Data-Driven Edu Chatbots

Donya Rooein
donya.rooein@polimi.it
Politecnico di Milano
Milan, Italy

ABSTRACT

There are many online courses and contents on the web, so each learner can find the best one and choose it. However, sometimes many online courses are poorly accessible due to the limits of the search engines on the web. The advent of intelligent systems, and online Chatbots, in particular, has brought improvement in various fields. Education Chatbots improve communication, increase productivity, and simplify learning interaction. This study aims to provide an intelligent Edu Chatbot with a high level of customization to learners who have different needs. This way, they can find their personalized learning path dynamically and their customized content without too much time and effort. This is precisely what e-learning needs because of the enormous amount of material on the web.

CCS CONCEPTS

• **Information systems → Specialized information retrieval**;
• **Applied computing → E-learning**.

KEYWORDS

Chatbots, E-learning, Artificial Intelligence, Interactive Tutoring Systems, Adaptive Learning

ACM Reference Format:
Donya Rooein. 2019. Data-Driven Edu Chatbots. In *Companion Proceedings of the 2019 World Wide Web Conference (WWW '19 Companion), May 13–17, 2019, San Francisco, CA, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3308560.3314191

1 INTRODUCTION

In recent years, digital technology has been evolving in various applications, and the interest in research on intelligent systems has become prominent. Design and implementation of Chatbots are one of most popular kinds of digital technologies. A Chatbot is a computer software that is able to answer some users' questions in a session [5]. The word "intelligent" before Chatbot in an interaction specifies the need for intelligent replies rather than just random ones [6]. Chatbots can be used in various applications such as education, commerce, medicine, and entertainment. Among the big IT companies that have developed Chatbots for industry and research can be named Apple Siri, Microsoft Cortana, Facebook M, and IBM Watson [5]

WWW '19 Companion, May 13–17, 2019, San Francisco, CA, USA

ACM ISBN 978-1-4503-6675-5/19/05.
https://doi.org/10.1145/3308560.3314191

Building Chatbots in education (Edu Chatbots) that can interact with learners through natural language like in other applications, had great importance and attracted the attention of researchers in recent years. The purpose of this research is to design and implement an educational framework to provide different educational pathways for learners independently from the content of the subject. The goal of this chatbot is to customize the learning pathway for the learners according to the many attributes . These attributes are driven by profiling the learners and their desires of learning.

In the traditional way of learning, instructors teach content despite the learners' background and goals for attending the course. In Online classes, while students can choose their proper course, they still can not go through a course in different learning pathways in different situations. There is room for variety concerning how an online course should offer diverse learning pathways.

In this study, we hope to find a learning framework for learners through Chatbots to enhance personalization in learning. The main focus of this research is modeling learner interaction with Chatbot and content modeling. The goal of this study is to drive the best learning pathway among various learning choices and bring a high level of flexibility and customization to e-learning for users, according to their goals and desires. There are some studies on the modular architecture for Chatbots which allow them to stay the same while swapping the teaching content [1]. Through this approach, we deploy the Chatbot in several ways at a lower cost. Although, this approach has never been developed on real-world problems. In this PhD thesis, we are going to look for solutions in this regard.

2 STATE OF THE ART

There are many studies related to e-learning system that shed some light on the traditional education system drawbacks [2]. Intelligent tutoring systems [1] are computer-based tutoring systems that tailored to the needs of individuals learners [4]. There is a study on possible effects of a conversational agent in an online portfolio in communication and interaction with a user. The collected data during 4 weeks demonstrated the lack of user satisfaction in completing class activities. However, they had positive feedback on virtual assistant as a social companion. The research introduced three main issues for further research: 1) Learner-developed conversational agents, 2) Improve the agents "intelligence" regarding the domain of knowledge, and 3) Further developed virtual characters that can meet users' humanistic and utilitarian expectations [7]. Rossi et al. in 2011 used a Chatbot in an open source learning management system. The system has been used to support online teachers work by developing a specific knowledge base to pedagogical aspects. The mentioned Chatbot could answer more difficult questions better [10].

[1] ITSs

Wong J.S. et al. used cognitive learning terms in online course materials to study the learning process [11]. In 2018, Clarizia F. et al. presented an educational Chatbot modeling in order to manage best answers to students [5]. They designed this model by using natural language processing techniques to give answers to students.

Nenkov et al. introduced a learning management system to model an intermediary between a student and a teacher [9]. In this study, an intelligent Chatbot was used to automate the interaction between student and teacher implemented using IBM Bluemix platform with IBM Watson technology. IBM Watson is a cognitive system that interacts between users and computers that improve human expertise services by using intelligent techniques such as natural language processing, Analytic, and Machine learning in the form of Chatbots [8].

In a recent study on highly personalized Chatbots in education Casola et al. designed an innovative methodology to provide online courses which are close to the learners‘ desires [3]. They called this methodology iMOOC that has helped the learner to find her learning pathway through the material of the online course by considering her background and profiling attributes. The iMOOC is a content structuring for adaptive education. In this study learners‘ features such as background, effort, needs, depth, and type of interest are taken into account to get the most suitable individualized learning experience. In addition, this structure arranged and organized each piece of contents most properly. To address these points and make adaptive education feasible in iMOOC, they determined items in content and tag them to reflect their properties and the subject of content [1]. Some of the key steps to make education content well adaptive are:

(a) itemized items in small chunks
(b) item meta tagging
(c) content tagging

This methodology by combining these steps can offer learner‘s own learning experience.

Politecnico di Milano [2] and IBM [3] as a joint effort have developed innovative technology in data driven Chatbots research area that named iCHAT [1]. In order to demonstrate the interaction between the user and Chatbot, iCHAT has four components. Conversation routes dynamically according to the information flow that shown in the Figure 1. First, each verified user can launch a stream of conversation with Chatbot through iCHAT interface. It is possible for a user to start several streams of conversation during one session. At the end of the session, all streams can be saved with iCHAT Long Term Memory to be resumed in the next session.

3 PROPOSED APPROACH

Due to the popularity of Chatbots in recent years, many studies have investigated the various aspects of this technology. With respect to current studies on Edu Chatbot, this research aims to study potential ways to have a more intelligent Chatbot in educational purpose. A number of motivations exist for relying on this research to design and develop an intelligent Chatbot through better users‘ understanding. According to previous studies on Online Open Courses

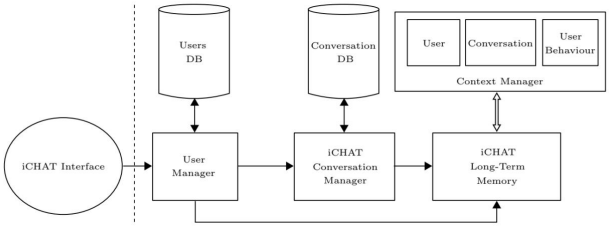

Figure 1: iCHAT‘s overall architecture [1]

(MOOCs), poor rate of completion of an online course is a fundamental challenge [3]. The idea of a model of learners‘ behaviors in the learning process that can identify their learning needs to simplify learning path is to investigate in this study.

This study has concentrated more on the question of understanding the proper learning pathways for each learner, whereas intelligent Chatbot must both understand and generate learning pathways in different conversational streams. Concentrated Study to offer a framework in this regard and implement as an intelligent Chatbot to provide Personalizing to learners is one of the research goals here. Analyzing users‘ data in the learning process is another open challenge in this study to provide a learning pathway dynamically based on the users‘ feedback. There is currently no possibility that you can change your learning pathway of an online course during the course. This feature brings high customization during an online course to the learner and provides her for adjusting the amount of content based on the situation.

Since the goal of learners might be affected by divergent constraints, our most general motivations arise from the fact that it is essential to provide a flexible learning pathway in accordance with these constraints to reduce the probability of not completing the course due to the lack of coordination of the contents of the courses with the student’s needs. There are some features which can enhance Chatbot‘s capabilities to communicate more effectively with users. For instance, we can mention them as an animated gesture supporting knowledge base, provide content from other knowledge sources, and multiple conversational strategies. Besides, by using technologies such as linguistic tools Chatbot enhance the language capabilities to be more like humans.

Another key aspect of this research will be the automation of the possible process in this methodology that is not met in the current process. The quality of the courses and overloads to tag them should be automated as much as possible to save time and accuracy of features related to them. So, this is necessary to model contents properly in this issue. In this respect, content modeling plays an important role in making the Chatbot intelligent. The framework in which the content is correctly modeled increases the accuracy of the Chatbot‘s responses to the user. The general approach for modeling this study is in Figure 2. The novelty concerning current approaches lies in the facet that Chatbot rather than being an expert on the subject is an expert on a body of content covering a subject. In other words, the Chatbot is tasked with efficiently and emphatically interacting with the learner.

[2]HOC-LAB
[3]The Italian research division

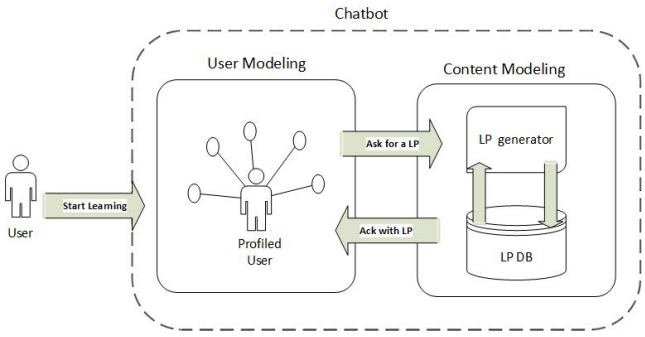

Figure 2: General overview of purposed framework for Chatbot

4 METHODOLOGY

We are currently developing a conversational framework that integrates the real-time aspect of the traditional education system between students and teachers and drive a customized learning pathway from user's profiling. We believe that such a system provides a robust platform on which to continue the development of online education with Chatbots in both academic and industry contents.

The primary research method for this study is to investigate real-world problems in both academic and industry and survey on literature review and conceptual modeling for Chatbots, with the specific focus on Edu Chatbots. A good depth research background is essential to find challenges and research gaps in the problem statement.

In the second stage of this study, there are two sides of the problem statement which should investigate. On the learner side, user profiling identification and classification through a structured approach is at a very beginning stage toward a user modeling. This study will review various types of user intent modeling and their characteristic. In this stage, the iChat architecture will evolve to consider as a part of a framework. On the other side, a study on content modeling is essential to find an efficient structured data on the content of the online courses. A classification method will develop to categorize constraint factors. For the purpose of design learning pathways and modeling, a semantic extraction approach in the industry is essential. Design framework will identify based on a comprehensive review of current industry practices and academic researches.

Third, we should validate our design through design validation approaches such as prototype and get assistance from expert opinions and apply our design in a real context to see the results. There is currently a prototype of iCHAT technology and this work will developing iCHAT to improve quality and adaptive learning in the context of both academic and industry.

Forth, develop and implement the Chatbot into the real online courses. Once our Chatbot is ready, the analytics begin with the collection of data. The data collection is a critical step in providing information due to answer the research questions.

Finally, all the effort and resource dedicated to the previous stages of methodology culminate in the ultimate step. With data, we can analyze whether this approach answers the research questions or not. Evaluate the implementation is in this step. It is possible to measure the rate of the online course completion and compare it with traditional e-learning techniques to evaluate this Edu Chatbot. A quantified approach to measure user engagement is the average time that she spent in a defined period on Chatbot.

During the first year of Ph.D. candidature, the project plan is:

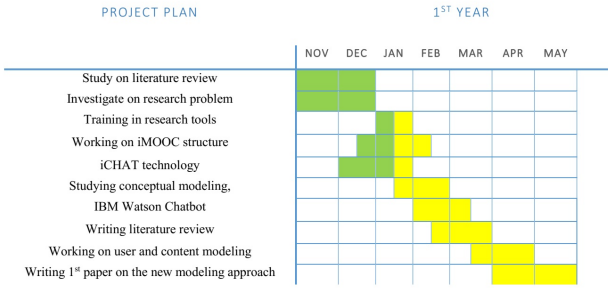

Figure 3: Project Plan

5 RESULTS

This research aims to make possibility into reality in Edu Chatbots context. Edu Chatbot can be used as a tutor, a student evaluator, for a question and answering or to communicate with a teacher. The possible result of this research might be to implement the Chatbot what can easy adapt to the current subject or course so that content manager could do without the requirement of development skills. Moreover, the ultimate Chatbot might be able to bring customization in a personal way by profiling learners. The Chatbot capabilities can expand by including it in other systems such as e-learning systems, a virtual environment or library system or by adding augmenting technology like text-to-speech technology, linguistic tools or animation.

Up to now, after passing two months of my doctoral thesis and studying on this topic, the necessary information has been collected from previous studies. By using this information and the insight of current technologies, the main question has identified for this research and how the outcome of this research can be practical and contribute Chatbot in the context of education. After being able to understand the user status as a set of features. The next introduced feature is the topology navigation. It is essential to understand the user's will, on the other hand, is important to provide a learning pathway that can adapt to those needs. To evaluate one learning pathway which has assigned to a learner, we can consider user's preference of content type, as well as feedback of what a learner did in the previous session such as the difficulty of the last exercise. Then use that information as a prior to create next learning pathway.

6 CONCLUSIONS AND FUTURE WORK

The study aims to further develope the educational Chatbots by reviewing previous works and working on current iCHAT technology and implementing it in the real world application. This research

will be done by finding out the place of Chatbots in education and complete Chatbot requirements in the education context to add pedagogical value in this research area.

A separate line of future research might be to look more at what the Chatbot could bring to the virtual world and if this could become a valuable pedagogical tool. Chatbot with the ability to individualize a service might go beyond in just education context. This idea examines the generalization of this Chatbot to the diverse application such as entertainment and treatment.

As one of the most important motivations for implementing Data Driven Edu Chatbot comes from using current iCHAT technology and iMOOC structure to design and implement a general framework in education and evaluate the efficiency of those purposed approaches.

In the end, we can categorize the current challenges to implement this Edu Chatbot in three major categories as follows:

- Drive important intents to implement in Chatbot as many possible capabilities which can support them (User modeling).
- Finer control on the processing of a sentence. It is essential to measure adjectives in a sentence for adaptive learning by mapping sentences into the specific entities (Content modeling).
- It should be considered in Chatbot to adjust with different styles. Talking in many different styles for Chatbots can be a plentiful challenge if the user interacts with the Chatbot with a small sentence. So, it is better to expand the Chatbot dialog tree in a more human way (Adaptive approach).
- Despite limitations in interacting with the interface, tools can be built to bridge the production gap and put our Chatbot into the efficient way which can reduce time and cost (efficiency).

7 ACKNOWLEDGMENTS

I would like to express my special appreciation and thanks to my advisors Professor Dr. Barbara Pernici and Dr. Paolo Paolini, you have been a tremendous mentor for me. I would like to thank you for encouraging me to start my Ph.D. in this research field and for allowing me to grow as a research scientist.

REFERENCES

[1] Damla Ezgi Akcora, Andrea Belli, Marina Berardi, Stella Casola, Nicoletta Di Blas, Stefano Falletta, Alessandro Faraotti, Luca Lodi, Daniela Nossa Diaz, Paolo Paolini, Fabrizio Renzi, and Filippo Vannella. 2018. Conversational Support for Education. In *Artificial Intelligence in Education*, Carolyn Penstein Rosé, Roberto Martínez-Maldonado, H. Ulrich Hoppe, Rose Luckin, Manolis Mavrikis, Kaska Porayska-Pomsta, Bruce McLaren, and Benedict du Boulay (Eds.). Springer International Publishing, Cham, 14–19.

[2] Xabier Basogain, Miguel Ángel Olabe, and Juan Carlos Olabe. 2017. Transition to a Modern Education System through e-Learning. In *Proceedings of the 2017 International Conference on Education and E-Learning*. ACM, 41–46.

[3] Stella Casola, Nicoletta Di Blas, Paolo Paolini, and Giuseppe Pelagatti. 2018. Designing and delivering MOOCs that fit all sizes. In *Society for Information Technology & Teacher Education International Conference*. Association for the Advancement of Computing in Education (AACE), 110–117.

[4] Sanket Dinesh Choksey. 2004. Developing an affordable authoring tool for intelligent tutoring systems. (2004).

[5] Fabio Clarizia, Francesco Colace, Marco Lombardi, Francesco Pascale, and Domenico Santaniello. 2018. Chatbot: An Education Support System for Student. In *International Symposium on Cyberspace Safety and Security*. Springer, 291–302.

[6] O. V. Deryugina. 2010. Chatterbots. *Scientific and Technical Information Processing* 37, 2 (01 Apr 2010), 143–147. https://doi.org/10.3103/S0147688210020097

[7] Aaron Doering, George Veletsianos, and Theano Yerasimou. 2008. Conversational agents and their longitudinal affordances on communication and interaction. *Journal of Interactive Learning Research* 19, 2 (2008), 251–270.

[8] Judith S Hurwitz, Marcia Kaufman, and Adrian Bowles. 2015. *Cognitive computing and big data analytics*. John Wiley & Sons.

[9] Nayden Nenkov, George Dimitrov, Yuriy Dyachenko, and Katerina Koeva. 2016. Artificial intelligence technologies for personnel learning management systems. In *Intelligent Systems (IS), 2016 IEEE 8th International Conference on*. IEEE, 189–195.

[10] Pier Giuseppe Rossi and Simone Carletti. 2011. MAPIT: a pedagogical-relational ITS. *Procedia Computer Science* 3 (2011), 820–826.

[11] Jian-Syuan Wong, Bart Pursel, Anna Divinsky, and Bernard J Jansen. 2016. An analysis of cognitive learning context in MOOC forum messages. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 1315–1321.