

事例とニューラルネットからの分類ルール抽出法

齊藤和巳 中野良平

NTT 情報通信処理研究所

本稿では、知識獲得ボトルネック解決のための試みとして分類ルールを自動抽出する2つの方法を提案する。1つの方法は、ヒュリスティクスを用いて探索空間の枝刈りを行いながら事例からルール候補を生成するフェーズと有用なルールを選択するフェーズからなる。この方法は2次元データ、医療診断問題、気象予想問題へ適用された。そして、そこで導入されたヒュリスティクスの効果を評価した。もう1つの方法は、事例の属性値を少しずつ変化させながらニューラルネットの出力レベルの変化からルール候補を生成するフェーズと有用なルールを選択するフェーズからなる。この方法は2次元データへ適用された。

Rule Extraction from Facts and Neural Networks

Kazumi Saito Ryohei Nakano

NTT Communications and Information Processing Laboratories

1-2356, Take, Yokosuka-shi, Kanagawa-ken, 238-03 Japan

We propose two rule extraction methods as one way of overcoming the knowledge acquisition bottleneck. One method consists of two phases: the first phase generates candidates for rules from facts while pruning the search space by the use of heuristics, and the second phase selects useful rules. The method was applied to a two-dimensional artificial problem, a medical diagnosis problem and a weather forecast problem, and effects of the heuristics introduced in the method were evaluated. The other method also consists of two phases: the first phase generates candidates for rules from changes in output levels of a neural network while changing an attribute value of facts gradually, and the second phase selects useful rules. This method was applied to a two-dimensional artificial problem.

1. はじめに

エキスパートシステム構築における最も重要な課題の一つに知識獲得がある。一般に、ヒューマンエキスパートからの知識（ルール）の抽出は困難で労の多いものであり、可能な範囲でその自動化が強く望まれる。

事例からのルール抽出（概念学習）法の目標は、次の3点に関して、それぞれに優れた手法を開発することであるとする。1点目は効率良くルールを抽出することである。問題のサイズに対して多項式オーダーの時間で抽出を行えることは勿論として、問題の規模が大きく多数の事例がある場合でも、妥当な計算量で抽出を行いたい。ID3[Q86]は情報量の期待値を用いて事例を分割して行くことにより、問題の規模が大きくなっても効率良く決定木（decision tree）を構築できる。2点目は簡潔なルールを抽出することである。AQアルゴリズム[M75]はAQシステム[MMHL86]や概念クラスタリング[MS83]の基本アルゴリズムとして用いられ、抽出結果のルールが人間に理解しやすくなるよう工夫されている。3点目は汎化されたルールを抽出することである。一般に、ルールが汎化されたものであるか知ること、は、難しい問題であるが、学習可能性理論[V84]（learnability theory）では、ある確率で学習結果を保証するのに必要な事例数を知ることができる。また、[V84][BEHW86]では、それぞれ具体的なルール抽出アルゴリズムが述べられている。しかし、以上に述べた手法はそれぞれに優れた方法であるが、ルール抽出の目標を全て満足しているとは言い難く、より優れた方法の開発が望まれる。

ニューラルネットを知識ベース、推論の基とするニューラルエキスパートシステム[N88]はルールに基づく従来のエキスパートシステムと比べると、高速でしぶとくノイズに強い推論ができる、短期間に開発できる、そして、データを追加学習して容易に進化できる等の魅力的な長所を有す。しかしながら、実用的なのは分類型問題に限られる。判断の正答性が説明できない、そして、獲得した専門知識が簡単には理解できない等の課題がある。それぞれの課題解決の試みには、説明能力を持つニューラルエキスパートシステムの研究[G88a]、ニューラルネットからの知識抽出法の提案[G88b][SN88]等がある。

本稿では、診断を始め応用の広い分類問題における分類ルールを自動抽出する方法について報告する。まず、ルールの抽出を行う枠組みについて説明する。次に、事例が多く存在する場合に有効となるルール抽出法（RF法：Rule Extraction from Facts）とその実験結果を示す。次に、事例はそれほど多くないがニューラルネットに領域知識等が与えられているときに有効となるルール抽出法（RN法：Rule Extraction from Neural Nets）とその実験結果を示す。

2. フレームワーク

分類ルールの抽出を行う枠組みについて説明する。一つ概念（concept）は、既に学習した他の概念、または、その概念の構成要素となる特徴からなる属性（attributes）によって表現される。例えば、病名という概念は（体温、血圧、性別、…）という属性によって表現できるとする。事例は分類したい概念に属す正の事例と属さない負の事例からなる。例えば、正の事例はその病名の患者であり、負の事例は非患者である。そして、各事例は属性空間内の点として位置付けられる。

一般に、属性を用いて概念を表現する形式はさまざまであり、それらは1つのクラスに対応している。ここではそのクラスとして、人間にとって最も直感的で理解しやすいと考えられる選言標準形（DNF：Disjunctive Normal Form）を考える。DNFは項（term）の論理和（or）として表現され、項はリテラル（literal）の論理積（and）として表現される。各リテラルは対応する1つの属性にのみ関係し、

その属性の値によって、リテラルの値は真または偽となる。また、主な学習アルゴリズムでは、属性値は2値（真または偽）または離散値であるとしている。しかし、広範にアルゴリズムを適用させるため、ここでは属性値を連続量に拡張したものを考える。即ち、1つのリテラルは、属性値に対して適当な実数の区間に含まれるか否かによって真偽を判定するものとする。例えば、 $x_1[0.2, 0.7]$ と表現する。

今後、エキスパートシステム構築における知識獲得という観点から、項のことをルールと呼び、DNFのことをルールの集合と呼ぶ。即ち、我々の目標は、事例とニューラルネットからルールの集合を求めることである。

3. 事例からのルール抽出

RF法はルール抽出の目標に対して次の特徴を有す。正の事例を基に次第にルールを拡張し、余分な探索を避けることにより、アルゴリズムの効率化を計っている。また、以下でも示すようにさまざまなヒューリスティクスを容易に組み込める構造であり、簡潔なルールを選択する戦略を開発し取り込むことが可能である。さらに、RF法は、少なくともある条件の下では、学習可能性を保証する Occam-algorithm [BEHW87] であることが証明可能であると考えられる。

RF法は、ルール候補を生成するフェーズと有用なルールを選択するフェーズからなる。第一フェーズでは、冗長なルール候補の大幅な枝刈りにより生成時間を短縮する戦略を組み込んでいる。第二フェーズはSC (Set_Covering) 問題に他ならず、それに有効なGアルゴリズム (Greedy Algorithm) をベースに、より良いルールを選択するためのヒューリスティクスを組み込んで改良したものをを用いる。ここでの第一フェーズに対応するものとして、[BEHW86] は可能な全てのルール候補を作り出すアルゴリズムを提案した。しかし、これは理論的な考察をするためのものであり、実用的とは言えない。また、AQアルゴリズムのスター作成法は、1つの正の事例に対して、負の事例を含まない最も一般的な複合体 (ルール) を求めるので、探索する複合体の数を制限したとしても、一般に、処理負荷が問題になると考えられる。

3.1 RF法でのルール候補の生成

RF法の第一フェーズのアルゴリズムは以下である。まず、数個の正の事例しか説明できないルール (初期ルール) を作成する。次に、このルールでは説明できない正の事例 (外の事例) を選び、もし可能ならばその事例をも説明できるようにルールの拡大を繰り返す。そして、もしどの外の事例に対してもルールを拡大することができなければ、それをルール候補とする。この操作を全ての正の事例が少なくとも1つ以上のルールで説明できるようになるまで繰り返す。ここで注意すべきことは、どのルールも負の事例を説明してはならないことである。

このアルゴリズムで重要となるのは、初期ルールの作成方法、及び、外の事例を選ぶ順番である。初期ルールの作成方法に関しては、適当な2つの正の事例を選び、それらの内で最も多くの正の事例を説明するルールを選び、ここで複数のルールが候補となる場合には、カバーする領域が広いルールを選ぶというヒューリスティクスを適用した。また、外の事例を選ぶ順番に関しては、どのルールにも説明されていない正の事例を優先して選ぶというヒューリスティクスを適用した。

3.2 SC問題

ルールの選択に用いるSC問題は「NP-困難」として知られているが、単純なGアルゴリズムにより、

かなり良い近似解を得られることが証明されている [GJ79] [J74]。また、学習アルゴリズムの一部として SC 問題を用いるアイデアは、例えば、[BEHW86] にも見ることができる。

G アルゴリズムは、最も多くの事例を説明するルールを選択し、そこで説明された事例を消去するという2つの処理を事例がなくなるまで繰り返す方法である。但し、元々の SC 問題では、説明する事例数が最大となるルールが複数ある場合には、どれを選択するかは任意であるが、ここではより有益なルールを抽出するため、初期状態において多くの事例を説明したルールを選び、それでも複数のルールが候補となる場合には、カバーする領域が広いルールを選ぶというヒューリスティクスを適用した。

以下に示す実験では、多くの事例が消去された後、説明できる事例数が最大となるルールが複数となることがしばしば起こった。このとき上記のヒューリスティクスを用いることにより、明らかに不適当であると考えられるルールの選択を防ぐことができた。また、実験では、十分多くの事例が消去された場合に処理を終了させたものもある。

3.3 RF法の2次元データへの適用

RF法の有効性を検証するため、RF法を2次元データへ適用した。事例は $[0, 1] \times [0, 1]$ から一様分布に従って400事例を選び、図1に示す集合Fに含まれる事例を正の事例、さもなければ負の事例とした。RF法を適用した結果、図1に示す5つのルールを抽出した。但し、ルールの番号は抽出された順番であり、一般に、番号が小さい方がより多くの事例を説明するルールである。また、図1に結果を表示したのもも示す。但し、白丸は正の事例、黒丸は負の事例を表わし、細線で囲まれた領域が正の事例が現われる領域（Fに対応する領域）であり、矩形枠が抽出したルールである。実験より、RF法を用いて良い近似ルールを得られることが分かる。

3.4 RF法の医療診断問題への適用

RF法の現実問題での有効性を検証するため、RF法を医療診断問題へ適用した。ここでは、216属性（各属性値は2値）を持つ300事例（正の事例数=130、負の事例数=170）から、筋収縮性頭痛であるか患者を診断するルールを抽出した。また、この問題では、ID3をそのまま適用できる

F = $\{(x_1, x_2) :$
 $((x_1 - 0.25)^2 + (x_2 - 0.25)^2 < 0.22) \text{ or}$
 $(0.1 < x_2 < 2x_1 - 0.1 \text{ and } 0.1 < x_2 < -2x_1 - 1.1) \text{ or}$
 $(0.5 < x_1 < 0.8 \text{ and } 0.5 < x_2 < 0.9)\}$

正の事例の領域

Rule1 : $x_1[0.50, 0.79] \text{ and } x_2[0.50, 0.89]$
 Rule2 : $x_1[0.08, 0.43] \text{ and } x_2[0.57, 0.87]$
 Rule3 : $x_1[0.20, 0.41] \text{ and } x_2[0.10, 0.37]$
 Rule4 : $x_1[0.15, 0.46] \text{ and } x_2[0.10, 0.22]$
 Rule5 : $x_1[0.22, 0.36] \text{ and } x_2[0.10, 0.45]$

抽出したルール

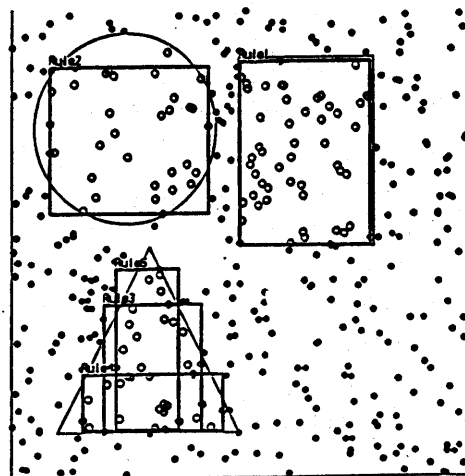


図1. RF法の2次元データへの適用

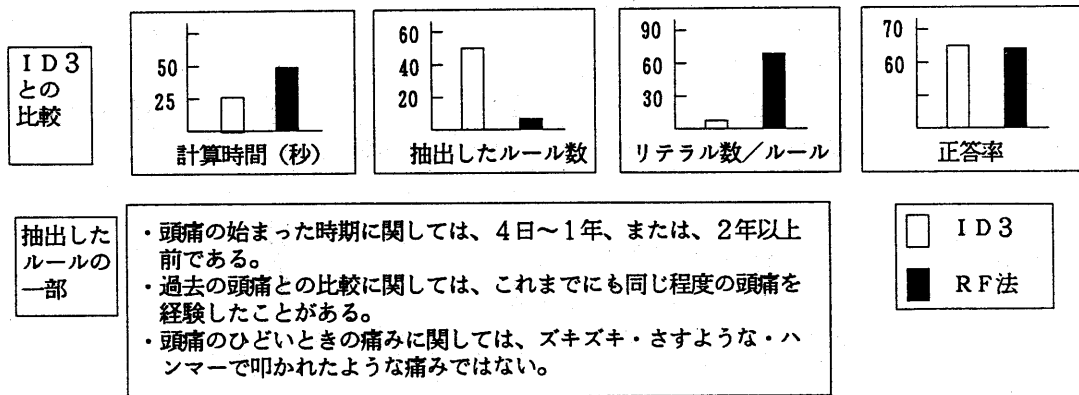


図2 RF法の医療診断問題への適用

ので、それぞれの手法を比較した。比較項目としては、アルゴリズムの効率、抽出したルールの簡潔さ、抽出したルールの一般性を評価するため、アルゴリズムの計算時間、抽出したルール数と1つのルールが持つリテラル数の平均値、トレーニングした事例とは異なる96事例における正答率を用いた。比較結果を図2に示す。実験より、ID3に対し、1つのルールが持つリテラル数の平均値は悪いが、抽出したルール数はかなり少なく、RF法の有望性を示すものであると考える。また、図2では、属性の意味を考慮し、RF法が最初に抽出したルールの一部を示す。

3.5 RF法の気象予想問題への適用

属性が連続量と離散値の混合した複数の属性からなる場合でも、RF法が有効であることを検証するため、RF法を気象予想問題へ適用した。即ち、図3に示す7つの属性（月、天気、風向は離散値、これら以外は連続量）を用い、1987年4～9月（183日）間で、大阪の天気から翌日の東京の天気が雨となるか予想するルールを抽出した（正の事例数=49, 負の事例数=134）。また、ここではより簡潔なルールを抽出するため、Decision List (if~then~else) 形式のルールを抽出するのにR

Mo: 月, We: 天気, Hi: 最高気温 (°C), Lo: 最低気温 (°C), Pr: 降水量 (mm)
Wd: 最大風速 (風向), Wv: 最大風速 (m/S), Hu: 平均湿度 (%)

```

if Mo ≤ 7, Hi ≥ 21.8, Lo ≤ 27.1, Hu ≤ 70, ESE ≤ Wd ≤ N, Wv ≤ 9.5 then False (45)
else if We = 快晴 or 晴れ, Hu ≤ 62, 4.7 ≤ Wv ≤ 8.5 then False (20)
else if Lo ≤ 22.2, Hu ≥ 61, N ≤ Wd ≤ W, Wv ≤ 5.9, Pr ≤ 2.5 then False (16)
else if We = 曇り or 霧雨 or 雨, Hi ≥ 22.7, Lo ≤ 25.7, 66 ≤ Hu ≤ 79, 5.4 ≤ Wv ≤ 8.8 then True (16)
else if Hi ≥ 19.9, Lo ≤ 26.7, Hu ≥ 66, Wv ≥ 6.9, Pr ≤ 25.7 then False (21)
else if Mo ≤ 8, We = 曇り or 霧雨 or 雨, 57 ≤ Hu ≤ 77, N ≤ Wd ≤ W then False (12)
else if Lo ≤ 24.3, Hu ≥ 58%, N ≤ Wd ≤ WSW, Pr ≤ 13.5 then True (14)
else if 13.5 ≤ Hi ≤ 26.7, Hu ≥ 65, 4.7 ≤ Wd ≤ 7.8, Pr ≤ 28 then False (11)
else if Mo ≥ 5, We = 曇り or 霧雨 or 雨, Pr ≤ 38.5 then True (9)
else if Hi ≤ 34.8, 49 ≤ Hu ≤ 64% then False (8)
else True (10)

```

図3 気象予想問題への適用

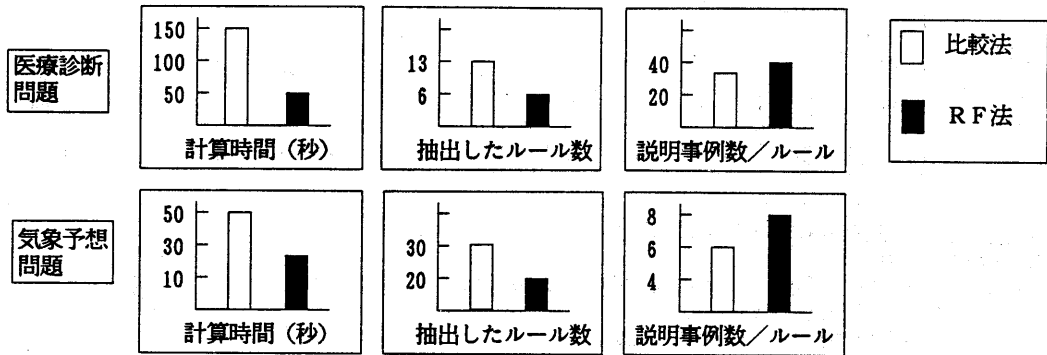


図4 RF法のヒューリスティクスの評価

RF法を用いた。即ち、全ての事例が消去されるまで、RF法を用いて正のルール集合と負のルール集合を求め、2つの集合から最も多くの事例を含むルールを選び、そのルールに含まれた事例を消去することを繰り返すことである。図3に適用結果を示す。但し、風向は最大風速を計測した方向であり、風向の不等号は時計回りである。また、各ルールに付随する数はそのルールで説明される事例数であり、このルール集合のトレーニングデータに対する正答率は99%以上である。実験の結果より、RF法に基づいたアルゴリズムを用いて、属性値が離散値でも連続量であっても、かなり簡潔なルール集合を求められることが分かる。

3.6 RF法のヒューリスティクスの評価

RF法で用いたヒューリスティクスが有効であることを検証するため、ヒューリスティクスを用いない方法とRF法を比較した。比較法では、全ての正の事例が少なくとも1つのルールに含まれるようになるまで以下の処理を繰り返す。まず、どのルールにも含まれない2つの事例を格納順に選び、初期ルールを作成する。そして、全ての正の事例に対して格納順にルールの拡張を行う。比較項目としては、アルゴリズムの効率、抽出したルールの簡潔さを評価するため、アルゴリズムの計算時間、抽出したルール数と1つのルールが説明する事例数の平均値を用いた。医療診断問題と気象予想問題にそれぞれの方法を適用した。結果を図4に示す。実験より、どの比較項目でもRF法が優り、そのヒューリスティクスが有効であることが分かる。

4. ニューラルネットからのルール抽出

事例に基づく知識処理の有力な手法であるニューラルネットは学習結果の正当性が説明できないという欠点があった。ここでは、ニューラルネットと事例から、if-then 型の分類ルールを抽出するRN法を提案する。既に、ニューラルネットからルールを抽出する方法[G88b]は存在するが、いかにして組み合わせの爆発を防ぐかが共通の重要な課題である。RN法では、ニューラルネットへのアクセス回数は、高々、属性数×正の事例数×負の事例数/探索幅、である。また、領域知識等がニューラルネットに与えられ学習が完了している場合には、少数の事例からでも有効なルールを抽出できる。さらに、属性値が離散値でも連続量でも良く、適用範囲が広い。

RN法は2フェーズからなり、第一フェーズにおいて、ニューラルネットと事例からルール候補を撰

動法と類似した考え方をを用いて作成し、第二フェーズでは、RF法と同じ方法を用いる。但し、ここで扱うニューラルネットは、例えば、バックプロパゲーションを学習アルゴリズムとする多層ネットワーク[RM86]である。

4.1 RN法でのルール候補の生成

RN法の第一フェーズでは、ニューラルネットをブラックボックスとみなし、各事例に摂動を与えることにより、その入出力関係からルール候補を作成する。具体的には、まず、ニューラルネットの出力に対し、正の事例とみなす値域（正の値域）と負の事例とみなす値域（負の値域）を決定する。次に、順次任意の正の事例を選択し、各属性毎に摂動を施すことによりルール候補の原形を作成する。即ち、このルール候補の原形は、1つの正の事例を基準とし、ニューラルネットの出力が正の値域となる属性の区間をリテラルとして持つルール（正のルール）である。しかし、1つの属性に摂動を施すときは他の属性の値を固定させるので、一般に、このルール候補の原形は負の事例を含んでしまう。そこで、そこに含まれる任意の負の事例に摂動を施すことにより、ニューラルネットの出力が負の値域となる属性の区間をリテラルとして持つルール（負のルール）を作成する。そして、正のルールから順次負のルールを引くことによりルール候補を作成する。ここで注意すべきことは、引いた結果のルールもルールの形式で表現できることである。

4.2 RN法の2次元データへの適用

RN法の有効性を検証するため、RF法の場合と同じ条件で2次元データにRN法を適用した。但し、ニューラルネットは、同じ400事例をバックプロパゲーションを用いて、誤差の減少がほとんどなくなるまで学習させた。RN法の適用例を図5に示す。図では、等高線がBPネットワーク（ $2 \times 30 \times 1$ ）の出力レベル、矩形枠が抽出したルールである。また、抽出に利用する事例は全領域からランダムに20事例を選んだ。この実験より、RN法は少数の事例からでも有効なルールを抽出できることが分かる。

F = {(x1, x2) :
 ((x1-0.25)2+(x2-0.25)2(0.22)or
 (0.1(x2(2x1-0.1 and 0.1(x2(-2x1-1.1)or
 (0.5(x1(0.8 and 0.5(x2(0.9))

正の事例の領域

Rule1 : x1[0.14, 0.43] and x2[0.56, 0.91]
 Rule2 : x1[0.12, 0.45] and x2[0.09, 0.26]
 Rule3 : x1[0.50, 0.78] and x2[0.50, 0.88]
 Rule4 : x1[0.06, 0.38] and x2[0.63, 0.87]
 Rule5 : x1[0.19, 0.36] and x2[0.11, 0.37]

抽出したルール

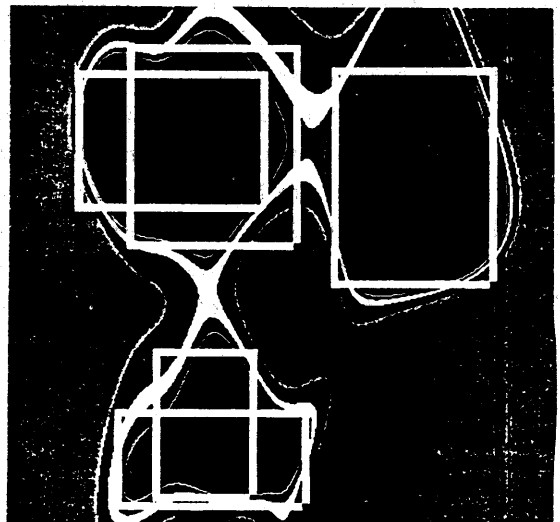


図5. RN法の2次元データへの適用

5. おわりに

本稿では、知識獲得ボトルネック解決のための試みとして分類ルールを自動抽出する方法について報告した。即ち、事例が多く存在する場合に有効となるRF法、及び、事例はそれほど多くないが、ニューラルネットに領域知識等が与えられているときに有効となるRN法を考案した。

今後は、さらに多くのケースに適用し2つの方法の有効性の検証を進めるとともに、各方法の改良と理論的な解析を進めていく予定である。

参考文献

- [BEHW86] Blumer, A., A. Ehrenfeucht, D. Haussler, M. Warmuth, "Classifying learnable geometric concepts with the Vapnik-Chervonenkis dimension", 18th ACM Symposium on the Theory of Computing, 1986.
- [BEHW87] Blumer, A., A. Ehrenfeucht, D. Haussler, M. Warmuth, "Occam's Razor", Information Processing Letters 24, 1987.
- [G88a] Gallant, S. I., "Connectionist Expert Systems", CACM, Vol. 31, No. 2, 1988.
- [G88b] Gallant, S. I., "Example-Based Knowledge Engineering with Connectionist Expert Systems", MIDOCON, 1988.
- [GJ79] Garey, M., D. Johnson, "Computers and Intractability", W. H. Freeman, 1979.
- [J74] Johnson, D. S., "Approximation Algorithms for combinatorial problems", Journal of Computer and Systems Sciences, Vol. 9, 1974.
- [M75] Michalski, R. S., "Synthesis of optimal and quasi-optimal variable-valued logic formula", International Symposium on Multi-valued Logic, 1975.
- [MMHL86] Michalski, R. S., I. Mozetic, J. Hong, N. Larvraç, "The multi-purpose incremental learning system AQ15 and its testing application to three medical domains", AAAI, 1986.
- [MS83] Michalski, R. S., R. E. Stepp, "Learning from Observation: Conceptual Clustering", Machine Learning, Springer-Verlag, 1983.
- [N88] 中野: ニューラルネットと知識情報処理, IFA研究分科会資料, No88-6, pp. 9-16, 1988.
- [Q83] Quinlan, J. R., "Learning Efficient Classification Procedures and their Application to Chess End Games", Machine Learning, Springer-Verlag, 1983.
- [RM86] Rumelhart, D. E., J. McClelland, "Parallel Distributed Processing", MIT Press, 1986.
- [SN88] Saito, K., R. Nakano, "Medical diagnostic expert system based on PDP model", IEEE ICNN, 1988.
- [V84] Valiant, L. G., "A Theory of the learnable", CACM, Vol. 27, No. 11, 1984.