

【特集】 「機械学習、それが人に及ぼさざる理由」

# 計算学習理論における学習

## Machine Learning in Computational Learning Theory

有村 博紀  
Hiroki Arimura

九州大学大学院システム情報科学研究院  
Department of Informatics, Kyushu University.  
arim@i.kyushu-u.ac.jp, http://www.i.kyushu-u.ac.jp/~arim/

**Keywords:** machine learning, human learning, data mining.

### 1. はじめに

#### 1.1 人間の知と機械の知

機械学習 (Machine Learning) とは、計算機による学習の研究である。筆者は、人工知能の研究からスタートして、計算学習理論と、データマイニングの研究を行ってきた。最近ではテキストやウェブデータなどを対象とした機械学習やデータマイニングの研究を行っている [Abe 02, Arimura 94, 有村 02, Asai 02a, Asai 02b]。特に、具体的な学習やデータマイニングの問題に対して、効率良い学習機構を開発したり、問題を解析したりすることに興味をもって研究している。

今回、企画者の野島久雄先生と今井むつみ先生から「人間の知と機械の知」という魅力的なテーマに関する討論会へのお誘いをいただいて、討論に参加した。ここで、筆者の担当は、機械学習である。人間の知能とは何かという壮大な問いに対して、計算機による学習と人間の学習を比較することで、その本質について考えようというのが、ここでのねらいということになる。

具体的には、討論のたたき台として、両先生から次の三つの問いをいただいた。

- 機械は学習する必要があるか?
- 人間の学習は、機械学習に有用か?
- 機械学習の研究から、人間の学習がわかるのか?

第一の問いには「イエス」と答えるとしても、正直に言って、一介の機械学習研究者にとって、これらの問いに答えるのはかなり難しい。なぜこれらの問いに答えるのが難しいかを、自分なりに考えてみると、次のような理由かと思われる。

筆者は、人間の知に興味はあるが、日頃から人間の知能の解明が中心だと考えたことはないし、人間の認知過程の研究をしているわけではない。本特集のテーマ「機械の学習、それが人に及ぼさざる理由を考える」に至っては、途方にくれるばかりである (変なことでも言おうものなら、世界中の機械学習研究者から石を投げられそうである)。

かといって、実用主義に徹して「機械学習の研究は、

人間の知能の本質的な解明には役立たないが、工学的技術として役に立つからそれでいい」と言い切るのにも抵抗がある。学習の本質の解明を目指して、人間の知能を解明すべく機械学習の研究を創始した先人たちに申し訳ない気がするのである。

これらの感想は、典型的かどうかはわからないが、機械学習の研究者にある程度共通する考えではないだろうか。この点で、認知科学の研究者とは、研究に対する考え方が少し異なっていると思う。

そこで本稿では、筆者がこれまでの研究で興味深く感じた事柄や事例を通して、機械学習が目指してきたものを振り返り、機械学習の研究が上記の問いに対してどのように答えられるかを筆者にわかる範囲で考えてみたい。本稿で触れたいトピックを表 1 にあげておく。以下、筆者にお付き合いいただければ幸いである。

表 1 機械学習の歴史 (本稿で扱うトピック)

<ul style="list-style-type: none"> <li>・ 先史: 統計学におけるモデル推定の研究。</li> <li>・ 1960s-70s : 人工知能             <ul style="list-style-type: none"> <li>- Samuel のチェッカプログラム</li> <li>- Plotkin の最小汎化</li> </ul> </li> <li>・ ニューロと統計 (70s - 80s)             <ul style="list-style-type: none"> <li>- パーセプトロン</li> <li>- モデル推定, 統計決定理論</li> </ul> </li> <li>・ 帰納推論 ('70s)</li> <li>・ 計算学習理論 ('80s -)             <ul style="list-style-type: none"> <li>- PAC 学習, 質問学習, オンライン予測学習</li> <li>- 単純さの理論</li> </ul> </li> <li>・ 応用機械学習の盛況 ('90s -)             <ul style="list-style-type: none"> <li>- 役に立つ理論の到来</li> <li>- ブースティング ('96)</li> <li>- サポートベクターマシン ('90s 後半)</li> </ul> </li> <li>・ データマイニング ('90s 後半)             <ul style="list-style-type: none"> <li>- NASA JPL のデータマイニング事例</li> <li>- WWW やe-コマースでの個人化</li> <li>- 新しいタイプのデータ工学?</li> </ul> </li> </ul>
--

#### 1.2 機械学習の研究は何をしているか

はじめに、「機械学習の研究が人間の学習機構の解明に役立つか?」という問いに関連して、機械学習の研究者がどのようなアプローチで研究していて、何を明らかにしようとしているかについて考えてみたい。

筆者自身は、人工知能と、計算学習理論、データマイニングの研究を行う中で、特に自分の研究スタイルを意識したことはないが、強いていえば一貫して数理的な・計

算機科学的な立場から研究を行ってきたと思う。これは、さまざまな問題の背後に個々の問題を越えた数理的構造があり、その数理的な構造を研究することで、問題の本質の解明に取り組むという考え方である。

例えば、A. Turing と萌芽期の計算機科学者たちが取り組んだ計算可能性 (computability) の理論は、「計算とは何か」という基本的な問いに数理的に答えようとした成果である。また、その後続く、計算量理論 (computational complexity) の発展も、効率良い計算の可能性と限界を明らかにしようとする努力だと考えられる [van Leeuwen 90]。これらは、計算の数理的な研究といってよい。計算学習理論の研究は、学習とは何かを探求すること、そして、何が学習できて、何が学習できないかを明らかにすることを目標としている。この過程で、学習の (計算としての) 本質が明らかになっていくという考えである。このアプローチの良い点は、計算可能性や計算量理論がそうであるように、それが計算機であるか、人間であるかにかかわらず、有限の主体 (計算機プログラムも計算資源も有限である) による、形式的な学習について研究するものだけということである。一方、人間の知の探求を考えたとき、このアプローチによる限界もあるだろう。これらの点については、以降で議論したい。

## 2. 機械学習とは？

機械学習とは何だろうか？ この問いに対して、一般的な答えを与えることは難しいが、次のようにいくつかの研究にその考え方を見ることができる。

- N. Wiener の著書「サイバネティックス」[Wiener 62] の中で議論されている、自分を変化させながら、環境に適応していく機械 [Wiener 62]。
- A. Samuel の自動チェッカプレーヤ [Samuel 59]。事前の学習により良い指し手を覚えて、人間とチェッカを指すプログラム。
- M. Gold による帰納推論 (inductive inference) の研究 [Gold 67]。入出力の無限列から極限において、未知の帰納関数 (プログラム) を推論する。
- 応用機械学習分野における例からの学習 (または概念学習) の研究 [Mitchell 86]。素朴ベイズや、最近傍法、決定木学習アルゴリズムなど [Quinlan 86]。
- D. Angluin による質問学習に代表される例からの多項式時間学習アルゴリズムの研究 [Angluin 92]。
- L. Valiant による PAC 学習 (Probably Approximately Correct Learning) や V. Vapnik による統計的学習の研究 [Valiant 84, Vapnik 98]。線形分類アルゴリズムや SVM アルゴリズムの研究など。

1962年にN. Wienerは、著書「サイバネティックス」の9章「学習する機械、増殖する機械」の中で、生物組織を特長づけるものとして、学習する能力と増殖する能

力をあげている [Wiener 62]。彼は、機械が学習することは可能か、また機械による学習は有用かといった問題を議論し、学習とは、過去の環境に基づいて、自分を変えていくことで環境に適応していくことだと結論している。これは、経験から学ぶという意味で、人間の日常的な学習に近い。

学習とは、隠された未知の規則をその具体例から見つけることだという見方もできる。与えられた仮説空間 (未知規則の候補集合) から、与えられた例と、さらに未知の例をうまく説明する規則を見つけることと考えるのである。また別の見方では、学習とは、具体例をより一般的な概念または規則に汎化することだとも考えられる。これらのさまざまな見方について、以下の章で見ていく。

## 3. 帰納推論

1960年代末から科学哲学や人間の言語獲得を動機付けとして、帰納推論 (inductive inference) の研究が行われた。Goldは、Solomonoffと並んで、計算学習理論の創始者とされる [Gold 67, Solomonoff 64]。Goldは、どのようなプログラムが例だけから原理的に推論可能かという古くからの問いを、計算可能性の理論の枠組みで形式化し、未知関数を入出力例の無限列から極限同定する問題として学習を定式化した。また、これをもとに言語の帰納推論についても議論した。

はじめに、先生は自然数上の一入力計算可能関数  $f$  を隠しもっている。例えば、 $f(x) = x^2 + 1$  とする。先生は、順に入力と出力の対を学習者に与えていく：

「入力 0 の答えは 1」,  
「入力 1 の答えは 2」,  
「入力 2 の答えは 5」,

...

学習者は、入出力対  $(n, f(n))$  を一つもらう度に、未知関数に対する仮説  $g$  を一つ出力する。ある時点  $n \geq 0$  以降で、学習者の出力が二度と変更されず、それが未知の関数のプログラムに一致していれば、学習は成功である。

学習者が、毎回、例を説明する仮説をでたらめに選んだのでは、学習はうまくいかない。Goldは、最初の関数から仮説を始めて、反例が来るたびに、またそのときだけに、仮説を更新していくことで、列挙可能な関数のクラスは帰納推論可能であることを示した。また、個々の関数の推論可能性を論じるのではなく、関数のクラスが推論可能かを議論することが大事なことを示した。

これらにより、通常のプログラム全体のクラスは (かなり広いクラスである)、上の列挙手法で推論可能なことを示している。有限の時間で当てられるものは少ないが、学習問題を極限で考えるとかなりのものがうまく当てられるといえる。

Gold の研究は、現在よく行われる機械学習の定式化とは、かなり異なっており、また単純だが、帰納推論を単なる哲学的な議論に終わらせることなく、数理的なモデルで厳密に議論可能であることを示したことに意義がある。認知科学も同様だと思うが、反証可能なやりかたで、実証を積み重ねることで、自然科学的な研究を進めようというわけである。また、学習の無限プロセスとしての見方も、後の研究に影響を与えた。

#### 4. 例からの学習または概念学習

機械学習を単純化した捉え方として、学習を、表面的な現象から、その背後にある未知の規則を獲得することだと考えることができる。別な言い方では、未知の規則をその具体例から学習するのである。これを例からの学習 (learning from examples) または概念学習 (concept learning) と呼ぶ。特に属性データからの学習は、1980年代から盛んに研究されるようになった [Mitchell 86, Valiant 84].

##### 4.1 概念学習とは

概念学習では、はじめに個体の全体集合  $U$  を想定し、その上で物体の集合である概念  $c \subseteq U$  が定まっているとする。概念を、それに属する個体全体からなる外延として考えるわけである。例えば、生物と非生物の全体の集合  $U$  を考えると、その中で「温血動物」という概念  $c_{\text{温血動物}}$  は、動物である個体全体のなす集合だと考え、「は虫類」とは、は虫類である個体全体のなす集合  $c_{\text{は虫類}}$  と考える。

学習は、次のように行われる。まず、未知の概念 (例えば「犬」) があると考えて、この概念に関する具体例 (または単に例) が学習者に次々と与えられる。例は任意の個体であり、それが未知概念に属するなら正例と呼ばれ、そうでなければ負例と呼ばれる。正か負かの区別を、例の分類ラベルという。学習者は、次々と与えられる具体例から、どのような個体が正例で、どのような個体が負例かを分類する規則を学習していく。学習の目標は、任意の個体が未知の概念に属するか否かを正確に分類する規則を獲得することである。

##### 4.2 属性データと分類規則

実際には、機械学習アルゴリズムへの入力である例集合は、表 2 のような表形式の属性データとして表される。表 2 の一番上の行において、一番右の名前が分類ラベル (ここでは「プレイ」) であり、残りが説明属性 (ここでは、「日光」、「気温」、「湿度」、「風速」) である。表の各行は、一つの例を表しており、各説明属性の値を並べて例を表している。説明属性の値は、真偽値の 2 値  $\{0, 1\}$  をとる場合や、有限個の離散値、実数値を取る場合などさまざまな場合がある。分類ラベルは、通常  $\{0, 1\}$  の 2

表 2 概念学習：入力例

id	日光	気温	湿度	風有り	プレイ
1	sunny	hot	high	false	0
2	sunny	hot	high	true	0
3	overcast	hot	high	false	1
4	rainy	mild	high	false	1
5	rainy	cool	normal	false	1
6	rainy	cool	normal	true	0
7	overcast	cool	normal	true	1
8	sunny	mild	high	false	0
9	sunny	cool	normal	false	1
10	rainy	mild	normal	false	1

値を考える。

分類規則は、「日光 = sunny」や「風速  $\geq 20$ 」のような属性値に関するテストを用いて、未知の例に対して分類ラベルを割り当てる。例えば、表 2 の例では、説明属性である「日光」と「気温」、「湿度」、「風有り」の値から、「プレイ」が真か偽かを予測するのである。代表的な予測または学習手法としては、決定木や決定規則などの論理式の学習手法や [Mitchell 86], 単純ベイズ法 (Naive Bayes) や最近傍法 (Nearest Neighbor) などの統計的学習手法が用いられる [Duda 00]. 規則として許される表現の全体を、仮説空間と呼ぶ。属性データを含む説明属性の数をデータの次元と呼び、多数の属性を含むデータを高次元データという。

##### 4.3 仮説の発見

多くの学習アルゴリズムに共通する基本的な戦略は、とりあえず手持ちのデータを正しく分類する規則を、仮説空間  $\mathcal{H}$  から仮説として見つけることである。これを無矛盾仮説発見戦略という。一般的には、ある規則が例集合 (訓練例) をすべて正しく分類したからといって、その規則が次に与えられる新しい例を正しく分類する保証はない。しかし、これは、直感的には良さそうな戦略である。

実際、広く用いられる確率的学習モデルの一つである PAC 学習の枠組みでは、仮説空間  $\mathcal{H}$  に含まれる規則の数が有限ならば、上記の無矛盾仮説発見戦略で計算される規則は、ある程度の (そう悪くない) 分類誤差の保証をもつ。さらに、仮説空間が潜在的に無限集合であっても、VC 次元と呼ばれる仮説空間の複雑さが有限ならば、それに比例する程度の分類誤差の保証をもつことが示されている [Cristianini 00, Vapnik 98].

##### 4.4 複雑な規則の発見

無矛盾な規則を見つけるという基本戦略が決まったところで、実際に例集合から規則を計算する。例えば、想定している規則が

$$(\text{風有り} = \text{false}) \Rightarrow (\text{プレイ} = 1)$$

のように非常に単純な規則だけならば、可能な規則を一つずつ列挙して、それが例を正しく分類するかテストすれ

ばよい。しかし、仮説空間が任意の選言標準形の  $n$  変数ブール式 (DNF) だとすると、大きな  $n$  に対して、 $2^{2^{(n)}}$  個あるすべてのブール式を列挙することは不可能に近い。そこで、大きな仮説空間をうまく扱う工夫が必要である。

例えば、C4.5 などの決定木学習では、任意のブール関数を表現可能な決定木と呼ばれる規則を用いて、分割統治法を用いて、効率良く規則を発見する。根から順に貪欲法で決定木を構築していくことで、属性数が大きなデータに対しても効率良く無矛盾な規則を構築することができる [Duda 00, Quinlan 86]。

また、AQ などの集合被覆方式の学習アルゴリズムでは、正例の一部を覆い、負例をまったく覆わないような単純なブール式 (例えば単項式) を見つけては、それが覆う正例を例集合から除き、また残りに対して同じ操作を繰り返す。こうして得られた単項式を選言で連結することで、

$$\begin{aligned} & ((\text{風有り} = \text{true}) \wedge (\text{日光} = \text{rainy})) \vee \\ & ((\text{風有り} = \text{false}) \wedge (\text{日光} = \text{sunny})) \\ \Rightarrow & (\text{プレイ} = 1) \end{aligned}$$

例集合に無矛盾な規則を効率良く見つける [Mitchell 86]。以上のような工夫で、高次元データにおける複雑な仮説を効率良く扱っている。

#### 4.5 過剰学習の問題

高次元データや、複雑な仮説空間では別の問題もある。前節の C4.5 や AQ などの学習アルゴリズムでは、制限しなければ、仮説空間が任意のブール関数を表現可能である。しかし、応用機械学習の分野では、データに対して複雑すぎる仮説空間は、データへの過剰適応を引き起こし、未知の例に関する予測精度を悪化させることが知られている。

そこで、上のアルゴリズムでは、できるだけサイズの小さな仮説を生成し、途中で生成を打ち切るような経験則が取り入れられており、過剰適応の問題を回避しようとしている。また、いったん構築した無矛盾仮説を刈り込んで、精度を保ったまま、同時に小さく少数のパラメータだけを含むような仮説に精錬する刈込みという手法が開発されている。このような仮説の「経済性」も、学習精度を保つうえで重要な問題である。

#### 4.6 マージン最大化原理

属性データというかなり単純なモデルを用いることで、おもしろいことがわかる。SVM のようなマージン最大化原理を用いた学習アルゴリズムも高次元性または過剰学習への一つの提案と考えられる [Cristianini 00]。

4.3 節で触れたように、仮説空間が有限サイズまたはその VC 次元が有限なら、統計学習理論 (PAC 学習) から、例に無矛盾な仮説の分類誤差を見積もることができる。しかし、この見積りは、任意の確率分布に対して成

立する最悪ケースの見積りであって、実際に期待できる誤差よりもかなり大きめに見積っている。現実には、いじわるな確率分布でデータが生成されることはあまりないと考えると、これは悲観的過ぎる見積りである。

それでは、より現実の性能に近い見積りは可能だろうか？ 今、 $n$  個の属性をもつデータを  $n$  次元空間にばらまかれたラベル 1 または 0 をもつ点と考えると、この空間の超平面であるような仮説 (線形規則) を考えよう。パーセプトロンや SVM などのよく用いられる学習アルゴリズムは、この種の仮説を用いる。

すると、正例と負例に無矛盾な仮説は、超平面に対応し、この空間のラベル 1 をもつ点と 0 をもつ点を二つのグループに分けるはずである。このとき、すべての例のうち、平面に最も近い点から平面までの距離を、マージン (margin)  $\gamma$  と呼ぶ。さらに、点のちらばりの最大距離 (半径) を  $R$  とおく。

このとき、4.3 節と同じように、マージンが大きいときにはマージン  $\gamma$  と  $R$  の言葉で、例に無矛盾な仮説の分類誤差を見積もることができて、マージンがいつも大きくてちらばりの半径が小さいならば分類誤差も小さくなることが示せる [Cristianini 00]。つまり、最悪事データの生成を考えるのではなく、やさしいデータが来るのがほとんどの状況で、常にマージンを小さくできる学習機械が使えれば、データの高次元性にもかかわらず、かなりうまく学習ができるといえるのである。

また、このマージン最大化は、無関係な属性を含めた空間の次元ではなく、平面に「近い」点だけの次元がかかわるようにできるという意味で、高次元性に対する一つの解を提案していて興味深い。最近人気のある SVM は、このマージン最大化の考えを採用し、カーネル法と組み合わせられて構成された学習アルゴリズムである [Cristianini 00, Vapnik 98]。

### 5. 構造的な一般化

学習の別の側面が、具体例の構造的な一般化である。Plotkin は、1970 年に一階述語論理の枠組みで、与えられた正例の集合から、共通の特徴を抽出して、より一般的な概念を得るための最小汎化手続きを研究した [Plotkin 70]。この最小汎化手続きは、具体例に共通する最大構造を抽出して、例を一般化するという最小汎化の考え方を、機械学習に持ち込んだ。

Plotkin の研究で、もう一つの大事な点は、仮説空間に汎化関係を定義して半順序構造をつくり、これをプログラムや規則の意味論的な順序と関係付けて、仮説空間の効率的な探索を実現することである。例えば、一階述語論理の代入に基づく包摂関係は、そのような汎化順序の一例であり、これを用いて最小汎化の計算が可能である。この汎化順序の考えは、一階述語論理や他の表現に拡張され、その後、Mitchell のバージョン空間や Shapiro の

モデル推論システムなどに発展している [Mitchell 86, Shapiro 82]. 筆者も、最小汎化を複数の概念による汎化に自然に拡張した極小多重汎化の研究をしており [Arimura 94], いくつかの概念の和からなる概念の正例からの効率良い学習に有効であることを示している.

## 6. データの高次元性

具体例から一般的な規則や知識を獲得する学習は、一般に難しい問題である。では、この学習の難しさはどこからくるのだろうか？

ちょっと乱暴だが、あえていえば、学習の難しさは

データ量×次元

に起因するといえる。

よくある説明は、次のとおりである。次元数  $n$  の属性データの個々の例は、 $n$  次元空間の点だとみなせる。したがって、一見すると、空間をいくつかの適当なセルに区切って、同じ分類ラベルをもつセルだけをラベル付けすれば、学習、すなわち、ラベルの予測は簡単に実現できそうである。しかし、次元数  $n$  に対して、総セル数は  $2^{O(n)}$  で増大するので、次元数  $n$  が大きければ、どんなに例数が多くてもほとんどのセルが空になってしまう。つまり、高次元データは、学習に限らず、人間の直感が（直感しか？）働かない世界なのである。したがって、このような疎なデータで、うまくラベル付けを行う規則を見つけることは難しい。

概念学習は、現実の複雑な環境を、表形式の属性データからの 2 値分類規則の学習として単純化したいわば、学習のプロトタイプ的なモデルである。しかし、その一方で、現実の学習問題で本質的な

- データ空間の高次元性（次元の呪い）
- 規則における組合せの膨大さ
- 汎化と過剰適応の問題
- 属性抽出の問題

といった学習の基本的な難しさをそのまま受け継いでいる。生産的な問いを発することが良い研究の枠組みだとすれば、その単純さにもかかわらず、現在の機械学習研究は、比較的良い問いを提出しているといえるのではない

表 3 成功した機械学習アルゴリズムの例

<ul style="list-style-type: none"> <li>・ 決定木学習                     <ul style="list-style-type: none"> <li>– CART (80's), ID3 ('88), C4.5 (90's)</li> <li>– 情報量基準最適化+分割統治法による局所探索</li> </ul> </li> <li>・ 線形分類規則                     <ul style="list-style-type: none"> <li>– パーセプトロン (70's)</li> <li>– Winnow 族 (乗算重み更新+多数決原理)</li> <li>– Boosting (誤りの増幅+多数決原理)</li> <li>– SVM (余白最大化原理+カーネル法)</li> </ul> </li> <li>・ 帰納論理プログラミング                     <ul style="list-style-type: none"> <li>– Progol (精密化による汎化)</li> </ul> </li> <li>・ 統計的手法                     <ul style="list-style-type: none"> <li>– 単純ベイズ法 (ベイズ推論の一次近似)</li> <li>– 最近傍学習法 (事例に基づく学習)</li> </ul> </li> </ul>
--

だろうか。

表 3 に、いくつかの成功した学習アルゴリズムをあげる。先に触れたように、これらのアルゴリズムは、主にデータと規則の高次元性に起因する学習の難しさを解決しようとした試みだといえる。

アルゴリズムについて、紙数の制限でこれ以上の説明はできないが、読者のみなさんには、これらの学習アルゴリズムが上記の困難をどのように解決しているかを考えていただくとおもしろいと思う。

## 7. おわりに

データと規則の高次元性のもとでの学習は、計算機だけでなく、人間にとっても難しい問題に違いない。紹介した機械学習アルゴリズムが採用している工夫と人間の学習の関係や、人間がどのように問題を解決または回避しているかは、興味がある問題である。

統計的学習については、[麻生 03] や情報論的学習理論の解説などを参照いただきたい。機械学習とデータマイニングに関しては、適当な文献と解説を参照していただければ幸いである [Angluin 92, Fayyad 02, Mitchell 86, Sutton 98]. 機械学習の応用の現状については、少し前になるが、Tom Mitchell の「機械学習は本当に役立っているか？」と題した解説があっただけでなく、本稿が、人間の学習と機械の学習に関して、考えるきっかけになれば幸いである。

## 謝辞

今回の討論の機会をいただいた今井先生と野島先生、および討論の参加者の皆様に感謝いたします。元田 浩先生、鷲尾 隆先生（阪大）をはじめとする科研費特定領域研究「アクティブマイニング」メンバ諸氏と、山本章博先生（北大）、平田耕一先生、篠原 武先生（九工大）、佐藤 健先生、宇野毅明先生（国立情報学研）、山西健司氏、竹内純一氏（NEC）、篠原歩先生、竹田正幸先生（九大）には、計算学習理論とデータマイニングに関して議論の機会をいただきました。また、鈴木英之進先生（横浜国大）には、過剰学習についてご教示いただきました。ここに謝意を表します。

## ◇ 参考文献 ◇

[Abe 02] Abe, K., Kawasoe, S., Asai, T., Arimura, H. and Arikawa, S.: Optimized Substructure Discovery for Semi-structured Data, *Proc. PKDD'02*, LNAI 2431, pp. 1-14, Springer-Verlag (2002)

[Angluin 92] Angluin, D.: Computational Learning Theory: Survey and selected bibliography, *Proc. 24th STOC*, pp. 351-369, ACM (1992)

[Arimura 94] Arimura, H., Ishizaka, H., Shinohara, T. and Otsuki, S.: A generalization of the least general generalization, *Machine Intelligence*, Vol. 13, pp. 59-85, Oxford

- University Press (1994)
- [有村 02] 有村博紀, 坂本比呂志: テキストマイニングにおける最適パターン発見, 応用数理, 特集「テキストマイニング」, Vol. 12, No. 4, pp. 32-44 (2002)
- [Asai 02a] Asai, T., Abe, K., Kawasoe, S., Arimura, H., Sakamoto, H. and Arikawa, S.: Efficient substructure discovery from large semi-structured data, *Proc. SDM' 02*, pp. 158-174, SIAM (2002)
- [Asai 02b] Asai, T., Abe, K., Kawasoe, S., Arimura, H., Sakamoto, H. and Arikawa, S.: Online algorithms for mining semi-structured data stream, *Proc. IEEE ICDM' 02* (2002)
- [麻生 03] 麻生英樹: 機械の学習と人間の学習—統計的学習の立場から—, 人工知能学会誌, Vol. 18, No. 5, pp. 526-530 (2003)
- [Cristianini 00] Cristianini, N. Shawe-Taylor, J.: *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Cambridge (2000)
- [Duda 00] Duda, R. O., Hart, P. E., Stork, D. G.: *Pattern Classification*, 2nd edition, Wiley (2000)
- [Fayyad 02] U. M. Fayyad, R. Uthurusamy, eds.: *Evolving Data Mining into Solutions for Insights*, *CACM*, Vol. 45, No. 8, pp. 28-32 (2002)
- [Gold 67] E. M. Gold: Language identification in the limit, *Information and Control*, Vol. 10, No. 5, pp. 447-474 (1967)
- [Mitchell 97] Mitchell, T. M.: Does machine learning really work?, *AI Magazine*, Vol. 18, No. 3, pp. 11-20 (1997)
- [Mitchell 86] Mitchell, T., Carbonell, J. G. and Michalski, R. S.: *Machine Learning: A Guide to Current Research*, Kluwer (1986)
- [Plotkin 70] Plotkin, G. D.: A note on inductive generalization, *Machine Intelligence*, Vol. 5, pp. 153-163 (1970)
- [Samuel 59] Samuel, A. L.: Some studies in machine learning, using the game of Checkers, *IBM Journal of Research and Development*, Vol. 3, pp. 210-229 (1959)
- [Shapiro 82] Shapiro, E. Y.: *Inductive inference of theories from facts*, Technical Report 2, Yale University (1981), Reprinted in Lassez, J.-L. and Plotkin, G. D. (eds.): *Computational Logic*, pp. 199-245, The MIT Press (1992); 有川節夫 訳: 知識の帰納的推論, 共立出版 (1986)
- [Solomonoff 64] Solomonoff, R. J.: A Formal Theory of Inductive Inference, *Information and Control*, Vol. 7, No. 1, pp. 1-22 (1964)
- [Sutton 98] Sutton, R. S. and Barto, A. G.: *Reinforcement Learning: An Introduction*, MIT Press (1998)
- [Quinlan 86] Quinlan, J. R.: Induction of Decision Trees, *Machine Learning*, Vol. 1, No. 1, pp. 81-106 (1986)
- [Valiant 84] Valiant, L. G.: *A Theory of the Learnable*, Vol. 27, No. 11, pp. 1134-1142, CACM (1984)
- [Vapnik 98] Vapnik, V.: *Statistical Learning Theory*, Wiley (1998)
- [van Leeuwen 90] van Leeuwen (ed.): *Handbook of Theoretical Computer Science*, Volume A, Algorithms and Complexity Elsevier (1990); 廣瀬 健, 野崎昭弘, 小林孝次郎 監訳: コンピューター基礎理論ハンドブック I, アルゴリズムと複雑さ, 丸善 (1990)
- [Wiener 62] N. ウィーナー 著, 池原正夫 ほか共訳: 学習する機械, サイバネティクス (第2版), 第9章, 岩波書店 (1962)

2003年7月1日 受理

## 著者紹介



有村 博紀 (正会員)

1990年九州大学大学院総合理工学研究科情報システム学専攻修士課程修了。同年同大学情報工学部知能情報工学科助手。同講師, 同助教を経て, 1996年同大学院システム情報科学研究科助教授。現在, 同大学院システム情報科学研究科情報理学部助教授。1996年ヘルシンキ大学訪問研究員。1999~2002年科学技術事業団さきがけ研究21「情報と知」領域研究員。博士(理学)。計算論的学習理論とデータマイニングの研究に従事。人工知能学会2000年度優秀論文賞などを受賞。ACM, 情報処理学会各会員。