Task-oriented spoken dialog system for second-language learning

Oh-Woog Kwon[1], Young-Kil Kim[2], and Yunkeun Lee[3]

Abstract. This paper introduces a Dialog-Based Computer Assisted second-Language Learning (DB-CALL) system using task-oriented dialogue processing technology. The system promotes dialogue with a second-language learner for a specific task, such as purchasing tour tickets, ordering food, passing through immigration, etc. The dialog system plays a role of a ticket agent, a waiter, or an immigration officer at each task and the learner completes the task through talking with the system in the second-languages. The system gives a new situation of the task at every trial, so that the learner could have a different experience and learn various expressions in the same task. Our DB-CALL system is based on a task-oriented dialog system consisting of a language understanding module using a structural support vector model and a dynamic dialog graph based dialog management module. Our task-oriented dialog system is trained from the intention-annotated real dialog scripts. The experiments for the performance of our proposed dialog system show the average task success rate of 85.52%, the average turn success rate of 85.32%, and the average turn length of 14.61. In a satisfactory survey targeting the subjects, the subjects agreed that our system makes them learn the second-language more efficiently and with less effort.

Keywords: computer-assisted second-language learning system, dialog-based CALL, task-oriented dialog system, language understanding, dynamic dialog graph based dialog management.

1. Introduction

Many methods from the field of natural language processing and dialog processing have been employed in CALL (Johnson & Valente, 2009; Kwon, Lee, Kim, & Lee,

1. Electronics and Telecommunications Research Institute, Daejeon, Korea; ohwoog@etri.re.kr
2. Electronics and Telecommunications Research Institute, Daejeon, Korea; kimyk@etri.re.kr
3. Electronics and Telecommunications Research Institute, Daejeon, Korea; yklee@etri.re.kr

How to cite this article: Kwon, O.-W., Kim, Y.-K., & Lee, Y. (2016). Task-oriented spoken dialog system for second-language learning . In S. Papadima-Sophocleous, L. Bradley & S. Thouësny (Eds), *CALL communities and culture – short papers from EUROCALL 2016* (pp. 237-242). Research-publishing.net. https://doi.org/10.14705/rpnet.2016.eurocall2016.568

2015; Wilske, 2015). The approaches engage the learner in a dialog and provide educational feedback, focusing on either form-based instruction or meaning-based instruction. Focus-on-form approaches tend to allow for relatively constrained input to provide more explicit feedback of linguistic forms and formal correctness about the learner's utterance (Kwon et al., 2015), whereas focus-on-meaning approaches tend to allow the learner to freely speak for providing opportunities for communication in the real world, but provide the less informative feedback (Wilske, 2015).

In this paper, we introduce a task-oriented spoken dialog system for providing the opportunities to use English in a situation that is similar to situations in real-life. A task-focused CALL system gives learners rich and realistic opportunities to practice achieving those tasks (Johnson & Valente, 2009). Our task-oriented spoken dialog system allows the learners to speak freely without a fixed scenario in a given task, then the learners can change the system initiative dialogue to their initiative dialogue. Our system provides no grammatical feedback and only provides the next recommended utterances for the learners after each system response. The proposed DB-CALL system focuses on meaning-based second language learning for the learners being familiar with English forms.

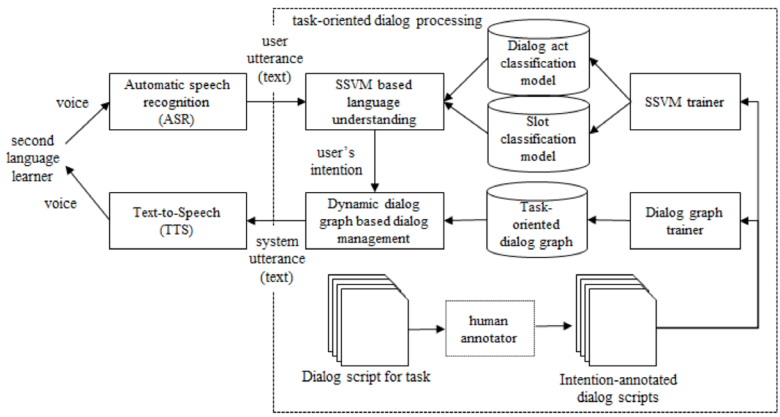
2. Task-oriented spoken dialog system for second-language learning

The system consists of Automatic Speech Recognition (ASR), task-oriented dialog processing, and Text-To-Speech (TTS) as shown in Figure 1. The ASR is optimized to recognize the English utterances of Korean learners as well as native speakers' utterances (Kwon et al., 2015) using a commercial English TTS engine to generate voices from system responses.

Our task-oriented dialog processing consists of a language understanding module using a Structural Support Vector Model (SSVM) (Lee & Jang, 2009) and a dynamic dialog graph based dialog management module. Our task-oriented dialog system uses the knowledge, which is trained from intention-annotated dialog scripts. Every dialog script is a real dialog between two persons who play a role of native (system) and learner (user) to accomplish the predefined task goal. We annotated every system and learner utterance with its intention using predefined dialog acts and slots. To classify user utterance with the predefined dialog acts and slots, the language understanding module trained dialog act classification model and slot classification model from user utterances of the intention-annotated corpus using SSVM, respectively. The dialog

management model automatically constructs the task-oriented dialog-graph from the intention-annotated corpus. The task-oriented dialog-graph consists of user and system intention nodes having slot history vectors and direct links between nodes with weights. The user and system nodes of the graph corresponds to the user and system intention of the dialog scripts. The edges between nodes present the dialog flow of the dialog scripts. The dialog management module finds the best user node similar to the intention of current user's utterance in the current dialog history, and then selects dynamically one among the system nodes linked by the best user node according to the level of difficulty for learning. We define the longer the length of dialogue is, the more difficult the level. Therefore, the management selects the system node on the shorter path into the terminal nodes when the level is easier. The dialog management module generates system responses using the example sentence of the selected system node. The management module memorizes the used frequency of each node during communication with a learner and selects the system node having the least frequency to provide the learner new situations.

Figure 1. Schematic diagram of the proposed task-oriented dialog system for second-language learning



3. Results and discussion

For the experiments, we constructed two task-oriented dialog systems, purchasing city tour bus tickets and passing through immigration. For each task, we firstly set up several situations, next gathered 100 dialog scripts between an English native (system) and a Korean English teacher (user), then English natives proofread the utterances of Korean teachers. Figure 2 shows an example of dialog script for

purchasing city tour bus tickets. Next, we annotated the utterances with system or user intentions. Finally, our task-oriented dialog system was trained from the intention annotated scripts.

Figure 2. Example of dialog script for purchasing city tour bus tickets

Task: Buying city tour bus tickets
Goal: Buy New York City Bus Tour tickets
Scenario condition: (1) 1 adult, 1 child (5 years old); (2) Take daytime tour; (3) Get two-day ticket; (4) Visit Statue of Liberty; (5) Get 1 audio guide, 1 Korean tour booklet; (6) Spend $120 or less

S: Hi there. Welcome to the New York City Bus Tour Center.
U: I want to buy tickets for me and my child.
S: What kind of tour would you like to take?
U: We would like to go on a tour during the day.
S: We have two daytime tours: the Downtown Tour and the All Around Town Tour.
U: Which tour goes to the Statue of Liberty?
S: You can visit the Statue of Liberty on either tour.
U: Can we get two-day tickets?
S: You can get two-day tickets for the All Around Town Tour.
U: Do you provide any guide materials?
S: We have an audio guide for 10 dollars and free tour booklets.
U: Do you have tour booklets in Korean?
S: Yes, we have Korean tour booklets.
U: How much are tickets for the All Around Town Tour?
S: Adult tickets are 60 dollars.
U: How much are tickets for children?
S: Children 5 years old and under pay 50 dollars.
U: Okay. I would like one adult ticket and one child ticket.
S: Is your child 5 years old?
U: Yes. I would like also an audio guide and a Korean tour booklet.
S: Okay. That will be 120 dollars.
U: Here you go.
S: Thank you. Enjoy the tour.

To evaluate the performance of the systems, we recruited 20 subjects: Group A with low proficiency levels (five subjects with Test of English for International Communication (TOEIC) scores lower than 500), Group B with middle proficiency levels (ten subjects with TOEIC scores between 500 and 800), and Group C with high proficiency levels (five subjects with TOEIC scores higher than 800). Each subject had a dialogue with the system to achieve a given mission. In the experiments, the subjects tried four missions for each task. If there are ASR errors, the subjects corrected the ASR errors by typing to only evaluate our proposed task-based dialog processing engine without ASR errors. Table 1 shows the results of the experiments. The experiments show the average task success rate of 85.52%, the average turn success rate of 85.32%, and the average turn length of 14.61. Contrary to our expectations, the success rates of Group C are lower than Group A and B. The reason is that the subjects of Group C tried the communication and intelligence abilities of the dialog system by speaking various utterances and driving the system to an unexpected situation, while the subjects of Group A and B made utmost efforts to accomplish their goals.

Table 1. The results of the experiments to evaluate the task-oriented dialog system

Measure	Subject group			
	Group A	Group B	Group C	All subjects
Avg. Task Success Rate (%)	93.75	89.90	72.92	**85.52**
Avg. Turn Success Rate (%)	86.71	87.09	82.17	**85.32**
Avg. Turn Length	14.71	13.80	15.32	**14.61**

We also investigated a satisfactory survey targeting the subjects about the second-language usefulness and efficiency. Table 2 shows the survey questionnaires and the average satisfactory scores of the subjects. The subjects indicated their level of agreement on each of the questionnaires. The levels of agreement are scored from one (strongly disagree) to six (strongly agree). From the results of the survey, most subjects agreed to that our system helps them learn English more efficiently and with less effort. Although the subjects of Group C did not faithfully take part in the experiments, the survey gained a better satisfaction score from Group C than those of the Group A and B. We assumed the reason is that our proposed task-oriented dialog system meets the requirement of the learners with high proficiency levels who want to exercise and acquire English through communicating spontaneously in authentic social situations.

Table 2. The results of a satisfactory survey

Survey questionnaires	Level of agreement (1 through 6)			
	Group A	Group B	Group C	All
Q1) The naturalness of the dialogue made the system so true to life.	4.20	4.42	4.63	4.41
Q2) The hints and suggestions are appropriate for the dialog context.	4.95	4.67	4.75	4.76
Q3) The dialog system helps you learn English more efficiently.	4.65	4.92	5.13	4.90
Q4) The ASR function helps you learn English more efficiently.	5.05	5.08	5.38	5.15
Q5) The TTS function helps you learn English more efficiently.	4.80	4.67	4.63	4.69
Q6) The system has some good features to help us learn English more efficiently and with less effort, compared to the previous methods.	5.15	5.25	5.25	5.23
Total	**4.80**	**4.83**	**4.96**	**4.86**

*Level of agreement : 1- strongly disagree, 2 - disagree, 3 - somewhat disagree, 4 - somewhat agree, 5 - agree, 6 - strongly agree

4. Conclusions

This paper described a DB-CALL system using a task-oriented dialog processing technology. The task-oriented dialog understanding and managing knowledge are automatically trained from the intention-annotated real-dialog scripts for accomplishing the task. In the experiments, the system showed a good performance with the average task success rate of 85.52% and the average turn success rate of 85.32%. Also, from the result of a satisfactory survey, we expected that our task-oriented dialog system might satisfy the needs of the second-language learners with high proficiency levels who want to communicate with natives in real-life situations.

5. Acknowledgements

This work was supported by the ICT R&D program of MSIP/IITP [R0126-15-1117, Core technology development of the spontaneous speech dialogue processing for language learning].

References

Johnson, W. L., & Valente, A. (2009). Tactical language and culture training systems: using AI to teach foreign languages and cultures. *AI Magazine, 30*(2), 72-83.

Kwon, O.-W., Lee, K., Kim, Y.-K., & Lee, Y. (2015). GenieTutor: a computer assisted second-language learning system based on semantic and grammar correctness evaluations. In F. Helm, L. Bradley, M. Guarda & S. Thouësny (Eds), *Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (pp. 330-335). Dublin Ireland: Research-publishing.net. https://doi.org/10.14705/rpnet.2015.000354

Lee, C., & Jang, M.-G. (2009). Fast training of structured SVM using fixed-threshold sequential minimal optimization. *ETRI Journal, 31*(2), 121-128. https://doi.org/10.4218/etrij.09.0108.0276

Wilske, S. (2015). *Form and meaning in dialog-based computer-assisted language learning*. Ph.D. thesis. Saarland University.

Research-publishing.net

Published by Research-publishing.net, not-for-profit association
Dublin, Ireland; Voillans, France, info@research-publishing.net

CALL communities and culture – short papers from EUROCALL 2016
Edited by Salomi Papadima-Sophocleous, Linda Bradley, and Sylvie Thouësny

Typeset by Research-publishing.net
Cover design by © Easy Conferences, info@easyconferences.eu, www.easyconferences.eu
Cover layout by © Raphaël Savina (raphael@savina.net)
Photo "bridge" on cover by © Andriy Markov/Shutterstock
Photo "frog" on cover by © Fany Savina (fany.savina@gmail.com)
Fonts used are licensed under a SIL Open Font License

ISBN13: 978-1-908416-43-8 (Paperback - Print on demand, black and white)
Print on demand technology is a high-quality, innovative and ecological printing method; with which the book is never 'out of stock' or 'out of print'.

ISBN13: 978-1-908416-44-5 (Ebook, PDF, colour)
ISBN13: 978-1-908416-45-2 (Ebook, EPUB, colour)

Legal deposit, Ireland: The National Library of Ireland, The Library of Trinity College, The Library of the University of Limerick, The Library of Dublin City University, The Library of NUI Cork, The Library of NUI Maynooth, The Library of University College Dublin, The Library of NUI Galway.

Legal deposit, United Kingdom: The British Library.
British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.

Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: décembre 2016.