

A Theoretical and Empirical Investigation into the Equivalence of Graph Neural Networks and the Weisfeiler-Leman Algorithm

From the faculty of Mathematics, Physics, and Computer Science for the purpose of obtaining the
academic degree of Bachelor of Sciences.

Eric Tillmann Bill

Supervision:

Prof. Dr. rer. nat. Christopher Morris

Informatik 6
RWTH Aachen University

Contents

1 Definition

Definition 1. For any graphs G, H we will denote $G \simeq_{1\text{WL}} H$ if the 1-WL isomorphism test can not distinguish both graphs. Note that due to the soundness of this algorithm, if $G \not\simeq_{1\text{WL}} H$, we always can conclude that $G \not\cong H$.

Definition 2. Let \mathcal{C} be a collection of permutation invariant functions from $\mathcal{X}^{n \times n}$ to \mathbb{R} . We say \mathcal{C} is **1-WL-Discriminating** if for all graphs $G_1, G_2 \in \mathcal{X}$ for which the 1-WL isomorphism test concludes non-isomorphic, there exists a function $h \in \mathcal{C}$ such that $f(G_1) \neq f(G_2)$.

Definition 3. Let \mathcal{C} be a collection of permutation invariant functions from $\mathcal{X}^{n \times n}$ to \mathbb{R} . We say \mathcal{C} is **GNN-Approximating** if for all permutation-invariant functions \mathcal{A} computed by a GNN, and for all $\epsilon \in \mathbb{R}$ with $\epsilon > 0$, there exists $h_{\mathcal{A}, \epsilon} \in \mathcal{C}$ such that $\|\mathcal{A} - h_{\mathcal{A}, \epsilon}\|_\infty := \sup_{G \in \mathcal{X}} |f(G) - h_{\mathcal{A}, \epsilon}(G)| < \epsilon$

2 Theorems

In this thesis we concentrate on finite graphs. We therefore, let \mathcal{X} be $\mathcal{X} := \{1, \dots, k\}$ for an arbitrary $k \in \mathbb{N}$.

Theorem 4. Let \mathcal{C} be a collection of functions computed by 1-WL+NN models, then \mathcal{C} is **1-WL-Discriminating**.

Theorem 5. Let \mathcal{C} be a collection of functions that is **1-WL-Discriminating**, then \mathcal{C} is also **GNN-Approximating**.

3 Proofs

Since we concentrate in this thesis on finite graphs, let \mathcal{X} be $\mathcal{X} := \{1, \dots, k\}$ for an arbitrary $k \in \mathbb{N}$.

3.1 1-WL+NN is 1-WL-Discriminating

Proof. We consider the collection \mathfrak{B}_k of permutation invariant functions, where every $\mathcal{B} \in \mathfrak{B}_k$ is of the form $\mathcal{B} = \text{MLP} \circ f$. Here MLP is an arbitrary multilayer perceptron mapping vectors from \mathbb{N}^K to \mathbb{R} and f the *counting-encoding* function. Further, let $G_1, G_2 \in \mathcal{X}^{n \times n}$ such that the 1-WL isomorphism test concludes non-isomorphic ($G_1 \not\cong G_2$). We denote with $(C_\infty)_G$ the final coloring computed by the 1-WL algorithm when applied on G .

Due to the 1-WL isomorphism test concluding $G_1 \not\cong G_2$, there exists a color $c \in \mathbb{N}$ such that $(C_\infty)_{G_1}(c) \neq (C_\infty)_{G_2}(c)$. If we now consider as MLP the following function $\text{MLP} : \mathbb{N}^K \rightarrow \mathbb{R}, v \mapsto W \cdot v$ with $W \in \mathbb{N}^{1 \times K}$ such that $W_{1,c} := 1$ and $W_{1,i} := 0$ for all $i \in [K] \setminus \{c\}$. Then we can conclude that $\mathcal{B}(G_1) \neq \mathcal{B}(G_2)$. Since G_1, G_2 are arbitrary, we can conclude the proof. \square

3.2 1-WL-Discriminating is also GNN-Approximating

Lemma 6. Let \mathcal{C} be a collection of permutation-invariant functions computed by 1-WL+NN. Further, $h_1, \dots, h_n \in \mathcal{C}$ and MLP a multilayer perceptron, then the function \mathcal{A} composed of

$\mathcal{A}(G^*) := \text{MLP}(h_1(G^*), \dots, h_n(G^*))$ is already contained in \mathcal{C} with $\mathcal{A} \in \mathcal{C}$, where we define the weight matrix as follows:

$$W_{i,j} := \begin{cases} 1, & \text{if } (i \bmod K) = j \\ 0, & \text{else} \end{cases}$$

Then we can build $\mathcal{A} :=$

Proof. Assume the above and let $f : \mathcal{X} \rightarrow \mathbb{N}^K$ be the encoding function used by the collection \mathcal{C} . Further, let $\text{MLP}_1, \dots, \text{MLP}_n$ be the multilayer perceptrons the functions h_1, \dots, h_n are composed of respectively. Then we define a single layer MLP M , with $M : \mathbb{N}^K \rightarrow \mathbb{N}^{n \cdot K}, v \mapsto W \cdot v$ with $W \in \mathbb{N}^{n \cdot K}$ \square

Lemma 7 (My Lemma). Let \mathcal{C} be a collection of permutation-invariant functions from $\mathcal{X}^{n \times n}$ to \mathbb{R} that is 1-WL-Discriminating, then for all $G \in \mathcal{X}^{n \times n}$, there exists a function $h_G \in \mathcal{C}$ such that for all $G^* : h_G(G^*) = 0$ if and only if $G \simeq_{1\text{WL}} G^*$.

Proof. For any $G_1, G_2 \in \mathcal{X}^{n \times n}$ with $G_1 \not\simeq_{1\text{WL}} G_2$ let $h_{G_1, G_2} \in \mathcal{C}$ be the function distinguishing them, with $h_{G_1, G_2}(G_1) \neq h_{G_1, G_2}(G_2)$. We define the function \bar{h}_{G_1, G_2} to work on $G^* \in \mathcal{X}^{n \times n}$ as follows:

$$\begin{aligned} \bar{h}_{G_1, G_2} &= |h_{G_1, G_2}(G^*) - h_{G_1, G_2}(G_1)| \\ &= \max(h_{G_1, G_2}(G^*) - h_{G_1, G_2}(G_1)) + \max(h_{G_1, G_2}(G_1) - h_{G_1, G_2}(G^*)) \end{aligned}$$

Note, that $h_{G_1, G_2}(G_1)$ in the formula is a fixed constant. With this function we can ensure, that if $G^* \simeq_{1\text{WL}} G : \bar{h}_{G_1, G_2}(G^*) = 0$, otherwise if $G^* \not\simeq_{1\text{WL}} G : \bar{h}_{G_1, G_2}(G^*) > 0$.

We can than define the function h_{G_1} working over all graphs $G^* \in \mathcal{X}^{n \times n}$ as follows:

$$h_{G_1}(G^*) := \sum_{G_2 \in \mathcal{X}^{n \times n}, G_1 \simeq_{1\text{WL}} G_2} \bar{h}_{G_1, G_2}(G^*).$$

We will quickly demonstrate that this function satisfies the condition of the lemma, let G_1 be fixed and G^* an arbitrary graph from $\mathcal{X}^{n \times n}$:

1. If $G_1 \simeq_{1\text{WL}} G^*$, then there exists n
2. If $G_1 \not\simeq_{1\text{WL}} G^*$, then

TODO: can be easily encoded in a single MLP layer. \square

Lemma 8. Let \mathcal{C} be a class of permutation-invariant functions from $\mathcal{X}^{n \times n}$ to \mathbb{R} so that for all $G \in \mathcal{X}^{n \times n}$, there exists $h_G \in \mathcal{C}$ satisfying $h_G(G^*) = 0$ if and only if $G \simeq G^*$ for all $G^* \in \mathcal{X}^{n \times n}$. Then for every $G \in \mathcal{X}^{n \times n}$, there exists a function $\phi_G \in \mathcal{C}$ such that for all $G^* \in \mathcal{X}^{n \times n}$: $\phi_G(G^*) = \mathbb{1}_{G \simeq_{1\text{WL}} G^*}$.

Proof. Following ??, for every $G \in \mathcal{X}^{n \times n}$ there exists $\tilde{h}_G \in \mathcal{C}$ such that for all $G^* \in \mathcal{X}^{n \times n}$: $\tilde{h}_G(G^*) = 0$ if and only if $G \simeq_{1\text{WL}} G^*$. Due to \mathcal{X} being finite, we can define for every graph G the constant:

$$\delta_G := \frac{1}{2} \min_{G^* \in \mathcal{X}^{n \times n}, G \not\simeq_{1\text{WL}} G^*} |\tilde{h}_G(G^*)| > 0.$$

Actually
in \mathcal{C}^{+1}

With this constant, we can use a so-called “bump” function working from \mathbb{R} to \mathbb{R} that will be similar to the indicator function. For parameter $a \in \mathbb{R}$ with $a > 0$ let:

$$\psi_a(x) := \max\left(\frac{x}{a} - 1, 0\right) + \max\left(\frac{x}{a} + 1, 0\right) - 2 \cdot \max\left(\frac{x}{a}, 0\right).$$

The interesting property of ψ_a is that it maps every value x to 0, except for x being drawn from the interval $(-a, a)$. We use this property to define for every graph $G \in \mathcal{X}^{n \times n}$ the function $\varphi_G(G^*) := \psi_{\delta_G}(h_G(G^*))$. We will quickly demonstrate that this function is equal to the indicator function, let G be fixed and G^* an arbitrary graph from $\mathcal{X}^{n \times n}$:

1. If $G \simeq_{1\text{WL}} G^*$, then $h_G(G^*) = 0$ resulting in $\varphi_G(G^*) = \psi_{\delta_G}(0) = 1$.
2. If $G \not\simeq_{1\text{WL}} G^*$ then $h_G(G^*) > 0$, such that $|h_G(G^*)| > \delta_G$ resulting in $\varphi_G(G^*) = 0$.

Note we can encode each φ_G via a single MLP layer, where δ_G is a constant and the max operator is replaced by the non-linear activation function ReLU of the layer. \square

Lemma 9. Let \mathcal{C} be a class of permutation-invariant functions from $\mathcal{X}^{n \times n}$ to \mathbb{R} so that for all $G \in \mathcal{X}^{n \times n}$, there exists $\varphi_G \in \mathcal{C}$ satisfying $\varphi_G(G^*) = \mathbb{1}_{G \simeq_{1\text{WL}} G^*}$ for all $G^* \in \mathcal{X}^{n \times n}$, then C^{+1} is universally approximating.

Proof. Let \mathcal{C} be a collection of permutation invariant functions from $\mathcal{X}^{n \times n}$ to \mathbb{R} . For any function \mathcal{A} computed by an GNN, we want to show that we can decompose it as follows for any $G^* \in \mathcal{X}^{n \times n}$:

$$\begin{aligned} \mathcal{A}(G^*) &= \left(\frac{1}{|\mathcal{X}^{n \times n} / \simeq_{1\text{WL}}(G^*)|} \sum_{G \in \mathcal{X}^{n \times n}} \mathbb{1}_{G \simeq_{1\text{WL}} G^*} \right) \cdot \mathcal{A}(G^*) \\ &= \frac{1}{|\mathcal{X}^{n \times n} / \simeq_{1\text{WL}}(G^*)|} \sum_{G \in \mathcal{X}^{n \times n}} \mathcal{A}(G) \cdot \mathbb{1}_{G \simeq_{1\text{WL}} G^*} \\ &= \sum_{G \in \mathcal{X}^{n \times n}} \frac{\mathcal{A}(G)}{|\mathcal{X}^{n \times n} / \simeq_{1\text{WL}}(G^*)|} \cdot \varphi_G(G^*) \end{aligned} \tag{0.1}$$

with $\mathcal{X}^{n \times n} / \simeq_{1\text{WL}}(G^*)$ denoting the set of all graphs G over $\mathcal{X}^{n \times n}$ that are equivalent to G^* according to the $\simeq_{1\text{WL}}$ relation.

Since \mathcal{A} is permutation-invariant, and GNNs are at most as good as the 1-WL algorithm in distinguishing non-isomorphic graphs, we can use the fact that for every graph $G, H \in \mathcal{X}^{n \times n}$ with $G \simeq_{1\text{WL}} H$: $\mathcal{A}(G) = \mathcal{A}(H)$. Therefore, we can decompose \mathcal{A} as outlined above. We can encode this decomposition in a single MLP layer with $\frac{\mathcal{A}(G)}{|\mathcal{X}^{n \times n} / \simeq_{1\text{WL}}(G^*)|}$ being and $\varphi_G \in \mathcal{C}$ encoding the indicator function. Important to note, we can only do this since \mathcal{X} is finite, making the overall sum finite and the size of $\mathcal{X}^{n \times n} / \simeq_{1\text{WL}}(G^*)$ well-defined. \square

1-WL
Discriminative
and gnn
approximation
better
style
and
command