



A Review of Conversational Agents in Education

Carlos Rodrigues^{1,2}, Arsénio Reis^{1,3} , Rodrigo Pereira¹, Paulo Martins^{1,3} ,
José Sousa^{1,3} , and Tiago Pinto^{1,3}  

¹ Universidade de Trás-Os-Montes E Alto Douro, Vila Real, Portugal
{a175393,a168798}@alunos.utad.pt, {ars,pmartins,jmsousa,
tiagopinto}@utad.pt

² Universidade Aberta, Aberta, Portugal

³ INESC-TEC, Vila Real, Portugal

Abstract. The use of mobile conversations is increasing all around the world. A conversational agent (CA) is mostly useful due to the fast response times and their simple nature. Recently, we have seen the development and increasing use of dialog systems on the Web. A conversational agent (CA) is a system capable of conversing with a user in natural language, in a way that it simulates a human dialog. Examples of CA can be found in several areas, including healthcare, entertainment, business, and education. In this paper a state of the art review of these dialog systems is presented, comprising different categories, different approaches and trends. The purpose of this work is to identify and compare the main existing approaches for building CA, categorizing them and highlighting the main strengths and weaknesses. Furthermore, it seeks to contextualize their use in an educational context and to discover the issues related to this task that may help in the choice of future investigations in the area of conversational natural language processing in educational context.

Keywords: Conversational agents · Education · Natural language processing

1 Introduction

Artificial Intelligence (AI) applied in education is expanding quickly. Some of the most popular AI technologies, like Conversational Agents are being used to support teaching and learning activities in the classroom or at home [1].

Conversational agents (CA) or dialog systems, also called chatbots or chatterbots, have become increasingly common. Language-based HMIs, like virtual assistants or chatbots, provide information without time-consuming queries. Moreover, they hide the complexity and size of the information behind [2]. Applications of chatbots range from personal assistants on cell phones, sales bots on e-commerce websites, information retrieval, helpdesk, customer support and digital assistants, teaching-learning process support, and others. These systems are intended to carry out coherent conversations with humans in text or speech or both, in natural language. The creation of chatbots dates back to the ELIZA conversational system, which emulated a psychotherapist [3]. Over the years, new Artificial Intelligence (AI) techniques have been applied to the construction

of these agents, so that examples of systems can be broadly categorized into three paradigms or “generations”: the first, based on the combination of patterns and grammar rules; the second, grounded in production rules and artificial neural networks; and the third, which makes use of AIML markup languages [4]. However, the development of intelligent conversation with agents is still an unsolved research problem that raises many challenges in the artificial intelligence community [5]. This paper aims to identify and compare the main existing approaches to build chatbots, categorize them, compare them and highlight the main strengths and weaknesses. It also seeks to contextualize their use in an educational context. The main goal is to discover the issues related to this task that may help in choosing future research in the area of conversational NLP in an educational context.

2 Categories of Conversational Agents

CA can be categorized according to different characteristics, such as the interaction type, the domain of application, its purpose and the response generation models. The considered characteristics can consider the main learning strategy of the CAs and the contextualization capabilities of the model. In general we can classify CAs, based on different aspects [6]:

- Mode of interaction
- Goals
- Design approach
- Knowledge domain
- Regardless of how the Response Generation is done, these Chatbots share the same basis: analyze what the user says, interpret that analysis, and finally provide a response.

3 Approaches in the Implementation of Conversational Agents

This section discusses how CA can be developed, highlighting rule-based CA and AI-based CA. In AI-based CA, a distinction will also be made between information retrieval CA and generative CA. The pros and cons of each approach are also discussed. It should be noted that it is possible to use combinations of different Models in order to produce as optimal results as possible.

3.1 Rule-Based

Initial CA were based in rules. These approaches are generally simpler, but less broad in scope due to the lack of capabilities in responding to difficult questions [7]. Rule-based CAs reply to queries through pattern matching. In this way, they are insensitive and unable to adapt to unknown patterns. In addition, pattern-matching rules can be difficult and time-consuming to produce and maintain. Pattern matching rules are specific to a domain, and not easily transferable among different contexts [7].

3.2 AI-Based

Unlike rule-based models, AI based approaches usually rely on ML models and extract information by learning from previous knowledge of through interaction with humans. In order to accomplish such task, it is required to train with an ML algorithm that can learn a model based on training samples. Using ML algorithms removes the need to manually outline and code new sample matching rules, making chatbots greater and much less depending on a specific domain knowledge. [7]. These models can be subcategorized into models based on Retrieval Information and Generator models.

Information Retrieval (IR) Based

Having a dataset of Question-Response(Q-R) pairs, the IR-based model will search the Q-R dataset for the pair (Q',R') that best matches Q and returns R as the answer to Q [5]. Through this process, it enables reflecting training samples. Many search baseline models have been proposed to accomplish this purpose. [8]. Various works have addressed *Term Frequency-Inverse Document Frequency* (TF-IDF) retrieval models as a way to create CAs. For example, in [9] this approach is used to create a model directed to customer assistance and suggestion of products. Authors propose the application of Rhetorical Structure Theory [10] as a way to represent the characterization of connections among different replies. Among the used open domain datasets that have been most widely used to create the dialogue systems for generalist IR-based chatbots are *WikiAnswers*, *Yahoo Answers*, and *Twitter conversations* [7].

Generator Model-Based

Generator templates, create new answers for sentences according to the human interaction. Completely new sentences can be generated to respond different queries. Accomplishing this requires such models to learn how identify text structure and syntax, which is a difficult task. Consequently, results may lack consistency and even elegance in the generated texts [11]. Generators are usually based on sentences drawn from conversations. The algorithm learns from the data it is given. Its goal is to enable algorithms to generate good, linguistically correct answers based on input texts. Such models are generally based on deep learning (DL) algorithms that consist of encoder/decoders. [12].

Standard Models

Sequence-to-Sequence (Seq-to-Seq) models are the standard for chatbot modeling [7]. These models are fit for machine language problems; however, they also present good performance in natural language creation. The typical approach is using encoders and decoders [12]. This type of approach has several advantages. It is able to learn from data of different natures, domains and contexts, i.e. different domains, rather than one specific domain. This model does not require domain-specific knowledge to yield valuable results, but can be adapted to work with other algorithms if domain-specific knowledge needs to be further incorporated. Hence becoming a straightforward, but dynamic model, which may applied to very distinct PLN problems [13]. However, the main problem is that the size of the contextual information is restricted to a single vector, which means that when the size of the input text increases, there is a much higher chance that information,

possibly relevant, will be lost. As a consequence, sequence models under-perform when analysing long sentences and often generate confounding responses. Additionally, Seq-to-Seq models address single response at each time, hence often outputting inconsistent conversational order. [11].

Transformers

Transformers are the new trend in automatic/intelligent language models [14]. Transformers learn how to measure the importance of different pieces of data/text. They also support training parallelism, which allows dealing with much bigger pieces of data than before. These models have given birth to some of the most famous pre-trained systems such as BERT (Bidirectional Encoder Representations of transformers) [15] and GPT (transformer pre-trained generator). These models have created, and have evolved, using large language datasets, such as the Wikipedia and Common Crawl corpuses. However, they can still be refined for ad-hoc problems [16]. Other models were developed to address specific challenges, e.g. Reformer [17] and Transformer XL [18].

4 Evaluation Methods

A variety of CA evaluation methods have been used (Table 1). These usually follow the ISO 9214 usability guidelines [19]. The most popular methods for evaluating CAs are those based on efficiency. Other methods used are those based on satisfaction and effectiveness [20].

Table 1. Methods for evaluating chatbot against ISO 9214 [20]

ISO	Effectiveness	Efficiency	Satisfaction
4.1	Performance	Functionality Humanity	Affection, Ethics, Behavior, Accessibility
4.2	Content Evaluation	Functionality evaluation	User satisfaction
4.3	IR Perspective	Language Perspective AI Perspective	User Perspective
4.4	Smart Conversation	Functionality	Chatbot Interface Chatbot Personality
4.5	Domain Coverage	Coherence Conversation breadth Engagement Depth of conversation	conversational UX Engagement

5 Conversational Agents in Education

Many works can be found on CAs applied in teaching and learning [21, 22], assessment [23], administrative service delivery [24], consulting [25] or research and development [26].

The main advantages of using CA in education include [27]: content delivery, for example the ability for teachers/tutors to provide information in an online platform; quick and easy access, stimulus and engagement of learners. CAs in education also allows providing instant support during individual learning by supporting learners to facilitate activities e.g. delivering homework and evaluations [1], replying e-mails [28], adaptable to students' actions and emotions [29], and fast responses to their queries [30]. Future paths for CA research in aspects related to education include the development of ethical and functionality principles and usability testing. This denotes that the framework for chatbot development and implementation as well as design and content functionality needs to be improved. [27].

6 Conclusions

The development and use of conversational agents is increasing rapidly in multiple application domains. These agents are emerging in the form of virtual assistants, chatbots and other language-based interfaces, interacting with humans as digital assistants, sales bots, customer supporter, among many others. This paper has analysed how these systems are able to carry out coherent conversations with humans in text or speech or both, using natural language. For this purpose, while focusing on the application of conversational agents in education, the paper has identified several of the most promising approaches for the implementation of conversational agents, has reviewed how the performance evaluation takes place, and identified some relevant paths for future research and development, which include the development of functionality and ethical principles in chatbots, and the improvement of usability testing.

Acknowledgment. This work was supported by the RD Project “Continental Factory of Future, (CONTINENTAL FoF) / POCI-01-0247-FEDER-047512”, financed by the European Regional Development Fund(ERDF), through the Program “Programa Operacional Competitividade e Internacionalização (POCI)/PORTUGAL 2020”, under the management of aicep Portugal Global – Trade Investment Agency.

References

1. Okonkwo, C.W., Ade-Ibijola, A.: Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence* **2**, 100033. ISSN: 2666-920X. <https://www.sciencedirect.com/science/article/pii/S2666920X21000278> (2021)
2. Ondáš, S., Pleva, M., Hládek, D.: How chatbots can be involved in the education process. In: 2019 17th International Conference on Emerging eLearning Technologies and Applications (ICETA), pp. 575–580 (2019)
3. Weizenbaum, J.: On-line user languages. *BIT Numer. Math.* **6**, 58–65 (1966)
4. Sgobbi, F.S., Nunes, F.B., Bos, A.S., Bernardi, G., Tarouco, L.M.R.: Interação com artefatos e personagens artificiais em mundos virtuais in *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)* **25**, 642 (2014)
5. Mnasri, M.: Recent advances in conversational NLP: towards the standardization of Chatbot building. arXiv preprint [arXiv:1903.09025](https://arxiv.org/abs/1903.09025) (2019)

6. Hussain, S., Sianaki, O., Ababneh, N.: 946–956 (Mar. 2019). ISBN: 978-3-319-98284-7
7. Caldarini, G., Jaf, S., McGarry, K.: A literature survey of recent advances in Chatbots. *Information* **13**, 41 (2022)
8. Banchs, R. E., Li, H.: IRIS: a chat-oriented dialogue system based on the vector space model. In: *Proceedings of the ACL 2012 System Demonstrations*, pp. 37–42 (2012)
9. Galitsky, B., Ilvovsky, D.: On a chatbot conducting virtual dialogues. In: *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pp. 2925–2928 (2019)
10. Mann, W.C., Thompson, S.A.: Rhetorical structure theory: toward a functional theory of text organization. *Text-interdisciplinary Journal for the Study of Discourse* **8**, 243–281 (1988)
11. Sojasingarayar, A.: Seq2seq ai chatbot with attention mechanism. arXiv preprint [arXiv:2006.02767](https://arxiv.org/abs/2006.02767) (2020)
12. Vinyals, O., Le, Q.: A neural conversational model. arXiv preprint [arXiv:1506.05869](https://arxiv.org/abs/1506.05869) (2015)
13. Shum, H.-Y., He, X.-D., Li, D.: From Eliza to XiaoIce: challenges and opportunities with social chatbots. *Front. Inf. Technol. Electron. Eng.* **19**(1), 10–26 (2018). <https://doi.org/10.1631/FITEE.1700826>
14. Vaswani, A.: et al., Attention is all you need. *Adva. Neural Inf. Process. Syst.* **30** (2017)
15. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: Bert: pre-training of deep bidirectional transformers for language understanding. arXiv preprint [arXiv:1810.04805](https://arxiv.org/abs/1810.04805) (2018)
16. Acheampong, F.A., Nunoo-Mensah, H., Chen, W.: Transformer models for textbased emotion detection: a review of BERT-based approaches. *Artif. Intell. Rev.* **54**, 5789–5829 (2021)
17. Kitaev, N., Kaiser, Ł., Levskaya, A.: Reformer: the efficient transformer. arXiv preprint [arXiv:2001.04451](https://arxiv.org/abs/2001.04451) (2020)
18. Dai, Z., et al.: Transformer-xl: attentive language models beyond a fixed-length context. arXiv preprint [arXiv:1901.02860](https://arxiv.org/abs/1901.02860) (2019)
19. Abran, A., Khelifi, A., Suryn, W., Seffah, A.: Usability meanings and interpretations in ISO standards. *Software Qual. J.* **11**, 325–338 (2003)
20. Casas, J., Tricot, M.-O., Abou Khaled, O., Mugellini, E., Cudre-Mauroux, P.: Trends methods in Chatbot evaluation 280–286 (Oct 2020)
21. Sinha, S., Basak, S., Dey, Y., Mondal, A.: *Emerging Technology in Modelling and Graphics*, pp. 55–60. Springer (2020)
22. Okonkwo, C.W., Ade-Ibijola, A.: Python-bot: a chatbot for teaching python programming. *Eng. Lett.* **29** (2020)
23. Durall, E., Kapros, E.: Co-design for a competency self-assessment chatbot and survey in science education. In: *International Conference on Human-Computer Interaction*, pp. 13–24 (2020)
24. Röhrig, C., Heß, D.: OmniMan: a mobile assistive robot for intralogistics applications. *Eng. Lett.* **27** (2019)
25. D’Silva, G., Jani, M., Jadhav, V., Bhoir, A., Amin, P.: *Advanced Computing Technologies and Applications*, pp. 1–9. Springer (2020)
26. Mckie, I.A.S., Narayan, B.: Enhancing the academic library experience with chatbots: an exploration of research and implications for practice. *J. Aust. Libr. Inf. Assoc.* **68**, 268–277 (2019)
27. Okonkwo, C.W., Ade-Ibijola, A.: Chatbots applications in education: a systematic review. *Computers and Education: Artificial Intelligence* **2**, 100033 (2021)
28. Molnár, G., Szűts, Z.: The role of chatbots in formal education. In: *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)*, pp. 000197–000202 (2018)
29. Graesser, A.C.: Conversations with AutoTutor help students learn. *Int. J. Artif. Intell. Educ.* **26**, 124–132 (2016)

30. Sreelakshmi, A., Abhinaya, S., Nair, A., Nirmala, S.J.: A question answering and quiz generation chatbot for education. In: 2019 Grace Hopper Celebration India (GHCI), pp. 1–6 (2019)