A chatbot dialogue model: Understanding human-chatbot collaboration in a complex task environment

Due to greater budget restrictions and the popularity of online learning education institutes are increasingly adopting massive open online courses. One disadvantage is however the lack of individualized student support, resulting in questionable learning outcomes. Chatbots could act as such an individualized scaffold for students and therefore be a useful tool to improve student learning success. Having this in mind, we present a design science research dialogue model for designing a chatbot and its dialogue in a problem-based learning scenario, intended for use in education and online courses especially. We also propose seven design goals for designing a chatbot dialogue on which we then evaluate the model.

Chatbot, dialogue model, problem solving, design science research
A chatbot dialogue model: Understanding human-chatbot collaboration in a complex task environment

Abstract. Due to greater budget restrictions and the popularity of online learning education institutes are increasingly adopting massive open online courses. One disadvantage is however the lack of individualized student support, resulting in questionable learning outcomes. Chatbots could act as such an individualized scaffold for students and therefore be a useful tool to improve student learning success. Having this in mind, we present a design science research dialogue model for designing a chatbot and its dialogue in a problem-based learning scenario, intended for use in education and online courses especially. We also propose seven design goals for designing a chatbot dialogue on which we then evaluate the model.

Keywords: Chatbot, dialogue model, problem solving, design science research
1 Introduction

The rise of massive open online courses (MOOC) provides means to economize and scale education in response to higher budget restrictions [1] and raises the number of possible enrolled students by an unprecedented amount, compared to traditional higher education courses [2]. However, a disadvantage of MOOCs is the lack of individual support which cannot be addressed economically while catering to a massive audience, often resulting in a questionable learning outcome [3]. Chatbots have the potential to solve this issue as a supporting tool to help the user learn how to solve a complex set of problems [4]. A chatbot is a conversational interface that is able to interact with users in a certain topic with natural language sentences [5]. In 2017 the use of chatbots in U.S. households like Amazon Echo or Google Home has increased by 128.9 per cent compared to the previous year. Future forecasts indicate that their use will likely rise so that by 2022 the majority of the U.S. households will have at least one smart assistant in their home [6]. Chatbots have been around since at least 1966 [7], yet their use in more complex settings currently still involves difficulties, as interactions between users and chatbots remain for the most part simple and usually follow a simple question answer approach [8]. With this ability chatbots could support the development of higher-order skills in a way which MOOCs are currently lacking [9], due to the absence of individualized support [4]. The application of advanced technologies such as deep learning [10] and the recently unveiled Google Duplex, a more technologically advanced chatbot, already hints in a direction where in the near future chatbots can act as collaboration partner in solving more complex tasks [11]. However, this remains a research field which is only beginning to unfold, with none or little research undertaken to how chatbot user interaction can be structurally developed with in a problem task scenario.

Therefore, the aim of this paper is to design a model on how chatbots and users can collaborate together through dialogue in order to solve a specific problem, while outlining the learning process of the user. By building upon the model, chatbot designers can be enabled to leverage problem-based learning theory and thus create an individualized learning scaffold for students. To that effect a design science research (DSR) approach is used, following the lifecycle model by Peffers et al. [12]. To visualize the model, the Business Process Model and Notation (BPMN) language is used. Beyond the practical application the paper is also expected to provide research insights from the chatbot viewpoint for fields related to technology-mediated learning, technology-enhanced scaffolds and, more broadly, human machine communication. In the next chapter the theoretical background is addressed, followed by a description of the research methodology. Afterwards it is shown how the DSR process was followed, including a presentation of the model, the formulation of design goals and the evaluation of the model based on them. In the final part it is discussed whether and how the model can be useful for the initial research question and where weaknesses might exist. Finally, it is concluded how the research can be expanded based on this paper and where there remain gaps which can be subject of future research.
2 Theoretical background

2.1 Chatbot

A chatbot is able to interact with users in a certain domain or a certain topic with natural language sentences, usually by answering a question of a user or introducing a new topic [5]. Such a chatbot comprises software including conversational agents, chatterbots and virtual assistants known as Amazon Alexa or Google Home [4]. Chatbots and users can interact with each other by text and speech [10]. In their literature review Gnewuch et al. [13] derive four design principles which enable the avoidance of several issues related to designing conversational agents, or more specific chatbots. Among those are securing an ample communicative ability of the chatbot, providing the chatbot with conversational abilities, providing perspicuous and flexible conversation flows combined with clarification, confirmation and error-handling abilities. They enable the chatbot to use social cues and be aware of context. They are however more of a technical nature and can serve as a technical boundary for learning-focused aspects. The role of conversational cues is explored by Vtyurina and Fourney [8] and found to be especially important in a task-related context, which goes beyond a simple question answer paradigm, supporting the design principle of Gnewuch et al. [13] mentioned above. Winkler and Söllner [4] extend on chatbot research by specifically including the education aspect in a technological, or more specific chatbot, environment and derive several important factors to enable chatbots to succeed in an education environment. Chatbots can also help students by promoting their own proactivity, therefore increasing the learning success. This goes along with the necessity to inquire instead of instruct [14] and the general effectiveness of inquiry learning [15]. In addition, chatbots can be used to improve a student’s motivation and self-efficacy and are also able to provide feedback to students about their performance [4]. Another factor is the chatbot’s context awareness to increase cognitive and affective learning outcomes, which includes for example the detection of knowledge gaps of students [16] and relational behavior such as caring and empathy [17]. Additionally, the chatbot should adapt to reading comprehension and, in general, prior knowledge [18].

2.2 Scaffolding problem-solving phases

Student difficulties in learning can be partly attributed to misconceptions or lack of domain-specific knowledge [19], [20], the bulk of it is however related to a lack of problem-solving skills [19]. Kim and Hannafin [14] describe five phases according to which problem solving is done by the user in a technology-enhanced scaffolding context. They describe in what order and how problem solving support should appear. The phases are (1) identification and engagement, (2) exploration, (3) reconstruction, (4) presentation and communication and (5) reflection and negotiation. In the first phase learners find or generate problems and externalize them. During the exploration phase learners probe problems, plan investigations, and test hypotheses. In the reconstruction phase learners generate and revise potential solutions and explanations. The
presentation and communication phase involves sharing of constructive feedback with peers and teachers regarding visualized or verbalized solutions and explanations. Finally, in the reflection and negotiation phase learners examine the processes and strategies used and revise their solutions and explanations [14].

3 Methodology

We use a DSR approach following the lifecycle model of Peffers et al. [12], which includes the following steps: (1) identify problem, (2) define solution objectives, (3) design and development, (4) demonstration, (5) evaluation and (6) communication. The approach is chosen to enable the creation of a design which can be iteratively evaluated with a well-established method [20]. For the visualization of the model the Business Process Model Notation (BPMN) [21] is used. In the design the problem solving phases according to Kim and Hannafin [14] are used as a general basis on which the model is built. First, to enable the model to support solving a problem in a well-founded order and way, and second, to deduct the necessary steps the chatbot has to take in order to improve the learning outcome. As a foundation for the evaluation of the model essential design goals are derived from requirements and issues found in the existing literature, as described in the previous chapter. For that matter a formal literature research has been conducted using the following keywords: “chatbot”, “amazon alexa”, “amazon echo”, “problem solving”, “problem based learning” and “smart personal assistants”. Altogether we screened the abstracts of at least ninety articles and ultimately chose eight to derive design goals from. Furthermore, correspondence with two active researchers in the field was undertaken to improve, evaluate and expand on these design goals by capturing insights from experts who have already or are in the process of undertaking multiple studies in the same research field. By following these design goals, issues and requirements related to problem solving and learning can be depicted for the chatbot context. The resulting evaluation objective for the model is therefore to maximize the degree to which the design goals can be depicted.

4 Designing the chatbot problem solving model

4.1 Problem identification

As outlined in the introduction, education institutions are increasingly using MOOCs, i.e. in response to budget restrictions [1]. One of the main disadvantages of MOOCs is the lack of individualization affecting the learning success [3]. Chatbots can be used to alleviate this problem by guiding students through a problem task [4], yet in research there are few examples on how to implement them in the context of learning. If there are, they are mostly restricted to simple question and response dialogue [22]. Furthermore, no studies could be found which specifically target the problem solving context and leverage established pedagogic models.
4.2 Objectives

The introductory and fundamental objective of this paper is to design a process model on how a chatbot and a user can collaborate together to solve a problem, while also improving the learning outcome of the user. In the following the objectives are described more specifically. An overview of the formulated design goals (DG) can be seen in figure 1. An appropriate inclusion and connection with existing pedagogic literature and their models, especially in the problem solving context, is necessary to make sure the chatbot can mimic or build upon good practice learning processes in the technology-enhanced scaffolding context (DG1). Proactivity and motivation are important factors for a student’s learning success [23]. As such promoting motivation and proactivity is an important means to improve the learning outcome and should be considered when developing a chatbot in a learning context [4]. This also goes in line with the five problem phases by Kim and Hannafin [14], where the learner is the main driver of the learning process and an inquiry-based learning approach is used, instead of an instruction one. If the collaboration is predominantly or only driven by the instructor/the chatbot learning opportunities are missed [24] (DG2). In addition, providing effective feedback is recognized as a key indicator to learning success both in general research and is also part of the latter problem solving phases of Kim and Hannafin [14]. Where possible, the chatbot should provide indicative feedback [4] and otherwise promote other means to get feedback such as directing to a more knowledgeable person (DG3). This was especially the concern of the aforementioned experts. As identified, providing related domain-specific information [25-26] through the chatbot is a prerequisite to empower the user to solve complex problems (DG4). The involvement of conversational cues [13] in a chatbot can initially be a technical problem. Because of the vast path of ways to transition between the problem solving task, it leads i.e. to the question how and when a user is going forward with the task, since the chatbot is completely reliant on user input to determine the status of that task. Beyond that, flexible reiterations might also be needed, when an error is noticed unexpectedly. Both imply that the chatbot designer cannot expect a linear collaboration process (DG5). Context awareness is another key indicator for an improved learning outcome directly identified by Winkler and Söllner [4] related to the individualization of the learning process (DG6). Since the main intended use of the derived model is to support and guide the development of a chatbot skill, the practitioner’s view is highly relevant in order to optimize the model to accommodate to the nuances in the field. The ease of use and simplicity is emphasized by both researchers as a key requirement,
recommending the use of a model language which is already well known to a large audience (DG7).

<table>
<thead>
<tr>
<th>Design Goals</th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>DG1</td>
<td>Incorporation of a structured problem solving approach and existing theory</td>
</tr>
<tr>
<td>DG2</td>
<td>Promote the user’s motivation and proactivity</td>
</tr>
<tr>
<td>DG3</td>
<td>Provide and promote feedback to the user</td>
</tr>
<tr>
<td>DG4</td>
<td>Provide related information to the user</td>
</tr>
<tr>
<td>DG6</td>
<td>Inclusion of context awareness</td>
</tr>
<tr>
<td>DG7</td>
<td>Ease the use and understanding of the model</td>
</tr>
</tbody>
</table>

Figure 1. Design goals for creating a chatbot in a problem-based learning context

4.3 Design and development

The chatbot dialogue model for problem-based learning uses the BPMN and displays the interaction between a chatbot and its user. The tasks of both chatbot and user are supposed to run in parallel where aligned and not specifically mentioned otherwise. The message flows indicate the collaboration between chatbot and user. It also does not differentiate between text or speech dialogue and assumes that it can be applied on both, since there are no fundamental differences between either which would invalidate the model’s generic assumptions. Due to the complexity there exist two layers of the model. In the first layer the general process is shown, while sub processes are used in the second layer to specify the procedure of important tasks in more detail. This is also used to increase the modularity, as chatbot designer might want to implement only those parts of the model, which are applicable to their own, perhaps less generic, problem set or are complementary to an existing one. There are the following five steps in the model: Onboarding, exploration, task execution, presentation and reflection. They are supposed to be roughly equivalent to the five problem solving phases of Kim and Hannafin [14]. Between the steps there are regularly confirmation or information BPMN tasks, to indicate a transition to/of the chatbot. The first layer of the model is visualized in figure 2. The visualization of the remaining sub processes of the second layer, as well as the first layer in bigger size, can be found in the addendum. In the onboarding step user and chatbot alike provide information to each other to determine whether a collaboration is possible and when the problem should be approached. More importantly the problem itself is defined. If the problem is already known to the user,
no further guiding by the chatbot is necessary. If not, the model outlines this - depending on the problem set that is implemented by the chatbot designer. In the second exploration step the role of the chatbot has a mainly passive and informational nature to support the problem solution exploration of the user in an inquiring way.

**Figure 2.** The model’s (reduced) first layer

The user probes the problem and, if necessary, conducts further research. Going forward, a proper solving approach is chosen. For the chatbot to assist the user, the problem is first structured. To assist research resources can be provided. If no research is needed due to low complexity or other reasons, or alternate approaches are helpful for understanding, other problem solving approaches can be provided by the chatbot. If the decision about the solving approach by the user doesn’t seem to be applicable, the chatbot can suggest the aforementioned problem solving approaches as an alternative solution. Finally, the chatbot defines a goal for later feedback and evaluation. In the task step the previously identified problem solving approach is executed. The process itself is internalized by the user and reiterations performed, if the initial results are not satisfactory. The chatbot is supposed to assist the user by taking care of easier tasks, including sharing necessary information, and making sure that the ongoing progress is proceeding correctly. Furthermore, if a task is relatively difficult in comparison to the user’s experience, more instructions should be given to reduce frustration/prevent loss of motivation. The chatbot is less present in the presentation step, since a presentation normally occurs between the user and one or more persons, ideally a classroom [14]. If the result is of quantitative nature, a chatbot might be able to provide reasonable support. If not at least one outside person should be consulted, ideally with more expertise. In traditional MOOC, like Coursera, there are already mechanics which link users with other users to get help and feedback [27]. Coursera also maintains “community mentors” who fulfill the role of a course expert [28]. With that context the chatbot could provide information on how to contact these persons or even connect the user directly with such a mentor, similar to a live chat functionality in ecommerce.
Based on that feedback the process can progress to the reflection step. In the final step the user self reflects on the chosen problem solving approach and how the implementation followed in order to learn from it. Furthermore, feedback is directed to the chatbot, which is then processed by the chatbot to improve a future next iteration. Before ending the process, the user has the chance to directly test the learning outcome by iterating through the process again from the exploration step onwards. If no reiteration is undertaken, the user ends the process by informing the chatbot. To support the user’s reflection the chatbot can provide direct feedback about the approach and the result, if possible, and also suggest alternative solutions. Finally, the user should also be motivated about the learning outcome, i.e. by providing comparison to other users and friends, praising the user on great success or stamina and also stifle proactivity by i.e. presenting more advanced topics.

4.4 Demonstration

A formal demonstration has been made illustrating the model and its process with multiple iterations of an example dialogue. The demonstration includes user interaction with task tracking and several non-straightforward user learning difficulties, which shall depict the complexity which is currently not part of current chatbot learning research [8]. These illustrations were then shown to both researchers whose iterative feedback became then part of the evaluation step. For example, it was specified that the model’s complexity is not beneficial, leading among other things to the division of first and second layer in the model. Furthermore, it was found, that the problem-based learning analogy was not likely to be easily comprehended by the model’s users. As a consequence the model’s own steps were aligned to the problem solving ones of Kim and Hannafin [14]. A future study is supposed to be conducted in which a technical implementation is made, using this model.

4.5 Evaluation

To evaluate whether the model can be validated in the first iteration according to Peffers et al. [12] the model is analyzed to which extent the content of the respective design goal is captured. To represent DG1, the incorporation of a structured problem solving approach is obligatory. The already presented problem solving phases of Kim and Hannafin [14] are the core of the model to which chatbot and user interaction are aligned. Every step of the model is closely related to a problem solving phase and includes the subtasks of the respective scaffold/chatbot or user side. By leveraging this the model can include a researched and established way of problem solving among learners which was previously linked to technology-enhanced scaffolding in a non-chatbot context. The promotion of the user’s motivation and proactivity with DG2 can be found in multiple steps of the model. In the exploration step support/inquiry is intended, not instruction, stifling proactivity. It is also possible to provide problem solving approaches as an example to help the user extrapolate other approaches. In addition, goals are defined to stifle motivation. In both exploration and task steps the chatbot is primarily in a passive, supportive role. In the task step those tasks which
provide meagre or no learning success or are currently too difficult for the user are taken care of by the chatbot, while those tasks which do are supposed to be internalized by the user. In the reflection step further motivational and proactivity inspiring information is offered. The feedback supply to accommodate DG3 is found mainly in the exploration and task steps. First, feedback is provided by providing alternate problem solving approaches, if necessary. Secondary, the problem status is tracked, meaning misunderstandings or more broadly errors can be either avoided or corrected on the spot. Feedback about the result can also be given in the presentation and reflection step. This is primarily due to the possible role of an outside expert. When possible, direct feedback by the chatbot can be given in the reflection step, mostly on quantitative results and by suggesting alternative solutions. Otherwise the chatbot can make sure that feedback is available to the user by directing him to other sources.

Additionally, the model should provide related information to the user according to DG4. In the model this is mainly captured in the exploration step by presenting research opportunities/problem solving approaches and in the task step by taking care of the easy (sub-)tasks and sharing information about it. In that way the lack of domain-specific knowledge or experience can be alleviated. For DG5 the involvement of a “conversation flow flexibility” is necessary. In the model transition between the steps are explicitly defined as BPMN tasks. This emphasizes how reliant the chatbot is on user input regarding transition and that, if no input is given, an inquiry must be made to the user. Regarding reiterations due to errors, miscalculations etc. two reiterations are included in the model; when the results of the task step have been made, deciding whether the preliminary results warrant going forward to the next steps and also at the end, deciding if the problem has been solved based on the set goals. Including the individualization is of particular importance to be able to provide the context awareness according to DG6. Therefore, the extent to which problem solving approaches are provided or harder tasks instructed, based on the user’s existing knowledge, is implied in the model. This DG is however perceived more as a “meta” DG and is therefore not specifically mentioned at every step. The chatbot designer should implicitly include it in the whole development process. An example is showing empathy which should be addressed at every single step. Regarding the ease of use and understanding of the model (DG7) the BPMN 2.0 modelling language is used as it is already known to a larger audience. This does not mean however that other languages could not have been used as well. Here BPMN was chosen due to own familiarity, researcher advice and popular use [21]. In the model’s case pools are chosen as means to separate chatbot- and user-triggered interaction. Alternatively, simple lanes could also be used. To lower the complexity and subsequently ease the understanding of the model two layers are separated by subprocesses. Each step of the model can therefore be looked at separately by the designer. The designer or user of the model is therefore not forced to comprehend the whole of the model to apply parts of it. In the second iteration of the evaluation an example dialogue was presented to both researchers, together with the model. Based on their experiences they were then asked about what changes they would now request and where they feel the model might include or exclude too much based on their own experiences. This goes in line with the procedure outlined in the previous part of the chapter.
5 Discussion

With this paper the question how chatbot and user can collaborate on a problem effectively in education has been investigated. As a result a model is proposed which is based on seven DG and was evaluated regarding how chatbot designers can be enabled to improve the learning outcomes of their educational/learning application by adhering to the problem solving phases of Kim and Hannafin [14]. In the evaluation it was found that all DG are depicted to some extent in the model. Most distinguishing is the support the chatbot can give when the task is actually done by providing additional information. On the other hand, the feedback part can be lacking, when, due to the qualitative nature of a result, no direct feedback can be given and outside persons have to fill the open gap. Nevertheless, the chatbot can alleviate this to some extent by directing the user to some type of mentor, if needed. While exchange with two active researchers was conducted regarding both design goal and the model itself, more chatbot experts, possibly from different backgrounds, should be involved in the process. In the future further research in this field is expected, which will possibly lead to a different set of DGs. Since the current model has a very generic and modular character, chatbot designers can leverage the current design and either customize it to their more specific use case or use only parts of the model. On the other hand, this means that chatbot designers cannot expect to take the generic model and adapt it to a more specific context without any own input. Limitations also include that Kim and Hannafin [14] only refer to a technical-enhanced classroom in their five problem solving phases. The role of a (missing) teacher has not been investigated in this paper and it remains unclear how the replacement or supplement of a chatbot affects the validity of the five problem solving phases. This is a topic for further research. Nevertheless, by relying on these phases, the model can build upon established learning research, focusing on an increased learning performance while utilizing chatbots.

6 Contribution

In the future chatbots have the growing potential to transform the educational landscape by creating individual learning experiences for students and reducing teacher workload on a global scale [4]. With this paper the current research is expanded by first insights on how to combine learning and chatbot practically. We enable chatbot designers and developers to model the interaction processes between a chatbot and the user in order to design their chatbot applications having that goal in mind. Going forward, research can extend and build upon this foundation. Beyond practical insights, the paper also provides research insights for fields related to technology-mediated learning, technology-enhanced scaffolding or more broadly human machine communication, where collaboration between human and machine is touched on, by applying traditional learning literature on a chatbot. For future research the methodic foundation should be extended according to the evaluation cycle of the DSR approach of Peffers et al. [12]. It is also prudent to test the model practically by developing an technical chatbot application with this model in mind. In a broader chatbot and learning context, Winkler
and Söllner [4] identify ten gaps in current research literature. With this paper, the gap “Integration of chatbots in different steps of the learning process with the help of learning theories” has been investigated.

References


**Addendum**

Complete illustrations are available here: https://www.dropbox.com/sh/218y7ewic12n8l6/AAAUsg2K4ADN6kob_gxL7E_4a?dl=0