2023-09-01

A Framework for Confusion Mitigation in Task-Oriented Interactions

Na Li
Technological University Dublin, d19125334@mytudublin.ie

Robert J. Ross
Technological University Dublin, robert.ross@tudublin.ie

Funder: Science Foundation Ireland

A Framework for Confusion Mitigation in Task-Oriented Interactions

Na Li
Technological University Dublin
na.li@tudublin.ie

Robert Ross
Technological University Dublin
robert.ross@tudublin.ie

Abstract

Confusion is a mental state that can be triggered in task-oriented interactions and which can if left unattended lead to boredom, frustration, or disengagement from the task at hand. Previous work has demonstrated that confusion can be detected in situated human-robot interactions from visual and auditory cues. Therefore, in the next step, we propose appropriate interaction structures in this study, which should be used to mitigate confusion. We motivate and describe this dialogue mechanism through an information state-style dialogue framework and policies, and also outline the approach we are taking to integrate such a meta-conversational goal alongside core task-oriented considerations in modern data-driven conversational techniques.

1 Introduction

While we have a keen common sense intuition of what it means to be confused, the concept theoretically has only had some study in affective sciences: From a positive perspective, confusion is an effective response that occurs in people willing to explore new knowledge or tasks, but it is also an epistemic emotion that is associated with cognitive impasses when people try to solve problems (Lodge et al., 2018). The effects of the confusion state have been studied in online learning and driver assistance (Grafsgaard et al., 2011; Atapattu et al., 2020; Hori et al., 2016), but to date, the amount of research on confusion focused on the dialogue domain has been limited. One potential reason for the limited systematic study of confusion in the dialogue community may be that confusion is arguably better detected and more relevant in physically embodied interactions such as with robotic systems, although in this domain, research to date has been limited. In previous research (Li and Ross, 2023a), we have shown that it is possible to systematically identify users in a state of confusion, at least in a controlled study. If we can directly detect confusion as a cognitive state in interactions, the question then becomes: How should we train or otherwise adjust our dialogue policy to mitigate that confusion? Certainly, some of this mitigation would factor into the design cycle where we measure user confusion during initial interactions and adjust task designs to reduce the potential for confusion, but we also need to allow for the fact that confusion will occur (particularly in educational or training settings (D'Mello et al., 2014)) and that the conversational policies deployed must be able to dynamically adjust to the user in a confused state.

Given this challenge, in this paper we present a policy framework for the mitigation of confusion in task-oriented interaction. The policy framework builds on some design concepts from classical information state (IS) and dialogue acts representations from Dynamic Interpretation Theory (DIT), and Dialogue Act Markup in Several Layers (DAMSL). Our intuition for designing our dialogue framework builds on IS dialogue models and related toolkits such as TrindiKit, which is a dialogue move engine toolkit and the IBiS system (issue-based dialogue system) (Traum and Larsson, 2003; Larsson, 2002).

We first outline a set of relevant atomic information state and dialogue acts specifications; we then outline an information state structure including dialogue moves, and formalise the detailed dialogue policies corresponding to the dialogue acts. Following that, we illustrate the proposed approach using several scenarios as case studies. While the approach is very much a classical perspective, this is simply a stepping stone for us to providing aligned behaviours in data-driven policies.

2 Information State & Framework Design

In this study, the information state represents cumulative additions from previous actions in dialogue, and also the mechanisms to trigger dialogue moves for activating a corresponding dialogue act.

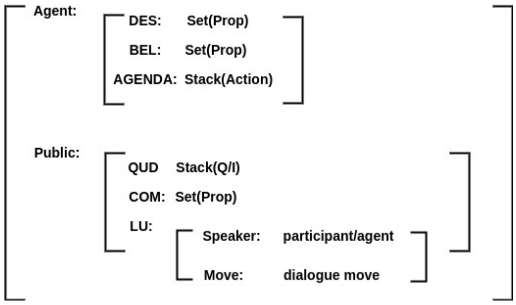

Figure 1: IS structure for confusion mitigation

Figure 1 presents an overview of the information state structure that we assume. The information state structure is typical of many other information state proposals such as IBiS, but for the sake of clarity, we briefly summarise for the unfamiliar reader. At a high-level the information state is split between a private grouping of state variables (Agent) which are internal to the agent and a public grouping of variables (Public) which the dialogue model assumes are shared between both agents. Within the private entities, the field /Agent/DES (desire) is a set of propositions (prop) that are used to capture the goals that the agent wishes to achieve. The field /Agent/BEL (believe) is a set of propositions that are directly correlated to the task that is taken to be true. Finally, the field /Agent/Agenda is a stack of plans which the agent intends to enact in order to achieve dialogue goals or otherwise lead to manipulate the mental state.

Turning to the public elements of the information state, the field /Public/QUD is a stack of questions under discussion (QUD). The QUD encompasses the ordering of unresolved questions or tasks to be confirmed that have been raised within the dialogue. The field /Public/COM includes a set of propositions that the user and the agent have committed to in the dialogue. It is not necessary for discourse participants to genuinely believe in those propositions, but discourse participants should have made a commitment to those statements for the objectives of the conversation. Finally, the field /Public/LU simply captures the last utterances in terms of the speakers and the specific dialogue moves associated with the utterances and the specific dialogue moves associated with the utterance.

Building on Larsson (2002)'s IBiS1 model, our dialogue moves are coarse-grained operations that trigger updates to the information state and the selection of relevant dialogue acts. Therefore, we designed ten dialogue moves and nine dialogue acts in our technical report (Li and Ross, 2023b), which can be applied across four information types (*i.e.*, statement, feedback, generic, and interface), to operationalise a policy to mitigate user confusion states. Our technical details include two tables (*i.e.*, Table 1 and Table 2) in Li and Ross (2023b) outline the general form of communication updates associated with these dialogue acts and the specific updates related to confusion states, respectively.

In that report, we detail a dialogue management process that is based on these definitions. A confusion detection model is assumed and integrated into the dialogue framework for real-time detection of the user's confusion states. Our model assumes semantically distinct levels of productive confusion, unproductive confusion, and non-confusion. When a confusion state is detected, this aspect of the dialogue policy becomes active. This structuring is in accordance with similar elements of communicative management in those moves and acts are selected to achieve the interaction goal of mitigating the user's confusion state. When an interlocutor is not manifesting confusion behaviours, the dialogue policy proceeds with those moves and acts associated with task progression as outlined. Moreover, we also present a task-oriented dialogue scenario in that report with associated updates of dialogues to help elucidate the policy presented.

3 Discussion & Outlook

In this paper, the proposed models and the underlying report have been designed and applied at a conceptual and empirical level in part of our human-avatar and human-robot studies. While the key motivators for these earlier studies were (a) whether confusion states can be induced; and (b) whether it is possible to detect confusion states extraverbally.

The policy presented here is to highlight one way in which we can identify and mitigate confusion as a pragmatic phenomenon that can be identified. While the benefit of a controlled dialogue flow remains important, we do recognise the importance of folding in the goals of embodied structured conversation with the naturalness and task-oriented appeal of integration with large language model-based solutions. Although the current proposal is still embryonic and not in a state where it can be systematically evaluated, we believe that the study of pragmatic effects in embodied systems presents an important next step for the study of the semantics and pragmatics of dialogues.

Acknowledgements

This publication has emanated from research conducted with the financial support of Science Foundation Ireland under Grant number 18/CRT/6183.

References

Thushari Atapattu, Katrina Falkner, Menasha Thilakaratne, Lavendini Sivaneasharajah, and Rangana Jayashanka. 2020. What do linguistic expressions tell us about learners' confusion? a domain-independent analysis in moocs. *IEEE Transactions on Learning Technologies*, 13(4):878–888.

Sidney D'Mello, Blair Lehman, Reinhard Pekrun, and Art Graesser. 2014. Confusion can be beneficial for learning. *Learning and Instruction*, 29:153–170.

Joseph F Grafsgaard, Kristy Elizabeth Boyer, and James C Lester. 2011. Predicting Facial Indicators of Confusion with Hidden Markov Models. Technical report.

Chiori Hori, Shinji Watanabe, Takaaki Hori, Bret A. Harsham, JohnR. Hershey, Yusuke Koji, Yoichi Fujii, and Yuki Furumoto. 2016. Driver confusion status detection using recurrent neural networks. In *2016 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6.

Staffan Larsson. 2002. Issue-based dialogue management.

Na Li and Robert Ross. 2023a. Hmm, you seem confused! tracking interlocutor confusion for situated task-oriented hri. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '23, page 142–151, New York, NY, USA. Association for Computing Machinery.

Na Li and Robert Ross. 2023b. Technical report: A framework for confusion mitigation in task-oriented interactions.

Jason M. Lodge, Gregor Kennedy, Lori Lockyer, Amael Arguel, and Mariya Pachman. 2018. Understanding Difficulties and Resulting Confusion in Learning: An Integrative Review. *Frontiers in Education*, 3.

David Traum and Staffan Larsson. 2003. The information state approach to dialogue management. *Springer Netherlands eBooks*.