

A job interview dialogue system with autonomous android ERICA

Koji Inoue, Kohei Hara, Divesh Lala, Shizuka Nakamura, Katsuya Takanashi, and Tatsuya Kawahara

Abstract We demonstrate a job interview dialogue with the autonomous android ERICA which plays the role of an interviewer. Conventional job interview dialogue systems ask only pre-defined questions. The job interview system of ERICA generates follow-up questions based on the interviewee's response on the fly. The follow-up questions consist of two kinds of approaches: selection-based and keyword-based. The first type question is based on selection from a pre-defined question set, which can be used in many cases. The second type of question is based on a keyword extracted from the interviewee's response, which digs into the interviewee's response dynamically. These follow-up questions contribute to realizing natural and trained dialogue.

1 Introduction

Spoken dialogue systems have been developed for various scenarios so far such as smartphone apps and smart speakers, and future systems are expected to handle more social interaction like human-human dialogues in our daily lives. For example, in a real-life job interview, interviewers make various kinds of questions directly to interviewees in order to elicit information necessary for determining acceptance. A spoken dialogue system is expected to play the role of an interviewer so that interviewees are able to practice answering against expected questions. To assist with interview practice, the system has to realize a job interview dialogue similar to real human-human dialogue.

We demonstrate a job interview dialogue with autonomous android ERICA [3, 4] in the role of an interviewer. ERICA looks like a human being and is able to generate various behaviors including non-verbal ones such as eye gaze and head nodding.

Graduate School of Informatics, Kyoto University, Japan,
e-mail: [inoue] [hara] [lala] [shizuka] [takanashi] [kawahara]@sap.ist.i.kyoto-u.ac.jp

Current spoken dialogue systems for job interview dialogue ask only pre-defined questions [2, 7, 1, 6]. Although it is important for interviewers to dynamically ask follow-up questions to know more about the current topic, only a small number of studies have been conducted on automatic generation of follow-up questions [8]. In this demonstration, ERICA generates follow-up questions based on how well the interviewee's response fulfills the previous question and also keywords extracted from that response. It is expected that these follow-up questions make the job interview dialogue more natural and stimulate some thought from the interviewee, which is required for a practical system for job interview training.

2 System configuration

While the basic dialogue flow is controlled with finite state transition, questions are generated based on interviewees' responses as explained below.

2.1 Scenario

In this demonstration, ERICA plays the role of the interviewer in a job interview. The dialogue content is independent of any business category or company, so questions from ERICA focus on the motivation and experience of interviewees. Therefore, the proposed system can be applied to interviewees with various background without modifying the list of questions.

2.2 Dialogue flow

The dialogue flow is controlled with finite state transitions as illustrated in Figure 1. The current job interview consists of several topics. Example topics are *reasons for application* and *accomplishments in college*. Each topic starts with a base question which is followed by follow-up questions. The base question is an open question on the topic. For example, when the topic is *reasons for application*, a base question is "What is the reason you applied for this job?" Topics and corresponding base questions are pre-defined. Next, based on an interviewee's response to the base question, the system generates a selection-based follow-up question. Finally, based on the response to the follow-up question, a keyword-based follow-up question is generated. Then, the current topic ends and proceeds to the next topic. The following section describes how to generate the above two kinds of follow-up questions in detail.

The other settings of ERICA's behaviors are as follows. ERICA generates head nodding when an interviewee is answering, based on a model for backchannel generation using prosodic features [5]. For the turn-taking behavior, ERICA waits for

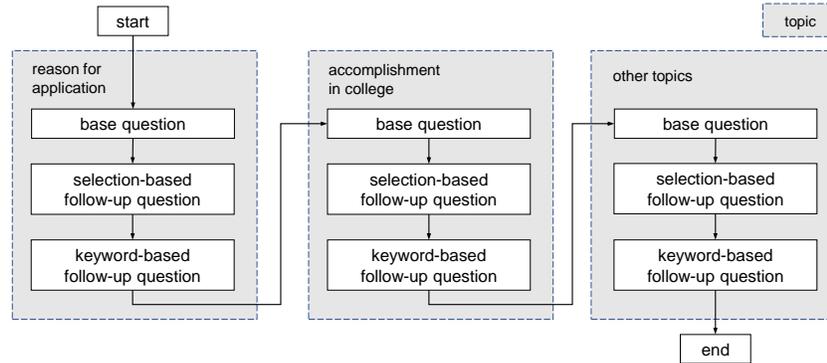


Fig. 1 Dialogue flow of job interview

four seconds silence to take the floor. ERICA should wait for longer than in other dialogue scenarios because the system has to make sure that an interviewee has definitely finished answering. Note that ERICA generates head nodding while ERICA waits for the user's turn to end to avoid an unnatural silence. If an interviewee's response contains only a few nouns, ERICA will reply with a phrase such as "Sorry, could you say it again?" to prompt the interviewee to continue the answering.

2.3 Generation of follow-up questions

We address how to generate the two kinds of follow-up questions below. At first, the system generates a selection-based follow-up question against a response to a base question. Next, a keyword-based follow-up question is generated against a response to the selection-based follow-up question.

2.3.1 Selection-based follow-up questions

We prepare several follow-up questions related to a base question, and the system selects an appropriate one based on the interviewee's response. This approach is robust against to any interviewee's responses because sentences of the questions are handcrafted. The prepared follow-up questions are categorized into three categories taking into account how well the response has fulfilled the base question.

- **category 1** (high priority) follow-up questions on which an interviewee must answer
- **category 2** (middle priority) follow-up questions on which an interviewee should answer
- **category 3** (low priority) other follow-up questions

Category 3 is a backup for categories 1 and 2. For example, for the topic *reason for apply*, follow-up questions can be as follows.

- **category 1** What is the specific reason why you want to work in our company?
- **category 2** What do you want to achieve after you enter our company?
- **category 3** Have you applied to any other companies?

The question of category 1 should be used when an interviewee’s response did not satisfy the base question. The question of category 2 should be used when the interviewee accounted for the base question but did not mention in detail related to the follow-up question. The question of category 3 should be used when the interviewee has answered the base question well. Category 1 has the highest priority to be asked, followed by categories 2 and 3.

The system selects an appropriate follow-up question from the above-mentioned prepared set based on the degree of fulfillment of an interviewee’s response. At first, if the number of nouns of an interviewee’s response is larger than a threshold, follow-up questions of category 1 are excluded from candidates, since these were likely already mentioned in the interviewee’s response. Next, for each follow-up question, the system checks if it has been addressed in the response. For this, in advance, we define a set of words to represent each follow-up question. The set contains words that appear in the sentence of the follow-up question itself. The set also includes words related to the follow-up question. For example, related words for a follow-up question “*Where do you see yourself 10 years in the future?*” are *future*, *years*, and so on. Then, if one of the words in an interviewee’s response to a base question is semantically close to one of the words in the above-mentioned list, the corresponding follow-up question is excluded from the candidates. In the above example, if an interviewee’s response includes *future*, the follow-up question will not be selected. To measure the semantic distance between words, we use word embedding (word2vec) and its cosine distance. Finally, among follow-up questions left in the set of candidates, the system selects one of them from the category with the highest priority. If there are several possible candidates in the same category, the system randomly selects one question.

2.3.2 Keyword-based follow-up questions

The system also generates another type of follow-up question based on a keyword contained in the interviewee’s response. Although the selection-based follow-up question can always be generated, their sentences are pre-defined. Therefore, it might be difficult for interviewees to feel that their responses are comprehended by the system. To deal with this issue, the system dynamically generates a follow-up question reflecting the response of an interviewee.

The system extracts a keyword from an interviewee’s response and fills the keyword in a template. First, the system calculates the inverse document frequency



Fig. 2 Snapshot of job interview dialogue with ERICA

(IDF) value that was trained with the text8 corpus¹ and the gensim software². We hypothesize that an appropriate keyword is a rarity so that it has a higher IDF value. A keyword is determined as having the highest score of the IDF values. If several nouns are continuously observed, they are regarded as a compound noun, and then it is processed as one noun. To calculate an IDF value for a compound noun, the maximum value is used from several IDF values of consisting nouns.

After extracting a keyword, the system fills the keyword in a pre-defined template to generate a sentence of a follow-up question. For example, when an extracted keyword is *spoken dialogue systems*, the system generates a follow-up question as “*You mentioned spoken dialogue systems, so could you explain them in more detail?*”.

3 Dialogue example

A dialogue example is introduced as below³. The topic is *reasons for application*. Note that **U** and **S** represent utterances of a user and ERICA, respectively. Fig. 2 shows a snapshot of a job interview with ERICA. We use a 16-channel microphone array to enhance the interviewee’s speech for automatic speech recognition so that the interviewee can talk with ERICA without holding a microphone (hands-free).

S1 What is the reason why you applied for this job? (**base question**)

U1 I have learned artificial intelligence in my university. In this company, I can apply my skills of AI technologies to the job and can obtain new skills. This is the reason why I applied for this job.

S2 I see. What kind of effort or preparation do you make to achieve your goal?
(**selection-based follow-up question**)

U2 I learn the latest technologies because this industry drastically changes. Recently, I am learning machine learning methods.

¹ <http://mattmahoney.net/dc/textdata.html>

² <https://radimrehurek.com/gensim/models/tfidfmodel.html>

³ Demo video is available at <https://youtu.be/JpT1SVp2zx8>

S3 I see. Could you explain about machine learning in detail?

(**keyword-based follow-up question**)

U3 I am especially interested in deep learning. It is reported that deep learning allows us to train a complicated model that can be applied to various classification tasks and achieve higher accuracy than conventional approaches.

S4 I see. Thank you.

4 Conclusion

We have described a demonstration of a job interview dialogue with the autonomous android ERICA. The proposed system generates follow-up questions based on an interviewee's response. The follow-up questions consist of two kinds of approaches: selection-based and keyword-based. While the former is based on a pre-defined set of questions, the latter dynamically reflects an interviewee's response by extracting a keyword which was used. These follow-up questions contribute to realizing a natural job interview dialogue with ERICA.

Acknowledgments

This work was supported by JST ERATO Ishiguro Symbiotic Human-Robot Interaction program (Grant Number JPMJER1401), Japan.

References

1. Ali, M.R., Crasta, D., Jin, L., Baretto, A., Pachter, J., Rogge, R.D., Hoque, M.E.: LISSA-Live interactive social skill assistance. In: *ACII*, pp. 173–179 (2015)
2. Hoque, M.E., Courgeon, M., Martin, J.C., Mutlu, B., Picard, R.W.: MACH: My automated conversation coach. In: *UbiComp*, pp. 697–706 (2013)
3. Inoue, K., Milhorat, P., Lala, D., Zhao, T., Kawahara, T.: Talking with ERICA, an autonomous android. In: *Sigdial*, pp. 212–215 (2016)
4. Kawahara, T.: Spoken dialogue system for a human-like conversational robot ERICA. In: *IWSDS* (2018)
5. Kawahara, T., Yamaguchi, T., Inoue, K., Takahashi, K., Ward, N.G.: Prediction and generation of backchannel form for attentive listening systems. In: *Interspeech*, pp. 2890–2894 (2016)
6. Kobori, T., Nakano, M., Nakamura, T.: Small talk improves user impressions of interview dialogue systems. In: *Sigdial*, pp. 370–380 (2016)
7. Smith, M.J., Ginger, E.J., Wright, K., Wright, M.A., Taylor, J.L., Humm, L.B., Olsen, D.E., Bell, M.D., Fleming, M.F.: Virtual reality job interview training in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders* **44**(10), 2450–2463 (2014)
8. Su, M.H., Wu, C.H., Huang, K.Y., Hong, Q.B., Huang, H.H.: Follow-up question generation using pattern-based seq2seq with a small corpus for interview coaching. *Interspeech* pp. 1006–1010 (2018)