Weisfeiler and Leman Go Loopy

Raffaele Paolino*, Sohir Maskey*, Pascal Welke, Gitta Kutyniok

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN — TU WIEN — NEURAL INFORMATION PROCESSING SYSTEMS

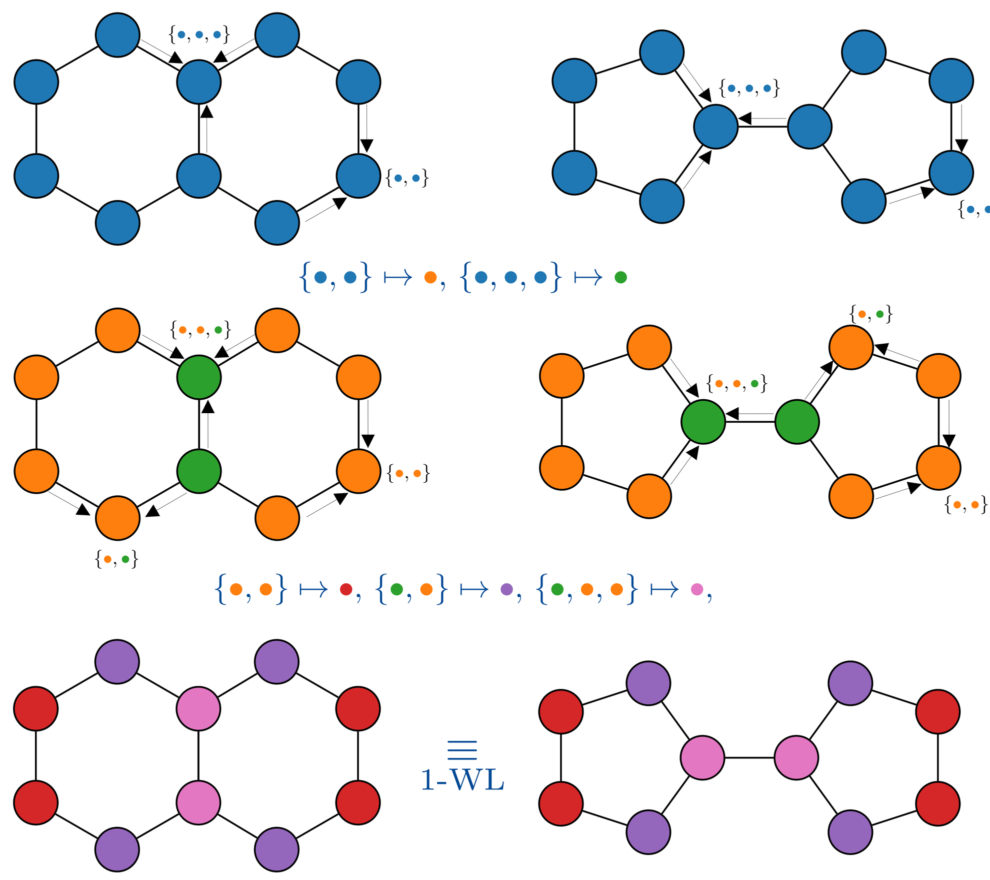
1. Motivation

↪ The most common paradigm of GNNs is **message passing**.

- Messages are collected from neighboring nodes.
- Messages are used to update the node features.

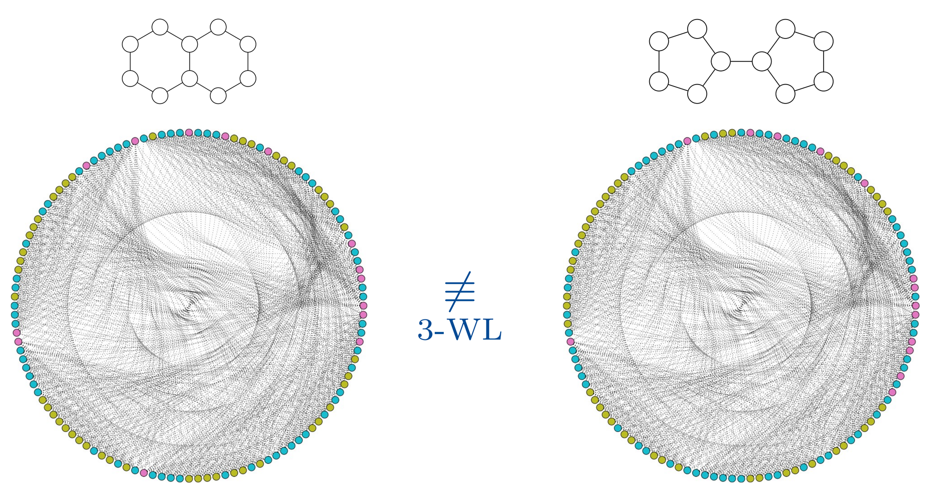
↪ 1-WL bounds expressive power of message passing GNNs.

- Scalable but not powerful enough.



↪ 3-WL applies message passing to an auxiliary graph.

- Powerful but not scalable.



↪ Easily distinguished by the counts of cycles.

	Pattern F				
Decalyn	10	22	0	78	0
Bycyclopentyl	10	22	0	78	20

↪ Need for more expressive and scalable neural architectures:

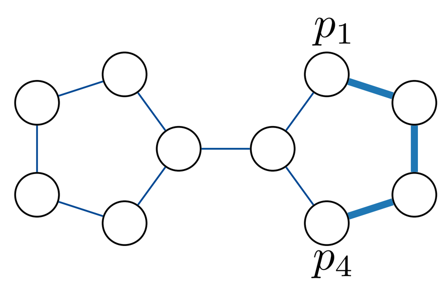
- Message Passing Neural Networks expressive power bounded by Weisfeiler-Leman test.
- Neural Networks based on higher-order Weisfeiler-Leman test present scalability issues.

↪ Ability to count important substructures:

- Other methods have limited cycle-counting powers.
- Explicit substructure counting is expensive and not flexible.

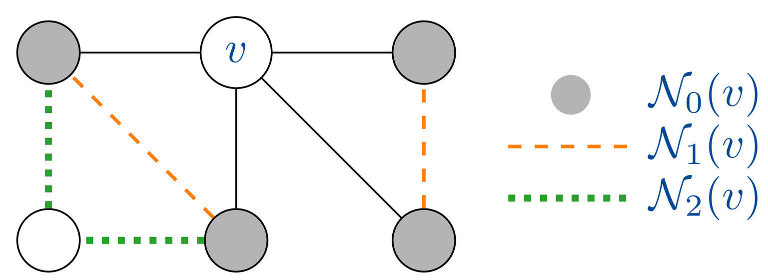
2. Loopy Weisfeiler Leman

Definition. Given a graph G, a simple path of length r is a collection $\mathbf{p} = \{p_i\}_{i=1}^{r+1}$ of $r+1$ distinct nodes such that consecutive nodes are adjacent.



Definition. Given a graph G and an integer $r \geq 1$, we define the r-neighborhood $\mathcal{N}_r(v)$ of $v \in V(G)$ as the set of all simple paths of length r between distinct direct neighbors of v which do not contain v, i.e.,

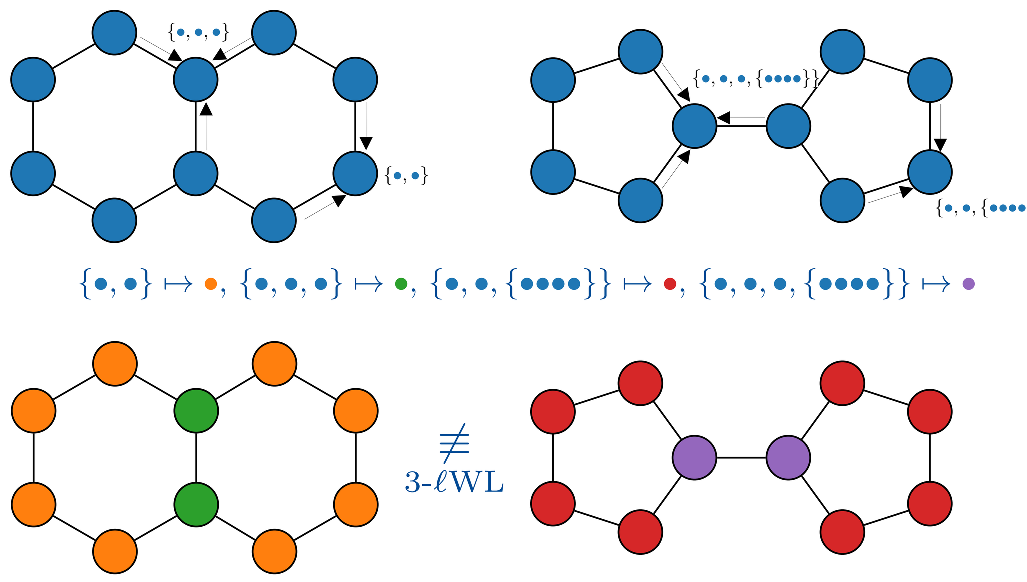
$$\mathcal{N}_r(v) \coloneqq \{\mathbf{p} \mid \mathbf{p}\ r\text{-path},\ p_1, p_{r+1} \in \mathcal{N}(v),\ v \notin \mathbf{p}\}.$$



Definition. We define the r-loopy Weisfeiler-Leman (r-ℓWL) test by the following color update:

$$c_r^{(t+1)}(v) \leftarrow \mathrm{HASH}_r\Bigg(c_r^{(t)}(v), \Big\{\Big\{c_r^{(t)}(\mathbf{p}) \mid \mathbf{p} \in \mathcal{N}_0(v)\Big\}\Big\}, \ \vdots \ \Big\{\Big\{c_r^{(t)}(\mathbf{p}) \mid \mathbf{p} \in \mathcal{N}_r(v)\Big\}\Big\}\Bigg),$$

where $c_r^{(t)}(\mathbf{p}) \coloneqq \left(c_r^{(t)}(p_1), c_r^{(t)}(p_2), \dots, c_r^{(t)}(p_{r+1})\right)$ is the sequence of colors of nodes in the path.



5. Summary

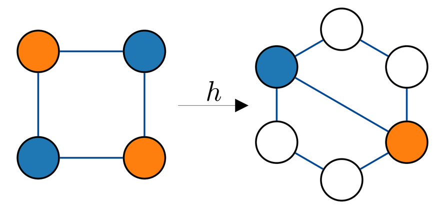
↪ Expressive:

- strictly more powerful than 1-WL;
- incomparable to k-WL and subgraph GNNs;
- more powerful than injecting subgraph-counts and homomorphism-counts as features;

↪ Scalable:

- preprocessing complexity $\mathcal{O}(N\, d^{r+2})$;
- linear complexity in the forward pass w.r.t. the number of edges and the number of paths in the r-neighborhoods;

3. Expressivity



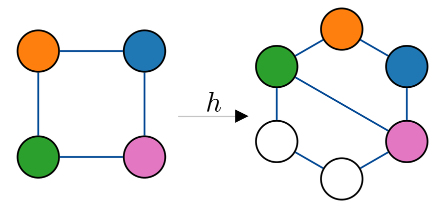
Definition. Let F and G be two graphs. A homomorphism from F to G is a map $h : V(F) \to V(G)$ such that $\{u, v\} \in E(F)$ implies $\{h(u), h(v)\} \in E(G)$. The set of homomorphisms from F to G is denoted by $\mathrm{Hom}(F, G)$, and its cardinality by $\mathrm{hom}(F, G) \coloneqq |\mathrm{Hom}(F, G)|$.

Counting homomorphisms is a complete isomorphic measure: $G \cong H$ if and only if $\mathrm{hom}(F, G) = \mathrm{hom}(F, H)\ \forall F$.

- 1-WL can homomorphism-count trees and forests.
- 3-WL can homomorphism-count graphs with tree width less than 3.

Theorem. Let $r \geq 1$. Then, r-ℓWL can homomorphism-count any graph in which every edge lies on at most one simple cycle of length at most $r+2$.

- r-ℓWL is more powerful than $\mathcal{F}$-Hom-GNNs, where $\mathcal{F} = \{C_3, \dots, C_{r+2}\}$.
- For any $k > 0$, 1-ℓWL is not less powerful than Subgraph k-GNN.



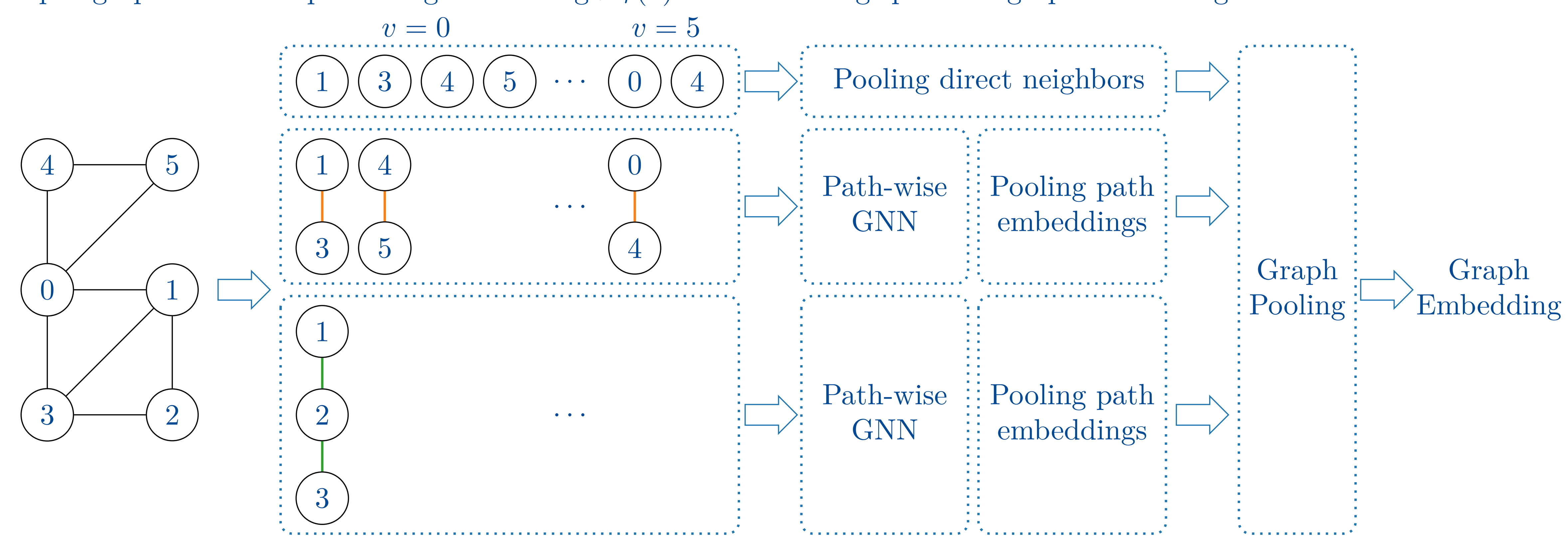
Definition. A subgraph isomorphism is an injective homomorphism.

Counting subgraph isomorphisms is the intuitive idea of counting how many times a motif appears in the graph.

- MPNN can subgraph-count paths of length up to 2.
- 3-WL can subgraph-count cycles of length up to 7.

Theorem. Let $r \geq 1$, r-ℓWL can subgraph-count all cycles with at most $r+2$ nodes. Moreover, $\forall k \in \mathbb{N}\ \exists r \in \mathbb{N}$ such that r-ℓWL is not less powerful than k-WL.

4. Experiments



ZINC

	Model	ZINC12K
1	5-ℓGIN	0.072 ± 0.002
2	DRFWL	0.077 ± 0.002
3	CIN	0.079 ± 0.006

	Model	ZINC250K
1	5-ℓGIN	0.022 ± 0.001
2	CIN	0.022 ± 0.002
3	I2-GNN	0.023 ± 0.001

QM9

	Model	μ
1	DTNN	0.244
2	DRFWL	0.346
3	5-ℓGIN	0.350

	Model	α
1	5-ℓGIN	0.217
2	DRFWL	0.222
3	I2-GNN	0.230

	Model	ϵ_{homo}
1	5-ℓGIN	0.00205
2	DRFWL	0.00226
3	DeepLRP	0.00254

	Model	U_0
1	5-ℓGIN	0.042
2	DRFWL	0.156
3	I2-GNN	0.211

	Model	R^2
1	5-ℓGIN	13.21
2	DRFWL	15.04
3	DTNN	17.0

	Model	ϵ_{lumo}
1	5-ℓGIN	0.00216
2	DRFWL	0.00225
3	DTNN	0.00267

Num. of distinguished pairs.

Model	B (60)	R (140)	E (100)	C (100)
3-WL	60	50	100	60
PPGN	60	50	100	23
NGNN	59	48	59	0
GSN	60	99	95	0
OSAN	52	41	82	2
4-ℓGIN	60	100	95	2

Test MAE for homomorphism- and subgraph-counts.

Model	$\mathrm{hom}(F,G)$			$\mathrm{sub}(F,G)$					
MPNN	0.300	0.233	0.254	0.358	0.208	0.188	0.146	0.261	0.205
Subgraph GNN	0.011	0.015	0.012	0.010	0.020	0.024	0.046	0.007	0.027
Local 2-GNN	0.008	0.008	0.010	0.008	0.011	0.017	0.034	0.007	0.016
Local 2-FGNN	0.003	0.005	0.004	0.003	0.004	0.010	0.020	0.003	0.010
r-ℓGIN	0.001	0.006	0.009	0.0005	0.0005	0.0003	0.0003	0.001	0.0004