Attribute exploration with background knowledge

Bernhard Ganter

1 Introduction

The purpose of this article is to describe a generalized version of a well known knowledge acquistion method, called *attribute exploration.* To get a rough idea of what these explorations are about, imagine you want to classify some collection G of items according to selected properties. For example, G could be a class of mathematical structures, e.g. groups, to be classified by structural properties like "commutative", "nilpotent", etc. Or G could consist of technical devices, car engines for example, and the attributes may reflect properties such as reliability, weight, price, and so on. But G might also be a set of persons, perhaps the students of your university, and the classifying attributes may be field of study, age, degree, etc. Attribute exploration then would help you to explore the implicational logic of these objects wrt. the attributes they exhibit.

Computer implementations of attribute exploration have existed since the early 80s. With growing experience there has been demand for generalizations of the approach. In some applications one would like to include uncertain knowledge, negated attributes (Burmeister [2]) or background information (Stumme [9]). It is our aim here to suggest a framework for doing all this. Our setting is somewhat related to that of Nowak [7], but the two approaches were developed independently and have not yet been integrated.

A specific feature of this knowledge acquisition technique is its mathematical rigour. We work with discrete data, not including probabilistic or approximative handling of knowledge. This may be seen as a limitation, but it also makes the results transparent and reliable. For the processing of mathematical knowledge this is an inalienable condition.

2 A lemma in Propositional Logic

When we discuss our approach with scholars having a background in logic, they often tend to translate our results to the language of propositional calculus. We therefore formulate one of our basic tools in these terms.

Let M be a finite set of propositional variables and let $\mathcal{L}(M)$ be the set of propositional formulae composed from M and from the connectors $\wedge$ and $\neg$ (other junctors like $\vee$ and $\rightarrow$ are introduced in the usual manner). For a set $X \subseteq M$ let $f_X : M \rightarrow \{T, F\}$ be the characteristic function defined by $f_X(m) = T :\iff m \in X$, and let $\bar{f}_X : \mathcal{L}(M) \rightarrow \{T, F\}$ denote the canonical extension of f_X to the formulae. We call X a model of the formula α, denoted $X \models \alpha$, if $\bar{f}_X(\alpha) = T$. For a set $\mathcal{F} \subseteq \mathcal{L}(M)$ of formulae and a set $\mathcal{S} \subseteq \mathfrak{P}(M)$ of subsets let

$$\operatorname{Mod} \mathcal{F} := \{X \subseteq M \mid X \models \alpha \text{ for all } \alpha \in \mathcal{F}\}$$

be the *set of models* of $\mathcal{F}$ and

$$\operatorname{Th} \mathcal{S} := \{\alpha \in \mathcal{L}(M) \mid X \models \alpha \text{ for all } X \in \mathcal{S}\}$$

be the *theory* of $\mathcal{S}$.

Any formula of the form $\neg a_1 \vee \neg a_2 \vee \ldots \vee \neg a_s \vee n_1 \vee n_2 \vee \ldots \vee n_t$, where $A := \{a_1, \ldots, a_s\}$ and $N := \{n_1, \ldots, n_t\}$ are sets of variables, is a *clause* and will be abbreviated by (A, N). A conjunction $(A, N_1) \wedge \ldots \wedge (A, N_r)$ of clauses with the same negated variables is called a *cumulated clause*. By the standard rules of propositional calculus, any cumulated clause can be written in the form

$$\bigwedge A \rightarrow \bigwedge_{i=1}^{r} \bigvee N_i,$$

and the distributive law implies that formulas

$$\bigwedge A \rightarrow \bigvee_{j=1}^{s} \bigwedge C_j,$$

where the C_j are sets of variables, also are equivalent to culumated clauses. A clause (A, N) with $|N| = 1$ is a (definite) *Horn clause*. A cumulated clause,

$$(A, \{n_1\}) \wedge \ldots \wedge (A, \{n_r\}),$$

consisting of Horn clauses is called an *implication* and is abbreviated by $A \rightarrow N$, where $N := \{n_1 \ldots, n_r\}$.

If $\mathcal{F}$ is some fixed set of formulae, then one can associate to every subset $X \subseteq M$ of the (finite set of) variables a cumulated clause α_X by

$$\alpha_X := X \rightarrow \bigvee \{\bigwedge (Y \setminus X) \mid Y \in \operatorname{Mod} \mathcal{F}, X \subseteq Y\}.$$

We note some immediate facts for later use:

Proposition 1 *$A \subseteq M$ is a model of α_X if and only if $X \not\subseteq A$ or there is some model Y of $\mathcal{F}$ satisfying $X \subseteq Y \subseteq A$. In particular, every model of $\mathcal{F}$ is a model of α_X, and X is a model of α_X if and only if X is a model of $\mathcal{F}$.*

We are going to show that every propositional theory $\mathcal{F}$ has a canonical basis consisting of cumulated clauses. We use the term *basis* for a set $\mathcal{B} \subseteq \mathcal{L}(M)$ of formulae which:

- is *non-redundant*, i.e. satisfies $\mathcal{B} \not\subseteq \mathrm{Th}(\mathrm{Mod}(\mathcal{B} \setminus \{\beta\}))$ for all $\beta \in \mathcal{B}$; and
- *generates* $\mathcal{F}$, i.e. satisfies $\mathcal{F} = \mathrm{Th}\,\mathrm{Mod}\,\mathcal{B}$.

It is clear that every theory has a non-redundant generating system, since we have assumed M to be finite. Our construction will produce such a system $\mathcal{B}$ to each $\mathcal{F}$ in a uniform manner; moreover, $\mathcal{B}$ will inherit all automorphisms of $\mathcal{F}$. This is why we call $\mathcal{B}$ a *canonical basis* of $\mathcal{F}$.

Definition 1 A set $P \subseteq M$ is *pseudo-closed* with respect to $\mathcal{F}$ if:

- P is not a model of $\mathcal{F}$, and
- for every pseudo-closed set $Q \subseteq P$, $Q \neq P$, there is a model X of $\mathcal{F}$ satisfying $Q \subseteq X \subseteq P$.

$\diamond$

Since M is finite this is a correct definition.

Lemma 1 *Let M be finite. To each theory $\mathcal{F} \subseteq \mathcal{L}(M)$ the set*

$$\mathcal{B} := \{\alpha_X \mid X \textit{ pseudo-closed }\}$$

is a non-redundant generating set for $\mathcal{F}$.

Proof Every pseudo-closed set X is a model of $\mathcal{B} \setminus \{\alpha_X\}$, but not of $\mathcal{B}$. So $\mathcal{B}$ is non-redundant. Now let P be some model of $\mathcal{B}$. Then P is a model of each α_Q, Q pseudo-closed. So for each $Q \subseteq P$ there is, according to proposition 1, a model Y of $\mathcal{F}$ satisfying $Q \subseteq Y \subseteq P$. Consequently P must either be a model of $\mathcal{F}$ or itself be pseudo-closed. But the latter is impossible since then P is not a model of $\mathcal{B}$. Therefore P must be a model of $\mathcal{F}$, which shows that

$$\mathrm{Mod}\,\mathcal{B} \subseteq \mathrm{Mod}\,\mathcal{F}.$$

The converse inclusion was already stated in proposition 1, so the lemma is proved. $\square$

The lemma generalizes a result of Duquenne and Guigues [5] on implications. If $\mathcal{F}$ consists of implications, then $\mathrm{Mod}\,\mathcal{F}$ is a closure system. Conversely, if $\mathrm{Mod}\,\mathcal{F}$ is a closure system, then each cumulated clause α_X with $X \notin \mathrm{Mod}\,\mathcal{F}$ is equivalent to the implication

$$X \to \mathcal{F}(X) \setminus X,$$

where $\mathcal{F}(X)$ denotes the closure of X.

So if $\mathcal{S}$ is a closure system on the finite set M and $X \mapsto \mathcal{S}(X)$ denotes the corresponding closure operator, then the above results specialize in the following manner: A subset $P \subseteq M$ is *pseudo-closed* with respect to $\mathcal{S}$, if P is not closed but contains the closure of every pseudo-closed proper subset. We obtain

Lemma 2 (Duquenne and Guigues) *For a finite closure system $\mathcal{S}$ the set of implications*

$$\mathcal{B} := \{P \to \mathcal{S}(P) \setminus P \mid P \text{ pseudo-closed } \}$$

is a non-redundant generating set for $\mathrm{Th}(\mathcal{S})$.

We introduce some basic notions and mention some elementary facts on implications that will play a rôle in the sequel. For a set $\mathcal{S} \subseteq \mathfrak{P}(M)$ of subsets (not necessarily a closure system) let

$$\mathrm{Imp}(\mathcal{S}) := \mathrm{Th}(\mathcal{S}) \cap \text{implications}$$

be the *implicational theory* of $\mathcal{S}$, consisting of all implications in the theory $\mathrm{Th}(\mathcal{S})$ of $\mathcal{S}$. $\mathrm{Imp}\,\mathrm{Mod}$ is a closure operator on sets of implications. $\mathrm{Imp}\,\mathrm{Mod}\,\mathcal{F}$ is called the implicational theory *generated by* $\mathcal{F}$. Lemma 2 shows that every implicational theory has a canonical generating set.

Implicational theories can be characterized by inference rules, e.g. by the *Armstrong rules* (see Maier [6])

$$(F_1) \quad \frac{}{X \to X}, \qquad (F_2) \quad \frac{X \to Y}{X \cup Z \to Y}, \qquad (F_6) \quad \frac{X \to Y, \quad Y \cup Z \to W}{X \cup Z \to W}.$$

A set $\mathcal{F}$ of implications is an implicational theory, i.e. $\mathcal{F} = \mathrm{Imp}(\mathrm{Mod}(\mathcal{F}))$, if and only if it is closed under (F_1), (F_2), and (F_6).

We have already mentioned that $\mathrm{Mod}\,\mathcal{F}$ is a closure system if $\mathcal{F}$ is a set of implications. The corresponding closure operator is denoted by

$$X \mapsto \mathcal{F}(X).$$

The closure operator and the implicational theory generated by $\mathcal{F}$ are closely related by

$$(A \to B) \in \mathrm{Imp}\,\mathrm{Mod}\,\mathcal{F} \iff B \subseteq \mathcal{F}(A).$$

It is easy to compute the closure of a set X: Check if there is an implication $A \to B \in \mathcal{F}$ with $A \subseteq X$, $B \not\subseteq X$. If so, replace X by $X \cup B$ and repeat the process. If not, X is closed. (See Wild [10] for a more detailed description of the algorithm.) This can be combined with a simple algorithm (described in [3]) to actually compute all $\mathcal{F}$-closed sets, given $\mathcal{F}$. The algorithm produces the closed sets in some lexicographic order; it can also be used to compute the pseudo-closed sets.

3 Attribute exploration

A *formal context* (G, M, I) consists of two sets G (of *objects*), M (of *attributes*), and a binary relation $I \subseteq G \times M$ (expressing for each object which attributes it has)[1]. Using for $A \subseteq G$ and $B \subseteq M$ the operators

$$\begin{aligned} A^I &:= \{m \in M \mid (g,m) \in I \text{ for all } g \in A\}, \\ B^I &:= \{g \in G \mid (g,m) \in I \text{ for all } m \in B\}, \end{aligned}$$

we introduce a *formal concept* of (G, M, I) as a pair (A, B) of sets satisfying

$$A \subseteq G, \quad B \subseteq M, \quad A^I = B, \quad A = B^I.$$

Concepts are ordered by

$$(A_1, B_1) \leq (A_2, B_2) \quad :\iff A_1 \subseteq A_2 \quad (\iff B_1 \supseteq B_2).$$

The set $\mathfrak{B}(G, M, I)$ of all concepts of (G, M, I) with this order is a complete lattice, called the *concept lattice* of (G, M, I).

An implication $A \to B$ on M is given by two subsets A and B of M; $X \subseteq M$ is a model of $A \to B$ in the sense of section 2 if and only if $A \subseteq X$ implies $B \subseteq X$. We say that $A \to B$ *holds* for an object $g \in G$ if g^I is a model of $A \to B$. $\mathrm{Imp}(G, M, I)$ denotes the implicational theory of (G, M, I), i.e. the set of those implications that hold for all $g \in G$.

Even for small contexts $\mathrm{Imp}(G, M, I)$ is huge, and it is desirable to have a small generating set at hand. A solution is offered by lemma 2, at least if M is a finite set of modest size. The difficulty remains of how to compute this basis in practice, when G is large, perhaps even infinite.

If $\mathcal{F} \subseteq \mathrm{Imp}(G, M, I)$ then the implicational theory generated by $\mathcal{F}$ is naturally contained in $\mathrm{Imp}(G, M, I)$. On the other hand, if $H \subseteq G$, then each implication that holds in (G, M, I) must also hold in the subcontext $(H, M, I \cap (H \times M))$. Combining the two cases, we get

$$\mathrm{Imp}(\mathrm{Mod}(\mathcal{F})) \subseteq \mathrm{Imp}(G, M, I) \subseteq \mathrm{Imp}(H, M, I \cap (H \times M)).$$

Thus if $\mathcal{F} \subseteq \mathrm{Imp}(G, M, I)$ and $H \subseteq G$ satisfy

$$\mathrm{Imp}(\mathrm{Mod}(\mathcal{F})) = \mathrm{Imp}(H, M, I \cap (H \times M)),$$

then this theory must also be equal to $\mathrm{Imp}(G, M, I)$.

In other words, if we find a good set of *examples*, i.e. a subset $H \subseteq G$ such that each implication from a generating set $\mathcal{B}$ of $\mathrm{Imp}(H, M, I \cap (H \times M))$ also holds in (G, M, I), then we know that $\mathcal{B}$ must generate $\mathrm{Imp}(G, M, I)$.

Attribute exploration organizes this process. During the interactive procedure two lists are kept: a set $\mathcal{F}$ of implications known to hold in (G, M, I) and a subcontext (H, M, J) of examples, where $H \subseteq G$ and $J := I \cap (H \times M)$. The iterated step is,

[1]See Ganter&Wille [3] for an introduction to Formal Concept Analysis.

right angled	$\rightarrow$	not equilateral
obtuse angled	$\rightarrow$	not equilateral, oblique angled
acute angled	$\rightarrow$	oblique angled
oblique angled, right angled	$\rightarrow$	all attributes
acute angled, obtuse angled	$\rightarrow$	all attributes
equilateral	$\rightarrow$	isoceles, oblique angled, acute angled
equilateral, not equilateral	$\rightarrow$	all attributes

Figure 1: An implicational base for the triangles.

Find the "smallest" set $X \subseteq M$ such that

$$X = \mathcal{F}(X) \neq X^{JJ}.$$

If no such X exists then the process terminates, and $\mathcal{F}$ generates $\mathrm{Imp}(G, M, I) = \mathrm{Imp}(H, M, J)$.

If any X with $X = \mathcal{F}(X) \neq X^{JJ}$ is found, then the question

$$X \rightarrow X^{JJ} \quad ?$$

is asked. The user must react either by confirming that the suggested implication holds in (G, M, I), in which case the implication is added to $\mathcal{F}$, or by giving an example $g \in G$ such that $X \subseteq g^I$, $X^{JJ} \not\subseteq g^I$, which then is included in H.

As an illustation we quote (from Burmeister[2]) a small example of an exploration concerning triangles in elementary geometry. For the attribute set $M := \{a, b, c, d, e, f, g\}$ with

a:	equilateral	d:	oblique angled	g:	right angled
b:	not equilateral	e:	acute angled		
c:	isoceles	f:	obtuse angled		

the exploration process terminates after seven implications and seven "examples". The result is shown in figure 1, except that we have simplified the premises of the implications where possible. (The pseudoclosed sets can be obtained by closing each premise wrt. all other implications of the list.) The examples can be read from the concept lattice in figure 2.

4 Implications plus background knowledge

We could use lemma 1 to generalize attribute exploration to arbitrary propositional theories. However, there are good reasons to use implicational logic. It does

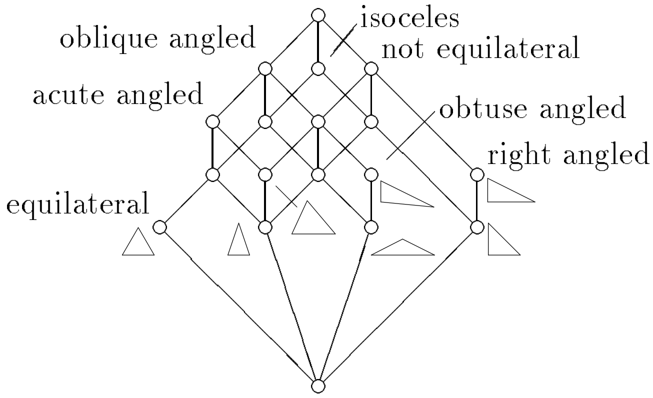


Figure 2: The concept lattice of the triangles.

not only help avoid combinatorial explosion; implicational statements are also particularly easy to comprehend by the human "operator".

But sometimes the limitation to implications is too strict, and a relaxation is desired. Suppose, for example, that you try to classify groups with respect to given properties. Implications then may be used to encode that "abelian" and "nonabelian" are mutually exclusive attributes, but they cannot express that each group is either abelian or nonabelian.

We suggest a cautious generalization of the exploration procedure in which we allow the user to provide arbitrary (propositional) "background knowledge". The exploration itself will remain implicational, i.e. only implications will be asked. Already this modest generalization causes serious complexity problems. We will just ignore these until the next section, where we discuss possible ways to get around these difficulties.

In the generalized version, all input from the user must be given in the form of clauses (A, N). This is certainly sufficient, since every propositional formula is equivalent to a conjuction of clauses. A clause (A, N) may be used as input in two different meanings:

- as a *universal statement* for objects:

  "Each object having all the attributes from A has at least one attribute from N.";

- as an *existential statement* for objects:

  "There is an object having all the attributes in A but none of the attributes from N."

Note that the second is just the negation of the first.

At any stage of the procedure let $\mathcal{U}$ denote the set of all the universal statements given as input, including the implications that have confirmed to be valid.

We are particularly interested in the implicational theory $\mathrm{Imp}\,\mathrm{Mod}\,\mathcal{U}$ generated by $\mathcal{U}$, and in the corresponding closure operator

$$X \mapsto \mathbf{U}(X) := \bigcap\{T \in \mathrm{Mod}\,\mathcal{U} \mid X \subseteq T\}$$

on M.

Let $\mathcal{E}$ be the set of existential statements given as input. These statements generalize (and replace) the specification of "examples" in the original exploration procedure. We no longer require that an example must be completely specified, and allow partial information instead.

To give a first impression of how such clauses may be applied, we come back to the example of the triangles (a more substantial example will be given below). The universal clauses that we use in this example are of particularly simple forms, namely

$(A, \mathbf{\emptyset})$ "Contradiction": There is no object having these attributes.

$(\mathbf{\emptyset}, A)$ "Exhaustion": Each object has an attribute from A.

$(A, \{m\})$ "Implication": The attributes in A imply m.

$(\{m\}, A)$ "Differentiation": $m^I = \bigcup_{a \in A}\{m, a\}^I$.

For the attributes of triangles the following clauses are immediate:

({equilateral, not equilateral}, ∅), (∅, {equilateral, not equilateral}),
({right angled, oblique angled}, ∅), (∅, {right angled, oblique angled})
({actue angled}, {oblique angled}), ({obtuse angled}, {oblique angled}),
({oblique angled}, {acute angled, obtuse angled}),
({acute angled, obtuse angled}, ∅).

A suggestive way to code such clauses is to write a list of all possible attribute combinations of the chosen subcollections in form of contexts, one for each subcollection:

	a	b
1	×	
2		×

	g	d	e	f
3	×			
4		×	×	
5		×		×

The context representation can be used both for universal and existential clauses; the existential ones ("attribute combinations that we know to occur") are conveniently specified as a subset of the universal ones ("attribute combinations we cannot exclude"). The seven clauses given above are not sufficient to generate the implicational base of figure 1. One implication has to be added:

$$\text{equilateral} \to \text{ isoceles, acute angled .}$$

Also, the rows of the two contexts are not sufficient as existential clauses. Examples have to be provided to demonstrate that "isoceles" is independent from each of "right angled", "acute angled", and "obtuse angled".

Clauses may contradict one another, and an implementation of the procedure as a computer program should check the consistency of the given data. This can be done using standard methods of propositional logic and will not be described here. We make the general assumption that $\mathcal{U}$ and $\mathcal{E}$ are *compatible*, i.e. that $\mathcal{U} \cup \mathcal{E}$ has a model. Note that $\mathcal{U}$ and $\mathcal{E}$ are compatible if and only if for every clause $(A, N) \in \mathcal{E}$ there is a a set $X \subseteq M$ that *realizes* (A, N), i.e. satisfying

$$X \in \operatorname{Mod}\mathcal{U}, A \subseteq X, \text{ and } X \cap N = \emptyset.$$

We will refer to such a family of sets as a *realizer of* $\mathcal{E}$.

The information obtained from $\mathcal{U}$ about the possible realizers may be used to extend the statements in $\mathcal{E}$. In fact, for a given existential clause (A, N) let

$$A^* := \bigcap\{T \in \operatorname{Mod}\mathcal{U}, A \subseteq T, T \cap N = \emptyset\}$$

and, with $\overline{T} := M \setminus T$,

$$N_* := \bigcap\{\overline{T} \in \operatorname{Mod}\mathcal{U}, N \subseteq \overline{T}, \overline{T} \cap A = \emptyset\}.$$

Then each X realizing (A, N) also realizes (A^*, N_*). We can associate two set operators with $\mathcal{E}$:

$$X \mapsto \mathbf{E}^\circ(X) := \bigcap\{\overline{N} \mid (A, N) \in \mathcal{E}, X \subseteq A\}$$

where $\overline{N} := M \setminus N$, and

$$X \mapsto \mathbf{E}(X) := \bigcap\{\overline{N_*} \mid (A, N) \in \mathcal{E}, X \subseteq A^*\}.$$

Clearly we have

$$\mathbf{U}(X) \subseteq \mathbf{E}(X) \subseteq \mathbf{E}^\circ(X).$$

Proposition 2 *Suppose that for some $X \subseteq M$ and some $m \in M$ the implication $X \to m$ is not a consequence of $\mathcal{U}$. Then $\mathcal{U} \cup \{X \to m\}$ is compatible with $\mathcal{E}$ if and only if $m \in \mathbf{E}(X) \setminus \mathbf{U}(X)$.*

Proof $\mathcal{U} \cup \{X \to m\}$ is compatible with $\mathcal{E}$ if and only if it is possible to choose a realizer $\mathcal{R}$ of $\mathcal{E}$ such that each $R \in \mathcal{R}$ respects $X \to m$. If $m \in \mathbf{E}(X) \setminus \mathbf{U}(X)$, this can be done, since for each $(A, N) \in \mathcal{E}$ with $X \not\subseteq A^*$ there is some set T realizing (A, N) with $X \not\subseteq T$, and for each $(A, N) \in \mathcal{E}$ with $X \subseteq A^*$ there is some T realizing (A, N) with $m \in T$, according to the definition of N_*.

Conversely, if $\mathcal{U} \cup \{X \to m\}$ is compatible with $\mathcal{E}$, then $m \in \mathbf{E}(X)$, because there must be a realizer $\mathcal{R}$ of $\mathcal{E}$ satisfying $\mathcal{R} \subseteq \operatorname{Mod}(\mathcal{U} \cup \{X \to m\})$. So for every $(A, N) \in \mathcal{E}$ there will be a $T \in \mathcal{R}$ containing A and disjoint from N, and if $X \subseteq T$ then $m \in T$. So $m \notin N_*$ by definition and thus $m \in \mathbf{E}(X)$. $\square$

The proposition states that $\mathcal{U}$ can be extended by some implication if and only if there is some $X \subseteq M$ such that $\mathbf{E}(X) \neq \mathbf{U}(X)$. Since $\mathbf{U}$ is a closure operator we may assume that $X = \mathbf{U}(X)$. This suggests the strategy for the generalized exploration procedure:

> Find the "smallest" set $X \subseteq M$ such that
>
> $$X = \mathbf{U}(X) \neq \mathbf{E}(X).$$
>
> If no such X exists then the process terminates, and $\operatorname{Imp}\operatorname{Mod}\mathcal{U}$ is the largest implicational theory containing all implicational consequences of $\mathcal{U}$, which is compatible with $\mathcal{E}$.
>
> If any X with $X = \mathbf{U}(X) \neq \mathbf{E}(X)$ is found, then the question
>
> $$X \to \mathbf{E}(X) \quad ?$$
>
> is asked. The user must react either by giving an admissible universal clause (A, N) such that $A \subseteq X$, $N \cap X = \emptyset$, which then is included in $\mathcal{U}$, or by giving an admissible existential clause (A, N) with $X \subseteq A$, $\mathbf{E}(X) \cap N \neq \emptyset$, which then is added to $\mathcal{E}$.

This is repeated until the process terminates. What we have described is the interactive version. There is, of course, a non-interactive variant of this: Rather than asking $X \to \mathbf{E}(X)$ as a question, this implication is then automatically added to $\mathcal{U}$.

How can $\mathbf{U}(X)$ be computed, given $\mathcal{U}$ and X? The following proposition gives a first answer which may however often be unsatisfactory, because computation is too complex:

Proposition 3 *If there is some $(A, N) \in \mathcal{U}$ satisfying $A \subseteq X$, $N \cap X = \emptyset$, then*

$$\mathbf{U}(X) = \bigcap_{n \in N} \mathbf{U}(X \cup \{n\}).$$

If no such $(A, N) \in \mathcal{U}$ exists then $X = \mathbf{U}(X)$.

Proof X is a model for each clause (A, N) with $A \not\subseteq X$ or $N \cap X \neq \emptyset$. If this condition is satisfied for all $(A, N) \in \mathcal{U}$ then $X \in \operatorname{Mod}\mathcal{U}$. Otherwise,

$$\begin{aligned}\mathbf{U}(X) &= \bigcap\{T \in \operatorname{Mod}\mathcal{U} \mid X \subseteq T\} \\ &= \bigcap_{n \in N} \bigcap\{T \in \operatorname{Mod}\mathcal{U} \mid X \cup \{n\} \subseteq T\}.\end{aligned}$$

$\square$

5 Approximation

The problem of the approach described above is its computational complexity. Apart from checking the consistency, the main problem is to compute the operators $\mathbf{U}$ and $\mathbf{E}$ when generating the questions. If this turns out to be a serious obstacle, one might replace the operator $\mathbf{E}$ by $\mathbf{E}^\circ$, which is easy to compute, and ask the question

$$X \to \mathbf{E}^\circ(X) \quad ?$$

instead. This is of course a partially "stupid" question because part of it could in principle automatically be answered by means of the data already present. But the original question $X \to \mathbf{E}(X)$ is contained as a subquestion. Thus replacing $\mathbf{E}$ by $\mathbf{E}^\circ$ may put more labour on the user, but the results remain correct and complete.

We can do something similar for $\mathbf{U}$ and introduce a closure operator $\mathbf{U}^\circ$ which is easier to compute and which satisfies $\mathbf{U}^\circ(X) \subseteq \mathbf{U}(X)$ for all $X \subseteq M$. The question

$$\mathbf{U}(X) \to \mathbf{E}(X) \quad ?$$

then will, if necessary, be replaced by

$$\mathbf{U}^\circ(X) \to \mathbf{E}^\circ(X) \quad ?$$

To introduce $\mathbf{U}^\circ$ we split the set $\mathcal{U}$ of universal statements into two parts $\mathcal{V}$ and $\mathcal{W}$ with

$$\mathrm{Th}\,\mathrm{Mod}\,\mathcal{U} = \mathrm{Th}\,\mathrm{Mod}(\mathcal{V} \cup \mathcal{W}),$$

such that $\mathcal{W}$ consists of implications only (and $\mathcal{V}$ is small). This guarantees

$$\mathrm{Imp}\,\mathrm{Mod}\,\mathcal{W} \subseteq \mathrm{Imp}\,\mathrm{Mod}\,\mathcal{U}$$

and consequently, if we define $\mathbf{U}^\circ$ to be the closure operator corresponding to $\mathcal{W}$,

$$\mathbf{U}^\circ(X) := \bigcap\{T \in \mathrm{Mod}\,\mathcal{W} \mid X \subseteq T\},$$

we obtain $\mathbf{U}^\circ(X) \subseteq \mathbf{U}(X)$ for all $X \subseteq M$.

Ideally, $\mathcal{W}$ generates the same implications as $\mathcal{U}$, and the equality $\mathbf{U}^\circ(X) = \mathbf{U}(X)$ holds. For example, $\mathcal{W}$ could be an implicational base for $\mathrm{Imp}\,\mathrm{Mod}\,\mathcal{U}$. The computation of $\mathbf{U}^\circ(X)$ is relatively easy, as mentioned in section 3 above.

It is worthwhile studying which implicational theories are closed with respect to consequences derived from a given set of clauses. This will be formulated in the theorem below. For a better understanding let us mention that in the examples we have studied so far, the "non-implicative" part $\mathcal{V}$ was relatively small, so that the consequences of $\mathcal{V}$ could be computed. Of particular interest are the *prime implicants* of $\mathcal{V}$, by which we mean the minimal clauses in $\mathrm{Th}\,\mathrm{Mod}\,\mathcal{V}$.

Clauses may generate implications from other implications. Well known instances of this are the rules of *contraposition* and of *exhaustion* in propositional logic. These are special cases of the following general (A,N)-*exhaustion rule*

$$F_{(A,N)} \qquad \frac{\forall_{n\in N}\ X\cup\{n\}\to m}{X\to m} \qquad \text{if } A\subseteq X.$$

Proposition 4 *If $\mathcal{F}$ is a family of formulae and (A,N) is a clause, then*

$$\operatorname{Imp}\operatorname{Mod}(\mathcal{F}\cup\{(A,N)\})$$

is closed under $F_{(A,N)}$.

Proof Suppose $X\cup\{n\}\to m$ is in $\operatorname{Imp}\operatorname{Mod}(\mathcal{F}\cup\{(A,N)\})$ for all $n\in N$. We claim that then $X\to m$ also is, i.e.

$$m\in\bigcap\{T\in\operatorname{Mod}(\mathcal{F}\cup\{(A,N)\})\mid X\subseteq T\}.$$

A subset T of M containing X (and thus A) is a model of (A,N) only if $n\in T$ for some $n\in N$, and therefore is a model of $\mathcal{F}\cup\{(A,N)\}$ only if $m\in T$. □

Theorem 1 *Let $\mathcal{V}$ be a set of clauses on the finite set M. A set $\mathcal{F}$ of implications on M satisfies*

$$\mathcal{F}=\operatorname{Imp}\operatorname{Mod}(\mathcal{F}\cup\mathcal{V})$$

if and only if $\mathcal{F}$ is closed under F_1, F_2, F_6 and under the (A,N)-exhaustion rules for all prime implicants (A,N) of $\mathcal{V}$.

Proof Is is immediate from the proposition that $\operatorname{Imp}\operatorname{Mod}(\mathcal{F}\cup\mathcal{V})$ is closed under all these rules. Suppose that, conversely, $\mathcal{F}$ is an implicational theory closed under exhaustion wrt. $\mathcal{V}$. We must prove that every model of $\mathcal{F}$ is an intersection of models of $\mathcal{F}\cup\mathcal{V}$. So let X be a model of $\mathcal{F}$ which is not a model of $\mathcal{V}$. Then from the set of $\mathcal{V}$-models containing X, i.e. from

$$\{Y\in\operatorname{Mod}\mathcal{V}\mid X\subseteq Y\},$$

a transversal T with $T\cap X=\emptyset$ can be selected, and the clause (X,T) is a consequence of $\mathcal{V}$. There must be some prime implicant (A,N) of $\mathcal{V}$ with $A\subseteq X$, $N\subseteq T$, and by the hypothesis $\mathcal{F}$ is closed under the (A,N)-exhaustion rule. This implies

$$X=\bigcap_{n\in N}\mathcal{F}(X\cup\{n\}),$$

because $m\in\mathcal{F}(X\cup\{n\})$ is equivalent to $(X\cup\{n\}\to m)\in\mathcal{F}$ for all $n\in N$, which by exhaustion implies that $(X\to m)\in\mathcal{F}$. Since X is closed, $m\in X$ follows. Therefore X is $\bigcap$-reducible. Each $\bigcap$-irreducible model of $\mathcal{F}$ therefore is a model of $\mathcal{V}$. □

6 Many valued attributes

A natural instance of the theory sketched above is that of many valued attributes. Let us recall that a *many valued context* is defined to be a quadruple

$$(G, M, W, I),$$

where G, M, and W are sets (of *objects*, *many valued attributes*, and *attribute values*, resp.), and $I \subseteq G \times M \times W$ is a ternary relation satisfying

$$(g, m, v) \in I, (g, m, w) \in I \quad \Rightarrow \quad v = w.$$

If $(g, m, w) \in I$ then *w is the value of the attribute m for the object g*, and the condition ensures there is always at most one value for each object-attribute pair. We will assume that the many valued contexts under consideration are *complete*, i.e. that for each $g \in G$ and each $m \in M$ there is precisely one value $m(g) \in W$ with $(g, m, m(g)) \in I$. $m(G) \subseteq W$ is the *value set* of the attribute m.

To distinguish many valued contexts from the formal contexts defined in §3, the latter will sometimes be referred to as *univalued* contexts.

Concepts, implications etc. for many valued contexts are defined in terms of thoses of univalued contexts. There is a standard method, called *plain conceptual scaling*, to transform a many valued context (G, M, W, I) into a univalued one: For each $m \in M$, one chooses a (univalued) context $\mathbb{S}_m := (G_m, M_m, I_m)$ with $m(G) \subseteq G_m$. $\mathbb{S}_m$ is called a *scale* for the attribute m; G_m is the set of *scale values* and M_m is the set of *scale attributes* of the respective scale. The context (G, N, J) defined by

$$N := \bigcup_{m \in M} \{m\} \times M_m$$

and

$$(g, (m, n)) \in J \quad :\Longleftrightarrow \quad (g, m, w) \in I \text{ and } (w, n) \in I_m$$

is the *derived context*.

As an example, we show the classification of the two-dimensional crystallographic point groups (i.e. stabilizers of a point p in the symmetry group of a planar ornament, see Hahn [4]) according to the rotations and mirror reflections they contain. The ten possible attribute value combinations are shown in figure 3, together with scales that describe the structure of the attribute values.

The derived context with respect to the scales given in figure 3 is shown in figure 4.

For a given set M of attributes and given scales $\mathbb{S}_m := (G_m, M_m, I_m), m \in M$, there is a "most general[2]" many valued context (G, M, W, I) with these value

[2] We silently identify indistinguishable objects, i.e. objects g and h with $m(g) = m(h)$ for all $m \in M$.

group	$\mathfrak{rot}_p$	$\mathfrak{m}_p$
1	–	–
2	2	–
3	3	–
4	4	–
6	6	–
m	–	1
$2m$	2	2
$3m$	3	3
$4m$	4	4
$6m$	6	6

$\mathfrak{rot}_p$:= order of rotation, $\mathfrak{m}_p$:= number of directions of mirror axes.

$\mathfrak{rot}_p$	–	2|	3|	4|	6|
–	×				
2		×			
3			×		
4		×		×	
6		×	×		×

$\mathfrak{m}_p$	$=0$	≥ 1	≥ 2	≥ 3	≥ 4	≥ 6
–	×					
1		×				
2		×	×			
3		×	×	×		
4		×	×	×	×	
6		×	×	×	×	×

Figure 3: Rotations and reflections in the two-dimensional crystallographic point groups, and scales for the two attributes.

	$\mathfrak{rot}_p$					$\mathfrak{m}_p$					
	–	2|	3|	4|	6|	$=0$	≥ 1	≥ 2	≥ 3	≥ 4	≥ 6
1	×					×					
2		×				×					
3			×			×					
4		×		×		×					
6		×	×		×	×					
m	×						×				
$2m$		×					×	×			
$3m$			×				×	×	×		
$4m$		×	×				×	×	×	×	
$6m$		×	×		×		×	×	×	×	×

Figure 4: The derived context obtained by plain scaling.

sets, given by

$$G := \bigtimes_{m \in M} G_m, \quad W := \bigcup_{m \in M} G_m, \quad (g, m, w) \in I :\iff g(m) = w.$$

The corresponding derived context

$$(\bigtimes_{m \in M} G_m, \bigcup_{m \in M} \{m\} \times M_m, J)$$

is called the *semiproduct* of the scales $\mathbb{S}_m$. Any subset $H \subseteq G$ induces a subcontext of the many valued context and also its derived context in the semiproduct. Therefore, exploration of many valued attributes with respect to a given plain scaling can be understood as "exploration within a semiproduct". We generalize and consider "exploration within a given context", that is, attribute exploration with the additional restriction that all models can be choosen only from a given context, called the *frame context*. This is covered by the general approach developed in §4, since the frame context can be defined using (universal) clauses. But if the frame context of modest size and accessible for computation, the process becomes quite obvious and it is unnecessary to store lists of (universal or existential) clauses. We then use three (univalued) contexts for the bookkeeping:

$\mathbb{P} := (G, M, I)$, the context of *potential* examples,

$\mathbb{E}_+ := (E, M, I_+)$, the first context of examples, encoding which attributes the given examples are known to have. For $e \in E$ we use the abbreviation,

$$e^+ := \{m \in M \mid (e, m) \in I_+\}.$$

$\mathbb{E}_? := (E, M, I_?)$, the second context of examples, encoding which attributes the given examples are not known not to have. For $e \in E$ we use the abbreviation,

$$e^? := \{m \in M \mid (e, m) \in I_?\}.$$

Initially, $\mathbb{P}$ will contain the frame context, e.g. the semiproduct corresponding to a many valued context. For every existential clause $(A, N) \in \mathcal{E}$ we include a new object $e \in E$ with $e^+ := A$ and $e^? := M \setminus N$.

Every universal clause $(A, N) \in \mathcal{U}$ causes modifications of $\mathbb{P}$, $\mathbb{E}_+$, and $\mathbb{E}_?$. From G all objects are removed that are not models of (A, N). In $\mathbb{E}_+$ and $\mathbb{E}_?$ the incidence relations are modified to encode the consequences of the clause. More precisely, we perform the following steps.

1. In $\mathbb{P}$, the set of objects G is replaced by the subset

$$\{g \in G \mid A \not\subseteq g^I \text{ or } g^I \cap N \neq \emptyset\},$$

and the incidence relation is restricted to this subset.

2. For each $e \in E$ let

$$e^* := \bigcap\{g^I \mid g \in G, e^+ \subseteq g^I \subseteq e^?\}.$$

I_+ is replaced by $I_+ \cup \{(e,m) \mid m \in e^*\}$.

3. For each $e \in E$ let

$$e_* := \bigcup\{g^I \mid g \in G, e^+ \subseteq g^I \subseteq e^?\}.$$

$I_?$ is replaced by $I_? \setminus \{(e,m) \mid m \notin e_*\}$.

It is not difficult to see that the operators **U** and **E** defined in §4 can be defined in terms of the contexts $\mathbb{P}$ and $\mathbb{E}_?$. We have

$$\mathbf{U}(X) = X^{II} \quad \text{and} \quad \mathbf{E}(X) = X^{??}.$$

So the iterated step in the exploration procedure is

> Find the "smallest" $X \subseteq M$ such that
>
> $$X = X^{II} \neq X^{??}.$$
>
> If no such X exists the exploration process terminates and $\operatorname{Imp}\mathbb{P}$ is the largest implicational theory of a subcontext of the frame context containig all implicational consequences of the universal clauses given, which is compatible with the existential clauses.
>
> If any X with $X = X^{II} \neq X^{??}$ is found, then the question
>
> $$X \to X^{??} \qquad ?$$
>
> is asked. The user must react by either giving an admissible universal clause (A,N) such that $A \subseteq X$, $N \cap X = \emptyset$, which then leads to a modification of $\mathbb{P}$, $\mathbb{E}_+$, and $\mathbb{E}_?$, as described above, or by giving an admissible existential clause (A,N) with $X \subseteq A$, $X^{??} \cap N \neq \emptyset$, which leads to an extension of the contexts $\mathbb{E}_+$ and $\mathbb{E}_?$ by a new object e with
>
> $$e^+ := \bigcap\{g^I \mid g \in G, A \subseteq g^I \subseteq M \setminus N\}$$
>
> and
>
> $$e^? := \bigcup\{g^I \mid g \in G, A \subseteq g^I \subseteq M \setminus N\}.$$

	$\mathfrak{rot}$	$\mathfrak{m}$	$\mathfrak{g}$	$\frac{\mathfrak{rot}}{\mathfrak{m}}$	$\mathfrak{rot}\Vert\mathfrak{m}$
p1	–	–	–	–	–
p2	2	–	–	–	–
pm	–	1	–	–	–
pg	–	–	1	–	–
cm	–	1	1	–	–
p2mm	2	2	–	2	–
p2mg	2	1	1	–	2
p2gg	2	–	2	–	–
c2mm	2	2	2	2	2
p3	3	–	–	–	–
p3m1	3	3	3	3	–
p31m	3	3	3	3	3
p4	4	–	–	–	–
p4mm	4	4	2	4	–
p4gm	4	2	4	2	4
p6	6	–	–	–	–
p6mm	6	6	6	6	–

$\mathfrak{rot}$: maximal order of a rotation,

$\mathfrak{m}$: number of directions of reflection axes,

$\mathfrak{g}$: number of directions of glide-reflection axes which are not reflection axes,

$\frac{\mathfrak{rot}}{\mathfrak{m}}$: maximal order of a rotation, the center of which is incident with a reflection axis,

$\mathfrak{rot}\Vert\mathfrak{m}$: maximal order of a rotation, the center of which is non-incident with a reflection axis, if there is a reflection.

Figure 5: Two-dimensional crystallographic groups.

7 An example

To demonstrate these methods we consider once more the classification of two-dimensional crystallographic groups. It is well known that there are exactly seventeen such groups and that they can be classified according to attributes describing rotations, reflections, glide-reflections and their combinations. A list is given in figure 5 (Hahn [4], see Wille [11] for a concept lattice). This table can be viewed as a many valued context. For a plain scaling, the scales from figure 3 may be used; more precisely, the $\mathfrak{rot}_p$-scale from figure 3 fits for the attributes $\mathfrak{rot}$, $\frac{\mathfrak{rot}}{\mathfrak{m}}$, and $\mathfrak{rot}\Vert\mathfrak{m}$, while the $\mathfrak{m}_p$-scale is suited for $\mathfrak{m}$ and $\mathfrak{g}$.

A proof that this list is complete is not very difficult. But it is tedious and therefore is skipped in most textbooks. We will outline how a proof could be "organized automatically" using the above techniques. But note that we do not intend to give an automatic proof! The mathematical argumentation will not be touched, the purpose of the algorithm is to support the case analysis.

A classification proof has to show that a) figure 5 contains all possibles cases and b) each entry corresponds to a unique solution. For brevity we treat only part a) here.

The example is particularly simple, insofar as the complete list of examples is given from the very beginning. This implies that the contexts $\mathbb{E}_+$ and $\mathbb{E}_?$, defined in the previous section, are both equal to the derived context of figure 5 and remain unchanged. Existential clauses will play no rôle. The task is to find a base for the propositional theory of the given examples. The propositions in such a base then require a proof; if all these are proved, the classification is complete.

We might for example use lemma 2 and apply the algorithm described in section 3 to $\mathbb{E}_+$ (confirming all suggested implications). This yields an implicational base for $\mathbb{E}_+$, and we could prove the classification by showing that each of these implications holds for all crystallographic groups. It turns however out that an implicational base of $\mathbb{E}_+$ consists of 57 implications. It is not a pleasant idea to organize a proof by dividing it into 57 cases.

If we use the "many-valued-contex"-approach as developed in section 6, the situation improves slightly. The "frame context" $\mathbb{P}$ is the semiproduct of the scales, and thus, having $6^2 \cdot 5^3 = 4500$ objects, is rather small. An implicational base of $\mathbb{E}_+$ within this frame consists of 31 implications, which is better, but still too large.

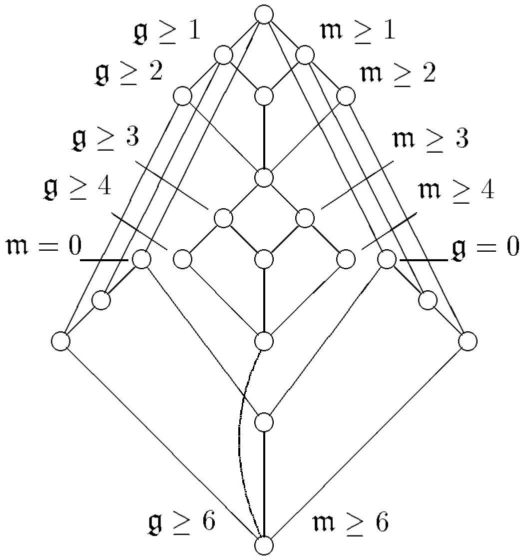
The way to proceed is to further reduce the size of the frame context by adding universal statements. Loosely speaking, we simplify the problem by "telling the algorithm what we know" and will, in return, be told what remains to be proved. The implementation[3] I use returns in addition a list of the cases which have not yet been excluded.

One usually states the most obvious propositions first, e.g., clearly every rotation has a center that either is incident with a mirror axes or not. Including this information reduces the frame context in our example from 4500 to 3270 objects and the size of the relative implicational base from 31 to 25.

A more systematic way to find propositions that are easy to prove is to restrict to small subsets of the attribute set. We give one example. Select from figure 5 only the columns $\mathfrak{m}$ and $\mathfrak{g}$, consider the derived subcontext (with the same scaling as before), its concept lattice and an implicational base, see figure 6. All these implications are, more or less obviously, true for crystallographic groups. They reduce the number of open implications to 18.

This may be repeated for other choices of attributes. The hope is that finally most of the trivial aspects are eliminated and the implicational base is reduced to the "essential" problems, i.e. to those that require a nontrivial argument. In our example, after comparing the attribute $\mathfrak{rot}$ with each other attribute, the number of open cases reduces to 36, and nine implications remain to be proved. These are given in figure 7.

[3]Sincere thanks to Rüdiger Krauße for his beautiful programs.



$$
\begin{array}{lcr}
\mathfrak{m} \geq 3 & \to & \mathfrak{g} \geq 2 \\
\mathfrak{g} \geq 3 & \to & \mathfrak{m} \geq 2 \\
\mathfrak{m} \geq 6 & \to & \mathfrak{g} \geq 6 \\
\mathfrak{g} \geq 6 & \to & \mathfrak{m} \geq 6 \\
\mathfrak{m} \geq 2, \mathfrak{g} \geq 1 & \to & \mathfrak{g} \geq 2 \\
\mathfrak{m} \geq 1, \mathfrak{g} \geq 2 & \to & \mathfrak{m} \geq 2 \\
\mathfrak{m} \geq 4, \mathfrak{g} \geq 3 & \to & \mathfrak{g} \geq 6 \\
\mathfrak{m} \geq 3, \mathfrak{g} \geq 4 & \to & \mathfrak{m} \geq 6
\end{array}
$$

Figure 6: The interplay of mirror- and glide-reflections. On the left the concept lattice (group names omitted), on the right an implicational base relative to the semiproduct.

$$
\begin{array}{rcl}
4|\mathfrak{rot}, \mathfrak{m} \geq 2, \mathfrak{g} \geq 2, 4|\frac{\mathfrak{rot}}{\mathfrak{m}} & \to & \mathfrak{rot}||\mathfrak{m} = - \\
4|\mathfrak{rot}, \mathfrak{m} \geq 2, \mathfrak{g} \geq 2, 2|\frac{\mathfrak{rot}}{\mathfrak{m}}, 4|\mathfrak{rot}||\mathfrak{m} & \to & \mathfrak{g} \geq 4 \\
4|\mathfrak{rot}, \mathfrak{g} \geq 2 & \to & \mathfrak{m} \geq 2 \\
2|\mathfrak{rot}, \mathfrak{m} \geq 4, \mathfrak{g} \geq 2, 2|\frac{\mathfrak{rot}}{\mathfrak{m}} & \to & \mathfrak{rot}||\mathfrak{m} = - \\
2|\mathfrak{rot}, \mathfrak{m} \geq 1, 2|\frac{\mathfrak{rot}}{\mathfrak{m}} & \to & \mathfrak{m} \geq 2 \\
2|\mathfrak{rot}, \mathfrak{g} \geq 1, \mathfrak{rot}||\mathfrak{m} = - & \to & \mathfrak{g} \geq 2 \\
\mathfrak{m} \geq 2, \mathfrak{g} \geq 2, \mathfrak{rot}||\mathfrak{m} = - & \to & \mathfrak{m} \geq 3 \\
\mathfrak{g} = 0 & \to & \mathfrak{rot}||\mathfrak{m} = - \\
4|\frac{\mathfrak{rot}}{\mathfrak{m}} & \to & \mathfrak{m} \geq 4
\end{array}
$$

Figure 7: These implications remain to be proved.

8 Conclusion

We are confident that the tool described above has some potential to support mathematical knowledge processing. It is apparent that there are many limitations: the method is restricted to a very specific kind of knowledge, and, due to complexity reasons, is feasible for projects of modest size.

But within these boundaries it may be useful and can handle more than trivial portions of mathematical knowledge. Using plain attribute exploration, Reeg and Weiß [8] have completely explored the implicational theory of 50 frequently used attributes of finite lattices. Baader [1] has solved a representation problem in Description Logics using the exploration technique. We have applied the method to find axiom systems for certain relational structures used in linguistic classification; we later found these characterizations in the literature, stated as theorems. (We have worked with yet another generalization, combining the approach with that of Zickwolff [12], thereby including some elements of predicate logic.)

As said before, we do not expect that this method could replace any creative mathematical work. But it may ease the handling of some steps in mathematical work that is considered trivial, but tedious.

References

[1] Franz Baader. *Computing a minimal representation of the subsumption lattice of all conjunctions of concepts defined in a terminology.* In: Gerard Ellis, Robert A. Levinson, Andrew Fall and Veronica Dahl (eds.): *Proceedings of the International KRUSE Symposium: Knowledge Retrieval, Storage and Use for Efficiency.* Springer Lecture Notes in Artificial Intelligence, 1995.

[2] Peter Burmeister. Merkmalimplikationen bei unvollständigem Wissen. In: Wilfried Lex, ed. *Arbeitstagung Begriffsanalyse und Künstliche Intelligenz.* Informatik–Bericht 89/3, TU Clausthal, 1991, 15–46.

[3] Bernhard Ganter and Rudolf Wille: *Formale Begriffsanalyse – Mathematische Grundlagen.* Springer Verlag, Heidelberg 1996.

[4] Theo Hahn (ed.): *International tables for crystallography.* D. Reidel, Dordrecht/Boston 1983.

[5] J.-L. Guigues and Vincent Duquenne. Familles minimales d'implications informatives resultant d'un tableau de données binaires. *Math. Sci. Humaines* 95 (1986), 5–18.

[6] David Maier. *The theory of relational data bases.* Computer Science Press, Rockville 1983.

[7] Chris Nowak. *Conceptual knowledge and preference.* PhD thesis, The University of Adelaide, 1996.

[8] Stefan Reeg and Wolfgang Weiß. *Properties of finite lattices.* Diplomarbeit, TH Darmstadt, 1990.

[9] Gerd Stumme. Attribute exploration with background implications and exceptions. In: Hans-Herrmann Bock, Wolfgang Polasek, editors. *Data analysis and information systems.* Springer, Berlin–Heidelberg–New York, 1996, pp. 457ff.

[10] Marcel Wild. A theory of finite closure spaces based on implications. *Advances in Mathematics* 108.1 (1994), 118–134.

[11] Rudolf Wille. Bedeutungen von Begriffsverbänden. In: Bernhard Ganter, Rudolf Wille, and Karl Erich Wolff (eds.). *Beiträge zur Begriffsanalyse.* B. I.–Wissenschaftsverlag, Mannheim, 1987, 161–211.

[12] Monika Zickwolff. *Rule exploration: first order logic in formal concept analysis.* Dissertation, TH Darmstadt, 1991.