Educational Chatbots: A Sustainable Approach for Customizable Conversations for Education

Donya Rooein[a], Paolo Paolini[b] and Barbara Pernici[c]
Poliotecnico di Milano, Milan, Italy

Keywords: Educational Chatbot, Adaptive Learning, Conversational Agents, Customizable Conversations.

Abstract: This paper proposes using chatbots as "tutors" in a learning environment; tutors who are not domain experts but helpers in guiding students through bodies of learning material. The most original contributions are the proposal that conversation should be content-independent (although chatbots speak about content); The production process should allow non-technical actors to customize chatbots and keep the costs of development and deployment low. We specifically discuss conversation customization, which is relevant, especially for learning applications, where users might have specific needs or problems. We achieve the features introduced above via extensive "configuration" (regarding direct programming), making the underlying technology novel and original. Experiments with teachers and students have shown that chatbots in education can be effective and that customization of conversations is relevant and valued by users.

1 INTRODUCTION

Chatbots are famous for being a valid alternative to traditional point-and-click interaction between humans and computers (Brandtzæg and Følstad, 2017). There are many studies who have investigated different roles of chatbot in education (Hwang and Chang, 2021), and this paper is about customizable (or even adaptive) conversations for educational chatbots. This study is an alternative approach to the typical one-size-fits-all chatbot solution for educational chatbot conversation design, which is relevant to education, where psychology and individual needs are of great importance.

Customization of conversational features may include the chatbot's loquacity, the topic of the chatbot's turns, the chatbot's style and wording, or the length of the chatbot's wording in each turn. Our idea is that customization of the conversation will make it more effective, i.e., better tuned to the user's profile (including students with special needs) and better tuned to contextual situations (e.g., the user is tired or in a hurry). Customization means that someone will make choices (explicit or implicit); in the future, we envision adaptive conversations, i.e., the chatbot will interpret data and decide on possible changes to conversational features.

There are many different roles for chatbots in education, such as a chatbot to support learning experience, an assistive tool, or a tutoring and mentoring role (Wollny et al., 2021). We propose to use chatbots as tutors, helping students through a body of content. A chatbot acting as a tutor is not a domain expert (say History or Computer Science); it must be able to sustain a good conversation in a learning experience and help the student move through the various learning elements[1]. There are several motivations for this choice: i) there is a vast amount of digital content already available (e.g., online courses, MOOCs[2], Learning objects, digital resources); rather than creating new content, chatbots can help at making better use of existing ones; ii) with respect to the traditional (interactive) interface, the conversation may add friendliness, empathy, and easiness to use all ingredients relevant for learning; iii) learning, especially in formal education, has an already established group of actors (e.g., authors, teachers, publishers, etc.); a new technology should improve the activities of these actors, rather than replace them.

Refining the above idea, we elaborated many

[a] https://orcid.org/0000-0002-0368-3084
[b] https://orcid.org/0000-0003-3486-5662
[c] https://orcid.org/0000-0002-2034-9774

[1] The learning element refers to each learning material unit, e.g., a theoretical description.

[2] Massive Open Online Course

Rooein, D., Paolini, P. and Pernici, B.
Educational Chatbots: A Sustainable Approach for Customizable Conversations for Education.
DOI: 10.5220/0011083200003182
In *Proceedings of the 14th International Conference on Computer Supported Education (CSEDU 2022) - Volume 1*, pages 314-321
ISBN: 978-989-758-562-3; ISSN: 2184-5026

high-level requirements as follows: i) the conversation afforded by the chatbot should have as much as possible a human flavor, providing functional and nonfunctional aspects; ii) the chatbot should conduct mainly a proactive conversations taking the lead in order to take the learner across the material effectively; iii) the content should be customizable according to the profile of the learner, her specific needs and also contextual situation; iv) the conversation itself should be customizable, in order to be accepted and effective for the learner; v) actors in educations should be empowered to control customization both of content and conversation since they are in charge of organizing learning; vi) learners should be able to further control and customize their learning experience. vii) the overall process for creating chatbots should be streamlined, effective, low cost, and with little IT personnel involvement. Consequently, a specific type of chatbot named by TalkyTutor was developed.

In this paper, we focus specifically upon the feature of the chatbot that allows the conversation to be customized. The basic idea is that the conversation is controlled by several configuration data that can be easily modified without programming. Initial experiments (with teachers and students) show that customizing conversation is very promising and appreciated by users. In Section 2, we discuss related work; in Section 3, we discuss how conversations are described via a configuration-driven approach and can be customized; in Section 4, we describe the initial experimentation, the related qualitative and quantitative assessment; in Section 5, we draw the conclusions and discuss future work.

2 RELATED WORK

2.1 Chatbots in Education

The concept of educational chatbots has its origins from intelligent tutoring systems, which address the idea of building a learning tool that is intelligent enough to understand learners' needs and proceed accordingly (Song et al., 2017). Authors in (Burkhard et al., 2021) presented a theoretical basis for the use of smart machines like chatbots in education by considering the role of teachers and focusing on the necessity for teachers to play an active role in the digital transformation.

From the broad application of conversational agents in education over time, we mention the Grasser et al. a study from 20 years ago who introduced tutoring systems as a conversational agent to help college students learn about computer literacy (Graesser et al., 2001). Heller et al. developed a chatbot designed by open source architecture of AIML[3] to improve student-content interaction in distance learning (Heller et al., 2005). In 2008, Kerry et al. worked on using conversational agents for self-assessment in e-learning (Kerry et al., 2008). Chatbots are often applied for organizational support to perform specific tasks, e.g. automated FAQ (Han and Lee, 2022). An intelligent teaching assistant (iTA) introduced by (Duggirala et al., 2021) to help students by providing detailed answers to their questions by using a generative model, extracted the relevant content from the top-ranked paragraph to generate the answer.

2.2 Chatbots and Conversation

According to McTear, despite all the progress in speech recognition and natural language understanding, but still chatbots suffer from the lack of the conversational abilities that make the interaction with them unnatural (McTear, 2018). Conversation by itself is a complex system, but the aspect of the conversation design plays a crucial role on the chatbot's effectiveness. Natural Conversation Framework (NCF) presented a new approach for implementation of the multi-turn conversation between chatbot and human (Moore and Arar, 2019). This approach uses a library of predefined UX patterns driven by natural human conversation. The conversation structure in chatbots are mainly hand-crafted, which means the whole conversation is embodied by content and developed by IT experts (Paek and Pieraccini, 2008). This approach makes chatbot development and maintenance expensive and not easily generalizable to other domains. On the other hand, statistical methods use machine learning algorithms to learn about an optimal dialogue strategy from the interaction with users. In addition, there are end-to-end dialogue systems that use deep neural networks to generate responses from the large corpora of dialogues (Serban et al., 2016).

2.3 Chatbots and Customization

If we look deeper into the past three years of research on educational chatbots, we can consider Tegos et al.'s work on a configurable design of chatbots for synchronous collaborative activities in MOOCs in a university setting relevant to our research (Tegos et al., 2019). They targeted MOOCs and integrated them using chatbots to support simple tasks such as collecting feedback from learners. A recent study(Rooein et al., 2020) has shown the use of chatbots also for training new employees in a factory by adaptive approach.

[3]Artificial Intelligence Markup Language

Authors in (von Wolff et al., 2019) and (Winkler et al., 2020) focused on the educational actors (students and teachers) to extract important features and investigated the requirements implementing a chatbot in university setting and also confirm the acceptance of the chatbots in academic environments.

All the mentioned approaches are highly involved by IT experts to develop and maintain the conversation module of the chatbots. Here, we start with hand crafted approach to design the conversation component of chatbots, and later by a configure driven approach make it more sustainable in the production and maintenance point of view and also delivering a more flexible conversation.

3 DESIGNING AND DEPLOYING ADAPTIVE CONVERSATIONS

In this section, we describe our approach for designing and implementing conversation machine in educational chatbots. The key aspect of this approach is that it does not support a specific conversation, but it allows designers, authors and teachers to model the conversations that they wish. And, not less important, it allows learners to customize the conversation of the chatbot during the learning experience. There are two very original aspects: i) conversation design is totally independent of the content that the chatbot will deliver; ii) a great deal of the conversation features is controlled via configuration rather than by programming; changing configuration data allows modifying the conversation at low (nearly zero) cost.

Following what we had said in the Section 1, the conversation that we envision for this chatbot should exhibit a number of relevant features:

- As close as possible to natural conversations (Moore and Arar, 2019).
- Proactive, in the sense that the chatbot leads the conversation.
- Responsive, in the sense that the chatbot properly reacts to user turns, that could be solicited (e.g. asking for feedback) or unsolicited.
- Customizable (and adaptive) in a number of ways.
- Designed by nontechnical actors. (i.e. authors, teachers, conversation experts, . . .)
- Be sustainable, both in terms of costs and time required.

In the following of this section we describe some of the components of TalkyTutor that allow to implement the above features: **Conversation Machine** (the strategy of the conversation of the chatbot), **Dialouge Categories**, **Turns** of the chatbot (what the chatbot says and when), **Wording, Templates, and variables** (the specific utterances of the chatbot and its formulation), **Intents** (what users may say and how it is interpreted), and **Loquacity** control (controlling how often the chatbots speaks and the length of its utterances).

3.1 Conversation Machine

Conversation Machine is the most complex piece of the machinery to generate a conversation between chatbot and human. This component in chatbot's architectures performs these tasks: i) it generates the conversation turns of the chatbot; ii) it understands the user turns of conversation, iii) it organizes the flow of turns, and iv) it calls other "engines" of the chatbot when needed (e.g. to fetch a new item of content) The conversation is modeled as a state machine. The turns of the chatbot corresponds to entering or leaving a state. All turns of the chatbots are made of messages from different categories.

The conversation machine, coupled with corresponding turns, defines the overall strategy of the conversation; a conversation machine defines a family of chatbots, in the sense that all the members of the family adopt a similar strategy of conversation (independently of the content being delivered).

A state machine is relatively complex; therefore, it needs a conversation designer to shape it. In addition, each state machine requires a new programming if they need more states than the default state machine, since proper hooking to various actions needs to be established (e.g., when to ask for a new piece of content). Once a state machine is created (defining a family of chatbots), several specific chatbots can be created, just via configuration with no programming.

3.2 Conversation Categories

As it was said above, the turns of the chatbot are classified into categories. Figure 2 shows an example of categories. Categories are not built-in by the technology; they are part of the configuration: They can be modified without programming, and are clearly dependent upon the design of the conversation machine. A category defines the overall semantic of a turn of the chatbot; "GREETINGS" for example means exchanging pleasantries (e.g. "nice to see you again"); "FORECAST" means to anticipate what will happen for completing the current learning path, etc. Classifying the turns of the chatbot into categories has a double purpose: helping designers to identify the need for a turn; and also helping them to put turns in

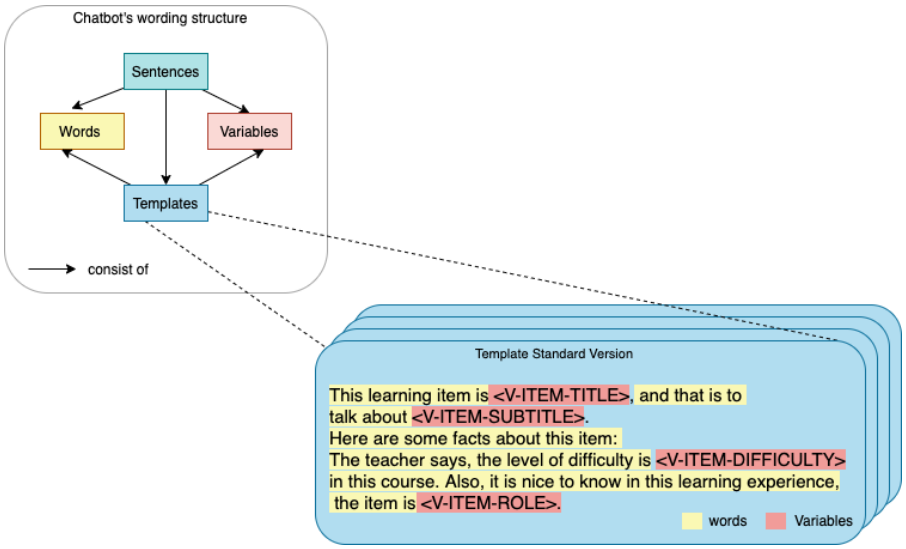

Figure 1: Structure of chatbot's wording and template example.

Acronym	Category	Description
GR	GREETINGS	Exchanging pleasanteries with the user
SU	SUPPORT	Helping the user with the chatbot
PR	PREVIEW	Describing incoming content
SM	SUMMARY	Summarizing what has been done
FC	FORECAST	Summarizing what needs to be done
AC	ACTION	For taking action (move on, break, suspend, stop)
FD	FEEDBACK	Getting the opinion of the use
RE	REINFORCEMENT	Gives a reinforcement to the user

Figure 2: An example categories for the turns of the chatbot.

the proper order. At the start of a session, for example, "GREETINGS" are at the beginning of a sequence; at the end of a session they should come at the end. These categories are attached to the states and arcs of state machine to describe the chatbot dialogues in the conversation with user. Also, these categories are used in the unsolicited turns to understand users and provide them the response. The chatbot for example could say "did you like the item?" ("FEEDBACK" category), and wait for the reply. Alternatively, the user (in a different turn), may say "I did not like it", which is interpreted as an unsolicited feedback.

3.3 Wording, Templates and Variables

An original and relevant feature of TalkyTutor is the possibility of easily controlling style and wording. Style define an overall way of speaking (for the chatbot); examples could be friendly, professional or empathic. The specific turns of the chatbot , as used by the conversation machine are defined via a "master table", that is used to debug and tune a specific family of chatbots. Designers, however, may specify what the chatbot actually says, by associating each turn in the master table with a specific sentence. Wording configuration tables allow designers to introduce their own wording for each specific turn, controlling both style and length of the turn. A relevant aspect of the state machine is that it is not dependent on the content; however, it can speak about individual items and the learning pathway currently being used. This is obtained by encoding templates (describing complex information) and variables (describing simple information). In the other word, each chatbot's category is consisted by one to many sentences, and each sentence contains several fixed wordings, templates and variables as shown in Figure 1. Templates and variables can be created by authors, without interfering with conversation design. Various versions of templates can be created (e.g., a standard, short or long length version). Figure 1 also shows an example of template for an item of content; it embeds some fixed words and simple variables derived from metadata associated to the content item. All chatbot's wording in the format of the fixed words, templates and variables are represented in tables.

We must remark again that style, wording and templates are controlled via configuration tables and therefore may be part of the customization by the various actors, without the need of programming.

3.4 Intent Recognition

The modular architecture brings the opportunity to use different intent detection services in the chatbot. Currently, TalkyTutor chatbots can select either an internal module for intent detection trained based on the classic BERT model developed by Google (Devlin et al., 2018) or inspecting external services such

as IBM Watson. In both cases, the chatbot only understands the range of defined intents in configuration data such as ”ask for help” or ”ask for summary”.

3.5 Loquacity Model

An additional issue is the loquacity of the chatbot: how many turns does it take? Should they be more or less? How verbose are these turns? It is clear that there is no predefined answer: in some situations, and for some users, very few and short turns can be suitable. In other situations, instead, more (longer) turns may be better. There is also another issue: some of the turns of chatbot are mandatory, in the sense that they can't be skipped in the conversation; other turns (e.g. a reinforcement message) could be skipped without spoiling the functionality. In order to cope with this, issue we have devised a way to control how many turns the chatbot takes. The turns defined in the conversation machine are the maximum: using trimming algorithms we can cut them down. The problem is that trimming has to be appropriate: e.g. i) never skipping mandatory turns; ii) making sure that no category is over represented; iii) making sure that no category is skipped forever. For the time being we designed a preliminary trimmed algorithm which is directly controllable via configuration; what can be controlled is the level of loquacity is directly related to the metadata of each wording's category (e.g, mandatory categories are not skippable in a different loquacity, but optional once may be skip on the different turn's of chatbot). The user decides if more or less turns are desired, and the algorithms do the job. Experimental testing (discussed in the next section) shows that users like to control the loquacity of their chatbot.

3.6 Customizing the Chatbot

The combination of the above can create very adaptive conversations.
The configuration of the conversation above described may require some conversation expertise but no programming at all. Redefining a state machine and corresponding Master Table (including the definition of transitions and rules to fire them) might be not easy. The amount of time could vary depending on the level of the modifications and it could take a few hours or few days. Modifying Style Tables and Alternatives is almost trivial. The current implementation of the above machinery has shown that we can control the conversational features of the chatbot by greatly reducing reprogramming time.

There is a final overall issue: who does what, in terms of design and configuration? Since most of the features discussed in this section have not direct counterpart in the literature, we did experiment with various possibilities, and the end we came up with rules shown in Figure 1, that seem reasonable for an education environment. These rules could be modified in the future, for different application realms.

In the next section we discuss as the above features were used in an empirical testing involving 12 teachers and 80 students (of school and higher education).

4 EXPERIMENTING WITH ADAPTIVE CONVERSATIONS

The approach and the technology described in the previous sections have been tested in this paper with two different bodies of content: “Advanced Computer Architecture” (ACA) in English and “La curtis” (medieval history) in Italian.

4.1 Experiment

The experimentation of evaluating the TalkyTutor runs over several months. Firstlt, we asked teachers were to use chatbots with two points of view: i) were they willing to adopt them for their students? ii) what could have been the reaction of their students? Overall, 12 teachers were involved, 7 from Higher education and 5 from schools. In terms of disciplines, 5 teachers were from humanities and 7 from STEM. A qualitative investigation was performed through structured interviews. Several changes were introduced based on this preliminary assessment (especially to the interface and adaptivity features).

Secondly, new experimentation was conducted with a focus group. The same group of teachers was asked to repeat their experience. Students (selected by their teachers in high school and invited by email for higher education) were involved with using the chatbot in their own environment (school or university classes). Overall, 81 students were asked to use the chatbot to simulate a learning session (30 minutes at least) and fill up a survey. A few students were also involved in a focus group for interviews. 33 students were from Higher education and 48 students from junior high school. Students were guided by the teachers that also provided local instructions.

For the experimentation a specific family of chatbot was created. Two different content were adapted from existing material: ACA, developed by a professor in higher education, and “La Curtis”, a course about Medieval History. Both courses were real, in

Table 1: Current rules for chatbot configuration.

Conversation machine:	This is created by conversation designers with the support of IT specialists. It is done only once for a family of chatbots.
Categories and turns:	They are created by experts in the conversation for education. They are defined only once for a family of chatbots.
Stule, wording and templates:	They are initially defined for a family of chatbots. They can be (optionally) modified by publishers, authors, and teachers.
User intents:	What users may say and how it is interpreted. They are initially created by designers but can easily expanded.
Loquacity of the chatbot:	It can be controlled dynamically by chatbot designers (teachers say they that they should also control it)

the sense that were created using text and slides of existing courses. "ACA" was delivered in English (with the chatbot speaking in English) to higher education students only; "La Curtis" was delivered in Italian (with chatbot speaking in Italian) to Higher Education students and Junior High school students. In the following of this section we briefly discuss i) the qualitative analysis of teacher reactions; ii) the quantitative analysis of the surveys filled by students; iii) a qualitative analysis of comments made by students.

Let us examine first the reactions of teachers. Several issues were discussed in the interviews, but we only present overall reactions and opinions about conversation adaptivity. Teachers, in general, were highly positive about having Talky tutor chatbots supporting learning by their students. Most teachers enjoyed their experience. Some teachers (higher education) found the conversation with the chatbot a little slow in pace. "In the beginning, I was suspicious; then I realized that it could work very well, especially for less motivated students" (school teachers).

Teachers did not sufficiently understand the need for conversational adaptivity at first. After a while, they started to like the idea. There is an interesting divide: i) school teachers think they should control style and loquacity for their students (adaptivity for individual needs); ii) higher education teachers think learners should be empowered.

To our surprise, wording customization is felt as significantly important by teachers; almost all of them declared that they would spend 2-3 hours to substitute standard sentences of the chatbot with their wording. There several motivations:

- Learners will recognize their teacher, improving their motivations to learn.
- The effect of presence (by the teacher) would be enhanced.
- The psychological impact will make the chatbot more persuasive.

"It will provide less feeling of speaking to a robot" (school teacher)

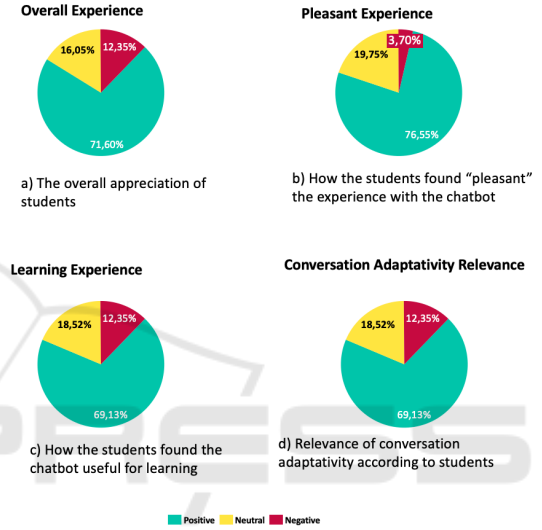

Figure 3: The of experiment with students.

Let us now discuss the analysis of the surveys filled by the students. Figure 3.a shows the overall appreciation of the experience with TalkyTutor chatbots. The majority (score 5 or 4) liked it; a minority did not like it (score 0 or 1). School students were slightly more positive than higher education students (by less than 2%).

Figure 3.b shows how pleasant was the use of the chatbot for the students. Again the vast majority found the experience pleasant (score 4 or 5) while a minority had a negative opinion (score 0 or 1). Again school students were slightly more positive. Figure 3.c shows how the students found the experience useful for learning. Figure 3.c demonstrates the role of adaptativity between students.

While compiling the surveys students could add comments. Table 2 shows a few comments by students. Some of them are from the (tiny) minority that do not appreciate chatbots.

Table 3 shows a few comments by students about the adaptivity of the conversation. A few comments were misplaced since they were referring to other

Table 2: Students about the idea of using TalkyTutor. **H**: **H**igher education; **S**: Junior High **S**chool.

No.	Level	Comment
1	H	It's cool to have a guide that helps you keeping track of learning and that can customize your learning experience, but in the worst case it's just a complex table of content, not that useful.
2	H	Being praised while learning it's a thing to not be underestimated in my opinion because it pushes the student to be more active and focused on the videos with respect to just watch them without any kind of interaction like in simply playing a video playlist.
3	H	It was fun to talk with the chatbot.
4	S	1It will help students with problems to how study a course in different ways.
5	S	It is an important alternative to traditional learning; I hope they will be introduced very soon at school.
6	S	Besides content, the chatbot should propose quizzes.
7	H	I prefer human relationships.
8	H	I don't like chatbots in general.

problems. Overall, we may say that teachers and students liked learning with chatbots and found the experience pleasant and (potentially) useful for learning. We are aware that the perception after 30 minutes could be different from actual usage over an extended period of time.

5 CONCLUSIONS AND FUTURE WORK

Let us summarize the most relevant contributions of this paper; first of all, we put forward the idea of making chatbot customizable in order to be tuned to the user profile and the context; this is specifically relevant when chatbots are used for a learning experience, where personalization is a great relevance (Cai et al., 2021). Next, we propose a specific role for chatbots in education: tutors leading users across content. Finally, we advocate the need to streamline the production process with two main goals: i) reducing costs and effort for deployment; ii) empowering non-technical actors to direct control the features of the chatbot.

Table 3: Students about the Conversation Adaptativity. **H**: **H**igher education; **S**: Junior High **S**chool.

No.	Level	Comment
1	H	I tried different modes and levels of loquacity and I appreciated a lot that we can personalize it based on our tastes.
2	S	It helps to make the chatbot an alternative to traditional learning.
3	S	It is important, otherwise, the chatbot could become boring and repetitive.
4	S	It can help to make chatbot speaks as human beings speak.
5	H	It is important and it should be further developed.
6	H	It is important to be able to speed up the interaction.
7	S	It is important to help to understand what the chatbot says.
8	H	Differences among styles could be stronger. Empathic style should be improved.

In order to make the above real, we have developed an original technology, where chatbots are shaped via extensive use of configuration data; somehow, we have developed a chatbot generator. We envision the production of a chatbot into the following steps:

(a) Create a family of chatbots, sharing a common conversation strategy.

(b) Instantiate a specific chatbot and plugging the content (adequately organized and with proper metadata).

(c) Customize the conversation using configuration data, controlling the turns of the chatbot, the style and the wording, the loquacity, etc.

It should be noted that steps "b" and "c" could be interchanged, and repeated several times. In addition, experimentation with teachers and students has shown that chatbots as tutors can be effective and that customization of conversation is perceived as important.

We currently developed a platform for building TalkyTutor chatbots and attaching the learning materials. The platform is standalone for now, but it should be delivered as a service to be more scalable in the future. From an application point of view, it is also essential to define actors' role in customization: who does control of what? Authors, publishers, teachers, and students find customization (of content and conversation) valuable, but they have different ideas

about who does what. Teachers at school, for example, would like to constrain conversation customization possibilities for their students tightly; they want to take basic choices, while students might have a different idea.

From the technical side, two issues are preeminent: i) to add wider choices of interfaces, including vocal ones and integration with Alexa style frameworks; ii) to further improve the production process, making it easier for non-technical actors to work on customization.

As a research issue, we are investigating how to move from customization to adaptation. There is a component in our architecture responsible for decision-making along with the conversation; at the moment, it takes the simple decisions about suspending or stopping the session rather than keep going. As a real tutor, it should take more important decisions: the content should decide if an item needs to be repeated or if the current pathway is appropriate or should be changed; for the conversation.

ACKNOWLEDGEMENTS

This work has been partially supported by a grant of EIT Digital and IBM Italy.

REFERENCES

Brandtzæg, P. B. and Følstad, A. (2017). Why people use chatbots. In Kompatsiaris, I., Cave, J., Satsiou, A., Carle, G., Passani, A., Kontopoulos, E., Diplaris, S., and McMillan, D., editors, *Internet Science - 4th International Conference, INSCI 2017, Thessaloniki, Greece, November 22-24, 2017, Proceedings*, volume 10673 of *Lecture Notes in Computer Science*, pages 377–392. Springer.

Burkhard, M., Seufert, S., and Guggemos, J. (2021). Relative strengths of teachers and smart machines: Towards an augmented task sharing. In *CSEDU (1)*, pages 73–83.

Cai, W., Grossman, J., Lin, Z. J., Sheng, H., Wei, J. T.-Z., Williams, J. J., and Goel, S. (2021). Bandit algorithms to personalize educational chatbots. *Machine Learning*, pages 1–30.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Duggirala, V. D., Butler, R. S., and Kashani, F. B. (2021). ita: A digital teaching assistant. In *CSEDU (2)*, pages 274–281.

Graesser, A. C., VanLehn, K., Rose, C. P., Jordan, P. W., and Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI Magazine*, 22(4):39.

Han, S. and Lee, M. K. (2022). Faq chatbot and inclusive learning in massive open online courses. *Computers & Education*, 179:104395.

Heller, B., Proctor, M., Mah, D., Jewell, L., and Cheung, B. (2005). Freudbot: An investigation of chatbot technology in distance education. In *EdMedia+ Innovate Learning*, pages 3913–3918. Association for the Advancement of Computing in Education (AACE).

Hwang, G.-J. and Chang, C.-Y. (2021). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*, 0(0):1–14.

Kerry, A., Ellis, R., and Bull, S. (2008). Conversational agents in e-learning. In *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pages 169–182. Springer.

McTear, M. (2018). Conversation modelling for chatbots: current approaches and future directions. In Berton, A., Haiber, U., and Minker, W., editors, *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung 2018*, pages 175–185. TUDpress, Dresden.

Moore, R. J. and Arar, R. (2019). *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*. Association for Computing Machinery, New York, NY, USA.

Paek, T. and Pieraccini, R. (2008). Automating spoken dialogue management design using machine learning: An industry perspective. *Speech communication*, 50(8-9):716–729.

Rooein, D., Bianchini, D., Leotta, F., Mecella, M., Paolini, P., and Pernici, B. (2020). Chatting about processes in digital factories: A model-based approach. In *Enterprise, Business-Process and Information Systems Modeling*, pages 70–84. Springer.

Serban, I., Sordoni, A., Bengio, Y., Courville, A., and Pineau, J. (2016). Building end-to-end dialogue systems using generative hierarchical neural network models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 30(1).

Song, D., Oh, E. Y., and Rice, M. (2017). Interacting with a conversational agent system for educational purposes in online courses. In *2017 10th international conference on human system interactions (HSI)*, pages 78–82. IEEE.

Tegos, S., Demetriadis, S., Psathas, G., and Tsiatsos, T. (2019). A configurable agent to advance peers' productive dialogue in moocs. In *International Workshop on Chatbot Research and Design*, pages 245–259. Springer.

von Wolff, R. M., Nörtemann, J., Hobert, S., and Schumann, M. (2019). Chatbots for the information acquisition at universities–a student's view on the application area. In *International Workshop on Chatbot Research and Design*, pages 231–244. Springer.

Winkler, R., Hobert, S., and Appius, M. (2020). Empowering educators to build their own conversational agents in online education.

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., and Drachsler, H. (2021). Are we there yet?-a systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence*, 4.