Vector Space Representations of Equational Theories

Luis Berlioz
Universidad Nacional
Autónoma de Honduras
luis.berlioz@unah.edu.hn

Paul-André Melliès
Université Paris Cité
CNRS, INRIA
mellies@irif.fr

1 Introduction

Building on the collaborative Equational Theories project initiated by Terence Tao fifteen months ago [1], and combining it with ideas coming from machine learning, finite model theory, and mathematical logic, we construct a *latent space of equational theories* (ETs) where each ET is embedded at a specific point determined by its statistical behavior with respect to a large sample of finite magmas. This experiment enables us to observe, for the first time, how implications flow, producing surprisingly oriented and well-structured chains of logical implications in the latent space of equational theories. The Equational Theories project describes *exhaustively* the full implication preorder $(G, \Rightarrow)$ between equations defined with *at most four instances* of the binary connective $\diamond$ on a given magma $(A, \diamond)$. This includes fundamental equational theories such as the *associativity law* (numbered as $\mathsf{Eqn}[4512]$ in the ETs project) which involves three variables x, y, z and four instances of the connective:

$$\forall x, y, z \in A,\ x \diamond (y \diamond z) = (x \diamond y) \diamond z,$$

the *commutativity law* ($\mathsf{Eqn}[43]$) $\forall x, y \in A, x \diamond y = y \diamond x$, with two variables x, y, and two instances of the connective and the *idempotence law* ($\mathsf{Eqn}[3]$) with one variable x and one instance of the connective: $x = x \diamond x$. One methodological benefit of the exhaustive approach advocated by Tao is that other, more exotic and largely unexplored, equational theories suddenly appear as meaningful. An example is the relationship between the *Obelix law*, $x = (y \diamond x) \diamond (y \diamond (y \diamond x))$ and the *Asterix law*, $x = (y \diamond (x \diamond (y \diamond x)))$. Indeed, the Asterix law implies the Obelix law for all *finite magmas*, but not for infinite ones.

Using the concept of *Stone pairing* [2, 3, 4], we experiment with the idea that mathematical concepts (in this case, equational theories) can be located as points in a latent space. The *Stone pairing* between an equational theory φ and a finite magma $(A, \diamond)$ is defined as follows:

$$\langle\, \varphi(\mathbf{x}) \mid A \,\rangle = \frac{\#\left\{\, (a_1, \ldots, a_n) \in A^n \mid \varphi(a_1, \ldots, a_n) \text{ holds} \,\right\}}{\# A^{\,n}}$$

where $\#A$ is the number of elements in A and φ is the formula corresponding to an equational theory in n variables. By way of illustration, observe that whenever $A \vDash \forall \mathbf{x}, \varphi(\mathbf{x})$ we have that $\langle\, \varphi(\mathbf{x}) \mid A \,\rangle = 1$. It follows in particular that

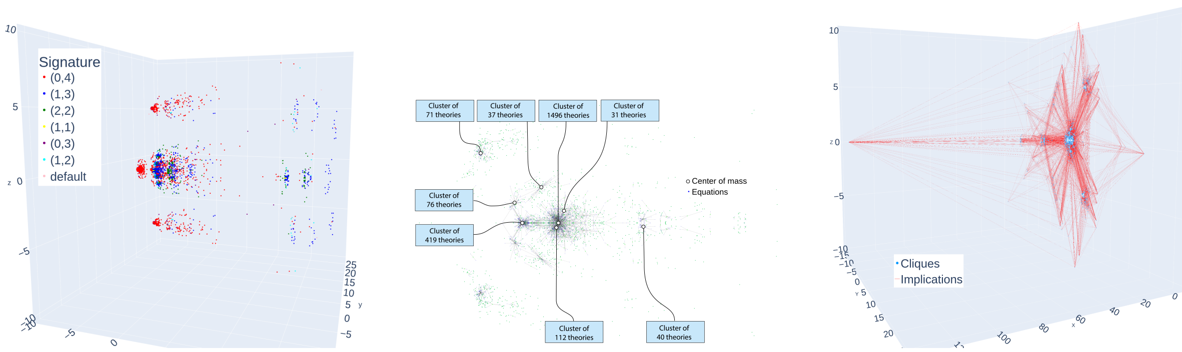
$$\langle\ x = x \mid A\ \rangle \ = \ 1 \quad \text{and} \quad \langle\ x = y \mid A\ \rangle \ = \ \frac{1}{\#A}$$

for every finite magma A, since there is a probability of one divided by the cardinality of A that two elements $a_1, a_2 \in A$ happen to be equal. Once the latent space of equational theories has been defined and constructed, we derive an *implication graph* $(G, \rightsquigarrow)$ from the full implication preorder $(G, \Rightarrow)$ made available online by the ET project.

Despite its simple technical and conceptual design, this empirical study at the crossroad of mathematical logic and machine learning reveals the existence of a well-organized hidden structure of the equational theories and the theorems between them. We believe that the discovery of this emerging piece of mathematical landscape and the exploration of its remarkable geometric and statistical properties provide a compelling demostration of the benefits of developing an experimental approach to logic.

2 The feature and latent spaces of equational theories

In order to generate a large number of magma features we proceed in the following way. We start by generating a large number $n = \#\mathsf{Mags}$ of finite magmas $A_1, \ldots, A_n$ of a fixed size N. Each magma A_ℓ for $1 \leq \ell \leq n$ is thus generated



(a) (b) (c)

Figure 1: In all the images, the first, second and third principal components are identified with the x, y and z axis respectively. (a) Scatter plot of all ETs, where every equation is colored according to its signature. (b) Visualizing the center of mass of each provably equivalent clique. Each ET is drawn connected to its corresponding center of mass. (c) View of the latent space of the 1415 vertices and 4824 edges of the implication graph $(G, \rightsquigarrow)$ modulo reversibility. An interactive version of the images is available ***here***.

as a two-dimensional array of size $N \times N$, where each entry is a randomly chosen value between 0 and $N-1$. This two-dimensional array describes the multiplication table of the magma $A_\ell = \{0, \ldots, N-1\}$ and thus completely characterizes it. We typically work with magmas of size $4 \leq N \leq 18$, and samples of sizes n between 20 000 and 100 000. The complete code for this is provided in the project's repo.

Next, we compute the Stone pairing of each random magma and each of the $\#\mathsf{Eqns} = 4694$ equations. The feature dataset is a two-dimensional array whose components are the Stone pairings $\langle\, \varphi_i \mid A_\ell \,\rangle$ for $1 \leq i \leq \#\mathsf{Eqns}$ and $1 \leq \ell \leq \#\mathsf{Mags}$.

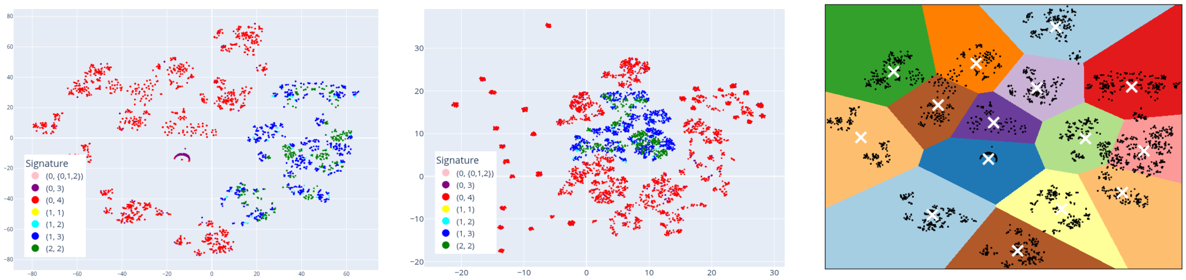
We apply principal component analysis (PCA) to the feature array in order to obtain a low dimensional vector representation of each ET. Mostly, we use three principal components for visualization purposes, but this value achieves an explained variance ratio of more than 80%. The PCA method produces the linear projection onto the subspace with the maximal variability. We call the resulting projection the *latent space of equational theories*.

3 Empirical properties of the latent space

By exploring the latent space and examining the positions of the 4694 equational theories, we observe that the latent space is surprisingly good at identifying and grouping ETs. In Fig. 1(a), for example, we color each equation according to its *signature*, i.e., the ordered pair containing the number of connectives on the left and right hand side of the equation. ETs with the same signature clearly cluster together, and the same is true for many other attributes of ETs. We also experimented with other common dimensionality reduction techniques, in Fig. 2 the plots produced with `t-SNE` and `UMAP` are shown using the same coloring by signature as in Fig. 1(a). Nevertheless, PCA conveys the most information about the structure of the data by unveiling properties like directionality, symmetry and preserving most of the original Euclidean distances.

Another interesting property of the latent space is given by provably equivalent ETs. These form equivalence classes (or cliques) as opposed to ETs for which only a strict implication holds. Fig. 1(b) shows all classes of provably equivalent ETs. To aid visualization, we also plot the center of mass of each class and all the members of the clique connected to it. We observe that logically equivalent ETs are seven times closer together on average than strict implications with respect to their Euclidean distance.

Even more striking is the evident planar symmetry of the ETs in latent space. We identified that dual pairs of ETs have approximately opposite values in their second component, i.e., for any pair of dual ETs, their respective latent space vectors are given by (x_1, x_2, x_3) and $(x_1, -x_2, x_3)$. The dual of an ET is the result of reverting every operation and renaming the variables, for example, the dual of $x = x \diamond y$ is $x = y \diamond x$ (see [1], pg. 10). Self-dual ETs have a second



(a) (b) (c)

Figure 2: Common dimensionality reduction methods for the same data as in figure 1. Each point is colored according to the signature of the corresponding ET. (a) Two-dimensional t-SNE plot produced using the scikit-learn implementation with `components` and `perplexity` parameters set to 2 and 30.0 respectively. (b) UMAP embeddings using the `umap-learn` package, with parameters `n_neighbors=10` and `min_dist=1.0`. (c) K-means clustering of the t-SNE reduced data, with number of clusters set to 15. Centroids are marked with a white X.

component of zero, as is to be expected. This motivated us to include the transpose of each random multiplication table in the training dataset. Hence, the Stone pairs of each ET and its dual are sampled in exactly the same values. This modification caused the pairs of dual equations to be exactly symmetrical up to floating point accuracy.

Implications $\mathsf{Eqn}[j] \implies \mathsf{Eqn}[k]$ are visualized as arrows connecting two ETs, between the vertices associated to $\mathsf{Eqn}[j]$ and $\mathsf{Eqn}[k]$ in the latent space. There is a strong tendency for true implications to point in the direction of the first principal component: almost 87% of the more than 8 million true implications form an acute angle with the first component. In Fig. 1(c), we present the resulting view of all the implications among equivalence cliques. This strong directionality prompted us to predict the truth value of an implication using deep learning with the principal components as input data. We used the same coupling method described in [1] pg. 58, where two randomly picked ETs, $\mathsf{Eqn}[j]$ and $\mathsf{Eqn}[k]$ are placed together in an ordered pair. Their first d principal components are joined to form a vector of length $2d$ to form the features of the model. The first and second halves represent the hypothesis and the conclusion of the implications respectively, and the labels are the truth values of the implication $\mathsf{Eqn}[j] \implies \mathsf{Eqn}[k]$. We were able to match the accuracy statistic reported by the ET paper by training on the first eight principal components (we obtained, accuracy: 0.9931, precision: 0.9866, F1-score: 0.9908). The model consisted of a one-dimensional ResNet with a convolutional one-dimensional block implemented in Pytorch[1].

The characteristic shape that emerges in the latent space is very persistent. It is recognizable even after removing a random sample of 80% of the 4694 ETs. Also, the shape is approximately preserved and the volume increases as the number of magmas increases. The reverse phenomenon is observed when sampling a constant number of magmas of increasing size. In future work we plan to investigate this phenomenon in greater depth, explaining its peculiar geometry, adding equations with more than four connectives, decomposing ETs based on their geometric properties and even including formulas with different quantifiers.

References

[1] Matthew Bolan, Jose Brox, Mario Carneiro, Martin Dvorak, Andres Goens, Harald Husum, Zoltan Kocsis, Alex Meiburg, Pietro Monticone, David Renshaw, et al. The equational theories project: Using lean and github to complete an implication graph in universal algebra. Technical report, Technical report, 2024.

[2] Mai Gehrke, Tomáš Jakl, and Luca Reggio. A duality theoretic view on limits of finite structures: Extended version. *Logical Methods in Computer Science*, Volume 18, Issue 1, January 2022.

[3] Jaroslav Nešetřil and Patrice Ossona de Mendez. A model theory approach to structural limits. *Commentationes Mathematicae Universitatis Carolinae*, 53(4):581–603, 2012.

[1]The training and eval loop was implemented with assistance from the Kimi K3 language model.

[4] Jaroslav Nešetřil and Patrice Ossona de Mendez. A unified approach to structural limits and limits of graphs with bounded tree-depth. *Memoirs of the American Mathematical Society*, 263(1272):0–0, January 2020.