Towards a unified framework of numerical analogies: Open questions and perspectives

Yves Lepage[1,*], Miguel Couceiro[2,3]

[1]IPS, Waseda University, Japan
[2]LORIA, Université de Lorraine, France
[3]INESC-ID, IST, Universidade de Lisboa, Portugal

Abstract
A recently-proposed framework for modeling analogies exploits tight relationships between analogies and generalized means, and unifies different models of numerical analogies. In this contribution we further exploit this framework and discuss pertaining questions such as invariance properties, the semantics and the expressive power of such analogies. We also explore further topics related to equivalence classes and invariance properties, generalized mean approximations, manifold representations, as well as algorithmic aspects and machine learning applications such as machine classification.

Keywords
Analogy, Analogical inference, Formalization, Case-Based Reasoning, Knowledge Representation

1. Introduction and motivation

Analogical reasoning (AR) is a remarkable capability of human thought that exploits parallels between situations of different nature to infer plausible conclusions, by relying simultaneously on similarities and dissimilarities. Machine learning (ML) and artificial intelligence (AI) have tried to develop AR but the early works lacked theoretical and formalization support or are limited to Booleans. This situation started to change two decades ago when researchers adopted the view of the so-called analogical proportions as statements of the form "a relates to b as c relates to d", usually denoted $a : b :: c : d$. Such proportions are at the root of the analogical inference mechanism, and several formalisms to study this mechanism have been proposed, which follow different axiomatic and logical approaches [1, 2, 3]. The study of relationships between two pairs of objects A and B, and of objects C and D, can focus on different aspects. For instance, one may want to judge whether the relationship between C and D is the same as that between A and B, and assess the quality of such relationships, or one may discuss the quality of attribute or relationship similarities [4], following the foundational work by Gentner [5]. One may also see A and C as problems, B as a solution to problem A, and ask whether the transposition of the ratio of A to B on C generates a D, and to what extent the

IARML@IJCAI'2024: Workshop on the Interactions between Analogical Reasoning and Machine Learning, at IJCAI'2024, August, 2024, Jeju, South Korea
*Corresponding author.
yves.lepage@waseda.jp (Y. Lepage); miguel.couceiro@loria.fr (M. Couceiro)
0000-0002-3059-4271 (Y. Lepage); 0000-0003-2316-7623 (M. Couceiro)
© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

generated D is a solution to problem C.

There are two basic tasks associated with AR. The first is *analogy making* that corresponds to the task of detecting and deciding whether a quadruple (a, b, c, d) corresponds to a valid analogical proportion. The second is *analogy solving* that refers to the task of finding or extrapolating, for a given triple a, b, c the value x such that $a : b :: c : x$ is a valid analogy. This task is typically addressed in the literature by retrieval and adaptation, i.e., defining an x from a pool of retrieved candidate solutions to be suitably adapted. In fact, analogy solving is somewhat related to case-based reasoning (CBR) [6, 7, 8] where, given a set P of problems, a set S of solutions and a set C of cases $(x, y) \in P \times S$, the CBR task is to find a solution y_t to a given target problem x_t. CBR basically consists in (1) selecting k source cases in the case base according to some criteria related to the target problem (retrieval step), and (2) reusing the k retrieved cases for proposing a target solution (adaptation step). For $k = 1$, the desired solution y_t corresponds to the solution of the analogical equation $x : y :: x_t : y_t$. For higher values of k, different models of analogy on P and S could be taken into account. For instance, when $k = 3$, the retrieval task consists in finding a triple of cases (x_1, y_1), (x_2, y_2) and (x_3, y_3), such that $x_1 : x_2 :: x_3 : x_t$ is valid and such that $y_1 : y_2 :: y_3 : y$ is solvable in y. In this setting the desired y_t would then be one of such solutions [9].

The latter idea was extended to analogy based classification [10] where objects are viewed as attribute tuples (instances) $\mathbf{x} = (x_1, \ldots, x_n)$. Here, analogical inference relies on the idea that if four instances $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$ are in analogical proportion for most of the attributes describing them, then it may still be the case for the other attributes. Similarly, if class labels are known for $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and unknown for $\mathbf{d}$, then one may infer the label for $\mathbf{d}$ as a solution of an analogical proportion equation. Theoretically, it is quite challenging to find and characterize situations where such an *analogical inference principle* (AIP) can be soundly applied, but there have been several efforts in this direction [11, 12, 13]. In fact, the latter culminated recently in a Galois theory of analogical classifiers [14] illustrated over classical numerical analogies. In fact, the principle of analogical inference has been integrated into various machine learning tasks, *e.g.* preference learning and recommendation [15, 16, 17], and used to solve difficult reasoning tasks such as academic aptitude tests, visual question answering and target sense verification [18, 19, 20, 21], as well as support computational creativity [22], analogical transfer [23, 24], and explainable AI [25, 26].

Recently, analogies and AR became quite popular due to the successes of deep learning together with distributional representations (embeddings). For instance, [27, 28] show that vector representations of quadruples respecting certain linear transformations satisfy common properties of analogies, whereas [29] unveils the potential analogies as a benchmark to evaluate the quality of embedding models. This exploration has extended to complex structures such as knowledge graphs (KG) over multimodal domains, to address tasks such as named entity recognition, link prediction, relation discovery (abduction) and KG bootstrapping and completion [30, 31, 32], by leveraging multimodal knowledge embeddings [33, 34, 35, 36]. Despite these impressive results by such analogy-based approaches in rather complex tasks, many works have questioned the retrieved analogical relations, their dependency and limitation with respect to the underlying representation model, and even the evaluation procedure [37, 38]. Many works have been advocating for foundational mathematical frameworks and experiments to gain a better understanding of the analogical capabilities of embedding models as well as recent large language models [39, 40, 41, 42].

In this paper we revisit the recent unified framework of numerical analogies proposed in [43]. It is rooted in the intuitive and classical idea of "analogy", where a quadruple (A, B, C, D) is said to constitute a valid analogy, denoted by $A : B :: C : D$, if the "mean of the *extremes* A and D is equal to the mean of the *means* [1] B and C". Here, we consider a generalized notion of mean that goes back to Hölder [44]. (see Section 2). Endowed with this notion, the authors of [43] introduced the so-called *analogies in power* p (see Section 2), and showed several noteworthy and rather surprising results, namely, that

1. any sequence of four increasing positive real numbers is an analogy in a unique suitable power,
2. any such analogy can be reduced to an equivalent arithmetic analogy, and
3. any analogical equation has a solution for increasing numbers.

In this position paper, we survey these and other results, and further explore the potential of the unifying framework proposed in [43]. In particular, we will address mathematical challenges revealed by this new analogy formalization, and discuss potential impacts in algorithmic approaches as well as downstream ML tasks. The paper is thus organized as follows. In Section 2 we briefly recall the framework proposed in [43] to formalize analogies. We then summarize a few surprising results in Section 3. The discussion on the open problems and perspectives is carried out in Section 4, and divided into mathematical challenges (Subsection 4.1) and ML impacts (Subsection 4.2).

Notation and terminology. We will use lower case letters a, b, c and d when the terms of the analogy are numbers. We will also use $\mathbb{R}$, $\mathbb{R}^*$, $\mathbb{R}^+ \setminus \{0\}$ and $\mathbb{C}$ to denote the sets of real numbers, real numbers without 0, positive numbers and complex numbers, respectively.

2. Generalized means and analogy

Recall that the p-*generalized mean* of real positive numbers $x_1, \ldots x_N$ is

$$m_p(x_1, \ldots x_N) = \lim_{r \to p} \sqrt[r]{\frac{1}{N} \sum_{i=1}^{N} x_i^r}$$

for all $p \in [-\infty, +\infty]$. Note that this notion subsumes the classical notions of:

- arithmetic mean:
$$m_1(x_1, \ldots x_N) = \frac{1}{N} \sum_{i=1}^{N} x_i;$$
- harmonic mean:
$$m_{-1}(x_1, \ldots x_N) = \frac{1}{\frac{1}{N} \sum_{i=1}^{N} \frac{1}{x_i}};$$

[1] In this paper the term *mean* will be understood in two different ways. The *means* in an analogy will refer to the terms B and C, while *generalized means* of two numbers will refer to the notion of an average. Hopefully, the context will be clear and will not lead to any ambiguity.

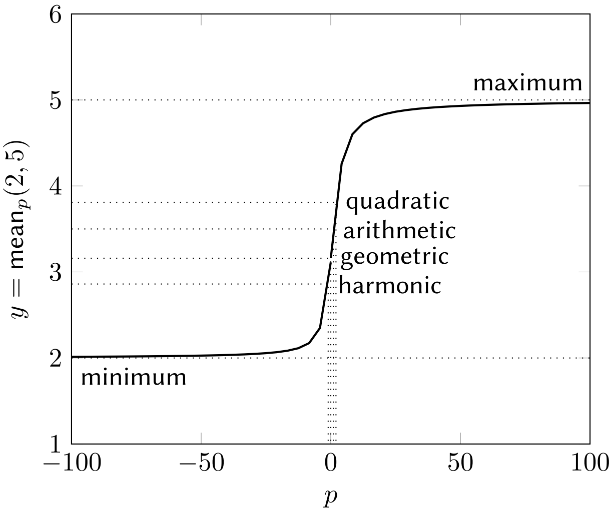

Figure 1: Visualization of generalized means of $a = 2$ and $d = 5$.

- root mean square:

$$m_2(x_1, \ldots x_N) = \sqrt[2]{\frac{1}{N}\sum_{i=1}^{N} x_i^2};$$

- geometric mean:

$$\lim_{p\to 0} m_p(x_1, \ldots x_N) = \sqrt[N]{\prod_{i=1}^{N} x_i};$$

- maximum:

$$\lim_{p\to +\infty} m_p(x_1, \ldots x_N) = \max(x_1, \ldots x_N);$$

- minimum:

$$\lim_{p\to -\infty} m_p(x_1, \ldots x_N) = \min(x_1, \ldots x_N).$$

The generalized mean of two numbers a and d reduces to

$$m_p(a,d) = \lim_{r\to p} \sqrt[r]{\frac{1}{2}(a^r + d^r)},$$

for all $p \in \mathbb{R}$, and even for p equal to $-\infty$ or $+\infty$. Figure 1 illustrates the curve obtained for the particular values $a = 2$ and $d = 5$. It is noteworthy that the generalized mean function of two numbers does not behave like an odd function with respect to p. Although defined above for positive reals, this can be generalized to the case when a, d and p are complex numbers, except in some undefined cases. Figure 2 shows the curve obtained for the generalized means of two complex numbers $-20 + i20$ and $50 + i70$.

Based on this notion of generalized means, the authors of [43] proposed the following unified notion of numerical analogy.

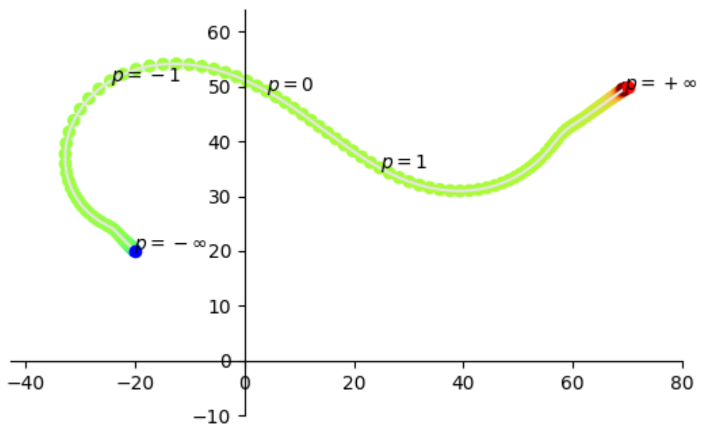

Figure 2: Generalized means for the two complex numbers $a = -20 + i20$ and $d = 50 + i70$. The color is supposed to reflect the transition from $p = -\infty$ in blue on a up to $p = +\infty$ in red on d, through all values in $\mathbb{R}$. The points around $p = 0$ are in green and they visibly occupy the main part of the curve. The point for $p = 0$. i.e., the geometric mean, is located near the vertical axis. The harmonic ($p = -1$) and the arithmetic means ($p = 1$) are also indicated.

Definition 1. *Let $a, b, c, d \in \mathbb{R}^+ \setminus \{0\}$, and $p \in \mathbb{R}$. The* analogy in power p *is defined as follows:*

$$a : b ::^p c : d \overset{def.}{\Leftrightarrow} m_p(a,d) = m_p(b,c) \;\Leftrightarrow\; \lim_{r \to p} \sqrt[r]{\frac{1}{2}(a^r + d^r)} = \lim_{r \to p} \sqrt[r]{\frac{1}{2}(b^r + c^r)}\,. \tag{1}$$

This definition generalizes the classical arithmetic, geometric, or harmonic analogies, as illustrated in Table 1. Note that when $p \in \mathbb{R}^*$, the limits in (1) are not necessary, neither is taking the p-th root, the one-half factor can be eliminated and demonstrations can then directly exploit the formula $a^p + d^p$ or the equality $a^p + d^p = b^p + c^p$. In the case $p = 0$, the simplified versions of the formulas for the geometric mean are used: $a \times d$ or $a \times d = b \times c$. Finally, in the two infinite cases, formulas with $\min$ and $\max$ are used.

The above defined relation between four terms defines a mathematical analogy in the expected sense of the term, *i.e.*, firstly, :: is a dependence relation (reflexivity and symmetry). Indeed, transitivity is also verified so that, :: is an equivalence relation in this case. Secondly, it is easy to verify that the permutation of the means $a : c :: b : d$ and the inverse of the ratios $b : a :: d : c$ hold, which implies in total eight equivalent forms, that classical formalizations of mathematical analogy allow writing for the same analogy, by playing on the position of the terms:

$$\begin{array}{llll} a : b :: c : d & b : a :: d : c & c : a :: d : b & d : b :: c : a \\ a : c :: b : d & b : d :: a : c & c : d :: a : b & d : c :: b : a \end{array}$$

Table 1
Some illustrative examples subsuming instances of classical models of numerical analogies.

Data	Equality of means	Analogy	Name of analogy
5, 7, 10, **12**	$\frac{1}{2}(\mathbf{5}+\mathbf{12}) = \frac{1}{2}(7+10)$	$\mathbf{5} : 7 ::^{1} 10 : \mathbf{12}$	Arithmetic
2, 4, 8, **16**	$\sqrt{\mathbf{2}\times\mathbf{16}} = \sqrt{4\times 8}$	$\mathbf{2} : 4 ::^{0} 8 : \mathbf{16}$	Geometric
8/5, 2, 4, **8**	$\dfrac{1}{\frac{1}{2}(\frac{1}{\frac{\mathbf{8}}{\mathbf{5}}}+\frac{1}{\mathbf{8}})} = \dfrac{1}{\frac{1}{2}(\frac{1}{2}+\frac{1}{4})}$	$\frac{\mathbf{8}}{\mathbf{5}} : 2 ::^{-1} 4 : \mathbf{8}$	Harmonic
2, 2, 5, **67**	$\min(\mathbf{2},\mathbf{67}) = \min(2,5)$	$\mathbf{2} : 2 ::^{-\infty} 5 : \mathbf{67}$	By minimum value

3. Main results of Lepage & Couceiro [43]

The main contributions of article [43] can be summarized in three rather surprising results. The first one states that analogies exist and are unique between any quadruple of positive real numbers in increasing order (the numbering of the results does not reflect the importance but the order afterwards).

Theorem 2. *Given four increasing positive real numbers, all different, we can always see an analogy between them and this analogy is unique.*

Note that the most common form of analogy between numbers is the arithmetic analogy, which is just the particular case $p = 1$. In addition to Theorem 2, the second main result states that any analogy between four positive reals can be thought of as an arithmetic analogy, as it can always be transformed into an arithmetic analogy.

Theorem 3. *Any analogy in p between four positive real numbers can be reduced to an arithmetic analogy.*

As vector representations in machine learning make use of real numbers, the main results are presented with respect to real numbers, but extensions are possible. In particular, the third main result states that every analogical equation of the form $A : B :: C : x$ is solvable over the complex numbers.

Theorem 4. *Conditioned on p, it is possible to solve any analogical equation with complex terms.*

This mathematical formalism to model analogies reveals several challenges, both of mathematical and of ML nature, that we will discuss in the next section.

4. Perspectives opened by this formalization

We will discuss some challenges and perspectives opened by this formalization of numerical analogies. We identify two main directions of future work that correspond to the following sections.

4.1. Mathematical aspects

Equivalence classes. From the definition of $::^p$, it is not difficult to verify that for, $p \neq 0, +\infty, -\infty$, the relation $::^p$ is reflexive, symmetric and transitive, and thus constitutes an equivalence relation. It follows from Theorem 2 that the corresponding equivalence classes are in bijection. Furthermore, the authors of [43] also showed that $::^p$ remains an equivalence relation in the cases $p = +\infty, -\infty$, and that the corresponding equivalence classes remain in bijection. However, this situation seems to change for the case $p = 0$.

It is thus natural to ask for a structural study of these equivalence classes. The answer to this question would shed some light into potential axiomatizations of $::^p$ and some of the challenges below. For instance, we could determine and transfer invariance results between equivalence classes (see below). Also, it would make clear the meaning of p (see Subsection 4.2). Furthermore, one may ask whether such a trichotomy with respect to p still holds in the complex case.

Invariance properties. Let $R \subseteq \mathbb{S}^n$ be an n-ary relation over $\mathbb{S} \in \{\mathbb{R}, \mathbb{R}^*, \mathbb{R}^+ \setminus \{0\}, \mathbb{C}\}$, and let $f : \mathbb{S} \to \mathbb{S}$. Function f is said to be a *homomorphism* of R, and relation R is said to be an *invariant*[2] under f if for every $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in R$, we have

$$f(\mathbf{a}) = (f(a_1), f(a_2), \ldots, f(a_n)) \in R.$$

For instance, it was observed in [45, 43] that any affine transformation (translation) leaves the analogy invariant for $p = 1$. Similarly, the multiplication by any positive scalar leaves analogies invariant for any real p.

These facts ask for the connection between p and invariance properties, *e.g.*, what are the classes of $::^p$ invariant under a certain class of functions or, conversely, given a p, what is the class of homomorphisms. These questions can be extended to higher order relations with an immediate application in the study of analogical classifiers (see Subsection 4.2). Also, taking the "functional" definition of analogy under which $a : b :: c : d$ holds whenever there is a transformation t such that $b = t(a)$ and $d = t(c)$, one may want to characterize the functions that leave such analogies invariant. Moreover, we propose to extend these questions to the richer realm of complex numbers.

Behavior of analogies on manifolds. Word embeddings are sets of vector representations of words based on the distributional hypothesis that the semantics of a word is characterized by its context. This has been implemented in efficient methods like Word2Vec and its successors. It has been claimed that, in such a space, analogies can be formalized by an arithmetic relation and theoretical arguments have been given in support of this [28, 46]. The Google analogy test

[2]Note that these notions naturally extend to any underlying set $\mathbb{S}$.

set [47] or BATS [39] are collections of analogies grouped by types of relations. For instance, the capital-common-country type contains analogies like: *Paris* : *France* :: *Berlin* : *Germany* or *Tehran* : *Iran* :: *Havana* : *Cuba.* We wonder whether significant information can be drawn from the observation of the various analogical powers that one can compute on each different dimension of each different analogy of the same type. Because the set of words contained in such a group of analogies can be seen as defining, by extension, a manifold in the representation space, an open question is whether observations on the analogical powers would allow to characterize this manifold.

Another way of addressing the above question is to have a view that separates the terms in an analogy. To illustrate with an example, given a number of points (*e.g.*, a certain number of masculine titles like *emperor*, *count*, *duke* and *king*) and some corresponding points (feminine titles: *empress*, *countess*, *duchess*), one may try to characterize not only each of the two manifolds for these two sets of words, but also the correspondance betweenthese manifolds, or even the manifold that separates the two sets of words, by reminiscence with classification and SVMs. How would that allow to deduce, on the basis of analogical powers, the fact that the last missing point in the list above should be *queen*?

4.2. Machine learning: semantics and practical aspects

Meaning of p. As suggested in the previous section about equivalence classes, a question is whether a meaning can be assigned to p. It comes naturally to mind that it could possibly reflect some strength of the analogy. This stems from the fact that the two powers $-\infty$ and $+\infty$ do not have exactly the same properties as the other powers. In particular, an analogical equation in $-\infty$ or $+\infty$ looks more relaxed than in other powers because, there may exist an infinite number of solutions. For instance $2 : 2 ::^{-\infty} 3 : x$ admits any value for x greater than 3 as a solution. Although it seems that this is a particular behaviour for the infinities, it also seems that, depending on the kind of data at hand, some analogical powers might have a heavier signification than, say other analogical powers. For instance, it could be the case in word embedding spaces, where arithmetic analogies are being felt as the most reasonable kind of analogy based on theoretical considerations [46], the dimensions on which the analogical power p departs greatly from 1 might be considered as less representative of the essence of the relationship expressed by the word analogy.

Algorithmic aspects In the current state of our implementation, the analogical power for a quadruple of real numbers is determined with a certain precision by dichotomic search. We implemented a Python package that wraps C functions.[3] For a million quadruples of increasing positive real numbers drawn at random, the determination of the analogical power with a precision of 10^{-4} takes 2.3 seconds on an Apple M1 processor (3.2 GHz). Acceleration being always wanted, practical as well as theoretical questions can be considered.

Firstly, in practice, one may inspect whether exploiting some properties of invariance, *e.g.*, by multiplication or reduction to arithmetic analogy, allows the use of tabular techniques to accelerate computation. Tables of pre-computed cases could be loaded into memory to allow

[3]http://lepage-lab.ips.waseda.ac.jp/projects/Kakenhi_Project_21K12038/ > Experimental results > Python packages

for fast access. In conjunction with such tabulation techniques, the use of quantization can also be considered.

Secondly, more theoretically, one can ask whether the use of analytic forms of approximating functions can be exploited. The analytic expression of the generalized mean function itself is reasonably simple, but its derivative and in particular the derivative of differences, as needed in analogy, have rather complex expressions that would entail heavy computations if one would envisage gradient techniques. By contrast, the "popularity" of the sigmoid function as an activation function in neural networks is to be found in its simple analytic properties, in particular the simplicity of its derivative $\sigma'(x) = \sigma(x) \times (1 - \sigma(x))$, and the existence of a central symmetry $\sigma(-x) = 1 - \sigma(x)$, something that the generalized mean function does not have. It is however undeniable that, despite the previous remark, the generalized mean function looks very much alike the sigmoid function. In the hope of faster computation to determine analogical powers, one can ask whether approximations of the generalized mean function by, *e.g.*, the sigmoid function, could lead to accelerations. This first entails the determination of suitable parameters to fit the sigmoid function onto the generalized mean function. The next task is to estimate the error under such an approximation and to quantify the trade-off between such an error and the acceleration, if any is obtained.

Relationships between invariance properties and classification As discussed in the Introduction, analogy solving and analogical transfer are tightly related to case-based reasoning (CBR) [8]. In particular, we mentioned the case where retrieval consists in finding a triple of cases such that analogy holds in the problem domain, and finding a solution consists in solving the corresponding analogical equation in the solution domain. This idea was extended to analogy based classification [10] where objects are viewed as attribute tuples (instances) $\mathbf{x} = (x_1, \ldots, x_n)$. Here, analogical inference relies on the idea that if four instances $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ are in analogy for most of the attributes describing them, then it may still be the case for the other attributes. Similarly, if class labels are known for $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and unknown for $\mathbf{d}$, then one may infer the label for $\mathbf{d}$ as a solution of an analogical equation. Theoretically, it is quite challenging to find and characterize situations where such an *analogical inference principle* (AIP) can be soundly applied. In case of Boolean attributes, a first step for explaining the analogical mechanism was to characterize the set of functions for which AIP is sound (*i.e.*, no error occurs) no matter which triplets of examples are used. In [12], it was shown 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 [13]. These results were extended in [11] to nominal domains when taking the *minimal model* of analogy, i.e., only patterns of the form $x : x :: y : y$ and $x : y :: x : y$.

The latter were unified into a Galois theory of analogical classifiers [14] by exploiting the notion of "polymorphism", in which formal models of analogy are put in a two-way correspondence with classifiers compatible with the analogical inference principle. To some extent, this reflects invariance properties of higher-order, in which homomorphisms may be n-ary, for $n \geq 2$. This asks for revisiting the Galois framework of [14] under Defintion 1, and obtaining explicit descriptions of the Galois sets of analogical classifiers, possibly parametrized by p.

5. Conclusion

In this position paper, we surveyed the recently proposed formalization of numerical analogy, more precisely on positive real numbers, with its definition and the three main results associated with it. This framework opens new avenues in two main directions. The first direction we identified concerns the mathematical aspects for machine learning pertaining to invariance properties and the behavior of analogies. The second one pertains to theoretical and practical aspects, such as the semantics, explainability considerations, and impacts on downstream ML tasks.

Acknowledgments

This work is the result of a collaboration made possible by the first author's sabbatical stay with the second author's organization. The authors would like to thank Waseda University Research Promotion Section and the LORIA International Chair for their funding.

This work benefited from a research grant from the Japan Society for the Promotion of Science, Kiban C, n° 21K12038, entitled "Theoretically grounded algorithms for the automatic extraction of analogy test sets in natural language processing" and was partially supported by the ANR project "Analogies: from theory to tools and applications" (AT2TA), ANR-22-CE23-0023.

References

[1] Y. Lepage, Analogy and formal languages, Electron. Notes Theor. Comput. Sci. 53 (2001) 180–191.
[2] L. Miclet, S. Bayoudh, A. Delhay, Analogical dissimilarity: Definition, algorithms and two experiments in machine learning, J. Artif. Intell. Res. 32 (2008) 793–824.
[3] H. Prade, G. Richard, Analogical proportions: Why they are useful in AI, in: Proc. IJCAI-21, 2021, pp. 4568–4576. Survey Track.
[4] P. D. Turney, Similarity of semantic relations, Computational Linguistics 32 (2006) 379–416.
[5] D. Gentner, Structure mapping: A theoretical model for analogy, Cognitive Science 7 (1983) 155–170.
[6] F. Badra, M. Lesot, Coat-apc: When analogical proportion-based classification meets case-based reasoning, in: ATA@ ICCBR 2022, volume 3389 of *CEUR Workshop Proceedings*, Nancy, France, 2022, pp. 43–56.
[7] F. Badra, M.-J. Lesot, E. Marquer, M. Couceiro, Some Perspectives on Similarity Learning for Case-Based Reasoning and Analogical Transfer, in: IARML@IJCAI 2023, volume 3492, CEUR-WS.org, 2023, pp. 16–29.
[8] J. Lieber, E. Nauer, H. Prade, When revision-based case adaptation meets analogical extrapolation, in: ICCBR 2021, volume 12877 of *LNCS*, Springer, 2021, pp. 156–170.
[9] Y. Lepage, E. Denoual, Purest ever example-based machine translation: detailed presentation and assessment, Machine Translation 19 (2005) 251–282.
[10] M. Bounhas, H. Prade, G. Richard, Analogy-based classifiers for nominal or numerical data, IJAR 91 (2017) 36–55.

[11] M. Couceiro, E. Lehtonen, L. Miclet, H. Prade, G. Richard, When nominal analogical proportions do not fail, in: SUM 2020, volume 12322 of *LNCS*, Springer, 2020, pp. 68–83.
[12] M. Couceiro, N. Hug, H. Prade, G. Richard, Behavior of Analogical Inference w.r.t. Boolean Functions, in: IJCAI 2018, 2018, pp. 2057–2063.
[13] M. Couceiro, N. Hug, H. Prade, G. Richard, Analogy-preserving functions: A way to extend boolean samples, in: IJCAI 2017, 2017, pp. 1–7.
[14] M. Couceiro, E. Lehtonen, Galois theory for analogical classifiers, Ann. Math. Artif. Intell. 92 (2024) 29–47.
[15] M. A. Fahandar, E. Hüllermeier, Learning to rank based on analogical reasoning, in: AAAI-18, 2018, pp. 2951–2958.
[16] M. A. Fahandar, E. Hüllermeier, Analogical embedding for analogy-based learning to rank, in: IDA 2021, volume 12695 of *LNCS*, Springer, 2021, pp. 76–88.
[17] M. Mitchell, Abstraction and analogy-making in artificial intelligence, Annals of the New York Academy of Sciences 1505 (2021) 79–101.
[18] Y. Bitton, R. Yosef, E. Strugo, D. Shahaf, R. Schwartz, G. Stanovsky, VASR: visual analogies of situation recognition, in: (AAAI 2023) and (IAAI 2023) and (EAAI 2023), AAAI Press, Washington, DC, USA, 2023, pp. 241–249.
[19] J. Peyre, I. Laptev, C. Schmid, J. Sivic, Detecting unseen visual relations using analogies, in: IEEE ICCV 2019, 2019, pp. 1981–1990.
[20] F. Sadeghi, C. L. Zitnick, A. Farhadi, Visalogy: Answering visual analogy questions, in: NIPS 2015, 2015, pp. 1882–1890.
[21] G. Zervakis, E. Vincent, M. Couceiro, M. Schoenauer, E. Marquer, An analogy based approach for solving target sense verification, in: NLPIR 2022, ACM, 2022, pp. 144–151.
[22] A. K. Goel, Computational design, analogy, and creativity, in: Computational Creativity - The Philosophy and Engineering of Autonomously Creative Systems, Springer, 2019, pp. 141–158.
[23] A. Cornuéjols, P.-A. Murena, R. Olivier, Transfer learning by learning projections from target to source, in: IDA 2020, volume 12080 of *LNCS*, Springer, 2020, pp. 119–131.
[24] S. Alsaidi, A. Decker, P. Lay, E. Marquer, P.-A. Murena, M. Couceiro, On the transferability of neural models of morphological analogies, in: AIMLAI 2021, volume 1 of *PKDD/ECML Workshops*, Bilbao/Virtual, Spain, 2021, pp. 76–89.
[25] E. Hüllermeier, Towards analogy-based explanations in machine learning, in: MDAI 2020, volume 12256 of *LNCS*, Springer, 2020, pp. 205–217.
[26] M. T. Keane, B. Smyth, Good counterfactuals and where to find them: A case-based technique for generating counterfactuals for explainable AI (XAI), in: ICCBR 2020, volume 12311 of *LNCS*, Springer, 2020, pp. 163–178.
[27] A. Drozd, A. Gladkova, S. Matsuoka, Word embeddings, analogies, and machine learning: Beyond king - man + woman = queen, in: COLING 2016, 2016, pp. 3519–3530.
[28] T. Mikolov, W. Yih, G. Zweig, Linguistic regularities in continuous space word representations, in: HLT-NAACL 2013, 2013, pp. 746–751.
[29] T. Schnabel, I. Labutov, D. M. Mimno, T. Joachims, Evaluation methods for unsupervised word embeddings, in: EMNLP 2015, 2015, pp. 298–307.
[30] A. Ushio, L. E. Anke, S. Schockaert, J. Camacho-Collados, BERT is to NLP what alexnet is to CV: can pre-trained language models identify analogies?, in: ACL/IJCNLP 2021, volume 1,

2021, pp. 3609–3624.

[31] N. Zhang, L. Li, X. Chen, X. Liang, S. Deng, H. Chen, Multimodal analogical reasoning over knowledge graphs, in: ICLR 2023, 2023, pp. 1–20.

[32] L. Jarnac, M. Couceiro, P. Monnin, Relevant entity selection: Knowledge graph bootstrapping via zero-shot analogical pruning, in: CIKM 2023, ACM, 2023, pp. 934–944.

[33] T. Trouillon, J. Welbl, S. Riedel, É. Gaussier, G. Bouchard, Complex embeddings for simple link prediction, in: ICML 2016, volume 48 of *JMLR*, 2016, pp. 2071–2080.

[34] R. Xie, Z. Liu, H. Luan, M. Sun, Image-embodied knowledge representation learning, in: IJCAI 2017, ijcai.org, 2017, pp. 3140–3146.

[35] Z. Wang, L. Li, Q. Li, D. Zeng, Multimodal data enhanced representation learning for knowledge graphs, in: IJCNN 2019, IEEE, 2019, pp. 1–8.

[36] M. Wang, S. Wang, H. Yang, Z. Zhang, X. Chen, G. Qi, Is visual context really helpful for knowledge graph? A representation learning perspective, in: MM 2021, ACM, 2021, pp. 2735–2743.

[37] T. Linzen, Issues in evaluating semantic spaces using word analogies, in: RepEval@ACL 2016, 2016, pp. 13–18.

[38] N. Schluter, The word analogy testing caveat, in: NAACL-HLT 2018, Volume 2 (Short Papers), 2018, pp. 242–246.

[39] A. Gladkova, A. Drozd, S. Matsuoka, Analogy-based detection of morphological and semantic relations with word embeddings: what works and what doesn't, in: SRW@HLT-NAACL 2016, 2016, pp. 8–15.

[40] Z. Bouraoui, S. Jameel, S. Schockaert, Relation induction in word embeddings revisited, in: COLING 2018, Association for Computational Linguistics, 2018, pp. 1627–1637.

[41] C. Allen, T. M. Hospedales, Analogies explained: Towards understanding word embeddings, in: ICML 2019, volume 97 of *Proceedings of Machine Learning Research*, PMLR, 2019, pp. 223–231.

[42] M. R. Petersen, L. van der Plas, Can language models learn analogical reasoning? Investigating training objectives and comparisons to human performance, in: EMNLP 2023, 2023, pp. 16414–16425.

[43] Y. Lepage, M. Couceiro, Any four real numbers are on all fours with analogy, 2024 (under review).

[44] O. L. Hölder, Grenzwerthe von Reihen an der Convergenzgrenze, Math. Annalen 20 (1882) 535–549.

[45] M. Couceiro, Y. Lepage, Analogie et moyenne généralisée, in: Journées d'Intelligence Artificielle Fondamentale - Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes - JIAF-JFPDA 2024, La Rochelle, France, 2024, p. no page number. URL: https://inria.hal.science/hal-04568339.

[46] O. Levy, Y. Goldberg, Neural word embedding as implicit matrix factorization, in: Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, K. Weinberger (Eds.), Proceedings of Advances in Neural Information Processing Systems 27 (NIPS 2014), Curran Associates, Inc., 2014, pp. 2177–2185. URL: https://papers.nips.cc/paper_files/paper/2014.

[47] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, in: Y. Bengio, Y. LeCun (Eds.), ICLR 2013, 2013. URL: http://arxiv.org/abs/1301.3781.