Available online at www.sciencedirect.com

ScienceDirect

Electronic Notes in Theoretical Computer Science 303 (2014) 37–57

Electronic Notes in Theoretical Computer Science

www.elsevier.com/locate/entcs

Exploring Users' Preferences in a Fuzzy Setting

Cynthia Vera Glodeanu[1,2]

Institute of Algebra
Technische Universität Dresden
Dresden, Germany

Abstract

We propose a new method for modelling users' preferences on attributes that contain more than one trait. Starting with a data set the users have to enter a sort of order on the attributes in form of formulas corresponding to their preferences. Based on this order they only receive the relevant formal concepts, i.e., "object-attribute clusters", where relevant corresponds to the users' point of view. The preference modelling is done within the framework of Formal Fuzzy Concept Analysis. This has numerous advantages. First, the relevant information is contained in a complete lattice, the *concept lattice*, that allows the users to browse among their preferences. This lattice may be used for further data analysis by applying different methods from Formal Concept Analysis. Second, we can investigate the computation of non-redundant bases for the entered formulas. Since the users are allowed to enter the formulas, these may be redundant. The base offers a better overview of the preferences and thus the formulas can be altered more easily.

Keywords: Formal Concept Analysis, fuzzy data, data reduction, $\mathbf{L}^*$-closure operators.

1 Introduction

Formal Concept Analysis [16] is an instrument for data analysis based on lattice theory. Starting with a set of *formal objects*, a set of *formal attributes* and an *incidence relation* indicating which object has which attribute, one obtains a *formal context* combining these three components. The context, in turn, allows for the computation of the *formal concepts*. These concepts are understood as units with a *conceptual extent* and a *conceptual intent*, an idea that can be found already in the *Logic of Port Royal* [11]. The extent of a concept contains all the objects shared by the attributes from its intent. Dually, the intent of a formal concept contains all the attributes that the objects from its extent have in common. The order on the concepts is given by the *subconcept-superconcept relation*. Together with this

[1] Supported by Grant No. IE120596, "Universal Algebra and its dual: monads and comonads, Lawvere theories and what?" of The Royal Society

[2] Email: `Cynthia-Vera.Glodeanu@tu-dresden.de`

http://dx.doi.org/10.1016/j.entcs.2014.02.003

relation the set of all concepts forms a complete lattice, the *concept lattice*, which represents the basis for further data analysis.

Since its early years, Formal Concept Analysis has developed into a research field in its own right with a thriving theoretical community and a rapidly expanding range of applications: analysis of domestic violence [29,30], genetics and biology [20,13,1,12], cancer studies [24], the management of Web 2.0 data [19], rough sets [15], non-metric factor analysis [10,14,7] and data mining [32,33], to name a few.

We will work with a fuzzy variant of Formal Concept Analysis [31,4,9] where the incidence relation is replaced by a fuzzy relation encoding the vagueness of the statements. All the above notions can be defined in such a setting. This generalisation of Formal Concept Analysis has been fruitful for the handling of real-world data. It found applications within a few years after its introduction in fields like social networks [22], the management of incomplete and inconsistent data sets [23,25], non-metric factor analysis [2], etc.

The advantage of concept lattices is that they contain all the information of the data set encoded by concepts. However, the lattice may become too big and may be hard to read. To overcome this problem, the fuzzy concept lattice constrained by hedges [9] was developed. However, this data reduction approach is of a different nature than ours as it does not take into account the users' preferences. In this paper we propose a method for reducing the size of the fuzzy concept lattice based on the preferences of users. Through this method users have stronger control over the information omitted from the data. We develop *users' preference formulas* for compound attributes, i.e., for qualities which include more than one trait. For instance, the notion “wealth” might be considered as a compound attribute consisting of “investment” and “fluency”. A person who is wealthy has to have high values on both investment and fluency. The users have to enter formulas reflecting their preferences. Based on these preferences the users only obtain the concepts relevant for them. We will briefly discuss formulas for non-compound attributes as well. The formulas are entered by the users and thus they may be redundant. Having a set of non-redundant formulas facilitates their further investigation and alteration. Therefore, we will investigate the computation of such sets of non-redundant formulas.

This article is an extended version of [17]. The new results give further insight into the connection between special kinds of formulas and closure operators and into the computation of non-redundant formula sets. Further, we illustrate the method on a real-world data set.

There is related work to ours from crisp Formal Concept Analysis. *Attribute dependency formulas* were introduced in [8]. These formulas impose constraints on the concepts as well, but were not designed for compound attributes. In this paper we will briefly discuss their fuzzification and see that they are not appropriate for our framework. A somehow different approach of modelling users' preferences within the framework of crisp Formal Concept Analysis was presented in [28]. Starting with the users' preferences on objects, one obtains a preference relation on concepts and afterwards on the attributes. The method embeds preference logic into the