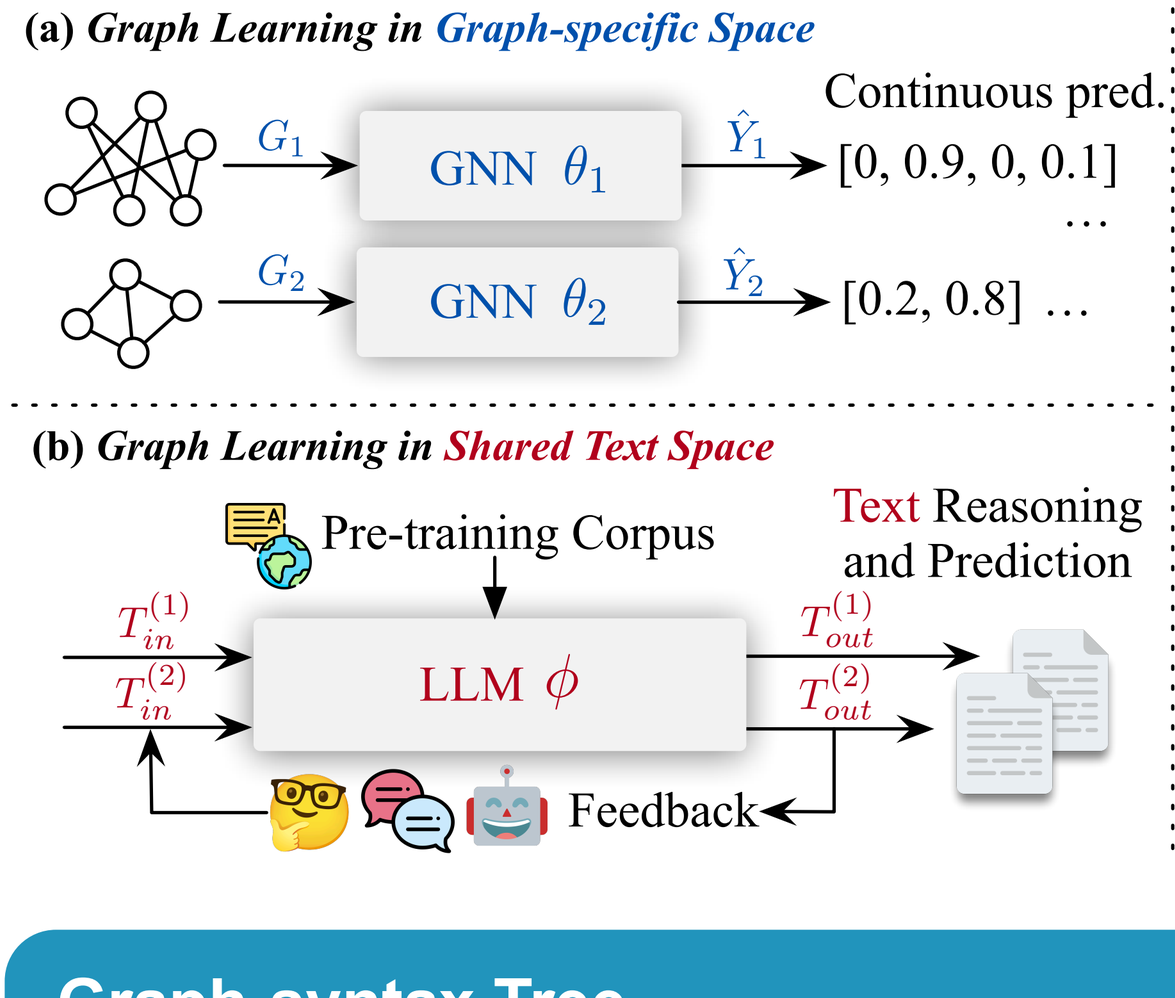
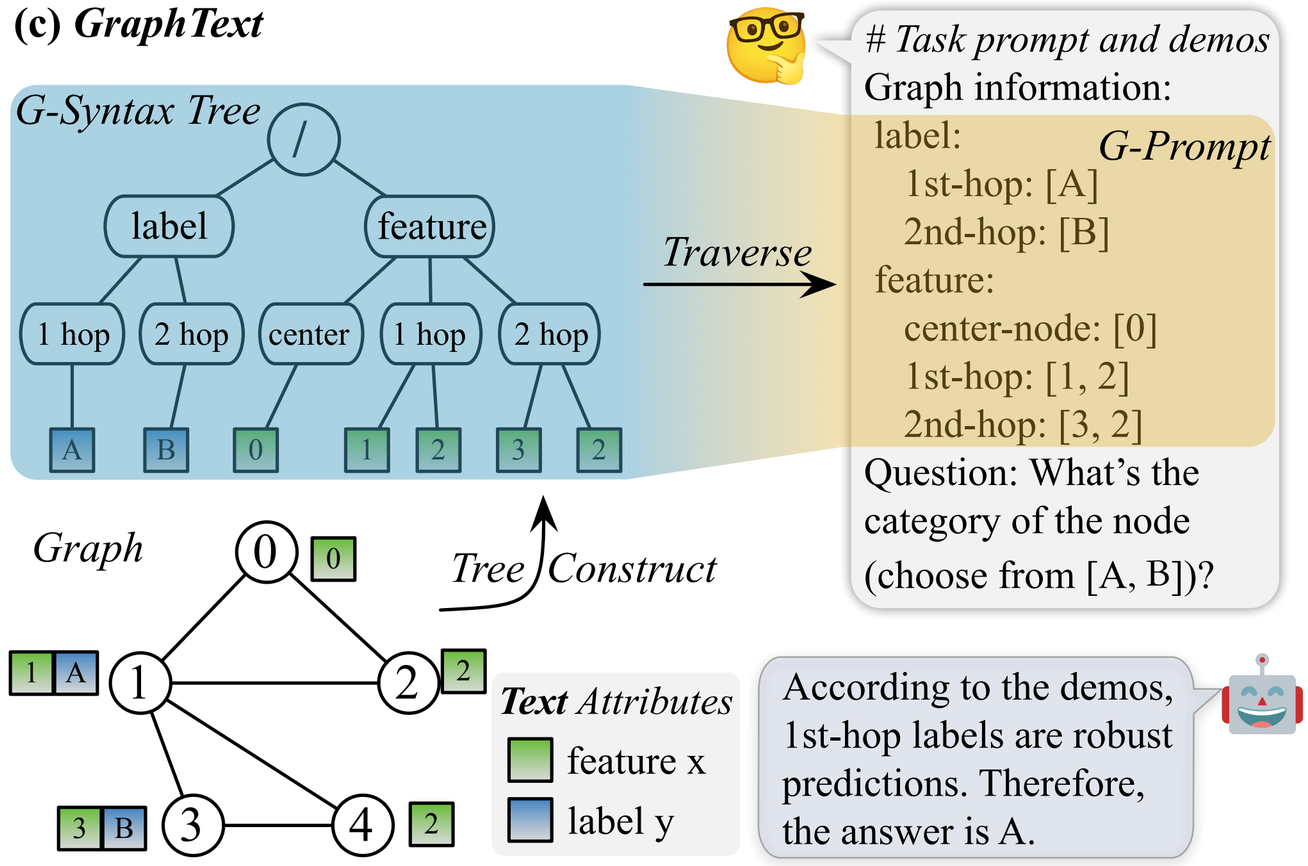
GraphText: Graph Reasoning in Text Space

Jianan Zhao, Le Zhuo, Yikang Shen, Meng Qu, Kai Liu, Michael Bronstein, Zhaocheng Zhu, Jian Tang

Mila | IBM | MIT | Université de Montréal | University of Oxford

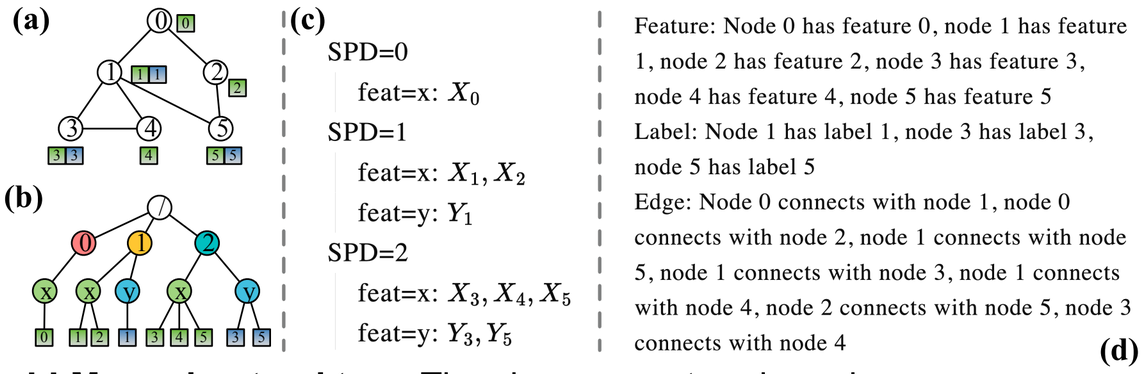
GraphText uses **tree** as an intermediary to bridge **structure** and **sequence**.





Graph-syntax Tree

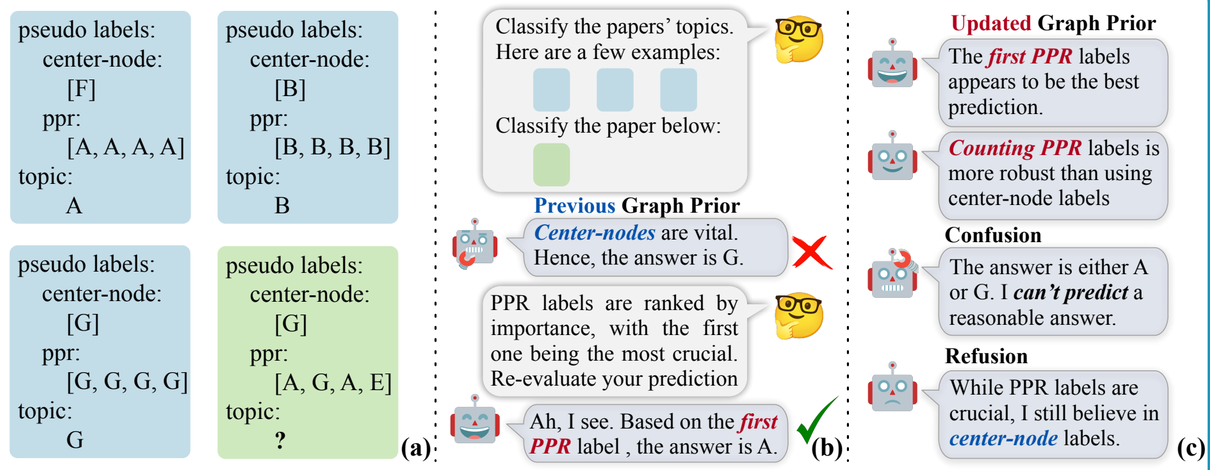
Structured prompt for graph reasoning.



- **LLMs understand tree**: They have seen tree-based corpus, e.g. code data, webpages (DOM tree), and XML files.
- Built-in **structured** prompt: Hierarchical tree structure (c) v.s. structure flattened in description (d).
- Easy incorporation of **graph inductive bias**: Design the structure and attributes and of the graph-syntax tree.

Interactive Graph Reasoning

Interactive graph reasoning in natural language.

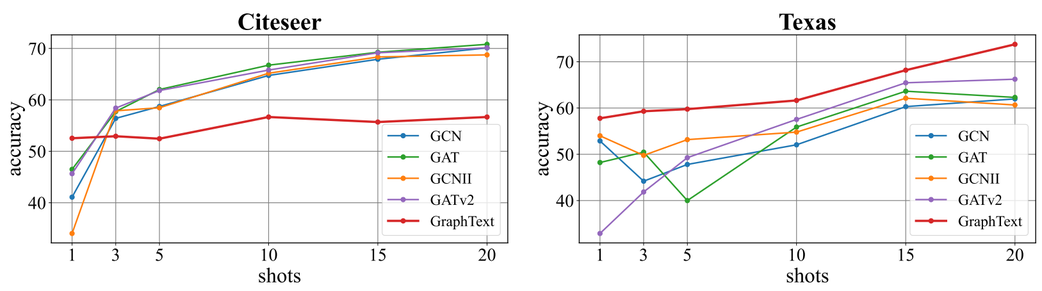


- **Interpretability**: generate and *clarify* predictions in natural language
- **Interactiveness**: LLMs can adapt their prior inductive bias based on the human feedback.

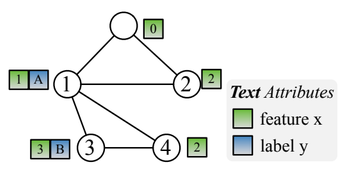
Experiments

Setting	Model	Cora	Citeseer	Texas	Wisconsin	Cornell
Supervised Learning	GCN	81.4	69.8	59.5	49.0	37.8
	GAT	80.8	69.4	54.1	49.0	**45.9**
	GCNII	81.2	69.8	56.8	51.0	40.5
	GATv2	**82.3**	**69.9**	**62.2**	**52.9**	43.2
In-Context Learning	NeighborText	26.3	13.7	5.4	9.8	21.6
	GML	38.5	28.4	10.8	23.5	21.6
	GraphML	49.9	28.9	16.2	33.3	29.7
	GraphText of+or	33.4	36.9	5.4	29.4	24.3
	GraphText of+sr	52.1	50.4	73.0	60.8	46.0
	GraphText sf+or	64.5	51.0	73.0	35.3	48.7
	GraphText sf+sr	**68.3**	**58.6**	**75.7**	**67.6**	**57.9**
Comparisons	Δ SFT-GCN	-13.1%	-11.2%	+16.2%	+18.6%	+20.1%
	Δ Best ICL	+18.4%	+29.7%	+59.5%	+34.3%	+28.2%

SOTA node classification performance, comparable to supervised baselines



GraphText surpasses supervised baselines when labels are few



(a) An example graph

(b) GraphText

```
label:
  1st-hop: [A]
  2nd-hop: [B]
feature:
  center-node: [0]
  1st-hop: [1, 2]
  2nd-hop: [3, 2]
```

(c) Rev. hierarchy

```
center-node:
  feature: [0]
1st-hop:
  label: [A]
  feature: [1, 2]
2nd-hop:
  label: [B]
  feature: [3, 2]
```

(d) w/o struct-sem

```
attributes:
  [A]
  [B]

  [0]
  [1,2]
  [3,2]
```

(e) Sequence

```
attributes:
  [A, B, 0, 1,
  2, 3, 2]
```

(f) Set

```
attributes:
  [2, 0, 3, B
  1, 2, A]
```

Model	Cora		Citeseer	
	Acc. %	Δ	Acc. %	Δ
GraphText	**68.3**	-	**58.6**	-
rev. hierarchy	68.3	-0 %	57.6	-1.7 %
w/o struct-sem	67.8	-0.5 %	56.3	-3.9 %
sequence	67.0	-1.3 %	53.0	-9.6 %
set	65.9	-2.4 %	56.4	-3.8 %

Ablations of graph-syntax trees

Model	Interaction	Accuracy	Reasoning		
			PPR	Center-node	Conf./Ref.
GPT-4	Before	73.3	73.3	26.7	0
	After	**100 (+26.7)**	100	0	0
ChatGPT	Before	26.7	26.7	53.3	20.0
	After	**63.6 (+36.9)**	72.7	18.2	9.1

Quantitative interactive graph reasoning performance on Cora node #2188

Paper

Code