

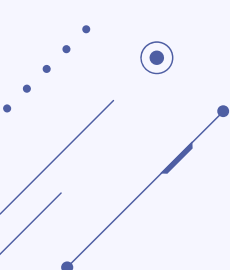
Approximation de la Graph Edit Distance (GED) par architecture transformeur

Réalisé par: Noé GILLE

Doctorant référant: Aldo MOSCATELLI



**UFR Sciences
et Techniques**



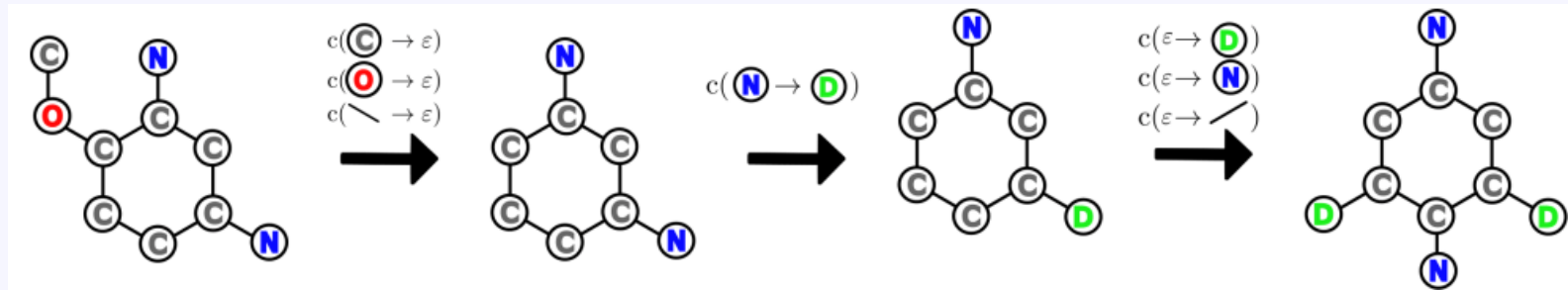
Plan

- I. GED et GNN
 - II. Pure transformers are powerful graph learners
 - III. Transformeur pour approximer la GED
 - IV. Résultats et perspectives
- 

Graph Edit Distance (GED)

- **Distance** entre deux graphes
- **Nombre d'opérations d'édits** pour passer d'un graphe à l'autre

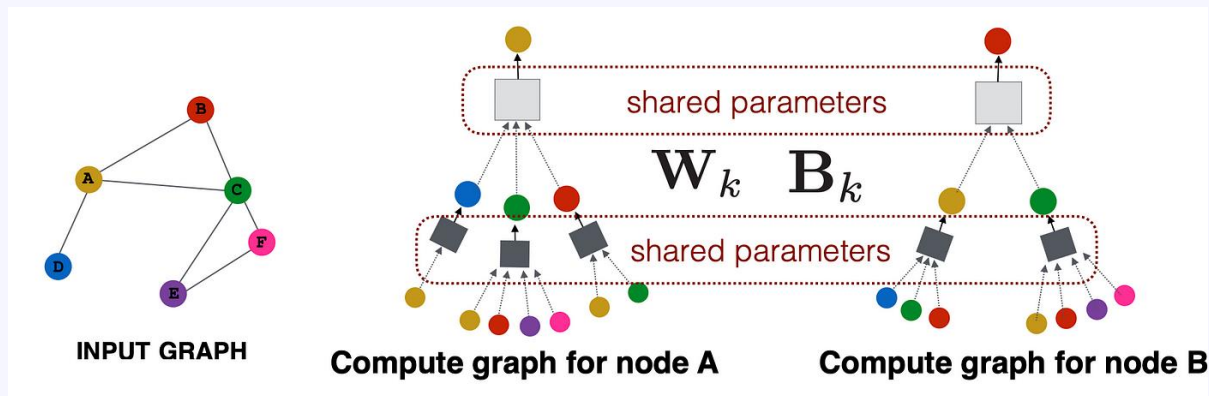
$$GED(g_1, g_2) = \min_{(e_1, \dots, e_k) \in \mathcal{P}(g_1, g_2)} \sum_{i=1}^k c(e_i)$$



* S. Adam & A. Moscatelli

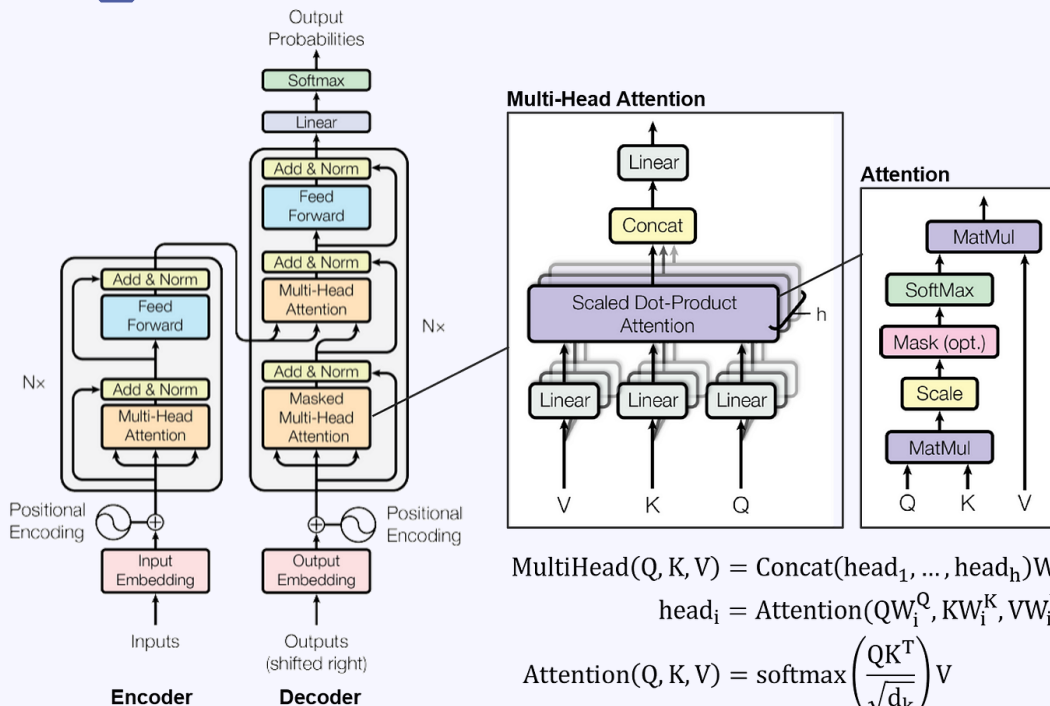
- Problème **NP-hard** $O(x^n)$
- **Applications** : Requêtage de base de données, KNN, Chimie computationnelle, ...

Graph Neural Networks (GNNs)



- **Invariance en permutation**
- **Global pooling:** Graphe de taille variable mais réponse de taille fixe
- **Message passing GNN** : A chaque itération, chaque nœud agrège de l'information sur ses voisins (Ex : GCN...)

Architecture transformeur



$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^O$$

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$$

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Architecture Transformeur

*Attention is all you need, Vaswani et al., 2017

Architecture transformeur

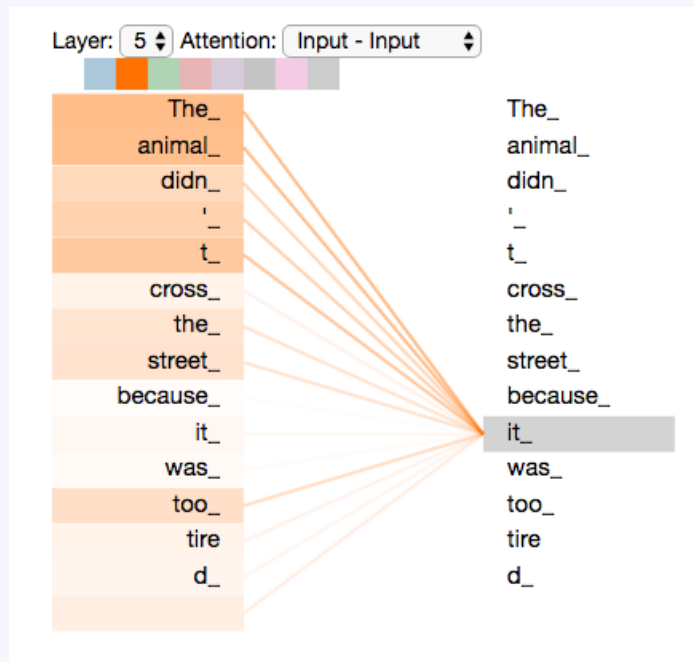
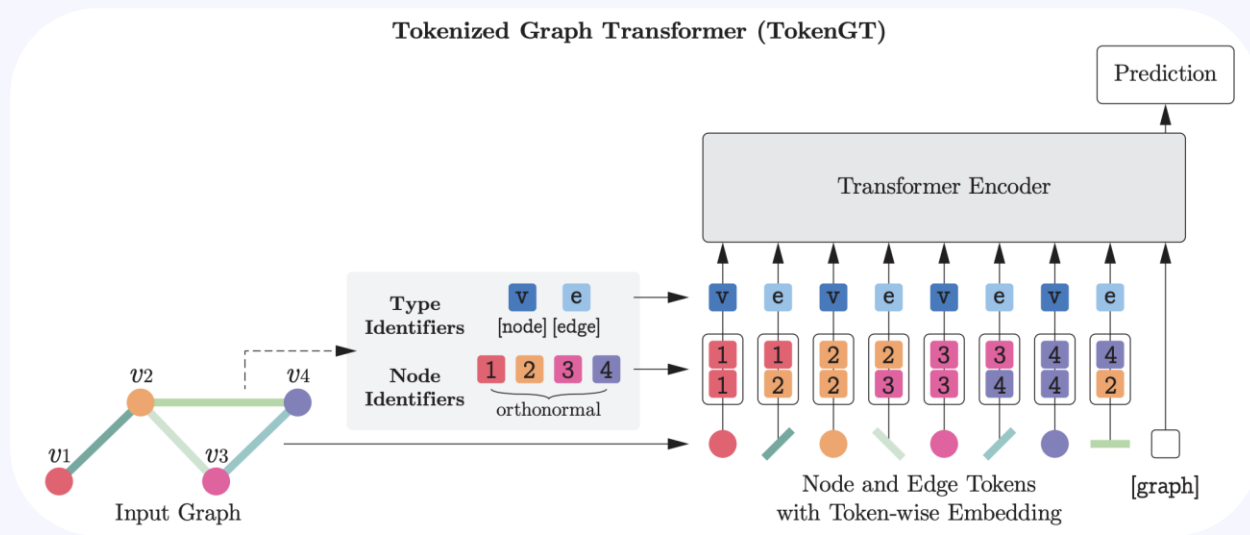


Illustration du mécanisme d'attention sur du texte

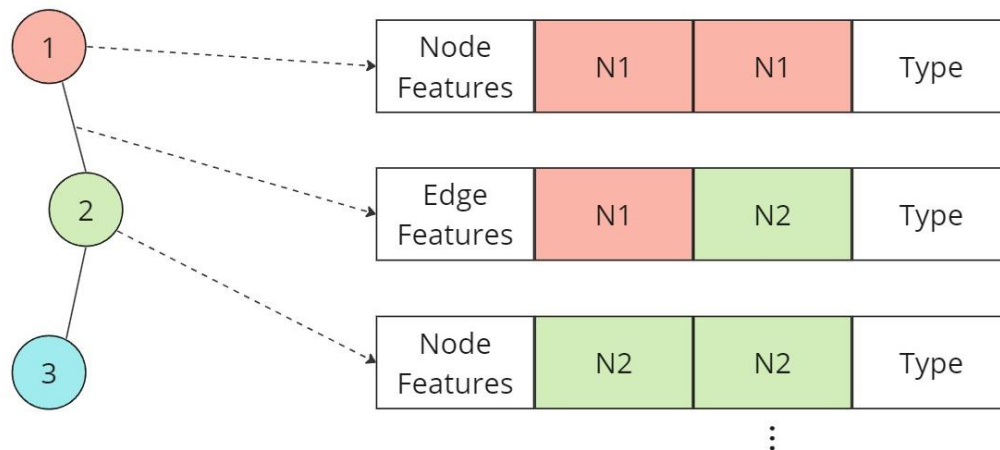
Tokenized Graph Transformer



Architecture Tokenized Graph Transformer

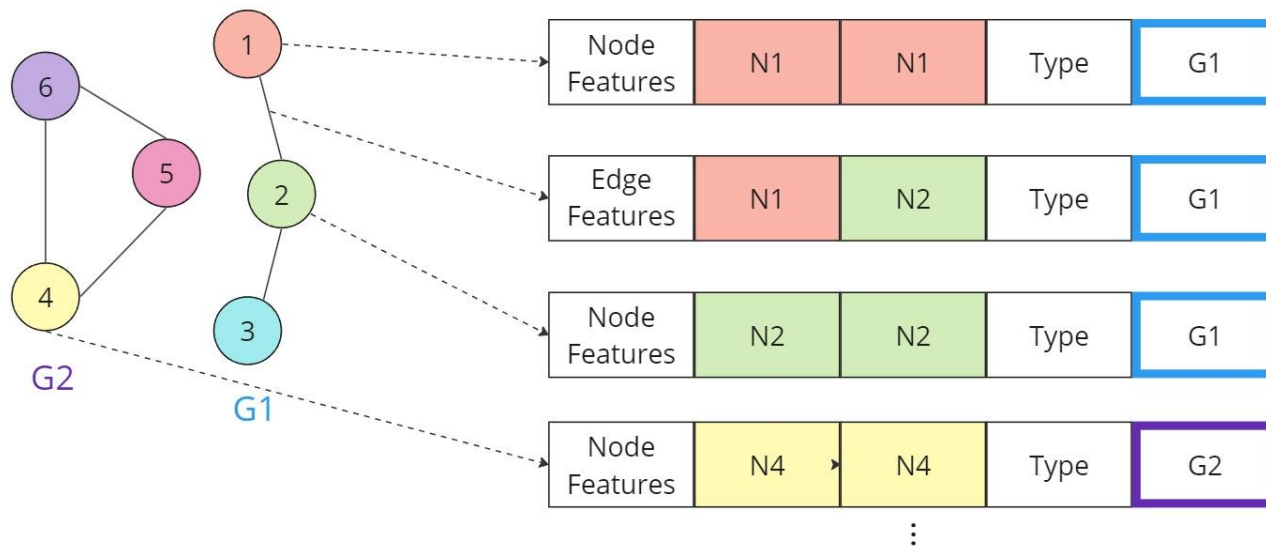
**Pure transformers are powerful graph learners, Kim et al, 2022*

Embedding



Création des Tokens dans TokenGT

Embedding



Création des Tokens dans TokenGT (modifié)

Embedding

L'identifiant de noeud encode la structure du graphe

- Orthogonalité $\Rightarrow |N| \leq d$
- Vecteurs propres du $L = I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$

Architecture siamoise :

- On traite les deux graphes ensemble
- **Symétrie de la distance** : $d(G1, G2) = d(G2, G1)$
- **L'Identifiant de graphe** est aussi **orthogonal**

Protocol experimental

Datasets utilisés:

Dataset	Nb graphes	Paires	Nœuds max	Features ?
AIDS700nef	700	490k	10	Oui
LINUX	1000	1M	10	Non
IMDBMulti	1500	2.25M	89	Non

Paramètres d'entraînements :

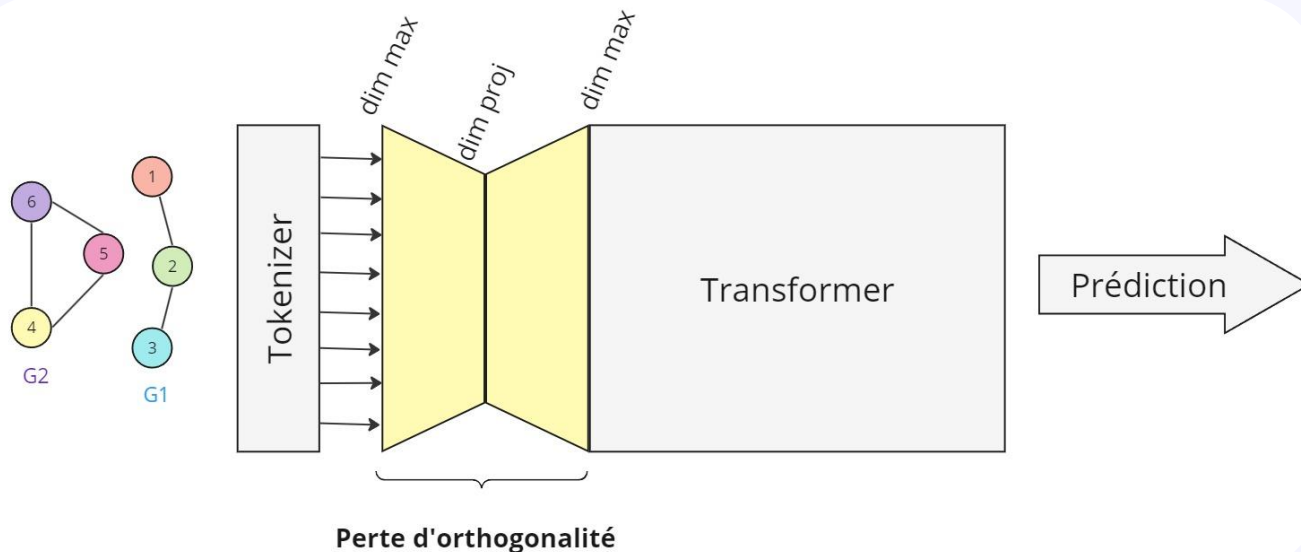
Model	Paramètres	heads	Dropout	epochs	Dimension Token	Batch size
AIDS700nef	121K	8	0,25	200	40	256
LINUX	121K	8	0,1	200	40	256
IMDBMulti	182K	8	0,1	?	356->20	6

Résultats

		GED		Rang			
		MSE	MAE	τ	ρ	p@10	p@20
LINUX	SimGNN	0,443	0,489	0,830	0,939	94,2%	93,3%
	GNOME	0,104	0,214	0,932	0,987	98,3%	99,1%
	TokenGT	0,069	0,052	0,910	0,969	99,8%	99,9%
AIDS700nef	SimGNN	1,075	0,816	0,690	0,843	42,1%	51,4%
	GNOME	0,555	0,490	0,863	0,959	81,3%	86,7%
	TokenGT	0,880	0,726	0,744	0,874	93,1%	91,9%
IMDBMulti	SimGNN	4389,06	28,082	0,770	0,878	75,9%	77,7%
	GNOME	33,433	2,976	0,953	0,994	87,7%	87,8%
	TokenGT (valid)	65,99	-	-	-	-	-

Expérience sur l'orthogonalité

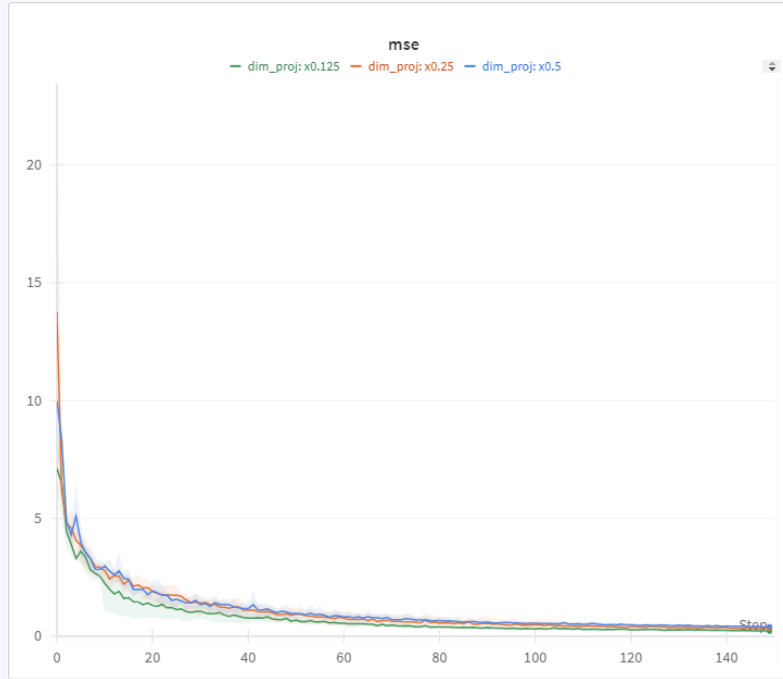
- L'**orthogonalité** est-elle strictement **nécessaire** ?



Principe de l'expérience

Expérience sur l'orthogonalité

- L'orthogonalité est-elle strictement **nécessaire** ?



Echelle linéaire

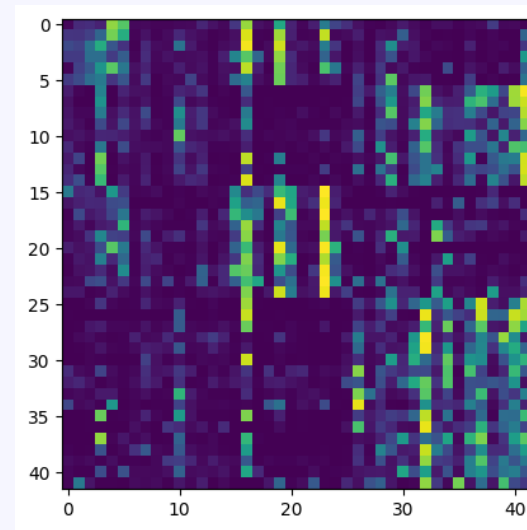
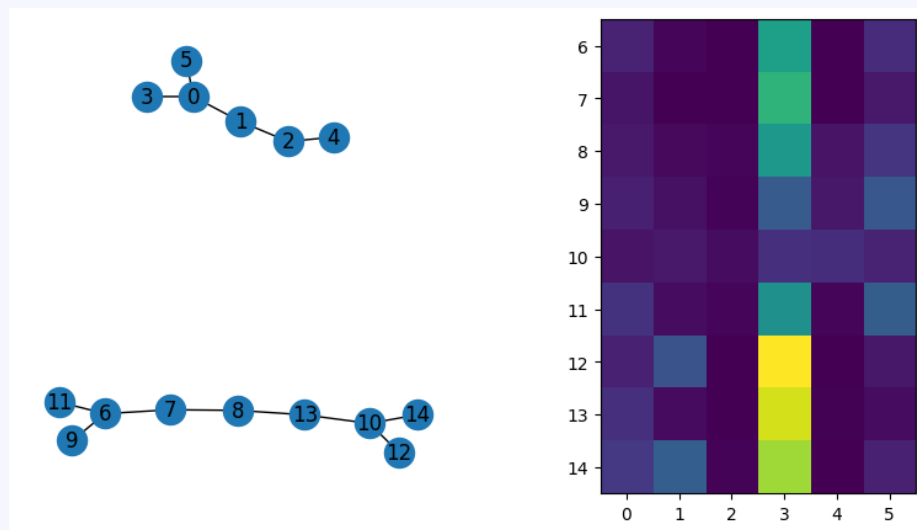


Echelle logarithmique

MSE (validation) sur LINUX avec plusieurs facteurs de projection

Recherche de l'explicabilité

- **Objectif** : Obtenir le **chemin d'édition**



Attention maps

Conclusion

- Le **mécanisme d'attention** est capable de traiter la structure complexe qui relie deux graphes lors du calcul de la **GED**
- De nombreuses **perspectives d'améliorations** sont envisageable par le choix du **Laplacien** ou de l'**hyper-paramétrisation** du modèle

Références

- [1] Yunsheng BAI et al. *SimGNN : A Neural Network Approach to Fast Graph Similarity Computation*. 2020. arXiv : 1808.05689 [cs.LG].
- [2] Jinwoo KIM et al. *Pure Transformers are Powerful Graph Learners*. 2022. arXiv : 2207.02505 [cs.LG].
- [3] Ashish VASWANI et al. *Attention Is All You Need*. 2023. arXiv : 1706.03762 [cs.CL].