Zahra Razavi

Department of Computer Science
University of Rochester
14627, Rochester, NY

[srazavi@cs.rochester.edu]

1 Research Interests

My research interests lie in **Natural Language Processing** and its applications, in particular topics related to **dialogue systems**, **conversation planning**, and leading natural flowing conversation. In order to construct a human-like conversational agent, I also do some research on **knowledge representation and reasoning**, including the use of **common sense knowledge** and on-line **knowledge extraction** from the user's input throughout the conversation.

1.1 Current Research

I am currently working with several collaborators on developing a dialogue system called LISSA which is designed to help people improve their social skills (Ali et. al., 2015). LISSA is an avatar that tries to lead a natural conversation with a user, while at the same time providing some feedback on the user's non-verbal behavior. The particular challenge I am working on is to enable the system to hold a natural open-ended conversation is the challenge I am working on. In order to conduct the conversation, LISSA needs to 1) understand what the user says and 2) provide a reasonable, appropriate reaction to the user's utterance. This is done through a structure called a schema which contains the knowledge needed by LISSA to carry out a high-level task such as making a small talk about general topics. A schema consists of a dynamically modifiable outline of the events LISSA expects to occur in the dialogue, plus a hash-table which stores the interpreted information gained during the conversation.

The first version of LISSA with the automated dialogue system has been evaluated in a speed-dating task, in which the output transcripts were compared to the ones from the same tests with Wizard-of-Oz dialogue manager. The results show almost no differences between the rating results of transcripts from automated and Wizard-of-Oz techniques, judged by multiple research assistants who did not know the source of the transcripts.

We are also running a study on using a variant of LISSA for helping teenagers with Asperger Syndrome to improve their communication skills. The exploratory study showed success in terms of user satisfaction with the system (Razavi et. al., 2016). More studies are ongoing in order to confidently evaluate the system.

1.2 Future Works

The future goal for the dialogue system I am working on is to make the dialogue even more natural and effective. An outline of ways of accomplishing this is as follows:

1- Reasoning based on the knowledge extracted from the user's input: As we collect some information from the user's utterances and add it to our knowledge base, we are planning to enable the system to make inferences from the available knowledge and use the results later during the conversation. This should give the user a greater sense of naturalness and of being understood by the system.

2- In support of the above technique, we plan to insert some world knowledge, so that the inference system could make more human-like inferences throughout the conversation.

3- Another possible improvement of the system is to integrate the verbal and nonverbal behaviors of the system. Currently, the feedback on the user's nonverbal behavior is shown with some screen icons while the conversation is going on. However, the more natural way is to make the avatar comment on the user's nonverbal behavior at some appropriate points during the conversation.

4- It is also possible to have the avatar show appropriate non-verbal behaviors based on its understanding of the users input. These non-verbal behaviors include head nods, smiling and backchannel utterances.

5- Automated turn-taking is a challenging issue in current spoken dialogue systems. Many systems have the user control the turn-taking by pressing a button; we do the same in LISSA. However, in a natural dialogue system, the system needs to recognize the end of a turn and take its own turn automatically.

2 Future of Spoken Dialog Research

Dialogue systems have recently been of special interest both in academy and industry. The successful deployment of the early versions of personal assistants (like Siri, Cortana, etc.) encouraged research in the area. However, the existing spoken dialogue systems are still command based; the area of open conversation leaves much room for further exploration explore. I believe that a major focus in dialogue systems research in 5 to 10 years will be to make systems more interactive, human-like and able to have open-ended conversation. In order to reach that point, the system should have a reasonable plan for each conversation, extract knowledge from the user's input and have a reasoning ability that it can use along with various kinds of world knowledge, allowing it infer new information as humans do during a conversation.

Besides functioning as personal assistants, spoken dialogue systems could play other roles, including companionship (e.g. Smith et. al., 2011). Companionable systems have many potential applications, such as helping people improve their skills (Ali et al., 2015), tutoring children (Foster et. al., 2010), helping doctors to diagnose some psychological disorders (DeVault et. al., 2014), and helping lonely people such as elderly people to overcome their feeling of isolation.

Another area which I believe will make a great stride is the integration of language with other modalities. Implementing spoken dialogue in virtual agents needs to combine vision, language, and non-verbal behaviors of the agent in a natural and human-like way. Robots need even more complicated integration, since they should make meaningful head and body movements in 3D space during the conversation.

3 Suggestions for discussion

- How to design a flexible dialogue plan for an open-ended conversation?
- How to extract knowledge from the user's inputs during the conversation? How to represent the extracted knowledge so that it can be used for reasoning during the conversation?
- How can we insert world knowledge into system and make it available throughout the conversation?
- How could we launch human-like reasoning based on the knowledge extracted from user inputs and on stored world knowledge to show more genuine understanding?
- How could we integrate verbal and non-nonverbal behaviors in virtual agents and robots with an automatic dialogue manager?

- How could we automate the turn-taking task in a conversation with a human, so that the machine could confidently recognize the end of turn and take the turn?

References

S. Zahra Razavi, M. Rafayet Ali, Tristram H. Smith, M. Ehsan Hoque and Lenhart L. Schubert. 2016. *Can chatting with a virtual human help Aspies improve conversational skills? An initial exploration,* Proceedings of the 16th International Conference on Intelligent Virtual Agents, LA, US (in press)

M. Rafayet Ali, Dev Crasta, Li Jin, et al. 2015. *LISSA- Live Interactive Social Skills Assistance,* In International Conference on Affective Computing and Intelligent Interaction (ACII) IEEE. 173-179.

David DeVault, Ron Artstein, Grace Benn, et al. 2014. *SimSensei Kiosk: A Virtual Human Interviewer for Healthcare Decision Support.* International Conference on Autonomous Agents and Multi-Agent Systems, 1061–1068.

Cameron Smith, Nigel Crook, Johan Boye, et al. 2011. *Interaction strategies for an affective conversational agent,* In International Conference on Intelligent Virtual Agents, Springer Berlin Heidelberg, 301-314.

Mary E. Foster, Katerina Avramides, Sara Bernardini, et al. 2010. *Supporting children's social communication skills through interactive narratives with virtual characters.* In Proceedings of the 18th ACM international conference on Multimedia, 1111-1114.

Biographical Sketch

Zahra Razavi is pursuing her PhD in Computer Science department at university of Rochester. As a second year PhD student, she is working in the Natural Language group under supervision of Professor Lenhart Schubert. She earned her Bachelor and Master in Electrical Engineering from Sharif University of Technology and University of Tehran respectively.