3F2-01

帰納論理プログラミングにおける背景知識の能動的学習法

Active Learning Method of Background Knowledge in Inductive Logic Programming

白井 秀範　松井 藤五郎　大和田 勇人
Hidenori Shirai　Tohgoroh Matsui　Hayato Ohwada

東京理科大学 理工学部 経営工学科
Department of Industrial Administration,Faculty of Science and Technology,Tokyo University of Science.

In this research, we propose a method of realizing active learning in Inductive Logic Programming (ILP). If the data is increased, although hypothesis will improve, the learning time also increases as the number of the data increases. Therefore, how to acquire high accuracy by a small number of data is an important research theme. This research solves it by selecting data required for acquiring high accuracy. Since the quantity of Backgroud knowledge data has serious influence for the accuracy of Hypothesis in ILP, it is carried out only for Background knowledge data. The selecting samples data is a view of Active Learning, and we call this method Active Inductive Learning Method (AILM). AILM obtains the high accuracy hypothesis under smaller number of data. In this paper, we describe the definite procedure and implementation of AILM as active learning.

1. はじめに

近年，機械学習の分野において，論理プログラムを用いた概念学習である帰納論理プログラミング (ILP)[Muggleton 99] は広い領域で盛んに研究が行われている．

ルールの精度を上げるためには，より大量のデータを元に学習すれば良いが，データが増えれば増えるほど，学習に要する時間はデータ数に比例して増加してしまう．したがって，“ いかに少ないデータで良い精度を得るか ”は重要な研究テーマである．本研究では，それを“ 良い精度を得るために最低限必要なデータを選び出す ”という方向で解決する．このデータの選択的抽出は能動的学習 [Cohn 94] の考え方であり，本研究ではその手法として能動的帰納学習法を提案する．また，ILP では背景知識のデータ量が仮説の精度に大きな影響を与えるので，本研究では背景知識にのみ能動的学習法を適用する．

従来の ILP では帰納推論を用いて仮説を求めるのみであったが，本研究では能動的学習を行うために発想推論 [Kakas 98] を用いている．それは，必要な背景知識を特定し初期の背景知識集合に追加するためであり，そうすることで，本来は高い精度でありながら背景知識が欠けているために選ばれなかった仮説を求めることが可能となる．

関連研究として，ILP 上で発想推論を行う手法に TCIE[Muggleton 00] がある．しかし，それはあくまで帰納推論で得られない知識を求めることが目的であるため，初期の背景知識中に発想推論で用いるルールが必要である．それに対し，本手法では発想推論に用いるルールを目標仮説として生成するため，初期条件が緩く，TCIE よりも多くの知識を求めることが可能である．また，本研究と同様に ILP 上で能動的学習を行う手法に CLML[Bryant 99] があるが，これは概念や応用分野についての提唱が主であり，能動的学習の根幹である知識の特定方法が示されていなかった．そこで，本手法では能動的学習法の手順を定式化にするだけでなく，知識の特定方法についても特に詳細に定義している．

以下，2 章で能動的帰納学習法の手順とアルゴリズムを示す．3 章では，本手法を実装したシステムである FALS を用いて評価実験を行い，有効性を示す．続いて 4 章で，関連研究と比較したあと，最後に 5 章で結論を述べる．

A: 白井秀範，東京理科大学 理工学部 経営工学科 大和田研究室，千葉県野田市山崎 2641，04(7124)1501，shirai@ia.noda.tus.ac.jp

2. 能動的帰納学習法

2.1 手順の概要

能動的帰納学習法 (Active Inductive Learning Method, AILM) は次の手順で行われる．

Step1　背景知識と事例を用意する
Step2　帰納推論を用いて仮説を生成する
Step3　探索結果から目標仮説を作成する
Step4　発想推論により目標背景知識を求める
Step5　目標背景知識の一部を質問する
Step6　真である知識を背景知識に追加する．
Step2 へ戻る

これらの手順を図に表したものが，図 1 である．

Step3 では Step2 で得られた仮説を求める際の探索結果を用いて，その仮説よりも精度が高いと予想される目標仮説を作成する．Step4 では，Step3 で求めた目標仮説と事例を用いて発想推論を行い目標背景知識を求める．

目標背景知識が元の背景知識集合に追加されれば，再び Step2 が実行されたときに目標仮説が得られる可能性は高い．しかし，目標背景知識は発想推論により得られたものであるから真である保証は無い．そこで，Step5 ではそれらの真偽を質問している．

2.2 具体例

簡単な例を用いて，能動的帰納学習法の具体例を示す。

Step1 では，初期条件として背景知識 $\mathcal{B}$ と正事例 E^+，負事例 E^- を以下のように与える．また，これらの集合を表記するための記号“ {} ”は今回省略している．

$$\begin{aligned}\mathcal{B} &= p(a) \leftarrow,\ \ p(b) \leftarrow,\ \ p(d) \leftarrow,\ \ q(a) \leftarrow,\\ &\quad\ q(c) \leftarrow,\ \ q(d) \leftarrow,\ \ r(b) \leftarrow,\ \ r(c) \leftarrow\\ E^+ &= s(a) \leftarrow,\ \ s(b) \leftarrow\\ E^- &= s(c) \leftarrow,\ \ s(d) \leftarrow\end{aligned}$$

Step2 では，これらを用いて帰納推論を実行し仮説 H を得る．このときの誤差許容率を 0.35 とする．

$$H = s(X) \leftarrow p(X) \quad \cdots (\dagger)$$

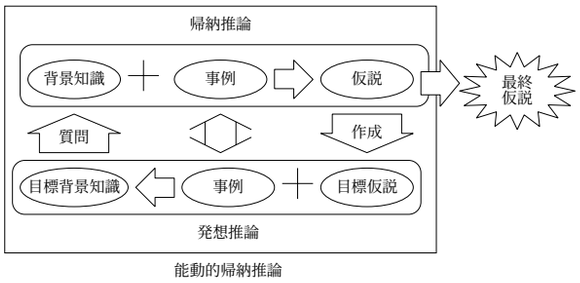

図 1: 能動的帰納学習法

Step3 では，Step2 で得られた仮説以外で MDL が最大の仮説を選択し，目標仮説とする．この例では事例数が少ないため，便宜上，仮説のリテラル数を無視している．よって，目標仮説は以下の仮説となる．

$$h\_target \ = s(X) \leftarrow p(X),\ r(X)$$

Step4 では，事例と目標仮説で発想推論をすることで、目標背景知識 B_target を求める．このとき，目標仮説は近似式 h_target' を用いる．近似式は，正事例 E^+ を一般化して仮説の帰結とし，p_target を仮説の条件部としたものである．ここで，p_target とは，目標仮説の条件部に来る述語が背景知識中に存在する数が一番少ない述語である．よって，この例の場合 p_target は以下のようになり，目標背景知識の近似式 h_target'，及び，目標背景知識 B_target は以下のようになる．

$$\begin{aligned} &p\_target \ \ = r(X) \leftarrow \\ &h\_target\text{'} = s(X) \leftarrow r(X) \\ &B\_target \ \ = r(a) \leftarrow,\ \ r(b) \leftarrow,\ \ r(c) \leftarrow,\ \ r(d) \leftarrow \end{aligned}$$

Step5 では，Step4 で求めた目標背景知識をもとに質問 Q をする．質問は，目標背景知識からそのときの背景知識を取り除いたものとする，したがって Q は以下のようになる．また，その質問の中で真実の知識を Q_true と表し，今回は以下のように仮定する．

$$\begin{aligned} Q \ \ \ \ \ \ &= r(a) \leftarrow,\ \ r(d) \leftarrow \\ Q\_true &= r(a) \leftarrow \end{aligned}$$

Step6 では，質問した知識 Q の中で真実の知識 Q_true だけを背景知識に追加する．その結果，最終的に背景知識 $\mathcal{B}$ は以下のようになる．

$$\begin{aligned} \mathcal{B} \ = &p(a) \leftarrow,\ \ p(b) \leftarrow,\ \ p(d) \leftarrow,\ \ q(a) \leftarrow,\ \ q(c) \leftarrow \\ &q(d) \leftarrow,\ \ r(a) \leftarrow,\ \ r(b) \leftarrow,\ \ r(c) \leftarrow \end{aligned}$$

Step6 の後は，Step2 へ戻り追加後の背景知識を用いて，再び帰納推論を実行する．許容誤差を 0.3 とすることで以下の仮説 H を得る．

$$H \ \ = s(X) \leftarrow p(X),\ q(X)$$

最終的に得られた仮説は，被覆する正事例が 2 で被覆する負事例が無く，精度は 100%である．それに対し，前に得られた仮説 (†) は，被覆する正事例は 2 であるが被覆する負事例が 1 であり，その精度は 66%である．よって，背景知識を追加する前に較べ，仮説の精度は 51%向上した．

この例では許容誤差を変化させているが，この例の背景知識数が極端に少ないために便宜上そのようにしている．通常の学習の場合には，許容誤差を変える必要はない．

2.3 アルゴリズム

この章では，先程示した能動的帰納学習法の手順について詳しく説明する．

まず，Step1 は初期条件として背景知識 $\mathcal{B}$ と事例 E を与える．次に Step2 では，以下の論理的設定を満たす仮説 H を逆判意法により求める．ここまでは従来の帰納推論手法と同じである．

$$\begin{cases} \mathcal{B} \cup H \models E^+ \\ \mathcal{B} \cup H \cup E^- \not\models \square \end{cases}$$

上記の式を満たす仮説 H は複数存在する．しかし，代表的な帰納推論システム：PROGOL[Muggleton 95] では最良の仮説のみを出力するために，記述長最小原理に基づいた以下の評価値 $MDL(h)$[古川 01] を用いている．

$$MDL(h) = p - n - c \quad (\ h \in \phi(h)\ )$$

ここで，p は仮説 h が被覆する正事例数，n は仮説 h が被覆する負事例数，c は仮説 h の条件部分のリテラル数である．また，$\phi(h)$ は仮説 h の探索空間を表している．この $MDL(h)$ が表す意味は，“ より多くの正事例のみを説明して負事例を説明せず，さらに仮説のリテラル数が少ない仮説が良い ”ということである．

Step3 からが本研究の本質となる能動的学習法の部分である．能動的帰納学習法は目標仮説を実際の仮説として実現することで，仮説の精度を向上させる方法である．目標仮説 h_target の定義は，Step2 で求めた仮説を H としたとき，以下の式で表される．

$$\begin{aligned} h' \ \ \ \ \ \ \ \ &= argmax_{h \in H} MDL(h) \\ h\_target &= argmax_{h \in \phi(h') - \{h'\}} MDL(h) \end{aligned}$$

ここで，$argmax MDL(h)$ とは $MDL(h)$ が最大値をとるような h を表している．この定義は二つのステップからなる．まず，Step2 で得られた仮説の中から MDL が最大となる一つの仮説 h' を選び，次にその仮説 h' を求めた探索空間 $\phi(h')$ の中から MDL が h' の次に最大となる仮説 h_target を求める．この仮説 h_target を本研究では目標仮説と呼んでいる．

次に Step4 では，目標仮説と Step1 で用いた事例を用いて発想推論を行っている．発想推論の論理的設定 [古川 01] は以下のようになる．O,T をそれぞれ観測された事実の集合，利用可能な既知の知識集合とする．また，$H = \{h_1, h_2, ..., h_n\}$ をあらかじめ用意された可能な仮説集合とする．そして，$T \not\models O$ のときに，

$$T \cup H\text{'} \models O,\ \ H\text{'} \subseteq H$$

の条件を満たす H' を H から選択する．能動的帰納学習法では，上記の式 H'，O，T をそれぞれ ILP の基本要素である仮説 H，事例 E，背景知識 $\mathcal{B}$ に対応させている．

発想推論を行う理由は，目標仮説を実現するような背景知識を求め，その真偽を確認するためである．しかし，目標仮説の条件部に含まれるリテラルは多数存在するため，目標仮説をそのまま用いて発想推論すると，複数の述語を横断的に質問してしまうことになる．なので，実際にはある特定の述語に着目して追加していった方が効率的であり，本手法ではその問題を解決するために目標仮説の条件部から目標述語を特定している．

目標述語 P_target の定義式は，$Num(p)$ を p が背景知識中に存在する個数としたとき，以下の式で表現出来る．ここでは，仮説 H 中に存在する述語を $Pred\{H\}$ と表記している．

$$P\_target = argmin_{p \in Pred\{H\_target\}} Num(p)$$

この式は，背景知識集合 $\mathcal{B}$ を目標仮説 H_target に含まれる述語 $Pred\{H\_target\}$ ごとに数えたときに一番少ない述語を目標述語 P_target とするという意味である．この目標述語を用いて目標仮説を近似することにより，目標仮説の役割を果たしつつ，特定の述語のみを質問することが可能となる．よって，この目標述語 P_target を用いた目標仮説の近似式 H_target' の定義式は以下のようになる．ここで，ILP で用いられる目標概念 (Target concept) を表す述語を T と表記する．

$$H\_target' = T(X, Y, \cdots, Z) \leftarrow P\_target(W, V, \cdots, U)$$

ここで，P_target の変数は目標仮説 H_target に含まれていたときの状態であるとする．したがって，Step4 で発想推論を行うことにより得られる目標背景知識 B_target は以下の式で表される．

$$B\_target = \{\ P\_target\ \theta \leftarrow \mid \forall e \in E\ [\ e = T\theta\ ]\ \}$$

この式は，目標背景知識 B_target が目標述語 P_target の変数に事例 E の定数部分 θ を代入することで得られることを示している．

このように得られた目標背景知識は本手法で自動的に作り出した知識であるため，その事実が成り立つ保障はない．よって Step5 ではその知識の真偽を質問という形式で調べている．また，質問する際に既に背景知識中に存在するものについては質問する必要がないので，目標背景知識から取り除いている．したがって，質問 Q の定義は以下のように表現できる．また，その質問により真であると確定した知識を Q_true と表現することにする．ここで，$\mathcal{B}$ は初期状態の背景知識を表している．

$$\begin{aligned} Q &= B\_target \setminus \mathcal{B} \\ Q\_true &= \{\ q \mid q \in Q,\ q = TRUE\ \} \end{aligned}$$

最後に，Step6 として先程定義した Q_true について背景知識集合 $\mathcal{B}$ に追加する．以下に背景知識追加の定義を示す．ここで，“ := ” の記号は代入を表している．

$$\mathcal{B} := \mathcal{B} \cup Q\_true$$

この追加は，その新しい背景知識集合を用いて帰納推論を行ったときに目標仮説が得られるようにするためであり，それにより仮説の精度を向上させることが可能となる．本手法はループ構造になっており，Step6 の後は再び Step2 から実行することになる．このループの終了条件は，ユーザが Step2 で精度の良い仮説を得た場合か，または Step5 で質問が無くなったときである．

2.4 能動的帰納学習法の利点

能動的帰納学習法では，背景知識の欠けている部分を効率的に埋めることにより仮説の精度を向上させる方法である．どの知識から追加するべきかを考える際に，本研究では目標仮説を用いている．それは，目標仮説は現在の仮説よりも精度が高い可能性があると考えるからである．この章では，なぜ目標仮説に注目したのかについて述べる．

2.3 にも示したように，代表的な帰納推論システムである PROGOL では MDL の値によって最適な仮説を探索している．つまり，出力される仮説は ILP の論理設定を満たす複数の仮説候補の中から，最も MDL 値の高い仮説である．しかしながら，そのように得られた仮説は，初期条件として与えたデータ集合に対しては最も良い仮説であるが，データ集合全体に対しても良いとは限らない．つまり，MDL 値がある程度高い仮説については，実際にはどれも適切に値する仮説なのである．

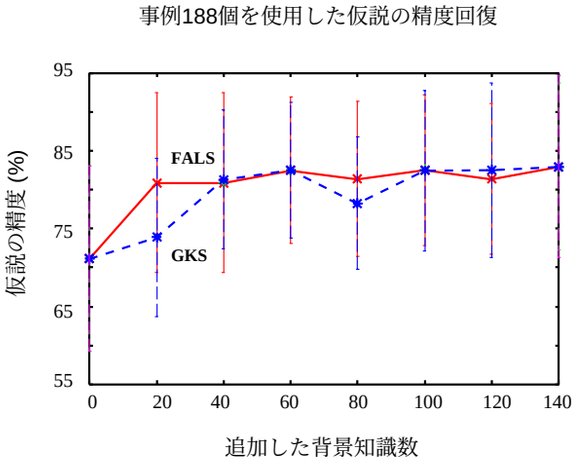

図 2: 背景知識の追加に対する仮説の精度向上

本手法では，MDL 値が二番目の仮説を目標仮説に設定している．それは，現在出力されている仮説が外部データに対して精度が低い場合，その仮説は初期状態のデータにしか適用されないような特殊な仮説であると考えられる．その場合，他に高い精度を出せそうな仮説を考えると，MDL 値が高いが一番でないために出力されなかった仮説となる．なぜなら，これらの仮説は本当は高い精度を実現する仮説であるにもかかわらず，背景知識が一部欠けているために MDL 値が低くなっている場合が考えられるからである．よって，本研究ではこのような仮説に焦点を当て，欠けている背景知識を追加することにより，本当に精度の良い仮説を導き出すことを目的としている．

3. 評価実験

能動的帰納学習システム FALS を実装し，実験を行った．FALS とは，First order theory Active Learning System の略であり，一階述語論理に基づく能動的学習システムという意味である．評価実験の目的は，能動的帰納学習法を実装したシステムである FALS が従来の帰納推 GKS[Mizoguchi 95] より，精度の高い仮説が得られるかどうかを検証することである．

FALS は現段階では実験段階にあり，能動的帰納学習法を忠実に再現したものではない．しかしながら，今回の実験データに関しては，本手法で追加する知識と FALS が追加する知識は一致しており，評価実験としては有効であると考えられる．このシステムの入力データは従来の帰納推論システム PROGOL と同じく，一階述語論理言語 Prolog[Lloyd 84] で記述している．また，FALS の比較対象として用いている GKS とは，PROGOL とほぼ同じアルゴリズムを用いた ILP システムであり，仮想の PROGOL と考えても差し支えない．

次に実験方法の手順について述べる．まず背景知識に含まれる事実をランダムに削除する．次に，FALS は提案された知識を追加するという方法で，GKS はランダムに背景知識を追加する．そのあと，その増加した背景知識と事例を用いて，それぞれ再学習し仮説を生成する．最後に，その得られた仮説を検定し精度を比較する．この操作を，追加する背景知識数を変えて繰り返し実行する．

実験データには ILP システムのベンチマークである

Mutagenesis[Debnath 91] を用いた．FALS では，質問に対する答えは，Yes か No の二択で行うことを前提としていた．しかし，今回の場合は，背景知識の数値が連続値であることを考慮し，質問として直接数値を答える形式にしている．つまり，背景知識の数値を質問し，その値を背景知識の定数としてそのまま追加する方法を採用している．

また，時間短縮の観点から，初期状態として背景知識から atm と bond を削除しているが，事例に関しては正事例 125 個，負事例 63 個，全て用いている．背景知識の述語には，lumo，logp，gteq，lteq の 4 種類があり，そのうち lumo と logp は事実として与えられ，gteq，lteq はルールとして与えられている．よって，今回の実験で変化させている背景知識は，lumo と logp の二つになる．初期の背景知識数はそれぞれ logp が 10 個，lumo が 100 個から始めている．

その結果を図 3. に示す．FALS の利点は，同じ精度を実現するために必要な背景知識が従来の帰納推論法に比べ少なく済むことであり，図 3. からもそれが分かる．ここで，仮説の精度には Cross-Validation 法による 10 分割を用いており，点線はそのときの精度の標準偏差を示したものである．この図より背景知識を 20 個追加したとき，GKS では 73%であるのに対し，FALS では 82%を達成している．よって，能動的帰納学習法により背景知識を追加した方が，ランダムに追加するよりも効率的であるということが出来る．

4. 関連研究

本章では，本研究に対する関連研究を取り上げる．

能動的学習法を ILP 上で実現する手法には，本研究の他に Bryant らが考案した CLML[Bryant 99] がある．CLML の目的は，仮説の精度を向上させることが目的である．その手順は，まず学習を行い仮説を生成する．次に能動的学習法により選ばれた実験をロボットが自動的に行い，その結果を分析する．最後にそのそのデータを基に再学習を行い，再び仮説を生成する．それらの手順を仮説の精度が向上するまで繰り返すというものである．実験の選定基準としては，仮説の精度向上につながるようなものや実験コストが少ないものを優先するとしているが，どのように実験を選んでいるのかといった，能動的学習の知識特定部分については具体的な記述がなかった．それに対して，本研究では，能動的学習法の手順を全て定式化し，知識の特定部分を含むそれらのアルゴリズムを詳しく示している．

次に，本手法では知識を特定する際に発想推論を用いているが，同じく ILP 上で発想推論を行う手法として，Muggleton らが考案した TCIE[Muggleton 00] がある．この手法の目的は，概念として提唱されていた発想推論を実在する ILP システムである PROGOL で実現することである．実際に PROGOL5.0 では発想推論を行うことが出来るようになったが，PROGOL をそのまま能動的学習法に使用するには多くの問題点がある．

まず，PROGOL5.0 では背景知識のルールと事例を用いて発想推論を行っているため，与えられた背景知識中にルールが存在しなければ発想推論を行うことが出来ない．また，発想推論により求めたい知識の述語を，前もって modeh 宣言しなければならない．つまり，始めから求められ知識をユーザ側が予想しておく必要がある．しかし，現実にはそれを予想するのは困難であり，さらに全ての述語に対し modeh 宣言を行ったとしても，得られる知識が一度に複数の述語に渡る可能性がある．したがって，初期の背景知識中にルールを必要とせず，さらに求める知識に対してその都度最適な述語を一つ自動的に選択する本手法の方が，能動的学習に適していると考えられる．

5. まとめ

本稿では，新しい ILP 手法である能動的帰納学習法を提案し，その実装である ILP システム FALS を用いてその有効性を検証した．能動的帰納学習法は，従来の帰納推論手法に背景知識の能動的学習を付け加えた手法である．それは，仮説の精度を向上させるために必要な背景知識から追加することで，効率的な学習を可能とするものであり，評価実験からもそのことが確認できた．

参考文献

[Bryant 99] Bryant, C., Muggleton, S., Page, C., and Sternberg, M.: Combining active learning with inductive logic programming to close the loop in machine learning (1999).

[Cohn 94] Cohn, D. A., Atlas, L., and Ladner, R. E.: Improving Generalization with Active Learning, *Machine Learning*, Vol. 15, No. 2, pp. 201–221 (1994).

[Debnath 91] Debnath, A., G. Schusterman and Hansch, C.: Structure-activity relationship of mutagenic aromatic and heteroaromatic nitro compounds. Correlation with molecular orbital energies and hydrophobicity., in *Journal of Medicinal Chemistry*, Vol. 34, pp. 786–797 (1991).

[古川 01] 古川康一：帰納論理プログラミング (2001).

[Kakas 98] Kakas, A., Kowalski, R., and Toni, F.: The role of abduction in logic programming (1998).

[Lloyd 84] Lloyd, J. W.: Foundations of Logic Programming (1984).

[Mizoguchi 95] Mizoguchi, and Ohwada, : Using Inductive Logic Programming for Constraint Acquisition in Constraint-based Problem Solving., *Proc.of the 5th International Workshop on ILP*, pp. 297–322 (1995).

[Muggleton 95] Muggleton, S.: Inverse Entailment and Progol, *New Generation Computing, Special issue on Inductive Logic Programming*, Vol. 13, No. 3-4, pp. 245–286 (1995).

[Muggleton 99] Muggleton, S.: Inductive Logic Programming, in *The MIT Encyclopedia of the Cognitive Sciences (MITECS)*, MIT Press (1999).

[Muggleton 00] Muggleton, S. H. and Bryant, C. H.: Theory Completion Using Inverse Entailment, *Lecture Notes in Computer Science*, Vol. 1866, pp. 130–?? (2000).