Reasoning with Logical Proportions

Henri Prade (1)
1. IRIT - University of Toulouse, 118 Rte de Narbonne, 31062 Toulouse Cedex 09, France
prade@irit.fr

Gilles Richard (2, 1)
2. British Institute of Technology and E-commerce
258-262 Romford Road, London E7 9HZ, UK
grichard@bite.ac.uk

Abstract

By logical proportion, we mean a statement that expresses a semantical equivalence between two pairs of propositions. In these pairs, each element is compared to the other in terms of similarities and/or dissimilarities. An example of such a proportion is the well known analogical proportion: *a is to b as c is to d*. Analogical proportions have been recently characterized in logical terms, but there are many other proportions that are worth of interest. Some of them can be related to the analogical pattern, others are related to semantical equivalence between conditional objects and express statements such as *a ressembles to b and differs from b in the same way as c with respect to d*. We show that there are 5 *direct* proportions, including the analogical one and 4 others having a conditional object flavor, where the change (if any) from a to b goes in the same direction as the change from c to d (if any), together with 5 *reverse* proportions obtained by switching c and d. Moreover, there exists only one auto-reverse proportion called *paralogy* and stating that *what a and b have in common, c and d have it as well*. It is then established that there is none other proportion than these ones (with the exception of 4 degenerated ones) that satisfies a natural "full identity" requirement. The paper proposes a structured and unified view of these logical proportions and discusses their characteristic properties. It extends previous works where only proportions related to analogy were considered. It also explores the use of these logical proportions in transduction-like inference, where new items are classified on the basis of already classified items without trying to induce a generic model, considering similarities and differences between items only. Taking advantage of different proportions, a transduction procedure is proposed.

Reasoning may handle generic pieces of knowledge that apply to a class of situations, as well as particular facts. Deductive reasoning goes from generic to generic; or applies generic to particular (e.g. "all men are mortal" and "Socrates is a man" then "Socrate is mortal"), while *induction* infers from particular to generic ("a, b, c are mortal men" then "all men are mortal"). Obviously, after an inductive step, which allows us to establish generic rules from a limited set of examples, we can apply these rules in a deductive way to a new case, "Socrate" for instance, and again deduce "Socrate is mortal": these generic rules are then used as a prediction tool to be applied to not previously seen data. Nevertheless, at least from a philosophical viewpoint, this way to reason is not absolutely satisfactory since, basically we go from particular to particular, but going through generic. Moreover, when we only have a limited set of observable situations, it is not a straightforward task to extract generic rules that apply to all possible cases.

In fact, there exists some well-known inference processes allowing to go from particular to particular in a direct way. This type of reasoning process is sometimes named as *transduction* (this word was coined by (Gammerman, Vovk, and Vapnik 1998), but the idea at least dates back to (Russell 1912)). Case-based reasoning is typically a transductive process since, starting from a limited sample of previously observed situations/cases, we try to match a current incomplete case with a similar, previously observed one in order to guess the missing part. More generally analogical reasoning puts in parallel two particular situations that hold between related items. A noticeable form of analogical reasoning is based on analogical proportions that are statements of the form *a is to b as c is to d*. A very specific instance of analogical proportion-based reasoning (which the name 'proportion' comes from) is the well known "rule of three" where in a numerical setting we extrapolate the value of a number x starting from the assumption that $\frac{a}{b} = \frac{c}{x}$ holds.

It has been recently advocated in (Miclet and Prade 2009; Prade and Richard 2010), that such extrapolation reasoning could be formalized for Boolean proportions as well. In the latter work, the authors highlight the existence of two other proportions, namely *reverse analogy* and *paralogy*, also suitable for inference purpose. In fact, a careful reexamination of this work has led us to consider other proportions first studied in this paper, which are related to non monotonic reasoning which may cover other inference cases of interest when used in a transduction procedure.

In this paper, we first study the existing proportions in detail. Through basic properties requirements, we identify those ones that may provide the basis for new inference principles, suitable for transduction. In the next section, we recall the 3 proportions related to analogy, and we highlight the existence of other noticeable proportions. Then we provide Boolean interpretations for all the previous proportions, and investigate their formal properties. In the second half of the paper, we examine how formal proportions could be the basis of a general transductive paradigm. Finally, we overview related works and we conclude by indicating some lines for further research.

Logical proportions

Informally speaking, logical proportions are expressions that involve 4 terms, here denoted a, b, c, d, and where the relation existing between a and b is equated to the relation between c and d. Generally speaking, a *logical* proportion establishes an equivalence between similarities / dissimilarities existing between a and b with similarities / dissimilarities existing between c and d. In the following, we choose to denote this equivalence by "::" (in agreement with current practice in analogical reasoning).

To start our investigation, the most suitable viewpoint is to consider the items a, b, c, d as representing sets of binary features belonging to a universe U (i.e. an item is viewed as the set of binary features in U possessed by the item). In that case, each item can be viewed as a subset of U. Then, the dissimilarity between a and b can be appreciated in terms of $a \cap \overline{b}$ and / or $\overline{a} \cap b$, where $\overline{a}$ denotes the complement of a in U, while the similarity is estimated by means of $a \cap b$ and / or $\overline{a} \cap \overline{b}$. Let us briefly review the results of (Prade and Richard 2009; 2010) where analogy, reverse analogy and paralogy have been fully investigated. This will be of great help to understand the place of the other proportions of interest.

Analogy, reverse analogy and paralogy

An analogical proportion, usually denoted $a : b :: c : d$, focuses on differences and should hold when the differences between a and b and between c and d are the same. This is formally defined (Miclet and Prade 2009) as:

$$a \cap \overline{b} = c \cap \overline{d} \text{ and } \overline{a} \cap b = \overline{c} \cap d \quad (A) \tag{1}$$

Then, two other proportions have been introduced (Prade and Richard 2009): *reverse analogy*, the sister proportion denoted $a \,!\, b :: c \,!\, d$, which is defined as:

$$a \cap \overline{b} = \overline{c} \cap d \text{ and } \overline{a} \cap b = c \cap \overline{d} \quad (R) \tag{2}$$

If instead of focusing on differences, we emphasize similarities, we get a new proportion denoted $a \,;\, b :: c \,;\, d$ called *paralogy*, whose definition is:

$$a \cap b = c \cap d \text{ and } \overline{a} \cap \overline{b} = \overline{c} \cap \overline{d} \quad (P) \tag{3}$$

Obviously the second part of this definition is equivalent to $a \cup b = c \cup d$. But, for the sake of homogeneity, we prefer to only use $\cap$ in the definition. The two following properties can be easily checked; see (Prade and Richard 2009; 2010) respectively:

Property 1 $a : b :: c : d$ *iff* $a \,!\, b :: d \,!\, c$ *iff* $a \,;\, d :: c \,;\, b$

Property 2 *When the analogical proportion holds for a tuple* (a, b, c, d)*, it also holds for a set of 7 other permutations of this tuple, obtained by exchanging the two mid terms or permuting the two first terms with the two last ones, or repeating these two operations alternatively. The two other proportions can be also equivalently written in 8 different ways.*

The first property establishes a strong link between the three proportions: we can go from a proportion to another one just by permuting some parameters. The second property tells us, in particular, that any proportion that holds for (a, b, c, d), can be rewritten equivalently through an appropriate permutation of (a, b, c, d) in such a way that the last term is a, b, c, or d. This remark will be of interest when dealing with the equation solving problem considered later in this paper. Let us now examine other possibilities for defining proportions.

Other interesting proportions

As explained, analogy and reverse analogy focus on dissimilarities, while paralogy refers to similarities. But other options become available if we accept to mix similarity and dissimilarity constraints.Starting from the definitions of analogical and paralogical proportions, we have several options:

- $a \cap b = c \cap d$ and $a \cap \overline{b} = c \cap \overline{d}$, denoted $b|a :: d|c$
- $a \cap b = c \cap d$ and $\overline{a} \cap b = \overline{c} \cap d$, denoted $a|b :: c|d$
- $\overline{a} \cap \overline{b} = \overline{c} \cap \overline{d}$ and $\overline{a} \cap b = \overline{c} \cap d$, denoted $\overline{b}|\overline{a} :: \overline{d}|\overline{c}$
- $\overline{a} \cap \overline{b} = \overline{c} \cap \overline{d}$ and $a \cap \overline{b} = c \cap \overline{d}$, denoted $\overline{a}|\overline{b} :: \overline{c}|\overline{d}$

As for reverse analogy, one may switch c and d to get:

- $a \cap b = c \cap d$ and $a \cap \overline{b} = \overline{c} \cap d$, denoted $b|a :: c|d$
- $a \cap b = c \cap d$ and $\overline{a} \cap b = c \cap \overline{d}$, denoted $a|b :: d|c$
- $\overline{a} \cap \overline{b} = \overline{c} \cap \overline{d}$ and $\overline{a} \cap b = c \cap \overline{d}$, denoted $\overline{b}|\overline{a} :: \overline{c}|\overline{d}$
- $\overline{a} \cap \overline{b} = \overline{c} \cap \overline{d}$ and $a \cap \overline{b} = \overline{c} \cap d$, denoted $\overline{a}|\overline{b} :: \overline{d}|\overline{c}$

This makes 8 new proportions. We use the notation $|$ for all of them, since the constraints that define them are reminiscent of the semantical equivalence between conditional objects. Indeed, it has been advocated by (Dubois and Prade 1994) that a rule "if a then b" is not a two-valued entity, but a three valued one that is called 'conditional object' and denoted $b|a$. To see it, consider a database containing item descriptions. If the rule "if a then b" is to be evaluated against this database, it clearly creates a partition of 3 kinds of situation, namely:

- the set of items, examples of the rule, that are characterized by $a \cap b$.
- the set of items, counter-examples of the rule, that are characterized by $a \cap \overline{b}$.
- the set of items for which the rule is irrelevant, that are characterized by $\overline{a}$.

Each case should be encoded by means of a different truth-value. The two first cases correspond to the usual truth-values "true" and "false" respectively. The third case corresponds to a third truth-value. A rule is thus modeled as a pair of disjoint sets representing its examples and its counter-examples, namely $(a \cap b,\ a \cap \overline{b})$. Interestingly enough, this tri-valued representation also tolerates non-monotonicity, and can offer a tri-valued semantics to nonmonotonic consequence relations obeying postulates of the Preferential entailment system of (Kraus, Lehmann, and Magidor 1990); see (Dubois and Prade 1994; Benferhat, Dubois, and Prade 1997). It is indeed intuitively satisfying to consider that a rule r_1 = "if a then b" is safer than a rule r_2 = "if c then d", if r_1 has more examples and less counterexamples than r_2 (in the sense of inclusion). The semantical equivalence between rules then write $a \cap b = c \cap d$ and $a \cap \overline{b} = c \cap \overline{d}$, which clearly entails $a = c$, which agrees with the intuition that in order to be equal two rules should have the same set of applicability situations. This idea of a rule as a "tri-event" actually goes back to (De Finetti 1936), where this equivalence is defined for the first time. This idea is also the backbone of De Finetti's approach to conditional probability (De Finetti 1974). The three-valued representation of a rule also casts new light on "paradoxes of confirmation", a question which goes back to Nicod and Hempel (Hempel 1965), where the rules "all ravens are black" and "all non black items are not ravens" (which might seem to be equivalent from a material implication point of view), have the same counter-examples (ravens that are not black), but not the same examples, and then can no longer be confirmed by the same observations, as soon as we leave the pure material implication understanding; see (Benferhat et al. 2008). Still, we have to keep in mind that we are not interested here in generic (default) rules, but we are rather putting in parallel and comparing particular items described in terms of binary features (even if these items may be the basis of inductive generalizations). Leaving the set interpretation, let us now examine the Boolean interpretation.

Boolean interpretation

It is easy to adopt a Boolean viewpoint, considering our proportions as Boolean operators. It is then sufficient to translate the set operators into Boolean connectors: $\cap$ becomes $\wedge$ and $\overline{a}$ is just the negation of a. Equality is changed into $\equiv$. Thus, the Boolean connective representing the (direct) analogical proportion $a : b :: c : d$ can be written as the logical formula:

$$a : b :: c : d = ((a \wedge \neg b) \equiv (c \wedge \neg d)) \wedge ((\neg a \wedge b) \equiv (\neg c \wedge d))$$

This formula is true for the 6 truth value assignments of a, b, c, d appearing in the left part of Table 1, and is false for the $2^4 - 6 = 10$ remaining possible assignments. The most "visual" way to understand the behavior of the proportions is to examine their truth tables that are given in Table 1, Table 2 and Table 3. It is worth noticing that all the 11 proportions, viewed as Boolean connectives in the above sense, hold true for only 6 cases, and this is why we have restricted Tables 1, 2, 3 to the 4-tuples corresponding to these cases. Let us note that both Table 2 and Table 3 highlight the fact that their corresponding proportions behave similarly inside each table on the 4 first lines, which are respectively the 4 first lines of direct analogy and reverse analogy tables (see Table 1). Moreover for all of them, the first two lines express that $T(a, a, a, a)$ holds, where T denotes the considered proportion; we come back to the meaning of this property later

Table 1: Analogy-related proportions

(Direct) Analogy $\mathbf{a : b :: c : d}$				*Rev. Analogy* $\mathbf{a \,!\, b :: c \,!\, d}$				*Paralogy* $\mathbf{a \,;\, b :: c \,;\, d}$			
a	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**
0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1
0	1	0	1	0	1	1	0	0	1	0	1
1	0	1	0	1	0	0	1	1	0	1	0
0	0	1	1	0	0	1	1	0	1	1	0
1	1	0	0	1	1	0	0	1	0	0	1

Table 2: Direct proportions (other than analogy)

$\mathbf{b \mid a :: d \mid c}$				$\mathbf{a \mid b :: c \mid d}$				$\mathbf{\overline{b} \mid \overline{a} :: \overline{d} \mid \overline{c}}$				$\mathbf{\overline{a} \mid \overline{b} :: \overline{c} \mid \overline{d}}$			
a	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
0	0	0	1	0	0	1	0	1	0	1	1	0	1	1	1
0	1	0	0	1	0	0	0	1	1	1	0	1	1	0	1

in this paper. The 3rd and 4th lines in Table 2 (resp. Table 3) express that the change from c to d goes in the same (resp. reverse) sense with respect to the change from a to b.

Then natural questions arise. How many proportions can be defined through 6 4-tuples of binary values which are the 4-tuples for which they are true, as in the above tables ? Does any set of 6 distinct 4-tuples of binary values give birth to a proportion? Generally speaking, are there proportions, other than the 11 proportions already identified, which are worth considering? Before exploring all the possibilities, let us examine the construction process of proportions first.

Definition 1 *A logical proportion is defined via a couple of equalities of the form* $\alpha \cap \beta = \gamma \cap \delta$, $\alpha \in \{a, \overline{a}\}, \beta \in \{b, \overline{b}\}, \gamma \in \{c, \overline{c}\}, \delta \in \{d, \overline{d}\}$*, and where*
- i) the equalities have to be distinct,
- ii) the order of the equalities is irrelevant.

This leaves $(16 \times 16 - 16)/2 = 120$ possibilities. Indeed, there are 4 choices for the left member of the first equality constraint ($a \cap b$, $a \cap \overline{b}$, $\overline{a} \cap b$, $\overline{a} \cap \overline{b}$), leading to $4 \times 4 = 16$ equalities, and to $16 \times 16 = 256$ pairs of equalities, from which we have to subtract 16 pairs where the two constraints are identical. Finally, we divide by 2 since the order of the equalities is irrelevant.

Besides, it is worth pointing out that there are sets of 6 4-tuple assignments associated to 1 in the truth table that do not correspond to any formal proportion (simply because, from a purely combinatorial viewpoint, there are many more such assignments than 120). For instance, it can be checked that the set $\{0000, 1111, 0011, 1100, 1010, 0110\}$ does not correspond to any formal proportion. As can be seen, this set of 6 assignments includes the first four lines of the tables of direct and reverse analogical proportions, which correspond to their common part (see Table 1), and just "mix" their two last lines, expressing respectively a direct and a

Table 3: Reverse proportions (other than rev. analogy)

$\mathbf{b} \mid \mathbf{a} :: \mathbf{c} \mid \mathbf{d}$				$\mathbf{a} \mid \mathbf{b} :: \mathbf{d} \mid \mathbf{c}$				$\overline{\mathbf{b}} \mid \overline{\mathbf{a}} :: \overline{\mathbf{c}} \mid \overline{\mathbf{d}}$				$\overline{\mathbf{a}} \mid \overline{\mathbf{b}} :: \overline{\mathbf{d}} \mid \overline{\mathbf{c}}$			
a	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0
1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1
0	0	1	0	0	0	0	1	1	0	1	1	0	1	1	1
0	1	0	0	1	0	0	0	1	1	0	1	1	1	1	0

reverse change when going from c to d, with respect to the change taking place in the pair (a, b).

All the 120 logical proportions do not seem equally interesting. Let us notice that the 11 proportions obtained in the two previous subsections are true for the assignments 0000 and 1111. In case of the analogical proportion, it can be read "a is to a as a is to a" (whatever the value of a). A proportion T that holds as true for $(0, 0, 0, 0)$ and $(1, 1, 1, 1)$ is said to satisfy the *full identity* property, which is formally expressed as $T(a, a, a, a)$. More generally, since our intended use of these proportions is based on the observation of similarities and differences, there is indeed no point to distinguish the pair (a, a) from the pair (a, a). In some sense, *full identity* is a very natural requirement and is a kind of minimal property expected to be satisfied. Indeed in such a case, identity holds both inside each pair and between the pairs. Obviously, this is not the case for the pairs (a, a) and (b, b) and, as we will see, we cannot expect $T(a, a, b, b)$ to hold for every proportion.

It does not come as a surprise that it is possible to define logical proportions following the same scheme of construction, as in the previous subsection, which do not satisfy *full identity*. For instance,

- $a \cap b = c \cap \overline{d}$ and $a \cap \overline{b} = c \cap d$ holds for 0000, but not for 1111.

- $a \cap b = c \cap d$ and $\overline{a} \cap \overline{b} = \overline{c} \cap d$ holds for 1 1 1 1, but not for 0 0 0 0.

- any proportion defined via $a \cap b = \overline{c} \cap \overline{d}$ does neither hold for 0000, nor for 1111.[1]

However as explained above, it makes sense to restrict ourselves to proportions satisfying *full identity*. Finding the number of such proportions is a rather tedious exercise and relies on a careful investigation about the diverse ways to build up proportions such that their two defining conditions of the form $\alpha \wedge \beta \equiv \gamma \wedge \delta$, where $\alpha \in \{a, \neg a\}, \beta \in \{b, \neg b\}, \gamma \in \{c, \neg c\}, \delta \in \{d, \neg d\}$ hold as true for entries $(a, b, c, d) = (1, 1, 1, 1)$ and $(a, b, c, d) = (0, 0, 0, 0)$. This leads to the following result:

Property 3 *There are exactly 15 logical proportions satisfying full identity. They include the 11 logical proportions previously examined, plus the 4 following ones:*
$\overline{a} \cap b = \overline{c} \cap d$ *and* $\overline{a} \cap b = c \cap \overline{d}$ *(1)*
$\overline{a} \cap b = \overline{c} \cap d$ *and* $a \cap \overline{b} = \overline{c} \cap d$ *(2)*
$a \cap \overline{b} = c \cap \overline{d}$ *and* $\overline{a} \cap b = c \cap \overline{d}$ *(3)*
$a \cap \overline{b} = c \cap \overline{d}$ *and* $a \cap \overline{b} = \overline{c} \cap d$ *(4)*

Table 4: The 4 generated proportions satisfying full identity

$1.\mathbf{b} \subseteq \mathbf{a}\ \mathbf{c} = \mathbf{d}$				$2.\mathbf{a} = \mathbf{b}\ \mathbf{d} \subseteq \mathbf{c}$				$3.\mathbf{a} = \mathbf{b}\ \mathbf{c} \subseteq \mathbf{d}$				$4.\mathbf{a} \subseteq \mathbf{b}\ \mathbf{c} = \mathbf{d}$			
a	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0
1	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0
1	0	1	1	1	1	1	0	1	1	0	1	0	1	1	1

We exhibit their truth table in Table 4. It can be easily seen that the first pair of constraints entails $b \subseteq a$ and $c = d$, the second $a = b$ and $d \subseteq c$, the third $a = b$ and $c \subseteq d$, and the fourth $a \subseteq b$ and $c = d$. So these proportions can be regarded as degenerated ones, but it is worth noticing that they cover a *classical entailment* situation inside one of the pairs each time. Note also that the tables of these four proportions share the same 4 first lines, lines that can be also found in the tables of direct analogical and reverse analogical proportions.

Thus, if we ignore these 4 "degenerated" proportions, we are back to the 11 logical proportions initially generated by equating similarities on the one hand, and dissimilarities on the other hand, between pairs.

Properties of logical proportions

Let T denote a Boolean proportion. Apart from the full identity property, $T(a, a, a, a)$ with $a \in \{0, 1\}$, which has been already discussed, there are other properties that are worth of interest for proportions. A property also related to the idea of identity, but stronger than full identity, is for $a, b \in \{0, 1\}$

- $T(a, a, b, b)$ (identity)

Clearly, *full identity* postulate simply forces T to hold for both assignments 0000 and 1111, while *identity* postulate forces T to also hold for assignments 0011 and 1100: this is definitely a stronger requirement. As already said, the 15 proportions of Tables 1 to 4 satisfy $T(a, a, a, a)$, while only direct analogy, reverse analogy, and the 4 "degenerated" proportions appearing in Property 3 satisfy *identity*.

This remark singles out paralogy which rather satisfies

- $T(a, b, a, b)$ (reflexivity)
- $T(a, b, b, a)$ (reverse reflexivity)

together with *full identity*. Clearly, *reflexivity* postulate forces T to hold for 0101 and 1010, while *reverse reflexivity* postulate forces T to hold for 0110 and 1001. Thus, direct analogy also satisfies *reflexivity* (and violates *reverse reflexivity*), while reverse analogy satisfies *reverse reflexivity* (and violates *reflexivity*). Paralogy violates *identity*.

Furthermore, the notation "::" suggests that a logical proportion should be symmetric in the sense that we could exchange the left hand side and the right hand side of "::" without changing the truth value of the proportion. Namely

[1] Let us note that there is a specific proportion defined by $a \cap b = \overline{c} \cap \overline{d}$ and $\overline{a} \cap \overline{b} = c \cap d$, which obviously does not hold for 1111 nor 0000. This proportion is obtained from the paralogy definition by switching the second half of the equalities. This is why we may call it *inverse paralogy*. It expresses that "what a and b have in common, c and d do not have it, and vice-versa". Due to lack of space, we will not investigate this proportion more deeply in this paper.

Table 5: Postulates

Analogy A	**Rev. analogy R**	**Paralogy P**
$a : b :: a : b$	$a \,!\, b :: b \,!\, a$	$a \,;\, b :: a \,;\, b$
$a{:}b :: c{:}d \Rightarrow c{:}d :: a{:}b$	$a!b :: c!d \Rightarrow c!d :: a!b$	$a;b :: c;d \Rightarrow c;d :: a;b$
$a{:}b :: c{:}d \Rightarrow a{:}c :: b{:}d$	$a!b :: c!d \Rightarrow c!b :: a!d$	$a;b :: c;d \Rightarrow a;b :: d;c$

Definition 2 *Let T denote a Boolean proportion. T satisfies the symmetry postulate iff $T(a,b,c,d) \Rightarrow T(c,d,a,b)$*

It can be checked that the direct proportions, including analogy, as well as reverse analogy and paralogy satisfy *symmetry*. However, the reverse proportions of Table 3 violates *symmetry* (in fact *symmetry* exchanges the proportions in the table).

Property 4 *Only 7 proportions satisfy full identity and symmetry: namely A, P, R and the 4 (direct) conditional proportions of Table 2.*

Another property which is worth discussing is independence with respect to encoding. The 0/1 coding is just a matter of convention and could be replaced with the reverse coding 1/0, i.e. for instance if a proportion holds for 1001, then it should hold for 0110 also. We call this property *mirroring* for short. Still, observe that, although apparently desirable for logical proportions, this very strong property is violated in classical reasoning for many connectives, since, e.g., $a \to b$ is equivalent to $\neg b \to \neg a$, and not to $\neg a \to \neg b$. This tends to indicate that the mirroring property is indeed interesting, but not necessarily compulsory for having the considered logical proportion useful. In other words, the very meaning of the proportions that satisfy *mirroring* does not change when we switch from one coding system to the other one. This can be formalized as follows.

Definition 3 *Let T denote a Boolean proportion. T satisfies the mirroring property iff $T(a,b,c,d) \Rightarrow T(\overline{a}, \overline{b}, \overline{c}, \overline{d})$.*

We have the following result:

Property 5 *Among the logical proportions satisfying full identity, only 3 proportions satisfy mirroring, namely analogy, reverse analogy and paralogy.*

This result highlights the fact that analogy, reverse analogy and paralogy play a particular role among all the proportions. These 3 proportions do not only satisfy *mirroring* (and *symmetry*), but are also characterized by permutation properties. In fact, the postulate-based definitions of analogy, reverse analogy and paralogy have been given in (Prade and Richard 2010) and are recalled in Table 5. Each proportion is defined via 3 postulates. Obviously, the second postulate is just the previous *symmetry* property. The third postulate specifies under which permutation the proportion is preserved (e.g., central permutation for analogical proportions). As a consequence of the postulates, we retrieve Property 1 that establishes a strong link between the three kinds of analogy-related proportion. In agreement with their third postulate, it appears that A satisfies $T(a,b,a,b)$ and $T(a,a,b,b)$ by central permutation, R satisfies $T(a,b,b,a)$ and $T(a,a,b,b)$ by permuting the first and the third component and symmetry, and finally P satisfies $T(a,b,a,b)$ and $T(a,b,b,a)$ by permuting the two last components. But we have the following property:

Property 6 *If A is an analogy, then $\not\models A(a,b,b,a)$;*
If R is a reverse-analogy, then $\not\models R(a,b,a,b)$;
If P is a paralogy, then $\not\models P(a,a,b,b)$.

It simply means that, for instance, we can build up a model of analogical proportion A (i.e. satisfying the 3 required postulates for analogy) such that $A(a,b,b,a)$ does not hold in the given model.

In the set and Boolean models, analogy and paralogy are transitive, reverse analogy is not transitive and satisfies the property:

$$a \,!\, b :: c \,!\, d \text{ and } c \,!\, d :: e \,!\, f \Rightarrow a : b :: e : f$$

This means that *transitivity*, when it holds for a proportion, is not a consequence of the corresponding postulates in Table 5, but a peculiarity proper to some interpretations such as the Boolean one; see (Prade and Richard 2010).

Equation solving with formal proportions

Let us now turn to an effective way of using these proportions, as we apply the rule of three for numerical proportions. The rule of three involves 3 items a, b, c and allows to find out a fourth one x such that the numerical proportion $\frac{a}{b} = \frac{c}{x}$ holds. This is exactly the kind of process we want to clone for the logical proportions, but with more freedom since we have the choice between different proportions. Given only 3 Boolean values a, b, c, is it possible to find out a proportion T among the 7 symmetrical proportions of Tables 1 and 2 (we mainly restrict ourselves to these ones in the following), and an item $x \in \{0,1\}$ such that $T(a,b,c,x)$ holds? In the case of a positive answer, a new question arises: is the solution unique? Given a proportion T, the problem of finding x such that $T(a,b,c,x)$ holds will be referred as "equation solving".

For example, starting from $1, 0, 0$ we can check that $1 \,!\, 0 :: 0 \,!\, 1$), i.e. a reverse analogy holds between these 4 items. If we rather consider a direct proportion in Table 2, the solution may not exist (e.g., for $(a,b,c) = (1,1,0)$), or the solution may be not unique for two of them. Indeed, taking for instance the proportion $b \mid a :: x \mid c$, $(a,b,c) = (0,0,0)$ may be completed both by $x = 0$ or by $x = 1$, and it is the same for $(a,b,c) = (0,1,0)$. This might be found as agreeing with the idea of implicit tolerance associated with default rules, however this is not specific of direct (or reverse) proportions as we shall see. If we focus on the "perfect" proportions of Table 1, namely analogy, reverse analogy and paralogy, the solution becomes unique, when it exists, as now restated in logical terms (these results have also their counterparts in a set interpretation setting; see (Prade and Richard 2009)).

Property 7 *The analogical equation $a : b :: c : x$ is solvable iff $((a \equiv b) \vee (a \equiv c)) = 1$. In that case, the unique solution is $x = a \equiv (b \equiv c)$ (Miclet and Prade 2009).*
The reverse analogical equation $a \,!\, b :: c \,!\, x$ is solvable iff $((b \equiv a) \vee (b \equiv c)) = 1$. In that case, the unique solution is $x = b \equiv (a \equiv c)$.
The paralogical equation $a \,;\, b :: c \,;\, x$ is solvable iff $((c \equiv b) \vee (c \equiv a)) = 1$. In that case, the unique solution is $x = c \equiv (a \equiv b)$.

Despite these apparently intricate conditions, it is generally easy to find out if there is a solution by matching the pattern (a, b, c) at hand with Table 1 above. For instance $(0, 1, 1)$ is solvable for reverse analogy and paralogy with 0. It can be easily deduced that:

Property 8 *For any 3-uple (a, b, c), there exists at least 2 distinct proportions T among A, R, and P such that $T(a, b, c, x)$ is solvable and the solution is the same.*

Let us remember that (Klein 1982), who is the author of the first attempt to give a logical-like reading of analogical proportions, used an implicit definition that in fact covers the 3 proportions together, since the unique solutions of their equations are the same (due to the associativity of $\equiv$). His initial proposal has been refined here by splitting it in several definitions, providing a much more accurate view of the analogical proportion, and of the related proportions.

Besides, the equation "find x such that $T(a, b, c, x)$ holds", may have no solution for the reverse proportions of Table 3, or several solutions for two of them, (as in the case of direct proportions of Table 2). It is exactly the same for the "degenerated" proportions of Table 4 that always lead to a unique solution for two of them (1 and 4), while the two others (2 and 3) may have several solutions in some cases.

Reasoning with logical proportions

Let us now investigate the potential use of these proportions to design a general inference mechanism where we apply a kind of "continuity" principle. When a logical proportion holds for some observable features describing items a, b, c, and d, we may assume that the same proportion still holds between the other features of d, and the corresponding, observable and known features of a, b, and c. This simple idea, when implemented as a prediction rule, provides the basis for a transductive engine.

The transductive reasoning problem

We focus here on a simple binary classification task. Each element of the problem universe is represented as a vector of n binary features, i.e. our input space is $X = \mathbb{B}^n$. Moreover, each element of X (representing one item, or several distinct items, in the problem universe) belongs to a class encoded by 0 or 1. In other words, we try to classify an incoming item and the output space is $\{0, 1\}$). A set S of already classified items is supposed to be given, which is then just a subset of $X \times \{0, 1\} = \mathbb{B}^{n+1}$, i.e., a set of labeled examples $(x, cl(x))$. In the following, it is assumed that the class of an element is unique. This means that $\nexists (x, c)$ and $(x, c') \in S$, with $c \neq c'$. This agrees with the existence of an underlying classifying function cl from X to $\{0, 1\}$, this function being only known for the elements in S. The basic idea underlying transductive reasoning is that one tries to predict the class of a new piece of data on the basis of the previously observed data S, without any attempt to guess a generic model for the observed data (which would be induction). In the approach proposed here, given a new instance d whose class is unknown, one looks for triples (a, b, c) of examples in the input space S such that some logical proportion T holds simultaneously between the binary features associated to (a, b, c, d) (i.e. T holds on the input space) in order to predict the class of d by applying the same proportion T to the output space $\{0, 1\}$ (i.e. the space of classes).

In the absence of external knowledge, the only thing that we have at our disposal for predicting the class of a new element d is the *observation of the behavior* of the data at hand, trying to relate the *variations of $cl(x)$ with the variation of the features of x.*

This first amounts to examine in what respect vectors are similar or different, and this is exactly what logical proportions are designed for. We use the following notations:

$$\langle a, b \rangle \;=\; \{j \in [1, n], a_j = b_j = 1\} \;\; \text{(agreement set)}$$

$$\rangle a, b \langle \;=\; \{j \in [1, n], a_j \neq b_j\} \;\; \text{(disagreement set)}$$

where $[1, n]$ denotes here an interval in the integers. Considering the Boolean lattice $\mathbb{B}$ as ordered with $0 < 1$, we can refine our definitions with:

$$\rangle a, b \rangle \;=\; \{j \in [1, n], a_j > b_j\}$$

$$\langle a, b \langle \;=\; \{j \in [1, n], a_j < b_j\}$$

Clearly, $\rangle a, b \langle = \rangle a, b \rangle \cup \langle a, b \langle$, $\langle a, b \rangle \cup \rangle a, b \langle = [1, n]$ and $\langle a, b \rangle = [1, n]$ iff $a = b$. Let us now consider four vectors a, b, c, d. If A, P, R stands respectively for analogical, paralogical and reverse analogical proportion, our previous definitions lead to:

$$A(a, b, c, d) = 1 \text{ iff } \rangle a, b \rangle = \rangle c, d \rangle \text{ and } \langle a, b \langle = \langle c, d \langle$$

$$P(a, b, c, d) = 1 \text{ iff } \langle a, b \rangle = \langle c, d \rangle \text{ and } \rangle a, b \langle = \rangle c, d \langle$$

$$R(a, b, c, d) = 1 \text{ iff } \rangle a, b \rangle = \langle c, d \langle \text{ and } \langle a, b \langle = \rangle c, d \rangle$$

For a new item b, one may consider that the obstacle to be classified as a previously labeled element a is just the set $\rangle a, b \langle$, since if this set is empty then we classify b as a. A first natural idea is to try to minimize this set, and more generally the distance between a and b, which is reminiscent of the principle underlying k-nn-like (k nearest neighbors-like) methods. Namely in that view, when $k = 1$, letting $|\ |$ denote cardinality, one takes

$$cl(b) = cl(argmin_x\ |\rangle b, x \langle|).$$

General inference principle

As already said, it is assumed that the items we are dealing with are vectors $a = (a_1, \ldots, a_n)$ of Boolean values that encode the values of n binary attributes associated to a situation. Starting from a 4-tuple of vectors (a, b, c, d), we consider that if there is a formal proportion that holds between the first p components of these vectors, then this proportion should hold for the last remaining components as well. This inference principle can be stated as below:

$$\frac{\forall i \in [1, p],\;\; T(a_i, b_i, c_i, d_i)}{\forall j \in [p+1, n],\;\; T(a_j, b_j, c_j, d_j)}$$

This is a generalized form of analogical reasoning, where we transfer knowledge from some components of our vectors to their remaining components. Another way to put it is to define a formal proportion T^k over $\mathbb{B}^k$ with

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

It is immediate that this extension T^k (whatever the value of k) has the properties of T and for the sake of simplicity, we denote T^k as T. Our inference principle can then be restated as:

$$\frac{\forall i \in [1,p], \;\; T(a_i, b_i, c_i, d_i)}{T(a, b, c, d)}$$

This simple rule should allow us to devise a transductive process that we examine below, letting $p = n$ and $k = n+1$. It is worth pointing out that properties such as *full identity* or *mirroring* are especially relevant in that perspective. Indeed, it is expected that in the case where d is such that it exists $a \in S$ with $\forall i \in [1,p], d_i = a_i$, then $T(a_i, a_i, a_i, d_i)$ holds. Thus, the approach includes the extreme particular case where we have to classify an item whose representation (in the input space) is completely similar to an already classified item. The mirroring property that expresses independence with respect to the encoding seems also very desirable since it ensures that whatever the convention used for the positive or the negative encodings of the value of each feature and of the class, one shall obtain the same classification result. However, mirroring is only a sufficient condition for ensuring this, as we shall see later in the paper. For the time being, we thus require both full identity and mirroring for the logical proportions that we are going to use. Thanks to Proposition 5, this means that we are only using the set of the three proportions A, R, and P, for the moment. Observe that these proportions are also symmetrical in agreement with the idea that the comparisons of a and b on the one hand and of c and d on the other hand play symmetrical roles.

Logical proportions-based transduction

In our approach, we apply a strategy very different from k-nn methods, since the new item d to be classified is not just compared with classified items on a one-to-one basis. For the sake of uniformity with previous sections, we continue to denote d the new item to be classified: d is known, but $cl(d)$ is unknown. Instead of looking for a c minimizing the disagreement set $\rangle c, d\langle$, we consider any 3-tuples $(a, b, c) \in S^3$ such that $\rangle a, b\langle = \rangle c, d\langle$. Thus we focus on the common features where the pairs (a, b) and (c, d) differ, i.e. in order to see if these changes of feature values (between a and b, and between c and d) are associated or not with classification changes.

As already said, when $\exists s \in S, |\rangle s, x\langle| = 0$, $T(s, s, s, x)$ holds for any proportion satisfying full identity, and the unique solution for having $T(cl(s), cl(s), cl(s), cl(x))$ to hold is $cl(s) = cl(x)$. Let us now examine first the case where $|\rangle a, b\langle| = 1$, i.e., they differ on only 1 feature. To better understand the strategy, let us first consider the following example where $n = 6$ and the class is in the last column:

	1	2	3	4	5	6	*cl*
a	1	1	0	1	0	0	1
b	1	1	0	1	1	0	0
c	1	0	0	1	1	1	0
d	1	0	0	1	0	1	?

We observe a variation for component 5 since we have $a_5 = 0, b_5 = 1, c_5 = 1, d_5 = 0$. The following questions have to be answered in order to predict the class for d:

- Do we have a proportion T (among A, P, R) such that $T(a, b, c, d)$ holds? In that case, only reverse analogy R works.
- Do we have a solution for the class equation $R(cl(a), cl(b), cl(c), x)$? We have 1 as a unique solution for reverse analogy, which leads us to conclude $cl(d) = 1$.

We have simply applied the previous inference principle: roughly speaking, when a proportion holds for the features, it should also hold for the class. Note that this mechanism agrees here with the fact that the comparison of $(a, cl(a))$ with $(b, cl(b))$ leads us to conclude that the only known reason for having $cl(a) \neq cl(b)$, is the difference between a and b on their 5th component. Since a similar difference is observed on c and d, this also plausibly leads to the conclusion that we should have $cl(c) \neq cl(d)$. If we have had $cl(a) = cl(b)$, we would have concluded $cl(c) = cl(d)$ by the proportion mechanism, in agreement with the idea in such a case that the change in feature value (here the 5th one) does not affect the class. Moreover, it can be checked that we cannot have situations such that $\rangle a, b\langle = \rangle a', b'\langle = \rangle a, a'\langle = \{j\}$, where $cl(a) = cl(b)$ while $cl(a') \neq cl(b')$ (otherwise there is no possible function underlying S). It means that in a common context where all the features values are equal on the 4 tuples, except for one feature, it cannot exist examples suggesting that there is no class change and other ones that there is.

The above example is indeed representative of a rather general situation, which amounts to say that a kind of "ceteris paribus" principle is at work here: If in a given context it is observed that a change in a feature value leads to a classification change, and that there does not exist another context where the same change has no impact on the classification, then one is led to assume that it would be true in any context (which agrees with the idea that the function underlying the examples should not be more complicate than required by the examples). This principle would still apply for more complex changes involving several features that would be repeatedly associated with the same classification change (or absence of change).

Still it is possible even in the case where $\rangle a, b\langle = \rangle c, d\langle = \{j\}$ that a proportion T holds for a, b, c, d, while $T(cl(a), cl(b), cl(c), cl(x))$ has no solution. Let us examine the different possible situations.

Indeed, in such case for $\forall i \neq j$, $(a_i, b_i, c_i, d_i) \in \{(1,1,1,1), (0,0,0,0), (1,1,0,0), (0,0,1,1)\}$. If only the two first 4-tuples appear among the n features, $T \in \{A, R, P\}$, while if one of the last two 4-tuples is obtained for some $i \neq j$, then $T \in \{A, R\}$ and P is excluded. Thanks to Property 8, we are sure to be able to complete with a unique class $cl(d)$ a given 4-tuple (a, b, c, d), for which $\exists! j$ such that $\rangle a, b\langle = \{j\}$, where $cl(a)$, $cl(b)$, $cl(c)$ are known, *only if* the proportion(s) that hold(s) for $i = 1, n$ (including $i = j$) has a solution for the corresponding class equation. This is illustrated by the following example.

There is one suitable proportion T such that $T(a, b, c, d)$

holds in the following example:

	1	2	3	4	5	cl
a	1	1	0	1	1	0
b	1	1	0	1	0	1
c	1	0	0	1	1	0
d	1	0	0	1	0	?

It is A, and this proportion makes the class equation with (a, b, c) solvable, and leads to $cl(d) = 1$. But if only one proportion, (say, e.g. A) is compatible with (a, b, c, d), it may be the case that $T(cl(s), cl(s), cl(s), cl(x))$ has solution only for $T = P$, or $T = R$, as in the example below.

	1	2	3	4	5	cl
a	1	1	0	1	1	1
b	1	1	0	1	0	0
c	1	0	0	1	1	0
d	1	0	0	1	0	?

It can be checked that this happens only if the same value of j is associated with two different classes in the contexts defined by $\{a_i = b_i \text{ for } i \neq j\}$ and $\{c_i = d_i \text{ for } i \neq j\}$. For instance, in the above example, for $j = 5$, $a_5 = 1 = c_5$, the item a is in class 1 in the context $a_1 = 1, a_2 = 1, a_3 = 0, a_4 = 1$, while c is in class 0 in the context $c_1 = 1, c_2 = 0, c_3 = 0, c_4 = 1$. In such a case, this means that the context, and not only j, influence the classification. This is why it seems preferable not to use such triples (a, b, c) for predicting the value of d (at least if we have better triples available; see however the section "More options" in the following for a possible way of inferring a class even in this case).

Towards an algorithm

The above analysis highlights an important point for designing a tranduction algorithm. Instead of considering any triple (a, b, c) in order to classify d, for each proportion T (among A, R, and P) we consider the set S_T of triples (a, b, c) such that the equation $T((a, cl(a)), (b, cl(b)), (c, cl(c)), (x, cl(x)))$ is solvable:

$$S_T = \{(a,b,c) \mid T((a,cl(a)),(b,cl(b)),(c,cl(c)),(x,cl(x))) \text{solvable}\}$$

Obviously the other triples are useless for our objective because, whatever the coming d, they cannot constitute a logical proportion with d. This preparation of the suitable set can be done offline. Generally, one finds more than one suitable triple (a, b, c) such that $T(a, b, c, d)$ then we have to choose among diverse solutions for the classes. Here, we have to decide what is the most "suitable" solution among all the potential options. Let us take an example where a, b, c, b', $c' \in S$:

	1	2	3	cl		1	2	3	cl
a	1	1	1	0	a	1	1	1	0
b	1	0	1	0	b'	1	1	0	1
c	0	1	1	1	c'	0	0	0	1
d	0	0	1	?	d	0	0	1	?

Clearly, on the left hand side, $(a, b, c) \in S_A$ and by analogical proportion we get $cl(d) = 1$, while on the right hand side, $(a', b', c') \in S_R$ and we get $cl(d) = 0$ by reverse analogy. However, observe that $\rangle a, b\langle = \rangle c, d\langle = \{2\}$, while $\rangle a, b'\langle = \{3\}$ and $\rangle c', d\langle = \{3\}$. Indeed, such a situation may take place only when the differences concern different features (a situation where the extrapolation is obviously bolder). In the above example, the first prediction ($cl(d) = 1$) may be preferred, since a, b, c, d are similar on at least one component (here the 3rd one) where we have 'full identity', just making the 4 situations closer (while there are pairwise identities only for a, b', and for c', d).

This leads us to another point: each set S_T can be broken down into a sequence of distinct subsets $S_T = \bigcup_{\sigma \in [1,n]} S_T^\sigma$ where $S_T^\sigma = \{(a, b, c) \in S_T \mid |\rangle a, b\langle| = \sigma\}$ (i.e. a and b differ on σ bits). When there is a choice between $(a, b, c) \in S_T^\sigma$ and $(a', b', c') \in S_{T'}^\tau$ to classify d, preference will be given to the class associated with the smallest of the indices σ and τ. We can view this principle as common sense since we choose to consider the most similar items to classify the new one.

Algorithm

Let us now translate all the previous remarks into an algorithm. In the following, we consider that we have at our disposal a procedure for solving a logical equation and more precisely: for each considered proportion T, a method $T.solveEq(a, b, c)$ gives the unique solution of the equation $T(a, b, c, x) = 1$ when this solution exists (and returns $false$ otherwise).

Below, we provide the algorithmic scheme defining a transductive engine based on logical proportions A, R, P. As general input, we have the set S of labeled examples and dynamically, we have to classify new data not belonging to S. It appears that a pre-process work of S can be done offline, before any observation of new data to be classified.

Offline work:

1) For $T \in \{A, P, R\}$, build up the set S_T :

$$S_T = \{(a,b,c) \mid T((a,cl(a)),(b,cl(b)),(c,cl(c)),(x,cl(x))) \text{solvable}\}$$

Note that these 3 sets are not necessarily disjoint. Obviously, here we keep a record of the solution of the class equation $T(cl(a), cl(b), cl(c), z)$ to avoid to call again the function $T.solveEq()$ during the online part of the algorithm.

2) For $\sigma = 1$ to n, for $T \in \{A, P, R\}$, build up the following sets:

$$S_T^\sigma = \{(a, b, c) \in S_T \mid |\rangle a, b\langle| = \sigma\}$$

For a given T, these sets are necessarily disjoint.

Online work:

1) input: $S \subset U \times \{0, 1\}, d \in U$ such that $(d, cl(d)) \notin S$

2) $\sigma \leftarrow 1$; //we start our loop. σ is the number of "differing" bits

$Cl1 \leftarrow 0; Cl0 \leftarrow 0;$

3) for each T, build up the set $S_{T,d}^\sigma$ equals to

$$\{(a, b, c) \in S_T^\sigma \mid T(a, b, c, d) \text{ and } \rangle a, b\langle = \rangle c, d\langle\}$$

4) if $\sum_{T\in\{A,P,R\}} \mid S^{\sigma}_{T,d} \mid = 0$

then $\{\sigma \leftarrow \sigma + 1; goto\ 3\}$ //we have no solution at this level

else {for $T \in \{A, P, R\}, for (a, b, c) \in S^{\sigma}_{T,d}$

if $T.solveEq(cl(a), cl(b), cl(c)) = 1$ then Cl1++ else Cl0++;

if $Cl1 > Cl0$ then $cl(d) = 1$ else

if $Cl1 < Cl0$ then $cl(d) = 0$ else

$cl(d) = undefined$

return $cl(d)$ }

Let us examine each step of this algorithm.

Step 1 obviously does not consider a d such that $(d, cl(d))$ belongs to S: in that case, it simply means the new item d to be classified has been already classified and there is no need to go further.

Step 2 initializes 3 counters:

- σ that defines in some sense the level of difference between d and a previously labeled item c. If we are not successful in classifying d at this level, the algorithm will go to the next level.
- $Cl1$ (resp. $Cl0$) which is the number of proportions classifying d as 1 (resp. 0).

Step 3 looks for all the context c differing from d on exactly the same σ bits and corresponding to a solvable proportion. At this step, the preprocessing of the initial set is useful to minimize the number of cases to be checked. At the end of this step, we have built up 3 sets $S^{\sigma}_{A,d}, S^{\sigma}_{P,d}, S^{\sigma}_{R,d}$. Roughly speaking, $S^{\sigma}_{T,d}$ is the set of triples (a, b, c) such that $T(a, b, c, d)$ holds and where the pairs $(a.b)$ and (c, d) differ in the same manner.

Step 4 is the final one: if the 3 previous sets are empty, it means that there is no solvable proportions involving d as last element at difference level σ: then we try to find out a solution at level $\sigma+1$ and are back to step 3. On the opposite, if at least one of these sets is not empty, then we can solve the class equation, get a solution 1 or 0. The 2 counters $Cl1$ for class 1 (resp. $Cl0$ for class 0) have to be updated. Then we implement a basic majority vote: for instance, if $Cl1 > Cl0$, it means that among the holding proportions, we have a majority of proportions giving 1 as solution for the $cl(d)$, then we decide to classify d as 1.

It appears here that, while most learners are able to generalize to any new input, this is not the case for our transductive learner. From time to time, the current form of our algorithm cannot decide between 1 and 0. Obviously, some improvements can be done, depending on the extra knowledge we have of the problem at hand. One of the simplest generalization we could consider is to have an order over the set of binary features indicating their relative importance for instance, as discussed in the "More options" section below.

Complexity

Le us provide a complexity estimation of the algorithm. From an algorithmic viewpoint, the equation solving process is quite straightforward and in fact linear in the size of the input, i.e. n. Then it appears that we have to explore the set S^3 which could reveal untractable when the size of S is huge. Nevertheless, we do not need to explore the whole set $S \times S \times S$ since we are only interested in triples (a, b, c) such that there is a T where the equation $T(a, c(a)), (b, c(b)), (c, c(c)), (x, cl(x)))$ is solvable. Obviously, this computation can be done offline, and then the size of the space to explore online can be reduced. This is the aim of the offline section of the algorithm. It is quite clear that we can take advantage of the properties of the considered logical proportions in terms of substitutions to subsequently reduce the search space: for instance if $A(a, b, c, d)$ holds (resp. does not hold) then 8 permutations of (a, b, c, d) will hold (resp. not hold) decreasing by a factor 8 the size of the space to be investigated.

Step 1 of the offline part is of the order of $3 \times |S|^3$

Step 2 of the offline work is of the order of $3 \times n \times |S|^3$ where n is the dimension of the input space U.

It then appears that the offline part is $(\mathcal{O}(n|S|^3))$, but obviously this is not satisfactory since the power 3 can quickly lead to huge numbers. At least, we remain polynomial at this stage.

Step 3 of the online part is of the order of $3 \times |S|^3$

Step 4 of the online part is in the range of $3 \times |S|^3$ multiplied by the time of the Boolean test $T.solveEq(cl(a), cl(b), cl(c)) = 1$ since, as previously explained,the solving process itself has been carried out during the offline part of the algorithm. This Boolean test is time constant since this is just the observation of the last element in the line starting with (a, b, c). Finally Step 4 is linear in $|S|^3$ as well.

It appears finally that the online part of our algorithm is $\mathcal{O}(|S|^3)$. There is obviously a need for optimisation. This can be done using the symmetry and permutation properties of our 3 proportions. We leave this issue for future works, but it is known from (Stroppa and Yvon 2005) that, in the case of analogical proportion, there are rooms for improvement.

Some previous similar experiences have shown the efficiency of an approach of this kind (see (Miclet, Bayoudh, and Delhay 2008) for instance). However the proposed approach differs from this work in several respects:

- We never consider approximate proportions, we consider only exact proportions. If there is no exact proportions, then we are not able to predict. Obviously, if we want to be able to predict whatever the context, we have to define an approximate prediction process. This could be done using techniques in (Miclet, Bayoudh, and Delhay 2008) for instance, or using the ideas outlined in the next section.
- Moreover, instead of considering a unique proportion (namely analogy), we allow 3 proportions to hold. Obviously, this will enlarge the range of situations, where we can decide and in some sense, somewhat balance the drawback coming from the fact that we consider only exact proportions (rather than approximate ones as in (Miclet, Bayoudh, and Delhay 2008)).

More options

The above proposed procedure makes use only of proportions A, R, or P. It may happen that the procedure is not able to recommend a particular class for a new item. In order to at least partially overcome these limitations, some further options may be considered.

- An ordering on the features may be available, either from experts, or from some preliminary analysis of the data (by means of some principal component analysis for instance). This ordering is supposed to rank-order the features according to their decreasing capability to induce a classification change when their value changes, in a large number of distinct contexts. Which such an ordering, one may expect to identify features that have no influence or almost no influence, and to neglect them in the procedure. This may help diminishing the number of situations where one cannot find a proportion that holds for the feature description part and that has a solution for extrapolating the unknown class, since the number of useful features to be considered will be reduced.
- Another interesting option when it does not exist any triple a, b, c such that componentwise a formal proportion of Table 1 holds, is to look for a proportion of Table 2 that tolerates some "variability' as in the following example (see feature 2, or 4 , vs. feature 3):

a	1	0	0	0	1
b	1	1	1	0	0
c	1	0	0	0	1
d	1	0	1	1	?

  Here the application of the proportion $b \mid a :: d \mid c$ will yield $cl(d) = 0$, while no proportion of Table 1 is applicable. Another example is provided by the simple self explanatory example below,

a	*lay eggs* (1)	*small* (0)	*fly* (1)	*bird* (1)
b	*lay eggs* (1)	*big* (1)	*notfly* (0)	*bird* (1)
c	*lay eggs* (1)	*small* (0)	*fly* (1)	*bird* (1)
d	*lay eggs* (1)	*big* (1)	*fly* (1)	?

  where the third proportion in Table 2 would allow us to conclude that d flies.
- Another situation of interest where the proposed procedure cannot conclude, is when, given a d to be classified, one cannot find a suitable triple (a, b, c) in S where a proportion T is applicable for all features, but where there exists a triple a, b, c such that, componentwise some formal proportion holds (not necessarily the same one for each component). For instance,

	1	2	3	4	cl
a	1	0	0	0	1
b	1	1	1	0	0
c	1	1	0	1	1
d	1	0	1	1	?
T	A, R, P	R, P	A, P	A, R	

  In that case, we see that there is no proportion T such that $T(a, b, c, d)$. Nevertheless, there still exists a unique solution $cl(d) = cl(c) \equiv (cl(a) \equiv cl(b))$ thanks to Property 7 (since any triple $cl(a), cl(b), cl(c)$ can be completed by two of the proportions of Table 1, due to Property 8). In the above example, the solution is indeed unique and equal to 0). Such an option is to be used only when no other option applies.

Related works

When it comes to infer new information from the comparison of cases, we are obviously led back to analogical reasoning, which has received much attention from a cognitive science point of view. This kind of reasoning has been mainly formalized in the setting of first order logic (Davies and Russell 1987; Melis and Veloso 1998), also taking inspiration from the idea of functional dependency. It can be stated as:

- two particular terms s and t share a common property P, i.e. $P(s)$ and $P(t)$ both hold,
- s satisfies an additional property Q, i.e. $Q(s)$ holds,
- then t should satisfy Q, i.e. $Q(t)$ holds.

See (Prade and Richard 2010) for a discussion of the relation with the analogical proportion approach. These approaches take only 2 terms into consideration: the source s and the target t. Analogical inference based on analogical proportion allows to consider the source and the target as pairs of terms (a, b) and (c, d). Instead of handling 2 items, we rather deal with 4 items and we have to break down the initial reasoning to take into account the relation between these items as an essential ingredient for an inferential step. Obviously, this makes also a difference with the case-based reasoning (Aamodt and Plaza 1994) where the new case at hand has to be matched to only one previously seen case. In (Miclet, Bayoudh, and Delhay 2008), a first use of Boolean analogical proportion has been designed and very successfully applied to a classification task. Nevertheless, they use only one type of proportion: our way to proceed generally provides more options, involving more 3-tuples (a, b, c) and allows to make a full use of the available information. Besides, after the pioneering preliminary investigation made in (Lepage 2001) for analogical proportions only, and its logical counterpart provided in (Miclet and Prade 2009), the more recent work by (Prade and Richard 2009) has introduced two new proportions (paralogy and reverse analogy) which capture intuitions different from the one underlying the standard analogical proportion, while (Prade and Richard 2010) has provided a detailed study of the core properties underlying these formal proportions. But, this work missed the other natural options, laid bare in our paper, coming from conditional objects, and did not discuss transduction. In view of transduction, this gives us more potentialities for exploiting every piece of information.

(Katona, Keszler, and Sali 2010) have recently proposed a distance between two databases, which remains zero between an initial database and this database completed with a new item, provided that this item leaves the existing functional dependencies unchanged. This seems relevant for our transduction problem, since the knowledge of the functional dependencies that determine the class from a subset of features, and which can be induced from a set of observed data, may help classifying a new item.

Conclusion

In this paper, we have investigated all the logical proportions involving 4 Boolean items that acknowledge a so-called full identity requirement. We have provided a unified approach for identifying these proportions. Relying only on the observation of existing similarities/disimilarities between pairs of items, these proportions capture common sense intuition, going from standard analogical proportion to conditional objects. It has been shown that among all the available proportions, only 11 seem to be particularly meaningful: they equate similarities on the one hand, and dissimilarities on another hand, between pairs. Moreover, we have established the uniqueness of analogy, reverse analogy and paralogy as the proportions satisfying an ultimate requirement reminiscent of the very meaning of the word ”proportion”: the code-independence (mirroring). Starting from a given proportion, it has been shown that a simple inference mechanism allows to deduce new information, in a way reminiscent of the well known ”rule of three”. This mechanism appears to be the basis of a transductive reasoning process, allowing us to complete information when only a limited sample of data is available. Since we make only use of exact proportions, our algorithm cannot always classified a new item. We have given diverse ways to overcome this issue. Another solution would be to accept imperfect proportions and to define a kind of distance to perfection, as it has been done in other approaches. Obviously, it remains to implement and experiment such a proportion-based transductive machinery, to experiment it on benchmarks and to compare the results with existing classification methods. Previous works by (Miclet, Bayoudh, and Delhay 2008) only based on pure analogical proportions, allow us to be rather optimistic.

References

Aamodt, A., and Plaza, E. 1994. Case-based reasoning; foundational issues, methodological variations, and system approaches. *AICom, 7 (1)* 39–59.

Benferhat, S.; Dubois, D.; Kaci, S.; and Prade, H. 2008. Modeling positive and negative information in possibility theory. *Int. J. of Intelligent Systems* 23:1094–1118.

Benferhat, S.; Dubois, D.; and Prade, H. 1997. Nonmonotonic reasoning, conditional objects and possibility theory. *Artificial Intelligence* 92:259–276.

Davies, T. R., and Russell, S. J. 1987. A logical approach to reasoning by analogy. In *IJCAI-87*, 264–270. Morgan Kaufmann.

De Finetti, B. 1936. La logique des probabilités. In *Congress International de Philosophie Scientifique*, 1–9. Paris, France: Hermann et Cie.

De Finetti, B. 1974. *Theory of Probability — A Critical Introductory Treatment*, volume 1 & 2. Chichester, United Kingdom: John Wiley & Sons.

Dubois, D., and Prade, H. 1994. Conditional objects as nonmonotonic consequence relationships. *IEEE Transactions on Systems, Man and Cybernetics* 24(12):1724–1739.

Gammerman, A.; Vovk, V.; and Vapnik, V. 1998. Learning by transduction. In *In Uncertainty in Artificial Intelligence*, 148–155. Morgan Kaufmann.

Hempel, C. 1965. Studies in the logic of confirmation. *In Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* 3–46.

Katona, G. O. H.; Keszler, A.; and Sali, A. 2010. On the distance of databases. In *Proc. 6th Inter. Symp. on Foundations of Information and Knowledge Systems (FoiKS'10), Sofia*, volume LNCS 5956, 76–93. Springer.

Klein, S. 1982. Culture, mysticism & social structure and the calculation of behavior. In *Proc. Europ. Conf. in AI (ECAI)*, 141–146.

Kraus, K.; Lehmann, D.; and Magidor, M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44:167–207.

Lepage, Y. 2001. Analogy and formal languages. In *Proc. FG/MOL 2001*, 373–378. (see also http://www.slt.atr.co.jp/ lepage/pdf/dhdryl.pdf.gz).

Melis, E., and Veloso, M. 1998. Analogy in problem solving. In *Handbook of Practical Reasoning: Comput. and Theor. Aspects*. OUP.

Miclet, L., and Prade, H. 2009. Handling analogical proportions in classical logic and fuzzy logics settings. In *Proc. ECSQARU*, 638–650. Springer, LNCS 5590.

Miclet, L.; Bayoudh, S.; and Delhay, A. 2008. Analogical dissimilarity: definition, algorithms and two experiments in machine learning. *JAIR, 32* 793–824, http://hal.inria.fr/inria–00070321/en/.

Prade, H., and Richard, G. 2009. Analogy, paralogy and reverse analogy: Postulates and inferences. In *Proc. 32nd Ann. Conf. on Artif. Intellig. (KI 2009), Paderborn*, volume LNAI 5803, 306–314. Springer.

Prade, H., and Richard, G. 2010. Analogical proportions: another logical view. In Bramer, M.; Ellis, R.; and Petridis, M., eds., *Research and Development in Intelligent Systems XXVI, Proc. 29th Ann. Inter. Conf. on Artif. Intellig. (SGAI'09), Cambridge, UK, 15-17 December 2009*, 121–134. Springer.

Russell, B. 1912. *The Problems of Philosophy. Chap. VI. On induction*. Home Univ. Libr.; Oxford Univ. Pr., 1959; http://www.ditext.B. com/russell/russell.html.

Stroppa, N., and Yvon, F. 2005. An analogical learner for morphological analysis. In *Proc. 9th Conf. Comput. Natural Language Learning (CoNLL-2005)*, 120–127.