Relational proportions between objects and attributes

Nelly Barbot[1], Laurent Miclet[1], and Henri Prade[2]

[1] Univ. Rennes, CNRS, IRISA, F-22305 Lannion, France,
nbarbot@irisa.fr, laurent.miclet@gmail.com,
[2] IRIT, CNRS & Univ. P. Sabatier, 31062 Toulouse Cedex 9, France,
prade@irit.fr

Abstract. Analogical proportions are statements of the form "A is to B as C is to D", where A, B, C, D are items of the same nature, or not. In this paper, we more particularly consider "relational proportions" of the form "object A has the same relationship with attribute a as object B with attribute b". We provide a formal definition for relational proportions, and investigate how they can be extracted from a formal context, in the setting of formal concept analysis.

Keywords: Analogy, analogical reasoning, analogical proportion, analogy in lattices, formal concept analysis.

1 Introduction

A statement such as "Carlsen is to chess as Mozart is to music" introduces Carlsen as a precocious virtuoso of chess, a quality that Mozart is well known to have concerning music. It relates two types of items, here people and activities. It is an example of what we call *relational proportions* which are statements of the form "object A has the same relationship with attribute a as object B with attribute b". This can be viewed as a special case of analogical proportions which are statements of the form "A is to B as C is to D". In the case where A, B, C, D are items which can be represented in terms of the same set of features, a formal definition has been proposed for analogical proportions in the setting of Boolean logic and then extended using multiple-valued logic for handling numerical features [2, 9], by stating that "A differs from B as C differs from D and B differs from A as D differs from C".

The nature of relational proportions suggests to handle them in the setting of formal concept analysis. This leads us to the question of defining analogical proportions between formal concepts. The paper first recalls the definition of analogical proportions in non distributive lattices, as already presented in [4]. Then it brings original material, firstly by studying the links between analogical proportions between formal concepts and analogical proportions between objects or attributes. It also shows how relational proportions can be obtained in a formal context from the identification of an analogical complex.

2 Analogical proportions: basics and formalization

Analogical proportions are usually characterized by three axioms. The first two axioms acknowledge the symmetrical role played by the pairs (x, y) and (z, t) in the proportion 'x is to y as z is to t', and enforce the idea that y and z can be interchanged if the proportion is valid, just as in the equality of two numerical ratios where means can be exchanged. This view dates back to Aristotle. A third (optional) axiom, called determinism, insists on the uniqueness of the solution $t = y$ for completing the analogical proportion in t: $(x : y :: x : t)$. These axioms are studied in [1].

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

1. Reflexivity of 'as': $(x : y :: x : y)$
2. Symmetry of 'as': $(x : y :: z : t) \Leftrightarrow (z : t :: x : y)$
3. Exchange of means: $(x : y :: z : t) \Leftrightarrow (x : z :: y : t)$

Then, thanks to symmetry, it can be easily seen that $(x : y :: z : t) \Leftrightarrow (t : y :: z : x)$ should also hold (exchange of the extremes). According to the first two axioms, four other formulations are equivalent to the canonical form $(x : y :: z : t)$. Finally, the eight equivalent forms of an analogical proportion are: $(x : y :: z : t)$, $(z : t :: x : y)$, $(y : x :: t : z)$, $(t : z :: y : x)$, $(z : x :: t : y)$, $(t : y :: z : x)$, $(x : z :: y : t)$ and $(y : t :: x : z)$.

With respect to this axiomatic definition of AP, Stroppa and Yvon [3] have given another definition, based on the notion of factorization when the set of objects is a commutative semigroups. From these previous works, Miclet *et al.* [4] have derived the following definitions in the lattice framework.

Definition 2. *A 4-tuple (x, y, z, t) of a lattice $(L, \vee, \wedge, \leq)^4$ is a Factorial Analogical Proportion (FAP) $(x : y :: z : t)$ iff:*

$$\begin{array}{ll} x = (x \wedge y) \vee (x \wedge z) & \quad x = (x \vee y) \wedge (x \vee z) \\ y = (x \wedge y) \vee (y \wedge t) & \quad y = (x \vee y) \wedge (y \vee t) \\ z = (z \wedge t) \vee (x \wedge z) & \quad z = (z \vee t) \wedge (x \vee z) \\ t = (z \wedge t) \vee (y \wedge t) & \quad t = (z \vee t) \wedge (y \vee t) \end{array}$$

Definition 3. *A 4-tuple (x, y, z, t) of $(L, \vee, \wedge, \leq)^4$ is a Weak Analogical Proportion (WAP) when $x \wedge t = y \wedge z$ and $x \vee t = y \vee z$. It is denoted $x : y \; WAP \; z : t$.*

In the case of a distributive lattice (e.g. a Boolean lattice), this alternative definition is equivalent to the FAP. But, in general, a FAP is a WAP and the converse is false, which explains the use of adjective "weak" [4].

Example 1. Let us consider a finite set Σ and the associated Boolean lattice $(2^{\Sigma}, \cup, \cap, \leq)$. When saying of subsets x, y, z, t of Σ that "x is to y as z is to t", we express that x differs from y in the same way as z differs from t. For example, if $x = \{a, b, e\}$ and $y = \{b, c, e\}$, we see that to transform x into y, we have to remove a and add c. Now, if $z = \{a, d, e\}$, we can construct t with the same operations, to obtain $t = \{c, d, e\}$. In more formal terms, with this definition, the following properties are asked to x, y, z and t (with $x \setminus y = x \cap \neg y$): $x \setminus y = z \setminus t$ and $y \setminus x = t \setminus z$, which are equivalent to $x \cap t = y \cap z$ and $x \cup t = y \cup z$. These relations linking x, y, z, t are clearly symmetrical, and satisfy the exchange of the means. Hence it is a correct definition of the AP in the Boolean setting [2].

The next proposition gives a simple example of FAP in a lattice.

Proposition 1. *Let y and z be two elements of a lattice, the proportion $y : y \vee z :: y \wedge z : z$ is a FAP. We call it a Canonical Analogical Proportion (CAP).*

Proof. The first equality of Definition 2, namely $y = (y \wedge (y \vee z)) \vee (y \wedge (y \wedge z))$, is true since the right member is equal to $(y) \vee (y \wedge z) = y$. The verification of the three other equalities of Definition 2 is similar, using the absorption laws.

3 Analogical proportions in FCA

In order to derive more specifically the AP notion in a Formal Concept Analysis framework (FCA), we first recall some basic elements of FCA, before studying the relations between several kinds of AP and their characterization in FCA.

3.1 Formal concept analysis

FCA starts with a binary relation R defined between a set $\mathcal{O}$ of objects and a set $\mathcal{A}$ of attributes. The tuple $(\mathcal{O}, \mathcal{A}, R)$ is called a *formal context*. The notation $(o, a) \in R$ or oRa means that object o has attribute a. We denote $o^\uparrow = \{a \in \mathcal{A} | (o, a) \in R\}$ the attribute set of object o and $a^\downarrow = \{o \in \mathcal{O} | (o, a) \in R\}$ the object set having attribute a. Similarly, for any subset $\mathbf{o}$ of objects, $\mathbf{o}^\uparrow$ is defined as $\{a \in \mathcal{A} | a^\downarrow \supseteq \mathbf{o}\}$. Then a *formal concept* is defined as a pair $(\mathbf{o}, \mathbf{a})$, such that $\mathbf{a}^\downarrow = \mathbf{o}$ and $\mathbf{o}^\uparrow = \mathbf{a}$. One calls $\mathbf{o}$ the *extension* of the concept and $\mathbf{a}$ its *intension*.

The set of all formal concepts is equipped with a partial order (denoted $\leq$) defined as: $(\mathbf{o}_1, \mathbf{a}_1) \leq (\mathbf{o}_2, \mathbf{a}_2)$ iff $\mathbf{o}_1 \subseteq \mathbf{o}_2$ (or, equivalently, $\mathbf{a}_2 \subseteq \mathbf{a}_1$). Then it is structured as a lattice, called the *concept lattice* of R.

Example 2. The concept lattice of the following context R is shown in Figure 1.

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
o_1		×	×	×			×	×	×
o_2	×		×	×	×		×		×
o_3	×	×		×		×		×	×
o_4	×	×	×		×	×			×
o_5	×	×	×	×	×	×	×	×	×

The following preliminaries are simple consequences of the definition of concept lattice and the Main Theorem of Formal Concepts [5, 6]. They allow for a quick demonstration of propositions in the next section.

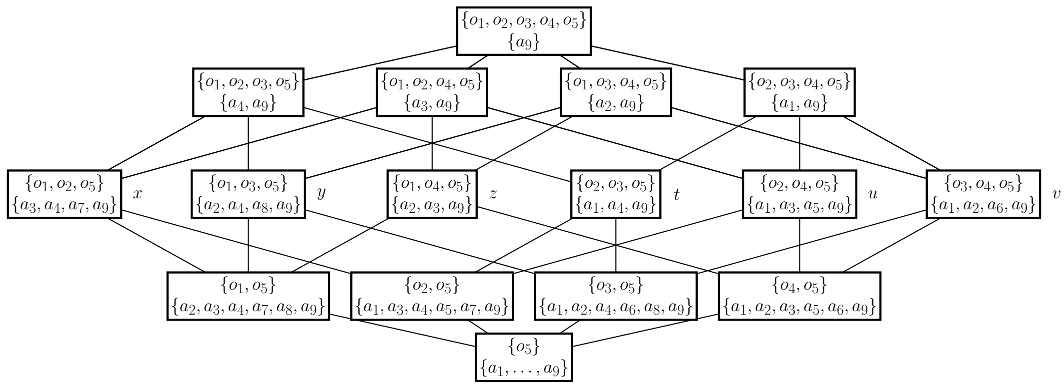

Fig. 1. The formal concept lattice of R (it is a Boolean lattice).

Preliminary 1 *Given two concepts* $x = (\mathbf{o}_x, \mathbf{a}_x)$ *and* $y = (\mathbf{o}_y, \mathbf{a}_y)$*, one has* $(\mathbf{o}_x \cup \mathbf{o}_y)^{\uparrow} = \mathbf{a}_x \cap \mathbf{a}_y$ *and* $(\mathbf{a}_x \cup \mathbf{a}_y)^{\downarrow} = \mathbf{o}_x \cap \mathbf{o}_y$*.*

Preliminary 2 *Given two concepts* $x = (\mathbf{o}_x, \mathbf{a}_x)$ *and* $y = (\mathbf{o}_y, \mathbf{a}_y)$*, one has* $\mathbf{o}_x \cup \mathbf{o}_y \subseteq \mathbf{o}_{x \vee y}$*,* $\mathbf{o}_x \cap \mathbf{o}_y = \mathbf{o}_{x \wedge y}$*,* $\mathbf{a}_x \cup \mathbf{a}_y \subseteq \mathbf{a}_{x \wedge y}$ *and* $\mathbf{a}_x \cap \mathbf{a}_y = \mathbf{a}_{x \vee y}$*.*

Preliminary 3 *Let* $\mathbf{o}$ *(resp.* $\mathbf{a}$*) be a subset of* $\mathcal{O}$ *(resp.* $\mathcal{A}$*), there exists at most one concept* x *such that* $\mathbf{o}_x = \mathbf{o}$ *(resp.* $\mathbf{a}_x = \mathbf{a}$*).*

3.2 Weak and strong analogical proportions in FCA

Since concepts are associated to a set of attributes and objects, the main objectives of this section are to relate the AP definitions with these sets and to study the links the AP on concept lattice and AP on object or attribute sets.

Proposition 2. *Let* x*,* y*,* z *and* t *be four concepts, one has:*
($x \vee t = y \vee z$ *iff* $\mathbf{a}_x \cap \mathbf{a}_t = \mathbf{a}_y \cap \mathbf{a}_z$*) and (*$x \wedge t = y \wedge z$ *iff* $\mathbf{o}_x \cap \mathbf{o}_t = \mathbf{o}_y \cap \mathbf{o}_z$*).*
As consequence, ($x : y \; WAP \; z : t$*) iff* $\mathbf{a}_x \cap \mathbf{a}_t = \mathbf{a}_y \cap \mathbf{a}_z$ *and* $\mathbf{o}_x \cap \mathbf{o}_t = \mathbf{o}_y \cap \mathbf{o}_z$*.*

Proof. From Preliminary 2, $x \vee t = y \vee z$ implies $\mathbf{a}_x \cap \mathbf{a}_t = \mathbf{a}_y \cap \mathbf{a}_z$ and conversely, $\mathbf{a}_x \cap \mathbf{a}_t = \mathbf{a}_y \cap \mathbf{a}_z$ implies $\mathbf{a}_{x \vee t} = \mathbf{a}_{y \vee z}$. Thus, $x \vee t = y \vee z$ using Preliminary 3. The proof of the second equivalence can be done in a similar manner.

Proposition 3. *Let* x*,* y*,* z *and* t *be four concepts, if (*$\mathbf{a}_x : \mathbf{a}_y :: \mathbf{a}_z : \mathbf{a}_t$ *or* $\mathbf{o}_x : \mathbf{o}_y :: \mathbf{o}_z : \mathbf{o}_t$*) then* $x : y \; WAP \; z : t$*.*

Proof. Let x, y, z and t be four concepts such that $\mathbf{a}_x : \mathbf{a}_y :: \mathbf{a}_z : \mathbf{a}_t$, or equivalently $\mathbf{a}_x \cap \mathbf{a}_t = \mathbf{a}_y \cap \mathbf{a}_z$ and $\mathbf{a}_x \cup \mathbf{a}_t = \mathbf{a}_y \cup \mathbf{a}_z$ (the APs between the subsets of attributes correspond to FAPs in the Boolean lattice of $(2^{\mathcal{A}}, \cup, \cap, \subseteq)$). Thanks to Proposition 2, $\mathbf{a}_x \cap \mathbf{a}_t = \mathbf{a}_y \cap \mathbf{a}_z$ is equivalent to $x \vee t = y \vee z$. Moreover, using Preliminary 1, we have $(\mathbf{a}_x \cup \mathbf{a}_t)^{\downarrow} = \mathbf{o}_x \cap \mathbf{o}_t$ and $(\mathbf{a}_y \cup \mathbf{a}_z)^{\downarrow} = \mathbf{o}_y \cap \mathbf{o}_z$. Thus, $\mathbf{a}_x \cup \mathbf{a}_t = \mathbf{a}_y \cup \mathbf{a}_z$ implies $\mathbf{o}_x \cap \mathbf{o}_t = \mathbf{o}_y \cap \mathbf{o}_z$. In the case where $\mathbf{o}_x : \mathbf{o}_y :: \mathbf{o}_z : \mathbf{o}_t$, the proof is similar since we also have $(\mathbf{o}_x \cup \mathbf{o}_t)^{\uparrow} = \mathbf{a}_x \cap \mathbf{a}_t$ and $(\mathbf{o}_y \cup \mathbf{o}_y)^{\uparrow} = \mathbf{a}_y \cap \mathbf{a}_z$.

Comments. The converse is false. Let us consider the following formal context

	a_1	a_2	a_3	a_4	a_5
o_1			×	×	
o_2	×		×		
o_3		×		×	
o_4	×	×			×

its concept lattice is displayed on Figure 2. Concepts $x = (\{o_1\}, \{a_3, a_4\})$, $y = (\{o_2\}, \{a_1, a_3\})$, $z = (\{o_3\}, \{a_2, a_4\})$ and $t = (\{o_4\}, \{a_1, a_2, a_5\})$ are in WAP, due to Proposition 2. However, the Boolean APs $\mathbf{a}_x : \mathbf{a}_y :: \mathbf{a}_z : \mathbf{a}_t$ and $\mathbf{o}_x : \mathbf{o}_y :: \mathbf{o}_z : \mathbf{o}_t$ are both false. The WAP between concepts is less restrictive than the AP between sets of attributes: in a WAP, objects are allowed to possess attributes which are not shared by any other object concerned in the WAP.

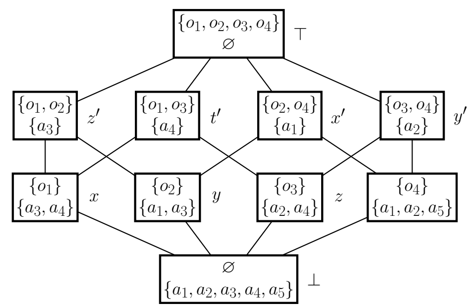

Fig. 2. In this lattice, x, y, z and t are in WAP but $\mathbf{a}_x : \mathbf{a}_y :: \mathbf{a}_z : \mathbf{a}_t$ and $\mathbf{o}_x : \mathbf{o}_y :: \mathbf{o}_z : \mathbf{o}_t$ are both false. Besides, x', y', z' and t' are in WAP and $\mathbf{o}_{x'} : \mathbf{o}_{y'} :: \mathbf{o}_{z'} : \mathbf{o}_{t'}$ is true, but $\mathbf{a}_{x'} : \mathbf{a}_{y'} :: \mathbf{a}_{z'} : \mathbf{a}_{t'}$ and the FAP $x' : y' :: z' : t'$ are both false.

We give now a proposition which leads us to a corollary in which is defined yet another analogical proportion between formal concepts, the strongest of all.

Proposition 4. *Let x, y, z and t be four concepts, if ($\mathbf{a}_x \cup \mathbf{a}_t = \mathbf{a}_y \cup \mathbf{a}_z$ and $\mathbf{o}_x \cup \mathbf{o}_t = \mathbf{o}_y \cup \mathbf{o}_z$) then the FAP $x : y :: z : t$ holds.*

Proof. $\mathbf{a}_x \cup \mathbf{a}_t = \mathbf{a}_y \cup \mathbf{a}_z$ implies that $\mathbf{a}_x = (\mathbf{a}_x \cap \mathbf{a}_y) \cup (\mathbf{a}_x \cap \mathbf{a}_z)$. It results that, using Preliminaries 1 and 2, $\mathbf{o}_x = (\mathbf{a}_x)^{\downarrow} = (\mathbf{a}_x \cap \mathbf{a}_y)^{\downarrow} \cap (\mathbf{a}_x \cap \mathbf{a}_z)^{\downarrow}$. Then,

$$\mathbf{o}_x = (\mathbf{a}_{x\vee y})^{\downarrow} \cap (\mathbf{a}_{x\vee z})^{\downarrow} = \mathbf{o}_{x\vee y} \cap \mathbf{o}_{x\vee z} = \mathbf{o}_{(x\vee y)\wedge(x\vee z)}$$

and Preliminary 3 permits to obtain $x = (x \vee y) \wedge (x \vee z)$.

In a same way, from $\mathbf{o}_x \cup \mathbf{o}_t = \mathbf{o}_y \cup \mathbf{o}_z$, we get that $\mathbf{o}_x = (\mathbf{o}_x \cap \mathbf{o}_y) \cup (\mathbf{o}_x \cap \mathbf{o}_z)$ and $\mathbf{a}_x = (\mathbf{o}_x)^{\uparrow} = (\mathbf{o}_x \cap \mathbf{o}_y)^{\uparrow} \cap (\mathbf{o}_x \cap \mathbf{o}_z)^{\uparrow}$. Then,

$$\mathbf{a}_x = (\mathbf{o}_{x\wedge y})^{\uparrow} \cap (\mathbf{o}_{x\wedge z})^{\uparrow} = \mathbf{a}_{x\wedge y} \cap \mathbf{a}_{x\wedge z} = \mathbf{a}_{(x\wedge y)\vee(x\wedge z)}.$$

Thus, $x = (x\wedge y)\vee(x\wedge z)$. All the equalities of Definition 2 are similarly checked.

Corollary 1. *Let x, y, z and t be four concepts, the following two conjunctions are equivalent:*

$$\mathbf{a}_x \cup \mathbf{a}_t = \mathbf{a}_y \cup \mathbf{a}_z \textit{ and } \mathbf{o}_x \cup \mathbf{o}_t = \mathbf{o}_y \cup \mathbf{o}_z$$
$$\mathbf{a}_x : \mathbf{a}_y :: \mathbf{a}_z : \mathbf{a}_t \textit{ and } \mathbf{o}_x : \mathbf{o}_y :: \mathbf{o}_z : \mathbf{o}_t$$

This characterizes a particular case of FAP between concepts that we call a Strong Analogical Proportion (SAP). It is denoted $x : y \; SAP \; z : t$. In other words, four concepts in analogical proportion on attributes and on objects are said to be in strong analogical proportion.

Proof. Let x, y, z and t be such that $\mathbf{a}_x \cup \mathbf{a}_t = \mathbf{a}_y \cup \mathbf{a}_z$ and $\mathbf{o}_x \cup \mathbf{o}_t = \mathbf{o}_y \cup \mathbf{o}_z$, Proposition 4 implies the FAP $x : y :: z : t$, and then $x : y \; WAP \; z : t$. Hence, using Proposition 2, we have $\mathbf{a}_x \cap \mathbf{a}_t = \mathbf{a}_y \cap \mathbf{a}_z$ and $\mathbf{o}_x \cap \mathbf{o}_t = \mathbf{o}_y \cap \mathbf{o}_z$. Consequently, $\mathbf{a}_x : \mathbf{a}_y :: \mathbf{a}_z : \mathbf{a}_t$ and $\mathbf{o}_x : \mathbf{o}_y :: \mathbf{o}_z : \mathbf{o}_t$. The converse is trivial.

Comments. From Corollary 1, $\mathbf{a}_x : \mathbf{a}_y :: \mathbf{a}_z : \mathbf{a}_t$ and $\mathbf{o}_x : \mathbf{o}_y :: \mathbf{o}_z : \mathbf{o}_t$ imply the FAP $x : y :: z : t$. However, the reciprocal is false. Let us consider the concept lattice displayed in Figure 2: we have the FAP $y : \top :: \bot : z$ (which is a CAP) but $\mathbf{o}_y \cup \mathbf{o}_z \neq \mathbf{o}_\top \cup \mathbf{o}_\bot$ and $\mathbf{a}_y \cup \mathbf{a}_z \neq \mathbf{a}_\top \cup \mathbf{a}_\bot$.

Example 3. In the Boolean lattice displayed in Figure 1, concepts x, y, u and v form a FAP but are not in SAP. Indeed, $\mathbf{a}_x : \mathbf{a}_y :: \mathbf{a}_u : \mathbf{a}_v$ does not hold. However, without changing the lattice, the formal context can be reduced to

	a_1	a_2	a_3	a_4
o_1		×	×	×
o_2	×		×	×
o_3	×	×		×
o_4	×	×	×	

and the reduced representation of x, y, u and v gives ($x : y \; SAP \; u : v$):

	a_1	a_2	a_3	a_4
$\mathbf{a}_x$			×	×
$\mathbf{a}_y$		×		×
$\mathbf{a}_u$	×		×	
$\mathbf{a}_v$	×	×		

	o_1	o_2	o_3	o_4
$\mathbf{o}_x$	×	×		
$\mathbf{o}_y$	×		×	
$\mathbf{o}_u$		×		×
$\mathbf{o}_v$			×	×

These observations stem from the fact that the FAP and WAP between concepts are directly related to the lattice whereas the Boolean AP between object or attribute sets directly depends on the formal context.

4 Formal concepts and relational proportion

4.1 From a RP to concepts in AP

In this section, we study if we can deduce from a relational proportion "A is the B of a", or "A is to a as B is to b", formal concepts in WAP and an analogical complex from this knowledge.

As an example, we have found in a web magazine[3] the following proportion "Massimiliano Alajmo is the Mozart of Italian cooking". The background knowledge allowing to understand this relational proportion is the following: music and Italian cooking are disciplines practiced by humans, such disciplines can be practiced with different levels of ability, Mozart is a musician and Mozart is a genius in music discipline. Since the quality "to be a genius" is not possessed by everybody, there must exist many "ordinary gifted" musicians. Then, the background knowledge can be expressed by the following formal context:

	a_1	a_2	a_3
o_1	×	×	
o_2	×		×

where o_1 stands for Mozart, o_2 for one of "ordinary gifted" musicians, a_1 is the attribute "practices music", a_2 "is a genius" and a_3 "has an ordinary ability".

Now, when the new data "Alajmo is the Mozart of Italian cooking" is introduced, the knowledge extends as follows: Alajmo practices Italian cooking, and he has something in common with Mozart that is not Italian cooking. The relational proportion is a reduced form of "Alajmo is to Italian cooking as Mozart is to music". Since Mozart has only the other attribute "Genius", Alajmo must have it. Moreover, since cooking is a discipline practiced by humans, there must exist some ordinary gifted Italian cook. At last, we must introduce the notion of non-genius in our universe. If we do not, we implicitly suppose that everybody is a genius for some activity. The knowledge is now as follows

	a_1	a_2	a_3	a_4
o_1	×	×		
o_2	×		×	
o_3		×		×
a_4			×	×

where o_3 stands for Alajmo, o_4 an ordinary gifted Italian cook and a_4 Italian cooking. This context is called the *analogical context*. Considering the associated concept lattice, the closest analogical proportion to "Alajmo is the Mozart of Italian cooking" is $(\{o_3\}, \{a_2, a_4\}) : (\{o_4\}, \{a_3, a_4\})\ WAP\ (\{o_1\}, \{a_1, a_2\}) : (\{o_2\}, \{a_1, a_3\})$ which translates into "Mozart is to some ordinary musician as Alajmo is to some ordinary cook".

More formally, from the relational proportion "o_1 is the o_2 of a", we can derive an analogical context as above. It is composed of objects o_1 and o_2, described by four attributes: a is possessed by o_1 and not by o_2, $\tilde{a}$ is possessed by o_2 and not by o_1, b is possessed both by o_1 and o_2 and $\tilde{b}$ is some attribute not possessed by o_1 nor o_2. Secondly we complete the context with two objects o_3 and o_4 that are the complements of o_2 and o_1 with respect to the four attributes. The resulted context is the analogical context where $a_1 = b$, $a_2 = a$, $a_3 = \tilde{a}$ and $a_4 = \tilde{b}$.

[3] http://www.slate.fr/story/43841/massimiliano-alajmo

4.2 Analogical complex

In the previous paragraph, it turns out that the analogical context is an interesting pattern, from which we can extract relational proportion. A more general definition of this pattern, named *analogical complex*, has been given in [7].

An analogical complex is a subcontext of a formal context described by:

		×	×
	×		×
×		×	
×	×		

associated with the binary matrix $AS = \begin{pmatrix} 0\,0\,1\,1 \\ 0\,1\,0\,1 \\ 1\,0\,1\,0 \\ 1\,1\,0\,0 \end{pmatrix}$. Matrix AS exhibits characteristic pattern of a Boolean analogical proportion [2] and is called an *analogical schema*. We write $AS(i,j)$ if its value at row i and column j is 1.

Definition 4. *Given a formal context* $(\mathcal{O}, \mathcal{A}, R)$*, a set of objects* $\mathbf{o} \subseteq \mathcal{O}$*,* $\mathbf{o} = \mathbf{o}_1 \cup \mathbf{o}_2 \cup \mathbf{o}_3 \cup \mathbf{o}_4$*, a set of attributes* $\mathbf{a} \subseteq \mathcal{A}$*,* $\mathbf{a} = \mathbf{a}_1 \cup \mathbf{a}_2 \cup \mathbf{a}_3 \cup \mathbf{a}_4$*, and a binary relation* R*, the subcontext* $(\mathbf{o}, \mathbf{a})$ *forms an analogical complex* $(\mathbf{o}_{1,4}, \mathbf{a}_{1,4})$ *iff*

1. *the binary relation is compatible with the analogical schema AS:*
   $\forall i \in [1,4],\ \forall o \in \mathbf{o}_i,\ \forall j \in [1,4],\ \forall a \in \mathbf{a}_j,\ ((o,a) \in R) \Leftrightarrow AS(i,j).$
2. *The context is maximal with respect to the first property (*$\oplus$ *denotes the exclusive or and* $\setminus$ *the set-theoretic difference):*
   $\forall o \in \mathcal{O} \setminus \mathbf{o}, \forall i \in [1,4],\ \exists j \in [1,4], \exists a \in \mathbf{a}_j, ((o,a) \in R) \oplus AS(i,j).$
   $\forall a \in \mathcal{A} \setminus \mathbf{a}, \forall j \in [1,4],\ \exists i \in [1,4], \exists o \in \mathbf{o}_i, ((o,a) \in R) \oplus AS(i,j).$

An analogical complex is complete if none of sets $\mathbf{a}_1, \ldots, \mathbf{a}_4, \mathbf{o}_1, \ldots, \mathbf{o}_4$ *are empty.*

Comments.

1. In order to simplify the notations, the Cartesian products $\mathbf{o}_1 \times \ldots \times \mathbf{o}_4$ and $\mathbf{a}_1 \times \ldots \times \mathbf{a}_4$ are respectively denoted $\mathbf{o}_{1,4}$ and $\mathbf{a}_{1,4}$.
2. In [7], it has been shown that the set of the analogical complexes of any formal context is itself structured as a lattice.

Example 4. Let us consider a subcontext, called SmallZoo, extracted from the Zoo data base [8], it has been shown in [7] that 24 analogical complexes (18 complete ones) can be derived, like the following complete one:

SmallZoo	hair	feathers	eggs	milk	airborne	aquatic	predator	toothed
	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7
o_0 aardvark	×			×			×	×
o_1 chicken		×	×		×			
o_2 crow		×	×		×		×	
o_3 dolphin				×		×	×	×
o_4 duck		×	×		×	×		
o_5 fruitbat	×			×	×			×
o_6 kiwi		×	×				×	
o_7 mink	×			×		×	×	×
o_8 penguin		×	×			×	×	
o_9 platypus	×		×	×		×	×	

		$\mathbf{a}_1$	$\mathbf{a}_2$			$\mathbf{a}_3$		$\mathbf{a}_4$
		a_5	a_0	a_3	a_7	a_1	a_2	a_4
$\mathbf{o}_1$	o_1					×	×	×
	o_2					×	×	×
$\mathbf{o}_2$	o_5		×	×	×			×
$\mathbf{o}_3$	o_8	×				×	×	
$\mathbf{o}_4$	o_7	×	×	×	×			

From the analogical complex structure, we derive a formal definition of a relational proportion.

Definition 5. *Let* $(\mathbf{o}_{1,4}, \mathbf{a}_{1,4})$ *be a complete analogical complex in a formal context, the following sets of objects and attributes are said to be in the* formal relational proportion $(\mathbf{o}_1$ *is to* $\mathbf{a}_3$ *as* $\mathbf{o}_2$ *is to* $\mathbf{a}_2)$*, and we write:* $(\mathbf{o}_1 \Updownarrow \mathbf{a}_3 \upuparrows\!\!\downdownarrows \mathbf{o}_2 \Updownarrow \mathbf{a}_2)$.

Comments.

1. The reduced form of the relational proportion would be ($\mathbf{o}_1$ is the $\mathbf{o}_2$ of $\mathbf{a}_3$).
2. From the same complex, we can extract the 4 following formal relational proportions $\left(\mathbf{o}_1 \Updownarrow \mathbf{a}_4 \upuparrows\!\!\downdownarrows \mathbf{o}_3 \Updownarrow \mathbf{a}_1\right)$, $\left(\mathbf{o}_2 \Updownarrow \mathbf{a}_4 \upuparrows\!\!\downdownarrows \mathbf{o}_4 \Updownarrow \mathbf{a}_1\right)$ and $\left(\mathbf{o}_3 \Updownarrow \mathbf{a}_3 \upuparrows\!\!\downdownarrows \mathbf{o}_4 \Updownarrow \mathbf{a}_2\right)$. Since the operator $\upuparrows\!\!\downdownarrows$ is commutative, it gives a total of 8, but permuting the extreme and the means in a relational proportion may lead to awkward phrasings.

Example 5. Let us take the complex from SmallZoo described above. It implies all attributes but a_6 (predator) and objects o_1 and o_2 (chicken and crow), o_5 (fruitbat), o_8 (penguin) and o_7 (mink). From this context, the RP in reduced form "a fruitbat is the mink of airborne animals" can be derived for instance, meaning that fruitbat and mink have hair, are toothed and produce milk, but that the mink is aquatic at the contrary of the fruitbat. Of course, the interest of such phrases has to be taken in context: the SmallZoo data base is supposed to be the only knowledge.

4.3 WAP and analogical complex

In this section we explore the links between WAP between concepts and complete analogical complex, and then the formal relational proportion.

First, we are interested in defining a non degenerated WAP, called *complete*, forbidding inclusion between two of its concepts. It is a key notion for building WAPs between concepts with a sound cognitive interpretation.

Definition 6. *Let us consider* $(x : y\ WAP\ z : t)$*, this WAP is complete when*

1. *either* $(\mathbf{a}_x \cap \mathbf{a}_y) \setminus \mathbf{a}_\cap$, $(\mathbf{a}_x \cap \mathbf{a}_z) \setminus \mathbf{a}_\cap$, $(\mathbf{a}_y \cap \mathbf{a}_t) \setminus \mathbf{a}_\cap$ *and* $(\mathbf{a}_z \cap \mathbf{a}_t) \setminus \mathbf{a}_\cap$ *are nonempty (called complete WAP through attributes),*
2. *or* $(\mathbf{o}_x \cap \mathbf{o}_y) \setminus \mathbf{o}_\cap$, $(\mathbf{o}_x \cap \mathbf{o}_z) \setminus \mathbf{o}_\cap$, $(\mathbf{o}_y \cap \mathbf{o}_t) \setminus \mathbf{o}_\cap$ *and* $(\mathbf{o}_z \cap \mathbf{o}_t) \setminus \mathbf{o}_\cap$ *are nonempty (called complete WAP through objects).*

where $\mathbf{a}_\cap = \mathbf{a}_x \cap \mathbf{a}_y \cap \mathbf{a}_z \cap \mathbf{a}_t$ *and* $\mathbf{o}_\cap = \mathbf{o}_x \cap \mathbf{o}_y \cap \mathbf{o}_z \cap \mathbf{o}_t$.

Proposition 5.
1. *A complete WAP is an* antichain *of concepts.*
2. *For a complete WAP through attributes,* $(x \vee y)$, $(x \vee z)$, $(y \vee t)$ *and* $(z \vee t)$ *are in antichain. Similarly, for a complete WAP through objects,* $(x \wedge y)$, $(x \wedge z)$, $(y \wedge t)$ *and* $(z \wedge t)$ *are in antichain.*
3. *A FAP in antichain forms a complete WAP through attributes and objects, and reciprocally.*

Proof. 1. Let us suppose that $(x : y \; WAP \; z : t)$ and $x \leq y$. From Preliminary 2, we get $\mathbf{a}_x \cap \mathbf{a}_y = \mathbf{a}_{x \vee y}$. Then $\mathbf{a}_x \cap \mathbf{a}_y = \mathbf{a}_y$ and using Proposition 2

$$\begin{aligned}\mathbf{a}_x \cap \mathbf{a}_z &= (\mathbf{a}_x \cap \mathbf{a}_y) \cap \mathbf{a}_z = \mathbf{a}_x \cap (\mathbf{a}_y \cap \mathbf{a}_z) \\ &= \mathbf{a}_x \cap (\mathbf{a}_x \cap \mathbf{a}_t) = \mathbf{a}_x \cap \mathbf{a}_t = \mathbf{a}_\cap .\end{aligned}$$

Thus, $\mathbf{a}_x \cap \mathbf{a}_z \setminus \mathbf{a}_\cap = \varnothing$ and $(x : y \; WAP \; z : t)$ is not a complete WAP.

2. From a complete WAP through attributes, $\mathbf{a}_{x \vee y} = \mathbf{a}_x \cap \mathbf{a}_y$ and three analog equalities hold. Due to this completeness, there is no inclusion between $\mathbf{a}_{x \vee y}$, $\mathbf{a}_{x \vee z}$, $\mathbf{a}_{z \vee t}$ and $\mathbf{a}_{y \vee t}$. The associated concepts are then in antichain.

3. Let us consider the FAP $x : y :: z : t$ where $\{x, y, z, t\}$ is an antichain. From Proposition 2, we have $x = (x \wedge y) \vee (x \wedge z)$ and $x = (x \vee y) \wedge (x \vee z)$ which are equivalent to $\mathbf{a}_x = \mathbf{a}_{x \wedge y} \cap \mathbf{a}_{x \wedge z}$ and $\mathbf{o}_x = \mathbf{o}_{x \vee y} \cap \mathbf{o}_{x \vee z}$ thanks to Preliminary 2. Similarly, $\mathbf{a}_y = \mathbf{a}_{x \wedge y} \cap \mathbf{a}_{y \wedge t}$ and $\mathbf{o}_x = \mathbf{o}_{x \vee y} \cap \mathbf{o}_{y \vee t}$ and

$$\begin{aligned}\mathbf{a}_\cap &= \mathbf{a}_x \cap \mathbf{a}_y \cap \mathbf{a}_z \cap \mathbf{a}_t \; = \; \mathbf{a}_x \cap \mathbf{a}_t \\ \mathbf{o}_\cap &= \mathbf{o}_x \cap \mathbf{o}_y \cap \mathbf{o}_z \cap \mathbf{o}_t \; = \; \mathbf{o}_x \cap \mathbf{o}_t\end{aligned}$$

due to Proposition 2 and the fact that a FAP is a WAP. Therefore, we have $\mathbf{a}_{x \wedge y} \setminus (\mathbf{a}_x \cap \mathbf{a}_t) \subseteq \mathbf{a}_x \cap \mathbf{a}_y \setminus \mathbf{a}_\cap$. Moreover, $\mathbf{a}_{x \wedge y} \setminus (\mathbf{a}_x \cap \mathbf{a}_t) = \mathbf{a}_{x \wedge y} \setminus \mathbf{a}_{x \vee t}$ is nonempty. Indeed, $\mathbf{a}_{x \wedge y} \setminus \mathbf{a}_{x \vee t} = \varnothing$ implies that $x \vee t \leq x \wedge y$ which is impossible since $\{x, y, z, t\}$ is an antichain. Similarly, we can prove that $\mathbf{o}_x \cap \mathbf{o}_y \setminus \mathbf{o}_\cap \neq \varnothing$.

Reciprocally, let us take a complete WAP through attributes and objects. From the previous properties, $\{x, y, z, t\}$ is an antichain, as well as $\{x \vee y, x \vee z, y \vee t, z \vee t\}$ and $\{x \wedge y, x \wedge z, y \wedge t, z \wedge t\}$. Therefore, these 12 concepts are distinct and it can be proved that they generate a Boolean sublattice. Because of the distributivity of this sublattice, the WAP $(x : y \; WAP \; z : t)$ is then a FAP.

In order to derive relational proportion from an analogical proportion between concepts, we consider a complete WAP through attributes (a similar reasoning can be done from a complete WAP through objects) and introduce a process to extract an analogical complex.

Due to the completeness, sets $\mathbf{a}_1 = (\mathbf{a}_z \cap \mathbf{a}_t) \setminus \mathbf{a}_\cap$, $\mathbf{a}_2 = (\mathbf{a}_y \cap \mathbf{a}_t) \setminus \mathbf{a}_\cap$, $\mathbf{a}_3 = (\mathbf{a}_x \cap \mathbf{a}_z) \setminus \mathbf{a}_\cap$ and $\mathbf{a}_4 = (\mathbf{a}_x \cap \mathbf{a}_y) \setminus \mathbf{a}_\cap$ are nonempty. We also define $\mathbf{o}_1 = \widetilde{\mathbf{o}_x}$ the set of objects proper to x (that appear in $\mathbf{o}_x$ but not in the objects of y, z and t) and similarly $\mathbf{o}_2 = \widetilde{\mathbf{o}_y}$, $\mathbf{o}_3 = \widetilde{\mathbf{o}_z}$ and $\mathbf{o}_4 = \widetilde{\mathbf{o}_t}$.

By construction, every object of $\mathbf{o}_1$ is in relation with every attribute of $\mathbf{a}_3 \cup \mathbf{a}_4$. It is also the case between $\mathbf{o}_2$ and $\mathbf{a}_2 \cup \mathbf{a}_4$, $\mathbf{o}_3$ and $\mathbf{a}_1 \cup \mathbf{a}_3$, $\mathbf{o}_4$ and $\mathbf{a}_1 \cup \mathbf{a}_2$. For all the other combinations, for instance $\mathbf{o}_1$ and $\mathbf{a}_1$, for any $o \in \mathbf{o}_1$, there exists $a \in \mathbf{a}_1$ such that o and a are not in relation. However, these properties do not guarantee that the subcontext $(\mathbf{o}_{1,4}, \mathbf{a}_{1,4})$ is an analogical schema, even if it is a closed schema. Indeed, it can exists an object $o \in \mathbf{o}_i$ in relation with an attribute $a \in \mathbf{a}_j$, where $(i, j) \in \{(1, 1), (1, 2), (2, 1), (2, 3), (3, 2), (3, 4), (4, 3), (4, 4)\}$. In such a case, either a or o is removed and this postprocessing permits to obtain an analogical schema. But this schema is not necessarily a complex, since the associated subcontext may be not maximal. Then a second postprocessing maximises the schema into complex, adding new attributes and/or objects chosen among those which do not appear in $\mathbf{a}_x \cup \ldots \cup \mathbf{a}_t$ nor $\mathbf{o}_x \cup \ldots \cup \mathbf{o}_t$. Finally, we check that the resulting analogical complexes are complete.

This method can lead to several complexes, according to the choices in both postprocessings. This set of complexes is a sub-lattice of the lattice of complexes.

Example 6. In SmallZoo, $x = (\{o_1, o_2, o_4\}, \{a_1, a_2, a_4\})$, $y = (\{o_5\}, \{a_0, a_3, a_4, a_7\})$, $z = (\{o_4, o_8\}, \{a_1, a_2, a_5\})$, $t = (\{o_7, o_9\}, \{a_0, a_3, a_5, a_6\})$ are concepts in complete WAP through attributes. At the beginning, $\mathbf{o}_1 = \{o_1, o_2\}$, $\mathbf{o}_2 = \{o_5\}$, $\mathbf{o}_3 = \{o_8\}$, $\mathbf{o}_4 = \{o_7, o_9\}$, $\mathbf{a}_1 = \{a_5\}$, $\mathbf{a}_2 = \{a_0, a_3\}$, $\mathbf{a}_3 = \{a_1, a_2\}$ and $\mathbf{a}_4 = \{a_4\}$ and, due to the relation between o_9 and a_2, the first postprocessing can remove (either o_9 or) a_2:

	a_5	a_0	a_3	a_1	a_2	a_4
o_1				×	×	×
o_2				×	×	×
o_5		×	×			×
o_8	×			×	×	
o_7	×	×	×			
o_9	×	×	×		×	

		$\mathbf{a}_1$	$\mathbf{a}_2$		$\mathbf{a}_3$	$\mathbf{a}_4$
		a_5	a_0	a_3	a_1	a_4
$\mathbf{o}_1$	o_1				×	×
	o_2				×	×
$\mathbf{o}_2$	o_5		×	×		×
$\mathbf{o}_3$	o_8	×			×	
$\mathbf{o}_4$	o_7	×	×	×		
	o_9	×	×	×		

After removing a_2, the right table is an analogical schema and we can check that it is maximal in SmallZoo. Note that if we had chosen to remove o_9, the postprocessings would have produced the analogical complex previously detailed in Example 4.

For example, from the complete analogical complex described above, we can derive the following relational proportion: "the chicken and the crow are to the feathers as the fruitbat is to the hair, the milk and the teeth". It makes sense when considering that all these animals share the attribute "airborne".

Likewise, another proportion from the same complex is "the fruitbat is to the airborne animals as the mink and the platypus are to the aquatic animals" (fruitbat, mink and platypus share the attributes hair and milk). The reduced form "the fruitbat is the mink of airborne animals" is the same as that of Example 5, Section 4.2, although the complexes involved are slightly different.

5 Conclusion

The paper has shown how relational proportions can be identified in a formal context. Relational proportions offer a basis for concise forms of explanations. Indeed, if B has some well-known features, the proportion "object A is to attribute a as object B is to attribute b" provides an argument for stating that "object A is the B of a", when A possesses these well-known features also, as in "Carlsen is the Mozart of chess". It is worth pointing out that two cognitive capabilities, namely conceptual categorization and analogical reasoning can be handled together in the setting of formal concept analysis. This introductory presentation has left aside the algorithmic side (based on the identification of formal complexes), which is discussed in the long version of the paper [10].

Our study of proportions between concepts explores a simple and fixed relation between concepts in a single lattice. It would be interesting to connect it with the general framework of Relational Concept Analysis (see e.g., [11]), and with a recent proposal based on antichains [12].

Following the pioneering work of Rumelhart and Abrahamson [13], a number of recent woks in computational linguistics (e.g., [14]) have been using a parallelogram-based modeling of analogical proportions in numerical settings, where words are represented by vectors of great dimension. Bridging this computational view of analogical proportions with the work presented here is certainly a challenging task for the future.

References

1. Lepage, Y.: De l'analogie rendant compte de la commutation en linguistique. Université Joseph Fourier, Grenoble (2001). www.slt.atr.jp/~lepage/pdf/dhdryl.pdf
2. Miclet, L., Prade, H.: Handling Analogical Proportions in Classical Logic and Fuzzy Logics Settings. In: 10th Eur. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, pp. 638–650, Verona (2009).
3. 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).
4. Miclet, L., Barbot, N., Prade, H.: From Analogical Proportions in Lattices to Proportional Analogies in Formal Contexts. In: 21st of Eur. Conf. on Artificial Intelligence, Prague (2014).
5. Bělohlávek, R.: Introduction to formal concept analysis. Technical report, Palacký University, (2008).
6. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer-Verlag (1998).
7. Miclet, L., Nicolas, J.: From formal concepts to analogical complexes. In: 12th Int. Conf. on Concept Lattices and their App., pp. 159–170, Clermont-Ferrand (2015).
8. Lichman, M.: UCI Machine Learning Repository. University of California, Irvine, School of Information and Computer Sciences (2013). http://archive.ics.uci.edu/ml
9. Dubois, D., Prade, H., Richard, G.: Multiple-valued extensions of analogical proportions, Fuzzy Sets and Systems v.292, 193–202 (2016).
10. Barbot, N., Miclet, L., Prade, H.: Analogy between concepts. Submitted.
11. Rouane-Hacene, M., Huchard, M., Napoli, A., Valtchev, P. : Relational concept analysis: mining concept lattices from multi-relational data. Annals of Mathematics and Artificial Intelligence, v. 67, 81-108 (2013).
12. Valverde-Albacete, F. J., Peláez-Moreno, C., Cabrera, I. P., Cordero, P., Ojeda-Aciego, M.: Formal independence analysis. In: 17th Int. Conf. Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations, Jesús Medina, et al. (ed.), CCIS 853, Springer, 596–608, Cádiz (2018).
13. Rumelhart, D. E., Abrahamson, A. A.: A model for analogical reasoning. Cognitive Psychol. v.5, 1–28 (1973).
14. Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J.: Distributed representations of words and phrases and their compositionality. Advances in Neural Information Processing Systems 26 (2013).