FRICODILE: Providing FRICO with Dialogue Capability

Prakash Jamakatel,[1] Rebecca de Venezia,[2] Christian Muise,[3] Jane Jean Kiam[1]
[1]University of the Bundeswehr Munich, Germany
[2]Queen's University, Canada

Abstract

Besides observing the pilot passively and providing contextual assistance in light of carrying out various flight tasks more efficiently, engaging the pilot in a dialogue creates an opportunity to go beyond passive monitoring of pilot's activities and flight status, by actively requesting for input information from the pilot, enriching therefore the knowledge of the pilot assistance system on the context. Such a dialogue has to be dynamic, concise and goal-directed, as it can be safety critical. In this paper, we demonstrate how FOND-planning-based dialogue agents can be applied to the use case, while also leveraging recent advances in natural language understanding and text-to-speech synthesis, in order to generate instances for the computation of plans as guidance to the pilot in an automated manner problem.
Video: youtu.be/1nOUfo2ENQk

Introduction

Pilot error is a probable cause in 85 % of general aviation crashes (Li et al. 2001). The laxer regulatory hurdles for obtaining pilot licenses (Council of European Union 2012) coupled with the missing assistance system (BFU 2022) have been identified as accountable factors. Previously, we demonstrated FRICO: An AI-Enabled **Fri**endly **Co**ckpit Assistance System (Jamakatel, Morchedi, and Kiam 2023), an assistance system capable of providing contextual guidance to pilots using Plan and Goal Recognition (PGR) as hierarchical planning (Höller et al. 2018) and probabilistic state estimation (Jamakatel et al. 2023). The guidance, either in the form of strategic plans (i.e. high-level tasks) or tactical plans (i.e. flight path), is *contextual*, as FRICO derived flight state and pilot's intents using passive in-cockpit sensing to obtain measurement data such as flight data, pilot's gaze, pilot's interactions with the instrument panel, etc.

To better integrate the pilot into the decision-making loop, we extend the ability of FRICO so that it can actively engage the pilot in a conversation to supplement the passively inferred context. This dialogue aids in making the passively inferred context richer while increasing the involvement of the pilot decision-making process. The goal-directedness of the dialogue and a guarantee that a rational interlocutor will reach the dialogue goal is essential in this safety-critical domain. LLMs-powered dialogue agent, because of their tendency to hallucinate, are ill-suited in such domains. Recently, Plan4Dial (De Venezia and Muise 2023), an open-source FOND-planning-based goal-directed dialogue-agent framework, was introduced. We use this Plan4dial framework to provide **FRICO** with **di**a**l**oug**e** and introduce FRICODILE in this paper based on the works published in (Jamakatel et al. 2024). Based on information obtained from the pilot through the dialogue and the passively inferred context, FRICODILE infers the context used to generate the problem instance to be solved with a hierarchical planner.

Dialogue Planning

In the dialogue planning framework, a dialogue tree is synthesized using automated planning. This forgoes the unpredictable results that might occur in end-to-end neural networks and is better maintainable than hand-coding complex dialogue trees (Muise et al. 2019). In this framework, the dialogue is abstracted using a PDDL model, which contains information only on how to drive the conversation based only on what the dialogue agent knows (Muise et al. 2019). Using Fully Observable Non-Deterministic (FOND) planning, the specification that an outcome should be selected from multiple possible outcomes in a dialogue turn can be modelled (De Venezia and Muise 2023). Once the dialogue is specified using a PDDL model, it can be solved using a FOND planner to obtain the corresponding dialogue tree as contingent plans.

System Overview

Figure 1 demonstrates our system-level approach for generating strategic plans based on dialogue. The modules marked in blue are taken over from FRICO (Jamakatel, Morchedi, and Kiam 2023). This includes the action recognition module, which infers the pilot's actions by combining various sensor data and the plan recognition module, which infers the pilot's current goal. This is referred to as the passive context.

The dialogue in FRICODILE is carried out primarily using speech channels. The audio signals received from the pilot are transcribed into text and then fed as input to the dialogue executor. The dialogue executor evaluates the utterances of the audio signal using a natural language understanding (NLU) system to extract what the pilot wants to achieve with the utterance (intent) and which parts of the utterances are important in the context (entities). The response is then determined using a pre-defined dialogue model. This response is then transformed into audio signals using a Text-To-Speech (TTS) system and then communicated to the pilot. Once the local goal of the conversation has been reached, the problem instance for strategic plan generation is set. Then strategic plans for various flight tasks based on the extended version of the previously published ultralight domain (Kiam and Jamakatel 2023) using a hierarchical planner is generated.

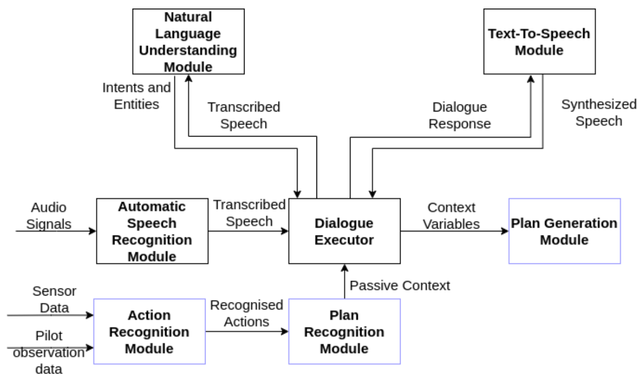

Figure 1: System Architecture of FRICODILE (Modules taken over from FRICO in blue)

Technical Implementation

We refer to (Jamakatel et al. 2023) for details on the technical implementation of the passive context recogniser. We describe our dialogue agent using declarative syntax (Figure 2) and generate a contingent plan for the dialogue using a FOND planner similar to (De Venezia and Muise 2023). This plan can be executed by the Hovor dialogue executor (De Venezia and Muise 2023). The audio signals from the pilots are first split into utterances using silence thresholds and subsequently forwarded to the Automated Speech Recogniser (ASR) module powered by Distil-Whisper[1] via a FastAPI-endpoint[2]. For NLU, a BERT-based system using the Huggingface Transformers library[3] is implemented. The response from the dialogue executor is transformed into speech using the coqui-TTS package using `ljspeech/vits` model[4]. For the generation of strategic plans, once the context (i.e. problem instance) is set, we use the ARIES planner[5] interface exposed by the Unified Planning library[6]. The generated strategic plan is then visualized and communicated to the pilot through a GUI.

[1] https://github.com/huggingface/distil-whisper

[2] https://fastapi.tiangolo.com/

[3] https://huggingface.co/docs/transformers

[4] https://github.com/coqui-ai/TTS

[5] https://github.com/plaans/aries

[6] https://github.com/aiplan4eu/unified-planning

```
get-location_for_smoke_driven:
  type: custom
  subtype: slot_fill
  condition:
    emergency_type:
      value: smoke_known_location_required
  parameters:
    action_name:  get-location_for_smoke_driven
    overall_intent: share-location
    entities:
      - location
    message_variants:
      - Whats your best guess where the smoke is coming from?
    fallback_message_variants:
      - Can you please repeat it?
    config_entities:
      location:
        clarify_message_variants:
          - Did you mean $location?
    additional_updates:
      - outcome:
          location:
            known: true
        response_variants:
          - noted the location
```

Figure 2: Declarative syntax based on (De Venezia and Muise 2023) used to describe the dialogue agent

Acknowledgments

This work is funded by the German Federal Ministry of Economic Affairs and Climate Action (Project MOREALIS).

References

BFU. 2022. Studie zur Flugsicherheit von Luftsportgeräten - Analyse von Unfällen und Störungen mit Luftsportgeräten in Deutschland in den Jahren 2000-2019. Technical Report BFU22-803.1, Bundesstelle für Flugunfalluntersuchung.

Council of European Union. 2012. Council regulation (EU) no 748/2012.

De Venezia, R., and Muise, C. 2023. Plan4dial: A dialogue planning framework.

Höller, D.; Behnke, G.; Bercher, P.; and Biundo, S. 2018. Plan and goal recognition as HTN planning. In *ICTAI*.

Jamakatel, P.; Bercher, P.; Schulte, A.; and Kiam, J. J. 2023. Towards intelligent companion systems in general aviation using hierarchical plan and goal recognition. In *HAI 2023*, 229–237. ACM.

Jamakatel, P.; de Venezia, R.; Muise, C.; and Kiam, J. J. 2024. A goal-directed dialogue system for assistance in safety-critical application. In *IJCAI*.

Jamakatel, P.; Morchedi, S.; and Kiam, J. 2023. FRICO: An AI-enabled friendly cockpit assistance system. In *ICAPS 2023 System Demonstrations*.

Kiam, J. J., and Jamakatel, P. 2023. Can HTN planning make flying alone safer? In *Proceedings of the 6th ICAPS Workshop on Hierarchical Planning (HPlan 2023)*, 44–48.

Li, G.; Baker, S. P.; Grabowski, J. G.; and Rebok, G. W. 2001. Factors associated with pilot error in aviation crashes. *Aviation, space, and environmental medicine* 72(1):52–58.

Muise, C.; Chakraborti, T.; Agarwal, S.; Bajgar, O.; Chaudhary, A.; Lastras-Montano, L. A.; Ondrej, J.; Vodolan, M.; and Wiecha, C. 2019. Planning for goal-oriented dialogue systems.