Designing a Virtual Client for Requirements Elicitation Interviews

Sourav Debnath✉ and Paola Spoletini✉

Kennesaw State University, Marietta, USA
sdebnath@students.kennesaw.edu, pspoleti@kennesaw.edu

Abstract. [Context and motivation] Role-playing offer experiential learning through the simulation of real-world scenarios; for this reason, it is widely used in software engineering education. In Requirements Engineering, role-playing is a popular way to provide students hands-on experience with requirements elicitation interviews. [**Problem]** However, managing a role-playing activity to simulate requirements elicitation interviews in a class is time consuming, as it often requires pairing students with student assistants or fellow classmates who act as either customers or requirement analysts as well as creating and maintaining the interview schedules between the actors. To make the adoption of role-playing activities in a class feasible, there is a need to develop a solution to reduce instructors' workload. [**Principal ideas]** To solve this problem we propose the use of VIrtual CustOmer (VICO), an intent-based, multimodal, conversational agent. VICO offers an interview experience comparable to talking to a human and provides a transcript of the interview annotated with the mistakes students made in it. The adoption of VICO will eliminate the need to schedule interviews as the students can interact with it in their free time. Moreover, the transcript of the interview allows students to evaluate their performance to refine and improve their interviewing skills. **[Contribution]** In this research preview, we show the architecture of VICO and how it can be developed using existing technologies, we provide an online rule-based initial prototype and show the practicality and applicability of this tool through an exploratory study.

Keywords: Requirements Elicitation Interview, Role-playing, Requirements Engineering Education and Training, Intelligent Agent.

1 Introduction

The goal of requirements elicitation is to discover requirements for a system by communicating with the stakeholders and exploring available information. Among the variety of available elicitation techniques, requirements elicitation interviews are the most used [1] [6] [11] and among the most effective [5] [6] [14]. While often perceived by students and young analysts as an easy technique to master, the success of requirements elicitation interviews depends on many (soft) skills, such as the ability to create a relationship with the interviewee to ease the process, formulate questions properly, and introspect or probe into the customers' needs. The importance of these skills and the level of effort required to acquire them are difficult to communicate in a convincing way through traditional lectures but are immediately perceived by students through

practice. For this reason, role-playing is a popular pedagogical approach to teach these skills [15].

As a class activity, role-playing for requirements elicitation interviews can be performed by pairing students with each other [16] or with student assistants acting as either customers or requirements analysts [9]. Despite the positive results obtained by using role-playing activities to teach requirements elicitation interviews [15] [16], such activities are not often adopted since they can be cumbersome to manage in a standard classroom setting. Indeed, scheduling students for the interview activity (with either peers or student assistants) and maintaining the schedule are time and resource consuming activities. Furthermore, role-playing does not provide all students with an experience of comparable quality: some students may be partnered with someone motivated and talented who plays her role well, while others may be partnered with someone who has no interest in the activity and so underperforms. Also, students can usually role-play only one or two interviews which may not provide them with enough experience to realize their mistakes and improve their interviewing skills.

To overcome these limitations, and consequently support the diffusion of role-playing activities to teach requirements elicitation interviews and to train young analysts, similarly to what is done in other fields [2], [10], [8], [12], [13] we propose the use of VIrtual CustOmer (VICO), an intent-based, multimodal, conversational agent, able to offer an interview experience comparable to the one provided by a human. In addition, VICO also provides the users with a transcript of the interview annotated with the mistakes that the user made in the interview. Given its double contribution, the adoption of VICO has the potential not only to resolve the problems connected to organize role-playing interviews, but also to allow students to evaluate and reason on their performance to refine and improve their interviewing skills. Our research towards the development of VICO has the goal to answer to the following research questions:

RQ1: Can the use of an agent-based solution statistically significantly improve interviewing skills of novice analysts?

RQ2: How comparable can an agent-based solution be to real human interaction in terms of (**2.1**) usability, (**2.2)** learning experience, (**2.3)** engagement?

Because of the ambitious nature of our research, we are planning to develop a series of prototypes with increasing level of intelligence and building on the evaluation of each prototype to develop the subsequent one using more advanced algorithms and architectures. In particular, we are planning to produce three prototypes to evaluate our idea and results in a timely manner.

In the rest of the paper, we present an overview of the envisioned final version of VICO (Section 2), we outline our research agenda (Section 3), and we then describe our first prototype of VICO, V_0 and the results of an exploratory study aimed to initially answer to RQ2 (Section 4). Finally, Section 5 concludes the paper.

2 An overview of VICO

VICO is defined as a multimodal intent-based conversational agent. Roughly speaking, this means that VICO can support different input and output modalities and is capable of processing natural language questions and building an adequate response to them.

VICO has two main goals: playing the role of a customer who wants to develop a product and is interviewed by a VICO's user, and producing a transcript annotated with the mistakes made by the user during the interview. To achieve these goals, VICO needs to be able to (1) listen to the interviewer and understand her questions, (2) appropriately react to them, generating a suitable answer, (3) communicate the answer to the interviewer, (4) keep track of the interview status, and (5) identify the mistakes the user made while conducting the interview.

The capabilities (1)-(3) are comparable to capabilities of virtual agents, such as Ellie, or the agents created with either TARDIS or Virtual People Factory, used for training purposes in specific fields. Ellie [8] is a virtual human, a fully animated face-to-face interviewer used to make automated healthcare decision after conversing with patients. It supports a pre-fixed number of sentences (and their variations) and can detect non-verbal behavior like distress, anxiety. This tool can show emotions like empathy and active engagement which builds trust with the patients. TARDIS [2] is, instead, a framework to create virtual agents that can be used to run social coaching workshops in which young adults are taught how to talk, behave and present themselves in job interviews. In this tool the scenes, dialogues and animations are predetermined by the authors. Similarly, to Ellie, also TARDIS agents have the capability to detect social cues of the participants and produce reactive animations in the virtual recruiter. Finally, Virtual People Factory [4] is a tool to create virtual humans whose conversational capabilities rely on a model built on real conversations. In [15], Virtual People Factory has been used to build virtual patients with stomachache capable of talking in natural language. The goal was to use these virtual humans to teach medical students how to interact with patients in realistic settings. To build the conversational model efficiently, Rossen et al. propose the use of crowdsourcing [15].

The above-mentioned tools suggest that constructing a virtual agent with human-comparable conversational capabilities using the existing technologies is a realistic goal. However, since these solutions are domain-specific and require building a specific model for each different type of human, adopting them in our case would require building a new model for each requirements elicitation interviewee. While our current prototype is also based on predefined knowledge, our plan for VICO is to create a tool that relies on a model independent from the domain of the interview built only on the structure of correct/incorrect interviews and correct/incorrect questions. The specific domain, separated by the model and provided as a knowledge base, will be used only to contextualize the answers of the agent, while their structure will be determined using an analysis of the question structure and the interview current state. To realize this separation, we will rely on the high-level architecture shown in Figure 1.

VICO's users interact with it online through a user interface with an Embodied Agent. The inputs from the users will be analyzed using a Speech Synthesizer; which will extract the intents and entities. Intents are categorizations of the user's input that identify

what the user wants from VICO and entities are keywords that determine what specific information a user is requesting. The intents and the entities are then combined with the interview status, a list of the past intents, entities, questions and answers of the current interview, received from the Interview Recorder. The combined information will then be sent to the answer generator. A part of the Speech Synthesizer could be built using DeepSpeech [7], an open-source speech recognition software.

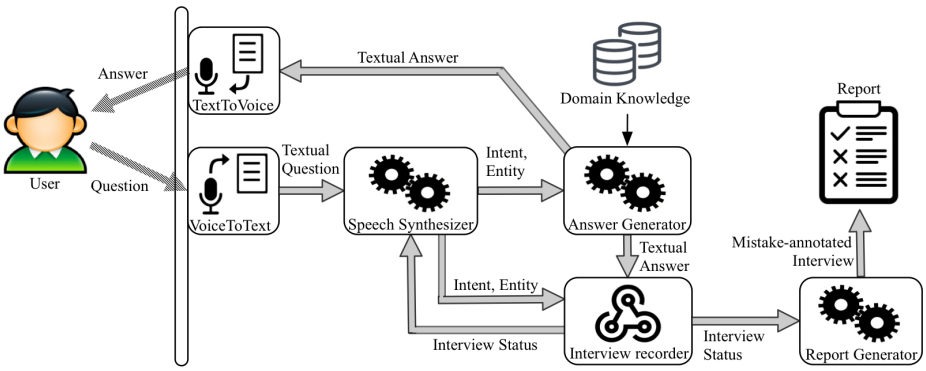

Fig. 1. Overview of VICO

The contextualized and structured information related to the current question are then sent to the Answer Generator that uses them and the domain knowledge in the knowledge base to formulate an appropriate answer to the question. This component is the core of VICO and will be developed to identify the structure of the response independently from the domain and to complete this response using the knowledge base separately. In this way, the only part to update, when a new domain for the interview is selected, is the domain base.

The response generated by the Answer Generator is then sent to the user through the Embodied Agent and to the Interview Recorder, an aggregator that keeps track of the interview from its beginning to its current state. The status, as explained above, is used in the Speech Synthesizer, and is also sent to the Report Generator, which analyzes the questions to identify the user's mistakes. The complete recorded interview annotated with the mistakes identified by Report Generator is provided to the user.

3 Research Agenda

Because of the ambitious nature of the proposed idea, before developing the agent with all the characteristics described in Section 2, we plan to develop two prototypes with increasing level of intelligence and, by building on the evaluation of each prototype, to construct the subsequent one using more advanced algorithms and architectures . This will allow us to evaluate our idea in a timely manner.

The goal of the first prototype, V_0, is to evaluate the potential of our idea. V_0 is designed as a basic online multiple-choice text adventure software that uses an open-source interactive non-linear story-telling tool to simulate a requirements elicitation interview. V_0 is evaluated through an exploratory study to investigate the perception of users while using V_0 and its potential impact on their performance. More details on V_0 and the results of the exploratory study are briefly V_0 presented in Section 4.

As intermediate prototype, we plan to develop a tool with similar capability of the tool described in Section 3, with the only difference that this prototype will have limited ability to formulate answers. Its Answer Generator will act exactly as the agents produced by TARDIS and Virtual People Factory and will use the input from the Speech Synthesizer to select the appropriate response from a knowledge corpus. This corpus will contain a comprehensive set of responses related to the domain of the interview. These responses are primarily extracted from role-played interviews and manual entries from domain experts. To evaluate this prototype, we plan to run two controlled experiments, one to answer to RQ1 and one to answer to RQ2. In the first experiment, the control group will only be taught how to run interviews and will be then asked to execute one as interviewer, while the experimental group will use the tool before running the interview. The interviews of both groups will be independently analyzed by two experts to measure the effectiveness of VICO in improving the quality of the interviews. During the analysis of the interviews, the experts will be blind with respect to the group in which participants belong. The second experiment, to evaluate RQ2, compares the impact of the training with a human with the impact of the training with VICO. So, the control group will be first role-playing with a human, while the experimental group will be role-playing with VICO, then they both run a second interview with a human. The participants in both groups will be asked to fill out a follow-up questionnaire, aimed to investigate how the groups perceived the effect of the training. Also, as done for the interview in the first experiment, the second interview of both groups will be analyzed to have an objective measure of the effectiveness of VICO compared with the human-based experience.

The final product of VICO will be evaluated in a similar way. Moreover, we will analyze the required time to build knowledge bases to evaluate the time needed to build new customers. Finally, we will ask members of our institution's industry advisory board to evaluate the tool with their younger analysts to explore the opportunity to use it in industry settings as a training tool for freshly hired analysts.

4 Initial Prototype and Exploratory study

The initial prototype[1] for VICO, V_0, simulates a virtual customer, who wants to build a software system for her ski resorts. It has been developed using Twine, an open-source tool for creating non-linear stories. When using V_0, at each step, the user is given three options of questions to ask the customer. Once the user selects a question, V_0 provides a suitable predefined answer. Each question has different levels of mistakes associated

[1]Available at http://www.interviewsim.com.s3-website.us-east-2.amazonaws.com/

with it. The questions and responses were developed by examining 80 interview recordings collected from role-playing interview activity and cover 12 types of mistakes, selected from the 34 mistake types described in [3]. At the end of the interview, the participants are given the full list of questions they asked and the associated mistakes.

We performed an initial exploratory study with 17 undergraduate junior students majoring in software engineering to evaluate the potential of V_0. The participants were divided into 2 groups. Both groups initially watched a 20 minutes video on how to conduct elicitation interviews [9]. Then the participants in the first group were given an initial description of a ski resort project and were asked to use the tool and analyze their mistakes. The second group skipped this step. Then, both groups were given the description of another project for a hair salon and performed interview as interviewers on this topic with a human fictional customer. The interviews were recorded and analyzed independently by the authors to identify the mistakes in the asked questions using an evaluation sheet provided in [9]. To ensure a unbiased review of the interviews the assignment of the participants to each group was performed by an external researcher and not shared with the authors. This preliminary informal analysis shows that the participants who used V_0 before conducting the interview with the customer made less mistakes than the participants who did not use the prototype. In particular, they avoided the majority of the mistake types embedded into V_0.

Moreover, 25 participants used V_0 and evaluated its engagement, helpfulness and level of difficulty, with the following results: 73% of the participants rated the engagement positively, 81% found V_0 helpful and 77% found it easy to use.

5 Conclusion

In this research preview, we introduced our idea and research agenda to develop a human-comparable agent-based customer to train analysts to perform requirements elicitation interviews and save the resources needed to manage the same activity with human participants. We have also presented our initial prototype, V_0, and the preliminary results obtained by using it. Notice that some of the user of V_0 sent us verbose positive feedback that encourages us to further evaluate V_0 and move forward with our research agenda.

Acknowledgment

The authors thank Kim Hertz for her support in the enrollment of the participants that guaranteed that the authors were blind with respect to the participants' assigned group, the graduate and undergraduate research assistants in the Tiresias Lab for beta testing V_0, and all the participants for their time. This work was partially supported by the National Science Foundation under grant CCF-1718377.

References

1. Agarwal, R., Tanniru, M.R.: Knowledge acquisition using structured interviewing: an empirical investigation, JMIS, vol. 7, no. 1, pp. 123–140, 1990.
2. Anderson, K., Andre, E., Baur, T., Bernardini, S., Chollet, M., Chryssafidou, E., Damian, I., Ennis, C., Egges, A., Gebhard, P., Jones, H., Ochs, M., Pelachaud, C., Porayska-Pomsta, K., Rizzo, P., Sabouret, N.:. The TARDIS framework: Intelligent virtual agents for social coaching in job interviews. In Proceedings of International Conference on Advances in Computer Entertainment (2013), 476–491.
3. Bano, M., Bano, M., Zowghi, D., \Ferrari, A., Spoletini, P., Donati, B.: Learning from Mistakes: An Empirical Study of Elicitation Interviews Performed by Novices. IEEE 26th International Requirements Engineering Conference (RE), 182–193 (2018).
4. Carnell, S., Lok, B., James, M.T., Su, J. K.: Predicting student success in communication skills learning scenarios with virtual humans, 9th International Conference on Learning Analytics & Knowledge. ACM , 436–440 (2019).
5. Coughlan, J., Macredie, R.D.: Effective communication in requirements elicitation: a comparison of methodologies. Requirements. Eng., 7(2):47–60 (2002).
6. Davis, A.M., Tubío, Ó.D., Hickey, A.M., Juzgado, N.J., Moreno, A.M.: Effectiveness of Requirements Elicitation Techniques: Empirical Results Derived from a Systematic Review. 14th IEEE International Requirements Engineering Conference, 179-188 (2006).
7. DeepSpeech: https://github.com/mozilla/DeepSpeech
8. DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., Georgila, A., Gratch, J., Hartholt, A., Lhommet, M., Lucas, G., Marsella, S., Morbini, F., Nazarian, A., Scherer, S., Stratou, G., Suri, A., Traum, D., Wood, R., Xu, Y., Rizzo, A., Morency, L.P.: SimSensei Kiosk: A virtual human interviewer for healthcare decision support, in Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), Paris, France, 5 to 9 May 2014 (International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 2014), pp. 1061–1068;
9. Ferrari, A., Spoletini, P., Bano, M., Zowghi, D.: Learning Requirements Elicitation Interviews with Role-playing, Self-assessment and Peer-review, IEEE 27th International Requirements Engineering Conference (2018).
10. Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R.: Creating rapport with virtual agents, in IVA '07: Proceedings of the 7th international conference on Intelligent Virtual Agents. Springer-Verlag, 2007, pp. 125–138
11. Hickey, A.M, Davis., A.M.: Elicitation technique selection: how do experts do it?, 11th IEEE International Requirements Engineering Conference,169-178 (2003).
12. Kleinsmith, A., Rivera-Gutierrez, D., Finney, G., Cendan, J., Lok, B.: Understanding empathy training with virtual patients. Comput. Human Behav. 52, 151–158.
13. Rossen, B., Lind, D., Lok, B.: Human-centered Distributed Conversational Modeling: Efficient Modeling of Robust Virtual Human Conversations, International Conference on Intelligent Virtual Agents, 474–481 (2009).
14. Sutcliffe, A., Sawyer, P.: Requirements elicitation: towards the unknown unknowns. In RE'13, pages 92–104. IEEE, 2013.
15. Svensson, R.B., Regnell, B.: Is role playing in requirements engineering education increasing learning outcome? Requirements Engineering, 22 (4), 475–489 (2017).
16. Zowghi, D., Paryani, S.: Teaching requirements engineering through role playing: Lessons learnt, 11th IEEE International Requirements Engineering Conference, 233–241 (2003).