

## 2. 応用上の課題に対する解決法

Learning from Examples: Application to Industry 2. Application Problems and their Solutions by Yasuhiro AKIBA (NTT Communication Science Laboratories), Hussein ALMUALLIM (The Department of Information and Computer Science, King Fahd University of Petroleum & Minerals) and Shigeo KANEDA (Graduate School of Policy and Management/Faculty of Engineering, Doshisha University).

秋葉 泰弘<sup>1</sup>    フセイン・アルモアリム<sup>2</sup>    金田 重郎<sup>3</sup>

1 NTTコミュニケーション科学研究所

2 サウジアラビア国立石油鉱物大学

3 同志社大学大学院総合政策科学研究科・同志社大学工学部

### 1. はじめに

例からの学習技術は、事例が持つクラス（例：故障診断における症状に対する故障部位名、制御におけるシステムの状態に対するアクション）に従って、事例の表現空間をあらかじめ分割しておき、クラス未知の事例が到着すると、その分割に基づいて事例のクラスを推定する技術である（詳細は、本連載の初回を参照されたい）。

当初、例からの学習技術は、単純で理想化された問題しか扱えなかった。実問題へ適用するには、以下に示す4つの課題の解決を要したからである。

- (1) 実応用で収集できる訓練事例の数は少なく、学習結果の精度が不十分になる。これに対しどのように学習性能を確保するのか？
- (2) 実問題では理想的な事例（属性値は離散値表現であり、属性値に誤りは含まれず、分類に必要な属性にも過不足はない）でない事例が含まれることがある。このような問題を含む事例をいかにして許容

するのか？

- (3) 例からの学習アルゴリズムは多数存在し、最高の精度を示す学習アルゴリズムは応用に依存して変化する。いかにして、最適な学習アルゴリズムを選択するのか？

- (4) 実問題に学習アルゴリズムを適用するためのツールをどうやって整備するのか？

今日、さまざまな研究・開発によりこれらの課題は解決されつつあり、例からの学習技術は、実問題を扱える一工学的手法として確立しつつある。

以下本稿では、前回解説した基本技術を前提に、上述の4課題に対してどのような手法が開発されてきたかを、2章～5章において順次紹介する（図-1参照）。最後に、6章で今後の展望を述べる。

### 2. 訓練事例の少なさに対する対策

#### 2.1 背景知識の利用

訓練事例の少なさによる正解率低下を防ぐよく知られた手法に、背景知識の利用がある<sup>3), 4), 6), 7), 20)</sup>。背景

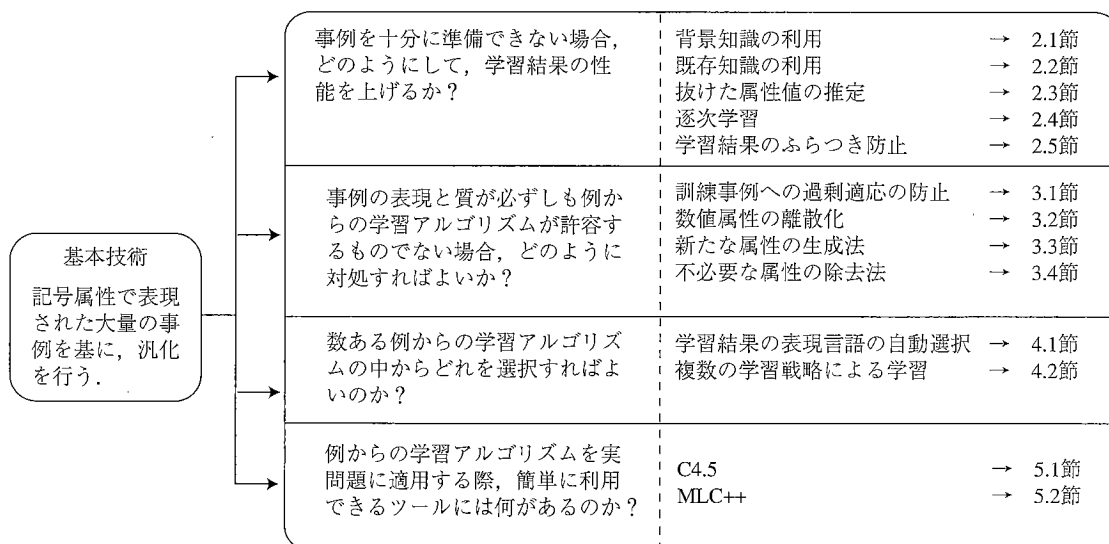


図-1 応用時のさまざまなテクニック

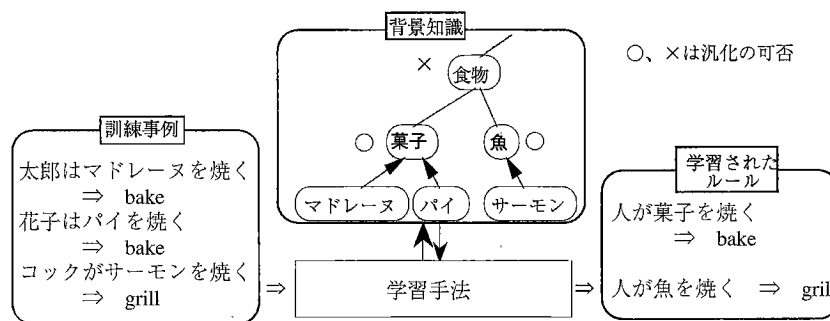


図-2 背景知識を利用した学習

知識は対象世界における概念の上下階層関係であり、自然言語処理では図-2に示すような木構造の背景知識を用いている<sup>\*</sup>。

背景知識は、事例の持つ属性値がノードの1つとなるように与えられる。背景知識を利用すれば、少数の事例からそれらをカバーする知識を生成できる。たとえば図-2では、事例からは「パイまたはマドレーヌを焼く時には、英訳はbake」という情報のみが得られる。背景知識を、パイとマドレーヌから上方向に進めれば、より一般的な知識「菓子を焼く時には英訳はbake」が導びかれる。

Almuallim<sup>6)</sup>は、背景知識を用いて事例をエンコーディングし直すことで、背景知識を利用した学習を実現した。Almuallimの手法では、ターゲット属性に対して、背景知識の各ノードに対応したダミー属性(0, 1の2値)を準備する。たとえば、ターゲット属性が形で、形に関する図-3の背景知識がある場合、任意の形、凸、多角形などがダミー属性となる。ダミー属性の個数は、背景知識が持つノード数に等しい。そして、事例の属性値から出発して最上端のルートに至る、背景知識上のパスに現われるノードに対応したダミー属性に「1」を入れ、その他のダミー属性に「0」を入れる。図-3では、属性値が楕円形なら、任意の形、凸、楕円形に対しては1、その他には0が入る。これにより、背景知識上より高いノードを用いた一般化が自然に実現できる。

## 2.2 既存知識と事例の融合

事例の少なさのほかの対処法は、人手作成のラフな既存知識(図-4参照)を用いて、事例の少なさを補う手法である<sup>2), 21), 22), 26)</sup>。既存知識とは、学習対象のラフなスケッチである。

秋葉<sup>1), 2)</sup>は、既存知識からそれがカバーしている事例(「仮事例」と呼ぶ)を逆に生成し、この仮事例と、現実世界の事例(「実事例」と呼ぶ)とを併せて訓練事例として、例からの学習アルゴリズムに入力するア

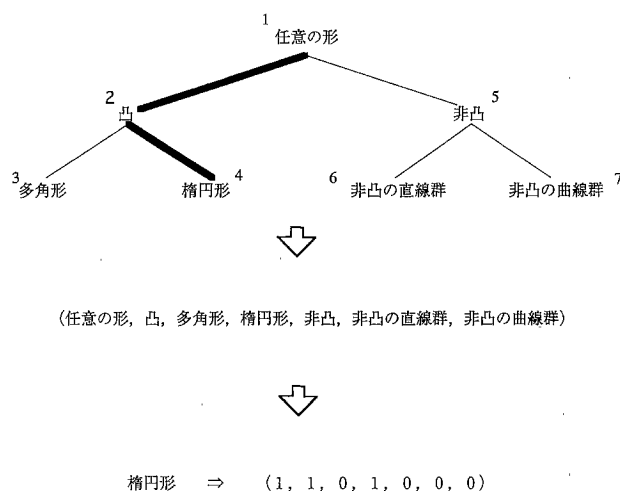


図-3 エンコーディング法

プローチを用いた。仮事例と実事例に付与した重みを調整することにより、既存知識と事例の双方をトータルに考慮した学習結果が得られる。この手法は汎用的であり、種々の例からの学習アルゴリズムに適用できる。

一方、Ourston<sup>21)</sup>, Pazzani<sup>22)</sup>, Quinlan<sup>26)</sup>らは、既存知識・事例を述語論理表現する手法を報告した。これらはTheory Revisionと呼ばれる技術分野に属し、事例に合致するよう既存知識を部分的に修正する。修正オペレータは、Ourston<sup>21)</sup>では、(1) ルールから不必要な条件の削除、(2) ルールの削除、(3) 新たなルールの追加、である。Theory Revisionは、既存知識に近い知識しか学習できない恐れがある。それに対し、秋葉の手法<sup>2)</sup>では、知識の全面的な作り直しが可能である。

## 2.3 Missing Valueの許容

実問題では、測定不可能などの理由で属性値が不明だったり、測定コスト大のために属性値が欠落する場合がある。これを「Missing Value」と呼ぶ。Missing Valueを持つ事例を訓練事例から除くと、事例数が減少する。そこで、Missing Valueの許容法が開発されており、決定木学習では、Quinlan<sup>4)</sup>やBreiman<sup>8)</sup>の手法がよく知られている。

<sup>\*</sup> 実際には、もっと大きな木構造を成す。文献1), 2)を参照。

<sup>\*\*</sup> この手法は、ある属性の属性値に曖昧性がある場合、すなわち、複数の属性値のうち1つだけが正しいがどれが正しいか分からないような場合にも、適用可能である。

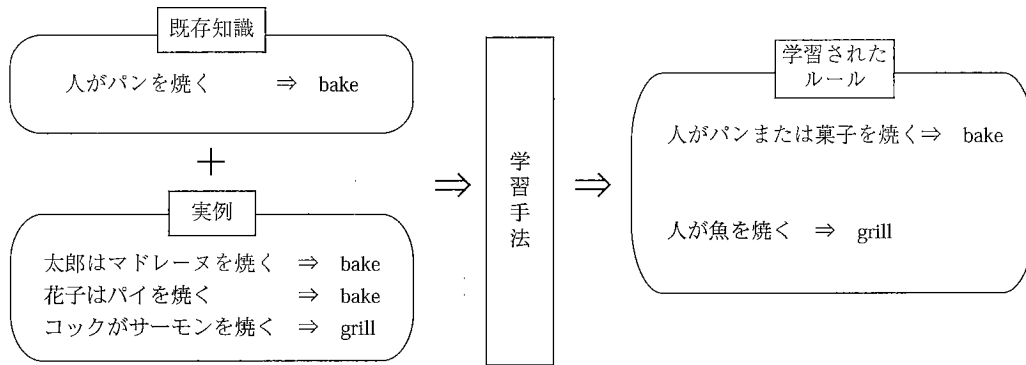


図-4 既存知識を利用した学習

Quinlanの手法では、決定木の作成に際し、1) 情報量期待値を計算するときにはMissing Valueを持つ事例を無視し、2) 事例を下位ノードに流す<sup>24)</sup>ときにはMissing Valueを持つ事例を、Missing Valueを持たない事例の比率に分割する。Quinlanは種々の方法を実験的に比較し<sup>24), 25)</sup>、この方法が一番よいと報告している。

Breimanの手法は、情報量期待値を計算する段階ではQuinlanと同様であるが、事例を下位のノードに流す<sup>23)</sup>方法が異なる。Missing Valueを持つ事例を、属性値が既知の属性の中で、最も相関が高い属性のみ流す。この方法は、相関の高い属性がある場合にはMissing Valueのある事例の流れをうまく模擬できるが、相関が低い場合には問題がある。その意味でQuinlanの手法の方が、Breimanの手法より汎用的である。

#### 2.4 訓練事例の逐次提供

以上の議論では、事例は一度に入手できるとした。一方、故障診断などのデータ入手を考えると、時間的経過を追って逐次的にデータが発生することも多い。この場合の最も簡単な対処法は、新規事例の追加ごとに全事例による学習を繰り返すことである。しかし、この手法は、1) 結果的に学習結果が変わらない場合でも、ゼロから学習するのと同じ学習時間が必要、2) 全事例をメモリに記憶するためメモリを浪費する、などの問題がある。

この問題に対処するため、決定木生成後に新規事例のみを決定木に流し<sup>23)</sup>、決定木を部分的に修正するアルゴリズムが研究されてきた。これらは逐次的修正可能なID3であるが、逐次的な事例追加のどの段階においてもID3と同一結果を保証したのは、UtgoffのID5R<sup>28)</sup>以降である。ID5Rでは、生成された決定木に最初の学習に利用した事例に対する情報を保持し、

<sup>23)</sup> 前回説明したように、決定木上の各ノード（葉を除く）には質問が対応付けられ、各ノードから伸びる各エッジには質問に対する回答が対応付けられている。事例を下位ノードに流すとは、質問に対する回答に応じて事例を上位ノードから下位ノードに受け渡すことをいう。また、この操作をルートノードから繰り返すことを事例を決定木に流すという。

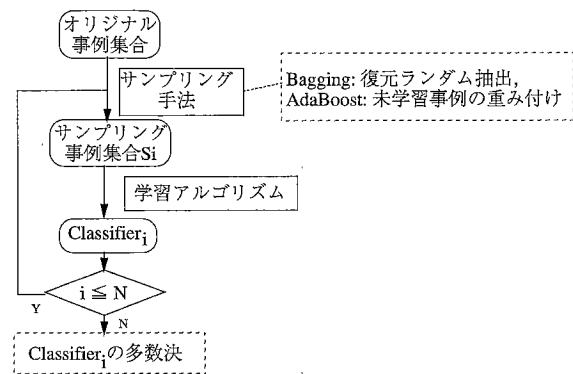


図-5 学習アルゴリズム安定化法

これと新しい事例の情報を併用して決定木を修正する。Utgoffは、Missing Valueや数値属性を許容するように、ID5Rを発展させたITIを提案している<sup>29)</sup>。これら、ID5系列の開発の経緯は、文献15)に詳しい。

#### 2.5 学習性能のばらつきの安定化

近年、学習アルゴリズムの不安定性が問題視されている。不安定性とは、訓練事例が少し変わると学習結果が大きく変化する現象をいう。事例数が少ないと不安定傾向が強く、不安定性は、学習結果の信頼性をゆるがす大きな問題である。

安定化する手法として、BreimanのBagging<sup>9)</sup>、FreundのAdaBoost<sup>13)</sup>が提案されている。以下、決定木学習を例に手法の概要を示すが、他の例からの学習アルゴリズムにも広く適用できる。

Baggingでは、母集団の大きさと同じ大きさの復元抽出サンプリングを行い、そのサンプリングデータを訓練事例として決定木学習を行う。この操作を複数回繰り返し、得られた複数の決定木の多数決によりクラス判別を行う(図-5参照)。

一方、AdaBoostでは、まず訓練事例をそのまま用いて(母集団の頻度分布で)決定木学習を行う。次に、その学習結果ではクラス判定を誤る事例の影響を強めた、新たな訓練事例(クラス誤り事例の分布を強めた分布)で、再度学習を行う。そしてこの決定木の再構

築を複数回行い、Baggingと同様多数決によるクラス判別を行う。

これら手法では、経験分布を基に経験分布に近いさまざまな確率分布を訓練事例上に張り、その確率分布ごとに決定木を作成している。安定化手法として提案されたのはBaggingのみであるが、その処理の類似性からみてAdaBoostなどのBoostingアプローチも、安定化手法と見せる。

### 3. 事例の質に対する対策

#### 3.1 過剰適応の防止

現実の事例では属性値に誤り(ノイズ)が含まれる。ノイズが存在すると、事例への過剰適応が生じる。たとえばID3で、あるノードまで6個の事例が辿りつき、その中の2個の事例はノイズにより誤って辿りついているとする。本来はノード展開を中止すべきであるがID3はさらに分割を続け、このノードの下にサブ決定木を展開する。この過剰適応現象を「過学習」(Over Fitting)と呼ぶ。過学習は未知事例への判別性能を低下させる。

上記問題を解決するため、決定木を刈り込み、コンパクト化する技術が発展しており、プルーニング(Pruning)と呼ばれる。プルーニングには、決定木生成時において、事例のクラスが1種類になる以前に展開を中断するPre-Pruningと、決定木作成後に枝を刈り込むPost-Pruningがある。一般にはPost-Pruningの方が高い判別性能が得られるとされ、種々の手法が提案されている。

Post-Pruningの代表的な方法<sup>4)</sup>では、あるノード配下のサブ決定木を刈り込んだ際の訓練事例に対する誤判定確率の増加を勘案し、訓練事例に対する判別性能を極力下げず、できるだけ大きく刈り込めるところから刈り込む。ただし、どんな応用にもベストの性能を示すPost-Pruning法はない<sup>19)</sup>。応用ごとに、最適なプルーニング手法やプルーニングパラメータを実験的に選択せねばならない。

#### 3.2 数値属性の許容

例からの学習アルゴリズムの中には、事例を表現する属性値として記号しか許容できないものがある。しかし実問題では、属性値として数値が現われる(これを「数値属性」と呼ぶ)。この場合、数値属性の記号化(離散化)が不可欠である。また数値属性を許容する学習アルゴリズムでも、数値属性の離散化により、学習時間の短縮や未知事例に対する正解率が向上した例が報告されている。

代表的な数値属性の離散化手法には、Fayyad<sup>12)</sup>とQuinlan<sup>4)</sup>がある。Fayyadの方法では、まず、離散化のターゲットとなる属性Aについて事例に現われたすべての属性値をソートする。次に、属性値の各々を

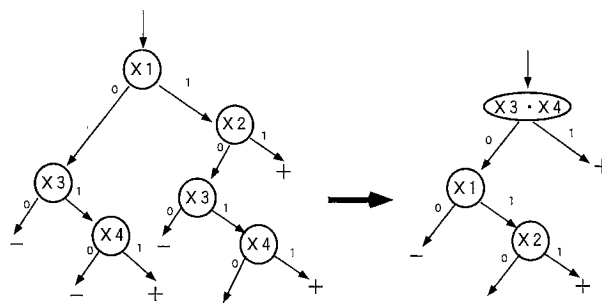


図-6 同じ部分木が複数現れる例 (Replication Problem)

閾値とする2分割を考え、これらを離散化の候補とする。そして、情報量期待値を最大とする閾値を1個選び、2分割を実行する。結果として得られた各分割区間について、ある種の分割終了条件(MDL基準<sup>54)</sup>に基づく)が満たされるまで、同様の分割を繰り返す。

一方Quinlanの方法は、決定木生成中の各ノードの展開時に、Fayyadの方法を局所的に実行するもので、ノードあたり1回だけ区間分割を行う。Quinlanの方法は適用範囲が決定木の学習に限定されるが、Fayyadの方法に比べて同等か僅かによい性能を期待できる。他の離散化手法は、Doughertyのサーベイ<sup>11)</sup>を参照されたい。

#### 3.3 新たな属性の生成

決定木学習である種の関数を表現すると、同じ部分木が多数現われ、表現量が大きくなることがあり、Replication Problemと呼ばれる。図-6にその例を示す。図-6左側は簡単な論理( $X1X2 + X3X4$ )の決定木であるが、下部に同一部分木( $X3X4$ )が何度も出現している。一方、複数属性を組み合わせ、別の属性値を生成し、事例に付加すると、決定木が簡単になることがある(図-6右は、 $X3X4$ を属性として追加した結果、単純化された)。これを、属性生成(Feature Construction)と呼ぶ。訓練事例の判別で同等の性能を示す場合、より簡単な学習結果の方が未知事例に対する判別性能が高い(これを「オッカムの剃刀の原理」と呼ぶ)。決定木においても、単純化により判別性能が向上する可能性がある。

属性生成機能を持つ決定木学習アルゴリズムの代表例は、Fringeに始まる一連のアルゴリズムである<sup>30)</sup>。Fringeでは、生成した決定木の葉部分に繰り返し現われる2項論理積関数(図-6の $X3X4$ )を複数種抽出し、属性として事例に追加し、再学習する。追加した属性候補中、意味のある属性候補のみが決定木に反映される。この手続きを何度も繰り返す。属性候補として取り出す関数は簡単であるが、繰り返しにより、論理積関数と他属性との論理和/積を生成でき、結果的

<sup>54</sup> MDL基準は、学習結果の良さを計るための原理で、事例をうまく説明できるルールが複数個あった場合、「ルールの記述量」と「ルールを基に事例を記述した際の事例の記述量」の和が小さいルールを好ましいとする。

に任意の関数を生成できる強力な手法である。なお、Fringeは「真」となる項のみ取り出していたが、その後「偽」となる項も取り出すなどの改良が加えられた<sup>30)</sup>。

### 3.4 不必要な属性の除去

機械学習では、学習アルゴリズムが必要な属性を自動選択するため、従来属性選択は不要と思われていた。これに対してAlmuallim<sup>5)</sup>は、簡単な属性選択手法を用いて、ID3にも属性選択が有効であると実験的に示した。これを契機に、属性選択の研究が開始された<sup>18)</sup>。たとえば、事例弁別能力に応じた重みを属性に与え、重みが閾値を超える属性のみを利用する方法などがある。これら初期の手法では、学習アルゴリズムを自由に選択できる。

しかし、学習アルゴリズムを特定しない場合、アルゴリズムに適した属性が選択されず、結果的に学習精度を下げる恐れがある。そこで、学習結果のテスト事例に対する性能を用いて、選択された属性の妥当性を評価するアプローチが現われた<sup>16)</sup>。性能推定にはクロスバリデーション (Cross Validation) 法<sup>5)</sup>を利用できる。

ただし属性選択では、属性の組合せをすべて試すのは不可能なため、山登り探索が必要である。具体的には、1) 属性をまったく選ばない状態からスタートし、1属性を追加した場合の性能向上を調べ、性能向上が最大の属性を追加する動作を繰り返す方法、2) 全属性を含める場合からスタートし、1属性ずつ除去する方法、3) 属性の追加/削除を各ステップごとに許容する方法、などがある。属性選択に関する研究は、Langleyのサーベイ<sup>18)</sup>に詳しい。

## 4. 学習アルゴリズムの選択

### 4.1 学習結果の表現言語の自動選択

最適な学習アルゴリズムを選択する1方法では、同一訓練データに複数の学習手法を適用し、最も高い精度を示した学習手法を採用する。この際、学習精度の予測にはクロスバリデーション法を利用する。この最適な学習手法の選択の有効性は、実験的にも確認されている。同様の手法は、学習手法における制御用パラメータ候補 (例: プルーニングの程度を表すパラメータ) からの最適値選択や、属性選択にも利用できる<sup>5)</sup>。

問題領域に対する予備知識があるときは、それを利用した方がより最適な学習手法の選択が可能となることがある<sup>27)</sup>。たとえば、1) 属性のほとんどが学習結果に影響を与える問題では、ニューラルネットワークがどち

らかとえば効果的であり、2) 属性の一部が学習結果に利用される場合や、学習結果が利用する属性が属性値に依存して変化する場合には、学習に要する時間の点から決定木が有利である。

前述の選択方法を多角的に行う手法もある<sup>14), 23)</sup>。これらは、最適なバイアス<sup>56)</sup>を上り探索で求める。まず、複数のバイアスで学習を試み、学習結果を正解率などで評価し、その中で最適なバイアスを選ぶ。選ばれたバイアスに基づき、次回学習を試みるバイアスを複数定め、同様の操作を繰り返す。この方面の研究は、まだ始まったばかりであり、今後の発展が楽しみである。

### 4.2 複数の学習戦略による学習

上記の考え方をさらに進めると、学習結果を決定木やニューラルネットのような、単一の表現形式での表現自体の妥当性が問題となる。この観点からBrodley<sup>10)</sup>は、決定木、線形判別関数、k-Nearest Neighbour法のいずれかをノードとする決定木を用いた学習アルゴリズムを提案した。提案手法は種々のデータにより評価されたが、いずれも、個別アルゴリズム (決定木、線形判別関数、k-Nearest Neighbour法) 中の最高精度を示したアルゴリズムの性能に等しいか、またはこれを超越する性能を得た。この研究は、既存の機械学習アルゴリズムの中から選別するだけでは、精度向上に限界があることを示している。

## 5. 学習アルゴリズムの実用ツール

### 5.1 C4.5

決定木学習のアルゴリズムの中で、特に実問題に広く適用されているのは、QuinlanのC4.5<sup>57)</sup>であり、ID3の改良版である。そのコード (C言語) は、C4.5のアルゴリズムの説明と共に、Quinlanの著作<sup>4)</sup>に添付され、著作を購入すれば自由にC4.5を利用できる。添付コードはよく練られており、アルゴリズムの詳細を理解する上でも、また例からの学習アルゴリズムにおける事例の管理方法の教科書としても優れる。自分でコーディングしようとする者には、格好の手本である。

C4.5で利用できる機能としては、決定木学習の基

<sup>55)</sup> 重複を許さずに、訓練事例をnグループにはほぼ等分割し、そのうちn-1グループを利用して学習を行い、残りの1グループでテストする。そして、テスト事例として選ぶグループを変えながら、この作業をn回行う。これにより、等価的に、すべての訓練事例をテスト事例として扱うことが可能となる。

<sup>56)</sup> 学習結果の表現法 (決定木、論理式、ニューラルネット等)、その表現に対する制約 (論理式の積項数の制限等)、学習時の制御パラメータ、学習戦略 (ID3を用いるか、CARTを用いるかなど)、利用する属性、などは、すべて、学習の前提条件であり、「バイアス (Bias)」と呼ばれる。学習結果/精度はバイアスに依存する。クロスバリデーション法は、このバイアス選択に利用できる。

<sup>57)</sup> C4.5にはさらに機能が付加され、C5.0として入手可能で、バグフィックスと若干の機能拡張 (2.5節で触れたAdaBoost流の安定化法など) がなされている。詳細情報は、<http://www.rulequest.com/>から入手可能である。また、C4.5と並び評される決定木学習アルゴリズム、CARTのパッケージ情報については、<http://www.salfordsystems.com/html/products.html>を参考にされたい。

本機能に加え、先の諸問題、(1) Missing Valueの取り扱い、(2) プルーニングの戦略、(3) 数値属性の離散化、以外に、(4) 決定木を利用した決定リスト学習、(5) 訓練事例の一部から事例を取り出して学習するウインド処理、(6) ノード上の属性値のグループ化、が付加されている。

オプションとしては、テスト選択基準の指定、グループ化の有無、各種パラメータの設定 (Pre-Pruning, Post-Pruning, ウインド処理) がある。デフォルト設定は、種々の応用に対して比較的うまく機能する設定がされている。もし、微妙な調整を必要とする場合には、次節で紹介する、MLC++のパラメータチューニング機能を利用してもよい。

## 5.2 MLC++

MLC++は、Kohaviらによって作られたC++のクラスライブラリーであり、例からの学習アルゴリズムを利用したり、コーディングするための環境を提供する<sup>17)</sup>。現在、学術研究での利用に限定して、ソースとバイナリーが開放されている。

MLC++では、決定木学習を始めとする、17種類もの学習アルゴリズムが提供される。したがって、例からの学習を適用してみたい、手元の応用に合致した手法を検討したいなどの要求に好適である。

また、その他の重要な機能として、

- (1) 学習アルゴリズムの正解率の推定 (Holdout<sup>18)</sup>, クロスバリデーション, Bootstrap),
  - (2) 学習に必要な属性を探す属性選択,
  - (3) 複数の学習学習結果を多数決するクラス決定,
  - (4) 数値データの離散化,
  - (5) 学習アルゴリズムにおける制御用パラメータの最適化 (正解率最大のパラメータを探索),
- などがある。これらの機能は、上述の学習アルゴリズムを内部で切り替えて利用している。

また、学習プログラムを書きたい人には、例からの学習のためのクラスライブラリーばかりでなく、ファイル入出力などの、関連手法を実現するC++のクラスライブラリーが提供されている。

## 6. おわりに

例からの学習技術を実問題に適用する際に解決すべき問題を明らかにし、解決手法を紹介した。例からの学習技術では、90年代に入ってから、実应用到するための要素技術が着実に整備されつつある。これは、80年代のToy Problemしか解けなかったことの反省からである。

例からの学習技術の中でID3などの決定木学習は、高い学習性能を持ちながらも学習時間が短い。この高

速性がある初めて、AdaBoostなどの学習性能安定化手法の適用も容易になり、クロスバリデーションによる学習アルゴリズム選択・パラメータチューニングも実用的手法となる。

例からの学習技術の発展には、さまざまな応用分野への適用結果を学習技術に帰還する必要がある。今後、応用の試みがさらに広がり、それらを基に、例からの学習技術がさらに発展するよう切に願う。

謝辞 日頃ご指導・ご鞭撻をいただき、NTTコミュニケーション科学研究所各位、ならびに、King Fahd University of Petroleum and Minerals各位に深謝いたします。

## 参考文献

- 1) 秋葉泰弘, 石井 恵, フェイン・アルモアリム, 金田重郎: 少ない事例から学習できる機械学習手法, NTT R&D, Vol.45, pp.49-58 (1996).
- 2) 秋葉泰弘, 石井 恵, フェイン・アルモアリム, 金田重郎: 人手作成ルールと事例に基づく英語動詞選択ルールの学習, 自然言語処理, Vol.3, No.3, pp.53-68 (1996).
- 3) フェイン・アルモアリム, 秋葉泰弘, 金田重郎: 木構造属性を許容する決定木学習, 人工知能学会誌, Vol.12, No.3, pp.421-429 (1997).
- 4) J.R.キンラン著, 古川康一監訳: AIによるデータ解析, トップラン, 東京 (1995).
- 5) Almuallim, H. and Dietterich, T. G.: Learning Boolean Concepts in the Presence of Many Irrelevant Features, Artificial Intelligence, Vol.69, pp.279-305 (1994).
- 6) Almuallim, H., Akiba, Y., Yamazaki, T. and Kaneda, S.: Induction of Japanese-English Translation Rules from Ambiguous Examples and a Large Semantic Hierarchy, 人工知能学会誌, Vol.9, No.5, pp.730-740 (1994).
- 7) Almuallim, H., Akiba, Y. and Kaneda, S.: An Efficient Algorithm for Finding Optimal Gain-Ratio Multiple-Split Tests on Hierarchical Attributes in Decision Tree Learning, <http://www1.doshisha.ac.jp/~skaneda/papers/aaai-96.html> (1997).
- 8) Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J.: Classification and Regression Trees, The Wadsworth & Brooks/Cole Advanced Books & Software, California (1984).
- 9) Breiman, L.: Bagging Predictors, Machine Learning, Vol.24, pp.123-140 (1996).
- 10) Brodley, C. E.: Recursive Automatic Bias Selection for Classifier Construction, Machine Learning, Vol.20, pp.63-94 (1995).
- 11) Dougherty, J., Kohavi, R. and Sahami, M.: Supervised and Unsupervised Discretization of Continuous Features, Proc. 12th International Conference, pp.194-202 (1995).
- 12) Fayyad, U. M. and Irani, K. B.: Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning, Proc. ICML93, pp.1022-1027 (1993).
- 13) Freund, Y. and Schapire R. E.: A Decision-Theoretic Generalization of On-line Learning and an Application to Boosting, Proc. EuroCOLT'95, pp.23-37 (1995).
- 14) Gordon, D. F.: Evaluation and Selection of Biases in Machine Learning, Machine Learning, Vol.20, pp.5-20 (1995).
- 15) Kalles, D. and Morris, T.: Efficient Incremental Induction of Decision Trees, Machine Learning, Vol.24, pp.231-242 (1996).
- 16) Kohavi, R. and Sommerfield, D.: Feature Subset Selection Using the Wrapper Method: Overfitting and Dynamic Search Space Topology, Proc. KDD-95, pp.192-197 (1995).
- 17) Kohavi, R. and Sommerfield, D.: Data Mining using MLC++ - A Machine Learning Library in C++, Proc. 8th International Conference on Tools with Artificial Intelligence (TAI96), pp.234-245 (1996).
- 18) Langley, P.: Selection of Relevant Features, Proc. AAAI Fall Symposium on Relevance, pp.171-182 (1994).
- 19) Mingers, J.: An Empirical Comparison of Pruning Methods for Decision Tree Induction, Machine Learning, Vol.4, pp.227-243 (1989).
- 20) Nunez, M.: The Use of Background Knowledge in Decision Tree Induction, Machine Learning, Vol.6, pp.231-250

<sup>18)</sup> ランダムに事例を2つの集合に分割し、片方の集合 (訓練事例と呼ぶ) で学習し、もう片方 (テスト事例と呼ぶ) で正解率を推定する方法。2対1にランダム分割するのが、普通である。

- (1991) .
- 21) Ourston, D. and Mooney, R.: Changing the Rules: A Comprehensive Approach to Theory Refinement, Proc. AAAI-90, pp.815-820 (1990) .
  - 22) Pazzani, M. and Kiblen, D.: The Utility of Knowledge in Inductive Learning, Machine Learning, Vol.9, No.1, pp.59-94 (1992) .
  - 23) Provost, F.J.: Inductive Policy: The Pragmatics of Bias Selection, Machine Learning, Vol.20, pp.35-61 (1995) .
  - 24) Quinlan, J. R.: Induction of Decision Trees, Machine Learning, Vol.1, pp.81-106 (1986) .
  - 25) Quinlan, J. R.: Unknown Attribute Values in Induction, Proc. 6th International Conference on Machine Learning, pp.164-168 (1989) .
  - 26) Quinlan, J.R.: Learning Logical Definitions from Relations, Machine Learning, Vol.5, pp.239-266 (1990) .
  - 27) Schaffer, C.: Selecting a Classification Method by Cross-Validation, Machine Learning, Vol.13, pp.135-143 (1993) .
  - 28) Utgoff, P.E.: Incremental Induction of Decision Trees, Machine Learning, Vol.4, pp.161-186 (1989) .
  - 29) Utgoff, P.E.: An Improved Algorithm for Incremental Induction of Decision Trees, Proc. ML94, pp.318-325 (1994) .
  - 30) Yang, D., Rendell, L. and Blix, G.: A Scheme for Feature Construction and a Comparison of Empirical Methods, Proc. IJCAI-91, pp.699-704 (1991) .

(平成9年11月18日受付)



秋葉 泰弘 (正会員)

1964年生。1988年早稲田大学教育学部数学科卒業。1990年同大学院理工学研究科数学専攻修士課程修了。同年、日本電信電話(株)入社。以来、機械学習、知識獲得、ニューラルネットの研究に従事。1995年人工知能学会全国大会優秀論文賞受賞。現在、NTTコミュニケーション科学研究所研究主任。  
e-mail: akiba@cslab.kecl.ntt.co.jp



フセイン・アルモアリム

1961年生。1984年東京工業大学工学部電子物理学科卒業。1986年同大学院工学研究科修士課程修了。1992年オレゴン州立大学計算機科学科博士課程修了。同年、NTT情報通信網研究所勤務。現在、サウジアラビア国立石油鉱物大学・計算機学科助教授。Ph.D. in Computer Science。機械学習、文字認識などの研究に従事。1991年AAAI Honorable Mention Award 受賞。1995年人工知能学会全国大会優秀論文賞受賞。AAAI会員。e-mail: hussein@ccse.kfupm.edu.sa



金田 重郎 (正会員)

1951年生。1974年京都大学工学部電気第二学科卒業。1976年同大学院電子工学専攻修士課程修了。同年、日本電信電話公社・武蔵野電気通信研究所入所。以来、誤り訂正符号、フォールトトレラント技術、知識獲得、エキスパートシステムなどの研究に従事。1997年より同志社大学大学院総合政策科学研究科・同志社大学工学部教授。工学博士、技術士(情報処理部門)。電子情報通信学会、IEEE、人工知能学会各会員。e-mail: skaneda@mail.doshisha.ac.jp