A Parameterized Algorithm for Exploring Concept Lattices

Peggy Cellier[1], Sébastien Ferré[1], Olivier Ridoux[1], and Mireille Ducassé[2]

[1] IRISA/University of Rennes 1
[2] IRISA/INSA, Campus universitaire de Beaulieu, 35042 Rennes, France
firstname.lastname@irisa.fr
http://www.irisa.fr/LIS/

Abstract. Formal Concept Analysis (FCA) is a natural framework for learning from positive and negative examples. Indeed, learning from examples results in sets of frequent concepts whose extent contains only these examples. In terms of association rules, the above learning strategy can be seen as searching the premises of exact rules where the consequence is fixed. In its most classical setting, FCA considers attributes as a non-ordered set. When attributes of the context are ordered, Conceptual Scaling allows the related taxonomy to be taken into account by producing a context completed with all attributes deduced from the taxonomy. The drawback, however, is that concept intents contain redundant information. In this article, we propose a parameterized generalization of a previously proposed algorithm, in order to learn rules in the presence of a taxonomy. The taxonomy is taken into account during the computation so as to remove all redundancies from intents. Simply changing one component, this parameterized algorithm can compute various kinds of concept-based rules. We present instantiations of the parameterized algorithm for learning positive and negative rules.

1 Introduction

Learning from examples is a fruitful approach when it is not possible to a priori design a model. It has been mainly tried for classification purposes [Mit97]. Classes are represented by examples and counter-examples, and a formal model of the classes is learned by a machine.

Formal Concept Analysis (FCA) [GW99] is a natural framework for learning from positive and negative examples [Kuz04]. Indeed, learning from positive examples (respectively negative examples) results in sets of frequent concepts with respect to a minimal support, whose extent contains only positive examples (respectively negative examples). In terms of *association rules* [AIS93, AS94], the above learning strategy can be seen as searching the premises of exact rules where the consequence is fixed. When augmented with statistical indicators like *confidence* and *support* it is possible to extract various kinds of concept-based rules taking into account exceptions [PBTL99, Zak04].

The input of FCA is a formal context that relates objects and attributes. FCA considers attributes as a non-ordered set. When attributes of the context

S.O. Kuznetsov and S. Schmidt (Eds.): ICFCA 2007, LNAI 4390, pp. 114–129, 2007.

are ordered, Conceptual Scaling [GW99] allows the attribute taxonomy to be taken into account by producing a context completed with all attributes deduced from the taxonomy. The drawback is that concept intents contain redundant information. In a previous work [CFRD06], we proposed an algorithm based on Bordat's algorithm [Bor86] to find frequent concepts in a context with taxonomy. In that algorithm, the taxonomy is taken into account during the computation so as to remove all redundancies from intents.

There are several kinds of association rules, and several related issues: e.g. find all association rules with respect to some criteria, compute all association rules with a given conclusion or premise. We propose a generic algorithm to address the above issues. It is a parameterized generalization of our previous algorithm. It learns rules and is able to benefit from the presence of a taxonomy. The advantage of taking a taxonomy into account is to reduce the size of the results. For example, the attributes of contexts about Living Things are intrinsically ordered [1]. For a target such as "suckling", a rule such as "Living Things"$\wedge$ "Animalia" $\wedge$ "Chordata" $\wedge$ "Vertebrata" $\wedge$ "Mammalia" $\rightarrow$ "suckling" is less relevant than the equivalent rule "Mammalia" $\rightarrow$ "suckling" where elements redundant with respect to the taxonomy have been eliminated. The presented algorithm can compute various kinds of concept-based rules by simply changing one component. We present two instantiations which find positive and negative rules. Positive rules predict some given target (e.g. predict a mushroom as poisonous), while negative rules predict its opposite (e.g. edible).

The contributions of this article are twofold. Firstly, it formally defines FCA with taxonomy (FCA-Tax) using the Logical Concept Analysis (LCA) framework [FR04], where the taxonomy is taken into account as a specific logic. Secondly, it specifies a generic algorithm which facilitates the exploration of frequent concepts in a context with taxonomy. Quantitative experiments show that taking a taxonomy into account does not introduce slowdowns. Furthermore, the pruning implemented by our algorithm related to the taxonomy can often improve efficiency.

In the following, Section 2 formally defines FCA with taxonomy (FCA-Tax). Section 3 presents the generalization of the algorithm described in [CFRD06] to filter frequent concepts in a formal context with taxonomy. Section 4 shows how to instantiate the algorithm to learn different kind rules. Section 5 discusses experimental results.

2 A Logical Framework for FCA with Taxonomy

In this section, we formally describe Formal Concept Analysis with Taxonomy (FCA-Tax) using the Logical Concept Analysis (LCA [FR04]) framework. We first present an example of context with taxonomy. Then we briefly introduce LCA, concept-based rules, and we instanciate LCA to FCA-Tax. A taxonomy describes how the attributes of the context are ordered and thus a taxonomy is a kind of logic where the subsumption relation represents this order relation.

[1] http://anthro.palomar.edu/animal/table_humans.htm