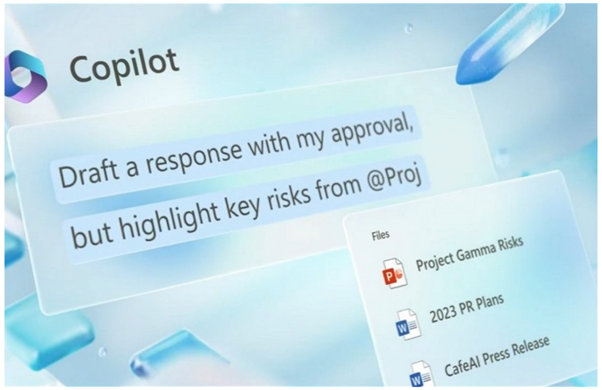
香港中文大學
The Chinese University of Hong Kong

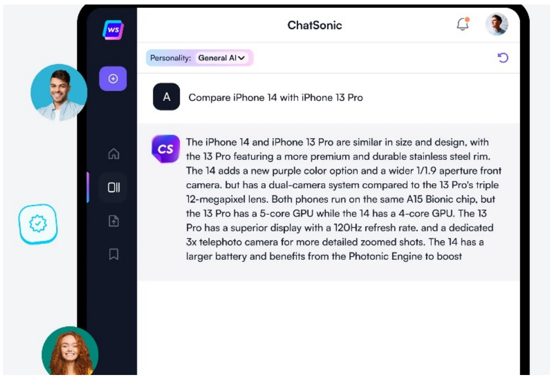
Does Graph Prompt Work? A Data Operation Perspective with Theoretical Analysis

Qunzhong WANG(王群中)
Department of Systems Engineering and Engineering Management
The Chinese University of Hong Kong

Artificial General Intelligence (AGI)

- **Artificial General Intelligence (AGI) has achieved huge success in NLP and CV areas.**
  - e.g., Copilot, ChatGPT, Midjourney

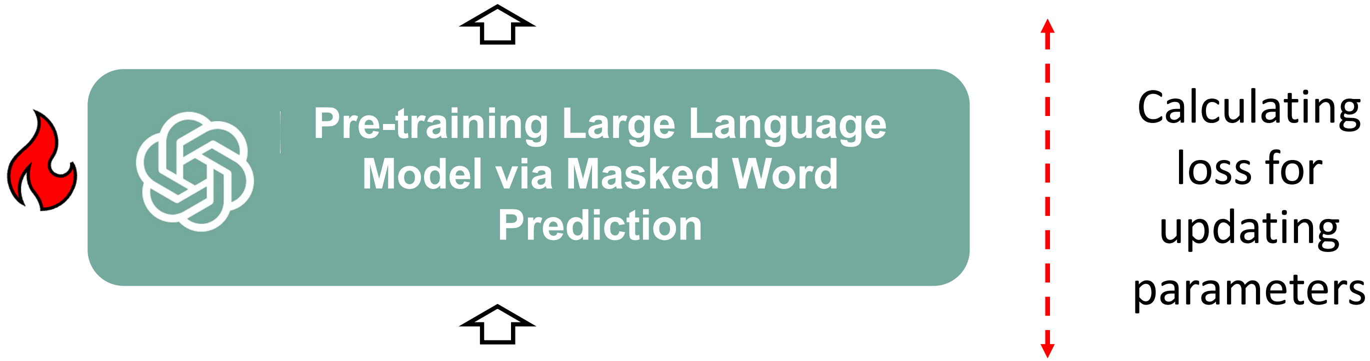




A Basic Workflow of AGI

- **Step 1: Pre-train a very large language model (LLM) via specific strategies.**
  - e.g., masked word prediction

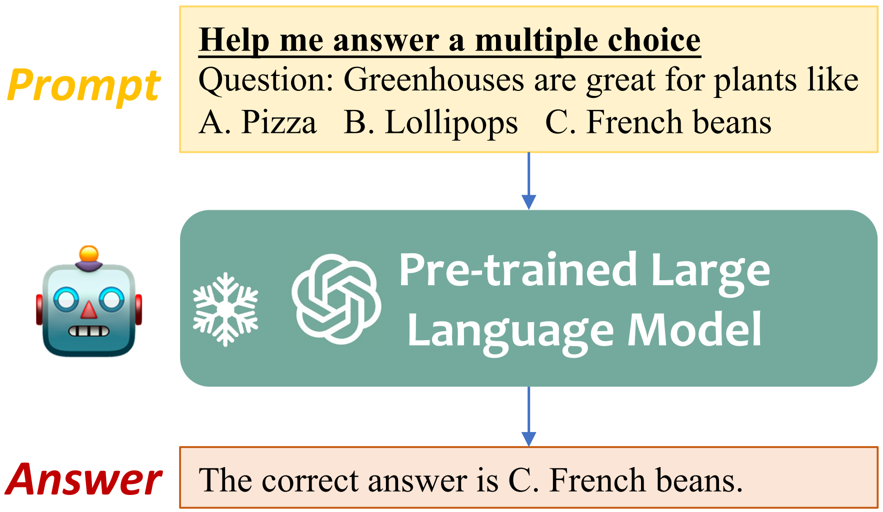
KDD2023 will **witness** many high-quality **papers**.



KDD2023 will **<Mask>** many high-quality **<Mask>.**

A Basic Workflow of AGI

➢ Step 2: Prompting a pre-trained LLM

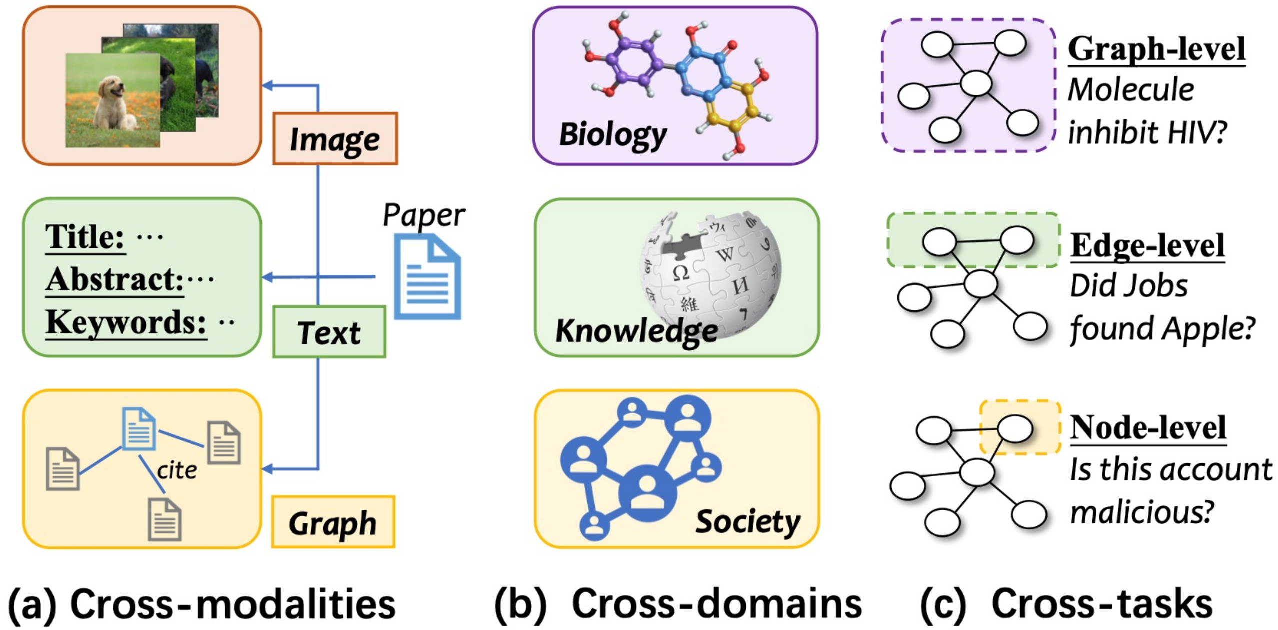


- A language prompt is a piece of text added to the beginning of an input text.
- The large language model can be pre-trained via next word prediction.

The question-answering task is reformulated to the word prediction task, which is consistent with the pre-training strategy, thus we do not need to tune LLM.

Graph AGI Still in the Early Stage

- **Why hard?**
  - Cross-modalities, cross-domains, cross-tasks
  - Social disputes: counterfactual outcomes, energy cost, etc.



Graph AGI Still in the Early Stage

- **How to solve these problems?**
  - Traditional fine-tuning approaches only adjust models, which is far from sufficient to solve graph problem as mentioned before.
  - To solve these problems, we need to focus on data-level operation, studying how to manipulate graph data directly
  - How to learn data manipulation strategy beyond manually designing?

We offer **Graph Prompt**, which is proved to be a **Parametric Approach** to simulate various graph data operations. (e.g. removing/adding links, nodes, subgraphs, changing features etc.)

Fine-tune vs Prompt

- **Fine-tune**
  - Need to tune the large pre-trained model (inefficient)
  - Do not change data
  - Limited task generalization
- **Prompt**
  - Freeze the large pre-trained model (efficient)
  - Have the capability of reformulating data
  - More general cross-tasks

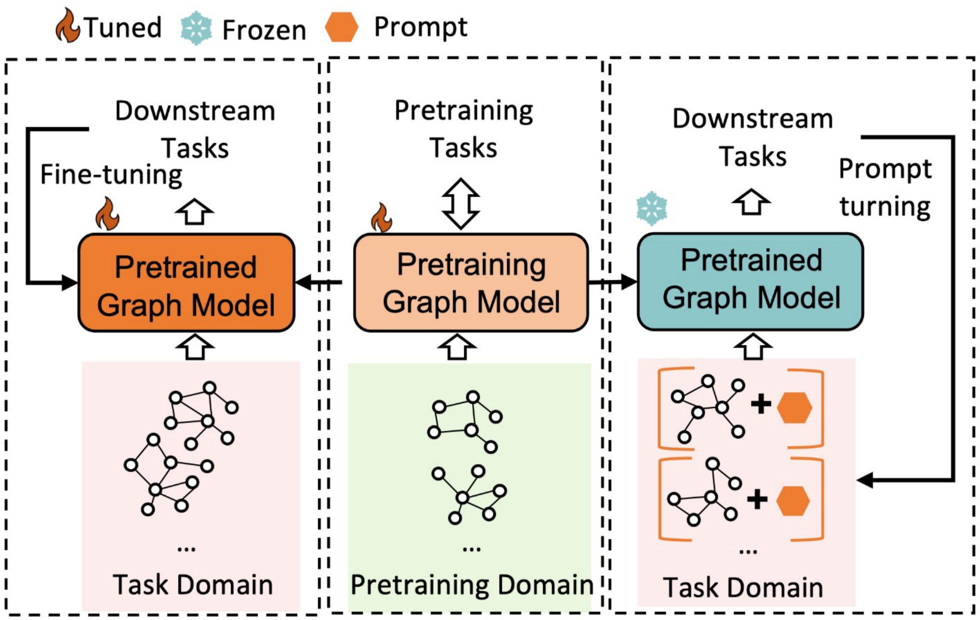


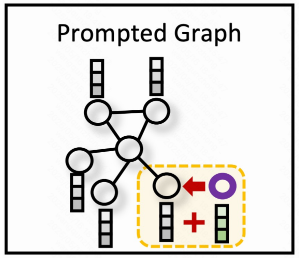
Figure 1: Fine-tuning, Pre-training, and Prompting.

Graph Prompt works perfectly

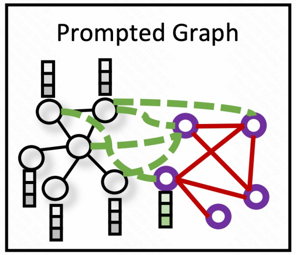
- **A great deal of work exist here:**
  - All in One: Multi-Task Prompting for Graph Neural Networks
  - Universal Prompt Tuning for Graph Neural Networks.
  - GPPT: Graph Pre-training and Prompt Tuning to Generalize Graph Neural Networks
  - GraphPrompt: Unifying Pre-Training and Downstream Tasks for Graph Neural Networks
- **In a nutshell, they implement graph prompts in different ways like features, subgraphs**
- **Yet to be answered: are Graph Prompts, as fine-tuning of graph data, powerful enough?**
- **In our paper, our direct response is "yes!"**

What Did we do

- **Framework of analysis: we selected "GPF" & "All in one" as two representatives**
  - GPF, from *Universal Prompt Tuning for Graph Neural Networks*.
    Core idea: train a prompt vector and add to each node.



  - Allinone, from the paper All in One: Multi-Task Prompting for Graph Neural Networks.
    core idea: train a prompt subgraph along with concatenation parameters to control the connection.



- **Our insight:**
  - By applying these modifications to the original graph, we aim to achieve significant degrees of freedom in the GNN's output
  - These degrees of freedom should be sufficient to cover arbitrary graph transformations.

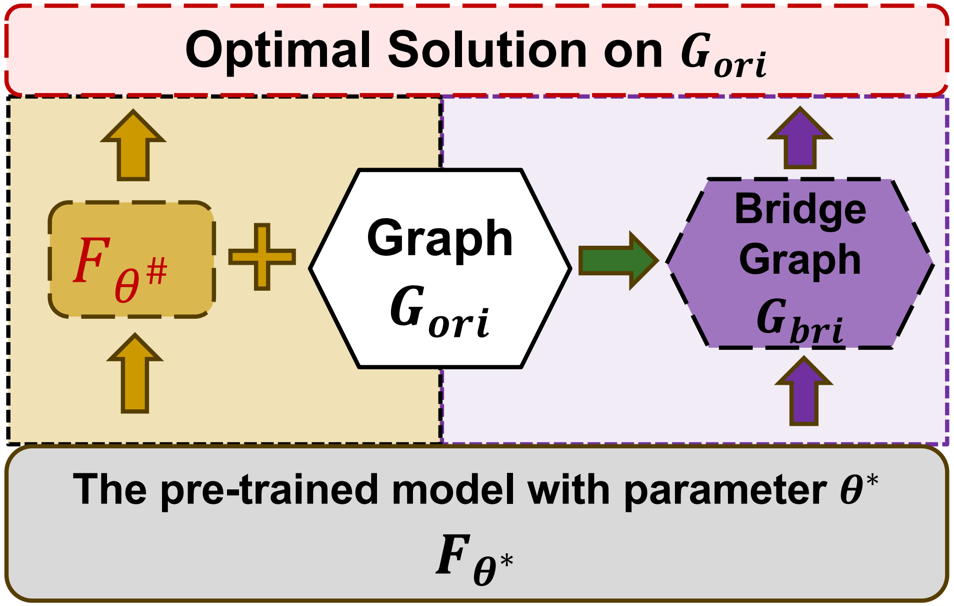
Why graph prompt works? A data operation perspective

Perspective from model tuning

$$F_{\theta^* \rightarrow \theta^\#}(G_{ori}) \rightarrow C(G_{ori})$$

Perspective from data operation

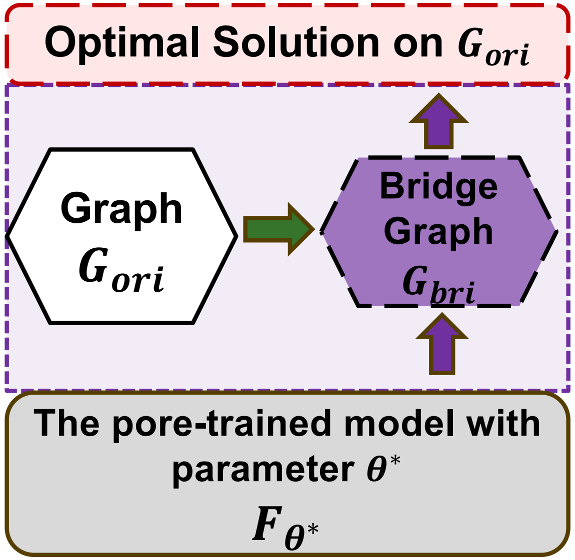
$$F_{\theta^*}(G_{bri}) = C(G_{ori})$$



- Before graph prompts, we only have one way to downstream tasks (fine-tune the pre-trained models)
- Graph prompting "***builds a new road to Rome***".
- The nature of graph prompt is to find a bridge graph for the given graph, filling the gap between F_{θ^*} and downstream tasks.

Why graph prompt works? A data operation perspective

- **Our 1st theory contribution:**
  - Bridge graphs always exist!
- **Our next question:**
  - *How difficult to find such a bridge graph using graph prompts?*



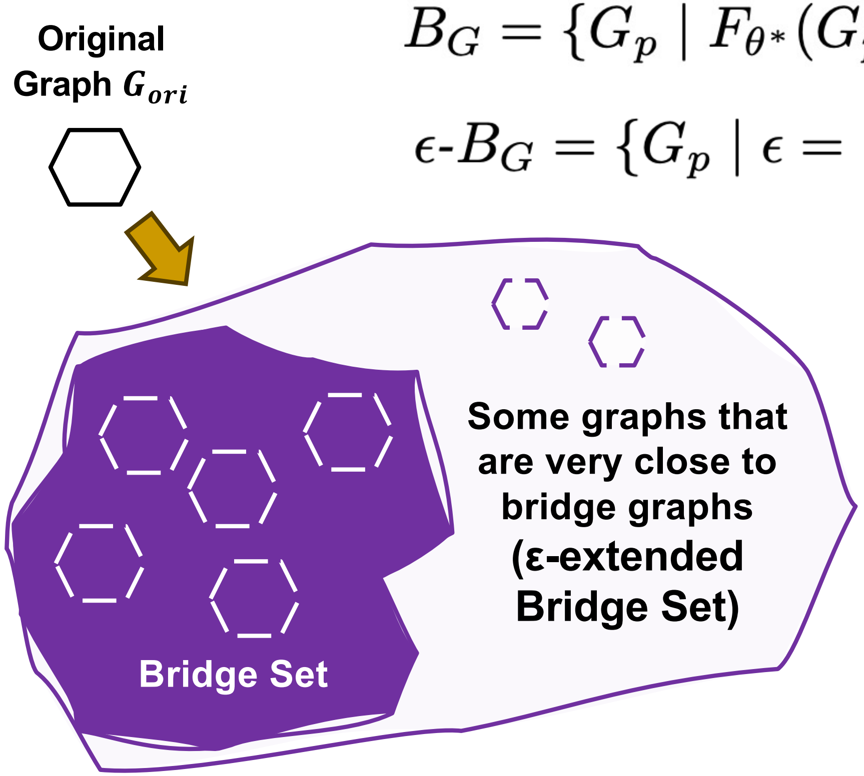
Theorem 1. *Let F_{θ^*} be a GNN model pre-trained on task T_{pre} with frozen parameters (θ^*); let T_{dow} be the downstream task and C is an optimal function to T_{dow}. Given any graph G_{ori}, $C(G_{ori})$ denotes the optimal embedding vector to the downstream task (i.e. can be parsed to yield correct results for G_{ori} in the downstream task), then there always exists a bridge graph G_{bri} such that $F_{\theta^*}(G_{bri}) = C(G_{ori})$.*

Measuring the difficulty of finding bridge graphs.

➢ Bridge Set and ε-extended Bridge Set

$$B_G = \{G_p \mid F_{\theta^*}(G_p) = C(G)\}$$

$$\epsilon\text{-}B_G = \{G_p \mid \epsilon = \|F_{\theta^*}(G_p) - C(G)\| \leq \epsilon^*\}$$



The effectiveness of graph prompting methods can be measured by whether they can uniformly project G_{ori} into the bridge set (without error), or at least map them into the extended bridge set with a small error.

Measuring the difficulty of finding bridge graphs.

- **Upper bound of the error on a single graph**
  - In full-rank models, the prompt can achieve zero approximation error.
  - In non-full-rank models, the error has a clearly defined upper bound, dependent on model expressiveness and data complexity.

Theorem 5. *For a GCN model F_θ, assume at least one layer's parameter matrix is not full rank, for GPF or All-in-One prompt, there exists an upper bound of ϵ such that for any input graph G, there exists an optimal ω where $P_\omega(G) \in \epsilon\text{-}B_G$, with $\epsilon \leq \mu(\theta^*) \cdot \lambda(G)$, where $\mu(\theta^*)$ and $\lambda(G)$ correspond to the model and graph G, respectively.*

Theorem 5 is proved in Appendix A.3.3, where the upper bound of the error ϵ can be further expressed as follows:

$$\mu(\theta^*)\lambda(G) = \sin(\Phi/2)\|C(G)\| \tag{4}$$

a measurement of the **model's expressiveness**

a measurement of the **data complexity**

Does that work for a batch of graphs?

- **For real-world scenarios, we often deal with a dataset.**
  - Already know work for a single graph is controllable.
  - For **n graphs**, having **n prompts** is feasible
  - but what is the minimum number of prompts required?
- **When only one vector GPF, or one node subgraph in Allinone**
  - Maybe not. You can’t expect a single vector to work miracles.
  - The core issue here is: for different graphs, the desired prompts are different, and a single prompt cannot meet everyone's requirements.

Theorem 6. *For a GCN model F_θ, for GPF with a single prompt vector or All-in-One with a single-token graph prompt, given a batch of graphs $\mathcal{G} = \{G_1, \cdots, G_i, \cdots, G_n\}$, the root mean squared error (RMSE) over $\{\epsilon_1, \cdots, \epsilon_n\}$ has a lower bound ϵ^o such that $\boxed{RMSE(\epsilon_1, \cdots, \epsilon_n) \geq \epsilon^o.}$*

Modeling the statistical distribution of errors

- **For the case of non-full-rank order, we have proven that the error has a strict upper bound.**
- **Now, we aim to model the data distribution for actual numerical values.**
  - Under reasonable assumptions, we have proven that the error distribution follows a chi distribution.

Theorem 8. *Given a GCN model F_θ with the last layer parameter matrix having rank $F - r$ (F is the graph embedding dimension, r is the rank lost), an input graph G, for the optimal ω, $P_\omega(G) \in \epsilon\text{-}B_G$. If the Graph Embedding Residuals follow the i.i.d. normal distribution, then ϵ follows a Chi distribution χ_r with r free variables.*

  - Further we can derive the mean and variance

Corollary 1 (Statistical Measures and Confidence Values of ϵ)**.** *The mean of ϵ is $\boxed{c\sqrt{2}\frac{\Gamma((r+1)/2)}{\Gamma(r/2)}}$, the variance is $\boxed{c^2\left(r - 2\frac{[\Gamma((r+1)/2)]^2}{[\Gamma(r/2)]^2}\right)}$, and confidence values can be obtained through $C_\chi^{r,p}$ using numerical methods or table lookup, where c is the scaling factor compared to the standard distribution, and r is the number of dimensions lost compared to a full-rank matrix.*

Extending to nonlinear models

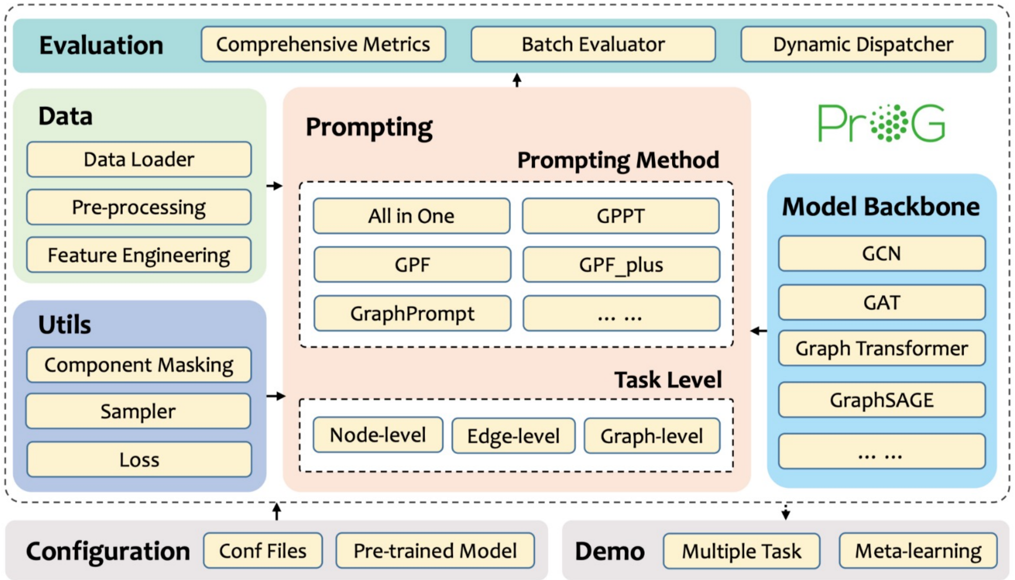
- **Moreover, we have extended the analytical framework from GCN to GAT, a type of nonlinear model.**
  - By "nonlinear," we mean that the aggregation parameters between different nodes depend on the values of the feature matrix
  - The core idea here is that in nonlinear models, aggregation is controlled by the features, and tunable parameters of prompts here is even more flexible.

Theorem 9. *Let F_θ be a GAT model. If any layer of the model has a full row rank parameter matrix, then for the All-in-One prompting framework, for any input graph G, there exists an optimal ω such that $P_\omega(G) \in B_G$. When the parameter matrix is not full rank, there is an upper bound $\mu(\theta) \cdot \lambda(G)$ making $P_\omega(G) \in \epsilon\text{-}B_G$, $\epsilon \leq \mu(\theta) \cdot \lambda(G)$.*

ProG

➢ **We develop a powerful tool to help researchers easily conduct various graph prompting approaches.**

A library built upon PyTorch to easily conduct single or multi-task prompting for pre-trained GNNs



https://github.com/sheldonresearch/ProG

ProG

Google

ProG github

https://github.com/sheldonresearch/ProG

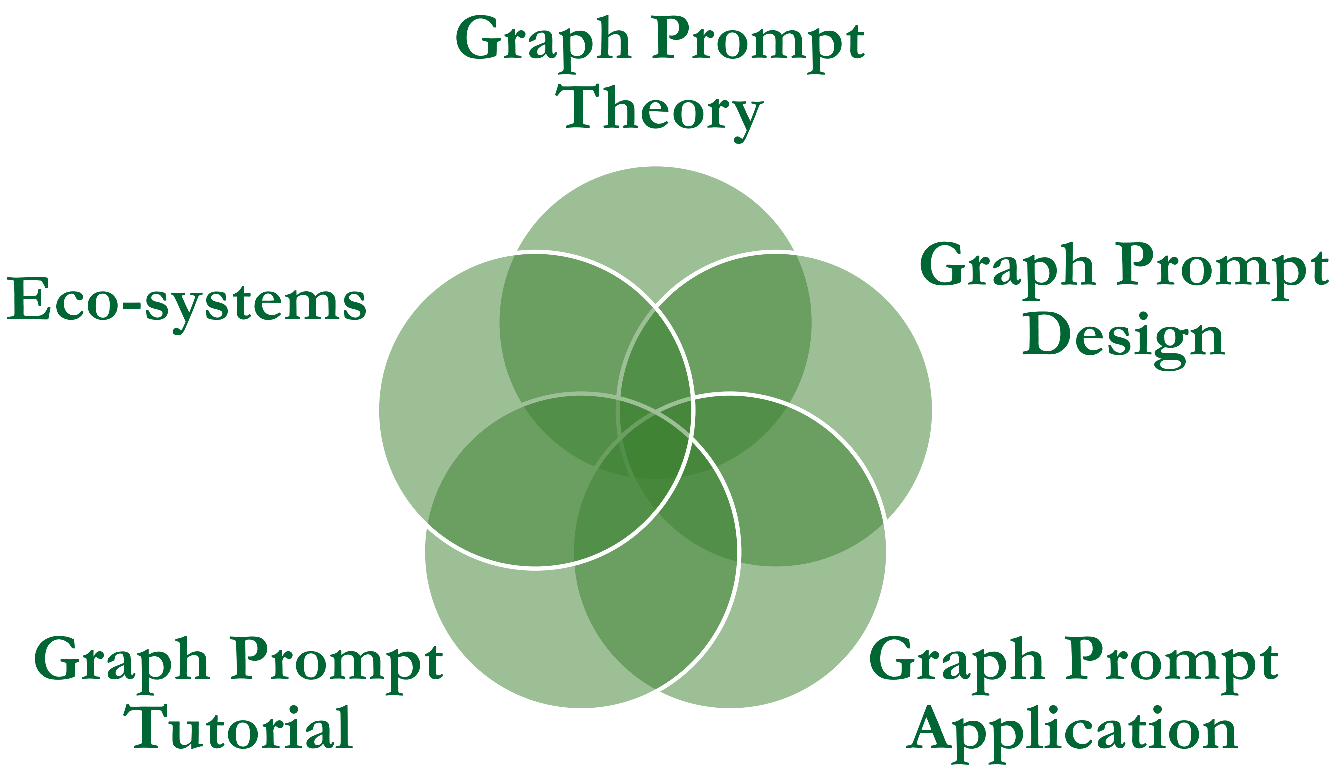
Scan the QR-code to visit our open project, ProG, a library built upon PyTorch to easily conduct single or multi-task prompting for pre-trained GNNs

A Full List of Our Works on Graph Prompts

(* equal contribution † corresponding author)

1. Theory Qunzhong Wang*, Xiangguo Sun*†, Hong Cheng. **Does Graph Prompt Work? A Data Operation Perspective with Theoretical Analysis**. arXiv. Paper
2. Framework Xiangguo Sun, Hong Cheng, Jia Li, Bo Liu, Jihong Guan. **All in One: Multi-Task Prompting for Graph Neural Networks**. SIGKDD 23. Paper
3. Framework Haihong Zhao*, Aochuan Chen*, Xiangguo Sun*†, Hong Cheng, Jia Li†. **All in One and One for Al Simple yet Effective Method towards Cross-domain Graph Pretraining**. SIGKDD 24. Paper
4. Framework Xi Chen, Siwei Zhang, Yun Xiong, Xixi Wu, Jiawei Zhang, Xiangguo Sun, Yao Zhang, Feng Zhao, Yulin Kang. **Prompt Learning on Temporal Interaction Graphs**. arXiv. Paper
5. Benchmark Chenyi Zi*, Haihong Zhao*, Xiangguo Sun†, Yiqing Lin, Hong Cheng, Jia Li. **ProG: A Graph Prom Learning Benchmark**. NeurIPS 2024. Paper
6. Tutorial Xiangguo Sun, Jiawen Zhang, Xixi Wu, Hong Cheng, Yun Xiong, Jia Li. **Graph Prompt Learning: A Comprehensive Survey and Beyond**. arXiv. Paper
7. Tutorial Jia Li, Xiangguo Sun, Yuhan Li, Zhixun Li, Hong Cheng, Jeffrey Xu Yu. **Graph Intelligence with Larg Language Models and Prompt Learning**. SIGKDD 24. Paper
8. Tutorial Yuhan Li*, Zhixun Li*, Peisong Wang*, Jia Li†, Xiangguo Sun, Hong Cheng, Jeffrey Xu Yu. **A Survey Graph Meets Large Language Model: Progress and Future Directions**. IJCAI 2024. Paper
9. Application Hengyu Zhang*, Chunxu Shen*, Xiangguo Sun†, Jie Tan, Yu Rong, Chengzhi Piao, Hong Cheng, Lingling Yi. **Adaptive Coordinators and Prompts on Heterogeneous Graphs for Cross-Domain Recommendations**. arXiv. Paper
10. Application Ziqi Gao, Xiangguo Sun, Zijing Liu, Yu Li, Hong Cheng, Jia Li†. **Protein Multimer Structure Prediction via PPI-guided Prompt Learning**. ICLR 2024. Paper
11. Application Jiahui Jin, Yifan Song, Dong Kan, Haojia Zhu, Xiangguo Sun, Zhicheng Li, Xigang Sun, Jinghui Zhang. **Urban Region Pre-training and Prompting: A Graph-based Approach**. arXiv. Paper
12. Application Yingying Wang, Yun Xiong, Xixi Wu, Xiangguo Sun, Jiawei Zhang. **DDIPrompt: Drug-Drug Interaction Event Prediction based on Graph Prompt Learning**. CIKM 2024. Paper

Our Research Framework on Graph Prompts



Our Research Framework on Graph Prompts

- ## Graph Prompting Theory
  - *Qunzhong Wang, Xiangguo Sun, Hong Cheng.* **Does Graph Prompt Work? A Data Operation Perspective with Theoretical Analysis**. ICML 2025
- ## Graph Prompt Design
  - *Xiangguo Sun, Hong Cheng, Jia Li, Bo Liu, Jihong Guan*. **All in One: Multi-Task Prompting for Graph Neural Networks.** KDD 23.
  - *Haihong Zhao, Aochuan Chen, Xiangguo Sun, Hong Cheng, Jia Li.* **All in One and One for All: A Simple yet Effective Method towards Cross-domain Graph Pretraining**. KDD 24.

Our Research Framework on Graph Prompts

- ## Graph Prompt Tutorial
  - *Chenyi Zi, Bowen Liu, Xiangguo Sun, Hong Cheng, Jia Li.* **Rethinking Graph Prompts: Unraveling the Power of Data Manipulation in Graph Neural Networks**. ICLR 2025 (BlogPosts)
  - *Xiangguo Sun, Jiawen Zhang, Xixi Wu, Hong Cheng, Yun Xiong, Jia Li.* **Graph Prompt Learning: A Comprehensive Survey and Beyond**. https://arxiv.org/abs/2311.16534
  - *Jia Li, Xiangguo Sun, Yuhan Li, Zhixun Li, Hong Cheng, Jeffrey Xu Yu.* **Graph Intelligence with Large Language Models and Prompt Learning.** SIGKDD 24.
  - *Yuhan Li, Zhixun Li, Peisong Wang, Jia Li, Xiangguo Sun, Hong Cheng, Jeffrey Xu Yu*. **A Survey of Graph Meets Large Language Model: Progress and Future Directions**. IJCAI 2024

Our Research Framework on Graph Prompts

- **Graph Prompt Benchmark**
  - *Chenyi Zi, Haihong Zhao, Xiangguo Sun, Yiqing Lin, Hong Cheng, Jia Li*. **ProG: A Graph Prompt Learning Benchmark**. NeurIPS 2024
- **Graph Prompt Application**
  - *Hengyu Zhang, Chunxu Shen, Xiangguo Sun, Jie Tan, Yu Rong, Chengzhi Piao, Hong Cheng, Lingling Yi*. **Adaptive Graph Integration for Cross-Domain Recommendation via Heterogeneous Graph Coordinators.** SIGIR 2025
  - *Ziqi Gao, Xiangguo Sun, Zijing Liu, Yu Li, Hong Cheng, Jia Li.* **Protein Multimer Structure Prediction via PPI-guided Prompt Learning.** ICLR 2024

Q&A

Thanks!

Graph Prompt Learning and Pre-training