Formal concept analysis over attributes with levels of granularity

Radim Bělohlávek, Vladimír Sklenář
Dept. Computer Science, Palacký University, Tomkova 40, CZ-779 00, Olomouc, Czech Republic
Email: {radim.belohlavek, vladimir.sklenar}@upol.cz

Abstract—Formal concept analysis (FCA) is a method of exploratory analysis of object-attribute data tables. The two main outputs are a hierarchical structure of clusters (so-called formal concepts) and a non-redundant basis of so-called attribute implications. An important topic in FCA is to cope with a possibly large number of resulting clusters. We propose a method to control the number of clusters by means of specification of a granularity level of attributes. A user selects an appropriate level of granularity of each attribute. If the corresponding set of clusters is too large, the user can select a lower level of granularity for appropriate attributes. The resulting set of clusters is then smaller and can be seen as a rougher version of the original set of clusters. If the corresponding set of clusters is too small, the user can select a finer level of granularity for appropriate attributes.

The resulting set of clusters is then larger and can be seen as a rougher version of the original set of clusters. The paper presents a preliminary study on this topic. We describe the motivations, the method, basic theoretical insight, and experiments demonstrating the method. Formal concept analysis (FCA) is a method of exploratory analysis of object-attribute data tables. The two main outputs are a hierarchical structure of clusters (so-called formal concepts) and a non-redundant basis of so-called attribute implications. An important topic in FCA is to cope with a possibly large number of resulting clusters. We propose a method to control the number of clusters by means of specification of a granularity level of attributes. A user selects an appropriate level of granularity of each attribute. If the corresponding set of clusters is too large, the user can select a lower level of granularity for appropriate attributes. The resulting set of clusters is then smaller and can be seen as a rougher version of the original set of clusters. If the corresponding set of clusters is too small, the user can select a finer level of granularity for appropriate attributes. The resulting set of clusters is then larger and can be seen as a rougher version of the original set of clusters. The paper presents a preliminary study on this topic. We describe the motivations, the method, basic theoretical insight, and experiments demonstrating the method.

I. INTRODUCTION AND PROBLEM SETTING

Formal concept analysis (FCA) [7], [8] is an exploratory method of analysis of tabular data describing objects and their attributes. One of the outputs of FCA is a hierarchical structure of clusters from the input data table. The clusters are called formal concepts and these are pairs \( \langle A, B \rangle \) consisting of a collection \( A \) of objects and a collection \( B \) of attributes. Formal concepts can be partially ordered by a natural subconcept-superconcept relation. The resulting partially ordered set, called a concept lattice, forms a complete lattice and can be visualized by a labeled Hasse diagram. Formal concepts \( \langle A, B \rangle \) result from the idea (going back to traditional Port-Royal logic) of a concept as consisting of its extent \( A \) and its intent \( B \). Alternatively, formal concepts can be thought of as maximal rectangles contained in the object-attribute data table. In the basic setting, the attributes are binary presence/absence attributes and the data table is a 0/1-matrix. More general attributes are handled by so-called conceptual scaling (see [8]), i.e. a suitable transformation of a general data table into a 0/1-data table which respects the meaning of attributes. Formal concepts are clusters of data drawn together by having common attributes. FCA has been applied in various fields, among others in software engineering, reengineering problems (redesign of hierarchical structures), text classification (analyzing e-mail collections, classification of library items), browsing retrieval and database views, psychology (study of development of concepts by children), civil engineering (system for checking dependencies in regulations), classification and systematizing of heuristic methods, physiology (color perception), preprocessing of data; see [7], [8] and the references therein.

A problem related to the direct user interpretation of a concept lattice is very often caused by the fact that the number of extracted formal concepts is not satisfactory. On the one hand, it may happen that the number of formal concepts is too large. Too large a number of formal concepts provides an overly fine granulation of the input objects and is difficult to interpret for the expert. On the other hand, it may happen that the number of formal concepts is too small. Too small a number of formal concepts does not provide a sufficiently fine granulation of the input objects for the expert.

In this paper, we present a method to control the number of extracted formal concepts by means of selecting levels of granularity of the attributes given in the input data. The paper is organized as follows. Section II presents preliminaries on formal concept analysis. In Section III we present our approach. Section IV contains illustrative examples. Section V presents conclusions and outlines future research.

II. PRELIMINARIES

Formal concept analysis deals with input data in the form of a table with rows corresponding to objects and columns corresponding to attributes which describes a relationship between the objects and attributes. The data table is formally represented by a so-called formal context which is a triplet \( \langle X, Y, I \rangle \) where \( I \) is a binary relation between \( X \) and \( Y \), \( \langle x, y \rangle \in I \) meaning that the object \( x \) has the attribute \( y \). For each \( A \subseteq X \) denote by \( A^\uparrow \) a subset of \( Y \) defined by
\[
A^\uparrow = \{ y \mid \text{for each } x \in A : \langle x, y \rangle \in I \}.
\]
Similarly, for \( B \subseteq Y \) denote by \( B^\downarrow \) a subset of \( X \) defined by
\[
B^\downarrow = \{ x \mid \text{for each } y \in B : \langle x, y \rangle \in I \}.
\]
That is, \( A^1 \) is the set of all attributes from \( Y \) shared by all objects from \( A \), and \( B^1 \) is the set of all objects from \( X \) sharing all attributes from \( B \). A formal concept in \( \langle X, Y, I \rangle \) is a pair \( \langle A, B \rangle \) of \( A \subseteq X \) and \( B \subseteq Y \) satisfying \( A^1 = B \) and \( B^1 = A \). That is, a formal concept consists of a set of \( A \) of objects which fall under the concept and a set \( B \) of attributes which fall under the concept such that \( A \) is the set of all objects sharing all attributes from \( B \) and, conversely, \( B \) is the collection of all attributes from \( Y \) shared by all objects from \( A \). This definition formalizes the traditional approach to concepts which is due to Port-Royal logic. The sets \( A \) and \( B \) are called the extent and the intent of the concept \( \langle A, B \rangle \), respectively. The set

\[
\mathcal{B}(X,Y,I) = \{\langle A,B \rangle \mid A^1 = B, B^1 = A\}
\]

of all formal concepts in \( \langle X, Y, I \rangle \) can be naturally equipped with a partial order \( \leq \) defined by \( \langle A_1,B_1 \rangle \leq \langle A_2,B_2 \rangle \) iff \( A_1 \subseteq A_2 \) (or, equivalently, \( B_2 \subseteq B_1 \)). That is, \( \langle A_1,B_1 \rangle \leq \langle A_2,B_2 \rangle \) means that each object from \( A_1 \) belongs to \( A_2 \) (or, equivalently, each attribute from \( B_2 \) belongs to \( B_1 \)). Therefore, \( \leq \) models the natural subconcept-superc concept hierarchy under which \( \text{dog} \) is a subconcept of \( \text{mammal} \), etc.

The structure of \( \mathcal{B}(X,Y,I) \) is described by the so-called main theorem of concept lattices [8]:

**Theorem 1:**

1. \( \mathcal{B}(X,Y,I) \) is under \( \leq \) a complete lattice where the infima and suprema are given by

\[
\wedge_{j \in J} \langle A_j, B_j \rangle = \left( \bigcap_{j \in J} A_j, \bigcup_{j \in J} B_j \right)^{11},
\]

\[
\vee_{j \in J} \langle A_j, B_j \rangle = \left( \bigcup_{j \in J} A_j, \bigcap_{j \in J} B_j \right) .
\]

2. Moreover, an arbitrary complete lattice \( V = \langle V, \leq \rangle \) is isomorphic to \( \mathcal{B}(X,Y,I) \) iff there are mappings \( \gamma : X \to V \), \( \mu : Y \to V \) such that

(i) \( \gamma(x) \) is \( \vee \)-dense in \( V \), \( \mu(y) \) is \( \wedge \)-dense in \( V \);

(ii) \( \gamma(x) \leq \mu(y) \) iff \( \langle x, y \rangle \in I \).

**III. GRANULARITY OF ATTRIBUTES**

In the basic setting of FCA, no further information except for the input data table \( T = \langle X, Y, I \rangle \) is taken into account. In this section, we present a possibility to have, instead of \( \langle X, Y, I \rangle \), a more structured input data which allows us to control the granularity of the object attributes. The main gain is that this way we get a means to control the number of formal concepts extracted from the input data.

Granulation is an important phenomenon performed successfully by humans in everyday life. Basically, granulation means considering a collection of pieces of the outer world as a whole—a granule. For example, looking at a person, we may distinguish her head, left arm, right arm, etc. Then, the head, left arm, right arm, etc., are granules for us. The granules might serve as basic units with which our reasoning is concerned. Depending on a particular situation, we might want to use finer or rougher granules, i.e. to increase or decrease a level of granularity. A finer granularity usually leads to a more precise reasoning at the cost of higher computational demands.

A rougher granularity leads to a less precise reasoning with the benefit of being less computationally demanding. For instance, we increase the level of granularity, if we distinguish nose, ears, eyes, etc., instead of distinguishing only a head.

The importance of the phenomenon of granulation and its role in data manipulation and reasoning has been repeatedly emphasized by Zadeh [12].

Granulation and granularity levels naturally appear in tabular data describing objects and their attributes. For instance, consider the data depicted in Tab. I. The tables describe objects \( a, \ldots, g \). The left table describes the objects by means of their attributes \( \text{L} \) (large), \( \text{R} \) (red), and \( \text{G} \) (green). The right table, however, uses attributes \( \text{LG} \) (light green) and \( \text{DG} \) (dark green) instead of a single attribute green. Attributes \( \text{IG} \) and \( \text{gG} \) provide a higher level of granularity for the description of the objects than a single attribute \( \text{G} \). Namely, attributes \( \text{IG} \) and \( \text{gG} \) may be seen as a refinement of \( \text{G} \).

The left table of Tab. II shows all the formal concepts extracted from the left table of Tab. I (denote this collection of formal concepts by \( B_1 \)), the right table of Tab. II shows all the formal concepts extracted from the right table of Tab. I (denote this collection by \( B_2 \)). As one can see from Tab. II, the increase of granularity level represented replacing attribute \( \text{G} \) by attributes \( \text{IG} \) and \( \text{gG} \) leads to an increase in the number of extracted formal concepts. Nevertheless, we can see that there is a natural relationship between \( B_1 \) and \( B_2 \). Namely, \( B_2 \) can be seen as “refinement” of \( B_1 \) in the sense that it contains finer concepts than \( B_1 \). For instance, instead of a “rougher” formal concept no. 6 from \( B_1 \) with its extent \( \{a,b,c\} \) and its intent \( \{L,G\} \), \( B_2 \) contains two finer formal concepts, namely,
Definition 2: Let \( X \) be a set of objects. A gl-tree (granularity-level tree) for attribute \( y \) is a rooted tree with the following properties:

- Each node of the tree is labeled by a symbol (denoted usually by \( y, y_1, z, \ldots \) and called an attribute); the root is denoted by \( y \);
- To each label \( z \) of a node there is associated a set \( \{ z \}_1^{\text{def}} \subseteq X \); objects from \( \{ z \}_1^{\text{def}} \) are considered as objects to which attribute \( z \) applies;
- If a node labeled by \( z \) has as its successors nodes labeled by \( z_1, z_2, \ldots, z_n \), then \( \{ z_1 \}_1^{\text{def}}, \{ z_2 \}_1^{\text{def}}, \ldots, \{ z_n \}_1^{\text{def}} \) is a partition of \( \{ z \}_1^{\text{def}} \).

Remark 1: (1) The fact that \( \{ z_1 \}_1^{\text{def}}, \{ z_2 \}_1^{\text{def}}, \ldots, \{ z_n \}_1^{\text{def}} \) is a partition of \( \{ z \}_1^{\text{def}} \) means that, first, each \( \{ z_i \}_1^{\text{def}} \) is non-empty; second, for \( i \neq j \), we have \( \{ z_i \}_1^{\text{def}} \cap \{ z_j \}_1^{\text{def}} = \emptyset \) (attributes \( z_i \) and \( z_j \) are disjoint); third, \( \{ z_1 \}_1^{\text{def}} \cup \cdots \cup \{ z_n \}_1^{\text{def}} = \{ z \}_1^{\text{def}} \) (attributes \( z_1, z_2, \ldots, z_n \) cover \( \{ z \}_1^{\text{def}} \)).

(2) The definition of a structure describing several levels of granularity may be more general. We work with the above definition for the sake of simplicity.

Example 1: Consider Tab. I. Then one may consider a simple gl-tree for attribute \( G \) with a root labeled by \( G \), two successors of the root, labeled by \( IG \) and \( dG \), and the corresponding sets of objects given by \( \{ G \}_1^{\text{def}} = \{ a, b, c, d, e \}, \{ IG \}_1^{\text{def}} = \{ a, b, d \}, \{ dG \}_1^{\text{def}} = \{ c, e \} \).

A selection of an appropriate level of granularity can be described by the following notion of a cut in a gl-tree:

Definition 3: A cut in a gl-tree for \( y \) is a set \( C = \{ y_1, \ldots, y_n \} \) of labels of nodes of the gl-tree such that for each leaf node \( u \), there exists exactly one node \( v \) on the path from the root to \( u \) such that the label of \( v \) belongs to \( C \).

Remark 2: (1) In other words, \( C \) is a cut if and only if \( \{ y_1 \}_1^{\text{def}}, \ldots, \{ y_n \}_1^{\text{def}} \) is a partition of \( \{ y \}_1^{\text{def}} \). In formally, a cut is a refinement of attribute \( y \) which can be obtained by moving down the paths of the tree, starting in the root.

(2) For example, \( \{ G \}_1^{\text{def}} \) and \( \{ IG, dG \}_1^{\text{def}} \) are the only cuts of the gl-tree from Example 1.

The relation of a refinement induces a partial order on the set of all cuts of a given gl-tree by putting for cuts \( C_1 = \{ y_1, \ldots, y_n \} \) and \( C_2 = \{ z_1, \ldots, z_m \} \),

\[ C_1 \leq C_2 \quad \text{iff} \quad \{ y_1 \}_1^{\text{def}}, \ldots, \{ y_n \}_1^{\text{def}} \] 

is a subpartition of \( \{ z_1 \}_1^{\text{def}}, \ldots, \{ z_m \}_1^{\text{def}} \). For instance, \( \{ IG, dG \}_1^{\text{def}} \subseteq \{ G \}_1^{\text{def}} \), cf. above.

Let now \( \langle X, Y, I \rangle \) be an input data table. Suppose that we have for each attribute \( y \in Y \) a gl-tree \( T_y \) for \( y \). Let for each \( y \in Y \), \( C_y \) be a cut in \( T_y \) and denote by

\[ C = \{ C_y \mid y \in Y \} \]

the collection of all these cuts. Each such a collection \( C \) induces a data table \( \langle X, Y_C, I_C \rangle \) such that

\[ Y_C = \bigcup_{y \in Y} C_y \]

and we put for each \( z \in Y_C \)

\[ (x, z) \in I_C \quad \text{iff} \quad x \in \{ z \}_1^{\text{def}}. \]

That is, \( \langle X, Y_C, I_C \rangle \) results from \( \langle X, Y, I \rangle \) by replacing each attribute \( y \in Y \) by the corresponding collection \( C_y \) of attributes (refinements of \( y \)).

Example 2: For the example from Tab. I, putting \( C_1 \) and \( C_2 \) as \( \langle X, Y, I \rangle \) and \( \langle X, Y_C, I_C \rangle \) of Tab. I and \( \langle X, Y_C, I_C \rangle \) of Tab. I, we have \( \langle X, Y_C, I_C \rangle \) and \( \langle X, Y_C, I_C \rangle \) according to intuition.

Denote the concept lattice corresponding to \( \langle X, Y_C, I_C \rangle \) by \( B(X, Y_C, I_C) \) or simply by \( B_C \).

Each collection \( C \) of cuts represents a selection of levels of granularity of the attributes under consideration. Now, the main question we are interested in the following: given two collections \( C_1 \) and \( C_2 \) of levels of granularity, what is the relationship between the corresponding concept lattices \( B(X, Y_{C_1}, I_{C_1}) \) and \( B(X, Y_{C_2}, I_{C_2}) \)? Due to the limited scope, we restrict ourselves to the condition when \( C_1 \) is a refinement of \( C_2 \), denoted by

\[ C_1 \leq C_2, \]

meaning that for each \( y \in Y \) we have \( C_1 y \leq C_2 y \) for the corresponding cuts \( C_1 y \subseteq C_2 y \) in \( C_2 \).

Theorem 4: If \( C_1 \leq C_2 \) then for each formal concept \( \langle A, B \rangle \in B(X, Y_{C_2}, I_{C_2}) \) there are formal concepts \( \langle A_k, B_k \rangle \in B(X, Y_{C_1}, I_{C_1}) \), \( k \in K \), such that \( A = \bigcup_{k \in K} A_k \).

Proof: Omitted due to lack of space.

The previous theorem says that if we refine our attributes then the extent (cluster of objects) of each formal concept from the concept lattice of the “finer attributes”. We omit further theoretical description of the relationships due to lack of space.

IV. ILLUSTRATIVE EXAMPLES

We now present illustrative examples. We use Hasse diagrams and label the nodes corresponding to formal concepts by boxes containing concept descriptions. For example, \( \{1, 3, 7\}, \{a, b\} \) is a concept with extent \( \{1, 3, 7\} \) and intent \( \{a, b\} \). Consider a data table described in Tab. III. The
<table>
<thead>
<tr>
<th>Accident</th>
<th>Cause</th>
<th>Day</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>speed</td>
<td>thursday</td>
<td>9-10</td>
</tr>
<tr>
<td>2</td>
<td>alcohol</td>
<td>friday</td>
<td>23-24</td>
</tr>
<tr>
<td>3</td>
<td>priority</td>
<td>saturday</td>
<td>9-10</td>
</tr>
<tr>
<td>4</td>
<td>priority</td>
<td>monday</td>
<td>9-10</td>
</tr>
<tr>
<td>5</td>
<td>brakes</td>
<td>saturday</td>
<td>10-11</td>
</tr>
<tr>
<td>6</td>
<td>steering</td>
<td>thursday</td>
<td>12-13</td>
</tr>
<tr>
<td>7</td>
<td>steering</td>
<td>sunday</td>
<td>10-11</td>
</tr>
<tr>
<td>8</td>
<td>brakes</td>
<td>monday</td>
<td>10-11</td>
</tr>
<tr>
<td>9</td>
<td>speed</td>
<td>monday</td>
<td>1-2</td>
</tr>
</tbody>
</table>

**TABLE III**

*Formal context given by accidents and their properties.*

The table represents data about nine car accidents (accident 1, ..., accident 9) and their three attributes (cause, day, time). Attribute cause describes the cause of the accident, attribute day describes the day the accident happened, and attribute time describes the time interval when the accident happened.

The accidents represent the objects, i.e., \( X \) has nine elements. We denote the accidents by their numbers only, i.e., \( X = \{1, \ldots, 9\} \). In order to obtain data tables with binary attributes, we may consider the following binary attributes and the corresponding gl-trees:

- Attributes related to cause: The most general attribute (root of the tree) will be called “cause”, it is the label of the root of the tree and we have \( \{\text{cause}\}^1 = X \).
- Attributes related to time: The most general attribute (root of the tree) will be called “time” and it has attributes “working day” and “weekend”, and “monday”, dots, “sunday” as refinements (we omit details).
- Attributes related to day: The most general attribute (root of the tree) will be called “day” and it has attributes “working day” and “weekend”, and “monday”, dots, “sunday” as refinements (we omit details).

Consider first a selection of granularity levels given by a collection \( C = \{C_c, C_d, C_t\} \) with \( C_c = \{\text{technical cause}, \text{driver fault}\} \), \( C_d = \{\text{working day}, \ldots, \text{weekend}\} \), \( C_t = \{\text{daytime}, \text{night}\} \). The corresponding concept lattice is depicted in Fig. 2.

Suppose now the user find the formal concepts too fine and their number too large. The user can select other granularity levels, e.g., those represented by a collection \( C = \{C_c, C_d, C_t\} \) with \( C_c = \{\text{technical cause}, \text{driver fault}\} \), \( C_d = \{\text{technical cause}\} \), \( C_t = \{\text{daytime}, \text{night}\} \). The corresponding concept lattice is depicted in Fig. 3.

Both of the concept lattices provide a classification of the accidents. One can see that the concept lattice corresponding to a smaller level of granularity has a less number of formal concepts and that these concepts can be seen as providing a rougher granulation of the set of objects and thus provides a rougher classification.

V. CONCLUSIONS AND FUTURE RESEARCH

The paper presents preliminary results on incorporating the idea of granulations and levels of granularity into formal concept analysis. The main practical effect of the presented approach is a possibility to control, in a parameterized way, the number of extracted formal concepts from input data and to control the granulation by means of the formal concepts.

The future research will be focused on the following topics:

- Relationships to conceptual scaling. In the framework of FCA, many-valued (non-binary) attributes are handled by means of so-called conceptual scaling [8]. There is an obvious connection between the levels of attribute granularity considered in this paper and different scalings of many-valued attributes which needs to be explored.
- In addition to formal concepts, the other output of FCA is represented by so-called attribute implications [8]. The effect of changing granularity levels of attributes on the extracted attribute implications and interesting subsets of attribute implications (like non-redundant bases) will be investigated.
- The basic setting of FCA was generalized to fuzzy attributes in several papers, see e.g. [1], [2], [6], [10]. Investigation of the topics presented in this paper is a natural way to continue the state of art.
- We presented only basic theoretical insight; the next step is to look at further relationships between concept lattices.
Fig. 2. Concept lattice of a data table from Tab. III with $C = \{C_c, C_d, C_t\}$ and $C_c = \{\text{brakes, steering, alcohol, speed, priority}\}$, $C_d = \{\text{monday, ..., sunday}\}$, $C_t = \{0-1, \ldots, 23-0\}$.

Fig. 3. Concept lattice of a data table from Tab. III with $C = \{C_c, C_d, C_t\}$ and $C_c = \{\text{technical cause, driver fault}\}$, $C_d = \{\text{working day, ..., weekend}\}$, $C_t = \{\text{daytime, night}\}$.
of data with a changed level of granularity.

- **Algorithms.** Design of algorithms enabling to compute a new concept lattice corresponding to a data with an increased level of granularity with the help of the original concept lattice.

**ACKNOWLEDGMENT**

Supported by grant No. 1ET101370417 of GA AV ČR, by grant No. 201/05/0079 of the Czech Science Foundation. R. Bělohlávek acknowledges support by by institutional support, research plan MSM 6198959214.

**REFERENCES**


