A new computational fabric for Graph Neural Networks

Graph Neural Networks (GNNs) typically align their computation graph with the structure of the input graph. But are graphs the right computational fabric for GNNs? A recent line of papers challenges this assumption by replacing graphs with more general objects coming from the field of algebraic topology, which offer multiple theoretical and computational advantages.

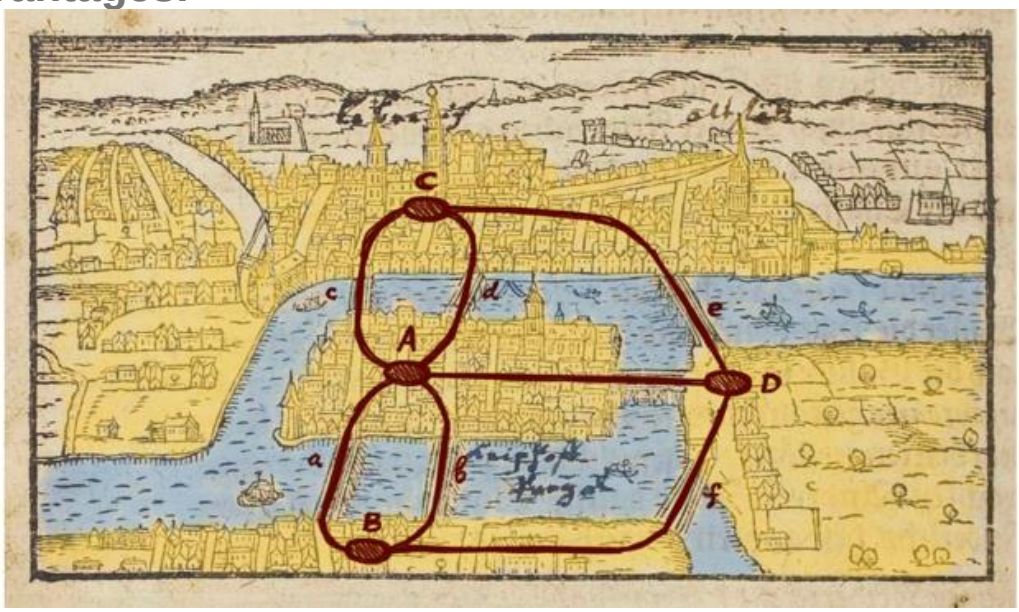


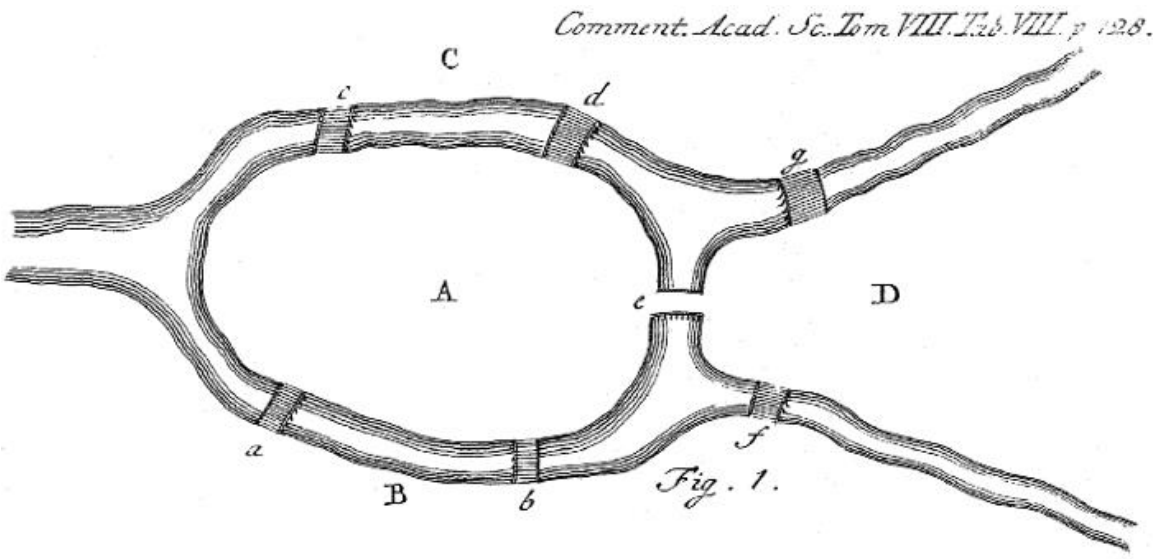
Image: based on Wikipedia.

This post was co-authored with Cristian Bodnar and Fabrizio Frasca and is based on the papers C. Bodnar, F. Frasca, et al., Weisfeiler and Lehman Go Topological: Message Passing Simplicial Networks (2021) ICML and C. Bodnar, F. Frasca et al., Weisfeiler and Lehman Go Cellular: CW Networks (2021)

NeurIPS. It is part of the series on [Graph Neural Networks through the lens of Differential Geometry and Algebraic Topology](). See also other posts from the series discussing [Neural Diffusion PDEs](), [graph rewiring with Ricci flows](), and [cellular sheaves]().

> "Topology! The stratosphere of human thought! In the twenty-fourth century, it might possibly be of use to someone." — Aleksandr Solzhenitsyn, In the First Circle *(1968)*

Graphs are used to model anything ranging from computer networks to particle interactions in the Large Hadron Collider. What makes graphs so ubiquitous is their discrete and combinatorial nature, allowing them to express abstract relations while remaining amenable to computations. One of the reasons for their popularity is the fact that graphs abstract out the geometry, i.e. where the nodes are positioned in space or how the edges are curved, leaving only a representation of how nodes are connected. The origins of graph theory stem from this very observation made by Leonhard Euler in 1741 in his work on *geometria situs* ("geometry of location") [1], in which he showed the famous problem of the Seven Bridges of Königsberg has no solution.

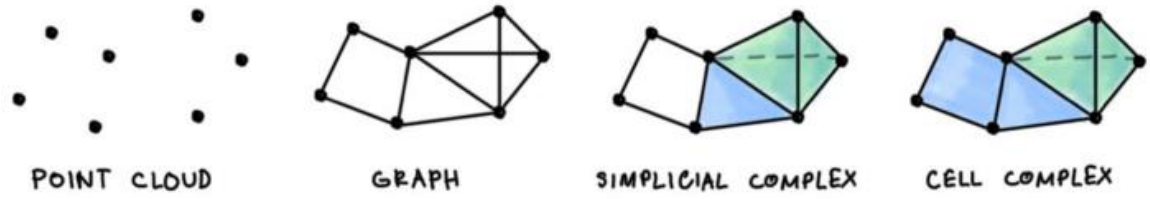


The Seven Bridges problem asks to find a cyclic walk through the city of Königsberg without crossing any bridge more than once. As Euler remarked, the exact shape of the city of Königsberg is irrelevant. What matters is how the different pieces of land (nodes of a graph) are connected to each other (edges). Euler showed that such a cycle exists if and only if all nodes have even degree. Only five of the original bridges survived until modern days. Image: Wikipedia.

Interestingly, Euler's discovery did not only mark the beginning of graph theory, but is also often regarded as the birth of topology. As with graphs, topologists are interested in those properties of a space that are independent of its particular shape or geometry [2]. The modern manifestation of these ideas emerged in 1895 with the "*analysis situs*", a seminal paper by Henri Poincaré [3]. His work ignited the interest in combinatorial descriptions of manifolds from which topological invariants could be more easily found and computed.

Leonhard Euler (1707–1783) and Henri Poincaré (1854–1912). Images: Wikipedia.

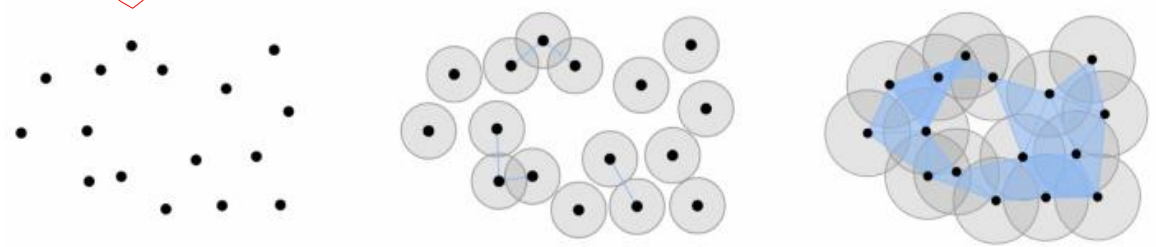
These combinatorial descriptions are known today as *cell complexes* [8] and can be thought of as higher-dimensional generalisations of graphs. Unlike graphs, which are formed of nodes and edges, cell complexes can also contain higher-dimensional structures or "cells": vertices are 0-cells, edges are 1-cells, 2D surfaces are 2-cells, and so on. To construct a cell complex, one can proceed hierarchically by glueing the boundary of a cell to other lower dimensional cells. In the particular case when the cells are formed of simplices (e.g. edges, triangles, tetrahedra, etc), these spaces are also called *simplicial complexes*.



Graphs can be seen as a set of vertices to which we attach edges (1-cells). Analogously, simplicial and cell complexes can be seen as graphs to which we attach 2-cells (shown in blue), 3-cells (green) and so on.

Topology in machine learning and data science

Contrary to the scepticism of Solzhenitsyn's character whose quote we chose to open this post [9], one did not have to wait for four hundred years to turn topology into a practical tool. Topological constructions such as simplicial complexes have been used in machine learning and data science under the umbrella of Topological Data Analysis (TDA), a class of methods that emerged in the 1990s attempting to analyse "the shape of the data" [4–7] in a way that is metric-insensitive and robust to noise. The roots of TDA can be traced back to the late 1920s, to the work of Leopold Vietoris [43], one of the most prolific topologists of the 20th century. However, these techniques had to wait for the birth of modern computing in order to be applied at scale.

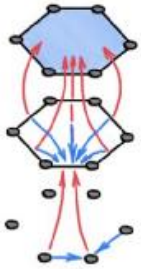


Given a point cloud, the intersections between closed balls of a fixed radius around each point yield a simplicial complex. By progressively increasing the radius of the balls, we obtain a nested sequence of simplicial complexes. Image: Bastian Rieck.

The workhorse of TDA is Persistent Homology (PH) [7], an approach for extracting topological features from point clouds. Given a dataset of points, PH creates a nested sequence of simplicial complexes, where each complex corresponds to a

certain scale at which the underlying point cloud is analysed. Then, it keeps track of the various topological features (e.g., connected components, cycles, or voids) that appear and disappear as the scale is progressively increased and one transitions from one complex in the sequence to the next. In the deep learning era, persistent homology has had a "second life" after it was shown that one can back-propagate through it, allowing thus to integrate the already established TDA apparatus in deep learning frameworks [10–17].

A more recent series of works has proposed a different use for simplicial and cell complexes in Geometric Deep Learning, as a richer underlying topological space to support the data and the computations performed on it. The first few works that took advantage of this perspective [18–22] proposed convolutional models as well as random walk approaches [42] operating on simplicial complexes. As shown in our papers [24,25], the convolutional models can be understood as specific instantiations of message passing on simplicial and cell complexes [23–25]. Because the computations are driven by the topological structure (i.e. the neighbourhood structure) of these spaces, we refer to this set of methods as *topological message passing*. In this framework, adjacent cells, potentially of different dimensions, are exchanging messages as depicted below.

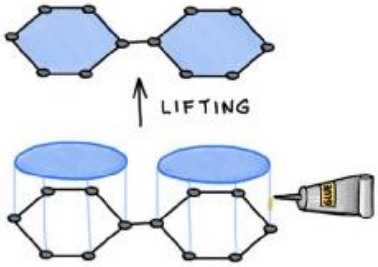


Sketch of topological message passing. Blue arrows depict "horizontal" information propagation between upper-adjacent cells, i.e. cells in the boundary of the same higher-dimensional cell. Red arrows depict "vertical" information propagation, whereby cells receive messages from lower-dimensional cells in their boundary. Summarising information from boundary cells into a coarser representation, this computation can be interpreted as a form of (differentiable) pooling.

Going beyond graphs in GNNs

Despite the rich structure provided by cell complexes, we cannot ignore that graphs are by far the most common topological objects in machine learning, and there are very few datasets that go beyond them. Nonetheless, we showed that one can still take advantage of these interesting topological spaces by transforming the input graph. We call the conversion of a graph into a higher-dimensional topological space a "lifting", in analogy to the eponymous concept from category theory. It is a transformation attaching higher-dimensional cells to the input graph by following certain rules. For instance, a graph can be lifted into a cell complex by attaching a higher-dimensional cell to each clique or cycle of the graph. By doing so, the graph is replaced with a different space that has more structure and can provide a better

computational fabric for GNNs than the original graph. In what follows, we discuss the specific advantages of this approach.

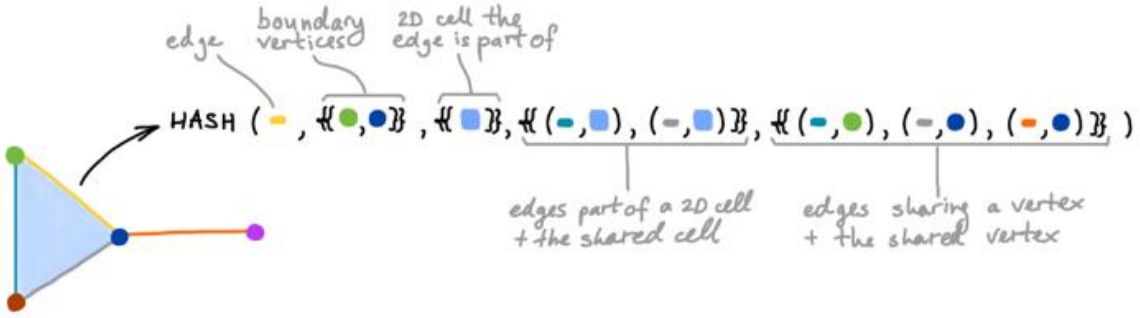


A higher dimensional cell complex can be constructed from a graph by (for instance) glueing the boundary of a 2D closed disk to an induced cycle in the graph.

Higher-order features and structures. GNNs typically adopt a node-centric view, where the data residing on the edges are treated only as auxiliary information augmenting the communication between the vertices. In topological message passing, *all* cells are first-class citizens. Irrespective of their dimension, they are assigned a specific representation that is evolved by exchanging messages with adjacent cells. This provides a recipe for explicitly modelling certain higher-order structures and the interactions between them. In particular, it provides a principled method to evolve the edge (i.e. 1-cell) features of the input graph, something that a large class of GNN models do not consider.

Higher-order interactions. Graphs are by definition *dyadic* ("pairwise") and cannot represent relations and interactions involving more than two objects [26]. This could be a problem when modelling complex systems characterised by

higher-order interactions: for instance, three reactants in a chemical reaction could interact concomitantly. In a cell complex, such situations could be encoded by connecting the reactants by a 2-cell (i.e. a “filled” triangle). Consequently, the computational flow of the model adapts to the presence of the higher-order interaction.

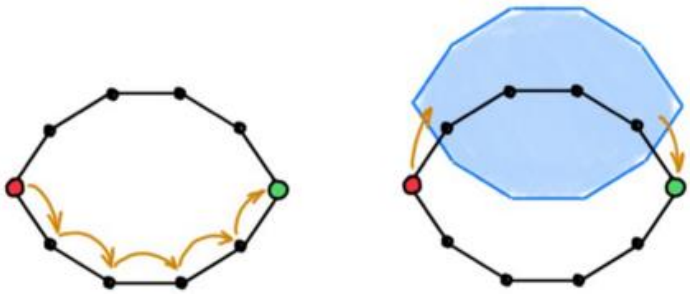


The Cellular Weisfeiler-Lehman (CWL) test [24,25] extends the classic WL test [27] to cell complexes. Each step of the algorithm perfectly hashes the colours of the neighbouring cells (potentially of different dimensions).

Expressive power. The expressive power of message passing GNNs is bounded by the Weisfeiler-Leman (WL) graph isomorphism test [27–29]. It is known that WL cannot detect certain graph substructures like triangles or cycles failing to distinguish between even very simple non-isomorphic graphs. In our papers [24,25], we introduced a cellular version of the WL test (CWL) that can be used to test the isomorphism of cell complexes. When this new test is paired with a graph lifting procedure like the one described above, we showed it can distinguish larger classes of graphs than the WL test. Under certain conditions, the topological message passing procedures

we proposed inherit the benefits of this test, leading to improved expressive power compared to standard GNNs [24,25,30].

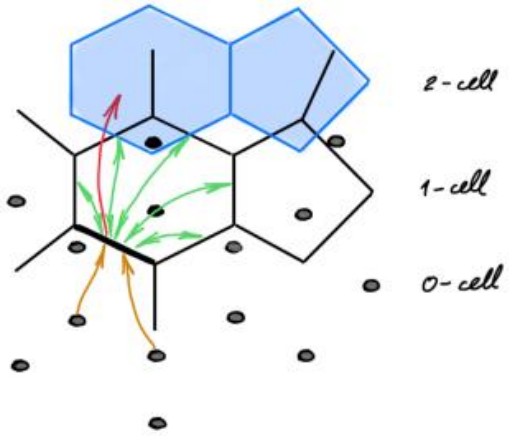
Under-reaching, oversmoothing and bottlenecks. Message Passing GNNs require n layers to make nodes that are n-hops away communicate. Therefore, when using just a few layers, nodes that are far apart cannot exchange messages, a phenomenon known as *under-reaching* [31,32]. In contrast, using too many layers could lead to *oversmoothing* [33,34] and messages could get lost in structural *bottlenecks* [32] of the graph. Cell complexes may alleviate these problems because the richer neighbourhood structure induced by the higher-dimensional cells creates shortcuts between nodes that can be very far apart. Therefore, information is propagated efficiently with a contained number of computational steps.



GNNs need many layers to make nodes that are far away communicate (left). Higher-dimensional cells change the underlying topological structure of the space by creating shortcuts (right). This allows distant nodes to exchange information in just a few message-passing steps.

Hierarchical modelling. The computations performed by topological message passing are naturally hierarchical with

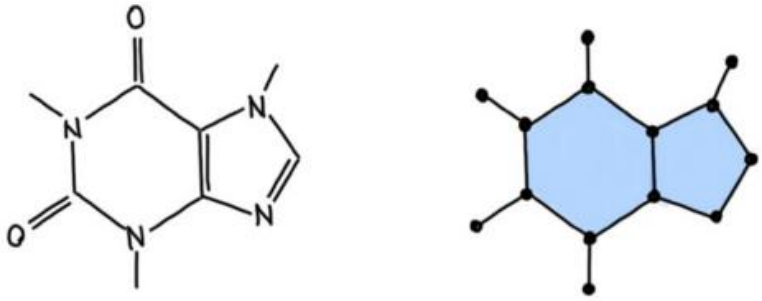
information flowing from lower dimensional cells to higher-dimensional cells and back. This could be seen as a form of "vertical" (and differentiable) pooling, as opposed to "horizontal" pooling in standard graph neural networks. This maintains the inductive bias of "compressing" regions of the graph without disregarding the fine-grained information of the input graph, which often hurts the performance of pooling-based GNNs.



Topological message passing allows the information to hierarchically between cells of different dimension.

Domain alignment. Certain applications naturally align with the structure of cell complexes. For instance, the atoms, bonds and chemical rings of a molecule can be represented as 0-cells, 1-cells, and 2-cells. The direct correspondence between the physical structure of the molecule and the cell complex representation naturally allows topological message passing to take advantage of the aforementioned properties. Using these representations, we showed that topological message passing achieves state-of-the-art results in molecular property prediction tasks [25]. Other well-aligned applications could include discrete manifolds (meshes) in

computer graphics applications, social networks (where cliques are particularly important), or spatial graphs such as Google maps (where the blocks between streets can be naturally represented as “cubical” cells).



The caffeine molecule is modelled as a two-dimensional cellular complex.

Where Topology and Differential Geometry meet

Topological message passing preserves many interesting connections to algebraic topology and differential geometry, allowing to exploit mathematical tools that have so far been under-explored in graph- and geometric deep learning.

The algebra of holes and orientation equivariance. In algebraic topology, one typically works with *oriented* simplicial complexes in which there is an arbitrary “orientation” in each simplex. For instance, for each edge, we pick a source and a target node and for each triangle, we pick an order to traverse its nodes. Once an orientation is chosen, we can perform interesting algebraic operations on the complex such as computing the boundary of certain simplices via a “boundary operator”. These

algebraic manipulations can also be used to find “holes” in the simplicial complex — regions with no boundary that are not on the boundary of something else. Under the hood, Persistent Homology relies on these computations to detect topological features.

The boundary operator applied to a 2-simplex yields a triangle. Applying the operator again to the triangle, yields zero because the triangle is a loop and therefore, it does not have a boundary.

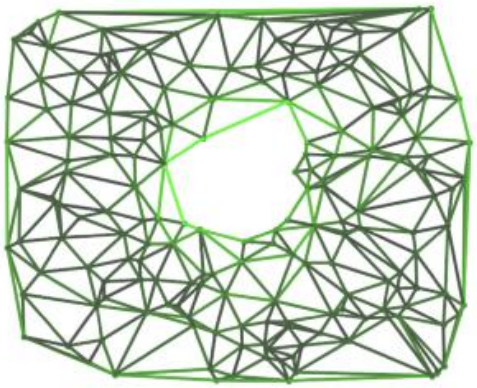
Topological message passing can be seen as a (non-linear) generalisation of algebraic operators such as the boundary operator. Therefore, it is necessary for topological message passing to behave similarly: we would like the outputs of our layers to “congruently” respond to changes in the orientation of the input complex. In other words, we would like our layers to be *orientation equivariant*. In our work [24], we studied how topological message passing can satisfy this property by picking the right kind of non-linearities and message passing functions. Concurrently, this has also been examined in a pure convolutional setting [20].

Distinguishing topological spaces. One of the first known topological invariants, the Euler characteristic [8,41], was originally used for the classification of platonic solids. We can define it as the alternating sum of the number of cells in each

dimension. It might come as surprising that if two cell complexes are homeomorphic, these sums will be identical, even if they are very different discretisations of the same space.

An interesting fact is that the readout operation of a topological message passing model can easily compute this topological invariant since it applies a permutation-invariant reduction over the cells of each dimension. Therefore this type of model is equipped by construction to distinguish between certain spaces that are *not* homeomorphic (i.e. have a different Euler characteristic). From a computational perspective, this can be seen as a generalisation of the WL test, where we are not simply interested in determining if two cell complexes are strictly the same, but if they are homeomorphic to each other.

Discrete Hodge Theory provides a more geometric interpretation for the topological properties of cell complexes. When the sign of the features associated with the *k*-cells depends on the orientation of the k-cell, the features can be seen mathematically as a discrete version of differential k-forms from differential geometry (i.e. k-dimensional volume elements that can be integrated) [45]. A Laplacian operator generalising the graph Laplacian, known as the *Hodge Laplacian* [21,44], can act upon these differential forms. It can be shown that the diffusion PDE based on this Laplacian converges in the limit to a signal that is related to the holes of the complex [44].



The diffusion PDE based on the Hodge Laplacian converges in the limit to the projection of the initial differential form onto the kernel of the Laplacian. This image shows how the zero-eigenvector of the Hodge Laplacian takes high values around the hole in the complex. Image: Andrei Popescu.

The first simplicial neural network models [18–20] are in fact convolutional models based on the Hodge Laplacian, which in turn have been inspired by topological signal processing [21,22]. More recently, convolutional models based on a version of this operator were used for solving NP-hard problems in computational algebraic topology [35].

Final thoughts

Are these just graphs in disguise? Recent papers [36,37] have argued that, amongst others, topological message passing approaches are nothing but a message passing GNN operating on a modified graph encoding the structure of the cell complex. This is certainly true for convolutional models, whose messaging passing computations involve pairs of cells. However, in its most general form, the message function allows higher-dimensional

cells to modulate the messages passed between lower-dimensional cells on their boundary. This is (in general) no longer possible via regular message passing on a graph because an edge connects exactly two nodes, while a 2-cell can connect an arbitrarily large number of edges.

In either case, the computations are driven by the topology of the underlying space that the data is attached to. We believe that adopting this topological perspective on message passing has benefits that go beyond purely computational considerations. Besides the valuable mathematical connections, it opens the research discourse to other mathematical and computational disciplines and favours positive cross-pollination between our often too monadic communities.

What is next for topological message passing? We anticipate two major future directions for topological message passing approaches. First, many of the architectures developed in GNNs over the years (such as attention mechanisms [38,39]) will likely be adopted in these new topological spaces, while making use of their specific characteristics. Second, additional mathematical objects and tools from the field of algebraic topology (including constructions such as cellular sheaves [40] that might sound exotic even to the most math-savvy ML researchers) will be adopted by the graph- and geometric deep learning communities. Such methods can both provide answers

to old problems as well as help address new ones, or, as Robert Ghrist has put it: “novel challenges necessitate novel math”.

[1] L. Euler, Solutio problematis ad geometriam situs pertinentis (1741), Enestrom 53, Euler Archive.

[2] Roughly, topology operates with the notion of a neighbourhood but not distance or angle, which are geometric constructions.

[3] H. Poincaré, Analysis situs (1895), Journal de l’École Polytechnique.

[4] P. Frosini, A distance for similarity classes of submanifolds of a Euclidean space (1990), Bulletin of the Australian Mathematical Society 42(3):407–415.

[5] H. Edelsbrunner et al., Topological persistence and simplification (2002), Discrete and Computational Geometry 28(4): 511–533.

[6] G. Carlsson, Topology and data (2009), Bulletin of the American Mathematical Society 46:255–308.

[7] H. Edelsbrunner and J. Harer. Computational Topology: An Introduction (2010), American Mathematical Society.

[8] A. Hatcher, Algebraic topology (2002), Cambridge University Press.

[9] In the original Russian: “Топология! Стратосфера человеческой мысли! В двадцать четвёртом столетии она, может быть, и понадобится кому-нибудь.” In the novel, this judgement is pronounced by the semi-autobiographic character Gleb Nerzhin, a mathematician struggling with the Faustian question of whether to work on applied research that would help Stalin’s regime. The author Aleksandr Solzhenitsyn was himself a mathematician by training, see a note on Robert Ghrist’s lecture from which we stole this quote.

[10] M. Horn, E. De Brouwer, et al. Topological Graph Neural Networks (2022), ICLR.

[11] R. Gabrielsson et al., A Topology Layer for Machine Learning (2020), AISTATS.

[12] M. Moor et al., Topological Autoencoders (2020), ICML.

[13] C. Hofer et al., Deep Learning with Topological Signatures (2017), NIPS.

[14] C. Hofer et al., Connectivity-Optimized Representation Learning via Persistent Homology (2019), ICML.

[15] C. Hofer et al., Learning Representations of Persistence Barcodes (2019), JMLR 20(126):1–45, 2019.

[16] C. Hofer et al., Graph Filtration Learning (2020), ICML

[17] C. Hofer et al., Topologically Densified Distributions (2020), ICML

[18] E. Bunch et al. Simplicial 2-complex Convolutional Neural Networks (2020), NeurIPS Workshop on Topological Data Analysis and Beyond.

[19] S. Ebli et al. Simplicial Neural Networks (2020), NeurIPS Workshop on Topological Data Analysis and Beyond.

[20] M. T. Roddenberry, N. Glaze and S. Segarra, Principled Simplicial Neural Networks for Trajectory Prediction (2021), ICML.

[21] S. Barbarossa et al. Topological Signal Processing over Simplicial Complexes (2020), IEEE Transactions on Signal Processing 68:2992–3007.

[22] S. Sardellitti et al. Topological Signal Processing over Cell Complexes (2021), Asilomar Conference on Signals, Systems, and Computers.

[23] M. Hajij et al. Cell Complex Neural Networks (2020), NeurIPS Workshop on Topological Data Analysis and Beyond.

[24] C. Bodnar, F. Frasca, Y. G. Wang, G. Montúfar et al. Weisfeiler and Lehman go Topological: Message Passing Simplicial Networks (2021), ICML.

[25] C. Bodnar, F. Frasca, et al. Weisfeiler and Lehman go Cellular: CW Networks (2021), NeurIPS.

[26] F. Battiston et al, The physics of higher-order interactions in complex systems (2021), Nature Physics 17:1093–1098.

[27] B. Weisfeiler and A. Leman, The reduction of a graph to canonical form and the algebra which appears therein (1968), Nauchno-Technicheskaya Informatsia 2(9):12–16. See a post about the relation of the WL test and GNNs. For an in-depth review (and historic notes) see C. Morris et al., Weisfeiler and Leman go Machine Learning: the story so far (2021) arXiv:2112.09992.

[28] K. Xu et al., How Powerful are Graph Neural Networks? (2019), ICLR.

[29] C. Morris et al., Weisfeiler and Leman Go Neural: Higher-order Graph Neural Networks (2019), AAAI.

[30] We note that orthogonal to our approach, a persistent version of the WL test was proposed by B. Rieck et al, A Persistent Weisfeiler-Lehman Procedure for Graph Classification (2019), ICML. At the same time, persistent homology was also used to make GNNs more expressive than the WL test [10].

[31] P. Barceló et al., The Logical Expressiveness of Graph Neural Networks (2020), ICLR.

[32] U. Alon and E. Yahav, On the bottleneck of Graph Neural Networks and its practical implications (2021), ICLR.

[33] K. Oono and T. Suzuki, Graph Neural Networks exponentially lose expressive power for node classification (2020), ICLR.

[34] C. Cai and Y. Wang, A note on over-smoothing for Graph Neural Networks (2020), arXiv:2006.13318.

[35] A. D. Keros et al., Dist2Cycle: A Simplicial Neural Network for Homology Localization (2021), AAAI.

[36] P. Veličković, Message passing all the way up (2022), ICLR Workshop on Geometrical and Topological Representation Learning.

[37] F. Jogl et al. Reducing learning on cell complexes to graphs (2022), ICLR Workshop on Geometrical and Topological Representation Learning.

[38] C. W. J. Goh et al. Simplicial Attention Networks (2022), ICLR Workshop on Geometrical and Topological Representation Learning.

[39] L. Giusti et al., Simplicial Attention Neural Networks (2022), arXiv:2203.07485.

[40] C. Bodnar, F. Di Giovanni et al., Neural Sheaf Diffusion: A Topological Perspective on Heterophily and Oversmoothing in GNNs (2022) arXiv:2202.04579. See also an accompanying post.

[41] L. Euler, Elementa doctrinae solidorum (1758), Novi Commentarii Academiae Scientiarum Petropolitanae:109–140.

[42] C. Hacker, k-simplex2vec: a simplicial extension of node2vec (2020), NeurIPS Workshop on Topological Data Analysis and Beyond.

[43] L. Vietoris, Über den höheren Zusammenhang kompakter Räume und eine Klasse von zusammenhangstreuen Abbildungen (1927), Mathematische Annalen 97:454–472.

[44] A. Muhammad and M. Egerstedt, Control using higher order Laplacians in network topologies (2006), International Symposium on Mathematical Theory of Networks and Systems.

[45] When restricted to edges, these signals have a nice interpretation as flows passing through a network. For instance, see T. Gebhart et al., Go with the Flow? A Large-Scale Analysis of Health Care Delivery Networks in the United States Using Hodge Theory (2021), IEEE Big Data 2021 as part of the Workshop on Applications of Topological Data Analysis to Big Data.

We are grateful to Ben Chamberlain and Bastian Rieck for proofreading this post. For additional articles about deep learning on graphs, see my other posts in Towards Data Science, subscribe to my posts and YouTube channel, get Medium membership, or follow me on Twitter.