

# From formal concepts to analogical complexes

Laurent Miclet<sup>1</sup> and Jacques Nicolas<sup>2</sup>

<sup>1</sup> Université de Rennes 1, UMR IRISA, Dyliss team, Rennes, France,  
miclet@univ-rennes1.fr

<sup>2</sup> Inria Rennes, France, jacques.nicolas@inria.fr

**Abstract.** Reasoning by analogy is an important component of common sense reasoning whose formalization has undergone recent improvements with the logical and algebraic study of the analogical proportion. The starting point of this study considers analogical proportions on a formal context. We introduce analogical complexes, a companion of formal concepts formed by using analogy between four subsets of objects in place of the initial binary relation. They represent subsets of objects and attributes that share a maximal analogical relation. We show that the set of all complexes can be structured in an analogical complex lattice and give explicit formulae for the computation of their infimum and supremum.

**Keywords:** analogical reasoning, analogical proportion, formal concept, analogical complex, lattice of analogical complexes

## 1 Introduction

Analogical reasoning [4] plays an important role in human reasoning. It enables us to draw plausible conclusions by exploiting parallels between situations, and as such has been studied in AI for a long time, e.g., [5, 9] under various approaches [3]. A key pattern which is associated with the idea of analogical reasoning is the notion of analogical proportion (AP), i. e. a statement between two pairs  $(A, B)$  and  $(C, D)$  of the form ‘ $A$  is to  $B$  as  $C$  is to  $D$ ’ where all elements  $A, B, C, D$  are in a same category .

However, it is only in the last decade that researchers working in computational linguistics have started to study these proportions in a formal way [6, 17, 19]. More recently, analogical proportions have been shown as being of particular interest for classification tasks [10] or for solving IQ tests [2]. Moreover, in the last five years, there has been a number of works, e.g., [11, 15] studying the propositional logic modeling of analogical proportions.

In all previous cases, the ability to work on the set of all possible analogical proportions is required, either for checking missing objects or attributes or for making informed recommendations or more generally ensuring the completeness and efficiency of reasoning. In practice the analysis of objects composed of binary attributes, such as those studied by Formal Concept Analysis, is an important and easy context where AP are used. The question is whether it is possible to obtain a good representation of the space of all AP by applying the principles of

FCA. A heuristic algorithm to discover such proportions by inspecting a lattice of formal concepts has been proposed in [14]. Moreover, a definition of an analogical proportion between formal concepts has been given in [13], as a particular case of proportions between elements of a lattice, studied also in [18].

In this paper, we are interested in a slightly different task involving a more integrated view of concept categorization and analogy: looking for the structure of the space of all AP. Our goal is to build an extension of formal concepts considering the presence of analogical proportions as the funding relation instead of the initial binary relation between objects and attributes. We call this extension *analogical complexes*, which isolate subcontexts in formal contexts with a certain structure reflecting the existence of a maximal analogical proportion between subsets of objects and subsets of attributes.

## 2 Basics on Analogical Proportion

**Definition 1 (Analogical proportion [7, 12]).** *An analogical proportion (AP) on a set  $X$  is a quaternary relation on  $X$ , i.e. a subset of  $X^4$  whose elements  $(x, y, z, t)$ , written  $x : y :: z : t$ , which reads 'x is to y as z is to t', must obey the following two axioms:*

1. *Symmetry of 'as':  $x : y :: z : t \Leftrightarrow z : t :: x : y$*
2. *Exchange of means:  $x : y :: z : t \Leftrightarrow x : z :: y : t$*

*In case of formal contexts, objects are described by boolean attributes. An AP  $(x, y, z, t)$  between four Boolean variables exists if the following formula is true:*

$$(x \wedge \neg y) \Leftrightarrow (z \wedge \neg t) \text{ and } (y \wedge \neg x) \Leftrightarrow (t \wedge \neg z)$$

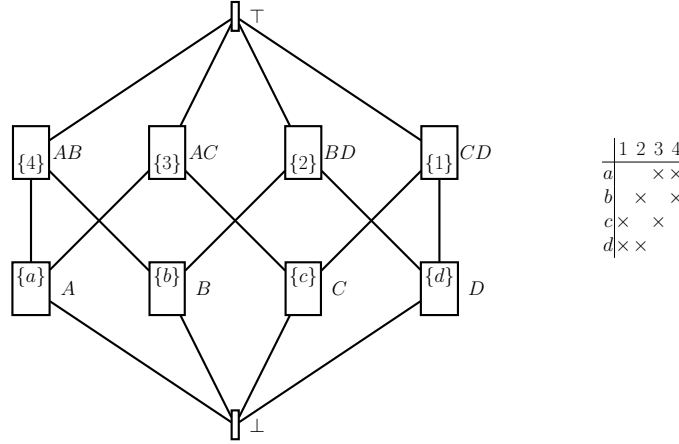
Basically, the formula expresses that the dissimilarity observed between  $x$  and  $y$  is the same as the dissimilarity between  $z$  and  $t$ . An equivalent formula is

$$x \neq y \Leftrightarrow (x = z \wedge y = t) \text{ and } x = y \Leftrightarrow z = t$$

It has 6 models of Boolean 4-tuples among the 16 possible ones. Note that this includes the trivial cases where  $x = y = z = t$ . Since we are only interested in this paper in non trivial analogical proportions, we further require that  $x \neq t$  and  $y \neq z$ . This reduces the number of possible Boolean 4-tuples in AP to four and it leads to the notion of analogical schema that we will use for the definition of analogical complexes.

**Definition 2 (Analogical schema).** *The binary matrix  $AS = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$  is called an analogical schema. We write  $AS(i, j)$  if the value at row  $i$  and column  $j$  of matrix  $AS$  is 1 (e.g.  $AS(1,3)$  and  $AS(1,4)$ ).*

The analogical schema may be seen as a formal context on four objects  $o_1, o_2, o_3, o_4$  that are in the non-trivial AP:  $o_1 : o_2 :: o_3 : o_4$ . The figure 1 shows the associated concept lattice. In this lattice,  $A \wedge D = B \wedge C$  and  $A \vee D = B \vee C$ . The figure also give names for each column and row profiles that we call object and attribute types: for instance the first column as type 1 and the second row as type b.



**Fig. 1.** Left: Concept lattice of an analogical schema (reduced labeling). Analogical schema with object and attribute types.

We use in this paper the zoo dataset proposed by R. Forsyth [8] for illustration purpose. We call smallzoo the formal context extracted from this database corresponding to attributes 2 to 9 and to the objects corresponding to the two largest classes 1 and 2. Moreover, this context has been clarified and we have chosen arbitrarily one object for each of the 10 different types of objects with different attribute profiles. The corresponding table is given below.

<b>smallzoo</b>	2	3	4	5	6	7	8	9	18
	hair	feathers	eggs	milk	airborne	aquatic	predator	toothed	type
1 aardvark	1	0	0	1	0	0	1	1	1
12 chicken	0	1	1	0	1	0	0	0	2
17 crow	0	1	1	0	1	0	1	0	2
20 dolphin	0	0	0	1	0	1	1	1	1
22 duck	0	1	1	0	1	1	0	0	2
28 fruitbat	1	0	0	1	1	0	0	1	1
42 kiwi	0	1	1	0	0	0	1	0	2
49 mink	1	0	0	1	0	1	1	1	1
59 penguin	0	1	1	0	0	1	1	0	2
64 platypus	1	0	1	1	0	1	1	0	1

The formal concept lattice is provided in figure 2, as computed by FCA Extension [16]. It contains 31 elements. The central elements (at least two objects and two attributes) are listed below:

$c(3)$	$(\{20; 49; 59; 64\}, \{7; 8\})$	$c(6)$	$(\{1; 20; 28; 49\}, \{5; 9\})$
$c(7)$	$(\{1; 20; 49; 64\}, \{5; 8\})$	$c(8)$	$(\{1; 20; 49\}, \{5; 8; 9\})$
$c(9)$	$(\{20; 49; 64\}, \{5; 7; 8\})$	$c(10)$	$(\{20; 49\}, \{5; 7; 8; 9\})$
$c(12)$	$(\{17; 42; 59; 64\}, \{4; 8\})$	$c(13)$	$(\{22; 59; 64\}, \{4; 7\})$
$c(14)$	$(\{59; 64\}, \{4; 7; 8\})$	$c(15)$	$(\{12; 17; 22; 42; 59\}, \{3; 4\})$
$c(16)$	$(\{17; 42; 59\}, \{3; 4; 8\})$	$c(17)$	$(\{22; 59\}, \{3; 4; 7\})$
$c(19)$	$(\{12; 17; 22\}, \{3; 4; 6\})$	$c(22)$	$(\{1; 28; 49; 64\}, \{2; 5\})$
$c(23)$	$(\{1; 28; 49\}, \{2; 5; 9\})$	$c(24)$	$(\{1; 49; 64\}, \{2; 5; 8\})$
$c(25)$	$(\{1; 49\}, \{2; 5; 8; 9\})$	$c(26)$	$(\{49; 64\}, \{2; 5; 7; 8\})$

*Example 1.* If one extracts in smallzoo the subcontext crossing (12, 28, 59, 49) - that is, (chicken, fruitbat, penguin, mink)- and (7, 2, 3, 6) -(aquatic, hair, feathers, airborne)-, it is clearly an analogical schema.

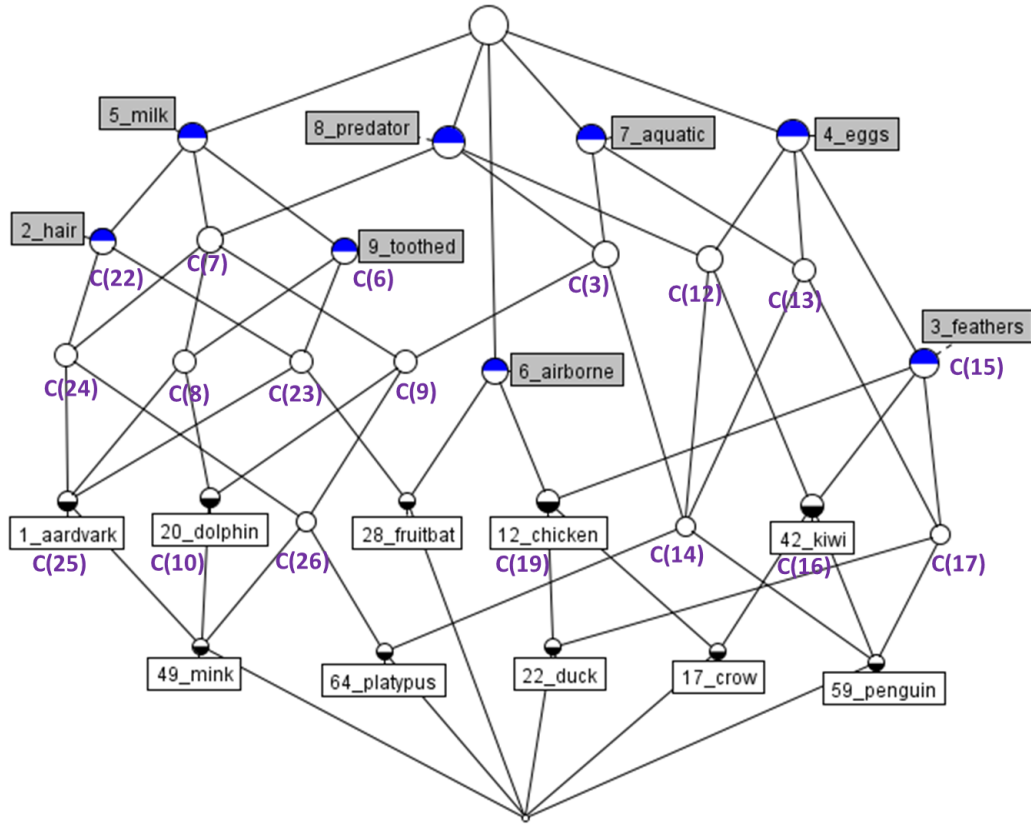
The 4-tuple (chicken : fruitbat :: penguin : mink) is an analogical proportion that finds a support using attributes (aquatic, hair, feathers, airborne). Each attribute reflects one of the four possible types of Boolean analogy. For instance, hair is false for chicken and penguin and true for fruitbat and mink whereas feathers is true for chicken and penguin and false for fruitbat and mink. The observed analogy can be explained thanks to this typology: the dissimilarity between chicken and fruitbat based on the opposition feather/hair is the same as the dissimilarity between penguin and mink and there are two other opposite attributes, airborne and aquatic, that explain the similarity within each 'is to' relation. Note that the analogical schema is fully symmetric and thus one could also in principle write AP between attributes: hair:feathers::aquatic: airborne.

### 3 An analogical complex is to an analogical proportion as a concept is to a binary relation

#### 3.1 Analogical complexes

A formal concept on a context  $(X, Y, I)$  is a maximal subcontext for which relation  $I$  is valid. We define analogical complexes in the same way: they are maximal subcontexts for which the 4-tuples are in AP. This requires to split objects and attributes in four classes.

**Definition 3 (Analogical complex).** *Given a formal context  $(X, Y, I)$ , a set of objects  $O \subseteq X$ ,  $O = O_1 \cup O_2 \cup O_3 \cup O_4$ , a set of attributes  $A \subseteq Y$ ,  $A = A_1 \cup A_2 \cup A_3 \cup A_4$ , and a binary relation  $I$ , the subcontext  $(O, A)$  forms an analogical complex  $(O_{1,4}, A_{1,4})$  iff*



**Fig. 2.** Formal concept lattice of formal context smallzoo. Drawing from Concept Explorer [20].

1. The binary relation is compatible with the analogical schema  $AS$ :  
 $\forall o \in O_i, i = 1..4, \forall a \in A_j, j = 1..4, I(o, a) \Leftrightarrow AS(i, j)$ .
2. The context is maximal with respect to the first property ( $\oplus$  denotes the exclusive or and  $\setminus$  the set-theoretic difference):  
 $\forall o \in X \setminus O, \exists j \in [1, 4], \exists a \in A_j, I(o, a) \oplus AS(i, j)$ .  
 $\forall a \in Y \setminus A, \exists i \in [1, 4], \exists o \in O_i, I(o, a) \oplus AS(i, j)$ .

The first property states that the value of an attribute for an object in a complex is a function of object type and attribute type (integer from 1 to 4) given by the analogical schema. The second property states that adding an object (resp. an attribute) to the complex would discard the first property for at least one attribute (resp. object) value. Note that the ways analogical schema or analogical complex are defined are completely symmetric. Thus the role of objects and attributes may be interchanged in all properties on analogical complexes.

*Example 2.* We extract two subcontexts from smallzoo, highlighting analogical schemas by sorting rows and columns.

	A <sub>1</sub>		A <sub>2</sub>			A <sub>3</sub>		A <sub>4</sub>
	a <sub>7</sub>	a <sub>8</sub>	a <sub>2</sub>	a <sub>5</sub>	a <sub>9</sub>	a <sub>3</sub>	a <sub>4</sub>	a <sub>6</sub>
O <sub>1</sub> o <sub>12</sub> (chicken)	0	0	0	0	0	1	1	1
O <sub>2</sub> o <sub>28</sub> (fruitbat)	0	0	1	1	1	0	0	1
O <sub>3</sub> o <sub>59</sub> (penguin)	1	1	0	0	0	1	1	0
O <sub>4</sub> o <sub>49</sub> (mink)	1	1	1	1	1	0	0	0

	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	A <sub>4</sub>
	a <sub>7</sub>			a <sub>6</sub>
O <sub>1</sub> o <sub>12</sub>	0			1
o <sub>17</sub>	0			1
o <sub>28</sub>	0			1
O <sub>2</sub> o <sub>12</sub>	0			1
o <sub>17</sub>	0			1
o <sub>28</sub>	0			1
O <sub>3</sub> o <sub>20</sub>	1			0
o <sub>49</sub>	1			0
o <sub>59</sub>	1			0
o <sub>64</sub>	1			0
O <sub>4</sub> o <sub>20</sub>	1			0
o <sub>49</sub>	1			0
o <sub>59</sub>	1			0
o <sub>64</sub>	1			0

These subcontexts are maximal in the sense that it is not possible to add an object or an attribute without breaking the analogical proportion. They are associated to the following analogical complexes:

$$((\{12\}, \{28\}, \{59\}, \{49\}), (\{7, 8\}, \{2, 5, 9\}, \{3, 4\}, \{6\}))$$

$$((\{12, 17, 28\}, \{12, 17, 28\}, \{20, 49, 59, 64\}, \{20, 49, 59, 64\}), (\{7\}, \emptyset, \emptyset, \{6\}))$$

The first example provides a strong analogical relation between four animals in the context smallzoo since it uses all attributes and all the types of analogy. Attribute clusters correspond to aquatic predators, toothed animals with hair and milk, birds (feathers and eggs) and flying animals (airborne). The second example shows some of the sets in analogical complexes can be empty. In such a case some sets may be duplicated. Among all complexes, those that exhibit all types of analogy are particularly meaningful: we call them complete complexes.

### 3.2 Complete analogical complexes (CAC)

**Definition 4.** A complex  $\mathcal{C} = (O_{1,4}, A_{1,4})$  is complete if none of its eight sets are empty.

By construction, if  $\mathcal{CA} = (O_{1,4}, A_{1,4})$  is a complete analogical complex and if  $\mathcal{A} = \bigcup_{i=1,4} A_i$ , the following formula holds:

$$\forall(o_1, o_2, o_3, o_4) \in O_{1,4}, \forall(a_1, a_2, a_3, a_4) \in A_{1,4}$$

$$(o_1^\uparrow \cap o_4^\uparrow) \cap \mathcal{A} = (o_2^\uparrow \cap o_3^\uparrow) \cap \mathcal{A} = \emptyset \text{ and } o_1^\uparrow \cup o_4^\uparrow = o_2^\uparrow \cap o_3^\uparrow = \mathcal{A}$$

The next proposition shows that CAC exhibits strong discrimination and similarity properties among pairs of objects and attributes. The similarity condition alone would lead to the concatenation of independent (non overlapping) formal concepts. The discrimination condition tempers this tendency by requiring the simultaneous presence of opposite pairs.

**Proposition 1.** *Let us define on a formal context  $FC = (X, Y, I)$  the relations:*

$$\text{discrimination}(o_i, o_j, a_k, a_l) = I(o_i, a_k) \wedge I(o_j, a_l) \wedge \neg I(o_i, a_l) \wedge \neg I(o_j, a_k).$$

$$\text{similarity}(o_i, o_j, a_k, a_l) = I(o_i, a_k) \wedge I(o_j, a_k) \wedge I(o_i, a_l) \wedge I(o_j, a_l).$$

*A complete analogical complex  $(O_{1,4}, A_{1,4})$  in  $FC$  corresponds to a maximal subcontext such that:*

1. *object pair discrimination (resp. similarity):  $\forall (o_i, o_j) \in O_i \times O_j, i \neq j, \exists (a_k, a_l) \in A_k \times A_l$  such that  $\text{discrimination}(o_i, o_j, a_k, a_l)$  (resp.  $\text{similarity}(o_i, o_j, a_k, a_l)$ );*
2. *attribute pair discrimination (resp. similarity):  $\forall (a_k, a_l) \in A_k \times A_l, k \neq l, \exists (o_i, o_j) \in O_i \times O_j$  such that  $\text{discrimination}(o_i, o_j, a_k, a_l)$  (resp.  $\text{similarity}(o_i, o_j, a_k, a_l)$ ).*

*Proof.* Since objects and attribute have a completely symmetrical role, it is sufficient to prove the proposition for object pairs. It proceeds easily by enumerating the possible type pairs with different elements. If objects have type 1 and 2 or 3 and 4, attributes allowing object pair discrimination have type  $b$  and  $c$  and attributes allowing object pair similarity have type  $a$  and  $d$ . If objects have type 1 and 3 or 2 and 4, attributes allowing object pair discrimination have type  $a$  and  $d$  and attributes allowing object pair similarity have type  $b$  and  $c$ . If objects have type 1 and 4 and if  $t_1 \in T_1 = \{a, b\}$  and  $t_2 \in T_2 = \{c, d\}$ , attributes allowing object pair discrimination have type  $t_1$  and  $t_2$  and attributes allowing object pair similarity have different types both in  $T_1$  or both in  $T_2$ . If objects have type 2 and 3 and if  $t_1 \in T_1 = \{a, c\}$  and  $t_2 \in T_2 = \{b, d\}$ , attributes allowing object pair discrimination have type  $t_1$  and  $t_2$  and attributes allowing object pair similarity have different types both in  $T_1$  or both in  $T_2$ .  $\square$

In case of incomplete complexes, some of these properties are no more relevant and a degenerate behaviour may appear: some of the sets may be identical. This fact allows to establish a new proposition on complete complexes:

**Proposition 2.** *In a complete analogical complex, side-by-side intersections of sets are empty.*

*Proof.* This property holds since when the intersection of two object (resp. attribute) sets in an analogical complex  $AC$  is not empty, then  $AC$  contains at least two empty attribute (resp. object) sets. This fact is a consequence of property 1.

Indeed, if an object belongs to two different types, their profiles must be the

same. The discrimination property ensures that the profile of two different object types differ by at least two different attribute with different types (e.g. if the object has type 1 and 3, attributes of type  $b$  and  $c$  should have different values). Thus it cannot exist attributes of the discriminant type (e.g. attributes of type  $b$  and  $c$  in the previous case) and the corresponding sets are empty. This completes the proof.

The converse of the proposition is not true: if all side-by-side intersections of sets differ, the complex is not necessary complete. For instance, consider the following context:

	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$
$o_1$	0	0	0	1	1	1
$o_2$	0	1	1	1	1	1
$o_3$	1	0	0	0	0	1
$o_4$	1	1	1	0	0	0
$o_5$	1	0	1	1	0	0

It contains the following not complete complex:

$$(\{o_1\}, \{o_2\}, \{o_3\}, \{o_4\}), (\{a_1\}, \{a_2, a_3\}, \emptyset, \{a_4, a_5\})$$

## 4 The lattice of analogical complexes

**Definition 5 (Partial Order on analogical complexes).** *Given two analogical complexes  $\mathcal{C}^1 = (O_{1,4}^1, A_{1,4}^1)$  and  $\mathcal{C}^2 = (O_{1,4}^2, A_{1,4}^2)$ , the partial order  $\leq$  is defined by*

$$\mathcal{C}^1 \leq \mathcal{C}^2 \quad \text{iff} \quad (O_i^1 \subseteq O_i^2 \text{ for } i = 1, 4 \quad \text{and} \quad A_i^2 \subseteq A_i^1 \text{ for } i = 1, 4).$$

$\mathcal{C}^1$  is called a sub-complex of  $\mathcal{C}^2$  and  $\mathcal{C}^2$  is called a super-complex of  $\mathcal{C}^1$

As for formal concepts, the set of all complexes has a lattice structure. Let us first define a derivation operator on analogical quadruplets:

**Definition 6 (Derivation on set quadruplets).**

*Let  $O = O_1 \cup O_2 \cup O_3 \cup O_4$  be a set of objects partitioned in four subsets, and  $A$  be a set of attributes. For all  $i$  and  $j \in [1, 4]$ , one defines  $O_i^{\prime j} = \{a \in A \mid \forall o \in O_i \ I(o, a) \Leftrightarrow AS(i, j)\}$*

*Let  $A = A_1 \cup A_2 \cup A_3 \cup A_4$  be a set of attributes partitioned in four subsets, and  $O$  be a set of objects. For all  $i$  and  $j \in [1, 4]$ , one defines  $A_i^{\prime j} = \{o \in O \mid \forall a \in A_i \ I(o, a) \Leftrightarrow AS(i, j)\}$*

*Finally, we define the derivation on quadruplets as follows:*

$$O'_{1,4} = \left( \prod_{j=1}^4 O_j^{\prime 1}, \prod_{j=1}^4 O_j^{\prime 2}, \prod_{j=1}^4 O_j^{\prime 3}, \prod_{j=1}^4 O_j^{\prime 4} \right)$$

$$A'_{1,4} = \left( \prod_{j=1}^4 A_j^{\prime 1}, \prod_{j=1}^4 A_j^{\prime 2}, \prod_{j=1}^4 A_j^{\prime 3}, \prod_{j=1}^4 A_j^{\prime 4} \right)$$



*Example 3.* Consider  $O = (\{12\}, \{28\}, \{59\}, \{49\})$ . One has:  $O_1^1 = \{a \in A \mid \neg I(12, a)\} = \{2, 5, 7, 8, 9\}$ ;  
 $O_2^1 = \{a \in A \mid \neg I(28, a)\} = \{3, 4, 7, 8\}$ ;  
 $O_3^1 = \{a \in A \mid I(59, a)\} = \{3, 4, 7, 8\}$ ;  
 $O_4^1 = \{a \in A \mid I(49, a)\} = \{2, 4, 5, 7, 8\}$  Then  $O'_1 = \bigcap_{j=1}^4 O_j^1 = \{7, 8\}$   
 Finally,  $O' = (\{7, 8\}, \{2, 5, 9\}, \{3, 4\}, \{6\})$ .

We exhibit a basic theorem for these complexes that naturally extends the basic theorem on concepts:

**Proposition 3.** *Given two analogical complexes  $\mathcal{C}^1 = (O_{1,4}^1, A_{1,4}^1)$  and  $\mathcal{C}^2 = (O_{1,4}^2, A_{1,4}^2)$ ,*

– *The join of  $\mathcal{C}^1$  and  $\mathcal{C}^2$  is defined by  $\mathcal{C}^1 \wedge \mathcal{C}^2 = (\mathcal{O}_{1,4}, \mathcal{A}_{1,4})$  where*

$$\forall i \in [1, 4] \mathcal{O}_i = O_i(\mathcal{C}^1) \cap O_i(\mathcal{C}^2)$$

$$\mathcal{A}_{1,4} = (A_1(\mathcal{C}^1) \cup A_1(\mathcal{C}^2), A_2(\mathcal{C}^1) \cup A_2(\mathcal{C}^2), A_3(\mathcal{C}^1) \cup A_3(\mathcal{C}^2), A_4(\mathcal{C}^1) \cup A_4(\mathcal{C}^2))''$$

– *The meet of  $\mathcal{C}^1$  and  $\mathcal{C}^2$  is defined by  $\mathcal{C}^1 \vee \mathcal{C}^2 = (\mathcal{O}_{1,4}, \mathcal{A}_{1,4})$  where*

$$\mathcal{O}_{1,4} = (O_1(\mathcal{C}^1) \cup O_1(\mathcal{C}^2), O_2(\mathcal{C}^1) \cup O_2(\mathcal{C}^2), O_3(\mathcal{C}^1) \cup O_3(\mathcal{C}^2), O_4(\mathcal{C}^1) \cup O_4(\mathcal{C}^2))''$$

*Proof.* The meet and the join are dual and one only needs to prove the proposition for the join. The ordering by set inclusion requires the set of objects  $\mathcal{O}_i$  of  $\mathcal{C}^1 \wedge \mathcal{C}^2$  to be included in  $O_i(\mathcal{C}^1) \cap O_i(\mathcal{C}^2)$  and its set of attributes  $\mathcal{A}_j$  to be included in  $A_j(\mathcal{C}^1) \cup A_j(\mathcal{C}^2)$ . Taking exactly the intersection of objects thus ensures the set of objects to be maximal. The corresponding maximal sets of attributes may be inferred using the derivation operator ' we have just defined. Another way to generate these sets is to apply the derivation operator twice on the union of sets of attributes.

*Example 4.* The complex lattice of smallzoo has 24 elements, including 18 complete complexes. It is sketched in figure 3.

In this lattice, for example, the join of the analogical complex numbered 9 and 12, which are as follows

$$9 = ((\{12\}, \{28\}, \{59, 64\}, \{20, 49\}), (\{7\}, \{9\}, \{4\}, \{6\}))$$

$$12 = ((\{12, 17\}, \{28\}, \{59\}, \{49, 64\}), (\{7\}, \{2, 5\}, \{3\}, \{6\}))$$

is number 15, namely:

$$15 = ((\{12\}, \{28\}, \{59\}, \{49\}), (\{7, 8\}, \{2, 5, 9\}, \{3, 4\}, \{6\}))$$

The resulting object sets are for each type the intersection of the two joined object sets. The resulting attribute sets contain for each type the union of the two joined attribute sets and may contain other elements with a correct profile on all objects. For instance,  $A_1(9 \wedge 12) = \{7, 8\}$  is made of the union of  $A_1(9)$

and  $A_1(12)$  ( $\{7\}$ ) plus attribute 8 since 8 has the right profile  $(0, 0, 1, 1)$  on  $O_{1,4}$  (that is,  $\neg I(12, 8)$ ,  $\neg I(28, 8)$ ,  $I(59, 8)$  and  $I(49, 8)$ ).

The meet of the analogical complexes numbered 9 and 12 is number 19, namely

$$19 = ((\{12, 17, 28\}, \{12, 17, 28\}, \{20, 49, 59, 64\}, \{20, 49, 59, 64\}), (\{7\}, \emptyset, \emptyset, \{6\}))$$

## 5 Conclusion

We have introduced a new conceptual object called analogical complex that uses a complex relation, analogical proportion, to compare objects with respect to their attribute values. Although this relation works on set quadruplets instead of simple sets like in formal concepts, we have shown that it is possible to keep the main properties of concepts, that is, maximality and comparison at the level of object or attribute pairs. The set of all complexes are structured within a lattice that contains two types of elements. The most meaningful ones only contain non empty sets and are a strong support for doing analogical inference. An interesting extension of this work would be to develop this inference process for analogical data mining in a way close to rule generation in FCA.

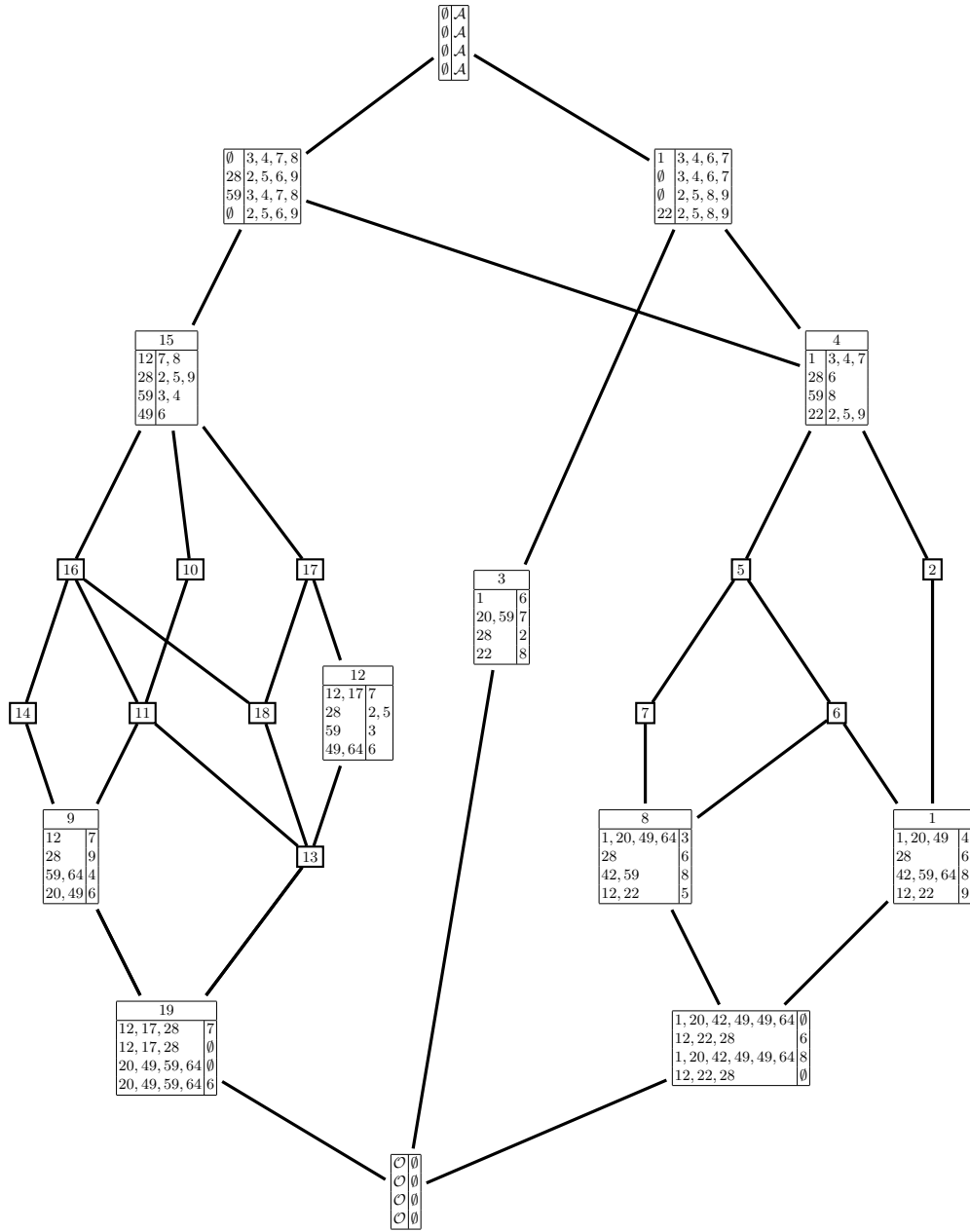
The degenerate case where some of the sets are empty is more frequent than in FCA where their presence is limited to the top or bottom of the lattice. The presence of a single empty set may reflect the lack of some object or attribute and is thus a possible new research direction for completing a knowledge base or an ontology. Particularly, analogy in a Boolean framework introduces a form of negation through the search of dissimilarities (discrimination) between objects.

We have written an implementation the search for complete analogical complexes, using the Answer Set Programming framework [1]. The properties of definition 3 are translated straightforwardly in logical constraints and the search of all complexes is achieved by an ASP solver looking for all solutions. The description of the ASP program would be beyond the scope of this paper but it can be seen as a relatively simple exercise of extension of the search for formal concepts by adding a few logical constraints. It is likely that most of the existing tools of FCA could be adapted the same way for analogical complex analysis. This would allow to include both categorization and analogy within common data mining environments.

## References

1. Brewka, G., Eiter, T., Truszczyński, M.: Answer set programming at a glance. *Commun. ACM* 54(12), 92–103 (Dec 2011), <http://doi.acm.org/10.1145/2043174.2043195>
2. Correa, W., Prade, H., Richard, G.: When intelligence is just a matter of copying. In: et al., L.D.R. (ed.) *Proc. 20th Europ. Conf. on Artificial Intelligence*, Montpellier, Aug. 27-31. pp. 276–281. IOS Press (2012)

3. French, R.M.: The computational modeling of analogy-making. *Trends in Cognitive Sciences* 6(5), 200 – 205 (2002)
4. Gentner, D., Holyoak, K.J., Kokinov, B.N.: *The Analogical Mind: Perspectives from Cognitive Science. Cognitive Science, and Philosophy*, MIT Press, Cambridge, MA (2001)
5. Hofstadter, D., Mitchell, M.: The Copycat project: A model of mental fluidity and analogy-making. In: Hofstadter, D., *The Fluid Analogies Research Group* (eds.) *Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*. pp. 205–267. Basic Books, Inc., New York, NY (1995)
6. Lepage, Y.: Analogy and formal languages. *Electr. Notes Theor. Comput. Sci.* 53 (2001)
7. Lepage, Y.: Analogy and formal languages. In: *Proc. FG/MOL 2001*. pp. 373–378 (2001), (see also <http://www.slt.atr.co.jp/lepage/pdf/dhdryl.pdf.gz>)
8. Lichman, M.: UCI machine learning repository (2013), <http://archive.ics.uci.edu/ml>
9. Melis, E., Veloso, M.: Analogy in problem solving. In: *Handbook of Practical Reasoning: Computational and Theoretical Aspects*. Oxford Univ. Press (1998)
10. Miclet, L., Bayouhd, S., Delhay, A.: Analogical dissimilarity: definition, algorithms and two experiments in machine learning. *JAIR*, 32 pp. 793–824 (2008)
11. Miclet, L., Prade, H.: Handling analogical proportions in classical logic and fuzzy logics settings. *Proc. 10th Eur. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'09)*, Verona (2009)
12. Miclet, L., Prade, H.: Handling analogical proportions in classical logic and fuzzy logics settings. In: *Proc. 10th Eur. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'09)*, Verona. pp. 638–650. Springer, LNCS 5590 (2009)
13. Miclet, L., Barbot, N., Prade, H.: From analogical proportions in lattices to proportional analogies in formal concepts. In: *ECAI - 21th European Conference on Artificial Intelligence*. Prague, Czech Republic (Aug 2014)
14. Miclet, L., Prade, H., Guennec, D.: Looking for Analogical Proportions in a Formal Concept Analysis Setting. In: *Amedeo Napoli, V.V. (ed.) Conference on Concept Lattices and Their Applications*. pp. 295–307. Nancy, France (Oct 2011)
15. Prade, H., Richard, G.: Homogeneous logical proportions: Their uniqueness and their role in similarity-based prediction. *Proc. of the 13th International Conference on Principles of Knowledge Representation and Reasoning KR2012* pp. 402 – 412 (2012)
16. Radvansky, M., Sklenar, V.: Fca extension for ms excel 2007, <http://www.fca.radvansky.net> (2010)
17. Stroppa, N., Yvon, F.: An analogical learner for morphological analysis. In: *Online Proc. 9th Conf. Comput. Natural Language Learning (CoNLL-2005)*. pp. 120–127 (2005)
18. Stroppa, N., Yvon, F.: Analogical learning and formal proportions: Definitions and methodological issues. *ENST Paris report* (2005)
19. Stroppa, N., Yvon, F.: Du quatrième de proportion comme principe inductif : une proposition et son application à l'apprentissage de la morphologie. *Traitement Automatique des Langues* 47(2), 1–27 (2006)
20. Yevtushenko, S.: System of data analysis "concept explorer". (in russian). In: *Proc. of the 7th national conference on artificial intelligence (KII-2000)*, Russia. pp. 127–134 (2000)



**Fig. 3.** Hasse diagram of the analogical complex lattice for formal context smallzoo. For reasons of space some nodes are not explicitly given.