Weisfeiler and Lehman Go Paths: Learning Topological Features via Path Complexes

Quang Truong and Peter Chin

Thayer School of Engineering, Dartmouth College

Table of contents

Motivation

Weisfeiler-Lehman Test

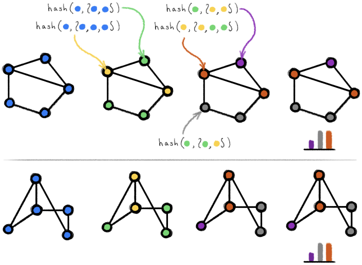


Figure 1: An illustration of the 1-WL test. Image is from M. Bronstein's blog [1].

1-WL test [22] is a simple algorithm to determine if two graphs are not isomorphic.

$$c_v^{(t+1)} = \textsc{Hash}\left(c_v^{(t)}, \left\{\!\!\left\{ c_w^{(t)} \mid w \in \mathcal{N}(v) \right\}\!\!\right\}\right)$$

[1] https://towardsdatascience.com/expressive-power-of-graph-neural-networks-and-the-weisefeiler-lehman-test-b883db3c7c49

If two graphs do not have the same histogram, they are not isomorphic.

However, the converse does not necessarily hold true. There exists a pair of graphs that are not isomorphic but still have the same color histogram.

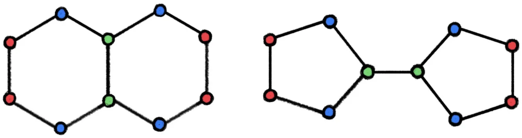


Figure 2: Two non-isomorphic graphs have a similar color histogram. Image is from M. Bronstein's blog [1].

[1] https://towardsdatascience.com/expressive-power-of-graph-neural-networks-and-the-weisefeiler-lehman-test-b883db3c7c49

Message-Passing Framework

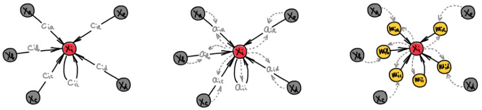


Figure 3: Different types of message passing. Image is from M. Bronstein's blog [2].

Message-passing framework [9] allows us to conceptualize Graph Neural Networks (GNNs) with propagated information from nodes along edges.

Vanilla GNNs are proven to be upper-bounded by 1-WL test [22] in terms of graph expressivity [23].

[2]https://thegradient.pub/graph-neural-networks-beyond-message-passing-and-weisfeiler-lehman/

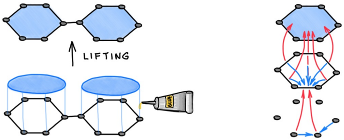


Figure 4: Lifting a graph to a regular cell complex and performing higher-order message-passing. Image is from M. Bronstein's blog [2].

In order to overcome the 1-WL test boundary, prior approaches incorporate topological structures in the message-passing procedure [2, 3].

MPSN [3]: Cliques $\Longleftrightarrow$ Simplices

CWN [2]: Cycles or Rings $\Longleftrightarrow$ 2-Cells

[2] https://thegradient.pub/graph-neural-networks-beyond-message-passing-and-weisfeiler-lehman/

A higher-order message-passing framework relies on relations that are not explicitly modeled by the vanilla message-passing framework.

Definition (Relations between members [2, 3])

For any member σ of K, there are four types of relations:

- Boundary $\mathcal{B}(\sigma) = \{\tau \mid \tau \prec \sigma\}$
- Co-boundary $\mathcal{C}(\sigma) = \{\tau \mid \sigma \prec \tau\}$
- Upper-adjacent neighborhood $\mathcal{N}_{\uparrow}(\sigma) = \{\tau \mid \sigma \prec \delta \ \wedge \ \tau \prec \delta\}$
- Lower-adjacent neighborhood $\mathcal{N}_{\downarrow}(\sigma) = \{\tau \mid \delta \prec \sigma \ \wedge \ \delta \prec \tau\}$

It is clear that we cannot lift graphs to higher-order spaces if certain substructures do not exist in the graphs.

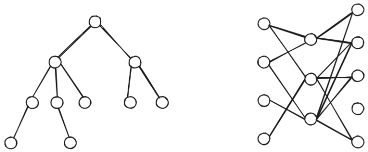


Figure 5: Examples of graphs without cliques, cycles, or rings.

Key problems:

1. A more generalized color refinement algorithm.
2. Theoretical connections with the current topological color refinement algorithms.
3. Practically effective and feasible.

Path Complex

Definition (Elementary path [12, 13])

Given a finite non-empty set V whose element is called vertex, an elementary p-path on set V is any sequence of vertices with length $p+1$. Elementary p-path is denoted by $e_{i_0 \dots i_p}$.

Definition (Boundary operator on elementary paths [12, 13])

Boundary operator on elementary $p-$paths is defined as:

$$\partial \, e_{i_0 \dots i_p} = \sum_{q=0}^{p} (-1)^q e_{i_0 \dots \hat{i}_q \dots i_p},$$

where $\hat{i}_q$ indicates the removal of the index i_q from the sequence $i_0 \dots i_p$.

Definition (Path complex [12, 13])

Given a finite non-empty set V, a path complex P is a non-empty collection of elementary paths such that for any sequence of vertices that belong to P, the truncated sequences, in which either the first vertex or the last vertex is removed, are also included in P.

We denote $P_p \subset P$ where P_p contains all paths with length p. Elements of P_p are called **allowed elementary p-paths**, while any sequences that do not exist in P_p are called **non-allowed elementary p-paths**.

Define $\mathcal{S}_p$ a space spanned by all simple paths with length p.

Define P a path complex with the highest dimension p such that for any dimension $k \leq p$, P_k contains all elementary k-paths that span $\mathcal{S}_k$, and boundary set of any elementary k-paths is restricted to elementary $(k-1)$-paths in $\mathcal{S}_{k-1}$.

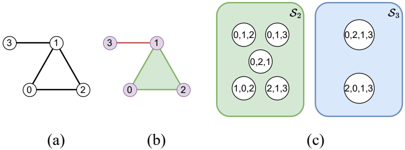
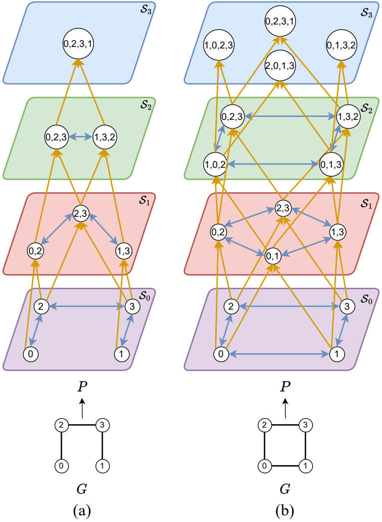


Figure 6: (a) Original graph; (b) Simplicial complex, which contains a 2-simplex, 4 1-simplices, and 4 0-simplices, arising from the original graph. (c) Simple path spaces $\mathcal{S}_2$ and $\mathcal{S}_3$ corresponding to the path complex arising from the original graph.



Despite its simplicity, the way we define Path Complex is sufficient to perform color refinement and generalize other topological Weisfeiler-Lehman tests.

Figure 7: Examples of path complexes arising from (a) a simple path with length of 3 and (b) a ring with size of 4. Blue arrows demonstrate upper-adjacent relations, while orange arrows demonstrate boundary relations.

Path Weisfeiler-Lehman Test

Theorem

PWL is at least as powerful as SWL [3] at distinguishing non-isomorphic graphs.

Theorem

PWL is at least as powerful as CWL(k-IC) [2] at distinguishing non-isomorphic graphs.

Corollary

PWL is strictly more powerful than WL at distinguishing non-isomorphic graphs.

Corollary

PWL is not less powerful than 3-WL at distinguishing non-isomorphic graphs.

Path Complex Networks

We can achieve maximal expressivity by extending GIN [23] to topological GNNs.

$$h_\sigma^{(t+1)} = \mathrm{MLP}^{(t)}_{\mathrm{UP},p}\left(m_{\mathcal{B}}^{(t)}(\sigma) \,||\, m_{\uparrow}^{(t)}(\sigma)\right)$$

$$m_{\mathcal{B}}^{(t)}(\sigma) = \mathrm{MLP}^{(t)}_{\mathcal{B},p}\left((1+\varepsilon_{\mathcal{B}})\, h_\sigma^{(t)} + \sum_{\tau \in \mathcal{B}(\sigma)} h_\tau^{(t)}\right)$$

$$m_{\uparrow}^{(t)}(\sigma) = \mathrm{MLP}^{(t)}_{\uparrow,p}\left((1+\varepsilon_{\uparrow})\, h_\sigma^{(t)} + \sum_{\substack{\tau \in \mathcal{N}_\uparrow(\sigma) \\ \delta \in \mathcal{C}(\sigma,\tau)}} \mathrm{MLP}^{(t)}_{M,p}\left(h_\tau^{(t)} \,||\, h_\delta^{(t)}\right)\right)$$

where σ is an elementary path (simplex for SIN [3] or cell for [2]).

Dataset	PROTEINS	NCI1	NCI109	IMDB-B
PK [18]	73.7 ± 0.7	82.5 ± 0.5	N/A	N/A
WL Kernel [19]	75.0 ± 3.1	86.0 ± 1.8 ⧫	N/A	73.8 ± 3.9
GSN [4]	76.6 ± 5.0	83.5 ± 2.0	N/A	77.8 ± 3.3 ⧫
pathGCN [7]	80.4 ± 4.2 ▲	83.3 ± 1.3	N/A	N/A
PathNN [16]	75.2 ± 3.9	82.3 ± 1.9	N/A	72.6 ± 3.3
SIN [3] †	76.4 ± 3.3	82.7 ± 2.1	N/A	75.6 ± 3.2 ●
CIN [2] †	77.0 ± 4.3	83.6 ± 1.4	84.0 ± 1.6 ●	75.6 ± 3.7
CAN [10]	78.2 ± 2.0	84.5 ± 1.6	83.6 ± 1.2	N/A
CIN++ [11]	80.5 ± 3.9 ⧫	85.3 ± 1.2 ▲	84.5 ± 2.4 ⧫	N/A
PIN (Ours)	78.8 ± 4.4 ●	85.1 ± 1.5 ●	84.0 ± 1.5 ▲	76.6 ± 2.9 ▲

Table 1: TUDataset Benchmarks [17]. The top-3 methods in each benchmark are denoted by ⧫ (1^{st} place), ▲ (2^{nd} place), and ● (3^{rd} place). Baselines are denoted by †.

Dataset	ZINC		OGBG-MOLHIV	
	No Edge Feat.	W/ Edge Feat.	Test ROC-AUC	Val. ROC-AUC
GCN [15]	0.469 ± 0.002	N/A	N/A	N/A
GAT [21]	0.463 ± 0.002	N/A	N/A	N/A
GatedGCN [5]	0.422 ± 0.006	0.363 ± 0.009	N/A	N/A
GIN [23]	0.408 ± 0.008	0.252 ± 0.014	77.07 ± 1.49	84.79 ± 0.68
PNA [6]	0.320 ± 0.032	0.188 ± 0.004	79.05 ± 1.32	85.19 ± 0.99
DGN [1]	0.219 ± 0.010	0.168 ± 0.003	79.70 ± 0.97	84.70 ± 0.47
HIMP [8]	N/A	0.151 ± 0.006	78.80 ± 0.82	N/A
GSN [4]	0.140 ± 0.006	0.115 ± 0.012	77.99 ± 1.00	**86.58 ± 0.84**
PathNN [16]	N/A	0.090 ± 0.004	79.17 ± 1.09	N/A
CIN [2] †	**0.115 ± 0.003**	0.079 ± 0.006	**80.94 ± 0.57**	N/A
CIN++ [11]	N/A	**0.077 ± 0.004**	80.63 ± 0.94	N/A
PIN (Ours)	0.139 ± 0.004	0.096 ± 0.006	79.44 ± 1.40	82.41 ± 0.96

Table 2: ZINC [20] and OGBG-MOLHIV [14] datasets. Bold texts indicate the best performance. Performance on ZINC is evaluated by Mean Squared Error, while performance on OGBG-MOLHIV is evaluated by ROC-AUC. Baseline is denoted by †.

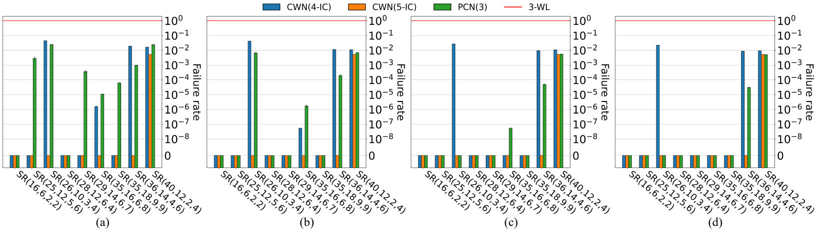


Figure 8: Failure rate comparison on SRG Families. (a) 3 message-passing (MP) layers. (b) 4 MP layers. (c) 5 MP layers. (d) 6 MP layers.

[4] found that counting 3-paths in a graph is not sufficient to distinguish non-isomorphic graphs in the SR experiment, but we found that message-passing between paths help us distinguish ALL strongly regular graphs in the experiment.

D. Beaini, S. Passaro, V. L'etourneau, W. L. Hamilton, G. Corso, and P. Lio'.
Directional graph networks.
In *International Conference on Machine Learning*, 2020.

C. Bodnar, F. Frasca, N. Otter, Y. Wang, P. Liò, G. F. Montufar, and M. Bronstein.
Weisfeiler and Lehman go cellular: CW networks.
In M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 2625–2640. Curran Associates, Inc., 2021.

- C. Bodnar, F. Frasca, Y. Wang, N. Otter, G. F. Montufar, P. Lió, and M. Bronstein.
  Weisfeiler and Lehman go topological: Message passing simplicial networks.
  In M. Meila and T. Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 1026–1037. PMLR, 18–24 Jul 2021.
- G. Bouritsas, F. Frasca, S. Zafeiriou, and M. Bronstein.
  Improving graph neural network expressivity via subgraph isomorphism counting.
  IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(1):657–668, 2023.
- X. Bresson and T. Laurent.
  Residual gated graph convnets, 2018.

G. Corso, L. Cavalleri, D. Beaini, P. Liò, and P. Veličković.
Principal neighbourhood aggregation for graph nets.
In *Advances in Neural Information Processing Systems*, 2020.

M. Eliasof, E. Haber, and E. Treister.
pathGCN: Learning general graph spatial operators from paths.
In K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvari, G. Niu, and S. Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 5878–5891. PMLR, 17–23 Jul 2022.

M. Fey, J. G. Yuen, and F. Weichert.
Hierarchical inter-message passing for learning on molecular graphs.
In *ICML Graph Representation Learning and Beyond (GRL+) Workhop*, 2020.

- J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl.
  Neural message passing for quantum chemistry.
  In *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ICML'17, page 1263–1272. JMLR.org, 2017.
- L. Giusti, C. Battiloro, L. Testa, P. D. Lorenzo, S. Sardellitti, and S. Barbarossa.
  Cell attention networks, 2022.
- L. Giusti, T. Reu, F. Ceccarelli, C. Bodnar, and P. Liò.
  Cin++: Enhancing topological message passing, 2023.
- A. Grigor'yan, Y. Lin, Y. Muranov, and S.-T. Yau.
  Homologies of path complexes and digraphs, May 2013.
  arXiv:1207.2834 [math].
- A. A. Grigor'yan, Y. Lin, Y. V. Muranov, and S.-T. Yau.
  Path Complexes and their Homologies.
  Journal of Mathematical Sciences, 248(5):564–599, Aug. 2020.

W. Hu, M. Fey, M. Zitnik, Y. Dong, H. Ren, B. Liu, M. Catasta, and J. Leskovec.
Open graph benchmark: Datasets for machine learning on graphs.
In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS'20, Red Hook, NY, USA, 2020. Curran Associates Inc.

T. N. Kipf and M. Welling.
Semi-supervised classification with graph convolutional networks.
In *International Conference on Learning Representations*, 2017.

G. Michel, G. Nikolentzos, J. Lutzeyer, and M. Vazirgiannis.
Path neural networks: Expressive and accurate graph neural networks.
In *Proceedings of the 40th International Conference on Machine Learning (ICML)*, 2023.

C. Morris, N. M. Kriege, F. Bause, K. Kersting, P. Mutzel, and M. Neumann.
Tudataset: A collection of benchmark datasets for learning with graphs.
In *ICML 2020 Workshop on Graph Representation Learning and Beyond (GRL+ 2020)*, 2020.

M. Neumann, R. Garnett, C. Bauckhage, and K. Kersting.
Propagation kernels: efficient graph kernels from propagated information.
Mach. Learn., 102(2):209–245, 2016.

N. Shervashidze, P. Schweitzer, E. J. van Leeuwen, K. Mehlhorn, and K. M. Borgwardt.
Weisfeiler-lehman graph kernels.
Journal of Machine Learning Research, 12(77):2539–2561, 2011.

T. Sterling and J. J. Irwin.
ZINC 15 – Ligand discovery for everyone.
Journal of Chemical Information and Modeling, 55(11):2324–2337, 2015.

P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio.
Graph Attention Networks.
In *International Conference on Learning Representations (ICLR)*, Feb. 2018.
arXiv: 1710.10903 version: 3.

B. Weisfeiler and A. Lehman.
A reduction of a graph to a canonical form and an algebra arising during this reduction.
Nauchno-Technicheskaya Informatsiya, 1968.

K. Xu, W. Hu, J. Leskovec, and S. Jegelka.
How powerful are graph neural networks?
In *International Conference on Learning Representations*, 2019.