

Introduction to the Special Issue on Dialog State Tracking

Jason D. Williams

JASON.WILLIAMS@MICROSOFT.COM

*Microsoft Research
One Microsoft Way
Redmond, WA, 98052, USA*

Antoine Raux

ARAUX@FB.COM

*Facebook
1 Facebook Way
Menlo Park, CA 94025, USA*

Matthew Henderson

MATTHEN@GOOGLE.COM

*Google
1600 Amphitheatre Parkway
Mountain View, CA 94043, USA*

Editor: David Schlangen

In the core of most task-oriented conversation systems is a component called a *dialog state tracker*, which estimates the user's goal given all of the dialog history so far. For example, in a weather information system, the dialog state might indicate the location the user is interested in (Seattle, London, Beijing), and for which date (today, tomorrow, this Saturday). Dialog state tracking is difficult because automatic speech recognition (ASR) and spoken language understanding (SLU) errors are common, and can cause the system to misunderstand the user. Even so, state tracking is crucial because the system relies on the estimated dialog state to choose actions – for example, which weather forecast to provide.

As conversational systems become a part of daily life – examples include Apple's Siri, Google Now, and Cortana from Microsoft – interest from the research community in dialog state tracking has grown considerably. Yet to date, work in dialog state tracking has been spread among a number of workshops and conferences. With this backdrop, this special issue has sought to provide thorough coverage of the state-of-the-art in this specialized field. Of 7 papers submitted, 4 were ultimately accepted for publication.

One of the main catalysts for increased in dialog state tracking recently has been a series of shared tasks called the Dialog State Tracking Challenge, and all papers in this special issue draw on these tasks. In these challenge tasks, detailed logs of human-computer dialogs are released, and teams build trackers that attempt to infer the correct dialog state at each turn. Results from all the systems that participate in each challenge have been published in special sessions at conferences and workshops (Williams et al., 2013; Henderson et al., 2014b,a). The first paper reviews work in dialog state tracking generally, then provides a cohesive review of the first three instances of the Dialog State Tracking Challenge, including a detailed description of the data, a compilation of results, and an analysis of the technical advances resulting from the first three challenges.

The second, third, and fourth papers explore emerging methods for dialog state tracking in some depth.

The second paper, *Spectral decomposition method of dialog state tracking via collective matrix factorization* by Julien Perez, observes that the task of dialog state tracking can be cast as a matrix factorization problem. This novel approach yields state-of-the-art performance on dialog data from the Second Dialog State Tracking Challenge.

The third paper, *Dialog History Construction with Long-Short Term Memory for Robust Generative Dialog State Tracking* by Byung-Jun Lee and Kee-Eung Kim, adopts an established generative-style state update, but estimates two of the core components using a “long short-term memory” (LSTM) neural network. The authors observe state-of-the-art performance on dialog data from the first and second Dialog State Tracking Challenge.

The fourth paper, *Recurrent Polynomial Network for Dialogue State Tracking* by Kai Sun, Qizhe Xie, and Kai Yu, introduces a computational recurrent network architecture based on polynomials. The advantage of this approach is that it allows expert knowledge to be encoded in the network, which substantially reduces the search space and thus the number of dialogs required to train the model. The authors observe state-of-the-art performance on dialog data from the second and third Dialog State Tracking Challenges.

Together these three papers provide a concrete illustration of the roles of machine learning – including both generative and discriminative methods – and the role of expert knowledge, for example hand-crafted rules.

We thank Prof David Schlangen, Editor-in-chief of this journal, for his constant support and attentiveness, as well as the editorial board of the Dialogue and Discourse journal for adopting this special issue. We also thank the reviewers very much for their careful work.

As this special issue goes to press, we note that the Dialog State Tracking Challenge series has continued, with the fourth Dialog State Tracking Challenge having recently concluded Kim et al. (2016), and planning for the fifth and sixth underway now.

This is an exciting period for spoken dialogue systems, both in terms of theoretical advances and commercial deployments. We hope this special issue provides a useful platform for its papers’ authors, and helps support the continuing advance of the state-of-the-art in this domain.

Jason D. Williams, Microsoft Research

Antoine Raux, Facebook

Mathew Henderson, Google

References

- Matthew Henderson, Blaise Thomson, and Jason D. Williams. The third dialog state tracking challenge. In *Proc IEEE Workshop on Spoken Language Technologies (SLT), South Lake Tahoe, Nevada, USA, 2014a*.
- Matthew Henderson, Blaise Thomson, and Jason D Williams. The Second Dialog State Tracking Challenge. In *Proc SIGdial Conf on Discourse and Dialogue, Philadelphia, USA, 2014b*.
- Seokhwan Kim, Luis Fernando DHaro, Rafael E. Banchs, Jason D. Williams, and Matthew Henderson. The fourth dialog state tracking challenge. In *Proc Intl Workshop on Spoken Dialog Systems (IWSDS), Saariselka, Finland, 2016*.

Jason D Williams, Antoine Raux, Deepak Ramachadran, and Alan Black. The Dialog State Tracking Challenge. In *Proc SIGdial Conf on Discourse and Dialogue, Metz, France, August 2013*.