Math. Appl. **6** (2017), 99–110
DOI: 10.13164/ma.2017.06

CLASSIFICATION TREES IN A BOX EXTENT LATTICE

LAURA VERES

Abstract. In this paper we show that, during an elementary extension of a context, each of the classification trees of the newly created box extent lattice can be obtained by modifying the classification trees of the box extent lattice of the original, smaller context. We also devise an algorithm which, starting from a classification tree of the box extent lattice of the smaller context $(H, M, I \cap H \times M)$, gives a classification tree of the extended context (G, M, I) which contains the new elements inserted. The efficiency of the method is given by the fact that it is sufficient to know the original context, the classification tree of the box extent lattice and its box extents while the knowledge of a new box extension of the extended context mesh elements is not required (except for one, which is the new element box extension).

1. Preliminaries: Box lattice, extent lattice

A context ([2]) is a triple (G, M, I) where G and M are sets and $I \subseteq G \times M$ is a binary relation. The elements of G and M are called objects and attributes of the context, respectively. The relation gIm means that the object g has the attribute m. A small context can be easily represented by a cross table, i.e., by a rectangular table with rows headed by the object names and the columns by the attribute names. A cross in the intersection of row g and column m, means that object g has attribute m. For all sets $A \subseteq G$ and $B \subseteq M$, we define

$A' = \{m \in M \mid g \ I \ m \text{ for all } g \in A\}$,

$B' = \{g \in G \mid g \ I \ m \text{ for all } m \in B\}$.

A concept of the context (G, M, I) is a pair (A, B) in which $A' = B$ and $B' = A$, and $A \subseteq G$, $B \subseteq M$. We call A the extent and B the intent of the concept (A, B). $\mathcal{L}(G, M, I)$ denotes the set of all concepts of the context (G, M, I), and $\mathcal{E}(G, M, I)$ denotes the set of all concept extents of the context.

$\mathcal{L}(G, M, I)$ can be endowed with the structure of a complete lattice defining the join and meet of concepts as follows ([2]):

$$\bigwedge_{i\in I}(A_i, B_i) = \left(\bigcup_{i\in I} A_i, \left(\bigcap_{i\in I} B_i\right)''\right),$$

$$\bigvee_{i\in I}(A_i, B_i) = \left(\left(\bigcap_{i\in I} A_i\right)'', \bigcup_{i\in I} B_i\right).$$

MSC (2010): primary 06B23; secondary 06B05, 06B15.

Keywords: concept lattice, extent partition, box extent, one-object extension, classification tree.

The lattice $(\mathcal{L}(G, M, I), \wedge, \vee)$ will be called the concept lattice of the context (G, M, I). Clearly, ordering the extents by $\subseteq$, we obtain a lattice $(\mathcal{E}(G, M, I), \subseteq)$ isomorphic with the concept lattice $\mathcal{L}(G, M, I)$.

An extent partition of a formal context (G, M, I) is a partition of G, all blocks of which are concept extents. Clearly, the trivial partition $\{G\}$ is an extent partition. Note that, since the intersection of extents always yields an extent, the common refinements of extent partitions are still extent partitions. Therefore, the extent partitions of (G, M, I) form a complete $\wedge$-subsemilattice of the partition lattice of G, and Thus, a complete lattice, which will be denoted by $\mathrm{Ext}(G, M, I)$. In particular, there is always a finest extent partition of the context denoted by $\pi_\Box$.

Example 1.1. Figure 1 shows an example of a context (G, M, I) with $G = \{a, b, c, d, e, f, g, h\}$, $M = \{1, 2, 3, 4, 5, 6\}$, and the corresponding concept lattice. A concept lattice is usually very big even for such a small context. It is enough to

K	1	2	3	4	5	6
a	X					
b				X	X	X
c	X				X	X
d		X	X		X	
e		X	X		X	X
f	X	X	X			
g			X	X	X	X
h		X		X		X

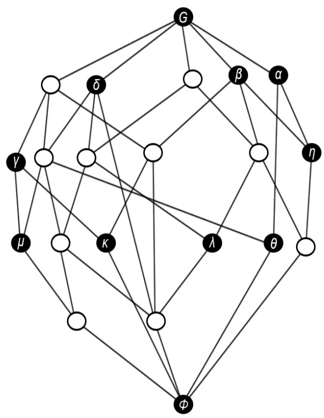

Figure 1. Example of context.

know the extent of a concept because the extent and the intent mutually determine each other. One extent, for example, is always the object set G itself or $\emptyset$. In the concept lattice, we denote some concepts with black circle and Greek letters because they are also box elements of the context. In the same picture, the reader can also see the corresponding box lattice of the context. These concepts (box elements) are

$\alpha = (\{a, c, f\}; \{1\})$, $\eta = (\{a, f\}; \{1, 2\})$,
$\beta = (\{a, d, e, f, h\}; \{2\})$, $\theta = (\{c\}; \{1, 5, 6\})$,
$\gamma = (\{b, g, h\}; \{4, 6\})$, $\lambda = (\{d, e\}, \{2, 3, 5\})$,
$\delta = (\{b, c, d, e, g\}; \{5\})$, $\kappa = (\{h\}; \{2, 4, 6\})$,
$\mu = (\{b, g\}; \{4, 5, 6\})$.

The notions of box element and box extent are defined after the example). Some extent partitions of the context are $\pi_1 = \{a, b, c, d, e, f, g\}$ or
$\pi_2 = \{\{a, c, f\}, \{b, g, h\}, \{d, e\}\}$.

In [6], the notion of a classification systems was introduced: Let L be a complete lattice. A nonempty set $S = \{a_j \mid j \in J\}$ of nonzero elements of L is called a classification system of L iff

1. $a_j \wedge a_k = 0$ for all $j \neq k$, $j, k \in J$, and
2. $x = \bigvee_{j \in J} (x \wedge a_j)$ for all $x \in L$.

The zero element of a complete lattice L and all elements that are contained in some classification system of L are called the box elements of L. The set of all box elements of L is denoted by $Box(L)$ and $(Box(L), \leq)$ is a poset obtained by restricting the partial order of L to the box elements. If every nonzero element of L is a join of atoms of L, then L is called an atomistic lattice.

In [5] the following was shown: If L is a complete lattice in which every element is a join of some completely join-irreducible elements, then $(Box(L), \leq)$ is a complete atomistic lattice.

In [3] the box extents of a context (G, M, I) where defined as the sets E belonging to some extent partition of (G, M, I) or $E = \varnothing''$. The set of all box extents of L is denoted by $\mathcal{B}(G, M, I)$. The box extents are ordered by inclusion and in [3] was proved that $\mathcal{B}(G, M, I)$ is a complete atomistic lattice which is a $\wedge$-subsemilattice of the concept lattice. (Note that if $\varnothing'' \neq \varnothing$, then $\{G\}$ is the only extent partition of the context (G, M, I)). Observe that each object $g \in G$ is contained in a smallest box extent denoted by $g^{\Box\Box}$, which is a class containing g of the finest extent partition $\pi_{\Box}$ of the context (G, M, I). In other words, E is a box extent iff $g \in E \Longrightarrow g^{\Box\Box} \in E$ or $g \notin E \Longleftrightarrow g^{\Box\Box} \cap E = \varnothing$.

Example 1.2. The box extents of the context in Example 1.1 are G,$\{a, c, f\}$; $\{a, d, e, f, h\}$; $\{b, g, h\}$; $\{b, c, d, e, g\}$; $\{a, f\}$; $\{c\}$; $\{d, e\}$; $\{h\}$; $\{b, g\}$.

In [4], the notion was introduced of CD-independent sets in an arbitrary poset, we will define it in a lattice as follows:

Let L be a bounded lattice. A set $X \subseteq L$ is called CD-independent if, for any $x, y \in X$, either $x \leq y$ or $y \leq x$ or $x \wedge y = 0$ (any elements of X are comparable or disjoint). Maximal CD-independent sets (with respect to $\subset$) are called CD-bases.

Let L be a lattice with the smallest element 0. A set $O = \{a_i | \, i \in I\}$, $I \neq \varnothing$ of nonzero elements of L is called an orthogonal system or disjoint set, if $a_i \wedge a_j = 0$, $i \neq j$. O is a complete orthogonal system, if there is no other orthogonal system O' of L containing O as a proper subset. Notice that if $S_1 = \{a_i | i \in I\}$, $I \neq \varnothing$ and $S_2 = \{b_j | j \in J\}$, $J \neq \varnothing$ are two orthogonal systems, then $S_1 \leq S_2$, if for each $i \in I$ there exists $j(i) \in J$ such that $a_i \leq b_{j(i)}$. We denoted by $Ort(L)$ the set of all orthogonal systems, $(Ort(L), \leq)$ is evidently a poset, moreover a lattice ([8]).

Theorem 1.3. ([4, 8]) *Let L be a lattice, and T a CD-base in L.*

(i) *Then, there exists a chain $\mathcal{C} = \{S_\lambda \mid \lambda \in \Lambda\}$ in $Ort(L)$, such that $T = \bigcup_{\lambda \in \Lambda} S_\lambda$.*
(ii) *Any CD-base T is the union of orthogonal systems (disjoint sets) belonging to a maximal chain.*

2. Classification trees and complete classification trees

Classification trees are used in many fields: lattice theory, data mining, group technology problems. Classification trees are used for clustering objects by their attributes, and they appear in some clustering problems originated in Group Technology. Several independent definitions are for classification trees in lattice theory, we will define them as a special chain and also we will show some construction theorems for them.

First, we will recall the notions of classification trees, relations between classification trees, orthogonal (disjoint) systems and CD-independent sets [8].

Let L be a lattice with the greatest element $1 \in L$ and $T \subseteq L \setminus \{0\}$, $T \neq \emptyset$ a subset of it. T is a directed tree, if $[t) \cap T$ is a chain for each $t \in T$ and $1 \in T$.

Now, strengthening the above condition, we obtain the notion of a classification tree:

Definition 2.1. A directed tree T is a classification tree, if $[x) \cap T$ is a chain (nonempty) for each $x \in L \setminus \{0\}$. H is a maximal classification tree, if there is no other classification tree which contains it as a proper subset.

Observe that, in the definition of a classification tree, the condition $[t) \cap T$ is a chain is replaced by a stronger one: $[x) \cap T$ is a chain (nonempty). As a consequence, any classification tree is also a directed tree but, in addition, for any $x, y \in X$, either $x \wedge y = 0$ or x, y are comparable. Hence, any classification tree is CD-independent, and any maximal classification tree is a CD-base. Moreover, in [4, 7] the following assertions are proved, which are also true for classification trees and CD-independent sets:

Proposition 2.2. *Let L be a bounded lattice and $T \subseteq L$ a nonempty subset of it. Then, the following assertions are equivalent:*

(i) *T is a maximal classification tree in L;*
(ii) *$T \cup \{0\}$ is a CD-base of L.*

Remark 2.3. (i) If L is an atomistic lattice and $A(L)$ the set of its atoms, then any CD-base as well as any maximal classification tree of L contains $A(L)$.

(ii) If T is a classification tree and we add some atoms to it, then the union still remains a classification tree. (If $S \subseteq A(L)$, then $T \cup S$ is a classification tree.) [7] Therefore, any maximal classification tree of L contains all the atoms of L.

The notion of contexts is applied in Group Technology problems. This engineering field exploits similarities between technological objects and divides them into relatively homogenous groups, extent partitions (classes) in order to optimize manufacturing processes. The concept of a classification tree also appears in the Group Technology. In Group Technology, by the term "classification tree", we refer to a tree constituted from extents belonging to a fixed context.

Definition 2.4. Let $K = (G, M, I)$ be a finite context and $\mathcal{E}(G, M, I)$ the lattice of concept extents belonging to it. Let $\mathcal{T} \subseteq \mathcal{E}(G, M, I)$ be a classification tree in the lattice $\mathcal{E}(G, M, I)$. $\mathcal{T}$ is called a complete classification tree if, for any maximal antichain $\{E_i |\ i \in I\} \subseteq \mathcal{T}$, we have $\bigcup_{i \in I} E_i = G$. $\mathcal{T}$ is a maximal complete classification tree if there is no other complete classification tree which contains it as a proper subfamily.

Remark 2.5. Since any pair A, B of sets in $\mathcal{T}$ is either comparable or disjoint, the maximal antichains of $\mathcal{T}$ must be extent partitions of G. Therefore, all elements of such a complete classification tree are box extents, ie. $\mathcal{T}$ is a classification tree in the lattice of box extents.

Proposition 2.6. *Let (G, M, I) be a finite context, $\mathcal{E}(G, M, I)$ the extent lattice of the context and $\mathcal{B}(G, M, I)$ the lattice of the box extents of the context. Then, the followings are equivalent:*

(i) *$\mathcal{T} \subseteq \mathcal{E}(G, M, I)$ is a complete classification tree;*
(ii) *$\mathcal{T}$ is a classification tree in $\mathcal{B}(G, M, I)$ and each maximal antichain of it $\{E_i|\ i \in I\} \subseteq \mathcal{T}$ is a complete orthogonal system in $\mathcal{B}(G, M, I)$.*

Proof. (i)$\Rightarrow$(ii) First, we will prove that $\mathcal{T}$ is a classification tree in the box extent lattice $\mathcal{B}(G, M, I)$. In view of Remark 2.5, each element $A \in \mathcal{T}$ is a box extent, hence $\mathcal{T} \subseteq \mathcal{B}(G, M, I)$. As $\mathcal{T}$ is finite, any $A \in \mathcal{T}$ must be a member in a maximal antichain $\{E_i|\ i \in I\} \subseteq \mathcal{T}$, ie. $A = E_k$ for some $k \in I$. Since the elements of an antichain are incomparable, for any two elements $E_i, E_j \in \mathcal{T}$, $i, j \in I$, we have $E_i \cap E_j = \emptyset$. From here it follows that any maximal antichain $\{E_i|\ i \in I\} \subseteq \mathcal{T}$ is a part of an orthogonal system in $\mathcal{B}(G, M, I)$, so it is itself an orthogonal system in $\mathcal{B}(G, M, I)$. As $\mathcal{T}$ is a complete classification tree in $\mathcal{E}(G, M, I)$, by definition $\bigcup\limits_{i \in I} E_i = G$. This implies that the orthogonal system $\{E_i|\ i \in I\}$ cannot be extended with a new element $E_0 \in \mathcal{B}(G, M, I)$, $E_0 \neq \emptyset$. Indeed, if this extension were possible, then $E_i \cap E_0 = \emptyset$ for all $i \in I$ would imply that $E_0 = G \cap E_0 = \left(\bigcup\limits_{i \in I} E_i \right) \cap E_0 = \emptyset$, a contradiction. Hence, $\{E_i|\ i \in I\}$ is a complete orthogonal system in $\mathcal{B}(G, M, I)$. The fact that $\mathcal{T}$ is a classification tree in $\mathcal{B}(G, M, I)$ is an obvious consequence of the fact that the 0-element and the 1-element and the meet-operation coincide in $\mathcal{E}(G, M, I)$ and $\mathcal{B}(G, M, I)$.

(ii)$\Rightarrow$(i) Assume that (ii) is satisfied. In order to prove (i), it is now sufficient to show that, for every maximal antichain $\{E_i|\ i \in I\} \subseteq \mathcal{T}$, we have $\bigcup\limits_{i \in I} E_i = G$. By our hypothesis, $\{E_i|\ i \in I\}$ is a complete orthogonal system. Let $g \in G$ and $g^{\square\square}$ be the box extents of it. Since $\{E_i|\ i \in I\} \cup g^{\square\square}$ is not an orthogonal system, we get that $g^{\square\square} \cap E_k \neq \emptyset$ for some $k \in I$. Since $g^{\square\square}$ is an atom in the lattice of box extents, this implies $g^{\square\square} \subseteq E_k$. Hence, $G \subseteq \bigcup\limits_{i \in I} E_i$. As $\bigcup\limits_{i \in I} E_i \subseteq G$, $\bigcup\limits_{i \in I} E_i = G$. Thus, (i) holds. $\square$

Corollary 2.7. *$\mathcal{T} \subseteq \mathcal{B}(G, M, I)$ is a maximal classification tree in $\mathcal{B}(G, M, I)$ if and only if $\mathcal{T}$ is a maximal complete classification tree in the extent lattice $\mathcal{E}(G, M, I)$.*

Proof. Let $\mathcal{T}$ be a maximal classification tree in $\mathcal{B}(G, M, I)$. Then, $\mathcal{T}$ contains all atoms of the lattice $\mathcal{B}(G, M, I)$. We are going to prove that $\mathcal{T}$ is a complete classification tree in $\mathcal{E}(G, M, I)$. Let S be a maximal antichain in $\mathcal{T}$. As $\mathcal{T}$ is a CD-independent set, S is an orthogonal system. We have to prove that S is a complete orthogonal system.

We denote by $A(\mathcal{B}(G,M,I))$ the set of all atoms of the lattice $\mathcal{B}(G,M,I)$. It is easy to see that $A(\mathcal{B}(G,M,I))$ is a complete orthogonal system. In [4], it was proved that an orthogonal system O is complete whenever $A(L) \leqslant O$. (where $A(L)$ is the set of all atoms of the lattice L and O is an orthogonal system). Thus, it is sufficient to prove that $A(\mathcal{B}(G,M,I) \leqslant S$. Indeed, we already know that $A(\mathcal{B}(G,M,I) \subseteq \mathcal{T}$. If this inequation is not satisfied, then there exists an atom a in $\mathcal{B}(G,M,I)$ which is not smaller than any element of S. Since $a \in \mathcal{T}$, $S \cup \{a\}$ is an antichain in $\mathcal{T}$, which is a contradiction. Thus, we have $A(\mathcal{B}(G,M,I)) \leqslant S$ and S is a complete orthogonal system. By Proposition 2.6, $\mathcal{T}$ is a complete classification tree in the lattice $\mathcal{E}(G,M,I)$. Now, let $\mathcal{F} \subseteq \mathcal{E}(G,M,I)$ be a complete classification tree which contains $\mathcal{T}$. Hence, using the characterization in Proposition 2.6, $\mathcal{F}$ is also a classification tree in the lattice $\mathcal{B}(G,M,I)$ and, by definition, $\mathcal{T}$ is a maximal classification tree in $\mathcal{B}(G,M,I)$ so thet we obtain $\mathcal{F} = \mathcal{T}$. Thus, $\mathcal{T}$ is a maximal complete classification tree in $\mathcal{E}(G,M,I)$.

Conversely, assume that $\mathcal{T} \subseteq \mathcal{E}(G,M,I)$ is a maximal complete classification tree. Hence, in view of Proposition 2.5, $\mathcal{T}$ is also a classification tree in the box extent lattice $\mathcal{B}(G,M,I)$ Thus, it is a subset of a maximal classification tree $\mathcal{M}$ of $\mathcal{B}(G,M,I)$. Therefore, $\mathcal{M}$ is a complete classification tree in the lattice $\mathcal{E}(G,M,I)$ and $\mathcal{T} \subseteq \mathcal{M}$ is maximal. We obtain $\mathcal{T} = \mathcal{M}$.

Summarizing, we deduce that $\mathcal{T}$ is a maximal classification tree in the lattice $\mathcal{B}(G,M,I)$. □

3. Classification trees in the lattice of box extents

In what follows we will consider a formal context $K = (G,M,I)$ with a concrete and fixed technical meaning, where G and M are finite and nonempty sets, G denotes a fixed set of technical objects, M denotes a fixed set of some possible, technically relevant properties, and for any $g \in G$ and any $m \in M$, gIm means that the object (part) g has the property m. Additionally, we assume that the context does not contain rows or columns filled with only zeros. A full zero column means that the corresponding property is not of any of the parts so that it is irrelevant; a full zero row corresponds to a part g possessing none of the properties from our list.

Let $K = (G,M,I)$ be a context and $K_H = (H,M,I \cap H \times M)$ a subcontext of it. We say that $K = (G,M,I)$ is a one-object extension of $K_H = (H,M,I \cap H \times M)$, where $H \subseteq G$, if there exists a $z \in G$ such that $H = G \setminus \{z\}$. In this section we use the results of [3], where the authors studied changes of the box extents of a one-object extension of the context. In [3], it was shown that the intersection of box extents is also a box extent. In [3], it was also proved that any extent partition of the subcontext is also an extent partition of K and the box extents of the subcontext are also box extents of K and the following propositions:

Proposition 3.1. [3] *If $\pi = \{A_k \mid k \in K\}$ is an extent partition of K then*

$$\pi_H = \{A_k \cap H \mid k \in K\} \setminus \{\emptyset\}$$

is an extent partition of the subcontext $(H,M,I \cap H \times M)$. If π is the finest extent partition of K, then π_H is called the restriction of the extent partition π and conversely. π_H is not necessarily the finest extent partition of H.

Corollary 3.2. [3] *If E is a box extent of (G, M, I), then $E \cap H$ is a box extent of $(H, M, I \cap H \times M)$.*

Corollary 3.3. [3] *$\mathcal{B}(G, M, I)$ is a complete atomistic lattice. The atomic box extents are the classes of the finest extent partition $\pi_\Box$.*

Proposition 3.4. [3] *If E is a box extent of (G, M, I) and $H = G \setminus \{z\}$ or some $z \in G$, then*

(i) *E is a box extent of $(H, M, I \cap H \times M)$with $E \cap\ z^{\Box\Box} = \emptyset$ or*
(ii) *$E \setminus \{z\}$ is a box extent of $(H, M, I \cap H \times M)$.*

Proposition 3.5. [3] *If $H = G \setminus \{z\}$, then A is a class of finest extent partition of $\pi_\Box$ of (G, M, I) if and only if*

(i) *either $A = z^{\Box\Box}$, or*
(ii) *A is a class, disjoint from $z^{\Box\Box}$, of the finest extent partition of the subcontext $(H, M, I \cap H \times M)$.*

Theorem 3.6. [3] *Let (G, M, I) be a context, E a box extent of the subcontext $(H, M, I \cap H \times M)$ with $H = G \setminus \{z\}$. Then,*

(i) *E is a box extent of (G, M, I) if and only if $z^{\Box\Box} \cap E'' = \emptyset$;*
(ii) *$E^* = E \cup \{z\}$ is a box extent of (G, M, I) if and only if $z^{\Box\Box} \setminus \{z\} \subseteq E$ and $(E \cup \{z\})'' = E \cup \{z\}$.*

In other words, in view of the theorem below, there are two possibilities:
1) E is also box extent in the new context iff $z^{\Box\Box} \cap E'' = \emptyset$;
2) or $E \cup \{z\}$ is a box extent in the new context iff $z^{\Box\Box} \setminus \{z\} \subseteq E$ and $(E \cup \{z\})'' = E \cup \{z\}$.

In what follows, we show that, during an elementary extension of the context, each of the classification trees of the newly created box extent lattice can be obtained by modifying the classification trees of the box extent lattice of the original, smaller context.

Proposition 3.7. *Let $K_H = (H, M, I \cap H \times M)$ be a subcontext of $K = (G, M, I)$ and $\mathcal{T}$ a classification tree in the box extent lattice $\mathcal{B}(G, M, I)$. Then, the set*

$$\mathcal{T}_H = \{E \cap H \mid E \in \mathcal{T}\}$$

is a classification tree in the box lattice $\mathcal{B}(K_H)$.

Proof. In [3] it was proved that if E is a box extent of the context (G, M, I), then $E \cap H$ is the box extent of the subcontext $(H, M, I \cap H \times M)$. Thus, for any $E \in \mathcal{T}$, $E \cap H$ is a box extent of the context K_H, so $\mathcal{T}_H \subseteq \mathcal{B}(K_H)$. Since $G \in \mathcal{T}$, $H = G \cap H \in \mathcal{T}_H$ and also H is the largest element of the box extent lattice $\mathcal{B}(K_H)$. We have to prove that $\mathcal{T}_H$ is a classification tree. It is sufficient to prove that $\mathcal{T}_H$ is a CD-independent set of the lattice $\mathcal{B}(K_H)$. Let $E_1 \cap H$ and $E_2 \cap H$ be two incomparable elements of $\mathcal{T}_H$. Then, $E_1 \cap H \neq \emptyset$ and $E_2 \cap H \neq \emptyset$ and $E_1, E_2 \in \mathcal{T}$ are also incomparable. Since $\mathcal{T}$ is classification tree and also a CD-independent set, we have $E_1 \cap E_2 = \emptyset$. Then, $(E_1 \cap H) \cap (E_2 \cap H) = E_1 \cap E_2 \cap H = \emptyset$. Thus, we proved that $\mathcal{T}_H$ is also a CD-independent set in the lattice $\mathcal{B}(K_H)$. Therefore, $\mathcal{T}_H$ is a classification tree in $\mathcal{B}(K_H)$. $\Box$

Next, we will show how to construct a classification tree after having a one-object extension of a context:

Theorem 3.8. *Let $K_H = (H, M, I \cap H \times M)$ be a subcontext of a finite context (G, M, I) such that $H = G \setminus \{z\}$, $z^{\Box\Box} \neq \{z\}$. Let $\mathcal{T}$ be a classification tree in the box extent lattice $\mathcal{B}(K_H)$. Then:*

(i) $\mathcal{T}^{(1)} = \{E \in \mathcal{T} | \ E \in \mathcal{B}(G, M, I)\}$ *is an order ideal in* $(\mathcal{T}, \subseteq)$ *and* $\mathcal{T}^{(2)} = \{E \in \mathcal{T} | \ E \cup \{z\} \in \mathcal{B}(G, M, I)\}$ *a finite chain in* $\mathcal{T}$ *and* $\mathcal{T}^{(1)} \cap \mathcal{T}^{(2)} = \emptyset$;

(ii) $\mathcal{T}^* = \mathcal{T}^{(1)} \cup \left\{E \cup \{z\} \mid E \in \mathcal{T}^{(2)}\right\}$ *is a classification tree in the lattice* $\mathcal{B}(G, M, I)$.

(iii) *If* $\mathcal{T}$ *contains all the atoms of the box extent lattice* $\mathcal{B}(K_H)$, *then* $\mathcal{T}^* \cup \{z^{\Box\Box}\}$ *is a classification tree in* $\mathcal{B}(G, M, I)$ *which contains all the atoms of it.*

Proof. (i) Let $E \in \mathcal{T}^{(1)}$ and $F \subseteq E$, $F \in \mathcal{T}$. Since E is a box extent of (G, M, I), in view of Theorem 3.6 we get $z^{\Box\Box} \cap E'' = \emptyset$. Since $F'' \subseteq E''$, we have $z^{\Box\Box} \cap F'' = \emptyset$, which in view of Theorem 3.6 means that $F \in \mathcal{T}^{(1)}$. Thus, $\mathcal{T}^{(1)}$ is an order ideal in $\mathcal{T}$. We have to prove now that $\mathcal{T}^{(2)}$ is a finite chain in $\mathcal{T}$. We take the set $\mathcal{C} = \left\{E \in \mathcal{T} | \ z^{\Box\Box} \setminus \{z\} \subseteq E\right\}$. Obviously, $H \in \mathcal{C}$, then $\mathcal{C} \neq \emptyset$. As $z^{\Box\Box} \setminus \{z\} \neq \emptyset$ for any $E \in \mathcal{C}$, we get $E \cap z^{\Box\Box} \neq \emptyset$. Thus, $\mathcal{T}^{(1)}$ and $\mathcal{C}$ have no common elements. Take $E \in \mathcal{C}$ and $F \in \mathcal{T}$ such that $E \subseteq F$. We have $z^{\Box\Box} \setminus \{z\} \subseteq F$, so $F \in \mathcal{C}$. Therefore, $\mathcal{C}$ is an order filter of $(\mathcal{T}, \subseteq)$ and, since H is finite and $\mathcal{C} \subseteq$, $\mathcal{B}(K_H)$ is lower bounded by it's minimal elements. We show that $\mathcal{C}$ has only one minimal element $E_1 \in \mathcal{C}$. Assume that $E_2 \in \mathcal{C}$ is a minimal element in $\mathcal{C}$ and $E_1 \neq E_2$. Since $E_1, E_2 \in \mathcal{T}$ and E_1, E_2 are incomparable, we have $E_1 \cap E_2 = \emptyset$, which is a contradiction because $z^{\Box\Box} \setminus \{z\} \subseteq E_1 \cap E_2$ and $z^{\Box\Box} \setminus \{z\}$ are nonempty by hypothesis. Thus, E_1 is the least element of $\mathcal{C}$ and $\mathcal{C}$ is equal to $[E_1) \cap \mathcal{T}$. Since $\mathcal{T}$ is a classification tree, $\mathcal{C}$ must be a chain. Consider now the set $\mathcal{T}^{(2)}$. By Theorem 3.6 any $E \in \mathcal{T}^{(2)}$ satisfies $z^{\Box\Box} \setminus \{z\} \subseteq E$. Hence $\mathcal{T}^{(2)} \subseteq \mathcal{C}$. Therefore, $\mathcal{T}^{(2)}$ is a finite chain and $\mathcal{T}^{(1)} \cap \mathcal{T}^{(2)} = \emptyset$.

(ii) As $\mathcal{T}^{(1)} \subseteq \mathcal{T}$, we have that $\mathcal{T}^{(1)}$ is a CD-independent set. Observe that $\left\{E \cup \{z\} \mid E \in \mathcal{T}^{(2)}\right\}$ is a chain in $\mathcal{B}(G, M, I)$. Indeed, let $E_1 \cup \{z\}$ and $E_2 \cup \{z\}$ be two elements of this set. Since $\mathcal{T}^{(2)}$ is a chain, we have $E_1 \subseteq E_2$ or $E_2 \subseteq E_1$ and $E_1 \cup \{z\} \subseteq E_2 \cup \{z\}$, or conversely $E_2 \cup \{z\} \subseteq E_1 \cup \{z\}$. We show that the set $\mathcal{T}^{(1)} \cup \left\{E \cup \{z\} \mid E \in \mathcal{T}^{(2)}\right\}$ is also CD-independent.

Take $E_1 \in \mathcal{T}^{(1)}$ and $E_2 \in \mathcal{T}^{(2)}$. As $E_1, E_2 \in \mathcal{T}$ and $\mathcal{T}$ is a classification tree, we have the following three cases: $E_1 \subseteq E_2$ or $E_2 \subseteq E_1$ or $E_1 \cap E_2 = \emptyset$. We have to show that in each case the sets E_1 and $E_2 \cup \{z\}$ are either comparable or disjoint.

Clearly, in the first case, $E_1 \subseteq E_2 \cup \{z\}$.

In the second case, $E_2 \subseteq E_1$ implies $E_2 \in \mathcal{T}^{(1)}$ because $\mathcal{T}^{(1)}$ is an order ideal in $\mathcal{T}$. However, this is impossible, because $E_2 \in \mathcal{T}^{(2)}$ and $\mathcal{T}^{(1)} \cap \mathcal{T}^{(2)} = \emptyset$.

Let us consider now the case $E_1 \cap E_2 = \emptyset$. Since $E_1 \in \mathcal{T}^{(1)}$, we have $E_1'' \cap z^{\Box\Box} = \emptyset$ and this implies also $E_1 \cap \{z\} = \emptyset$. Therefore, we obtain $E_1 \cap (E_2 \cup \{z\}) = \emptyset$.

Thus, we have proved that $\mathcal{T}^* = \mathcal{T}^{(1)} \cup \left\{E \cup \{z\} \mid E \in \mathcal{T}^{(2)}\right\}$ is CD-independent.

Since $G = H \cup \{z\}$ and $H \in \mathcal{T}$ (and $G \in \mathcal{B}(G, M, I)$), we obtain $G \in \mathcal{T}^*$. As G is the greatest element of the lattice $\mathcal{B}(G, M, I)$, in view of Remark 2.5 ,$\mathcal{T}^*$ is a classification tree in $\mathcal{B}(G, M, I)$.

(iii) Observe that because $z^{\Box\Box}$ is an atom in the lattice $\mathcal{B}(G,M,I)$, if we add it to the classification tree $\mathcal{T}^*$, then in view of Proposition 2.2 (ii) $\mathcal{T}^* \cup \{z^{\Box\Box}\}$ remains a classification tree in $\mathcal{B}(G,M,I)$.

Finally, assume that $\mathcal{T}$ contains all the atoms of the lattice $\mathcal{B}(K_H)$. We have to show that the classification tree $\mathcal{T}^* \cup \{z^{\Box\Box}\}$ contains all the atoms of the lattice $\mathcal{B}(G,M,I)$. Evidently, it contains the atom $z^{\Box\Box}$, too. On the other hand, the atoms of $\mathcal{B}(G,M,I)$ are blocks of the finest extent partition $\pi_\Box$ of (G,M,I). Then, in view of Proposition 3.5 all the other atoms of $\mathcal{B}(G,M,I)$ which are different from $z^{\Box\Box}$ are also atoms of the lattice $\mathcal{B}(K_H)$, and belong to $\mathcal{T}$. Therefore, this atoms belong to $\mathcal{T}^{(1)}$ by the construction of $\mathcal{T}^{(1)}$. Since $\mathcal{T}^{(1)} \subseteq \mathcal{T}^* \cup \{z^{\Box\Box}\}$, $\mathcal{T}^* \cup \{z^{\Box\Box}\}$ contains all the atoms of $\mathcal{B}(G,M,I)$. $\Box$

Theorem 3.9. *Let $K_H = (H, M, I \cap H \times M)$ be a subcontext of the finite context $K = (G,\, M, I)$ such that $H = G \setminus \{z\}$, $z^{\Box\Box} \neq \{z\}$. Then,*

(i) *For each classification tree $\mathcal{T}_G \subseteq \mathcal{B}(G,M,I)$ there exists a classification tree $\mathcal{T}_H \subseteq \mathcal{B}(K_H)$ such that the equality $\mathcal{T}_G = \mathcal{T}_H^*$ is satisfied.*

(ii) *If $\mathcal{T}_G \subseteq \mathcal{B}(G,M,I)$ is a maximal classification tree, then, in $\mathcal{B}(K_H)$, there is also a maximal classification tree $\mathcal{M}$ such that $\mathcal{T}_G = \mathcal{M}^*$.*

Proof. (i) In view of Proposition 3.7, $\mathcal{T}_H = \{E \cap H \mid E \in \mathcal{T}_G\}$ is a classification tree in $\mathcal{B}(K_H)$. We show that the classification tree

$$\mathcal{T}_H^* = \mathcal{T}_H^{(1)} \cup \left\{E \cup \{z\} \mid E \in \mathcal{T}_H^{(2)}\right\} = \{E \in \mathcal{T}_H \mid E \in \mathcal{B}(G,M,I)\} \cup$$
$$\{E \cup \{z\} \mid E \in \mathcal{T}_H,\, E \cup \{z\} \in \mathcal{B}(G,M,I)\}$$

assigned to $\mathcal{T}_H$ by Theorem 3.8 is equal to $\mathcal{T}_G$. ($\mathcal{T}_H^{(1)}$, $\mathcal{T}_H^{(2)}$ were defined in Theorem 3.8)

Let $E \in \mathcal{T}_G$ be arbitrary. Then, $F = E \cap H \in \mathcal{T}_H$ by the definition of $\mathcal{T}_H$. We have only two cases:

(1) if $E \subseteq H$, then $E = F$;

(2) if $E \nsubseteq H$, then $E = F \cup \{z\}$.

In the first case, $F \in \mathcal{T}_H^{(1)}$ so we have $E = F \in \mathcal{T}_H^{(1)} \subseteq \mathcal{T}_H^*$.

In the second case, as $F \cup \{z\} \in \mathcal{B}(G,M,I)$ we have $E = F \cup \{z\} \in \mathcal{T}_H^*$.

In both cases, $\mathcal{T}_G \subseteq \mathcal{T}_H^*$ because $E \in \mathcal{T}_H^*$.

Conversely, let $E \in \mathcal{T}_H^*$ be arbitrary. By the construction of $\mathcal{T}_H^*$, we have two cases either $E \in \mathcal{T}_H^{(1)} \subseteq \mathcal{T}_H$ or $E = F \cup \{z\}$, $F \in \mathcal{T}_H^{(2)} \subseteq \mathcal{T}_H$.

Then, by the definition of $\mathcal{T}_H$, there exists an $A \in \mathcal{T}_G$ such that, in the first case, we have $A \cap H = E$ while $A \cap H = F$ in the second case.

First we show that $z \in A$ is not possible. Indeed, in view of Theorem 3.6, as $E \in \mathcal{B}(K_H)$, $E \cap z^{\Box\Box} = E'' \cap z^{\Box\Box} = \emptyset$. On the other hand, from $z \in A$ we would get $z^{\Box\Box} \subseteq A$ so $z^{\Box\Box} \setminus \{z\} \subseteq A \cap H = E$. In combination with the first result, this would imply $z^{\Box\Box} \setminus \{z\} = \emptyset$,i.e., $z^{\Box\Box} = \{z\}$, contradiction.

Thus, in the case of $A \cap H = E$, we have $z \notin A$. Therefore, $A \subseteq H$ and $E = A \cap H = A \in \mathcal{T}_G$.

In case two, $A \cap H = F$. Observe that $A \subseteq H$ would imply $F = A \in \mathcal{B}(G,M,I)$ which means that $F \in \mathcal{T}_H^{(1)}$. However, this is not possible because, by definition, $F \in \mathcal{T}_H^{(2)}$ and $\mathcal{T}_H^{(2)} \cap \mathcal{T}_H^{(1)} = \emptyset$. Therefore, $A \nsubseteq H$, which means $z \in A$. Then, $A = F \cup \{z\} = E$. Thus, we get $E = A \in \mathcal{T}_G$.

Since, in both cases, we proved $E \in \mathcal{T}_G$, it follows $\mathcal{T}_H^* = \mathcal{T}_G$.

(ii) Assume that $\mathcal{T}_G$ is a maximal classification tree in the lattice $\mathcal{B}(G, M, I)$. Since $\mathcal{T}_G = \mathcal{T}_H^*$, if $\mathcal{T}_H$ is a maximal classification tree in $\mathcal{B}(K_H)$, we are done. If $\mathcal{T}_H$ is not maximal, then, in $\mathcal{B}(K_H)$, there exists a maximal classification tree $\mathcal{M}$ in $\mathcal{B}(K_H)$ such that $\mathcal{T}_H \subseteq \mathcal{M}$. Then, we have $\mathcal{T}_G = \mathcal{T}_H^* \subseteq \mathcal{M}^*$. In view of Theorem 3.8 $\mathcal{M}^*$ is also a classification tree in $\mathcal{B}(G, M, I)$. As $\mathcal{T}_G$ is a maximal classification tree by hypothesis, we get $\mathcal{T}_G = \mathcal{M}^* = \mathcal{T}_H^*$. □

4. An algorithm for tree-construction in the box extent lattice

In chapter 3 of the article, we saw that the box extents of a context are not really changed after a one-object extension or reduction. We add an object to the context which has attributes from an existing M attribute set of the context. We showed that, during a one-object extension of the context, each of the classification trees of the newly created box extent lattice can be obtained by modifying the classification trees of the box extent lattice of the original, smaller context.

Based on Theorem 3.8 we devise an algorithm that, starting from a classification tree of the box extent lattice of the smaller context $(H, M, I \cap H \times M)$, provides a classification tree of the extended context (G, M, I) that contains the new elements inserted. The effectiveness of this method is given by the fact that it is sufficient to know the original context, the classification tree of the box extent lattice, and its box extents. We do not need a new box extent of the extended context mesh elements (except for the one that is the box extent of the new element).

To construct the classification trees of the box extent lattice, we will use a recursive ORTOFA algorithm presented in [8]. For this algorithm, we need the original context K and the box extents, which are contained in the matrix DS. The next step is to determine $z^{\square\square}$. For finding the extent partition $E = z^{\square\square}$, containing the new inserted element, we will use Algorithm 2 presented by A. Körei in [3].

The function `LISTA_FA` first searches for the elements of the order ideal part of our classification tree, ie., we check if $z^{\square\square}$ is smaller than any box extent stored in the matrix L. If this condition is satisfied, we put the element in the matrix $S1$. After that, using the KOBJ(KTUL), we verify that the elements of $S1$ are in fact box extents, and compare them with the function. In the last step, we find the chain part of our classification tree.

In the algorithm we used two functions:

- `BENNE(A,B)` – find out whether matrix A of the box extents contains the elements of B;
- `BERAK(A,V,m)` – a procedure that inserts in A the vector V as a new $(m+1)$-th row.

In the function `KOBJ(KTUL)`, the function `KTUL` finds the common properties of the object set, and the function `KOBJ` finds the common objects of the attribute set and uses these two functions one after the other to get the set A'' of the object set A.

We use the following matrices:

- $DS(m \times n)$ contains the box extents,
- $L(m \times n)$ contains a copy of the box extents,

$S1(m \times (n+1))$ contains the elements above to $z^{\square\square}$,
$S(m \times n)$ contains the chain above to $z^{\square\square}$ and $F(m \times n)$ stores the elements below.

```
LISTA_FA(DS)
L ⟵ DS
S1 ⟵ ∅
S ⟵ ∅
F ⟵ ∅
k ⟵ 0
for i ⟵ 1 to m
     do  BENNE ⟵ true
        for j ⟵ 1 to n
            do if z□□[j]>L[i][j]
                 then BENNE ⟵ false
       if BENNE
           then k++
                BERAK(S1,L[i],k)
 for i ⟵ 1 to k
     do S1[i][n+1] ⟵ 1
 A ⟵ S1^T
 A ⟵ KOBJ(KTUL(A,K),K)  */here K means the extended context
 A ⟵ A^T
 l ⟵ 0
 for i ⟵ 1 to k
     do   dext ⟵ true
          for j ⟵ 1 to (n+1)
              do if A[i][j] ≠ S1[i][j]
                     then dext ⟵ false
          if dext
              then l++
                   BERAK(S,S1[i],l)
 h ⟵ 0
 for i ⟵ 1 to m
     do if BENNE(L[i],z□□)
            then h++
                 BERAK(F,L[i],h)
 return S,F

BERAK(A,V,m)
m++
for t ← 1 to n
    do A[m][t] ⟵ V[t]
return A,m

BENNE(A,B)
bent ← true
for i ← 1 to n
```

```
3.      do if A[i] > B[i]
4.            then bent ← false
5.  return bent
```

The method is efficient as it is sufficient to know the original context, the classification tree of the box extent lattice and its box extents, without the need to know the new box extents of the extended context mesh elements (except for the one that is the box extent of the new element).

The main parts of the process are the dual matrix operations operating cycles. The run-time of this algorithm is $O(n^2)$ polynomial time, which means the worst run time for a sufficiently large $n = \max\{n, m, k\}$. However, the method contains many one-line instruction (decisions, variable value increase with 1) if the value of n is not too big, then the running time is linear, $O(n)$, because the number of steps, the executable instructions, is n. Summing up, the worst running time is a second-order polynomial in time ([1]).

References

[1] T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein, *Introduction to Algorithms*, 3rd ed., MIT Press and McGraw-Hill, 2009.

[2] B. Ganter and R. Wille, *Formal Concept Analysis. Mathematical Foundations*, Springer Verlag, Berlin, 1999.

[3] B. Ganter, A. Körei and S. Radeleczki, *Extent partitions and context extensions*, Math. Slovaca, **63** (2013), 693–706.

[4] E. K. Horváth and S. Radeleczki, *Notes on CD-independent subsets*, Acta Sci. Math. **78** (2012), 3–24.

[5] A. Körei and S. Radeleczki, *Box elements in a concept lattice*, in: B. Ganter and L. Kwuida (eds.), Contribution to ICFCA 2006, Verlag Allgemeine Wissenschaft, Dresden, 2006, 41–56.

[6] S. Radeleczki, *Classification systems and the decomposition of a lattice into direct products*, Math. Notes **1** (2000), 145–156.

[7] L. Veres, *Generalization of classification trees for a poset*, Creat. Math. Inform. **17** (2008), 72–77.

[8] L. Veres, *Classification trees and orthogonal systems*, microCAD 2007 International Scientific Conference, Vol. G, Miskolc, 2007, 63–68.

Laura Veres, Institute of Mathematics, University of Miskolc, 3530 Miskolc-Egyetemvaros, Hungary

e-mail: matvlaura@gmail.com