
Articles

2023-08-04

Technical Report: 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

Follow this and additional works at: <https://arrow.tudublin.ie/creaart>



Part of the [Computer Engineering Commons](#), [Linguistics Commons](#), and the [Other Social and Behavioral Sciences Commons](#)

Recommended Citation

Li, N., & Ross, R. J. (2023). Technique Report: A Framework for Confusion Mitigation in Task-Oriented Interactions. Technological University Dublin. DOI: 10.21427/C3R6-RB43

This Technical Report is brought to you for free and open access by ARROW@TU Dublin. It has been accepted for inclusion in Articles by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, vera.kilshaw@tudublin.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 International License](#).
Funder: Science Foundation Ireland

Technical Report

A Framework for Confusion Mitigation in Task-Oriented Interactions

Na Li and Robert Ross
Technological University Dublin
{na.li, robert.ross}@tudublin.ie

August, 2023

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. Since previous work has demonstrated that confusion can be detected in embodied situated interactions from visual and auditory cues, in this technique report, we propose appropriate interaction structures which should be used to mitigate confusion. We motivate and describe this dialogue mechanism through an information state-style policy with examples, 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. While the current policy design is a starting point, we believe it raises some interesting challenges for the integration of a reusable meta-conversational policy with highly data-driven approaches which have been enabled by large language models.

Keywords: Confusion mitigation, situated dialogues, dialogue framework, information state, dialogue acts, dialogue policy

1 Introduction

Confusion is an effective response that occurs in people willing to explore new knowledge or understand something, but it is also an epistemic emotion that is associated with cognitive impasses while people try to solve problems or issues [Lodge et al. \(2018\)](#). The effects of confusion state have been studied in the online learning system and driver assistance [Grafsgaard et al. \(2011\)](#); [Atapattu et al. \(2020\)](#); [Hori et al. \(2016\)](#), but the amount of research on confusion focused on the dialogue domain to date has been limited.

In this technique report, we present a policy framework for minimising confusion in task-oriented interaction. Specially, 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 as we present it here is very much a classical perspective, this is simply a stepping stone for us

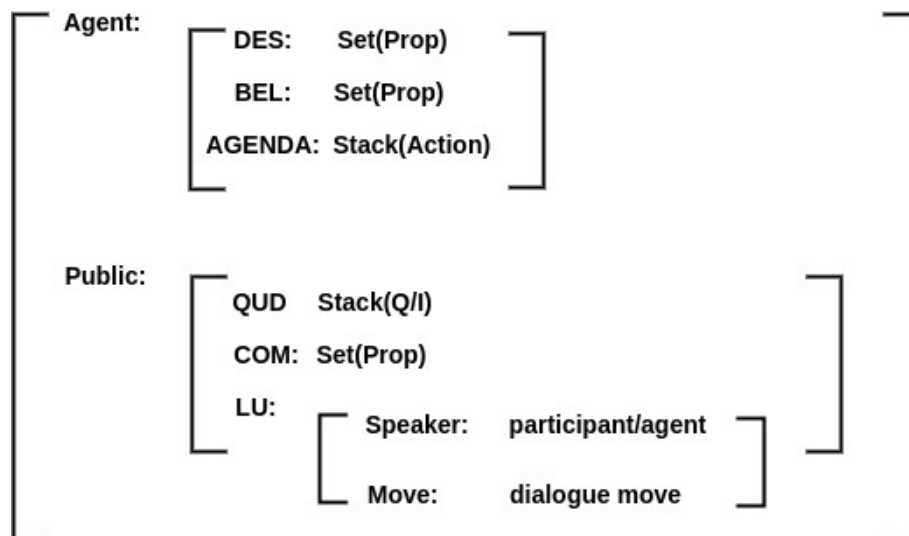
to providing aligned behaviours in dialogue policies with a data-driven confusion detection model. Therefore, in the final section we outline the approach we are taking to this development.

2 Dialogue Framework for Confusion Mitigation

In this section, we present a semi-formalised dialogue framework including our information state architecture, dialogue acts, and dialogue moves for confusion mitigation. It should be noted that this work assumes the existence of the confusion state estimation algorithm from our previous confusion detection study (Li and Ross, 2023).

2.1 Information State Definition

In the context of our research, 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. Our focus is on the aspects of this that relate to the monitoring and subsequent mitigation of a particular mental state, *i.e.*, the confusion level; however, the general principles shown should be compatible with a range of specific theories of task-orientated interaction, as a basis building upon Larsson (2002)’s IBiS1 information state.



*DES: Desire, BEL: Believe, QUD: Question Under Discussion, COM: Commitment, LU: Last Utterance, Prop: proposition

Figure 1: Information State 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 from the literature, 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 is a set of propositions that are used to capture the goals that the agent wishes to achieve. The field /Agent/BEL 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.

2.2 Move and Act Definitions

Again, building on Larsson (2002)'s IBiS1 model, we must define a set of dialogue moves and dialogue acts. The dialogue moves are coarse grained operations which trigger updates to information state as well as the selection of relevant dialogue acts. For the purpose of explaining our current model, we introduce seven basic dialogue moves as follows:

1. **ask(q)**, where q is a question asked or a task requested.
2. **answer(a)**, where a is the answer provided.
3. **restate(q/i)**, where q or i are, respectively, a question asked or information provided relevant to a task.
4. **ack(a)**, where a is the agent who agrees that the user's answer is correct.
5. **confirm(help)**, the agent needs to receive a confirmation from the user to help them.
6. **affirm(q)**, where q is the agent who admits that the question or the task is difficult.
7. **downdate(q)**, where q is the agent that reduces the difficulty of the question or the task asked by the agent.

as well as three more moves that are specific to the management of confusion state conditions:

1. **inforExten(q)**, the agent provides extent information on the question to help the user overcome their confusion.
2. **inforSply(q)**, the agent provides more information to the user to help answer the question.
3. **Freetalk(sbj)**, where sbj is a new subject with which the agent has a casual conversation with the user.

Turning to the finer-grained dialogue acts, we propose nine dialogue acts which can apply across four information types (*i.e.*, statement, feedback, generic, and interface), to operationalise a policy to mitigate user confusion states:

1. **Restatement** (statement): The agent repeats the original task or question.
2. **Feedback request** (feedback): The agent asks for the participant’s feedback or their response.
3. **Information extension** (statement): The agent provides more information to expand on the question or the task that has already been raised.
4. **Information supplement** (statement): The agent provides comprehensive information for participants to quickly and easily understand.
5. **Confirmation** (generic): The agent confirms whether the user needs help to overcome their confusion.
6. **Affirm** (generic): The agent acknowledges that the question is challenging and might cause the participant to become confused.
7. **Acknowledge** (generic): The agent acknowledges that the user is correct with positive feedback.
8. **Subject change** (statement): The agent changes a straightforward question or a topic.
9. **Notify confusion** (interface): The agent detects user’s confusion state in real-time, that is, productive confusion, unproductive confusion, or non-confusion.

Table 1 outlines the general form of communication updates associated with these dialogue acts.

Table 1: Communicative Update Rules associated with Dialogue Acts

Dialogue Acts	Communication Rules
Restatement	Repeat the question/information either at the same speed or more slowly.
Feedback request	Option 1: Ask the participant whether they can follow what the agent has said.
	Option 2: Ask the participant whether it is difficult for them to answer this question.
	Option 3: Ask the participant whether they want to continue to answer this question by themselves.
Information extension	Provide more explanations to fix the issued questions or lost information.
Information supplement	Provide the full information / question in different ways to easily understand without confusion.
Confirmation	Get a help confirmation from the participant to overcome their confusion.
Affirm	Affirm that this information or question has issues or is difficult leading to the participant cannot answer.
Acknowledge	Acknowledge the participant’s response correctly to remove the participant’s source of confusion.
Subject change	Option 1: Raise a simple question that the participant can answer without confusion.
	Option 2: Bring up another interesting topic to arise the participant’s engagement.

Figure 2 illustrates an extract of the 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 that moves and dialogue acts are selected to achieve

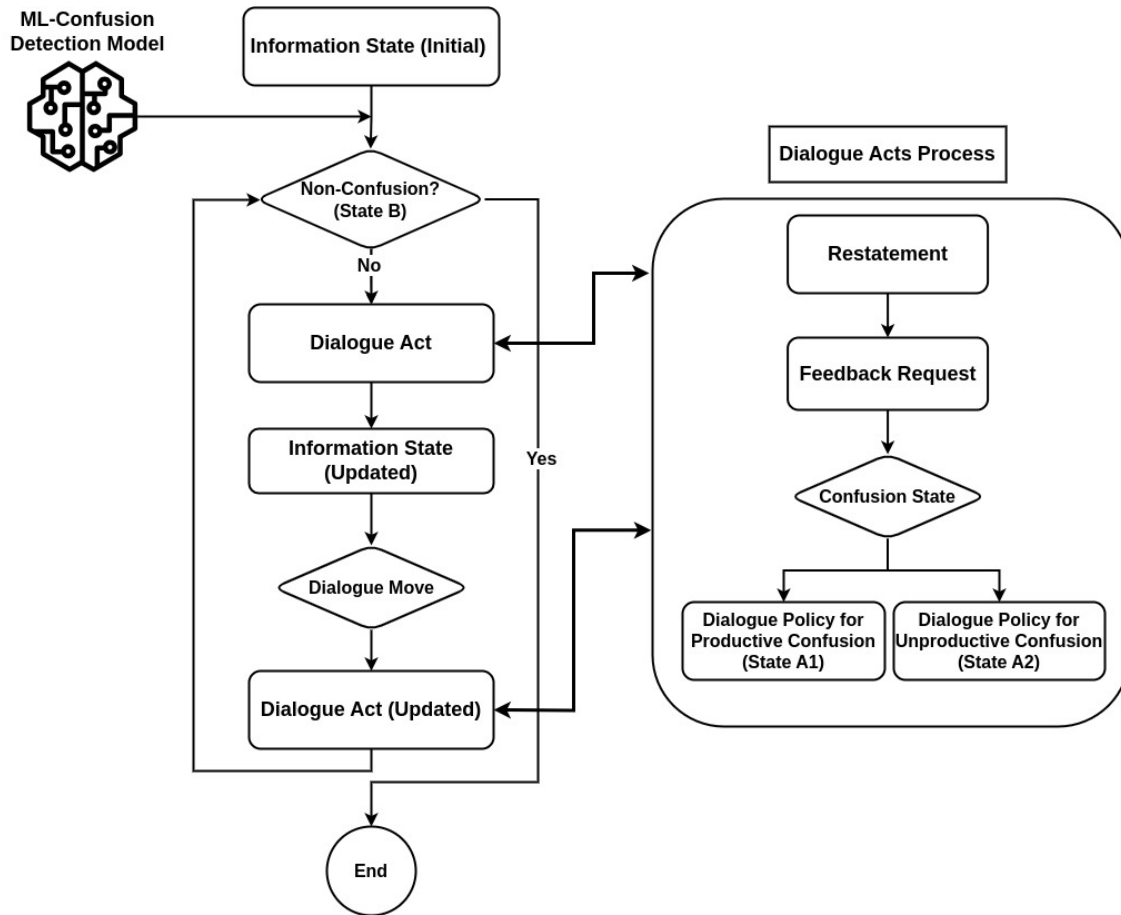


Figure 2: The working process of the dialogue framework for confusion mitigation

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 for example in Larsson (2002)’s IBiS1 model.

According to our basic dialogue policy design (Li and Ross, 2022), we can drill into the specifics of this approach by examining in Table 2 the specific updates associated with confusion indicators. Specifically, we indicate the dialogue acts used to resolve the user’s current confusion state. A key point here is that this policy makes policy choices based on (a) whether the participant is likely in a productive or unproductive confusion state; and (b) what is most likely to be the induction (cause) of confusion, *i.e.*, complex information, insufficient information, or others. Note that in our earlier work we highlight contradictory information and inconsistent feedback as being potential testable inductions for confusion states. However, the current model is focused only on what can be operationalised into a working system, and the reality is that it is difficult to estimate if a user’s confusion might be due to contradictory information or false feedback, as by definition it is hard for a system to determine if it itself has made a mistake in its interactions with the user. Thus, we generalise the two induction categories Contradictory Information and False Feedback to a single - general induction category.

Verbal signalling would certainly be able to identify false feedback and contradictory information, which are two causes of confusion. However, it should be noted that our modelling is currently not attempting to model explicit verbal signalling of confusion causes from the user. Based on studies in the literature, we believe this

Table 2: Dialogue policy for mitigating different confusion states

Policy of productive confusion mitigation		
Confusion Inductions	Dialogue Acts	Communicative Detail
Complex infor*	Restatement	The agent will reintroduce the complex information step by step.
	Feedback request	The agent will ask whether the participant is clear on the issue or question.
	Confirmation	The agent will confirm whether the participant needs a help to answer this question or complete the task.
	Information supplement	The more and extra information will be told to the participant.
Insufficient infor	Affirm	The agent will affirm the question / information is insufficient.
	Confirmation	The agent will confirm whether the participant needs help to answer this question or complete the task.
	Information extension	The agent will provide the lost part of the question / information.
	Information supplement	The more and extra information for this question or task will be told to the participant.
Others	Acknowledge	The agent will show positive and correct feedback.
	Affirm	The agent will affirm the question / information has issues.
	Confirmation	The agent will confirm whether the participant needs help to answer this question or complete the task.
	Information extension	The agent will correct the question / information to be consistent.
	Information supplement	The more and extra information will be told to the participant.
	Feedback request	The agent will ask the participant whether they want to continue to answer this question or complete this task by themselves.
	Subject change	The agent will talk about a simple question or task with positive feedback.
Policy of unproductive confusion mitigation		
Confusion Inductions	Dialogue Acts	Communicative Detail
Complex infor* Insufficient infor Others	Affirm	The agent will confirm that the question / task is difficult which has led to the participant being confused.
	Information supplement	The more and extra information for this question or task will be told to the participant.
	Feedback request	The agent will ask whether participants want to continue to answer this question or task by themselves.
	Acknowledge	The agent will show a positive and correct response or feedback.
	Confirmation	The agent will confirm whether the participant needs help to answer this question or complete the task.
	Subject change	Option 1: The agent will ask a straightforward question or talk about a simple task to avoid confusion states. Option 2: The agent will talk about another interesting topic <i>e.g.</i> favourite food, movie, <i>etc.</i>

* Infor: Information

to be the less interesting manifestation of confusion and that truly natural systems should be able to focus on picking up on such communicative indications without the need for direct verbal interventions by the user if not needed.

3 Dialogue Example: Simulating a Question and Answer Dialogue

To help elucidate the policy presented, in this section, we present a task-oriented dialogue scenario with associated updates of dialogues. This scenario is focused around a set of verbal understanding tasks that we have used in large scale Human-Robot Interaction studies (Li and Ross, 2023) to induce and monitor confusion states in users. In this scenario below, productive confusion is labelled “State A1”, unproductive confusion is labelled “State A2”, while “State B” refers to non-confusion state. The task-oriented dialogue scenario is as follows.

SYSTEM: I will ask you an interesting word problem. If you do not know the answer, please ask me to repeat it again, or you can give up answering anytime.

USER: okey.

SYSTEM: Jessica plans to make 6 litres of blackberry soda and 2 litres of lemon-lime soda for the class party. She also plans to make 18 batches of biscuits and 50 chocolate bars. She will buy 40 bags of crisps and 60 bags of candy. Due to her budget issue, she can only buy 10 batches of biscuits and make 2 more litres of blueberry soda. One litre of soda will serve 5.5 students at the party. The question is how many students can Jessica serve her soda to?

...: <The system observes user entering confused state (A1).>

SYSTEM: Hey my friend, do you need me to repeat this question?

USER: yes, please.

SYSTEM: <The system repeats the question.>

USER: Sorry, I don't know the answer.

...: <System observes continued confused state (A1)>

SYSTEM: <The system repeats the question slowly.>

USER: Sorry, I still don't know the answer.

...: <The user does not respect. System detects user's confusion state (A2)>

SYSTEM: I am sorry, maybe this question is too difficult, do you need my help to answer it?

USER: Yes, please.

...: <The user wants to solve this problem, and the system detects the user confusion state (A1)>

SYSTEM: <The system restates this word problem again with the extra information supplied step by step, slowly and briefly.>

First, Jessica plans to make 6 litres of blackberry soda and 2 litres of lemon-lime soda for the class party. Second, she also plans to make 2 more litres of blueberry soda. One litre of soda will serve 5.5 students at the party. The question is how many students can Jessica serve her soda to?

USER: The answer is 55.

SYSTEM: Yes, you are correct.

...: <The system observes user enters a non-confusion state (B)>

In this dialogue scenario, the first turn in the dialogue involves the agent posing a verbal problem to the user. The specific nature of the task here is not relevant – rather that is part of some embodied task-oriented interaction. In the second turn, the user acknowledges the problem given, and thereafter the system continues to monitor the user’s state. After a period, the system determines that the user is displaying indications of having entered productive confusion State A1. The system then restates the task to the user. Following a period, the system observes that the user remains in a confused state (A1), at which time the agent attempts to provide more help to answer the question.

However, the system first observing that the user has entered state A2, at which time the agent asks the user whether need help to continually answer the question, and the agent affirms that this question is difficult. Following the time that the system observes, the user re-enters in a confused state (A1). In the next step, with the help of the system, the user solves this problem with the correct feedback from the system. Finally, the user enters the non-confusion state (B). Such a policy is equivalent to providing more information or changing a topic and moving the conversation on, but may, of course, not always be appropriate depending on the specific nature of the task. Following seven detailed steps through of the associated updated information state policy with dialogue acts.

1. <ask_question_to_usr>

Information state:

[Agent/BEL: {NULL}]

[Agent/DES: {nonconfusion(usr)}]

[Agent/AGENDA: <notify confusion(usr)>]

[Public/QUD: <?a.wordproblem(agent)>]

[Public/COM: {State A1}]

[Public/LU: {Speaker: agent; Move: {answer(wordproblem(usr))}]

Dialogue Act: {notify confusion(usr)}

2. <ask_question_to_usr_A1>

Information state 2.1:

[Agent/BEL: {State A1(usr), tohelp(agent)}]

[Agent/DES: {nonconfusion(usr), tohelp(agent)}]

[Agent/AGENDA: <restate(wordproblem(agent))>]

[Public/QUD: <?a.wordproblem(agent)>]

[Public/COM: {State A1}]

[Public/LU: {Speaker: usr; Move: {confirm(help)}]

Dialogue Act 2.1: {confirmation(usr)}

Information state 2.2:

[Agent/BEL: {State A1(usr), tohelp(agent)}]

[Agent/DES: {nonconfusion(usr), tohelp(agent)}]

[Agent/AGENDA:

<inforExten(wordproblem(agent))>]

[Public/QUD: <?a.wordproblem(agent)>]

[Public/COM: {State A1}]

[Public/LU: {Speaker: agent; Move: {restate(wordproblem(agent))}]

Dialogue Act 2.2: {restatement(agent)}

3. <ask_question_to_usr_insufficient>

Information state:

[Agent/BEL: {State A1(usr), tohelp(agent)}]

[Agent/DES: {nonconfusion(usr), tohelp(agent)}]
 [Agent/AGENDA: <inforSply(wordproblem(agent))>]
 [Public/QUD: <?a.wordproblem(agent)>]
 [Public/COM: {State A1}]
 [Public/LU: {Speaker: usr; Move: {inforExten(wordproblem(agent))}}]
Dialogue Act: {information extension(agent)}
4. <ask_question_to_usr_complex>
Information state:
 [Agent/BEL: {State A1(usr), tohelp(agent)}]
 [Agent/DES: {nonconfusion(usr), tohelp(agent)}]
 [Agent/AGENDA: <cofirm(answer(wordproblem)), affirm(wordproblem(agent))>]
 [Public/QUD: <?a.wordproblem(agent)>]
 [Public/COM: {State A1}]
 [Public/LU: {Speaker: agent; Move: {inforSply(wordproblem(agent))}}]
Dialogue Act: {information supplement(agent)}
5. <ask_question_to_usr_complex_2>
Information state:
 [Agent/BEL: {State A2(usr), tohelp(agent)}]
 [Agent/DES: {nonconfusion(usr), tohelp(agent)}]
 [Agent/AGENDA: <downdate(wordproblem)>]
 [Public/QUD: <?a.wordproblem(agent)>]
 [Public/COM: {State A2}]
 [Public/LU: {Speaker: usr; Move: (answer(wordproblem)), affirm(wordproblem(agent))}]
Dialogue Act: {feedbackrequest(usr), affirm(agent)}
6. <ask_question_to_usr_state A2>
Information state:
 [Agent/BEL: {State A2(usr), tohelp(agent)}]
 [Agent/DES: {nonconfusion(usr), tohelp(agent)}]
 [Agent/AGENDA: <notify confusion(usr), ack(answer(usr))>]
 [Public/QUD: <?a.wordproblem(agent)>]
 [Public/COM: {State A2}]
 [Public/LU: {Speaker: agent; Move: {downdate(wordproblem)}}]
Dialogue Act: {notify confusion(usr)}
7. <ask_question_to_usr_state B>
Information state:
 [Agent/BEL: {State B(usr)}]
 [Agent/DES: {nonconfusion(usr)}]
 [Agent/AGENDA: <NULL>]
 [Public/QUD: <?a.wordproblem(agent)>]
 [Public/COM: {State B}]
 [Public/LU: {Speaker: usr; Move: {ack(answer(usr))}}]
Dialogue Act: {notify confusion(usr)}

4 Discussion

The models proposed in this report have at a conceptual and empirical level been validated and used in part of a series of human-avatar and human-robot based studies. The policy as presented here has shown that we can identify and mitigate confusion as a pragmatic phenomenon which can be identified extra-verbally. We chose an information state style description of the policy as it is mostly transparent and can

easily be mapped to concrete implementation approaches and dialogue management methods such as dialogue flows and frame-based methods. Therefore, we also see the proposal as highly reusable across domains and even modalities of interaction, *e.g.*, game with a virtual agent [Heyselaar et al. \(2017\)](#).

Although the benefit of a controlled dialogue flow remains important, we do however 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. To this end, one particular fruitful branch of future work will be to integrate these policies alongside large language model driven systems through a process of model alignment.

5 Conclusion

In this technique report, we presented a semi-formalised framework of task-oriented dialogue for confusion mitigation, which is expressed through the concept of information states and the IBS1 model in particular. This policy is a rationalisation and operationalisation of an earlier proposal which simply focused on what confusion states may be induced in a user. To illustrate, we included a dialogue example based on concrete human-robot interactions in our task domain. While the current proposal is, of course, still embryonic and not at a state where it can be systematically evaluated, we argue that investigating pragmatic effect in embodied systems constitutes a crucial advancement in understanding 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. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

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.
- Joseph F Grafsgaard, Kristy Elizabeth Boyer, and James C Lester. 2011. Predicting Facial Indicators of Confusion with Hidden Markov Models. Technical report.
- Evelien Heyselaar, Peter Hagoort, and Katrien Segaert. 2017. [In dialogue with an avatar, language behavior is identical to dialogue with a human partner](#). *Behavior Research Methods*, 49.
- 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. 2022. [Dialogue policies for confusion mitigation in situated hri](#). In *Proceedings of the 26th Workshop on the Semantics and Pragmatics of Dialogue - Poster Abstracts*, Dublin, Ireland. SEMDIAL.
- Na Li and Robert Ross. 2023. [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.
- 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.