10th International Conference on Higher Education Advances (HEAd'24)
Universitat Politècnica de València, València, 2024
DOI: https://doi.org/10.4995/HEAd24.2024.17366

Using ChatGPT for Creating Computer Science Exercises – a Case Study from the Real World

Antje Neve, Dieter Pawelczak
Department of Electrical and Computer Technology and Communication Technology, University of the Bundeswehr Munich, Germany.

How to cite: Neve, A; Pawelczak, D. 2024. Using ChatGPT for Creating Computer Science Exercises – a Case Study from the Real World. In: 10th International Conference on Higher Education Advances (HEAd'24). Valencia, 18-21 June 2024. https://doi.org/10.4995/HEAd24.2024.17366

Abstract

Using ChatGPT (Chat Generative Pre-trained Transformer) to support the tedious work of creating programming exercises seems obvious and helpful. This paper explains in detail the experience of working with the the artificial teaching assistant ChatGPT. The usage for the creation of lab assignments for a bachelor and a master level course is described, followed by the evaluation of the powerful yet not flawless tool for this specific use case. The article details a workflow that shows the steps for creating content and the necessity of feedback loops through the expertise and competence of the teachers. This article provides insights into the advantages, disadvantages and risks and possibilities that occur when working with this technology.

Keywords: *AI;computer science; teaching experience, STEM education.*

1. Introduction

ChatGPT (Chat Generative Pre-trained Transformer) has turned the usage of artificial intelligence systems upside down since November 2022 when it was published in version 3.5 by OpenAI (Roose, 2023). It enables anybody to use a Large Language Model (LLM) with natural language requests and receive structured text, source code, summaries or solutions to all kinds of queries (OpenAI, 2022). LLMs are part of the lives of students and teachers nowadays and they offer a chance for more efficient work but also require working with the students to teach them to assess the output and gain hands-on experience on the advantages and disadvantages that the models offer. The critical points need to be examined and students need to be made aware of the risks, chances and possibilities of these models (Domenech, 2023).

Kasneci et al. (2023) discuss in detail in which domains ChatGPT can be helpful for teachers and lecturers: planning of courses, support while/in learning a new (programming) language or in the evaluation and assessment of teaching material as well as the creation of multiple choice questions. They also mention that the development of personalized course materials are a

common use case for teachers. Despite the positive effects of reducing the work load of teachers for preparation of course materials or generating feedback reports to students (Mosaiyebzadeh et al, 2023), the use of ChatGPT also provides risks for teachers and learners. It is obvious, that copyright issues may cause troubles as ChatGPT does not reveal its references and even worse hallucinates output and references (Zuccon et al., 2023). Major implications can arise as ChatGPT's generated output might be biased and foster inappropriate stereotypes or lead to unfairness with respect to ethnic minorites (Kasneci et. al, 2023). According to Denny et al. (2024) generated programming solutions might also show bad programming habits and contain security flaws.

On the other hand, Students can use LLMs as never-tiring assistants for research assignments, receive support for the structuring and summary of texts and get help in assessing a problem from different view points (Kasneci et al., 2023). To get more insights into the advantages, disadvantages and risks and possibilities that occur when working with this promising technology, we performed a case study about the usage of ChatGPT as a teacher assistant in the bachelor degree programme *Applied Computer Science and Communication Technology* and master degree *Computer Aided Engineering* at the University of Bundeswehr Munich during lecture periods spring and fall/autumn 2023. ChatGPT was used for the creation of computer science lab exercises including sample solutions. The following sections detail the intention of the usage of ChatGPT, present the prompts and results and evaluate the actual benefit. By example it illustrates how ChatGPT can help teachers and students but also how its output needs to be carefully processed and evaluated before presenting it to students.

2. Usage of ChatGPT

2.1. Intent

Computer Science education usually includes programming exercises. These vary according to the courses content and teaching goals. There are usually programming classes, teaching algorithms, data structures and often the specifics of a particular language such as C/C++, Python or Java. There are also subjects such as Internet of Things (IoT) or Robotics and Sensors which often times convey a large variety of topics such as sensors, communication stacks or the principles of object detection in a robot (ACM and IEEE, 2020). These subjects usually contain labs with assignments that also include programming tasks. The large variety of frameworks, programming languages and last but not least hardware that runs the programs makes it challenging to keep up with the developments while providing the students with a broad overview of available technologies and teaching methods and not only technology stacks. The usage of ChatGPT appears to be a promising help in creating up-to-date exercises and solutions.

2.2. Example 1 – Lab for IoT Course – Undergraduate Level

For the class Internet of Things in a bachelor level computer science course, a lab to learn about the REST API (Representational State Transfer Application Programming Interface) in the programming language Python was created with the help of ChatGPT. REST APIs are widely used for example to request and process information from interfaces over the internet.

The method to work with ChatGPT was as follows:

1. Tell ChatGPT: "You are a professor. Create a lab for a computer science course to show the usage of a REST API in Python."
2. Request: "Create a sample solution for this lab. "
3. Optional: Create a formated output for the lab assignment description, e.g. a LaTeX, HTML or markup document.

Output of the first step: The specific task was the to create a small program to request weather information such as temperature and humidity for a given location and print this information. The output was structured into different subtasks including how to hand in the assignment.

For the second step ChatGPT's output included the advice to read the API documentation of the chosen REST API *OpenWeatherMap* (OpenWeather, 2024) which is freely available as well as sample code snippets that worked and gave a solution for request the temperature of a given location.

```
response=requests.get(f"http://api.openweathermap.org/data/2.5/
       weather?q={city}&appid={api_key}")
```

The temperature in the sample solution was in Kelvin and the code snippet included a simple calculation to achieve the conversion.

```
# convert from Kelvin to Celsius
temperature = data["main"]["temp"] - 273.15
```

However, ChatGPT didn't follow its own advice to read the API carefully. The *OpenWeather-Map* API includes a switch to use metric units which results in the temperature being in Celsius and makes the previous calculation unnecessary.

```
url=f"http://api.openweathermap.org/data/2.5/weather?q={city}
       &appid={api_key}&units=metric&lang=en"
```

The use of ChatGPT for the creation of a lab and a sample solution seems easy enough and produces output that compiles without errors, but at the same time, the sample solution doesn't live up to ChatGPT's own standards. It doesn't use the possibilities that the Open Weather Map API offers, but instead "hallucinates" an unnecessary calculation. This doesn't actually teach

the students well. The authors used the output to adapt the assignment and discuss with students about the correct usage of the API and make them aware of possible mistakes.

2.3. Example 2 –Application for Robotics and Sensors Course – Graduate Level

After an initially positive experience with ChatGPT as a learning tool assistant, it was used for lesson planning and lab creation in a graduate level course for Robotics and Sensor Systems. It is important and creates more in-depth output to give ChatGPT context such as the role of the prompting person (professor) or the audience of desired output (graduate level students of computer science). For the course Robotics and Sensors – a master level class – the teaching goals included gaining knowledge about object detection in general. Classes of detected object can include cars or pedestrians as in driver assistant system applications or faces and gestures. Naturally hands-on experience leads to more in-depth knowledge. Similar to the IoT domain, the topic of robotics and sensors includes a wide range of hardware, software and a large variety of software frameworks and libraries. For a teacher it can be very time consuming to keep up with the toolchains. In this master class it was decided to specifically study and apply the recognition of faces and gestures. To achieve this, two different libraries – OpenCV (OpenCV, 2024) and Mediapipe (Mediapipe, 2024) were chosen by the teacher. Both libraries are very popular and include hundreds of functions which are well documented. ChatGPT was used to create several labs on different levels – an introduction lab to get aquainted with the object detection and computer vision library OpenCV, an advanced lab to learn about filters to track a specific point in a detected face with the same library. For the third lab, ChatGPT was initially prompted to provide information and guidance on the specific topic of gesture recognition. ChatGPTs output included naming different application areas such as gesture recognition, sign language interpretation, virtual reality applications to immersively control hand or finger move–ments, medical applications or usage in automotive applications. Its response also included the potential of the technology such as the support of people with special needs. To practically explore gesture recognition the Mediapipe framework was used which includes a large amount of functions amongst them methods to recognize hands, fingers and finally gestures such as the victory sign. The output of the lab assignments by ChatGPT included the requirements for installing the programming language and respective packages, the assignments and a sample solution – the authors added teaching goals, references to APIs and specific classifier files and an adapted sample solution.

It can be noted that ChatGPTs output was more promising the more specific the prompts were and the more context they contained. As in the previous example ChatGPT proved to be a helpful assistant with structured output even in the desired format (LaTeX) for the lab assign-ments. It supported the teacher and enabled the usage of more frameworks. As a result an exploration of a wider range of practical applications was easier to prepare and to include in the lessons. However, ChatGPT didn’t include any valid references (compare e.g. Zuccon et al.,

2023) such as web pages for further reading or links to API documenation. The code that ChatGPT produced compiled and was correct. The authors assume that this is caused by the popularity of the chosen libraries and their assumingly big share in ChatGPT's training data. Generating the correct code, was an iterative process. Because the author had used the libraries quite extensively before a quality assessment of ChatGPTs output was possible.

2.4. Established Workflow

Figure 1 shows the workflow that was developed and used for the specific use case of computer science labs described above in section 2.2. and 2.3. At first, an idea for the lab is "discussed" with ChatGPT. Secondly, the sample solution for the lab is generated by ChatGPT and tested and validated by the teacher. Thirdly, a lab assignment document is generated in the desired output format. The fourth step includes a refinement adding references and teaching goals to the ouput. The first two steps require an iterative feedback loop led by the teacher to mitigate hallucination effects from ChatGPT. The teacher must be familiar with the libraries used in ChatGPT's output and needs to be competent enough to assess the quality of the generated source code.

3. Evaluation

Our experience with using ChatGPT confirms the findings of Kasneci et al., 2023, that in general, LLMs can provide teachers with output that can support lesson planning. In our case it enables a more diverse usage of libraries and frameworks for lab assignments. The output is structured and well formulated, its generation saves time and enriches the lessons. However, ChatGPT has no concept of right or wrong, there might be copyright issues and the data from LLMs can be biased due to unfair training data. Another risk of using LLMs for lesson preparation is that the output can be shallow. More alarming risks include that ChatGPTs output can be plain false and/or include incorrect parts, also known as "hallucinating". This can lead to a false sense of security and thus to superficial learning and a false confidence on the side of the students (Denny et al., 2024). When using LLMs in the classroom their advantages, disadvantages and the above mentioned risks need to be conveyed clearly to the students.

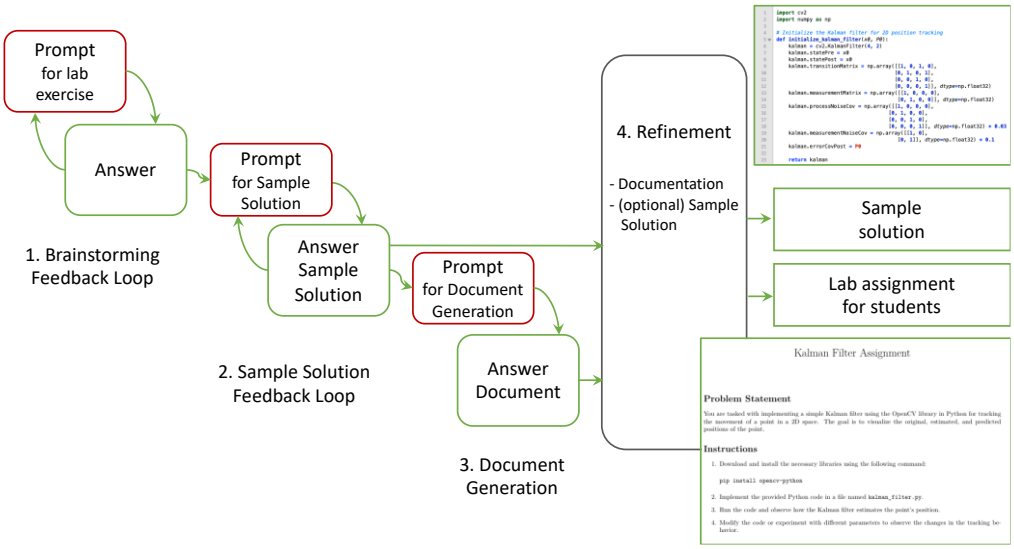

Figure 1. From an idea to a lab assignment - workflow for the lab generation with ChatGPT

4. Conclusion and Outlook

The authors used ChatGPT to create one new lab assignment on bachelor level and three new labs on master level. The labs include the actual assignment description in a directly usable output format (LaTeX), a sample solution in Python and additional teaching materials. Confirming the experience of others, ChatGPT proves to be a powerful yet not flawless tool. It can enrich lessons, bring more variety and be helpful to the teacher in the tedious tasks to create lab assignments. It is essential to critically assess ChatGPT's output and have sufficient experience and background know-how in the relevant topic. This is necessary to be able to foster and convey the critical thinking regarding this tool and its output to the students. Students benefit on various levels and teachers also need to stay up to date and validate output and enhance their own knowledge. Neither teachers nor students shall rely too much on the model. The Golden Rule of Usage for LLM Models shall be applied: *"Don't use the output of an LLM if you cannot verify its correctness." (Markgraf, 2023)*

In the future it will remain a challenge to find the balance of necessary study work to shape the ability in a certain domain like software development and therefore also the ability to critically assess output of LLM models and the use of tools like LLM models to enable broader views, easier access and adapted learning.

References

ACM and IEEE (2020). Computer Curricula 2020 - CC2020. Paradigms for Global Computing Education. *Tech. Rep.*, Association for Computing Machinery (ACM) and IEEE Computer Society. DOI: 10.1145/3467967

Denny, P., Prather, J., Becker, B. A., Finnie-Ansley, J., Hellas, A., Leinonen, J., Luxton-Reilly, A., Reeves, B. N., Santos, E. A., & Sarsa, S. (2024). Computing Education in the Era of Generative AI, Communications of the ACM, 67 (3). DOI: 10.1145/3624720

Domenech, J. (2023). ChatGPT in the classroom: friend or foe?, 9th International Conference on Higher Education Advances (HEAd'23), Universitat Politècnica de València, València, 2023, DOI: 10.4995/HEAd23.2023.16179

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, Volume 103, DOI: 10.1016/j.lindif.2023.102274.

Markgraf, Goetz. (2023) Software Entwicklung ohne Software Entwickler? KI-gestützte Softwareentwicklung – ein Erfahrungsbericht, Der Softwerker, Codecentric online: https://info.codecentric.de/hubfs/Softwerker/Softwerker%20Spezial%20Generative%20KI/Softwerker_Spezial_GenKI_online.pdf (in German)

MediaPipe Library. Online: https://developers.google.com/mediapipe (accessed April, 2024)

Mosaiyebzadeh, F., Pouriyeh, S., Parizi, R., Dehbozorgi, N., Dorodchi, M., & Macêdo Batista, D. (2023). Exploring the Role of ChatGPT in Education: Applications and Challenges. In Proc. of the 24th Annual Conf. on Information Technology Education (SIGITE '23). Kennesaw State University, Georgia, USA, October 11-14, pp. 84-89. DOI: 10.1145/3585059.3611445

OpenAI (2022). Introducing ChatGPT: Optimizing Language Models for Dialogue. November 30, 2022. Online available: https://openai.com/blog/chatgpt/ (accessed April 24th, 2024)

OpenCV Library. Online: https://opencv.org/ (accessed April 24th, 2024)

OpenWeatherMap. Online: https://openweathermap.org/api (accessed April 24th, 2024)

Roose, K. (2023). Don't ban ChatGPT in schools. Teach with it. The New York Times, 2023/01/12

Zuccon, G., Koopman, B., & Shaik, R. (2023). ChatGPT Hallucinates when Attributing Answers. In Proc. of the Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval in the Asia Pacific Region (SIGIR-AP '23). Beijing, China, November 26-29, pp. 46-51. DOI: 10.1145/3624918.3625329