Weisfeiler-Lehman Meets Gromov-Wasserstein

Samantha Chen[1], Sunhyuk Lim[2], Facundo Mémoli[3], **Zhengchao Wan**[1] and Yusu Wang[1]

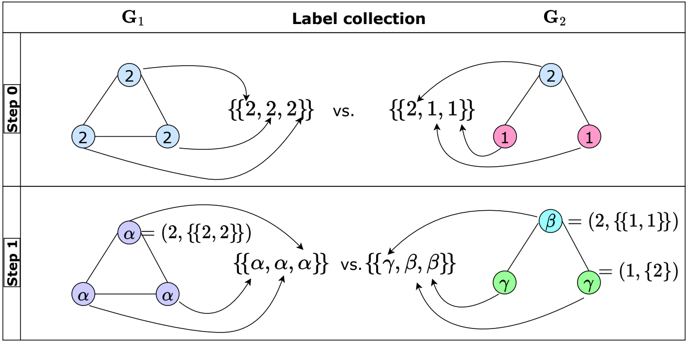
[1] University of California San Diego
[2] Max Planck Institute for Mathematics in the Sciences
[3] The Ohio State University

Weisfeiler-Lehman Graph Isomorphism Test

- A procedure to test whether two labeled graphs (i.e., graphs with node features, called labels here) are potentially isomorphic or not
- Interesting connection to graph neural networks (GNNs)

Weisfeiler-Lehman Graph Isomorphism Test

- A procedure to test whether two labeled graphs (i.e., graphs with node features, called labels here) are potentially isomorphic or not
- Interesting connection to graph neural networks (GNNs)



Weisfeiler-Lehman Graph Isomorphism Test

Note

WL test is only an isomorphism indicator. It does not offer a meaningful distance between graphs.

Weisfeiler-Lehman Graph Isomorphism Test

Note

WL test is only an isomorphism indicator. It does not offer a meaningful distance between graphs.

Goal

Define a **distance** between graphs with labels in **metric-spaces**[a] so that the distance has the **same** discriminative power as the WL test.

[a] we compare not just graph combinatorial structures, but also node features (called labels here)

Weisfeiler-Lehman Graph Isomorphism Test

Note

WL test is only an isomorphism indicator. It does not offer a meaningful distance between graphs.

Goal

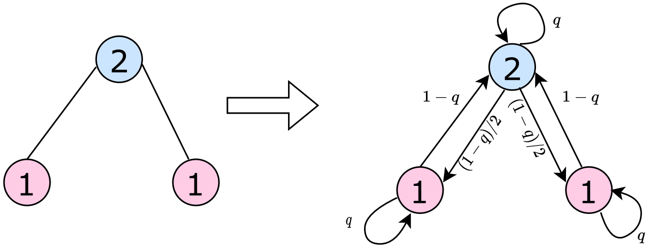
Define a **distance** between graphs with labels in **metric-spaces**[a] so that the distance has the **same** discriminative power as the WL test.

[a] we compare not just graph combinatorial structures, but also node features (called labels here)

- Wesifeiler-Lehman graph kernels (Shervashidze et al., 2011)
  - only for categorical labels
- Wasserstein Weisfeiler-Lehman (WWL) graph kernels (Togninalli et al., 2019)
  - for both categorical and Euclidean labels
  - less discriminative than the WL test when considering Euclidean labels

The Weisfeiler-Lehman distance

- Labeled graph $\rightarrow$ Labeled measure **Markov chain** (LMMC) $(X, m^X, \mu_X, \ell_X : X \rightarrow Z)$



The Weisfeiler-Lehman distance

- Labeled graph → Labeled measure Markov chain (LMMC)
- Multisets → Probability measures

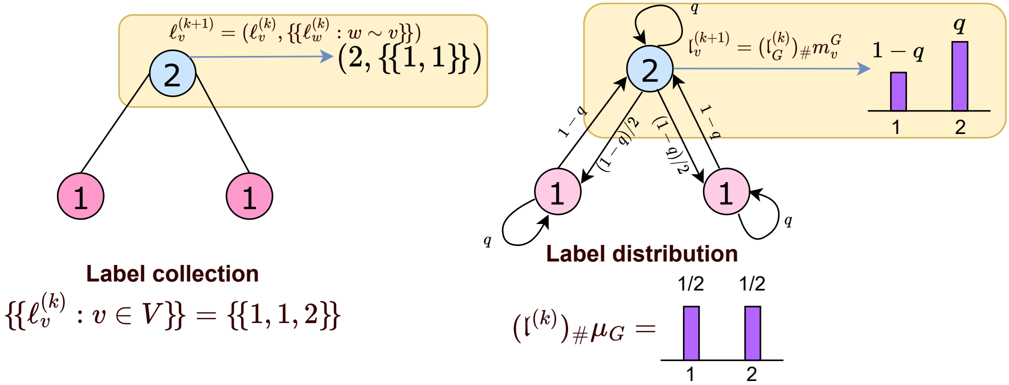


Figure: Update node labels as local label distributions through transition probabilities.

The Weisfeiler-Lehman distance

- Labeled graph → Labeled measure Markov chain (LMMC)
- Multisets → Probability measures

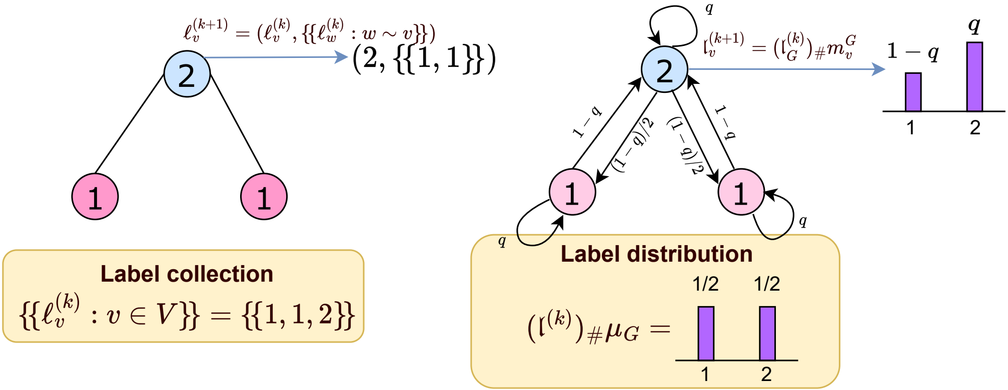
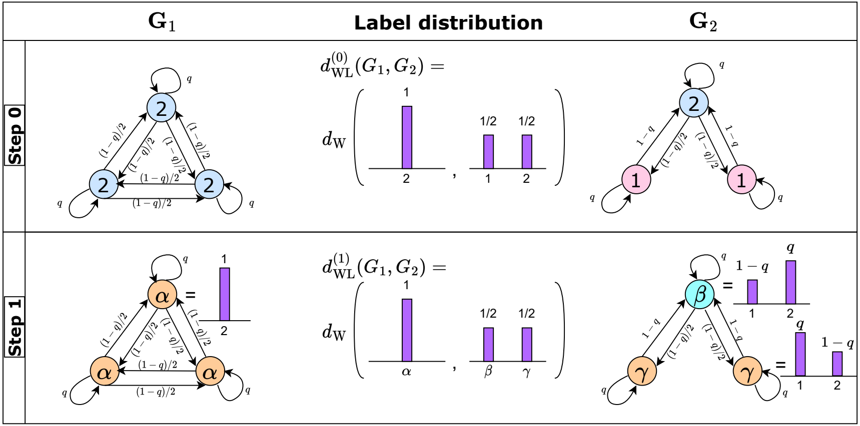


Figure: Generate the label distribution via the stationary distribution μ_G

The Weisfeiler-Lehman distance

- Labeled graphs → Labeled measure Markov chains (LMMCs)
- Multisets → Probability measures
- Comparing label distributions via Optimal Transport; $d_{\mathrm{WL}} = \sup_{k\geq 0} d_{\mathrm{WL}}^{(k)}$



The Weisfeiler-Lehman distance

Theorem (Informal)

- d_{WL} defines a pseudo-distance on the collection of LMMCs.
- $d_{\mathrm{WL}} = 0$ iff the two labeled graphs pass the WL test (up to a mild change of labels).

The Weisfeiler-Lehman distance

Theorem (Informal)

- d_{WL} defines a pseudo-distance on the collection of LMMCs.
- $d_{\mathrm{WL}} = 0$ iff the two labeled graphs pass the WL test (up to a mild change of labels).

Proposition

d_{WL} admits efficient lower bounds (including the WWL graph kernel distance)

The Weisfeiler-Lehman distance

Theorem (Informal)

- d_{WL} defines a pseudo-distance on the collection of LMMCs.
- $d_{\mathrm{WL}} = 0$ iff the two labeled graphs pass the WL test (up to a mild change of labels).

Proposition

d_{WL} admits efficient lower bounds (including the WWL graph kernel distance)

Method	MUTAG	PROTEINS	PTC-FM	PTC-MR	IMDB-B	IMDB-M	COX2
$d_{\mathrm{WL}}^{(k)}$	**92.1 ± 6.3**	63.0 ± 3.5	62.2 ± 8.5	56.2 ± 6.3	**70.0 ± 4.3**	**41.3 ± 4.8**	76.1 ± 5.5
$d_{\mathrm{WLLB}}^{(k)}$	87.3 ± 1.9	**66.2 ± 2.2**	**62.5 ± 8.5**	**57.8 ± 6.8**	69.9 ± 2.5	40.6 ± 3.8	**81.2 ± 5.3**
WWL	85.1 ± 6.5	64.7 ± 2.8	58.2 ± 8.5	54.3 ± 7.9	65.0 ± 3.3	40.0 ± 3.3	76.1 ± 5.6

Further Theoretical Justifications

Markov chain neural networks (MCNNs)

- We identify a neural network structure for LMMCs.
- MCNNs reduce to **standard GNNs** when restricted to labeled graphs.
- MCNNs have the **same** discriminative power as the WL distance and thus the WL test.
- MCNNs are **universal** w.r.t. the WL distance.

Further Theoretical Justifications

Markov chain neural networks (MCNNs)

- We identify a neural network structure for LMMCs.
- MCNNs reduce to **standard GNNs** when restricted to labeled graphs.
- MCNNs have the **same** discriminative power as the WL distance and thus the WL test.
- MCNNs are **universal** w.r.t. the WL distance.

Weisfeiler-Lehman meets Gromov-Wasserstein

- Define a variant of **Gromov-Wasserstein** distance d_{GW} between Markov chain metric spaces.
- d_{GW} is a **proper** distance, i.e., $d_{\mathrm{GW}} = 0 \Longleftrightarrow$ isomorphism between Markov chain metric spaces.
- d_{WL} is **stable** w.r.t. d_{GW} (the ground truth distance between Markov chains).

Thank You

Please see our paper for more details:

https://arxiv.org/abs/2202.02495