HAL
archives-ouvertes.fr

When nominal analogical proportions do not fail

Miguel Couceiro, Erkko Lehtonen, Laurent Miclet, Henri Prade, Gilles Richard

▶ To cite this version:

HAL Id: hal-02864071

https://hal.archives-ouvertes.fr/hal-02864071v2

Submitted on 30 Jul 2020

When nominal analogical proportions
do not fail⋆

Miguel Couceiro[1], Erkko Lehtonen[2], Laurent Miclet[3], Henri Prade[4], and
Gilles Richard[4]

[1] Université de Lorraine, CNRS, Inria N.G.E., Loria, France,
`miguel.couceiro@loria.fr`
[2] Centro de Matemática e Aplicações, Faculdade de Ciências e Tecnologia,
Universidade Nova de Lisboa, Portugal, `e.lehtonen@fct.unl.pt`
[3] IRISA, University of Rennes, France, `laurent.miclet@gmail.com`
[4] IRIT, CNRS & Université Paul Sabatier, France, {`prade,richard`}`@irit.fr`

Abstract. Analogical proportions are statements of the form "**a** is to **b** as **c** is to **d**", where **a**, **b**, **c**, **d** are tuples of attribute values describing items. The mechanism of analogical inference, empirically proved to be efficient in classification and reasoning tasks, started to be better understood when the characterization of the class of classification functions with which the analogical inference always agrees was established for Boolean attributes. The purpose of this paper is to study the case of finite attribute domains that are not necessarily two-valued, i.e., when attributes are nominal. In particular, we describe the more stringent class of "hard" analogy preserving (HAP) functions $f: X_1 \times \cdots \times X_m \to X$ over finite domains $X_1, \ldots, X_m, X$ for binary classification purposes. This description is obtained in two steps. First we observe that such AP functions are *almost affine,* that is, their restriction to any $S_1 \times \cdots \times S_m$, where $S_i \subseteq X_i$ and $|S_i| \leq 2$ $(1 \leq i \leq m)$, can be turned into an affine function by renaming variable and function values. We then use this result together with some universal algebraic tools to show that they are essentially unary or *quasi-linear,* which provides a general representation of HAP functions. As a by-product, in the case when $X_1 = \cdots = X_m = X$, it follows that this class of HAP functions constitutes a clone on X, thus generalizing several results by some of the authors in the Boolean case.

1 Introduction

An analogy establishes a parallel between two situations, which are similar in many respects and dissimilar in others. If such a parallel holds to some extent, there are pairs (a, b) and (c, d) such as "a is to b in situation 1 as c is to d in

---

situation 2" [7]. Analogical proportions are of the form "a is to b as c is to d". It is only recently that researchers have provided representations of this quaternary relation in different settings [10, 12], or algorithms for finding an item d, given a, b, and c, for building an analogical proportion from these three items when it is possible [9]. The items considered in this paper are supposed to be represented by vectors of attribute values.

Analogical inference relies on the idea that if four items a, b, c, d are in analogical proportion for each of the n attributes describing them, it may still be the case for another attribute. For instance, if class labels are known for a, b, c and unknown for d, then one may infer the label for d as a solution of an analogical proportion equation [16]. Obviously, analogical inference rule is not a sound rule, and the effectiveness of analogical classifiers [2, 11] looks quite mysterious. From a theoretical viewpoint it is quite challenging to find and characterize situations where such an inference rule can be applied in a sound way. In case of Boolean attributes, a first step for explaining this state of facts was to characterize the set of functions for which analogical inference is sound, i.e., no error occurs, no matter which triplets of examples are used. In [4], it was proved that these so-called "analogy-preserving" (AP) functions coincide exactly with the set of affine Boolean functions. Moreover, when the function is close to being affine, it was also shown that the prediction accuracy remains high [5]. When attributes are valued on finite domains, which we refer to as the "*nominal case*" (it includes the Boolean case), the problem of identifying the AP functions had remained a challenging open problem until now. This paper aims to solving this problem in the context of binary classification problems and to providing a complete description of the more stringent class of "hard" AP functions.

The paper is organized as follows. Section 2 provides the necessary background on analogical proportions and analogical inference in the Boolean and in the nominal cases. Then we introduce the notion of analogy-preserving functions on which analogical inference never fails, and discusses an illustrative example in the nominal case, which emphasizes the linkage of analogical proportions with trees cataloguing items according to the values of the attributes used for describing them. Later, a local description of hard analogy preserving functions is given in terms of almost affine functions, which is then extended to a global description given in terms of the notion of quasi-linear functions.

2 Background

Postulates An analogical proportion is a 4-ary relation, denoted $a : b :: c : d$, between items a, b, c, d, supposed to obey the following 3 postulates (e.g., [10]):
  - $\forall a, b, \ a : b :: a : b$ (*reflexivity*)
  - $\forall a, b, c, d, \ a : b :: c : d \rightarrow c : d :: a : b$ (*symmetry*)
  - $\forall a, b, c, d, \ a : b :: c : d \rightarrow a : c :: b : d$ (*central permutation*)
The repeated and alternate application of the two last postulates entail that an analogical proportion $a : b :: c : d$ has 8 equivalent forms: $a : b :: c : d = c : d :: a : b = c : a :: d : b = d : b :: c : a = d : c :: b : a = b : a :: d : c = b : d :: a : c = a : c :: b : d$. Some immediate consequences can be observed:

i) $\forall a, b, \; a : a :: b : b \; (identity)$

ii) $\forall a, b, c, d, \; a : b :: c : d \rightarrow d : b :: c : a \; (extreme \; permutation)$

iii) $\forall a, b, c, d, \; a : b :: c : d \rightarrow b : a :: d : c \; (inside \; pair \; reversing)$

iv) $\forall a, b, c, d, \; a : b :: c : d \rightarrow d : c :: b : a \; (complete \; reversal)$

Boolean case Let us now assume for a while that a, b, c, d denote Boolean variables, i.e., their values belong to the set $\{0, 1\}$. This may be thought of as encoding the fact that a given property is true or false for the considered item. Since items are usually described in terms of several properties, this modeling of analogical proportions is then extended to tuples in a component-wise manner as recalled later. As shown in [13], the minimal Boolean model obeying the analogical proportion postulates makes $a : b :: c : d$ true only for the six patterns

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

while $a : b :: c : d$ is false for the other ten patterns of values for the four variables a, b, c, d. This is the truth table of a quaternary connective that can be logically expressed as $a : b :: c : d = ((a \wedge \neg b) \equiv (c \wedge \neg d)) \wedge ((\neg a \wedge b) \equiv (\neg c \wedge d))$ [12].

It can be seen on this expression that the analogical proportion "a is to b as c is to d" formally states that "a differs from b as c differs from d and b differs from a as d differs from c". It means that $a = b \Leftrightarrow c = d$, and that $a \neq b \Leftrightarrow c \neq d$ (with the further requirement that both truth value changes are in the same direction (either from 1 to 0, or from 0 to 1, when going from a to b, and from c to d). So, the analogy is as much a matter of dissimilarity as a matter of similarity.

Nominal case In the nominal case, attributes are supposed to take their values on finite domains (which are not necessarily ordered). For instance, the attribute domain of `color` may be the set $\{blue, red, yellow\}$. Let s and t be two values in such a finite domain X. It follows from reflexivity and central permutation postulates that $s : t :: s : t$ and $s : s :: t : t$ should hold. By the symmetry postulate, s and t play the same role. Note also that s and t are not necessarily distinct. This leads to a minimal model of analogical proportion for nominal values, which can be stated as follows: $a : b :: c : d$ holds if and only if

$$(a, b, c, d) \in \{(s, t, s, t), (s, s, t, t) \mid s, t \in X\}.$$

This clearly covers the Boolean case as a particular case, leading to the 6 lines seen in the Boolean truth table. If $|X| = n$, we obviously have n^4 tuples (a, b, c, d). Among them, we have i) n valid analogies of type $s : s :: s : s$, ii) $n(n-1)$ of type $s : t :: s : t$ with $s \neq t$, and iii) $n(n-1)$ of type $s : s :: t : t$ with $s \neq t$.

Hence, a total of $n(2n-1)$, which shows that the number of valid analogies increases with the square of the cardinality of the underlying set. For instance, for an attribute such as `color` whose values belong to, e.g., $X = \{blue, red, yellow\}$, we have only 15 valid analogies among 81 combinations. For instance, $red : yellow :: red : yellow$ holds, but it is not the case for $red : yellow :: red : blue$. Following the definition, an analogical proportion that holds with nominal attribute values involves *at most two distinct values*. This remark will have important consequences in the theoretical part of this paper. So, in the nominal case,

$a : b :: c : d$ is false if and only if $|\{a, b, c, d\}| \geq 3$ or $(a, b, c, d) \in Neg$, where

$$Neg = \{(s, t, t, s), (s, s, s, t), (s, s, t, s), (s, t, s, s), (t, s, s, s) \mid s, t \in X, s \neq t\}.$$

Representing objects with a single Boolean or nominal attribute is usually not expressive enough. In general, items are represented by *tuples* of values, each component being the value of an attribute, either Boolean or nominal. Extending the definition of analogy to nominal tuples of the form $\mathbf{a} = (a_1, \ldots, a_m)$ belonging to a Cartesian product $X = X_1 \times \cdots \times X_m$ can be done component-wise as follows:
$$\mathbf{a} : \mathbf{b} :: \mathbf{c} : \mathbf{d} \text{ holds} \Leftrightarrow \forall i \in [1, m], a_i : b_i :: c_i : d_i \text{ holds}.$$
Throughout the paper, the attribute domains $X_1, \ldots, X_m$ are assumed to be finite sets with at least two elements each.

Analogical inference In the Boolean case, the problem of finding an $x \in \{0, 1\}$ such that $a : b :: c : x$ holds, does not always have a solution. Indeed, neither $0 : 1 :: 1 : x$ nor $1 : 0 :: 0 : x$ has a solution (since 0111, 0110, 1000, 1001 are not valid patterns for an analogical proportion). In fact, a solution exists if and only if $(a \equiv b) \vee (a \equiv c)$ holds. When a solution exists, it is unique and is given by $x = c \equiv (a \equiv b)$. This corresponds to the original view advocated by S. Klein [8], who however applied the latter formula even to the cases $0 : 1 :: 1 : x$ and $1 : 0 :: 0 : x$, where it yields $x = 0$ and $x = 1$ respectively.

In the nominal case, the situation is similar. The analogical proportion $a : b :: c : x$ may have no solution ($s : t :: t : x$ has no solution as soon as $s \neq t$), and otherwise (if $a = b$ or $a = c$) the solution is unique, and is given by $x = b$ if $a = c$ and $x = c$ if $a = b$. Namely, the solutions of $s : t :: s : x$, $s : s :: t : x$, and $s : s :: s : x$ are $x = t$, $x = t$, and $x = s$, respectively.

This motivates the following inference pattern first proposed in [16]

$$\frac{\forall i \in \{1, \ldots, m\}, \ a_i : b_i :: c_i : d_i \text{ holds}}{a_{m+1} : b_{m+1} :: c_{m+1} : d_{m+1} \text{ holds}}$$

which enables us to compute d_{m+1}, provided that $a_{m+1} : b_{m+1} :: c_{m+1} : x$ has a solution. This pattern expresses a rather bold inference which amounts to saying that if the representations of four items are in analogical proportion on m attributes, they should remain in analogical proportion with respect to their labels. Note that, we can restrict ourselves to binary labels, since a multiple class prediction can be obtained by solving a series of binary class problems.

In this paper, we adopt a completely different viewpoint: instead of adding constraints for ensuring the soundness of analogical inference, we want to characterize contexts where this inference is valid, without adding any further constraints. In the next section, we proceed with a discussion on analogical inference, and we present the notion of analogy-preserving functions.

3   Analogy-preserving functions

In the analogical inference pattern that was introduced in the previous section, we implicitly assume that there is a dependency that links labels to the values

of the m attributes. More precisely, there is some unknown function f such that $e_{m+1} = f(e_1, \ldots, e_m)$, for any item $\mathbf{e} = (e_1, \ldots, e_m)$. Such a function f can be thought of as a classifier that associates to each item, a (unique) class based on the values of the m attribute values describing it.

Since the solutions of analogical equations (when they exist) are unique, the previous pattern can be also written as follows:

$$\begin{array}{ccccc c} a_1 & \cdots & a_i & \cdots & a_m & f(\mathbf{a}) \\ b_1 & \cdots & b_i & \cdots & b_m & f(\mathbf{b}) \\ \hline c_1 & \cdots & c_i & \cdots & c_m & f(\mathbf{c}) \\ \hline d_1 & \cdots & d_i & \cdots & d_m & f(\mathbf{d}) \end{array}$$

where $\mathbf{a} = (a_1, \ldots, a_m)$, $\mathbf{b} = (b_1, \ldots, b_m)$, $\mathbf{c} = (c_1, \ldots, c_m)$ and $\mathbf{d} = (d_1, \ldots, d_m)$.

Remark 3.1. Note that in the nominal case, each column i has at most two distinct elements belonging to the attribute domain X_i of attribute i.

As previously highlighted, the conclusions obtained by analogical inference are brittle. This means here that for a given $\mathbf{d} = (d_1, \ldots, d_i, \ldots, d_m)$ for which we want to evaluate $f(\mathbf{d})$, there may exist several triplets $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ such that $f(\mathbf{a}) : f(\mathbf{b}) :: f(\mathbf{c}) : x$ is solvable, maybe leading to different solutions. In that case, at least from a theoretical viewpoint, it is clear that applying the analogical inference principle for a given $\mathbf{d}$ will not give a unique value to predict $f(\mathbf{d})$. To cope with real-life situations, one generally uses a majority vote for computing a plausible $f(\mathbf{d})$. But an interesting particular case is when all the analogical predictions are the same whatever the triplets. This will be the case as soon as the function f is analogy-preserving, a notion we now formally define.

3.1   Basic notions and motivation

In the following, $X_1, \ldots, X_m, X$ denote finite sets with cardinality at least 2.

Definition 3.2. *Let* $\mathbf{X} = X_1 \times \cdots \times X_m$. *A function* $f \colon \mathbf{X} \to X$ *is* analogy-preserving *(AP for short) if for every* $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbf{X}$,

$\mathbf{a} : \mathbf{b} :: \mathbf{c} : \mathbf{d}$ *and* $solvable(f(\mathbf{a}), f(\mathbf{b}), f(\mathbf{c})) \Rightarrow sol(f(\mathbf{a}), f(\mathbf{b}), f(\mathbf{c})) = f(\mathbf{d})$,

where $solvable(f(\mathbf{a}), f(\mathbf{b}), f(\mathbf{c}))$ *means that there exists an* x *such that* $f(\mathbf{a}) : f(\mathbf{b}) :: f(\mathbf{c}) : x$ *holds, and* $sol(f(\mathbf{a}), f(\mathbf{b}), f(\mathbf{c}))$ *is the unique solution for* x.

Note that if f is AP, there cannot exist $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{a}', \mathbf{b}', \mathbf{c}'$ such that

1. $\mathbf{a} : \mathbf{b} :: \mathbf{c} : \mathbf{d}$ and $\mathbf{a}' : \mathbf{b}' :: \mathbf{c}' : \mathbf{d}$,
2. $solvable(f(\mathbf{a}), f(\mathbf{b}), f(\mathbf{c}))$ and $solvable(f(\mathbf{a}'), f(\mathbf{b}'), f(\mathbf{c}'))$, and
3. $sol(f(\mathbf{a}), f(\mathbf{b}), f(\mathbf{c})) \neq sol(f(\mathbf{a}'), f(\mathbf{b}'), f(\mathbf{c}'))$.

In other words, AP functions are exactly those for which analogical inference never fails. Let us denote by **AP** the set of all AP functions. The following examples suggest that the class of AP functions is quite large and diverse.

Example 3.3. Consider the class of essentially unary[1] functions $f \colon \mathbf{X} \to X$, i.e., of the form $f(\mathbf{x}) = \varphi(x_i)$, for some map $\varphi \colon X_i \to X$. It is not difficult to see that such functions are AP. This corresponds to the simplest example of classifiers since the predicted classes are then determined by the value of a single attribute.

Example 3.4. Consider now the class of *injective* functions $f \colon \mathbf{X} \to X$, i.e., of functions that satisfy the condition: if $\mathbf{x} \neq \mathbf{y}$, then $f(\mathbf{x}) \neq f(\mathbf{y})$. Again, such functions are AP. The key observation is that if $\mathbf{a}, \mathbf{b}, \mathbf{c}$ are pairwise distinct, then so are $f(\mathbf{a}), f(\mathbf{b}), f(\mathbf{c})$ and the condition *solvable*$(f(\mathbf{a}), f(\mathbf{b}), f(\mathbf{c}))$ in the definition of AP functions does not hold. Thus injective functions are AP. However, injective functions are of little interest in classification since the number of class labels is expected to be smaller than the number of items.

Example 3.5. Using the same argument, we can relax the previous example to obtain other classes of AP functions. For instance, let $\mathbf{1} = (1, \ldots, 1) \in \{0, 1\}^m$ and consider the class of *pseudo-Boolean functions* $f \colon \{0, 1\}^m \to X$ satisfying the following condition: the kernel[2] of f is the form $\ker f = \{(\mathbf{a}, \mathbf{a} \oplus \mathbf{1}) \mid \mathbf{a} \in \{0, 1\}^m\}$, where $\oplus$ denotes addition modulo 2. Again, it is not difficult to verify that such functions are AP. These are examples of "reflexive" functions [6], i.e., functions satisfying the condition that for every $\mathbf{x} \in \{0, 1\}^m$, $f(\mathbf{x} \oplus \mathbf{1}) = f(\mathbf{x})$.

Example 3.6. Examples 3.4 and 3.5 can be generalized as follows. Recall that the set $\mathbb{B} := \{0, 1\}$ constitutes a 2-element field with the operations $\oplus$ (addition modulo 2) and $\otimes$ (multiplication modulo 2). For any natural number m, the set $\mathbb{B}^m$, equipped with scalar multiplication and addition of vectors, is a vector space over $\mathbb{B}$. Let V be a fixed subspace of the vector space $\mathbb{B}^m$. Any function $f \colon \mathbb{B}^m \to X$ such that $\ker f$ is the set of affine spaces that are translations of V is AP. Examples 3.4 and 3.5 correspond to the cases where V is the trivial subspace and the 1-dimensional subspace $\{\mathbf{0}, \mathbf{1}\}$, respectively.

In view of Remark 3.1 we will focus on the following subclass of AP functions.

Definition 3.7. *An AP function $f \colon \mathbf{X} \to X$ is called a* hard AP (HAP) *function if for all $S_i \subseteq X_i$ with $|S_i| \leq 2$ ($1 \leq i \leq m$) it holds that $|\mathrm{Im}\, f|_S| \leq 2$, where $S := S_1 \times \cdots \times S_m$. We denote the class of HAP functions by* **HAP**.

Remark 3.8. Observe that **HAP** contains all essentially unary functions. If X is a 2-element set, then **HAP** = **AP**.

3.2   ANF representations and affine functions

In this section we recall some well-known facts about the simplest interesting case of functions, namely, the Boolean functions.

---

[1] An argument x_i is said to be *inessential* in $f \colon \mathbf{X} \to X$ if for all $(a_1, \ldots, a_m) \in \mathbf{X}$, $a_i' \in X_i$, we have $f(a_1, \ldots, a_m) = f(a_1, \ldots, a_{i-1}, a_i', a_{i+1}, \ldots, a_m)$. Otherwise, x_i is said to be *essential* in f. The number of essential arguments of f is called the *essential arity* of f.

[2] Recall that the *kernel* of f is $\ker f := \{(\mathbf{a}, \mathbf{b}) \in \{0, 1\}^m \times \{0, 1\}^m \mid f(\mathbf{a}) = f(\mathbf{b})\}$.

There are several formalisms to represent Boolean functions, such as the classical DNF and CNF representations. However, in the analogical framework the algebraic representation of Boolean functions turns out be more relevant than the former classical representations based on the standard logical operators $\vee$ (disjunction) and $\wedge$ (conjunction). Following [15, 17], each Boolean function $f\colon \{0,1\}^m \to \{0,1\}$ (of *arity m*) is uniquely represented by a multilinear polynomial called the algebraic normal form of f that we recall below.

Let $\mathbb{B}$ be the 2-element field over $\{0,1\}$ with its 2 usual operators $\oplus$ (addition modulo 2) and $\otimes$ (multiplication modulo 2). Note that they correspond respectively to the *exclusive or* and to the *conjunction* in logical terms. Equipped with scalar multiplication (which coincides here with the multiplication modulo 2) over $\mathbb{B}$ and addition, the set $\mathbb{B}[x_1, \ldots, x_m]$ of polynomials on the m indeterminates $x_1, \ldots, x_m$ is a vector space over $\mathbb{B}$.

A (*multilinear*) *monomial* is a term of the form $\mathbf{x}_I := \prod_{i \in I} x_i$, for some (possibly empty) finite set of positive integers I with the convention that 1 is the empty monomial $\mathbf{x}_\emptyset$. The size $|I|$ is called the *degree* of $\mathbf{x}_I$, denoted $d(\mathbf{x}_I)$. A (*multilinear*) *polynomial* is a sum of monomials

$$\sum_{I \subseteq \{1, \ldots, m\}} \omega_I \cdot \mathbf{x}_I$$

where each ω_I belongs to $\mathbb{B}$ (addition is understood as addition modulo 2). Note that the monomials 0 and 1 are just $0 \cdot \mathbf{x}_\emptyset$ and $1 \cdot \mathbf{x}_\emptyset$, respectively. The *degree* of a polynomial is then the maximum degree among the degrees of its monomials.

An *algebraic normal form* (ANF) of a Boolean function f of arity m is simply a multilinear polynomial in $\mathbb{B}[x_1, \ldots, x_m]$ that represents it:

$$f(x_1, \ldots, x_m) = \sum_{I \subseteq \{1, \ldots, m\}} \omega_I \cdot \mathbf{x}_I.$$

It is well known that the ANF representation of a Boolean function is unique (see, e.g., [6]), and thus we can define the *degree* $d(f)$ of a Boolean function f as the degree of the polynomial that represents it. Note that the constant 0 and 1 functions are the only Boolean functions of degree 0, whereas *projections* (that correspond to the selection of a single attribute and that are represented by variables x_i) and their negations (that are represented by polynomials of the form $x_i \oplus 1$) are the only functions of degree 1.

A Boolean function $f\colon \mathbb{B}^m \to \mathbb{B}$ is said to be *affine* if $d(f) \leq 1$, i.e., there exist $\omega_0, \omega_1, \ldots, \omega_m \in \mathbb{B}$ such that

$$f(x_1, \ldots, x_m) = \sum_{i=1}^m \omega_i \cdot x_i + \omega_0.$$

The set of affine functions of arity m is denoted by $\mathcal{L}_m$, so that $\mathcal{L} = \bigcup_{m \geq 0} \mathcal{L}_m$ is the set of all affine functions. If $\omega_0 = 0$, then such an affine function is said to be *linear*. Thus, affine functions are either linear functions or their negations.

Our interest in this class of affine functions is motivated by the characterization of AP Boolean functions (i.e., in case when items are described by Boolean

attributes) [4]: AP Boolean functions are exactly those Boolean functions that are affine, i.e., $\mathbf{AP} = \mathcal{L}$.

In the following sections, we will generalize this result to the case of nominal attributes. This is not a straightforward extension of the Boolean case as we shall see. Before doing that, we provide an illustrative example that puts nominal analogical proportions in another perspective, and that reveals the close relationship of analogical proportions and taxonomic trees, as recently suggested in the Boolean case [1].

3.3   An illustrative example

In the illustrative example below, items are assumed to be described by means of three attributes (i.e., $m = 3$), numbered from 1 to 3, namely: $1 = \texttt{shape}$, $2 = \texttt{color}$, and $3 = \texttt{weight}$, where respectively $X_1 = \{circle$ (c), $square$ (s)$\}$, $X_2 = \{blue$ (b), red (r), $yellow$ (y)$\}$, and $X_3 = \{light$ (l), $heavy$ (h)$\}$. Due to space constraints, we chose a small example, with a non-Boolean nominal attribute, namely, X_2 with $|X_2| = 3$. There are two class labels referred to by 0 and 1.

Table 1 enumerates the 12 items, $\mathbf{a}, \mathbf{a}', \mathbf{b}, \ldots, \mathbf{f}'$, that can be distinguished on the basis of the three attributes above. Moreover, we consider 4 ways of classifying them into the two classes, each of which corresponding to the 4 functions g_1, g_2, g_3, g_4.

items	shape	color	weight	g_1	g_2	g_3	g_4
$\mathbf{a}$	c	b	l	1	1	1	1
$\mathbf{a}'$	c	b	h	1	1	0	0
$\mathbf{b}$	c	r	l	1	0	0	1
$\mathbf{b}'$	c	r	h	0	0	0	0
$\mathbf{c}$	c	y	l	0	1	0	1
$\mathbf{c}'$	c	y	h	0	1	0	0
$\mathbf{d}$	s	b	l	1	0	1	1
$\mathbf{d}'$	s	b	h	1	0	0	0
$\mathbf{e}$	s	r	l	1	1	0	1
$\mathbf{e}'$	s	r	h	0	1	0	0
$\mathbf{f}$	s	y	l	0	0	0	1
$\mathbf{f}'$	s	y	h	0	0	0	0

Table 1. Items, attributes, and classifications.

Clearly, g_1 is not an AP function since: i) $\mathbf{a} : \mathbf{a}' :: \mathbf{b} : \mathbf{b}'$ holds, ii) $1 : 1 :: 1 : x$ is solvable (just take $x = 1$), iii) but $g_1(\mathbf{b}') = 0$ is not the solution of $1 : 1 :: 1 : x$.

The function g_2 looks more promising, since the 4-tuple $\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}'$ is associated with $1 : 1 :: 0 : 0$, which holds as an analogical proportion. However, looking at $\mathbf{a}, \mathbf{b}', \mathbf{d}, \mathbf{e}'$ we again have an analogical proportion on the three attributes. However this is associated with $1 : 0 :: 0 : 1$ which is not an analogical proportion. Nonetheless, $1 : 0 :: 0 : x$ is not a solvable proportion: as such, it

cannot be considered as a counter-example for proving that g_2 is not AP. In fact, by an exhaustive search we can see that there is no counter-example in the table showing that g_2 is not AP. Thus, g_2 is AP: class 1 can be described by the equivalence c $\equiv$ ¬r (since ¬r = b ∨ y and ¬c = s here), and class 0 can be described by c $\oplus$ ¬r.

For g_3, we can consider the following four tuples: $\mathbf{c} = (c, y, l)$ with $g_3(\mathbf{c}) = 0$, $\mathbf{c}' = (c, y, h)$ with $g_3(\mathbf{c}') = 0$, $\mathbf{d} = (s, b, l)$ with $g_3(\mathbf{d}) = 1$, and $\mathbf{d}' = (s, b, h)$ with $g_3(\mathbf{d}') = 0$. In this case, we have $\mathbf{c} : \mathbf{c}' :: \mathbf{d} : \mathbf{d}'$ and $solvable(g_3(\mathbf{c}), g_3(\mathbf{c}'), g_3(\mathbf{d}))$, but $g_3(\mathbf{c}) : g_3(\mathbf{c}') :: g_3(\mathbf{d}) : g_3(\mathbf{d}')$ does not hold. This shows that g_3 is not AP. For g_3, class 1 corresponds to the blue light objects, which clearly corresponds to a monomial of degree 2.

The situation is simpler for g_4, where class 1 corresponds to the light objects. It is not difficult to see that it is essentially unary, and thus an AP function.

What happens with these different classification functions is better understood by looking at classification trees, which is the topic of the next subsection.

3.4   Taxonomic trees

A table describing all the possible items that can be distinguished in terms of a set of nominal attributes can be straightforwardly associated with a taxonomic tree, taking the attributes in some order. The tree corresponding to Table 1, with two binary attributes and one ternary one, is given in Figure 1. At the third level, we retrieve the $2 \cdot 2 \cdot 3 = 12$ items from $\mathbf{a}$ to $\mathbf{f}'$. They can be encoded by following the path from the root, using a standard convention: at each level the edges are numbered from the left from 0 to 1, or to 2. Thus, for example, $\mathbf{b}'$ is associated with 011, corresponding to attribute values c, r, h; see Figure 1.

A large number of analogical proportions are hidden between the leaves of such a taxonomic tree. Namely, in our example with 12 items, there are exactly 30 *distinct* analogical proportions on the three attributes (where all the elements in the 4-tuples are distinct). For instance, we have $\mathbf{a} : \mathbf{a}' :: \mathbf{b} : \mathbf{b}'$, or $\mathbf{c} : \mathbf{c}' :: \mathbf{d} : \mathbf{d}'$. This can be checked by observing that here these analogical proportions involving 3 attributes are either

- of the form $uxt : uyt :: vxt : vyt$ (with one constant attribute value), or
- of the form $uxt : vxz :: uyt : vyz$ (with no constant attribute value),

where t, u, v, x, y, z are attribute values, one by attribute in each tuple (such as tux), since an analogical proportion can involve at most 2 distinct values for each attribute. The ordering of attributes has no special meaning, but is the same in each tuple. The 2 above patterns remain the same under symmetry. Note that

$$uxt : uxt :: vxt : vxt \quad \text{and} \quad uxt : vyt :: uxt : vyt$$

are not considered, since they hold trivially by identity or reflexivity. Note also that the form $uxt : vxz :: uyt : vyz$ is the same as $uxt : uyt :: vxz : vyz$ by central permutation (even if the number of constant attributes in the first and second pairs of tuples vary from two to one).
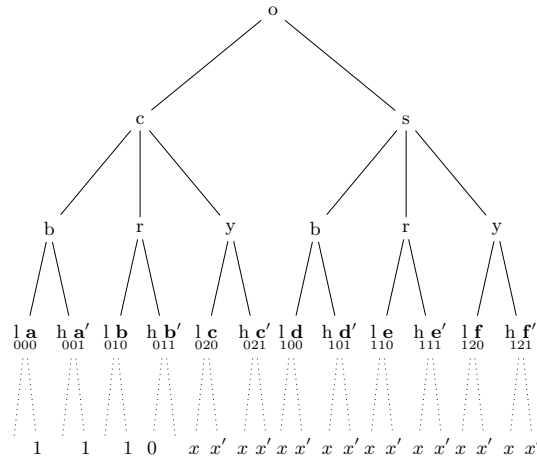
o

c   s

b   r   y   b   r   y

$\underset{000}{\text{l }\mathbf{a}}$ $\underset{001}{\text{h }\mathbf{a}'}$ $\underset{010}{\text{l }\mathbf{b}}$ $\underset{011}{\text{h }\mathbf{b}'}$ $\underset{020}{\text{l }\mathbf{c}}$ $\underset{021}{\text{h }\mathbf{c}'}$ $\underset{100}{\text{l }\mathbf{d}}$ $\underset{101}{\text{h }\mathbf{d}'}$ $\underset{110}{\text{l }\mathbf{e}}$ $\underset{111}{\text{h }\mathbf{e}'}$ $\underset{120}{\text{l }\mathbf{f}}$ $\underset{121}{\text{h }\mathbf{f}'}$

1   1   1 0 x x' x x' x x' x x' x x' x x' x x' x x'

Fig. 1. Example of classification tree.

Reading this taxonomic tree horizontally, there is quite a number of analogical proportions that hold between 4-tuples of items. Assume for a short while that we only have a tree induced by 3 *binary* attributes. Then each of the two forms $uxt : uyt :: vxt : vyt$ and $uxt : vxz :: uyt : vyz$ yields 6 analogical proportions, i.e., in this case we have 12 distinct non-trivial analogical proportions. Indeed, considering the first form $uxt : uyt :: vxt : vyt$, there are $3 \cdot 2 = 6$ possible choices of value for t in case of 3 binary attributes, the possible instantiations of ux, uy, vx, vy being all equivalent due to analogical proportion postulates. For the second form $uxt : vxz :: uyt : vyz$, we can observe that uxt and uyt (as well as vxz and vyz) differ only on one attribute value. There are 6 possible instantiations for this attribute in case of 3 binary attributes, the possible instantiations of the two remaining binary attributes being all equivalent.

In case of two binary attributes and one ternary attribute as in the example, a similar counting can be made. For the first form, we now have $1 \cdot 3 + 2 \cdot 2 = 7$ possible instantiations for t. Moreover when t is not a value of the ternary attribute, we have 3 possible ways of instantiating ux, uy, vx, vy. Altogether the first form then yields $1 \cdot 3 + 2 \cdot 2 \cdot 3 = 15$ analogical proportions. For the second form there are 3 ways of instantiating the "changing attribute" if it is the ternary one; and $2 \cdot 2$ if it is a binary attribute, in this later case there are 3 possible ways of instantiating the remaining attributes. Again, we get 15 analogical proportions, and a total of 30 distinct analogical proportions.

As suggested by the above example, the number of analogical proportions increases rapidly with the cardinalities of the attribute domains and with the number of levels in the tree. This suggests how important the presence of analogical proportions in a classification process is.

Given one of our functions g_i, nothing forbids to consider its value as another attribute (an n-ary one if there are n classes). So instead of considering $\mathbf{a}$, we consider $(\mathbf{a}, g_i(\mathbf{a}))$ as an $(m + 1)$-tuple. We can then carry on the building of the tree as the dotted part in Figure 1 associated with the example. But now,

a choice has to be made between the different classification options for each item corresponding to a path from the root to a leaf: the classification option is related to the value of g_i (in our case 0 or 1). If the class assignment is not "well-balanced" the resulting function will not be AP, as in Fig. 1 where we can clearly observe that $0001 : 0011 :: 0101 : 0110$ does *not* hold (on the 4th digit).

4 Local description of HAP functions: almost affine functions

Recall that $\mathbb{B}$ denotes the 2-element field with $\oplus$ (addition modulo 2) and $\otimes$ (multiplication modulo 2). In the Boolean case, i.e., when $X_1 = \cdots = X_m = X = \mathbb{B}$, the class of AP functions was completely described in [4], where the following theorem was proved:

Theorem 4.1. *A function $f \colon \mathbb{B}^m \to \mathbb{B}$ is AP if and only if it is affine. In particular, the class of AP functions constitutes a clone of Boolean functions, i.e., it contains all projections and it is closed under compositions.*

To extend this result to HAP functions in the arbitrary nominal case (see Subsection 3.1), we shall make use of the following useful observations. From the definition of HAP functions it follows that the restriction $f|_S$ of a HAP function f to any subset $S := S_1 \times \cdots \times S_m \subseteq \mathbf{X} = X_1 \times \cdots \times X_m$ with $|S_i| \leq 2$ ($1 \leq i \leq m$) must have at most 2 distinct values. Therefore, every such function can be thought of as a Boolean function by a suitable renaming of variable and function values. Thus, from Theorem 4.1, we have the following corollary.

Corollary 4.2. *Let X be a set, let $S_1, \ldots, S_m$ be 2-element sets, and set $S := S_1 \times \cdots \times S_m$. A function $f \colon S \to X$ is HAP if and only if there exist $\varphi^f \colon \{0,1\} \to X$ and $\varphi_i^f \colon S_i \to \{0,1\}$ such that*

$$f|_S = \varphi^f(\varphi_1^{f,S}(x_1) \oplus \cdots \oplus \varphi_m^{f,S}(x_m) \oplus c). \tag{1}$$

Remark 4.3. Note that the term $\oplus c$ can be encoded into φ^f so that (1) can be simplified into

$$f|_S = \varphi^f(\varphi_1^{f,S}(x_1) \oplus \cdots \oplus \varphi_m^{f,S}(x_m)). \tag{2}$$

We will generalize these results by introducing the notion of "almost affine" functions, and show that HAP functions are exactly the almost affine functions.

Definition 4.4. *A function $f \colon \mathbf{X} \to X$ is* almost affine *if for any $S_i \subseteq X_i$ with $|S_i| \leq 2$, and $S := S_1 \times \cdots \times S_m$, there exist $\varphi^{f,S} \colon \{0,1\} \to \operatorname{Im}(f|_S)$ and $\varphi_i^{f,S} \colon S_i \to \{0,1\}$ such that $f|_S = \varphi^{f,S}(\varphi_1^{f,S}(x_1) \oplus \cdots \oplus \varphi_m^{f,S}(x_m))$.*

Theorem 4.5. *A function $f \colon \mathbf{X} \to X$ is HAP if and only if it is almost affine.*

Proof. By Corollary 4.2, we know that every HAP function is almost affine. Moreover, by definition, every restriction $f|_S$ of an almost affine function f is AP on $S \subseteq \mathbf{X}$. Therefore, f is HAP. $\qquad\square$

As we will see, this description is quite useful. However, it has some limitations since it requires a local inspection on each subset

$$S = S_1 \times \cdots \times S_m \subseteq \mathbf{X} = X_1 \times \cdots \times X_m.$$

In the next section we provide a global description of almost affine functions, and thus a description of HAP functions.

5 Global description of HAP functions: quasi-linear functions

In the previous section, we showed that the class of HAP functions $f \colon \mathbf{X} \to X$ coincides exactly with the class of almost affine functions. In this section we will show that the HAP functions are either essentially unary or quasilinear.

Definition 5.1. *A function $f \colon \mathbf{X} \to X$ is* quasilinear *if there exist $\varphi \colon \{0,1\} \to X$ and $\varphi_i \colon X_i \to \{0,1\}$ ($1 \leq i \leq m$) such that $f = \varphi(\varphi_1(x_1) \oplus \cdots \oplus \varphi_m(x_m))$.*

We are going to make use of Jablonski's Fundamental Lemma (see, e.g. [14]).

Lemma 5.2 (Jablonski's Fundamental Lemma).

1. *Let $f \colon X^m \to X$ be a function that has at least two essential arguments and $|\mathrm{Im}(f)| = \ell > 2$. Then there exist sets $S_i \subseteq X$ ($1 \leq i \leq m$) with $|S_i| \leq 2$ such that for $S := S_1 \times \cdots \times S_m$, $|\mathrm{Im}(f|_S)| \geq 3$.*
2. *More generally, let $f \colon \mathbf{X} \to X$ be a function that has at least two essential arguments and $|\mathrm{Im}(f)| = \ell > 2$. Then for any k with $2 < k \leq \ell$, there exist sets $S_i \subseteq X_i$ ($1 \leq i \leq m$) with $|S_i| \leq k-1$ such that for $S := S_1 \times \cdots \times S_m$, $|\mathrm{Im}(f|_S)| \geq k$.*

Lemma 5.3. *Let $f \colon \mathbf{X} \to X$. Assume that for all $S_i \subseteq X_i$ with $|S_i| \leq 2$ ($1 \leq i \leq m$), and $S := S_1 \times \cdots \times S_m$, we have that $|\mathrm{Im}(f|_S)| \leq 2$. Then f is essentially unary or $|\mathrm{Im}(f)| \leq 2$.*

Proof. Suppose, to the contrary, that f has at least two essential arguments and $|\mathrm{Im}(f)| = \ell \geq 3$. By Lemma 5.2, item 2, there exist sets $S_i \subseteq X_i$ ($1 \leq i \leq m$) with $|S_i| \leq 2$ such that for $S := S_1 \times \cdots \times S_m$, $|\mathrm{Im}(f|_S)| \geq 3$. We have reached a contradiction. $\qquad\square$

In other words, Lemma 5.3 asserts that an almost affine function f is either essentially unary or has a range of at most two elements.

Proposition 5.4. *A function $f \colon \mathbf{X} \to X$ is almost affine if and only if it is essentially unary or quasilinear.*

Proof. Assume that $f \colon \mathbf{X} \to X$ is almost affine and has at least two essential arguments. Then $|\mathrm{Im}(f)| \leq 2$ by Lemma 5.3.

We are going to show that for all $S_i \subseteq X_i$ ($1 \leq i \leq m$) and $S := S_1 \times \cdots \times S_m$, there exist maps $\varphi^{f,S} \colon \{0,1\} \to X$ and $\varphi_i^{f,S} \colon S_i \to \{0,1\}$ such that

$$f|_S = \varphi^{f,S}\big(\varphi_1^{f,S}(x_1) \oplus \cdots \oplus \varphi_m^{f,S}(x_m)\big).$$

The claim holds whenever $|S_i| \leq 2$ $(1 \leq i \leq m)$ by definition.

We proceed with an inductive argument. Assume that the claim holds for all sets $S_i \subseteq X_i$ $(1 \leq i \leq m)$ with $|S_i| \leq k_i$ for some $k_1, \ldots, k_m$ with $2 \leq k_i \leq |X_i|$. We will show that if $j \in \{1, \ldots, m\}$ and $k_j < |X_j|$, then the claim holds also for all sets $S_i \subseteq X_i$ $(1 \leq i \leq n)$ with $|S_i| \leq k_i$ $(i \neq j)$ and $|S_j| = k_j + 1$.

So, let $S_i \subseteq X_i$ $(1 \leq i \leq m)$ with $|S_i| \leq k_i$ $(i \neq j)$ and $|S_j| = k_j + 1$, and write $S := S_1 \times \cdots \times S_m$. Assume that $a, b \in S_j$, $a \neq b$. Let $S_j' := S_j \setminus \{b\}$ and $S_j^* := \{a, b\}$, and let

$$
\begin{aligned}
S' &:= S_1 \times \cdots \times S_{j-1} \times S_j' \times S_{j+1} \times \cdots \times S_m, \\
S^* &:= S_1 \times \cdots \times S_{j-1} \times S_j^* \times S_{j+1} \times \cdots \times S_m, \\
T &:= S_1 \times \cdots \times S_{j-1} \times \{a\} \times S_{j+1} \times \cdots \times S_m = S' \cap S^*, \\
T' &:= S_1 \times \cdots \times S_{j-1} \times \{b\} \times S_{j+1} \times \cdots \times S_m = S \setminus S'.
\end{aligned}
$$

By the inductive hypothesis, there exist maps $\varphi^{f,S'}$, $\varphi_i^{f,S'}$ $(1 \leq i \leq m)$ and φ^{f,S^*}, φ_i^{f,S^*} $(1 \leq i \leq m)$ such that $f|_{S'} = \varphi^{f,S'}\big(\varphi_1^{f,S'}(x_1) \oplus \cdots \oplus \varphi_m^{f,S'}(x_m)\big)$, and $f|_{S^*} = \varphi^{f,S^*}\big(\varphi_1^{f,S^*}(x_1) \oplus \cdots \oplus \varphi_m^{f,S^*}(x_m)\big)$.

Let now $\varphi^{f,S} := \varphi^{f,S'}$, $\varphi_i^{f,S} := \varphi_i^{f,S'}$ for $i \neq j$, and define $\varphi_j^{f,S} : S_j \to \{0,1\}$ as the extension of $\varphi_j^{f,S'} : S_j' \to \{0,1\}$ that satisfies the condition $\varphi_j^{f,S}(a) = \varphi_j^{f,S}(b)$ if and only if $\varphi_j^{f,S^*}(a) = \varphi_j^{f,S^*}(b)$.

It remains to show that $f|_S = \varphi^{f,S}\big(\varphi_1^{f,S}(x_1) \oplus \cdots \oplus \varphi_m^{f,S}(x_m)\big)$. Let $\mathbf{x} = (x_1, \ldots, x_m) \in S$. If $\mathbf{x} \in S'$ then

$$
\varphi^{f,S}\big(\varphi_1^{f,S}(x_1) \oplus \cdots \oplus \varphi_m^{f,S}(x_m)\big) = \varphi^{f,S'}\big(\varphi_1^{f,S'}(x_1) \oplus \cdots \oplus \varphi_m^{f,S'}(x_m)\big) = f(\mathbf{x}).
$$

Assume now that $\mathbf{x} \in S \setminus S' = T'$. Then $x_j = b$, so $\mathbf{x} \in S^*$. If $\varphi_j^{f,S^*}(a) = \varphi_j^{f,S^*}(b)$, then also $\varphi_j^{f,S}(a) = \varphi_j^{f,S}(b)$, and we have

$$
\begin{aligned}
&\varphi^{f,S}\big(\varphi_1^{f,S}(x_1) \oplus \cdots \oplus \varphi_j^{f,S}(b) \oplus \cdots \oplus \varphi_m^{f,S}(x_m)\big) = \\
&\varphi^{f,S}\big(\varphi_1^{f,S}(x_1) \oplus \cdots \oplus \varphi_j^{f,S}(a) \oplus \cdots \oplus \varphi_m^{f,S}(x_m)\big) = \\
&\varphi^{f,S^*}\big(\varphi_1^{f,S^*}(x_1) \oplus \cdots \oplus \varphi_j^{f,S^*}(a) \oplus \cdots \oplus \varphi_m^{f,S^*}(x_m)\big) = \\
&\varphi^{f,S^*}\big(\varphi_1^{f,S^*}(x_1) \oplus \cdots \oplus \varphi_j^{f,S^*}(b) \oplus \cdots \oplus \varphi_m^{f,S^*}(x_m)\big) = f(\mathbf{x}).
\end{aligned}
$$

If $\varphi_j^{f,S^*}(a) \neq \varphi_j^{f,S^*}(b)$, i.e., $\varphi_j^{f,S^*}(a) \oplus 1 = \varphi_j^{f,S^*}(b)$, then also $\varphi_j^{f,S}(a) \oplus 1 = \varphi_j^{f,S}(b)$, and we have

$$
\begin{aligned}
&\varphi^{f,S}\big(\varphi_1^{f,S}(x_1) \oplus \cdots \oplus \varphi_j^{f,S}(b) \oplus \cdots \oplus \varphi_m^{f,S}(x_m)\big) = \\
&\varphi^{f,S}\big(\varphi_1^{f,S}(x_1) \oplus \cdots \oplus \varphi_j^{f,S}(a) \oplus 1 \oplus \cdots \oplus \varphi_m^{f,S}(x_m)\big) = \\
&\varphi^{f,S^*}\big(\varphi_1^{f,S^*}(x_1) \oplus \cdots \oplus \varphi_j^{f,S^*}(a) \oplus 1 \oplus \cdots \oplus \varphi_m^{f,S^*}(x_m)\big) = \\
&\varphi^{f,S^*}\big(\varphi_1^{f,S^*}(x_1) \oplus \cdots \oplus \varphi_j^{f,S^*}(b) \oplus \cdots \oplus \varphi_m^{f,S^*}(x_m)\big) = f(\mathbf{x}).
\end{aligned}
$$

Therefore $\varphi^{f,S}\big(\varphi_1^{f,S}(x_1) \oplus \cdots \oplus \varphi_m^{f,S}(x_m)\big) = f(\mathbf{x})$ for all $\mathbf{x} \in S$. $\qquad\square$

Example 5.5. Note that both g_2 and g_4 of Subsection 3.3 are in fact HAP functions, since both are quasilinear. Indeed, g_2 and g_4 can be represented as

$$g_2(x_1, x_2, x_3) = \varphi(\varphi_1(x_1) \oplus \varphi_2(x_2) \oplus \varphi_3(x_3))$$

with

$$
\begin{aligned}
&\varphi_1 \colon \{c, s\} \to \{0, 1\}, && c \mapsto 0, && s \mapsto 1, \\
&\varphi_2 \colon \{b, r, y\} \to \{0, 1\}, && b \mapsto 1, && r \mapsto 0, && y \mapsto 1, \\
&\varphi_3 \colon \{l, h\} \to \{0, 1\}, && l \mapsto 0, && h \mapsto 0, \\
&\varphi \colon \{0, 1\} \to \{0, 1\}, && 0 \mapsto 0, && 1 \mapsto 1
\end{aligned}
$$

and

$$g_4(x_1, x_2, x_3) = \psi(\psi_1(x_1) \oplus \psi_2(x_2) \oplus \psi_3(x_3))$$

with

$$
\begin{aligned}
&\psi_1 \colon \{c, s\} \to \{0, 1\}, && c \mapsto 0, && s \mapsto 0, \\
&\psi_2 \colon \{b, r, y\} \to \{0, 1\}, && b \mapsto 0, && r \mapsto 0, && y \mapsto 0, \\
&\psi_3 \colon \{l, h\} \to \{0, 1\}, && l \mapsto 1, && h \mapsto 0, \\
&\psi \colon \{0, 1\} \to \{0, 1\}, && 0 \mapsto 0, && 1 \mapsto 1.
\end{aligned}
$$

Moreover, g_4 is essentially unary because

$$g_4(x_1, x_2, x_3) = \gamma(x_3)$$

with $\gamma \colon \{l, h\} \to \{0, 1\}$, $l \mapsto 1$, $h \mapsto 0$.

We have seen that the class of HAP functions on X coincides with the class of almost affine functions on X. In the case when $X_1 = \cdots = X_m = X$, this class is exactly Burle's clone of quasilinear functions [3], thus generalizing the last assertion of Theorem 4.1.

Proposition 5.6. *For every finite X, the class of HAP functions on X constitutes a clone, i.e., it contains every projection on X and it is closed under forming compositions: if $f \colon X^n \to X$ and each $g_i \colon X^m \to X$ is HAP on X, then so is $f' = f(g_1, \ldots, g_n) \colon X^m \to X$.*

6   Conclusion

The above results describe the class of hard analogy-preserving functions over finite domains, including the Boolean case as a particular case. The case of non-finite domains remains open. Still it is an important step towards a better understanding of analogical inference. The analogy-preserving functions are those for which analogical inference never fails for predicting their values. Still the set of situations where analogical inference gives good predictions is much larger, as

shown by experiments, since a good prediction does not require that *all* predictions given by triplets are the same, but that a majority of triplets give the good prediction. However, these theoretical results contribute to a better understanding of analogical inference and show that applying analogical proportion-based inference amounts to enforcing linearity as much as possible, at least in a local manner. Analogical proportions are pervasive, as shown by their abundance in taxonomic trees, and are thus an important notion for reasoning from data.

References

1. Barbot, N., Miclet, L., Prade, H., Richard, G.: A new perspective on analogical proportions. In: Proc. 15th Eur. Conf. Symb. and Quantit. Approaches to Reas. with Uncert. (ECSQARU'19), Belgrade. pp. 163–174. Springer, LNCS 11726 (2019)
2. Bounhas, M., Prade, H., Richard, G.: Analogy-based classifiers for nominal or numerical data. Int. J. Approx. Reasoning **91**, 36–55 (2017)
3. Burle, G.A.: Classes of k-valued logic which contain all functions of a single variable. Diskret. Analiz **10**, 3–7 (1967), (Russian)
4. Couceiro, M., Hug, N., Prade, H., Richard, G.: Analogy-preserving functions: A way to extend Boolean samples. In: Proc. 26th Int. Joint Conf. on Artificial Intelligence, IJCAI'17, Melbourne,-Aug. 19-25. pp. 1575– 1581 (2017)
5. Couceiro, M., Hug, N., Prade, H., Richard, G.: Behavior of analogical inference w.r.t. Boolean functions. In: Lang, J. (ed.) Proc. 27th Int. Joint Conf. on Artificial Intelligence, IJCAI'18, July 13-19, Stockholm. pp. 2057–2063. ijcai.org (2018)
6. Crama, Y., Hammer, P.: Boolean Functions - Theory, Algorithms, and Applications, Encyclopedia of mathematics and its applications, vol. 142. Cambridge University Press (2011)
7. Hesse, M.: On defining analogy. Proc. of the Aristotelian Soc. **60**, 79–100 (1959)
8. Klein, S.: Analogy and mysticism and the structure of culture (and Comments and Reply). Current Anthropology **24 (2)**, 151–180 (1983)
9. Langlais, P., Patry, A.: Translating unknown words by analogical learning. In: Joint Conf. on Empirical Methods in Natural Language Processing (EMNLP) and Conf. on Computat. Natural Language Learning (CONLL). pp. 877–886. Prague (2007)
10. Lepage, Y.: Analogy and formal languages. In: Proc. FG/MOL. pp. 373–378 (2001)
11. Miclet, L., Bayoudh, S., Delhay, A.: Analogical dissimilarity: definition, algorithms and two experiments in machine learning. JAIR **32**, 793–824 (2008)
12. Miclet, L., Prade, H.: Handling analogical proportions in classical logic and fuzzy logics settings. In: Proc. 10th Eur. Conf. Symb. and Quantit. Approaches to Reas. with Uncert. (ECSQARU'09),Verona. pp. 638–650. Springer, LNCS 5590 (2009)
13. Prade, H., Richard, G.: Analogical proportions: From equality to inequality. Int. J. Approx. Reasoning **101**, 234–254 (2018)
14. Salomaa, A.: On essential variables of functions, especially in the algebra of logic. Ann. Acad. Sci. Fenn. Ser. A I **339**, 1–11 (1963)
15. Stone, M.H.: The theory of representation for Boolean algebras. Trans. of the American Mathematical Society **40(1)**, 37–111 (1936)
16. Stroppa, N., Yvon, F.: Du quatrième de proportion comme principe inductif : une proposition et son application à l'apprentissage de la morphologie. Traitement Automatique des Langues **47**(2), 1–27 (2006)
17. Zhegalkin, I.I.: On the technique of calculating propositions in symbolic logic. Mat. Sb. **43**, 9–28 (1927)