

帰納論理プログラミングにおける能動的背景知識選択法の論理的再構築

Logical restructuring Active Selecting Method of Background Knowledge in ILP

辻 裕貴 松井 藤五郎 大和田 勇人
Yuki Tsuji Tohogoroh Matsui Hayato Ohwada

東京理科大学大学院 理工学研究科 経営工学専攻

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 Background knowledge data has serious influence for the accuracy of Hypothesis in ILP, it is carried out only for Background knowledge data.

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、これに対し“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。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能動的選択法を適用する。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。

従来研究として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 [白井 03] がある。しかしこの手法は発想推論を行う際に、仮説に現れる述語を制限している。これに対し本論文では発想推論時に制限せず、帰納推論により得られたすべての仮説に対し発想推論を行う事を提案する。このように本手法では発想推論時に制限を行わないため、より論理的な枠組みで帰納推論と発想推論を組み合わせることができる。

2. 能動的背景知識選択法

本章では帰納論理プログラミングに能動的学習を取り入れた能動的背景知識選択法を説明する。

2.1 従来手法

従来の能動的背景知識選択法の概要を以下に示す。

1. 与えられた背景知識と事例を用いて仮説を帰納推論により求める。
2. 得られた仮説から目標仮説を生成する。

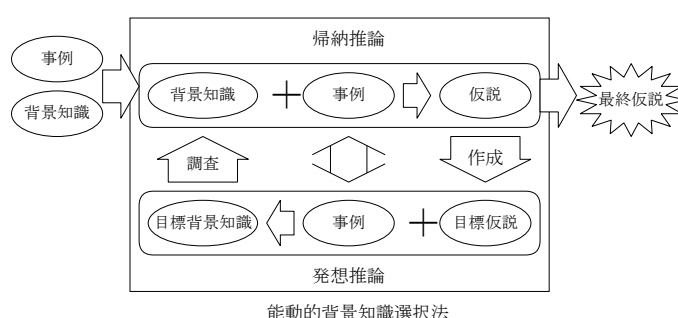


図 1: 能動的背景知識選択法の概要

3. その目標仮説と与えられた事例を用いて発想推論を行い、事実の集合である目標背景知識を求める。
4. その目標背景知識集合内のそれぞれの事実を調査する。
5. 事実調査の結果を背景知識に追加する。
6. 手順 1 に戻る。
これらの手順を手順 2 で目標仮説が生成できなくなるまで繰り返す。

図 1 は以上の手順を示している。

2.2 問題点

手順 2 において目標仮説を生成する際の基準というものに論理的な根拠がなく、非常にヒューリスティックであるということであるという事は、帰納推論と発想推論を組み合わせた論理体系において問題がある。調査すべき事実の選定は枠組みの最終段階、つまり帰納推論で得られた全仮説に対し発想推論を行った後に行うべきである。

3. 能動的背景知識選択法の再構築

本章では従来手法の問題点を解決できるよう、能動的背景知識選択法の再構築を提案する。

3.1 提案手法

提案手法の概要は以下の通りである。

連絡先: 辻裕貴, 東京理科大学大学院理工学研究科経営工学専攻大和田研究室, 千葉県野田市山崎 2641, 04(7124)1501, j7405626@ed.noda.tus.ac.jp

1. 与えられた背景知識と事例を用いて仮説を帰納推論により求める。
2. 得られた仮説と与えられた事例を用いて発想推論を行い、事実の集合である目標背景知識群を求める。
3. その中で調査する背景知識を選択する。
4. 事実調査の結果を背景知識に追加する。
5. 手順 1 に戻る。
一定以上の精度が得られるか、手順 2 で得られた背景知識群がすべて調査済みになるまで続ける。

表 1: 背景知識と事例

| B | | | \mathcal{E}^+ | \mathcal{E}^- |
|-----------|----------------------------------|-----------|-----------------|-----------------|
| $p(1, 8)$ | $q(3, 7)$ | $r(1, y)$ | $s(1)$ | $s(2)$ |
| $p(2, 8)$ | $q(4, 2)$ | $r(2, n)$ | $s(5)$ | $s(3)$ |
| $p(5, 7)$ | $q(5, 2)$ | $r(4, y)$ | $s(6)$ | $s(4)$ |
| $p(7, 8)$ | $q(6, 2)$ | $r(8, n)$ | $s(7)$ | |
| $p(8, 8)$ | $gteq(X, Y) \leftarrow X \geq Y$ | | $s(8)$ | |

従来手法とは手順 2 以降の発想推論フェイズが違うものとなっている。

まず手順 1 で、ILP と同様の手法である帰納推論を用いて仮説を求める。もし背景知識が不完全であれば、つまり、いくつか未定義な事実がある場合には、帰納推論によって得られた仮説も不完全なものになる。それに対し、もし背景知識が完全であれば、得られた仮説よりも、より適切な仮説が得られる可能性がある。よって、我々は背景知識を満たそうとする。

背景知識を追加するために、手順 2 では手順 1 で得られた仮説 \mathcal{H} と、与えられている事例 \mathcal{E} を用いて発想推論を行い、事実 (ファクト) の集合である目標背景知識群 B を生成する。それは以下の定義式を満たす最小の節集合となる。

$$\mathcal{H} \cup B \models \mathcal{E}$$

次に手順 3 では、目標背景知識群からどの事実に関して調査するか選択し、手順 4 で選択された事実について実際に調査を行う。調査方法は、異なる領域ごとに異なる方法を採用。それは、医学検査、化学実験、データベース検索、専門家への質問等である。調査の結果は背景知識に追加され、その増加した背景知識と与えられた事例を用いて再び帰納推論を行い、仮説を生成する。

これらの手順は、推論によって得られた仮説の精度が要求された精度を超えた時点、又は、背景知識に追加する事実が無くなった時点まで、繰り返す。

3.2 従来との違い

従来手法との違いは手順 2 で発想推論を行う際に、目標仮説を改めて生成せず、帰納推論によって得られた仮説をそのまま使うという点である。こうすることで帰納推論から発想推論への流れにヒューリスティックな要素がなく論理的なものとなっている。こうして得られた目標知識群から、手順 3 で実際に調査したい項目を選ぶことで次の帰納推論に繋げる。

3.3 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。まず、初期の背景知識 B と事例 \mathcal{E} が表 1 のように与えられるとする。ILP における目標概念を $s/1$ 、述語 p, q, r は、それぞれの事例に対する属性値を表している。始めに、 B と \mathcal{E} を用いて帰納推論を行い、仮説 \mathcal{H} を生成する。この際、得られる仮説は

$$\mathcal{H} = \{ s(X) \leftarrow p(X, 8), \quad s(X) \leftarrow q(X, 2) \}$$

である。この仮説は、5 つの正事例全てを被覆するが、2 つの負事例も被覆している。それゆえ、 $B \cup \mathcal{H}$ の精度は $\frac{6}{8} = 0.75$ である。これらの仮説に対し発想推論を行うと

$$B = \{ p(1, Y1), p(2, Y2), \dots, p(8, Y8) \\ q(1, Y1), q(2, Y2), \dots, q(8, Y8) \}$$

表 2: 表 1 に示した事例に対する \mathcal{E} と B 内の述語 p, q, r を表形式で示す。

| id | p | q | r | s | id | p | q | r | s |
|----|-----|-----|-----|-----|----|-----|-----|-----|-----|
| 1 | 8 | (6) | y | + | 5 | 7 | 2 | (n) | + |
| 2 | 8 | (9) | n | - | 6 | (2) | 2 | (y) | + |
| 3 | (6) | 7 | (y) | - | 7 | 8 | (4) | (y) | + |
| 4 | (1) | 2 | y | - | 8 | 8 | (3) | n | + |

となる。例えばここで q に関して追加すると決定すると新しい事実 (即ち、 $q(1, 6), q(2, 9), \dots, q(8, 3)$) を獲得し、それらを B に追加する。

そして改めて帰納推論を行い、

$$\mathcal{H} = \{ s(X) \leftarrow p(X, Y), q(X, Z), gteq(Y, Z), \\ s(X) \leftarrow q(X, 2) \}$$

という仮説を得る。新しい仮説の精度は、 $\frac{7}{8} = 0.875$ である。ここで先ほどのループをさらに続けると次は p に関する事実が追加され

$$\mathcal{H}_{\text{target}} = \{ s(X) \leftarrow p(X, Y), q(X, Z), gteq(Y, Z) \}$$

といった仮説が最終的に得られる。この仮説は全ての事実を証明できる。

4. まとめ

本稿では、ILP における能動的背景知識選択法の再構築を提案した。能動的背景知識選択法は、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が、それをより論理的な枠組みで構築する事ができた。

参考文献

- [Muggleton 99] Muggleton, S.: Inductive Logic Programming, in The MIT Encyclopedia of the Cognitive Sciences (MITECS), MIT Press(1999).
- [Kakas 98] Kakas, A., Kowalski, R., and Toni, F.: The role of abduction in logic programming (1998).
- [白井 03] 白井 秀範, 松井 藤五郎, 大和田 勇人帰納論理プログラミングにおける背景知識の能動的選択法第 17 回 人工知能学会全国大会 (2003)
- [古川 01] 古川康一: 帰納論理プログラミング (2001).