Can Graph Descriptive Order Affect Solving Graph Problems with LLMs?

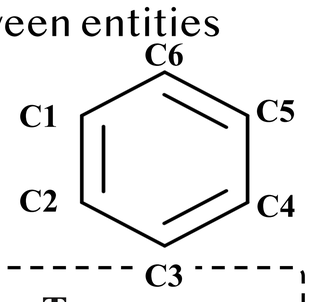
Yuyao Ge, Shenghua Liu, Baolong Bi, Yiwei Wang,
Lingrui Mei, Wenjie Feng, Lizhe Chen, Xueqi Cheng
Institute of Computing Technology, Chinese Academy of Sciences
University of California, Merced
Email: yuyao.ge.work@gmail.com

Paper

Motivation

- In real-world scenarios, the order of textual descriptions significantly influences how humans interpret relationships between entities.
- To better understand, let's abstract relationships between entities into graph representations:

For a standard benzene ring, which of the following two descriptions is easier for humans to understand?



Description One

C1-C2, C2-C3, C3-C4, C4-C5, C5-C6, C6-C1

Humans can immediately recognize that **this is a benzene ring!**

Description Two

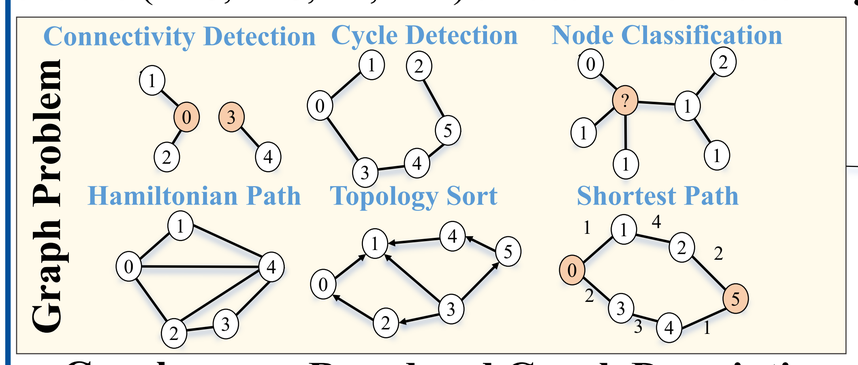
C3-C4, C1-C2, C5-C6, C2-C3, C4-C5, C6-C1

Humans need to reorganize the information to understand what it is.

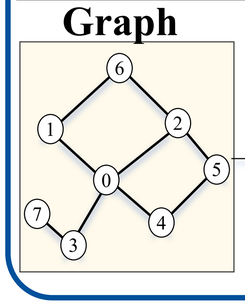
- The order of graph description has a significant impact on human understanding of its structure. Does this effect also exist in LLMs?

Graph Description Generation Framework

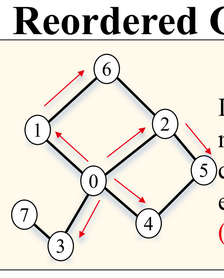
We designed six types of graph tasks to assess how four graph traversal orders (DFS, BFS, PR, PPR) affect LLMs' reasoning performance.



Graph



Reordered Graph Description



Reordered in BFS order
In an undirected graph, (i, j) means that node i and node j are connected with an edge, and the edges are: (0, 1), (0, 2), (0, 3), (0, 4), (1, 6), (2, 5), …

Prompt

<Description with BFS>:
In an undirected graph, (i, j) means that node i and node j are connected with an edge, and the edges are: (0, 1), (0, 2), (0, 3), (0, 4), (1, 6), (2, 5), …

<Cycle Detect>:
Q: Is there a cycle in this graph?

<Zero-short CoT>:
Let's think step by step.

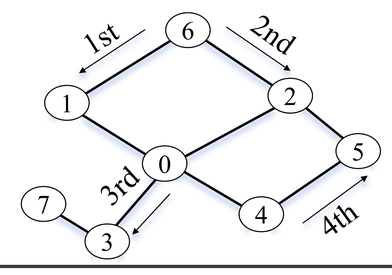
LLM: $\mathcal{M}$

Orders

Random The edges $\mathcal{E}$ of the graph are shuffled randomly.
Breadth-First Search (BFS) Order Sequentialize the edegs with BFS.
Depth-First Search (DFS) Order Sequentialize the edges with DFS.
PageRank (PR) Order Nodes are sorted in descending order by their PageRank scores $PR(v)$, where $PR(v) = \alpha \sum_{u \in \mathcal{N}^{-1}(v)} \frac{PR(u)}{|\mathcal{N}(u)|} + (1-\alpha)$. For each node, starting with the highest-ranked, edges to its neighbors $u \in \mathcal{N}(v)$ are added to the edge list $\mathcal{L}_{\mathcal{PR}}$.
Personalized PageRank (PPR) Order For PPR, the ranking is computed as $PR_S(v) = \alpha \sum_{u \in \mathcal{N}^{-1}(v)} \frac{PR_S(u)}{|\mathcal{N}(u)|} + (1-\alpha) \cdot e_v$.

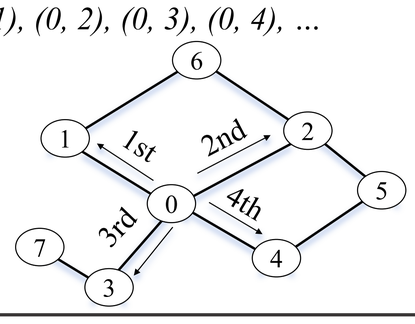
Random Order

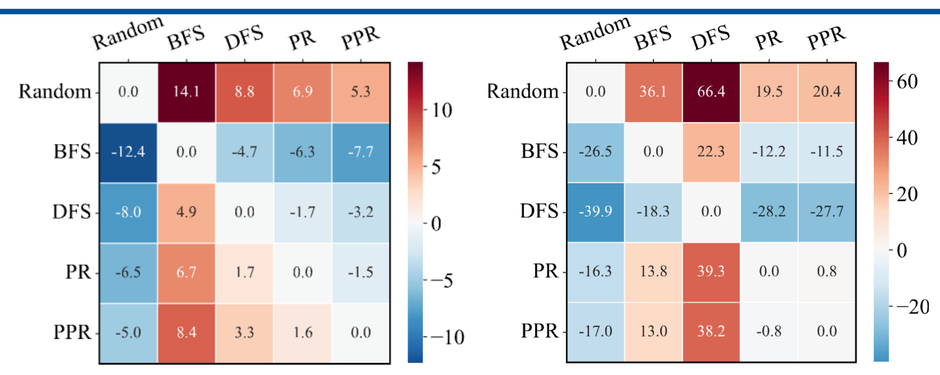
In an undirected graph, (i, j) means that node i and node j are connected with an edge, and the edges are:
(6, 1), (6, 2), (0, 3), (4, 5), …



BFS Order

In an undirected graph , (i, j) means that node i and node j are connected with an edge, and the edges are:
(0, 1), (0, 2), (0, 3), (0, 4), …





(a) Connectivity (b) Hamilton Path

- Task characteristics determine optimal ordering strategy
  - Connectivity task need local view
  - Hamilton Path need global view

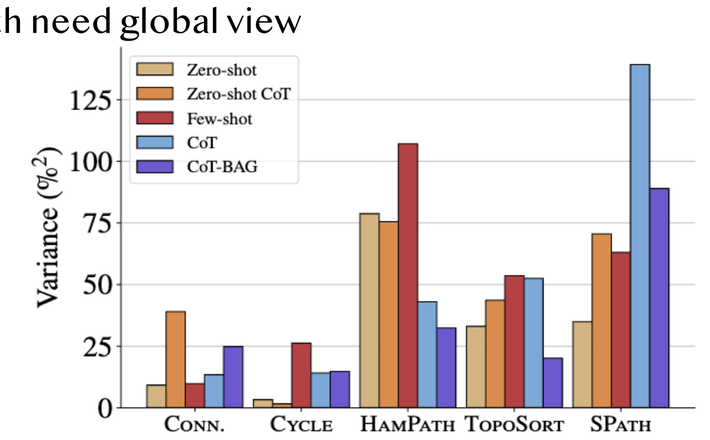
Results

Task	Order	Zero-shot	Zero-shot CoT	Few-shot	CoT	CoT-BAG	Avg.
Conn.	Random	73.93(-)	70.71(-)	81.07(-)	83.93(-)	82.14(-)	78.36(-)
	BFS	82.14(↑11.11)	87.50(↑23.74)	89.29(↑10.14)	92.50(↑10.21)	95.71(↑16.52)	89.43(↑14.13)
	DFS	79.29(↑7.25)	82.14(↑16.16)	87.14(↑7.49)	88.21(↑5.10)	89.29(↑8.70)	85.21(↑8.75)
	PR	77.86(↑5.32)	83.57(↑18.19)	85.71(↑5.72)	84.29(↑0.43)	87.50(↑6.53)	83.79(↑6.93)
	PPR	76.79(↑3.87)	81.07(↑14.65)	83.93(↑3.53)	84.64(↑0.85)	86.07(↑4.78)	82.50(↑5.29)
Cycle	Random	51.79(-)	53.57(-)	65.36(-)	75.71(-)	76.07(-)	64.50(-)
	BFS	55.71(↑7.57)	56.07(↑4.67)	79.29(↑21.31)	86.07(↑13.68)	86.43(↑13.62)	72.71(↑12.73)
	DFS	52.14(↑0.68)	53.93(↑0.67)	73.21(↑12.01)	79.29(↑4.73)	81.07(↑6.57)	67.93(↑5.31)
	PR	55.36(↑6.89)	56.43(↑5.33)	70.36(↑7.65)	80.36(↑6.14)	83.21(↑9.39)	69.14(↑7.20)
	PPR	54.29(↑4.83)	55.00(↑2.67)	70.00(↑7.10)	79.29(↑4.73)	80.00(↑5.17)	67.72(↑4.99)
HamPath	Random	10.71(-)	15.36(-)	40.00(-)	46.07(-)	45.36(-)	31.50(-)
	BFS	20.00(↑86.74)	20.71(↑34.83)	57.86(↑44.65)	58.57(↑27.13)	57.14(↑25.97)	42.86(↑36.05)
	DFS	33.93(↑216.81)	37.50(↑144.14)	67.50(↑68.75)	63.93(↑38.77)	59.29(↑30.71)	52.43(↑66.44)
	PR	15.00(↑40.06)	19.29(↑25.59)	48.93(↑22.32)	55.00(↑19.38)	50.00(↑10.23)	37.64(↑19.50)
	PPR	16.43(↑53.41)	18.93(↑23.24)	50.00(↑25.00)	53.93(↑17.06)	50.36(↑11.02)	37.93(↑20.41)
TopoSort	Random	28.93(-)	31.07(-)	58.21(-)	56.07(-)	60.36(-)	46.93(-)
	BFS	43.21(↑49.36)	40.36(↑29.90)	67.14(↑15.34)	61.43(↑9.56)	65.00(↑7.69)	55.43(↑18.11)
	DFS	42.14(↑45.66)	48.93(↑57.48)	77.86(↑33.76)	74.29(↑32.50)	72.86(↑20.71)	63.21(↑34.71)
	PR	35.36(↑22.23)	35.71(↑14.93)	71.07(↑22.09)	58.21(↑3.82)	65.36(↑8.28)	53.14(↑13.24)
	PPR	37.14(↑28.38)	39.64(↑27.58)	72.50(↑24.55)	58.93(↑5.10)	66.43(↑10.06)	54.93(↑17.05)
SPath	Random	20.00(-)	25.00(-)	26.07(-)	38.93(-)	40.71(-)	30.14(-)
	BFS	35.36(↑76.80)	42.50(↑70.00)	45.36(↑73.99)	67.50(↑73.39)	65.71(↑61.41)	51.29(↑70.15)
	DFS	32.14(↑60.70)	34.29(↑37.16)	45.00(↑72.61)	58.57(↑50.45)	57.14(↑40.36)	45.43(↑50.71)
	PR	30.36(↑51.80)	43.93(↑75.72)	38.93(↑49.33)	43.93(↑12.84)	48.93(↑20.19)	41.21(↑36.74)
	PPR	32.50(↑62.50)	44.64(↑78.56)	42.14(↑61.64)	45.36(↑16.52)	49.64(↑21.94)	42.86(↑42.18)

- Ordered descriptions consistently outperform random baseline.
  - Complex tasks benefit more from ordering.
  - Different prompting methods maintain ordering benefits

Sampling	Order	CORA Acc.	Citeseer Acc.	Pubmed Acc.
Ego	Random	70.00(–)	67.33(–)	72.00(–)
	BFS	72.00(+2.86)	68.67(+1.99)	74.00(+2.78)
	DFS	71.33(+1.90)	68.66(+1.98)	77.33(+7.40)
	PR	75.33(+7.61)	71.33(+5.94)	82.67(+14.82)
	PPR	73.33(+4.76)	69.33(+2.97)	77.33(+7.40)
Forest Fire	Random	79.33(–)	68.67(–)	69.99(–)
	BFS	82.67(+4.21)	71.33(+3.87)	74.00(+5.73)
	DFS	81.33(+2.52)	70.00(+1.94)	76.00(+8.59)
	PR	83.33(+5.04)	71.33(+3.87)	76.00(+8.59)
	PPR	82.00(+3.36)	70.67(+2.91)	74.67(+6.69)

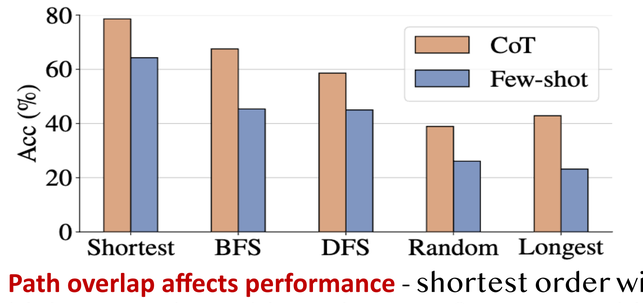
Probability-based orders outperform traversal-based orders.



Simpler tasks show greater robustness, while complex tasks are more sensitive to description order.

Path Overlapping or Better Understanding?

Two New Orders: Shortest Path Order (edges ordered by the actual shortest path) and Longest Path Order (edges ordered by the longest path from source to target).



- **Path overlap affects performance** - shortest order with highest overlap achieves best performance, while longest with lowest overlap performs worst.
- **Ordering enhances understanding beyond overlap** - DFS has higher overlap than BFS, yet BFS outperforms DFS

Conclusion

- Graph Descriptive Order greatly affects LLMs' Ability of Solving Graph Problems
- Simple Tasks are more robustness to descriptive order, while complex tasks are more sensitive.
- There is no specific order that has the best performance across all tasks, as task characteristics determine the optimal ordering strategy.
- Structured description ordering can enhance LLMs' understanding of graphs.