Analogical proportions: From equality to inequality

Henri Prade, Gilles Richard

Official URL

DOI : https://doi.org/10.1016/j.ijar.2018.07.005

Analogical proportions: From equality to inequality ☆

Henri Prade [a,b,*], Gilles Richard [a]

[a] IRIT, CNRS & Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 09, France
[b] QCIS, University of Technology, Sydney, Australia

Keywords:
Analogical proportion
Similarity Dissimilarity
Boolean logic Multiple-valued logic
Algorithmic complexity

A B S T R A C T

Analogical proportions are statements of the form *a* is to *b* as *c* is to *d*. Such expressions compare the pair (*a*, *b*) with the pair (*c*, *d*). Previous papers have developed logical modelings of such proportions both in Boolean and in multiple-valued settings. They emphasize a reading of the proportion as "the way *a* and *b* differ is the same as *c* and *d* differ". The ambition of this paper is twofold. The paper first provides a deeper understanding and further justifications of the Boolean modeling, before introducing analogical inequalities, where "as" is replaced by "as much as" in the comparison of two pairs. From an abstract viewpoint, analogical proportions are supposed to obey at least three postulates expressing reflexivity, symmetry, and stability under central permutation. Nevertheless these postulates are not enough to determine a single model and a minimality condition has to be added as shown in this paper. These models are organized in a complete lattice based on set inclusion. This leads us to discuss lower and upper approximations of the minimal model. Apart from being minimal, this model can also be evaluated in terms of Kolmogorov complexity via an expression reflecting the intended meaning of analogy. We show that the six Boolean patterns of the minimal model that make Boolean analogy true minimize this expression. Besides, analogical proportions extend to 4-tuples of Boolean vectors. This enables us to explain why analogical proportions also reads in terms of similarity (rather than difference, i.e., dissimilarity): *a* and *d* share the same presence or absence of features as *b* and *c*. Moreover, we establish a link between analogical proportion and Hamming distances between components of the proportion. We also emphasize that analogical proportions are pervasive in any comparison of two vectors *a* and *d* that implicitly induce the existence of "intermediary" vectors *b* and *c* forming together such a proportion. The similarity reading and the dissimilarity reading of a Boolean analogical proportion are no longer equivalent in the multiple-valued setting, where they give birth to two distinct options that are recalled. These options are also discussed with respect to their capability to handle so-called "continuous" logical proportions of the form *a* is to *b* as *b* is to *c* involving some idea of "betweenness". In all the previously investigated issues, the pairs involved in the 4-tuples were compared via equalities of similarities or equalities of dissimilarities. This observation suggests to also consider statements of the form "*a* is to *b* at least as much as *c* is to *d*", leading to the concept of "analogical inequalities". Thus, instead of expressing equality between differences or similarities, as it is the case for analogical proportions, it is also interesting to express inequalities between such differences or similarities. Starting from the modeling of analogical proportions, we investigate the logical modeling of analogical inequalities, both in the Boolean and in the multiple-valued cases, and discuss their potential use in relation with some recent related work in computer vision.

☆ This paper is part of the Virtual special issue on the 14th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2017), Edited by Alessandro Antonucci, Laurence Cholvy and Odile Papini.

* Corresponding author at: IRIT, CNRS & Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 09, France.
E-mail addresses: prade@irit.fr (H. Prade), richard@irit.fr (G. Richard).

https://doi.org/10.1016/j.ijar.2018.07.005

1. Introduction

Comparative thinking plays a key role in our assessment of reality. This has been recognized for a long time. Making comparison is closely related to similarity judgment [45] and analogy making [12,17]. Analogical proportions, i.e., statements of the form a is to b as c is to d, usually denoted $a : b :: c : d$, provides a well-known way for expressing a comparative judgment between two pairs (a, b) and (c, d); see, e.g., [14,15]. Such a statement suggests that the comparison (in terms of similarity and dissimilarity) of the elements of pair (a, b) yields the same kind of result as when comparing the elements of pair (c, d) [39].

Analogical proportions constitute a key notion for formalizing analogical inference by relying on the following principle: if such proportions hold on a noticeable subset of known features used for describing the four items, the proportion may still hold on other features as well, which may help guessing the unknown values of d on these other features from their values on a, b, and c. The interest of such inference mechanism has been recently pointed out in machine learning for classification problems [2,28,4], and in visual multiple-class categorization tasks for handling pieces of knowledge about semantic relationships between classes. More precisely in this latter case, analogical proportions are used for expressing analogies between pairs of concrete objects in the same semantic universe and with similar abstraction level, and then this gives birth to constraints that serve regularization purposes [18]. Besides, the power of analogical proportion-based inference has been also illustrated on the solving of IQ tests [6].

Different formal modelings of analogical proportions have been proposed in the last decades. Quite early, a theory of analogical reasoning, where elements are represented as points in multidimensional Euclidean spaces, and analogical proportions are represented by parallelograms in such spaces, has been proposed in [41]. This geometric view is at work in the above-mentioned reference in visual categorization. An empirical modeling of analogy making, where the fourth term d of an analogical proportion $a : b :: c : d$ to be completed is obtained by minimization of the difference of changes between a and b and between c and d is at work in the programs ANALOGY [10] and COPYCAT [16]. Later on, a machine learning-oriented view where analogical proportions are represented in terms of Kolmogorov algorithmic complexity has been presented in [5]. A similar, but simplified modeling, still expressing that a and b differ as c and d differ, can be found in [1], where the complexities of the target and source universes have not to be taken into account, since they are identical in this latter case. Quite more recently, a set of various algebraic modelings of analogical proportions have been introduced and discussed in [25,29,30,46]. Following these works, a logical modeling has been proposed [31,32]. This logical modeling makes clear that the analogical proportion holds if and only if a differs from b as c differs from d and vice-versa. This fits quite well with what is suggested by the usual reading of the proportion that states that "a is to b as c is to d", where "a is to b" (resp. "c is to d") refers to an implicit pairwise comparison, and the central "as" to an identity. This leads to a Boolean truth table for $a : b :: c : d$ which makes the expression true for six patterns of values of the 4-tuple a, b, c, d among $2^4 = 16$ possible patterns. It can easily be checked that the expected postulates (reflexivity, symmetry, formal permutation) are satisfied by the modeling. However, one may wonder if other modelings would make sense for an analogical proportion, and if other justifications could be found for the above-mentioned modeling. This is one of the goals of this paper.

The paper first investigates new justifications of the Boolean expression of an analogical proportion. First, starting from the core postulates supposed to be satisfied by an analogical proportion, and agreed by everybody for a long time, we exhibit all the Boolean models compatible with them. There are several ones, but the smallest model is the basic Boolean expression of an analogical proportion previously proposed. This smallest model is indeed characterized by the six expected Boolean patterns. Another understanding of analogical proportion, in terms of *similarity*, can be stated as "what a and d have in common (positively and negatively), b and c have it also". It corresponds to a Boolean formula that turns to be equivalent to the one induced by the difference-based reading, since the same truth-table is obtained in both cases, as observed for about ten years now [32]. We also provide a direct proof and an intuitive explanation of this fact.

Moreover, we try to evaluate the cognitive significance of the proposed Boolean modeling of an analogical proportion in terms of algorithmic complexity (i.e., Kolmogorov complexity) and show that it is also minimal among all Boolean patterns with respect to an algorithmic complexity-based definition of an analogical proportion. Indeed algorithmic complexity measures a kind of universal information content of a Boolean string. Despite its inherent uncomputability, there exist powerful tools for computing good approximations. Kolmogorov complexity has been proved to be of great value in diverse applications: for example, in distance measures [3] and classification methods, plagiarism detection, network intrusion detection [13], and in numerous other applications [27].

As already said, analogical proportions express the identity of the results of the comparisons of two pairs. We may wonder if an inequality instead of an equality would make sense as well and would be useful for expressing constraints of the form "a is to b as much as c is to d". In fact, constraints of the same kind, but stated in terms of distances, have been shown to be useful for categorization tasks in computer vision for representing pieces of knowledge stating *relative* comparisons between quadruplets of images, feature by feature [23,24]. Interestingly enough, it has been also recently noticed that similar relations in terms of comparison of pairs were also present in multiple criteria analysis for expressing, for instance, that the “difference” between two evaluation vectors on a criterion is smaller than (i.e., does not compensate) the “difference” between the vectors on the rest of the criteria [34]. This recent emergence of the interest for inequality constraints between pairs of items motivates the introduction of the notion of “analogical inequalities” and their formal study, in relation with the Boolean as well as the multiple-valued modeling of analogical proportions.

2. Summary

The paper is based on two conference papers [37,38] dealing respectively:

– on the one hand, with logical and algorithmic complexity justifications for the Boolean modeling of analogical proportions, and

– on the other hand, with the notion of “analogical inequalities”.

The paper gathers this material in a unified way. It also substantially extends it in various directions in particular regarding diverse aspects of Boolean proportions not previously considered: i) direct proof of the equivalence between the 2 basic expressions of analogical proportions (Subsection 3.4); ii) Investigation of the validity of the vectorial extension for these formulas (Subsection 5.1); iii) Operational explanation of the equivalence between these formulas (Subsection 5.2); iv) Link between analogical proportions and Hamming distance (Subsection 5.3); v) Emergence of analogical proportions from items comparison (Subsection 5.4); vi) Formal discussion of the modeling of continuous analogical proportions (Subsection 6.2). Moreover, the presentation of the results of the conference paper has been somehow improved by giving more details and providing extensive discussions of related works.

We now describe the general organization and contents of the article. We first start in Section 3 with the postulates governing analogical proportions and study the Boolean modelings of analogy[1] compatible with them. This investigation lays bare a lattice of Boolean models for analogical proportions, which includes a minimal model corresponding to the one introduced in [31] and subsequently developed, as well as another model of particular interest proposed a long time ago by Klein [20,21]. While the minimal model makes the proportion true for six input patterns, Klein's operator is true for eight. We also study lower approximations of the minimal model which are true for four input patterns only and which only satisfy a part of the postulates. The minimal model can be naturally associated with a syntactic Boolean formula expressing exactly that the difference of a with b is the same as the difference of c with d and vice-versa. We investigate other remarkable, logically equivalent, expressions of the analogical proportion. In particular, we establish that the analogical proportion also fits with the statement that “what a and d have in common (positively and negatively), b and c have it also”. The discussion of the meaningful syntactic forms is important for determining the forms that are appropriate for an extension to Boolean vectors, and then to multiple-valued logic.

In Section 4, we first briefly review the main definition and theorems of Kolmogorov complexity. Once we have the main tools, we are in a position to give a Kolmogorov complexity-based definition of analogy and to experiment in order to empirically validate the minimal model-based definition. Section 5 is devoted to analogical proportion defined for Boolean vectors. This setting makes easier the understanding of the equivalence between the difference-based and the similarity-based readings of analogy. Moreover, we establish links between vectorial analogical proportion and constraints expressed in terms of Hamming distances between components of the proportion. We also point out that analogical proportions are pervasive in any comparison of two vectors, say a and d, that implicitly induce the existence of “intermediary” vectors b and c forming together a proportion $a : b :: c : d$. In Section 6, two meaningful extensions of Boolean definitions to multiple-valued proportions, agreeing respectively with the difference-based and the similarity-based readings of analogy, are recalled and put in the perspective of the previous sections. Then continuous analogical proportions, i.e., statements of the form “a is to b as b is to c” which are trivialized in the Boolean setting, but makes sense for multiple-valued proportions, are discussed. With all this necessary background on Boolean and multiple-valued proportions, we are equipped for finally introducing “analogical inequalities” in Section 7, in the Boolean and then in the multiple-valued settings. In both cases, we establish some expected properties, including transitivity. Finally, we discuss the potential interest of such inequalities, on the basis of some recent related works.

3. Boolean analogical proportions

At the time of Aristotle, the idea of analogical proportion originated from the notion of numerical proportion. In that respect, the *arithmetic* proportion between 4 integers a, b, c, d, which holds iff $a - b = c - d$, is a good prototype of the idea of analogical proportion, since we can read it as “a differs from b as c differs from d”, which perfectly fits with “a is to b as c is to d”. This arithmetic proportion can be easily extended over a vector space like $\mathbb{R}^n$ by keeping the same definition. It simply means the proportion holds in $\mathbb{R}^n$ iff the points a, b, c, d are the vertices of a parallelogram in a plan. In the following, we denote an analogical proportion between four ordered elements a, b, c, d as $a : b :: c : d$. When considering Boolean interpretation where $a, b, c, d \in \{0, 1\}$, it is tempting to carry on with the same definition as $\{0, 1\} \subset \mathbb{R}$, with the inevitable drawback that difference is not an internal operator in $\mathbb{B} = \{0, 1\}$. Nevertheless, if we draw the truth table (16 lines) corresponding to this definition, we get Table 1 highlighting that only 6 among $2^4 = 16$ lines are valid proportions.

[1] In this paper, when there is no ambiguity, we may use the words ‘analogy’, or ‘proportion’, in place of ‘analogical proportion’.

Table 1
Boolean valuations for $a : b :: c : d$.

a	b	c	d	$a : b :: c : d$	a	b	c	d	$a : b :: c : d$
0	0	0	0	**1**	1	0	0	0	0
0	0	0	1	0	1	0	0	1	0
0	0	1	0	0	1	0	1	0	**1**
0	0	1	1	**1**	1	0	1	1	0
0	1	0	0	0	1	1	0	0	**1**
0	1	0	1	**1**	1	1	0	1	0
0	1	1	0	0	1	1	1	0	0
0	1	1	1	0	1	1	1	1	**1**

3.1. Postulate perspective

Taking inspiration from numerical proportions and from an abstract viewpoint, analogy can be viewed as a quaternary relation R, supposed obeying the following three postulates (e.g., [25,29]):

1. $\forall a, b, R(a, b, a, b)$ (reflexivity);
2. $\forall a, b, c, d, R(a, b, c, d) \to R(c, d, a, b)$ (*symmetry*);
3. $\forall a, b, c, d, R(a, b, c, d) \to R(a, c, b, d)$ (*central permutation*).

Some basic properties can be deduced by proper applications of symmetry and central permutation:

- $\forall a, b, R(a, a, b, b)$ (identity);
- $\forall a, b, c, d, R(a, b, c, d) \to R(b, a, d, c)$ (*inside pair reversing*);
- $\forall a, b, c, d, R(a, b, c, d) \to R(d, b, c, a)$ (*extreme permutation*).

In fact, another less standard axiom expected from a *natural* analogy is:

$$\forall a, b, x, \quad R(a, a, b, x) \to (x = b) \ \ (unicity)$$

All these properties fit with the intuition of what may be an analogical proportion. It can be also easily checked that they are satisfied by the arithmetic proportion $a - b = c - d$ (as well as with the geometric proportion $\frac{a}{b} = \frac{c}{d}$).

If we now focus on $\mathbb{B} = \{0, 1\}$ as interpretation domain, R should be interpreted as a subset of $\mathbb{B}^4$. Removing the empty set, leaves $2^{16} - 1$ candidate models. It is straightforward to get a basic model:

- By applying reflexivity, we see that $0101, 1010$ should belong to the relation and $0000, 1111$ as well since we may have $a = b$, and
- Central permutation then leads to add 0011 and 1100.

Thus, we get the model $\Omega_0 = \{0000, 1111, 0101, 1010, 0011, 1100\}$, which is also stable under symmetry. Obviously Ω_0 is the *smallest* Boolean model for analogy as every model should include it as a subset.

3.2. Lattice of Boolean models

In fact, there is more than one model of analogy in a Boolean interpretation. We have the following result:

Property 1. *There are exactly 8 models of analogy (i.e., satisfying the three first postulates) over $\mathbb{B}$. There are exactly 2 models of analogy satisfying the three first postulates plus unicity.*

Proof. Any model should include Ω_0. Let us note that a bigger model should necessarily have an even cardinality due to the following facts:

– To be bigger than Ω_0, it should contain a string s containing both 0 and 1.

– Thanks to symmetry or central permutation axioms, it should contain the symmetric *cdab* of $s = abcd$ and the central permutation *acbd* of s: necessarily, one of these 2 strings is different from s (otherwise, we get $a = b = c = d$).

So we have to look for models of cardinality 8, 10, 12, 14 and 16. Obviously $\mathbb{B}^4$ of cardinality 16 is a model, the biggest one. Due to the axioms, we have to add to Ω_0 subsets of $\mathbb{B}^4$ that are stable w.r.t. symmetry and central permutation. We have exactly:

– one subset with 2 elements:

$S_2 = \{1001, 0110\}$

– two subsets with 4 elements:

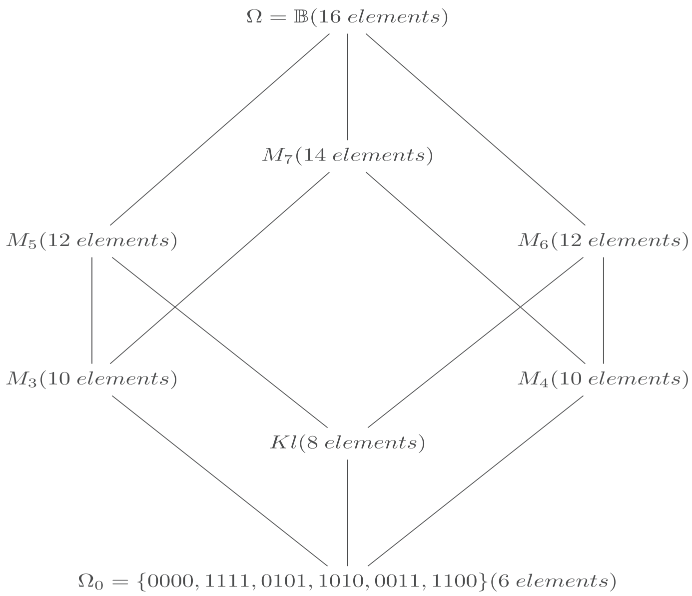

Fig. 1. The lattice of Boolean models of analogy.

i) $S_3 = \{1110, 1101, 1011, 0111\}$;
ii) $S_4 = \{0001, 0010, 0100, 1000\}$.

Since every model has to be built by adding to Ω_0 one of the previous subsets, we get the following models for analogy in $\mathbb{B}$:

1) one model with 6 elements: Ω_0 (the smallest one),
2) one model with 8 elements:
$Kl = \Omega_0 \cup S_2 = \{0000, 1111, 0101, 1010, 0011, 1100, 0110, 1001\}$
As previously mentioned, this model is due to S. Klein [20].
3) two models with 10 elements:
– $M_3 = \Omega_0 \cup S_3 = \{0000, 1111, 0101, 1010, 0011, 1100, 1110, 1101, 1011, 0111\}$,
– $M_4 = \Omega_0 \cup S_4 = \{0000, 1111, 0101, 1010, 0011, 1100, 0001, 0010, 0100, 1000\}$,
4) two models with 12 elements:
– $M_5 = M_3 \cup S_2 = \{0000, 1111, 0101, 1010, 0011, 1100, 1110, 1101, 1011, 0111, 0110, 1001\}$,
– $M_6 = M_4 \cup S_2 = \{0000, 1111, 0101, 1010, 0011, 1100, 0001, 0010, 0100, 1000, 0110, 1001\}$,
5) one model with 14 elements:
– $M_7 = M_3 \cup S_4 = M_4 \cup S_3 = \Omega_0 \cup S_3 \cup S_4 = \{0000, 1111, 0101, 1010, 0011, 1100, 1110, 1101, 1011, 0111, 0001, 0010, 0100, 1000\}$,
6) one model with exactly 16 elements: $\Omega = \Omega_0 \cup S_2 \cup S_3 \cup S_4 = \mathbb{B}$.

Finally Ω_0 and Kl satisfy unicity but M_3 (containing 1100 and 1101) and M_4 (containing 0000 and 0001) do not satisfy it. This achieves the proof. □

The set of models is a lattice with bottom element Ω_0 and top element $\mathbb{B}$, see Fig. 1. As can be seen, 8 models fit with the axioms in the Boolean case, including the 6-patterns model Ω_0 and the 8-patterns model Kl due to Klein. However, it is natural to privilege the smallest model, the minimal one that just accounts for the axioms and nothing more. Besides, it is worth noticing that 4 models are *code independent*, which means that they are stable if 0 and 1 are exchanged. For a given model M, this is formally expressed as:

$$(a, b, c, d) \in M \implies (\neg a, \neg b, \neg c, \neg d) \in M.$$

They are Ω_0, Kl, M_7 and Ω. Although code independency is not a postulate, it seems to be a desirable property for analogical proportion since the encoding by 0 or 1 is a matter of convention (depending on if one privileges a property or its opposite).

3.3. *Upper and lower approximations of analogical proportions*

As can be seen, Klein's model can be considered as an upper bound of Ω_0. It was the first logical view proposed for analogical proportion [20]. S. Klein suggested that an analogical proportion would hold as soon as a, b, c are completed by d taken as $d = c \equiv (a \equiv b)$. This amounts to define it as $A_K(a, b, c, d) \triangleq (a \equiv b) \equiv (c \equiv d)$. Then $0 : 1 :: 1 : 0$ and $1 : 0 :: 0 : 1$ would become valid analogical proportions and leads to the model denoted Kl in the previous subsection. The validity of such patterns may be advocated on the basis of a functional view of analogy where $a : f(a) :: b : f(b)$ sounds indeed valid, taking the negation in $\mathbb{B}$ for f. But, this is debatable since then $A_K(a, b, c, d) \leftrightarrow A_K(b, a, c, d)$ (which does not fit with intuition). It turns out that $a : b :: c : d \rightarrow A_K(a, b, c, d)$.

While $A_K(a, b, c, d)$ is an upper approximation of $a : b :: c : d$ true for 8 patterns, one may look for lower approximations that are true for 4 patterns only (taking into account code independency). There are 3 such approximations, given below, followed by the patterns they validate[2]:

- $(a \equiv b) \wedge (c \equiv d)$ $\begin{array}{|cccc|}\hline 1&1&1&1\\ \hline 0&0&0&0\\ \hline 1&1&0&0\\ \hline 0&0&1&1\\ \hline\end{array}$;
- $(a \equiv c) \wedge (b \equiv d)$ $\begin{array}{|cccc|}\hline 1&1&1&1\\ \hline 0&0&0&0\\ \hline 1&0&1&0\\ \hline 0&1&0&1\\ \hline\end{array}$;
- $(a \not\equiv d) \wedge (b \not\equiv c)$ $\begin{array}{|cccc|}\hline 1&1&0&0\\ \hline 0&0&1&1\\ \hline 1&0&1&0\\ \hline 0&1&0&1\\ \hline\end{array}$.

Obviously, these lower approximations that are missing two of the patterns of the minimal model Ω_0, while preserving code independency, obey only a subset of the three postulates. All are symmetrical. Only the third approximation satisfies central permutation, but misses to satisfy "a is to a as a is to a" (a special case of reflexivity and identity, satisfied by the first two approximations).

We have seen that the analysis of the Boolean models that are compatible with the postulates of analogical proportion offers a kind of intrinsic justification that only the minimal model true for 6 patterns (and false for the 10 others) is reasonable. Before providing another justification in terms of algorithmic complexity in the next section, we revisit and discuss different syntactic propositional forms that are true for these patterns only, and that lay bare various aspects of the meaning of Boolean analogical proportion.

3.4. Boolean formulas for analogical proportion

Instead of relying on the "arithmetical" definition of analogy ($a - b = c - d$), we may look for a propositional logic definition of $a : b :: c : d$. Since analogy is a matter of comparison, it is natural to use basic comparative indicators as building blocks, as done in [35,36]. These indicators are naturally associated to any pair of Boolean variables (a, b):
– $a \wedge b$ and $\neg a \wedge \neg b$ are *positive similarity* and *negative similarity* indicators respectively: $a \wedge b$ (resp. $\neg a \wedge \neg b$) is true iff only both a and b are true (resp. false);
– $a \wedge \neg b$ and $\neg a \wedge b$ are *dissimilarity* indicators: $a \wedge \neg b$ (resp. $\neg a \wedge b$) is true iff only a (resp. b) is true and b (resp. a) is false.

Then analogical proportion $a : b :: c : d$ could be defined by formula (1) [32] as below:

$$a : b :: c : d =_{def} (a \wedge \neg b \equiv c \wedge \neg d) \wedge (\neg a \wedge b \equiv \neg c \wedge d) \tag{1}$$

This formula reads "a differs from b as c differs from d and b differs from a as d differs from c", which fits with the expected meaning of analogy. Formula (1) is obviously stable w.r.t. negation, namely

$$a : b :: c : d \rightarrow \neg a : \neg b :: \neg c : \neg d$$

making clear the code independency of analogical proportion. As can be noticed, the presence of dissimilarity indicators in (1) exactly fits with patterns 1010 and 0101, while the four other patterns cover the case of no difference between a and b and between c and d.

Taking inspiration of a well-known property of arithmetical proportions: $a - b = c - d$ is equivalent to $a + d = b + c$ (or from a similar property for geometric proportion: $\frac{a}{b} = \frac{c}{d}$ equivalent to $ad = bc$ for non-zero numbers), one could ask if there

[2] There are 3 companion approximations that involve the two additional patterns of A_K

$$(a \equiv d) \wedge (b \equiv c) \begin{array}{|cccc|}\hline 1&1&1&1\\ \hline 0&0&0&0\\ \hline 1&0&0&1\\ \hline 0&1&1&0\\ \hline\end{array}; \quad (a \not\equiv b) \wedge (c \not\equiv d) \begin{array}{|cccc|}\hline 1&0&0&1\\ \hline 0&1&1&0\\ \hline 1&0&1&0\\ \hline 0&1&0&1\\ \hline\end{array}; \quad (a \not\equiv c) \wedge (b \not\equiv d) \begin{array}{|cccc|}\hline 1&1&0&0\\ \hline 0&0&1&1\\ \hline 1&0&0&1\\ \hline 0&1&1&0\\ \hline\end{array}.$$

would be a counterpart in Boolean logic, involving a and d on one side and b and c on the other side. This is indeed the case: the following two formulas, which are clearly equivalent (the first one (1) exhibiting only comparison indicators, while formula (2) has no negation), have the same truth tables [32] as formula (1).

$$(a \wedge d \equiv b \wedge c) \ \wedge \ (\neg a \wedge \neg d \equiv \neg b \wedge \neg c) \ \ (A) \,;\, (a \wedge d \equiv b \wedge c) \ \wedge \ (a \vee d \equiv b \vee c) \ \ (B) \tag{2}$$

Note that code independency is obvious on (1), as well as the 3 postulates of analogy on the two expressions.

Formulas (1) and (2) provide two completely different views of the idea of analogy, respectively in terms of dissimilarity and in terms of similarity. Their logical equivalence is far from being obvious at first glance, without resorting to truth tables, as often done. This is why, we give below a simple proof in syntactic terms, using several steps.

Property 2. $((a \wedge \neg b \equiv c \wedge \neg d) \wedge (\neg a \wedge b \equiv \neg c \wedge d)) \equiv ((a \wedge d \equiv b \wedge c) \ \wedge \ (a \vee d \equiv b \vee c))$ *is a tautology.*

Proof.

i) (1) implies $a \to (b \vee c)$.

Since (1) holds, the first equivalence in (1) $a \wedge \neg b \equiv c \wedge \neg d$ holds. Thus $(a \wedge \neg b) \to c$ holds, which is equivalent to $a \to (b \vee c)$. This gives the result.

ii) (2) implies $(a \wedge \neg b \equiv c \wedge \neg d)$.

Starting from (2), we suppose $a \wedge \neg b$ holds. Considering the first equivalence in expression (2): for this equivalence to hold, we need d not to hold, i.e., $\neg d$ to hold. Considering now the second equivalence in expression (2), since b does not hold, we need c to hold. Then we have w implies $(a \wedge \neg b \to c \wedge \neg d)$. Proving that (2) implies $(c \wedge \neg d \to a \wedge \neg b)$ is done similarly.

iii) (1) implies (2).

(1) implies $(d \to (b \vee c))$ comes from extreme permutation property and statement i). Then (1) implies $((a \vee d) \to (b \vee c))$. Code independency of (1) and De Morgan laws ensure that (1) implies $((b \wedge c) \to (a \wedge d))$ as well. Observing that exchanging a and c on the one hand and b and d on the other hand leave (1) unchanged thanks to symmetry postulate, it makes clear that (1) implies $((a \wedge d) \equiv (b \wedge c))$ and $((a \vee d) \equiv (b \vee c))$. This is the expected result.

iv) (2) implies (1).

We start from statement ii). Code independency applied to expression (2) shows that (2) implies $(\neg a \wedge b \equiv \neg c \wedge d)$ and we are done. □

In light of expression (2), "a is to b as c is to d" can now be read "what a and d have in common, b and c have it also (both positively and negatively)", which, however, is a less straightforward reading of the idea of analogy than the one associated with expression (1).

As said above, expression (2) can also be viewed as the logical counterpart of a well-known property of geometrical proportions: the product of the means is equal to the product of the extremes. Interestingly enough, Piaget [33, pp. 35–37] named *logical proportion* any logical expression between four propositional formulas a, b, c, d for which (2) is true. Apparently, and strangely enough, Piaget never related this expression to the idea of analogy.

Besides, by observing the truth table of the minimal model of analogical proportion, it is easy to check that formulas other than (1) or (2) can be obtained by combining lower approximations by means of disjunction. Three options are possible, which yield noticeable, equivalent, disjunctive expressions of the analogical proportion:

Property 3. *The 3 following formulas are equivalent to formula (1) (and then (2) as well):*

$$((a \not\equiv d) \wedge (b \not\equiv c)) \vee ((a \equiv b) \wedge (c \equiv d))$$

$$((a \equiv b) \wedge (c \equiv d)) \vee ((a \equiv c) \wedge (b \equiv d))$$

$$((a \equiv c \wedge b \equiv d)) \vee ((a \not\equiv d) \wedge (b \not\equiv c))$$

Nevertheless, we shall see in Section 5 that these formulas, involving disjunction as top connector instead of conjunction, are not really suitable to model analogy in higher dimension when considering analogy between 4-tuples of vectors of Boolean features.

We now investigate if another justification in favor of the minimal model Ω_0 could be obtained. Instead of focusing on the word "minimal" in terms of size, it could be helpful to consider the word "minimal" in the context of information theory. We investigate this option in the following section.

4. Complexity-based justification of the Boolean model

As far as we know, the work of [5] was the first to establish a link between analogical inference and information theory, starting from a machine learning perspective. The aim was to integrate analogical reasoning in the global landscape of predicting values from observable examples. A clear definition of "information content" comes from Kolmogorov complexity theory, also known as Algorithmic Complexity Theory. Kolmogorov complexity is not a new concept and the theory has been designed many years ago: see for instance [27] for an in-depth study. Let us briefly recall the main ideas of this theory.

4.1. Kolmogorov complexity as a starting point

We need the help of a universal Turing machine denoted U. Then p denotes a program running on U. Two situations can happen: i) either p does not stop for the input x, or ii) p stops for the input x and outputs a finite string y. In that case, we write $U(p,x) = y$. The Kolmogorov complexity [27] of y w.r.t. x is then defined as:

$$K_U(y/x) = \min\{|p|, U(p,x) = y\}.$$

$K_U(y/x)$ is the size of the shortest program able to reconstruct y with the help of x. The Kolmogorov complexity [27] of y is just obtained with the empty string ϵ:

$$K_U(y) = \min\{|p|, U(p,\epsilon) = y\}.$$

Given a string s, $K_U(s)$ is an integer which, in some sense, is a measure of the information content of s: instead of sending s to somebody, we can send p from which s can be recovered as soon as this somebody has the machine U. K_U enjoys a lot of properties among which is a kind of universality: this complexity is independent of the underlying Turing machine as we have the invariance theorem [27]:

Property 4. *If $U1$ and $U2$ are two universal Turing machines, there exists a constant c_{U1U2} such that for all strings s: $|K_{U1}(s) - K_{U2}(s)| < c_{U1U2}$, where $K_{U1}(s)$ and $K_{U2}(s)$ denote the algorithmic complexity of s w.r.t. $U1$ and $U2$ respectively.*

This theorem guarantees that complexity values may only diverge by a constant c (e.g. the length of a compiler or a translation program) and for huge complexity strings, we can denote K without specifying the Turing machine U. It can also be shown that [27]:

Property 5. $\forall x, y, K(xy) = K(x) + K(y/x) + \mathcal{O}(1)$.

Unfortunately K has been proved as a non-computable function [27]. But in fact, K or an upper bound of K can be estimated in diverse ways that we investigate now.

4.2. Complexity estimation

The first well known option available to estimate K is via lossless compression algorithm. For instance *bzip* approximates better than *gzip*, and the PAQ family is still better than *bzip*2. Due to the invariance theorem, when the size of s is huge, using compression will provide a relatively stable approximation as the constant c in the theorem can be considered as negligible. It is obviously not the case when the size of s becomes small. When s is short, compression is not a valid option. On another side, the constant c can prevent from providing stable approximations of $K(s)$. Luckily, the works of [7,8,43] give means of providing sensible values for the complexity of short strings (i.e. less than 10 bits). This job has been done by the Algorithmic Nature Group (https://algorithmicnature.org/). They have developed a tool OACC for Online Algorithmic Complexity Calculator (http://www.complexitycalculator.com/) allowing to estimate the complexity of short strings. The authors derived their approach from a Levin's theorem [26,8] establishing the exact connection between $m(s)$ and $K(s)$, where $m(s)$ is a semi-measure known as the Universal Distribution defined as follows [44]: $m(s) = \Sigma_{p:U(p,\epsilon)=s} 2^{-|p|}$.

Property 6. *There exists a constant c depending only on the underlying Turing machine such that: $\forall s, |-log_2(m(s)) - K(s)| < c$.*

Rewriting the formula as $K(s) = -log_2(m(s)) + \mathcal{O}(1)$, shows that estimating K could also be done via estimating $m(s)$. Estimating $m(s)$ becomes realistic when s is short as we have to estimate the probability for s to be the output of a short program. Considering simple Turing machines as described in [43], over a Boolean alphabet $\{0,1\}$ and a finite number n of states $\{1,\ldots,n\}$ plus a special Halt state denoted 0, there are exactly $(4n+2)^{2n}$ such Turing machines. Using clever optimizations [43], running these machines for $n=4$ and $n=5$ becomes realistic and provides an estimation of $m(s)$ and ultimately of $K(s)$. In the following, we denote $K'(s)$ this OACC estimation of $K(s)$.

4.3. Short chains complexity estimation

Some properties are expected from a complexity calculator machinery to be in accordance with a cognitive process:

1. There is no way to distinguish strings of length 1 and it is absolutely clear that $K(0) = K(1)$ should hold whatever the considered universal Turing machine.

Table 2
Complexity of 1 bit and 2 bits chains with OACC.

x	$K(x)$
0	3.5473880692782100
1	3.5473880692782100

x_1x_2	$K(x_1x_2)$
00	$5.4141012345247104 = a$
01	$5.4141040197301500 = b$
10	$5.4141040197301500 = b$
11	$5.4141012345247104 = a$

2. An important point is to be able to distinguish the 4 strings of length 2: $00, 11, 10, 01$ and we expect the following properties: $K(00) = K(11) < K(01) = K(10)$;
3. In terms of n bits strings, we expect $0\dots0$ and $1\dots1$ to be the simplest ones and to have the same complexity.

Observing the tables in [8], it appears that the properties above are satisfied, namely:

– Whatever the number of states of a 2-symbols Turing machine, $K'(0) = K'(1)$.

– Whatever the number of states of a 2-symbols Turing machine, $K'(00) = K'(11) = a$, $K'(01) = K'(10) = b$ and $a < b$.

– Whatever the number of states of a 2-symbols Turing machine, and for strings of length less than or equal to 10 (short strings) $K'(0\dots0) = K'(1\dots1) = a$ and a is the minimum value among the set of values.

Then the estimation of K via K' coming from the OACC estimator is a suitable candidate for our purpose. But before going further, we have first to check that OACC validate the above conditions. As we can check by examining Table 2 and column 4 of the final table in Section 4.5, these basic cognitive evidences are confirmed with the OACC tool. So we can start from OACC to check the properties required to validate the analogical hypothesis that we propose in the next section.

4.4. A Kolmogorov formula to measure the quality of an analogy

When stated with a machine learning perspective, the problem of analogical inference is as follows: for a given x_3, predict x_4 such that the *target* pair (x_3, x_4) is in the *same* relation that another given *source* pair (x_1, x_2) considered as an example. The pair (x_3, x_4) is the target pair which is partially known. In the case of classification where the 2nd element in a pair is the label, it amounts to predict the label of x_3 having only one classified example (x_1, x_2) at hand.

A functional view amounts to considering a hidden function f such that $x_2 = f(x_1)$ and we have to guess $x_4 = f(x_3)$. This functional view is the one developed in [5]: the problem of analogical inference strictly fits with a regression problem but with only one example. Ruling out any statistical models, this approach needs a brand new formalization that the authors extract from algorithmic complexity theory. Instead of trying to find regularities among a large set of observations (statistical approach), they consider the very meaning of each of the 3 observables $x_1, x_2 = f(x_1)$ and x_3. We start from this philosophy, but we depart from it as below:

– We focus on the Boolean case where the 3 objects under consideration are Boolean vectors. So we do not have to care about the change between the source domain representation and the target domain representation: these 2 domains are identical. The cost of this representation change is null in terms of algorithmic complexity.

– To be in line with the machine learning minimal assumption that there exists some unknown probability distribution P from which the data are drawn, we do not consider that x_2 is a (hidden) function of x_1. We just have a probability of observing x_2 having already observed x_1 which is more general than associating a fixed x_2 with every given x_1. It could be the case that for another x_2' we still have $x_1 : x_2' :: x_3 : x_4$.

As a consequence, we start from the following intuitions:

1. For $x_1 : x_2 :: x_3 : x_4$ to be accepted as a valid analogy, it is clear that the way we go from x_1 to x_2 should not be very different from the way we go from x_3 to x_4 (but it has not to be a functional link). We suggest to measure this expected proximity with the difference $|K(x_2/x_1) - K(x_4/x_3)|$. Considering $K(x_2/x_1)$ as the *difficulty* to build x_2 from x_1, the previous expression $|K(x_2/x_1) - K(x_4/x_3)|$, when small, tells us that it is not more difficult to build x_4 from x_3 than to build x_2 from x_1, and vice versa. This is what we call the atomic view of analogy. But this is obviously not enough.
2. In fact, the previous formula does not tell anything about the link between the pair (x_1, x_2) and the pair (x_3, x_4). For $x_1 : x_2 :: x_3 : x_4$ to be accepted as a valid analogy, the *difficulty* to apprehend the string x_1x_2 from the string x_3x_4 should be close to the difficulty to apprehend x_3x_4 from x_1x_2. We suggest to measure this expected proximity with the difference $|K(x_1x_2/x_3x_4) - K(x_3x_4/x_1x_2)|$. This difference is obviously symmetric and is linked to the symmetry of an analogy.
3. Above all, the global picture has to be "simple" i.e. telling that $x_1 : x_2 :: x_3 : x_4$ is a valid analogy should not be too disturbing, at least from a cognitive viewpoint. This means that the occurrence of the string $x_1x_2x_3x_4$ in this order should be highly plausible. We suggest to measure this plausibility with $K(x_1x_2x_3x_4)$ which is the size of the shortest program producing the binary string $x_1x_2x_3x_4$ from a universal Turing machine.

Following the ideas of [5], we use the sum as aggregator operator and denote $k(x_1x_2x_3x_4)$ the following formula measuring, in some sense, the quality of an analogy:

$$|K(x_2/x_1) - K(x_4/x_3)| + |K(x_1x_2/x_3x_4) - K(x_3x_4/x_1x_2)| + K(x_1x_2x_3x_4)$$

Using OACC http://www.complexitycalculator.com/

k(abcd) = |K(b/a) – K(d/c)| + |K(cd/ab) – K(ab/cd)| + K(abcd)

	A	B		C	D	
abcd	K(b/a)	K(d/c)	K(abcd)	k(cd/ab)	K(ab/cd)	|A-B|+|C-D| +K(abcd)
0000	1.8667131652	1.8667131652	11.2174683967	5.8033671621	5.8033671621	11.2174683967
1111	1.8667131652	1.8667131652	11.2174683967	5.8033671621	5.8033671621	11.2174683967
0101	1.8667159505	1.8667159505	11.7002793293	6.2861753096	6.2861753096	11.7002793293
1010	1.8667159505	1.8667159505	11.7002793293	6.2861753096	6.2861753096	11.7002793293
0001	1.8667131652	1.8667159505	11.5731249872	6.1590237527	6.3435937193	11.757697739
1110	1.8667131652	1.8667159505	11.5731249872	6.1590237527	6.3435937193	11.757697739
0111	1.8667159505	1.8667131652	11.5731249872	6.1590209675	6.3435965045	11.7577033094
1000	1.8667159505	1.8667131652	11.5731249872	6.1590209675	6.3435965045	11.7577033094
0011	1.8667131652	1.8667131652	11.8099819092	6.3958806747	6.3958806747	11.8099819092
1100	1.8667131652	1.8667131652	11.8099819092	6.3958806747	6.3958806747	11.8099819092
0100	1.8667159505	1.8667131652	11.757697739	6.3435937193	6.1590237527	11.9422704908
1011	1.8667159505	1.8667131652	11.757697739	6.3435937193	6.1590237527	11.9422704908
0010	1.8667131652	1.8667159505	11.757697739	6.3435965045	6.1590209675	11.9422760613
1101	1.8667131652	1.8667159505	11.757697739	6.3435965045	6.1590209675	11.9422760613
1001	1.8667159505	1.8667159505	12.0548412692	6.6407372495	6.6407372495	12.0548412692
0110	1.8667159505	1.8667159505	12.0548412692	6.6407372495	6.6407372495	12.0548412692

Fig. 2. OACC estimations.

This leads us to postulate that the "best" x_4 we are looking for to build a valid analogy $x_1 : x_2 :: x_3 : x_4$ is the one minimizing this expression. So, we have:

$$x_4 = argmin_u k(x_1 x_2 x_3 u)$$

Let us see if we can, at least from an empirical viewpoint, validate this model. Let us note that we can have several values minimizing such an expression.

4.5. Validation in the Boolean setting

As we are not in a position to prove something at this stage, let us just investigate now the empirical evidence for our formula. One point to start with is to check if this formula holds in the very basic Boolean case. Considering x_1, x_2, x_3, x_4 as Boolean values, we have to check how the 6 cases of valid analogical proportions actually behave w.r.t. the formula $k(x_1x_2x_3x_4)$. Thus, we have to estimate formula $k(x_1x_2x_3x_4)$ for every $x_1x_2x_3x_4 \in \mathbb{B}^4$. The point is that our strings are very short: only 4 bits. So, as explained in Sections 4.2 and 4.3, we have to rely on OACC instead of a compression estimation.

On top of that, we have to consider, not only pure Kolmogorov complexity K but also complexity w.r.t. a given string as in $K(x_3x_4/x_1x_2)$. Generally, it is quite clear that $K(xy) \leq K(x) + K(y/x)$: roughly speaking, we can build a program whose output is xy by concatenating a program whose output is x to a program taking x as input and providing y as output. It is more difficult to get a more precise bound. Thanks to Theorem 2: $K(xy) = K(x) + K(y/x) + \mathcal{O}(1)$, which shows that we can approximate $K(y/x)$ with $K(xy) - K(x)$. As we now have all the tools needed to approximate formula k, it remains to use OACC to compute the estimation. Fig. 2 reports the results of this computation. As can be seen for the 6 patterns of the model Ω_0 of analogical proportion, the unique solution of equation $a:b::c:x$ always corresponds to a string $abcx$ that minimizes expression k w.r.t. the other option $abc\overline{x}$ (where $\overline{x} = \neg x$), e.g. $k(1111) < k(1110)$. Besides 0101 is simpler than 0110 despite the fact that in the second case there is also an underlying function such that $x_2 = f(x_1)$ and $x_4 = f(x_3)$: the negation. Note that 0110 and 1001 exhibit the highest complexity as estimated by OACC. It eliminates Kl. As there is no known convergence result regarding K and that we cannot estimate the constant in the formula $K(s) = -log_2(m(s)) + \mathcal{O}(1)$, these experiences should only be considered as adding a bit of credibility to the smallest model.

Nevertheless, in real life application, attributes or features are not necessarily Boolean and a graded extension of analogical expression is needed. This is the object of the next section to investigate how to extend these formulas.

5. Extension to Boolean vectors

Representing objects with a single Boolean value is not generally sufficient to handle real situations, and items are usually represented by *vectors* of Boolean values, each component being the value of a binary attribute. So we have to extend the Boolean case to Boolean vectors. This can be done componentwise, as illustrated by the following example "a calf (A) is to a cow (B) as a foal (C) is to a mare (D)".[3] The attributes have been chosen in order to exhibit (vertically) each of the 6 patterns that make an analogical proportion true.

[3] Often used by Laurent Miclet in his oral presentations.

	mammal	young	equine	adult female	bovine	adult male
A: calf	1	1	0	0	1	0
B: cow	1	0	0	1	1	0
C: foal	1	1	1	0	0	0
D: mare	1	0	1	1	0	0

5.1. Analogical proportions between vectors

A way to move to Boolean vectors would be to extend the operators of Boolean calculus to vectors and to consider the derived definitions. With that perspective, we have in $\mathbb{B}^n$:

- $\neg\vec{a} =_{def} (\neg a_1, \ldots, \neg a_n)$
- $\vec{a} \vee \vec{b} =_{def} (a_1 \vee b_1, \ldots, a_n \vee b_n)$
- $\vec{a} \wedge \vec{b} =_{def} (a_1 \wedge b_1, \ldots, a_n \wedge b_n)$
- $\vec{a} \equiv \vec{b} =_{def} (a_1 \equiv b_1, \ldots, a_n \equiv b_n)$

The truth value of a Boolean vector $\vec{a}$ is just $\bigwedge_{i=1}^{n} a_i$. As a consequence, $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$ has a truth value which is just $\bigwedge_{i=1}^{n} a_i : b_i :: c_i : d_i$. It is quite easy to see that the three postulates of analogy are still valid with this vectorial definition. It is also the case for code independency. Moreover the equivalence between

$$(\vec{a} \wedge \neg\vec{b} \equiv \vec{c} \wedge \neg\vec{d}) \wedge (\neg\vec{a} \wedge \vec{b} \equiv \neg\vec{c} \wedge \vec{d})$$

and

$$(\vec{a} \wedge \vec{d} \equiv \vec{b} \wedge \vec{c}) \ \wedge \ (\vec{a} \vee \vec{d} \equiv \vec{b} \vee \vec{c})$$

is still valid. But if we consider the vectorial version of

$$((a \equiv b) \wedge (c \equiv d)) \vee ((a \equiv c) \wedge (b \equiv d)),$$

coming from Property 3 where $\vee$ is the top operator, it is definitely not equivalent to $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$. Let us start from an example to get the picture. With:

$$\vec{a} = (0, 1, 0), \vec{b} = (1, 1, 0), \vec{c} = (0, 1, 1), \vec{d} = (1, 1, 1),$$

we have $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$ (which is better visualized in Table 3) despite the fact that $((\vec{a} \equiv \vec{b}) \wedge (\vec{c} \equiv \vec{d})) \vee ((\vec{a} \equiv \vec{c}) \wedge (\vec{b} \equiv \vec{d}))$ does not hold. This comes from the fact that the top operator $\vee$ in the formula does not commute with the $\wedge$ operator involved in the truth value definition i.e.:

$$\bigwedge_{i=1}^{n}((a_i \equiv b_i) \wedge (c_i \equiv d_i)) \vee ((a_i \equiv c_i) \wedge (b_i \equiv d_i)) \not\Rightarrow$$
$$(\bigwedge_{i=1}^{n}(a_i \equiv b_i) \wedge \bigwedge_{i=1}^{n}(c_i \equiv d_i)) \vee (\bigwedge_{i=1}^{n}(a_i \equiv c_i) \wedge \bigwedge_{i=1}^{n}(b_i \equiv d_i))$$

which is the exact definition of:

$$((\vec{a} \equiv \vec{b}) \wedge (\vec{c} \equiv \vec{d})) \vee ((\vec{a} \equiv \vec{c}) \wedge (\vec{b} \equiv \vec{d}))$$

As a consequence, we will not consider the equivalences shown in Property 3 as valid candidates for any extension of analogical proportion.

This leads to rest on the initial equivalent expressions (1) and (2) as being the most suitable one to accurately capture the intuitive meaning of analogical proportion.

Table 3
Analogical proportion $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$ between vectors.

$\vec{a}$	0	1	0
$\vec{b}$	1	1	0
$\vec{c}$	0	1	1
$\vec{d}$	1	1	1

Table 4
Pairing pairs $(\vec{a}, \vec{b})$ and $(\vec{c}, \vec{d})$.

	$\mathcal{A}_1$	...	$\mathcal{A}_{i-1}$	$\mathcal{A}_i$	...	$\mathcal{A}_{j-1}$	$\mathcal{A}_j$	...	$\mathcal{A}_{k-1}$	$\mathcal{A}_k$	...	$\mathcal{A}_{r-1}$	$\mathcal{A}_r$	...	$\mathcal{A}_{s-1}$	$\mathcal{A}_s$	...	$\mathcal{A}_n$
$\vec{a}$	1	...	1	0	...	0	1	...	1	0	...	0	1	...	1	0	...	0
$\vec{b}$	1	...	1	0	...	0	1	...	1	0	...	0	0	...	0	1	...	1
$\vec{c}$	1	...	1	0	...	0	0	...	0	1	...	1	1	...	1	0	...	0
$\vec{d}$	1	...	1	0	...	0	0	...	0	1	...	1	0	...	0	1	...	1

Table 5
Pairing $(\vec{a}, \vec{d})$ and $(\vec{b}, \vec{c})$.

	$\mathcal{A}_1$	...	$\mathcal{A}_{i-1}$	$\mathcal{A}_i$	...	$\mathcal{A}_{j-1}$	$\mathcal{A}_j$	...	$\mathcal{A}_{k-1}$	$\mathcal{A}_k$	...	$\mathcal{A}_{r-1}$	$\mathcal{A}_r$	...	$\mathcal{A}_{s-1}$	$\mathcal{A}_s$	...	$\mathcal{A}_n$
$\vec{a}$	1	...	1	0	...	0	1	...	1	0	...	0	1	...	1	0	...	0
$\vec{d}$	1	...	1	0	...	0	0	...	0	1	...	1	0	...	0	1	...	1
$\vec{b}$	1	...	1	0	...	0	1	...	1	0	...	0	0	...	0	1	...	1
$\vec{c}$	1	...	1	0	...	0	0	...	0	1	...	1	1	...	1	0	...	0

5.2. Playing with analogical proportions between vectors

It is important to notice that the four vectors in an analogical proportion are of the same nature, since they refer to the same set of features. Then symmetry just means that comparing the results of the comparisons of the two vectors inside each pair of vectors $(\vec{a}, \vec{b})$ and $(\vec{c}, \vec{d})$ does not depend on the ordering of the two pairs. Thus the repeated applications of symmetry followed by central permutation yield 8 equivalent forms of the analogical proportion: $(\vec{a}:\vec{b}::\vec{c}:\vec{d}) = (\vec{c}:\vec{d}::\vec{a}:\vec{b}) = (\vec{c}:\vec{a}::\vec{d}:\vec{b}) = (\vec{d}:\vec{b}::\vec{c}:\vec{a}) = (\vec{d}:\vec{c}::\vec{b}:\vec{a}) = (\vec{b}:\vec{a}::\vec{d}:\vec{c}) = (\vec{b}:\vec{d}::\vec{a}:\vec{c}) = (\vec{a}:\vec{c}::\vec{b}:\vec{d})$. Table 4 pictures the situation, where the components of the vectors have been suitably reordered in such a way that the attributes for which one of the 6 patterns characterizing the analogical proportion is observed, have been gathered, e.g., attributes $\mathcal{A}_1$ to $\mathcal{A}_{i-1}$ exhibits the pattern 1111. In the general case, some of the patterns may be absent.

Table 4 shows that building the analogical proportion $\vec{a}:\vec{b}::\vec{c}:\vec{d}$ is a matter of pairing the pair $(\vec{a}, \vec{b})$ with the pair $(\vec{c}, \vec{d})$. More precisely, on attributes $\mathcal{A}_1$ to $\mathcal{A}_{j-1}$, the four vectors are equal; on attributes $\mathcal{A}_j$ to $\mathcal{A}_{r-1}$, $\vec{a} = \vec{b}$ and $\vec{c} = \vec{d}$, but $(\vec{a}, \vec{b}) \neq (\vec{c}, \vec{d})$; on attributes $\mathcal{A}_r$ to $\mathcal{A}_n$, $\vec{a} = \vec{c}, \vec{b} = \vec{d}$, and $\vec{a} \neq \vec{b}$. In other words, on attributes $\mathcal{A}_1$ to $\mathcal{A}_{r-1}$ $\vec{a}$ and $\vec{b}$ agree and $\vec{c}$ and $\vec{d}$ agree as well. This contrasts with attributes $\mathcal{A}_r$ to $\mathcal{A}_n$, for which we can see that $\vec{a}$ differs from $\vec{b}$ as $\vec{c}$ differs from $\vec{d}$ (and vice-versa). We recognize the meaning of the formal definition of the analogical proportion as described by expression (1).

Let us now pair the vectors differently, namely considering pair $(\vec{a}, \vec{d})$ and pair $(\vec{b}, \vec{c})$, as in Table 5. First, we can see that $\vec{a}:\vec{d}::\vec{b}:\vec{c}$ *does not hold* due to attributes $\mathcal{A}_s$ to $\mathcal{A}_n$. Obviously, we continue to have $\vec{a} = \vec{b} = \vec{c} = \vec{d}$ for attributes $\mathcal{A}_1$ to $\mathcal{A}_{j-1}$, while on the rest of the attributes the values inside each pair differ (in the four possible different ways). Then it should not come as a surprise that we recover, in the vectorial case, expression (2), since the expression holds either when $a = d = b = c$, or when $a \neq d$ and $b \neq c$:

$$((a \wedge d) \equiv (b \wedge c)) \wedge ((a \vee d) \equiv (b \vee c)) \tag{2}$$

5.3. Link with Hamming distance

Since we have a relation between elements of $\mathbb{B}^n$, namely the analogical proportion, it is quite natural to consider this relation with regard to standard metric on $\mathbb{B}^n$. A classical one over $\mathbb{B}^n$ is Hamming distance defined as follows:

$$H_n(\vec{x}, \vec{y}) = |\{i \in [1, n] : x_i \neq y_i\}|$$

Note that this distance satisfies:

$$H_{n+1}(\vec{x}, \vec{y}) = H_n(\vec{x}/n, \vec{y}/n) + H_1(x_{n+1}, y_{n+1})$$

where $\vec{x}/n$ denotes the projection of $\vec{x}$ over $\mathbb{B}^n$. We write H when there is no ambiguity about the dimension. A first link between analogical proportion and Hamming distance can be easily checked:

Property 7. *$\forall \vec{a}, \vec{b}, \vec{c}, \vec{d} \in \mathbb{B}^n$ such that $\vec{a}:\vec{b}::\vec{c}:\vec{d}$, we have:*
$H(\vec{a}, \vec{b}) = H(\vec{c}, \vec{d})$, $H(\vec{a}, \vec{c}) = H(\vec{b}, \vec{d})$, $H(\vec{a}, \vec{d}) = H(\vec{b}, \vec{c})$.

Proof. We prove the first equality $H(\vec{a}, \vec{b}) = H(\vec{c}, \vec{d})$ by induction over n. The equality is obviously satisfied for $n = 1$. Now we have to note that, due to the definition of analogy on $\mathbb{B}^{n+1}$, given $\vec{a}, \vec{b}, \vec{c}, \vec{d} \in \mathbb{B}^{n+1}$,

$$\vec{a}:\vec{b}::\vec{c}:\vec{d} \text{ iff } \vec{a}/n:\vec{b}/n::\vec{c}/n:\vec{d}/n \text{ and } a_{n+1}:b_{n+1}::c_{n+1}:d_{n+1}$$

The induction hypothesis leads to:

$$H(\vec{a}/n, \vec{b}/n) = H(\vec{c}/n, \vec{d}/n) \text{ and } H(a_{n+1}, b_{n+1}) = H(c_{n+1}, d_{n+1})$$

Adding these 2 equations leads to the expected result. The same reasoning applies to the remaining 2 equalities. □

The reverse property does not hold in general: for instance $0 : 1 :: 1 : 0$ does not hold. This comes from the fact that Hamming distance is symmetric but analogy does not allow to reverse only one side of the proportion: when $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$ holds, it is unlikely that $\vec{b} : \vec{a} :: \vec{c} : \vec{d}$ holds. Obviously, Property 7 does not hold in $\mathbb{R}^n$ when equipped with Euclidean distance: a parallelogram $\vec{a}, \vec{b}, \vec{c}, \vec{d}$ does not always satisfy $d(\vec{a}, \vec{d}) = d(\vec{b}, \vec{c})$ where d is the Euclidean distance. A second link can still be established:

Property 8. *$\forall \vec{a}, \vec{b}, \vec{c}, \vec{d}$, such that $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$, we have:*
$H(\vec{a}, \vec{b}) + H(\vec{a}, \vec{c}) = H(\vec{a}, \vec{d})$ and similarly: $H(\vec{d}, \vec{c}) + H(\vec{d}, \vec{b}) = H(\vec{d}, \vec{a})$.

Proof. The first equality can be easily proved by induction over n. The property is true for $n = 1$. Since we have

$$\vec{a} : \vec{b} :: \vec{c} : \vec{d} \text{ iff } \vec{a}/n : \vec{b}/n :: \vec{c}/n : \vec{d}/n \text{ and } a_{n+1} : b_{n+1} :: c_{n+1} : d_{n+1}$$

we can apply the induction hypothesis to get $H(\vec{a}/n, \vec{b}/n) + H(\vec{a}/n, \vec{c}/n) = H(\vec{a}/n, \vec{d}/n)$ and $H(a_{n+1}, b_{n+1}) + H(a_{n+1}, c_{n+1}) = H(a_{n+1}, d_{n+1})$. And we are done by adding these 2 equations. Then the second equality is just deduced from the first one using Property 7. □

These results strengthen the geometrical interpretation of analogical proportion as a parallelogram in the suitable vector space, viewing a, b, c, d as points in $\mathbb{R}^n$ and vectors defined by pairs (a, b), (a, c), (a, d), etc.

5.4. Equations and the induction of analogical proportions by comparing two items

It is also interesting to consider analogical proportions from an equational point of view, as done by S. Klein [20] in his pioneering work. He was the first to propose a way to solve the equation between Boolean vectors $\vec{a} : \vec{b} :: \vec{c} : \vec{x}$ where $\vec{x}$ is unknown.

In $\mathbb{B}$, the equation $a : b :: c : x$ has a unique solution $x = c \equiv (a \equiv b)$ provided that $(a \equiv b) \vee (a \equiv c)$ holds. Indeed neither $0 : 1 :: 1 : 0$ nor $1 : 0 :: 0 : 1$ holds true in the minimal model of analogy, departing from Klein's model with which the equation has always a solution. This process can be extended componentwise to vectors. In that case, for instance, the following equation $010 : 100 : 011 : \vec{x}$ has for unique solution the vector $(1, 0, 1)$ which is not among the three previous vectors $(0, 1, 0), (1, 0, 0), (0, 1, 1)$. We can also observe in Table 1 that analogical proportions for vectors are *creative* (an informal quality usually associated with the idea of analogy) as they may involve 4 distinct vectors. Namely, one may have $\vec{a}$, $\vec{b}$, $\vec{c}$ all different while $\forall i \in [1, n], d_i = c_i \equiv (a_i \equiv b_i)$ and $\exists i\ d_i \neq a_i$, $\exists j\ d_j \neq b_j$, and $\exists k\ d_k \neq c_k$. For instance, in Table 1, we can take $(d_1, d_2) = (0, 0), with (a_1, a_2) = (1, 1), (b_1, b_2) = (1, 0), (c_1, c_2) = (0, 1)$.

Another equation of interest is $\vec{a} : \vec{x} :: \vec{y} : \vec{d}$ where $\vec{x}$ and $\vec{y}$ are unknown. As we shall see, starting from two Boolean vectors $\vec{a}$ and $\vec{d}$, one can find two "intermediary" vectors such that the equation holds. This shows how pervasive the concept of analogical proportion is, since it is implicitly present as soon as we compare two vectors $\vec{a}$ and $\vec{d}$. Let us first consider the equation in the Boolean set $\mathbb{B}$. As soon as we get two elements $a, d \in \mathbb{B}$, it is obvious that we can find two other elements (not necessarily distinct) such that $a : b :: c : d$, the pair (b, c) establishing a bridge between a and d. For instance to bridge 0 and 1 we can still use 0 and 1 to build $0 : 0 :: 1 : 1$ or 1 and 0 to build $0 : 1 :: 0 : 1$.

Consider now two *distinct* Boolean vectors $\vec{a}$ and $\vec{d}$ in $\mathbb{B}^n$ (if $\vec{a} = \vec{d}$, the solution is trivially $\vec{x} = \vec{y} = \vec{a}$). It is then possible to find two other vectors $\vec{b}$ and $\vec{c}$ such that $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$. Beyond the easy solutions $\vec{x} = \vec{a}$ and $\vec{y} = \vec{d}$ (or $\vec{x} = \vec{d}$ and $\vec{y} = \vec{a}$), it is generally possible to find two distinct vectors $\vec{b}$ and $\vec{c}$, themselves being distinct from $\vec{a}$ and $\vec{d}$, such that $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$ holds.

Indeed, let $Agr(\vec{a}, \vec{d})$ be the set of indices where $\vec{a}$ and $\vec{d}$ agree and $Dis(\vec{a}, \vec{d})$ the set of indices where the two vectors differ. Namely $Agr(\vec{a}, \vec{d}) = \{i \mid a_i = d_i\}$ and $Dis(\vec{a}, \vec{d}) = \{i \mid a_i \neq d_i\}$. Clearly, the two solution vectors $\vec{b}$ and $\vec{c}$ of equation $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$ should be such that

$$\forall i \in Agr(\vec{a}, \vec{d}), a_i = b_i = c_i = d_i$$

(all equal to 1 or all equal to 0). For the components in $Dis(\vec{a}, \vec{d})$ we have two options for preserving an analogical proportion componentwise

$$\forall i \in Dis(\vec{a}, \vec{d})(b_i = a_i \text{ and } c_i = d_i) \text{ or } (b_i = \neg a_i \text{ and } c_i = \neg d_i).$$

For instance, take $\vec{a} = (0, 1, 1, 0), \vec{d} = (0, 0, 1, 1)$. We have $Agr(\vec{a}, \vec{d}) = \{1, 3\}$ and $Dis(\vec{a}, \vec{d}) = \{2, 4\}$. Then $\vec{b} = (0, 1, 1, 1)$ and $\vec{c} = (0, 0, 1, 0)$ make $\vec{a} : \vec{b} :: \vec{c} : \vec{d}$ true. This is a solution among others.

Property 9. *As soon as* $Dif(\vec{a}, \vec{d})$ *contains at least two indices, there are solutions of equation* $\vec{a} : \vec{x} :: \vec{y} : \vec{d}$ *where the four vectors* $\vec{a}, \vec{x}, \vec{y}, \vec{d}$ *are distinct.*

This is exemplified by the above example. The creation of $(\vec{b}, \vec{c})$ from $\vec{a}$ and $\vec{d}$ has been recently illustrated in [19] on images, using a non-Boolean logic approach. Their work deals with the transfer of various visual attributes like color, texture, style, etc., from one image a to another image d. Using analogical proportions, the authors are the first to suggest a general method able to deal with any type of attributes. To transfer attributes from an image a to an image d, their solution is to formulate the mapping as a problem of building an analogical proportion between images $a : b :: c : d$, where b and c are unknown and have to be estimated. The assumption is that $b = \phi(a)$ and $d = \phi(c)$. To estimate the operator ϕ, the authors consider the features extracted from several convolutional neural networks. At the end of the process, the operator ϕ will be applied to c in order to get d.

Some well-known apps or services like Prisma (https://prisma-ai.com/), Google Deep Style (http://www.deepstylegenerator.com/), Ostagram (https://www.ostagram.me/) are able to transform photos and videos into works of art using the styles of famous artists like Van Gogh or Picasso. They mainly use a mix of classical AI with convolutional neural networks [11].

In comparison, the analogical approach of [19] is capable of higher quality content-specific stylization that better preserves structures.

6. Multiple-valued analogical proportions: A refresher

The aim of this section is mainly to provide the necessary background for the next section on analogical inequalities since they are weakened forms of Boolean or multiple-valued analogical proportions. It is an opportunity to emphasize the influence of the two equivalent syntactic forms of analogical proportion found meaningful in the Boolean case which lead to two multiple-valued extensions that are no longer equivalent. It is also the occasion to discuss the particular case of continuous analogical proportions, where the two central elements are equal, which are not often considered.

We assume that attributes are now valued in $[0, 1]$, possibly after renormalization, rather than in $\mathbb{B}$. Then we expect that the graded extension of $a : b :: c : d$ from $\mathbb{B}^4$ to $[0, 1]^4$ takes intermediary values between 0 and 1 (e.g., $0.9 : 0 :: 1 : 0$ can be neither 1 nor 0, but should rather have a high value since 0.9 is close to 1. For simplicity, in the sequel, a symbol a denotes both a variable and its truth-value. Note that using a simple difference is not a satisfactory option since $[0, 1]$ is not closed under the difference operator. Using absolute value of difference will not solve the problem; indeed we lose the orientation part of the information, since $|a - b| = |c - d|$ does not entail $a - b = c - d$. Since the Boolean model Ω_0 is the one that we choose for analogical proportion, we can consider that the Boolean formulas (1) and (2) given in subsection 3.4 are a good starting point to be extended over graded truth values using multiple-valued connectives. Let us recall these two formulas below:

$$(a \wedge \neg b \equiv c \wedge \neg d) \wedge (\neg a \wedge b \equiv \neg c \wedge d) \tag{1}$$

$$(a \wedge d \equiv b \wedge c) \;\wedge\; (a \vee d \equiv b \vee c) \tag{2}$$

6.1. Conservative and liberal extensions

Despite the fact that (1) and (2) are equivalent in Boolean logic, it will appear that this is not the case in a multiple-valued framework. The multiple-valued logic extension of the two formulas requires the choice of connectives for the external conjunction and the two equivalences. As described in [40,35,9], extending these formulas can be done in the following way:

i) the central, external conjunction $\wedge$ is taken as equal to the minimum for simplicity.
ii) $s \equiv t$ is taken as $\min(s \rightarrow_L t, t \rightarrow_L s)$ where $\rightarrow_L$ is Łukasiewicz implication, defined by $s \rightarrow_L t =_{def} \min(1, 1 - s + t)$, for $\mathcal{L} = [0, 1]$ and thus $s \equiv t =_{def} 1 - |s - t|$. There are two arguments in favor of this choice. First, with this definition, $s \equiv t$ takes the truth value 1 if and only if $s = t$, which is fully in the spirit of exact comparisons; moreover this index is explicitly related to the usual distance between numerical values (expressed by the absolute value of the difference between s and t).
iii) replacing the four expressions of the form $s \wedge \neg t$ by the bounded difference $\max(0, s - t) =_{def} 1 - (s \rightarrow_Ł t)$, which is associated to Łukasiewicz implication, using $1 - (\cdot)$ as negation.

This being said, we can now extend the two formulas (1) and (2). We start with the last one.

A straightforward extension of expression (2) is obtained by also taking minimum for the internal conjunction and maximum for the internal disjunction [35]. This is referred to as the *conservative* extension [9]. For instance $a \wedge d$ leads to $min(a, d)$ and $(a \wedge d \equiv b \wedge c)$ to $1 - |\min(a, d) - \min(b, c)|$. Globally, this yields:

$$a : b ::_C c : d = \min(1 - |\max(a, d) - \max(b, c)|, 1 - |\min(a, d) - \min(b, c)|) \tag{3}$$

It obviously coincides with $a:b::c:d$ on $\{0,1\}$ and code independency is preserved under the form $a:b::_C c:d = (1-a):(1-b)::_C (1-c):(1-d)$. Then it is clear that:

Property 10. *$a:b::_C c:d = 1$ if and only if $\min(a,d) = \min(b,c)$ and $\max(a,d) = \max(b,c)$.*

It means that only patterns of the form (x,y,x,y) or (x,x,y,y) where $x,y \in [0,1]$ and possibly $x=y$ make $a:b::_C c:d$ fully true, generalizing the six cases of Table 1 and replacing 0 and 1 by x and y. In these patterns, a,b,c,d take values on a binary set $\{x,y\} \subset [0,1]$ only, and the degree of change from a to b, if any, must be exactly the same as the one from c to d, in *direction* (the first requirement enforcing the same amplitude of change). This is clearly a conservative view of graded analogy that remains close in spirit to the Boolean case.

Besides, it is obvious as well from the expression (3) that:

Property 11. *$a:b::_C c:d = 0$ if and only if $|\min(a,d) - \min(b,c)| = 1$ or $|\max(a,d) - \max(b,c)| = 1$.*

In other words, the only patterns fully falsifying the analogical proportion are then of the form $1:0::_C x:1$ or $0:1::_C x:0, \forall x \in [0,1]$ (and the other patterns obtained from these two by symmetry and central permutation). Thus, $a:b::_C c:d = 0$ if and only if there is a maximal difference between a and b, while a and d are equal, whatever the remaining term (and the like for ones obtained by symmetry and central permutation). We can see that the two extreme patterns of the form $1:0::_C x:1$, namely $1:0::0:1$ and $1:0::1:1$, have truth-value 0. This confirms that it is a conservative extension of the false Boolean cases.

Transitivity, as expressed in the following property, is still valid:

Property 12. *$a:b::_C c:d = 1$ and $c:d::_C e:f = 1$ implies $a:b::_C e:f = 1$.*

Proof. The first thing to see is that following Property 10, $a:b::_C c:d = 1$ is equivalent to $\{a,d\} = \{b,c\}$. Then, we start from $a:b::_C c:d = 1$ and $c:d::_C e:f = 1$. We get that $\{a,d\} = \{b,c\}$ and $\{c,f\} = \{d,e\}$. Since $\{a,d\} = \{b,c\}$, we distinguish two cases:

- $c = d$: Then necessarily $a = b$ and $e = f$: in that case $a:b::e:f$ holds and implies $a:b::_C e:f = 1$.
- $c = a$: In that case, $b = d$, then $a = c$ and $\{c,f\} = \{a,f\} = \{d,e\} = \{b,e\}$ which means $a:b::_C e:f = 1$. □

We now consider the multiple-valued extension of formula (1), corresponding to the so-called *liberal* view. It is given by [40,9]:

$$a:b::_Ł c c:d = \begin{cases} 1- \mid (a-b)-(c-d) \mid, \\ \quad \text{if } a \geq b \text{ and } c \geq d, \text{ or } a \leq b \text{ and } c \leq d \\ 1 - \max(|a-b|, |c-d|), \\ \quad \text{if } a \leq b \text{ and } c \geq d, \text{ or } a \geq b \text{ and } c \leq d \end{cases} \tag{4}$$

It coincides with $a:b::c:d$ on $\{0,1\}$. Because $|a-b| = |(1-a)-(1-b)|$, it is easy to prove that code independency still holds under the form: $a:b::_Ł c:d = (1-a):(1-b)::_Ł (1-c):(1-d)$.

Property 13. *$a:b::_Ł c:d = 1$ if and only if $a-b = c-d$.*

For instance, $0.2:0.4::_L 0.6:0.8 = 0.2:0.4::_L 0.4:0.6 = 1$. This contrasts with the conservative definition for which $0.2:0.4::_C 0.6:0.8 = 0.2:0.4::_C 0.4:0.6 = 0.8$. This means that the analogical proportion is valid when the changes from a to b and from c to d have the same direction and amplitude, and there is no longer any limitation on the number of distinct values a,b,c,d involved in fully true analogical pattern.

Property 14. *$a:b::_Ł c:d = 0$ if and only if*
i) $a-b = 1$ and $c \leq d$, or ii) $b-a = 1$ and $d \leq c$, or iii) $a \leq b$ and $c-d = 1$, or iv) $b \leq a$ and $d-c = 1$.

Thus, $a:b::_Ł c:d = 0$ only when the change inside one of the pairs (a,b) or (c,d) is maximal, while the other pair shows either no change or a change in the opposite direction.

It is an immediate consequence of Property 13, that this liberal view of analogical relation is transitive:

$$a:b::_Ł c:d = 1 \text{ and } c:d::_Ł e:f = 1 \text{ implies } a:b::_Ł e:f = 1$$

6.2. Continuous analogical proportions

A particular case of analogical proportions are *continuous* analogical proportions, which are statements of the form "a is to b as b is to c". It can be easily seen that the unique solutions of equations $1 : x :: x : 1$ and $0 : x :: x : 0$ are respectively $x = 1$ and $x = 0$, while $1 : x :: x : 0$ or $0 : x :: x : 1$ have no solution in the Boolean case. So continuous analogical proportions are trivialized in the Boolean setting. This is no longer the case in the multiple-valued case if we choose the proper extension.

Indeed let us consider the particular case of continuous analogical proportions in the two extensions. Let us first try the conservative one. We get

$$a : b ::_C b : c = \min(1 - |\max(a, c) - b|, 1 - |\min(a, c) - b|).$$

As a consequence $a : b ::_C b : c = 1$ if and only if $a = b = c$, which confirms the closeness of the conservative extension with the Boolean case, and makes it unsuitable for our purpose. Now let us consider the liberal extension. We get

$$a : b ::_Ł b : c = \begin{cases} 1- \mid a + c - 2b \mid, \\ \quad \text{if } a \geq b \text{ and } b \geq c, \ \text{ or } a \leq b \text{ and } b \leq c \\ 1 - \max(\mid a - b \mid, \mid b - c \mid), \\ \quad \text{if } a \leq b \text{ and } b \geq c, \ \text{ or } a \geq b \text{ and } b \leq c \end{cases} \tag{5}$$

Now $a : b ::_Ł b : c = 1$ if and only if $b = (a + d)/2$. This confirms that this extension captures the idea of betweenness implicit in statements of the form "a is to b as b is to c". In fact, the equation $a : x :: x : c$ expresses a way of "creating" of an intermediary item between a and c. This can be illustrated by the well-known example of a centaur, an imaginary being half man, half horse corresponding to the continuous analogical proportion "man is to centaur as centaur is to horse" (i.e., centaur is a *mid* term between man and horse). Indeed centaur can be seen as a solution of equation $\vec{man} : \vec{x} :: \vec{x} : \vec{horse}$, as can be checked on the following simplified two-component `human-like` and `horse-like` representation:

	human-like	*horse-like*
$\vec{man}$	1	0
$\vec{centaur}$	1/2	1/2
$\vec{centaur}$	1/2	1/2
$\vec{horse}$	0	1

where $1 : 1/2 ::_Ł 1/2 : 0$ and $0 : 1/2 ::_Ł 1/2 : 1$ obviously hold.

In this example, 'centaur' is exactly in the middle between 'man' and 'horse'. One might like to express statements of the form "a is more to b than b is to c". This would be a particular case of analogical inequality. In the following section, we investigate a general way to move from analogical proportions to analogical inequalities. We first start with the Boolean case.

7. Analogical proportions: from equality to inequality

In this section, we propose a logical modeling for expressions of the form "a is to b at least as much as c to d", first in the Boolean case, and then in the multiple-valued case. We denote this expression by $a : b \ll c : d$. We take inspiration from complete orderings, such as the one on $\mathbb{R}$, where $a = b$ is equivalent to the conjunction $(a \leq b) \wedge (b \leq a)$. We can then see the ordering relation as a weakening of the equality. In the logical setting, equality is replaced with equivalence (sameness) and weakening equivalence is done with implication since we have:

$$a \equiv b \text{ is equivalent to } (a \to b) \wedge (b \to a).$$

7.1. Boolean case

Starting from the Boolean expression (1) of analogical proportion, we replace the two symbols $\equiv$ expressing sameness by two material implications $\to$ for modeling the fact that the result of the dissimilarity of c and d is larger or equal to the result of the comparison of a and b. Namely, we obtain

$$a : b \ll c : d =_{def} ((a \wedge \neg b) \to (c \wedge \neg d)) \wedge ((\neg a \wedge b) \to (\neg c \wedge d)) \tag{6}$$

It is easily derivable from this definition that the following expected properties hold:

- $a : b \ll a : b$
- $a : b :: c : d \to a : b \ll c : d$
- $a : b :: c : d \equiv ((a : b \ll c : d) \wedge (c : d \ll a : b))$
- $(a : b \ll c : d) \equiv (\neg a : \neg b \ll \neg c : \neg d)$

Table 6
Boolean valuations for $a : b \ll c : d$.

a	b	c	d	$a : b \ll c : d$	a	b	c	d	$a : b \ll c : d$
0	0	0	0	**1**	1	0	0	0	0
0	0	0	1	**1**	1	0	0	1	0
0	0	1	0	**1**	1	0	1	0	**1**
0	0	1	1	**1**	1	0	1	1	0
0	1	0	0	0	1	1	0	0	**1**
0	1	0	1	**1**	1	1	0	1	**1**
0	1	1	0	0	1	1	1	0	**1**
0	1	1	1	0	1	1	1	1	**1**

Table 7
Boolean valuations for sum and product-based inequalities between pairs (a, d) and (b, c) and for $a : b \ll_C c : d$.

a	b	c	d	Table 6	$+$	$\cdot$	m/M	a	b	c	d	Table 6	$+$	$\cdot$	m/M
0	0	0	0	**1**	1	1	1	1	0	0	0	0	0	1	*1*
0	0	0	1	**1**	0	1	1	1	0	0	1	0	0	0	0
0	0	1	0	**1**	1	1	*0*	1	0	1	0	**1**	1	1	1
0	0	1	1	**1**	1	1	1	1	0	1	1	0	0	0	0
0	1	0	0	0	1	1	0	1	1	0	0	**1**	1	1	1
0	1	0	1	**1**	1	1	1	1	1	0	1	**1**	0	0	*0*
0	1	1	0	0	1	1	0	1	1	1	0	**1**	1	1	1
0	1	1	1	0	1	1	*1*	1	1	1	1	**1**	1	1	1

Namely, $a : b \ll c : d$ is weaker than $a : b :: c : d$, while $a : b :: c : d$ holds if and only if both $a : b \ll c : d$ and $c : d \ll a : b$ hold; moreover, code independency is preserved.

As expected, a kind of transitivity also holds:

Property 15. $(a : b \ll c : d) \wedge (c : d \ll e : f) \rightarrow (a : b \ll e : f)$.

Proof. It follows from Definition 6 applying the commutativity of $\wedge$ and the transitivity of $\rightarrow$. □

The relation $\ll$ considered as a relation over pairs of elements, i.e. a relation over $\mathbb{B}^2$, is a pre-order but not an order since anti-symmetry does not hold.

The truth table of $a : b \ll c : d$ is given in Table 6. As can be seen $a : b \ll c : d$ holds true for the 6 patterns that makes analogical proportion true, plus the 4 patterns 0001, 0010, 1110, 1101. These latter patterns correspond to the 4 situations where $a \equiv b$ and $c \not\equiv d$. In these 4 situations a and b are indeed strictly closer than c and d, and these are the only cases in $\{0, 1\}$. Since the 4 situations where $a \equiv b$ and $c \equiv d$ are among the patterns making $a : b :: c : d$ true, we have

$$a : b \ll c : d \equiv (a : b :: c : d) \vee (a \equiv b) \tag{7}$$

It is also worth noticing that the central permutation property of analogical proportion now fails since 0010 and 1101 are true while 0100 and 1011 are false. This may be unexpected at first glance since the arithmetic proportion inequality, $a - b \leq c - d$, equivalent to $a + d \leq b + c$, satisfies central permutation in the numerical case. Similarly, $a \cdot d \leq b \cdot c$, associated to geometric proportion inequality, satisfies central permutation. However, they do not fully agree together, nor with Table 6, as can be seen in Table 7 where columns '$+$' and '$\cdot$' stand respectively for conditions:

(i) $a + d \leq b + c$, and
(ii) $a \cdot d \leq b \cdot c$.

Note also that conditions (i) and (ii) are not code independent. In particular, (i) or (ii) may hold true, while $a : b \ll c : d$ is false. This the case for patterns 0100, 0110, 0111 for both (i) and (ii), and for 1000 for (ii). This simply means that $a : b \ll c : d$ as defined here is not a genuine counterpart of (i) or (ii) on $\mathbb{R}$. Columns 'm/M' will be commented in the next subsection.

Lastly, note that the quaternary relation $a : b \ll c : d$ induces a ternary relation (just as a continuous analogical proportion of the form $a : b :: b : c$ is a particular case of analogical proportion [42]). It can be seen that $a : b \ll b : c$ is true only for the four patterns 0000, 0001, 1110 and 1111, and false for the four other patterns. It expresses that the dissimilarity between b and c is greater or equal to the one between a and b.

7.2. Multiple-valued case

Both the conservative and the liberal multiple-valued extensions of analogical proportion might have counterparts for multiple-valued analogical inequalities. Let us examine these two options.

Conservative multiple-valued analogical inequalities. Equations (1) (and (3)) suggest that $a:b \ll c:d$ should hold true (at degree 1) as soon as the following two conditions hold:

$$\text{(iii) } \min(a,d) \le \min(b,c) \text{ and } \max(a,d) \ge \max(b,c)$$

(since $\min(1-a, 1-d) \le \min(1-b, 1-c) \Leftrightarrow \max(a,d) \ge \max(b,c)$). Note that still condition (iii) is code independent. However, conditions (iii) do not agree with Table 6, as can be seen in Table 7 where columns 'm/M' stand for these conditions (iii). Indeed (iii) yields results that do not fit with Table 6 on 4 patterns: 0010, 1101 where (iii) is false, and 0111 and 1000 where (iii) is true. The last two patterns are especially troublesome since they do not seem to fit at all with the intuitive reading "a is to b at least as much as c to d". In case we would try the direct counterpart of Equations (2), i.e.,

$$\text{(iv) } \min(a,d) \le \min(b,c) \text{ and } \max(a,d) \le \max(b,c),$$

we would be also in trouble since (iv) is true for patterns 0100 et 0111 while $a:b \ll c:d$ is false for them; moreover (iv) is not code independent.

Although, one may propose an inequality version of Definition 3 as a graded counterpart of (iii), namely

$$a:b \ll_C c:d = \min(1 - \min(0, \min(a,d) - \min(b,c)), 1 - \min(0, \max(b,c) - \max(a,d))),$$

the meaning of this expression would be debatable due to the above-mentioned unsatisfactory behavior on 0111 and 1000. We now consider the other option.

Liberal multiple-valued analogical inequalities. Using (4), the expression (6) can be extended to the multiple-valued case, still keeping min for extending the central $\wedge$, $1-|s-t|$ for the $\equiv$ symbol, and the four expressions of the form $s \wedge \neg t$ as the bounded difference $\max(0, s-t)$. The resulting expression is then

$$a:b \ll_{Ł} c:d = \begin{cases} \min(1, 1-((b-a)-(d-c)) & \text{if } a \le b \text{ and } c \le d \\ \min(1, 1-((a-b)-(c-d)) & \text{if } a \ge b \text{ and } c \ge d \\ 1-(b-a) & \text{if } a \le b \text{ and } c \ge d \\ 1-(a-b) & \text{if } a \ge b \text{ and } c \le d \end{cases} \quad (8)$$

This expression coincides with $a:b \ll c:d$ and Table 6 on $\mathbb{B}^4$. Thus $a:b \ll_{Ł} c:d = 1$ can be read "c is more dissimilar from d than a is from b".

It is worth noticing that $a:b \ll_{Ł} c:d$ does not exactly amount at comparing absolute value distances, as in the constraint $|a-b| \le |c-d|$. Indeed it can be checked that we may have $a:b \ll_{Ł} c:d = 0$, while $|a-b| \le |c-d|$ holds (taking $a=d=0$ and $b=c=1$). Moreover $a:b \ll_{Ł} c:d$ provides a graded estimate of the extent to which the numerical constraint $a-b \le c-d$ is satisfied.

It can be checked that the following expected properties still hold

- $a:b \ll_{Ł} c:d = a:b \ll c:d$ when $a,b,c,d \in \{0,1\}$;
- $a:b \ll_{Ł} a:b = 1$;
- $a:b ::_{Ł} c:d \le a:b \ll_{Ł} c:d$;
- $a:b ::_{Ł} c:d = \min((a:b \ll_{Ł} c:d), (c:d \ll_{Ł} a:b))$;
- $(a:b \ll_{Ł} c:d) = ((1-a):(1-b) \ll_{Ł} (1-c):(1-d))$ (code independency).

In particular, $a:b \ll_{Ł} c:d = 1$ if and only if

- $a=b$, or
- $|b-a| \le |d-c|$ if $a \le b$ and $c \le d$, or $b \le a$ and $d \le c$.

A kind of transitivity (similar to the Boolean case) holds for this graded relation.

Property 16. *Transitivity holds for $\ll_{Ł}$ relation, i.e.,*

$$(a:b \ll_{Ł} c:d = 1) \wedge (c:d \ll_{Ł} e:f = 1) \implies (a:b \ll_{Ł} e:f = 1)$$

Proof. Using the above characterization of $a:b \ll_{Ł} c:d = 1$, we have different cases to consider.

1. Case $a=b$: in that case, $a:b \ll_{Ł} e:f = 1$ obviously holds.
2. Case where $|b-a| \le |d-c|$ and $a \le b$ and $c \le d$. Since $(c:d \ll_{Ł} e:f = 1)$, we necessarily have $|d-c| \le |f-e|$ and $e \le f$. We deduce $|b-a| \le |f-e|$ and $a \le b$ and $e \le f$, which is the exact definition of $a:b \ll_{Ł} e:f = 1$.
3. Case where $|b-a| \le |d-c|$ and $b \le a$ and $d \le c$: a similar reasoning applies to show that still $a:b \ll_{Ł} e:f = 1$. □

Moreover $a : b \ll_{Ł} c : d = 0$ if and only if

- $|b - a| = 1$ and $|d - c| = 0$, or
- $b - a = 1$ and $c \geq d$, or
- $a - b = 1$ and $c \leq d$.

Lastly, continuous analogical inequalities define the following graded comparative ternary relation:

$$a : b \ll_{Ł} b : c = \begin{cases} \min(1, 1 + (a + c) - 2b) & \text{if } a \leq b \leq c \\ \min(1, 1 + 2b - (a + c)) & \text{if } a \geq b \geq c \\ 1 - (b - a) \quad \text{if } a \leq b \text{ and } b \geq c \\ 1 - (a - b) \quad \text{if } a \geq b \text{ and } b \leq c \end{cases} \tag{9}$$

Note that $a : b \ll_{Ł} b : c = 1$ if and only if $a = b$ or $b \leq (a + c)/2$ (resp. $b \geq (a + c)/2$) if $a \leq b \leq c$ (resp. $a \geq b \geq c$), i.e., if and only if b is closer (in the broad sense) to a than to c. It means that the difference between b and c is greater or equal to the one between a and b and the differences are oriented in the same way (when non-zero).

Besides, all the definitions introduced in this section apply to a single attribute. Just as in the case of the analogical proportion, the definitions of analogical equality in the Boolean and in the multiple-valued cases could be extended to multiple attribute descriptions by applying them in a component-wise manner, attribute per attribute. If necessary, a global evaluation may be obtained by taking the $\wedge$ or the min of the truth values obtained for each considered attribute.

7.3. Related work about analogical inequality

We have formally introduced analogical inequalities in the two previous subsections. They aim to modeling statements of the form "*a* is to *b* at least as much as *c* to *d*". Examples of similar statements are not many in the literature. Formal models are still rarer. In the introduction we mention the fact that in multiple criteria analysis one may need to express that the "difference" between two evaluation vectors on a criterion is smaller than the "difference" between the vectors on the rest of the criteria [34]. Still in this work, such statements are not represented as such. In the following, we discuss other works, where a restricted model is put in use.

Potential interest for analogical inequalities might be rather found in the area of image classification. Indeed the authors of [23,24] investigate the use of constraints resembling analogical inequalities to learn a distance d between images by considering discriminative dissimilarity constraints when rich information between data is available. The novelty of this approach (the so-called Qwise approach) is that they consider quadruplet of images instead of focusing on pairs or triplets of images in their discriminative constraints, namely "image a is closer to image b than c is to image d" which is formalized as $d(a, b) \leq d(c, d)$. Then the issue is to learn d. They make use of standard convex optimization techniques for approximating d. The authors consider a training set TS of quadruplets images (a, b, c, d) supposed to satisfy $d(a, b) \leq d(c, d)$ and the distance to be learned is expressed as depending on a matrix ω in such a way that $d_{\omega}(a, b) = \omega^T \phi(a, b)$, where ϕ is considered as the difference $x - y$ where x (resp. y) stands for the representation of image a (resp. b). The final learning scheme of parameters ω is not far from an SVM scheme except that the loss function deals with quadruplets. The learned distance is then used with a k-nearest neighbor-like algorithm for classification purposes. The authors experiment Qwise on different datasets and their results show that they improve standard techniques of about 3% in terms of classification accuracy.

It should be clear that the behavior of a constraint such as $d(a, b) \leq d(c, d)$, is not exactly the same as the one of $a : b \ll_{Ł} c : d$. Indeed $d(a, b) = d(b, a)$, while a and b are not exchangeable in our definition. One may think that this may matter. When comparing the members of pairs (a, b) and (c, d) with a relation which is a matter of degree, saying that the difference between a and b is less than the difference between c and d is not the same as saying that the difference between b and a is less than the difference between c and d. Still this has to be experimented in practice.

Interestingly enough, another proposal based on analogical proportions, instead of analogical inequalities, but still dealing with image classification can be found in a work published about the same time [18]. It appears that analogical proportions can bring valuable information beyond the class labels themselves, as soon as we have some extra knowledge about the classes semantic relationships. Let us consider an example of the authors to catch up the idea. Analogical proportion *leopard : cat = wolf : dog* is valid since, in both pairs (*leopard, cat*) and (*wolf, dog*), the first class *lives in the wild*, *has fangs*, and *is more aggressive*, etc. But, in terms of image classification, it can be easier to distinguish between *dog* and *wolf* due to their distinct natural backgrounds than to distinguish between leopard and cat if we assume the training set consists of only close-up images. In [18], such analogical proportions between classes are discovered by starting from an attribute-based representation of classes. Each class is then a Boolean vector and the analogical proportion is estimated via a parallelogram "score". An exhaustive (unsupervised) search by enumerating all quadruplets of classes leads to keep as relevant proportions only the top-scoring quadruplets. Then a projection matrix is learned (via stochastic gradient descent or similar method) which projects all labeled elements of the training set onto the corresponding class, enforcing a large margin constraint for each instance (i.e., each instance should be closed to its corresponding label instance and far from the other ones). Experimenting on AWA (Animals with Attributes) containing 50 classes [22] and ImageNet datasets, the authors exhibit better performance than standard deep learning-based classification methods.

These two pieces of work highlight the fact that the idea of analogical proportion, considered from a viewpoint of equality or inequality, can bring a better classification accuracy for an image, as soon as we can get a kind of "analogical" meta-knowledge.

8. Conclusion

In this paper, we have first investigated the concept of analogical proportion, starting from the underlying traditional postulates, and investigating the Boolean models compatible with them. We have shown that a complete lattice of models is obtained. The smallest model (in terms of set inclusion) can also be justified in terms of Kolmogorov complexity. The 4-tuples of elements (a, b, c, d) of this minimal model are the ones which minimize an expression involving Kolmogorov complexity of several combinations of the items a, b, c, d. Solving the minimization problem associated to this expression might be the basis of a constructive process for getting analogical proportions.

Besides, we have discussed different noticeable syntactic logically equivalent expressions whose semantics corresponds to the smallest model for Boolean analogical proportion. All these expressions are not equally suitable for an extension to Boolean vectors. Two of them, which vectorially extend well, respectively emphasize the similarity and the dissimilarity aspects of analogical proportions, corresponding to two very different readings, namely "a and b differ as c and d differ" and "what a and d have in common, b and c have it also (both positively and negatively)". The paper also shows the good match between vectorial analogical proportions and Hamming distance-based characterizations. Analogical proportions are also considered from an equational point of view, emphasizing, for the first time, the interest of equation $\vec{a} : \vec{x} :: \vec{y} : \vec{d}$. Indeed given two items $\vec{a}$ and $\vec{d}$, described by vectors of Boolean features differing on at least two features, it is always possible to find distinct $\vec{x}$ and $\vec{y}$ making with them an analogical expression.

Moreover, we have recalled the similarity-based and the dissimilarity-based multiple-valued extensions of Boolean analogical proportions. Despite the fact that they still share some natural properties (like code independence and transitivity), they are no longer equivalent and encode different views of analogy. The similarity-based view restricts the number of valid patterns since only two different numbers can appear in an analogical pattern. On the contrary, the dissimilarity-based view is more flexible and focuses on the difference between graded truth values. This provides the basis for introducing and investigating the idea of analogical inequality viewed as a relaxation of the notion of analogical proportion relying on the idea of equality, both in the Boolean case and in the multiple-valued case. It appears that this proper extension does not just amount to comparing differences (or distances) between the elements of two pairs, but, as in the case of the analogical proportion, it also takes into account the orientation of the variations when going from a to b, and from c to d. Moreover, it also provides a graded estimate of the extent to which "c is more different from d than a is from b". This enables us to turn such a statement into a soft constraint, where the threshold corresponding to the minimal amount to which the constraint should hold might be a matter of learning in practice.

It is not coming as a coincidence that recent papers [18,23,24,19], originated from the machine learning and computer vision community, make use of analogical proportions and analogical inequalities to boost classification accuracy or to improve transfer between images. Nevertheless, analogical proportions and their derived concepts have still to be carefully investigated in order to exploit their full power not only for classification purposes, but for artificial intelligence at large.

References

[1] M. Bayoudh, H. Prade, G. Richard, Evaluation of analogical proportions through Kolmogorov complexity, Knowl.-Based Syst. 29 (2012) 20–30.
[2] S. Bayoudh, L. Miclet, A. Delhay, Learning by analogy: a classification rule for binary and nominal data, in: Proc. Int. Joint Conf. on Artificial Intelligence, IJCAI07, 2007, pp. 678–683.
[3] C.H. Bennett, P. Gács, M. Li, P. Vitányi, W.H. Zurek, Information distance, CoRR, arXiv:1006.3520, 2010.
[4] M. Bounhas, H. Prade, G. Richard, Analogy-based classifiers for nominal or numerical data, Int. J. Approx. Reason. 91 (2017) 36–55.
[5] A. Cornuéjols, Analogy as minimization of description length, in: G. Nakhaeizadeh, C. Taylor (Eds.), Machine Learning and Statistics: The Interface, Wiley, 1996, pp. 321–336.
[6] W. Correa Beltran, H. Prade, G. Richard, Constructive solving of Raven's IQ tests with analogical proportions, Int. J. Intell. Syst. 31 (11) (2016) 1072–1103.
[7] J.P. Delahaye, H. Zenil, On the Kolmogorov–Chaitin complexity for short sequences, CoRR, arXiv:0704.1043, 2007.
[8] J.P. Delahaye, H. Zenil, Numerical evaluation of algorithmic complexity for short strings: a glance into the innermost structure of randomness, Appl. Math. Comput. 219 (1) (2012) 63–77.
[9] D. Dubois, H. Prade, G. Richard, Multiple-valued extensions of analogical proportions, Fuzzy Sets Syst. 292 (2016) 193–202.
[10] T.G. Evans, A program for the solution of a class of geometric-analogy intelligence-test questions, in: M.L. Minsky (Ed.), Semantic Information Process-ing, MIT Press, Cambridge, MA, 1968, pp. 271–353.
[11] Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, Image style transfer using convolutional neural networks, in: The IEEE Conf. on Computer Vision and Pattern Recognition, CVPR, 2016, pp. 2414–2423.
[12] D. Gentner, K.J. Holyoak, B.N. Kokinov, The Analogical Mind: Perspectives from Cognitive Science, Cognitive Science and Philosophy, MIT Press, Cam-bridge, MA, 2001.
[13] S. Goel, S.F. Bush, Kolmogorov complexity estimates for detection of viruses in biologically inspired security systems: a comparison with traditional approaches, Complexity 9 (2) (2003).
[14] M. Hesse, On defining analogy, Proc. Aristot. Soc. 60 (1959) 79–100.
[15] M. Hesse, Models and Analogies in Science, 1st ed., Sheed & Ward, London, 1963, 2nd augmented ed., University of Notre Dame Press, 1966.

[16] D. Hofstadter, M. Mitchell, The Copycat project: a model of mental fluidity and analogy-making, in: D. Hofstadter, The Fluid Analogies Research Group (Eds.), Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought, Basic Books, Inc., New York, NY, 1995, pp. 205–267.
[17] D. Hofstadter, E. Sander, Surfaces and Essences: Analogy as the Fuel and Fire of Thinking, Basic Books, Paris, 2013, French version: L'Analogie, Cœur de la Pensée – Odile Jacob.
[18] S.J. Hwang, K. Grauman, F. Sha, Analogy-preserving semantic embedding for visual object categorization, in: Proc. 30th Int. Conf. on Machine Learning, vol. 28, JMLR.org, 2013, pp. III-639–III-647.
[19] L. Jing, Y. Lu, H. Gang, K. Sing Bing, Visual attribute transfer through deep image analogy, in: ACM Trans. Graph. (Proc. SIGGRAPH), vol. 36, 2017.
[20] S. Klein, Culture, mysticism & social structure and the calculation of behavior, in: Proc. 5th Europ. Conf. in Artificial Intelligence (ECAI'82), Orsay, France, 1982, pp. 141–146.
[21] S. Klein, Analogy and mysticism and the structure of culture (and Comments & Reply), Curr. Anthropol. 24 (2) (1983) 151–180.
[22] Ch.H. Lampert, H. Nickisch, S. Harmeling, Learning to detect unseen object classes by between-class attribute transfer, in: IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR), June 20–25, Miami, 2009.
[23] M.T. Law, N. Thome, M. Cord, Quadruplet-wise image similarity learning, in: IEEE Int. Conf. on Computer Vision (ICCV), Sydney, 2013, pp. 249–256.
[24] M.T. Law, N. Thome, M. Cord, Learning a distance metric from relative comparisons between quadruplets of images, Int. J. Comput. Vis. 121 (1) (2017) 65–94.
[25] Y. Lepage, Analogy and formal languages, in: L.S. Moss, R.T. Oehrle (Eds.), Proc. Joint Meeting of the 6th Conf. on Formal Grammar and the 7th Conf. on Mathematics of Language, Electron. Notes Theor. Comput. Sci. 53 (2002) 180–191.
[26] L. Levin, Laws of information conservation (non-growth) and aspects of the foundation of probability theory, Probl. Inf. Transm. 10 (1974) 206–210.
[27] M. Li, P. Vitanyi, An Introduction to Kolmogorov Complexity and Its Applications, 3rd ed., Springer Verlag, 2008.
[28] L. Miclet, S. Bayoudh, A. Delhay, Analogical dissimilarity: definition, algorithms and two experiments in machine learning, J. Artif. Intell. Res. 32 (2008) 793–824.
[29] L. Miclet, A. Delhay, Relation d'analogie et distance sur un alphabet défini par des traits, Technical Report 1632, IRISA, July 2004.
[30] L. Miclet, A. Delhay, Analogical Dissimilarity: Definition, Algorithms and First Experiments in Machine Learning, Technical Report RR-5694, INRIA, July 2005.
[31] L. Miclet, H. Prade, Logical definition of analogical proportion and its fuzzy extensions, in: Proc. Ann. Meet. North Amer. Fuzzy Inform. Proc. Soc. (NAFIPS), New York, IEEE, 2008, pp. 1–6.
[32] L. Miclet, H. Prade, 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, in: Lect. Notes Comput. Sci., vol. 5590, Springer, 2009, pp. 638–650.
[33] J. Piaget, Logic and Psychology, Manchester Univ. Press, 1953.
[34] M. Pirlot, H. Prade, G. Richard, Completing Preferences by Means of Analogical Proportions, Springer, 2016, pp. 135–147.
[35] H. Prade, G. Richard, From analogical proportion to logical proportions, Log. Univers. 7 (4) (2013) 441–505.
[36] H. Prade, G. Richard, Homogeneous and heterogeneous logical proportions, IfCoLog J. Log. Appl. 1 (1) (2014) 1–51.
[37] H. Prade, G. Richard, Analogical inequalities, in: A. Antonucci, L. Cholvy, O. Papini (Eds.), Proc. 14th Europ. Conf. (ECSQARU'17) on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, Lugano, July 10–14, in: Lect. Notes Comput. Sci., vol. 10369, Springer, 2017, pp. 3–9.
[38] H. Prade, G. Richard, Boolean analogical proportions – axiomatics and algorithmic complexity issues, in: A. Antonucci, L. Cholvy, O. Papini (Eds.), Proc. 14th Europ. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'17), Lugano, July 10–14, in: Lect. Notes Comput. Sci., vol. 10369, Springer, 2017, pp. 10–21.
[39] H. Prade, G. Richard, From the structures of opposition between similarity and dissimilarity indicators to logical proportions, in: G. Dodig-Crnkovic, R. Giovagnoli (Eds.), Representation and Reality in Humans, Other Living Organisms and Intelligent Machines, Springer, 2017, pp. 279–299.
[40] H. Prade, G. Richard, Multiple-valued logic interpretations of analogical, reverse analogical, and paralogical proportions, in: Proc. 40th IEEE Int. Symp. on Multiple-Valued Logic (ISMVL'10), Barcelona, 2010, pp. 258–263.
[41] D.E. Rumelhart, A.A. Abrahamson, A model for analogical reasoning, Cogn. Psychol. 5 (2005) 1–28.
[42] S. Schockaert, H. Prade, Completing rule bases using analogical proportions, in: H. Prade, G. Richard (Eds.), Computational Approaches to Analogical Reasoning – Current Trends, Springer, 2014.
[43] F. Soler-Toscano, H. Zenil, J.P. Delahaye, N. Gauvrit, Correspondence and independence of numerical evaluations of algorithmic information measures, Computability (2) (2013) 125–140.
[44] R.J. Solomonoff, A formal theory of inductive inference. Part I and II, Inf. Control 7 (1) (1964) 1–22, 224–254.
[45] A. Tversky, Features of similarity, Psychol. Rev. 84 (1977) 327–352.
[46] F. Yvon, N. Stroppa, Formal Models of Analogical Proportions, Technical Report D008, Ecole Nationale Supérieure des Télécommunications, Paris, France, 2006.