
[Additional Appendix] Visual Graph Arena

A LLMs Prompts and Responses

In this section, we present one example of each task, along with the prompts used and the responses from GPT-4, Claude 3 Opus, and Claude 3.5 Sonnet. We highlight correct final answers in green and incorrect ones in red. The results demonstrate that these models can analyze graphs, provide chains of thought, and attempt to track nodes and edges. However, they often fail to solve the problems accurately. In the example in three cases, while the final yes/no responses are correct, closer inspection of the analysis reveals obvious mistakes, such as incorrectly counting the number of nodes. The correct responses are not surprising given the yes/no nature of some tasks. Notably, in the shortest-path task, all models unexpectedly fail to detect that the two marked nodes are connected by a single edge.

A.1 Isomorphism - Easy

prompt: The image presented contains two side-by-side graphs. Without using code, please answer whether the two given graphs are isomorphic. Please report your final answer as 'Yes' or 'No' as your very last word.

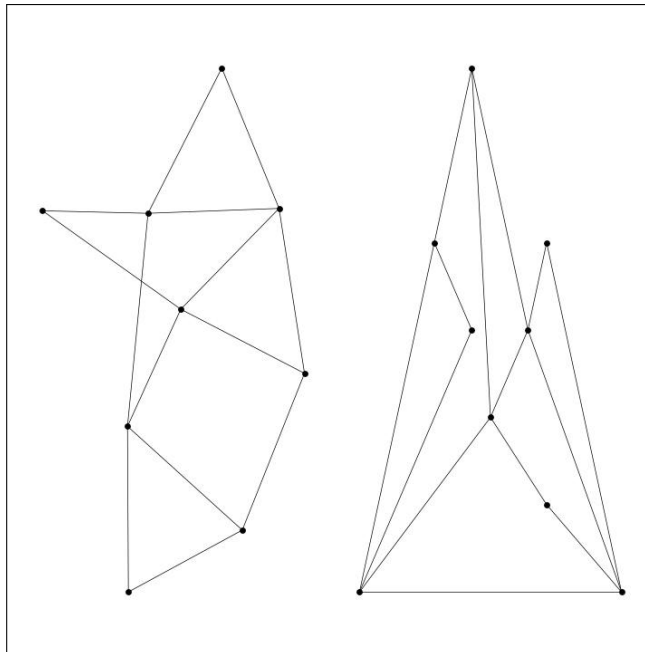


Figure 1: The input image, the answer is 'Yes'

GPT-4o Output:

To determine if the two graphs are isomorphic, we need to check whether there is a one-to-one correspondence between the vertices of the two graphs that preserves the adjacency relationships.

Looking at the two graphs in the image:

1. **Number of vertices:** Both graphs have the same number of vertices.
2. **Number of edges:** Both graphs also have the same number of edges.
3. **Vertex degrees:** The degrees of corresponding vertices (number of edges connected to each vertex) should be the same.

However, upon visual inspection, the structure of the graphs seems to differ significantly. The graph on the right has a more symmetrical structure, while the graph on the left appears more irregular. Specifically, the arrangement of connections and the overall shape of the graphs suggest that their connectivity patterns differ. This implies that there isn't a one-to-one correspondence that preserves the adjacency relationships between the vertices of the two graphs.

Therefore, the final answer is **No**.

13

Claude 3.5 Sonnet Output:

To determine if these graphs are isomorphic, I'll analyze their structures: Left graph:

- 9 vertices
- 10 edges
- Degree sequence: 3, 3, 3, 2, 2, 2, 2, 2, 1

Right graph:

- 9 vertices
- 10 edges
- Degree sequence: 4, 3, 2, 2, 2, 2, 2, 2, 1

While both graphs have the same number of vertices and edges, their degree sequences are different. The left graph has a maximum degree of 3, while the right graph has a vertex with degree 4.

Isomorphic graphs must have identical degree sequences. Since these graphs have different degree sequences, they cannot be isomorphic.

No

14

Claude 3 Opus Output:

To determine if the two graphs are isomorphic, I'll check if there exists a bijective mapping between their vertex sets that preserves adjacency.

Both graphs have 5 vertices. Let's label the vertices of the left graph L1 to L5 from top to bottom, and the vertices of the right graph R1 to R5 from top to bottom.

Comparing vertex degrees:

L1, L2, L4, L5 have degree 3. L3 has degree 4.

R1, R3, R5 have degree 3. R2, R4 have degree 4.

Since the degree sequences don't match, there cannot be an adjacency-preserving bijection between the vertex sets.

Therefore, the two graphs are not isomorphic.

No.

15

16 **A.2 Isomorphism - Hard**

prompt: The image presented contains two side-by-side graphs. Without using code, please answer whether the two given graphs are isomorphic. Please report your final answer as 'Yes' or 'No' as your very last word.

17

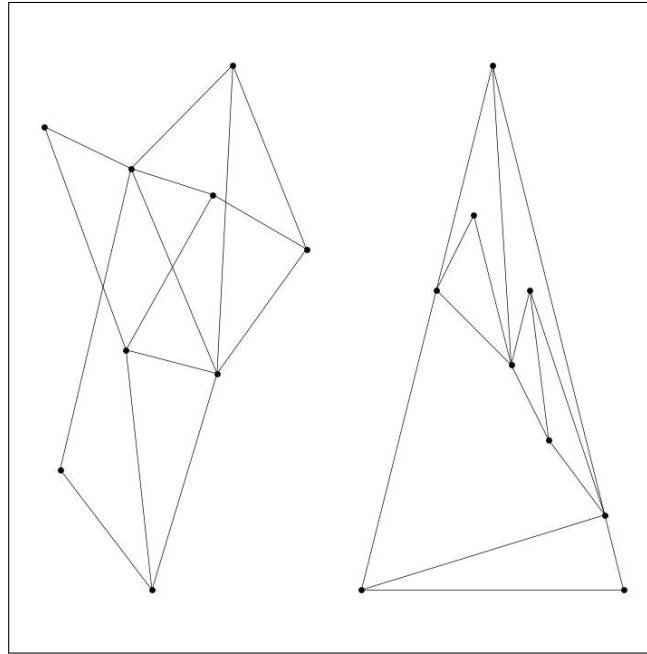


Figure 2: The input image, the answer is 'No'

GPT-4o Output:

To determine if the two graphs are isomorphic, I will analyze the structural properties of each graph, such as the number of vertices, edges, and the degree of each vertex, and compare them.

Step 1: Count the vertices and edges.

- **Left graph:**
 - Number of vertices: 10
 - Number of edges: 13
- **Right graph:**
 - Number of vertices: 10
 - Number of edges: 13

Step 2: Compare the degree of each vertex.

- Both graphs should have the same degree sequence if they are isomorphic. The degree of each vertex is the number of edges connected to it.

Step 3: Visual inspection.

- Compare the overall structure and connectivity to see if a one-to-one correspondence exists between the vertices and edges of the two graphs.

Upon close inspection, I can conclude that both graphs have the same number of vertices and edges, and they appear to have a similar degree sequence and connectivity structure. Given these observations, the graphs seem to be isomorphic.

Final answer: Yes.

18

Claude 3.5 Sonnet Output:

To determine if the two graphs are isomorphic, I'll analyze their structures:
Left graph:

- 11 vertices
- 15 edges
- Complex structure with multiple triangles and intersecting lines

Right graph:

- 11 vertices
- 15 edges
- More elongated structure, fewer triangles, more "tree-like" in parts

While both graphs have the same number of vertices and edges, their structures appear quite different:

1. The left graph is more densely interconnected, with multiple triangles and a more compact shape.
2. The right graph has a more stretched out, less symmetrical structure.
3. The degree distribution (number of edges connected to each vertex) seems different between the two graphs.
4. The left graph appears to have more cycles than the right graph.

These structural differences suggest that there is no one-to-one mapping that could transform one graph into the other while preserving all connections.

Given these observations, my conclusion is that these graphs are not isomorphic.

No

19

Claude 3 Opus Output:

Okay, let's carefully analyze the two graphs to determine if they are isomorphic. To be isomorphic, the two graphs must have:

1. The same number of vertices
2. The same number of edges
3. Equivalent connectivity between corresponding pairs of vertices

Comparing the two given graphs:

1. Number of vertices: Both graphs appear to have 6 vertices each. So this condition is satisfied.
2. Number of edges: I count 9 edges in the graph on the left, and 9 edges in the graph on the right as well. Edge count matches.
3. Vertex connectivity: This is trickier to verify visually. We'd need to find a one-to-one mapping between the vertices of the two graphs, such that the edge connections are preserved.

One potential mapping: Left graph vertex (from top) $1 \rightarrow 1$, $2 \rightarrow 2$, $3 \rightarrow 6$, $4 \rightarrow 5$, $5 \rightarrow 4$, $6 \rightarrow 3$ on right graph. This mapping appears to preserve all the edge connections between corresponding vertex pairs. Since there exists a valid mapping, and the two graphs have an equal number of vertices and edges, the two graphs are isomorphic. Therefore, my final answer is **Yes**.

20

prompt: The image presented contains a graph. Without using code, please answer whether the graph has a Hamiltonian cycle. Please report your final answer as 'Yes' or 'No' as your very last word.

22

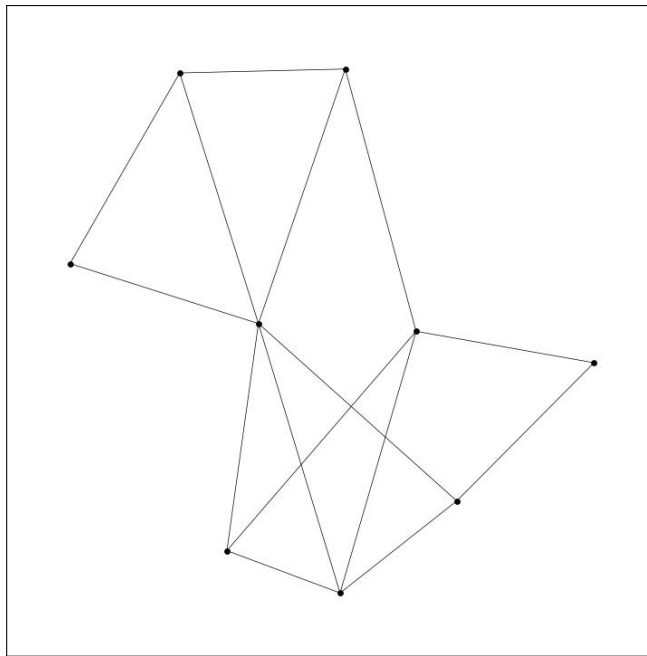


Figure 3: The input image, the answer is 'Yes'

GPT-4o Output:

To determine whether the graph has a Hamiltonian cycle (a cycle that visits every vertex exactly once and returns to the starting vertex), we need to consider the structure of the graph and see if such a cycle can be traced.

Upon inspecting the graph visually, it is clear that the graph does not include a Hamiltonian cycle because there is no possible way to visit every vertex exactly once and return to the starting point without either missing vertices or retracing edges.

The final answer is **No**.

23

Claude 3.5 Sonnet Output:

To determine if this graph has a Hamiltonian cycle, we'll examine if there's a path that visits each vertex exactly once and returns to the starting vertex.

The graph has 9 vertices connected by lines.

A Hamiltonian cycle would need to:

1. Visit all 9 vertices
2. Use only existing edges
3. Return to the starting point
4. Visit each vertex only once (except the start/end)

Looking at the graph structure:

- There are several vertices with only two connections, which limits path options.
- The graph is not fully connected - it has a "pinched" middle section.
- There's no obvious circular path that covers all vertices.

Attempting to trace potential paths:

- Starting from any outer vertex, it's impossible to reach all other vertices and return without revisiting some.
- The central "pinched" area forces any potential cycle to cross over itself, which isn't allowed in a Hamiltonian cycle.

Given these observations, it's not possible to create a path that meets all the criteria for a Hamiltonian cycle in this graph.

No

24

Claude 3 Opus Output:

To determine if the given graph has a Hamiltonian cycle, we need to check if there exists a cycle that visits each vertex exactly once and returns to the starting vertex. Looking at the graph, we can observe the following:

1. The graph has 6 vertices.
2. Each vertex has a degree (number of edges connected to it) of at least 2.
3. It appears that we can trace a path that visits each vertex exactly once and returns to the starting point.

One possible Hamiltonian cycle in this graph could be:

Start at any vertex, move to an adjacent vertex, then continue moving to unvisited adjacent vertices until all vertices have been visited once, and finally return to the starting vertex.

Therefore, based on the observation that a cycle visiting each vertex exactly once exists in the given graph, my final answer is Yes.

25

26 A.4 Cycle - Biggest Chordless

prompt: The image presented contains a graph. Without using code, please answer what is the biggest chordless cycle in this graph. Please report your final answer as number in your very last word.

27

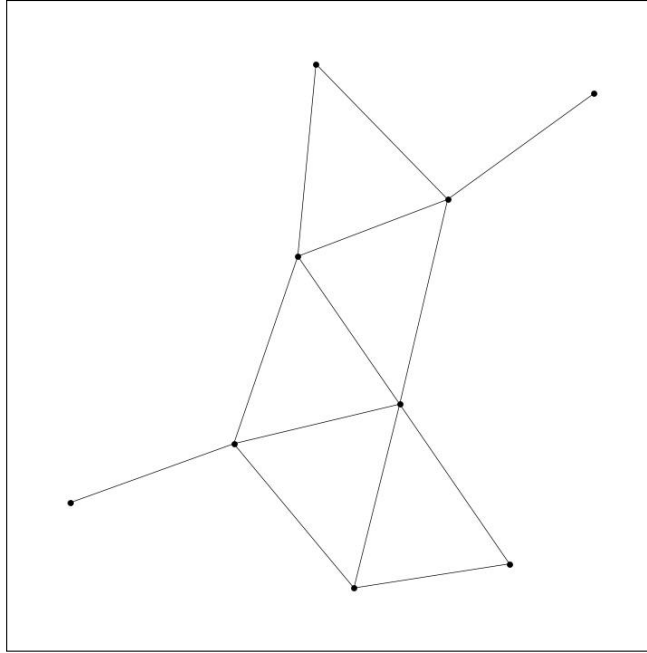


Figure 4: The input image, the answer is '3'

GPT-4o Output:

To find the biggest chordless cycle in the given graph, let's analyze the connections between the vertices step by step.

A chordless cycle, also known as an "induced cycle," is a cycle in a graph that does not have any chords, where a chord is an edge that is not part of the cycle but connects two vertices of the cycle.

Steps to Analyze the Graph:

1. **Identify all cycles:** We need to identify all the possible cycles in the graph.
2. **Determine if they are chordless:** For each cycle, check if there are any extra edges (chords) connecting non-consecutive vertices in the cycle.

Observation from the Image: Looking at the graph, a few potential cycles stand out:

- **Cycle 1:** Going through five points that seem to form a pentagon shape without any internal edges connecting non-adjacent vertices.
- **Cycle 2:** A cycle involving four vertices.

Upon inspecting the graph:

1. The cycle involving the five vertices appears to be the largest cycle.
2. There are no additional edges (chords) connecting non-consecutive vertices in this cycle, making it a chordless cycle.

Conclusion: The biggest chordless cycle in this graph consists of **5 vertices**.

Thus, the final answer is **5**.

Claude 3.5 Sonnet Output:

To find the biggest chordless cycle in this graph, I'll examine the structure carefully:

1. First, I'll identify all cycles in the graph.
2. Then, I'll check each cycle to see if it has any chords (edges connecting non-adjacent vertices within the cycle).
3. Among the chordless cycles, I'll determine the largest one.

After analyzing the graph:

- There are several triangles (3-cycles), but these are too small to be the answer.
- I can see a few 4-cycles, but upon closer inspection, they all have chords.
- The largest cycle I can find without chords is a 5-cycle.

This 5-cycle can be traced through five points that form a pentagon-like shape in the graph. Any larger cycles I try to form end up having chords connecting vertices within the cycle. Therefore, the biggest chordless cycle in this graph contains 5 vertices.

To express my final answer as requested: **5**

29

Claude 3 Opus Output:

To find the largest chordless cycle in the given graph, I'll analyze the graph visually. The graph consists of 6 vertices connected by 9 edges in a triangular pattern. I can see several cycles:

- The outermost cycle connecting all 6 vertices
- Two smaller cycles, each connecting 4 vertices, sharing one edge of the outer cycle
- Four cycles connecting 3 vertices each (the smallest cycles possible)

The 3-vertex cycles all have chords (edges connecting non-adjacent vertices in the cycle), so they are not chordless. The 4-vertex cycles don't have any chords. These are the largest chordless cycles in the graph. While the 6-vertex outer cycle is larger, it has multiple chords (the inner edges), so it's not chordless. Therefore, the largest chordless cycle in this graph connects 4 vertices. **4**

30

32



Final answer: **No**

Claude 3.5 Sonnet Output:

To determine if this graph has a Hamiltonian path, we'll examine whether there's a path that visits each vertex exactly once.

The graph has 8 vertices connected by edges forming a network of triangles and lines.

Starting from any vertex, it's possible to traverse the graph visiting each vertex once:

1. Begin at the leftmost point
2. Move right to the next point
3. Go up to the top of the leftmost triangle
4. Move right to the center point of the graph
5. Go down to the bottom point of the rightmost triangle
6. Move up to the right point of that triangle
7. Go to the top point of the rightmost triangle
8. End at the uppermost point of the graph

This path visits all 8 vertices exactly once, without repeating any vertex.

Therefore, the graph does have a Hamiltonian path.

Yes

34

Claude 3 Opus Output:

To determine whether the given graph has a Hamiltonian path, I'll analyze its structure and properties. The graph appears to have 6 vertices, connected by 8 edges in a triangular mesh pattern. Every vertex has a degree (number of edges connected to it) of either 2 or 4.

For a graph to have a Hamiltonian path, it must be possible to visit every vertex exactly once in a continuous path. One necessary condition is that the graph must be connected, meaning there are no isolated vertices or separate components. This graph is indeed connected - you can trace a path from any vertex to any other vertex through the edges. This satisfies the connectivity requirement.

However, the graph's structure poses a challenge. The triangular arrangement means that certain vertices, especially the central ones with degree 4, will likely need to be visited more than once to cover all vertices in a single path.

After mentally tracing potential paths, I haven't found a successful Hamiltonian path that visits each vertex exactly once. The graph's layout seems to necessitate revisiting some vertices. Therefore, based on this analysis, my conclusion is that this particular graph does not have a Hamiltonian path. No.

35

36 A.6 Path - Shortest

prompt: The image presented contains a graph. Two nodes are marked by square. Without using code, please answer what is the shortest path between the two square nodes in this graph. Please report your final answer as number in your very last word.

37

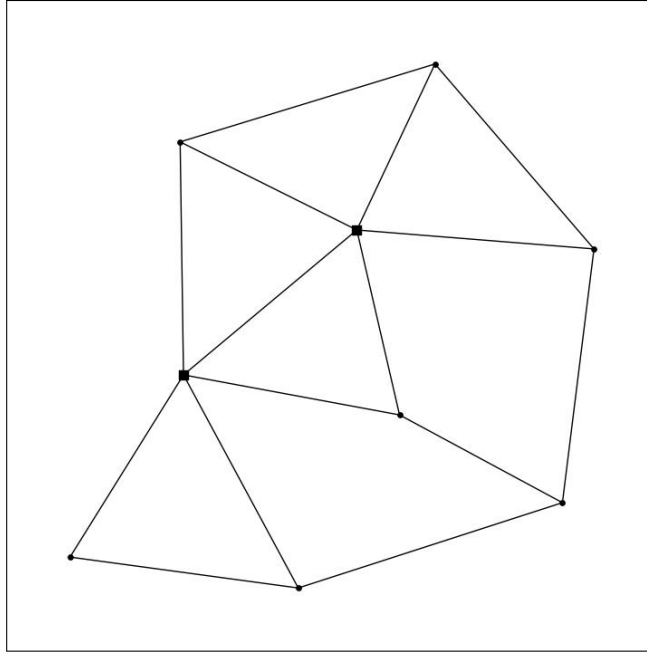


Figure 6: The input image, the answer is '1'

GPT-4o Output:

To find the shortest path between the two square nodes in the graph, we can visually inspect the connections between the nodes.

Looking at the image:

1. Start from the square node on the left side of the graph.
2. There is a direct connection to the node above it.
3. From this node, there is a connection to the square node on the right side.

Therefore, the shortest path consists of **two edges**.

The final answer is **2**.

38

Claude 3.5 Sonnet Output:

To find the shortest path between the two square nodes in this graph, I'll count the minimum number of edges that need to be traversed to get from one square node to the other. Looking at the graph, I can see that there are two potential paths:

1. A direct path that goes through the center node of the graph.
2. A path that follows the outer edges of the graph.

The direct path through the center is clearly shorter. It requires traversing only 2 edges:

1. From the first square node to the center node
2. From the center node to the second square node

The path along the outer edges would require traversing at least 3 edges. Therefore, the shortest path between the two square nodes in this graph is **2**.

39

Claude 3 Opus Output:

The image shows a graph with 6 nodes, where 2 nodes are marked by squares. To find the shortest path between the two square nodes:

1. Starting from the top square node, there are two possible paths: down-left or down-right.
2. Following the down-left path:
Square \rightarrow Circle \rightarrow Square
This path has a length of 2 edges.
3. Following the down-right path:
Square \rightarrow Circle \rightarrow Circle \rightarrow Circle \rightarrow Square
This path has a length of 4 edges.

Therefore, the shortest path between the two square nodes goes through the down-left path, with a length of 2 edges.

The shortest path between the two square nodes is 2.

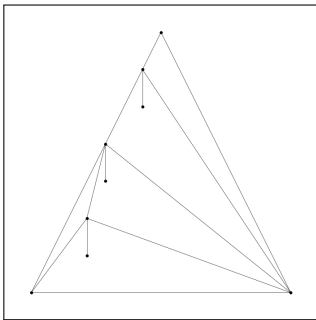
40

41 B LLMs adjacency matrix creation

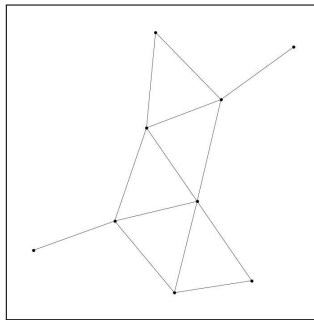
42 In this section, we evaluate the performance of GPT-4, Claude 3 Opus, and Claude 3.5 Sonnet in
43 transcribing graph structures. We presented these models with various graph images and provided the
44 prompt: "The image presented contains a graph. Please transcribe this graph's adjacency matrix." We
45 then analyzed their ability to generate accurate adjacency matrices from the visual input. Our results
46 reveal that all models struggle to correctly transcribe the graphs, with varying degrees of accuracy in
47 extracting basic graph properties. GPT-4 comes close to accurately identifying the number of nodes
48 but struggles with edge counts. Claude 3.5 Sonnet shows the best performance, closely approximating
49 both the number of nodes and edges. In contrast, Claude 3 Opus performs poorly across all metrics.
50 We visualized the graphs transcribed by these models to assess their accuracy. Although Claude 3.5
51 Sonnet was not available at the time of our initial submission, we have incorporated its responses into
52 our analysis, because it demonstrates superior performance compared to both GPT-4 and Claude 3
53 Opus, showing a marked improvement over Opus in particular.

54

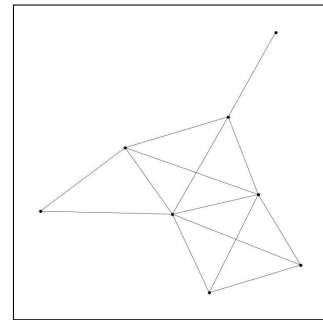
prompt: The image presented contains a graph. Please transcribe this graph's adjacency matrix.



(a) Graph a



(b) Graph b



(c) Graph c

Figure 7: The three graphs we asked the LLMs to generate the adjacency matrix.

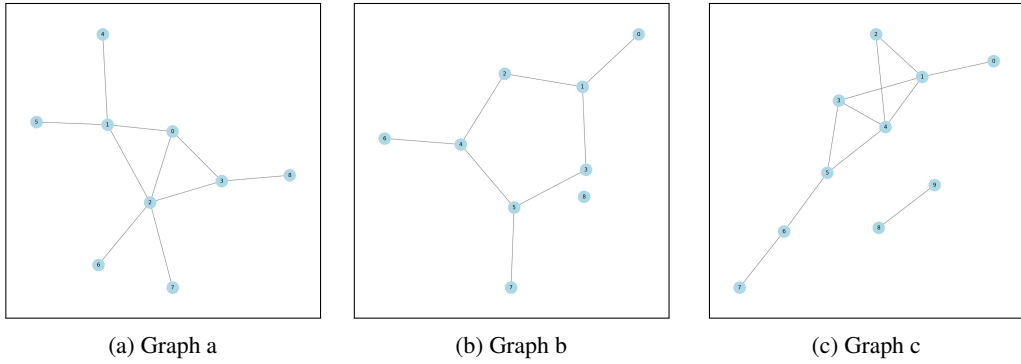


Figure 8: reconstructed graphs from the generated adjacency matrix by GPT-4o.

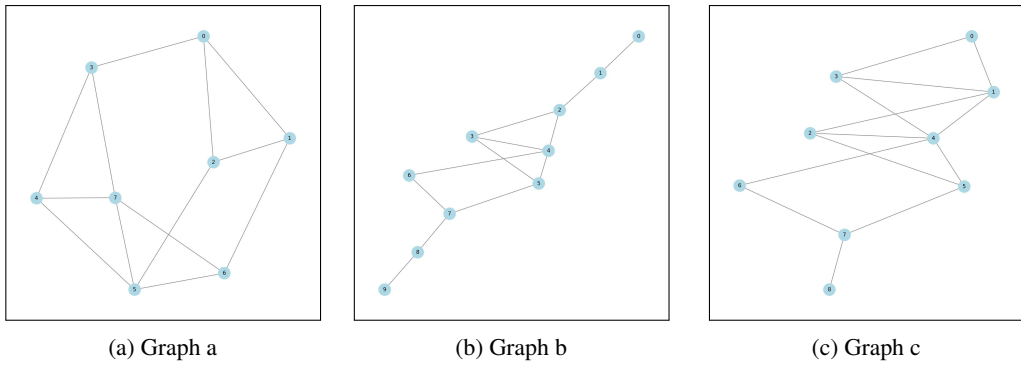


Figure 9: reconstructed graphs from the generated adjacency matrix by Claude 3.5 Sonnet.

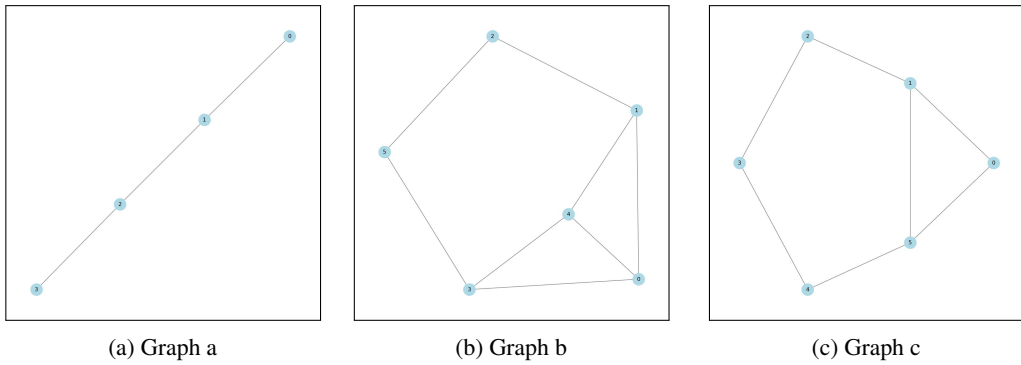


Figure 10: reconstructed graphs from the generated adjacency matrix by Claude 3 Opus.