Inference of mixed information in Formal Concept Analysis

P. Cordero, M. Enciso, A. Mora, and J. M. Rodriguez-Jiménez

University of Málaga, Andalucía Tech, Spain,
e-mail: {amora,jmrodriguez}@ctima.uma.es, {pcordero,enciso}@uma.es

Abstract. Negative information can be considered twofold: by means of a negation operator or by capturing the absence of information. In this second approach, a new framework have to be developed: from the syntax to the semantics, including the management of such generalized knowledge representation. In this work we traverse all these issues in the framework of formal concept analysis, introducing a new set of inference rules to manage mixed (positive and negative) attributes.

Summary of the work

We focus on Formal Concept Analysis which represents the information as a binary relation (named formal context) between objects as rows and attributes as columns. From this table, mining techniques to extract knowledge are well know and it is possible to compute a concept lattice and also a sets of implications representing the same knowledge. We are particularly interested in implications because they allow a symbolic manipulation by using logic.

Normally, the formal context stores that "an object has an attribute". In this work, we consider interesting not only this observation, represented with positive attributes, but also that "an object has not an attribute" (negative attribute), that is, not only the presence but also the absence of a given property.

This idea already appeared in data mining area, where some works [2, 8] considered the negation of attributes inside of implications: the use of positive and negative attributes appeared in basket market analysis, when rules considering that "if a customer usually buys a given product then he does not buy other specific product". In these former works, the negative information negatively impacts on the data mining method because some cut mechanisms used became not valid. In our opinion, this situation can be resolved if a suitable treatment with positive and negative attributes are added.

In FCA, as far as we know, only some works consider the management of both positive and negative attributes in FCA. In [1], the authors join the formal context with its complementary and use the classical FCA methods of data mining to retrieve implications with positive and negative attributes. Unfortunately, such an approach cannot take advantage of the semantics of the enriched framework. In 2012, Missaoui et.al. [3] have approached the generation of mixed implications from two given sets of implications only with positive and negative attributes respectively. We have followed this line: in [7], we extended the

classical FCA framework with new derivation operators constituting a Galois connection for the treatment of negative and positive information in FCA and in [5, 6] we proposed some mining algorithms to derive directly mixed implications.

In this work, we progress in this line by proposing a set of inference rules to manage efficiently mixed implications. This logic is based on Simplification paradigm presented in [4]. The new set of rules to manipulate implications with positive and negative attributes is presented here.

Acknowledgments

References

1. Gasmi, G., Yahia, S.B., Nguifo, E.M., Bouker, S.: Extraction of association rules based on literalsets. LNCS 4654, 293–302 (2007)
2. Mannila, H., Toivonen, H., Verkamo, A.I.: Efficient algorithms for discovering association rules. In: KDD Workshop. pp. 181–192 (1994)
3. Missaoui, R., Nourine, L., Renaud, Y.: Computing Implications with Negation from a Formal Context. Fundamenta Informaticae 115(4), 357–375 (dec 2012)
4. ngel Mora, Enciso, M., Cordero, P.: Closure via functional dependence simplification. International Journal of Computer Mathematics 89, 510–526 (2012)
5. Rodríguez-Jiménez, J., Cordero, P., Enciso, M., Mora, A.: A generalized framework to consider positive and negative attributes in formal concept analysis. In: CLA. CEUR Workshop Proceedings, vol. 1252, pp. 267–278. CEUR-WS.org (2014)
6. Rodríguez-Jiménez, J., Cordero, P., Enciso, M., Mora, A.: Data mining algorithms to compute mixed concepts with negative attributes: an application to breast cancer data analysis. Mathematical Methods in the Applied Sciences (2016)
7. Rodriguez-Jimenez, J., Cordero, P., Enciso, M., Rudolph, S.: Concept lattices with negative information: a characterization theorem. Information Sciences http://dx.doi.org/10.1016/j.ins.2016.06.015 (2016)
8. Srikant, R., Agrawal, R.: Mining quantitative association rules in large relational tables. In: Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data. pp. 1–12. SIGMOD '96, ACM, New York, NY, USA (1996)