A brief introduction to the concept of GENEO and its use for XAI

Patrizio Frosini

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

Knowledge Discovery and Data Mining Laboratory
November 21, 2024

Outline

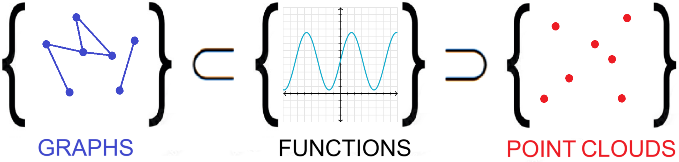
Assumption 1: Data are often represented by functions

Many types of data can be represented as functions:

Images, electrocardiograms, computerized tomography scans, and more.

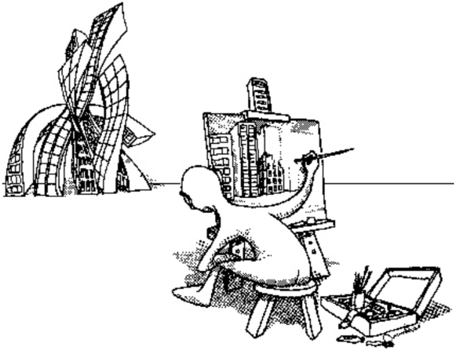
Additionally:

- 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

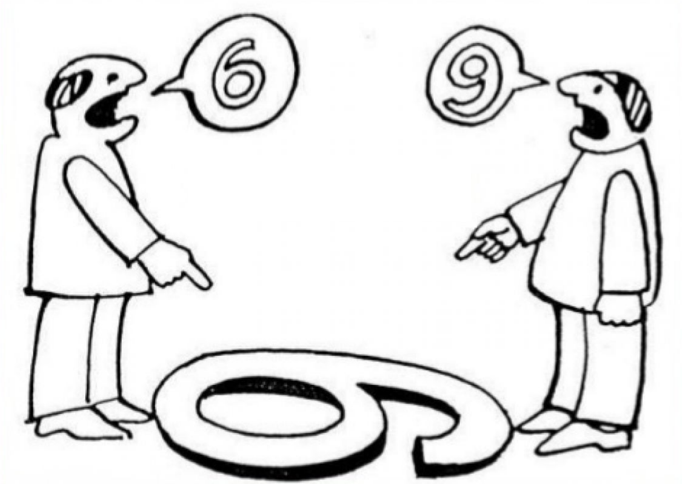
Data have no meaning without an observer to interpret them.



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.



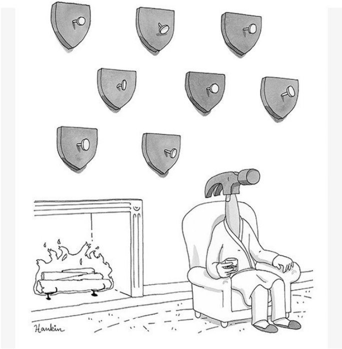
Assumption 4: Observers are important

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's start by defining perception pairs

Let us consider

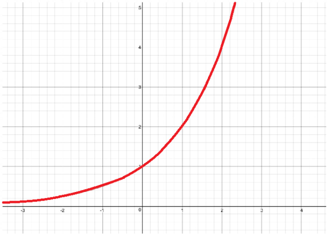
1. A collection Φ of functions from a set X to $\mathbb{R}^k$;
2. A group G of bijections $g : X \to X$ such that $\varphi \in \Phi \implies \varphi \circ g \in \Phi$ for every $\varphi \in \Phi$.

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

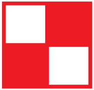
ADMISSIBLE AS AN ELECTROCARDIOGRAM



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:

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

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

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

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

A mathematical theory has been formulated to analyze and describe perception pairs.

For example:

- 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.

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

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

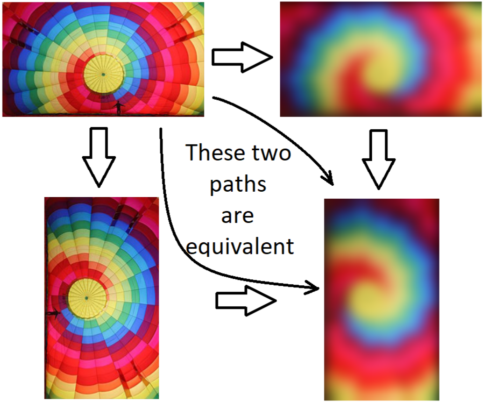
for every $\varphi \in \Phi, g \in G$ is called a *Group Equivariant Operator (**GEO**)*.

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.

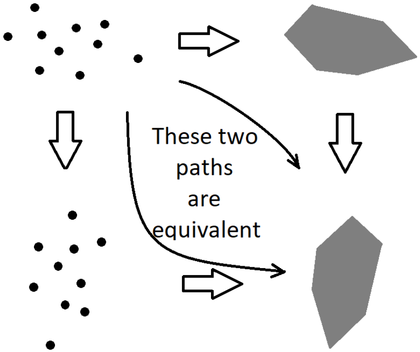
An example of GENEO

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.

These properties are quite useful in applications.

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".

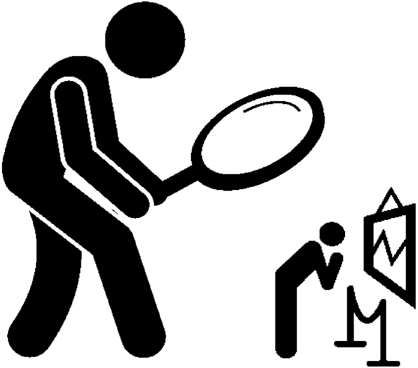
Three key observations (2)

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

$$\approx \begin{array}{c} (\Psi, K) \\ \uparrow \\ (F, T) \\ | \\ (\Phi, G) \end{array}$$

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

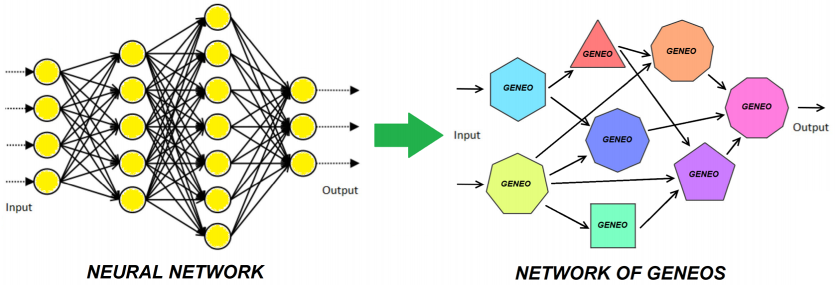
How can we build GENEOs?

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.

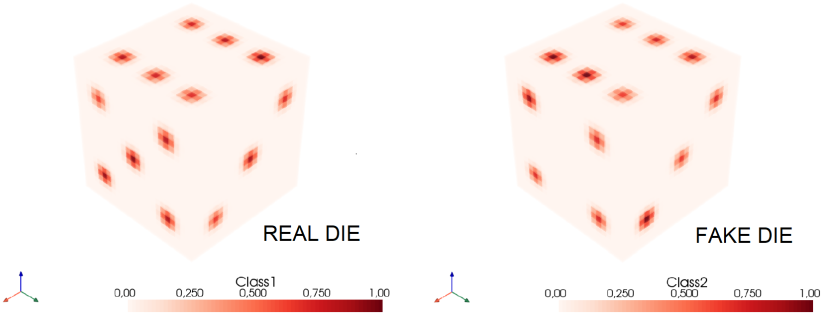
The main point in the approach based on GENEOs

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?

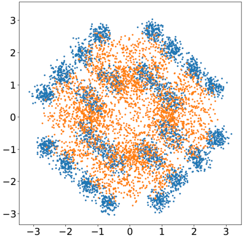
An example of use: comparison between real dice and fake dice.



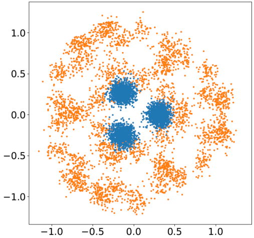
(Experiment and computations by Giovanni Bocchi)

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

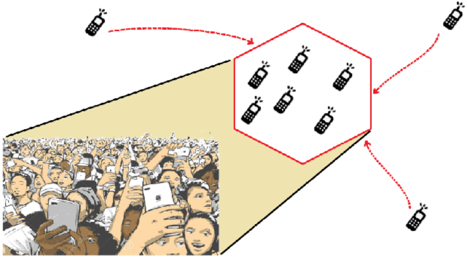
For more details about the use of GENEOs in Machine Learning, you can have a look at this paper:

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

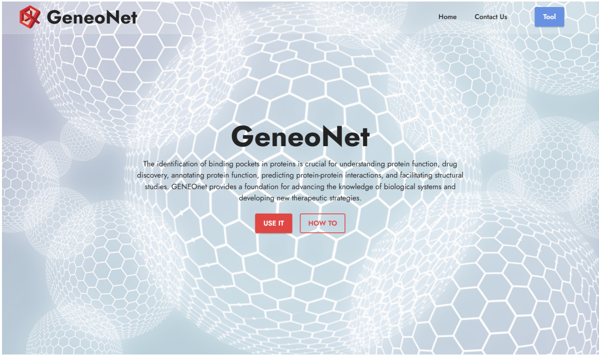
Horizon Europe (HORIZON)
Call: HORIZON-CL4-2023-HUMAN-01-CNECT
Project: 101135775-PANDORA
Funding: approximately 9 million euros.

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"

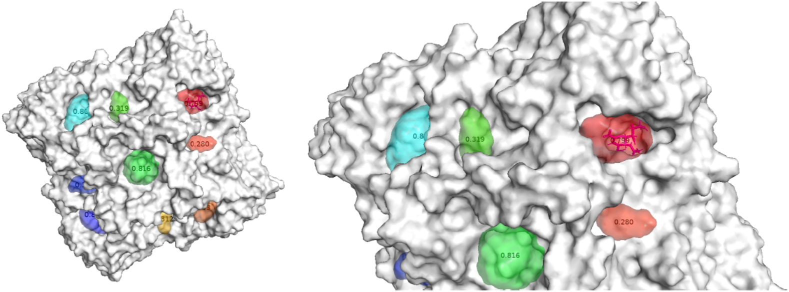
[5] Dompé Farmaceutici SpA

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.

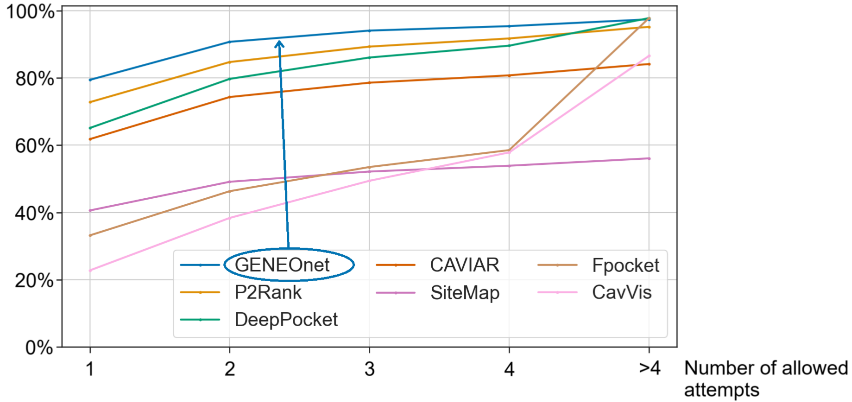


a) b)

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

Percentage of correct answers when allowing n=1,2,3,4 attempts.



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

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.

Basic idea

How can we transform our informal idea into a precise mathematical model?

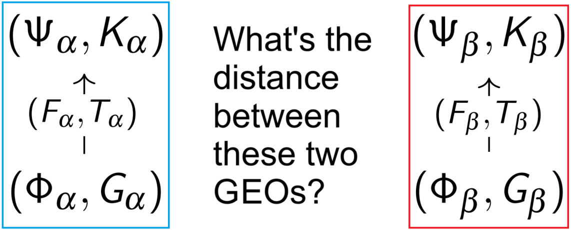
Let us begin by formalizing property 1.

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.

An extended pseudo-metric for *ALL* GEOs

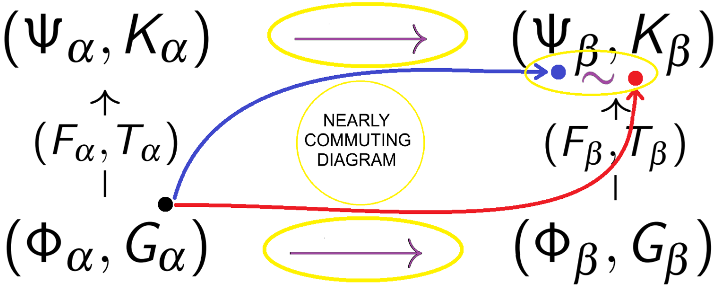
We have to introduce a pseudo-metric between GEOs that remains well-defined even when the GEOs operate on different domains and produce outputs in distinct codomains. This is a non-trivial challenge.



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 there exist two horizontal GENEOs that make this diagram nearly commutative, with the same holding true in the opposite direction:

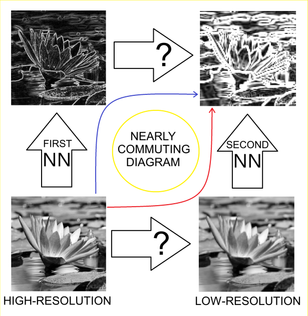


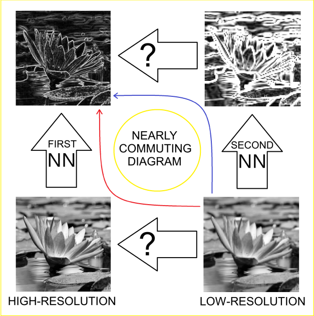
$$\begin{array}{ccc} (\Psi_\alpha, K_\alpha) & \longrightarrow & (\Psi_\beta, K_\beta) \\ \uparrow (F_\alpha, T_\alpha) & & \uparrow (F_\beta, T_\beta) \\ (\Phi_\alpha, G_\alpha) & \longrightarrow & (\Phi_\beta, G_\beta) \end{array}$$

We can measure the non-commutativity of each diagram by a **cost function**.

An example

Suppose we have two neural networks for edge detection in images, represented as GEOs.



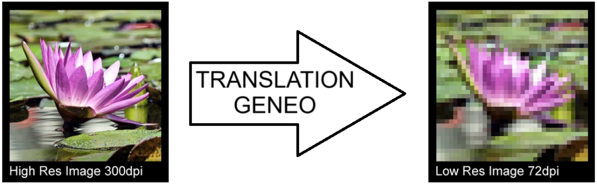


The two neural networks are considered close if there exist two pairs of horizontal GENEOs that make these diagrams nearly commutative.

An extended pseudo-metric for *ALL* GEOs

To formalize our new pseudo-metric d_E between GEOs, let us consider the category $\mathbf{S}_{all}$ 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 *translation GENEOs*. These morphisms describe the possible "logical correspondences" between data represented by different perception pairs.

For example, a translation 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. Therefore,

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

To proceed with the definition of our pseudo-metric on $\mathscr{G}$, we need to specify which logical correspondences between data we consider admissible. To this end, let us consider a small subcategory $\mathbf{S}$ of the category $\mathbf{S}_{all}$.

$\mathscr{G}$ will be the set of GEOs/functional agents where we will define our pseudo-metric, while the morphisms in **S** will be the translation GENEOs considered admissible.

An extended pseudo-metric for *ALL* GEOs

We can formalize the new pseudo-metric d_E on $\mathscr{G}$ as the infimum, over all admissible GENEOs, of the maximum between:

- the cost of the pair of GENEOs from GEO1 to GEO2;
- the cost of the pair of GENEOs from GEO2 to GEO1.

$$d_E = \inf_{\substack{(L,P)\\(M,Q)\\(L',P')\\(M',Q')}} \max\left(cost \begin{array}{ccc} (\Psi_\alpha, K_\alpha) & \xrightarrow{(M,Q)} & (\Psi_\beta, K_\beta) \\ \uparrow (F_\alpha, T_\alpha) & & \uparrow (F_\beta, T_\beta) \\ (\Phi_\alpha, G_\alpha) & \xrightarrow{(L,P)} & (\Phi_\beta, G_\beta) \\ \text{GEO1} & & \text{GEO2} \end{array} ,\ cost \begin{array}{ccc} (\Psi_\beta, K_\beta) & \xrightarrow{(M',Q')} & (\Psi_\alpha, K_\alpha) \\ \uparrow (F_\beta, T_\beta) & & \uparrow (F_\alpha, T_\alpha) \\ (\Phi_\beta, G_\beta) & \xrightarrow{(L',P')} & (\Phi_\alpha, G_\alpha) \\ \text{GEO2} & & \text{GEO1} \end{array} \right)$$

An extended pseudo-metric for *ALL* GEOs

Proposition

d_E 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 $d_E((F_\alpha, T_\alpha), (F_\beta, T_\beta))$ measures the *cost* of changing (F_α, T_α) into (F_β, T_β).

When $d_E((F_\alpha, T_\alpha), (F_\beta, T_\beta))$ 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.

Back to the basic idea of explanation

Let us recall our informal idea.

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.

The formalization of 1 is completed using the pseudo-metric d_E.
How about the formalization of 2?

Complexity of GEOs

Let us assume a (possibly infinite) set $\Gamma = \{(F_i, T_i) : (\Phi_i, G_i) \to (\Psi_i, K_i)\}$ of GEOs is given. We will say that Γ is our internal library.
For each GEO $(F_i, T_i) \in \Gamma$ we arbitrarily choose a value c_i representing the complexity $\mathrm{comp}((F_i, T_i))$ of (F_i, T_i).

Let us now consider the closure of Γ, i.e., the minimal set $\bar{\Gamma}$ such that

- $\bar{\Gamma} \supseteq \Gamma$;
- $\bar{\Gamma}$ is closed under composition (i.e., if $(F, T), (F', T') \in \bar{\Gamma}$ are composable, then $(F', T') \circ (F, T) \in \bar{\Gamma}$);
- $\bar{\Gamma}$ is closed under direct product (i.e., if the GEOs $(F, T), (F', T') \in \bar{\Gamma}$, then $(F, T) \otimes (F', T') \in \bar{\Gamma}$).

The complexity of each GEO in $\bar{\Gamma}$ is the **minimal total cost** of building such an operator through compositions and direct products.

Back to the basic idea of explanation

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:

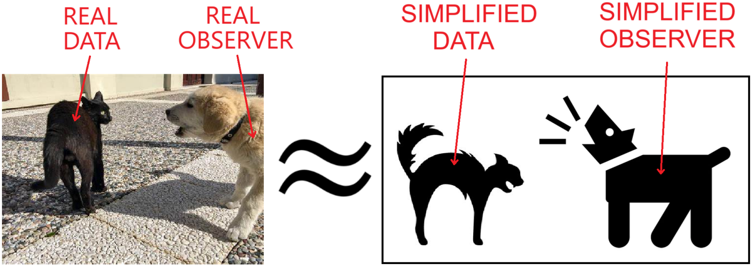
1. C perceives A and B as similar to each other;
2. C perceives B as less complex than A.

The formalization of this idea is now complete.

We believe that this general mathematical model of explainability, based on precise operator theory, could benefit XAI.

A mathematical concept of explanation

In summary, the pseudo-metric d_E 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 of complexity less than k represented by a GEO (F_β, T_β) when $d_E((F_\alpha, T_\alpha),(F_\beta, T_\beta)) \leq \varepsilon$.



Intelligence and contradiction

Representing observers as GENEOs results in another key implication, which can be interpreted as a form of "principle of contradiction".

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 10 (2009) 297–315

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Does intelligence imply contradiction?

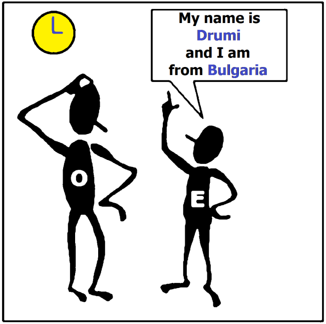
Action editor: Vasant Honavar

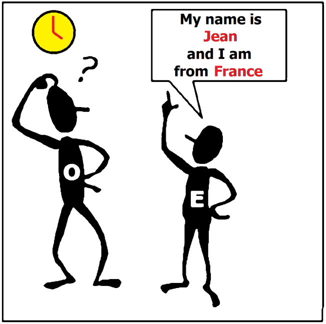
P. Frosini *

Department of Mathematics and ARCES, University of Bologna, I-40126 Bologna, Italy

Intelligence and contradiction

What do we mean by contradiction?





An entity E is said to be contradictory for an observer O if it reacts differently at different times under the same internal and external conditions, according to O's judgment.

Intelligence and contradiction

In an appropriate framework, this statement can be proven:

Every sufficiently intelligent entity is contradictory

Ludwig Josef Johann Wittgenstein

Tractatus
Logico-Philosophicus

By
LUDWIG WITTGENSTEIN

With an Introduction by
BERTRAND RUSSELL, F.R.S.

NEW YORK
HARCOURT, BRACE & COMPANY, INC.
LONDON: KEGAN PAUL, TRENCH, TRUBNER & CO., LTD.
1922

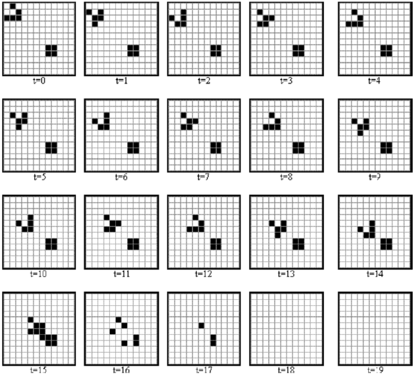
Intelligence and contradiction

Equivalently, we can say that

> The behavior of any sufficiently intelligent entity is unpredictable.

Intelligence and contradiction

How can we prove that?
We can use an approach based on cellular automata.



https://playgameoflife.com/

Intelligence and contradiction

Sketch of proof:

- An observer is identified (understood as a GENEO that transforms the functions representing the states of the cellular automaton into functions describing the perceived entity and its surrounding environment).
- The intelligence of an entity is defined as its ability to survive in the environment, as judged by the observer.
- It is shown that there exists a threshold for intelligence (dependent on the number of states the observer can associate with the entity and its environment), beyond which the observed entity necessarily appears contradictory to the chosen observer.

In this model, contradictoriness and unpredictability are not limitations of intelligent structures but necessary conditions for developing complex intellectual behaviors.

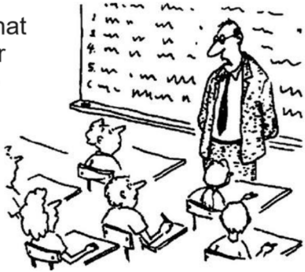
Intelligence and contradiction

Theorem. Let E be an entity with a finite lifespan and assume that its environment is deterministic. If the intelligence of E is greater than the product of the cardinalities of the sets P_{ent} and P_{ENV} the entity must necessarily be contradictory.

ATTENTION! The theorem does not assert that intelligent entities must change their behavior (this fact is obvious) but that they must do so without the observer understanding why.

P_{ent} = set of states of E recognized by the observer.

P_{ENV} = set of environmental states recognized by the observer.



"I expect you all to be independent, innovative, critical thinkers who will do exactly as I say!"

Intelligence and contradiction

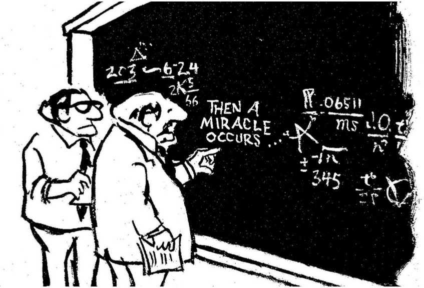
A precise formulation of this approach can be found here:

P. Frosini, Does intelligence imply contradiction?, Cognitive Systems Research, vol. 10 (2009), n. 4, 297-315.

(A synthetic and beautiful slideshow of this paper has been made by Mattia G. Bergomi. It is available at the link *https://mgbergomi.github.io/Contradiction/*.)

Intelligence and contradiction

According to our mathematical framework based on GENEOs, an agent A appears unpredictable to a fixed observer if the *"intelligence"* of A exceeds a threshold determined by the product of the number of states the observer can perceive in the agent and its environmental context. This implies that, to achieve predictability of behavior, it is necessary to choose models where the aforementioned threshold is greater than the desired intelligence value.



Summary

To sum up, GENEOs are novel mathematical tools designed to approximate equivariant neural networks using a compositional approach. They are particularly useful when prior knowledge about the expected behavior of the neural network is available. GENEOs are generally interpretable, making them potentially beneficial for explainable artificial intelligence (XAI) and helpful in elucidating certain properties of intelligent systems.

THANKS FOR YOUR ATTENTION!