A gentle introduction to GENEOs and their use for XAI

Patrizio Frosini

Department of Computer Science, University of Pisa
patrizio.frosini@unipi.it

October 23, 2024

Outline

Let us start with a puzzle

Some epistemological assumptions

Some basics on the theory of GENEOs

GENEOs and XAI

Let us start with a puzzle

Some epistemological assumptions

Some basics on the theory of GENEOs

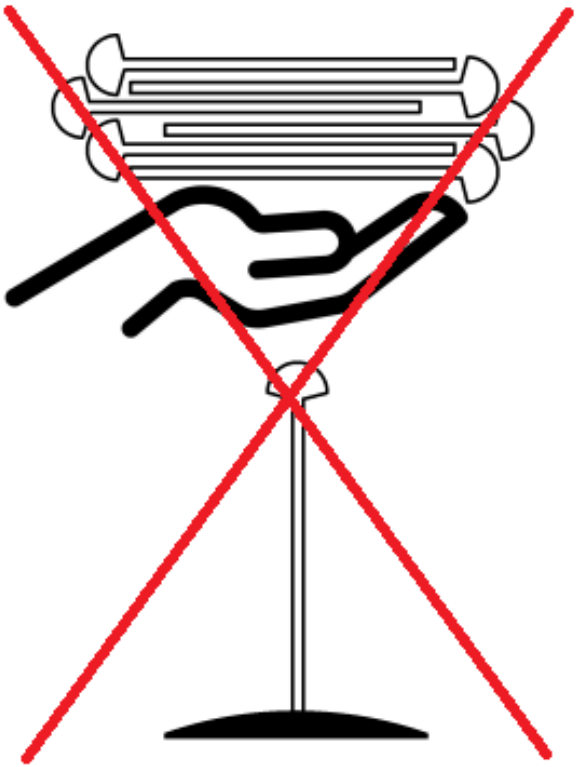
GENEOs and XAI

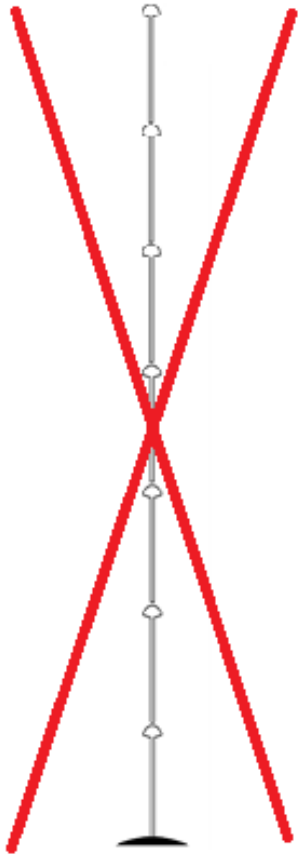
Let us start with a puzzle

How can we balance six peripheral poles on a central one without using glue or any tricks?

Let us start with a puzzle

No tricks or unlikely solutions allowed!





Let us start with a puzzle

The solution may not be trivial, but some ideas might help us find it:

- Balance is often linked to symmetry. Perhaps we should consider symmetrical arrangements.
- Since we can't use glue, we might prioritize positions where the poles intersect, as the intersections create friction and, in turn, stability.
- We should aim for structures with a low center of gravity. To achieve this, it's better to think of designs where the poles face downward.

Something like that? Yes, but unfortunately this structure is not a solution.

Let us start with a puzzle

The configuration space of this type is much smaller than the space of all possible positions. By exploring this space and creating different configurations, we can eventually find one that works.

What do we learn from the solution to this puzzle?

The puzzle we just examined teaches us that it is important to select the space of ideas in which it is appropriate to operate. In practice, this means learning to focus only on the information relevant to our problem, without even considering the rest.

How can we mathematically formalize the concept of an observer focusing on relevant information?

I will demonstrate a way to do this, which comes from Topological Data Analysis. It is based on the theory of Group Equivariant Non-Expansive Operators (GENEOs).

To begin, we will lay the groundwork by discussing the fundamental axioms of this theory.

Let us start with a puzzle

Some epistemological assumptions

Some basics on the theory of GENEOs

GENEOs and XAI

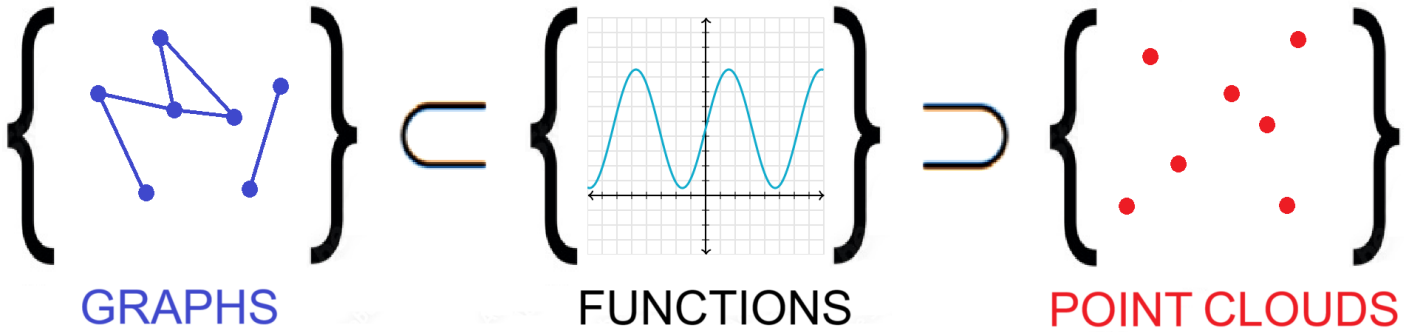
Assumption 1: Data are often represented by functions

Many types of data can be represented as functions:

Images, electrocardiograms, computerized tomography scans, and more.

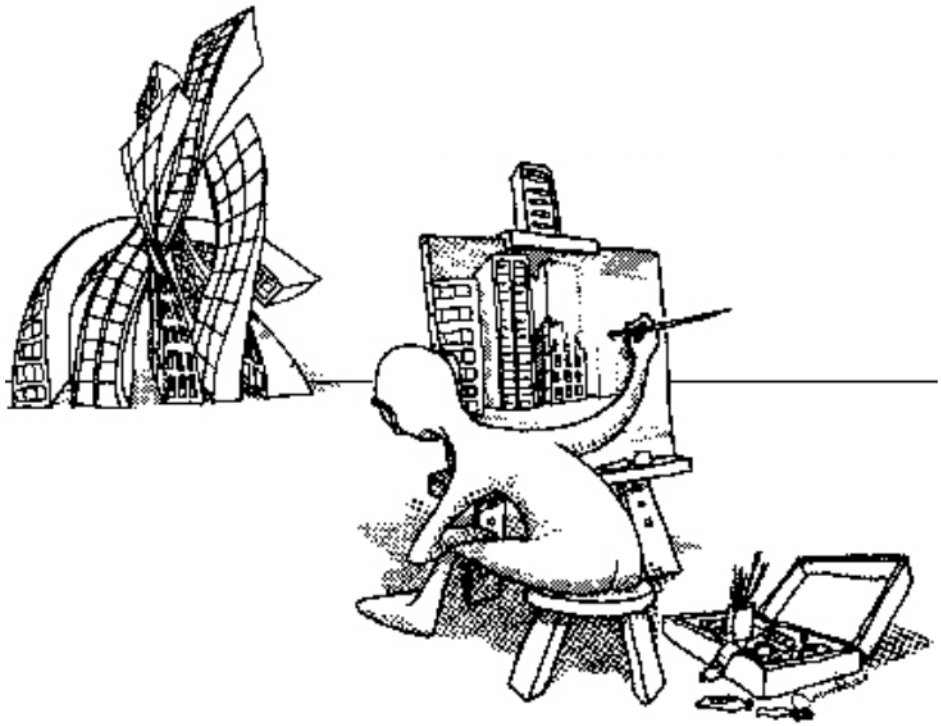
Additionally:

- A point cloud C in $\mathbb{R}^n$ (where C is equivalent to the function $d_C : \mathbb{R}^n \to \mathbb{R}$ that expresses the distance from C).
- A graph Γ (where Γ is equivalent to its adjacency matrix, which can be interpreted as a function).



Assumption 2: Data are processed by observers

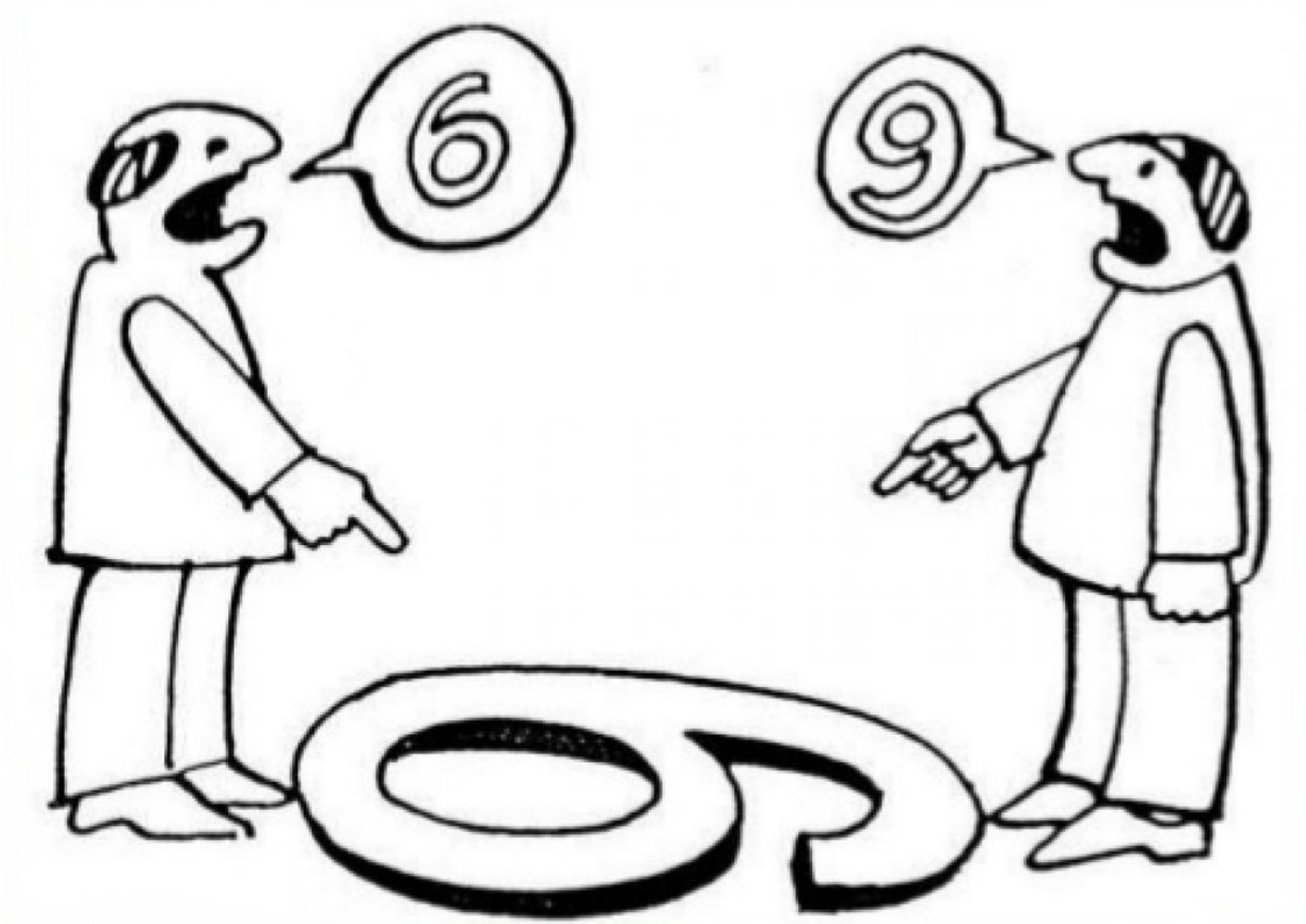
Data have no meaning without an observer to interpret them.



An observer is an agent that transforms data while preserving their symmetries.

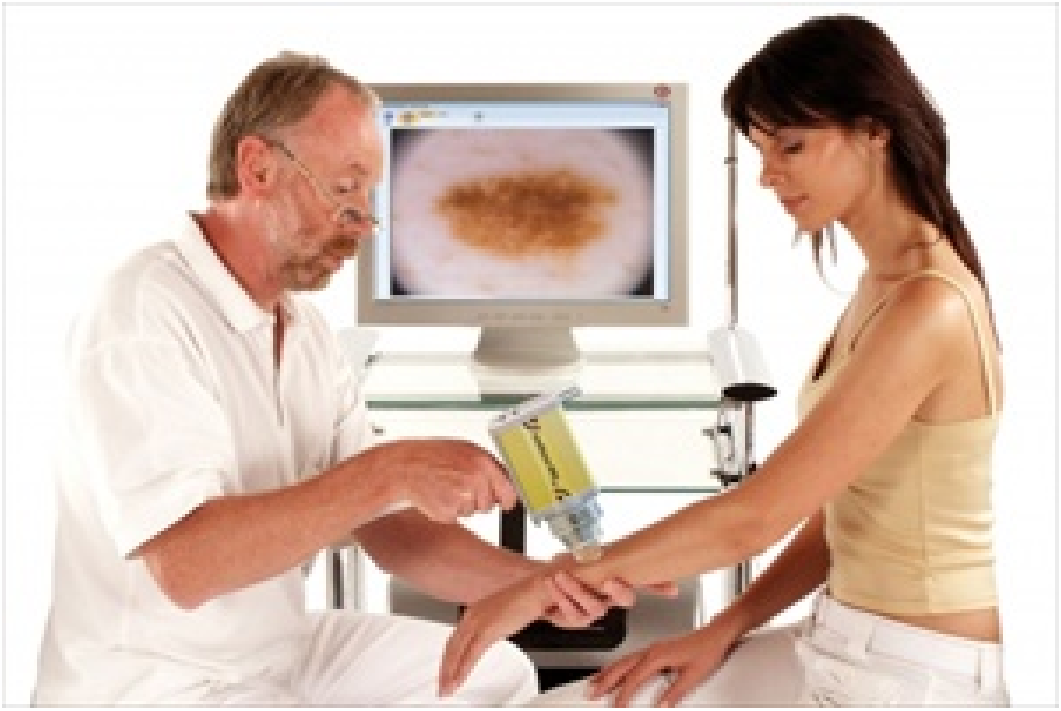
Assumption 3: Observers are variables

Data interpretation strongly depends on the chosen observer.



Assumption 4: Observers are important

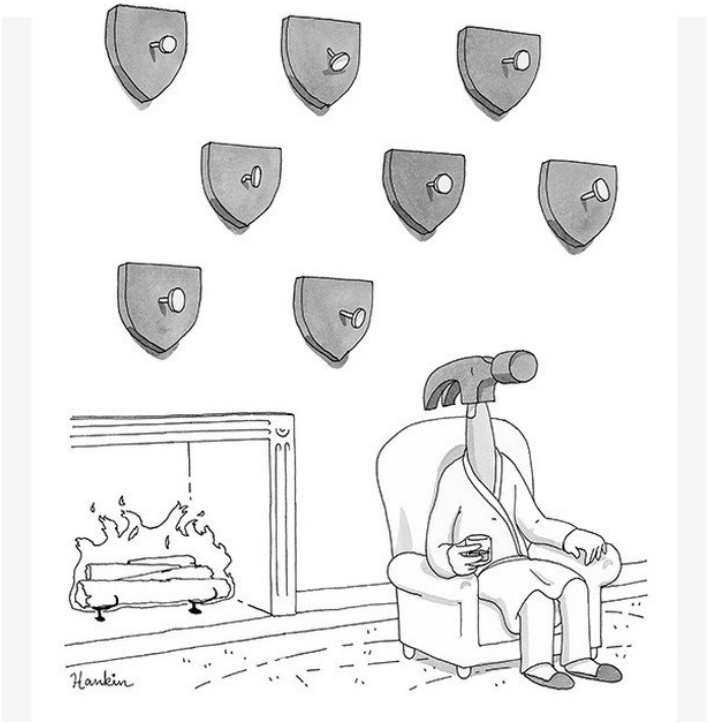
We are rarely directly interested in the data, but rather in how observers react to their presence.



Consequently, we should focus more on the properties of the observers than on the properties of the data.

Assumption 5: There is no structure in the data

Generally speaking, data lack inherent structure. Instead, the structure of data reflects the observer's own structure.



The shape is not in the data but in the eyes of the observer.

Let us start with a puzzle

Some epistemological assumptions

Some basics on the theory of GENEOs

GENEOs and XAI

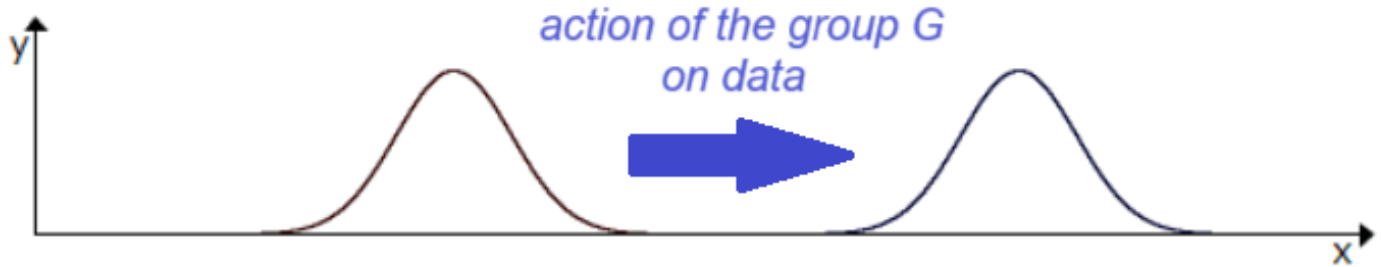
Let's start by defining perception pairs

Let us consider

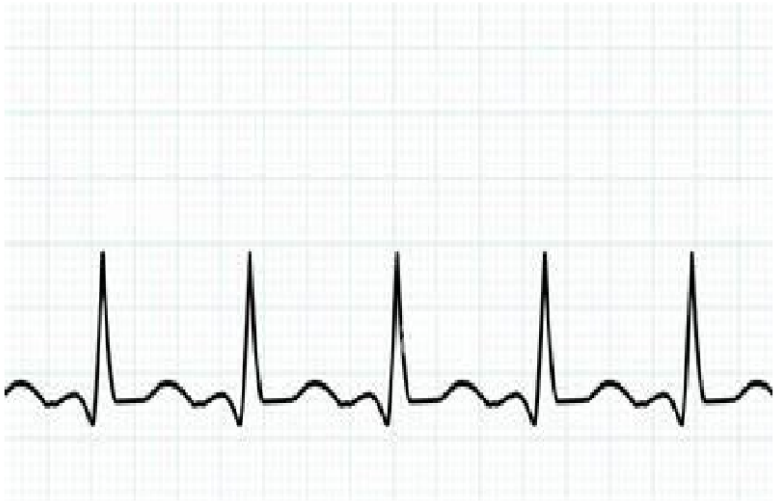
1. A collection Φ of functions from a set X to $\mathbb{R}$;
2. A group G of bijections $g : X \to X$ such that $\varphi \in \Phi \implies \varphi \circ g \in \Phi$ for every $\varphi \in \Phi$.

We say that (Φ, G) is a perception pair.

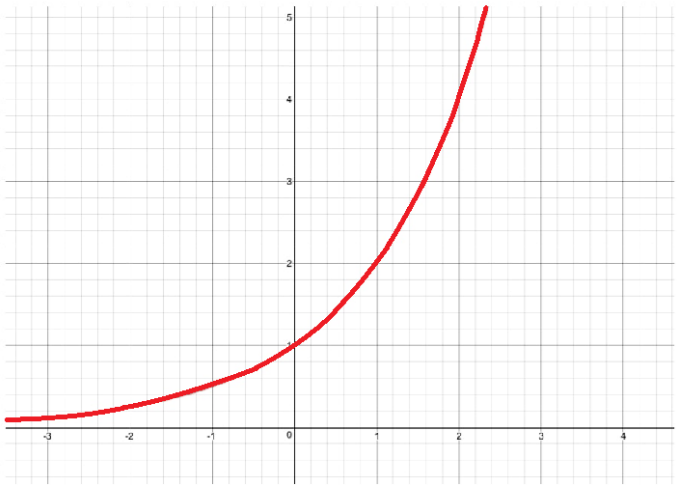
The choice of a perception pair states which data can be considered as legitimate measurements (the functions in Φ) and which group represents the admissible symmetries between data (the group G).



Admissible and not admissible data



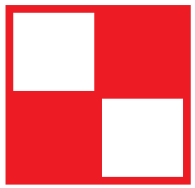
ADMISSIBLE AS AN ELECTROCARDIOGRAM



NOT ADMISSIBLE AS AN ELECTROCARDIOGRAM

6

ADMISSIBLE AS A NUMBER



NOT ADMISSIBLE AS A NUMBER

What metric can we consider on Φ, X and G?

We endow Φ with the sup-norm metric:

$$D_\Phi(\varphi_1, \varphi_2) = \sup_{x \in X} |\varphi_1(x) - \varphi_2(x)|.$$

NB: What other metric could we put on Φ, given that X is not endowed with any measure or structure?

Then, we endow X with the pseudo-metric

$$D_X(x_1, x_2) = \sup_{\varphi \in \Phi} |\varphi(x_1) - \varphi(x_2)|.$$

We recall that a pseudo-metric is just a metric d without the property $d(x_1, x_2) = 0 \implies x_1 = x_2$.

Finally, we put on G the pseudo-metric

$$D_G(g_1, g_2) := \sup_{\varphi \in \Phi} D_\Phi(\varphi \circ g_1, \varphi \circ g_2).$$

Some mathematical properties

- Every function $\varphi \in \Phi$ is non-expansive and hence continuous.
- 1. If Φ is compact and X is complete, then X is compact.
  2. If Φ is compact and G is complete, then G is compact.
  3. If Φ is totally bounded, we can always assume that Φ, X, and G are compact.
- G is a topological group for the topology induced by D_G, and the action of G on Φ by composition on the right is continuous.
- Any Φ-preserving bijection is an isometry.

A mathematical theory has been developed on this topic.

GEOs and GENEOs

Let us assume that two perception pairs (Φ, G), (Ψ, K) are given.

Each pair $(F : \Phi \to \Psi, T : G \to K)$ s. t. T is a homomorphism and

$$F(\varphi \circ g) = F(\varphi) \circ T(g)$$

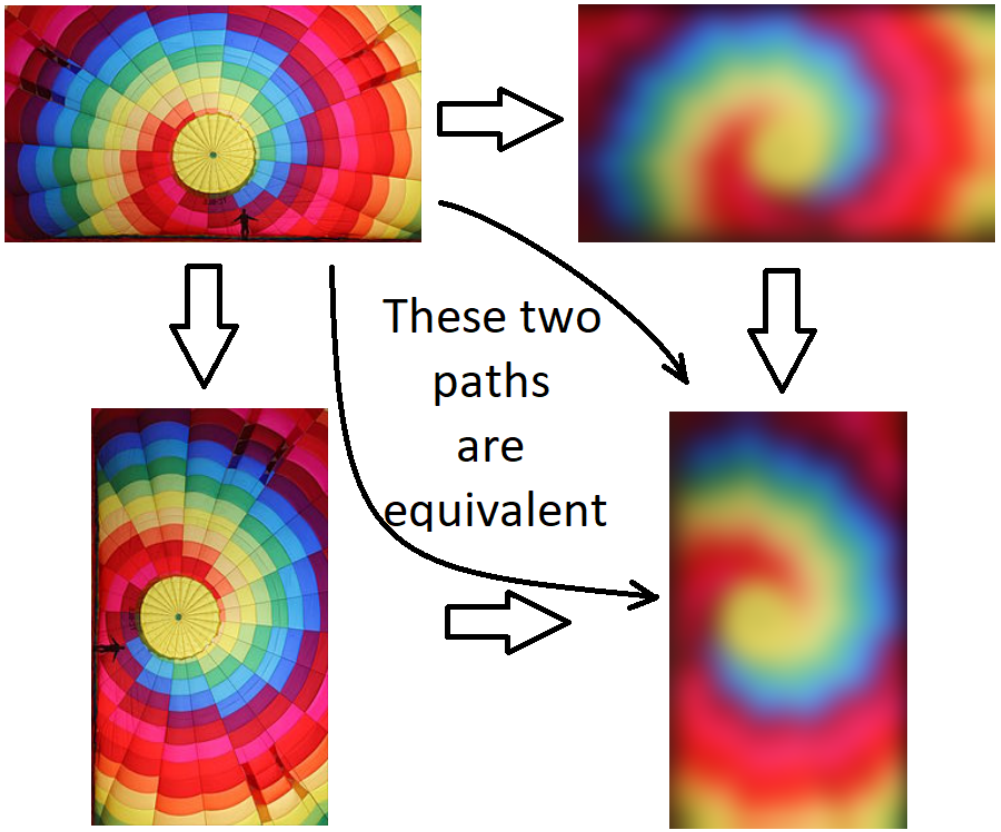
for every $\varphi \in \Phi, g \in G$ is called a *Group Equivariant Operator (***GEO***)*.

If F is also non-expansive (i.e., $D_\Psi(F(\varphi_1), F(\varphi_2)) \leq D_\Phi(\varphi_1, \varphi_2)$ for every $\varphi_1, \varphi_2 \in \Phi$), then (F, T) is called a *Group Equivariant Non-Expansive Operator (***GENEO***)*.

GEOs and GENEOs represent observers in our setting.

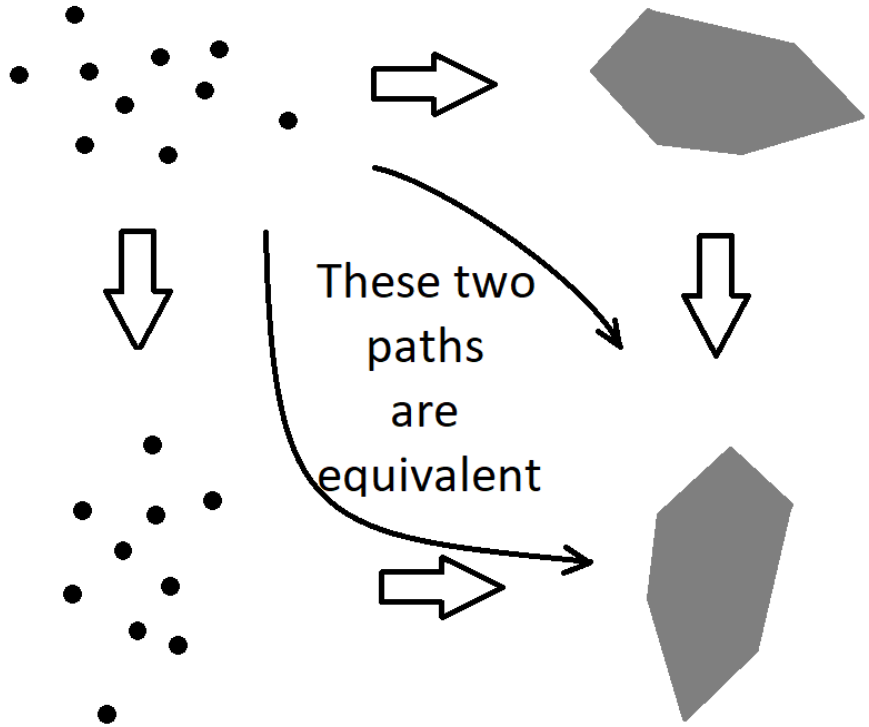
An example of GENEO

When we blur an image by applying a convolution with a rotationally symmetric kernel whose mass is less than 1 in L^1, we are applying a GENEO (here, we are considering the group of isometries).



Another example of GENEO

When we compute the convex hull of a cloud of points, we are applying a GENEO (here, we are considering the group of isometries).



Good news for applications

A metric can be naturally defined on the space of GENEOs between two fixed perception pairs (Φ, G) and (Ψ, K), given a fixed homomorphism T between the transformation groups G and K.

The following result holds.

Theorem

- *If the input and output spaces of admissible data are compact, then the space of GENEOs is also compact.* *(NOT TRUE FOR GEOS!)*
- *If the output space of admissible data is convex, then the space of GENEOs is also convex.*

Good news for applications

As a consequence,

- If the input and output spaces of admissible data can be approximated with arbitrarily small error, then the **space of observers** has the same property.
- If the output space of admissible data is convex, then the **space of observers** is also convex.

Three key observations (1)

- While the input space Φ of data is often non-convex (and hence averaging data does not make sense), the assumption of convexity of the output space Ψ implies the convexity of the space of observers and allows us to consider the "average of observers".

Three key observations (2)

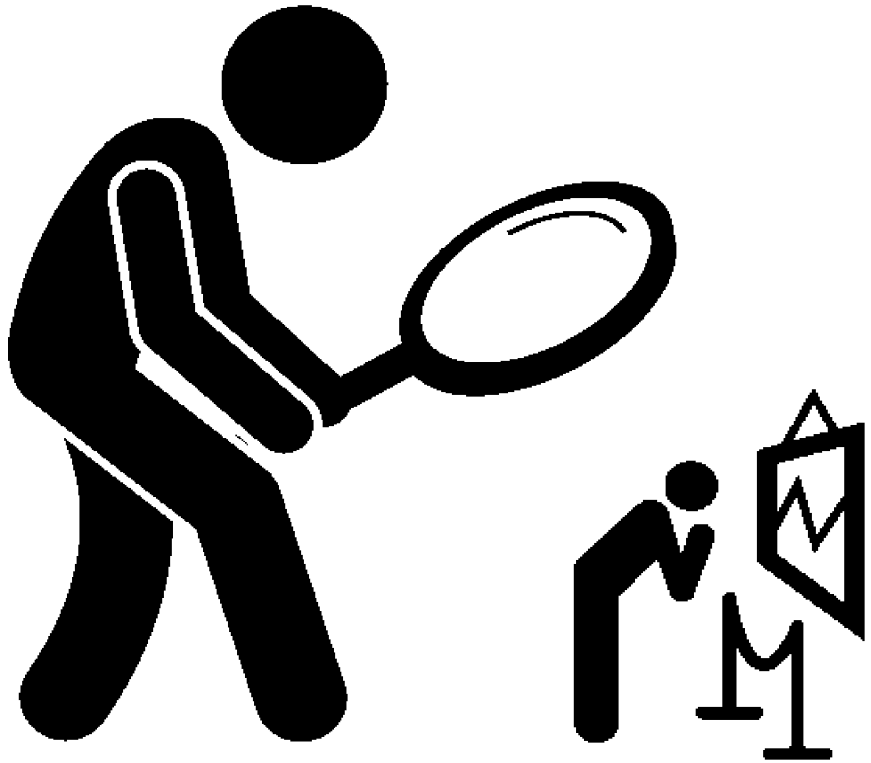
Our main goal is to develop a robust geometric and compositional theory for approximating an ideal observer through GENEOs.

$\approx$

$$(\Psi, K)$$
$$\uparrow$$
$$(F, T)$$
$$|$$
$$(\Phi, G)$$

Three key observations (3)

GENEOs are functions and can be taken as inputs of higher-level GENEOs. Data obtained through measuring instruments can be seen as GENEOs of level 0. Therefore, hierarchies of GENEOs can be considered.



Construction of GENEOs

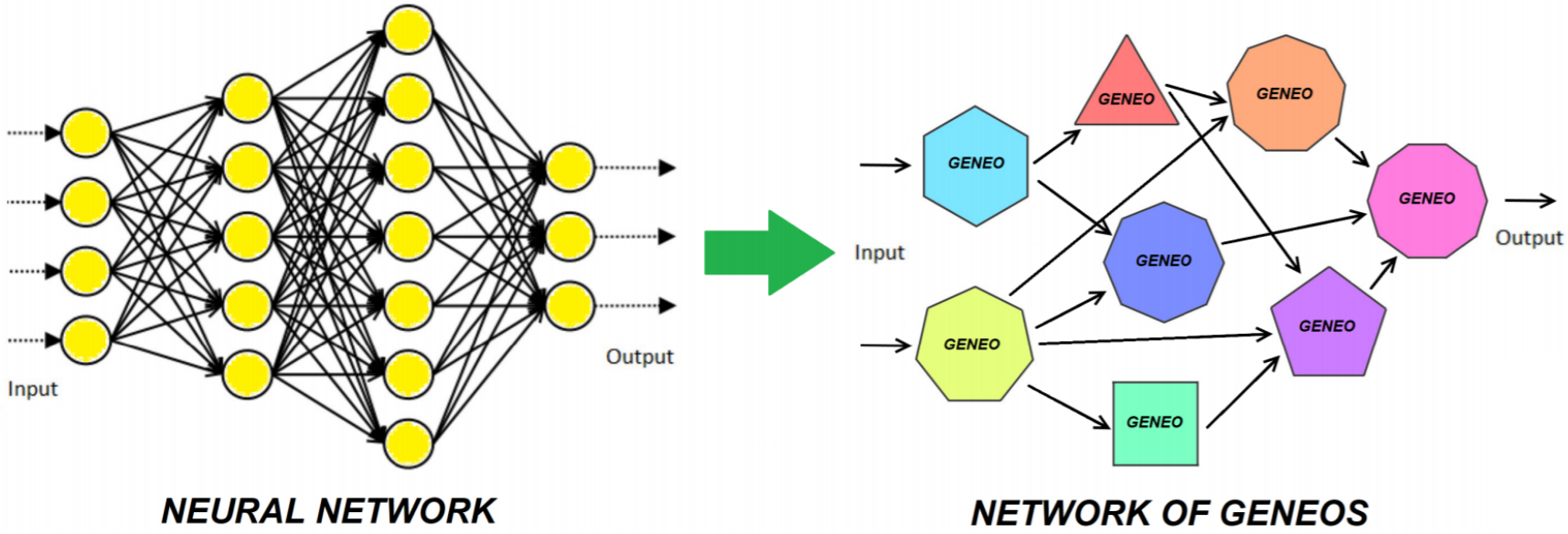
How can we build GENEOs?

The space of GENEOs is closed under composition, computation of minimum and maximum, translation, direct product, and convex combination. (However there is much more than this...)

GENEOs are like LEGO bricks that can be combined together to form more complex GENEOs.

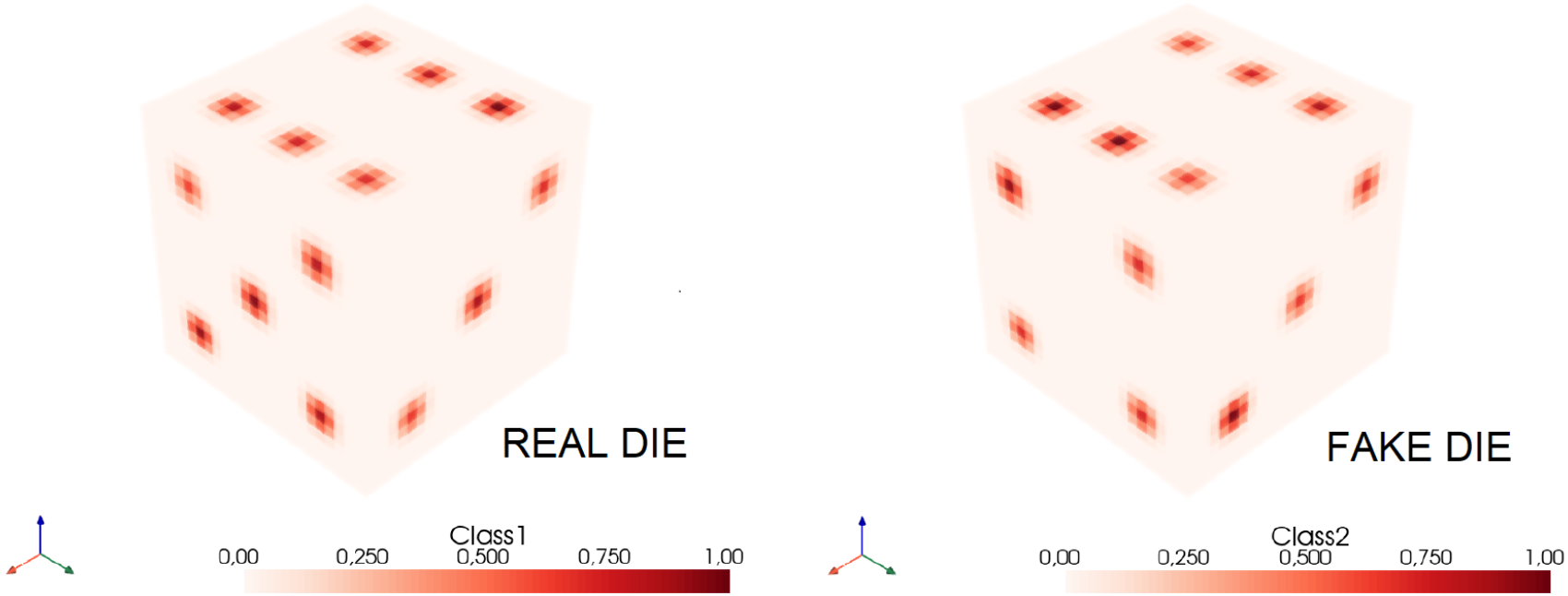
The main point in the approach based on GENEOs

In perspective, we are looking for a good compositional theory for building efficient and transparent networks of GENEOs.
Some preliminary experiments suggest that replacing neurons with GENEOs could make deep learning more transparent and interpretable and speed up the learning process.



What happens when we apply GENEOs to our data?

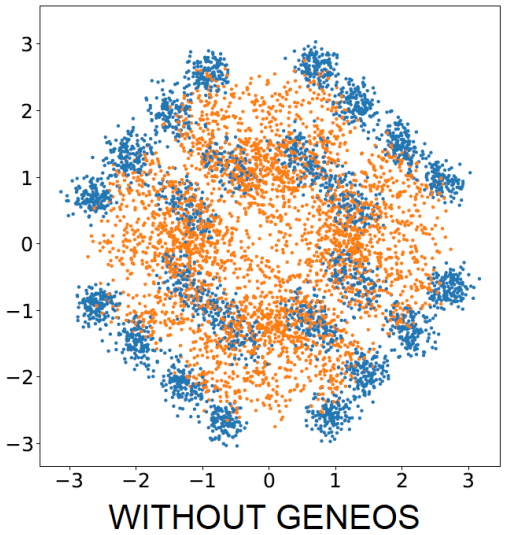
An example of use: comparison between real dice and fake dice.



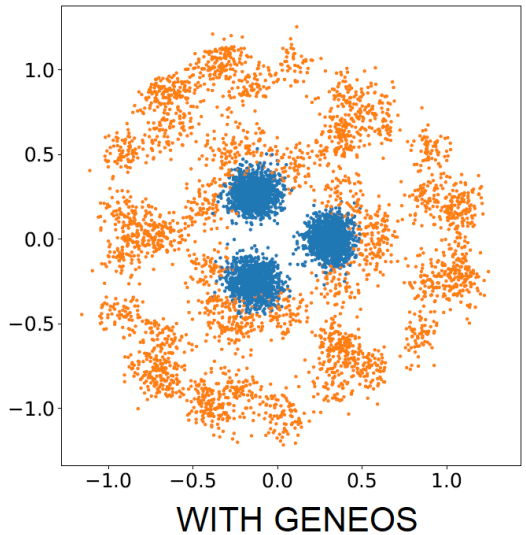
(Experiment and computations by Giovanni Bocchi)

What happens to data when we apply GENEOs?

We produced 10000 dice (a training set of size 7000 and a test set of size 3000), then we applied PCA to the test set and to the test set transformed by a suitable GENEO, optimized on the training set:



WITHOUT GENEOS



WITH GENEOS

For each die the first two principal components are plotted. Blue points are associated with **real dice**, while orange ones with **fake dice**. The GENEO we use was built by a convex combination of 3 GENEOs defined by permutant measures.

GENEOs and Machine Learning

More details about the theory of GENEOs are available in this paper:

nature machine intelligence

ARTICLES

https://doi.org/10.1038/s42256-019-0087-3

Towards a topological–geometrical theory of group equivariant non-expansive operators for data analysis and machine learning

Mattia G. Bergomi[1], Patrizio Frosini[2,3*], Daniela Giorgi[4] and Nicola Quercioli[2,3]

vol. 1(9) (2019), 423–433.

https://rdcu.be/bP6HV

GENEOs and Machine Learning

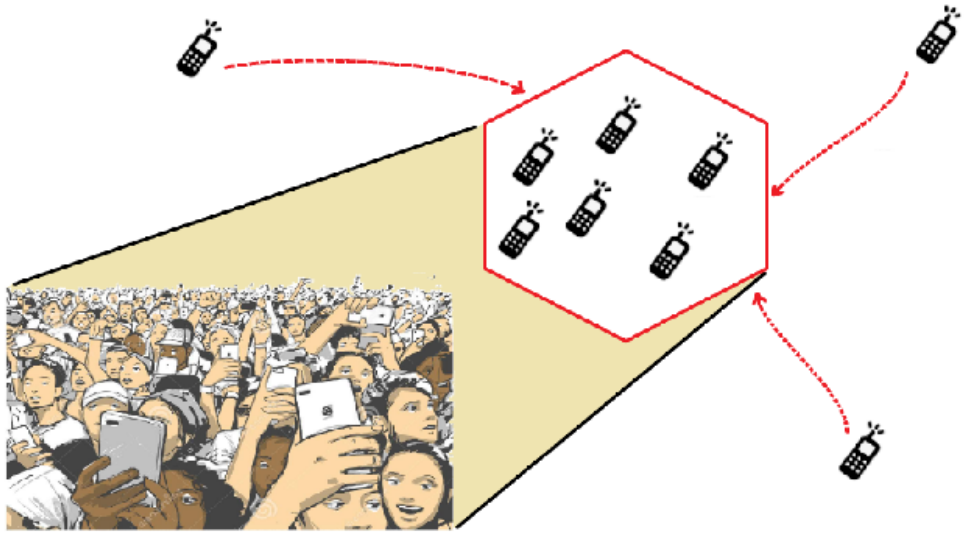
For more details about the use of GENEOs in Machine Learning, you can have a look at this paper:

https://ems.press/journals/mag/articles/10389352

Current research projects (I)

CNIT / WiLab - Huawei Joint Innovation Center (JIC)

Project on GENEOs for 6G



WILAB HUAWEI

Current research projects (II)

PANDORA

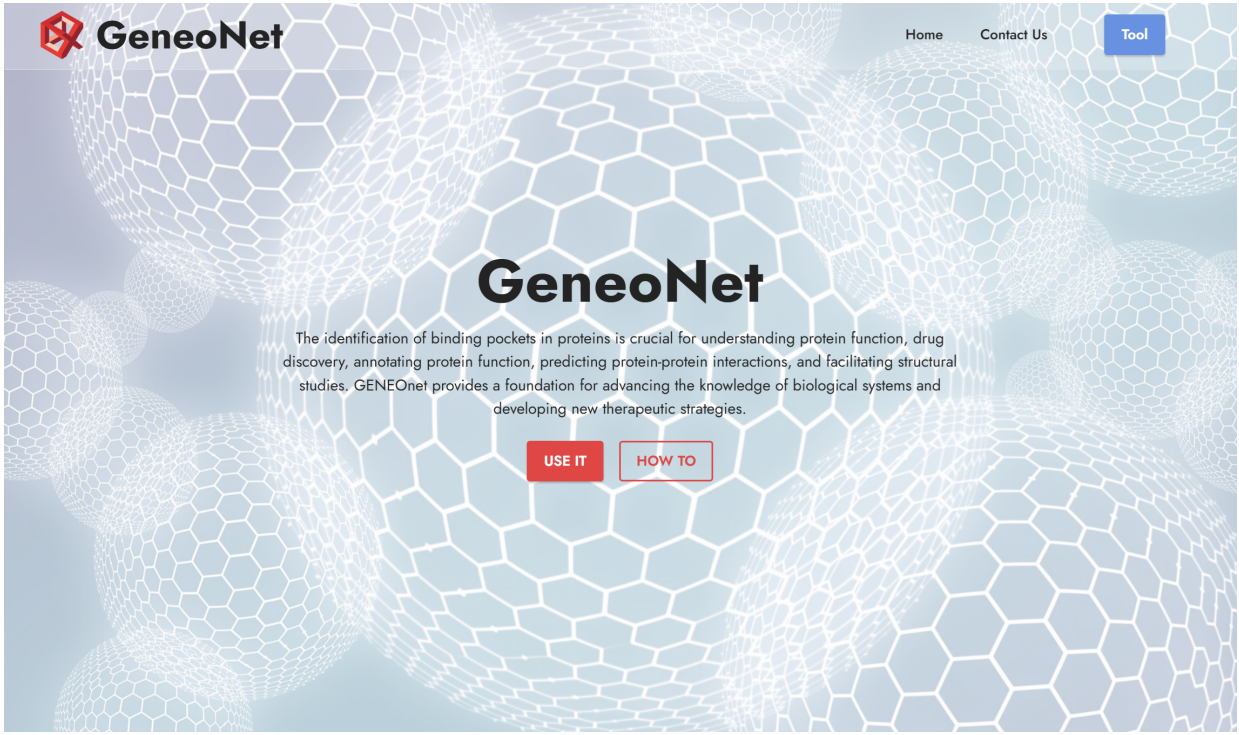
Horizon Europe (HORIZON)
Call: HORIZON-CL4-2023-HUMAN-01-CNECT
Project: 101135775-PANDORA
Funding: approximately 9 million euros.

Task 3.3 - Leveraging domain knowledge for explainable learning:
This task aims to investigate the use of domain knowledge in the development of explainable AI models. Tools like GENEOs for applications in TDA and ML and new theoretical methods of GENEOs for explainable AI will be used.

The project has received funding form the European Union's Horizon Europe Framework Programme (Horizon) under grant agreement No 101135775

https://pandora-heu.eu/consortium/

Current research projects (III)



The GENEOnet webservice represents the outcome of our partnership with Italian Pharmaceutical Company Dompé Farmaceutici S.p.A.: https://geneonet.exscalate.eu/

Finding pockets in proteins by applying GENEOs

GENEOnet: A new machine learning paradigm based on Group Equivariant Non-Expansive Operators. An application to protein pocket detection.

Giovanni Bocchi [1], Patrizio Frosini [2], Alessandra Micheletti [1], Alessandro Pedretti [3] Carmen Gratteri [4], Filippo Lunghini [5] , Andrea Rosario Beccari [5] and Carmine Talarico [5]

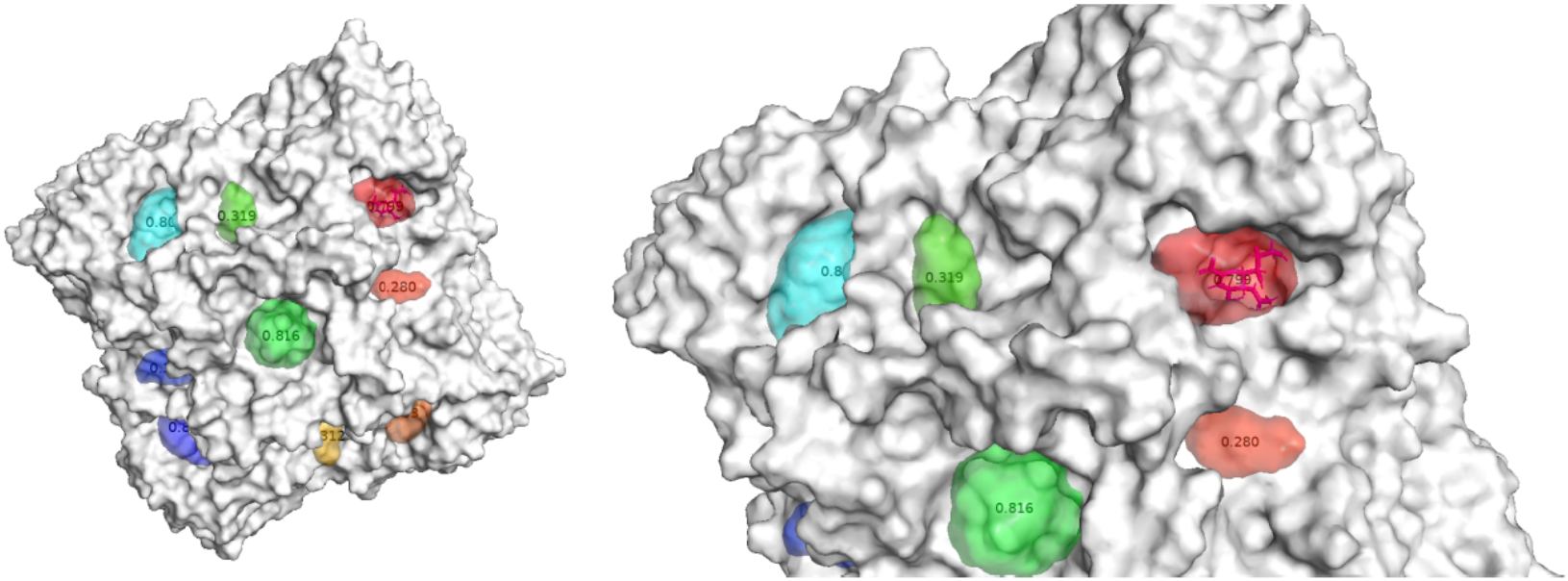
[1] Department of Environmental Science and Policy, Università degli Studi di Milano
[2] Department of Mathematics, Università degli Studi di Bologna
[3] Department of Pharmaceutical Sciences, Università degli Studi di Milano
[4] Dipartimento di Scienze della Salute, Università degli Studi "Magna Græcia di Catanzaro"
[5] Dompé Farmaceutici SpA

https://arxiv.org/ftp/arxiv/papers/2202/2202.00451.pdf

Updated results of this research have been presented at xAI-2024 (The 2nd World Conference on eXplainable Artificial Intelligence). Giovanni Bocchi has produced the data shown in these slides.

Finding pockets in proteins by applying GENEOs

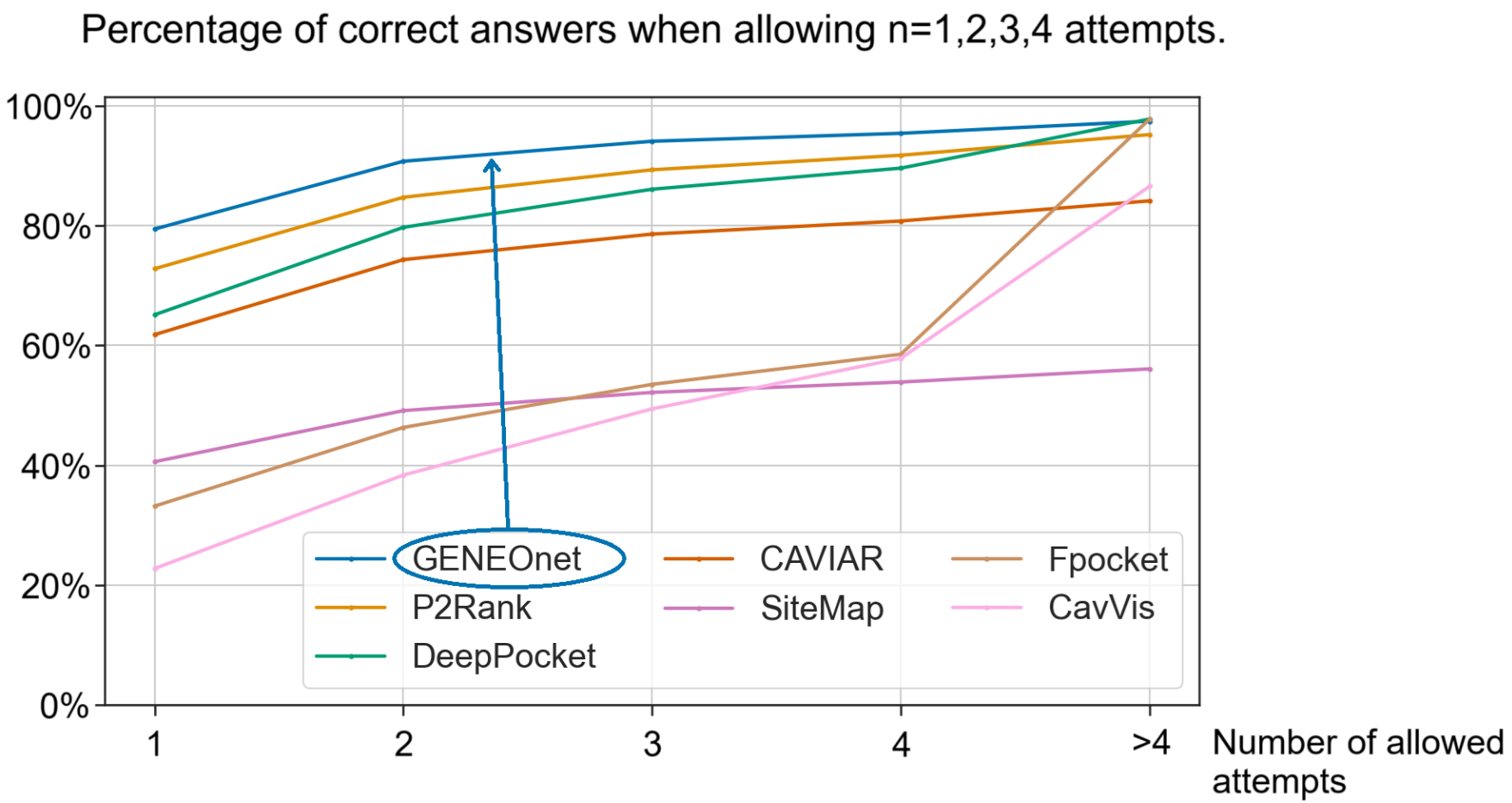
GENEOs can be used for the detection of druggable protein pockets.



a) b)

Model predictions for protein 2QWE. In Figure a) the global view of the prediction is shown, where different pockets are depicted in different colors and are labelled with their scores. In Figure b) the zoomed of the pocket containing the ligand is shown.

Results



Please note that GENEOnet uses 17 parameters, while a CNN such as DeepPocket requires 665122 parameters.

Let us start with a puzzle

Some epistemological assumptions

Some basics on the theory of GENEOs

GENEOs and XAI

Basic idea

How can we mathematically and generally formalize the concept of an explanation provided by an agent, viewed as a functional operator?

Informal idea: We could say that the action of an agent A is explained by another agent B from the perspective of an agent C if:

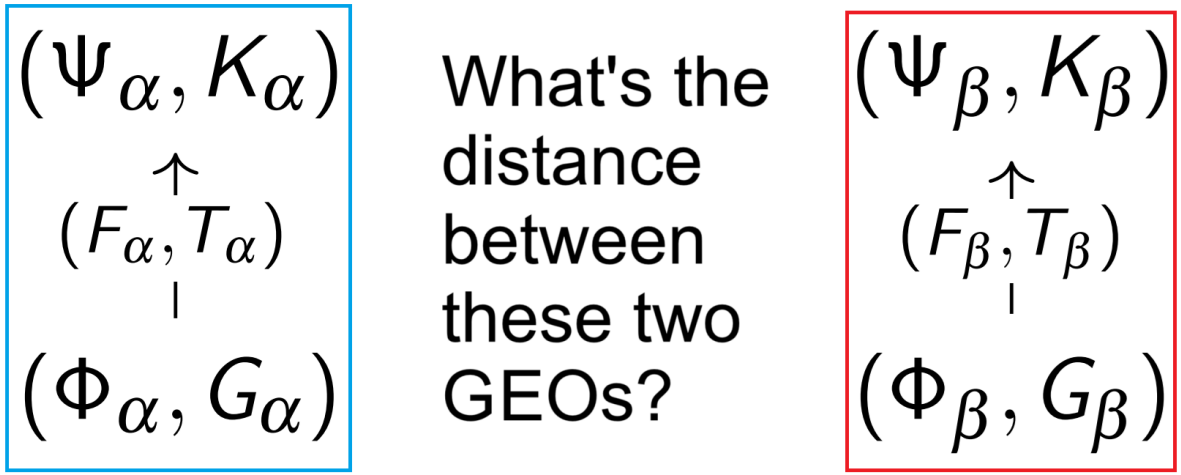
1. C perceives A and B as similar to each other;
2. C perceives B as less complex than A.

Note that if A and B are represented as GEOs, they are functions and can therefore be treated as inputs to a higher-level GEO C.

E.g., let's consider two neural networks represented as two GEOs.

An extended pseudo-metric for *ALL* GEOs

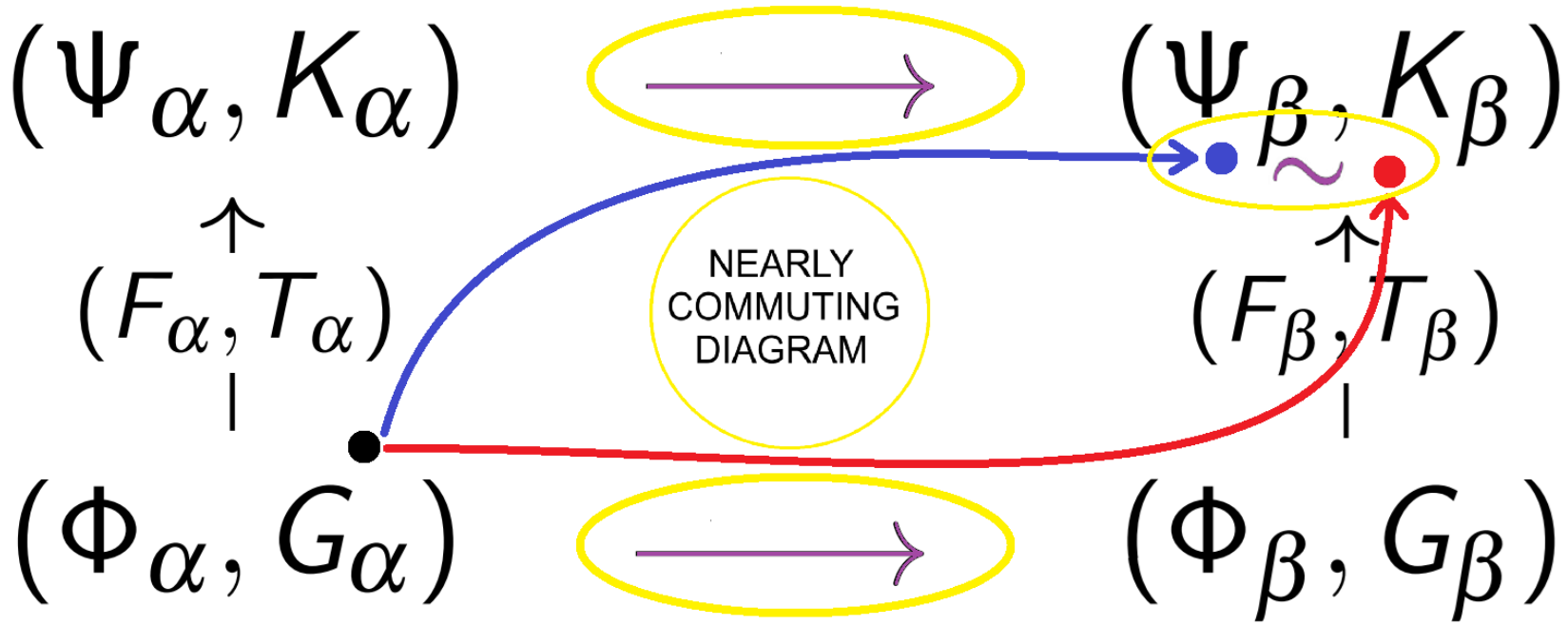
The first step in formalizing this idea is to introduce a metric between GEOs that remains well-defined even when the GEOs operate on different domains and produce outputs in distinct codomains. This is a non-trivial challenge.



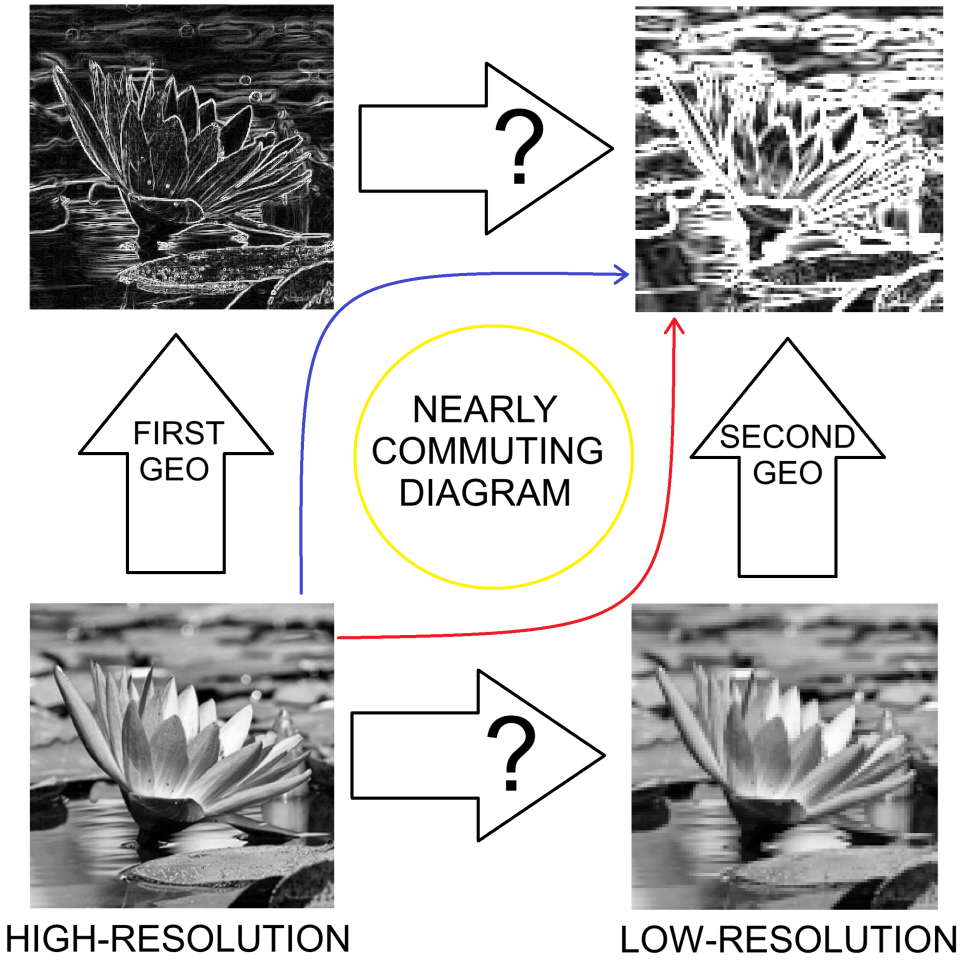
In other words, what does it mean for two GEOs to behave approximately the same way?

An extended pseudo-metric for *ALL* GEOs

Informally speaking, two GEOs are considered similar if we can find two horizontal GENEOs that make this diagram nearly commutative:



An example

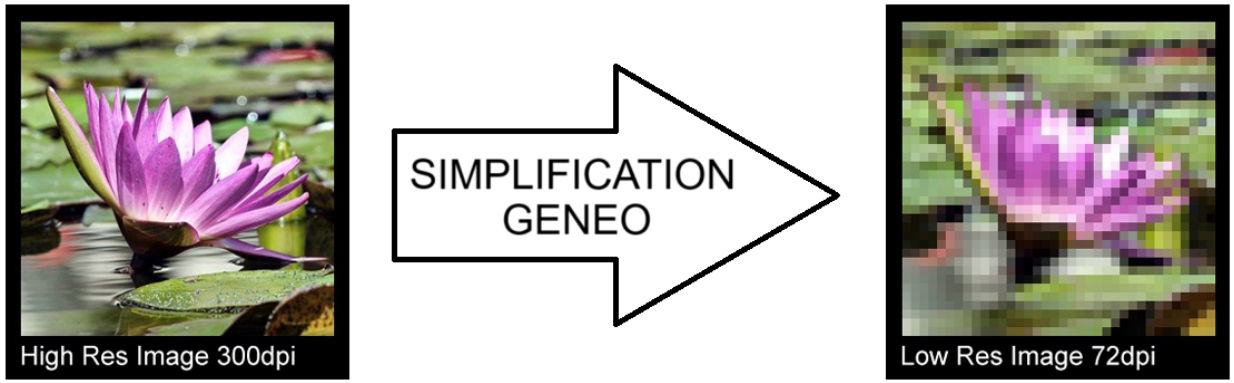


In the next slides we will formalize this idea.

An extended pseudo-metric for *ALL* GEOs

Let $\mathbf{S}_{all}$ be the category whose objects are all perception pairs, and whose morphisms $(F, T) : (\Phi, G) \to (\Phi', G')$ are GENEOs.
The morphisms in $\mathbf{S}_{all}$ are called *simplification GENEOs*. These morphisms describe the possible "logical correspondence" between data represented by different perception pairs.

For example, a simplification GENEO might transform high-resolution images into low-resolution images.



An extended pseudo-metric for *ALL* GEOs

Let us choose a set $\mathscr{G}$ of GEOs (we do not require that they have the same input or output perception pairs). Therefore,

$$\mathscr{G} = \{(F_\alpha, T_\alpha) : (\Phi_\alpha, G_\alpha) \to (\Psi_\alpha, K_\alpha)\}_{\alpha \in A}.$$

We do not require that the elements of $\{(\Phi_\alpha, G_\alpha)\}_{\alpha \in A}$ are distinct. The same applies to $\{(\Psi_\alpha, K_\alpha)\}_{\alpha \in A}$.

Let $\mathbf{S}$ be a small subcategory of the category $\mathbf{S}_{all}$.

$\mathscr{G}$ will be the set of GEOs/observers where we will define our pseudo-metric, while $\mathbf{S}$ will represent the collection of simplification GENEOs considered admissible.

An extended pseudo-metric for *ALL* GEOs

For any pair of GEOs

$$\Big((F_\alpha, T_\alpha) : (\Phi_\alpha, G_\alpha) \to (\Psi_\alpha, K_\alpha), (F_\beta, T_\beta) : (\Phi_\beta, G_\beta) \to (\Psi_\beta, K_\beta)\Big)$$

in $\mathscr{G}$, let us choose a set of quadruples of GENEOs

$$\mathscr{Q}(\alpha,\beta) \subseteq \left\{ \Big((L^i_{\alpha,\beta}, P_{\alpha,\beta}), (M^i_{\alpha,\beta}, Q_{\alpha,\beta}), (L^i_{\beta,\alpha}, P_{\beta,\alpha}), (M^i_{\beta,\alpha}, Q_{\beta,\alpha})\Big) \right\}_{i \in I}$$

in $\mathbf{S} \times \mathbf{S} \times \mathbf{S} \times \mathbf{S}$, such that $\mathscr{Q}(\beta,\gamma) \circ \mathscr{Q}(\alpha,\beta) = \mathscr{Q}(\alpha,\gamma)$ for any α and β, and [SEE NEXT SLIDE]

An extended pseudo-metric for *ALL* GEOs

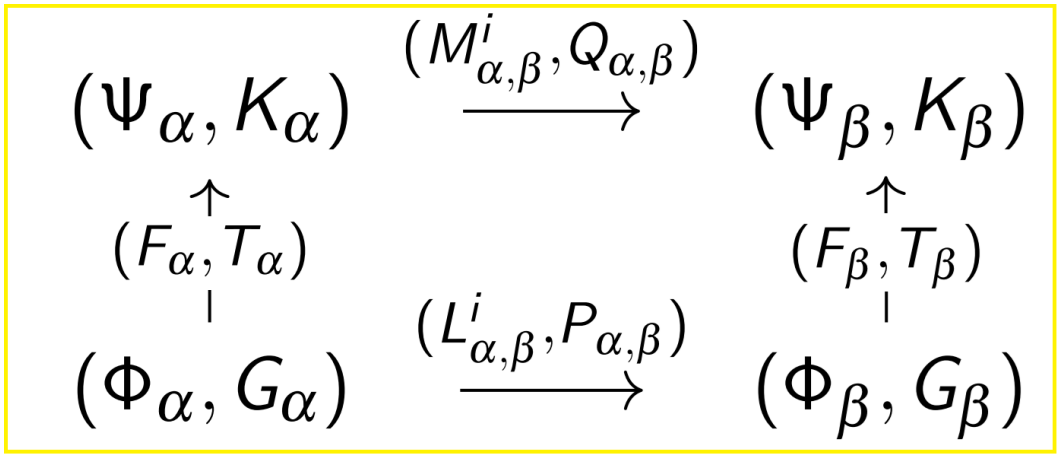
[CONTINUE FROM THE PREVIOUS SLIDE]

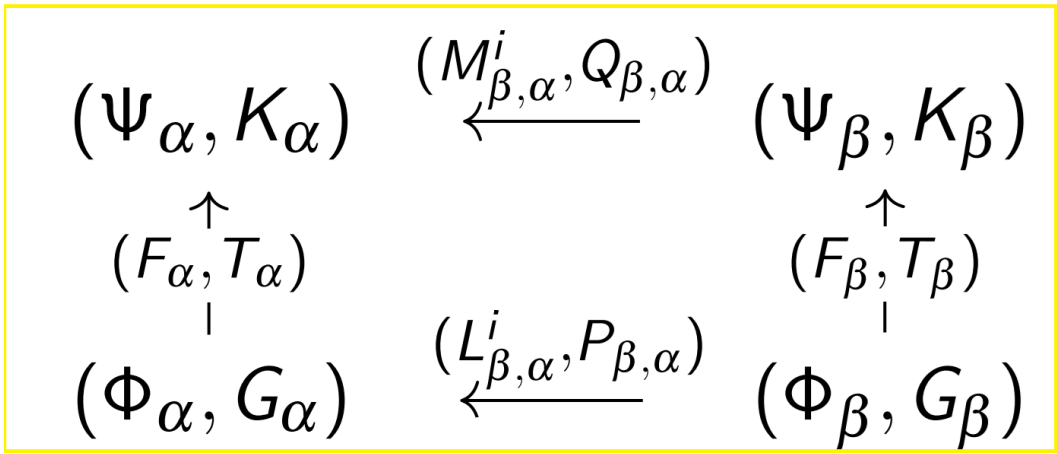
- $(L^i_{\alpha,\beta}, P_{\alpha,\beta})$ is a GENEO from (Φ_α, G_α) to (Φ_β, G_β)
- $(M^i_{\alpha,\beta}, Q_{\alpha,\beta})$ is a GENEO from (Ψ_α, K_α) to (Ψ_β, K_β)
- $(L^i_{\beta,\alpha}, P_{\beta,\alpha})$ is a GENEO from (Φ_β, G_β) to (Φ_α, G_α)
- $(M^i_{\beta,\alpha}, Q_{\beta,\alpha})$ is a GENEO from (Ψ_β, K_β) to (Ψ_α, K_α)
- $\bigcup_i L^i_{\alpha,\beta}(\Phi_\alpha) = \Phi_\beta$, $\bigcup_i M^i_{\alpha,\beta}(\Psi_\alpha) = \Psi_\beta$, $\bigcup_i L^i_{\beta,\alpha}(\Phi_\beta) = \Phi_\alpha$, $\bigcup_i M^i_{\beta,\alpha}(\Psi_\beta) = \Psi_\alpha$ *(data completeness assumption)*

for any $i \in I$.

An extended pseudo-metric for *ALL* GEOs

This is summarized in the figure below:





An extended pseudo-metric for *ALL* GEOs

To measure the non-commutativity of our diagrams, we need to define the following cost functions:

$$\mathrm{cost}\left(L^i_{\alpha,\beta}, M^i_{\alpha,\beta}\right) = \sup_{\varphi\in\Phi_\alpha} D_{\Phi_\alpha}\left(M^i_{\alpha,\beta}\circ F_\alpha(\varphi), F_\beta\circ L^i_{\alpha,\beta}(\varphi)\right)$$

$$\mathrm{cost}\left(P_{\alpha,\beta}, Q_{\alpha,\beta}\right) = \sup_{g\in G_\alpha} D_{K_\beta}\left(Q_{\alpha,\beta}\circ T_\alpha(g), T_\beta\circ P_{\alpha,\beta}(g)\right)$$

$$\mathrm{cost}\left(L^i_{\beta,\alpha}, M^i_{\beta,\alpha}\right) = \sup_{\varphi\in\Phi_\beta} D_{\Phi_\beta}\left(M^i_{\beta,\alpha}\circ F_\beta(\varphi), F_\alpha\circ L^i_{\beta,\alpha}(\varphi)\right)$$

$$\mathrm{cost}\left(P_{\beta,\alpha}, Q_{\beta,\alpha}\right) = \sup_{g\in G_\beta} D_{K_\alpha}\left(Q_{\beta,\alpha}\circ T_\beta(g), T_\alpha\circ P_{\beta,\alpha}(g)\right)$$

The non-expansiveness of GENEOs and the assumption of data completeness imply that $P_{\alpha,\beta}$, $Q_{\alpha,\beta}$, $P_{\beta,\alpha}$, $Q_{\beta,\alpha}$ are non-expansive.

An extended pseudo-metric for *ALL* GEOs

Now, we can define our pseudo-distance between GEOs.

Definition

If $(F_\alpha, T_\alpha), (F_\beta, T_\beta) \in \mathscr{G}$, then we can consider the value $d((F_\alpha, T_\alpha), (F_\beta, T_\beta))$ defined as follows

$$\inf_{i \in I} \max \Big\{ \mathrm{cost}\Big(L^i_{\alpha,\beta}, M^i_{\alpha,\beta}\Big), \mathrm{cost}\Big(P_{\alpha,\beta}, Q_{\alpha,\beta}\Big), \mathrm{cost}\Big(L^i_{\beta,\alpha}, M^i_{\beta,\alpha}\Big), \mathrm{cost}\Big(P_{\beta,\alpha}, Q_{\beta,\alpha}\Big) \Big\}$$

if $\mathscr{Q}(\alpha, \beta) \neq \emptyset$, and ∞ otherwise.

This value measures the non-commutativity of our diagrams.

An extended pseudo-metric for *ALL* GEOs

Proposition

d is an extended pseudo-distance.

This statement does not hold for **expansive** operators.

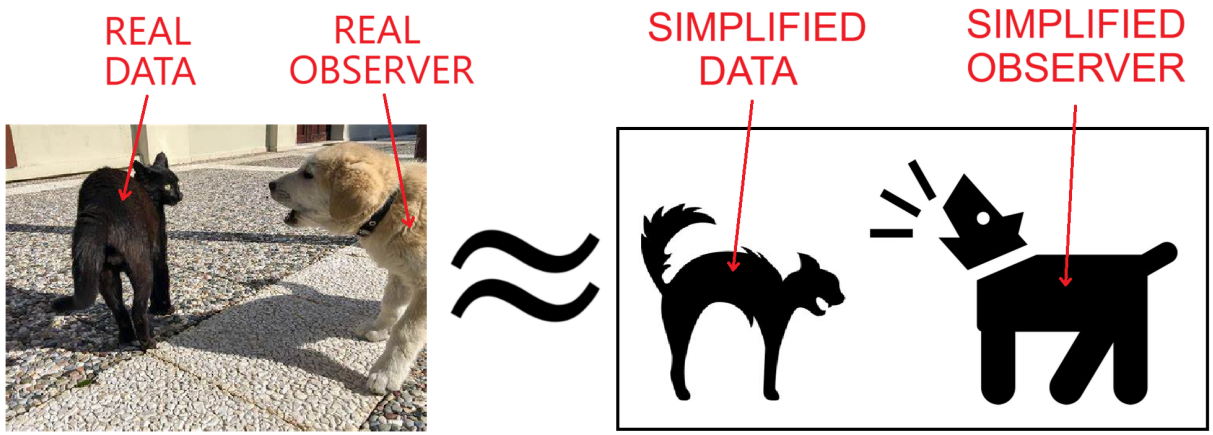
The non-expansiveness of GENEOs is a key component of our theory.

In simple terms, the value $d((F_\alpha, T_\alpha),(F_\beta, T_\beta))$ measures the *cost* of simplifying the data and the transformation groups.

When $d((F_\alpha, T_\alpha),(F_\beta, T_\beta))$ is small, it indicates that the GEOs (F_α, T_α) and (F_β, T_β) act approximately in the same way on the data they process.

A mathematical concept of explanation

In summary, the pseudo-metric d enables us to introduce a precise mathematical concept of *explanation*. Specifically, we can define it as follows: The action of an agent represented by a GEO (F_α, T_α) is *explained at a level* ε by the action of another agent represented by a GEO (F_β, T_β) when $d((F_\alpha, T_\alpha),(F_\beta, T_\beta)) \leq \varepsilon$.



A general mathematical model of explainability grounded in a precise operator theory could be beneficial for XAI.

Summary

To sum up, GENEOs are

- novel mathematical tools designed to approximate observers
- particularly useful when some knowledge about the behavior of observers is available
- generally interpretable
- potentially useful for explainable artificial intelligence (XAI).

THANKS FOR YOUR ATTENTION