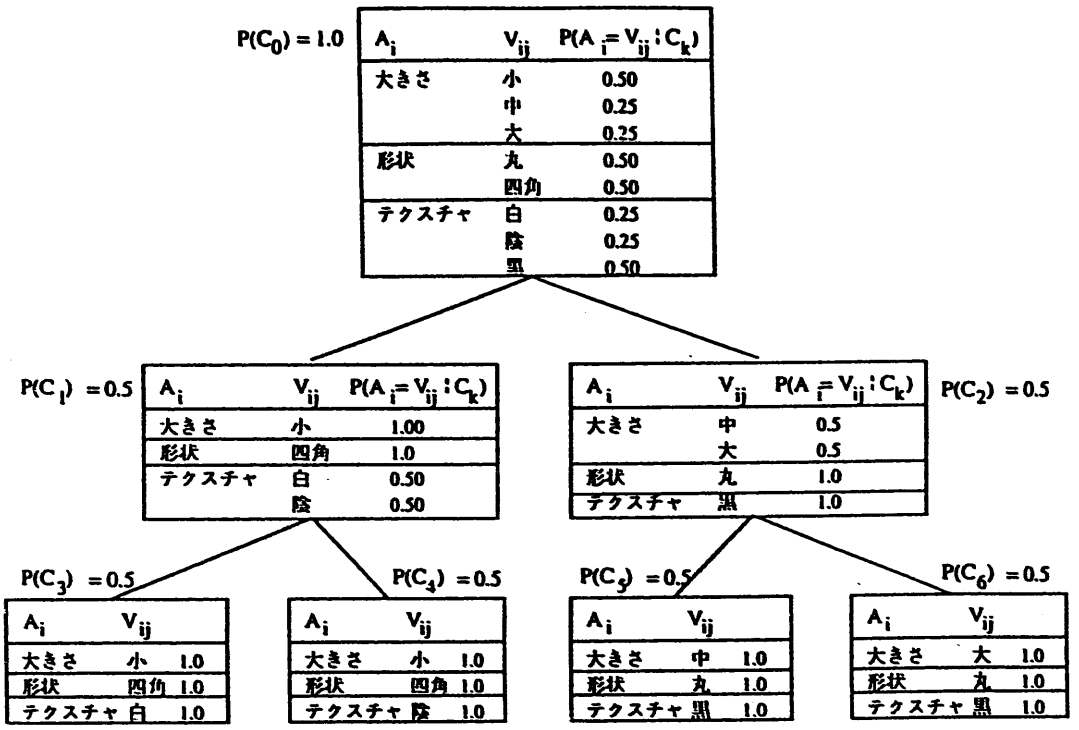



Fig. 5.3 Fisher's COBWEB

$C_k$ の概念を代表するものとしての有効性（category validity）」を表す指標.

- クラス間非類似性（inter-class dissimilarity）$P_d = P(C_k|A_i = V_{ij})$：ある事例において特徴 $A_i = V_{ij}$ が認められたときその事例がクラス $C_k$ に属する条件付き確率であり，「特徴値 $V_{ij}$ がクラス $C_k$ の概念を示唆する手がかりとしての有効性（cue validity）」を表す指標.

- **概念階層の評価指標**　Gluck と Corter は情報の伝達や推論に最も効率の良い，出現する事例に適応した概念クラスを構成するために，クラス内類似性とクラス間非類似性から計算されるカテゴリー有用度（category utility）の指標を考案している [15].

概念クラス $C_k(k = 1, 2, \cdots, n)$ が存在するとき次の式，

$$U_{ki} = \sum_j P(A_i = V_{ij}) P(C_k|A_i = V_{ij}) P(A_i = V_{ij}|C_k) \qquad (5.1)$$

は，属性 $A_i$ に注目し，クラス内類似性 $P_s$ およびクラス間非類似性 $P_d$ の積に，$A_i$ がとり得る個々の属性値の重要性の重み $P(A_i = V_{ij})$ を掛けたものである．す

なわち，よく現れる特徴（$A_i = V_{ij}$）ほど，$U_{ki}$ を強めることになる．この式(5.1)はBayesの法則，

$$P(A_i = V_{ij})P(C_k|A_i = V_{ij}) = P(C_k)P(A_i = V_{ij}|C_k) \tag{5.2}$$

によって，

$$U_{ki} = P(C_k)\sum_j P(A_i = V_{ij}|C_k)^2 \tag{5.3}$$

となる．式(5.3)において $P(A_i = V_{ij}|C_k)$ は，概念クラス $C_k$ の属性 $A_i$ の値を $V_{ij}$ と予測する確率であり，その予測の正しい確率が，やはり $P(A_i = V_{ij}|C_k)$ で与えられるとすると，$P(A_i = V_{ij}|C_k)^2$ は概念クラス $C_k$ を考える場合に実際に属性値を正しく予測できる属性の数の期待値を表わすことになる．そこでカテゴリーの質の評価指標であるカテゴリー有用度として次の式を定義する．

$$CU = \frac{\sum_{k=1}^{n} P(C_k)[\sum_i \sum_j P(A_i = V_{ij}|C_k)^2 - \sum_i \sum_j P(A_i = V_{ij})^2]}{n} \tag{5.4}$$

$k$は注目している部分構造におけるクラス

ここで

$$\sum_i \sum_j P(A_i = V_{ij}|C_k)^2 - \sum_i \sum_j P(A_i = V_{ij})^2 \tag{5.5}$$

ここで $P(A_i = V_{ij} \mid C_k)^2$ は観測例を既存クラス $C_k$ に置くことによって達成されるクラス内類似性と他の競合クラスからの非類似性を総合した評価尺度で，これからルートレベルでの（カテゴリを全く作らない場合の）分布に対応する項（$P(A_i = V_{ij})^2$）を差し引くことによって，観察例の当該カテゴリークラスへの付加の妥当性を判断する指標と考える．また $n$ は親カテゴリから派生している子カテゴリの総数を表し，$P(C_k)$ の項は観察例をより大きなカテゴリへ引き込むようなバイアスを与える役割を果たす[3]．

カテゴリー有用値は前節で述べた「基底レベルのカテゴリーが推論能力が最も高いレベル」という事実を反映した尺度になっている．例えば，いまあるオブジェクトが「動物」のカテゴリーに属することを知ったとして，そこから正しく推論できる属性の数はそんなには多くない．もちろんもし確かな推論ができるのであれば，そ

[3]このことは同時に観察の与えられる順序によって生成されるクラス構造が異なるものになることの原因にもなる．

の適用範囲は極めて大きい．一方，そのオブジェクトが「コマドリ」のカテゴリーに属することを知ったとしてそこから得られる推論は数多く考えられるが，逆にほんの少数のオブジェクトに対してしか当て嵌めらない．さらにオブジェクトの観察頻度のみならず，新たな事例を既存の概念クラスに割り当てる場合には，徒に多数の競合する概念クラスとの照合を避けるためにも，部分的に観察された特徴からそのオブジェクトの帰属する概念クラスをなるべくユニークに同定できること，すなわち $P(C_k|A_i = V_{ij})$ の大きな特徴から構成される概念クラスが望ましい．例えば，基底レベルでの「鳥」は，それがそのカテゴリーに帰属することを知らされた場合になるべく多くの属性を予測することができ，かつその適用範囲もできるだけ広いものであるし，他の「動物」のサブクラスから識別する際にも最も効果的でバランスのとれたレベルのカテゴリーをさすことになる．このことから式(5.1)の指標は人間の範疇形成時ならびに利用時おける効率性を反映したカテゴリー生成を誘導する指標になっている．

- 概念構造の更新操作　概念形成は教示例として与えられる観察の系列から逐次的に最も的確な既存クラスへの割当を決めながら実行される．このときの割当を決める際に用いられるのが上述のカテゴリー有用値で，付加された新たな事例（＝分類したい事例）をクラス $C_k$ に仮に帰属させ，そのクラス内の属性値分布を更新した上でカテゴリー有用値を算出する．クラス生成の手順は，単一の事例からなるクラスから開始して，これに随時新たな事例が付加される度にカテゴリーの再編を図5.4に示す4つのクラス操作の適用を上式の評価基準に基づいて決定し実行する．

  1. 既成クラスへの付加（incorporating-into-existing-class）　まず新事例を既成クラスの各々に一時的に帰属させ，そのもとでのクラスの各属性値の分布を更新し，そのもとでカテゴリー有用値を算出する．
  2. 新たな単一事例クラスの生成（creating-singleton）　またこれとは別に，新事例のみで新たにクラスを生成する場合のカテゴリー有用値も算出する．既有の事例群に対して矛盾を来すような新事例が観察された際にはこのように自動的に検知されて識別される．
  3. 複数の既成クラスの統合（class-merge）　ホストクラスの候補のうち高いカテゴリー有用値をもつ上位2つのクラスについて，図に示すようなクラス

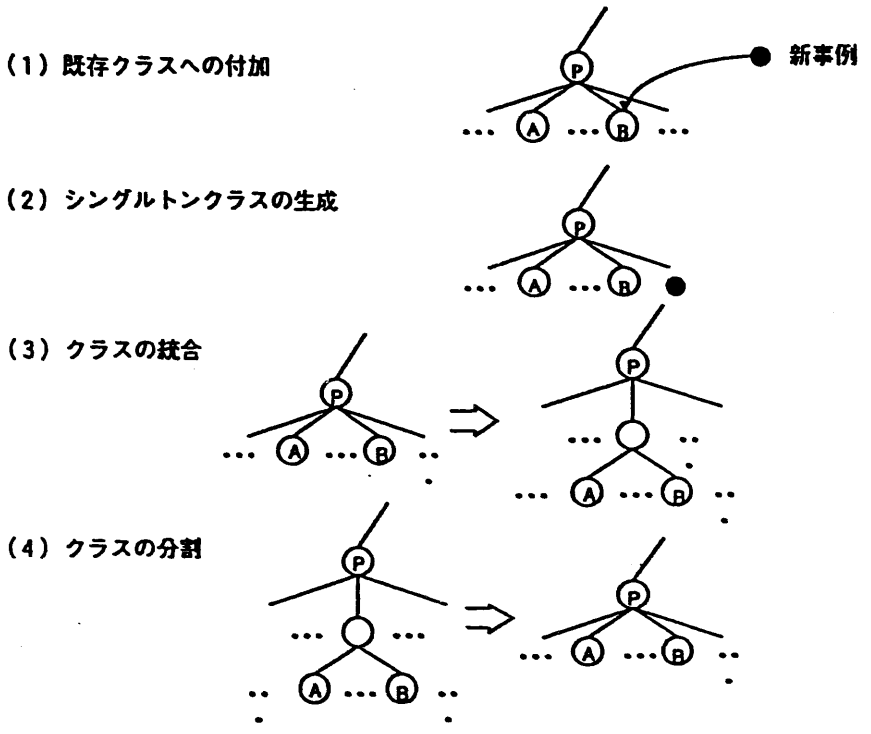



Fig. 5.4 概念構造の更新操作

統合を試み，このもとで試算したカテゴリー有用値がそれ以外の既成クラスのカテゴリー有用値より上回る場合には，この統合されたクラスに新事例を付加する．

4. 既成クラスの解体（class-splitting） 上記の手続きとは逆に，上位クラスを解体してその下位クラス群を直接帰属させる．

まず上記の 1. 2. については算出された値から，最大の値をとる既成クラス（もしくは新クラス）を決定し（＝ホスト・クラスと呼ばれる）新事例の付加を終了する．このとき同時に上記 3. 4. のクラス操作も合わせて考える．これらの両操作は本節冒頭部でも述べたように概念構造を探索する際の逐次的山上り法の局所性の影響を抑えるための 1. 2. の操作に対する双方向オペレータとして用意されている．しかし最適な概念構造の探索という意味では，これらの操作だけでは十分にロバストな結果を生みだす保証はなく，上記 4 つのクラス操作以外の操作子の導入や，互いに類似したものが続かないように多様性を維持しながら供与する方策が考えられているが，十分な解決策にはなりえていない．

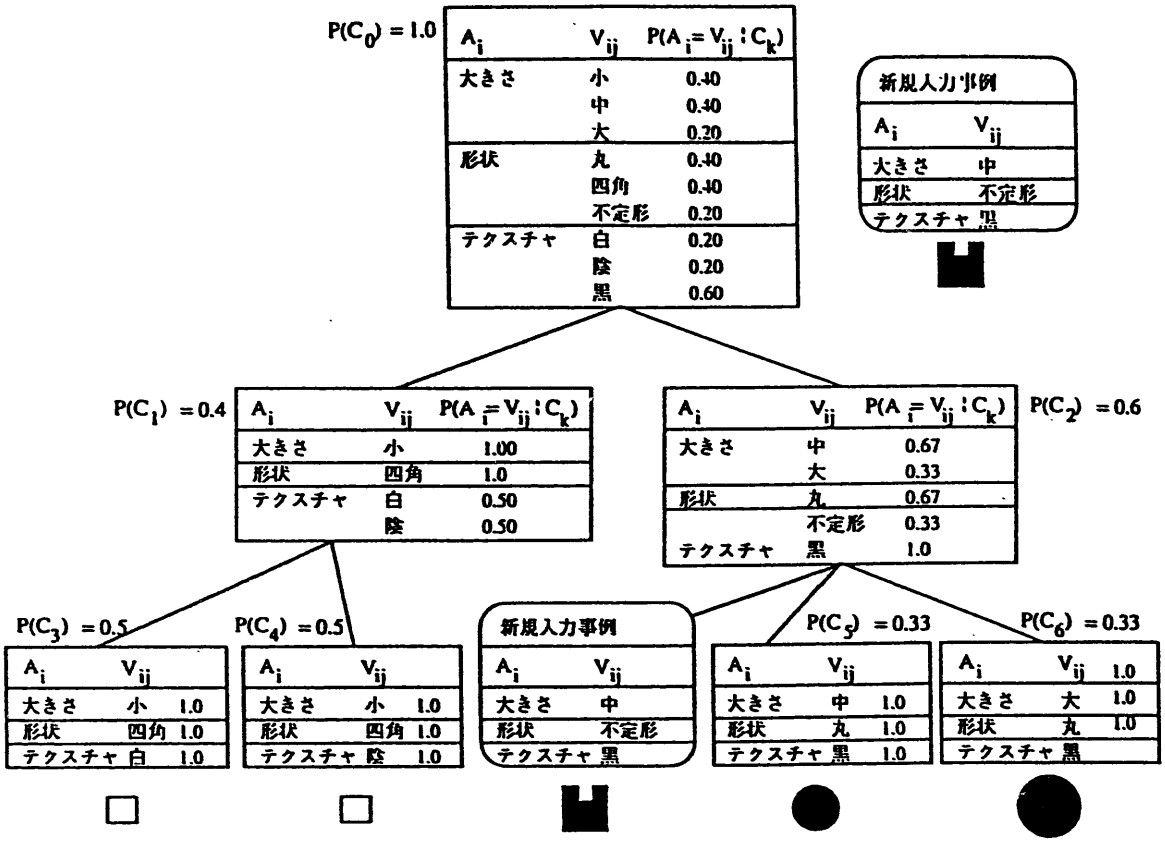



Fig. 5.5 新規事例を組み込んだ後の概念階層

## 5.3 概念形成手法に基づく推論

形成された概念構造に基づく具体的な推論手順は本質的に概念階層を構築する際の学習のフェーズと同じである．ここでは前節末尾に述べたCOBWEBを例にとり以下に説明する．いま新たな観察事象として属性—値の集合で表わされた入力事例を考える．この入力事例は概念形成時に使用された事例のようにすべての属性についてその値がわかっている必要はなく，むしろ不完全な記述でしか与えられない場合が通常である．システムはこの既知属性を使って，概念階層内の各概念クラスに対して上述のカテゴリー有用値を算出し，トップダウンにルートノードから順次階層を降りながらそのリーフに位置する事例の中で最も適合する過去の事例を同定する．そしてこの事例を構成している属性—値集合から，入力事例に欠落していた属性の値を参照し，この値を新規事例に対する推論の結果として出力する．なお終端ノードの事例まで階層を降りずに，その途中の

概念クラスで照合を終了し，このクラス内の記述を参照することによって推論結果を出力することも可能である．この場合，求める属性の値が複数の異なる値の間で分布している際には最も高い生起確率を有する値を求め，これを推論の結果として出力する．

上記の手順による推論は，例え入力事例で欠落している属性がどれであっても，また複数の属性が欠落していても，既に構築されている概念階層を使うことで可能になる．ただし，推論される値の正確さはいずれの属性についての値を推論するかによって偏りが生じ，他属性への依存度の度合いと推論の正確さの間には正の相関が見られる．なおこのような依存度としては，例えば個々の症例を事例と考え複数の兆候群とその原因となっている病名を属性と考えると，明らかに病名属性のもつ他属性への依存度は，ある兆候属性がもつ他属性への依存度よりも高いものになる．これは観測されているデータ群が一見平板な互いに関係を有しない属性の集合に見えても，観察される兆候の大部分が病名との間で因果関係で結ばれているような意味構造を背後に有していることによる．一方個々の兆候属性は他属性のあるものとは因果性を有するものの，全ての属性との間で因果的に結ばれるわけではないことから，前者ほどに高い依存度をもつには至らない．

## 5.4 概念形成手法の拡張

### 5.4.1 数値（連続値）属性の取扱い

実際問題への適用を計るためには，前節例題で述べたような離散値属性のみならず数値量を扱う必要がある．機械学習でこのような数値量を扱う方法としては，まず連続量を離散量に変換して扱う方法，いま一つの方法は数値属性の値の分布を正規分布に従うと仮定して分割する際の確率計算を正規分布の確率密度関数で代用する考え方で，COBWEB を拡張した CLASSIT と呼ばれるシステムで採用されている [12]．すなわち上式の $P(A_i = V_{ij} \mid C_k)^2, P(A_i = V_{ij})^2$ は各々 $1/\sigma_{ik}, 1/\sigma_{ip}$ に置き換えたものが用いられる．ここで $\sigma_{ik}$ はクラス $C_k$ における属性 $A_i$ の値の標準偏差であり，$\sigma_{ip}$ は親クラスでの標準偏差を表す．なお各属性についてその値の分布は正規分布

$$p_i(x) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left(\frac{-(x-\bar{V}_i)^2}{2\sigma_i^2}\right) \tag{5.6}$$

に従うことを前提とする．すなわち，

$$\sum_j P(A_i = V_{ij})^2 \Leftrightarrow \int \frac{1}{\sigma_i^2 2\pi} \exp\left(\frac{-(x-\bar{V}_i)}{\sigma_i}\right)^2 dx = \frac{1}{\sigma_i}\frac{1}{2\sqrt{\pi}} \tag{5.7}$$

であることから，定数である$\frac{1}{2\sqrt{\pi}}$を省略し，これを式に代入したものがカテゴリ有用値として用いられる．ここで$\bar{V}_i$は第$i$属性のとる値の平均値である．概念形成はそのクラス内の属性の値の分布の標準偏差が最小化されるようにクラス分割が進行する．

また属性値の分布が正規分布に従わない場合には以下の式

$$(\int_{-\frac{d_i}{2}\sigma_i}^{\frac{d_i}{2}\sigma_i} p_i(x)dx)^2 \tag{5.8}$$

において$p_i(x)$を属性毎に定義して算出された値を用いる[4]．

### 5.4.2 構造化された値を持つ事例からの概念形成

概念形成への入力となる事例表現については，平板な属性—値の対からなる集合としてではなく，各属性の値がさらに別の属性—値対で表現されるような構造的な値をとる場合も多い．LABYRINTH はこのような複雑な事例記述を入力とする場合の概念形成システムである [39]．LABYRINTH では，各事例は複数の要素の$part-of$関係で記述される．この要素としてはその値が観察から得られるような属性で記述された原始オブジェクトと，その値に他のオブジェクトが入るような構造化オブジェクトの２種類がある．たとえばカップを記述する場合には，その本体部と取っ手部がそれぞれ原始オブジェクトということになる．

まず事例が与えられるとこれを$part-of$関係に従って分解し，原始／構造化オブジェクトそれぞれの部分について COBWEB に従って個別に概念形成を実行する．まず原始オブジェクトについての概念階層を参照し，それが既存の概念中のいずれの概念クラスに帰属させられるかを判定する．ここで返される概念クラスが今度は構造化オブジェクトを記述する属性の値として採用され，このもとで構造化オブジェクトとしての概念形成が進行する．LABYRINTH ではこの構造化オブジェクトの属性値を決める際に，値の

[4]積分区間を常識のように設定するのは平均値付近で凝集されたクラスを意図的に生成し，親ノードでの値の分布状態に対して子ノードクラスでの分布状態の違いをより鮮明にコントラストをつける目的で導入される．$d_i$はこのために標準偏差の区間に対してどの位の比率をもった区間を積分区間とするかを決めるユーザの指定するパラメータである．

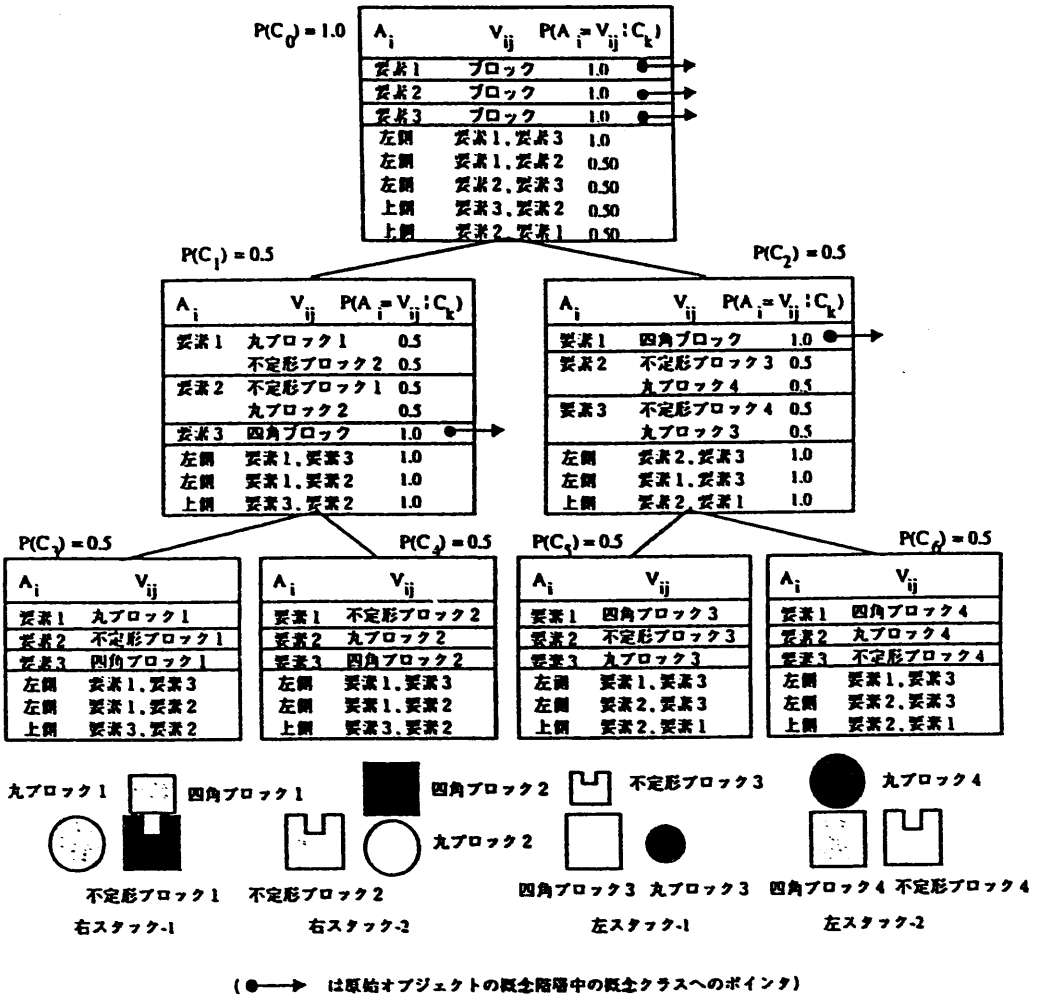



Fig. 5 6 構造化された値を持つ事例からの概念形成

一般化（climbing the generalaization tree）に類する操作を原始オブジェクトの概念木上で試みることにより構造化オブジェクトに関する予測性能を上げる工夫を行なっている．LABYRINTH 以外にも，例えば RESEARCHER, MERGE のシステムでは構造記述された複雑なデバイス上での概念形成を試みており，Levinson はグラフ構造表現された事例からのクラスタリング手法を提案し，遺伝子化学の分野に応用している [24]．また Segen は情報理論尺度（MDL）を用いたグラフのクラスタリング手法を提案している [36]．

### 5.4.3 不完全データからの概念形成

他の学習手法同様，データにノイズが含まれる場合，すなわち属性値に誤りが混入している場合には概念形成による推論精度は大きく影響を受ける．これは誤った事例を

オーバーフィッティング（over-fitting）してしまうことが原因で，とくに終端ノードで推論を出力する場合には誤った推論結果となる．そこで事例を終端ノードまで分類せずに，その途上のクラスにおいて推論を出力するのが得策となるが，この際どのレベルまでで概念形成を終了するか，いわば概念階層の枝刈り（プルーニング）が必要になる．このような枝刈りの機構としては以下のような方法が提案されている．

1. $\chi$-二乗検定を付加した概念形成　概念階層を拡張していく際に統計的な検定をかけ，その結果によってあるレベル以下への拡張を断念（プルーニング）することが提案されている．すなわちあるクラスでのある属性の値の分布が，それより下位におけるクラスでのその属性の値の分布に比べて有意な差（これはユーザの指定する confidence threshold により設定される）がない場合には，上位のクラスのレベルで最も頻度の高い値を予測値として出力する[5]．D.H. Fisher はこのような confidence threshold を連続的に変化させて実験を行なっている．confidence threshold が 0% というのは，わずかの違いでも有意な差として認識することを表し，従来の COBWEB のアルゴリズムと同じである．ノイズを含まない場合には，トレーニング初期において，confidence threshold は小さくても良好な予測精度を示すが，逆にこの段階では検定を導入することによって逆に予測精度に悪影響が見られる．そして十分な訓練を積んだ段階では confidence threshold をどのように変更しても大差ない結果となる．これに対しノイズが増えるに従って，最も高い予測精度を出すような confidence threshold は徐々に高くなっていく．すなわちプルーニングの有効性がより顕著になる．Fisher and Schlimmer は，$\chi$-二乗検定によってもたらされる予測精度の改善の度合いは，予測させる属性の他属性への依存度の大きさによっても異なり，依存度の高い属性について予測させる場合より，依存度の小さい属性を予測させる場合の方がそのメリットは大きい結果を示している．
2. 過去の分類結果を加味した概念形成　概念形成の過程において各クラス（ノード）には，各々の属性について過去の分類でその値を正しく予測できたかどうかについての情報が記憶されている．またそのクラスの直下のサブクラスにおいて正しい予測ができたかどうかについての情報も合わせて記憶されており，ある属性の値を推

[5]このようなプルーニングポイントは属性によって異なる．

論する必要が生じた際には，各クラスに記憶されているこの属性に関する過去の予測の正しさについてのカウンタを参照し，それ以上下へ降りても予測精度がもはや上昇しないようなレベルのクラスを見いだしてそのクラスに格納されている属性の値分布の中で最も高い確率をもつ値をその予測属性値として推論する．この手法は予めノイズのレベルや訓練量によってconfidence thresholdを調整する必要がなく，過去の分類結果を蓄積するだけで同等の効果が得られる．この考え方を拡張したものとして，規範値（normative value）を用いて予測をどの時点で行なうかを動的に制御する手法が提案されている．このような規範値を用いた分類の制御はそれ以上分類を進めても無意味な時点を決定するカットオフ戦略となっている．

### 5.4.4 文脈の組織化

1. 探索の制御知識の獲得　探索過程において成功に導いたオペレータがどのような文脈で使われたかについての知識（探索の制御知識）の獲得を目的として，類似文脈をクラスタリングしてそこから有効なオペレータ系列を見いだすための概念形成の利用が考えられる．B. Carlson, J. Weinberg and D. Fisherは，プッシュダウンオートマトンを用いた構文解析を例に，構文解析時の分岐点でいずれの構文を想定して解析するべきかについての制御知識を，その解析時の文脈の類似性に基づいて獲得することを試みている[3].
2. 説明構造からの概念形成　J.P. Yoo and D. Fisherは機械学習の分野で演繹学習の手法として知られる説明に基づく学習（explanation-based learning）において顕在化するユーティリティ問題（utility problem）[6]を回避するための試みとして，過去の説明生成に用いられた部分構造を有効に再利用して行くために説明構造からの概念形成を試みており，探索の効率が格段に上昇する結果を示している[43].
3. プランニング知識の概念形成　H. Yang and D. Fisherは，古典的なプランニングシステムであるSTRIPSで用いられるオペレータについて概念形成を試みている[42]．各オペレータは前提条件，付加リスト，削除リストから記述され，これらのリストを埋める述語を属性とする事例として記述することにより，オペレータの適

[6]領域理論（domain theory）と呼ばれる知識の探索を効率化するために獲得した知識が，その副作用として無駄なバックトラックの機会を生み出すなどして，むしろその後の問題解決の効率を改悪する結果に陥ってしまう問題.

用文脈をその類似性に従ってまとめ上げ，類似の状況で類似のオペレータが利用可能なように組織化するための方法論を提示している．また同様の試みがDAEDALUSのシステムにおいても試みられている[21]．

# 第６章　プラントのトレンドデータからの異常事象の概念学習

## 6.1　はじめに

本研究では前章で述べた概念形成（concept formation）の手法をプラントで観測される多数のプロセス変数の時系列データに対して適用し，プラント状態の概念カテゴリの自動生成を試みる．ここで入力となる教師データは，異常の発生に伴い時間とともに推移するプラント状態量の時系列データであるが，本研究ではこれらの教師データから，ある時区間における類似の複数の状態を包含する一般記述としての状態（ステート）概念，ならびにこれら状態概念の間の遷移パターンの帰納的な学習を行い，プラントの異常挙動に対応する概念カテゴリを，多面的・多階層（抽象度のマルチレベル）の構造を有した概念として生成し，これに基づいてオペレータの熟練度や経験，タスク環境（緊急性の違い）に依存した情報の能動的・選択的提示を可能にするための方法論の確立を目指すものである [33],[38].

## 6.2　自動化システムの概念形成・分類機能

分類（classification）は人間の情報処理の中でも最も基本的な認知活動の一つである [5]. 実際第一世代エキスパートシステムのうち代表的なシステムが行っている問題解決は「分類」のタスクであるのみならず，ドメインの違いを越え「分類」という汎用的なタスクレベルでの知識システム構築の必要性は Chandrasekaran の類型タスク（generic task）[4] や Clancey のヒューリスティック・クラシフィケーション（heuristic classification）[6] の概念として唱えられている.

一方このような分類概念を自動生成する試みとしては，統計的手法では主成分分析あるいは因子分析，やクラスタ分析，数量化 III 類法などが，また機械学習の帰納学習では識別木（discrimination tree）の自動生成に関する研究が活発に進められている．しかしこれらの手法が上述のようなインタフェースの満たすべき要件に叶ったものである

かとなると問題を残す．まず前者の統計的手法の多くは，事例（サンプル）の数としては多数を扱えるという利点があるものの，事例を記述する空間の次元があまりに大きくなると，その結果の解釈が極めて困難なものとなる．また分析の結果への意味づけや解釈といった作業は解析者のノウハウに任せられることから，オンライン・リアルタイムに分類から認識，決定行動へと結び付ける機構を実現することができない．一方後者の識別木の考え方はシャノンの情報量を指標としてカテゴリ同定を効率的に行うための属性の調べる順序が木に構成され，分類対象の事例のもつ属性値に基づいて，各々の分岐を逐次的に辿っていくことによってその分類結果が得られる．しかしここでの問題もやはり事例を構成する次元 $n$ が大きくなると，識別木を辿るのに要する時間と調べなければならない空間は，各々$O(n), O(q^2)$ のオーダーで拡大する（ここで $q$は一つの属性の値域の要素数，すなわち各分岐での枝別れの数である）ことになり，決して人間にとって親和性の高い知識とは言い難い．

ところで人間の認識の背景にある分類概念の構造はこのように対象が大規模になってもそれに伴って複雑化するものとは考えにくい．なぜなら人間の場合認知処理に費やすことのできる資源にはおのずと上限があるという前提，そして例え認識対象が複雑なものに変わっても，知覚から認識分類，行動決定に要する時間はさほど影響を受けずにとりあえずの行動を起こせるという事実である．このような現実の背景にあると考えられる機構は「階層的に構造化された概念の組織化」である．これは認知科学分野ではスキーマ（図式）[1] と呼ばれている構造体で，全体―部分，抽象―具象の両関係を介した階層的な構造を有し，これによる広狭さまざまな多面的な視点を供与できる．また構成される階層自体の構造は，分類事例（サンプル）数の増加や事例記述の次元数の増加に対して直ちに影響を受けることの少ないロバストなものである（Miller の "magical number 7 ± 5"）．本研究ではこのようなスキーマ構築の手法として概念形成（concept formation）手法を用いる．

## 6.3 対象プラントと時系列データの仕様

### 6.3.1 熱量調整プラント

本研究でプラント異常概念構築対象として取り上げるプラントは，都市ガス製造プラントの一部である熱量調整プラントである．図6.1に都市ガスプラントの全体図を示す．図中，ANG，BNG，それぞれアラスカ，ブルネイ産の天然ガスであり，SNGはナフサとLPGを原料として作り出された天然ガスとほぼ同質のガスである．この熱量調整プラントでは，それぞれ異なった熱量を持つ産出国別の天然ガス(ANG,BNG,SNG)及びLPGを混合することにより，所定の熱量・燃焼性に調整し，さらに圧力調整され，一度球形ホルダに(図中ball)保存された後，付臭を施し(図中smell)，都市ガスとして消費者へと供給するプラントである．

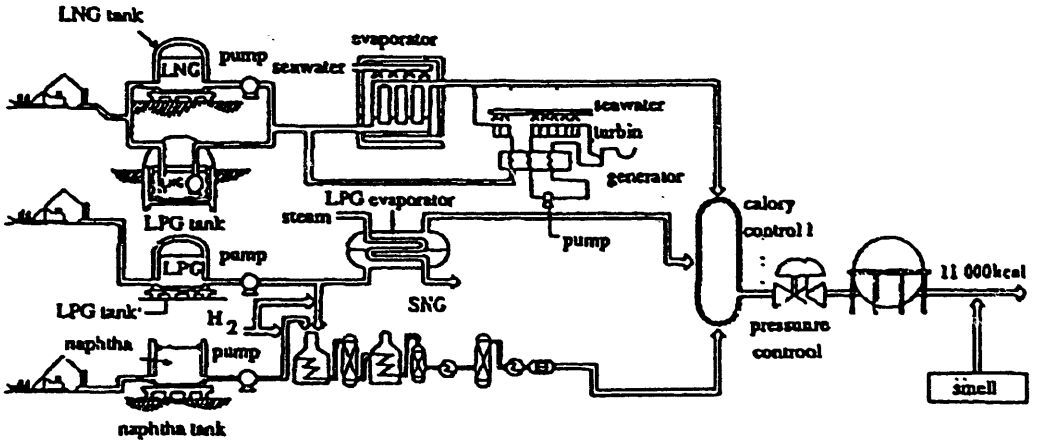



Fig. 6.1 都市ガスプラントの全体像

このプラントには，配管内におけるガスの圧力・流量・温度・熱量及びバルブの状態を検出する各種計装が適所に備え付けられており，各計装により検出されるデータは時々刻々と集中管理センターへ伝達される．この集中管理センターへの入力データに異常が見受けられない平常時においては，プラントは自動化システムにより自動的に運転されているが，設備の系統切替作業などの非定常作業，異常監視及び観測されない対処等の作業においては常時複数駐在しているオペレータの役割となる．

### 6.3.2 プラント計装データ仕様

本研究で用いているデータは熱量調整プラントで起こり得る異常を想定して作られてたシミュレーション・データであり，プラントのオペレータの訓練・育成に用いられているものである．以下にデータの仕様を挙げる．

プラント異常データ総数　異常事例104種，115パターン．

異常イベント例　配管詰まり，配管破損，配管漏洩，等．

計装の数　各故障データにつき49種．圧力計，流量計，温度計，バルブ開閉計，等．

計装データ仕様　各計装データは，サンプリング間隔2秒で128回のサンプリング，計256秒のデータ長をもつ時系列データ．また計測値は各々計装データ毎に正規化された連続値．

なおこのデータは異常イベントの生起前から生起後までの各センサの振舞いが示されており，異常に対するオペレータの操作等は加えられていない．図6.2に異常ケース「アラスカガス系統シャットダウン」の場合のプラント挙動を例示する[19].

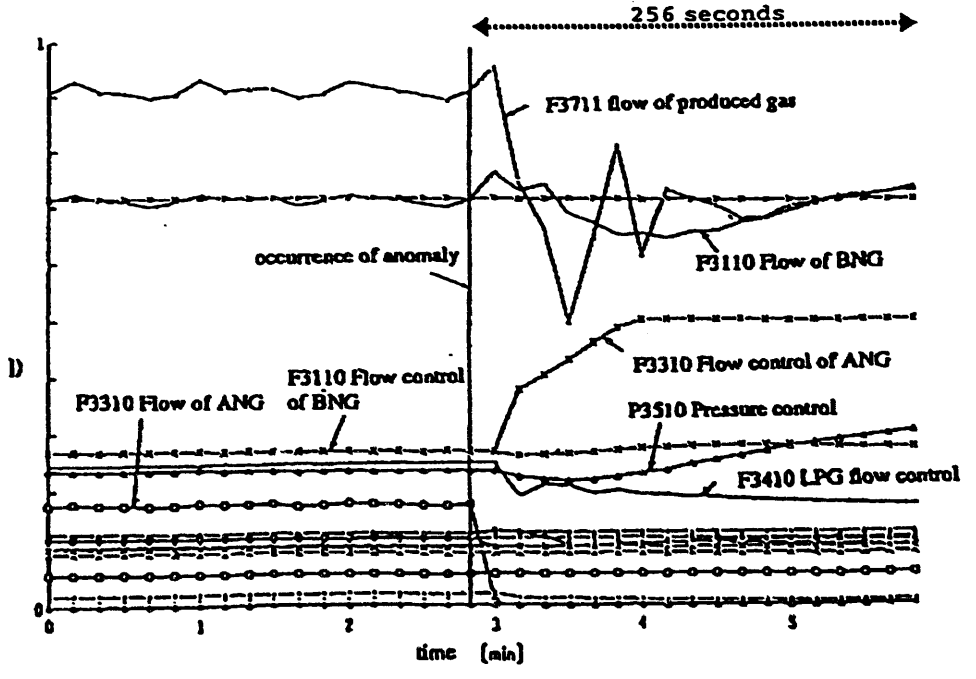



Fig. 6.2 異常「アラスカガス系統シャットダウン」のプラント挙動データ

## 6.4 プラント異常データの静的パターン概念の形成

### 6.4.1 プラント異常データへの COBWEB の適用

経験を積んだオペレータのプラント異常の認識においては，現在変化を示している複数のセンサの全体的な推移を観察して，その全体的なトレンドあるいは推移パターンから直ちに過去の類似異常ケースを想起して，現時点でのプラントの状態の把握ならびに異常原因の特定を行い，同時に今後の推移状態を予測下においた上で特定のセンサデータの出方を監視して異常を確定するという直観的な診断の形態が多用される．本研究ではこのようなタイプの診断を支援するシステムとして，プラントにおいて想定される諸異常について，プラント内部に設置された計装類を通じて送られてくる状態量ベクトルに関して前節で述べた概念形成の手法を用いたプラントの異常状態に対する多階層な概念記述を有するプラント異常概念の構築を試みる．

### 6.4.2 プラント挙動データの前処理と適用結果

本研究では異常発生時における各々の計装データの挙動をデータ長の時間区間内で現われたトレンドに関する形状特徴を抽出して記述し，一つの異常状態に対するプラント挙動（＝事例）を計装の一つ一つを属性として，またそのトレンド（形状）を値として表現する．以下この前処理をコード化と呼ぶことにする．ここでは以下の２種類のコード化を試みる．

1. 変化の有無に着目した２値によるコード化　異常発生時より約２分間において変化が観測された計装データ対し「変化あり (1)」の属性値を与え，全く観測されなかった計装データに対し「変化無し (0)」の属性値を与える．
2. 変化性状を細分類した多値によるコード化　属性値として「定常 (1)」,「単調増加 (2)」,「単調減少 (3)」,「山形 (4)」,「谷形 (5)」,「その他 (6)」の６種類を用意し，これらによりプラント挙動の記述を行う．

異常モデルの構築を試みる異常事例数は 100 事例であり，各々の事例において，49 種の全計装データに対しコード化アルゴリズムによりコード化した属性記述を作成し，これらに COBWEB を適用した結果を以下に示す．なお（）内の数字は各レベルで形成

されたクラス概念数のうち１つの事例が１つのクラスを形成しているクラス（singleton class）の数である．

| 階層の深さ | ２値によるコード化 | 多値によるコード化 |
|---|---|---|
| レベル１ | 2(0) | 4(0) |
| レベル２ | 5(0) | 12(1) |
| レベル３ | 13(2) | 34(16) |
| レベル４ | 30(9) | 52(38) |
| レベル５ | 59(44) | 32(28) |
| レベル６ | 40(35) | 12(10) |
| レベル７ | 10(10) | 4(3) |
| レベル８ | -(-) | 3(2) |
| レベル９ | -(-) | 2(2) |

Table 6.1 ２種類のコード化により各レベルで構成された概念クラス数（singleton classの数）

この結果より，同一事例群に対して，２値／６値の異なるコード化を行っているにもかかわらず，形成された階層概念の構造は属性値の増加に対してロバストな構造が得られていることがわかる．

## 6.5 動的挙動に対する概念形成手法

### 6.5.1 概念構造への時間の取り込み

前節での事例記述とそれに基づく分類では，異常の発生からその影響伝播の終了時点までのトレンドを静的なパターンとして分類していることにほかならない．すなわち事例とその上位概念としての概念クラスの間の概念関係，事例―クラス関係，もしくはサブクラス―クラス関係という，集合概念の包含関係（具象――般関係）のみに基づいて形成されたものである．しかし実際のプラント制御分野においては，このような影響伝播を待って判断を下すという状況は稀であり，むしろ熟練オペレータは，時々刻々と変化するプラント状態に対し，いち早く今後の推移を的確に予測させるような部分観測としての兆候を見いだしてプラントで発生している異常状態を同定し，その分類に基づ

いてその時々の判断を下し行動する状況が普通である．このようなプラントにおいて進行中のプロセスと同時並行でのリアルタイムでの分類を可能にするためには，新たに時間という要因を取り入れた概念形成，すなわちいま一つの概念関係である部分—全体関係を導入し，これに具象——一般関係を絡めた概念の形成手法を開発する必要がある [?]．

本研究ではプラント異常挙動の概念形成にあたり，以下の2種類の概念単位を考える．

- スキーマ：49種類の計装データの時間的な推移パターンとして表現される「プラント挙動」(behavior) に対応する概念クラスで，その外延集合はプラントシミュレータより得られる100の異常ケースを要素とする集合である．スキーマはその抽象度に応じて階層構造を成し「抽象—具象 (IS-A)」関係を介して関係づけられる．
- ステート：経時的に得られるプラント挙動を，ある特徴的なイベントの発生（例えば計装データのあるものが「定常状態から下降や上昇を始める」あるいは「下降や上昇から定常状態に復帰する」といったイベント）を境界として分割することによって得られる部分的時区間での個々の「プラント状態」(state) に対応する概念クラスである．計装データ数に対応する49次元の状態ベクトルとして表現される．スキーマ同様，ステートもその抽象度に応じて階層構造を成し「抽象—具象 (IS-A)」関係を介して関係づけられる．またスキーマとの関係においては「全体—部分 (PART-OF)」関係を構成する．

図6.3はこのステートの階層構造を図示したもので，状態を表すステートのクラスが各々のスキーマのクラスに内包されながら形成される．一つの異常ケースのプラント挙動は，これを構成する幾つかのステート概念の時系列として表現され，各々のステート概念はさらに細部的なステート概念の時系列として再帰的に記述される．図6.3に対応づけるならば，図中太線で示したのがスキーマクラス間の具象——一般関係であり，各クラスにはその水平方向に展開した階層に示されるようなスキーマ—ステート間の部分—全体関係が構築される．なおこの階層はそれぞれのスキーマクラスに固有の構造をとることになる．この図ではステート概念に関する具象——一般関係は明示的には表現されていないが，上位のスキーマクラスを構成するステート概念は，下位のスキーマクラスを構成するステート概念を包含している．前章で述べた概念形成での階層構造と同様に，階層の末端部でのリーフ部分には，概念形成に用いられる個々の事例が対応する．すなわちスキー

マ階層の末端では，プラント異常の各々の具体的な挙動としてのステート系列が対応し，ステート階層の末端では個々のステート記述の一つ一つが対応することになる．

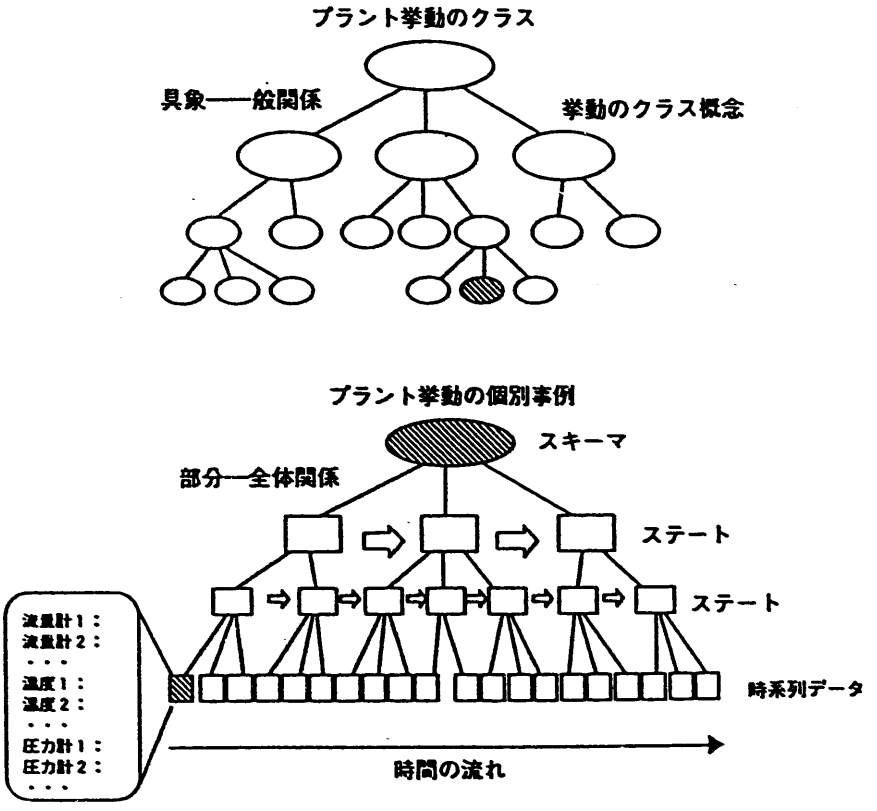



Fig. 6.3 スキーマ階層とステート階層

## 6.6 概念形成手法によるスキーマ階層の構築

### 6.6.1 プラント挙動の事例表現

まずプラント異常データの状態分割を行ない，各々の「状態」における状態記述を得る．ここで状態記述とは各々の状態における各計装の部分的時区間内での振舞いを記述したものであり，複数の状態記述の連鎖系列として1つのプラント異常の事例が表現される．

ここで本研究対象に対し適用した状態分割及び，状態記述の抽出手法について以下に示す (図 6.4).

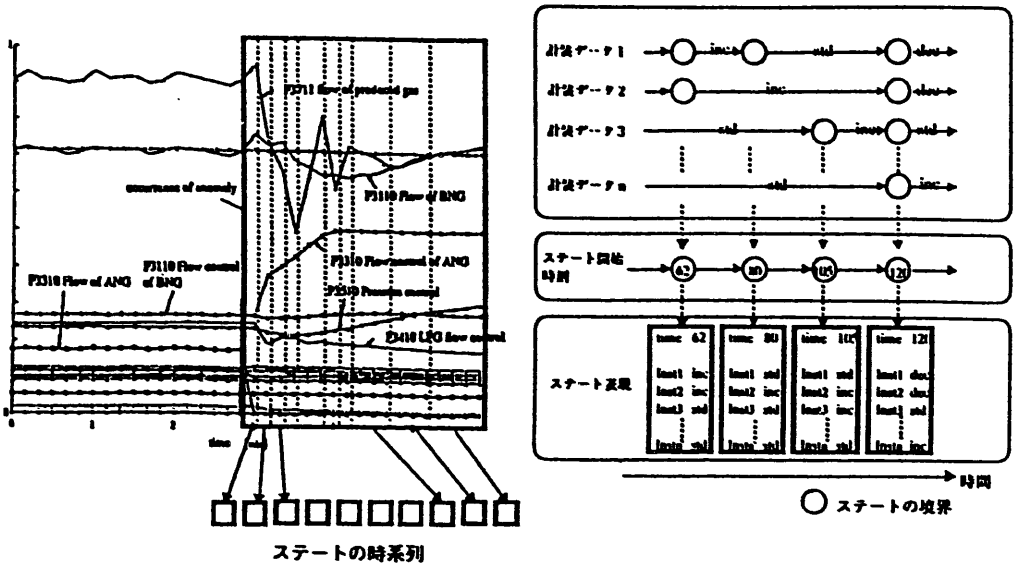



Fig. 6.4 プラント挙動の事例表現

1. 各計装の示すグラフ化されたデータの傾きが大きく変化した時間をイベント生起時間とする．
2. 個々のプラント挙動の事例において，一つ一つの計装データについて上述のようなイベントの生起時点を抽出し，これをすべての計装データについて和集合を取ることによって得られる順序づけされた時点の並びを構成して，これを各々の「状態」の生起時刻とする．また1つの状態生起時刻から次の状態生起時刻までの各計装データの振舞いを属性値としてベクトル表現したものを「ステート」記述とする．本研究では計装の振舞いを「下降」,「定常」,「上昇」の3値により記述する．なお各ステート記述には対応するイベント生起時刻も一つの属性（連続値）として付加されている．

### 6.6.2 静的パターンの概念形成との相違点

1. 構造を有する事例記述からの概念形成　一つ一つの事例であるプラント挙動は，決まった数の[属性—値]対で画一的かつ平坦に記述できるもののでなく，1つの事例が複数のステートの連鎖として表現されており，またこの系列長は事例によって異なる．
2. 構成される概念構造の評価　新規の挙動事例が新たに入力されたときに，既存のス

キーマクラスの一つ一つに仮想的に帰属させその上でカテゴリの評価指標（カテゴリ有用値）を算出して，これを最大にするスキーマクラス（あるいは新規の挙動事例のみからの単独クラス）を決定する．ただしこのとき上記項目1.の理由から，新規事例を構成するステート概念の一つ一つが，既存スキーマのもとに構成されているステート概念のいずれに帰属するかを判定する必要がある．なおこの対応は1対1に決まるものではなく，通常多対1の関係で対応づけられたり，あるいは既存のステート概念に帰属させるよりそれ自身で新たな単独のステートクラスを構成するのが妥当な場合もありうる．

以上の相違点を考慮し，プラント挙動の動的な推移概念を内包したスキーマ階層の構築アルゴリズムとして，スキーマ階層を形成する過程，ステート階層を形成する過程の両者に分け，各々で概念形成手法を再帰的に適用することによって動的挙動の概念形成を行う[38]．なおスキーマクラス，ステートクラスいずれにおいても，既成クラスへの付加，単独クラスの生成，統合，分割のクラス操作は前章で述べた手順と全く同様に行う．

## 6.7 プラント異常概念形成の結果

本実験では熱量調整プラント内部における ANG ライン，BNG ライン，SNG ラインにおける異常事例21事例，計装数15に対し異常モデルの構築を行なった．連続値属性の概念形成では各属性について標準偏差を用いて CU 値を算出する．従ってある属性の値の分布が同一値になるようなクラスが生成される場合 $\sigma_i = 0$ となり CU 値が $\infty$ となる．これを避けるため非負の$\sigma_{min}$をパラメータとして指定する．本実験においては標準偏差$\sigma$の最低値は 1.0，0.5，0.05，0.01，0.005 の5種類の設定で行なった．各設定毎における概念クラスの形成結果について以下にまとめる．

これより$\sigma_{min}$の値により形成される概念クラスの構造が変化することが分かる．すなわち，$\sigma_{min}$が小さい程，わずかの属性値の違いでも検知されて1つのスキーマより生成する「子」スキーマが多くなり，末広がりの傾向の強い概念クラスが形成される．一方$\sigma_{min}$を大きく設定すると，かなり異なる事例同志でもルート近くで同一クラスにとりこまれ，レベル深くに達するまで分化されないような概念クラスが形成されている．

| 階層の深さ | $\sigma_{min}$=1.0 | $\sigma_{min}$=0.5 | $\sigma_{min}$=0.05 | $\sigma_{min}$=0.01 | $\sigma_{min}$=0.001 |
|---|---|---|---|---|---|
| レベル1 | 2 | 3 | 5 | 3 | 3 |
| レベル2 | 5 | 6 | 11 | 9 | 8 |
| レベル3 | 11 | 10 | 14 | 14 | 15 |
| レベル4 | 7 | 13 | 6 | 7 | 10 |
| レベル5 | 5 | 6 | - | - | - |
| レベル6 | 3 | - | - | - | - |

Table 6.2 動的挙動の概念形成アルゴリズムによる結果

## 6.8 形成されたプラント挙動概念に基づくリアルタイム診断

### 6.8.1 プラント異常の同定アルゴリズム

概念形成手法は，事例を逐次的に，かつトップダウン的に，最もその事例に似た事例の集合より構成されるクラスを選択していくという手続きにより，学習（概念形成）のみならず既存の概念クラスに対しある新たな事例がどの概念クラスに最も当てはまるかという「同定」すなわち診断に利用することが可能である．

本研究対象である熱量調整プラント内の一部であるANGラインで想定される異常に対し前章で述べた手法によりプラント異常クラスを形成し，そのプラント異常クラスに対する異常事例の同定実験として，ANGライン関連の異常事例7事例（各事例の記述は8個の計装に限定）について事例の入力順をランダムに変更した5回のプラント異常クラスの形成を行ない，その平均値による評価を行なったところ，88.6%の正解率であった．さらに学習事例群を一度の学習にランダムに3回入力し学習させて構築を行なったプラント異常クラスにおいては，正解率は94.3%に向上した．

### 6.8.2 リアルタイム診断

実際のオペレータによるプラント制御においては，時々刻々とプラントより送られてくるデータに基づいて，すなわち異常発生から現在までに得られている情報のみに基づいて，その時々の異常の同定を行なうのが通常である．

前節で述べた同定アルゴリズムでは異常の伝播の終了までの状態記述が全て揃わな

くとも，部分的な観測情報のみからでも事例のスキーマクラスへの同定を行うことが可能となり，その時点までの観測から最も似ていると評価された事例の含まれるスキーマを提示することが可能となる．

本研究対象である熱量調整プラントにおけるANGラインの異常事例「アラスカ系外ライン詰まりの可能性1 (001-20)」に対し時系列同定を試みた結果を以下に示す．なお本研究では2秒を単位時系列として扱っているため，本論文における以下の記述において，「時間50」であれば実際のプラントでは100秒経過したことを意味する．

この異常事例は6つの状態記述から構成されている．もとのデータを図6.5に，変換された状態ベクトル系列を図6.6に示す． 以下に，この事例における各状態が入力され

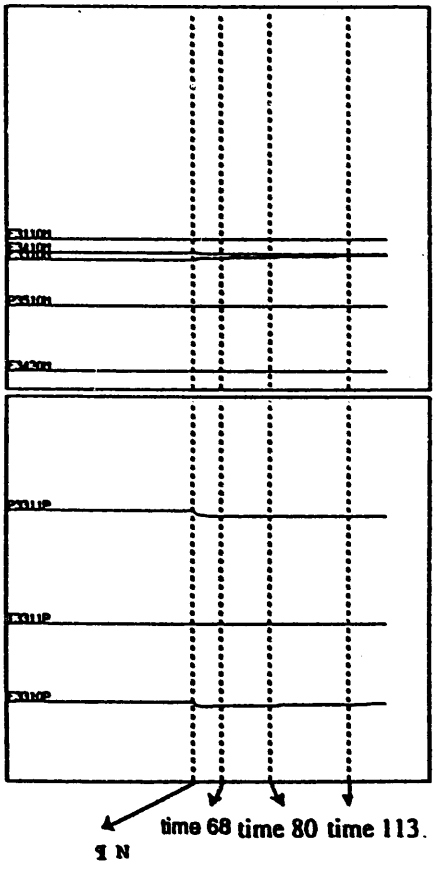



Fig. 6.5 異常の同定を行った時系列データ「アラスカ系ライン詰り1」

た時の時系列同定が行なわれた結果を時間の進行とともに図6.7(a)～(c)に示す．各々の時点での異常同定は図6.5中に記した各々の時点までに観察された状態から同定を試みた

| time: | time 62 (124sec) | time 68 (136sec) | time 80 (160sec) | time 95 (190sec) | time 113 (226sec) | time 125 (250sec) |
|---|---|---|---|---|---|---|
| instrumentation 1: | dec | dec | dec | dec | dec | dec |
| instrumentation 2: | dec | inc | inc | inc | inc | inc |
| instrumentation 3: | std | std | std | std | std | std |
| instrumentation 4: | dec | std | inc | inc | inc | std |
| instrumentation 5: | inc | inc | inc | inc | inc | inc |
| instrumentation 6: | dec | dec | dec | dec | dec | std |
| instrumentation 7: | std | std | std | std | std | std |
| instrumentation 8: | dec | dec | dec | std | inc | std |

Fig. 6.6 異常の同定を行った時系列データ「アラスカ系ライン詰り１」のステート表現

際に，既に形成されているスキーマ階層の各抽象度レベルでもっとも高い *CU* 値を呈したスキーマ概念クラスの外延集合が異常の第一候補 (candidate) として，さらに２番目に高い *CU* 値を呈したスキーマが第２候補として提示されている．なお，各スキーマのスキーマ階層内での位置関係をツリー構造として提示したものを併せて記す．

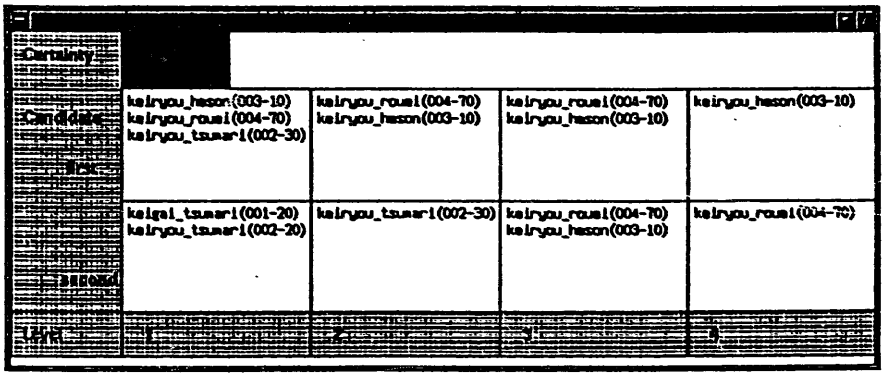
状態１が入力:時間 62　プラントの異常が発生し，現時点での異常診断システムのプラント状態の判断が階層的に示される．

状態２が入力:時間 68　先の状態に現在生起した新たな状態を追加したことによる判断が示されている．現時点では正解に至っていない．

状態３が入力:時間 80　現時点で異常診断システムは正解を得ている．

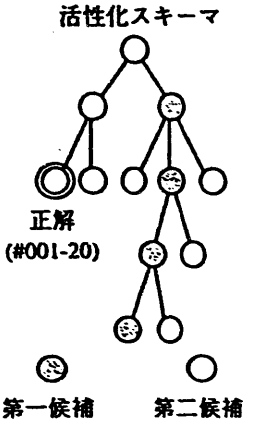
状態４が入力:時間 95～状態６が入力:125　異常診断システムとしては新たな「状態」の検出を行なっているが，状態４で得た結果の繰り返しとなっている．

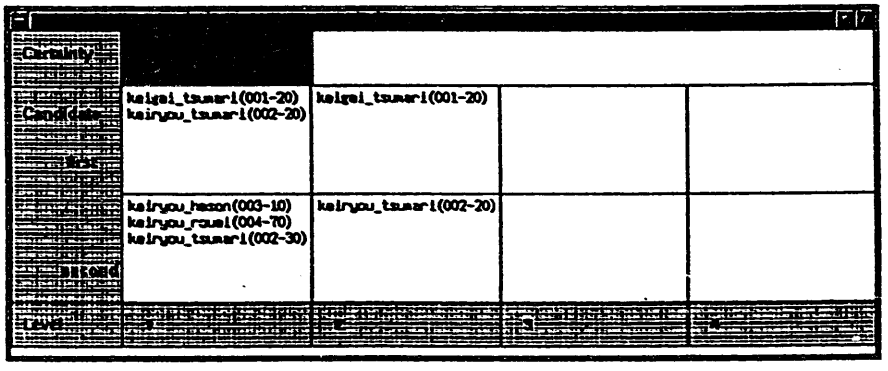
以上の結果より，本システムは時系列的に同定を行なうことが可能であり，またある時点までの部分的な観測のみからでも正しい同定結果を導き出す機能を有していることがわかる．このようにしてプラント挙動の観測初期，すなわち「状態」入力の少ない段階では，比較的抽象度の高いレベルの情報を提示することによって幅広い視点の提供と，それらの可能性への注意を喚起することが肝要であり，また観測が進み多くの「状態」入力から異常候補も確定してきた段階においては，より具体的な焦点を絞り込んだ情報の提示により，選択的に注意を喚起することは有用な実時間での支援環境を提供するものと考えられる．




cf. ***_hason: breagage, ***_tsumari: blockade, ***_rouei: leakage

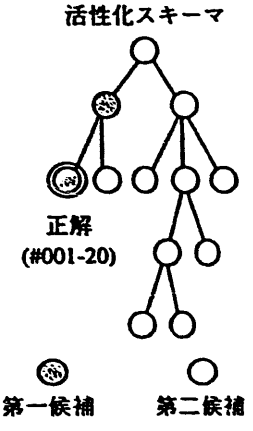
(a) 時刻 68

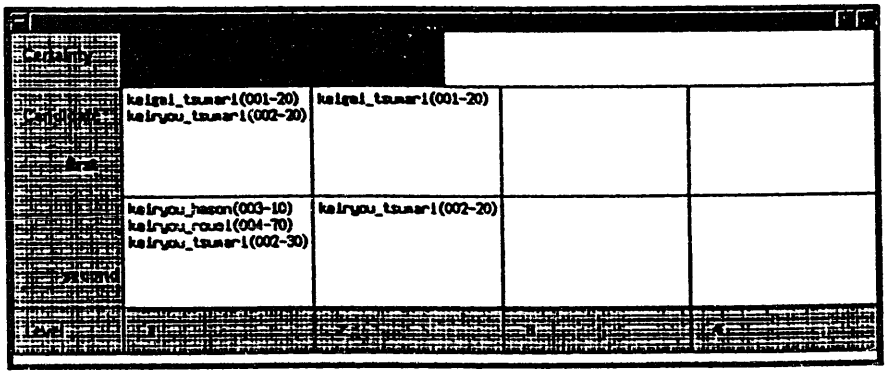



cf. ***_hason: breagage, ***_tsumari: blockade, ***_rouei: leakage

(b) 時刻 80




cf. ***_hason: breagage, ***_tsumari: blockade, ***_rouei: leakage

(c) 時刻 113

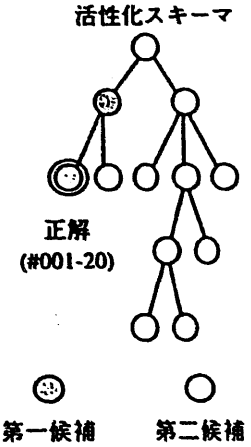



Fig. 6.7「アラスカ系ライン詰り１」に対するリアルタイム診断結果

# 第７章　まとめ

以上本報告では，大規模知識ベースの構築時において必須の技術課題となる矛盾の検知とその解消による知識の獲得学習の手法について概説した．前半では演繹学習や類推学習の分野を中心に展開されている矛盾検知とその解消による知識学習の考え方についてまとめ，後半では帰納学習の一つである概念学習の手法を中心に，過去に集積された事例集合と矛盾をきたすような新規事例が新たに入力された場合に，これを既存知識に同化するべく事例記憶をダイナミックに再組織化していくための手法を中心にまとめた．いずれの手法も基本的は，恒常的に正しい知識ベースを設計するという方針を破棄し，常に観察され入力されてくる事例に対して能動的な既存知識による解釈を試み，それが破綻した場合を「矛盾」として認識し，その上で既存知識の更改や既存事例の再組織化を通して「矛盾の解消」を計り自らの内部に取り込み同化していくことによる学習を実現するものである．このような常に産出を続け「未知」と「既知」の境界を継時的に自己決定していくことのできる「生命的」な情報処理の形態は，事前に完全な形での知識ベースの構築が困難な対象，あるいは想定外事象への自律プラントの対応を考えていくとき，不可欠な設計指針となるものと考える．

最後に本報告書をまとめるにあたり動燃事業団の尾崎禎彦氏（現在三菱電機株式会社）ならびに吉川信治氏に深謝申し上げます．

## 参考文献

[1] Bartlett, F.C.: Remembering: A Study in Experimental and Social Psychology, Cambridge Univ. Press, 1932.

[2] Carbonell, J.G.: Learning by Analogy: Formulation and Generalizing Plans from Past Experiences, in *Machine Learning: An Artificial Intelligence Approach*, Michalski, R. S., Carbonell, J. G., and Mitchell, T. M. (eds.), Morgan Kaufmann, pp.41-81, 1983.

[3] Carlson, J., Weinberg, J. and Fisher, D.: Search Control, Utility, and Concept Induction, *Proc. of the 7th Int. Conf. on Machine Learning*, 85-92, 1990.

[4] Chandrasekaran, B.: Generic Tasks in Knowledge-Based Reasoning: High-Level Building Blocks for Expert Eystem Design, *IEEE Expert*, 1-3, pp.23-30, 1986.

[5] Chandrasekaran, B. and Goel, A.: From Numbers to Symbols to Knowledge Structures: Artificial Intelligence Perspectives on the Classification Task, *IEEE Transactions on Systems,Man,and Cybernetics*, 18-3, pp.415–424,1988.

[6] Clancey, W.J.: Heuristic Classification, *Artif. Intell.*, 27-3, pp.289-350, 1985.

[7] Cheeseman, P. et al.: AutoClass: A Bayesian Classification System, *Proc. of the 5th Int. Conf. on Machine Learning*, pp.54-64, 1988.

[8] DARPA: Proc. of Workshop on Case-Based Reasoning, 1988, 1989.

[9] Ditterich T. G., and Michalski, R. S.: A Comparative Review of Selected Methods for Learning from Examples, in *Machine Learning: An Artificial Intelligence Approach*, Michalski, R. S., Carbonell, J. G., and Mitchell, T. M. (eds.), Morgan Kaufmann, pp.41-81, 1983.

[10] Feigenbaum, E.A. and Simon, H.A.: Epam-Like Models of Recognition and Learning, *Cognitive Science*, 8, pp.305-336, 1984.

[11] Fisher, D.H.: Knowledge acquisition via incremental conceptual ciustering, *Machine Learning*, 2, pp.139–172, 1987.

[12] Gennari, J.H., Langrey, P. and Fisher, D.: Model of incremental concept formation, *Machine Learning*, 5, pp.11-63, 1990.

[13] Gentner, D. et al.: Systematicity and Surface Similarity in the Development of Analogy, *Cognitive Science*, 10, pp.277-300, 1986.

[14] Gentner, D.: Sturucture Mapping: A Theoretical Framework for Analogy, *Cognitive Science*, 7, pp.155-170, 1983.

[15] Gluck, M. and Corter, J.: Information, uncertainty and the utility of categories, *Proc. of the Seventh Annual Conf. on Cognitive Science Society*, Lawrence Erlbaum, Irvine, CA, pp.283-287, 1985.

[16] Hammond, K.J.: Case-Based Planning, Academic Press, San Diego, CA, 1989.

[17] Hanson, S.J. and Bauer, M.: Conceptual Clustering, Categorization, and Polymorphy, *Machine Learning*, 3, pp.343-372, 1989.

[18] Koton,P.: Reasoning about Evidence in Causal Explanation, Proc. AAAI-88, pp.256-261, 1988.

[19] 小山: 都市ガス製造プラントの運転支援システム, 計測自動制御学会編：ニューロ・ファジイ・AIハンドブック, オーム社, 1994.

[20] Laird, J.E. et al.: Soar: An Architecture for General Intelligence, *Artificial Intelligence*, 33, pp.1-64, 1987.

[21] Langley, P., Thompson, K., Iba, W.F., Gennari, J. and Allen, J.A.: An Integrated Architecture for Autonomous Agents, Technical Report of Dept. of Information and Computer Sci., Univ. of California, Irvine, TR 89-28, 1989.

[22] Lebowitz, M.: Generalization from natural language text, *Cognitive Science*, 7, pp.1-40, 1983.

[23] Lebowitz, M.: Integrated learning: Controlling explanation, *Cognitive Science*, 10, pp.219-240, 1986.

[24] Levinson, R.: A Self-Organizing Retrieval System for System for Graphs, *Proc. of the 4th National Conf. on Artificial Intelligence*, pp.203-206, 1984.

[25] Michalski, R.S. and Stepp, R.: Learning from observation: Conceptual clustering, in R.S. Michalski et al. (Eds.), *Machine Learning*, 1983.

[26] Mitchell, T. M.: Generalization as Search, *Artificial Intelligence*, 18, pp.203-226, 1982.

[27] Mitchell, T.M. et al.: Explanation-Based Generalization: a unifying view, *Machine Learning*, 1, 47/80, 1986.

[28] Norman, D.A.: Categorization of action slips, *Psychological Review*, 88, pp.1-15, 1981.

[29] Rajamoney, S. and DeJong, G.: The Classification, Detection and Handling of Imperfect Theory Problem, *Proc. of IJCAI-10*, pp.205-207, 1987.

[30] Rosch, E. and Mervis, C.: Family Resembrance: Studies in the Internal Structure of Categories, *Cognitive Psychology*, 7, pp.573-605, 1975.

[31] 椹木，片井，岩井：オペレータの操作履歴からの制御則の自動チューニングと知的ファジィ制御，計測自動制御学会論文集，26, pp.854-861, 1990.

[32] 岩井, 片井, 椹木, 坂口, 福森: 知能システム工学, 計測自動制御学会, 1991.

[33] Sawaragi, T., Takada, Y., Katai, O. and Iwai, S.: Realtime Decision Support System for Plant Operators Using Concept Formation Method, *Preprints of International Federation of Automatic Control (IFAC) 13th World Congress*, Vol.L, pp.373-378, San Francisco, 1996.

[34] Schank, R.: Explanation Patterns: Understanding Mechanically and Creatively, LEA, 1986.

[35] Schank, R.: Dynamic Memory: A Theory for Reminding and Learning in Computer and People, Cambridge University Press, 1982.

[36] Segen, J.: Graph Clustering and Model Learning by Data Compression, *Proc. of the 7th Int. Conf. on Machine Learning*, pp.93-100, 1990.

[37] Smith , F.E. and Medin, D.L.: Categories and Concepts, Harvard University Press, Cambridge, MA, 1981.

[38] 高田: プラント時系列データの概念学習に基づくマンマシン協調診断, 平成６年度京都大学大学院工学研究科精密工学専攻修士論文.

[39] Thompson, K. and and Langley, P.: Concept Formation in Structured Domains, in Fisher, D.H. et al. (Eds.), *Concept Formation: Knowledge and Experience in Unsupervised Learning*, San Mateo, Morgan Kaufmann, 1991.

[40] Winston, P.H.: Learning and Reasoning by Analogy, *Comm. of ACM*, 23-12, pp.689-703, 1980.

[41] Winston, P.H.: Learning New Priciples from Precedents and Exercises, *Artificial Intelligence*, 19-3, pp.321-350, 1982.

[42] Yang, H. and Fisher, D.: Conceptual Clustering of Mean-Ends Plans, *Proc. of the 6th Int. Conf of Machine Learning*, 232-234, 1989.

[43] Yoo, J. and Fisher, D.: Concept Formation over Problem Solving, in Fisher, D.H. et al. (Eds.), *Concept Formation: Knowledge and Experience in Unsupervised Learning*, San Mateo, Morgan Kaufmann, 1991.