Multimodal Contextual Dialog State Tracking As Bayesian Specific Signal Transduction

Birkner, Joseph
joseph.birkner@tum.de

Karimi, Negin
negin.karimi@tum.de

Dolp, Andreas
andreas.dolp@tum.de

Airiian, Wagram
wagram.airiian@tum.de

Kharchenko, Alona
unicorn@roboy.org

Hostettler, Rafael
rh@gi.ai

April 2019

1   Abstract

A major challenge in the development of Natural Language Dialog Systems is to determine the intent of a user utterance, and to map the intent of an utterance U to a certain dialog application state T. While recent work in this area focuses on embedding these variables as Neural-Network generated latent representations, we hypothesize that a symbolic approach to Dialog State tracking might deliver higher utility with reduced development effort: By observing dialog system state as words out of a formal language over signals in the application context, with application states acting as contextual non-terminals, we set up a basic formal framework for dialog state propagation. Futhermore, we propose the notion of constraint-based Bayesian state specificity as utility to resolve conflicts between overlapping application states. We implement our system in the open-source library "Ravestate". Experiments with the implemented system both in text- and speech based scenarios with additional video input show very robust contextual behavior, while operating fully causally explainable and transparently.

2 Introduction

3 Related Work

4 State Tracking As Signal Transduction

4.1 Terminology

States $t \in T$ (Processes, Transition Functions, Transducers, Non-Terminals)
Properties $p \in P$ (Data, Channels)
Signals $c \in C$ (Constraints, Chunks)
Spikes $\hat{s_c} \in \hat{S}$
Activations $\hat{a_t} \in \hat{A}$
CausalGroup spike equivalence class $[\hat{s_c}]$. 2

4.2 Simple Example

4.3 Core Protocol

4.3.1 EMIT(G, C)

4.3.2 WIPE(G, C)

4.3.3 ACQUIRE(G, A, C)

4.3.4 WITHDRAW(G, A, C)

4.3.5 •

5 Bayesian Specificity As State Utility

6 Experiments

7 Conclusion

8 Future Work

9 References