GGEN 41006—17/3/2004—ADMINISTRATOR—97859

International Journal of General Systems, 2004 Vol. 00 (0), pp. 1–12

Taylor & Francis
Taylor & Francis Group

FORMAL CONCEPT ANALYSIS WITH HIERARCHICALLY ORDERED ATTRIBUTES

RADIM BĚLOHLÁVEK[a,b,*], VLADIMÍR SKLENÁŘ[a] and JIŘÍ ZACPAL[a]

[a]*Department of Computer Science, Palacký University, Tomkova 40, CZ-779 00 Olomouc, Czech Republic;* [b]*Institute of Research and Applications of Fuzzy Modeling, University of Ostrava, 30 Dubna 22, 701 03 Ostrava, Czech Republic*

(Received 22 October 2003; In final form 28 December 2003)

Formal concept analysis is a method of exploratory data analysis that aims at the extraction of natural clusters from object–attribute data tables. The clusters, called formal concepts, are naturally interpreted as human-perceived concepts in a traditional sense and can be partially ordered by a subconcept–superconcept hierarchy. The hierarchical structure of formal concepts (so-called concept lattice) represents a structured information obtained automatically from the input data table.

The present paper focuses on the analysis of input data with a predefined hierarchy on attributes extending thus the basic approach of formal concept analysis. The motivation of the present approach derives from the fact that very often, people (consciously or unconsciously) attach various importance to attributes which is then reflected in the conceptual classification based on these attributes. We define the notion of a formal concept respecting the attribute hierarchy. Formal concepts which do not respect the hierarchy are considered not relevant. Elimination of the non-relevant concepts leads to a reduced set of extracted concepts making the discovered structure of hidden concepts more comprehensible. We present basic formal results on our approach as well as illustrating examples.

Keywords: Formal concept analysis, Concept lattice, Hierarchy of attributes, Clustering

1. INTRODUCTION AND PROBLEM SETTING

Finding interesting patterns in data has traditionally been a challenging problem. In the pre-computer era, the extent of efficiently analyzable data was small and the patterns looked for in the data were simple patterns easily recognizable and graspable by humans. With the invention of computers, the situation changed rapidly. Computers made it possible to analyze large amounts of data as well as to look for new kinds of patterns in the data. Formal concept analysis (Ganter and Wille, 1999) is an example of a method for finding patterns and dependencies in data which can be run automatically. The patterns looked for are called formal concepts. The attractiveness of formal concept analysis derives mainly from the fact that formal concepts are interpretable as natural concepts well-understood by humans.

*Corresponding author. Address: Department of Computer Science, Palacký University, Tomkova 40, CZ-779 00 Olomouc, Czech Republic. E-mail: radim.belohlavek@upol.cz

ISSN 0308-1079 print/ISSN 1563-5104 online © 2004 Taylor & Francis Ltd
DOI: 10.1080/03081070410001679715

Both foundations and applications (classification, software (re)engineering, document and text organization, etc.) of formal concept analysis are well-documented (Ganter and Wille, 1999 and the references therein).[†]

In its basic setting, 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

$$A^{\uparrow} = \{y \,|\, \text{for each } x \in A : \langle x,y\rangle \in I\}.$$

Similarly, for $B \subseteq Y$ denote by $B^{\downarrow}$ a subset of X defined by

$$B^{\downarrow} = \{x \,|\, \text{for each } y \in B : \langle x,y\rangle \in I\}.$$

That is, $A^{\uparrow}$ is the set of all attributes from Y shared by all objects from A (and similarly for $B^{\downarrow}$). 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 $A^{\uparrow} = B$ and $B^{\downarrow} = A$. That is, a formal concept consists of a set 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 (Arnauld and Nicole, 1662). The sets A and B are called the extent and the intent of the concept $\langle A,B\rangle$, respectively. For example, for the concept dog, the extent consists of all dogs while the intent consists of all attributes common to all dogs. The set $\mathcal{B}(X,Y,I) = \{\langle A,B\rangle | A^{\uparrow} = B, B^{\downarrow} = A\}$ of all formal concepts in $\langle X,Y,I\rangle$ can be naturally equipped with a partial order $\leq$ defined by

$$\langle A_1,B_1\rangle \leq \langle A_2,B_2\rangle \ \text{iff}\ A_1 \subseteq A_2 \ (\text{or, equivalently, } B_2 \subseteq B_1).$$

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–superconcept hierarchy under which dog is a subconcept of mammal.

Under $\leq$, $\mathcal{B}(X,Y,I)$ happens to be a complete lattice, called a concept lattice, the basic structure of which is described by the so-called main theorem of concept lattices (Ganter and Wille, 1999; Wille, 1982).

THEOREM 1

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

$$\bigwedge_{j\in J}\langle A_j,B_j\rangle = \left\langle \bigcap_{j\in J} A_j, \left(\bigcup_{j\in J} B_j\right)^{\downarrow\uparrow}\right\rangle, \bigvee_{j\in J}\langle A_j,B_j\rangle = \left\langle \left(\bigcup_{j\in J} A_j\right)^{\uparrow\downarrow}, \bigcap_{j\in J} B_j\right\rangle. \tag{1}$$

(2) Moreover, an arbitrary complete lattice $\mathbf{V} = \langle V,\leq\rangle$ 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 $\bigvee$-dense in V, $\mu(Y)$ is $\bigwedge$-dense in V;
  (ii) $\gamma(x) \leq \mu(y)$ iff $\langle x,y\rangle \in I$.

[†]WWW: http://www.upriss.org.uk/fca/fca.html, http://www.kvocentral.org/resources/fca.html

Formal concept analysis thus treats both the individual objects and the individual attributes as distinct entities for which there is no further information available except for the relationship I saying which objects have which attributes. However, more often than not, both the set of objects and the set of attributes are supplied by an additional information. Further processing of the input data (formal context) should therefore take the additional information into account. Particularly, conceptual clustering should take the additional information into account in such a way that only those concepts which are in an appropriate sense compatible with the additional information, are considered relevant. For example, some attributes may be relevant (or relevant to some degree) with respect to a particular kind of decisions while some may be not. When processing a respective formal context in order to get some support for the decisions in question, the attributes which are not relevant to the decision may be disregarded. In the end, this may result in a simplification of the overall processing.

Therefore, it seems reasonable to assume that instead of $\langle X,Y,I\rangle$, the input data consists of a richer structure $\langle X,Y,I,\ldots\rangle$ where $\ldots$ may contain further sets (elements of which may represents goals, weights of relevance, etc.), relations, functions, etc. In this paper, we consider one particular kind of additional information which has the form of a hierarchy of attributes expressing their relative importance. The hierarchy enables us to eliminate formal concepts which are not compatible with the hierarchy. An important effect of the elimination is a natural reduction of the size of the resulting conceptual structure making the structure more comprehensible. A formal treatment of our approach is presented in the second section. The third section presents illustrative examples and discussion.

2. CONCEPT LATTICES OF CONTEXTS WITH HIERARCHICALLY ORDERED ATTRIBUTES

When considering a collection of attributes that can be observed on certain objects, people naturally consider some attributes more relevant than others. For example, having animals as objects, one may recognize several attributes like “warm-blooded”, “predator”, “gray color”, etc. Most of experts will probably agree on that “gray color” is less important than both “predator” and “warm-blooded”. Depending on the situation, experts may furthermore consider “predator” less important than “warm-blooded”. It should be, however, noted that the relative importance of attributes depends solely on the particular situation in question and is left to expert’s discretion. The relative importance of attributes influences naturally the expert’s classification of objects and, in particular, the formation of natural concepts. In the above example, for instance, the concept “warm blooded animals” will be probably considered natural by experts. On the other hand, “gray animal” will be considered not natural, although “existing” in the data. The formal reason for this is that “gray animal” covers the attribute “gray color” without covering the more important attributes “predator” and “warm-blooded”. From this point of view, concepts “warm-blooded predator” and “warm-blooded predator which is gray” should be considered natural which is in accordance with intuition: both “warm-blooded predator” and “warm-blooded predator which is gray” are particularizations of the general concept “warm-blooded animal” which result by adding less and less important attributes. The above discussion leads to the following formal treatment.

Definition 2 A formal context with hierarchically ordered attributes (a HA-context, for short) is a structure $\langle X,Y,I,\trianglelefteq\rangle$ (written also $\langle X,\langle Y,\trianglelefteq\rangle,I\rangle$) where $\langle X,Y,I\rangle$ is a formal context and $\trianglelefteq$ is a binary relation on Y.

Our primary interpretation of $\trianglelefteq$ is the following: $y_1 \trianglelefteq y_2$ means that y_2 is at least as important as y_1 in the sense discussed above. That is, we are mostly interested in situations where $\trianglelefteq$ is a quasiorder (i.e. a reflexive and transitive relation) or where $\trianglelefteq$ somehow represents a quasiorder. Repeating the above-mentioned arguments in terms of formal concepts, any set (category) A of objects can be narrowed using y_1 by considering only those objects from A which have the attribute y_1, i.e. by considering $A \cap \{y_1\}^{\downarrow} = \{x \in A \mid \langle x, y_1\rangle \in I\}$. However, condition $y_1 \trianglelefteq y_2$ may express an expert opinion according to which it makes sense to use y_1 for selecting a subset (subcategory) of A only if all objects from A have the attribute y_2 (i.e. $A \subseteq \{y_2\}^{\downarrow}$). In other words, y_1 may be applied only after the application of y_2. That is, $\trianglelefteq$ represents a restriction on how categories of objects can be formed—only categories which are compatible with $\trianglelefteq$ are considered relevant. This justifies the following definition.

Definition 3 For a HA-context $\langle X,\langle Y, \trianglelefteq\rangle,I\rangle$, a formal concept $\langle A,B\rangle \in \mathscr{B}(X,Y,I)$ is called compatible with $\trianglelefteq$ if for each $y_1,y_2 \in Y$, $y_1 \in B$ and $y_1 \trianglelefteq y_2$ implies $y_2 \in B$.

Compatible concepts are thus certain formal concepts from $\mathscr{B}(X,Y,I)$ satisfying the above-discussed restriction with respect to $\trianglelefteq$. The set of all formal concepts from $\mathscr{B}(X,Y,I)$ which are compatible with $\trianglelefteq$ will be denoted by $\mathscr{B}(X,\langle Y, \trianglelefteq\rangle,I)$, i.e.

$$\mathscr{B}(X,\langle Y,\trianglelefteq\rangle,I) = \{\langle A,B\rangle \in \mathscr{B}(X,Y,I) \mid \text{ for each } y_1,y_2 : y_1 \in B, y_1 \trianglelefteq y_2 \text{ implies } y_2 \in B\}.$$

It is obvious that $\mathscr{B}(X,\langle Y,\mathrm{id}_Y\rangle,I) = \mathscr{B}(X,Y,I)$, i.e. if $\trianglelefteq$ is the identity on Y then any formal concept of $\mathscr{B}(X,Y,I)$ is compatible with $\trianglelefteq$. The same holds true for $\trianglelefteq = \emptyset$ (the restriction formulated by $\trianglelefteq$ is empty), i.e. $\mathscr{B}(X,\langle Y,\emptyset\rangle,I) = \mathscr{B}(X,Y,I)$. More generally, we can proceed as follows. For a formal context $\langle X,Y,I\rangle$ denote by $\preceq_Y$ the binary relation defined on Y by

$$y_1 \preceq_Y y_2 \quad \text{if and only if for each } x \in X : \langle x, y_1\rangle \in I \text{ implies } \langle x, y_2\rangle \in I. \tag{2}$$

In other words, $y_1 \preceq_Y y_2$ if and only if $\{y_1\}^{\downarrow} \subseteq \{y_2\}^{\downarrow}$, i.e. $\langle \{y_1\}^{\downarrow}, \{y_1\}^{\downarrow\uparrow}\rangle \leq \langle \{y_2\}^{\downarrow}, \{y_2\}^{\downarrow\uparrow}\rangle$ in the subconcept–superconcept hierarchy. Obviously, $\preceq_Y$ is a quasiorder (reflexive and transitive relation) on Y. We have the following theorem.

Theorem 4 $\mathscr{B}(X,\langle Y,\trianglelefteq\rangle,I) = \mathscr{B}(X,Y,I)$ if and only if for each $y_1,y_2 \in Y$, $y_1 \trianglelefteq y_2$ implies $y_1 \preceq_Y y_2$.

Proof Let $\mathscr{B}(X,\langle Y,\trianglelefteq\rangle,I) = \mathscr{B}(X,Y,I)$ and take any $y_1,y_2 \in Y$ with $y_1 \trianglelefteq y_2$. Consider $\langle \{y_1\}^{\downarrow}, \{y_1\}^{\downarrow\uparrow}\rangle \in \mathscr{B}(X,Y,I)$. Since $\mathscr{B}(X,\langle Y,\trianglelefteq\rangle,I) = \mathscr{B}(X,Y,I)$, we have $\langle \{y_1\}^{\downarrow}, \{y_1\}^{\downarrow\uparrow}\rangle \in \mathscr{B}(X,\langle Y,\trianglelefteq\rangle,I)$ and so, by definition, $y_2 \in B$, i.e. y_2 is shared by all objects shared by y_1 showing $y_1 \preceq_Y y_2$.

Conversely, suppose that $y_1 \trianglelefteq y_2$ implies $y_1 \preceq_Y y_2$. Take any $\langle A,B\rangle \in \mathscr{B}(X,Y,I)$, any $y_1 \in B$ and assume $y_1 \trianglelefteq y_2$. $y_1 \in B$ means that y_1 is shared by all objects from A and so since $y_1 \trianglelefteq y_2$ implies $y_1 \preceq_Y y_2$, y_2 shares all objects from A as well, whence $y_2 \in B$. Thus, $\langle A,B\rangle \in \mathscr{B}(X,\langle Y, \trianglelefteq\rangle,I)$ completing the proof. □

The next theorem shows a natural result saying that the more restrictions, the less formal concepts satisfying the restrictions.

Theorem 5 If $\trianglelefteq_1 \subseteq \trianglelefteq_2$ then $\mathscr{B}(X,\langle Y,\trianglelefteq_2\rangle,I) \subseteq \mathscr{B}(X,\langle Y,\trianglelefteq_1\rangle,I)$.

Proof Assume $\langle A,B\rangle \in \mathscr{B}(X,\langle Y, \trianglelefteq_2\rangle,I)$, $y_1 \in B$, and $y_1 \trianglelefteq_1 y_2$. Since $\trianglelefteq_1 \subseteq \trianglelefteq_2$, we have $y_1 \trianglelefteq_2 y_2$ and so $y_2 \in B$ by definition of $\mathscr{B}(X,\langle Y,\trianglelefteq_2\rangle,I)$. This shows that $\langle A,B\rangle \in \mathscr{B}(X,\langle Y,\trianglelefteq_1\rangle,I)$. □

TABLE I Formal context from Example 1

	y_1	y_2	y_3
x_1	1	1	0
x_2	1	0	0
x_3	0	1	1

We will need the following lemma.

LEMMA 6 For HA-contexts $\langle X,\langle Y, \unlhd_j\rangle,I\rangle$ ($j \in J$) we have

$$\mathscr{B}(X, \langle Y, \cup_{j\in J} \unlhd_j \rangle, I) = \bigcap_{j\in J} \mathscr{B}(X, \langle Y, \unlhd_j\rangle, I).$$

Proof Since $\unlhd_j \subseteq \cup_{j\in J} \unlhd_j$ for each $j \in J$, Theorem 5 gives

$$\mathscr{B}(X, \langle Y, \cup_{j\in J} \unlhd_j \rangle, I) \subseteq \mathscr{B}(X, \langle Y, \unlhd_j\rangle, I)$$

for each $j \in J$ and so

$$\mathscr{B}(X, \langle Y, \cup_{j\in J} \unlhd_j \rangle, I) \bigcap_{j\in J} \mathscr{B}(X, \langle Y, \unlhd_j\rangle, I).$$

On the other hand, let $\langle A, B\rangle \in \bigcap_{j\in J} \mathscr{B}(X, \langle Y, \unlhd_j\rangle, I)$. Let $y_1 \in B$ and $\langle y_1, y_2\rangle \in \cup_{j\in J} \unlhd_j$. Then $\langle y_1, y_2\rangle \in \unlhd_j$ for some $j \in J$. Since we assumed $\langle A, B\rangle \in \mathscr{B}(X, \langle Y, \unlhd_j\rangle, I)$, we get $y_2 \in B$. This means that $\langle A, B\rangle \in \mathscr{B}(X, \langle Y, \cup_{j\in J} \unlhd_j \rangle, I)$, completing the proof. □

Example 1 The following example shows that the dual form of Lemma 6 does not hold. Consider the formal context given in Table I. Consider furthermore the relations $\unlhd_1$ and $\unlhd_2$ on Y given by $\unlhd_1 = \{\langle y_3, y_1\rangle\}$ and $\unlhd_2 = \{\langle y_2, y_1\rangle\}$. Then

$$\mathscr{B}(X, \langle Y, \unlhd_1\rangle, I) \subset \mathscr{B}(X, \langle Y, \unlhd_2\rangle, I) \subset \mathscr{B}(X, \langle Y, \unlhd_1 \cap \unlhd_2\rangle, I)$$

and so

$$\mathscr{B}(X, \langle Y, \unlhd_1\rangle, I) \cup \mathscr{B}(X, \langle Y, \unlhd_2\rangle, I) \neq \mathscr{B}(X, \langle Y, \unlhd_1 \cap \unlhd_2\rangle, I) \text{ (See Fig.1).}$$

Given a formal context $\langle X,Y,I\rangle$ and a binary relation $\unlhd$ on Y, a natural question arises for what binary relations Q on Y we have $\mathscr{B}(X, \langle Y, \unlhd\rangle, I) = \mathscr{B}(X, \langle Y, Q\rangle, I)$, i.e. what Q are restrictive to the same extent as $\unlhd$. We will answer the question with respect to the operations of transitive closure and reflexive-transitive closure. Namely, given a binary relation $\unlhd$ on attributes specifying their relative importance, the transitive closure of $\unlhd$

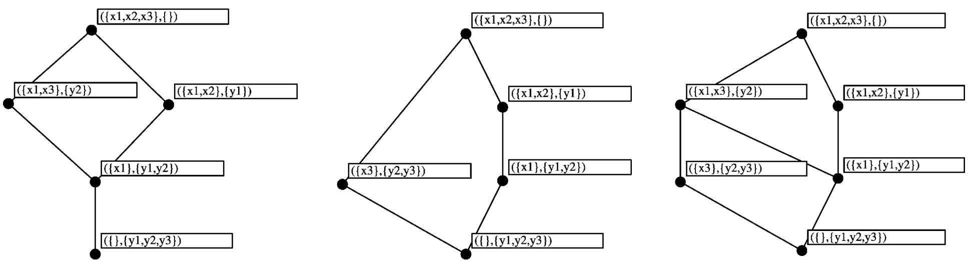

FIGURE 1 Concept lattices $\mathscr{B}(X, \langle Y, \unlhd_1\rangle, I)$, $\mathscr{B}(X, \langle Y, \unlhd_2\rangle, I)$ and $\mathscr{B}(X, \langle Y, \unlhd_1 \cap \unlhd_2\rangle, I)$ (from left to right) from Example 1.

represents an intuitively sound extension of $\trianglelefteq$. For a binary relation R, the transitive closure and the reflexive-transitive closure will be denoted by R^+ and R^*, respectively. We need the following lemma.

LEMMA 7 For an HA-context $\langle X, \langle Y, \trianglelefteq \rangle, I \rangle$ we have $\mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq^+ \rangle, I)$ and $\mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq^* \rangle, I)$.

Proof First, we show $\mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq^+ \rangle, I)$. Since $\trianglelefteq \subseteq \trianglelefteq^+$, Theorem 5 yields $\mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I) \supseteq \mathscr{B}(X, \langle Y, \trianglelefteq^+ \rangle, I)$. Conversely, let $\langle A, B \rangle \in \mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I)$, $y_1 \in B$ and $y_1 \trianglelefteq^+ y_2$. We have to show $y_2 \in B$. As $y_1 \trianglelefteq^+ y_2$, the well-known description of transitive closure yields that there are $z_1, \ldots, z_n \in Y$ such that $y_1 \trianglelefteq z_1, z_1 \trianglelefteq z_2, \ldots, z_{n-1} \trianglelefteq z_n, z_n \trianglelefteq y_2$. Taking into account that $\langle A, B \rangle \in \mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I)$, i.e. $\langle A,B \rangle$ is compatible with $\trianglelefteq$, we get $y_1 \in B$ and $y_1 \trianglelefteq z_1$ implies $z_1 \in B$ which, together with $z_1 \trianglelefteq z_2$ implies $z_2 \in B$. Repeating this argument, we finally get $y_2 \in B$ showing that $\langle A,B \rangle$ is compatible with $\trianglelefteq^+$, i.e. $\langle A, B \rangle \in \mathscr{B}(X, \langle Y, \trianglelefteq^+ \rangle, I)$.

The fact $\mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq^* \rangle, I)$ follows from the first fact, Lemma 6, and Theorem 4 by $\mathscr{B}(X, \langle Y, \trianglelefteq^* \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq^+ \cup \mathrm{id}_Y \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq^+ \rangle, I) \cap \mathscr{B}(X, \langle Y, \mathrm{id}_Y \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq^+ \rangle, I) \cap \mathscr{B}(X, Y, I) = \mathscr{B}(X, \langle Y, \trianglelefteq^+ \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I)$. □

For a binary relation Q on Y, we denote by Q^2 the composition $Q \circ Q$, i.e. $\langle y_1, y_2 \rangle \in Q^2$ iff there is some $y \in Y$ such that $\langle y_1, y \rangle \in Q$ and $\langle y, y_2 \rangle \in Q$. Now we have the following theorem.

THEOREM 8 For an HA-context $\langle X, \langle Y, \trianglelefteq \rangle, I \rangle$ we have

$$\mathscr{B}(X, \langle Y, \trianglelefteq - \trianglelefteq^2 \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq^+ \rangle, I).$$

Furthermore, for each binary relation Q on Y satisfying $\trianglelefteq - \trianglelefteq^2 \subseteq Q \subseteq \trianglelefteq^+$ we have $\mathscr{B}(X, \langle Y, Q \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I)$.

Proof It is well-known that $\trianglelefteq - \trianglelefteq^2$ (the transitive reduction of $\trianglelefteq$) satisfies $(\trianglelefteq - \trianglelefteq^2)^+ = \trianglelefteq^+$. Therefore, the first part follows directly by Lemma 7.

For the second part, Theorem 5 yields $\mathscr{B}(X, \langle Y, \trianglelefteq - \trianglelefteq^2 \rangle, I) \supseteq \mathscr{B}(X, \langle Y, Q \rangle, I) \supseteq \mathscr{B}(X, \langle Y, \trianglelefteq^+ \rangle, I)$. The assertion therefore follows by the first part. □

Theorem 8 shows natural bounds (in terms of transitive reduction and closure) on relation Q which are equally restrictive as $\trianglelefteq$.

Analogously, one can show the following theorem for reflexive and transitive closure and reduction.

THEOREM 9 For an HA-context $\langle X, \langle Y, \trianglelefteq \rangle, I \rangle$ we have

$$\mathscr{B}(X, \langle Y, \trianglelefteq - \trianglelefteq^2 - \mathrm{id}_Y \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq^* \rangle, I).$$

Furthermore, for each binary relation Q on Y satisfying $\trianglelefteq - \trianglelefteq^2 - \mathrm{id}_Y \subseteq Q \subseteq \trianglelefteq^*$ we have $\mathscr{B}(X, \langle Y, Q \rangle, I) = \mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I)$.

The restriction of the subconcept–superconcept hierarchy $\leq$ which is defined on $\mathscr{B}(X,Y,I)$ makes $\langle X, \langle Y, \trianglelefteq \rangle, I \rangle$ itself a partially ordered set $\langle \mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I), \leq \rangle$. The following theorem shows that $\langle X, \langle Y, \trianglelefteq \rangle, I \rangle$ is itself a complete lattice which is a reasonable substructure of the whole concept lattice $\mathscr{B}(X,Y,I)$.

THEOREM 10 $\mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I)$ equipped with $\leq$ is a complete lattice in which arbitrary suprema coincide with suprema in $\mathscr{B}(X,Y,I)$, i.e. it is a complete $\bigvee$-sublattice of $\mathscr{B}(X,Y,I)$.

Proof Take any $\langle A_i, B_i \rangle \in \mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I)$. We claim that $\langle A, B \rangle = \langle (\bigcap_i B_i)^{\downarrow}, \bigcap_i B_i \rangle$ belongs to $\mathscr{B}(X, \langle Y, \trianglelefteq \rangle, I)$. Indeed, if $y_1 \trianglelefteq y_2$ for some $y_1, y_2 \in B$, then $y_1, y_2 \in B_i$ for each $i \in I$.

TABLE II Formal context from Example 2

	y_1	y_2	y_3	y_4
x_1	1	0	1	1
x_2	0	0	1	0
x_3	0	0	0	1

$\langle A_i, B_i\rangle \in \mathscr{B}(X, \langle Y, \trianglelefteq\rangle, I)$ now implies that $y_2 \in B_i$ (for each i) and so $y_2 \in B$. Since $\langle A, B\rangle$ is the supremum of $\langle A_i,B_i\rangle$'s in $\mathscr{B}(X,Y,I)$ and since $\mathscr{B}(X,\langle Y, \trianglelefteq \rangle,I)$ is a subset of $\mathscr{B}(X,Y,I)$, we conclude that $\langle A,B\rangle$ is the supremum of $\langle A_i,B_i\rangle$'s in $\mathscr{B}(X,\langle Y, \trianglelefteq \rangle,I)$ as well.

Observe, furthermore, that $\langle Y^{\downarrow},Y\rangle \in \mathscr{B}(X,\langle Y, \trianglelefteq \rangle,I)$. Now, $\mathscr{B}(X,\langle Y, \trianglelefteq \rangle,I)$ is a partially ordered set with a least element $\langle Y^{\downarrow},Y\rangle$ in which there exist arbitrary suprema from which it follows that $\mathscr{B}(X,\langle Y, \trianglelefteq \rangle,I)$ is a complete lattice. □

The following example shows that infima in $\mathscr{B}(X,\langle Y, \trianglelefteq \rangle,I)$ do not generally coincide with infima in $\mathscr{B}(X,Y,I)$.

Example 2 Consider the formal context from Table II and a relation $\trianglelefteq$ given by $\trianglelefteq = \{\langle y_1, y_2\rangle\}$. Then for $\langle A_1, B_1\rangle = \langle\{x_1, x_2\}, \{y_3\}\rangle$ and $\langle A_2, B_2\rangle = \langle\{x_1, x_3\}, \{y_4\}\rangle$ from $\mathscr{B}(X,\langle Y, \trianglelefteq \rangle,I)$, their infimum in $\mathscr{B}(X,Y,I)$ is $\langle(B_1 \cup B_2)^{\downarrow}, (B_1 \cup B_2)^{\downarrow\uparrow}\rangle = \langle\{x_1\}, \{y_1, y_3, y_4\}\rangle$ which does not belong to $\mathscr{B}(X,\langle Y, \trianglelefteq \rangle,I)$ (Fig. 2). The infimum of $\langle A_1, B_1\rangle$ and $\langle A_2, B_2\rangle$ in $\mathscr{B}(X,\langle Y, \trianglelefteq \rangle,I)$ is $\langle\emptyset,\{y_1,y_2,y_3,y_4\}\rangle$.

3. EXAMPLES AND DISCUSSION

We now present illustrative examples. We assume that the reader is familiar with Hasse diagrams which will be used for visualization of concept lattices and attribute hierarchies. We label the nodes corresponding to formal concepts by boxes containing concept descriptions. For example, ({1,3,7},{3,4}) is a description of a concept the extent of which consists of objects 1, 3, and 7, and the intent of which consists of attributes 3 and 4.

Consider the formal context $\langle X,Y,I\rangle$ in Table III. The context contains plants and animals as the objects (labeled 1–12) and some of their properties as the attributes (labeled 1–8). The concept lattice $\mathscr{B}(X,Y,I)$ corresponding to formal concept $\langle X,Y,I\rangle$ contains 22 formal concepts and is depicted in Fig. 3. The formal concepts of $\mathscr{B}(X,Y,I)$ represent all concept-clusters that are present in the data. No attention is paid to importance or relative importance of attributes. Figure 4 shows the induced quasiorder $\preceq_Y$ (which is in fact a partial order in our example) as defined by Eq. (2). Let us now consider some attribute hierarchies $\trianglelefteq$ and the corresponding constrained concept lattices $\mathscr{B}(X,\langle Y, \trianglelefteq \rangle,I)$.

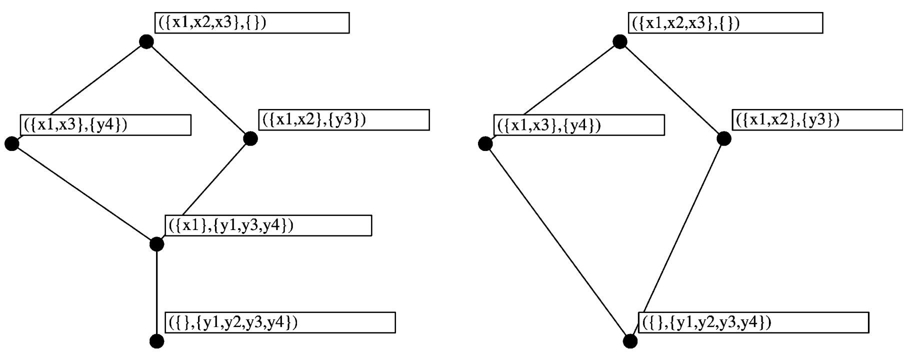

FIGURE 2 Concept lattices $\mathscr{B}(X,Y,I)$ (left) and $\mathscr{B}(X,\langle Y, \trianglelefteq \rangle,I)$ from Example 2.

GGEN 41006—17/3/2004—ADMINISTRATOR—97859

TABLE III Formal context given by plants and animals and their properties

		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1	Beech	0	0	1	0	0	0	1	1
2	Pond lily	0	0	0	0	1	0	1	0
3	Carp	1	0	0	1	0	0	0	0
4	Sheep	1	1	1	1	0	0	0	0
5	Yew	0	0	1	0	0	1	1	1
6	Culex	1	0	1	0	0	0	0	0
7	Raspberry	0	0	1	1	0	0	1	0
8	Dog	1	1	1	0	0	0	0	0
9	Newt	1	0	0	0	1	0	0	0
10	Adder	1	0	1	0	1	1	0	0
11	Apple tree	0	0	1	1	0	0	1	1
12	Deer	1	0	1	1	0	0	0	0

Attributes: 1—animal, 2—domestic, 3—terrestrial, 4—food, 5—protected, 6—poisonous, 7—plant, 8—tree.

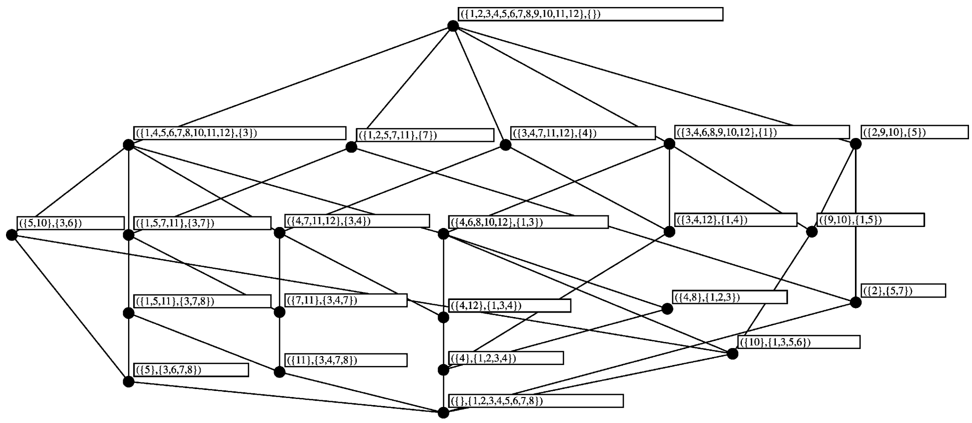

FIGURE 3 Concept lattice corresponding to the context from Table III.

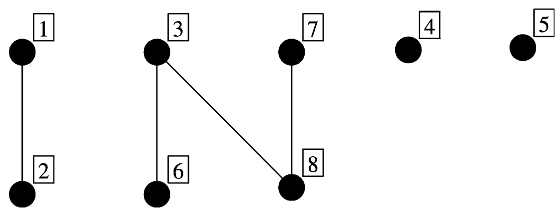

FIGURE 4 Partial order $\preceq_Y$ induced from the context from Table III by Eq. (2).

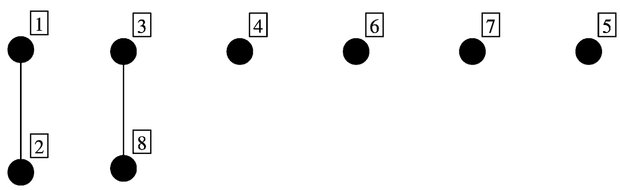

FIGURE 5 Attribute hierarchy $\trianglelefteq_1$ which is included in $\preceq_Y$.

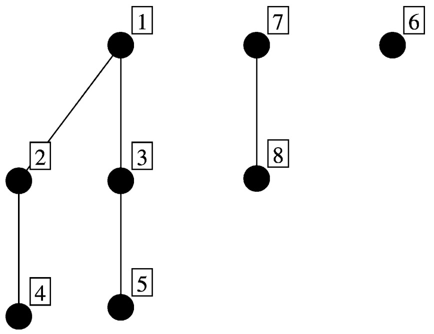

FIGURE 6 Attribute hierarchy $\trianglelefteq_2$.

First, consider a binary relation $\trianglelefteq_1$ on Y shown in Fig. 5. As it is easily seen, that $\trianglelefteq_1$ is included in $\preceq_Y$. Therefore, by Theorem 4, $\mathscr{B}(X,\langle Y, \trianglelefteq_1\rangle,I)$ coincides with $\mathscr{B}(X,Y,I)$, i.e. $\trianglelefteq_1$ imposes no constraints.

Second, consider hierarchy $\trianglelefteq_2$ of attributes depicted in Fig. 6. $\trianglelefteq_2$ expresses an expert opinion according to which attribute 1 (being animal) is more important than attribute 3 (being terrestrial) which is more important than attribute 5 (being protected), etc. Figure 7 shows the corresponding constrained concept lattice $\mathscr{B}(X,\langle Y,\trianglelefteq_2\rangle,I)$. $\mathscr{B}(X,\langle Y,\trianglelefteq_2\rangle,I)$ contains only those formal concepts from $\mathscr{B}(X,Y,I)$ which are compatible with $\trianglelefteq_2$ in the sense of Definition 3. For example, formal concept from $\mathscr{B}(X,Y,I)$ labeled ({1,4,5,6,7,8,10,11,12},{3}) is not compatible with $\trianglelefteq_2$ since its intent contains attribute 3 without containing the more important attribute 1. Therefore, ({1,4,5,6,7,8,10,11,12},{3}) is considered not relevant and is not contained in $\mathscr{B}(X,\langle Y,\trianglelefteq_2\rangle,I)$. Note that situations like

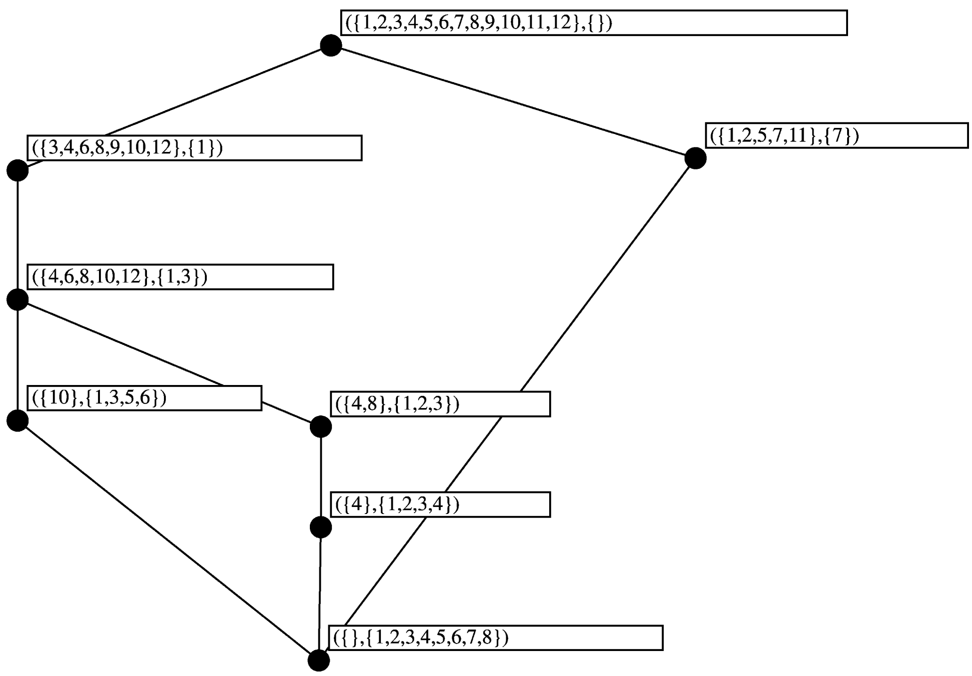

FIGURE 7 Constrained concept lattice $\mathscr{B}(X,\langle Y,\trianglelefteq_2\rangle,I)$.

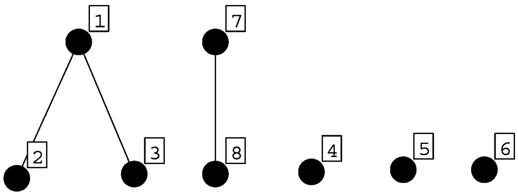

FIGURE 8 Attribute hierarchy $\trianglelefteq_3$.

this one occur very frequently. Namely, formal concept ({1,4,5,6,7,8,10,11,12},{3}) contains both terrestrial animals and terrestrial plants. However, an expert may be interested in such a classification scheme which first separates animals and only within the class of animals selects the terrestrial ones. From this point of view, the concept "terrestrial organisms" (i.e. concept ({1,4,5,6,7,8,10,11,12},{3}) in our context) is not natural and can be discarded. Formal concept analysis of data with hierarchical attributes thus provides a formal treatment of this problem.

Third, let us weaken $\trianglelefteq_2$ and consider hierarchy $\trianglelefteq_3$ from Fig. 8, Since $\trianglelefteq_3$ is contained in $\trianglelefteq_2$, Theorem 5 yields that $\mathscr{B}(X,\langle Y,\trianglelefteq_2\rangle,I) \subseteq \mathscr{B}(X,\langle Y,\trianglelefteq_3\rangle,I)$ as can be seen in Fig. 9. On the other hand, strengthening $\trianglelefteq_2$ and considering, for instance, the hierarchy $\trianglelefteq_4$ depicted in Fig. 10, we obtain a constrained concept lattice $\mathscr{B}(X,\langle Y,\trianglelefteq_4\rangle,I)$ depicted in Fig. 11. Again, we may check $\mathscr{B}(X,\langle Y,\trianglelefteq_4\rangle,I) \subseteq \mathscr{B}(X,\langle Y,\trianglelefteq_2\rangle,I)$ which follows from Theorem 5.

By Theorem 10, each of the constrained concept lattices discussed above is itself a complete lattice which is a $\bigvee$-sublattice of $\mathscr{B}(X,Y,I)$.

The above examples illustrate both of the two main contributions of the present approach. First, the present approach makes it possible to work with a natural kind of information very often accompanying a given set of attributes—the hierarchy of attributes. Considering only formal concepts which are compatible with a given attribute hierarchy $\trianglelefteq$ (i.e. instead of $\mathscr{B}(X,Y,I)$, considering only $\mathscr{B}(X,\langle Y,\trianglelefteq\rangle,I)$) can be seen as keeping only formal concepts that

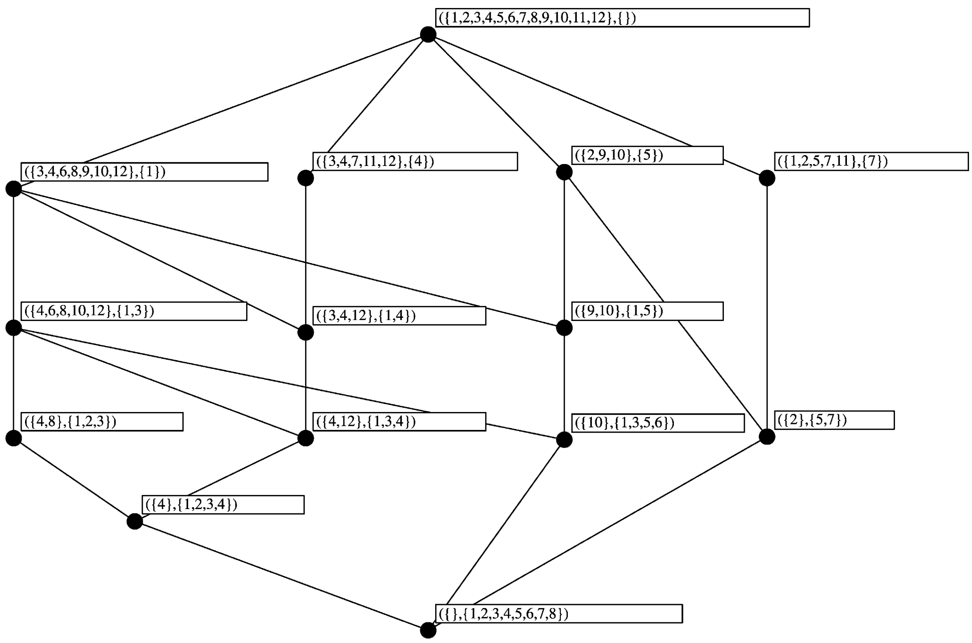

FIGURE 9 Constrained concept lattice $\mathscr{B}(X,\langle Y,\trianglelefteq_3\rangle,I)$.

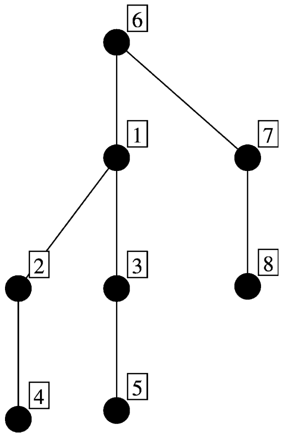

FIGURE 10 Attribute hierarchy $\trianglelefteq_4$.

reflect the additional information specified by $\trianglelefteq$ (and are in this sense natural) and filtering out formal concepts that do not obey $\trianglelefteq$. Second, an important side-effect of the present approach is the reduction of the number of formal concepts extracted from a given formal context. It is well-known that for large formal contexts, the resulting concept lattice may contain a large number of formal concepts. This can make the concept lattice hardly directly graspable by humans and reduction techniques like factorization and decomposition of the concept lattice have to be applied (Ganter and Wille, 1999). Since the present state of art of formal concept analysis does not offer a way to cope with attribute hierarchies, the reduction techniques have to be applied even if the user is able to formulate a natural attribute hierarchy. From this point of view the present approach offers a direct way to reduce the number of formal concepts by keeping only concepts compatible with an attribute hierarchy.

From a more general point of view, the present approach is but one particular case of a more general idea of constraining concept lattices by natural conditions resulting from an additional information $\mathscr{C}$ accompanying the formal context $\langle X, Y, I\rangle$. Future research will be directed to the investigation of other natural forms of $\mathscr{C}$ and the corresponding constraining rules. The next step is to consider constraining rules by which a formal concept $\langle A, B\rangle$ satisfies a constraint if it is true that whenever the attribute y belongs to B, then at least one of the attributes $y_1, \ldots, y_n$ belongs to B. Clearly, for $n = 1$, this becomes just the restrictions considered in this paper.

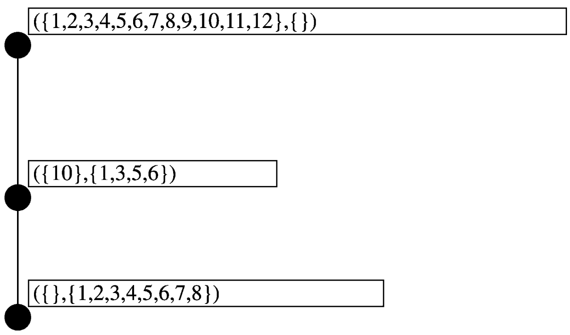

FIGURE 11 Constrained concept lattice $\mathscr{B}(X,\langle Y, \trianglelefteq_4\rangle,I)$.

Acknowledgements

The research of the first author was supported by grant No. 201/02/P076 of the GAČR.

References

Arnauld, A. and Nicole, P. (1662 : 1996) *La logique ou l'art de* penser English translation: *Logic or the Art of Thinking* (Cambridge University Press, Cambridge, MA).

Birkhoff, G. (1967) *Lattice Theory*, 3rd Ed. (AMS Coll. Publ. 25, Providence, RI).

Ganter, B. and Wille, R. (1999) *Formal Concept Analysis: Mathematical Foundations* (Springer-Verlag, Berlin).

Wille, R. (1982) "Restructuring lattice theory: an approach based on hierarchies of concepts", In: Rival, I., ed., *Ordered Sets* (Reidel, Dordrecht), pp 445–470.

Vladimír Sklenář is an assistant professor in the Department of Computer Science, Palacky University, Olomouc (Czech Republic). He graduated from Palacky University in 1981. His main interests are in software engineering, object oriented technologies, formal concept analysis. He is the author of several papers in these areas.

Jiří zacpal is an assistant professor in the Department of Computer Science, Palacky University, Olomouc (Czech Republic). He graduated from Palacký University in 1994. His main interests are in formal concept analysis, clustering, data mining. Jiri Zacpal collaborated on projects involved in processing of fuzzy data at the Faculty of Physical Culture of the Palacký University.

Radim Bělohlávek graduated in 1994 from Palacky University, Olomouc. In 1998 he received his Ph.D. in Computer Science from the Technical University of Ostrava, in 2001 he received a Ph.D. in Algebra from Palacký University, Olomouc. In 1992–93 he obtained a scholarship and spent 8 months at the University of Bern (Switzerland). In 1999 he obtained a post-doc position at the State University of New York at Binghamton, in 2000 he obtained the NATO Advanced Fellowship. Since 2001, Radim Bělohlávek is the Head of Department of Computer Science, Palacky University, Olomouc (Czech Republic). His interests are in pure and applied algebra and logic (fuzzy logic, universal algebra, lattices), processing of indeterminate information (soft computing, uncertainty calculi, data analysis), and foundations of mathematics and computing. He published 1 book and over 30 journal papers in these areas.

Journal: GGEN

Article no.: 41006

Author Query Form

COPY FOR AUTHOR

Dear Author,

During the preparation of your manuscript for typesetting some questions have arisen. These are listed below. Please check your typeset proof carefully and mark any corrections in the margin of the proof or compile them as a separate list. This form should then be returned with your marked proof/list of corrections to Alden Multimedia.

Disk use

In some instances we may be unable to process the electronic file of your article and/or artwork. In that case we have, for efficiency reasons, proceeded by using the hard copy of your manuscript. If this is the case the reasons are indicated below:

☐ Disk damaged ☐ Incompatible file format ☐ LaTeX file for non-LaTeX journal

☐ Virus infected ☐ Discrepancies between electronic file and (peer-reviewed, therefore definitive) hard copy.

☐ Other: ....

We have proceeded as follows:

☐ Manuscript scanned ☐ Manuscript keyed in ☐ Artwork scanned

☐ Files only partly used (parts processed differently:)

Bibliography

If discrepancies were noted between the literature list and the text references, the following may apply:

☐ The references listed below were noted in the text but appear to be missing from your literature list. Please complete the list or remove the references from the text.

☒ *Uncited references*: This section comprises references which occur in the reference list but not in the body of the text. Please position each reference in the text or, alternatively, delete it. Any reference not dealt with will be retained in this section.

Manuscript page/line	Details required	Author's Response
	Kindly check the modified running title. Kindly check the authors' affiliations. Author, this reference is not cited in the text. Please add or delete from reference list. Reference 2: Birkhoff, 1967. Kindly resupply the picture of the author "Vladimir Sklenar".	