

Document Version

Final published version

Citation (APA)

Rothkrantz, C., Fitrianie, S., & Rothkrantz, L. (2025). Intelligent Mathematical Tutor Based on ChatGPT or DeepSeek. In T. Vassilev, & V. Markova (Eds.), *Proceedings of the 2025 International Conference on Computer Systems and Technologies (CompSysTech)* IEEE. <https://doi.org/10.1109/CompSysTech65493.2025.11136856>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

**Green Open Access added to [TU Delft Institutional Repository](#)
as part of the Taverne amendment.**

More information about this copyright law amendment
can be found at <https://www.openaccess.nl>.

Otherwise as indicated in the copyright section:
the publisher is the copyright holder of this work and the
author uses the Dutch legislation to make this work public.

Intelligent Mathematical Tutor based on ChatGPT or DeepSeek

Cyril Rothkrantz

Mathematics

Veurs Lyceum

Leidschendam, the Netherlands

c.j.l.rothkrantz@gmail.com

Siska Fitrianie

Dept. of Intelligent Systems

Delft University of Technology

Delft, the Netherlands

s.fitrianie@tudelft.nl

Leon Rothkrantz

Faculty of Transportation Science

Czech Technical University of Prague

Prague, Czech Republic

l.j.m.rothkrantz@tudelft.nl

Abstract—Many secondary school students in the Netherlands require additional support to prepare for their mathematics exams. This paper presents the design of a Massive Open Online Course (MOOC) featuring training materials tailored for the mathematics exams, integrated with a digital intelligent tutor. The tutor, which students can activate as needed, is embedded within the course content. Two versions of the tutor were developed: one powered by ChatGPT and the other by DeepSeek. To evaluate their effectiveness, both versions were tested on a set of problems from W4Kangoeroe, an annual mathematics competition for primary and secondary school students. The results indicate that while both chatbots are capable of generating correct mathematical solutions, their approaches often differ significantly from those provided by experienced mathematics teachers. This highlights the potential of AI-driven tutors to assist students but also underscores the need for further refinement to align more closely with educational standards and teaching methodologies.

Keywords— MOOC, intelligent tutor, ChatGPT, DeepSeek, W4Kangoeroe

I. INTRODUCTION

At many secondary schools in the Netherlands, approximately half of the students require individual support to fully grasp all topics covered in their mathematics exams. While teachers typically deliver lessons to the entire class and offer additional exam preparation sessions, they often lack the time to provide personalized assistance. Although specialized tutors or institutes are available, their services come at a cost [8]. To address this gap, we designed a specialized Massive Open Online Course (MOOC) focused on mathematics, enabling students to access a digital tutor for help with specific assignments. However, creating a tutor capable of addressing all potential student challenges is no simple task. With the recent release of tools like ChatGPT and DeepSeek, which claim to solve school-level mathematical problems, we decided to test their capabilities.

The MOOC was implemented using Moodle, which features an OpenAI ChatBlock [https://moodle.org/plugins/block_openai_chat]. This block allows users to access 24/7 chat support via OpenAI's GPT or DeepSeek. Students can activate the digital tutor with a single click to seek help or verify solutions. The course includes a comprehensive set of problems from Kangoeroe [<https://www.w4kangoeroe.nl/>], an annual international mathematics contest for primary and secondary school students. Students can request support from either the ChatGPT or DeepSeek tutor to solve these problems or check their solutions.

In our experiment, we analysed the tutors' solutions to a selected set of problems and compared them with solutions provided by teachers, including two mathematician authors of this paper. One solution is discussed in detail to illustrate how

the chatbots approach mathematical problems and how their methods differ from those of teachers. While the chatbots generate correct solutions, they tend to be more formal and lengthy compared to the concise, teacher-provided solutions, which serve as annotations to the chatbots' responses.

The MOOC incorporates assignments from W4Kangoeroe, categorized into six competition levels: wizFUN, wizKID, wizSMART, wizBRAIN, wizPROF, and wizEXPERT. Each level includes problems, solutions, and teacher solutions available on the W4Kangoeroe website. Our project aims to transform this content into a structured learning resource within the MOOC. Students can access these assignments and activate ChatGPT or DeepSeek for assistance. To guide students effectively, we developed a didactic approach based on the teaching methods used at TU Delft. In this approach, teachers or assistants highlight key aspects of the chatbots' solutions, such as data presentation, algorithms, and computations, mirroring the concise and practical nature of teacher-provided solutions. This digital adaptation replicates the traditional classroom support system, where teachers or assistants circulate to help students individually or in groups, ensuring a seamless learning experience.

The objectives of this paper are as follows:

1. To investigate whether it is feasible to develop a digital tutor for solving mathematical assignments from W4Kangoeroe, tailored for secondary school students, using ChatGPT and DeepSeek.
2. To create a functional prototype of the digital tutor as outlined in the first objective.
3. To ensure that the solutions generated by the digital tutor are clear, easy to understand, and immediately usable for students.
4. To compare the solutions produced by the ChatGPT and DeepSeek-based tutors with those provided by human mathematics teachers.
5. To integrate these digital tutors, powered by ChatGPT and DeepSeek, into a specialized MOOC for practical implementation.

II. RELATED WORK

ChatGPT and DeepSeek are both advanced AI chatbot technologies. ChatGPT, developed by OpenAI, was initially designed for conversational interactions and offered free access, though it now operates on a paid subscription model. In contrast, DeepSeek is an open-source tool created by a Chinese company, specializing in large language models. Preliminary observations suggest that DeepSeek may outperform ChatGPT in mathematics and coding tasks, though no formal comparative studies have yet been

conducted. In this paper, we investigate this potential by developing a mathematical tutor using both ChatGPT and DeepSeek and evaluating their performance on a series of mathematical assignments.

Jančařík (2003) [1] conducted a study validating the use of chatbots for mathematics tutoring, focusing on how students interact with AI when solving problems independently, without external motivation. The research prioritized interaction processes over problem generation, and the authors developed a custom chatbot capable of basic user dialogues. Meanwhile, Phung (2023) [2] compared GPT-3.5 and GPT-4 on Python programming tasks, finding that GPT-4 outperformed GPT-3.5 and achieved results comparable to those of human users. The study also proposed future directions for enhancing GPT-4's applications. Additionally, research on AI in online learning platforms [3] compared traditional MOOCs with Korbit, a platform offering personalized and active learning experiences. Korbit demonstrated superior learning outcomes, and our model integrates several of its key features.

Yunianto et al. (2024) [4] explored ChatGPT's ability to solve geometric problems, noting that earlier versions required manual parameter inputs, whereas the latest version can process images, identify objects, and provide feedback. Oonk's PhD thesis (2009) [5] examined the integration of theory and practice in mathematics teacher education, emphasizing the importance of a structured learning environment. Freudenthal (1978) [6] advocated for applied mathematics in education, an approach we incorporated by using real-life assignments in our experiments. At TU Delft, mathematics courses combine lectures and workgroups, a method known as *colstructie*, which has evolved since 2000 with new instructional techniques [7].

Rothkrantz (2015) [7] introduced an inquiry-based didactic model, demonstrating that encouraging students to think aloud and ask questions enhances their understanding. The Freudenthal Institute (IOWO) [8] focused on realistic mathematics education, developing and testing teaching methods in collaboration with schools. Over time, their research shifted toward didactics and digital mathematics education. Sofroniou and Poutas (2016) [9] evaluated group work in tutorials, identifying its benefits and limitations. These findings are relevant to our study.

Rothkrantz (2017a) [10] developed new didactic models for mathematics teaching, highlighting the role of emotions in academic success. The study found that students who enjoyed their learning environment performed better, and these models were tested using educational games. In 2003, a modern version of Eliza was created, incorporating speech recognition and facial expression analysis to detect emotional states [11]. Eliza employs non-directive dialogue, inspired by Rogerian therapy, to encourage students to elaborate on their problems. Finally, Moodle introduced 24/7 chat support via ChatGPT and the Assistant API [12], a feature we implemented in our Moodle application, as discussed in Section III.

III. THE USE OF CHATBOXES IN MATHEMATICS AT SECONDARY SCHOOLS IN THE NETHERLANDS

The Dutch Association of Mathematics Teachers plays a central role in all teaching activities, including the development of new curricula and innovative didactic approaches. In recent years, chatbots have become a prominent topic of discussion at meetings, in newsletters, and

in EUCLIDES, a bi-monthly magazine published by the Association. ChatGPT, in particular, has been a frequent subject of debate, especially regarding its impact on homework, exams, and student learning outcomes.

The Freudenthal Institute [8], a leading research centre for mathematics education and didactics for primary and secondary schools, has also engaged in discussions on the use of chatbots. Researchers at the Institute have explored ways to integrate chatbots into regular teaching and learning practices in mathematics. However, their focus has primarily been on language-related educational activities rather than mathematics-specific applications.

Despite the growing interest, extensive scientific studies on the use of chatbots in education remain limited. Most discussions occur on web forums, websites, and similar platforms. For instance, Alex McFarland recently published an article on Unite.AI, highlighting the best AI-powered math tools. He argues that these tools are transforming the landscape of mathematical education by empowering students and educators to address complex challenges. Below, we list seven of the most popular tools currently used by students:

1. *Julius*: Julius AI [<https://julius.ai/home/ai-math/>] is an advanced math tutor designed to assist students and professionals in solving mathematical problems with ease.
2. *Photomath*: This widely used mobile app [<https://photomath.com>] employs computer vision and artificial intelligence to deliver instant solutions to math problems.
3. *MathGPTPro*: A state-of-the-art AI-powered app, MathGPTPro [<https://www.mathgptpro.com/>] supports both homework assistance and broader mathematical education through teaching and learning activities.
4. *Mathway*: A robust math problem solver, Mathway [<https://www.mathway.com/Algebra>] offers instant solutions to a wide range of problems, from basic arithmetic to advanced calculus, including problem-solving and logical reasoning.
5. *Symbolab*: An advanced platform, Symbolab [<https://www.symbolab.com>] uses AI to provide step-by-step solutions to diverse mathematical problems, enhancing mathematics education.
6. *GeoGebra*: An interactive software suite for learning and teaching mathematics and science, GeoGebra [<https://www.geogebra.org>] focuses on geometric problems and is suitable for students from primary school to university level.
7. *MathPapa*: An online and mobile algebra calculator and AI educational application, MathPapa [<https://www.mathpapa.com>]/<https://www.mathpapa.com>] is designed to help students learn and practice algebraic concepts.

Recently, EUCLIDES conducted a survey among mathematics teachers regarding the use of mathematical apps by students [13]. One key finding was the rapid increase in the adoption of such apps by students. The survey revealed that students prioritize apps that are easy to use and free of charge, often overlooking the quality, mathematical correctness, and rigor of the tools. In the near future, we plan to conduct a

large-scale survey to evaluate the use of these and similar tools by secondary school students, aiming to provide deeper insights into their effectiveness and impact on learning outcomes.

IV. MODEL AND ARCHITECTURE OF A TUTOR

Moodle currently offers a ChatGPT plugin that enables seamless integration of generative AI capabilities into e-learning platforms. Once installed, the plugin can be activated and configured within specific courses or activities (Fig. 1). In this study, we integrated the ChatGPT plugin (version 3.0.1-2025021700) with the GPT-4 model into a Moodle-based e-learning application to enhance the functionality of an AI-driven tutor.

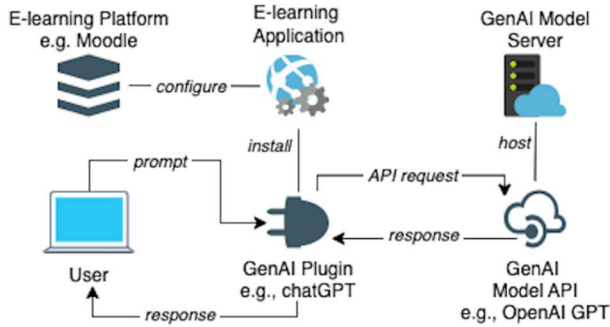


Fig. 1. Building a GenAI plugin on an e-learning application

An alternative approach to integrating generative AI capabilities into an e-learning platform involves deploying a large language model (LLM) on a dedicated server (Fig. 2). In this study, we implemented this method by incorporating DeepSeek into a Moodle-based e-learning environment. The deployment process commenced with the configuration of a server equipped with sufficient computational resources to accommodate the high processing demands of the LLM. Subsequently, all necessary dependencies and supplementary libraries required for model execution were installed. Once the environment was properly configured, the DeepSeek LLM (version deepseek-r1 8b) was retrieved and deployed using the Ollama framework. The integration with the e-learning platform was facilitated by exposing the LLM through an API, which was designed to process requests from the application. This API enabled real-time interactions.

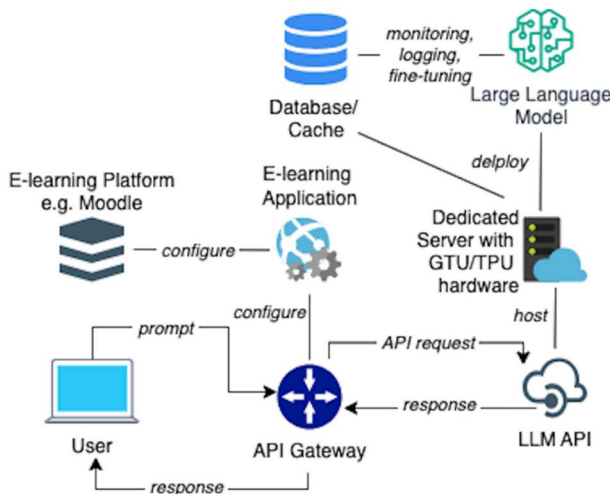


Fig. 2. Deploying an LLM on a dedicated server to support a GenAI feature on an e-learning application

Within both environments, both students and instructors can access the chatbots directly from the course interface. The chatbots serve multiple functions, including generating responses to student inquiries and assisting instructors in producing supplementary learning materials. By leveraging advanced natural language processing capabilities, such features in an e-learning application can facilitate context-aware interactions, making it a valuable tool for both synchronous and asynchronous learning.

V. DIDACTIC MODEL

For many years, mathematics education has been conducted using a parallel teaching approach. Theoretical lectures are delivered alongside practical sessions, where students apply the theory to assignments, either individually or in groups. During these practical sessions, more experienced or academically stronger students often take the lead in solving problems and explaining solutions to their peers. If the group as a whole struggles to solve a problem, an instructor intervenes by providing hints or presenting a complete solution. This approach has proven effective, as students learn to solve similar problems through a process of imitation and guided practice.

The didactic model of combining theory with practical sessions has been extensively researched and discussed in the PhD thesis of Oonk [5]. Supporting evidence for this approach can also be found in the work of Freudenthal [6]. Their experiments demonstrated that only a small subset of highly gifted and motivated students can solve mathematical problems independently. Most students require additional support and practice to achieve proficiency. Over time, however, students become capable of solving similar problems, which enhances their motivation and self-confidence. In our Massive Open Online Course (MOOC), we adopted a similar didactic approach. Students using the MOOC can practice assignments and request tutor support anonymously. Through repeated practice, they develop a stronger sense of mastery in solving the presented problems.

In our MOOC, assignments are initially presented to students without assistance. If a student is unable to solve a problem, they can request help by activating either ChatGPT or DeepSeek. Students are then encouraged to study the solutions provided by the chatbots and summarize them in a way they would explain to a fellow student. If a student still struggles, he can request further assistance from the tutor, which provides the teacher's solution to the problem. The student is then asked to compare the teacher's solution with those generated by the chatbots. By repeating this process across multiple assignments, students are expected to gradually improve their problem-solving abilities. If challenges persist, students can seek additional help from a peer or instructor via email. This structured approach ensures that students receive the necessary support while fostering independent learning and critical thinking skills.

VI. EXPERIMENT WITH CHATGPT AND DEEPSEEK

In Section IV, we introduced two digital tutors designed for mathematics education at secondary schools. Our goal is to determine which of the two performs better. However, at this stage, only preliminary studies are available to evaluate the performance of ChatGPT and DeepSeek. Initial findings suggest that ChatGPT excels in text analysis, while DeepSeek shows stronger capabilities in mathematics and coding. It is important to note that the design and coding of both chatbots

are not open-source. To objectively compare their performance, we propose evaluating them using a benchmark. For this purpose, we selected W4Kangoeroe, a collection of mathematical problems tailored for primary and secondary school students. In Section A, we provide an overview of W4Kangoeroe, and in Section B, we analyze the performance of the chatbots based on this benchmark.

A. W4Kangoeroe

In the Netherlands, there is a renowned research institute dedicated to the development and didactics of mathematics, named after the prominent mathematician Freudenthal [8]. In 2005, members of the Institute, along with mathematics teachers, proposed the idea of organizing an annual mathematics contest for students ranging from group 3 in primary schools to level 6 in secondary schools. The contest features assignments presented as multiple-choice questions, each offering five possible solutions, including the correct one.

The W4Kangoeroe contest started in 2005. Last year, the contest attracted 140,000 participants from 230 schools (Fig. 3). The exams are structured into six levels of increasing complexity: wizFUN, wizKID, wizSMART, wizBRAIN, wizPROF, and wizEXPERT. Each contest is composed of 25 assignments and for most assignments the answers and proofs are presented. To help students prepare, extensive sets of practice exercises are available. As a proof of concept, we have implemented some of these exercises into our system. With the necessary permissions, we plan to incorporate all the exercises into our MOOC to further support student learning and preparation.

Worldwide Mathematical Game W4 Kangoeroe



Fig. 3. Sticker when a school participates in the Kangoeroe game ([Home :: W4Kangoeroe](#))

B. Application of ChatGPT and DeepSeek on a Testset

The wizEXPERT category is designed for the highest levels of secondary school (equivalent to university level) and was introduced in 2022. For our experiment, we analyzed assignments from the years 2022 to 2025. In this paper, we focus on examples from the wizEXPERT category, specifically selecting assignments number 15 and 20 from the 2024 contest. In addition to presenting the problems, we also provide the proofs. The solutions were generated by mathematics teachers, as well as by ChatGPT and DeepSeek. The results from ChatGPT and DeepSeek are displayed in two columns for comparison. In Section C, we provide a detailed discussion and analysis of these results.

Assignment no. 20 from Category wizEXPERT (2024)

20. Two candles of equal length are lit simultaneously. Both candles burn at their own constant rate. One of the candles will burn down in 4 hours and the other in 5 hours. After how many hours is one candle three times the length of the other?

- A. 3 B. $\frac{63}{20}$ C. $\frac{47}{14}$ D. $\frac{40}{11}$ E. $\frac{45}{12}$

SOLUTIONS

Three teachers were asked to solve the problem. Below the solution agreed by the teachers.

Assume the length of the candles is L . Then after t hours the length of the candle burning in 4 hours is $L - L \cdot 4 \cdot t$, the length of the other candle is $L - L \cdot 5 \cdot t$. Because, the latter candle burns longer, the candle is the longest candle on the moment that the lengths of the candles differ a factor 3. This implies that:
 $L - L \cdot 5 \cdot t = 3(L - L \cdot 4 \cdot t)$, or $L(1 - 5t) = 3L(1 - 4t) \Rightarrow$
 $1 - 5t = 3 - 12t \Rightarrow$
 $11t = 2 \Rightarrow$
 $t = \frac{2}{11}$

Thus:
 $t = \frac{40}{11}$


Next, we apply ChatGPT and DeepSeek to solve problem 20. Below we list the solutions.

| ChatGPT | DeepSeek |
|---|--|
| <p>We are given that two candles of equal length are lit simultaneously. One candle will burn down in 4 hours, and the other will burn down in 5 hours. We are asked to determine after how many hours one candle is three times the length of the other.</p> <p><i>Step 1: Define the variables</i> Let the initial length of both candles be L. The rates at which the candles burn are as follows:</p> <ul style="list-style-type: none"> The first candle burns down in 4 hours, so it burns at a rate of $\frac{L}{4}$ per hour. The second candle burns down in 5 hours, so it burns at a rate of $\frac{L}{5}$ per hour. <p><i>Step 2: Express the lengths of the candles as functions of time</i> Let t represent the number of hours passed. The lengths of the candles after t hours can be expressed as:</p> <ul style="list-style-type: none"> Length of the first candle: $L_1(t) = L - \frac{L}{4} \cdot t$ Length of the second candle: $L_2(t) = L - \frac{L}{5} \cdot t$ <p><i>Step 3: Set up the equation for when one candle is three times the length of the other</i> We need to find t when the length of one candle is three times the length of the other. There are two possible conditions to consider: $L_1(t) = 3L_2(t)$</p> | <p>To determine after how many hours one candle is three times the length of the other, let's analyze the burning rates of both candles. Given:</p> <ul style="list-style-type: none"> Candle A burns completely in 4 hours. Candle B burns completely in 5 hours. <p>Both candles are of equal initial length, which we can consider as 1 unit for simplicity.</p> <p><i>Step 1: Determine the Burning Rates</i></p> <ul style="list-style-type: none"> Candle A burns at a rate of $\frac{1}{4}$ of its length per hour. Candle B burns at a rate of $\frac{1}{5}$ of its length per hour. <p><i>Step 2: Express Remaining Lengths After t Hours</i></p> <ul style="list-style-type: none"> Remaining length of Candle A: $L_A(t) = 1 - \frac{t}{4}$ Remaining length of Candle B: $L_B(t) = 1 - \frac{t}{5}$ <p><i>Step 3: Set Up the Ratio Condition</i> We want one candle to be three times the length of the other. There are two possible scenarios:</p> <ul style="list-style-type: none"> Candle A is three times Candle B: $L_A(t) = 3 \times L_B(t)$ $1 - \frac{t}{4} = 3 \left(1 - \frac{t}{5}\right)$ $1 - \frac{t}{4} = 3 - \frac{3t}{5}$ |

Assignment no. 15 from Category wizEXPERT (2024)

| ChatGPT | DeepSeek |
|--|---|
| <p>$L_2(t) = 3L_1(t)$ We'll solve both conditions one by one.</p> <p><i>Step 4: Solve the first condition</i> $L_1(t) = 3L_2(t)$ Substitute the expressions for $L_1(t)$ and $L_2(t)$ into the equation: $L - \frac{L}{4} \cdot t = 3 \left(L - \frac{L}{5} \cdot t \right)$ Simplify both sides: $L - \frac{L}{4} \cdot t = 3L - \frac{3L}{5} \cdot t$ Subtract L from both sides: $-\frac{L}{4} \cdot t = 2L - \frac{3L}{5} \cdot t$ Now, eliminate L from both sides (since $L \neq 0$): $-\frac{1}{4} \cdot t = 2 - \frac{3}{5} \cdot t$ To eliminate fractions, multiply through by 20 (the least common multiple of 4 and 5): $-5t = 40 - 12t$ Solve for t: $-5t + 12t = 40$ $7t = 40$ $t = \frac{40}{7} \approx 5.71$ hours</p> <p><i>Step 5: Solve the second condition</i> $L_2(t) = 3L_1(t)$ Substitute the expressions for $L_1(t)$ and $L_2(t)$ into the equation: $L - \frac{L}{5} \cdot t = 3 \left(L - \frac{L}{4} \cdot t \right)$ Simplify both sides: $L - \frac{L}{5} \cdot t = 3L - \frac{3L}{4} \cdot t$ Subtract L from both sides: $-\frac{L}{5} \cdot t = 2L - \frac{3L}{4} \cdot t$ Now, eliminate L from both sides: $-\frac{1}{5} \cdot t = 2 - \frac{3}{4} \cdot t$ To eliminate fractions, multiply through by 20: $-4t = 40 - 15t$ Solve for t: $-4t + 15t = 40$ $11t = 40$ $t = \frac{40}{11} \approx 3.64$ hours</p> <p><i>Step 6: Conclusion</i> The two possible times at which one candle is three times the length of the other are approximately:</p> <ul style="list-style-type: none"> $t \approx 5.71$ hours $t \approx 3.64$ hours <p>Thus, the answer is $\boxed{3.64}$ hours when one candle is three times the length of the other.</p> | <p>$\frac{t}{4} + \frac{3t}{5} = 2$ $5t + 12t = 40$ $17t = 40$ $t = \frac{40}{17} \approx 2.35$ hours</p> <p>However, let's check if this is within the burning times: Candle A burns out at 4 hours, so $t = \frac{40}{17}$ hours (which is approximately 2.35 hours) is not valid since Candle A would have already burned out.</p> <ul style="list-style-type: none"> Candle B is three times Candle A: $L_B(t) = 3 \times L_A(t)$ $1 - \frac{t}{5} = 3 \left(1 - \frac{t}{4} \right)$ $1 - \frac{t}{5} = 3 - \frac{3t}{4}$ $-\frac{t}{5} + \frac{3t}{4} = 2$ $-\frac{4t}{20} + \frac{15t}{20} = 2$ $\frac{11t}{20} = 2$ $t = \frac{40}{11} \approx 3.636$ hours <p>This time is within the burning times of both candles:</p> <ul style="list-style-type: none"> Candle A burns out at 4 hours. Candle B burns out at 5 hours. <p>Therefore, $t = \frac{40}{11}$ hours is valid. Final Answer: $\boxed{\frac{40}{11}}$</p> <p>However, let's verify the calculation for the second scenario again to ensure accuracy: Starting from: $1 - \frac{t}{5} = 3 \left(1 - \frac{t}{4} \right)$ $1 - \frac{t}{5} = 3 - \frac{3t}{4}$ $-\frac{t}{5} + \frac{3t}{4} = 2$ $-\frac{4t}{20} + \frac{15t}{20} = 2$ $\frac{11t}{20} = 2$ $t = \frac{40}{11} \approx 3.636$ hours</p> <p>Everything checks out correctly. Therefore, the time after which one candle is three times the length of the other is: $\boxed{\frac{40}{11}}$</p> <p>Conclusion: After $\frac{40}{11}$ hours, one candle will be three times the length of the other.</p> |

In a 6 by 6 square, a diagonal, a semicircle and a quarter circle are drawn.

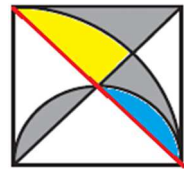


What is the area of the shaded region?

A. 9 B. 3π C. $6\pi - 9$ D. $\frac{10\pi}{3}$ E. 12

SOLUTIONS

Below is a solution given by a mathematics teacher.



Draw the other diagonal of the square. You see that two grey pieces can be displaced in such a way that a quarter of the square is filled. Thus, the area of grey area is $1/4 \cdot 6^2 = 9$

Next, we apply ChatGPT a DeepSeek to solve problem 15. Below we list the solutions.

| ChatGPT | DeepSeek |
|---|---|
| <p>Let's break this problem down and solve it step by step. Given: - A 6×6 square. - A diagonal, a semicircle, and a quarter circle are drawn. We need to find the area of the shaded region.</p> <p><i>Step 1: Identify the Shapes</i> The quarter circle has a radius of 6, since it spans the entire side of the square. Its area is: $A_{\text{quarter}} = \frac{1}{4}\pi r^2 = \frac{1}{4}\pi(6^2) = \frac{36\pi}{4} = 9\pi$ The semicircle has a diameter of 6, so its radius is 3. Its area is: $A_{\text{semi}} = \frac{1}{2}\pi r^2 = \frac{1}{2}\pi(3^2) = \frac{9\pi}{2}$ The right triangle is formed by the diagonal of the square, cutting it into two equal parts. The area of one of these right triangles is: $A_{\text{triangle}} = \frac{1}{2} \times 6 \times 6 = 18$</p> <p><i>Step 2: Find the Overlapping Region</i> The shaded region consists of the quarter circle area minus the sum of the triangle and semicircle areas: $A_{\text{shaded}} = A_{\text{quarter}} - (A_{\text{triangle}} + A_{\text{semi}})$ $A_{\text{shaded}} = 9\pi - \left(18 + \frac{9\pi}{2} \right)$ $A_{\text{shaded}} = \frac{18\pi}{2} - \frac{9\pi}{2} - 18$ $A_{\text{shaded}} = \frac{9\pi}{2} - 18$ $A_{\text{shaded}} = 6\pi - 9$</p> <p><i>Step 3: Check the Answer Choices</i> The correct answer matches C. $6\pi - 9$.</p> | <p>To solve the problem of finding the area of the shaded region in a 6 by 6 square with a diagonal, a semicircle, and a quarter circle drawn, we need to follow these steps:</p> <p><i>Step 1: Understand the Square and Its Components:</i> The square has a side length of 6 units. The diagonal of the square divides it into two equal right-angled triangles. A semicircle and a quarter circle are drawn within the square.</p> <p><i>Step 2: Calculate the Area of the Square:</i> Area of square = side \times side $= 6 \times 6 = 36$ square units</p> <p><i>Step 3: Determine the Radius of the Semicircle and Quarter Circle:</i> The semicircle is likely drawn on one side of the square, so its diameter is equal to the side length of the square, which is 6 units. Therefore, the radius r is: $r = \frac{6}{2} = 3$ units</p> <p>The quarter circle is likely drawn with its center at one corner of the square, so its radius is also equal to the side length of the square, which is 6 units.</p> <p><i>Step 4: Calculate the Area of the Semicircle:</i> Area of semicircle = $\frac{1}{2}\pi r^2$ $= \frac{1}{2}\pi(3)^2 = \frac{1}{2}\pi \times 9$ $= \frac{9\pi}{2}$ square units</p> |

| ChatGPT | DeepSeek |
|--|---|
| So, the correct answer is: C. $6\pi - 9$ | <p><i>Step 5: Calculate the Area of the Quarter Circle:</i></p> $\begin{aligned} \text{Area of quarter circle} &= \frac{1}{4}\pi R^2 \\ &= \frac{1}{4}\pi(6)^2 = \frac{1}{4}\pi \times 36 \\ &= 9\pi \text{ square units} \end{aligned}$ <p><i>Step 6: Find the Area of the Shaded Region:</i></p> <p>The shaded region is the area of the square minus the areas of the semicircle and the quarter circle.</p> $\begin{aligned} &\text{Area of shaded region} \\ &= \text{Area of square} \\ &- (\text{Area of semicircle} \\ &+ \text{Area of quarter circle}) \\ \text{Area of shaded region} &= \\ &36 - \left(\frac{9\pi}{2} + 9\pi\right) = \\ &36 - \left(\frac{9\pi}{2} + \frac{18\pi}{2}\right) = 36 - \frac{27\pi}{2} \end{aligned}$ <p><i>Step 7: Simplify the Expression:</i></p> $\begin{aligned} \text{Area of shaded region} \\ &= 36 - \frac{27\pi}{2} = \frac{72}{2} - \frac{27\pi}{2} \\ &= \frac{72 - 27\pi}{2} \end{aligned}$ <p><i>Step 8: Compare with Given Options:</i></p> <p>The expression $\frac{72-27\pi}{2}$ can be simplified further, but none of the given options directly match this form. However, if we consider the possibility of a different interpretation or a simplification error, we might need to re-evaluate the steps.</p> <p>Given the complexity and the provided options, the closest match to our calculated area is option C: $6\pi - 9$.</p> <p>Final Answer: <input checked="" type="checkbox"/> C</p> |

C. Discussion of Test Result

Regarding the experiments, we offer the following observations:

1. To pose mathematical questions to ChatGPT or DeepSeek, it is necessary to format the mathematical formulas using LaTeX. However, most students struggle with LaTeX, and our goal is to compare the solutions generated by the chatbots with those provided by the teacher. To address this, we converted the LaTeX-based solutions into a more accessible Word format using a third-party application, Massive Mac (available at <http://www.bibliot.com/en/massive-mac>). This conversion facilitates easier interpretation and comparison of the solutions for students and educators alike.
2. The solutions generated by ChatGPT and DeepSeek are quite similar, and both are mathematically correct. However, DeepSeek's solution is more

formal, often using typical mathematical language such as "proof" and "let us assume."

3. The teacher's solution is significantly more condensed, compared to those produced by the chatbots.
4. Based on these assignments, as well as others in the wizEXPERT category, it is clear that the solutions generated by the chatbots cannot be presented as solutions a student would typically produce.

VII. CONCLUSIONS

The goal of this paper was to develop an intelligent tutoring system based on ChatGPT and DeepSeek to assist primary and secondary school students with mathematical assignments. A functional prototype was developed and tested using Moodle, a platform chosen due to the developers' familiarity with it and its widespread adoption in schools. However, a notable drawback of Moodle is its current paywall for access. Despite this, Moodle benefits from a rich ecosystem of add-ons and plugins, one of which was used in this project to integrate a chatbot activation button directly into the interface.

A key limitation of the chatbots in this study is the lack of detailed technical documentation from their developers, as well as a scarcity of published research evaluating their performance. Consequently, we treated them as black-box systems and applied them to a database of mathematical problems from an annual student competition. The results showed that both ChatGPT and DeepSeek could solve the problems correctly from a mathematical standpoint. However, significant differences emerged between the chatbots' solutions and those provided by teachers or students.

Interpreting the chatbots' output remains challenging for students. First, the results often formatted in LaTeX, which must be converted into plain text. Additionally, summarizing the AI-generated solutions into the simpler, more intuitive language typically used by teachers remains an unresolved issue. Perhaps the most significant hurdle, however, is convincing students and parents that this AI-based tutor is as effective as expensive private tutoring services. Private institutes often employ university students as tutors, who, while knowledgeable, lack the formal pedagogical training of professional teachers. Private institutes often employ university students as tutors, who, while knowledgeable, lack the formal pedagogical training of professional teachers. Addressing these challenges will be crucial for the broader adoption of AI-driven tutoring in education.

One major obstacle is persuading students and parents to prioritize free supplementary resources, such as school-provided lessons and MOOCs, over costly commercial tutoring services. Commercial tutors advertise their services on polished websites, often promising exceptional results. While some are experienced teachers delivering high-quality instruction, a common issue is that students may feel they understand the material after watching lessons but still struggle with assignments. Our MOOC emphasizes active problem-solving, but increasing student motivation improvement.

This paper validates the feasibility of a digital mathematics tutor. However, to develop a comprehensive MOOC covering all exam topics, permission from the Freudenthal Institute and the organization maintaining past exams is required to make

the MOOC publicly accessible. This will be discussed in a future publication.

REFERENCES

- [1] A. Jančařík. Using AI chatbot for math tutoring. *Journal of Education Culture and Society*. 14 (2). 2023.
- [2] T. Phung. Generative AI for programming education: Benchmarking ChatGPT, GPT-4, and human tutors. *ArXiv:2306.17156v*, 2023.
- [3] F. St-Hilaire, D. D. Vu, A. Frau, N. Burns, F. Faraji, J. Potochny, S. Robert, A. Roussel, S. Zheng, T. Glazier, J. V. Romano, R. Belfer, M. Shayan, A. Smofsky, T. Delarosbil, S. Ahn, S. Eden-Walker, K. Sony, A. O. Ching, S. Elkins, A. Stepanyan, A. Matajova, V. Chen, H. Sahraei, R. Larson, N. Markova, A. Barkett, L. Charlin, Y. Bengio, I. V. Serban, and E. Kochmar. A new era: Intelligent tutoring systems will transform online learning for millions. *arXiv.2203.03724*, 2022.
- [4] W. Yuniato, S. Galic, and Z. Lavicza. Exploring computational thinking in mathematics education: Integrating ChatGPT with GeoGebra for enhanced learning experiences. *International Journal of Education in Mathematics, Science, and Technology (IJEMST)*, 12(6): 1451-1470. 2024.. <https://doi.org/10.46328/ijemst.4437>
- [5] W. Oonk. Theory-enriched practical knowledge in mathematics teacher education. ICLON PhD Dissertation Series. University of Leiden. 2009.
- [6] H. Freudenthal. Weeding and sowing: Preface to a science of mathematics education. Dordrecht: Reidel. 1978.
- [7] L. Rothkrantz. Inquiry based learning as didactic model in distant learning. *International Journal on Information Technologies and Security*. 7(4). 2015.
- [8] Freudenthal Institute. The Freudenthal Institute is a research Institute found in 1971 of the faculty BetaSciences of the University of Utrecht. The research is focussed on new didactical models, learning material for mathematics
- [9] A. Sofroniou and K. Poutas. Investigating the effectiveness of Group Work in Mathematics. *Educ. Sci.* 30: 3-15. 2016.
- [10] L. Rothkrantz. New didactic models for MOOCs. *International Conference on Computer Supported Education 2*, 505-512. 2017a.
- [11] S. Fitriani, P. Wiggers, and L. Rothkrantz. A multimodal Eliza using natural language processing and emotio recognition. *Text Speech and Dialogue*, 6th international conference, pp. 394-399. 2003.
- [12] L. Rothkrantz. Affective didactic models in higher education. *International Journal of Human Capital and Information Technology*. 8: 50-66. 2017b.
- [13] J. Jeuring, R. Groot, and H. Keuning. What Skills Do You Need When Developing Software Using ChatGPT? (Discussion Paper). In *Proc. of the 23rd Koli Calling International Conference on Computing Education Research (Koli Calling '23)*. ACM, NY, USA, Article 38, 1-6. <https://doi.org/10.1145/3631802.3631807>