Modeling Graph Structures: Graph Isomorphism and GNN

2022/09/23

Li Peng-Hsuan

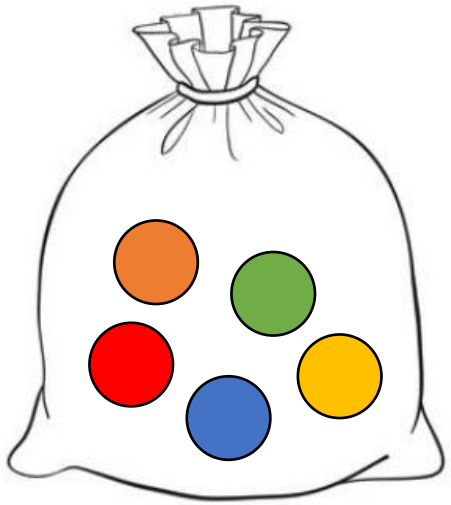
Items

- (-2019) Graph Isomorphism, Weisfeiler-Leman, and GNN
- (-2021) Beyond Weisfeiler-Leman: Heuristic and Theoretical Approaches
- (2022) A Sound Heuristic: Hierarchy of Local Isomorphism and GraphSNN

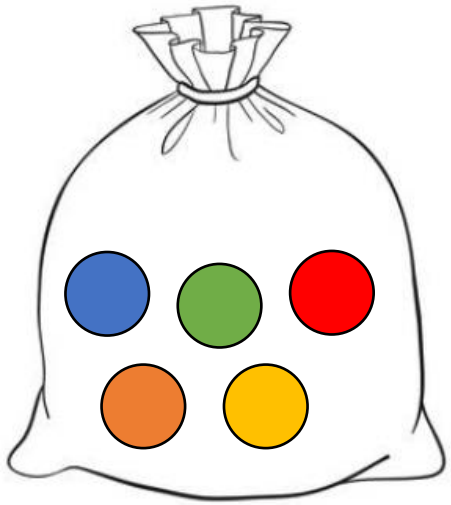
Items

- (-2019) Graph Isomorphism, Weisfeiler-Leman, and GNN
- (-2021) Beyond Weisfeiler-Leman: Heuristic and Theoretical Approaches
- (2022) A Sound Heuristic: Hierarchy of Local Isomorphism and GraphSNN

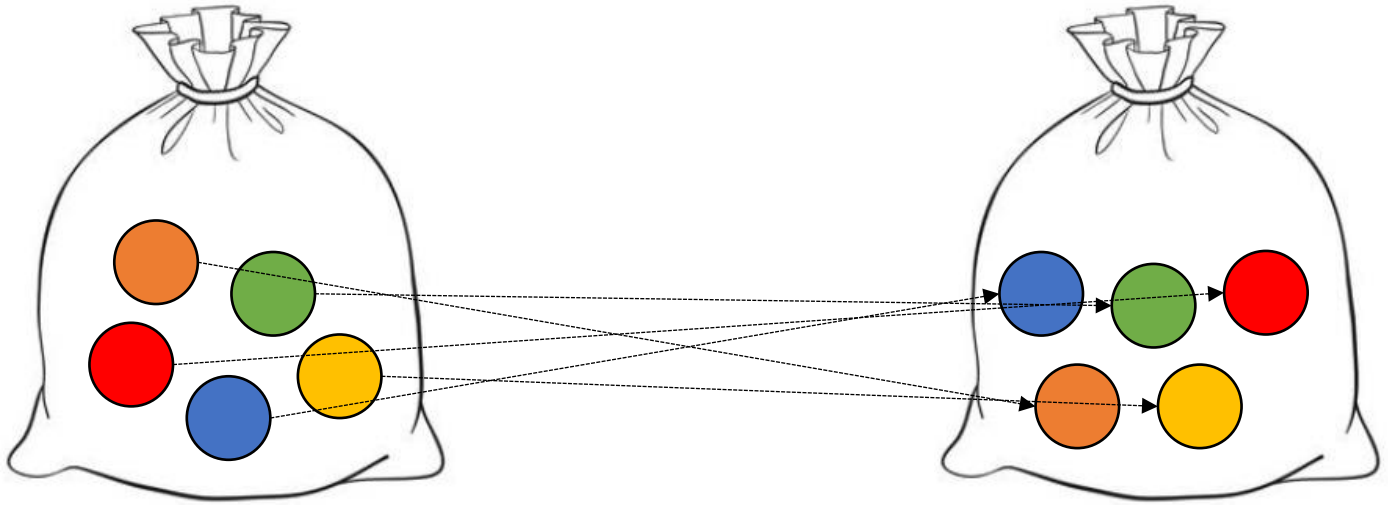
Comparing Sets



?
=

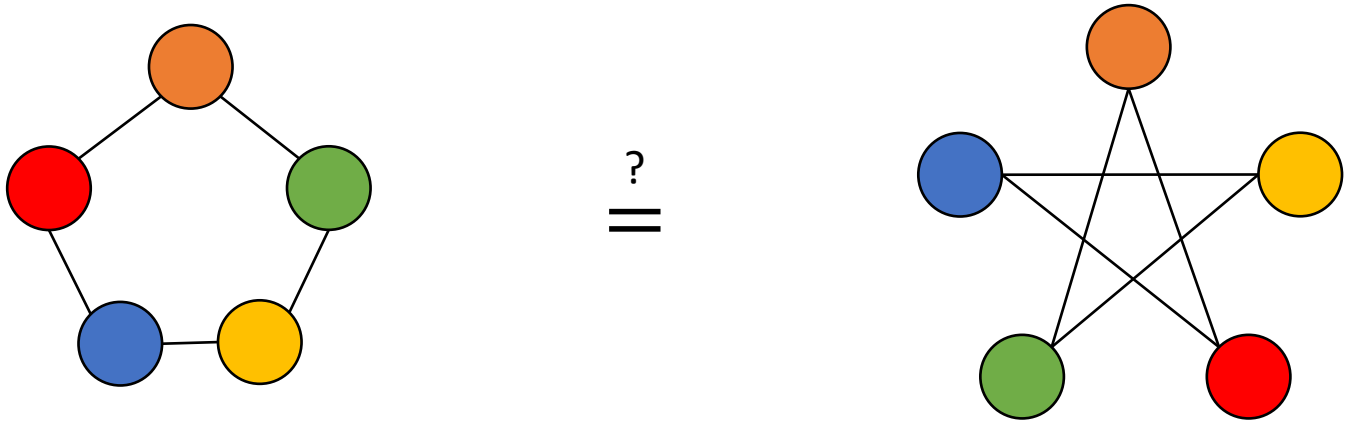


The Same Set

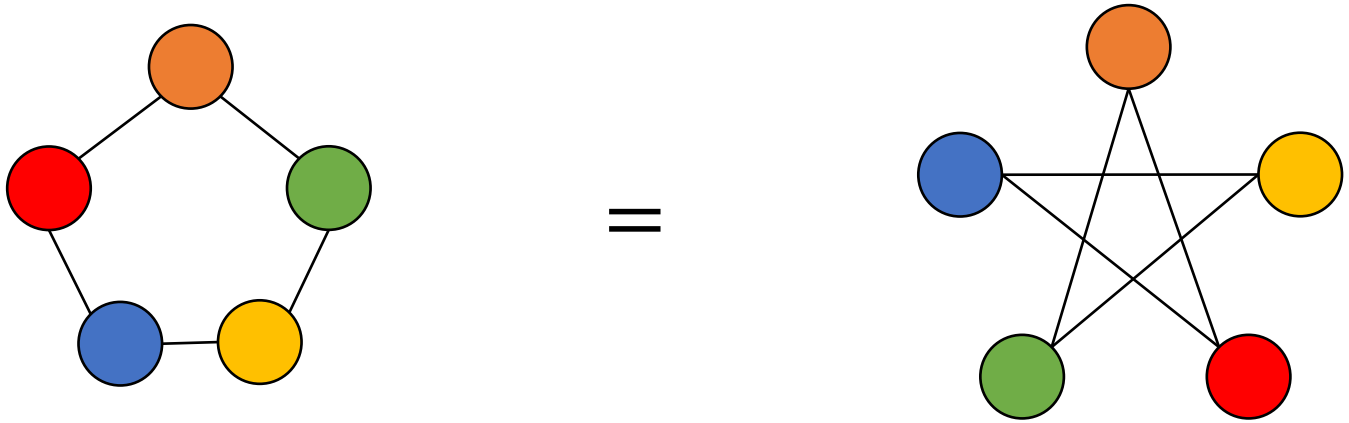


“=”: a bijective mapping s.t. member labels are preserved

Comparing Graphs



Graph Isomorphism



“=”: a bijective mapping s.t. nodes (labels), edges (labels) are preserved

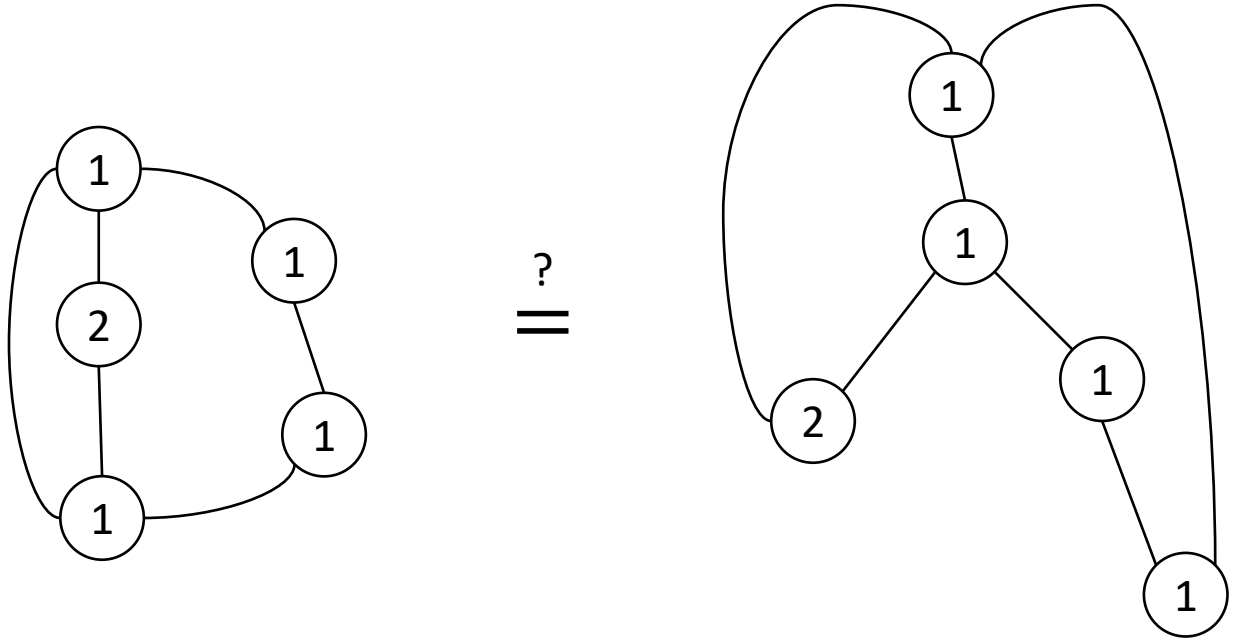
Graph Isomorphism Test

- Complexity
  - NP-completeness not known
  - The current algorithm with the best claim:
    - Quasi-polynomial time $2^{O((\log n)^c)}$ with $c = 3$

(Babai 2015), (Helfgott 2017)

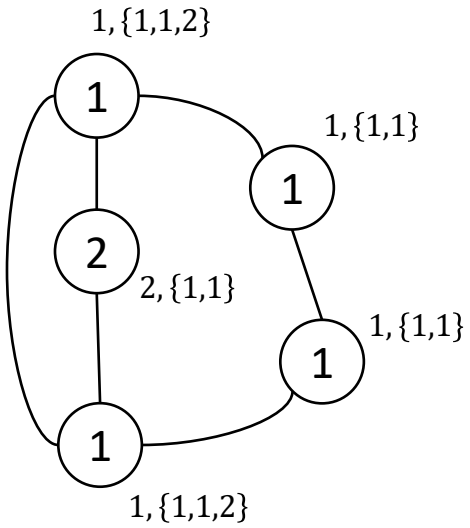
Weisfeiler-Leman

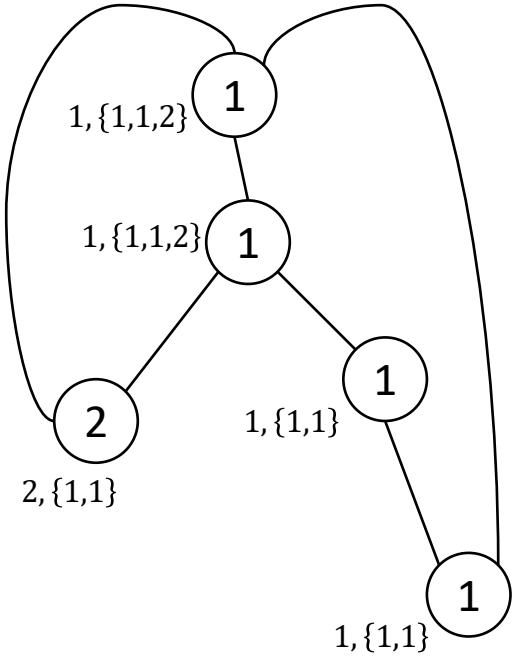
- Original



Weisfeiler-Leman

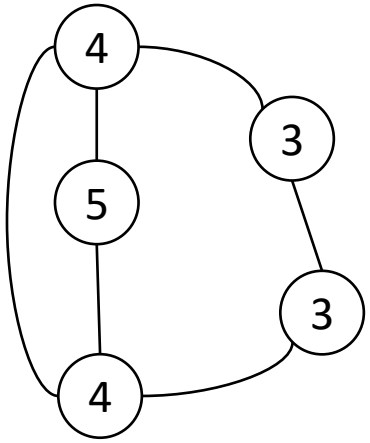
- Collect neighbor

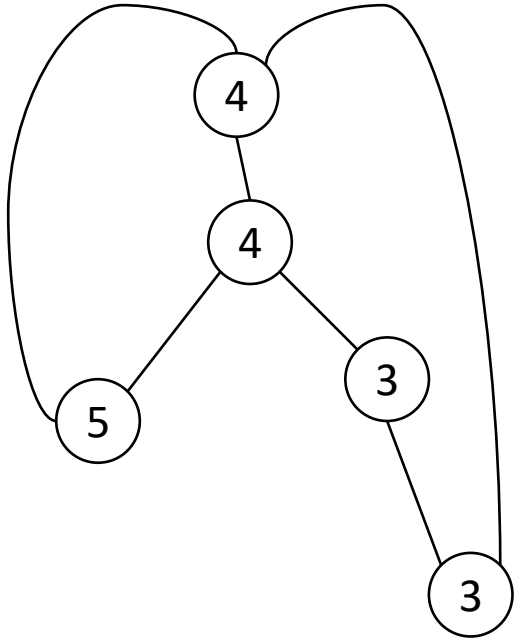




Weisfeiler-Leman

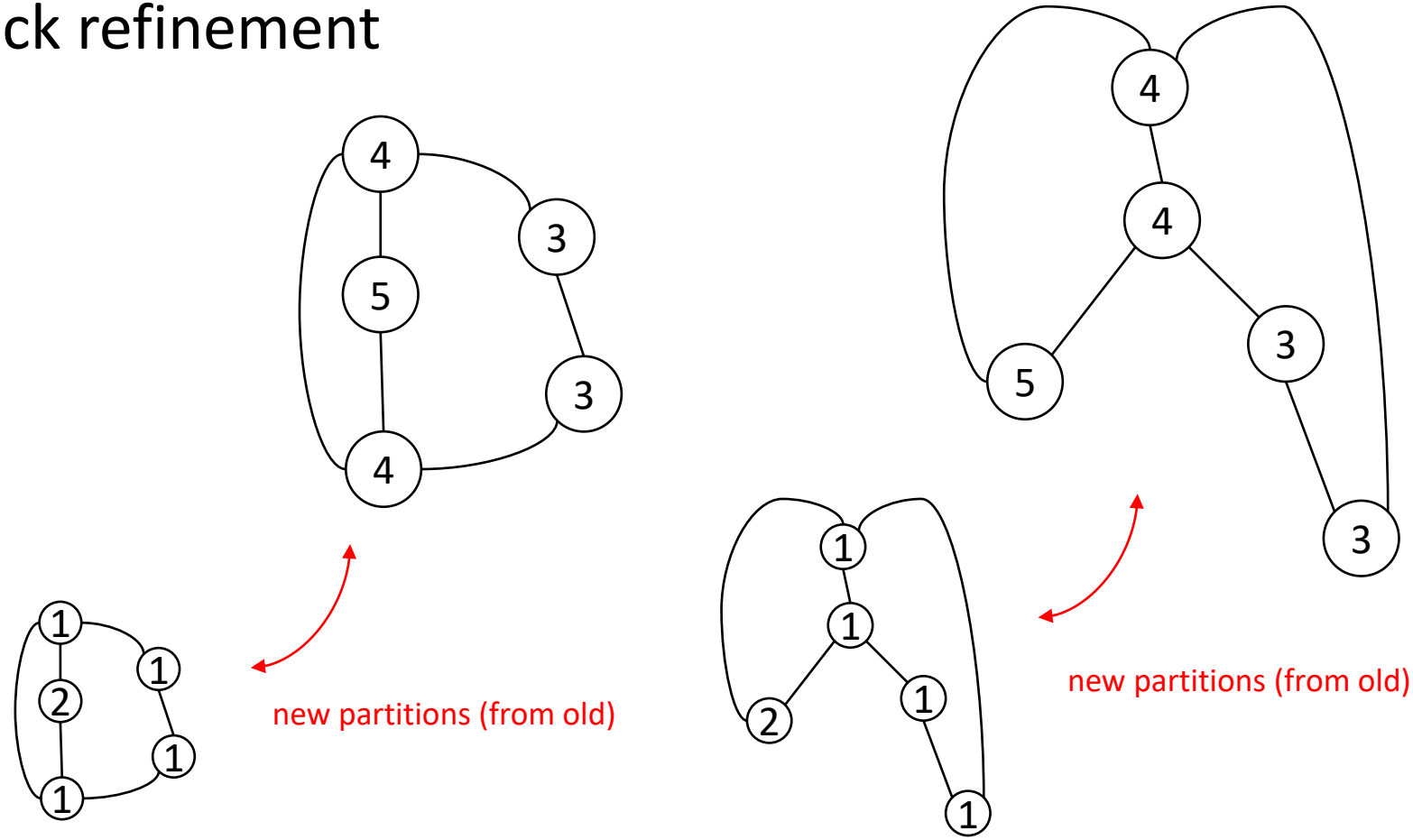
- Re-label





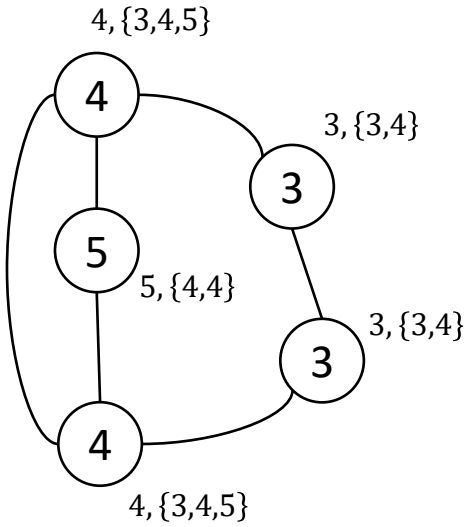
Weisfeiler-Leman

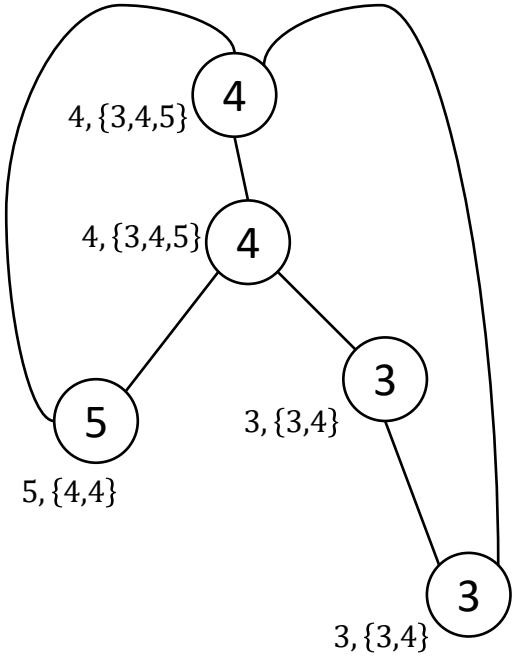
- Check refinement



Weisfeiler-Leman

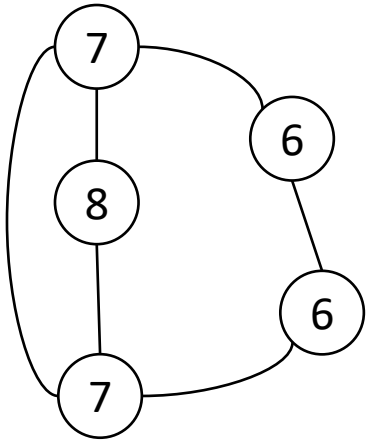
- Collect neighbor

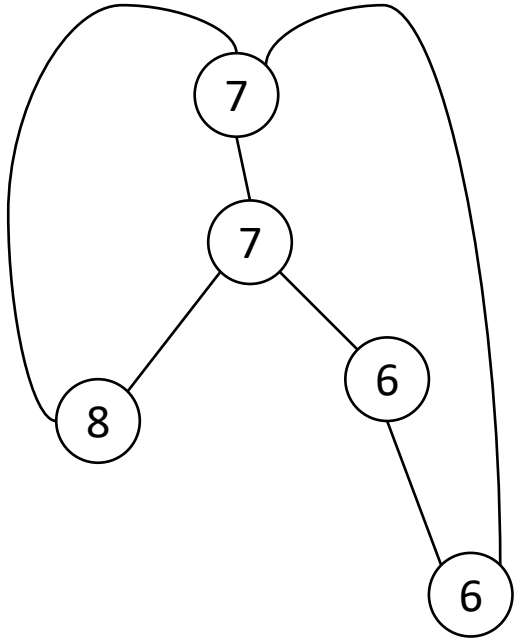




Weisfeiler-Leman

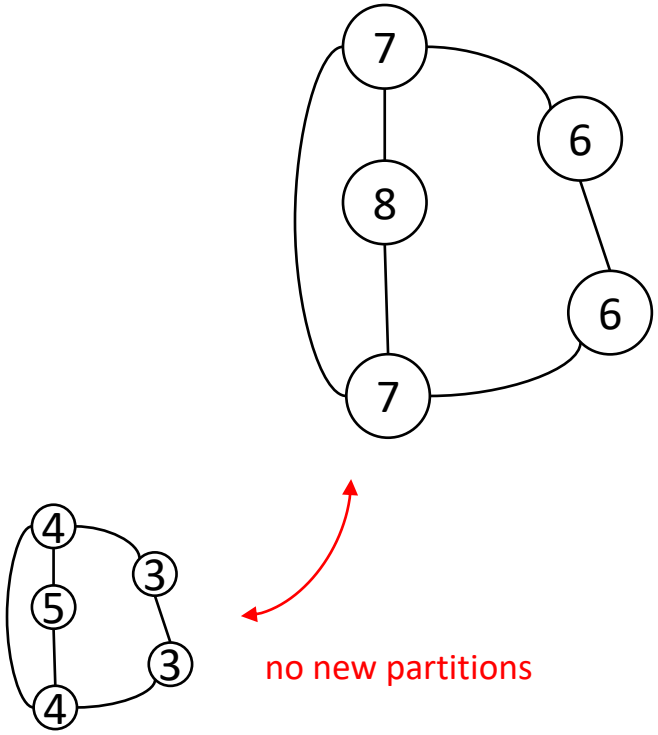
- Re-label



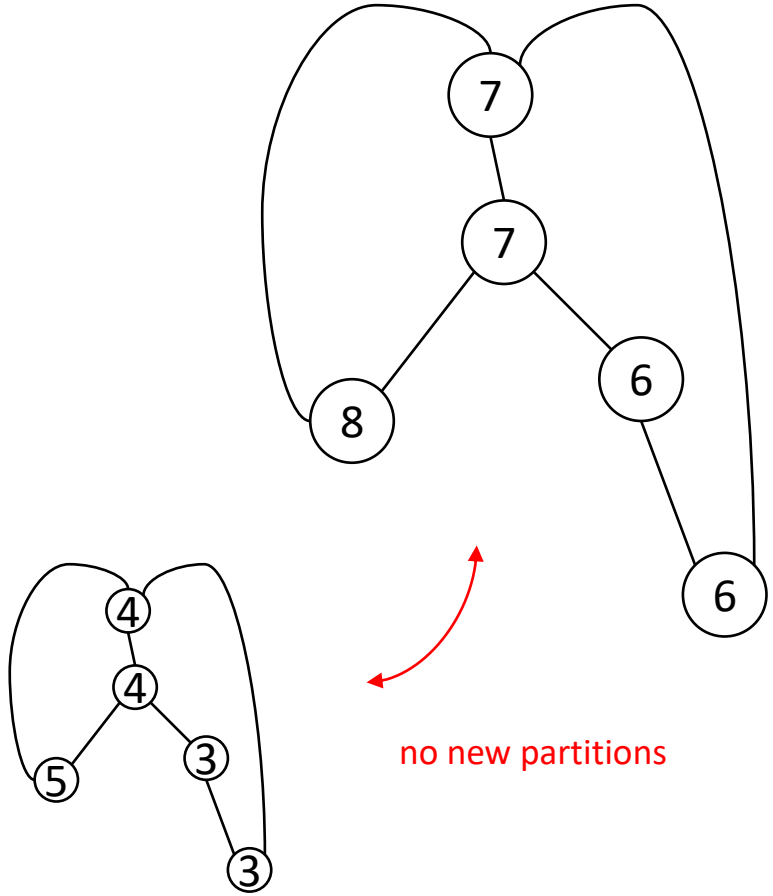


Weisfeiler-Leman

- Check refinement



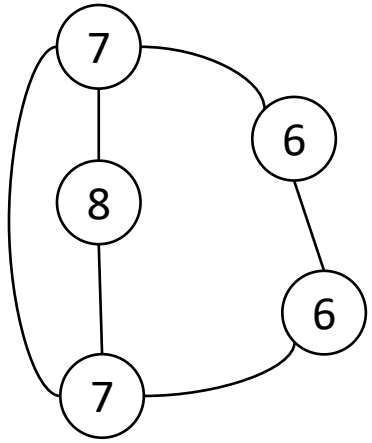
no new partitions

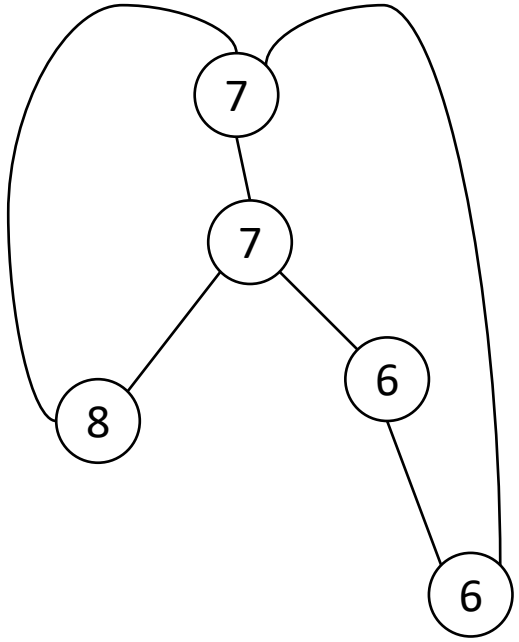


no new partitions

Weisfeiler-Leman

- Count final labels

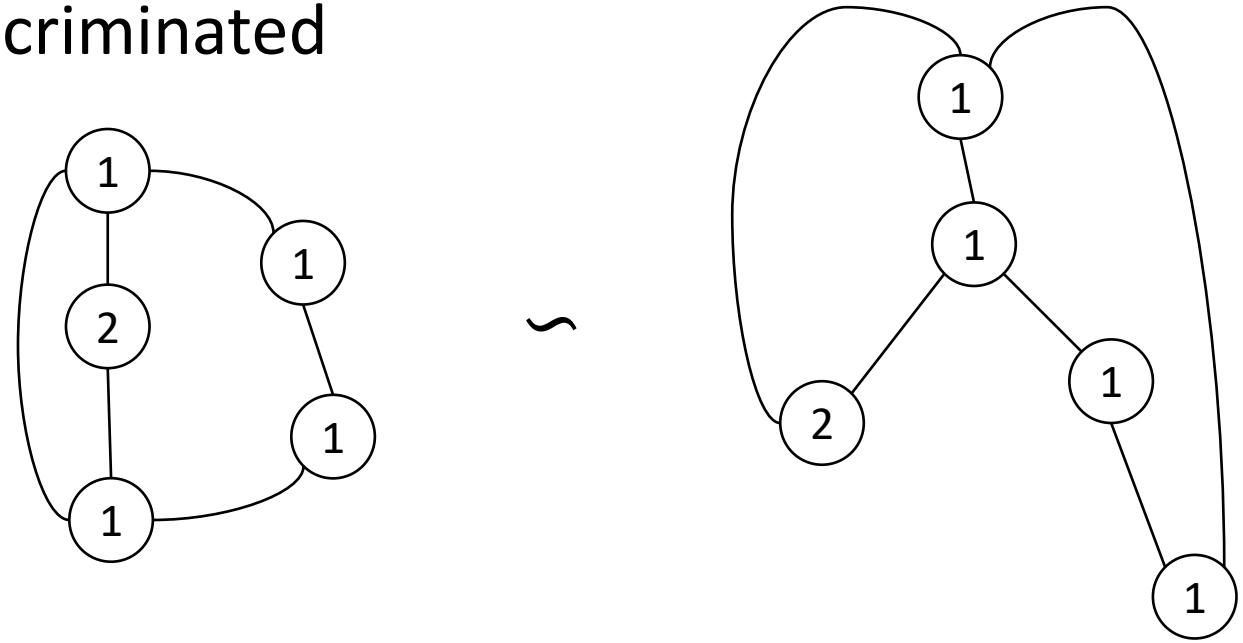




$\{6: 2, 7: 2, 8: 1\}$ = $\{6: 2, 7: 2, 8: 1\}$

Weisfeiler-Leman

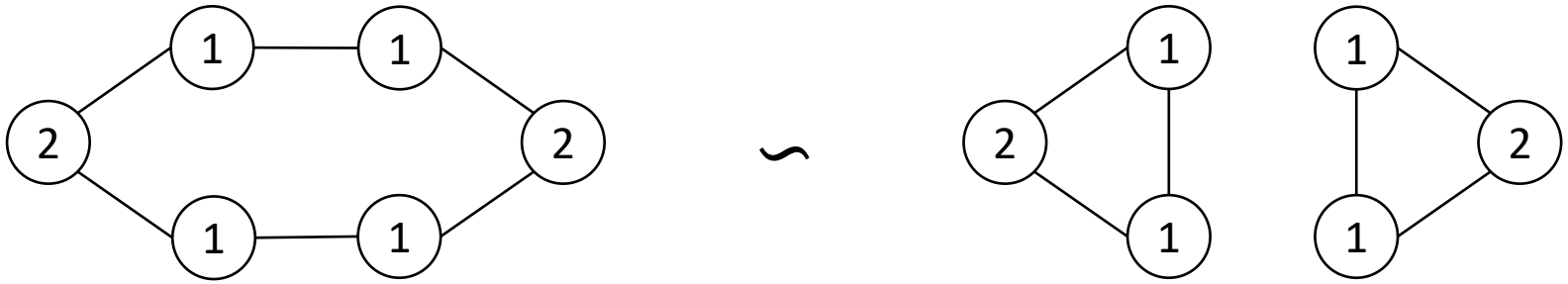
- Result: not discriminated



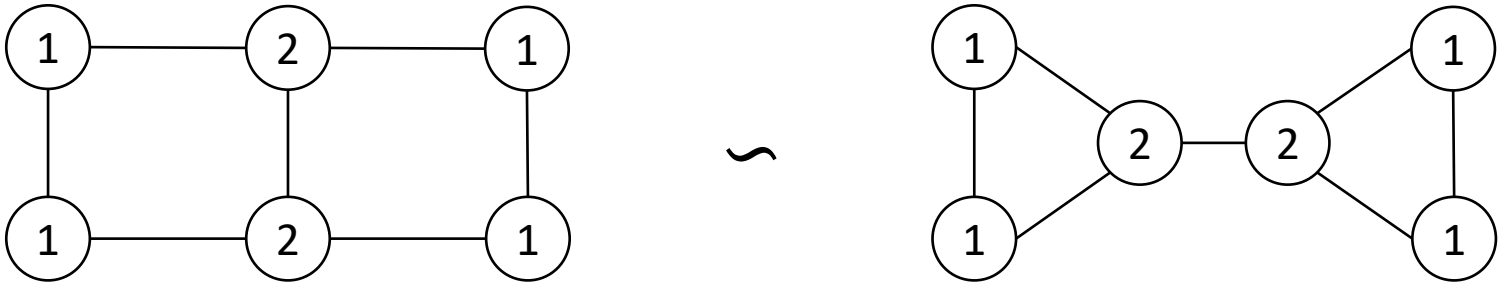
Weisfeiler-Leman

- Very fast
  - $O(n + e)$
- Sound
  - Isomorphic graphs won’t be discriminated
- Incomplete
  - Some non-isomorphic graphs are not discriminated

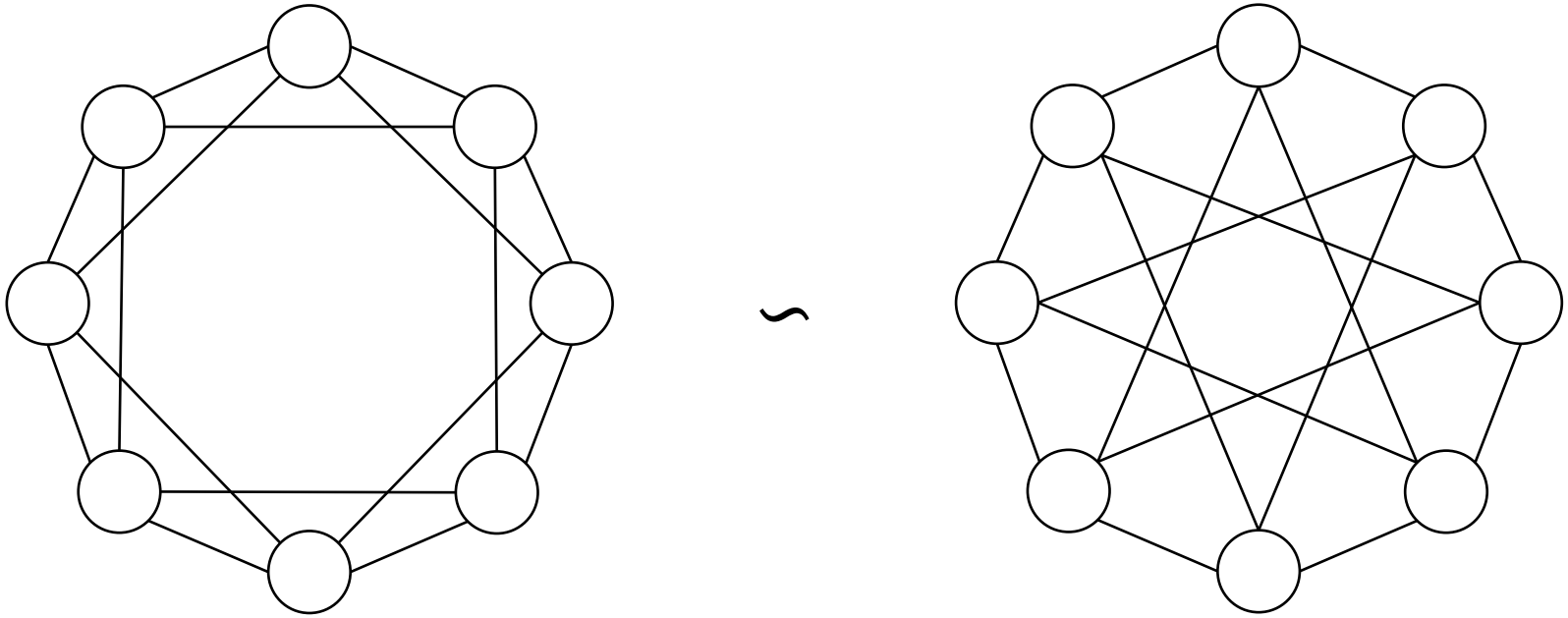
Weisfeiler-Leman: Limitations



Weisfeiler-Leman: Limitations



Weisfeiler-Leman: Limitations



GNN

- Graph Neural Networks (GNNs)
  - Each layer t maps each node and its neighbors from $R^{d_{t-1}}$ to R^{d_t}
  - (A message-passing framework)

GNN

- The power of GNNs in modeling graph structures
  - GNNs are at most as powerful as Weisfeiler-Leman
  - GCN, GraphSAGE, GAT are less powerful than Weisfeiler-Leman
  - GIN meets the simple requirements to be as powerful as Weisfeiler-Leman

$$h_v^{(k)} = \mathrm{MLP}^{(k)}\left(\left(1+\epsilon^{(k)}\right)\cdot h_v^{(k-1)} + \sum_{u\in\mathcal{N}(v)} h_u^{(k-1)}\right)$$

Xu et al. *How Powerful are Graph Neural Networks?* ICLR 2019.

Items

- (-2019) Graph Isomorphism, Weisfeiler-Leman, and GNN
- (-2021) Beyond Weisfeiler-Leman: Heuristic and Theoretical Approaches
- (2022) A Sound Heuristic: Hierarchy of Local Isomorphism and GraphSNN

Beyond Weisfeiler-Leman: Heuristics

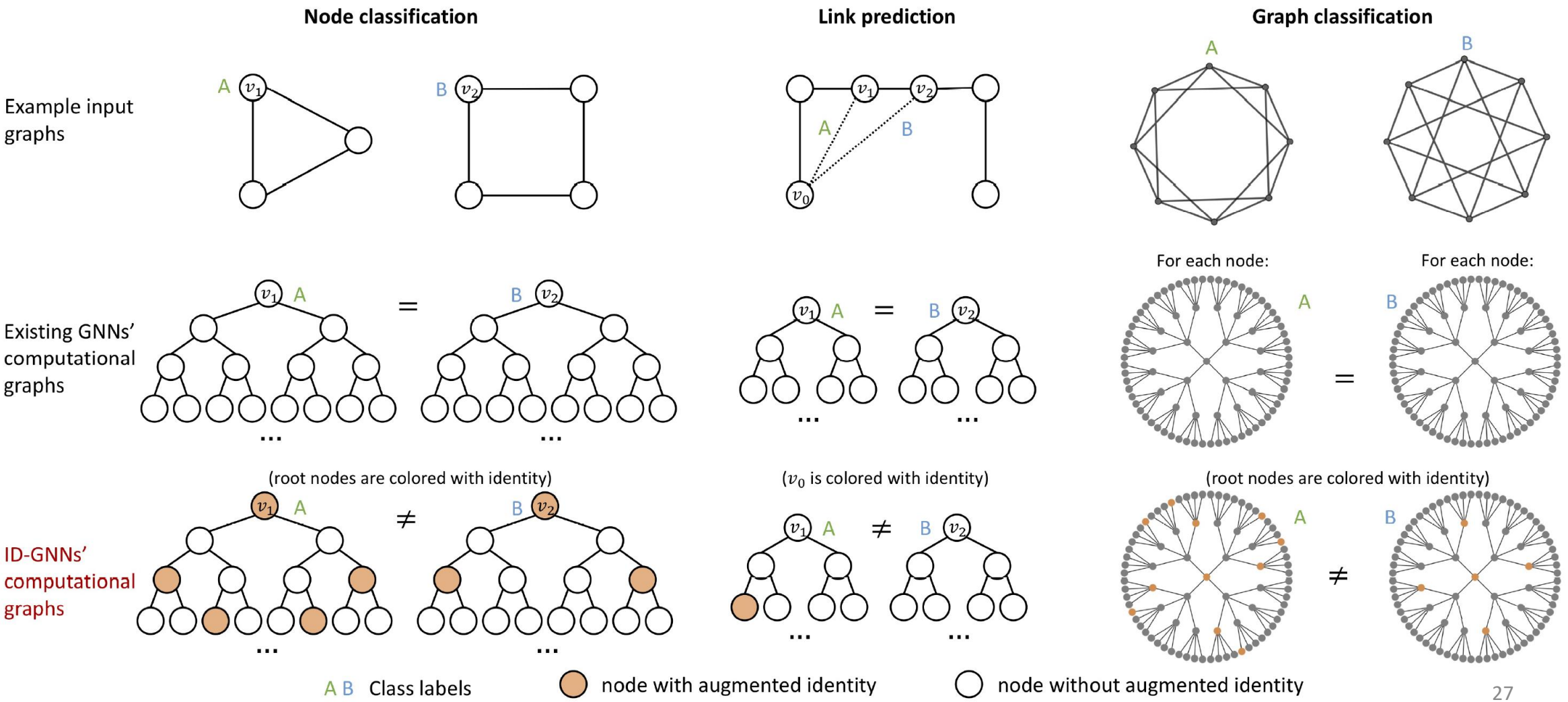
- Heuristic approaches
  - Add domain specific features
  - Add triangle, clique, cycle features
  - Add node identifiers
    - Arbitrary fixed id (for modeling one big graph)
    - Random id (for generalizability to unseen graphs)

Beyond Weisfeiler-Leman: Heuristics

- Heuristic approaches
  - ID-GNN
    - Fast
      - Add cycle count as node feature
    - Full
      - Individually expand a subgraph for each node
      - In each subgraph, distinguish the center node (and its duplicates) from other nodes

You et al. *Identity-aware Graph Neural Networks.* AAAI 2021.

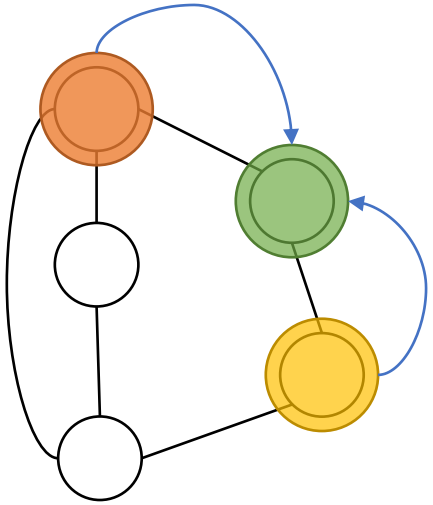
ID-GNN


Node classification
Link prediction
Graph classification
Example input graphs
A
v_1
B
v_2
v_0
Existing GNNs' computational graphs
=
...
For each node:
(root nodes are colored with identity)
(v_0 is colored with identity)
≠
ID-GNNs' computational graphs
A B Class labels
node with augmented identity
node without augmented identity

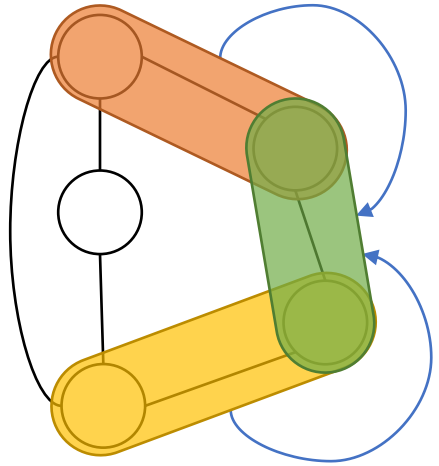
ID-GNN

		Node classification: real-world labels		Edge classification: link prediction				Graph classification: real-world labels			
		Cora	CiteSeer	ScaleFree	SmallWorld	ENZYMES	PROTEINS	ENZYMES	PROTEINS	BZR	ogbg-molhiv
GNNs	GCN	0.848±0.01	0.709±0.01	0.796±0.01	0.709±0.00	0.651±0.01	0.659±0.01	0.547±0.01	0.695±0.02	0.844±0.04	0.747±0.02
	SAGE	**0.868±0.01**	**0.726±0.01**	0.541±0.00	0.512±0.00	0.546±0.01	0.582±0.01	0.542±0.01	0.692±0.01	0.852±0.04	0.758±0.01
	GAT	0.857±0.01	0.716±0.01	0.500±0.00	0.500±0.00	0.478±0.01	0.491±0.01	**0.555±0.02**	**0.723±0.00**	0.848±0.03	0.742±0.01
	GIN	0.858±0.01	0.719±0.01	**0.802±0.01**	**0.722±0.01**	**0.654±0.01**	**0.667±0.00**	0.553±0.02	0.721±0.00	**0.856±0.02**	**0.762±0.03**
ID-GNNs Fast	GCN	0.851±0.02	0.715±0.00	0.856±0.03	0.719±0.00	0.649±0.01	0.671±0.01	0.600±0.01	***0.741±0.02***	0.807±0.02	0.772±0.02
	SAGE	0.866±0.02	***0.742±0.01***	***0.898±0.01***	0.743±0.02	0.671±0.04	0.701±0.01	***0.639±0.00***	0.724±0.03	0.835±0.06	**0.780±0.01**
	GAT	**0.870±0.02**	0.719±0.02	0.731±0.02	0.537±0.00	0.490±0.01	0.502±0.01	0.619±0.03	0.715±0.03	0.848±0.05	0.740±0.01
	GIN	0.864±0.01	0.719±0.01	0.837±0.01	**0.759±0.01**	**0.718±0.02**	**0.724±0.00**	0.567±0.01	0.723±0.01	**0.864±0.03**	0.755±0.02
ID-GNNs Full	GCN	0.863±0.01	0.719±0.01	0.771±0.04	0.798±0.03	0.838±0.01	0.878±0.02	**0.586±0.04**	0.715±0.02	0.881±0.04	0.769±0.01
	SAGE	0.875±0.01	**0.730±0.02**	0.741±0.01	0.724±0.03	0.819±0.01	0.863±0.01	0.547±0.02	0.721±0.01	0.864±0.02	***0.783±0.02***
	GAT	***0.878±0.01***	0.729±0.01	0.749±0.01	0.742±0.03	0.824±0.01	0.859±0.03	0.567±0.05	**0.738±0.01**	***0.881±0.04***	0.739±0.01
	GIN	0.851±0.00	0.725±0.01	**0.815±0.01**	***0.810±0.03***	***0.846±0.01***	***0.886±0.02***	0.544±0.02	0.730±0.03	0.852±0.03	0.756±0.00
Best ID-GNN over best GNN		1.0%	1.6%	9.6%	8.7%	19.2%	21.9%	8.3%	1.8%	2.5%	2.0%

Beyond Weisfeiler-Leman: Theoretical



Weisfeiler-Leman / GNN



k-WL / k-GNN

Morris et al. *Weisfeiler and Leman Go Neural: Higher-order Graph Neural Networks*. AAAI 2019.

k-WL / k-GNN

- Just like Weisfeiler-Leman / GNN, but
  - “Nodes” → k-node subgraphs
  - “Neighbors” → subgraphs with a (k-1)-node intersection

k-WL / k-GNN

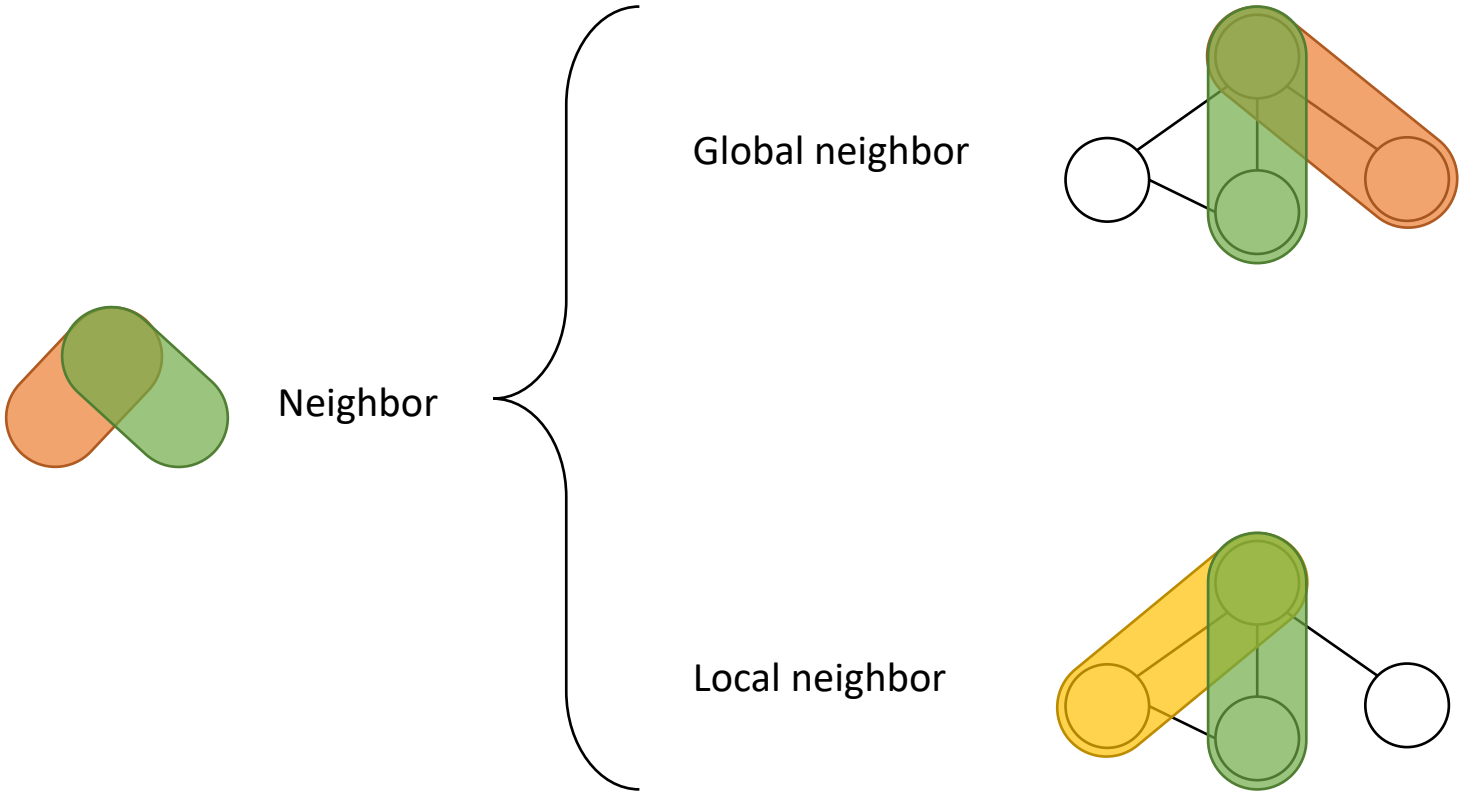
- Power in modeling graph structures
  - $\text{Weisfeiler–Leman} \equiv 2\text{–WL} \sqsupset 3\text{–WL} \sqsupset 4\text{–WL} \sqsupset \cdots$
- Complexity
  - $O(n^k)$

k-WL / k-GNN

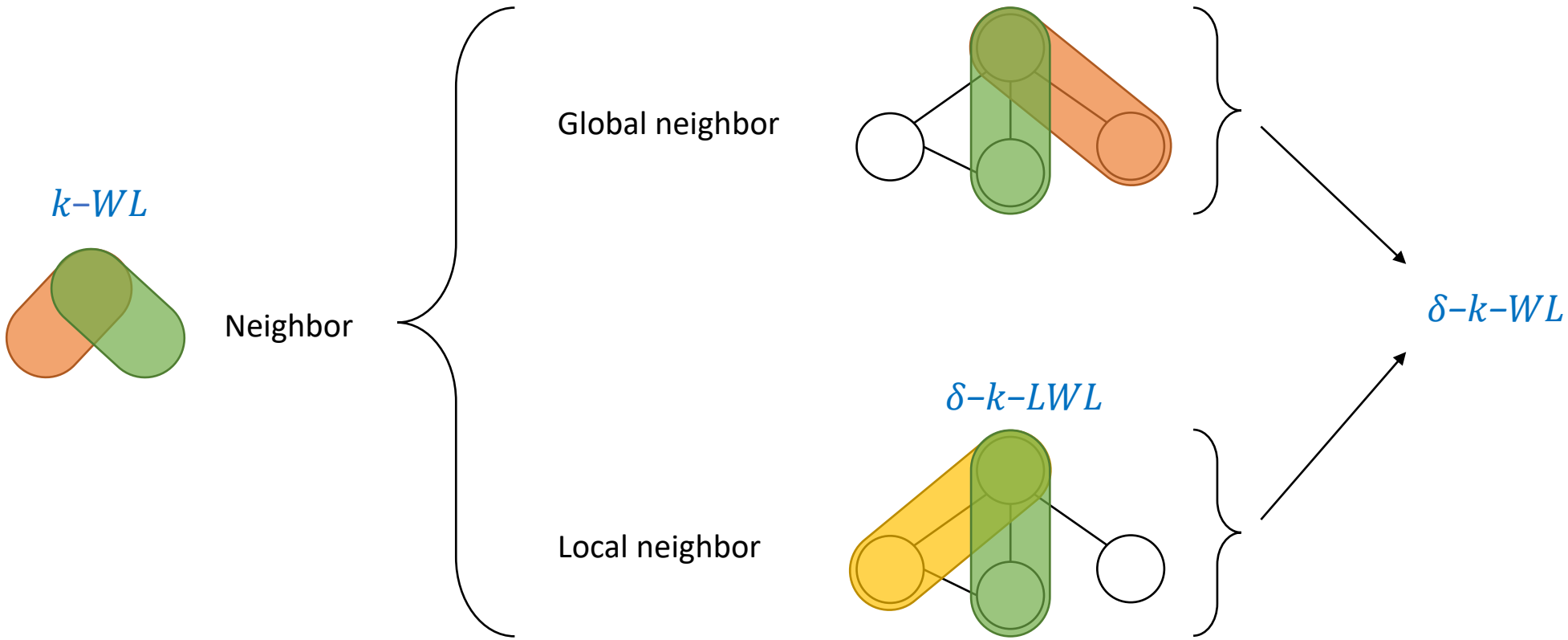
- 3-WL
  - Power in modeling graph structures
    - $\text{Weisfeiler–Leman} \equiv 2\text{–WL} \sqsupset 3\text{–WL}$
    - Distinguishes all planar graphs
  - Complexity
    - $O(n^2)$

Maron et al. *Invariant and Equivariant Graph Networks*. ICLR 2019.

k-WL / k-GNN

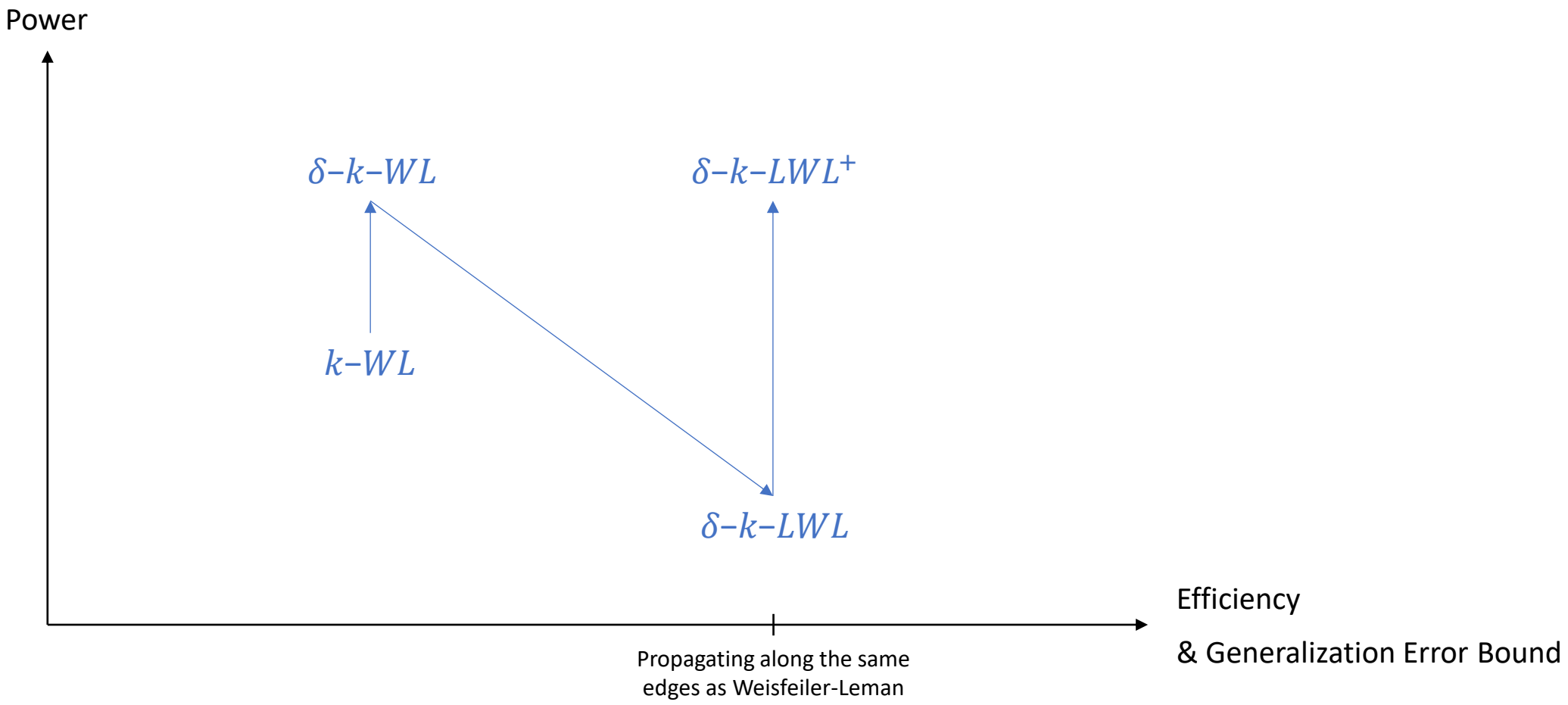


$\delta-k-LWL$ / $\delta-k-LGNN$



Morris et al. *Weisfeiler and Leman go sparse: Towards scalable higher-order graph embeddings*. NeurIPS 2020.

$\delta\text{-}k\text{-}LWL^{+}$ / $\delta\text{-}k\text{-}LGNN^{+}$



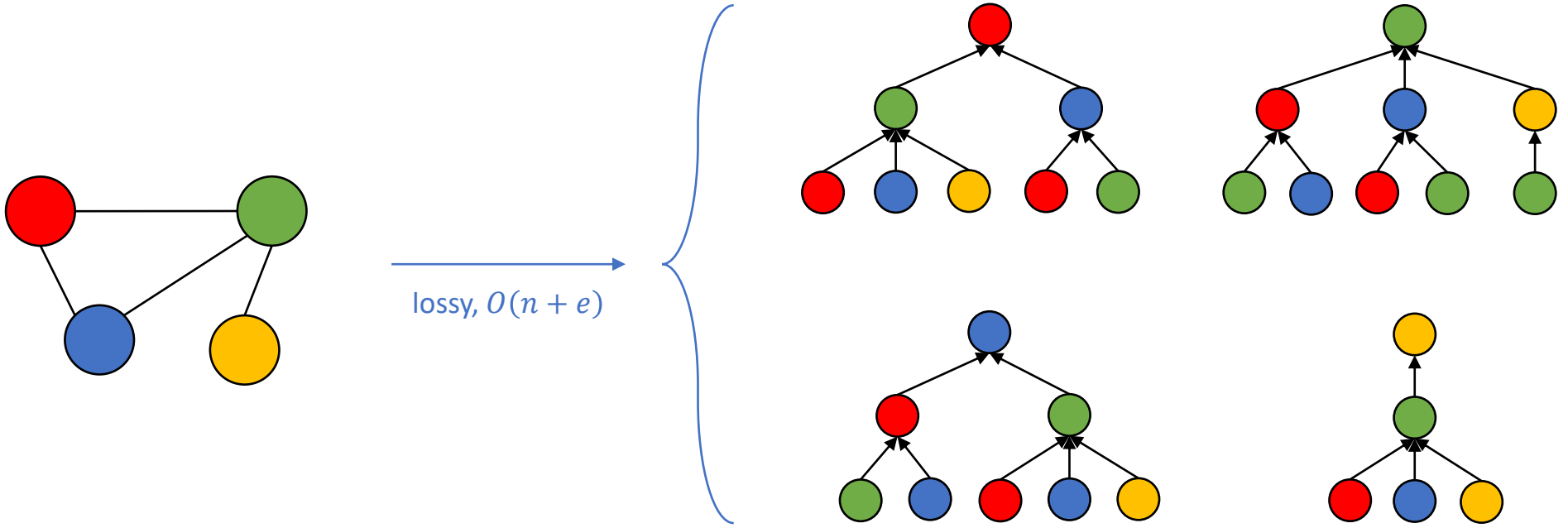
$\delta\text{-}k\text{-}LWL^{+}$ / $\delta\text{-}k\text{-}LGNN^{+}$

	Method	Dataset							
		ENZYMES	IMDB-BINARY	IMDB-MULTI	NCI1	NCI109	PTC_FM	PROTEINS	REDDIT-BINARY
Baseline	GR	29.7 ±0.6	58.9 ±1.0	39.0 ±0.8	66.1 ±0.4	66.3 ±0.2	61.3 ±1.1	71.2 ±0.6	60.0 ±0.2
	SP	40.7 ±0.9	58.5 ±0.4	39.4 ±0.3	74.0 ±0.3	73.0 ±0.4	61.3 ±1.3	75.6 ±0.5	84.6 ±0.3
	1-WL	50.7 ±1.2	72.5 ±0.5	50.0 ±0.5	84.2 ±0.3	84.3 ±0.3	**62.6** ±2.0	72.6 ±1.2	72.8 ±0.5
	WLOA	56.8 ±1.6	72.7 ±0.9	50.1 ±0.7	84.9 ±0.3	85.2 ±0.3	61.8 ±1.5	73.2 ±0.6	88.1 ±0.4
Neural	Gin-0	38.8 ±1.7	72.7 ±0.9	49.9 ±0.8	78.5 ±0.5	76.7 ±0.8	58.2 ±3.3	71.3 ±0.9	89.8 ±0.6
	Gin-ε	39.4 ±1.7	72.9 ±0.6	49.6 ±0.9	78.6 ±0.3	77.0 ±0.5	57.7 ±2.0	71.1 ±0.8	90.3 ±0.3
Global	2-WL	36.7 ±1.7	68.2 ±1.1	48.1 ±0.5	67.1 ±0.3	67.5 ±0.2	62.3 ±1.6	75.0 ±0.8	OOM
	3-WL	42.3 ±1.1	67.8 ±0.8	47.0 ±0.7	OOT	OOT	61.5 ±1.7	OOM	OOM
	δ-2-WL	37.5 ±1.2	68.1 ±1.1	47.9 ±0.7	67.0 ±0.5	67.2 ±0.4	61.9 ±0.9	75.0 ±0.4	OOM
	δ-3-WL	43.0 ±1.4	67.5 ±1.0	47.3 ±0/9	OOT	OOT	61.2 ±2.0	OOM	OOM
Local	δ-2-LWL	56.6 ±1.2	73.3 ±0.5	50.2 ±0.6	84.7 ±0.3	84.2 ±0.4	60.3 ±3.2	75.1 ±0.3	89.7 ±0.4
	δ-2-LWL^{+}	52.9 ±1.4	75.7 ±0.7	62.5 ±1.0	**91.4** ±0.2	**89.3** ±0.2	**62.6** ±1.6	**79.3** ±1.1	**91.1** ±0.5
	δ-3-LWL	**57.6** ±1.2	72.8 ±1.2	49.3 ±1.0	83.4 ±0.2	82.4 ±0.4	61.3 ±1.6	OOM	OOM
	δ-3-LWL^{+}	56.8 ±1.2	**76.2** ±0.8	**64.2** ±0.9	82.7 ±0.5	81.9 ±0.4	61.3 ±2.0	OOM	OOM

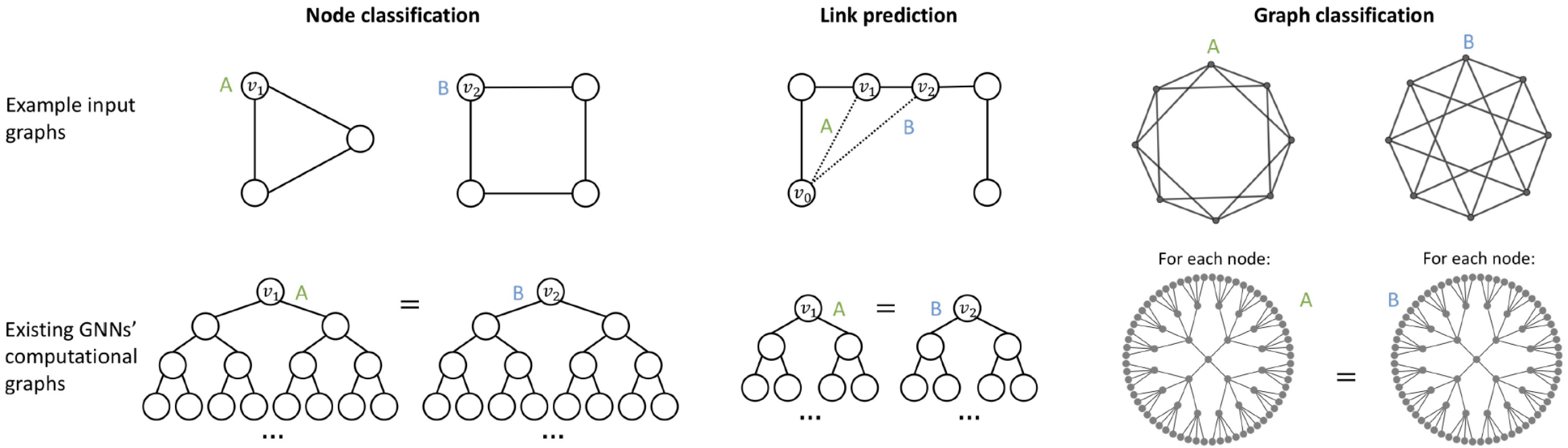
Items

- (-2019) Graph Isomorphism, Weisfeiler-Leman, and GNN
- (-2021) Beyond Weisfeiler-Leman: Heuristic and Theoretical Approaches
- (2022) A Sound Heuristic: Hierarchy of Local Isomorphism and GraphSNN

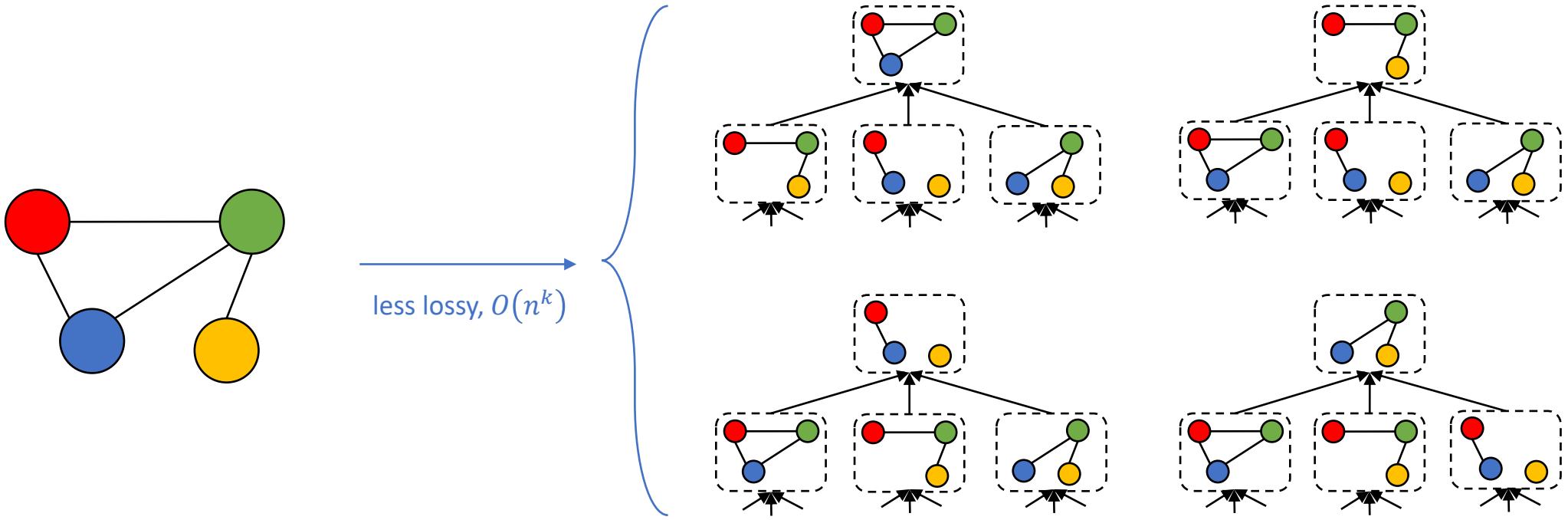
Weisfeiler-Leman / GNN



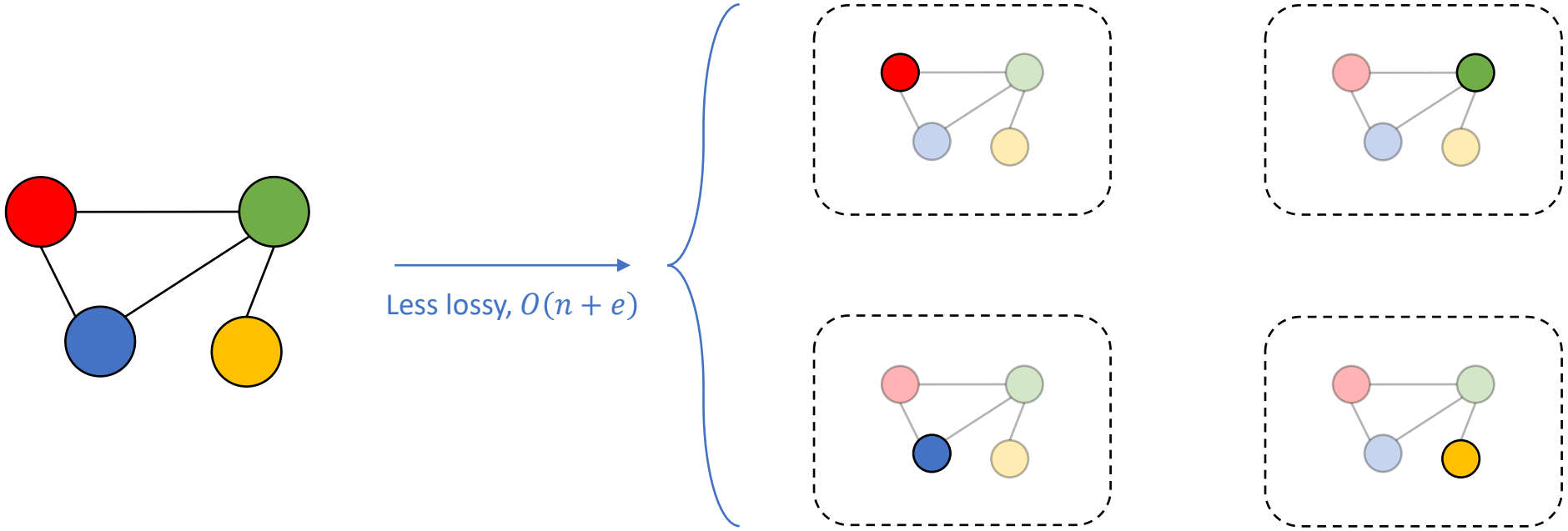
Weisfeiler-Leman / GNN



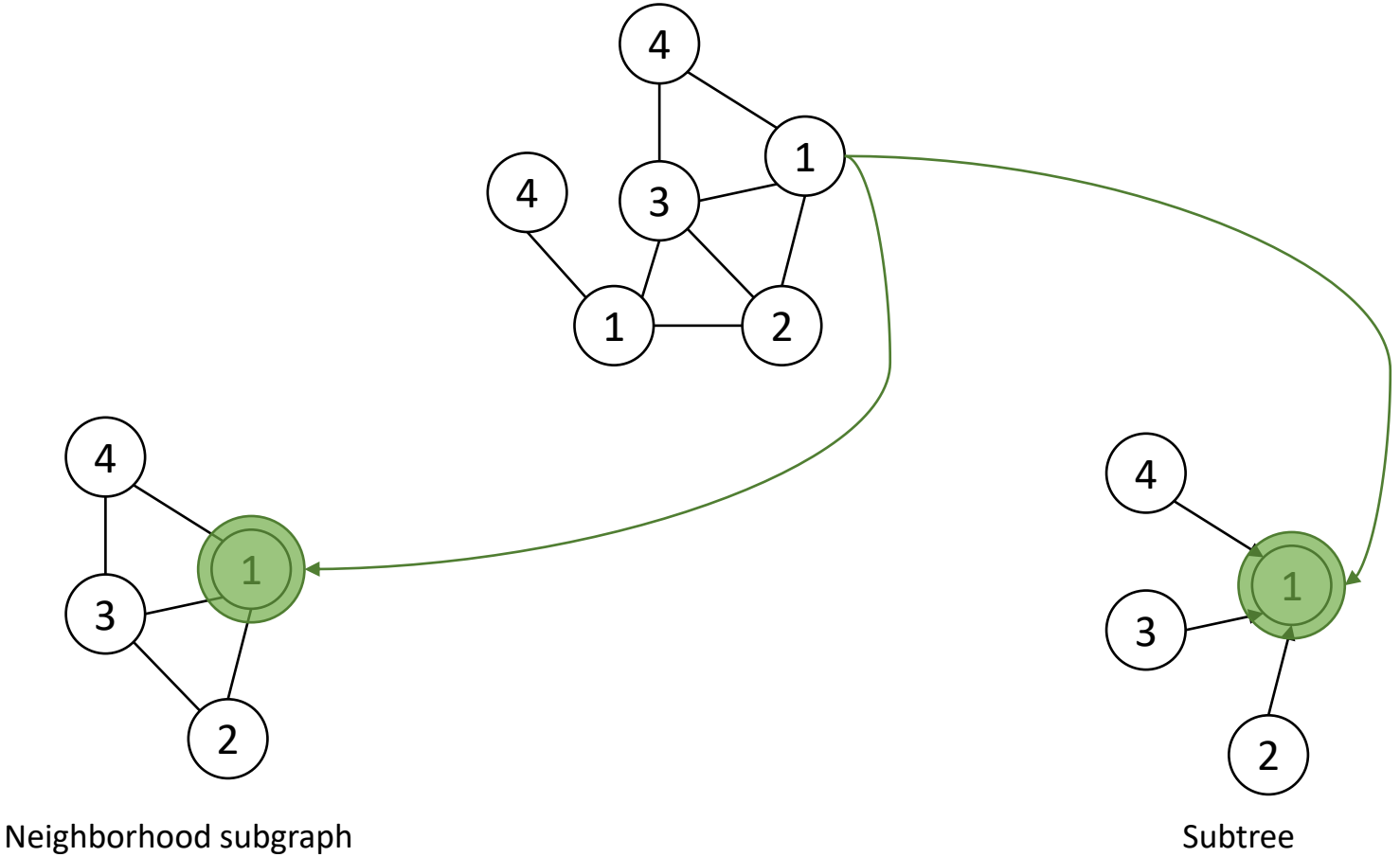
k-WL / k-GNN



Modeling Local Structures



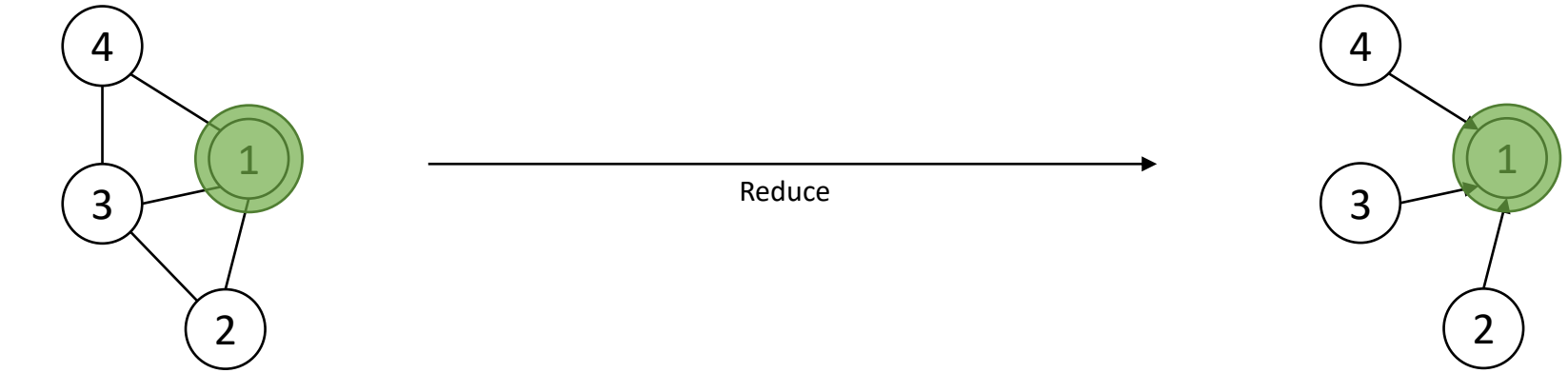
Modeling Local Structures



Neighborhood subgraph

Subtree

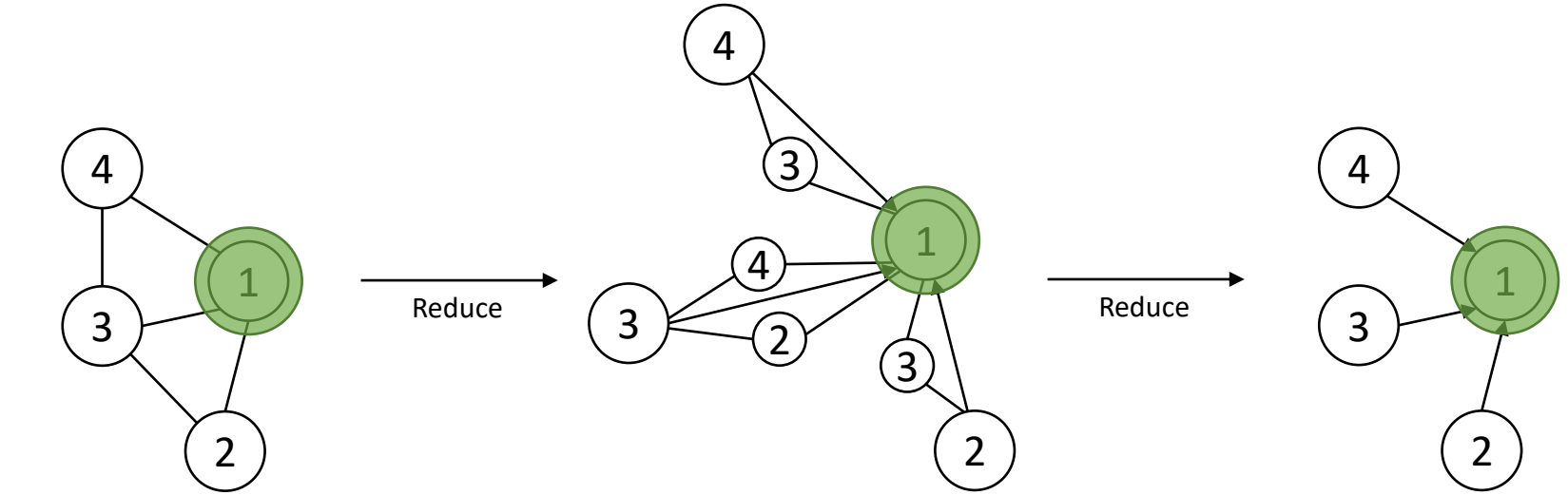
Modeling Local Structures



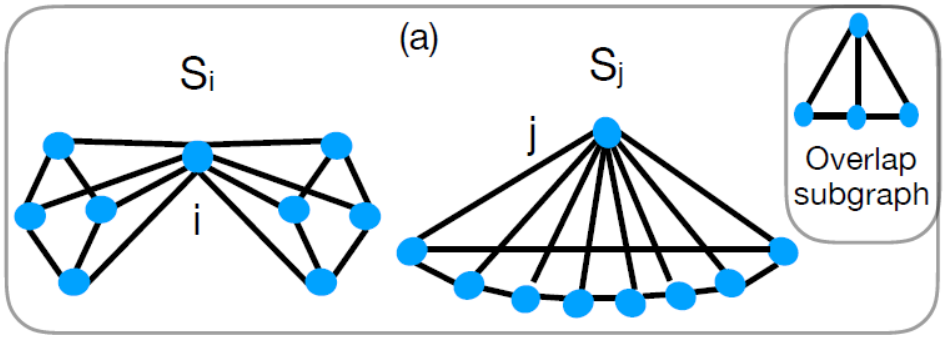
Neighborhood subgraph

Subtree

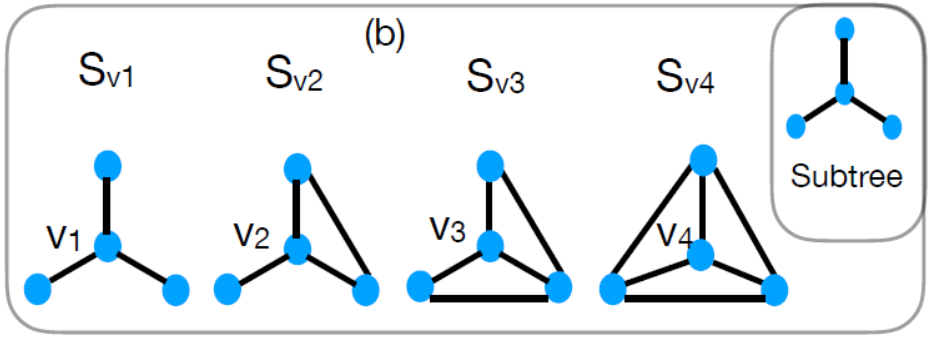
Hierarchy of Local Isomorphism



Hierarchy of Local Isomorphism

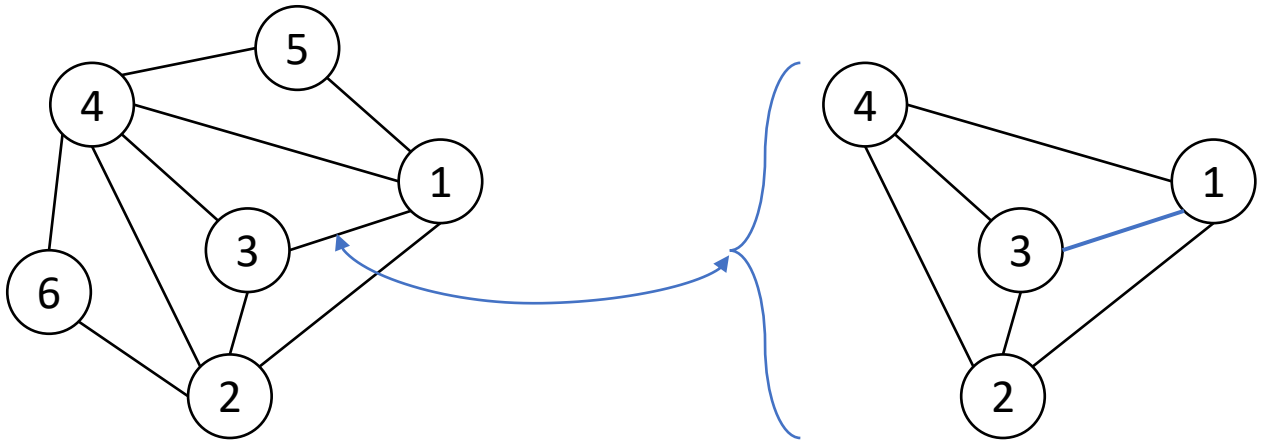


Overlap-isomorphic



Subtree-isomorphic

Edge ⟷ Overlap Subgraph



GraphSNN

- For each edge
  - Let S be its overlap subgraph with n nodes, e edges
  - Compute the overlap subgraph feature $\omega(S) \equiv \frac{e}{(n-1)n} n^{\lambda}$

    → $\omega \propto \text{density} \cdot \text{order}$

    → λ: the weight of order vs. density

GraphSNN

- For each edge
  - Let S be its overlap subgraph with n nodes, e edges
  - Compute the overlap subgraph feature $\omega(S) \equiv \frac{e}{(n-1)n} n^{\lambda}$

    → $\omega \propto \text{density} \cdot \text{order}$

    → λ: the weight of order vs. density
- Incorporate (pre-computed) ω into GNN propagation

GraphSNN

- The power of GraphSNN in modeling graph structures
  - Strictly more powerful than Weisfeiler-Leman / GNN
  - While having the same complexity

GraphSNN

Method	Cora	Citeseer	Pubmed	NELL	ogbn-arxiv
GCN	81.5 ± 0.4	70.3 ± 0.5	79.0 ± 0.5	66.0 ± 1.7	71.74 ± 0.29
GraphSNN$_{GCN}$	$\mathbf{83.1 \pm 1.8}$	$\mathbf{72.3 \pm 1.5}$	$\mathbf{79.8 \pm 1.2}$	$\mathbf{68.3 \pm 1.6}$	$\mathbf{72.20 \pm 0.90}$
GAT	83.0 ± 0.6	72.6 ± 0.6	78.5 ± 0.3	-	-
GraphSNN$_{GAT}$	$\mathbf{83.8 \pm 1.2}$	$\mathbf{73.5 \pm 1.6}$	$\mathbf{79.6 \pm 1.4}$	-	-
GIN	77.6 ± 1.1	66.1 ± 1.5	77.0 ± 1.2	61.5 ± 2.3	-
GraphSNN$_{GIN}$	$\mathbf{79.2 \pm 1.7}$	$\mathbf{68.3 \pm 1.5}$	$\mathbf{78.8 \pm 1.3}$	$\mathbf{63.8 \pm 2.7}$	-
GraphSAGE	79.2 ± 3.7	71.6 ± 1.9	77.4 ± 2.2	63.7 ± 5.2	71.49 ± 0.27
GraphSNN$_{GraphSAGE}$	$\mathbf{80.5 \pm 2.5}$	$\mathbf{72.7 \pm 3.2}$	$\mathbf{79.0 \pm 3.5}$	$\mathbf{66.3 \pm 5.6}$	$\mathbf{71.80 \pm 0.70}$

GraphSNN

Method	MUTAG	PTC-MR	PROTEINS	D&D	BZR	COX2	IMDB-B	RDT-M5K
WL	90.4 ± 5.7	59.9 ± 4.3	75.0 ± 3.1	79.4 ± 0.3	78.5 ± 0.6	81.7 ± 0.7	73.8 ± 3.9	52.5 ± 2.1
RetGK	90.3 ± 1.1	62.5 ± 1.6	75.8 ± 0.6	81.6 ± 0.3	-	-	71.9 ± 1.0	-
GNTK	90.0 ± 8.5	$\mathbf{67.9 \pm 6.9}$	75.6 ± 4.2	75.6 ± 3.9	83.6 ± 2.9	-	76.9 ± 3.6	-
P-WL	90.5 ± 1.3	64.0 ± 0.8	75.2 ± 0.3	78.6 ± 0.3	-	-	-	-
WL-PM	87.7 ± 0.8	61.4 ± 0.8	-	78.6 ± 0.2	-	-	-	-
WWL	87.2 ± 1.5	66.3 ± 1.2	74.2 ± 0.5	79.6 ± 0.5	84.4 ± 2.0	78.2 ± 0.4	74.3 ± 0.8	-
FGW	88.4 ± 5.6	65.3 ± 7.9	74.5 ± 2.7	-	85.1 ± 4.1	77.2 ± 4.8	63.8 ± 3.4	-
DGCNN	85.8 ± 1.7	58.6 ± 2.5	75.5 ± 0.9	79.3 ± 0.9	-	-	70.0 ± 0.9	48.7 ± 4.5
CapsGNN	86.6 ± 6.8	66.0 ± 1.8	76.2 ± 3.6	75.4 ± 4.1	-	-	73.1 ± 4.8	52.9 ± 1.5
†GraphSAGE	85.1 ± 7.6	63.9 ± 7.7	75.9 ± 3.2	72.9 ± 2.0	-	-	72.3 ± 5.3	50.0 ± 1.3
†GIN	89.4 ± 5.6	64.6 ± 7.0	75.9 ± 2.8	-	-	-	75.1 ± 5.1	57.5 ± 1.5
†GraphSNN (S)	$\mathbf{91.57 \pm 2.8}$	66.70 ± 3.7	$\mathbf{76.83 \pm 2.5}$	$\mathbf{81.97 \pm 2.6}$	$\mathbf{88.69 \pm 3.2}$	$\mathbf{82.86 \pm 3.1}$	$\mathbf{77.86 \pm 3.6}$	$\mathbf{58.43 \pm 2.3}$
†GraphSNN (R)	$\mathbf{91.24 \pm 2.5}$	66.96 ± 3.5	$\mathbf{76.51 \pm 2.5}$	$\mathbf{82.46 \pm 2.7}$	$\mathbf{88.97 \pm 2.9}$	$\mathbf{83.13 \pm 3.5}$	$\mathbf{76.93 \pm 3.3}$	$\mathbf{58.51 \pm 2.7}$
GraphSNN (S)	$\mathbf{94.70 \pm 1.9}$	$\mathbf{70.58 \pm 3.1}$	$\mathbf{78.42 \pm 2.7}$	$\mathbf{83.92 \pm 2.3}$	$\mathbf{91.12 \pm 3.0}$	$\mathbf{86.28 \pm 3.3}$	$\mathbf{78.51 \pm 2.8}$	$\mathbf{59.86 \pm 2.6}$
GraphSNN (R)	$\mathbf{94.14 \pm 1.2}$	$\mathbf{71.01 \pm 3.6}$	$\mathbf{78.21 \pm 2.9}$	$\mathbf{84.61 \pm 1.5}$	$\mathbf{91.88 \pm 3.2}$	$\mathbf{86.72 \pm 2.9}$	$\mathbf{77.87 \pm 3.1}$	$\mathbf{60.23 \pm 2.2}$

GraphSNN

Method	ogbg-molhiv	ogbg-moltox21	ogbg-moltoxcast	ogbg-ppa	ogbg-molpcba
GIN	75.58±1.40	74.91±0.51	63.41±0.74	68.92±1.00	22.66±0.28
GIN+VN	75.20±1.30	76.21±0.82	66.18±0.68	70.37±1.07	27.03±0.23
GSN	77.99±1.00	-	-	-	-
PNA	79.05±1.30	-	-	-	28.38±0.35
ID-GNN	78.30±2.00	-	-	-	-
Deep LRP	77.19±1.40	-	-	-	-
GraphSNN	78.51±1.70	75.45±1.10	65.40±0.71	70.66±1.65	24.96±1.50
GraphSNN+VN	**79.72±1.83**	**76.78±1.27**	**67.68±0.92**	**72.02±1.48**	**28.50±1.68**

GraphSNN

	Method	MUTAG	PTC-MR	PROTEINS	BZR	IMDB-B
GSN	GSN-e	90.6 ± 7.5	$\mathbf{68.2 \pm 7.2}$	76.6 ± 5.0	-	77.8 ± 3.3
	GSN-v	92.2 ± 7.5	67.4 ± 5.7	74.5 ± 5.0	-	76.8 ± 2.0
ID-GNNs	ID-GNN Fast	$\mathbf{96.5 \pm 3.2}$	61.9 ± 5.4	$\mathbf{78.0 \pm 3.5}$	86.4 ± 3.0	-
	ID-GNN Full	93.0 ± 5.6	62.5 ± 5.3	77.9 ± 2.4	88.1 ± 4.0	-
Ours	GraphSNN	91.57 ± 2.8	66.70 ± 3.7	76.83 ± 2.5	$\mathbf{88.69 \pm 3.2}$	$\mathbf{77.86 \pm 3.6}$
k-WL GNNs	1-GNN$_{NT}$	82.7 ± 0.0	51.2 ± 0.0	-	-	69.4 ± 0.0
	1-GNN	82.2 ± 0.0	59.0 ± 0.0	-	-	71.2 ± 0.0
	1-2-3-GNN$_{NT}$	84.4 ± 0.0	59.3 ± 0.0	-	-	70.3 ± 0.0
	1-2-3-GNN	86.1 ± 0.0	60.9 ± 0.0	-	-	74.2 ± 0.0
Ours	GraphSNN	$\mathbf{87.30 \pm 3.1}$	$\mathbf{61.63 \pm 2.8}$	$\mathbf{74.01 \pm 3.2}$	$\mathbf{82.72 \pm 3.9}$	$\mathbf{74.81 \pm 3.5}$