

Combining Machine Learning and Reasoning: The Example of Analogical Proportions

Henri Prade

IRIT, CNRS & Université de Toulouse, France

work in collaboration, in particular with
Myriam Bounhas, Gilles Richard

Abstract

After a few general remarks on research in machine learning and reasoning, I will use a specific research topic-based on analogical proportions, as an example. An analogical proportion is a quaternary relation linking four items a, b, c, d and stating that " a is to b as c is to d ". Starting from the fundamental postulates these proportions should satisfy, we shall present their Boolean and multi-valued models. Next, we will examine various connections to machine learning: i) the use of analogical proportions in classification; ii) the use of proportions related to analogical proportions to propose a new approach to clustering; iii) the definition of analogical proportions between probability distributions and their potential use for extrapolating probability distributions. Finally, the section on reasoning will focus on the use of analogical proportions in creativity and on the constructive solving of Raven's tests.

On AI research

- **Deep learning**: produces remarkable results, but requires a **large number of examples** and consumes a **lot of energy** (due to the intensive use of **sum and product operations** on floating-point numbers)
- Need to **combine Learning and Reasoning**
- Need for more **qualitative** approaches

My two lines of research in qualitative modeling

- **Possibilistic** reasoning and learning

Ismail Baaj, Henri Prade: A Sugeno Integral View of Binarized Neural Network Inference. CoRR abs/2604.17967 (2026)

- **Analogical reasoning and learning based on analogical proportions**

- Ismail Baaj, Zied Bouraoui, Antoine Cornuéjols, Thierry Denoeux, Sébastien Destercke, Didier Dubois, Marie-Jeanne Lesot, João Marques-Silva, Jérôme Mengin, Henri Prade, Steven Schockaert, Mathieu Serrurier, Olivier Strauss, Christel Vrain: **Synergies between machine learning and reasoning** - An introduction by the Kay R. Amel group. Int. J. Approx. Reason. 171: 109206 (2024)

My two lines of research in qualitative modeling

- **Possibilistic** reasoning and learning

Ismail Baaj, Henri Prade: A Sugeno Integral View of Binarized Neural Network Inference. CoRR abs/2604.17967 (2026)

- Analogical reasoning and learning based on **analogical proportions**

- Ismail Baaj, Zied Bouraoui, Antoine Cornuéjols, Thierry Denoeux, Sébastien Destercke, Didier Dubois, Marie-Jeanne Lesot, João Marques-Silva, Jérôme Mengin, Henri Prade, Steven Schockaert, Mathieu Serrurier, Olivier Strauss, Christel Vrain: **Synergies between machine learning and reasoning** - An introduction by the Kay R. Amel group. Int. J. Approx. Reason. 171: 109206 (2024)

Contents

- Postulates and Models
 - Postulates for analogical proportions
 - Boolean model and its consequences
 - Multiple-valued analogical proportions
- Machine Learning
 - Classification
 - Clustering
 - Analogical proportions between probability distributions
- Reasoning
 - Creativity
 - Solving Raven tests

Postulates for analogical proportions

- Analogical proportion

“*a* is to *b* as *c* is to *d*”

denoted $a : b :: c : d$

whatever the nature of the items a, b, c and d

- **Postulates**

$a : b :: a : b$ (reflexivity)

$a : b :: c : d \Rightarrow c : d :: a : b$ (symmetry)

$a : b :: c : d \Rightarrow a : c :: b : d$ (central permutation)

- Perhaps **Aristotle** was inspired by the properties of
 - **arithmetic** proportion $a - b = c - d$
 - **geometric** proportion $\frac{a}{b} = \frac{c}{d}$

Boolean model

- $a : b :: c : d =$

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

“ a differs from b as c differs from d (and vice-versa)”

“if a does not differ from b , c does not differ from d (and vice-versa)”

$$0 : 0 :: 0 : 0$$

$$1 : 1 :: 1 : 1$$

$$0 : 1 :: 0 : 1$$

- $1 : 0 :: 1 : 0$

$$0 : 0 :: 1 : 1$$

$$1 : 1 :: 0 : 0$$

Consequences of the Boolean model

- *code independence*

$$(a : b :: c : d) \Rightarrow \neg a : \neg b :: \neg c : \neg d$$

- *transitivity*:

$$(a : b :: c : d) \wedge (c : d :: e : f) \Rightarrow a : b :: e : f$$

$\Rightarrow$ analogical proportions are thus
equivalence relations between pairs (a, b) and (c, d)

- *central transitivity*:

$$(a : b :: c : d) \wedge (a : b' :: c' : d) \Rightarrow b' : b :: c : c'$$

- *unicity of the solution, when it exists*:

$$a : b :: c : d, a : b :: c : x \Rightarrow x = d$$

Multi-attribute analogical proportions

$$\vec{a} : \vec{b} :: \vec{c} : \vec{d} =_{def} \forall i \in [1, n], a_i : b_i :: c_i : d_i$$

	<i>mammal</i>	<i>bovine</i>	<i>equine</i>	<i>adult</i>	<i>young</i>
cow	1	1	0	1	0
calf	1	1	0	0	1
mare	1	0	1	1	0
foal	1	0	1	0	1

“the cow is to the calf as the mare is to the foal”

4 items a, b, c, d are distinct only if there are

at least **one attribute** with pattern of the form $s : t :: s : t$
 and **one attribute** with a pattern of the form $s : s :: t : t$
 with $s \neq t$

Vectorial / parallelogram view (see Ref. 9)

	sexM	sexF	powerposition	ordinaryposition	human	god
Man(M)	1	0	0	1	1	0
King(K)	1	0	1	0	1	0
Woman(W)	0	1	0	1	1	0
Queen(Q)	0	1	1	0	1	0

Table: The man/king/woman/queen example

- $\overrightarrow{MQ} = \overrightarrow{MK} + \overrightarrow{MW}$
- $\overrightarrow{Q} = \overrightarrow{M} + \overrightarrow{MK} + \overrightarrow{MW}$
- vectors with component values $-1, 0$, or 1
- the **parallelogram** view
(Rumelhart & Abrahamson, 1973) !
- could apply to word-to-vec embeddings

Graded analogical proportion - 1 (see Ref. 6)

- Numerical attributes, values *normalized* in $[0, 1]$
- graded extensions of analogical proportions
- 2 options which make sense

Option 1

$$a : b ::_L c : d = \begin{cases} 1 - |(a - b) - (c - d)|, & \text{if } a \geq b \text{ and } c \geq d, \text{ or } a \leq b \text{ and } c \leq d \\ 1 - \max(|a - b|, |c - d|), & \text{if } a \leq b \text{ and } c \geq d, \text{ or } a \geq b \text{ and } c \leq d \end{cases}$$

- Coincides with $a : b :: c : d$ on $\{0, 1\}$
- Equal to 1 iff $(a - b) = (c - d)$
- $a : b ::_L c : d = 0$ when the change in (a, b) or in (c, d) is *maximal*, while the other pair exhibits no change, or a change in the *opposite* direction

Graded analogical proportion - 2 (see Ref. 6)

Second option:

- $a : b ::_C c : d =$

$$\min(1 - |\max(a, d) - \max(b, c)|, 1 - |\min(a, d) - \min(b, c)|)$$

- Coincides with $a : b :: c : d$ on $\{0, 1\}$

- $a : b ::_C c : d = 1$

$$\Leftrightarrow \min(a, d) = \min(b, c) \text{ and } \max(a, d) = \max(b, c)$$

Only the valuations (s, s, t, t) , (s, t, s, t) (s, s, s, s)
make the analogical proportion true at degree 1 !!

- $a : b ::_{\mathbf{C}} c : d = 1 \Rightarrow a : b ::_{\mathbf{L}} c : d = 1$

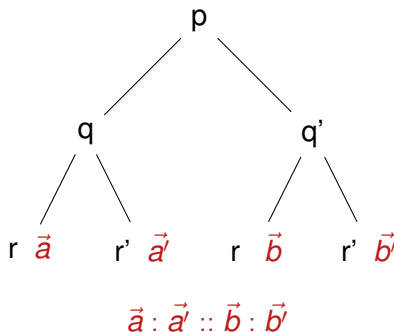
- for instance, $0 : 0.5 ::_{\mathbf{L}} 0.5 : 1 = 1$,
while $0 : 0.5 ::_{\mathbf{C}} 0.5 : 1 = 0.5$

Analogical proportions and taxonomies

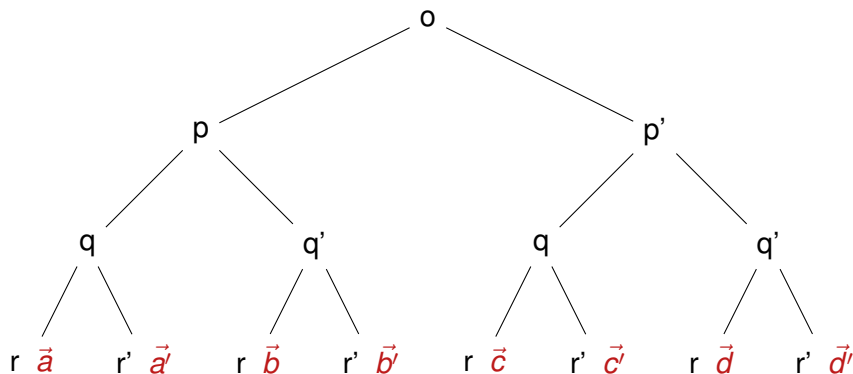
- When using **pairs de mutually exclusive properties**
 $(p \wedge p' = \perp, q \wedge q' = \perp, r \wedge r' = \perp)$,
 and building the 4 items $\vec{a}, \vec{a}', \vec{b}, \vec{b}'$
 respectively described with the 6 properties (p, q, r, r', q', p') as

	p	q	r	r'	q'	p'
$\vec{a}$	1	1	1	0	0	0
$\vec{a}'$	1	1	0	1	0	0
$\vec{b}$	1	0	1	0	1	0
$\vec{b}'$	1	0	0	1	1	0

We obtain the taxonomic tree



Introducing one more level in the tree



Example

Table 1 $p(\text{animal}) p(\text{canid}) q(\text{domest.}) r(\text{young}) r'(\text{adult}) q'(\text{wild}) p'(\text{suidae}) o'(\text{plac.})$

$\vec{a}$ chiot (puppy)	1	1	1	1	0	0	0	0
$\vec{a}'$ chien (dog)	1	1	1	0	1	0	0	0
$\vec{b}$ louveteau (cub)	1	1	0	1	0	1	0	0
$\vec{b}'$ loup (wolf)	1	1	0	0	1	1	0	0
$\vec{c}$ porcelet (piglet)	1	0	1	1	0	0	1	0
$\vec{c}'$ porc (pig)	1	0	1	0	1	0	1	0
$\vec{d}$ marcassin	1	0	0	1	0	1	1	0
$\vec{d}'$ sanglier (boar)	1	0	0	0	1	1	1	0

12 analogical proportions !:

$$\begin{aligned}
 &(\vec{a} : \vec{a}' :: \vec{b} : \vec{b}'), (\vec{a} : \vec{a}' :: \vec{c} : \vec{c}'), (\vec{a} : \vec{b} :: \vec{c} : \vec{d}), (\vec{a}' : \vec{b}' :: \vec{c}' : \vec{d}'), \\
 &(\vec{c} : \vec{c}' :: \vec{d} : \vec{d}'), (\vec{b} : \vec{b}' :: \vec{d} : \vec{d}') (\vec{b} : \vec{b}' :: \vec{c} : \vec{c}'), (\vec{a} : \vec{a}' :: \vec{d} : \vec{d}'), \\
 &(\vec{a} : \vec{b}' :: \vec{c} : \vec{d}'), (\vec{a}' : \vec{b} :: \vec{c}' : \vec{d}), (\vec{a} : \vec{b} :: \vec{c}' : \vec{d}'), (\vec{a}' : \vec{b}' :: \vec{c} : \vec{d})
 \end{aligned}$$

Analogical equation

- The equation $a : b :: c : x$ may have no solution in $\mathbb{B} = \{0, 1\}$
Neither $0 : 1 :: 1 : x$, nor $1 : 0 :: 0 : x$ have a solution
- when it exists (iff $(a \equiv b) \vee (a \equiv c)$ holds), it is unique
- $x = c \equiv (a \equiv b)$
- applies to Boolean vectors: find $\vec{x} = (x_1, \dots, x_n)$ s.t. $\vec{a} : \vec{b} :: \vec{c} : \vec{x}$ holds :
 $\Rightarrow n$ equations $a_i : b_i :: c_i : x_i$

General analogical inference



$$\frac{\forall i \in \{1, \dots, p\}, \quad a_i : b_i :: c_i : d_i \text{ holds}}{\forall j \in \{p + 1, \dots, n\}, \quad a_j : b_j :: c_j : d_j \text{ holds}}$$

(Stroppa, Yvon, 2005)

- Analogical reasoning amounts to find fully informed **triplets** $(\vec{a}, \vec{b}, \vec{c})$ suitable for inferering the missing values of a **not fully informed** item $(\vec{d})$
- When *several triplets* lead to distinct conclusions a *voting* procedure can be used
- Can be extended to **numerical** analogical proportions

Classification: direct application of inference principle

- Finding the class $cl(\vec{x})$ of $\vec{x}$ ($cl(\vec{x})$ nominal attribute)
- *Successfully applied* to multi-class problems, for Boolean, nominal, and / or numerical attributes

- analogical classifiers *always yield*

exact predictions when the classification is governed by an *affine Boolean function* (this includes *X-or* functions) and *only in this case*

but good results are obtained in other cases !

- *analogical proportions emphasize linearity*
- not all triplets are equally good !!

- choose *competent* pairs for building triplets (Ref. 2)

- prefer triplets that are equal on some attributes (Ref.3)

Classification: direct application of inference principle

- Finding the class $cl(\vec{x})$ of $\vec{x}$ ($cl(\vec{x})$ nominal attribute)
- *Successfully applied* to multi-class problems, for Boolean, nominal, and / or numerical attributes
- analogical classifiers **always yield**

exact predictions when the classification is governed by an **affine Boolean function** (this includes **X-or** functions) and *only in this case*

but good results are obtained in other cases !

- **analogical proportions emphasize linearity**
- not all triplets are equally good !!

- choose *competent* pairs for building triplets (Ref. 2)

- prefer triplets that are equal on some attributes (Ref.3)

Classification: direct application of inference principle

- Finding the class $cl(\vec{x})$ of $\vec{x}$ ($cl(\vec{x})$ nominal attribute)
- *Successfully applied* to multi-class problems, for Boolean, nominal, and / or numerical attributes
- analogical classifiers **always yield**

exact predictions when the classification is governed by an **affine Boolean function** (this includes **X-or** functions) and *only in this case*

but good results are obtained in other cases !

- **analogical proportions emphasize linearity**
 - not all triplets are equally good !!
- choose *competent* pairs for building triplets (Ref. 2)
 - prefer triplets that are equal on some attributes (Ref.3)

2 proportions associated to the analogical proportion

- *reverse analogy*: $\text{Rev}(a, b, c, d) \triangleq$

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

“*b is to a as c is to d*”

- *paralogy*: $\text{Par}(a, b, c, d) \triangleq$

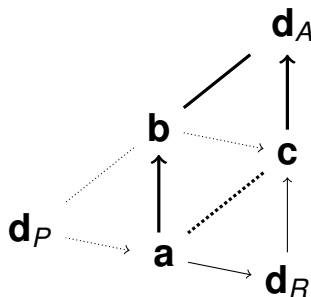
$$((a \wedge b) \equiv (c \wedge d)) \wedge ((\neg a \wedge \neg b) \equiv (\neg c \wedge \neg d))$$

what *a* and *b* have in common (positively or negatively), *c* and *d* have it also, and conversely

$$\text{Rev}(b, a, c, d) \Leftrightarrow \text{Ana}(a, b, c, d) \Leftrightarrow \text{Par}(c, b, a, d))$$

	Ana				Rev				Par			
•	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	0	1	1	0	0	1	1	1	0	0	1
	1	1	0	0	1	1	0	0	0	1	1	0
	0	1	0	1	0	1	1	0	0	1	0	1
	1	0	1	0	1	0	0	1	1	0	1	0

A geometric illustration



3 parallelograms

Inverse paralogy

- One exchanges the *positive* and *negative* similarity indicators of the pair (c, d) in the definition of the paralogy:

$$\text{Inv}(a, b, c, d) \triangleq$$

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

- “what a and b have in common, c and d have not it and conversely”: a form of “*orthogonality*”
- $\text{Ana}(a, b, c, d) \leftrightarrow \text{Inv}(a, \neg b, \neg c, d)$.
- Unique proportion *stable under all the permutations* of two terms:

$$\text{Inv}(a, b, c, d) \Leftrightarrow \text{Inv}(b, a, c, d)$$

$$\Leftrightarrow \text{Inv}(a, c, b, d) \Leftrightarrow \text{Inv}(c, b, a, d)$$

Analogy, Reverse Analogy, Paralogy, Inverse Paralogy

A			
0	0	0	0
1	1	1	1
0	0	1	1
1	1	0	0
0	1	0	1
1	0	1	0

R			
0	0	0	0
1	1	1	1
0	0	1	1
1	1	0	0
0	1	1	0
1	0	0	1

P			
0	0	0	0
1	1	1	1
0	1	1	0
1	0	0	1
0	1	0	1
1	0	1	0

I			
0	0	1	1
1	1	0	0
0	1	1	0
1	0	0	1
0	1	0	1
1	0	1	0

The cores of the homogeneous logical proportions

$$\underline{\underline{A}}(a, b, c, d) = [(\neg a \wedge b) \oplus (c \wedge \neg d)] \wedge [(a \wedge \neg b) \oplus (\neg c \wedge d)]$$

$$\underline{\underline{R}}(a, b, c, d) = [(\neg a \wedge b) \oplus (\neg c \wedge d)] \wedge [(a \wedge \neg b) \oplus (c \wedge \neg d)]$$

$$\underline{\underline{P}}(a, b, c, d) = [(a \wedge b) \oplus (\neg c \wedge \neg d)] \wedge [(\neg a \wedge \neg b) \oplus (c \wedge d)]$$

$$\begin{aligned} \underline{\underline{I}}(a, b, c, d) &= [(a \wedge b) \oplus (c \wedge d)] \wedge [(\neg a \wedge \neg b) \oplus (\neg c \wedge \neg d)] \\ &= [(a \wedge b) \oplus (c \wedge d)] \wedge [(a \vee b) \oplus (c \vee d)] \end{aligned}$$

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

Link between Inverse Paralogy/Paralogy and Clustering Ref.4

The notion de clusters relies on two basic principles:

- ① A **strong similarity** or a **proximity** between the objects belonging to the same cluster ;
- ② A **marked difference** or a **significant distance** from objects belonging to **other clusters**.

$$\underline{\underline{I}}(a, b, c, d) = [(a \wedge b) \oplus (c \wedge d)] \wedge [(a \vee b) \oplus (c \vee d)]$$

- The patterns (0, 0, 1, 1) and (1, 1, 0, 0) that make $\underline{\underline{I}}(a, b, c, d)$ true correspond to two pairs of objects belonging to **two clusters**
- $\underline{\underline{I}}$ measures a kind of **orthogonality** between the pairs (a, b) & (c, d)

$$\underline{\underline{P}}(a, b, c, d) = [(a \wedge b) \oplus (\neg c \wedge \neg d)] \wedge [(\neg a \wedge \neg b) \oplus (c \wedge d)]$$

- The configurations (0, 0, 0, 0), (1, 1, 1, 1) that make $\underline{\underline{P}}(a, b, c, d)$ true reflect the belonging of the pairs (a, b) and (c, d) to **the same cluster**
- $\underline{\underline{P}}$ rather measures the **coherence** between the pairs (a, b) and (c, d)

Link between Inverse Paralogy/Paralogy and Clustering Ref.4

The notion de clusters relies on two basic principles:

- ① A **strong similarity** or a **proximity** between the objects belonging to the same cluster ;
- ② A **marked difference** or a **significant distance** from objects belonging to **other clusters**.

$$\underline{\underline{I}}(a, b, c, d) = [(a \wedge b) \oplus (c \wedge d)] \wedge [(a \vee b) \oplus (c \vee d)]$$

- The patterns (0, 0, 1, 1) and (1, 1, 0, 0) that make $\underline{\underline{I}}(a, b, c, d)$ true correspond to two pairs of objects belonging to **two clusters**
- $\underline{\underline{I}}$ measures a kind of **orthogonality** between the pairs (a, b) & (c, d)

$$\underline{\underline{P}}(a, b, c, d) = [(a \wedge b) \oplus (\neg c \wedge \neg d)] \wedge [(\neg a \wedge \neg b) \oplus (c \wedge d)]$$

- The configurations (0, 0, 0, 0), (1, 1, 1, 1) that make $\underline{\underline{P}}(a, b, c, d)$ true reflect the belonging of the pairs (a, b) and (c, d) to **the same cluster**
- $\underline{\underline{P}}$ rather measures the **coherence** between the pairs (a, b) and (c, d)

Link between Inverse Paralogy/Paralogy and Clustering Ref.4

The notion de clusters relies on two basic principles:

- ① A **strong similarity** or a **proximity** between the objects belonging to the same cluster ;
- ② A **marked difference** or a **significant distance** from objects belonging to **other clusters**.

$$\underline{\underline{I}}(a, b, c, d) = [(a \wedge b) \oplus (c \wedge d)] \wedge [(a \vee b) \oplus (c \vee d)]$$

- The patterns (0, 0, 1, 1) and (1, 1, 0, 0) that make $\underline{\underline{I}}(a, b, c, d)$ true correspond to two pairs of objects belonging to **two clusters**
- $\underline{\underline{I}}$ measures a kind of **orthogonality** between the pairs (a, b) & (c, d)

$$\underline{\underline{P}}(a, b, c, d) = [(a \wedge b) \oplus (\neg c \wedge \neg d)] \wedge [(\neg a \wedge \neg b) \oplus (c \wedge d)]$$

- The configurations (0, 0, 0, 0), (1, 1, 1, 1) that make $\underline{\underline{P}}(a, b, c, d)$ true reflect the belonging of the pairs (a, b) and (c, d) to **the same cluster**
- $\underline{\underline{P}}$ rather measures the **coherence** between the pairs (a, b) and (c, d)

Extension to numerical data

1) Numerical values: $(x, y, z, t) \in [0, 1]^4$, $\underline{I}(x, y, z, t)$ and $\underline{P}(x, y, z, t) \in [0, 1]$

$$\underline{I}(x, y, z, t) = \min(|\min(x, y) - \min(z, t)|, |\max(x, y) - \max(z, t)|)$$

- $\underline{I}(x, y, z, t) = 1 \Rightarrow x=y=0$ & $z=t=1$, or $x=y=1$ & $z=t=0$
- If $x = y$ and $z = t$, $\underline{I}(x, y, z, t) = |x - z|$, distance between the pairs

$$\underline{P}(x, y, z, t) = 1 - \max(|x - y|, |z - t|, |x - z|, |y - t|)$$

- $\underline{P}(x, y, z, t) = 1$ as soon as x, y, z and t are equal.
- $\underline{P}(x, y, z, t)$ decreases as soon as one value deviates from the others.

2) Pairs of vectors (x, y) and (z, t) :

$$\mathcal{O} = \frac{1}{n} \sum_{i=1}^n \underline{I}(x_i, y_i, z_i, t_i)$$

Orthogonality

$$\mathcal{P} = \frac{1}{n} \sum_{i=1}^n \underline{P}(x_i, y_i, z_i, t_i)$$

Relational coherence

Extension to numerical data

1) Numerical values: $(x, y, z, t) \in [0, 1]^4$, $\underline{I}(x, y, z, t)$ and $\underline{P}(x, y, z, t) \in [0, 1]$

$$\underline{I}(x, y, z, t) = \min(|\min(x, y) - \min(z, t)|, |\max(x, y) - \max(z, t)|)$$

- $\underline{I}(x, y, z, t) = 1 \Rightarrow x=y=0$ & $z=t=1$, or $x=y=1$ & $z=t=0$
- If $x = y$ and $z = t$, $\underline{I}(x, y, z, t) = |x - z|$, distance between the pairs

$$\underline{P}(x, y, z, t) = 1 - \max(|x - y|, |z - t|, |x - z|, |y - t|)$$

- $\underline{P}(x, y, z, t) = 1$ as soon as x, y, z and t are **equal**.
- $\underline{P}(x, y, z, t)$ **decreases** as soon as one value deviates from the others.

2) Pairs of vectors (x, y) and (z, t) :

$$\mathcal{O} = \frac{1}{n} \sum_{i=1}^n \underline{I}(x_i, y_i, z_i, t_i)$$

Orthogonality

$$\mathcal{P} = \frac{1}{n} \sum_{i=1}^n \underline{P}(x_i, y_i, z_i, t_i)$$

Relational coherence

Extension to numerical data

1) Numerical values: $(x, y, z, t) \in [0, 1]^4$, $\underline{I}(x, y, z, t)$ and $\underline{P}(x, y, z, t) \in [0, 1]$

$$\underline{I}(x, y, z, t) = \min(|\min(x, y) - \min(z, t)|, |\max(x, y) - \max(z, t)|)$$

- $\underline{I}(x, y, z, t) = 1 \Rightarrow x=y=0$ & $z=t=1$, or $x=y=1$ & $z=t=0$
- If $x = y$ and $z = t$, $\underline{I}(x, y, z, t) = |x - z|$, distance between the pairs

$$\underline{P}(x, y, z, t) = 1 - \max(|x - y|, |z - t|, |x - z|, |y - t|)$$

- $\underline{P}(x, y, z, t) = 1$ as soon as x, y, z and t are **equal**.
- $\underline{P}(x, y, z, t)$ **decreases** as soon as one value deviates from the others.

2) Pairs of vectors (x, y) and (z, t) :

$$\mathcal{O} = \frac{1}{n} \sum_{i=1}^n \underline{I}(x_i, y_i, z_i, t_i)$$

Orthogonality

$$\mathcal{P} = \frac{1}{n} \sum_{i=1}^n \underline{P}(x_i, y_i, z_i, t_i)$$

Relational coherence

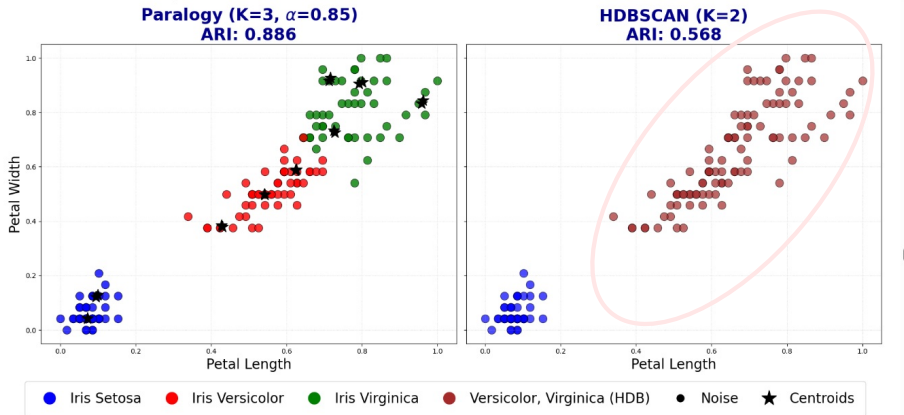
Ideas at work in the clustering algorithm

- The algorithm uses **pairs of points, rather than individual points**, to assign the data to groups as they form
- Idea: iteratively generate and refine **mini-clusters**
 - i) by identifying "orthogonal" pairs of points based on the orthogonality measure,
 - ii) by assigning these pairs to preliminary mini-clusters, and finally
 - iii) by merging mini-clusters whose relational coherence between pairs is high enough
- **No need to know the number of clusters**
- **Competes with k-means and HDBSCAN**
- **HDBSCAN**: may introduce noise (points left outside clusters)

Clustering procedure based on $\mathcal{O}$ & $\mathcal{P}$

- 1 Data normalization.
- 2 Construction of the nearest-neighbor pairs $\mathbb{P}$.
- 3 Calculation of the matrices $\mathcal{P}$ and $\mathcal{O}$.
- 4 Creation of mini-clusters using the orthogonality $\mathcal{O}$.
- 5 Assignment of pairs to mini-clusters.
- 6 Calculation of centroids μ .
- 7 Merging of mini-clusters guided by paralogy ($\mathcal{P}$).

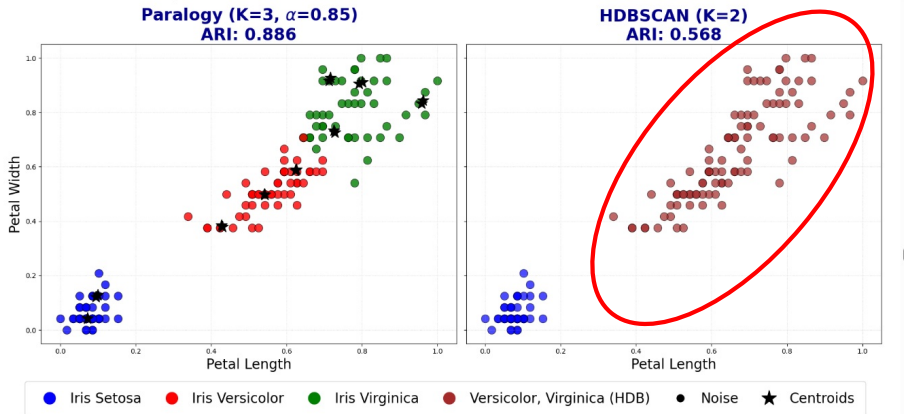
Results : Iris



Paralogy: captures the allometric relationships between petal length and width.

HDBSCAN: Treats Versicolor and Virginica as a single dense cluster

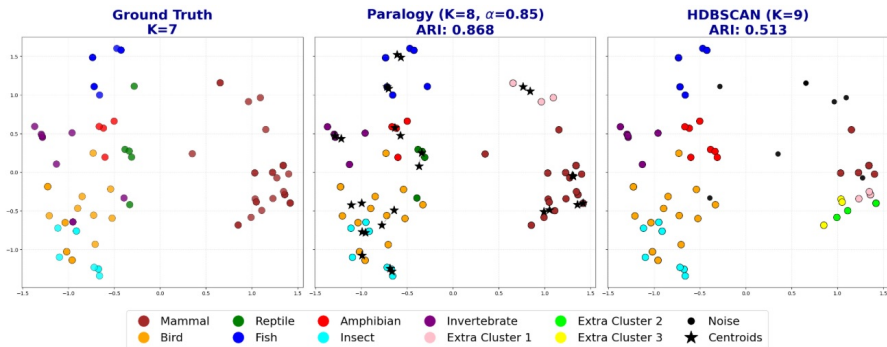
Results : Iris



Paralogy: captures the allometric relationships between petal length and width.

HDBSCAN: Treats Versicolor and Virginica as a **single dense cluster**

Results : Zoo



Paralogy: Correctly preserves biological classes and atypical cases such as the platypus

HDBSCAN: a) Fragments mammals, b) creates arbitrary clusters, c) unnecessarily merges amphibians and reptiles, and d) introduces noise

Analogical proportion on probability distributions (Ref. 10)

- a, b, c, d are probability distributions on

$$X = \{x_1, \dots, x_n\} \quad a_i = a(x_i), \quad b_i = b(x_i), \quad c_i = c(x_i), \quad d_i = d(x_i)$$

$$\sum_{i=1,n} a_i = 1, \quad \sum_{i=1,n} b_i = 1, \quad \sum_{i=1,n} c_i = 1, \quad \sum_{i=1,n} d_i = 1$$

Definition

These distributions are in *arithmetic* analogical proportion, $a : b ::_{ari} c : d$

if for all i we have $a_i - b_i = c_i - d_i$

Proposition

Let a, b, c be three **normalized** probability distributions, if $a : b ::_{ari} c : d$ and if

$0 \leq c_i + b_i - a_i \leq 1, \forall i$, then d is **normalized**

Arithmetic analogical proportion. Example

- $a_1 = 0.1; a_2 = 0.2; a_3 = 0.4; a_4 = 0.3.$
 $b_1 = 0.3; b_2 = 0.3; b_3 = 0.2; b_4 = 0.2.$
 $c_1 = 0.2; c_2 = 0.2; c_3 = 0.2; c_4 = 0.4.$
 $d_1 = 0.4; d_2 = 0.3; d_3 = 0; d_4 = 0.3.$
- given 2 distributions a and b , it is **always possible to find** a distribution c s.t. $\forall i, 0 \leq c_i + b_i - a_i \leq 1$

Arithmetico-geometric analogical proportion

Definition

Let a, b, c, d be four probability distributions over the same domain. These distributions are in *arithmetico-geometric* analogical proportion, denoted $a : b ::_{arigeo} c : d$ if $\forall i, a_i - b_i = c_i - d_i$ and $\frac{a_i}{b_i} = \frac{c_i}{d_i}$.

- **Proposition** There exist probability distributions a, b, c, d such that $a : b ::_{arigeo} c : d$ holds. They are such that for each component i ,
 - cas 1: either $a_i = b_i$ and $c_i = d_i$,
 - cas 2: or $a_i = c_i$ and $b_i = d_i$and these are the only solutions.

Arithmetico-geometric analogical proportion

Definition

Let a, b, c, d be four probability distributions over the same domain. These distributions are in *arithmetico-geometric* analogical proportion, denoted $a : b ::_{arigeo} c : d$ if $\forall i, a_i - b_i = c_i - d_i$ and $\frac{a_i}{b_i} = \frac{c_i}{d_i}$.

- **Proposition** There **exist** probability distributions a, b, c, d such that $a : b ::_{arigeo} c : d$ holds. They are such that for each component i ,
 - cas 1: either $a_i = b_i$ and $c_i = d_i$,
 - cas 2: or $a_i = c_i$ and $b_i = d_i$and these are the only solutions.

Example

- $a_1 = 0.1, a_2 = 0.3, a_3 = 0.2, a_4 = 0.3, a_5 = 0.1$

$$b_1 = 0.1, b_2 = 0.4, b_3 = 0.2, b_4 = 0.2, b_5 = 0.1$$

$$c_1 = 0.1, c_2 = 0.3, c_3 = 0.3, c_4 = 0.3, c_5 = 0$$

$$d_1 = 0.1, d_2 = 0.4, d_3 = 0.3, d_4 = 0.2, d_5 = 0$$

Proposition

Given 2 different probability distributions a and b , there *always exist* two probability distributions c and d such that $a : b ::_{arigeo} c : d$ holds.

If there exists at least one i such that $a_i = b_i$, the four distributions are different

Example

- $a_1 = 0.1, a_2 = 0.3, a_3 = 0.2, a_4 = 0.3, a_5 = 0.1$

$$b_1 = 0.1, b_2 = 0.4, b_3 = 0.2, b_4 = 0.2, b_5 = 0.1$$

$$c_1 = 0.1, c_2 = 0.3, c_3 = 0.3, c_4 = 0.3, c_5 = 0$$

$$d_1 = 0.1, d_2 = 0.4, d_3 = 0.3, d_4 = 0.2, d_5 = 0$$

Proposition

Given 2 different probability distributions a and b , there **always exist** two probability distributions c and d such that $a : b ::_{\text{arigeo}} c : d$ holds.

If there exists at least one i such that $a_i = b_i$, the four distributions are **different**

A nice property

Proposition

Let a, b, c, d be four probability distributions forming an arithmetico-geometric analogical proportion $a : b ::_{arigeo} c : d$. Then we have:

$$KL(a, b) = KL(c, d)$$

where KL is the Kullback-Leibler divergence.

- $KL(a, b) = \sum_{i=1, n} a_i \log \frac{a_i}{b_i}$
- **Property** If $a : b ::_{arigeo} c : d$ then
 - i) $KL(a, c) = KL(b, d)$; ii) $KL(b, a) = KL(d, c)$;
 - iii) $KL(d, b) = KL(c, a)$.

Another nice property

Proposition

Let a, b, c, d be four probability distributions forming an arithmetic analogical proportion $a : b ::_{\text{ari}} c : d$.

Then we have:

$$TV(a, b) = TV(c, d)$$

where TV is the *total variation distance*.

$$TV(a, b) = \frac{1}{2} \sum_{i=1}^n |a_i - b_i|$$

Predicting probability distributions (see Ref. 11)

- **Experimental evidence** (Movielens, S Traffic Accidents) when an analogical proportion holds between four distinct profiles in a population, a corresponding proportion, arithmetic (or even combining arithmetic and geometric proportions), often holds between associated probability distributions

$$\frac{\vec{a} : \vec{b} :: \vec{c} : \vec{d}}{p_{\vec{a}} : p_{\vec{b}} ::_{ari} p_{\vec{c}} : p_{\vec{d}}}$$

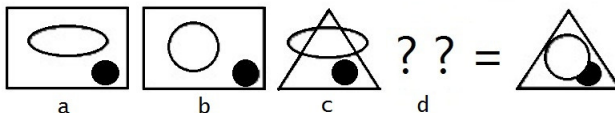
Creativity (see Ref. 1)



$$\frac{\vec{a}, \vec{b}, \vec{c}}{\vec{d} = (\vec{a} \equiv \vec{b} \equiv \vec{c})}$$

- $\vec{d}$ is **different** from $\vec{a}, \vec{b}, \vec{c}$

Puzzle



Th. Evans (1963)

Example encoded with 5 Boolean predicates

Rectangle(R), *Black dot*(BD), *Triangle*(T)

Circle(C), *Ellipse*(E) (in this order)

	R	BD	T	C	E
<i>a</i>	1	1	0	0	1
<i>b</i>	1	1	0	1	0
<i>c</i>	0	1	1	0	1
<i>x</i>	0	1	1	1	0

Guided creativity

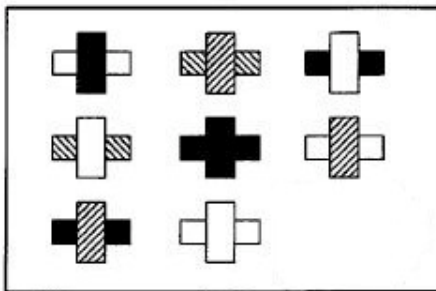
- a (small) set $\mathcal{P}$ of pairs $(\vec{a}_i, \vec{b}_i)$ such that $\vec{a}_i$ is regarded as "better" than $\vec{b}_i$
- enlarge $\mathcal{P}$ using an operation of a logic of pairs

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

- The operation $\wedge \vee$ has the merit of “cumulating” the acquisition of features by preserving patterns $(0, 1)$.

Raven tests : the solution is built, not chosen (Ref. 5, 1)

The IQ test: We are dealing with 3×3 matrices with **8 cells** containing drawings; we must work out what the correct content of the **9th cell** is, *from among 8 suggested solutions*



(a,b):f(a,b)::(c,d):f(c,d)

$(pi[1, 1], pi[1, 2]) : pi[1, 3] :: (pi[2, 1], pi[2, 2]) : pi[2, 3]) ::$

$(pi[3, 1], pi[3, 2]) : pi[3, 3])$

$(pi[1, 1], pi[2, 1]) : pi[3, 1] :: (pi[1, 2], pi[2, 2]) : pi[3, 2]) ::$

$(pi[1, 3], pi[2, 3]) : pi[3, 3])$

	1	2	3
1	WB	GG	BW
2	GW	BB	WG
3	BG	WW	?i?ii

?i?ii = **GB**

- for the horizontal bars:

(W,G) : B :: (G, B) : W (horizontal analysis)

(W,G) : B :: (W,W) : ?i (horizontal analysis)

(W,G) : B :: (G, B) : W (vertical analysis)

(W,G) : B :: (B,W) : ?i (vertical analysis)

- for the vertical bars:

(B,G) : W :: (W, B) : G

(B,G) : W :: (G,W) : ?ii

(B,W) : G :: (G, B) : W

(B,W) : G :: (W,G) : ?ii

Conclusion

- power of analogical proportions
- importance of **comparative reasoning**
- variety of applications
- to be developed
weak analogical proportions
(when central permutation does not hold)

Not covered by the talk

- **logical proportions** (see Ref. 8), composition of analogical proportions, pair comparators, logic of pairs
- analogical proportions are closely related to (weak) *multi-valued dependencies* in DB
- analogical proportions and Formal Concept Analysis (see Barbot, Miclet, Prade, AIJ, 2019)
"Carlsen is to chess what Mozart is to music"
- analogical proportions in *preference learning*
- analogical proportions and Kolmogoroff complexity
- rule base completion
- application to explanation, recommendation
- analogical querying to a database

Main references - 1

1. S. D. Afantenos, H. Prade, G. Richard, L. Cortez Bernardes: Analogical proportions and creativity: A preliminary study. ICCC 2024:289-298
2. M. Bounhas, H. Prade: Analogy-based classifiers: An improved algorithm exploiting competent data pairs. Int. J. Approx. Reason. 158: 108923 (2023)
3. M. Bounhas, H. Prade: Revisiting analogical proportions and analogical inference. Int. J. Approx. Reason. 171: 109202 (2024)
4. M. Bounhas, H. Prade: From logical proportions to a new clustering algorithm. SUM 2026, Athens, Oct. 21-23, LNAI, Springer, to appear
5. W. Correa Beltran, H. Prade, G. Richard: Constructive solving of Raven's IQ tests with analogical proportions. Int. J. Intell. Syst. 31(11): 1072-1103 (2016)

Main references - 2

6. D. Dubois, H. Prade, G. Richard: Multiple-valued extensions of analogical proportions. *Fuzzy Sets Syst.* 292: 193-202 (2016)
7. S. Lim, H. Prade, G. Richard: Classifying and completing word analogies by machine learning. *Int. J. Approx. Reason.* 132: 1-25 (2021)
8. H. Prade, G. Richard: From analogical proportion to logical proportions. *Logica Universalis* 7(4): 441-505 (2013)
9. H. Prade, G. Richard: Multiple analogical proportions. *AI Commun.* 34(3): 211-228 (2021)
10. H. Prade, G. Richard: Analogical proportions between probabilities. *ECSQARU* 2025: 148-163
11. H. Prade, G. Richard: From numerical proportions to analogical proportions between probabilities. *CoRR* abs/2606.23029 (2026)