Abstracting Probabilistic Relational Model by Specialisation

Hiroaki Watanabe HW3@DOC.IC.AC.UK
Department of Computing, Imperial College London, 180 Queen's Gate, London SW7 2AZ, UK

Stephen Muggleton SHM@DOC.IC.AC.UK
Department of Computing, Imperial College London, 180 Queen's Gate, London SW7 2AZ, UK

Keywords: logical specialisation, knowledge abstraction, EM algorithm

Abstract

Knowledge abstraction is an essential activity of human intelligence. This paper introduces a logic-based specialisation algorithm for relational domains under uncertainty. The specialisation algorithm takes a pair of existentially quantified conjunctions of first-order literals and computes a most general specialisation of the pair by adopting Plotkin's least general generalisation algorithm. The representation and specialisation algorithm provide a natural way to upgrade non-relational probabilistic models to relational probabilistic models.

1. Introduction

Knowledge abstraction is an essential activity of human intelligence. Inductive Logic Programming (ILP) shows that first-order logic based abstractions can be achieved not only by *generalisation* but also by *specialisation* (Nienhuys-Cheng & de Wolf, 1997). In Statistical Relational Learning (SRL) much work has been done mainly from a generalisation point of view. More studies need for characterising SRL from a specialisation point of view.

The aim of the present paper is to introduce a logic-based specialisation algorithm for relational domains under uncertainty. The specialisation algorithm takes a pair of existentially quantified conjunctions of first-order literals and computes a most general specialisation of the pair by adopting Plotkin's least general generalisation algorithm (Plotkin, 1971). In this paper, we apply such a specialisation algorithm for learning from relational dynamic world. Since we also analyse the computational complexity of our specialisation algorithm, the results of this study could be useful for non-relational probabilistic reasoning communities with large data.

Presented at *ILP-MLG-SRL*, Leuven, Belgium, 2009.

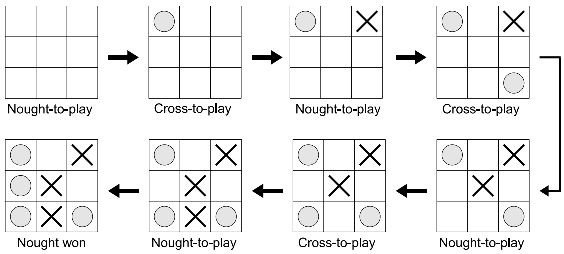

Figure 1. A sequence of plays: a Nought won case

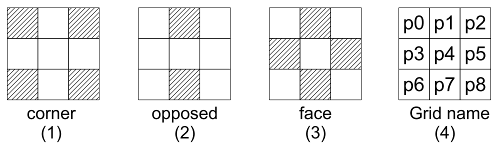

Figure 2. Relations between grids of the game board

2. Representing Observations in Relational Domain

A natural way to express observations in dynamic relational world is to capture the world in a sequence of snapshots. Each snapshot can contain relations whose truth value may change in a different snapshot. For example, let us consider a sequence of plays in Noughts and Crosses (or Tic Tac Toe) domain as shown in Figure 1. Assume that we introduce the three relations as shown in Figure 2. In Figure 1, the fourth board contains the following set of four relations:

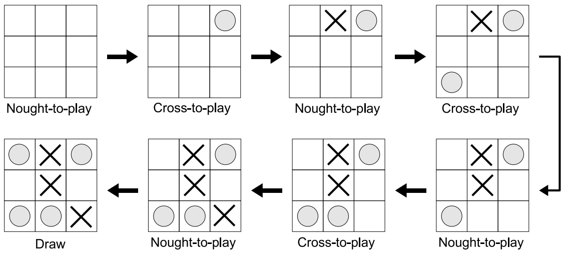

Figure 3. A sequence of plays: a draw case

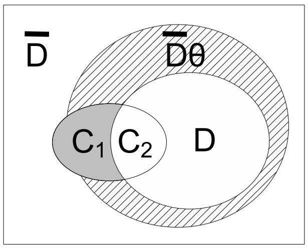

Figure 4. Subsumption Order for Negated Knowledge

```
{corner(1,p0), corner(2,p2), corner(1,p8),
                              opposed(p0,p8)}
```

in which (a) $p0, p2$ and $p8$ express the locations of the grids (Figure 2 (4)) and (b) "*nought*" and "*cross*" are expressed as 1 and 2 respectively.

These observed relations can be expressed in a conjunction of literals. For example, the fourth snapshot can be represented in the following conjunction.

$$F_1 \;:\quad corner(1,p0) \wedge corner(2,p2) \wedge corner(1,p8) \wedge opposed(p0,p8)$$

3. Specialising Sequences of Snapshots

In the previous section, we propose a representation of the snapshot in a conjunction of literals. Our interest is to develop a simple but general approach to find common patterns between a pair of snapshots. Let us assume we observe another sequence of plays which results a draw as shown in Figure 3. The fourth boards in Figure 3 can be expressed as follows.

$$F_2 \;:\quad corner(1,p2) \wedge face(2,p1) \wedge corner(1,p6) \wedge opposed(p2,p6)$$

An obvious common strategy on the fourth boards in Figure 1 and Figure 3 is "*a nought is newly placed at the opposed corner of the existing nought*". Can we extract such knowledge from F_1 and F_2?

Our approach is to compute most general specialisation (mgs) of a pair of conjunctions of literals. The knowledge representation we introduced in the previous section is a ground case of an existentially quantified conjunction of literal (ECOL). Now we define our Machine Learning task as follows.

Given: A pair of ECOLs, (F_i, F_j).
Find: Most general specialisation of (F_i, F_j).

To the best of our knowledge, mgs of a pair of ECOLs has not been studied in ILP although it is a natural representation of observations.

4. Theoretical Analysis

We show $\overline{lgg(\overline{F}_a, \overline{F}_b)}$ is the msg of the given ECOLs in this section. First of all, let us review the θ-subsumption order between two clauses briefly. We represent a clause, $l_1 \vee ... \vee l_m$ as a set of literals, $\{l_1, ..., l_m\}$. θ-subsumption order is introduced as follow.

Definition: θ-Subsumption Order
Let C and D be clauses. If there exist substitution θ such that $C\theta \subseteq D$, C is more general than D, $C \succeq_\theta D$, under θ-subsumption order.

Let F_a and F_b be ECOLs. We define a new generality order $\succeq_\exists$ between F_a and F_b.

Definition: e-Subsumption Order
Let F_a and F_b be ECOLs. e-Subsumption Order, $\succeq_\exists$, is defined associated with F_a and F_b as follows.

$$F_a \succeq_\exists F_b \quad \stackrel{\text{def}}{=} \quad F_a\theta \subseteq F_b$$

The following theorem shows a relation between the Plotkin's θ subsumption order and our existential subsumption order.

Theorem 1: θ-subsumption and e-subsumption
Given two clauses C and D, let $\overline{C}$ and $\overline{D}$ be the negated C and D respectively. Then

$$C \succeq_\theta D \Leftrightarrow \overline{C} \preceq_\exists \overline{D}.$$

Proof. It is clear that both $\overline{C}$ and $\overline{D}$ are ECOLs. Regarding $C \succeq_\theta D \Rightarrow \overline{C} \preceq_\exists \overline{D}$, we assume there exists a substitution θ such that

$$C\theta \subseteq D. \tag{1}$$

Now our aim is to show

$$\overline{D}\theta \subseteq \overline{C} \tag{2}$$

from (1). Figure 4 shows a general Venn diagram of the models of C and D where $C = C_1 \cup C_2$. The Figure shows that (2) holds iff C_1 is empty. Here, C_1 should be empty from (1). Regarding $C \succeq_\theta D \Leftarrow \overline{C} \preceq_\exists \overline{D}$, our aim is to show (1) from (2). Now C_1 is empty because of (2). Then (1) is always true. □

Theorem 2 (Most General Specialisation of a pair of ECOLs)
Given two ECOLs F_a and F_b, there exists the most general specialisation of F_a and F_b.

Proof. Let F_{com} be some ECOL such that

$$F_a \succeq_\exists F_{com} \quad \text{and} \quad F_b \succeq_\exists F_{com}$$

From the above Theorem 1, this can be stated as follows.

$$\overline{F}_a \preceq_\theta \overline{F}_{com} \qquad \overline{F}_b \preceq_\theta \overline{F}_{com}$$

Note that $\overline{F}_a$, $\overline{F}_{com}$, and $\overline{F}_b$ are all clauses now. As Plotkin shows (Plotkin, 1971), there exist the least general generalisation of $\overline{F}_a$ and $\overline{F}_b$.

$$\overline{F}_a \preceq_\theta lgg(\overline{F}_a, \overline{F}_b) \preceq_\theta \overline{F}_{com}$$

$$\overline{F}_b \preceq_\theta lgg(\overline{F}_a, \overline{F}_b) \preceq_\theta \overline{F}_{com}$$

The above two formulae are equivalent with

$$F_a \succeq_\exists \overline{lgg(\overline{F}_a, \overline{F}_b)} \succeq_\exists F_{com}$$

$$F_b \succeq_\exists \overline{lgg(\overline{F}_a, \overline{F}_b)} \succeq_\exists F_{com}$$

where $\overline{lgg(\overline{F}_a, \overline{F}_b)}$ is the most general specialisation of F_a and F_b in ECOL. □

The above proof shows that the following algorithm computes the most general specialisation of the pair of ECOLs.

MGS(F_a,F_b)
Input: a pair of ECOLs (F_a, F_b)
Output: Most General Specialisation of F_a and F_b
1: Compute the negations of F_a and F_b, $\overline{F}_a$ and $\overline{F}_b$.
2: Take lgg of $\overline{F}_a$ and $\overline{F}_b$, $lgg(\overline{F}_a, \overline{F}_b)$
3: Logicaly negate the result of the lgg, $\overline{lgg(\overline{F}_a, \overline{F}_b)}$ and output it.
4: Exit.

For example, mgs of F_1 and F_2 can be calculated as

$$F_3 \ : \ \exists XYZ \quad corner(1, X) \wedge corner(1, Y) \wedge corner(Z, p2) \wedge opposed(X, Y)$$

whereas the mgs of the ECOLs for the third snapshots in Figure 1 and Figure 3 can be computed as follows.

$$F_4 \ : \ \exists XY \quad corner(1, X) \wedge corner(Y, p2)$$

An abstracted snapshot, F_4, can be interpreted as

- a nought is placed at one corner and
- there is some mark at $p2$.

Then a state transition happens from F_4 to F_3 where

- two noughts are placed at opposed corners and
- something is placed at $p2$.

In F_3 we could find that a nought is newly placed at the opposite corner of the existing nought.

5. Conclusions

In this paper, the new specialisation algorithm for a pair of ECOLs is presented. This algorithm has been implemented in a PILP system, called *Cellist*, for learning logically extended Probabilistic Automata, Probabilistic Logic Automata (PLA) (Watanabe & Muggleton, 2009; Watanabe, 2009). Since ECOL is a natural representation for describing a set of relations, it should be straight forward to extend non-relational probabilistic models to relational probabilistic models. PLA is such a case; Probabilistic Automata is extended to PLA by simply embedding ECOL into a node in Probabilistic Automata. We expect this ECOL based specialisation approach smoothly bridges ILP and existing non-relational probabilistic models.

References

Nienhuys-Cheng, S., & de Wolf, R. (1997). Foundations of inductive logic programming.

Plotkin, G. (1971). *Automatic methods of inductive inference.* Doctoral dissertation, Edinburgh University, UK.

Watanabe, H. (2009). *A learning theory approach for probabilistic relational learning (draft).* Doctoral dissertation, Imperial College London, UK.

Watanabe, H., & Muggleton, S. (2009). Approximative generalisation: a PILP approach (submitted to ilp2009).