



Principles for the Design of an Educational Voice Assistant for Learning Java

Carlos Delgado Kloos^(✉), Carlos Alario-Hoyos,
Pedro J. Muñoz-Merino, Cristina Catalán Aguirre,
and Nuria González Castro

Universidad Carlos III de Madrid,
Av. Universidad, 30, 28911 Leganés, Madrid, Spain
{cdk, calario, pedmume}@it.uc3m.es,
crcatala@pa.uc3m.es, nurigonz@db.uc3m.es

Abstract. Conversational agents, be they text- or voice-powered, are acquiring a level of maturity that makes them useful for natural and smart interactions. In this paper, we explore some design principles for voice-commanded assistants for educational use, in particular, of one designed to train Java concepts as a complement to a MOOC (Massive Open Online Course) about programming.

Keywords: Conversational agents · Voice assistants · VUI · MOOCs · Java teaching · Design decisions

1 Voice Assistants for Education

We have seen improvements in the human-machine interface of computing systems over the last years and with increased speed in recent times. Apart from graphical user interfaces with touch features and multiple sensors, recently voice recognition technology is making a breakthrough.

Conversational agents with natural language understanding and machine learning allow it to maintain quite reasonable conversations in a very natural way without the user having to follow strict predefined commands. Voice assistants like Amazon Alexa, Google Assistant or Siri embodied in smartphones or smart speakers are entering the life of many users. From getting information (weather, traffic, news, etc.) to automating the home (lights, music, plugs, etc.), a wealth of use cases is coming up. This opens also interesting new opportunities for education that have not been fully explored yet.

MOOC (Massive Open Online Course) platforms like edX or Coursera showed the world that technology was ripe for more learning scenarios that had not been harnessed before. Cloud computing as the basis for video hosting, interactive quizzes, forums, and embedded tools allowed bringing educational experiences around the world. Their openness has brought new opportunities to disadvantaged populations. MOOCs are massive, but they are also personal. Learners can view videos at their own pace, do quizzes when they want, but can education be made even more personal?

Voice assistants are personal assistants in the sense that they allow learners to be involved in a useful learning conversation. In the literature, there exist some initiatives

of the use of conversational agents for education. The web site for Google Assistant [1] references many agents developed for educational purposes. Often, they just cover small goals. The work of Demetriadis et al. [2] uses conversational agents to interact together with learners in MOOC discussion forums. Our approach differs from the one of Demetriades in the sense that our assistant should be used in a one-to-one conversation with each learner, rather than being an additional partner with a multitude of learners in a forum.

2 Java-PAL

We have developed an educational program on the edX platform with a series of three MOOCs to learn programming in Java. The program is available both in English [3] and in Spanish [4]. Combined they have had close to half a million registrations in several runs so far. Learners have come from all over the world: 210 countries have been reported. Apart from the rich resources already offered through the platform, we wanted to offer the learner an additional learning experience through a voice assistant.

Java-PAL is an application with which students learn about the basic concepts of Java programming shown in the first of our three Java MOOCs. One of the main purposes of Java-PAL is to offer our learners concepts of Java programming based on what they have already studied [5, 6]. Java-PAL is programmed with Actions on Google. A prototype is available on smartphones and also on smart speakers like Google Home. Later on, we envisage to improve the present version with videos to be shown on devices like Google Home Hub and compatible devices.

3 Design Principles

Apart from the technical design decisions, which are described elsewhere [5, 6], we want to focus in this paper on some basic underlying design decisions that we have followed for Java-PAL. Usage modes of a web-based learning tool is quite different from one based on voice interactions.

In this paper, we summarize some principles we have inferred from our present experience. Figure 1 shows 5 of the 6 design principles:

- User-Friendliness with a Quality Conversation. A quality conversation is essential to engage the learner. We want to encourage the learner to be comfortable with the assistant and continue exploration. Style and content of the conversation have to be properly designed. Among others, messages of encouragement should be considered for the persona underlying the assistant.
- Overview for Conceptual Navigation. When learning about one concept, we want to offer the learner related concepts. Therefore, an underlying concept map is useful in order in order to connect to related concepts.
- Flexibility with Several Interaction Modes. Along the same lines, we want to allow the learner to be able to have different exploration modes. Such can be the direct request for the explanation of a concept, or for an example of a concept.

Alternatively, the learner should be able to let the assistant take the initiative and pose questions.

- Personalization to Learner Preferences. The assistant should be aware of each specific learner and their history and preferences to personalize the interaction individually.
- Adaptation to Device Affordances. Due to the proliferation of devices with different features, it is key to recognize the affordances and deliver the best experience in the context of the affordances available.



Fig. 1. 5 design principles

Apart from the first principle, a quality conversation, the other four represent adaptation to different aspects. The learner preferences principle adapts to the learner, the device affordances principle adapts to the device, the conceptual navigation principle adapts to the topic under study, and the interaction mode principle adapts to the mood and need at each moment.

There is a sixth principle. Such a tool is never the definitive one. A tool has to be maintained, adapted, evolved to changing specifications and insights. Evolvability and extensibility need to be considered from the very beginning of the design process. This takes us the principle No. 6.

- Extensibility for Future Evolution. Although we have designed our assistant initially for the teaching of Java, we should be able to adapt it to other MOOCs by changing content as easily as possible. Also, new devices, new usage modes, new pedagogies, etc. could all imply the evolution of our conversational agent.

Let's analyze these principles in closer detail.

3.1 User-Friendliness with a Quality Conversation

In a similar way as there are theories to design a user-friendly GUI (graphical user interface) [10], a VUI (voice user interface) needs to follow some guidelines. Grice [11] defines four proposals to maximize several aspects of a conversation. The assistant should maximize the amount of information provided (maxim of quantity) at each interaction, but only to the level that is relevant (maxim of relevance). It should also communicate clearly (maxim of manner) and cooperate with the user (maxim of quality). The assistant should engage with the user in a cooperative conversation.

When designing a conversational agent, it is necessary to give it a persona, a personality. It is very important not to design a cold relationship between learner and assistant, but to have one which is warm and welcoming. May the learner fail often or give all answers right, scheduling messages of encouragement is always right. Misunderstandings need to be handled well, time-outs be well thought of, in summary, the flow of conversation needs to be right.

3.2 Overview for Conceptual Navigation

In contrast to studying with the MOOC, where the learning sequence plays a preeminent role, we expect the usage of Java-PAL to be more sporadic, exploratory and playful. The learner might want to jump from one concept to a related one. For example, if the user is learning about expressions we can suggest to look at the concept of variable or at different kinds of operators, because they have a direct relationship. Or if the learner doesn't fully understand the concept of a statement, we can recommend him to look at concepts such as conditional statement and repetition.

The concepts and the relationships between them can be defined with a concept map. Our first approach was to use existing terminology that can be found in ontology literature, such as the work in [7] or [8], but it was difficult to establish the relationships between the concepts in which our Java MOOC is based on. Moreover, the main purpose of the application is not related to the kind of relationship between the concepts, but with the existence of a relationship itself. Thanks to the ontology defined (see Fig. 2), we are able to offer the learners related concepts based on their educational needs.



Fig. 2. Ontology of concepts

3.3 Flexibility with Several Interaction Modes

Learning that happens through dialog (also called dialogic learning [9]) has been proved to be more effective than education based on monologs. Learners take a more active role when they are involved in a conversation. We also wanted to provide our assistant with more than one mode of operation. Offering the learner different operation modes can make the dialog more refreshing and fluent.

There is a first mode where users can ask for definitions or examples of Java concepts that they need for clarification and understanding, or just the ones related to the questions they are not able to answer correctly. In this mode, users take the initiative of asking anything, but then the assistant can lead the teaching conversation offering related concepts. Then there is a second mode, where the assistant takes the initiative, and asks learners questions related to the concepts studied, to check their

knowledge. Finally, we found it interesting to have a feedback mode for self-development, where learners will be given the results of their performance during the quizzes, so they can know their strong and weak points, while also giving some encouragement and congratulations.

There are also some extra interesting features that we considered important to make the Java-PAL more appealing to the end users. They will be able to change some personalized settings, like the name to be called and the number of questions they would like to be asked. We wanted our Java-PAL to be able to adapt to possible or future changes in the users requirements; we are opened to listen to the users feedback, and add some other properties in the next versions.

3.4 Personalization to Learner Preferences

Personalized learning provides each learner with different specific user experience depending on many different parameters that conform a user model [12]. Personalization can help to improve the learning process. This learner profile can be obtained by different means: (1) the users introduce their preferences, feelings, behaviours, etc. directly, e.g. with a form; (2) there is an automatic detection of these features through students' interactions with the learning platform.

Adaptive learning for hypermedia systems can adapt the contents or the presentation [12]. In our voice assistant we are mainly focused on the personalization of contents, specifically in the adaptation of questions to students and in the adaptation of the recommended resources if a student fails a question. The personalization can be provided taken into account a user model (e.g., a basic one might only be based on the student knowledge on different skills) and the specific contents and their relationships (given by a proposed ontology).

The student model can include interactions from different platforms (e.g., by the Learning Management System but also with the voice assistant) and the data from different sources can be combined to detect some student feature. An ontology might also be used to provide interoperability at the semantic level among these different systems.

The content model given by the ontology provides some extensibility and interoperability at the semantic level. Concepts that are used in different platforms (e.g., in a MOOC or in a voice assistant) have a common meaning and relationships with other concepts based on the proposed ontology.

3.5 Adaptation to Device Affordances

Voice assistants are increasingly being integrated into a larger number of devices. Traditionally, voice assistants have been used only through desktop/laptop computers or mobile devices (smartphones, tablets), where the user has a screen to facilitate oral communication, including images or videos when needed, and a keyboard, to write in case there is a misunderstanding in the communication between user and voice assistant.

However, 2018 was the year of smart speakers (and a turning point for voice assistants), with Google Home, Amazon Echo or Apple HomePod at the fore-front.

These devices only allow an oral communication, since they do not bring any built-in screen or keyboard, forcing to simplify the communication between user and assistant. The next generation of smart devices will once again incorporate a screen, as is the case with Google Home Hub, allowing richer interactions between voice assistant and user.

But that is not all, there are hybrid situations in which the main source of communication with a voice assistant may be an oral conversation, but accompanied by small visual indications. This would be the case while driving in the car, as much visual information can turn into a distraction and cause loss of contact with the road, but small visual indications can help to better understand and follow the oral conversation.

All in all, a voice assistant for education must take into account the device with which the learner works, and its context, and adapt its interactions accordingly. For example, in an educational environment where the user gets questions, explanations, and examples from a smart speaker connected to a voice assistant, these need to be short and concrete, without the possibility of attaching an explanatory image or video. Nevertheless, in other contexts, in which it is possible to include visual support, more complex questions, explanations and examples can be used.

3.6 Extensibility for Future Evolution

Extensibility refers to the fact that a minimum number of changes should be done in order to make the application to work in other contexts, e.g., to other courses of a different domain, including new sources of data or with courses with different questions.

The voice assistant application should be as extensible as possible. There will be a part of the application that will be common. There will also be some parts of the application that should be replaced depending on the context.

The common parts include: (1) the semantic models with the ontologies that can model the user or even the contents if they are in the same domain; and (2) some algorithms for adaptation and personalization.

Among the changing parts are the ontology (it should be adapted to the domain but the defined relationships should be used in any contexts), the own questions and the annotations of the different resources (questions and contents).

4 Conclusion

As Anant Agarwal, CEO of edX, says, education will be omni-channel. Learners will want to use multiple devices that are synchronized to offer a convenient learning experience. Laptops and smartphones each have their specific features and moments for which they are the best fit. Now voice assistants have entered the scene with their limitations and promises offering an interesting complement to the existing landscape.

We have presented in this paper some decisions that have guided the design of a voice assistant for learning Java. We believe that voice assistants can be an important added value to existing learning methods and that we are just scratching the surface of what is possible with voice-operated applications.

Acknowledgments. This work has been partially funded by FEDER/Ministerio de Ciencia, Innovación y Universidades - Agencia Estatal de Investigación/Smartlet project (TIN2017-85179-C3-1-R). In addition, this work has been partially funded by the e-Madrid-CM project with grant No. S2018/TCS-4307, which is funded by the Madrid Regional Government (Comunidad de Madrid), by the Fondo Social Europeo (FSE) and by the Fondo Europeo de Desarrollo Regional (FEDER).

References

1. <https://assistant.google.com/explore/c/3/?hl=en-US>. Accessed May 2019
2. Demetriadis, S., et al.: Conversational agents as group-teacher interaction mediators in MOOCs. In: Learning with MOOCs 2018, Madrid, Spain, pp. 43–46. IEEE (2018)
3. Kloos, C.D., et al.: Introduction to Java Programming, professional certificate program in English, edX. www.edx.org/professional-certificate/uc3mx-introduction-java-programming. Accessed May 2019
4. Kloos, C.D., et al.: Introducción a la programación en Java, professional certificate program in Spanish, edX. www.edx.org/es/professional-certificate/uc3mx-introduccion-a-la-programacion-en-java. Accessed May 2019
5. Aguirre, C.C., Kloos, C.D., Alario-Hoyos, C., Muñoz-Merino, P.J.: Supporting a MOOC through a conversational agent. Design of a first prototype. In: 20th International Symposium on Computers in Education (SIIE), Cádiz, Spain, pp. 1–6. IEEE (2018)
6. Kloos, C.D., Catalan, C., Muñoz-Merino, P.J., Alario-Hoyos, C.: Design of a conversational agent as an educational tool. In: Learning with MOOCs 2018, Madrid, Spain, pp. 27–30. IEEE (2018)
7. John, S.: Development of an Educational Ontology for Java Programming (JLEO) with a hybrid methodology derived from conventional software engineering process models. *Int. J. Inf. Educ. Technol.* **4**(4), 308–312 (2018)
8. Boyce, S., Pahl, C.: Developing domain ontologies for course content. *Int. Forum Educ. Technol. Soc.* **10**(3), 275–288 (2007)
9. Flecha, R.: *Sharing Words: Theory and Practice of Dialogic Learning*. Rowman & Littlefield, Lanham (2000)
10. Nielsen, J.: *Designing Web Usability: The Practice of Simplicity*. New Riders Publishing, San Francisco (1999)
11. Grice, H.P.: Logic and conversation. In: Cole, P., Morgan, J. (eds.) *Syntax and Semantics*, vol. 3, pp. 22–40. Academic Press, Cambridge (1975)
12. Brusilovsky, P.: Methods and techniques of adaptive hypermedia. *User Model. User-Adap. Interact.* **6**(2–3), 87–129 (1996)