Interactive Exploration over Concept Lattices with LatViz

Mehwish Alam[1] and **Thi Nhu Nguyen Le**[2] and **Amedeo Napoli**[3]

Abstract. In this demo paper, we introduce `LatViz`, a new tool which allows the construction, the display and the exploration of concept lattices. `LatViz` proposes some remarkable improvements over existing tools and introduces various new functionalities focusing on interaction with experts, such as visualization of pattern structures (for dealing with complex non-binary data), AOC-posets (the irreducible elements of the lattice), concept annotations, filtering based on various criteria and visualization of implications. This way the user can effectively perform interactive exploratory knowledge discovery as often needed in knowledge engineering, and especially in ontology engineering.

1 INTRODUCTION

For facing the analysis of large amounts of complex data, there is a need for knowledge discovery and especially classification, where the resulting classes can be made available to the user for exploration, interpretation, and knowledge representation. In this research work, we use Formal Concept Analysis (FCA [6]), as a basis for classification. The purpose of this demo paper is to introduce `LatViz`, a new tool for drawing concept lattices and allowing interactive exploration. `LatViz` provides many new functionalities, such as the visualization of pattern structures and AOC-posets, concept annotation, concept lattice filtering, and visualization of implications. This way a user can effectively perform an interactive exploration over a concept lattice built from complex data from the web of data (RDF data), numerical or graph data.

A demo of `LatViz` is available on-line through `http://latviz.loria.fr/`. Moreover, a companion paper will be issued in July 2016 [2].

2 FUNCTIONALITIES OF LATVIZ

`LatViz` provides basic functionalities for building a concept lattice, and implements two main algorithms for building a concept lattice from a binary context. Facilities are also present for dealing with the so-called "pattern structures", that generalize FCA in the sense that any kind of data can be processed, e.g. numbers, sequences, trees and graphs [5]. This is a noticeable ability of `LatViz`, as most of the tools for displaying concept lattices can only work on binary data. At present, `LatViz` can work on complex data types such as numbers and RDF data [1, 3]. An example of pattern concept lattice based on intervals is shown in Figure 1 (30 hotels $\times$ 4 numerical attributes).

`LatViz` provides several interactive operations for reducing the exploration space of the expert. One is the display of the so-called AOC-posets, i.e. sets of "Attribute- and Object-Concepts" [4], where only irreducible elements of the concept lattice are shown. For example, a concept lattice including 31 concepts with a related AOC poset only including 20 concepts is shown in Figure 2. This set of irreducible elements can be seen as the core of the concept lattice and includes the whole set of information attached to the concept lattice.

`LatViz` allows the creation of a concept lattice level-wise by interaction, i.e. "breadth-first display strategy". When an expert clicks on the top concept, `LatViz` computes and displays the first level. After that the expert can select a given concept for continuing the exploration, then `LatViz` computes the next level for that concept (see Figure 3). Dually, it is possible to apply a "depth-first display strategy", allowing the expert to only highlight sub/super concepts of a selected concept. In this way, the expert can only examine the part of the concept lattice in which (s)he is interested. Then, there can be a "top-down" or a "bottom-up exploration" depending on the interest of the expert. Actually, each concept is based on two main facets, the intent which is the set of attribute and the extent which is the set of objects related to a concept. The top-down exploration materializes the interest on intents (attributes are inherited top-down) while the bottom-up exploration focuses on extents (extents are inherited bottom-up).

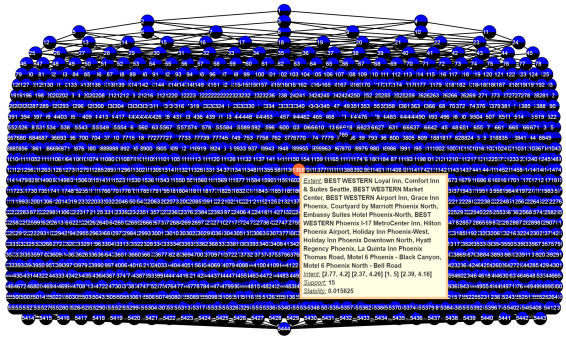

Figure 1: The full lattice (30 hotels $\times$ 4 attributes).

One important functionality, especially useful for information retrieval, is the possibility of using filtering. When a concept lattice is explored, expert is allowed to set several filtering criteria such as stability and lift (concept interest measures), extent size, intent size and finally specific object or attribute names. A concept lattice can also

[1] Laboratoire d'Informatique de Paris-Nord, Université Paris 13, Paris, France, email: alam@lipn.univ-paris13.fr

[2] LORIA (CNRS – Inria Nancy Grand Est – Université de Lorraine) BP 239, Vandoeuvre-les-Nancy, F-54506, France, email: thi-nhu-nguyen.le@loria.fr

[3] LORIA (CNRS – Inria Nancy Grand Est – Université de Lorraine) BP 239, Vandoeuvre-les-Nancy, F-54506, France, email: amedeo.napoli@loria.fr

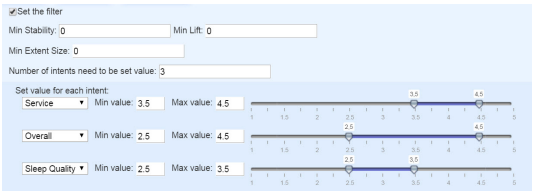

Figure 4: Filtering on "service", "overall", and "sleep quality" attributes.

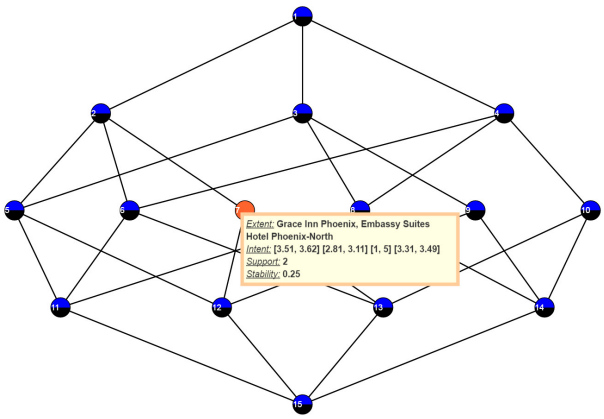

Figure 5: The filtered concept lattice.

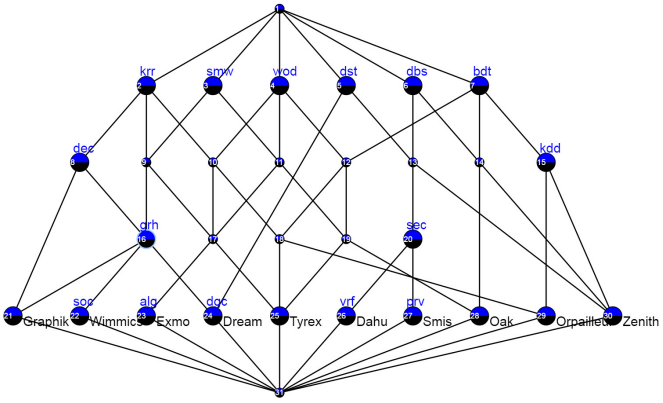

Figure 2: The concept lattice of teams. The concepts of the AOC poset are emphasized.

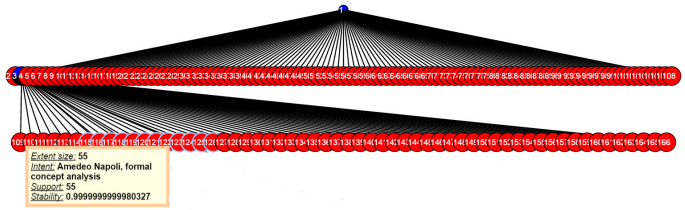

Figure 3: The display of a large concept lattice level by level.

be filtered by specifying the number of attributes to be considered, the upper and the lower limits for the intervals in the intent of each attribute along with stability, lift and extent size.

For example, we considered a data table about hotels with 30 hotels and 4 attributes, namely "service", "overall", "rooms", and "sleep quality". For each attribute we have a rating between 1 (minimal) and 5 (maximal). The filtering constraints and the resulting concept lattice are displayed on Figures 4 and 5. It should be noticed that the initial concept lattice includes 5444 concepts while the filtered one only includes 15!

Finally, `LatViz` proposes to visualize implications (rules with a confidence of 100%) with the help of scatter plots, where along the x-axis rules are ranked by increasing support and along the y-axis rules are ranked by increasing lift (as confidence is constant). Such kind of display helps the expert to single-out the rules w.r.t. support and lift, but other measures can be easily used if needed.

`LatViz` is still a tool under development and an experiment with potential users is actually running for understanding potential limitations and improvements. Moreover, other functionalities should still be implemented such as: scaling procedures, facilitating and extending the use of pattern structures (e.g. sequences, trees, graphs), introducing new interest measures for concepts, alternative display of implication and association rules, and the so-called nested line diagrams which are based on particular decompositions of the data tables.

3 TECHNICAL DATA

`LatViz` can be tested at `http://latviz.loria.fr/` with instructions at `http://latviz.loria.fr/about.html`. The requirements for users are standard: to be able to access one of the following browsers: Chrome, Firefox, IE (version 9 or after), Opera, and to have a network connection to a remote server.

REFERENCES

[1] Mehwish Alam, Aleksey Buzmakov, Amedeo Napoli, and Alibek Sailanbayev, 'Revisiting pattern structures for structured attribute sets', in *Proceedings of the 12th International Conference on Concept Lattices and Their Applications (CLA)*, pp. 241–252, (2015).

[2] Mehwish Alam, Nguyen Thi Nhu Le, and Amedeo Napoli, 'Latviz: A new practical tool for performing interactive exploration over concept lattices', in *Proceedings of 13th International Conference on Concept Lattices and Their Applications (CLA), Moscow, July 2016*, (2016).

[3] Mehwish Alam, Matthieu Osmuk, and Amedeo Napoli, 'RV-Xplorer: A Way to Navigate Lattice-Based Views over RDF Graphs', in *Proceedings of the 12th International Conference on Concept Lattices and Their Applications (CLA)*, pp. 23–34, (2015).

[4] Anne Berry, Alain Gutierrez, Marianne Huchard, Amedeo Napoli, and Alain Sigayret, 'Hermes: a simple and efficient algorithm for building the AOC-poset of a binary relation', *Annals of Mathematics and Artificial Intelligence*, **72**, 45–71, (2014).

[5] Bernhard Ganter and Sergei O. Kuznetsov, 'Pattern structures and their projections', in *ICCS*, pp. 129–142, (2001).

[6] Bernhard Ganter and Rudolf Wille, *Formal Concept Analysis: Mathematical Foundations*, Springer, 1999.