

## 解 説



# EXACT 学 習†

—質問からの概念学習—

榊原 康文†† 西野 哲朗†††

## 1. はじめに

計算論的学習理論という研究分野がここ数年盛んになってきている。これは機械学習についてその基礎理論を研究する分野である。一般に機械学習とは、計算機にプログラミングによって陽にプログラムを与える以外の方法によって、計算機がプログラムを獲得することを指す。本解説で取り上げる帰納推論という方法は、陽にプログラムを与える以外の方法の一つであり、また機械学習において最も活発に研究されている重要な方法である。帰納推論は、プログラムへの入力とその入力に対するプログラムの出力からなる幾つかの具体的な入出力ペアが与えられたとき、そのプログラム自身を推論する方法である。Sloan は彼の学位論文<sup>17)</sup>の中で、このような機械学習の問題を厳密にまた理論的に議論・研究するためには、少なくとも以下の機械学習問題を特徴付ける7つの項目に対してまず正しく答えられなければならないと主張している。

1. What is being learned?
2. From what is it learned?
3. What a priori knowledge does the learner begin with?
4. How is what is learned represented?
5. By what method is it learned?
6. How well is it learned?
7. How efficiently is it learned?

この中で2番目の項目が、この解説で取り上げる学習問題を特徴付けるものである。「何から学習されるか?」。先ほど帰納推論という学習方法は、

入出力の具体例から学習すると説明したが、この具体例の与え方にも自由度がある。どのような具体例をどのようにして与えるかについて、最近 Angluin はまったく新しい方法を提案し、定式化した<sup>2)</sup>。それは学習機械が教師への質問を用いて概念を学習するというものである。質問を用いることによって、適切な具体例や具体例についての情報を得ることができる。Angluin は、このような教師と学習者のパラダイムにおける概念学習の問題を形式的に自然に定義することに成功し、また多項式時間学習という概念学習における計算効率の基準を導入した。この解説では、このような質問を使う機械学習において、どのような質問が機械が学習する際に有効であるかという、質問の種類と多項式時間学習可能性との関係について、最新の肯定的・否定的両方の結果を織り交ぜながら解説していく。また質問とは異なる方法で具体例を与える代表的な学習モデルである「極限における同定」や「PAC 学習」などとの関連についても考察し、EXACT 学習において示される結果をそれらの学習モデルに適用する一般的方法についても述べる。

## 2. 概念学習、質問の種類、多項式時間学習

この解説では、その出力が0か1だけであるプログラムのみを学習対象として考える。このようなプログラムはすべての入力を0か1に分類するもの、すなわち世の中のを白か黒に分類するプログラムであり、このプログラムによって1と(黒と)分類されるものの全体を**概念**と呼ぶ(図-1)。概念を学習対象とする機械学習を一般に**概念学習**と呼ぶ。たとえば、入力として、アルファベット 26 文字と空白を勝手に任意の長さだけ並べた文字列すべての集合を考える。このとき、英語の文として(意味は考えないで単に文法的に)正しい文字列とそうでない文字列に分類するプログ

† Exact Learning—Concept Learning from Queries by Yasubumi SAKAKIBARA (International Institute for Advanced Study of Social Information Science, FUJITSU LABORATORIES LTD.) and Tetsuro NISHINO (Department of Information Science, Tokyo Denki University).

†† (株)富士通研究所国際情報社会科学研究所

††† 東京電機大学

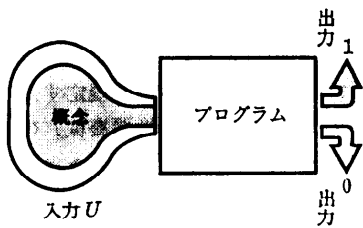


図-1 概念を定義するプログラム

ラムは、英語というものに対応する一つの概念を定義している。このように言語学習は一つの概念学習である。

前にも述べたように、陽にプログラムを与える以外の学習機械に具体例を与える方法の一つとして、最近 Angluin は全く新しい方法を提案し、定式化した<sup>2)</sup>。それは質問を用いる方法で、獲得すべきプログラムに関するいく種類かの質問をとおして、そのプログラムに関する情報が伝達されるようになっており、学習機械はそれらの質問を用いてそのプログラムを推論する(図-2)。しかしあらゆる種類の質問(たとえば「そのプログラムは何ですか?」など)が許されるわけではなく、Angluin は次のような6つの質問を基本的なものとして考えている。

それらの質問について述べるために、いくつかの記法に関する準備をする。プログラムに与えられる入力の全体集合を  $U$  とする。概念は  $U$  の任意の部分集合となる。学習すべき未知の概念を  $c_*$  と表す。概念を定義するプログラムを  $p$  とまたプログラムの列を  $p_1, p_2, \dots$  と書くこととする。またプログラム  $p$  が定義する概念を  $c(p)$  と表す。明らかに  $c(p) \subseteq U$  である。これらの記法を用いて概念学習に使われる6つの質問を定義する。

1. 所属性 (membership). 入力 は  $U$  の要素  $x$  であり、出力は  $x \in c_*$  ならば *yes*,  $x \notin c_*$  ならば *no* である。

2. 等価性 (equivalence). 入力 はプログラム  $p$  であり、出力は  $c(p) = c_*$ , すなわち入力されたプログラム  $p$  が未知の概念を定義しているならば *yes*,  $c(p) \neq c_*$  ならば *no* である。さらに、答が *no* のときは、 $c(p) \oplus c_*$  の任意の要素  $x$  が返される。(ここで、 $c(p) \oplus c_*$  は  $c(p)$  と  $c_*$  の対称差、すなわち、 $c(p) \oplus c_* = (c(p) - c_*) \cup (c_* - c(p))$ .)

3. 部分性 (subset). 入力 はプログラム  $p$  であり、出力は  $c(p) \subseteq c_*$ , すなわち入力されたプロ

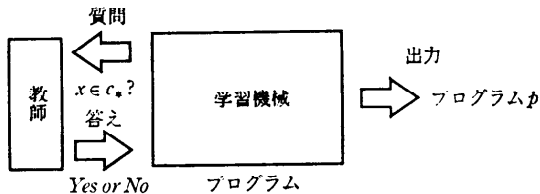


図-2 質問を用いた学習

ラム  $p$  の定義する概念が未知の概念に含まれているならば *yes*, そうでなければ *no* である。さらに、答が *no* のときは、 $c(p) - c_*$  の任意の要素  $x$  が返される。

4. 包含性 (superset). 入力 はプログラム  $p$  であり、出力は  $c(p) \supseteq c_*$ , すなわち入力されたプログラム  $p$  の定義する概念が未知の概念を含んでいるならば *yes*, そうでなければ *no* である。さらに、答が *no* のときは、 $c_* - c(p)$  の任意の要素  $x$  が返される。

5. 非交差性 (disjointness). 入力 はプログラム  $p$  であり、出力は  $c(p) \cap c_*$  が空ならば *yes*, そうでなければ *no* である。さらに、答が *no* のときは、 $c(p) \cap c_*$  の任意の要素  $x$  が返される。

6. 全体性 (exhaustiveness). 入力 はプログラム  $p$  であり、出力は  $c(p) \cup c_* = U$  ならば *yes*, そうでなければ *no* である。さらに、答が *no* のときは、 $c(p) \cup c_*$  に属さない任意の要素  $x$  が返される。

所属性質問以外の質問で返される要素  $x$  を反例と呼ぶ。反例の選択は任意であると仮定する。(たとえば、等価性質問の場合、反例は  $c(p) \oplus c_*$  中の要素であればどれでも良い。) 特に *yes*, *no* だけを答えて反例は返さないようなタイプの質問も考えられるが、このような質問を制限された質問と言うことにする。質問を用いた学習においては、許される質問の種類が学習機械の学習能力に大きな影響を与える。これらの影響を明らかにしようと、6つの質問のうちの幾つかを組み合わせたり、または単独で使用したりする種々のケースを解析することによって、各学習対象ごとに、学習機械が効率良く正しく学習するためにはどの質問とどの質問が十分であるか・不十分であるかを解明する研究が活発に行われている。その際に、計算機科学の分野において今日までに確立されている理論、たとえば計算の理論、計算量理論、アルゴリズムの理論、形式言語理論などにおける数学

的道具を用いて学習アルゴリズムを設計・解析したり、学習可能性・不可能性を研究していくのが、計算論的学習理論なのである。

ところで学習機械が学習対象を「正しく」学習する、あるいは「効率良く」学習するとは学習機械がどのように振る舞ったときを言うのであろうか？ やや数学的にはなるがこれらに対する厳密な定義を、先ほど紹介した Sloan が主張する機械学習問題のための7項目と照らし合わせながら与えることにする。まず学習機械は今日の計算機上で実行可能なプログラムとする。その入力に使われる学習モデルによって異なるが、本解説の EXACT 学習モデルにおいては、質問に答えてくれるオラクル（神託、教師）から与えられる。これは Sloan の7項目の2番目を規定する。（ほかの学習モデルについては後ほど紹介する。）出力はある概念を定義するプログラムである。次に学習機械が学習する「学習対象」を有限または無限個の概念の集まり（概念のクラス）とし、 $C$ で表す。すなわち  $C = \{c_1, c_2, \dots\}$  ( $c_i$  は概念)。学習される未知の概念  $c_*$  はこの  $C$  から選ばれるとする。この  $C$  は、7項目の1番目を規定する。学習機械が出力するプログラムのクラスを  $P$  と表す。すなわち  $P = \{p_1, p_2, \dots\}$  ( $p_i$  はプログラム)。  $P$  は仮設空間とも言う。この  $P$  は、7項目の4番目を規定する。各プログラム  $p$  に対して、そのプログラムの大きさを計るある適切な方法（たとえば、プログラムの長さなど）を仮定し、その方法によって決定されるプログラムの大きさをサイズと呼ぶ。このプログラムのサイズを用いて、未知の概念  $c_*$  の複雑さというものを定義する。未知の概念  $c_*$  の複雑さは、 $c(p) = c_*$  であるプログラム  $p$  の中で最もサイズが小さいプログラムのサイズである。学習機械が上にあげた質問の幾つかを用いて概念のクラス  $C$  を厳密に (exactly) 学習するとは、学習機械が学習中に発するすべての質問に対して ( $c_*$  に関して) 正しい答が返されるときに、学習機械が有限時間内に必ず停止し、かつ  $c(p) = c_*$  となるあるプログラム  $p$  を出力することを言う。これは7項目の6番目を規定する。またこのような学習機械  $A$  が効率良く学習するとは、 $A$  の実行中の任意の時点において、 $A$  によってそれまでに費やされた計算時間が  $n$  と  $m$  の多項式で押さえられることを言う。ここで、 $n$  は  $c_*$  の複雑さ、

$m$  はその時点までに与えられた反例の最大の長さとする。（反例が与えられていない場合には、 $m = 0$ 。）これは7項目の7番目を規定する。このようにプログラム（ここでは学習機械）の実行時間があるパラメータの多項式（の大きさ）で押さえられることは、現在の計算機上で効率良く実行できるプログラムが満たすべき条件として広く認識されている基準である。

### 3. 極小最適教師と学習アルゴリズム

所属性質問は、質問された要素が未知の概念  $c_*$  の要素であるか否かという1ビットの情報を返す。情報源がある分野の専門家であれば、学習機械が尋ねるさまざまなケースを分類するために、その専門家に質問することは妥当なことであろう。しばしば所属性質問は、学習機械が効率的な学習を達成するのに重要な役割を演じる。Angluin は所属性質問と等価性質問に答えることのできる能力をもった教師のことを、「極小最適教師 (Minimally Adequate Teacher)」と呼んでいる。極小最適教師は、学習機械にとって最低限必要であり、最も妥当であると考えられる能力をもった教師として提案された。Angluin<sup>1)</sup> は、この極小最適教師から正則言語のクラスを効率良く厳密に学習する学習機械（学習アルゴリズム）を示している。ここで正則言語とは、形式言語理論において定義される有限状態機械（有限オートマトン）によって認識される文字列の集合を指す。正則言語のクラスは、形式言語理論における Chomsky の階層において最も単純な言語のクラスであるが、さまざまな興味深い特徴をもつ基本的で重要な言語のクラスである。詳しくは文献 8), 22) を参照されたい。学習理論においても、正則言語を学習対象とする学習問題は最も基本的でかつ最も重要な問題として、活発に研究されてきている<sup>13)</sup>。一方、等価性質問だけからでは、どんな学習機械も正則言語のクラスを効率良く厳密に学習することはできないことが理論的に証明されている<sup>4)</sup>。このことは、所属性質問が効率的な学習にとっていかに重要であるかということを示す一つの事実として象徴的である。

さて Angluin によって示された正則言語を学習する学習機械は、質問を用いた学習における一つの基本的原理を示すものとして重要であると考え

られるので、ここでその簡単な紹介をすることにする。これは Sloan の 7 項目の 5 番目に答えるものである。まずあるアルファベット (記号の有限集合)  $\Sigma = \{a, b\}$  を固定し、全体集合  $U$  は  $\Sigma$  中の記号を並べてできる有限の長さの文字列全体の集合とする。極小最適教師は学習機械がこれから学習する正則言語  $c_*$  (これを「未知の正則言語」と呼ぶ) を知っており、学習者は初めアルファベット  $\Sigma$  だけを知っている。学習機械は、極小最適教師に所属性質問と等価性質問を繰り返すことによって、未知の正則言語を受理する最小状態オートマトンを出力する。またその実行時間は、最小状態オートマトンの状態数と学習機械の実行中に教師から返される反例の最大長の多項式で押さえられる。学習機械は、教師に質問をすることによって得られるデータ、すなわちある有限個の文字列とそれが未知の正則言語の要素であるか否かの情報を蓄積するために、内部に観測表と呼ばれるデータ構造をもつ。観測表は、行と列が文字列によってラベル付けされた行列で、行のラベルの集合は接頭語に関して、列のラベルの集合は接尾語に関して閉じており、表の各成分には 0 または 1 が記入される。観測表の各成分の意味は、次のように定められる。行のラベルである文字列  $s$  に列のラベルである文字列  $e$  を接続した文字列  $s \cdot e$  が未知の正則言語の要素であれば、 $s$  行  $e$  列の成分に 1 を記入、そうでなければ 0 を記入する。観測表中の各成分がこのような意味をもつようにするために、学習機械は教師に対して所属性質問を行い、その返答にしたがって観測表中の成分を埋めていく。この観測表が「無矛盾」であり、「閉じている」とき、(これら二つの条件については、文献 1) を参照されたい。) 観測表に含まれるデータからそのデータに無矛盾な最小状態オートマトンを、観測表の大きさの多項式時間で合成することができる。無矛盾で閉じた観測表とその表から合成された有限オートマトンの例を図-3 に示す。

この観測表の性質を利用するために、学習機械は教師に所属性質問を繰り返しながら無矛盾で閉じた観測表を構築していく。無矛盾で閉じた観測表がいったん得られると、学習機械はその表から有限オートマトンを合成し、教師に対して等価性質問をする。オートマトンが未知の正則言語を正

	$\lambda$	$a$	$b$
$\lambda$	1	0	0
$a$	0	1	0
$b$	0	0	1
$bb$	1	0	0
$ab$	0	0	0
$abb$	0	1	0
$aa$	1	0	0
$ba$	0	0	0
$bba$	0	1	0
$bbb$	0	0	1
$aba$	0	0	1
$abba$	1	0	0
$abbb$	0	0	0

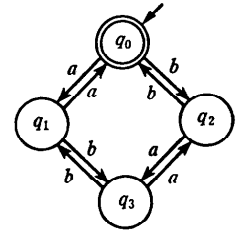


図-3 無矛盾で閉じた観測表と有限オートマトン

しく受理するものであれば、学習は終了し、学習機械は停止する。そうでない場合には、教師から返された反例とそのすべての接頭語を観測表に新たに加えて、上の手続きを再び繰り返す。

以上が Angluin によって与えられた正則言語に対する学習機械の大きな記述であるが、この優れた研究に触発されて、最近さまざまな学習対象が極小最適教師から効率良く厳密に学習可能であるという報告が数多く成されている。Angluin の学習機械は、有限オートマトンを学習する機械であると言い換えることもできる。有限オートマトンは有限の長さをもつ文字列を対象とし、その文字列の集合がもつ正則性を表現するシステムであるが、そこで用いられる基本的な考え方はより広い範囲の対象にまで広げることができる。たとえば、木構造を対象とするオートマトンとして木オートマトンが知られている。木オートマトンは、有限個のノードをもつ木を入力データとし、それらの木の集合がもつ正則性を表現するシステムである。(木オートマトンについては、文献 22) を参照。) この木オートマトンに対して、Angluin による学習機械を拡張することによって、木オートマトンによって受理される木の集合が極小最適教師から効率良く厳密に学習できることが示されている<sup>15), 16)</sup>。そしてその応用として、構文木の構造的記述に関する質問に答えることのできる極小最適教師から文脈自由言語のクラスが効率良く厳密に学習可能であることが示され<sup>16)</sup>、さらに論理型プログラムのある部分クラスが極小最適教師から効率良く厳密に学習可能であることが示されて

いる<sup>15)</sup>。文脈自由言語は一般に、文脈自由文法と呼ばれる形式的な文法によって定義される。ある文脈自由言語に属する文字列を、それを定義する文脈自由文法によって生成する過程を表したものが構文木と呼ばれるものであり、その構文木中の内部ノードに付けられたラベルを取り去った（または単一のラベルにすべて置き換えた）ものが構文木の構造的記述と呼ばれるものである。したがって構造的記述には、文字列の中に含まれるもの以上の文法に関する情報が詰まっており、構造的記述に関する質問に答えることのできる極小最適教師は、Angluin の正則言語の学習問題において定義された純粋な極小最適教師とは異なるものである（図-4）。また、木オートマトンに対する学習機械が、弱単項二階論理式と呼ばれる論理式のあるクラスを極小最適教師から効率良く厳密に学習するのに適応可能であることが示されている<sup>11)</sup>。

正則言語が極小最適教師から効率的に学習可能であることが判明した現在、次なる興味は、Chomsky の階層において一クラス上の言語である文脈自由言語のクラスが極小最適教師から効率的に学習可能であるか否かを調べることである。文脈自由言語が実用的な観点からも大変重要な言語のクラスであるということを考えると、これは現在最も重要な未解決の学習問題である。上の前者の研究は、この問題に対する部分的な解を与えるものである。もちろんこの問題を解くためには、構造的記述に関する質問に答えることのできる極小最適教師を仮定することはできない。肯定的解を得るには、極小最適教師から文脈自由言語のクラスを効率良く厳密に学習する学習機械が存在することを示せばよい。逆に否定的解を得るには、そのような学習機械が存在しないことを示せばよいが、これを示すための手掛りとして、最近、EXACT 学習モデルを計算量理論と関連付ける興味深い研究がなされた<sup>20)</sup>。

一方、文脈自由言語全体ではないが、そのある部分クラスであれば極小最適教師から効率良く厳密に学習可能であるという結果も幾つか報告され

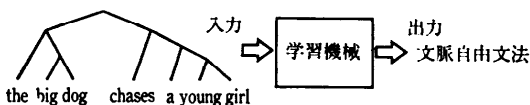


図-4 構造的記述からの文脈自由言語の学習

ている。One-counter 言語<sup>6)</sup>、Even linear 言語<sup>16)</sup>、そして単純決定性言語<sup>9)</sup>は皆、文脈自由言語の部分クラスであるが正則言語よりも真に大きなクラスであり、これらの言語のクラスは極小最適教師から効率良く厳密に学習可能であることが示されている。また括弧言語と呼ばれるやはり正則言語よりも真に大きい文脈自由言語の部分クラスが、極小最適教師から効率良く厳密に学習可能であることが、文献 16) の結果の簡単な系として導くことができる。これらの結果は、Angluin の正則言語に対する結果を拡張するものである。

さらに Chomsky の階層において一クラス上の言語である文脈依存言語の部分クラスで、文脈自由言語よりも真に大きなクラスの効率的な学習可能性に関する研究も幾つか成されている。形式言語を表現する（文脈自由文法などの）生成文法とは異なる表現システムとして、Smullyan による EFS (elementary formal system の略) がある。EFS は近年、形式言語の学習に対する統一的枠組を与える表現システムとして注目されており<sup>9)</sup>、また論理型プログラム (PROLOG など) との関係も深い<sup>21)</sup>。この EFS を形式言語の学習問題に対する表現システムとして用いることによって、文脈依存言語の部分クラスが極小最適教師とある特別な補助情報から効率的に厳密に学習可能であることが示されている<sup>14)</sup>。この結果は、効率的な学習のためには、学習対象を表現する表現システムの選択も非常に重要であるということを示している。

また最近、前に述べた木オートマトンに対する学習アルゴリズムを木変換機に対するアルゴリズムに拡張することにより（木の変換に関する質問も許す）、自然言語の構文を記述するために提案された語彙機能文法<sup>10)</sup>のある部分クラスが効率良く厳密に学習できることが示されている<sup>12)</sup>。語彙機能文法のこの部分クラスが定義する言語のクラスも、文脈依存言語の部分クラスでかつ文脈自由言語のクラスよりも真に大きい。したがって、そのような言語のクラスに対する、形式文法の表現を用いた学習アルゴリズムも得られたことになる。

今までは形式言語の学習問題に対する結果を主に説明してきたが、極小最適教師からはほかの学習対象も効率的に学習可能であることが示されて

いる。たとえば、ブール関数を定義するブール論理式のクラスである**単項 DNF 式**、 **$k$ -term DNF 式**、 **$k$ -clause CNF 式**は皆、極小最適教師から効率的に厳密に学習可能である<sup>2)</sup>。また  **$k$ -CNF 式** や  **$k$ -DNF 式**は、等価性質問だけから効率的に厳密に学習可能であることも示されている<sup>2)</sup>。

これまで所属性質問と等価性質問からの学習問題を考えてきたが、ほかの質問（たとえば、包含性質問など）やそれらとの組合せを用いた学習問題についても幾つかの結果が得られている。パターン言語と呼ばれる形式言語のあるクラスは、古くから学習対象として研究されてきた。（正確には、Gold による極限における同定の学習モデルにおいて、正の例だけから学習可能なクラスである。詳しくは本特集の解説 2. を参照。）いま  $\Sigma$  を定数記号の有限アルファベット、 $X$  を変数記号のアルファベットとする。パターンとは、 $\Sigma \cup X$  上の有限文字列である。パターン  $p$  の言語とは、 $p$  中の変数記号を、空でない定数文字列で置き換えて得られる、アルファベット  $\Sigma$  上のすべての文字列の集合である。同じ変数記号は同じ文字列で置き換える。たとえば、 $\Sigma = \{1, 2, \dots, 9\}$ 、 $X = \{x, y\}$ 、 $p = 12x5yy3x$  のとき、 $p$  の言語は、文字列 1205121230 や 128857777388 を含み、1253 や 122251031 を含まない。このパターン言語のクラスは、包含性質問だけを用いて効率的に厳密に学習可能である<sup>2)</sup>。しかもパターン言語を学習する学習機械は、包含性質問から返される反例を必要としない。一方、極小最適教師（さらに部分性質問をプラスしても）からはパターン言語を効率良く厳密に学習することはできないことが示されている<sup>2)</sup>。極小最適教師からではだめだが、ほかの質問を用いると効率良く学習可能なクラスが存在するという事は、大変興味深くまた驚くべきことである。

#### 4. 計算効率の下限を示す一般的手法と 否定的結果

前章では、さまざまな学習対象が効率的に学習可能であるという意味において肯定的な結果について説明してきたが、この章ではある質問だけしか利用できない場合、どんな学習機械も効率的に厳密に学習できない学習対象が存在するという否定的結果とそれを示すための一般的方法について

説明する。

まずはじめに、学習される概念のクラス  $C$  と学習機械が出力するプログラムのクラス（仮設空間） $P$  が定義する概念のクラスとは同じであり、すなわち  $C = \{c(p) | p \in P\}$  であり、そして質問に入力されるプログラムも  $P$  の要素であると仮定する。以下で示される学習機械の計算量の下限に関する結果は、この要請に強く依存している。

次のような簡単な学習対象を考える。  $2^n$  個の長さ  $n$  の 2 進文字列全体の集合を  $U$  とする。学習される概念は  $U$  の要素一つだけからなる集合とし、概念のクラス  $C$  はそれらの集合、すなわち  $C = \{\{x\} | x \in U\}$  とする。学習機械の計算量の下限（または否定的結果）を示す場合には、教師は学習機械からの質問に親切に答えるのではなく、学習機械が学習を達成するのをできるだけ遅らせようと、意地悪な反例をわざと選んだりするなどの振る舞いをするものと仮定して話を進める。このような教師は、教師というよりも学習機械にとっては敵と考えられるので、**敵対者**と呼ぶことにする。以下の敵対者は、等価性、所属性、部分性および非交差性質問を用いる任意の学習機械に対して、最悪の場合  $2^n - 1$  回の質問を行うことを余儀なくさせる。

敵対者は長さ  $n$  の 2 進文字列の集合  $S$  を保持している。初期段階で  $S$  は  $2^n$  個の長さ  $n$  の 2 進文字列全部を含んでいる。  $S$  が少なくとも二つの異なる文字列を含む間、敵対者は以下の操作を続ける。学習機械からの次の質問が文字列  $x$  に関する所属性質問の場合には *no* と答える。次の質問が単一要素集合  $\{x\}$  に関する等価性あるいは部分性質問の場合には *no* と答え、反例  $x$  を返す。次の質問が単一要素集合  $\{x\}$  に関する非交差性質問の場合には *yes* と答える。いずれの場合にも、 $x$  が  $S$  の要素であるならば、それを削除する。敵対者の応答が、 $S$  に残っている（任意の）一つの  $x$  に対して  $\{x\}$  を未知の概念としていることは容易に分かるだろう。また、各質問ごとに高々 1 個の要素が  $S$  から取り除かれる。したがって、正しい推測をするためには、アルゴリズムは少なくとも  $2^n - 1$  回の質問を行わなければならない。包含性質問が使える場合には、一回の質問で敵対者が隠している概念を暴露することができるだろう。

さてこの例にもあるように、概念のクラスによっては、ある種の質問しか利用できない場合、学習機械はしらみ潰し探索を余儀なくされる。ここでは、そのような概念のクラスの部分クラスに関する幾つかの性質を与えるために、上の手法を一般化する。以下の結果は、すべて Angluin<sup>2)</sup> による。

**補題 1** 概念のクラス  $C$  が、相異なる概念  $c_1, \dots, c_n$  と、相異なるインデックスの任意の対  $i, j$  に対して以下の条件を満足する概念  $c_n$  を含むと仮定する。(図-5 参照.)

$$c_i \cap c_j = c_n$$

このとき、制限された等価性質問と所属性および部分性質問を用いて、各概念  $c_i$  を厳密に学習する任意の学習機械は、最悪の場合、少なくとも  $N-1$  回の質問を行わなければならない。

この補題が、正則言語のクラスに適用できることを示そう。長さ  $n$  の 2 進文字列一つだけからなる集合は、 $n+2$  個の状態をもつ有限オートマトンによって認識できる。 $n+2$  状態有限オートマトンによって認識される正則言語のクラスは空集合を含むので、各  $c_i$  に対して 2 進文字列一つだけからなる集合を、 $c_n$  に対して空集合をそれぞれ対応させることによって、補題 1 の条件を満たすことができる。長さ  $n$  の 2 進文字列は全部で  $2^n$  個あるので、この正則言語のクラスを、制限された等価性、所属性、部分性質問を用いて厳密に学習する任意の学習機械は、最悪の場合、少なくとも  $2^n-1$  回の質問を行わなければならない。

同様の議論を展開することによって、補題 1 から、 $k$ -CNF 式と  $k$ -term DNF 式と単調 DNF 式に対しても、それを制限された等価性、所属性、部分性質問を用いて厳密に学習する任意の学習機械が、最悪の場合、指数オーダの回数の質問をしなければならないことを示すことができる。

“交わり集合”  $c_n$  自身が概念クラス  $C$  の要素でない場合には、上の結果を制限されていない等価性質問に対して成立するように強めることができる。

**補題 2** 概念のクラス  $C$  が、相異なる概念  $c_1, \dots, c_n$  を含み、また相異なるインデックスの任意の対  $i, j$  に対して以下の条件を満足する  $C$  に属さない集合  $c_n$  が存在すると仮定する。

$$c_i \cap c_j = c_n$$

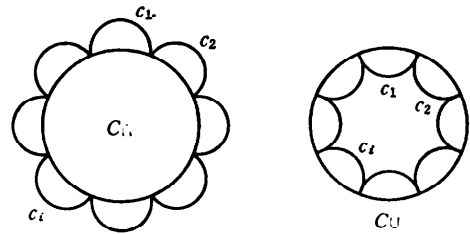


図-5 補題 1 と 3 が仮定する概念クラスのイメージ図

このとき、等価性、所属性および部分性質問を用いて、各概念  $c_i$  を厳密に学習する任意の学習機械は、最悪の場合、少なくとも  $N-1$  回の質問を行わなければならない。

補題 2 は、パターン言語のクラスを厳密に学習する学習機械が必要とする質問回数の指数的下限を得るために適用することができる。一方、等価性、所属性および包含性質問に関する、補題 1 および 2 と双対な、以下のような結果が知られる。

**補題 3** 概念のクラス  $C$  が、相異なる概念  $c_1, \dots, c_n$  と、相異なるインデックスの任意の対  $i, j$  に対して以下の条件を満足する概念  $c_U$  を含むと仮定する。(図-5 参照.)

$$c_i \cup c_j = c_U$$

このとき、制限された等価性、所属性および包含性質問を用いて、各概念  $c_i$  を厳密に学習する任意の学習機械は、最悪の場合、少なくとも  $N-1$  回の質問を行わなければならない。

この補題は、 $k$ -DNF 式と  $k$ -clause CNF 式に対して適用することができる。

**補題 4** 概念のクラス  $C$  が、相異なる概念  $c_1, \dots, c_n$  を含み、また相異なるインデックスの任意の対  $i, j$  に対して以下の条件を満足する  $C$  に属さない集合  $c_U$  が存在すると仮定する。

$$c_i \cup c_j = c_U$$

このとき、等価性、所属性および包含性質問を用いて、各概念  $c_i$  を厳密に学習する任意の学習機械は、最悪の場合、少なくとも  $N-1$  回の質問を行わなければならない。

上で述べたように補題 1 から、制限された等価性質問だけを用いる場合、どんな学習機械も正則言語のクラスを効率良く厳密に学習することはできないことを示すことができる。Angluin<sup>3),4)</sup> は上の議論をさらに発展させることによって、(制限されない)等価性質問だけからでは、どんな学習機械も効率良く学習することができないことを

示すための一般的手法を示している。ここでは詳しい議論は行わないが(詳しくは、文献 4)を参照)、Angluin は概念のクラスがもつ特徴として、**approximate fingerprint property** という性質を定義し、「概念のクラス  $C$  がこの property をもつならば、等価性質問だけを用いて  $C$  を効率的に厳密に学習する学習機械は存在しない」という定理を導いている。そして、この定理を用いることによって、正則言語のクラス、文脈自由言語のクラス、木オートマトンによって受理される木集合のクラス、CNF 式のクラス、DNF 式のクラス、単調 DNF 式のクラスなどが、等価性質問だけを用いては効率的に厳密に学習できないことを示している。

### 5. PAC 学習、極限における同定への 変換手法

以上述べてきた Angluin による EXACT 学習(質問からの概念学習)のモデルのほかに、代表的な学習モデルとして、本特集においても解説されている Valiant による PAC 学習モデル<sup>19)</sup>と Gold による極限における同定のモデル<sup>7)</sup>が知られている。一見、これらの三つの学習モデルは、お互いに全く独立していて何の関連性もないように思われるが、それぞれの学習モデルが誕生してきた歴史的背景をみても、それらは密接な関係にあり、各モデル間の比較やある学習モデルからほかの学習モデルへの学習機械の変換方法などについても活発に研究されている。ここでは PAC 学習モデルと極限における同定モデルについて簡単に定義を与え(各モデルの詳細については、本特集の解説 2. および 5. を参照)、本解説においてこれまでに示されてきた EXACT 学習モデルにおける学習機械をこれらの学習モデルに変換する一般的方法について考察する。

Valiant によって導入された PAC 学習モデル<sup>19)</sup>では、まず全体集合  $U$  上にある固定された確率分布  $D$  を仮定する。この確率分布  $D$  は任意に選ばれ、また学習機械にはこの  $D$  に関する情報は一切与えられない。学習機械は、入力を取らないサンプリングオラクル  $EX()$  と呼ばれるブラックボックスにアクセスすることができる。 $EX()$  が呼ばれると、分布  $D$  に従ってある要素  $x \in U$  が独立に取り出され、その要素  $x$  が未知の概念  $c_*$

の要素であるか否かを示す情報とともに学習機械に返される。学習の成功基準は二つの実数、正確さのパラメータ  $\epsilon$  と信頼性のパラメータ  $\delta$  に関して定義されている。ただし、 $0 < \epsilon \leq 1$ ,  $0 < \delta \leq 1$ 。二つの概念  $c_1$  と  $c_2$  の間の確率分布  $D$  に関する異なりぐあいを

$$d(c_1, c_2) = \sum_{x \in c_1 \oplus c_2} \text{Pr}_D(x)$$

で定義する。ここで、 $\text{Pr}_D(x)$  は確率分布  $D$  に従って要素  $x$  が取り出される確率を表す。 $EX()$  の一回の呼び出しにおいて、 $c_1$  および  $c_2$  のどちらか一方にだけ属している要素が得られる確率はちょうど  $d(c_1, c_2)$  である。学習機械が概念のクラス  $C$  を PAC (probably approximately correctly) 学習するとは、学習機械が常に停止して次の条件を満たすようなプログラム  $p$  を出力することを言う。

$$\Pr(d(c(p), c_*) \leq \epsilon) \geq 1 - \delta.$$

すなわち、 $\delta$  で保証された高い確率で、推測されたプログラム  $p$  が定義する概念と未知の概念との間には、 $\epsilon$  で保証される範囲内の違いしかないということである。またこのような学習機械  $A$  が効率良く学習するとは、 $A$  の計算時間が  $n, m, 1/\epsilon$  と  $1/\delta$  の多項式で押さえられることを言う。ここで、 $n$  は  $c_*$  の複雑さ、 $m$  は  $A$  の実行中に  $EX()$  から返された要素の最大の長さとする。

等価性質問を用いて概念のクラス  $C$  を厳密に学習する学習機械は、等価性質問の代わりに  $EX()$  を呼び出すことにより PAC 学習するように修正することができる。このアイデアでは、学習機械は  $p$  に関する等価性質問を行う代わりに、何回か  $EX()$  を呼び出し、 $EX()$  によって返された要素  $x$  とその分類が  $p$  と矛盾しないかどうか(すなわち、プログラム  $p$  が要素  $x$  を  $EX()$  と同じように分類するか否か)を調べる。矛盾が起こった場合には、学習機械は等価性質問によって  $no$  が返され、反例として  $x$  が与えられたものとして先に進む。 $p$  が与えられたすべての要素と無矛盾である場合には、等価性質問によって  $yes$  が返されたものとして処理する。

等価性質問に対して  $yes$  が返され、学習機械が停止してプログラム  $p$  が出力されたと仮定する。

$i$  番目の等価性質問において、学習機械が

$$q_i = \left\lceil \frac{1}{\epsilon} \left( \ln \frac{1}{\delta} + i \ln 2 \right) \right\rceil$$

回  $EX()$  を呼び出したときに、 $d(c(p), c_*) \geq \epsilon$  なる



プログラム  $p$  を出力する確率は高々  $(1-\varepsilon)^{q_i}$  である。したがって、学習の各ステージにおいて、学習機械が  $d(c(p), c_*) \geq \varepsilon$  なるプログラム  $p$  を出力する確率は、高々

$$\begin{aligned} \sum_{i=1}^{\infty} (1-\varepsilon)^{q_i} &\leq \sum_{i=1}^{\infty} e^{-\varepsilon q_i} \\ &\leq \sum_{i=1}^{\infty} \frac{\delta}{2^i} \\ &\leq \delta \end{aligned}$$

である。したがって、修正された学習機械は  $C$  を PAC 学習する。しかも等価性質問を用いて厳密に学習する学習機械が効率良く学習するならば、修正された学習機械も効率良く PAC 学習する。この結果から、極小最適教師から効率良く厳密に学習可能な概念のクラスはすべて、所属性質問を用いて効率良く PAC 学習可能であることが言える。

また極小最適教師は等価性質問に答えることのできる能力を仮定しているが、このような能力を仮定することには議論が分かれるところであろう。たとえば、正則言語に関しては、所属性決定問題と等価性決定問題はともに計算理論的に決定可能であり、その実行時間もオートマトンのサイズの多項式で押さえられるが、文脈自由言語の場合には、その等価性決定問題は計算理論的に決定不能であり、したがって文脈自由言語に対する極小最適教師を計算機上に実現することはできない。このような場合には、上で示したように、等価性質問の代用となるのは、サンプリングオラクル  $EX()$  を用いた確率的な等価性テストであろう。

それでは逆はどうであろうか？ すなわち、 $EX()$  オラクルの使用と PAC 学習の基準を、等価性質問の使用と厳密学習によって置き換えることができるのか？ もしも効率を保存しようとするならば、答えはノーである。それを示す一つの例が文献 2) に示されている。このように効率的な学習に関しては、EXACT 学習モデルと PAC 学習モデルの間には差がある。

次に、Gold によって導入された極限における同定の学習モデル<sup>7)</sup>との関係について考察する。極限における同定では学習を無限のプロセスとみなす。そこでは学習機械の究極あるいは極限における行動が学習の成功の基準として使われる。学

習機械に対しては、提示 (presentation) と呼ばれる例の無限列が与えられる。ここで例とは、全体集合  $U$  のある要素  $x$  と、それが未知の概念  $c_*$  の要素であるか否かを示す情報とからなる対を指す。また一つの提示において、全体集合  $U$  の各要素は少なくとも 1 回は現れると仮定される。それ以外の例の与え方・例を与える順序 (たとえば、同じ例が何回か繰り返して与えられるとか長さの短い順に与えられるとか) などについては任意である。学習機械は提示が与えられると、その提示中に次の新しい例が与えられるたびにプログラムを推測し、出力する。学習機械  $A$  が概念のクラス  $C$  を極限において同定 (identification in the limit) するとは、 $c_*$  についての任意の提示  $x_1, x_2, x_3, \dots$  に対して、 $A$  によって出力されるプログラムの無限列  $p_1, p_2, p_3, \dots$  が次の条件を満足することを言う：

$c(p_i) = c_*$  かつ  $p_i = p_{i+1} = p_{i+2} = \dots$   
を満足するある数  $i$  が存在する。

極限における同定のモデルにおいて効率的学習の基準を定義する場合、幾つかの議論と定義があるが、ここでは Pitt によって定義されたもの<sup>13)</sup>を紹介する。学習機械  $A$  が時点  $i$  において間接的な誤予測 (implicit error of prediction) をするとは、 $A$  によって時点  $i$  において出力されたプログラム  $p_i$  が  $i+1$  番目の例に対して正しくないときを言う。 $C$  を極限において同定する学習機械  $A$  が効率良く学習するとは、 $A$  による間接的な誤予測の回数が  $n$  の多項式で押さえられ、かつ  $i$  番目の例を受け取ってから  $i$  番目の推測であるプログラムを出力するまでの  $A$  の計算時間が、 $n$  と  $m_1 + m_2 + \dots + m_i$  の多項式で押さえられることを言う。ここで、 $n$  は  $c_*$  の複雑さ、各  $m_j (1 \leq j \leq i)$  は  $j$  番目の例の長さとする。

等価性質問を用いて概念のクラス  $C$  を厳密に学習する学習機械は、等価性質問の代わりに提示が与えられることによって極限において同定するように修正することができる。しかし等価性質問を用いて効率良く厳密に学習する学習機械をすべて、効率良く極限において同定する学習機械に修正できるわけではない。等価性質問を用いて効率良く厳密に学習する学習機械では、その推測を変更する回数が  $n$  と反例の長さの多項式で表されるかもしれない。この意味において、EXACT 学習

モデルは極限における同定モデル（というよりは、Pitt の定義）よりも効率的学習の基準に関する制限が緩い。したがって少なくとも以下のことが言える。等価性質問を用いて概念のクラス  $C$  を効率良く厳密に学習する学習機械が存在しないならば、 $C$  を効率良く極限において同定する学習機械も存在しない。この結果から、正則言語のクラス、文脈自由言語のクラス、CNF 式のクラスなどを効率良く極限において同定することはできないことが示される。

## 6. おわりに

人工知能の分野における機械学習の研究では、ややもすると問題が ill-defined のまま扱われ、そこで示される学習メカニズムや結果に対する評価も経験的なものによることが多い。本解説で示した手法に基づく研究では、教師と学習者のパラダイムを形式的に自然に定義することによって、帰納推論による学習という機械学習の一つの問題を well-defined な問題に導き、何が学習者の効率的な学習のために必要であるかを理論的に導き出すことに成功している。機械学習の研究をしっかりとした理論的基盤の上で健全に行っていくためには、本解説で示されたアプローチは有効であり、また大変重要である。

一方、形式言語理論がコンパイラなどの実際のシステムを構築するのに役だったように、EXACT 学習研究が実際的に有効であることを示すための応用分野を開拓することが望まれている。EXACT 学習の応用に関しては、幾つかの最近の例が解説、本号別稿に紹介されている。

最後に、本解説が、多くの研究者がこの分野に興味をもち、さらに、この方面の優れた研究を行う上での一助になれば幸いである。文献 2) には多くの未解決問題が示されている。

## 参 考 文 献

- 1) Angluin, D.: Learning Regular Sets from Queries and Counter-Examples, *Information and Computation*, 75, pp. 87-106 (1987).
- 2) Angluin, D.: Queries and Concept Learning, *Machine Learning*, 2, pp. 319-342 (1988).
- 3) Angluin, D.: Equivalence Queries and Approximate Fingerprints. In *Proceedings of 2nd Workshop on Computational Learning Theory*, pp. 134-145, Morgan Kaufmann (1989).
- 4) Angluin, D.: Negative Results for Equivalence Queries, *Machine Learning*, 5, pp. 121-150 (1990).
- 5) Arikawa, S., Shinohara, T. and Yamamoto, A.: Elementary Formal Systems as a Unifying Framework for Language Learning. In *Proceedings of 2nd Workshop on Computational Learning Theory*, pp. 312-327, Morgan Kaufmann (1989).
- 6) Berman, P. and Roos, R.: Learning One-Counter Languages in Polynomial Time. In *Proceedings of 28th IEEE Symposium on Foundations of Computer Science*, pp. 61-67, IEEE Computer Society Press (1987).
- 7) Gold, E. M.: Language Identification in the Limit, *Information and Control*, 10, pp. 447-474 (1967).
- 8) Hopcroft, J. E. and Ullman, J. D.: *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley (1979). (野崎, 高橋, 町田, 山崎共訳: オートマトン言語理論計算論 I, II, サイエンス社).
- 9) Ishizaka, H.: Polynomial Time Learnability of Simple Deterministic Languages, *Machine Learning*, 5, pp. 151-164 (1990).
- 10) Kaplan, R. M. and Bresnan, J.: Lexical Functional Grammar: A Formal System for Grammatical Representation. In Bresnan, J., editor, *The Mental Representation of Grammatical Relations*, MIT Press (1982).
- 11) Nishino, T.: Learning Logic Formulas and Programs Based on Their Models, Research Report TDU-IS-19, Department of Information Science, Tokyo Denki University (1990).
- 12) Nishino, T.: Mathematical Analysis of Lexical-Functional Grammars—Complexity, Parsability, and Learnability—. In *Proceedings of Seoul International Conference on Natural Language Processing* (1990).
- 13) Pitt, L.: Inductive Inference, DFAs, and Computational Complexity. In *Proceedings of AII-89 Workshop on Analogical and Inductive Inference (Lecture Notes in Computer Science, 397)*, pp. 18-44, Springer-Verlag (1989).
- 14) Sakakibara, Y.: On Learning Smullyan's Elementary Formal Systems: Towards an Efficient Learning for Context-Sensitive Languages, *Advances in Software Science and Technology*, 2, pp. 79-101 (1990).
- 15) Sakakibara, Y.: Inductive Inference of Logic Programs Based on Algebraic Semantics, *New Generation Computing*, 7, pp. 365-380 (1990).
- 16) Sakakibara, Y.: Learning Context-Free Grammars from Structural Data in Polynomial Time, *Theoretical Computer Science*, 76, pp. 223-242 (1990).
- 17) Sloan, R. H.: *Computational Learning Theory: New Models and Algorithms*, PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Tech-

- nology (1989), Technical Report MIT/LCS/TR-448.
- 18) Takada, Y.: Grammatical Inference for Even Linear Languages Based on Control Sets, *Information Processing Letters*, 28, pp. 193-199 (1988).
- 19) Valiant, L. G.: A Theory of the Learnable, *Communications of the ACM*, 27, pp. 1134-1142 (1984).
- 20) Watanabe, O.: A Formal Study of Learning via Queries. In *Proceedings of 17th International Colloquium on Automata, Languages and Programming (Lecture Notes in Computer Science, 443)*, pp. 139-152, Springer-Verlag (1990).
- 21) Yamamoto, A.: Elementary Formal Systems as a Logic Programming, Technical Report 12, RIFIS, Kyushu University (1989).
- 22) 小林孝次郎, 高橋正子: オートマトンの理論 (電子計算機基礎講座 11) 共立出版 (1983).  
(平成2年10月12日受付)



榊原 康文 (正会員)

1983年東京工業大学理学部情報科学科卒業。1985年同大学院理工学研究科情報科学専攻修士課程修了。同年富士通(株)(現在、(株)富士通研究所)国際情報社会科学研究所入所、現在に至る。計算論的学習理論、形式言語理論、論理型プログラムなどの究に従事。EATCS 日本ソフトウェア科学会、人工知能学会、LA 学会各会員。



西野 哲朗 (正会員)

昭和34年生。昭和57年早稲田大学理工学部数学科卒業。昭和59年同大学院理工学研究科博士前期課程修了。同年日本アイ・ビー・エム(株)入社。昭和62年東京電機大学理工学部情報科学科助手。この間、属性文法、自然言語処理、計算論的学習理論の研究に従事。オートマトンと形式言語理論、計算量理論に興味をもっている。電子情報通信学会、日本ソフトウェア科学会、CAI 学会各会員。

