GraphGPT: A Self Supervised-Learning for Intuitive, Logical, and Visual Education

Nicholas X. Wang
BASIS Independent Silicon Valley, San Jose, USA
nicholas@nicholasxwang.dev

Abstract—**In this paper, we introduce GraphGPT, a novel machine-learning approach designed for intuitive, logical, and visual education. GraphGPT is inspired by the observation that scientific graphs help learners visualize concepts, while causal graphs enhance the logic of responses. By combining the masked prediction capabilities of GPT with transform prediction using causal graphs, we can improve explanation accuracy. GraphGPT generates logical steps alongside diagrams, significantly enhancing intuition and visual presentation compared to ChatGPT. Experimental results show that GraphGPT outperforms ChatGPT by 25% in accuracy for multiple-choice questions. Furthermore, a vast majority of participants in the subjective tests confirmed that GraphGPT surpasses ChatGPT in all four evaluation categories. Many highlighted the diagrams and graphical explanations as key factors, giving GraphGPT a notable intuitive advantage. Our work is available at https://github.com/nicholasxwang/GraphGPT.**

I. Introduction

Education shapes the future of our society, making it a central focus of social impact. However, a recent report by Collegeboard [1] revealed that 60-70% of students scored a three or below on AP exams for STEM courses such as Physics, Macroeconomics, and Calculus. After school, students often have limited resources to address concepts they didn't understand in class. Since the emergence of ChatGPT [2], many students have been using it to answer their questions outside of school. ChatGPT is free, which has led to its extensive usage recently.

However, it has been noted that ChatGPT primarily focuses on language understanding and generation rather than solving math and logic problems [3]. Consequently, many researchers [4-10] are working on advancing large language models (LLMs), such as various versions of GPTs (e.g., GPT-3, GPT-4), to improve their mathematical problem-solving capabilities. On the other hand, intuitive education using advances in LLMs is emerging. For example, Khan Academy [11] created a tool called Khanmigo built on GPT-4. This tool aims to help students find answers through step-by-step interactions and develop their logical reasoning skills. However, these efforts are still in their nascent stages [12], and ChatGPT often struggles to justify its answers with clear logical steps, providing wordy and unintuitive explanations for beginners. Additionally, these efforts have not yet incorporated visual outcomes and interactions into the intuitive education process. Many STEM subjects require strong logic, understanding, and the development of intuition. Explaining concepts using only text is often insufficient. The saying "a picture is worth a thousand words" emphasizes that diagrams and graphs can be more intuitive than paragraphs of text. Therefore, an AI solution for intuitive education that includes visual elements would be highly valuable. This approach could be effectively applied in fields such as Microeconomics, Physics, Chemistry, Biology, Algebra, Geometry, and Calculus.

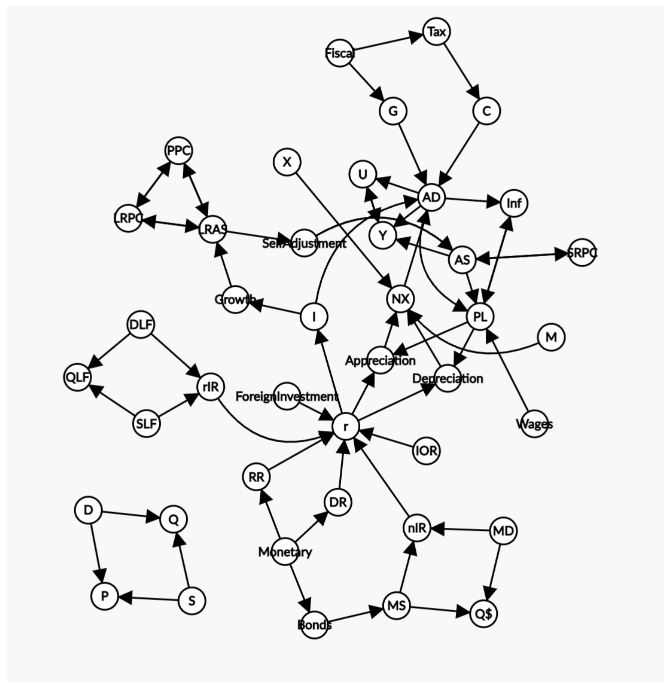


Fig. 1: An example of a causal graph.

In many STEM subjects, cause-and-effect relationships drive logical reasoning, making the study of causality essential. A causal or directed graph [13-16] can define the relationships among all associated components. It has been observed that causal graphs can act as a bridge between a question and an answer. Initially, the causal graph can be mapped from the question, and the logic for the answer can be derived step-by-step on the causal graph. This method corresponds to both text and visual presentations for each step, making the answer more intuitive for students. Based on this observation, we propose a self-supervised learning approach for intuitive education. The pretext task is to create "free" training data by generating a causal graph from the input question. Then, the causal graph and the original text question are embedded and sent into a GPT model for training. Masked prediction is used to predict a preliminary answer, which is further processed to generate step-by-step diagrams as well as a text answer. It is important to emphasize that the diagram drawing is driven by commands contained in the preliminary answer, which are learned from the causal graph. The use of scientific graphs aids in visualizing concepts, while causal graphs enhance logical coherence in responses. This dual approach integrates textual and visual elements to provide more intuitive answers for students.

Our main contribution is in threefold:

(1) Proposed a novel self-supervised learning approach that combines masked prediction with transform prediction using causal graphs, significantly enhancing explanation accuracy.

(2) By employing causal relationship analysis and causal graph representations, GraphGPT can generate coherent sequences of logical steps alongside diagrams. To our knowledge, this marks a pioneering application of Generative AI in this domain.

(3) Furthermore, we have developed a comprehensive mini-dataset and causal graph specifically tailored for economic course questions, answers, and explanations. This initiative represents a pioneering effort in AI data training for this subject area.

II. GraphGPT

In this paper, the examples primarily focus on Microeconomics, yet our theories and technology are applicable beyond this field. Macroeconomics, a crucial social science, studies the behavior of national economies. Economic entities are predominantly curves and variables, exhibiting two primary trajectories: increase and decrease (as shown in Table 1). Economic graphs depict relationships between two variables plotted on x- and y-axes, where each point represents a coordinate defined by two variables. Typically, the equilibrium point signifies the intersection of two curves. However, under specific conditions, this point may not necessarily be at equilibrium. For instance, the graph depicting Supply and Demand models the relationship between Price and Quantity of a product. The equilibrium between supply and demand determines the price and quantity. A shift in demand alters this equilibrium, subsequently affecting price and quantity. In addition to graphs, fundamental relationships exist such as "`price increase => inflation" and "more imports => less GDP.`" Macroeconomic problems can intuitively be addressed using Economic Graphs and Economic Relationships, rather than relying on rote memorization of facts and properties. Representing these relationships, a Causal Graph provides a straightforward method (As shown in Fig. 1).

	INCREASE	**DECREASE**
Curves	Move right	Move left
Variables	Increase	Decrease

TABLE I: Economic Rules

Imagine you're a high school student studying Macroeconomics. Sometimes, when you're given a question and its solution in text (Fig. 2), it might be hard to grasp how each step connects to the final answer. What you really need is a step-by-step explanation that's intuitive and comes with visual representation, as shown in Fig. 3. This kind of intuitive answer not only helps you understand the concept better but also makes it easier to remember it and apply it in the long run.

Let's denote by T the question that needs to be answered, A the final answer, S the logical steps needed to reach the solution, D the diagrams or scientific graphs that pair with S, N the total number of logical steps, thus $S = \{S_1, S_2, \dots S_N\}$ and $D = \{D_1, D_2, ...D_N\}$. As shown in Fig. 3, $N = 4$, and the diagram represents all 5 steps of the D's stacked on top of each other but with arrows to differentiate. The problem at hand can be formulated as the joint maximization of probability of A, S, D given T, that is,

$$\max P(A, S, D|T) \tag{1}$$

To simplify the problem, we define a list of commands C that can be utilized to derive S and D. The syntax of commands C is straightforward and follows the format of {`<Operation> <Operand>`}. The operations available include `NEW` (initiating a graph), `LEFT` or `RIGHT` (moving direction of a curve), and `INCREASE` or `DECREASE` (trending of a quantity). The operands can be one of five types: graph, curve, variable, point, and Economic Policy (e.g., Fiscal). This approach allows for a structured way to manipulate and analyze graphs and curves, facilitating clearer understanding and application across various subjects.

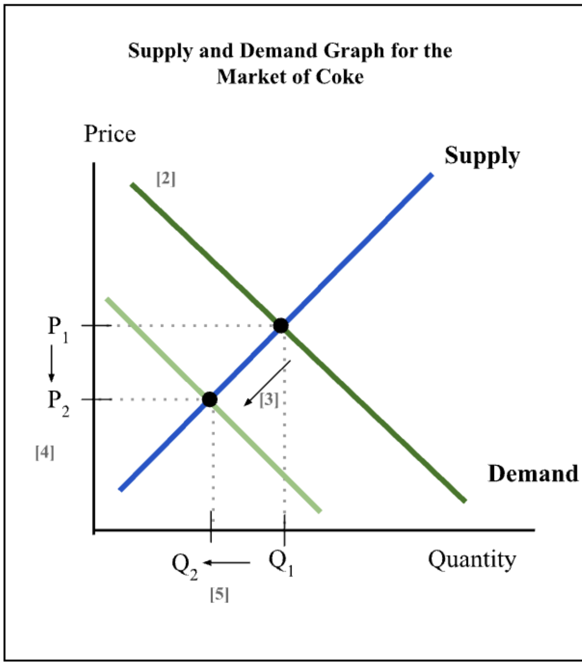
Hence, problem (1) can be rewritten as:

$$\max P(A, C|T) \tag{2}$$

which can be resolved with a neural network, as shown in Fig. 4. There are multiple approaches to consider, but a straightforward one is to employ a GPT model like GPT-3.5. By feeding it with question-answer pairs extracted from a textbook and utilizing masked prediction in unsupervised learning, we can effectively resolve the problem (2). To the best of our knowledge, the current state-of-the-art in GPT models, including the latest GPT-4o, does not demonstrate sufficient capability to directly generate C from T. Therefore, an intermediate representation is necessary to bridge this gap. A causal graph emerges as a viable option because it delineates internal reasoning logic, thereby preventing GPT from generating nonexistent relationships between quantities through hallucination. The primary contribution of the causal graph lies in establishing the correct sequence for the command list. Let G represent the internal graph associated with a question T, where V denotes the set of vertices and E the set of edges, defining $G = (V, E)$ as a directed graph. This structure allows for at most one edge from any vertex to another, capturing specific relationships between them. Vertices without a direct relationship do not have connecting edges. For subjects like Macroeconomics, a master Causal graph can be manually constructed, which is a one-time

Question: Recently, the price of Pepsi, a substitute for Coke, has decreased. How is the price and quantity of Coke affected?
(A) Quantity Increases, Price Increases
(B) Quantity Increases, Price Decreases
(C) Quantity Decreases, Price Increases
(D) Quantity Decreases, Price Decreases **(Correct Answer)**
(E) Not enough information

Fig. 2: Sample question and answer.



[1] S: If Pepsi price drops, demand of Coke drops (as customers prefer Pepsi)
D: `null;` C: `[NULL "null"]`

[2] S: This supply/demand graph would represent the supply and demand for Coke.
D: Initial supply/demand graph (as shown in left image) with dark green and blue line;
C: `[NEW "Supply/Demand"]`

[3] S: The curve of Demand move left, thus equilibrium point (Q_1, P_1) shifts leftwards to (Q_2, P_2)
D: add light green line; C: `[LEFT "D"]`

[4] S: Coke has to decrease price: $P_1 \rightarrow P_2$
D: add label and arrow; C: `["P" DECREASE]`

[5] S: Quantity decreases as a result: $Q_1 \rightarrow Q_2$
D: add label and arrow; C: `["Q" DECREASE]`

Conclusion: the answer is option (D)

Fig. 3: Sample intuitive explanation.

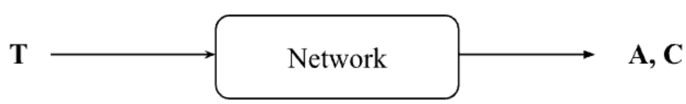


Fig. 4: A neural network to resolve question-answer pair training.

effort. G then functions as a sub-graph of this master Causal graph, represented by lists of vertex indices V and edge indices E. A typical syntax for G 's representation might look as: {`\START": start vertex index, \END": end node index, \EDGES": list of edges in the path from start vertex to end vertex`}. Multiple starting and ending points enable the representation of various paths. Using graph theory, paths can be derived in a straightforward manner from a directed graph once the starting and ending vertices are specified.

The NLP understanding of T and generation of G can be achieved by using a GPT model. It was observed that the accuracy is very high, that is:

$$P(G|T) \approx 1 \tag{3}$$

Once G is obtained, the pair $(T + G)$, $(A + C)$ is fed into a GPT model, which aims to:

$$\max P(A, C|T, G) \tag{4}$$

Considering Eq. (3), we have:

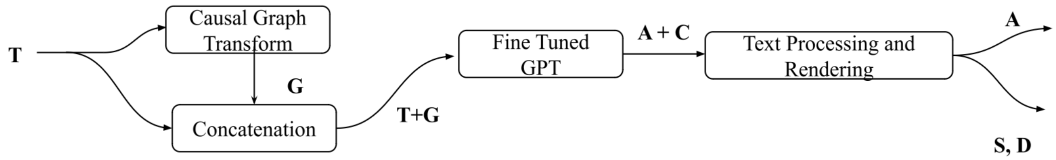


Fig. 5: Proposed GraphGPT network architecture.

$$P(A,C|T,G) = \frac{P(A,C,G|T)}{P(G|T)} \approx P(A,C,G|T)$$
$$= P(A,C|T)P(G|A,C,T) \approx P(A,C|T) \quad (5)$$

as $P(G|A,C,T)$ is at least as large as $P(G|T)$ and therefore also approximated to 1, according to Eq. (3). Hence, the maximization in Eq. (2) can be achieved by maximizing Eq. (4). Based on Eq. (4) and the preceding analysis, the network architecture of GraphGPT can be structured as shown in Fig. 5. The `Causal Graph Transform` module converts T into G. This transformed G is then concatenated with T and input into a `Fine-Tuned GPT` model to train and generate A and C. These outputs undergo processing in a `Text Processing and Rendering` module. Here, A is separated from C, and the commands C are used to derive S and D. This step ensures that the process is presented in a format that is more intuitive for humans, complemented by diagrams to enhance user understanding.

III. Experiments

We implemented GraphGPT by leveraging OpenAI's GPT-3.5-Turbo API, fine-tuning the model with our proprietary dataset. We also developed a master causal graph to enhance our solution. Text concatenation and post-processing functionalities were implemented using Python scripts, while diagram rendering utilized HTML's `<svg>` tag with coordinates generated programmatically in Python. Our dataset comprises over 100 new questions and solutions covering AP Macroeconomics Units 1 through 5, derived from textbooks and mock exams. To validate our approach, we compared our implementation against the publicly available ChatGPT 3.5, evaluating both the accuracy and intuitive performance of GraphGPT. As shown in Fig. 6, it was demonstrated that the accuracy of GraphGPT outperforms ChatGPT by around 25%.

The logical steps generated in the explanation were compared with the ground truth; 90% of the questions had logical steps that matched the ground truth while two had slight errors. Many of the examples with correct explanations that still led to incorrect answers were due to the lack of ability to connect the reasoning to the correct answer option or understanding the wording at the beginning incorrectly. Furthermore, the diagram generated in the explanation was compared with the ground truth, and we found that the accuracy of the graph rendering is 100%.

In our subjective tests, we enlisted 12 volunteers to participate, including 9 individuals without prior economic experience and 3 with economic backgrounds. Each participant reviewed 3 questions along with their solutions generated by GraphGPT and ChatGPT. Subsequently, participants completed a questionnaire evaluating: the answer quality, usefulness of step-by-step derivation, effectiveness of using diagram, and the overall intuitiveness of the solution. Scores from the questionnaire were averaged and normalized within a range of 0 to 10 for comprehensive analysis. As indicated in Fig. 7, GraphGPT received significantly higher scores in all categories compared to ChatGPT. The inclusion of diagrams in GraphGPT garnered substantial recognition, leading participants to award GraphGPT much higher scores in the Overall Intuitive Level category.

The feedback from participants strongly emphasized their preferences, revealing that many identified as visual learners. For example, one participant commented, "The innovation has more graphical explanations along with a sense of a written explanation. It's mainly because I am a visual learner and I learn through graphs along with images rather than through seeing text." Another participant straightforwardly remarked, "I appreciate the diagrams." Additionally, one participant noted, "Without a diagram, it makes the text more cramped and less pleasant to read," while another expressed a similar sentiment: "Showing the change on a graph is much more intuitive than just writing it in text."

IV. Conclusion

In this paper, we introduced GraphGPT, a self-supervised learning approach aimed at enhancing intuitive and visual education through scientific graphical representations and logical steps. Beyond Macroeconomics, our exploration encompassed subjects like Calculus, Microeconomics, Physics, Chemistry, and Biology, revealing the widespread utility of causal graphs across formal, social, and natural scientific disciplines. This versatility positions GraphGPT for broader adoption. As GPT technology continues to advance, we foresee ongoing improvements in GraphGPT's accuracy and capabilities. The rapid evolution in NLP holds tremendous potential to further enhance GraphGPT's effectiveness.

Acknowledgements

We express our gratitude to Professor Aggelos Katsaggelos for his mentorship on this paper. We also thank Alex Chen for verifying the accuracy and sophistication of GraphGPT's economics knowledge. Additionally, we appreciate the many anonymous volunteers who assisted in completing the subjective surveys.

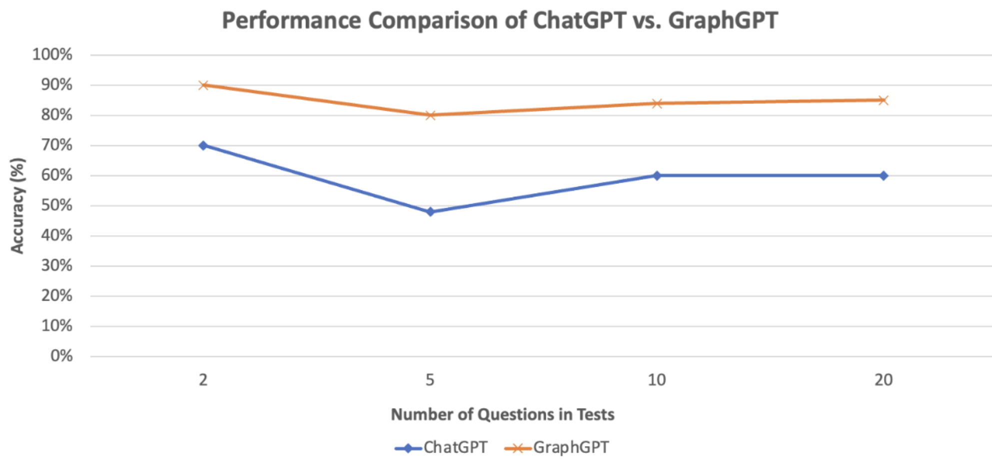


Fig. 6: Accuracy performance comparison of GraphGPT vs. ChatGPT.

	Overall Score	**Overall Score Inexper-ienced group**	**Overall Score Experie-nced Group**	**Textual Quality Score**	**Step by Step Score**	**Diag-ram Score**	**Overall Intuitive Level Score**
GraphGPT	8.4	8.2	8.7	8.1	9.1	8.5	8.0
ChatGPT	6.4	6.4	6.1	6.7	8.1	5.0	5.7

Fig. 7: Scores obtained from the subjective tests.

References

[1] "2023 AP Score Distributions." Collegeboard, apstudents.collegeboard.org/about-ap-scores/score-distributions/2023.

[2] "ChatGPT." OpenAI, chat.openai.com/. (2022)

[3] Zvornicanin, Enes. "Why Is ChatGPT Bad at Math?" Baeldung, 11 Jun. 2024, www.baeldung.com/cs/chatgpt-math-problems.

[4] Wei, Jason, et al. "Chain-of-thought prompting elicits reasoning in large language models." Advances in neural information processing systems 35 (2022): 24824-24837.

[5] Zhang, Beichen, et al. "Evaluating and improving tool-augmented computation-intensive math reasoning." Advances in Neural Information Processing Systems 36 (2024).

[6] Imani, Shima, Liang Du, and Harsh Shrivastava. "Math-prompter: Mathematical reasoning using large language models." arXiv preprint arXiv:2303.05398 (2023).

[7] Yu, Longhui, et al. "Metamath: Bootstrap your own mathematical questions for large language models." arXiv preprint arXiv:2309.12284 (2023).

[8] Ahn, Janice, et al. "Large language models for mathematical reasoning: Progresses and challenges." arXiv preprint arXiv:2402.00157 (2024).

[9] Scarlatos, Alexander, and Andrew Lan. "Tree-based representation and generation of natural and mathematical language." arXiv preprint arXiv:2302.07974 (2023).

[10] Trinh, Trieu H., et al. "Solving olympiad geometry without human demonstrations." Nature, 625.7995 (2024): 476-482

[11] "Khanmigo." Khan Academy, www.khanmigo.ai/. (2023)

[12] Kshetri, Nir. "The economics of generative artificial intelligence in the academic industry." Computer, 56.8 (2023): 77-83.

[13] Giammei, Lorenzo. "An Integrated Approach to Causality: The Role of Causal Graphs." Sapienza Università Di Roma, 2022, https://web.uniroma1.it/memotef/sites/default/files/Giammei _onlinefirst_Phd_2022.pdf.

[14] Schölkopf, Bernhard, et al. "Toward causal representation learning." Proceedings of the IEEE, 109.5 (2021): 612-634.

[15] Kaddour, Jean, et al. "Causal machine learning: A survey and open problems." arXiv preprint arXiv:2206.15475 (2022).

[16] Jin, Zhijing, et al. "Can large language models infer causation from correlation?." arXiv preprint arXiv:2306.05836 (2023).