

# Towards Meta-Level Descriptive ILP

Pedro Torres\*<sup>†</sup> and Simon Colton\*

\*Department of Computing, Imperial College, London, UK. sgc.ptorres@doc.ic.ac.uk

<sup>†</sup> Supported by Fundação para a Ciência e a Tecnologia, (SFRH/BD/12437/2003).

**Abstract.** We introduce a formal setting for a first-order production rule approach to descriptive learning and we consider searching for sets of production rules as opposed to a direct search for concepts. We sketch the PARROT algorithm which generates a basic set of production rules capable of producing a set of user-given concepts and present some preliminary ideas on how to enrich this set.

## 1 Introduction

Descriptive Inductive Logic Programming approaches aim to find interesting relationships in unclassified relational data. Claudien [4] and WARMR [3] both work in this setting, using a declarative language bias to define the search space. HR [1] forms clausal theories consisting of classification rules and association rules and has recently been described in ILP terms [2]. Unlike Claudien and WARMR, which use language biases, HR employs production rules to define the search space and various measures of interestingness to direct the search.

In a study involving algebraic structures [5], we extended our theory formation approach with a generic first-order production rule. By writing explicit definitions of first-order production rules, the user could add to the set of production rules used by HR. We extend this approach here, and we investigate how a search for sets of production rules could be carried out. If such a search can be efficiently implemented, this opens up the possibility of automatically generating production rule based theory formation systems tailored to a specific domain. Such tailoring will be enabled by the user presenting core background theory and interesting concepts derived from them, with our system able to generate production rule sets that produce the interesting concepts from the core ones, and hopefully additional interesting material.

## 2 First-order representation of production rules

Let  $\{\tau_i\}_{i=1}^N$  be a collection of sets, which we will call *types*. We define a *concept* to be a relation defined between any  $n$  types. We call  $n$  the *arity* of the concept. We will here consider concepts of the form  $\{(x_1, \dots, x_n) \in \tau_1 \times \dots \times \tau_n \mid \phi\}$ , where  $\phi$  is some first-order logic formula for which the only free variables are  $x_1, \dots, x_n$  and any other variables are bounded. The above concept will be written as  $[x_1, \dots, x_n] : \phi(x_1, \dots, x_n)$ . We write  $\mathcal{C}$  for the set of all concepts.

A *production rule* is a function  $\pi : \mathcal{D} \rightarrow \mathcal{C}$ , where  $\mathcal{D} \subset \mathcal{C}^q$ , for some  $q$ , called the *arity* of  $\pi$ . A production rule can be defined by specifying the transformation

it operates on the logical formulas defining the input concepts. Suppose we have  $x \in \tau_1$ ,  $y \in \tau_2$  and  $z \in \tau_3$ . We can define, for example,  $\pi$  as a function carrying out the mapping  $([x, y] : \phi(x, y), [y, z] : \varphi(y, z)) \mapsto [x, z] : \exists y. (\phi(x, y) \wedge \varphi(y, z))$ .

From an initial set of background concepts  $\{C_1, \dots, C_k\}$ , by recurrently applying production rules with appropriate domains, we can define new concepts, e.g.  $C = \pi_6(\pi_4(\pi_1(C_1, C_2), \pi_2(C_2, C_3)), \pi_5(\pi_3(C_4)))$ . Given a set of initial concepts  $\mathcal{C}$  and a set of production rules  $\Pi$ , we recursively define  $\Pi^*(\mathcal{C})$  as the set of all concepts which: (i) are in  $\mathcal{C}$  or (ii) can be written as  $\pi(C_1, \dots, C_n)$  where  $\pi$  is in  $\Pi$  and every  $C_i$  is in  $\Pi^*(\mathcal{C})$ . The set  $\Pi^*(\mathcal{C})$  can be intuitively described as the set of all concepts that can be produced from the concepts in  $\mathcal{C}$  by using the production rules in  $\Pi$ . If we fix  $\mathcal{C}$ , the set  $\Pi$  defines exactly which concepts are in the search space. The essential difference between language biases and sets of production rules is that the former usually let the user directly restrict the syntactical form of all the concepts formed, whereas the latter restrict the stepwise syntactical construction of new concepts from old ones.

One way of making conjectures in a production rule based system is to compare concepts produced. Suppose we have two concepts of the same arity and types for which all the available examples which satisfy one also satisfy the other. In these circumstances, it is possible to conjecture that they are equivalent. Similar situations may lead to implication and non-existence conjectures. The association rules that such a system may come up with are therefore all of the form  $[\forall \mathbf{x}. C_1(\mathbf{x}) \rightarrow C_2(\mathbf{x})]$ ,  $[\forall \mathbf{x}. C_1(\mathbf{x}) \leftrightarrow C_2(\mathbf{x})]$  or  $[\forall \mathbf{x}. \neg C_1(\mathbf{x})]$  where  $\mathbf{x}$  is some set of variables and  $C_1$  and  $C_2$  are concepts in  $\Pi^*(\mathcal{C})$ .

It is possible to define a conjecture as a 0-arity concept. In this light, conjecture making techniques are simply particular production rules of arity 2 (they take two concepts) which yield concepts of arity 0 (i.e. conjectures). By using these definitions, concepts and conjectures can be treated at the same level. Furthermore, we can define an order relation on the space of sets of production rules. Given a set of initial concepts,  $\mathcal{C}$ , and two sets of production rules,  $\Pi_1$  and  $\Pi_2$ , we say that  $(\Pi_1 \leq_{\mathcal{C}} \Pi_2)$  iff  $\Pi_1^*(\mathcal{C}) \subseteq \Pi_2^*(\mathcal{C})$ . In this case,  $\Pi_2$  is said to be *richer* than  $\Pi_1$ .

### 3 Searching for production rules

The main idea we explore is that, instead of searching for concepts with a fixed set of production rules, we can start by analysing examples of previously formed (interesting) theories in a chosen domain and build a set of production rules tailored to that particular domain.

In order to perform the search for sets of production rules, we have sketched three high-level algorithms: (i) the EXPANDER algorithm for computing  $\Pi^*(\mathcal{C})$ , up to some defined depth, from user-defined  $\mathcal{C}$  and  $\Pi$ ; (ii) the PARROT algorithm which takes a large set of formed concepts,  $\mathcal{C}$ , all ultimately constructed from some elementary concepts  $\mathcal{C}_0$ , and defines a set of production rules,  $\Pi_0$ , such that  $\Pi_0^*(\mathcal{C}_0) \supseteq \mathcal{C}$ ; (iii) the ENRICHER algorithm, which takes a set of production rules and iteratively enriches this set.

EXPANDER is a simple algorithm which considers all legal sequential application of production rules up to some prescribed depth. In the PARROT algorithm, each concept,  $C$ , in the user-given set of concepts,  $\mathcal{C}$ , can be expressed as a first order formula making use of some old concepts  $\mathcal{C}_0(C)$ . It is therefore possible to craft a production rule which takes the concepts  $\mathcal{C}_0(C)$  and produces  $C$ . If this is done for all the concepts in  $\mathcal{C}$ , we get a set of production rules,  $\Pi_0$ , which can generate all of the concepts in  $\mathcal{C}$  from the elementary concepts they are based on. Since we have put concepts and conjectures on the same level, the above reasoning means that, if the user includes conjectures (or, possibly, theorems) on the set of concepts  $\mathcal{C}$ , the production rule set given by PARROT will also be able to make those conjectures. If the set of concepts,  $\mathcal{C}$ , is very restricted, the production rule set output by PARROT may not be able to generate much more than  $\mathcal{C}$  itself. If this is the case, we will want to enrich that set.

A simple way to enrich a set of production rules,  $\Pi$ , is to remove a production rule  $\pi$  and add two separate production rules  $\pi_1$  and  $\pi_2$  such that, for every  $C$  in the domain of  $\pi$ , we have  $\pi(C) = \pi_1(\pi_2(C))$ . In this example we have considered rules of arity 1 but this *decomposition* idea can be generalised to higher arities. Given a set of production rules, there is a considerable number of possible decompositions enriching that set. To facilitate the analysis of these decompositions we have defined a set of so called *primitive* production rules which are closely based on the syntax of first order logic. By using the primitive production rules, we can write different composite production rules using a common vocabulary and search for matching subparts.

The ENRICHER algorithm is given a set of production rules and performs a best-first search through all decompositions. It uses EXPANDER to compute the concepts which each particular decomposition adds and calculates a score for each decomposition based on measures defined by the user. For example, using this algorithm, the user can ask for an enrichment which favours concepts with higher applicability (i.e. higher proportion of which satisfy the concept).

In the longer version of this paper we plan to give details of the three theoretical algorithms mentioned above and toy examples of their application. This will include the complete definition of the primitive production rules, which play an important role in the ENRICHER algorithm, and the analysis of some of their properties.

## References

1. S. Colton. *Automated Theory Formation in Pure Mathematics*. Springer, 2002.
2. S. Colton and S. Muggleton. Mathematical applications of Inductive Logic Programming. *Machine Learning (online first)* 2005.
3. Luc Dehaspe and Hannu Toivonen. Discovery of frequent DATALOG patterns. *Data Mining and Knowledge Discovery*, 3(1):7–36, 1999.
4. L. De Raedt and L. Dehaspe. Clausal discovery. *Machine Learning*, 26:99–146, 1997.
5. P. Torres and S. Colton. Using model generation in automated concept formation. In *Proceedings of the Automated Reasoning Workshop*, 2006.