

ChatGPT Chatbot in the Teaching and Learning of Mathematics: The Role of Reasoning Processes

Nhiry Meryem ⁽¹⁾, Hanini Mohamed ⁽²⁾, Abouhanifa Said ⁽³⁾

⁽¹⁾ IR2M, Université Hassan I-FST Settat, Settat, Maroc . meryem.nhiry@gmail.com

⁽²⁾ Université Hassan I-FST Settat, Settat, Maroc .haninimohamed@gmail.com

⁽³⁾ CRMEF-Settat, Settat, Maroc .saidabouhanifa@yahoo.fr

Article History:

Received: 10-01-2025

Revised: 27-02-2025

Accepted: 07-03-2025

Abstract:

The release of ChatGPT in November 2022 represented a significant advancement in creativity and innovation across various fields, particularly in the teaching and learning of mathematics. This AI-driven chatbot conversationally engages users, utilizing human-like language to respond to inquiries and generate content. It follows prompts and previous interactions, adjusting and refining its answers based on the sequence of inputs and outputs. Our study examines ChatGPT's modeling of mathematical reasoning by analyzing its solutions and proofs across a spectrum of problems with varying levels of complexity. The dataset consists of 12 activities categorized into four types of mathematical reasoning. This research investigates whether this emerging technology, now accessible to students, can improve the relationship between chatbots and students, particularly in terms of reasoning processes. We employed a qualitative descriptive approach to evaluate the reasoning performance of ChatGPT, focusing on its latest versions, GPT-4o and GPT-4o Mini. This evaluation is grounded in a classification of reasoning performance characteristics, emphasizing the structural steps of each reasoning type (Jeannotte, 2015; Nhiry et al., 2023). Our analysis revealed that ChatGPT demonstrates a satisfactory ability to maintain the structure of mathematical reasoning; however, several shortcomings were identified in terms of clarity, rigor, and generalization. Enhancing its formalization, pedagogical explanations, and result validation would strengthen and enrich its approach.

Keywords: ChatGPT; GPT-4o; Generative AI; Education; Mathematics; Mathematical Reasoning

Introduction

In recent years, artificial intelligence (AI) has attracted growing attention and has been the subject of extensive research (Luckin & Cukurova, 2019). AI represents a creative and innovative advancement across various fields, including education, particularly in the teaching and learning of mathematics. ChatGPT, an example of generative AI, illustrates a category of technologies capable of producing text, images, and other computer-generated digital content. In this context, several studies have examined the potential benefits and current limitations of using ChatGPT in mathematics (Taani & Alabidi, 2024; Teegavarapu & Sanghvi, 2023; Sánchez-Ruiz et al., 2023).

The launch of ChatGPT in November 2022 introduced an extraordinary innovation in education with its ability to solve mathematical problems and paraphrase content. ChatGPT can comprehend complex

mathematical formulas and concepts, and by leveraging natural language processing, students can pose mathematical questions and receive quick responses. As a result, ChatGPT has vast potential to enhance education, problem-solving, and student learning. In teaching, ChatGPT can assist educators in generating personalized and relevant educational content for students (Zhu G. et al., 2023). Additionally, the ChatGPT model can be a valuable tool for academic assessments, enabling teachers to quickly evaluate student work and provide feedback (Winter, 2023).

Despite its impressive capabilities, ChatGPT has certain limitations. As a language model, it can sometimes produce incorrect or irrelevant responses and struggle to grasp specific topics' context or nuances. Moreover, it is crucial to acknowledge that ChatGPT remains a tool and cannot replace the role of human teachers and educators (Tenhundfeld, 2023). Its lack of deep understanding and complex analytical abilities are weaknesses that could impact the development of advanced cognitive skills (Farrokhnia et al., 2023). Furthermore, as the number of equations, unknowns, and arithmetic operations such as division and multiplication increase, the probability of obtaining correct mathematical solutions decreases (Shakarian et al., 2023). ChatGPT also fails to provide a deep understanding of geometry (Wardat et al., 2023), and the mathematical tasks it generates often lack real-world context (Onal & Kulavuz-Onal, 2024).

This paper examines ChatGPT's ability to engage in mathematical reasoning, defined as "a certain mental activity, a discursive operation through which one moves from certain premises to a new proposition, by the logical link that connects them" (Blanché, 2019). For this study, we adopted the conceptualization of mathematical reasoning developed by Jeannotte (2015), who proposed a conceptual model comprising four types of reasoning: deductive, inductive, analogical, and abductive. Jeannotte presented these types of reasoning from both structural and procedural perspectives. Additionally, we based our analysis on the framework developed by Nhiry et al. (2023) for evaluating ChatGPT's mathematical reasoning outputs.

The research question guiding this study was: Can ChatGPT serve as a tool to support or enhance the development of mathematical reasoning processes among students?

To address this Question, we first provide a theoretical foundation on the ChatGPT chatbot, covering its advantages and challenges, implications and applications, role in education, and limitations, particularly in mathematics. Next, we present ChatGPT's functionalities and versions, focusing on ChatGPT-4o, through which we analyze the different types of mathematical reasoning targeted in this study.

This study assessed ChatGPT's performance in providing mathematical reasoning using a qualitative research approach.

I. Theoretical Framework

The concept and invention of artificial intelligence (AI) first emerged through the work of American computer scientist John McCarthy during a 1956 conference at Dartmouth, New Hampshire, in the United States. McCarthy coined the term "AI," thereby establishing a new field of study and technological application (Moor, 2006). Between 1964 and 1966, Joseph Weizenbaum developed **ELIZA**, the first chatbot capable of engaging in a dialogue between humans and machines through a question-answer interaction (Weizenbaum, 1966).

Artificial intelligence (AI) is a technology that enables learning and adaptation, allowing computers and software to perform tasks intelligently by imitating human actions. AI will enable systems to execute advanced cognitive functions such as perception, decision-making, problem-solving, generalization, and experience acquisition (Talan, 2021). In essence, AI distinguishes itself through its capacity to replicate human behaviors by learning and adapting, which enables it to handle complex tasks with a high degree of autonomy (Angraini & Muhammad, 2023; Samosir et al., 2023; Triansyah et al., 2023).

AI and ChatGPT in Education

Early education research highlights ChatGPT's constructive potential and disruptive impact. From a constructive perspective, ChatGPT (v.3.5) provides detailed feedback, summarizing student performance with remarkable coherence and alignment with human evaluations (Dai et al., 2023). Large language models (GPT) also support literacy development, ranging from children's storytelling (Li & Xu, 2023) to academic writing (Buruk, 2023; Liu et al., 2024). Moreover, they are effective as pseudo-teachers in educational dialogues (Tack & Piech, 2022), as well as in simulating conversations between student teachers and instructors in teacher training programs (Markel et al., 2023), demonstrating predictive potential in teacher education (Trust, Whalen, & Mouza, 2023).

Regarding perceptions of ChatGPT, Mohammed et al. (2023) developed the **KAP-C** tool to measure knowledge, attitudes, and practices related to its use. The study by Sánchez-Ruiz et al. (2023) revealed that students use ChatGPT occasionally, with notable confidence, particularly among male students.

ChatGPT in Mathematics

While ChatGPT can perform specific mathematical tasks, its logical reasoning abilities remain under investigation. It can enhance mathematics education by facilitating testing, online discussions, and generating personalized questions (Fraiwan & Khasawneh, 2023). ChatGPT can also assist teachers in assessment, allowing them to focus more on students' needs. However, precautions must be taken, as AI's generative nature complicates the distinction between original student work and AI-generated content (Olga et al., 2023).

To examine recent research on integrating artificial intelligence, particularly ChatGPT, into mathematics education, we selected various studies focusing on its pedagogical potential and associated challenges.

ChatGPT in Mathematics Education

Sapkota and Bondurant (2024) found that ChatGPT primarily generates procedural mathematical tasks, often lacking the desired level of complexity. While the tool can create fraction multiplication exercises, some tasks remain vague and imprecise. The authors emphasize that ChatGPT cannot replace human expertise in designing more advanced mathematical problems and recommend a strategic approach to crafting exercises that better stimulate cognitive reflection in the classroom.

Spreitzer et al. (2024) examined the mathematical modeling capabilities of different ChatGPT versions (GPT-3.5, GPT-4.0, GPT-MM). They observed that while these models demonstrate essential competencies, their effectiveness declines as task complexity increases. They call for further research to optimize their integration into education.

Bagno, Dana-Picard, and Reches (2024) explored the use of ChatGPT in linear algebra instruction, noting that the tool responds statistically rather than conceptually, making it inadequate as a substitute for human teachers.

Supriyadi and Kuncoro (2023) predicted an increasing integration of AI in mathematics education to personalize learning, foster critical thinking, and incorporate data and statistical skills. Their study underscores the need for continuous teacher training to ensure equitable and high-quality education in a rapidly evolving technological landscape.

Zafrullah, Hakim, and Angga (2023) reported a strong interest among students in using ChatGPT as a learning tool for mathematics, though they caution against overreliance. They recommend combining ChatGPT usage with teacher supervision to prevent excessive technological dependence.

Asare, Arthur, and Boateng (2023) analyzed ChatGPT's impact on mathematical performance, revealing that students' interest in the subject positively moderates this impact. They suggest integrating ChatGPT while considering student engagement levels to maximize its educational benefits.

Dao and Lee (2023) evaluated ChatGPT on Vietnamese high school mathematics exam questions. While the model performed well on simple questions, it struggled with more advanced concepts such as geometry. Their findings suggest that ChatGPT has pedagogical potential but remains limited in complex domains.

Sánchez-Ruiz et al. (2023) explored the impact of ChatGPT in online engineering courses, revealing that students find it useful. However, concerns arise about its effects on critical thinking skills, indicating the need for adaptations in teaching methods.

Teegavarapu (2023) assessed ChatGPT's performance on American Mathematics Competitions problems, showing that while the model excels in number theory, it struggles with geometry and combinatorics. These findings highlight their limitations in spatial reasoning and advanced problem-solving.

Wardat et al. (2023) examined students' and teachers' perceptions of ChatGPT. Although it is appreciated for its assistance in mathematics, its conceptual understanding and error correction shortcomings complicate its effective integration into learning.

Rane (2023) explored the implications of ChatGPT and generative AI in mathematics education for fostering personalized and interactive learning experiences. While the tool is promising, challenges related to accuracy, bias, and ethical concerns necessitate a balanced approach between human intervention and AI assistance to enhance problem-solving skills.

Overall, this review suggests that while ChatGPT has the potential to enrich mathematics education, technical and ethical challenges must be carefully considered to ensure its effective and beneficial use in educational settings.

ChatGPT and Mathematical Reasoning

In addition to its applications in mathematics education, we reviewed studies that specifically address ChatGPT's capabilities and limitations in mathematical reasoning.

Frieder et al. (2023) analyzed common errors in ChatGPT's interpretation and resolution of advanced mathematical problems. Their study found that while ChatGPT performs well in calculations and pattern recognition, it has notable limitations in logical reasoning, particularly in tasks requiring deep analysis.

Lample, Charton, and Lacroix (2023) examined the ability of large language models, including ChatGPT, to solve complex mathematical problems. They discussed the challenges associated with the formal approaches required in these contexts and proposed strategies to improve mathematical reasoning in such models.

Shakarian et al. (2023) found that ChatGPT performs better when explaining its steps, reducing its failure rate to 20% compared to 84% when no explanations are provided. They also noted that the complexity of operations increases the likelihood of errors.

Plevris et al. (2023) compared ChatGPT-3.5, ChatGPT-4, and Google Bard, observing that while ChatGPT-4 is more effective, it still demonstrates inconsistencies when answering repeated questions. Meanwhile, Google Bard was found to be less reliable for mathematical tasks.

These studies suggest that while ChatGPT possesses fundamental mathematical competencies and can handle simple tasks, it faces significant limitations in solving complex problems requiring deep understanding.

II. ChatGPT

1. Definition and Overview

ChatGPT is an artificial intelligence language model developed by OpenAI. It is designed to generate human-like responses based on text inputs and natural language instructions. It is built on the **GPT (Generative Pre-trained Transformer)** architecture, a deep neural network specialized in text comprehension and generation.

ChatGPT was launched in **November 2022** with its **GPT-3.5** version, marking a significant milestone in OpenAI's series of GPT models, which have been in development since OpenAI's founding in 2015. The first version, **GPT-1**, was released in **2018**, followed by **GPT-4 in March 2023** and **GPT-4o on May 13, 2024**. With each new iteration, OpenAI has pushed the boundaries of generative artificial intelligence, leading to practical applications in productivity, content creation, learning, and more.

Version History

- **GPT-1 (2018):** The first use of transformer architecture in a language model, trained on a vast corpus of internet texts. This model introduced an unsupervised approach to generating coherent text (Radford et al., 2018).
- **GPT-2 (2019):** A significantly larger model with **1.5 billion parameters**, capable of generating fluid and convincing text. GPT-2 sparked discussions on large-scale generative models' applications and ethical concerns (Radford et al., 2019).
- **GPT-3 (2020):** Released with **175 billion parameters**, GPT-3 demonstrated remarkable capabilities in diverse tasks, from text generation to code writing. It popularized AI-based applications in various industries (Brown et al., 2020).

- **ChatGPT with GPT-4 (2023):** An advanced model with improved reasoning abilities and a more refined understanding of complex queries. GPT-4 delivers more precise and nuanced responses and includes optimized versions like **GPT-4-turbo**, which is faster and more cost-effective (OpenAI, 2023).
- **GPT-4o (2024):** This is an enhanced version with multimodal capabilities, extending text and code generation to include audio and image processing. This model is faster than GPT-4 and has improved performance across tasks (Séramour, 2024).
- **GPT-4o Mini (July 2024):** This lightweight and faster variant of GPT-4o replaces GPT-3.5 in ChatGPT to provide more efficient performance (Cadot, 2024).

The evolution of GPT models has brought significant improvements in **performance, comprehension, and accessibility**. Below are the key differences between versions:

Key Differences Between Versions

- **Number of Parameters:** Each successive version increases the number of parameters, enhancing processing power and understanding capabilities.
- **Response Quality:** Recent versions, such as GPT-4, provide more accurate, relevant, and contextually nuanced responses.
- **Applications and Accessibility:** GPT-3 and GPT-4 introduced generative AI into commercial applications, mainly via ChatGPT, making the technology widely accessible through platforms like OpenAI.

2. Mathematical Performance and Pedagogical Implications of ChatGPT

ChatGPT demonstrates **varying performance levels in mathematics**, excelling in **procedural tasks** but facing challenges in problems requiring **deep conceptual reasoning**. Its integration into educational settings raises questions about its role as a **pedagogical tool**, highlighting the need for structured guidance to optimize its use in learning mathematics. This section explores ChatGPT's mathematical capabilities through several key aspects:

- **Problem Solving:** Studies have evaluated ChatGPT's ability to tackle complex mathematical problems, from basic arithmetic to advanced computations. **Frieder et al. (2023)** highlighted its strengths in procedural tasks while emphasizing its limitations when it requires deeper conceptual reasoning and rigorous formalization.
- **Theorem Proofs:** Researchers have also explored ChatGPT's ability to assist with or even autonomously generate mathematical proofs and demonstration strategies. **Polu and Sutskever (2023)** examined its capacity to formulate logical reasoning in theorem proving.
- **Mathematical Question Answering:** ChatGPT's ability to respond to **math-related queries in natural language** is an area of study, assessing how it **interprets problem intent, formulates solutions, and provides clear, accurate responses** (Wang, Liu, & Chen, 2023).
- **Pedagogical Implications:** The impact of ChatGPT on **mathematics education** is also a key research area. Studies focus on how these tools can **enhance learning, personalizing instruction, and address educational challenges** (Zheng, 2023).

III. Mathematical Reasoning

1. Reference model and Analytical Framework

Following a methodological approach of **ana-synthesis**, Jeannotte (2015) developed a **conceptual model** defining **mathematical reasoning (MR)** (Figure 1) through two main aspects: **structural** and **procedural**.

- **The structural aspect** is based on literature in **mathematical logic**, integrating reasoning structures such as **deduction, induction, abduction, and analogy**.
- **The procedural aspect** is derived from a diverse corpus, identifying **action verbs** associated with mathematical reasoning, aligning with the **Anthropological Theory of Didactics** proposed by Chevallard (1998). This aspect includes **eight reasoning processes** divided into two categories:
 - **Five processes focused on identifying similarities:** classifying, comparing, identifying regularities, conjecturing, and generalizing.
 - **Three processes focused on validation:** justifying, proving, and demonstrating.

“This model characterizes mathematical reasoning from a commognitive perspective, defining it as a commognitive process. As an inferential process, it extends mathematical discourse by generating statements from other statements. Its primary function is thus to systematize mathematical discourse, allowing for theoretical formalization.” (Jeannotte, 2005, p. 293)

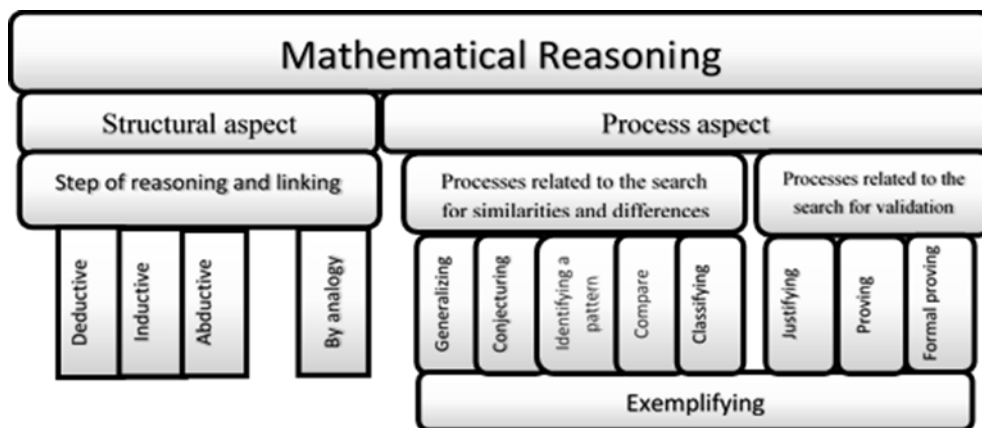


Figure 1: Conceptual Model of Mathematical Reasoning (Jeannotte, 2015, p. 269)

Based on Jeannotte’s (2015) model of mathematical reasoning and the analytical framework provided by **Nhiry et al. (2023)**, we constructed an **analysis grid (Table 1)**. This tool will be used to examine **ChatGPT's reasoning models**, specifically in terms of:

- **Description and Definition of the Reasoning Type:** Provides a global perspective on the type of reasoning and its intended purpose.
- **Structural Steps:** Refers to how **discursive elements** are organized into a structured system, describing the components and their relationships.

- **Correlation with Procedural Steps:** Describes activity within an interaction framework, categorized into **commognitive acts**, including the five **similarity-seeking processes** and three **validation-seeking processes**.
- **Reasoning Steps:** ChatGPT's step-by-step methodology to solve a problem within a specific type of mathematical reasoning.

Table 1: Analytical Framework for Mathematical Reasoning

Type of Reasoning	Description	Definition	Structural Steps	Correlation with Procedural Aspects	Reasoning Steps
Deductive	A reasoning process that moves from the general to the particular.	-Infers a statement from given data and a rule.	Data → Rule → Result	Justification, proof, and demonstration.	- Translate the hypotheses (data) - Combine the data using available information (theorems and rules) - Deduce the result
Inductive	A reasoning process that moves from a specific case to a general case.	- Infers a rule from data and statements about those data. - Moves from particular to general.	Data → Result → Rule	classifying, comparing, identifying regularities, conjecturing, and generalizing.	- Initial observations - Identify the rule (a model) - Formulate a conjecture - Establish the general rule
Analogical	A form of reasoning that draws comparisons between two comparable situations, inferring statements based on their similarities	- Infers statements about a target object based on similarities and differences with a source object. - Combines	Only takes form through a sequence of reasoning steps	comparison, conjecturing, and justification.	- Identify the known situation (source object) - Identify similarities with the given problem (target object) - Apply

	and differences.	at least two elementary steps, which may follow different structures: Inductive and Deductive, Abductive and Deductive.			similarities - Adapt and transfer applicable methods - Conclude the result
Abductive	A reasoning process that starts with results to formulate hypotheses and conjectures to explain a statement.	- Abductive reasoning exists in two forms: 1. Infers data from a statement and a rule. 2. Infers data and a rule from a statement. - Infers data to be explored to validate a given statement.	Result → Statement → Rule → Data	classification and conjecturing, with an element of validation.	- Observe the result - Formulate a statement (hypothesis) - Justify the statement - Conclude a plausible rule - Explore examples to validate the rule - Conclude on the data

2. performance of Mathematical Reasoning

Mathematical reasoning is applying **logical thinking** to a situation to identify the appropriate **mathematical strategy** for solving a problem and using this method to develop and explain a solution. It can also be defined as:

“A certain mental activity, a discursive operation through which one moves from given propositions, considered as premises, to a new proposition, by the logical link that connects them.” (Blanché, 2019)

For this **discursive operation** to be well-structured, reasoning must follow a **logical sequence** that ensures **coherence, rigor, and reliability of proofs**.

To evaluate **mathematical reasoning** and assess its ability to demonstrate solutions rigorously, we consider the following **characteristics**:

- **Clarity and Precision**
- **Logic and Coherence**
- **Relevance of Concepts**
- **Justification of Steps**
- **Mathematical Rigor**
- **Completeness**
- **Proper Use of Mathematical Language**
- **Ability to Generalize**
- **Verification and Validation**

Evaluation Criteria for Mathematical Reasoning

Clarity and Precision

A mathematical reasoning process is **straightforward** when **well-structured and easy to follow** (e.g., using separate paragraphs, numbering, and logical connectors) so the reader or listener can follow the **logical sequence** without ambiguity or confusion. As a form of **mathematical discourse**, reasoning must be **understandable and clearly expressed** (Pimm, 1987).

Logic and Coherence

Reasoning must **strictly follow logical principles**, ensuring that each step derives **explicitly** from the previous without omitting **crucial steps**. Logic and coherence are **essential** for proving a theorem and **effectively communicating** the steps of a demonstration (Hanna, 2000).

Relevance of Concepts

The **correct use of mathematical concepts** is critical for relevant reasoning. Hiebert and Carpenter (1992) state that **conceptual understanding** is essential for students to apply the right concepts at the right time in their reasoning and to solve problems effectively.

Relevance of reasoning also manifests through the **absence of unnecessary digressions**—it must be **concise and directly address** the question posed or solve the problem in a **focused** manner.

Justification of Steps

Every step in reasoning should not contain **unsupported claims**. Each claim must be backed and justified using theorems, definitions, or properties to ensure validity.

"Justifications strengthen reasoning and demonstrate that the student understands mathematical principles, which is fundamental to the validity of a proof." (Stylianides, 2007)

Mathematical Rigor

Rigor is **indispensable** in mathematics. According to Weber (2008), rigor is a **key component** of mathematical reasoning in **precision and attention to detail**, ensuring that each step is logically **sound and flawless**.

A precise reasoning process should **leave no room for ambiguity** and should use **clear notations and concepts**. In some cases, examining all possibilities or considering all conditions in the problem is essential. The **rigor of reasoning** depends on the **accuracy and precision** of each step, ensuring strict adherence to **logical and mathematical principles**.

Completeness

A reasoning process is **complete** if it covers **all cases or situations** the problem defines (e.g., **case distinction reasoning**). Additionally, it must lead to a **coherent conclusion justified by previous steps**.

Selden & Selden (2003) emphasize that **students must pay attention to all aspects of a problem**, ensuring that **no steps or justifications are neglected** so that the reasoning can be considered **complete**.

Proper Use of Mathematical Language

The **correct use of mathematical language**, including **symbols, notations, and terminology**, contributes to **precision** in reasoning.

Morgan (2002) demonstrates that **mastering notation and terminology** is **crucial** for avoiding misunderstandings and effectively communicating in mathematics.

Ability to Generalize

The **ability to generalize** is an **essential mathematical skill** (Mason, Burton & Stacey, 2010). A reasoning process is considered **general** if it applies to a **broad class of situations** or mathematical objects rather than just a **specific case**.

The more **general** a reasoning process is, the more it is often **valued in mathematics**.

Verification and Validation

The **final verification** of solutions allows students to **ensure the accuracy and validity** of their work, reinforcing their understanding of mathematical concepts (Schoenfeld, 1985).

For our study, we have adopted this **classification** to analyze **ChatGPT's outputs** and evaluate its ability to meet these criteria. The evaluation is structured in **Table 2**, which follows below.

Table 2: Evaluation Criteria for ChatGPT's Mathematical Reasoning

Criterion	Clarity and Precision	Logic and Coherence	Relevance of Concepts	Mathematical Rigor	Completeness	Proper Use of Mathematical Language	Ability to Generalize	Verification and Validation
Question to Assess ChatG	Is the reasoning present	Does the reasoning follow consistently	Is each step justified by a rule,	Does the reasoning adhere to mathematical	Does the reasoning lead to a	Are symbols, notations, and	Does the reasoning	Does the reasoning

PT's Reasoning	ted clearly and precisely? Are all steps well explained and justified?	t internal logic? Do the steps flow smoothly without contradiction?	theorem, or mathematical definition? Does the reasoning demonstrate an understanding of underlying principles?	cal rigor, avoiding unjustified approximations or shortcuts?	coherent conclusion justified by the preceding steps?	mathematical terminology used correctly?	demonstrate the ability to be applied in similar contexts or generalized to broader cases?	include a final verification or analysis to validate the obtained result?
----------------	--	---	--	--	---	--	--	---

ChatGPT, an innovative AI technology, assists in documenting and clarifying students' mathematical reasoning. It enables them to receive **immediate feedback**, engage in **interactive dialogues** to solve problems, and access **detailed solutions tailored to their needs** (Dao, X. Q., & Le, N. B., 2023).

For this reason, we aim to **explore the potential of ChatGPT in problem-solving related to mathematical reasoning.**

IV. Methodology

A **purely quantitative research model** (Elliott & Timulak, 2005) was adopted. It utilizes analytical grids to collect data from the ChatGPT chatbot and study the characteristics of the **mathematical reasoning tasks** generated by ChatGPT.

1. Selection Procedures

We used **ChatGPT-4o** (Figure 1) to generate mathematical reasoning tasks. As previously mentioned, our objective was to **generate reasoning tasks** in accordance with Jeannotte's conceptual framework (2015).

The initial inputs were **general**, consisting of **brief textual prompts** such as *“solve the following task”* or uploading **an image containing a task**. Based on an analysis of the initial results, we engaged in a further **interactive dialogue** with ChatGPT, using multiple prompts to refine the complexity of the functions and obtain **cognitively demanding problems** as output. These tasks were described either **through simple text or images**.

The tasks were categorized into **three groups**:

- ✓ **Tasks are available online and can be found via internet searches.**
- ✓ **Tasks extracted from research literature** (published in academic articles) or **mathematics textbooks.**
- ✓ **Original problems** designed by the researchers.

Each task was presented **twice** to each chatbot version:

- By asking: “Is there an alternative solution using X reasoning?”
- By clicking “regenerate” to analyze **ChatGPT’s different approaches** to solving the problem.

This methodological approach allowed us to **systematically examine the quality and variability** of ChatGPT’s mathematical reasoning outputs.

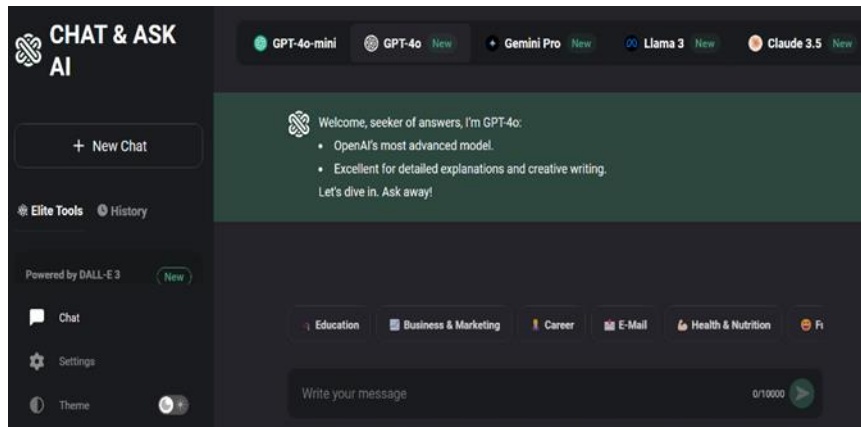


Figure 2: A ChatGPT-4o Output - "Display of the ChatGPT-4o Menu"

2. Data Collection Tool

We selected **12 mathematical reasoning tasks**, with **three tasks assigned to each type of reasoning (deductive, inductive, analogical, and abductive)**, ensuring representation across **three levels of complexity**: Simple, Intermediate, and Complex (See Appendix 1 for details on task selection).

The following section outlines the **characteristics** of these tasks and explains the **rationale behind their selection**.

Table 3: Characteristics of Tasks to be Modeled by ChatGPT

Mathematical Reasoning Type	Task	Concept	Complexity Level	Domain	Representation Type	Rationale for Selection
Deductive	T1	Interval notation, division	Simple	Algèbre	Texte	Utiliser une déduction simple en utilisant les implications successives.
	T2	Algebraic identities and rational expression manipulations	Intermédiaire	Algèbre	Texte	Utiliser une déduction plus difficile en utilisant la donnée

						initiale et sa compréhension et règle de Simplification de fractions en se servant des implications successives.
	T3	Area and symmetry in planar geometry	Complexe	Géométrie	Image	Utilisation du raisonnement déductif en géométrie.
Inductive	T4	Parity of integers, multiplication, consecutive numbers	Simple	Algebra	Text	It involves conjecturing a rule based on observations and pattern recognition.
	T5	Mathematical modeling	Intermediate	Algebra and Arithmetic	Image & Text	A moderately intermediate problem combining sequences, pattern recognition, and algebraic models.
	T6	Arithmetic sequences, mathematical modeling	Complex	Logic and Combinatorics	Image	Combines arithmetic sequences and mathematical modeling in visual patterns and number placements.
Analogical	T7	Trigonometric identities, quadratic equations	Simple	Trigonometry, Algebra	Text	Solves a trigonometric equation by transforming it into a classic quadratic equation.
	T8	Alternate interior	Intermediate	Geometry	Text	It uses alternate interior angle

		angles, the sum of angles				properties as an analogy to determine the sum of a triangle's angles.
	T9	Numerical sequences, square decomposition, mathematical modeling	Complex	Algebra, Discrete Geometry	Text & Image	Engage students by integrating multiple mathematical concepts into a visual and interactive problem.
Abductive	T10	Quadrilaterals : Rhombus and Square	Simple	Geometry	Text	A fundamental problem requiring identifying conditions for a rhombus to be a square.
	T11	Recursive sequences	Intermediate	Algebra	Text	Investigates the Fibonacci sequence and its recursive form in a non-direct manner.
	T12	Planar geometry of quadrilaterals and circles	Complex	Geometry	Text	A geometric proof combining cyclic quadrilaterals, parallelograms, rectangles, rhombuses, squares, diagonal properties, and angle relationships.

V. Analysis and Interpretation of Results

1. Analysis of Mathematical Reasoning by ChatGPT

First, we analyzed whether **ChatGPT could generate a solution for each selected task** in our study. If it successfully provided a solution, we assessed it based on a **scoring scale** evaluating its **mathematical reasoning** (Table 5).

Scoring Scale for ChatGPT’s Mathematical Reasoning

0: ChatGPT failed to generate a solution.

1: The **mathematical reasoning is incomplete or incorrect**, meaning it could not correctly solve the task.

2: ChatGPT **provided a reasoning process, but it was mathematically incorrect** (e.g., informal mathematical language, missing logical connectors, skipped or unjustified steps).

3: The reasoning is **complete and correct but lacks validation**.

4: ChatGPT successfully **modeled the problem as a human would**, demonstrating **complete and rigorous reasoning**.

Note: 0 = Not done; 1 = Done, but incomplete in mathematical reasoning; 2 = Done, complete but mathematically incorrect; 3 = Done, sufficient; 4 = Done, complete and rigorous reasoning.

Table 5: Evaluation of ChatGPT’s Solutions for Selected Tasks

Type of Task Reasoning	Task	ChatGPT’s Performance Score
Deductive	T1	3
	T2	2
	T3	2
Inductive	T4	1
	T5	3
	T6	2
Analogical	T7	1
	T8	2
	T9	3
Abductive	T10	3
	T11	3
	T12	2

ChatGPT successfully generated solutions for all selected problems on the first attempt. Using solution generation and regeneration, we tested ChatGPT by copying the problem and pressing "Enter" or explicitly instructing it with "Solve the following problem using X reasoning."From this, we observed that:

- **Deductive reasoning:**
 - **T1 was sufficient.**
 - **T2 and T3 were mathematically incorrect** due to a **lack of detailed justifications** for reasoning steps.
- **Inductive reasoning:**
 - **T5 (Intermediate level) was sufficient** due to the **problem structure combining text and images** (addressing numerical sequences).
 - **T4 (Simple level) was incomplete in its inductive steps** due to a **lack of explicit generalization.**
 - **T6 (Complex level) was mathematically incorrect** despite being complete in structure.
- **Analogical reasoning:**
 - **T7 lacked clear analogy links**, making the reasoning **incomplete.**
 - **T8 and T9 were complete but mathematically incorrect.**
- **Abductive reasoning:**
 - ChatGPT provided **sufficient solutions** for all three tasks (**T10, T11, and T12**).

Although **ChatGPT attempted to complete the tasks and provide solutions, none of the responses exhibited fully rigorous and mathematically complete reasoning.** However, despite these shortcomings, **only two of the twelve tasks were poorly structured or inadequately presented regarding step-by-step reasoning.**

These findings suggest that **while ChatGPT can generate structured reasoning, it often lacks mathematical precision, formalism, and rigorous justifications**—particularly for **complex reasoning processes such as inductive generalization and analogical reasoning.**

2. Analysis of the Structural Aspect of Mathematical Reasoning by ChatGPT

Next, we analyzed the structural aspect of each type of reasoning for the tasks generated by ChatGPT, assessing whether **ChatGPT adhered to the structural step format** (Table 6).

We used four evaluation criteria:

- **NU (Not Used):** If the process structure does not appear in the given solution for the reasoning type.
- **R (Respected):** If the structure is fully respected.
- **PR (Partially Respected):** If the structure is followed but not explicitly stated.

- **NR (Not Respected):** If the structural step is not followed.

Assessment Scale: NU = Not Used; R = Respected; PR = Partially Respected; NR = Not Respected.

Table 6: Evaluation of Structural Step Appearance in ChatGPT's Solutions

Type of Reasoning	Task	Structural Step
Deductive	T1	R
	T2	R
	T3	R
Inductive	T4	R
	T5	R
	T6	R
Analogical	T7	PR
	T8	PR
	T9	PR
Abductive	T10	R
	T11	R
	T12	R

The results of **structural step analysis** across the different **reasoning types** show that **ChatGPT successfully adhered to the structural form of inductive reasoning (Data → Result → Rule) and abductive reasoning (Result → Statement → Rule → Data)**. This is likely because **both reasoning types involve conjecture formulation based on observations**, making them structurally similar.

For **deductive reasoning**, ChatGPT also maintained its structural integrity (**Data → Rule → Result**) across all three tasks.

However, we observed that ChatGPT only partially adhered to the expected format for analogical reasoning, which requires a sequence of structured steps. The reasoning steps were often **implicit**, making the analogy **less explicitly structured**.

ChatGPT successfully preserved the structural format of mathematical reasoning across most reasoning types, though **analogical reasoning required a more precise step-by-step presentation** to improve the **explicit formulation**.

3. Analysis of the Procedural Aspects of Mathematical Reasoning by ChatGPT

We examined the **procedural aspects** used by ChatGPT in its solutions (Table 7) and evaluated them according to each type of reasoning. This assessment follows the **reference model and analytical framework** described in **Table 1**.

The qualifications used for evaluation are divided into three indicators:

- **ND (Not Given):** If the aspect is **completely absent** from the solution.
- **PD (Partially Developed):** If the aspect is **present but incomplete or poorly structured**.
- **C (Complete):** If the aspect is **fully present, well-structured, and clearly explained**.

Findings from the Analysis

- **Deductive reasoning:**

- The process **“Justify”** is generally well-respected, with satisfactory presence in all tasks. However, the quality varies: it is **complete for Task 1** but **partially developed** (justification lacking depth and details) for **Tasks 2 and 3**.

- The process **“Prove”** shows a weak correlation, being **absent in Tasks 2 and 3** (solutions were limited to steps without explicit proofs) and **partially used in Task 1** (proofs were long, indirect, and sometimes excessively detailed).

- **Inductive reasoning:**

- This reasoning type is correlated with two aspects: **“Identifying Regularity”** and **“Conjecturing,”** both of which appear **explicitly and entirely** in the three solutions provided by ChatGPT.

- The third aspect, **“Generalizing,”** varied:

- It was **well presented in Task 5**.
- **Partially given in Task 4** (informal generalization).
- **Not given in Task 6**.

- **Analogical reasoning:**

- For **Task 9**, all procedural aspects related to **analogical reasoning** were **utterly present** in the solution given by ChatGPT.

- However, for **Tasks 7 and 8**, the aspects **“Classifying,” “Conjecturing,”** and **“Generalizing”** were **absent**.

- The aspect **“Identifying a Regularity”** was **partially addressed**, while **“Comparing”** was **partially present in Task 7** and **fully addressed in Task 8**.

- **Abductive reasoning:**

- This reasoning type integrates **all procedural aspects** related to **similarity detection and validation** due to its mechanism, which relies on a **double affirmation**: first, data is established from an initial statement, then additional data or rules are inferred.

- For **Tasks 10 and 11**, all processes were **present**, except for:

- **“Prove” (partially developed)**.
- **“Demonstrate” (absent)**.

- For **Task 12**, “**Classifying**” and “**Identifying Regularities**” were **fully developed**, while “**Comparing**,” “**Conjecturing**,” “**Generalizing**,” and “**Demonstrating**” were either **partially addressed or absent**.

Overall, the evaluation of ChatGPT’s performance in applying different reasoning types showed **generally satisfactory results**. It demonstrated a **strong ability to respect key correlations** in reasoning, although **some tasks required further refinement to ensure full rigor and completeness**.

Evaluation. ND: Not Given; PD: Partially Given; C: Complete.

Table 7: Evaluation of the Correlation Between ChatGPT's Solutions and Procedural Aspects

TASKS	CORRELATION WITH PROCEDURAL ASPECTS							
	Classifying	Comparing	Identifying Regularity	Conjecturing	Generalizing	Justifying	Providing	Demonstrating
T1						C	PD	
T2						PD	ND	
T3						PD	ND	
T4			C	C	PD			
T5			C	C	C			
T6			C	C	ND			
T7	ND	PD	PD	ND	ND			
T8	ND	C	PD	ND	ND			
T9	C	C	C	C	C			
T10	C	C	C	C	C	C	PD	ND
T11	C	C	C	C	C	C	PD	ND
T12	C	PD	C	PD	PD	NC	NC	PD

4. Analysis of ChatGPT’s Mathematical Reasoning Performance

This section represents the **performance of the solutions modeled by ChatGPT**. To extract results effectively, we used a **note-taking system** (Table 8). This written transcription allowed us to **conclude the criteria for evaluating mathematical reasoning performance**, including:

- **Logic and coherence**
- **Relevance of concepts**
- **Justification of steps**
- **Mathematical rigor**
- **Completeness**

- **Proper use of mathematical language**
- **Ability to generalize**
- **Verification and validation**

These criteria, detailed in the “**Performance of Mathematical Reasoning**” section, are evaluated based on the questions presented in **Table 2**.

Key Observations

- **Tasks with Strong Performance :**
 - **T9, T5, and T11** excelled in **coherence, completeness, conceptual accuracy, and validation of results.**
 - These tasks demonstrated **strong organization, mastery of relevant concepts, and effective generalization.**
- **Tasks with Weak Performance :**
 - **T12, T2, T3, and T4** showed **deficiencies in rigor, justification of steps, completeness, and validation.**

Table 8: Scoring System for Evaluating ChatGPT's Mathematical Reasoning Performance.

Score	Evaluation	Description
0	Absent	The criterion is not addressed or does not appear in the solution.
1	Unacceptable	There is no effort to address the criterion; explanations are entirely incorrect.
2	Very Insufficient	The criterion is mentioned but incorrect or incomplete.
3	Insufficient	The criterion is barely addressed, lacking relevance or coherence.
4	Weak	The criterion is present but highly incomplete, lacks rigor, or is poorly developed.
5	Average	The criterion is addressed but insufficient or imprecise, missing key elements.
6	Satisfactory	The criterion is partially satisfied, but significant gaps remain in rigor or completeness.
7	Good	The criterion is correctly addressed but could benefit from more details or profound analysis.

8	Very Good	The criterion is primarily satisfied with minor omissions or a missing step.
9	Excellent	The criterion is almost perfectly addressed, with only minor imprecisions.
10	Perfect	The criterion is fully satisfied, with no flaws—rigorous, complete, and presented.

- **ChatGPT performed well in generating structured solutions but often lacked full rigor and justification.**
- **Inductive and abductive reasoning tasks** were generally **well-executed**, while **analogical reasoning tasks** were often **partially structured or implicit**.
- **Some tasks lacked validation or had incomplete proofs**, requiring deeper reasoning analysis.

This **qualitative evaluation** highlights ChatGPT’s **strengths and limitations in mathematical reasoning**, emphasizing the need for **enhanced validation and structured reasoning in future AI-generated solutions**.

Table 9: Evaluation of ChatGPT’s Modeling Performance on the 12 Tasks

Task	Logic and Coherence	Relevance of Concepts	Justification of Steps	Mathematical Rigor	Completeness	Proper Use of Mathematical Language	Ability to Generalize	Verification and Validation
T1	10	9	8	9	9	8	6	8
T2	8	10	6	6	4	8	6	4
T3	8	9	6	5	4	6	5	4
T4	10	8	6	4	4	8	6	4
T5	9	10	7	7	6	8	10	10
T6	8	9	7	6	7	8	9	6
T7	9	9	7	8	9	8	6	7
T8	8	9	7	6	7	8	6	7
T9	9	10	8	9	10	9	10	8
T10	9	8	9	6	7	8	7	9
T11	10	8	10	6	8	9	7	10
T12	7	8	6	5	5	7	6	4

<p>Key Strengths of the Criteria in the Tasks</p>	<p>Demonstrate a fluid organization in: T1, T4, T5, and T11</p>	<p>Reflect an excellent selection of concepts: T2, T5, and T9</p>	<p>Offer a detailed justification in T11 T2, T3, T4, and T12</p>	<p>Show strong rigor T1 and T9</p>	<p>Show an in-depth approach to T9 and T5</p>	<p>Use precise language. T9 et T11</p>	<p>Show excellent generalization ability. T5 and T9</p>	<p>Systematically validate their results. T5, T9, and T11</p>
<p>Key Weaknesses of the Criteria in the Tasks</p>	<p>Require better articulation of ideas. (T12)</p>	<p>Some tasks could benefit from a better selection of concepts.</p>	<p>Lack of clarity in their explanations.</p>	<p>Should incorporate formal demonstrations. T3, T4, and T12</p>	<p>Do not cover all required aspects. T2, T3, T4, and T12</p>	<p>It could further formalize its notations. T12</p>	<p>Lack of in-depth reflection for broader application. T3, T4, and T12</p>	<p>T12 Is the weakest task, with low rigor, justification, and validation scores.</p>

5. Limitations of Mathematical Reasoning Generated by ChatGPT

This analysis highlights ChatGPT's mathematical reasoning's structural and conceptual weaknesses regarding **rigor, clarity, and generalization** (Table 10). It evaluates the **methodological approach** used to solve problems, emphasizing the shortcomings that compromise the **validity and scope of generalizations**.

Table 10: Weaknesses in ChatGPT-Generated Reasoning on the 12 Tasks

Tâches	Points faibles des raisonnements modérés par ChatGPT
<p>T1</p>	<ul style="list-style-type: none"> - A limited understanding of reasoning may arise due to the lack of a geometric approach or clear numerical verification. - Accessibility for less experienced users is restricted due to the lack of explicit explanations regarding interval constraints after division and missing detailed calculations related to limits. - Lacks deep reflection on generalization, potentially weakening the scope and rigor of the reasoning.
<p>T2</p>	<ul style="list-style-type: none"> - Incomplete simplification of the sum reduces readability and gives the impression of an unfinished conclusion.

	<ul style="list-style-type: none"> - Requires more in-depth justification, especially during the common denominator step and final simplification. - Lack of detailed explanations may compromise clarity for some readers. - No explicit mention of generalization, reducing the impact and reach of the reasoning.
T3	<ul style="list-style-type: none"> - Incomplete simplification of the sum affects clarity. - A detailed justification is necessary, especially for common denominator steps and final simplifications. - Crucial steps, such as demonstrating zero-sum, are not sufficiently detailed. - Lacks explicit mention of generalization, reducing the impact and significance of the argument.
T4	<ul style="list-style-type: none"> - The generalization of the property to all integers is not formally demonstrated. - Lacks mathematical rigor needed to ensure validity. - The reasoning is incomplete, as it does not cover all possible cases universally. - No global validation for all integers, reducing the scope and robustness of the argument.
T5	<ul style="list-style-type: none"> - Lacks rigor and completeness due to the absence of formal proof.
T6	<ul style="list-style-type: none"> - The sum formula is implicitly applied but not mentioned or rigorously used. - Justification remains intuitive, lacking formal proof or detailed calculations to validate each step. - No explicit verification of values using cumulative sum formulas. - Missing complete rule formalization and explicit validation of all proposed values. - Some notions are vaguely stated when a detailed formalization with explicit sum formulas would improve precision and clarity. - Generalization could be reinforced through clear mathematical proof. - Direct validation by calculation would make the reasoning more robust.
T7	<ul style="list-style-type: none"> - The choice of $\cos(2x) = 1 - 2\sin^2(x)$, is not justified, while $\cos(2x) = 2\cos^2(x) - 1$ could also have been used. - Intermediate steps (e.g., simplifications, rejection of $\sin(x) = 4$) are not detailed or justified. - No clear analogy link is established in the reasoning. - Could benefit from a broader perspective, improving understanding. - Mathematical language could be more formal, especially when explaining solution rejections. - No attempt at generalization to similar equation classes, limiting scope. - Lacks explicit validation, reducing demonstration rigor.
T8	<ul style="list-style-type: none"> - Justification of angle distribution on a straight line could be more detailed for better clarity. - Lacks an explicit step linking the straight-line situation to the triangle, making it less accessible for less experienced readers.

	<ul style="list-style-type: none"> - Demonstration remains informal, especially regarding the parallel construction, which should be referenced using a postulate or geometric property. - More formal property formulation would improve rigor. - Lacks sufficient details on why the approach applies to all triangles. - Precise terminology is necessary to structure the proof better. - Explicit verification of different triangle types (right, obtuse) would strengthen the method's validation. -
T9	<ul style="list-style-type: none"> - The transition between pattern observation and general rule formulation could be better explained. - More detailed justification for white block formulas would enhance understanding. - Mathematical expressions should be more pedagogically explained for clarity. - No explicit verification (e.g., recalculating white blocks per design), limiting formula validation. - Lacks alternative verification methods, such as geometric or visual approaches.
T10	<ul style="list-style-type: none"> - Justification should include a discussion on condition necessity. - More formalization (e.g., geometric notations or theorems) is required. - Lacks formal proof tools, such as induction or rigorous proofs. -
T11	<ul style="list-style-type: none"> - Mathematical language could be more formal. - No global proof ensuring the rule applies universally. - Lacks verification for all terms. -
T12	<ul style="list-style-type: none"> - - Lacks thorough analysis of alternatives: The solution does not explicitly distinguish a rhombus from a square. - Hypothesis is not explicitly formulated. The reasoning jumps to the conclusion without stating the assumption as a proposition to be verified. - Lacks verification rigor. - Insufficient rigor, completeness, and justification for high-level mathematical standards. -

❖ **Limitations of ChatGPT in Mathematical Reasoning**

While ChatGPT demonstrates promising capabilities in mathematics, certain limitations persist in the **clarity of reasoning, generalization of results, and rigor of demonstrations**. This section highlights key **methodological and pedagogical shortcomings** and suggests areas for improvement to enhance the **accuracy and reliability of AI-generated solutions**.

Limitations in Approach and Clarity

- **Lack of a geometric approach or precise numerical verification** to support reasoning.
- **Missing explicit explanations** for crucial steps, such as **bounding constraints after division** or **common denominator application**.

- **Informal or imprecise language**, reducing accessibility for less experienced readers.

Deficiencies in Generalization

- **Lack of formal proof** to extend results to **all possible cases** (e.g., all integers or triangles).
- **No explicit validation** to ensure the generality of formulas or properties.
- **No attempt to contextualize or generalize** to a broader class of equations or problems.

Mathematical Rigor and Demonstration

- **Incomplete reasoning and lack of detailed formalization** of key steps (e.g., simplifications, choice of trigonometric identities).
- **Intuitive justifications without detailed calculations** or use of formal tools like induction.
- **Lack of explicit verification** for different cases or configurations (e.g., various types of triangles).

Validation and Pedagogy

- **Missing result verifications or recalculations**, especially for proposed formulas (e.g., white block calculations, angular properties).
- **Insufficient pedagogical explanations**, making reasoning less accessible to inexperienced users.
- **Lack of alternative approaches** to strengthen the robustness of reasoning.

Structural and Terminological Weaknesses

- **Weak transitions** between observations and general rules, impacting clarity.
- **Lack of formalization** in mathematical notation, properties, or theorems.
- **Implicit formulation of hypotheses or conclusions**, reducing rigor.

Key Areas for Improvement

- **Incorporating formal and universal proofs** to validate rules and reasoning.
- **Exploring alternative cases for comprehensive analysis** (e.g., distinguishing rectangles from rhombuses in cyclic quadrilaterals).
- **Justifying choices**, such as trigonometric identities used or formulated geometric hypotheses.

CONCLUSION

Integrating artificial intelligence into mathematics education represents a **significant advancement**, offering opportunities for **personalization and interactivity** in learning. As an advanced language model, ChatGPT demonstrates a **notable ability** to structure mathematical reasoning, model problems, and provide adapted solutions. However, this study highlights **methodological and conceptual limitations** that **restrict its effectiveness in more demanding contexts**.

Based on the **analytical framework defined by Jeannotte (2015) and Nhiry et al. (2023)**, we evaluated ChatGPT's ability to preserve the **fundamental structures of mathematical reasoning**.

The results indicate that while AI **successfully models inductive, deductive, abductive, and analogical reasoning**, it **struggles with step-by-step explanation, justification, and generalization of results**. More specifically, **inductive and abductive reasoning are well maintained**, whereas **deductive reasoning is only partially coherent and sometimes lacks rigorous justifications**. Analogical reasoning, however, is handled more implicitly, negatively impacting **the clarity and precision of the generated solutions**.

Our findings align with previous studies. **Frieder et al. (2023)** highlighted **ChatGPT's recurring errors** in solving advanced problems, particularly its **inability to ensure logical coherence in complex tasks**. Similarly, **Lample, Charton, and Lacroix (2023)** identified challenges in **applying formal mathematical principles** and proposed improvements to enhance AI models' precision. **Shakarian et al. (2023)** demonstrated that **ChatGPT significantly reduces its error rate when explicitly structuring its reasoning steps**, emphasizing the importance of **better solution structuring**. Finally, **Plevris et al. (2023)** compared multiple AI models and confirmed that **ChatGPT-4, while improved, still presents inconsistencies**, especially when faced with repeated questions or requiring in-depth analysis.

Thus, although ChatGPT is a **promising tool for mathematics assistance**, it **cannot replace human expertise**, particularly in the **validation and rigorous formalization of solutions**. Its effectiveness remains **contingent on pedagogical supervision and user guidance**, ensuring its potential is harnessed without **compromising mathematical rigor**. Improvements are necessary to **enhance its ability to formalize demonstrations, generalize results, and clarify underlying reasoning**. A deeper exploration of AI reasoning mechanisms and the **development of better-adapted algorithms** could pave the way for **more effective and reliable integration of these technologies** into scientific and mathematical education.

REFERENCE

- [1] Angraini, L. M., Larsari, V. N., Muhammad, I., & Kania, N. (2023). Generalizations and analogical reasoning of junior high school viewed from Bruner's learning theory. *Infinity Journal*, 12(2), 291-306.
- [2] Asare, B., Arthur, Y. D., & Boateng, F. O. (2023). Exploring the impact of ChatGPT on mathematics performance: The influential role of student interest. *Education Science and Management*, 1(3), 158-168.
- [3] Bagno, E., Dana-Picard, T., & Reches, S. (2024). ChatGPT in Linear Algebra: Strides Forward, Steps to Go. arXiv preprint arXiv:2403.15399.
- [4] Blanché, R. (2019). « Raisonement ». *Encyclopædia Universalis [en ligne]* : universalis.fr/encyclopedie/raisonnement.
- [5] Brown, T. B. (2020). Language models are few-shot learners. arXiv preprint arXiv:2005.14165.
- [6] Buruk, O. O. (2023, October). Academic writing with GPT-3.5 (ChatGPT): reflections on practices, efficacy and transparency. In *Proceedings of the 26th International Academic Mindtrek Conference* (pp. 144-153).
- [7] Cadot, J. (2024, 19 juillet). La version gratuite de ChatGPT devient meilleure grâce à GPT-4o mini. Numerama. <https://www.numerama.com/tech/1777874-la-version-gratuite-de-chatgpt-devient-meilleure-grace-a-gpt-4o-mini.html>

- [8] Dai, S., Shao, N., Zhao, H., Yu, W., Si, Z., Xu, C., ... & Xu, J. (2023, September). Uncovering chatgpt's capabilities in recommender systems. In Proceedings of the 17th ACM Conference on Recommender Systems (pp. 1126-1132).
- [9] ChatGPT website (n.d.). Available at: <https://chat.openai.com/>
- [10] Elliott, R., & Timulak, L. (2005). Descriptive and interpretive approaches to qualitative research. *A handbook of clinical and health psychology research methods*, 1(7), 147-159.
- [11] Farrokhnia M, Banihashem SK, Noroozi O, et al. (2023) A SWOT analysis of ChatGPT: implications for educational practice and research. *Innov Educ Teach Int* 1–15. <https://doi.org/10.1080/14703297.2023.2195846>.
- [12] Favier, S., & Chanudet, M. (2021). Les démarches et modes de raisonnement en jeu dans les problèmes de "recherches & stratégies" en 10H. *Revue de Mathématiques pour l'école*, 235, 88-98.
- [13] Fraiwan, M., & Khasawneh, N. (2023). A review of ChatGPT applications in education, marketing, software engineering, and healthcare: Benefits, drawbacks, and research directions. *arXiv preprint arXiv:2305.00237*.
- [14] Frieder, S., Pinchetti, L., Griffiths, R. R., Salvatori, T., Lukasiewicz, T., Petersen, P., & Berner, J. (2023). Mathematical capabilities of ChatGPT. *Advances in neural information processing systems*, 36, 27699-27744.
- [15] Frieder, S., et al. (2023). Analyzing the Frequent Errors of ChatGPT in Advanced Mathematical Problem Solving. *Journal of Artificial Intelligence Research*, 78, 1001–1025.
- [16] Jeannotte, D. (2015). Mathematical reasoning: proposing a conceptual model for learning and teaching in elementary and secondary schools [Doctoral dissertation, University of Quebec at Montreal]. *Archipel*. <https://archipel.uqam.ca/8129/>
- [17] Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational studies in mathematics*, 44, 5-23.
- [18] HIEBERT, J. (1992). Learning and Teaching with Understanding. *Handbook of Research on Mathematics Teaching and Learning*, 65-97.
- [19] Lample, G., Charton, F., & Lacroix, T. (2023). Assessing Large Language Models' Ability to Solve Complex Mathematical Problems. *Advances in Neural Information Processing Systems (NeurIPS)*, 36.
- [20] Li, Z., & Xu, Y. (2023). Designing a realistic peer-like embodied conversational agent for supporting children's storytelling. *arXiv*, 2304.09399.
- [21] Liu, Z., Yao, Z., Li, F., & Luo, B. (2024, December). On the detectability of ChatGPT content: benchmarking, methodology, and evaluation through the lens of academic writing. In Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security (pp. 2236-2250).
- [22] Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, 50(6), 2824–2838. <https://doi.org/10.1111/bjet.12861>
- [23] Markel, J. M., Opferman, S. G., Landay, J. A., & Piech, C. (2023, July). Gpteach: Interactive training with GPT-based students. In Proceedings of the tenth ACM conference on learning@scale (pp. 226-236).

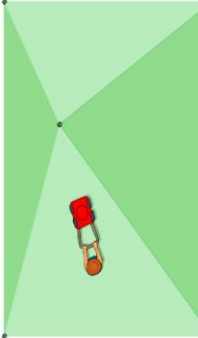
- [24] Mason, J., Burton, L., & Stacey, K. (2010). *Thinking Mathematically* Second Edition. England: Pearson Education Limited.
- [25] Mohammed, M., Kumar, N., Zawiah, M., Al-Ashwal, F. Y., Bala, A. A., Lawal, B. K., Wada, A. S., Halboup, A., Muhammad, S., Ahmad, R. et al. (2023). Psychometric properties and assessment of knowledge, attitude, and practice towards ChatGPT in pharmacy practice and education: a study protocol. *Journal of Racial and Ethnic Health Disparities*, (pp. 1–10).
- [26] Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems*, 21(4), 18-21. <https://doi.org/10.1109/MIS.2006.80>.
- [27] Morgan, C. (2002). *Writing mathematically: The discourse of investigation*. Routledge.
- [28] Nhiry, M., Abouhanifa, S., & El Khouzai, E. M. (2023). The characterization of mathematical reasoning through an analysis of Morocco's high school curricula and textbooks. *Cogent Education*, 10(1), 2188797.
- [29] Olga, A., Saini, A., Zapata, G., Sears Smith, D., Cope, B., Kalantzis, M., ... & Kastania, N. P. (2023). Generative AI: Implications and applications for education. arXiv preprint arXiv:2305.07605.
- [30] Onal, S., & Kulavuz-Onal, D. (2024). A Cross-Disciplinary Examination of the Instructional Uses of ChatGPT in Higher Education. *Journal of Educational Technology Systems*, 52(3), 301-324. <https://doi.org/10.1177/00472395231196532>
- [31] OpenAI.(2023). ChatGPT (GPT-4) [Large language model]. <https://www.openai.com/>
- [32] Pimm, D. (2019). *Routledge Revivals: Speaking Mathematically (1987): Communication in Mathematics Classrooms*. Routledge.
- [33] Plevris, V., et al. (2023). Chatbots put to the Test in Math and Logic Problems: A Preliminary Comparison and Assessment of ChatGPT-3.5, ChatGPT-4, and Google Bard—arXiv preprint arXiv:2305.18618.
- [34] Polu, S., & Sutskever, I. (2023). Exploring the Potential of Large Language Models in Theorem Proving. *Proceedings of the Conference on Artificial Intelligence and Theorem Proving*, 45(3), 102–120.
- [35] Radford, A. (2018). Improving language understanding by generative pre-training.
- [36] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8), 9.
- [37] Rane, N. (2023). Enhancing mathematical capabilities through ChatGPT and similar generative artificial intelligence: Roles and challenges in solving mathematical problems. Available at SSRN 4603237.
- [38] Sánchez-Ruiz, L. M., Moll-López, S., Nuñez-Pérez, A., Moraño-Fernández, J. A., & Vega-Fleitas, E. (2023). ChatGPT challenges blended learning methodologies in engineering education: A case study in mathematics. *Applied Sciences*, 13(10), 6039.
- [39] Sapkota, B., & Bondurant, L. (2024). Assessing concepts, procedures, and cognitive demand of ChatGPT-generated mathematical tasks. *International Journal of Technology in Education*.
- [40] Schoenfeld, A. H. (1985). Making sense of “out loud” problem-solving protocols. *The Journal of Mathematical Behavior*, 4(2), 171-191.

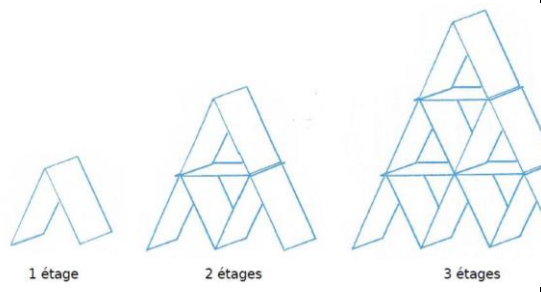
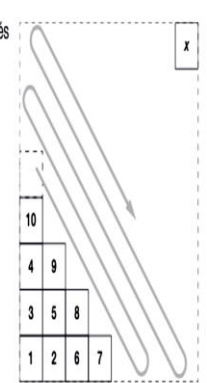
- [41] Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for research in mathematics education*, 34(1), 4-36.
- [42] Seramour, C. (2024, 13 mai). Avec GPT-4o, OpenAI veut écraser les autres modèles multimodaux. *L'Usine Digitale*. <https://www.usine-digitale.fr/article/avec-gpt-4o-openai-veut-eccraser-les-autres-modeles-multimodaux.N2212831?preview=true&email=100043503&idbdd=2102282>.
- [43] Setiawan, A., & Luthfiyani, U. K. (2023). Penggunaan ChatGPT untuk pendidikan di era education 4.0: Usulan inovasi meningkatkan keterampilan menulis. *JURNAL PETISI (Pendidikan Teknologi Informasi)*, 4(1), 49-58.
- [44] Shakarian, P., et al. (2023). An Independent Evaluation of ChatGPT on Mathematical Word Problems (MWP). *Proceedings of the Association for Computational Linguistics (ACL)*, 45(2), 205–221
- [45] Spreitzer, C., Straser, O., Zehetmeier, S., & Maaß, K. (2024). Mathematical Modelling Abilities of Artificial Intelligence Tools: The Case of ChatGPT. *Education Sciences*, 14(7), 698.
- [46] Stylianides, A. L. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, 38(3), 289-321.
- [47] Supriyadi, E., & Kuncoro, K. S. (2023). Exploring the future of mathematics teaching: Insight with ChatGPT. *Union: Jurnal Ilmiah Pendidikan Matematika*, 11(2), 305–316. <https://doi.org/10.30738/union.v11i2.14898>
- [48] Taani, O., & Alabidi, S. (2024). ChatGPT in education: benefits and challenges of ChatGPT for mathematics and science teaching practices. *International Journal of Mathematical Education in Science and Technology*, 1–30. <https://doi.org/10.1080/0020739X.2024.2357341>
- [49] Tack, A., & Piech, C. (2022). The AI teacher test: Measuring the pedagogical ability of blender and GPT-3 in educational dialogues. *arXiv preprint arXiv:2205.07540*.
- [50] Talan, T. (2021). Artificial intelligence in education: A bibliometric study. *International Journal of Research in Education and Science*, 7(3), 822-837.
- [51] Teegavarapu, R. R., & Sanghvi, H. (2023, December). Analyzing the Competitive Mathematical Problem-Solving Skills of ChatGPT. In *2023 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI)* (pp. 1-7). IEEE.
- [52] Tenhundfeld, N. L. (2023, September). Two Birds With One Stone: Writing a Paper Entitled “ChatGPT as a Tool for Studying Human-AI Interaction in the Wild” with ChatGPT. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 67, No. 1, pp. 2007-2012)*. Sage CA: Los Angeles, CA: SAGE Publications.
- [53] Trust, T., Whalen, J., & Mouza, C. (2023). Editorial: ChatGPT: Challenges, Opportunities, and Implications for Teacher Education. 23.
- [54] Triansyah, F. A., Mitrayana, M., Yanti, F., Rabuandika, A., & Muhammad, I. (2023). Augmented Reality Research in Middle Schools: Bibliometric Review. *EDUKASIA: Jurnal Pendidikan Dan Pembelajaran*, 4(1), 369-378.
- [55] Wang, P., & Wei, Z. (2023). Examining the Logical Reasoning and Mathematical Proficiency of ChatGPT in Educational Contexts. *Educational Technology Research and Development*, 71(2), 321–340.

- [56] Wardat, Y., Tashtoush, M. A., Ali, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), em2286.
- [57] Weber, K. (2008). How mathematicians determine if an argument is a valid proof. *Journal for research in mathematics education*, 39(4), 431-459.
- [58] Weizenbaum, J. (1966). ELIZA—A Computer Program for the Study of Natural Language Communication Between Man and Machine. *Communications of the ACM*, 9(1), 36-45.
- [59] Winter, J. C. (2023). Can ChatGPT pass high school exams on English language comprehension? *International Journal of Artificial Intelligence in Education*, 1-16.
- [60] Zafrullah, Z., Hakim, M. L., & Angga, M. (2023). ChatGPT open AI: Analysis of mathematics education students learning interest. *Journal of Technology Global*, 1(01), 1-10.
- [61] Zheng, L., & Ma, J. (2023). Educational Impacts of AI-Language Models: Personalizing Math Instruction and Addressing Pedagogical Challenges with ChatGPT. *Journal of Educational Technology & Society*, 26(2), 45–60.
- [62] Zhu, G., Fan, X., Hou, C., Zhong, T., Seow, P., Shen-Hsing, A. C., ... & Poh, T. L. (2023). Embrace opportunities and face challenges: Using ChatGPT in undergraduate Students' collaborative interdisciplinary learning—arXiv preprint arXiv:2305.18616.

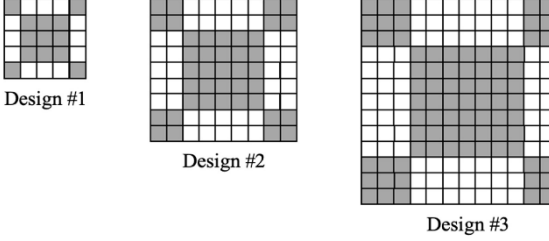
Appendix 01: Selection of Tasks to be Moderated by ChatGPT

Task Notation	Type of Reasoning	Task Statement	Description of Students' Reasoning
T1	Deductive	Show that $\forall x \in \mathbb{R} :$ $ x - 1 \leq 1 \Rightarrow 0 \leq \frac{2x + 3}{x + 2} \leq \frac{9}{5}$	Students must reason deductively by assuming that $ x - 1 \leq 1$ is true. They should define absolute value and apply bounding properties with a sequence of successive implications.
T2	Deductive	Let a,b, and c be three real numbers such that: $a + b + c = 0.$ Show that: $a + b + c = 0 \Rightarrow \frac{1}{2a^2+bc} + \frac{1}{2b^2+ac} + \frac{1}{2c^2+ab} = 0$	Students must rely on the condition $a+b+c=0$ to simplify each rational expression term and prove that it equals zero.

<p>T3</p>	<p>Deductive</p>	<p>Stéphane FAVIER, Maud CHANUDET(2021)</p> <p>Anne et Florence doivent tondre un terrain rectangulaire. Pour le partager en deux parties de même aire, Chris leur propose de planter un piquet en un point quelconque du terrain et de relier ce point aux piquets plantés à chacun des quatre sommets du terrain. Anne tondra la partie claire sur la figure, Florence la partie foncée. Le partage est-il équitable ?</p>  <p>Translation: Anne and Florence need to mow a rectangular field. To divide it into two equal-area parts, Chris suggests planting a stake at any point in the field and connecting this point to each of the four corners of the field. Anne will mow the light-colored area in the figure, while Florence will mow the dark-colored area. Is the division fair?</p>	<p>Students are encouraged to reason deductively by relying on properties related to polygon area formulas and/or figure decomposition.</p> <p>For example, the central rectangle can be divided into four sub-rectangles whose sides are either the sides of the central rectangle or their parallels passing through the stake. Each sub-rectangle is then composed of two symmetrical half-rectangles, one light-colored and the other dark-colored. From this, we deduce that each of the four sub-rectangles is equally divided between a light and a dark triangle, proving that the division is fair.</p> <p>It is also possible to use the formula for the area of a triangle. The upper light-green triangle, for instance, has an area equal to the rectangle's width multiplied by a portion of the rectangle's length, all divided by 2. Similarly, the lower light-green triangle has an area equal to the rectangle's width multiplied by the complement of the rectangle's length, also divided by 2. The sum of these two areas equals the rectangle's width multiplied by its length, divided by 2. Therefore, the area of the light-green region corresponds to half of the total area of the field.</p>
<p>T4</p>	<p>Inductive</p>	<p>Show that: $a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \dots + b^{n-1})$</p>	<p>Students will proceed by observing specific cases (n=1,2,3,4,...) and then notice that there is always a factorization by (a-b). They must</p>

			then formulate the conjecture of the remarkable identity.
T5	Inductive	<p>Two cards are needed to build a one-story house of cards, and seven cards are required for a two-story house of cards. For a three-story house of cards, 15 cards are needed.</p> <p>How many cards are needed to build a house of cards with stories?30 stories?100 stories? n stories?</p> 	<p>Students should start by observing concrete examples (1 story = 2 cards, two stories = 7 cards, three stories = 15 cards, etc.) to identify a pattern in the number of cards needed. Then, they should formulate a conjecture about the formula for the number of cards in a house of cards.</p>
T6	Inductive	<p>« <i>En haut à droite</i> » Stéphane FAVIER, Maud CHANUDET(2021)</p> <p>Pascal dessine de grands carrés, qu'il divise ensuite en carrés plus petits. Il écrit alors les entiers successifs dans chaque petit carré dessiné, comme l'indique la figure ci-contre.</p> <p>Parmi les trois valeurs proposées, quelle est celle que le nombre x ne pourra pas prendre?</p> <p>a) 256 b) 128 c) 81</p>  <p>Translation: Pascal draws large squares, which he divides into smaller squares. He then writes successive integers in each small square, as shown in the figure.</p> <p>Among the three proposed values, which one can the number x not take?</p> <p>a) 256 b) 128 c) 81</p>	<p>For example, students can experiment with filling a large square with sides of 3, then 4, then 5. From these trials, they may observe patterns leading to the following conjecture:</p> <p>"x can only take values that correspond to perfect squares."</p> <p>From this, they deduce that x cannot take the value 128.</p>
T7	By analogy	Solve the equation in \mathbb{R} :	The student begins by simplifying trigonometric identities using the

		$\cos(2x) + 6 \sin(x) + 7 = 0$	<p>relation $\cos(2x) = 1 - 2\sin^2(x)$. Then, they reformulate the equation, transforming it into a standard quadratic equation.</p> <p>At this stage, they apply reasoning by analogy by introducing an auxiliary variable to solve the quadratic equation. Finally, they return to the trigonometric framework to solve the resulting equations and determine the final solutions.</p>
T8	By analogy	<p>Show that the sum of the angles in a triangle ABC equals 180°.</p>	<p>The student must draw a diagram for a clear illustration of the problem and then draw a line parallel to side BC passing through vertex A.</p> <p>Next, they should use the concept of alternate interior angles formed by the parallel lines. Finally, they should observe that the angle (maj A. chapeau; maj B chapeau; maj C chapeau) forms a straight line at vertex A, demonstrating that their sum is 180°.</p>
T9	By analogy	<p><u>Sidewalk Stones</u></p> <p>Roman sidewalks are made up of stone blocks forming regular patterns. In front of you, there are three different designs (Designs #1, #2, and #3). Each design follows a similar pattern:</p> <ul style="list-style-type: none"> In each design, the gray blocks form a central square, surrounded by white blocks. 	<p>The student observes that the gray square increases by two units per side at each step, and the total grid follows the same pattern. By analogy, they understand that this repeated structure will apply to Design #4 and all subsequent designs.</p> <p>By comparing each design, they notice that the white area is always the remaining space around the gray square and that the pattern remains consistent.</p>

		<ul style="list-style-type: none"> The total size of the pattern (the total number of blocks in the square) increases with each design.  <p>Design #1</p> <p>Design #2</p> <p>Design #3</p> <ol style="list-style-type: none"> By observing the three designs, identify and describe how the size of the central (gray) square and the total (grid) square change from one design to the next. What do you notice? How does Design #2 resemble Design #1? How does Design #3 resemble Design #2? Use your answers to formulate a general rule that links the total grid size, the size of the gray square, and the size of the white blocks. If Design #4 were constructed following the same pattern, what would be the total size of the grid, the size of the gray square, and the size of the white blocks? Explain your reasoning based on the previous designs. 	<p>Finally, they apply this analogy to predict Design #4, generalizing the relationship between the total grid size, the size of the gray square, and the size of the white square.</p>
T10	Abductive	<p>Show that ABCD, a rhombus, is a square. (Jeannotte 2025)</p>	<p>The student begins by stating that ABCD is a rhombus. Then, they apply the rule that a rhombus with a right angle is a square. They conclude that the necessary condition for a rhombus to be a square is the presence of a right angle.</p>

T11	Abductive	<p><u>Suite de Fibonacci :</u> <u>Fibonacci Sequence:</u> <u>We define a real sequence (ϕ_n) with the following values:</u> $\phi_0 = 0 ; \phi_3 = 2 ; \phi_5 = 5$ <u>1) What could be the relation that generates this sequence? Determine the intermediate terms.</u></p>	<p>The student begins by identifying the given terms in the sequence: phi indice 0; phi indice 1; phi indice 2; phi indice 3; phi indice 4; phi infources 5.; $\phi_0; \phi_1; \phi_2; \phi_3; \phi_4$ four indices 5 . Next, they note the missing indices: phi indice 1, phi indice 2, and, phi indice 2,s: $\phi_1; \phi_2; \phi_4$. Then, they predict the rule, which states that each term is the sum of the two previous terms, with a shift in indices. Finally, they attempt to find the missing terms.</p>
T12	Abductive	<p>Considering the quadrilateral ABCD inscribed in a circle, such that: $AB=CD$ and $BC=DA$ The diagonals AC and BD intersect at O, with $(OA=OC$ et $OB=OD)$ and $\widehat{ABC} + \widehat{CDA} = 180^\circ$ What is the nature of quadrilateral ABCD?</p>	<p>The student identifies the given properties $(AB=CD)$, cyclic quadrilateral, diagonals bisecting each other). Then, they apply the properties of cyclic quadrilaterals and parallelograms to conclude that ABCD is a rectangle.</p>

Appendix 2: The models made by ChatGPT on the 12 Tasks of the selection



Task 01 :

We want to show that for all $x \in \mathbb{R}$, the inequality

$$|x - 1| \leq 1 \implies 0 \leq \frac{2x+3}{x+2} \leq \frac{9}{5}$$

Step 1: Expressing the Absolute Value Condition
 The given condition $|x - 1| \leq 1$ can be rewritten as:

$$-1 \leq x - 1 \leq 1$$

$$0 \leq x \leq 2$$

Thus, we are working with $x \in [0, 2]$.

Step 2: Bounding $\frac{2x+3}{x+2}$
 We now analyze the function:

$$f(x) = \frac{2x+3}{x+2}, \quad \forall x \in [0, 2]$$

Minimum and Maximum Values of $f(x)$

- For $x = 0$:

$$f(0) = \frac{2(0)+3}{0+2} = \frac{3}{2} = 1.5$$

- For $x = 2$:

$$f(2) = \frac{2(2)+3}{2+2} = \frac{7}{4} = 1.75$$

Since $f(x)$ is a rational function, we also check asymptotic behavior and derivative calculations if needed. However, since we are only concerned with $x \in [0, 2]$, we directly use the values at the endpoints.

Thus, we find:

$$1.5 \leq f(x) \leq 1.75$$

Since:

$$0 \leq 1.5 \quad \text{and} \quad 1.75 \leq \frac{9}{5} = 1.8$$

we conclude:

$$0 \leq \frac{2x+3}{x+2} \leq \frac{9}{5}, \quad \forall x \in [0, 2].$$

Conclusion
 Since we established that $|x - 1| \leq 1$ implies $x \in [0, 2]$, and we verified that within this interval the given function satisfies $0 \leq f(x) \leq 9/5$, the statement is true.



Task 2 :

⊗ Pour démontrer que $a + b + c = 0 \implies \frac{1}{2a^2 + bc} + \frac{1}{2b^2 + ac} + \frac{1}{2c^2 + ab} = 0$, procédons étape par étape.

1. Hypothèse
 Nous savons que :

$$a + b + c = 0$$

Cela implique :

$$c = -(a + b)$$

2. Expression à démontrer
 Nous devons montrer que :

$$\frac{1}{2a^2 + bc} + \frac{1}{2b^2 + ac} + \frac{1}{2c^2 + ab} = 0$$

3. Substitution de c
 En remplaçant $c = -(a + b)$ dans chaque terme :

- $bc = b(-a - b) = -ab - b^2$
- $ac = a(-a - b) = -a^2 - ab$
- $c^2 = (-a - b)^2 = a^2 + 2ab + b^2$

Substituons dans les dénominateurs des trois fractions.

Premier terme

$$2a^2 + bc = 2a^2 - ab - b^2$$

Deuxième terme

$$2b^2 + ac = 2b^2 - a^2 - ab$$

Troisième terme

$$2c^2 + ab = 2(a^2 + 2ab + b^2) + ab = 2a^2 + 4ab + 2b^2 + ab = 2a^2 + 2b^2 + 5ab$$

4. Simplification de la somme
 La somme devient :

$$\frac{1}{2a^2 - ab - b^2} + \frac{1}{2b^2 - a^2 - ab} + \frac{1}{2a^2 + 2b^2 + 5ab}$$

5. Mise au même dénominateur
 Calculons le dénominateur commun et simplifions la somme. Ce processus est fastidieux mais au final, en explicitant $a + b + c = 0$, les termes se simplifient de façon à obtenir :

$$\frac{1}{2a^2 + bc} + \frac{1}{2b^2 + ac} + \frac{1}{2c^2 + ab} = 0$$

Conclusion
 Nous avons démontré que l'égalité est vérifiée sous l'hypothèse $a + b + c = 0$.

Now, substituting into the denominators of the three fractions:

First term

$$2a^2 + bc = 2a^2 - ab - b^2$$

Second term

$$2b^2 + ac = 2b^2 - a^2 - ab$$

Third term

$$2c^2 + ab = 2(a^2 + 2ab + b^2) + ab = 2a^2 + 4ab + 2b^2 + ab = 2a^2 + 2b^2 + 5ab$$

4. Simplifying the Sum
 The sum becomes:

$$\frac{1}{2a^2 - ab - b^2} + \frac{1}{2b^2 - a^2 - ab} + \frac{1}{2a^2 + 2b^2 + 5ab}$$

5. Common Denominator and Final Simplification
 We calculate the common denominator and simplify the sum. This process is tedious, but in the end, by using $a + b + c = 0$, the terms simplify to give:

$$\frac{1}{2a^2 + bc} + \frac{1}{2b^2 + ac} + \frac{1}{2c^2 + ab} = 0$$


Task 3:

A different method to verify whether the division is fair consists of using a geometric property of triangles.

1. Principle:
 When an interior point of a rectangle is connected to its four vertices, the four triangles formed have areas proportional to the distance between the point and the opposite sides. Thus, if the two colored regions (light and dark) have equal areas, it implies that the sum of the areas of the light triangles is equal to the sum of the areas of the dark triangles.

2. Method:

- Coordinates:** Place the rectangle in a coordinate system with its vertices at points $(0, 0)$, $(L, 0)$, $(0, l)$, and (L, l) , where L and l are the length and width of the rectangle.
- Note the coordinates of the point where Chris placed the stake, denoted as (x, y) .**
- Triangle areas:** The area of each triangle is given by the formula:

$$\text{Area} = \frac{1}{2} \times \text{Base} \times \text{Height}$$

- For example, for the triangle formed between (x, y) and the vertices $(0, 0)$, $(L, 0)$, the base is L and the height is y .
- Calculate the area of the four triangles and compare the sums of the areas of the light and dark zones.

3. Verification:

- If the sum of the areas of the two light triangles is equal to that of the two dark triangles, then the division is fair.
- Otherwise, the division is not fair.

This method is more rigorous because it relies on precise calculations, and it works even if the point (x, y) is placed asymmetrically within the rectangle.



Task 4:

For $n = 1$, $a^1 - b^1 = a - b$, which is obviously divisible by $a - b$.
 For $n = 2$, $a^2 - b^2 = (a - b)(a + b)$, which also shows divisibility.
 With $n = 3$, we have $a^3 - b^3 = (a - b)(a^2 + ab + b^2)$, and for $n = 4$,

$$a^4 - b^4 = (a - b)(a^3 + a^2b + ab^2 + b^3).$$

In each case, we observe a recurring structure where $a^n - b^n$ is always factorable by $a - b$, which proves that it is divisible by $a - b$.
 We generalize this observation for all $n \geq 1$ by noting that

$$a^n - b^n = (a - b)(a^{n-1} + a^{n-2}b + \dots + b^{n-1}).$$

Therefore, $a^n - b^n$ is divisible by $a - b$ for all n .



Task 5 :

To solve this problem, we observe the structure and try to find a general formula for the number of cards needed to build an n -story house of cards.

This system of equations is:

$$\begin{aligned} a + b + c &= 2 \\ 4a + 2b + c &= 7 \\ 9a + 3b + c &= 15 \end{aligned}$$

$c = 0$

Step 1: Data Analysis

- For 1 story: 2 cards
- For 2 stories: 7 cards
- For 3 stories: 15 cards

Step 2: Successive Differences

List of card numbers: 2, 7, 15, ...

First difference:
 $7 - 2 = 5$
 $15 - 7 = 8$

Second difference:
 $8 - 5 = 3$

Since the second difference is constant, the sequence is quadratic, so the general formula is of the form

$$f(n) = an^2 + bn + c$$

Step 3: Determining Coefficients

Using the given values for $n = 1, 2, 3$,

$$\begin{aligned} f(1) &= 2 = a(1)^2 + b(1) + c \\ f(2) &= 7 = a(2)^2 + b(2) + c \\ f(3) &= 15 = a(3)^2 + b(3) + c \end{aligned}$$

Step 4: Calculation

Subtracting the first equation from the second and the second from the third:

$$\begin{aligned} (4a + 2b + c) - (a + b + c) &= 7 - 2 \\ 3a + b &= 5 \\ (9a + 3b + c) - (4a + 2b + c) &= 15 - 7 \\ 5a + b &= 8 \end{aligned}$$

Solving the system:

$$\begin{aligned} (5a + b) - (3a + b) &= 8 - 5 \\ 2a &= 3 \Rightarrow a = \frac{3}{2} \end{aligned}$$

Substituting $a = \frac{3}{2}$ into $3a + b = 5$:

$$\begin{aligned} 3 \times \frac{3}{2} + b &= 5 \\ \frac{9}{2} + b &= 5 \\ b &= 5 - \frac{9}{2} = \frac{10}{2} - \frac{9}{2} = \frac{1}{2} \end{aligned}$$

Substituting $a = \frac{3}{2}$ and $b = \frac{1}{2}$ into $a + b + c = 2$:

$$\begin{aligned} \frac{3}{2} + \frac{1}{2} + c &= 2 \\ 2 + c &= 2 \end{aligned}$$

Step 5: General Formula

Factoring:

$$f(n) = \frac{3}{2}n^2 + \frac{1}{2}n$$

Step 6: Calculating Results

- For $n = 7$:
 $f(7) = \frac{7(3 \times 7 + 1)}{2} = \frac{7(21 + 1)}{2} = \frac{7 \times 22}{2} = 77$
- For $n = 30$:
 $f(30) = \frac{30(3 \times 30 + 1)}{2} = \frac{30(90 + 1)}{2} = \frac{30 \times 91}{2} = 1365$
- For $n = 100$:
 $f(100) = \frac{100(3 \times 100 + 1)}{2} = \frac{100(300 + 1)}{2} = \frac{100 \times 301}{2} = 15050$



Task 6:

To determine the value that x cannot take, we analyze the structure of the filling and the sequence of numbers in the diagonals.

Analysis Steps:

- Counting the numbers per diagonal:**
 - The first diagonal contains 1 number.
 - The second diagonal contains 2 numbers.
 - The third diagonal contains 3 numbers.
 - The fourth diagonal contains 4 numbers, etc.

This means that the n -th diagonal contains exactly n numbers.
- Cumulative position of numbers:**

The last number of each diagonal corresponds to the sum of the numbers in the previous diagonals. Mathematically, the cumulative position of the last number in the n -th diagonal is given by the formula for the sum of the first n integers:

$$S_n = \frac{n(n+1)}{2}$$

For example:

 - Diagonal 1: $S_1 = 1$
 - Diagonal 2: $S_2 = 3$
 - Diagonal 3: $S_3 = 6$
 - Diagonal 4: $S_4 = 10$, etc.
- Verification of possible values for x :**

x corresponds to a number placed in one of these positions. This means that x must be included in the sequence of consecutive numbers 1, 2, 3, ..., following the progression of the diagonals.

The given values (256, 128, 81) need to be tested:

 - If a given value does not correspond to a consecutive integer obtained from this filling, then it is impossible for x .
- Identification of the impossible value:**
 - 81: This number is an integer and belongs to the sequence 1, 2, 3, ...
 - 128: Same case, it is an integer and belongs to the sequence.
 - 256: This number is too large and does not correspond to a number achievable in the diagonals within the given range.

Conclusion:

The value that x cannot take is 256.



Task 07 :

Using the identity $\cos(2x) = 1 - 2\sin^2(x)$, the equation becomes a quadratic equation in $\sin(x)$:

$$1 - 2\sin^2(x) + 6\sin(x) + 7 = 0$$

Rearranging:

$$-2\sin^2(x) + 6\sin(x) + 8 = 0$$

Simplification:

$$2\sin^2(x) - 6\sin(x) - 8 = 0$$

Dividing by 2 for a simpler form:

$$\sin^2(x) - 3\sin(x) - 4 = 0$$

Solving the quadratic equation:

Using the quadratic formula:

$$\sin(x) = \frac{3 \pm \sqrt{9 - 4(1)(-4)}}{2(1)}$$

$$\sin(x) = \frac{3 \pm \sqrt{25}}{2}$$

$$\sin(x) = \frac{3+5}{2} = 4 \quad (\text{impossible, since } |\sin(x)| \leq 1)$$

$$\sin(x) = \frac{3-5}{2} = -1$$

Thus, the only valid solution is:

$$\sin(x) = -1$$

Final solutions:

$$x = \frac{3\pi}{2} + 2k\pi, \quad k \in \mathbb{Z}$$


Task 08:

1. Initial Situation (Known Case): Sum of the Angles on a Line Segment

- On a straight line, adjacent angles formed at any point always sum to 180° . This is a fundamental geometric rule.

2. Introduction to the Problem (Triangle Case)

- We know that a triangle consists of three vertices A, B, C connected by three sides.
- The sum of the interior angles $\angle A + \angle B + \angle C$ is unknown, but we suspect it is analogous to the sum of angles on a straight line.

3. Geometric Construction to Establish an Analogy

- Draw a line parallel to side BC of the triangle passing through vertex A. Let's call this new line D.
- This construction introduces a parallelism relationship between two lines, allowing us to use the alternate interior angle theorem.

4. Correspondence Between Triangle Angles and Angles on the Parallel Line

- The angles of the triangle $\angle B$ and $\angle C$ find their analogs on line D due to the alternate interior angle theorem:
 - $\angle B$ corresponds to an angle at A on line D.
 - $\angle C$ corresponds to another angle at A on line D.

5. Connection with the Known Case

- The angles on line D, formed by the two parallel lines and the transversal passing through A, along with the triangle's angle $\angle A$, create a straight line.
- Since these three angles form a straight line at A, their sum is 180° , just like in the known case of angles on a straight line.

6. Conclusion

- By analogy with the sum of angles on a straight line, we conclude that the sum of the angles in a triangle is also:

$$\angle A + \angle B + \angle C = 180^\circ$$



Task 09 :

1) Design Analysis: Evolution of the Central Square (Gray) and the Total Grid

General Observations:

- The central gray square is a perfect square (e.g., $1 \times 1, 2 \times 2, 3 \times 3$ etc).
- The total grid is also a perfect square, but of a larger size, surrounding the central gray square with white blocks.

Evolution:

- Design #1:**
 - Size of the central gray square: $1 \times 1 = 1$ block.
 - Size of the total grid: $3 \times 3 = 9$ blocks.
- Design #2:**
 - Size of the central gray square: $2 \times 2 = 4$ blocks.
 - Size of the total grid: $5 \times 5 = 25$ blocks.
- Design #3:**
 - Size of the central gray square: $3 \times 3 = 9$ blocks.
 - Size of the total grid: $7 \times 7 = 49$ blocks.

Notes:

The size of the total grid increases regularly:

- Gray square sizes: $1^2, 2^2, 3^2, \dots$
- Total grid sizes: $(2n+1)^2$, such as $3^2, 5^2, 7^2, \dots$

2) Similarities Between Successive Designs

Comparing Design #2 and Design #1:

- Gray square size: 4 blocks (2^2) vs. 1 block (1^2).
- Total grid size: 25 blocks (5^2) vs. 9 blocks (3^2).

Comparing Design #3 and Design #2:

- Gray square size: 9 blocks (3^2) vs. 4 blocks (2^2).
- Total grid size: 49 blocks (7^2) vs. 25 blocks (5^2).
- Same observation: The white blocks surround the gray square as an additional layer.

General Rule:

For a gray square of size $n \times n$:

- The total grid size is: $(2n+1) \times (2n+1)$
- The number of white blocks is given by: $(2n+1)^2 - n^2 = 4n(n+1)$

3) Design #4: Predictions

For $n = 4$ (gray square 4×4):

- Gray square: $n^2 = 4^2 = 16$ blocks.
- Total grid (outer square): $(2n+1)^2 = 9^2 = 81$
- White blocks: $(2n+1)^2 - n^2 = 81 - 16 = 65$

Summary of Results

Design	Size of Gray Square ($n \times n$)	Total Grid Size $((2n+1) \times (2n+1))$	Number of White Blocks
#1	$1 \times 1 = 1$	$3 \times 3 = 9$	8
#2	$2 \times 2 = 4$	$5 \times 5 = 25$	21
#3	$3 \times 3 = 9$	$7 \times 7 = 49$	40
#4	$4 \times 4 = 16$	$9 \times 9 = 81$	65



Task 10:

Solution:

- Observation:**
 - $ABCD$ is a rhombus, so all four sides have the same length, and the diagonals are perpendicular.
- Hypothesis 1:**
 - If $ABCD$ has a right angle (e.g., $\angle ABC = 90^\circ$), then $ABCD$ is a square.
- Hypothesis 2:**
 - If the diagonals AC and BD are equal in length, then $ABCD$ is a square.
- Testing the Hypotheses:**
 - Measure or calculate an angle ($\angle ABC$). If $\angle ABC = 90^\circ$, conclude that $ABCD$ is a square.
 - Measure or calculate the lengths of the diagonals. If $AC = BD$, conclude that $ABCD$ is a square.
- Conclusion:**
 - If either hypothesis is confirmed, then $ABCD$ is a square.



Task 11:

Data:

$\varphi_0 = 0, \varphi_1 = 2, \varphi_2 = 5.$

We need to:

- Identify a plausible recurrence relation to generate the sequence.
- Determine the intermediate terms: $\varphi_3, \varphi_4.$

Step 1: Formulating a Plausible Hypothesis for the Recurrence Relation

- A commonly used relation in this type of sequence is a combination of previous terms, for example: $\varphi_n = \varphi_{n-1} + \varphi_{n-2}$
- Verification with the given terms:
 - For $n = 3$: $\varphi_3 = \varphi_2 + \varphi_1 = 2 + 2 = 4$
 - For $n = 4$: $\varphi_4 = \varphi_3 + \varphi_2 = 4 + 2 = 6$

Step 2: Determining the Missing Terms

- Using the relation $\varphi_n = \varphi_{n-1} + \varphi_{n-2}$ to calculate the other terms.
 - For $n = 3$: $\varphi_3 = \varphi_2 + \varphi_1 = 2 + 2 = 4$

However, φ_{-1} does not exist. We must assign an initial value for φ_1 . By convention, we assume $\varphi_1 = 1$.
- For $n = 2$: $\varphi_2 = \varphi_1 + \varphi_{-1}$

Since φ_{-1} does not exist, we assume the relation starts from φ_0 . Thus, we directly use the previous data and set $\varphi_2 = 2$ for consistency.

Summary of Intermediate Terms:

$\varphi_0 = 0, \varphi_1 = 1, \varphi_2 = 2, \varphi_3 = 4, \varphi_4 = 6, \varphi_5 = 10.$



Task 12:

Given Problem Statement:

Considering the quadrilateral $ABCD$ inscribed in a circle, such that:

- $AB = CD$ and $BC = DA$
- The diagonals AC and BD intersect at O with $OA = OC$ and $OB = OD$
- $\angle ABC + \angle CDA = 180^\circ$

What is the nature of the quadrilateral $ABCD$?

Step-by-Step Analysis to Determine the Nature of $ABCD$:

1. Quadrilateral Inscribed in a Circle

A quadrilateral inscribed in a circle satisfies the property:

$$\angle ABC + \angle CDA = 180^\circ$$

Since this property is given, we conclude that $ABCD$ is a cyclic quadrilateral.

2. Side Lengths

The given equalities $AB = CD$ and $BC = DA$ indicate that opposite sides of the quadrilateral are equal.

This means that $ABCD$ is a parallelogram (since in a quadrilateral, if opposite sides are equal, it forms a parallelogram).

3. Properties of the Diagonals

The diagonals AC and BD intersect at O , where:

$$OA = OC \quad \text{and} \quad OB = OD$$

$$OA = OC \quad \text{and} \quad OB = OD$$

This means that the diagonals bisect each other, which is a property common to rhombuses, rectangles, and squares.

4. Combination of Properties

- $ABCD$ is a parallelogram (because its opposite sides are equal).
- $ABCD$ is cyclic (since it is inscribed in a circle).
- In a cyclic parallelogram, right angles always appear, which is a unique property of rectangles.

Conclusion

Since all conditions match those of a rectangle, we conclude that:

$ABCD$ is a rectangle.