From the Structures of Opposition Between Similarity and Dissimilarity Indicators to Logical Proportions

A General Representation Setting for Capturing Homogeneity and Heterogeneity

Henri Prade and Gilles Richard

Abstract Comparative thinking plays a key role in our appraisal of reality. Comparing two objects or situations A and B, described in terms of Boolean features, may involve four basic similarity or dissimilarity indicators referring to what A and B have in common (positively or negatively), or to what is particular to A or particular to B. These four indicators are naturally organized into a cube of opposition, which includes two classical squares of opposition, as well as other noticeable squares. From the knowledge of one situation A, it is possible to recover the description of another one, B, provided that we have enough information about the comparison between A and B. Then comparison indicators between A and B can be equated with comparison indicators between two other situations C and D. A conjunction of two such comparisons between pairs (A, B) and (C, D) gives birth to what is called a logical proportion. Among the 120 existing logical proportions, 8 are of particular interest since they are independent of the encoding (positive or negative) used for representing the situations. Four of them have remarkable properties of homogeneity, and include the analogical proportion "A is to B as C is to D", while the four others express heterogeneity by stating that "there is an intruder among A, B, C and D, which is not X" (where X stands for A, B, C or D). Homogeneous and heterogeneous logical proportions are of interest in classification, anomaly detection tasks and IQ test solving.

H. Prade (✉) · G. Richard
IRIT, Université Paul Sabatier, 118 route de Narbonne, 31062
Toulouse Cedex 09, France
e-mail: prade@irit.fr

G. Richard
e-mail: richard@irit.fr

G. Dodig-Crnkovic and R. Giovagnoli (eds.), *Representation and Reality in Humans, Other Living Organisms and Intelligent Machines*, Studies in Applied Philosophy, Epistemology and Rational Ethics 28, DOI 10.1007/978-3-319-43784-2_14

1 Introduction

The role of comparison in our perception of reality has been recognized for a long time. Making comparison is closely related to similarity judgment [24] and analogy making [9]. In particular, comparison applied to numerical quantities is a matter of differences and ratios, which, by equating differences or ratios, leads to the idea of a proportion. Ancient Greek mathematicians, such as Archytas of Tarentum, or his follower Eudoxus of Cnidus, already studied mathematical proportions, including arithmetic, geometric and harmonic ones.

Numerical proportions involve four terms A, B, C and D. Arithmetic (resp. geometric) proportions state the equality of two differences (resp. ratios) between two ordered pairs (A, B) and (C, D) of numbers. Harmonic proportions combine geometric and arithmetic comparisons by stating that $A/D = (A - B)/(C - D)$. Associated with each type of proportion is a particular *mean* operation, obtained by taking $B = C$ as an unknown number in the equality stating the proportion.

Aristotle, following Eudoxus of Cnidus, does not only refer to geometric proportions, but also considers comparative relations between four symbolic (i.e. nonnumerical) terms that form in modern words an *analogical proportion*, namely a statement of the form "A is to B as C is to D" expressing an identity of relation between the ordered pairs (A, B) and (C, D) of symbols (where now A, B, C and D refer to objects or situations). The use of pictorial tests based on analogical proportions has often been used in psychological studies on the development of human thought [15].

However, it is only recently that analogical proportions have been cast in a logical setting, when A, B, C and D refer to four situations described in terms of Boolean features [12]. The logical expression of an analogical proportion exactly states that "A differs from B as C differs from D, and B differs from A as D differs from C". Thus an analogical proportion equates the dissimilarity of A with respect to B (resp. of B with respect to A) with the dissimilarity of C with respect to D (resp. of D with respect to C). This modeling provides a way of extrapolating D from A, B and C, as in the numerical setting, and to validate a computational procedure first discovered by Sheldon Klein [11], where the analogical proportion is supposed to hold, in a pointwise manner, for each feature, between the four Boolean values corresponding to A, B, C and D. However, A, B and C cannot always be completed with a D in order to make a valid analogical proportion. But if it exists, D is unique.

As already mentioned, completing an analogical proportion where the last item is missing is the basis of psychological tests. This mechanism may also be useful for solving more sophisticated IQ tests such as Raven's progressive matrices tests. Such tests are built from a 3×3 array with eight pictures in the first eight cells while the ninth picture has to be completed. The analogical proportion mechanism enables us to directly build the missing picture from the eight given pictures [3], rather than choosing it from among a set of candidate pictures as in the real test.

Then, the logical expression of an analogical proportion has been shown to be an important particular case of the remarkable notion of *logical proportion* which has been recently introduced [17, 19, 20]. However, the phrase "logical proportion" was first coined by [14] in a more restrictive sense for naming a quaternary relation between Boolean propositions, inspired from geometrical proportions. It turns out that the relation defined by Piaget is one of the noticeable expressions of an analogical proportion [19], a fact remained unnoticed by Piaget. The general form of a logical proportion is a conjunction of two equivalences between *comparison indicators* expressing similarities or dissimilarities pertaining to pairs of Boolean variables encoding the values of a considered feature for the pairs (A, B) and (C, D).

A form of comparative reasoning, simpler than the extrapolation of D from A, B and C, amounts to reconstruct, feature by feature, of the description of an object B from the value of some comparison indicators with respect to another object A, knowing A. This latter type of reasoning can be used as a starting point for providing a new introduction to the idea of a logical proportion, from the notion of comparison indicators. Besides, a natural question is then to identify the logical proportions that have extrapolative power, i.e. which lead to a unique solution for the fourth situation, knowing the three others, and assuming that a particular logical proportion holds for each feature between the four situations.

The paper is organized as follows: In Sect. 2, we investigate how we can compare an object A with an object B, thanks to four types of comparative information. We show that this gives birth to a cube of opposition, which includes two classical squares of opposition, and exhaustively exhibits the relations between the possible values of the four comparison indicators. These indicators are also shown to be necessary and sufficient for describing the respective configurations of two sets. Moreover, we determine, by easy Boolean calculations, in what cases the description of an object A may be recovered from the description of another object B and the knowledge of the values of two comparative indicators relating them. This establishes a link with logical proportions where two comparison indicators pertaining to a pair (A, B) are respectively equated with two comparison indicators pertaining to a pair (C, D). In Sect. 3, we provide a short background on logical proportions, highlighting some of their remarkable properties, and identifying eight logical proportions of particular interest: four homogeneous ones that include the analogical proportion and are based on equivalences between indicators of the same nature (referring either to similarity or dissimilarity), and four heterogeneous ones that encode the presence of an intruder among four items. In Sect. 4, we investigate the way to solve Boolean equations involving logical proportions, and identify the proportions suitable for extrapolation. It turns out that this is exactly the eight previously mentioned ones. In Sect. 5, after listing and discussing problems of interest in relation to extrapolation, we present a generic transduction rule, which, when combined with the equation solving process, provides a general basis for extrapolation. Lastly, we briefly discuss the respective merits of homogeneous and heterogeneous logical proportions for classification and prediction, or intruder detection. The contents of this article is partly based on a workshop paper [21].

2 Indicators as Comparative Descriptors

Given a collection $\mathcal{U}$ of Boolean properties, it is common practice to describe an object A by a set a of properties that this object satisfies. Thus, a is a representation of A, where $a \subseteq \mathcal{U}$. Note that $a = \emptyset$ or $a = \mathcal{U}$ are not forbidden, since A may have none of the properties in $\mathcal{U}$, or all of them. Having all the characteristics of A w.r.t. $\mathcal{U}$ means that we have complete knowledge of a (w.r.t. the conceptual space induced by $\mathcal{U}$).

When comparing objects A and B, one looks for their similarities and dissimilarities. More precisely, there are only four types of comparative information between two objects A and B with respect to $\mathcal{U}$ (in the following $\overline{a}$ denotes the set complement of set a):

- The set of properties that A and B share: $a \cap b$
- The set of properties that A has but that B does not have: $a \cap \overline{b}$
- The set of properties that B has but that A does not have: $\overline{a} \cap b$
- The set of properties that neither A nor B have: $\overline{a} \cap \overline{b}$

Thus, these four expressions provide comparative information regarding A and B, and we call them *set indicators* or *indicators* for short. We have two types of indicators:

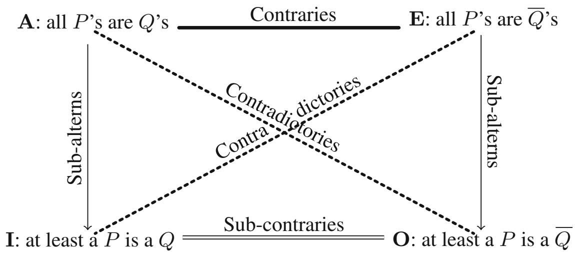
- $a \cap b$ and $\overline{a} \cap \overline{b}$, telling us about properties that both A and B have, or that both A and B do not have; they are called *similarity indicators*.
- $a \cap \overline{b} = a \setminus b$ and $\overline{a} \cap b = b \setminus a$, telling us about properties that only one among A and B has; they are called *dissimilarity indicators*.

We notice that the union of the four indicators is just $\mathcal{U}$ and the intersection of any two indicators is empty. This is reminiscent of the well-known work of Amos Tversky [24], taking into account the common features, the specificities of A w.r.t. B and the specificities of B w.r.t. A, respectively, modeled by $a \cap b$, $a \setminus b$ and $b \setminus a$ in order to define a global measure of similarity. However, here, we are rather interested in keeping track of in what respect items are similar and in what respect they are dissimilar using Boolean indicators in a logical setting.

2.1 The Cube of Opposition of Comparison Indicators

Aristotle and his followers noticed that universally and existentially quantified statements, which are encountered in syllogisms, can be organized into a *square of opposition*. More precisely, consider a statement (**A**) of the form "all P's are Q's", which is negated by the statement (**O**) "at least a P is not a Q", together with the statement (**E**) "no P is a Q", which is clearly in even stronger opposition to the first statement (**A**). These three statements, together with the negation of the last statement, namely (**I**) "at least a P is a Q", give birth to the square of opposition [13], traditionally

Fig. 1 Square of opposition



denoted by the letters **A**, **I** (affirmative half) and **E**, **O** (negative half), pictured in Fig. 1 (where $\overline{Q}$ stands for "not Q").

As can be checked, noticeable relations hold between the four vertices of the square:

1. **A** and **O** are the negation of each other, as are **E** and **I**;
2. **A** entails **I**, and **E** entails **O** (we assume that there is at least a P to avoid existential import problems);
3. **A** and **E** cannot be true together, but may be false together;
4. **I** and **O** cannot be false together, but may be true together.

Negating the predicates, i.e. changing P into $\overline{P}$ and Q into $\overline{Q}$, leads to another similar square of opposition **a**, **i**, **e**, **o**, based on "not-P's" assumed to constitute a non-empty set. Then the eight statements, **A**, **I**, **E**, **O**, **a**, **i**, **e**, **o** may be organized in what may be called a *cube of opposition* [6] as in Fig. 2. The front facet and the back facet of the cube are traditional squares of opposition, where the thick nondirected segment relates the contraries, the double thin nondirected segments the subcontraries, the diagonal nondirected segments the contradictories and the vertical unidirected segments point to subalterns, and express entailments.

Assuming that there are at least a P and at least a not-P entails that there are at least a Q and at least a not-Q. Then, we have that **A** entails **i**, **a** entails **I**, **e** entails **O**, and **E** entails **o**. Note also that the vertices **a** and **E**, as well as **A** and **e**, cannot be

Fig. 2 Cube of opposition

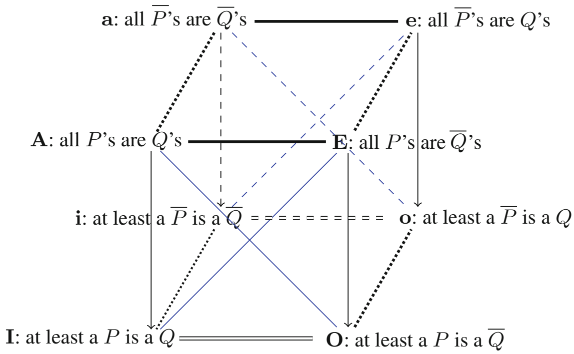
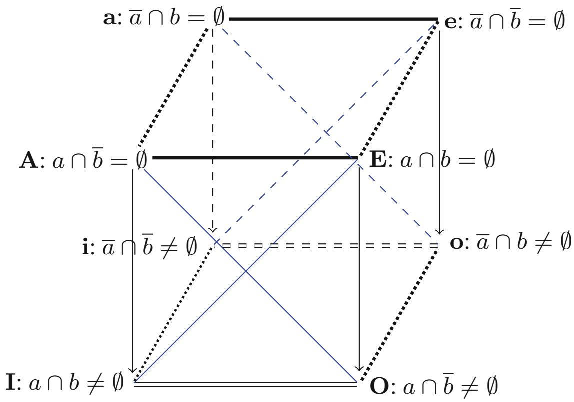


Fig. 3 Cube of opposition of comparison indicators



true together, while the vertices **i** and **O**, as well as **I** and **o**, cannot be false together. Lastly note that there is no logical link between **A** and **a**, **E** and **e**, **I** and **i**, or **O** and **o**.

Going back to the set of properties a and b describing two situations A and B, the cube of opposition may be rewritten in terms of the four set comparison indicators, which may be empty or nonempty. The four indicators appear in the side facets of the cube. The nonclassical squares of opposition of the side facets (with a different display of the vertices) were already discussed in [19], without referring to any cube. See Fig. 3. Note that we assume $a \neq \emptyset$, $\overline{a} \neq \emptyset$, $b \neq \emptyset$ and $\overline{b} \neq \emptyset$ here to avoid the counterpart of the existential import problems, since now the situations A and B play symmetric roles in the statements associated with the vertices of the cube.

2.2 Possible Configurations of Two Subsets in Terms of Set Indicators

Let us here denote by s, t, u, v the four set indicators pertaining to a and b: $s = a \cap b$, $u = a \cap \overline{b}$, $v = \overline{a} \cap b$ and $t = \overline{a} \cap \overline{b}$. Considering that a set indicator can be empty or not, we have $2^4 = 16$ configurations that we describe in Table 1. To each indicator $i \in \{s, t, u, v\}$, we can associate a Boolean variable i' defined as: $i' = 1$ if $i \neq \emptyset$ and $i' = 0$ otherwise.

As can be seen in Table 1, lines 1 and 2 correspond to situations of overlapping without inclusion, with coverage ($a \cup b = \mathcal{U}$) or not. Lines 3, 4, 5 and 6 correspond to situations of inclusion, with coverage or not. Lines 7 and 8 correspond to situations of equality, with coverage or not. Lines 9 and 10 correspond to situations of nonoverlapping, with coverage or not. The last six lines correspond to pathological situations where a or b are empty, with coverage or not. The four indicators s, t, u, v are thus jointly necessary for describing all the possible situations pertaining to the

Table 1 Respective configurations of two subsets

	Configuration	$s = a \cap b \neq \emptyset$	$u = a \cap \overline{b} \neq \emptyset$	$v = \overline{a} \cap b \neq \emptyset$	$t = \overline{a} \cap \overline{b} \neq \emptyset$
1	$a \cap b \neq \emptyset; a \not\subseteq b;$ $b \not\subseteq a; a \cup b \neq \mathcal{U}$	1	1	1	1
2	$a \cap b \neq \emptyset; a \not\subseteq b;$ $b \not\subseteq a; a \cup b = \mathcal{U}$	1	1	1	0
3	$b \subset a \subset \mathcal{U}$	1	1	0	1
4	$b \subset a; a = \mathcal{U}$	1	1	0	0
5	$a \subset b \subset \mathcal{U}$	1	0	1	1
6	$a \subset b; b = \mathcal{U}$	1	0	1	0
7	$a = b \subset \mathcal{U}$	1	0	0	1
8	$a = b = \mathcal{U}$	1	0	0	0
9	$a \cap b = \emptyset;$ $a \cup b \neq \mathcal{U}$	0	1	1	1
10	$a \cap b = \emptyset;$ $a \cup b = \mathcal{U}$	0	1	1	0
11	$a \subset \mathcal{U}; b = \emptyset;$	0	1	0	1
12	$a = \mathcal{U}; b = \emptyset$	0	1	0	0
13	$a = \emptyset; b \subset \mathcal{U}$	0	0	1	1
14	$a = \emptyset; b = \mathcal{U}$	0	0	1	0
15	$a = b = \emptyset;$ $\mathcal{U} \neq \emptyset$	0	0	0	1
16	$a = b = \emptyset = \mathcal{U}$	0	0	0	0

relative position of two subsets a and b, possibly empty, in a referential. Then the following properties can also be checked in Table 1: $\max(s', t', u', v') = 1$.

This means that the four set indicators cannot be simultaneously empty, except if the referential $\mathcal{U}$ is empty, which corresponds to line 16 in Table 1. In logical terms we have: $s' \vee t' \vee u' \vee v' \equiv \top$. If $a \neq \emptyset$, $b \neq \emptyset$, $a \neq \mathcal{U}$, $b \neq \mathcal{U}$, we have

$$\max(n(u'), n(v')) \leq \min(s', t'),$$

where $n(x) = 1$ if $x = 0$ and $n(x) = 0$ if $x = 1$. A counterpart of this property holds in possibility theory and has also been noticed in formal concept analysis [7]. This corresponds to lines 1, 2, 3, 5, 7, 9 and 10 in Table 1. This also corresponds to the five possible configurations of two nonempty subsets a and b (lines 1, 3, 5, 7, 9), first identified by Gergonne [8, 10] when discussing syllogisms, plus two configurations (lines 2, 10) where $a \cup b = \mathcal{U}$ (but where $a \neq \mathcal{U}$, $b \neq \mathcal{U}$).

It should also be noticed that a subset a can be defined from another subset b, with the help of some of its positive and negative similarities s and t to a, and its differences u and v, in many different ways, where b appears or not:

$$a = s \cup u\,; \qquad a = \overline{v \cup t}\,;$$

$$a = s \cup (\overline{b} \cap \overline{t})\,; \qquad a = \overline{v \cup (\overline{b} \cap \overline{u})} = \overline{v} \cap (b \cup u) = u \cup (b \cap \overline{v}).$$

2.3 Recovering a Subset from Another Subset and Comparative Information

It may be the case that we have no direct information at our disposal about A, but only some comparative information with respect to another known object B (i.e. where b is known w.r.t. $\mathcal{U}$). Obviously, if we know that A is identical to the known object B, then $a = b$ and we are done. On the other hand, if we know that A is the exact opposite of B, then we are done again with $a = \mathcal{U} \setminus b = \overline{b}$. But, in general, we may only have information about partial match and/or partial mismatch. In the following, we assume that the information is complete (with respect to $\mathcal{U}$) about some of the above comparison indicators. For instance, we know *all* the properties in $\mathcal{U}$ that A and B share. But, we may also take into account dissimilarities as well when the information is available. We now discuss how to get accurate information regarding A starting from B and from some of the previous types of comparative information.

It is quite clear that knowing b and only one comparison indicator is not enough to recover complete knowledge of a. For instance, there is not a unique subset a such that $s = a \cap b$ and b are given. We only know that $s \subseteq a \subseteq s \cup \overline{b}$.

In order to get an accurate view of A starting from comparative information w.r.t. B, we need to know at least two comparison indicators. So our problem can now be stated as follows: what are the pairs of indicators which could lead to complete knowledge of A? More formally, what are the pairs of indicators (i_1, i_2) such that a is a function of b, i_1, i_2?

As we have four indicators, we have four choices for i_1 and then three remaining choices for i_2, leading to a total of 12 options. But as the problem is entirely symmetrical with respect to i_1 and i_2, we only have six cases to consider:

1. $i_1 = a \cap b$ and $i_2 = a \cap \overline{b}$: in that case, we compute $a = (a \cap b) \cup (a \cap \overline{b}) = i_1 \cup i_2$ as the unique solution.
2. $i_1 = a \cap b$ and $i_2 = \overline{a} \cap b$: in that case, the solution is not unique for a, i.e. we do not have enough information for computing a. We only know that $i_1 \subseteq a \subseteq \overline{i_2}$.
3. $i_1 = a \cap b$ and $i_2 = \overline{a} \cap \overline{b}$: we start from $a = (a \cap b) \cup (a \cap \overline{b}) = i_1 \cup (a \cap \overline{b})$, but $a \cap \overline{b} = (a \cup b) \cap \overline{b} = \overline{i_2} \cap \overline{b} = \overline{i_2 \cup b}$, leading to a unique $a = i_1 \cup \overline{i_2 \cup b}$.
4. $i_1 = a \cap \overline{b}$ and $i_2 = \overline{a} \cap b$: starting from $a = (a \cap b) \cup (a \cap \overline{b}) = (a \cap b) \cup i_1$, since $b = (a \cap b) \cup (\overline{a} \cap b) = (a \cap b) \cup i_2$, we have $(a \cap b) = b \setminus i_2$, leading to the unique solution $a = i_1 \cup (b \setminus i_2)$.
5. $i_1 = a \cap \overline{b}$ and $i_2 = \overline{a} \cap \overline{b}$: this case is the dual of case 2 where b is replaced with $\overline{b}$. Similarly, the solution is still not unique. We still have $i_1 \subseteq a \subseteq \overline{i_2}$.
6. $i_1 = \overline{a} \cap b$ and $i_2 = \overline{a} \cap \overline{b}$: obviously $\overline{a} = i_1 \cup i_2$ then $a = \overline{i_1 \cup i_2}$.

Table 2 a is not computable

b	1	0	0	1	1	0
$b \cap a$	1	0	0	1	0	0
$b \cap \overline{a}$	0	0	0	0	1	0
a	1	?	?	1	0	?

Table 3 A unique a is computable

b	1	0	0	1	1	0
$b \cap a$	1	0	0	1	0	0
$\overline{b} \cap \overline{a}$	0	1	0	0	0	0
a	1	0	1	1	0	1

As we can see, there are only four pairs of indicators leading to a unique a (we recognize the expressions of the previous subsection), while the two remaining pairs are not informative enough to uniquely determine the set a.

Let us consider an example where $\mathcal{U}$ contains six properties such that every object can be represented as a six digit Boolean vector. For instance, object B is represented by the vector $b = 100110$, which means that B has the first, the fourth and the fifth property and not the others. Another object A is unknown, but we have some comparative information with respect to B: for instance we know $b \cap a$, i.e. the properties that B and A share: $b \cap a = 100100$. We also know $b \cap \overline{a} = 000010$. We are exactly in case 2 above, and we cannot recover A, as three components are not deducible, as shown in Table 2. This results from the fact that we have no information for A regarding the properties that B does not have. But in the case where we know $\overline{b} \cap \overline{a} = 010000$ (in place of $b \cap \overline{a}$), we can recover A as in Table 3, where the set d is just $(b \cap a) \cup \overline{(\overline{b} \cap \overline{a})} \cup b$ (as in case 3 above). It could also happen that the given disjoint subsets $i_1 \subset \mathcal{U}$ and $i_2 \subset \mathcal{U}$ are themselves indicators for other objects C and D different from A and B. A related question then arises: given two disjoint subsets i_1 and i_2, can we always find a pair of objects (C, D) such that i_1 and i_2 are the values of indicators for the pair (C, D)?

The answer to this question is positive and can be checked as follows: Considering only similarity indicators for instance, let us look for a pair (C, D) such that $c \cap d = i_1$ and $\overline{c} \cap \overline{d} = i_2$, knowing that $i_1 \cap i_2 = \emptyset$. The second equation is equivalent to $c \cup d = \overline{i_2}$, and the problem can be restated as an equivalent one: Knowing that $i_1 \subseteq \overline{i_2}$, find two sets c and d such that: $c \cap d = i_1$ and $c \cup d = \overline{i_2}$. It is well known that this problem has solutions (generally not unique). This remark allows us to change the setting of our initial problem, since it appears that having the values of two indicators for a pair (A, B) is equivalent to having two equalities between indicators for the pair (A, B) and indicators for another pair (C, D).

Since the power set $2^{\mathcal{U}}$ is a Boolean lattice, we can reword the problem in a Boolean setting where $\cap$ is translated into $\wedge$, $\cup$ into $\vee$, and $=$ into $\equiv$. We keep the complementation overbar, as a compact notation for negation. Then, equivalences between indicators lead to the notion of logical proportions that we investigate in the next section.

3 Brief Background on Logical Proportions

Adopting a propositional logic view, for a given pair of Boolean variables (a, b), we have four distinct indicators:

- $a \wedge b$ and $\overline{a} \wedge \overline{b}$, called *similarity indicators*
- $a \wedge \overline{b}$ and $\overline{a} \wedge b$, called *dissimilarity indicators*

Following what we did in the subset-based analysis above, comparing a pair of variables (a, b) with another pair (c, d) is done via a pair of equivalence between indicators.

Note that a, b, c and d no longer denote subsets of properties that respectively hold in situations A, B, C and D. Rather, they encode Boolean variables pertaining respectively to situations A, B, C and D, which refer to one particular property on the basis of which these situations are compared. In other words, variables a, b, c and d encode respectively if a particular considered property holds true, or not, in situations A, B, C and D. So we need as many 4-tuples of Boolean variables (a, b, c, d) of this kind as there are properties used for describing situations A, B, C and D.

3.1 The 120 Logical Proportions

As shown in the introductory discussion, it makes sense to consider the values of two indicators to be in a position to compute a Boolean variable related to them. As a consequence, it is legitimate to consider all the conjunctions of two equivalences between indicators: such a conjunction is called a *logical proportion* [16, 17]. More formally, let $I_{(a,b)}$ and $I'_{(a,b)}$ [1] (resp. $I_{(c,d)}$ and $I'_{(c,d)}$) denote two indicators for (a, b) (resp. (c, d)).

Definition 1 A logical proportion $T(a, b, c, d)$ is the conjunction of two distinct equivalences between indicators of the form $(I_{(a,b)} \equiv I_{(c,d)}) \wedge (I'_{(a,b)} \equiv I'_{(c,d)})$.

It can be easily checked that there are 16 distinct equivalences relating a comparison indicator pertaining to a and b to a comparison indicator pertaining to c and d. Their truth tables are given in Table 4, exhibiting, in each case, the ten patterns of truth values of a, b, c and d that make the corresponding equivalence true. It is false

[1]Note that $I_{(a,b)}$ (or $I'_{(a,b)}$) refers to one element in the set $\{a \wedge b, \overline{a} \wedge b, a \wedge \overline{b}, \overline{a} \wedge \overline{b}\}$, and should not be considered as a functional symbol: $I_{(a,b)}$ and $I_{(c,d)}$ may be indicators of two different kinds. Still, we use this notation for the sake of simplicity.

Table 4 Patterns making true the 16 equivalences between indicators. The conjunctive combination of $\underline{X}$ and $\overline{X}$ yields the logical proportion X

$\underline{\mathbf{A}}$	$\overline{\mathbf{A}}$	$\underline{\mathbf{R}}$	$\overline{\mathbf{R}}$	$\underline{\mathbf{P}}$	$\overline{\mathbf{P}}$	$\underline{\mathbf{I}}$	$\overline{\mathbf{I}}$
$a\bar{b} \equiv c\bar{d}$	$\bar{a}b \equiv \bar{c}d$	$a\bar{b} \equiv \bar{c}d$	$\bar{a}b \equiv c\bar{d}$	$ab \equiv cd$	$\bar{a}\bar{b} \equiv \bar{c}\bar{d}$	$ab \equiv \bar{c}\bar{d}$	$\bar{a}\bar{b} \equiv cd$
0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0	0 0 1 1	0 0 1 1
1 1 1 1	1 1 1 1	1 1 1 1	1 1 1 1	1 1 1 1	1 1 1 1	1 1 0 0	1 1 0 0
0 0 1 1	0 0 1 1	0 0 1 1	0 0 1 1	0 1 0 1	0 1 0 1	0 1 0 1	0 1 0 1
1 1 0 0	1 1 0 0	1 1 0 0	1 1 0 0	1 0 1 0	1 0 1 0	1 0 1 0	1 0 1 0
0 1 0 1	0 1 0 1	0 1 1 0	0 1 1 0	0 1 1 0	0 1 1 0	0 1 1 0	0 1 1 0
1 0 1 0	1 0 1 0	1 0 0 1	1 0 0 1	1 0 0 1	1 0 0 1	1 0 0 1	1 0 0 1
0 0 0 1	0 0 1 0	0 0 1 0	0 0 0 1	0 0 0 1	0 1 1 1	0 0 0 1	0 1 0 0
0 1 0 0	1 0 0 0	0 1 0 0	1 0 0 0	0 0 1 0	1 0 1 1	0 0 1 0	1 0 0 0
0 1 1 1	1 0 1 1	0 1 1 1	1 0 1 1	0 1 0 0	1 1 0 1	0 1 1 1	1 1 0 1
1 1 0 1	1 1 1 0	1 1 1 0	1 1 0 1	1 0 0 0	1 1 1 0	1 0 1 1	1 1 1 0

$\underline{\mathbf{H_a}}$	$\overline{\mathbf{H_a}}$	$\underline{\mathbf{H_b}}$	$\overline{\mathbf{H_b}}$	$\underline{\mathbf{H_c}}$	$\overline{\mathbf{H_c}}$	$\underline{\mathbf{H_d}}$	$\overline{\mathbf{H_d}}$
$a\bar{b} \equiv cd$	$\bar{a}b \equiv \bar{c}\bar{d}$	$\bar{a}b \equiv cd$	$a\bar{b} \equiv \bar{c}\bar{d}$	$ab \equiv c\bar{d}$	$\bar{a}\bar{b} \equiv \bar{c}d$	$ab \equiv \bar{c}d$	$\bar{a}\bar{b} \equiv c\bar{d}$
0 0 0 1	0 0 0 1	0 0 0 1	0 0 0 1	0 0 0 1	0 0 0 1	0 0 1 0	0 0 1 0
0 0 1 0	0 0 1 0	0 0 1 0	0 0 1 0	0 1 0 0	0 1 0 0	0 1 0 0	0 1 0 0
0 1 0 0	0 1 0 0	1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0	1 0 0 0
1 0 1 1	1 0 1 1	0 1 1 1	0 1 1 1	0 1 1 1	0 1 1 1	0 1 1 1	0 1 1 1
1 1 0 1	1 1 0 1	1 1 0 1	1 1 0 1	1 0 1 1	1 0 1 1	1 0 1 1	1 0 1 1
1 1 1 0	1 1 1 0	1 1 1 0	1 1 1 0	1 1 1 0	1 1 1 0	1 1 0 1	1 1 0 1
0 0 0 0	0 0 1 1	0 0 0 0	0 0 1 1	0 0 0 0	0 1 1 0	0 0 0 0	0 1 0 1
0 1 0 1	1 0 0 1	1 0 0 1	0 1 0 1	0 0 1 1	1 0 1 0	0 0 1 1	1 0 0 1
0 1 1 0	1 0 1 0	1 0 1 0	0 1 1 0	0 1 0 1	1 1 0 0	0 1 1 0	1 1 0 0
1 1 0 0	1 1 1 1	1 1 0 0	1 1 1 1	1 0 0 1	1 1 1 1	1 0 1 0	1 1 1 1

for the six missing patterns. Note that conjunctions are omitted in the expressions of comparison indicators in Table 4 for the sake of a compact writing. The meaning of the names given to these 16 equivalences will be made clear in the next two subsections.

Since we have to choose 2 distinct equivalences among $4 \times 4 = 16$ possible ones for defining a logical proportion, we have $\binom{16}{2} = 120$ such proportions, and it has been shown that they are all semantically distinct. Consequently, if two proportions are semantically equivalent, they have the same expression as a conjunction of two equivalences between indicators (up to the symmetry of the conjunction). Then a logical proportion is just a particular Boolean formula involving four variables and as such, has a truth table with 16 lines. It has been shown [19] that an equivalence between 2 indicators has exactly ten lines leading to truth value 1 in its truth table, and that any logical proportion has exactly six lines leading to truth value 1 in its truth table.

First of all, depending on the way the indicators are combined, different types of logical proportions are obtained [16, 19]. There are:

- **4** *homogeneous* proportions that involve only dissimilarity, or only similarity indicators: they are called, *analogy* (A), *reverse analogy* (R), *paralogy* (P) and *inverse paralogy* (I) (see definitions below).
- **16** *conditional* proportions defined as the conjunction of an equivalence between similarity indicators and of an equivalence between dissimilarity indicators, as, e.g. $((a \wedge b) \equiv (c \wedge d)) \wedge ((a \wedge \overline{b}) \equiv (c \wedge \overline{d}))$, which expresses that the two conditional objects[2] $b|a$ and $d|c$ have the same examples (first condition) and the same counter-examples (second condition).
- **20** *hybrid* proportions obtained as the conjunction of two equivalences between similarity and dissimilarity indicators, as in the following example: $(a \wedge b \equiv \overline{c} \wedge d) \wedge (\overline{a} \wedge \overline{b} \equiv c \wedge \overline{d})$.
- **32** *semihybrid* proportions for which one half of their expressions involve indicators of the same kind, while the other half requires equivalence between indicators of opposite kinds, as, e.g. $(a \wedge b \equiv c \wedge d) \wedge (\overline{a} \wedge \overline{b} \equiv c \wedge \overline{d})$.
- **48** *degenerated* proportions whose definition involves three distinct indicators only.

3.2 The Four Homogeneous Logical Proportions

When we consider the logical proportions with homogeneous equivalences only, i.e. the ones that involve only dissimilarity, or only similarity indicators, we get the four logical proportions mentioned above, which are listed below with their expression:

[2] A conditional object $b|a$ can take three truth values: true if $a \wedge b$ is true, false if $a \wedge \neg b$ is true, not applicable if a is false; it may be intuitively thought of as representing the rule "if a then b" [5].

Table 5 Analogy, reverse analogy, paralogy and inverse paralogy truth tables

A	**R**	**P**	**I**
0 0 0 0	0 0 0 0	0 0 0 0	1 1 0 0
1 1 1 1	1 1 1 1	1 1 1 1	0 0 1 1
0 0 1 1	0 0 1 1	1 0 0 1	1 0 0 1
1 1 0 0	1 1 0 0	0 1 1 0	0 1 1 0
0 1 0 1	0 1 1 0	0 1 0 1	0 1 0 1
1 0 1 0	1 0 0 1	1 0 1 0	1 0 1 0

- *Analogy*: $A(a, b, c, d)$, defined by $((a \wedge \overline{b}) \equiv (c \wedge \overline{d})) \wedge ((\overline{a} \wedge b) \equiv (\overline{c} \wedge d))$
- *Reverse analogy*: $R(a, b, c, d)$, defined by $((a \wedge \overline{b}) \equiv (\overline{c} \wedge d)) \wedge ((\overline{a} \wedge b) \equiv (c \wedge \overline{d}))$
- *Paralogy*: $P(a, b, c, d)$, defined by $((a \wedge b) \equiv (c \wedge d)) \wedge ((\overline{a} \wedge \overline{b}) \equiv (\overline{c} \wedge \overline{d}))$
- *Inverse paralogy*: $I(a, b, c, d)$, defined by $((a \wedge b) \equiv (\overline{c} \wedge \overline{d})) \wedge ((\overline{a} \wedge \overline{b}) \equiv (c \wedge d))$

Reverse analogy expresses that "a differs from b as d differs from c", and conversely, paralogy expresses that "what a and b have in common, c and d have also". The inverse paralogy expresses a complete opposition between the pairs (a, b) and (c, d). The truth tables of these homogeneous proportions are recalled in Table 5, where we only show the lines leading to truth value 1. The next proposition, easily deducible from the definition, establishes a link between analogy, reverse analogy and paralogy (while inverse paralogy I is not related to the three others through a simple permutation):

Proposition 1 *$A(a, b, c, d) \leftrightarrow R(a, b, d, c)$ and $A(a, b, c, d) \leftrightarrow P(a, d, c, b)$.*

Clearly, the truth table of **A** (resp. **R**, **P** and **I**) is obtained as the conjunction of the truth tables of equivalences $\underline{\mathbf{A}}$ and $\overline{\mathbf{A}}$ (resp. $\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$, $\underline{\mathbf{P}}$ and $\overline{\mathbf{P}}$, and $\underline{\mathbf{I}}$ and $\overline{\mathbf{I}}$), given in Table 4. It is not the aim of this paper to investigate the whole set of properties of this class of Boolean formulas, nevertheless we recall the basic behavior that can be expected from these proportions, in Table 6.

To conclude this section, let us note that the analogical proportion is the Boolean counterpart of the numerical proportion $\frac{a}{b} = \frac{c}{d}$. As such, it is expected that this proportion will enjoy a property similar to the well-known rule of three for numerical proportions: knowing three of the variables, we can compute a 4th one in order to build up a proportion. This is obviously linked to our initial problem, where we want to derive some missing information. We will investigate this issue in Sect. 4.

Table 6 Properties of logical proportions

Property	Definition	# of proportions	Homogeneous
Full identity	$T(a, a, a, a)$	15	*A, R, P*
Reflexivity	$T(a, b, a, b)$	6	*A, P*
Reverse reflexivity	$T(a, b, b, a)$	6	*R, P*
Sameness	$T(a, a, b, b)$	6	*A, R*
Symmetry	$T(a, b, c, d) \to T(c, d, a, b)$	12	*A, R, P, I*
Central permutation	$T(a, b, c, d) \to T(a, c, b, d)$	16	*A, I*
All permutations	$\forall i, j, T(a, b, c, d) \to T(p_{i,j}(a, b, c, d))$	1	*I*
Transitivity	$T(a, b, c, d) \wedge T(c, d, e, f) \to T(a, b, e, f)$	54	*A, P*
Code independency	$T(a, b, c, d) \to T(\overline{a}, \overline{b}, \overline{c}, \overline{d})$	8	*A, R, P, I*

3.3 The Four Heterogeneous Logical Proportions

A remarkable property shared by the four homogeneous logical proportions is their independency with respect to a positive or negative encoding of the features, namely $T(a, b, c, d)$ holds true if and only if $T(\overline{a}, \overline{b}, \overline{c}, \overline{d})$ holds true. Indeed it is the same for the evaluation of these proportions to describe, e.g. the size of objects *A*, *B*, *C* and *D* by indicating if they are large or not, or by indicating if they are small (understood as “not large”) or not. There are only 8 logical proportions among the 120 that are code independent [18]. The four others are the heterogeneous proportions that we present now.

While homogeneous proportions are defined through equivalences between comparison indicators of the same type (either similarity or dissimilarity), heterogeneous proportions are hybrid proportions defined between indicators of opposite types, while the code independency property is preserved. Their logical expression is given below:

$$\mathbf{H_a} : ((a \wedge \overline{b}) \equiv (c \wedge d)) \wedge ((\overline{a} \wedge b) \equiv (\overline{c} \wedge \overline{d})),$$
$$\mathbf{H_b} : ((\overline{a} \wedge b) \equiv (c \wedge d)) \wedge ((a \wedge \overline{b}) \equiv (\overline{c} \wedge \overline{d})),$$
$$\mathbf{H_c} : ((a \wedge b) \equiv (c \wedge \overline{d})) \wedge ((\overline{a} \wedge \overline{b}) \equiv (\overline{c} \wedge d)),$$
$$\mathbf{H_d} : ((a \wedge b) \equiv (\overline{c} \wedge d)) \wedge ((\overline{a} \wedge \overline{b}) \equiv (c \wedge \overline{d})).$$

Their truth tables are shown in Table 7. Clearly, the truth table of $\mathbf{H_x}$ is obtained as the conjunction of the truth tables of equivalences $\underline{\mathbf{H_x}}$ and $\overline{\mathbf{H_x}}$ given in Table 4. It is stunning to note that these truth tables exactly involve the eight missing tuples of the homogeneous tables, i.e. those ones having an odd number of 0 and 1. The meaning of $\mathbf{H_x}$ is easy to grasp: “there is an intruder among the four values, which is not *x*”.

Table 7 H_a, H_b, H_c, H_d: the six patterns that make them true

$\mathbf{H_a}$	$\mathbf{H_b}$	$\mathbf{H_c}$	$\mathbf{H_d}$
1 1 1 0	1 1 1 0	1 1 1 0	1 1 0 1
0 0 0 1	0 0 0 1	0 0 0 1	0 0 1 0
1 1 0 1	1 1 0 1	1 0 1 1	1 0 1 1
0 0 1 0	0 0 1 0	0 1 0 0	0 1 0 0
1 0 1 1	0 1 1 1	0 1 1 1	0 1 1 1
0 1 0 0	1 0 0 0	1 0 0 0	1 0 0 0

Heterogeneous proportions satisfy obvious permutation properties. For instance, in $\mathbf{H_a}$, b and c, b and d, and c and d can be exchanged. Besides, note that, if we change d into $\overline{d}$ (and vice versa) in $\mathbf{H_a}$, $\mathbf{H_b}$, $\mathbf{H_c}$ and $\mathbf{H_d}$, are changed in **A**, **R**, **P** and **I**, respectively. It should come as surprise that they satisfy the same association properties as the homogeneous ones: for instance, any combination of two or three heterogeneous proportions is satisfiable by 4-tuples, and the conjunction $H_a(a, b, c, d) \wedge H_b(a, b, c, d) \wedge H_c(a, b, c, d) \wedge H_d(a, b, c, d)$ is not satisfiable. These facts contribute to make the heterogeneous proportions the perfect dual of the homogeneous ones.

3.4 Extension to $\mathbb{B}^n$

It is unlikely that we can represent objects in practice with only one property. In standard datasets, objects are represented by Boolean vectors, and it makes sense to extend the previous framework to $\mathbb{B}^n$. The simplest way to do this is to consider a definition involving componentwise proportions as follows (where T denotes any proportion and $\vec{a}, \vec{b}, \vec{c}$ and $\vec{d}$ are in $\mathbb{B}^n$):

$$T(\vec{a}, \vec{b}, \vec{c}, \vec{d}) \text{ iff } \forall i \in [1, n], T(a_i, b_i, c_i, d_i),$$

where $\vec{x} = (x_1, \ldots, x_n)$ for $x = a$, b, c and d. All the previous properties of Boolean proportions remain valid for their counterpart in $\mathbb{B}^n$: For instance, if $T = A$, then the *central permutation* property (see Table 6) is still valid:

$$T(\vec{a}, \vec{b}, \vec{c}, \vec{d}) \rightarrow T(\vec{a}, \vec{c}, \vec{b}, \vec{d}).$$

Table 8 A non-univocal proportion

0	0	0	0
0	1	0	1
1	0	1	0
1	0	1	1
1	1	1	0
1	1	1	1

4 Equation Solving Process

As said in the introduction, the main problem we want to tackle is to compute missing information starting from existing information. In the context of logical proportions, the equation solving problem can be stated as follows:

Given a logical proportion T and three Boolean values a, b, c, does a Boolean value x such that $T(a, b, c, x) = 1$ exist, and in that case, is this value unique?

This is an *equation solving problem*. First of all, it is easy to see that there are always cases where the equation has no solution. Indeed, the triple a, b, c may take $2^3 = 8$ values, while any proportion T is true for only six distinct valuations, leaving at least two cases with no solution. For instance, when we deal with analogy A, the equations $A(1, 0, 0, x) = 1$ and $A(0, 1, 1, x) = 1$ have no solution.

When the equation $T(a, b, c, x) = 1$ is solvable, a proportion T which has a unique solution will be said to be *d-univocal*. The following proportion is not d-univocal, as can be seen from its truth table (Table 8), despite the fact that it satisfies full identity, reflexivity, symmetry and transitivity: $((\overline{a} \wedge \overline{b}) \equiv (\overline{c} \wedge \overline{d})) \wedge ((\overline{a} \wedge b) \equiv (\overline{c} \wedge d))$.

We have the following result:

Proposition 2 *There are exactly 64 d-univocal proportions (including the homogeneous ones). They are the 4 homogeneous proportions, 8 conditional proportions, 12 hybrid proportions, 24 semihybrid proportions and 16 degenerated proportions.*

When moving to $\mathbb{B}^n$, the previous result remains valid; i.e. for 64 proportions, the equation $T(\vec{a}, \vec{b}, \vec{c}, \vec{x}) = 1$ has at most one solution.

In a similar manner, one may define proportions that are a-, b- or c-univocal. If we impose the proportion to be univocal with respect to two positions, we get 32 solutions. With respect to three positions, we get 16 solutions. Lastly, we have the following result:

Proposition 3 *There are only eight logical proportions that are a-, b-, c- and d-univocal: they are the four homogeneous and the four heterogeneous logical proportions.*

This result should not be a surprise, since a proportion that is not x-univocal is necessarily true both for a pattern with an odd number of 1 and 0, and for a pattern with an even number of 1 and 0. This excludes the homogeneous and the heterogeneous proportions. When considering only the homogeneous proportions A, R, P, I, we have a more detailed result [17–19]:

Proposition 4
The analogical equation $A(a, b, c, x)$ *is solvable iff* $(a \equiv b) \lor (a \equiv c)$ *holds.*
The reverse analogical equation $R(a, b, c, x)$ *is solvable iff* $(b \equiv a) \lor (b \equiv c)$ *holds.*
The paralogical equation $P(a, b, c, x)$ *is solvable iff* $(c \equiv b) \lor (c \equiv a)$ *holds.*
In each of the three above cases, when it exists*, the unique solution* x *is given by* $c \equiv (a \equiv b)$ *(or equivalently* $x = a \equiv b \equiv c$*).*
The inverse paralogical equation $I(a, b, c, x)$ *is solvable iff* $(a \not\equiv b) \lor (b \not\equiv c)$ *holds. In that case, the unique solution* x *is given by* $c \not\equiv (a \not\equiv b)$.

As we can see, the first three homogeneous proportions A, R, P behave similarly. Still, the conditions of their solvability differ. Moreover, it can be checked that at least two of these proportions are always simultaneously solvable. Besides, when they are solvable, there is a common expression that yields the solution. This again points out a close relationship between A, R and P. This contrasts with proportion I, which in some sense behaves in an opposite manner.

5 Inference Based on Logical Proportions

From an inference perspective, the equation solving property is the main tool: when we have three known objects A, B, C and another one D whose properties are unknown, but we have the information that $T(A, B, C, D)$ for a given d-univocal proportion, then we can compute D.

5.1 Existence of a Logical Proportion Linking Four Boolean Vectors

Now we are back to our initial framework of objects represented as a collection of Boolean properties. In that case, an object A is represented as a Boolean vector and belongs to $\mathbb{B}^n$, where n is just the cardinal of $\mathcal{U}$ (the whole set of properties). We can then provide a list of relevant problems we may need to solve:

1. Given four objects A, B, C, D, is there a proportion T, among the 120 proportions, such that $T(a, b, c, d)$?
2. If the answer to question 1 is "Yes", can we exhibit such a proportion?
3. If the answer to question 1 is "Yes", is such a proportion unique?

4. Given three objects A, B, C and a proportion T, can we compute an object D such that $T(A, B, C, D)$? (i.e. is T a d-univocal proportion?)

Reasoning about the first problem is relatively straightforward. As soon as we have the Boolean representation a, b, c, d of A, B, C, D, we have to consider the n valuations (a_i, b_i, c_i, d_i) corresponding to the components of a, b, c and d. Among these n valuations, some can be identical: so let us denote by m the number of distinct valuations among these n valuations. Obviously, if $m > 7$, there is no proportion such that $T(a, b, c, d)$, since only six valuations can lead to a given proportion being true.

If $m = 1$, we have exactly 45 candidate proportions. If $m = 2$, we have exactly 15 candidate proportions. If $m = 3$, we have three or six candidate proportions. If $m = 4$, the landscape is a bit different. Obviously, we cannot get more than six proportions, but some combinations lead to zero candidate proportions. In fact, we can have zero, one, three or six candidate proportions. For instance:

Lemma *A logical proportion cannot satisfy the class of valuation* $\{0111, 1011, 1101, 1110\}$ *or the class* $\{1000, 0100, 0010, 0001\}$.

Proof It is enough to show that this is the case for an equivalence between indicators. So let us consider such an equivalence $l_1 \wedge l_2 \equiv l_3 \wedge l_4$. If this equivalence is valid for $\{0111, 1011\}$, it means that its truth value does not change when we switch the truth value of the two first literals from 0 to 1: there are only two indicators for a and b satisfying this requirement: $a \wedge b$ and $\overline{a} \wedge \overline{b}$. On top of that, if this equivalence is still valid for $\{1101, 1110\}$, it means that its truth value does not change when we switch the truth value of the two last literals from 0 to 1: there are only two indicators for c and d satisfying this requirement: $c \wedge d$ and $\overline{c} \wedge \overline{d}$. Then the equivalence $l_1 \wedge l_2 \equiv l_3 \wedge l_4$ is just $a \wedge b \equiv c \wedge d$, $a \wedge b \equiv \overline{c} \wedge \overline{d}$, $a \wedge b \equiv \overline{c} \wedge \overline{d}$ or $\overline{a} \wedge \overline{b} \equiv \overline{c} \wedge \overline{d}$. None of these equivalences satisfies the whole class $\{0111, 1011, 1101, 1110\}$. The same reasoning applies for the other class. □

From a practical viewpoint, as soon as we observe this class appearing as a part of the n valuations (a_i, b_i, c_i, d_i), we know that there is no suitable proportion. If $m = 5$ or $m = 6$, we get zero or one candidate proportion.

5.2 Induction with Proportions

We can now adopt a viewpoint, similar in some sense to the k-nearest neighbors philosophy, where we infer unknown properties of a partially known object D starting from the knowledge we have about its other specified properties. This induction principle can be stated as follows (where J is a subset of $[1, n]$):

$$\frac{\forall i \in [1, n] \setminus J, T(a_i, b_i, c_i, d_i)}{\forall i \in J, T(a_i, b_i, c_i, d_i)}.$$

This can be seen as a continuity principle assuming that, if it is known that a proportion holds for some attributes, this proportion should still hold for the other attributes. It extends the inference principle defined in the case of the analogical proportion [22, 23], to d-univocal proportions.

The application of this inference principle to predict missing information for a given object D is straightforward. If we are able to find a triple of known objects A, B, C and a d-univocal proportion T such that $T(a_i, b_i, c_i, d_i)$ holds true for the attributes i belonging to $[1, n] \setminus J$, we then solve the equations $T(a_j, b_j, c_j, x_j)$ for every $j \in J$, and the (unique) solution is considered to be the value of d_j.

A particular case of this inductive principle has been implemented for classification purposes, i.e. when there is only one missing piece of information, which is the class of the object. Different approaches have been implemented, mainly using homogeneous proportions, all of them relaxing the induction principle to allow some flexibility: one may cite [1, 2] using analogical proportion only, using the four homogeneous proportions. In terms of accuracy, it appears that the results are similar when using $\mathbf{A}, \mathbf{R}, \mathbf{P}$, while $\mathbf{I}$ exhibits a different behavior. The prediction of missing values, using $\mathbf{A}$, has recently been experimented with successfully [4].

This leads us to the question of choosing the most suitable proportion for a given task. We can also consider other proportions among the 60 remaining univocal ones. These proportions express different kinds of regularities, which may be more or less often encountered in datasets. While the three homogeneous proportions $\mathbf{A}, \mathbf{R}, \mathbf{P}$, which play similar roles up to permutations, appear to be suitable for classification or for completing missing values, the involutive paralogy $\mathbf{I}$ seems to be of a different nature by expressing orthogonality between situations. The four heterogeneous proportions play similar roles with respect to the presence of intruder values. Clearly, if situations A, B and C are very close, enforcing $\mathbf{H_a}$, $\mathbf{H_b}$ or $\mathbf{H_c}$ for each feature will lead to a D very different from the three other situations. On the contrary, if A, B and C are quite different, D will not be more different from each of them. The respective roles of the univocal logical proportions is a topic for further research. Generally speaking, the proposed approach tends to indicate the possibility of transposing the idea of proportional reasoning from numerical settings to Boolean or nominal worlds.

6 Conclusion

We have outlined a systematic discussion of the interest of logical comparison indicators, and of their potential value for reasoning tasks such as classification or completion of missing information. We have emphasized the role of logical proportions in these reasoning tasks, exploiting similarity and dissimilarity indicators. In particular, we have identified the prominent place of homogeneous proportions and of heterogeneous proportions. They capture very different semantics, and appear to be basic tools for classification, intruder detection [20], and other reasoning tasks [3].

References

1. Bayoudh, S., Miclet, L., Delhay, A.: Learning by analogy: a classification rule for binary and nominal data. In: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'07), pp. 678–683 (2007)
2. Bounhas, M., Prade, H., Richard, G.: Analogical classification: a new way to deal with examples. In: Schaub, T., Friedrich, G., O'Sullivan, B. (eds.) Proceedings of the 21st European Conference on Artificial Intelligence (ECAI'14), Prague, 18–22 Aug. Frontiers in Artificial Intelligence and Applications, vol. 263, pp. 135–140. IOS Press (2014)
3. Correa, W., Prade, H., Richard, G.: When intelligence is just a matter of copying. In: Proceedings of the 20th European Conference on Artificial Intelligence, Montpellier, 27–31 Aug, pp. 276–281. IOS Press (2012)
4. Correa Beltran, W., Jaudoin, H., Pivert, O.: Estimating null values in relational databases using analogical proportions. In: Laurent, A., Strauss, O., Bouchon-Meunier, B., Yager, R.R. (eds.) Proceedings of the 15th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'14), Part III, Montpellier, 15–19 July. Communication in Computer and Information Science, vol. 444, pp. 110–119. Springer (2014)
5. Dubois, D., Prade, H.: Conditional objects as nonmonotonic consequence relationships. IEEE Trans. Syst. Man Cybern. **24**, 1724–1740 (1994)
6. Dubois, D., Prade, H.: From Blanché's hexagonal organization of concepts to formal concept analysis and possibility theory. Logica Universalis **6**(1–2), 149–169 (2012)
7. Dubois, D., Prade, H.: Possibility theory and formal concept analysis: characterizing independent sub-contexts. Fuzzy Sets Syst. **196**, 4–16 (2012)
8. Faris, J.A.: The Gergonne relations. J. Symbolic Logic 207–231 (1955)
9. Gentner, D., Holyoak, K.J., Kokinov, B.N.: The Analogical Mind: Perspectives from Cognitive Science. Cognitive Science, and Philosophy, MIT Press, Cambridge, MA (2001)
10. Gergonne, J.D.: Essai de dialectique rationnelle. Ann. Math. Pures Appl. 189–228 (1817)
11. Klein, S.: Culture, mysticism & social structure and the calculation of behavior. In: Proceedings of the 5th European Conference in Artificial Intelligence (ECAI'82), Orsay, France, pp. 141–146 (1982)
12. Miclet, L., Prade, H.: Handling analogical proportions in classical logic and fuzzy logics settings. In: Proceedings of the 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU'09), Verona, pp. 638–650. Springer, LNCS 5590 (2009)
13. Parsons, T.: The traditional square of opposition. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Summer 2015 edn. (2015)
14. Piaget, J.: Logic and Psychology. Manchester University Press (1953)
15. Piaget, J.: The Development of Thought: Equilibration of Cognitive Structures (trans. Rosin, A.). Viking, Oxford, UK (1977)
16. Prade, H., Richard, G.: Logical proportions—typology and roadmap. In: Hüllermeier, E., Kruse, R., Hoffmann, F. (eds.) Computational Intelligence for Knowledge-Based Systems Design: Proceedings of the 13th International Conference on Information Processing and Management of Uncertainty (IPMU'10), Dortmund, 28 June–2 July. LNCS, vol. 6178, pp. 757–767. Springer (2010)
17. Prade, H., Richard, G.: Reasoning with logical proportions. In: Lin, F.Z., Sattler, U., Truszczynski, M. (eds.) Proceedings of the 12th International Conference on Principles of Knowledge Representation and Reasoning, KR 2010, Toronto, 9–13 May, pp. 545–555. AAAI Press (2010)
18. Prade, H., Richard, G.: Homogeneous logical proportions: their uniqueness and their role in similarity-based prediction. In: Brewka, G., Eiter, T., McIlraith, S.A. (eds.) Proceedings of the 13th International Conference on Principles of Knowledge Representation and Reasoning (KR'12), Rome, 10–14 June, pp. 402–412. AAAI Press (2012)
19. Prade, H., Richard, G.: From analogical proportion to logical proportions. Logica Universalis **7**(4), 441–505 (2013)

20. Prade, H., Richard, G.: Homogenous and heterogeneous logical proportions. If CoLog J. Logics Appl. **1**(1), 1–51 (2014)
21. Prade, H., Richard, G.: Proportional reasoning in a Boolean setting. In: Booth, R., Casini, G., Klarman, S., Richard, G., Varzinczak, I.J. (eds.) International Workshop on Defeasible and Ampliative Reasoning (DARe@ECAI'14), Prague, 19 Aug. CEUR Workshop Proceedings, vol. 1212 (2014)
22. Stroppa, N., Yvon, F.: An analogical learner for morphological analysis. In: Online Proceedings of the 9th Conference on Computational Natural Language Learning (CoNLL-2005), pp. 120–127 (2005)
23. Stroppa, N., Yvon, F.: Analogical learning and formal proportions: Definitions and methodological issues. Technical report, ENST, June 2005
24. Tversky, A.: Features of similarity. Psychol. Rev. **84**, 327–352 (1977)