Juliana Miehle

Institute of Communications Engineering
Dialogue Systems
Ulm University
Albert-Einstein-Allee 43
89081 Ulm, Germany

juliana.miehle@uni-ulm.de
http://nt.uni-ulm.de/miehle

1 Research Interests

In the area of **Spoken Dialogue Systems** (SDS), my research interest lies in rendering those systems more flexible while at the same time allowing a high degree of user-adaptiveness. While current systems are task-dependent or plan-based, I want to integrate ontologies which model domain information. Using ontology-based reasoning, the current domain state may be detected and forwarded to the **Dialogue Management** module. However, the decision process of the Dialogue Manager should not only rely on the domain information, but also take into account the user's current state including their culture and emotion. Therefore, two steps have to be performed in the Dialogue Manager: first, a state has to be modelled which takes into account all relevant information concerning the domain and the dialogue state and the user state. Second, an appropriate system action has to be generated. To accomplish this second part, I am interested in **Statistical Dialogue Modelling** and **Machine Learning** in order to utilize a methodology which is able to deal with the uncertainty emerging from the state modelling.

1.1 IQ-adaptive Statistical Dialogue Management using Gaussian Processes

In my Master Thesis, I examined the impact of the incorporation of the Interaction Quality metric (IQ) during the process of learning a policy for a Spoken Dialogue System. The Interaction Quality metric is thereby a measure for the user's satisfaction as described in (Schmitt and Ultes, 2015) which is automatically estimated at each dialogue turn. Since the trained policy decides which actions are taken next, the dialogue flow is adapted to the user's satisfaction. For the process of learning, Gaussian Process based Reinforcement Learning has been used as proposed in (Gašić and Young, 2014).

Afterwards, I investigated whether the incorporation of the IQ metric is beneficial. Therefore, different learning strategies with and without the IQ metric have been used to train different policies. Then, the performance of all trained policies has been evaluated regarding dialogue completion, task success, the average length of a dialogue and the average IQ value at the end of a dialogue.

In summary, one can conclude that the results concerning the incorporation of the Interaction Quality metric into the process of training a policy have great promise for the adaptation of the dialogue to the user's satisfaction and the use of the IQ metric as a measure for it. Both the use of the final IQ value at the end of a dialogue to decide whether the dialogue has been successful or not and the use of the current IQ value in every dialogue turn to calculate the immediate reward with which the policy is trained yield good results.

For future work on the subject of IQ-adaptive dialogue, the same adaptation techniques should be tested with real users as they might give new insight by showing unseen behaviour. Furthermore, other alternatives regarding the incorporation of the Interaction Quality metric may be applied and tested, e.g. the integration of the IQ in the belief space. It has to be analysed whether the IQ value may be used as a measure of the task success and which way is the best to incorporate it into the process of learning a policy.

1.2 Cultural Communication Idiosyncrasies in Human-Computer Interaction

In (Miehle et al., 2016), we investigated whether the cultural idiosyncrasies found in human-human interaction may be transferred to human-computer interaction. With the aim of designing a culture-sensitive Spoken Dialogue System, we designed a user study creating a dialogue in a domain that has the potential capacity to reveal cultural differences. The dialogue contained different options for the system output according to cultural differences. We conducted an on-line survey on the user's preference concerning the different options among Germans and Japanese to investigate whether the supposed differences may be applied in human-computer interaction.

Our results show that there are indeed differences, but not all results are consistent with the cultural models for human-human interaction. This suggests that the communication patterns are not only influenced by the culture, but also by the dialogue domain and the user emotion. Moreover, it is shown that not all cultural idiosyncrasies that occur in human-human interaction may be applied for human-computer interaction.

In this work, only one specific dialogue has been considered. To get a more general view and exclude effects which may depend rather on the domain than on the culture, in future work other dialogues from different domains should be examined. Furthermore, we have to identify how the defined cultural idiosyncrasies may be implemented in the Dialogue Management to design a culture-sensitive Spoken Dialogue System.

1.3 Future Work

While I made some first investigations in the two fields of Statistical Dialogue Management and Cultural Communication Idiosyncrasies in Human-Computer Interaction, I now want to use these results and address the Dialogue Management decision process which should take into account the user state as well as the dialogue state. For the dialogue state, I plan to integrate ontologies. My first aim is to model a state which takes into account all relevant information. Afterwards, I plan to use machine learning approaches to generate appropriate system actions.

2 Future of Spoken Dialog Research

In my opinion, one of the major aspects regarding the future research on Spoken Dialogue Systems is to find means to render the systems more flexible that one Spoken Dialogue System may be used for more domains and for larger domains without the need of being trained for every single sub-domain. Furthermore, the systems need to be developed further to companion systems, supporting their user with any task they have, thereby taking into account their current situation, their cultural and social background and their current emotional state. To accomplish this, we have to make use of paralinguistic information identified by psycholinguists that facilitates natural language understanding. The use of affective speech, gestures, body language, back-channel responses, and timing facilitates human-human communication and should therefore also be used for human-computer interaction. For dialogues to be successful between humans and machines, our Spoken Dialogue Systems have to be able to interpret all of these cues and react in an appropriate way.

3 Suggestions for Discussion

- Statistical Dialogue Modelling: Are pure Statistical Dialogue Modelling and Machine Learning methods really the answer?
- Reward Modelling: How do we tell the Spoken Dialogue System during training what is good or bad?
- Evaluation: Are there golden rules for the evaluation of adaptive Spoken Dialogue Systems? How do we measure the success of adaptation?

References

Juliana Miehle, Koichiro Yoshino, Louisa Pragst, Stefan Ultes, Satoshi Nakamura, and Wolfgang Minker. 2016. Cultural Communication Idiosyncrasies in Human-Computer Interaction. In *Proceedings of the SIGDIAL 2016 Conference*. Association for Computational Linguistics.

Milica Gašić and Steve Young. 2014. Gaussian Processes for POMDP-Based Dialogue Manager Optimization In *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(1):28-40.

Alexander Schmitt and Stefan Ultes. 2015. Interaction Quality: Assessing the Quality of Ongoing Spoken Dialog Interaction by Experts—And How It Relates to User Satisfaction. In *Speech Communication*, 74:12-36.

Biographical Sketch

Juliana studied Electrical Engineering at Ulm University (Germany) with concentration in ”Communication and System Technology”. In 2013, she received her Bachelor of Science (B. Sc.) and in 2015, she graduated on ”IQ-adaptive Statistical Dialogue Management using Gaussian Processes” and got her Master of Science (M. Sc.). After doing an internship at Nara University of Science and Technology (Japan), she joined the Dialogue Systems Group at Ulm University under the supervision of Prof. Dr. Dr.-Ing. Wolfgang Minker in 2016 as a research assistant and PhD student. Her topic is centred around user-adaptiveness and the integration of ontologies into Statistical Dialogue Management.