MIRIAM: A Multimodal Interface for Explaining the Reasoning Behind Actions of Remote Autonomous Systems

Hastie, Helen; Garcia, Francisco Javier Chiyah; Robb, David; Laskov, Atanas ; Patron, Pedro

Published in:
ICMI '18 Proceedings of the 20th ACM International Conference on Multimodal Interaction

DOI:
10.1145/3242969.3266297

Publication date:
2018

Document Version
Peer reviewed version

MIRIAM: A Multimodal Interface for Explaining the Reasoning Behind Actions of Remote Autonomous Systems

Helen Hastie
Francisco J. Chiyah Garcia
David A. Robb
Heriot-Watt University
Edinburgh, UK
h.hastie,fjc3,d.a.robb@hw.ac.uk

Atanas Laskov
Pedro Patron
SeeByte Ltd
Edinburgh, UK
atanas.laskov,pedro.patron@seebyte.com

ABSTRACT

Autonomous systems in remote locations have a high degree of autonomy and there is a need to explain *what* they are doing and *why*, in order to increase transparency and maintain trust. This is particularly important in hazardous, high-risk scenarios. Here, we describe a multimodal interface, MIRIAM, that enables remote vehicle behaviour to be queried by the user, along with mission and vehicle status. These explanations, as part of the multimodal interface, help improve the operator's mental model of what the vehicle can and can't do, increase transparency and assist with operator training.

CCS CONCEPTS

• **Human-centered computing → Interaction design**; • **Computing methodologies → Natural language generation**;

KEYWORDS

Multimodal output, natural language generation, autonomous systems, trust, transparency, explainable AI

ACM Reference Format:

Helen Hastie, Francisco J. Chiyah Garcia, David A. Robb, Atanas Laskov, and Pedro Patron. 2018. MIRIAM: A Multimodal Interface for Explaining the Reasoning Behind Actions of Remote Autonomous Systems. In *2018 International Conference on Multimodal Interaction (ICMI '18), October 16–20, 2018, Boulder, CO, USA.* ACM, New York, NY, USA, 2 pages.

Figure 1: An IVER3 autonomous underwater vehicle (right, centre of frame) launched from a small boat starts its mission in Loch Linnhe, Scotland, while operators on shore monitor mission progress with the MIRIAM interface

1 INTRODUCTION

We present for demonstration MIRIAM (Multimodal Intelligent inteRactIon for Autonomous systeMs): a prototype multimodal interface that provides voice assistance for operators of remote unmanned autonomous systems [6]. Robots and autonomous systems are increasingly being operated remotely in hazardous environments such as in the nuclear or oil and gas domains [5]. Typically, remote robots instill less trust than those co-located [1, 4], which can be problematic with respect to making decisions in high-stakes scenarios. We propose that one way to manage trust and user confidence is through greater transparency. Thus, the system we present here is able to describe behaviour of remote autonomous underwater vehicles (AUVs) (Figure 1) in terms of *what* the system is doing [6], *why* it is doing certain behaviours and give explanations to describe *how* it works [2]. We propose that natural language explanations can help formulate clear and accurate mental models of autonomous systems and robots, which in turn can enable more informed decision-making.

We will demonstrate MIRIAM as shown in Figure 2, which consists of a commercial user interface called SeeTrack provided by industry partners SeeByte Ltd, combined with a chat interface, which gives status, mission updates, alerts and importantly explanations of behaviour. This multimodal interface has been trialled with real operators and has been shown to increase their situation awareness over the graphical interface alone [8].

With regards explanations, we will demonstrate MIRIAM's ability to produce two types of explanations, as illustrated in Figure 3. Namely, we produce *why* explanations to provide a trace or reasoning and *why not* to explain *how* it works, elaborating on the system's control method or autonomy strategy [3]. Lim et al. (2009) [7] show that both these types explanations increase understanding and, therefore, are important with regards the user's mental model.

Furthermore, for explanations, the data received from the vehicles is used to steadily build a knowledge base and generate accurate explanations in a given context. This knowledge base is used in

ICMI '18, October 16–20, 2018, Boulder, CO, USA

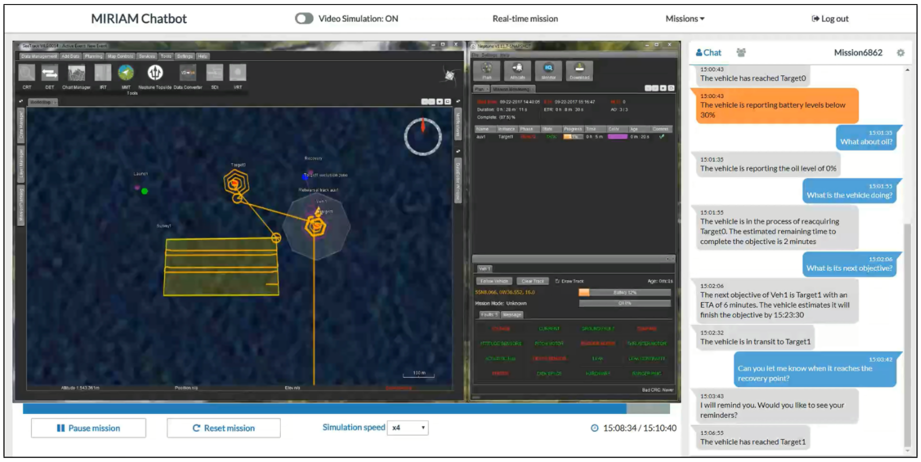

Figure 2: The multimodal interface with SeeTrack interface showing the simulated predicted path of the vehicle on the left and the chat interface on the right where explanations of vehicle behaviour appear

conjunction with an interpretable model of autonomy. For example, if a *why* request is made, the autonomy model is checked against the current mission status and history and the possible reasons are determined, along with a confidence value based on the information available at that point in the mission.

MIRIAM uses a rule-based NLP Engine that contextualises and parses the user's input for intent, formalising it as a semantic representation. It is able to process both static and dynamic data, such as names and mission-specific words. For example, it is able to reference dynamic objects such as"auv1", the particular name given to a vehicle in the mission plan, without the requirement to hard-code this name into the system. It can handle anaphoric references over multiple utterances e.g. "Where is Vehicle0?" ... "What is its estimated time to completion?". It also handles ellipsis e.g."What is the battery level of vehicle0?" ..."What about vehicle1?".

2 DEMONSTRATION

MIRIAM has been trialled and is demo-able with a real AUV vehicle (Figure 1). For conference demonstration purposes, however, we will demonstrate a mission in simulation. Conference delegates will be able to interact through speech or typed chat about a simulated autonomous underwater vehicle's (AUV) activity, as illustrated in Figure 2. The multimodal interface shows the vehicle track and progress on the left-hand-side through SeeByte's SeeTrack interface and includes a chat-interface on the right-hand-side. As well as explanations, other natural language interactions we will demonstrate include asking about the vehicle's current status, its plan and its current objectives, estimated time of arrival at specific locations, previous activities, mission progress, hardware fault diagnosis, and estimated time of completion of a specified objective. The system is mixed initiative in that it pro-actively sends warnings about important events, such as vehicle faults, critical battery status or a change in objectives. MIRIAM recognises different levels of importance and will pin critical alerts. The interface can be customised for various user-preferences and allows for the creation of reminders through natural language. A video showing examples of the interaction in the demo is at http://www2.macs.hw.ac.uk/miriam/interface

Chat A: Why	Chat B: Why not
User: *Why* is the vehicle spiralling up? **System:** This could be due to the following reasons: -The vehicle is performing a GPS fix (medium confidence) -The vehicle is transiting to its safe plan depth (5m) (low confidence) .... **User:** Can you confirm now *why* it is spiralling up? **System:** The vehicle is performing a GPS fix (high confidence)	**User:** *Why* is the vehicle *not* doing a GPS fix? **System:** The vehicle is not allowed to do GPS fixes during its transit to Point0 (high confidence)

Figure 3: Example dialogue explaining (A) giving the likely reasons why the vehicle might be exhibiting a behaviour and (B) explaining general behaviour

ACKNOWLEDGMENTS

We gratefully thank our funders: UK MOD Dstl ACC101939; RAEng/Leverhulme Trust (Hastie/LTSRF1617/13/37); EPSRC EP/R026173/1.

REFERENCES

[1] W. Bainbridge, J. Hart, E. S. Kim, and B. Scassellati. 2008. The Effect of Presence on Human-Robot Interaction. *Proc. IEEE Int. Symp. on Robot and Human Interactive Communication* (2008), 701–706.
[2] F. J. C. Garcia, D.A. Robb, X. Liu, A. Laskov, P. Patron, and H. Hastie. 2018. Explain Yourself: A Natural Language Interface for Scrutable Autonomous Robots. In *Proc. Explainable Robotic Systems Workshop, HRI'18*.
[3] S. Gregor and I. Benbasat. 1999. Explanations From Intelligent Systems: Theoretical Foundations and Implications for Practice. *MIS Quarterly vol 23 No. 4* (1999), 497.
[4] Helen Hastie, Xingkun Liu, and Pedro Patron. 2017. Trust triggers for multimodal command and control interfaces. In *Proc. ICMI'17*. 261–268.
[5] H. Hastie, K. Lohan, M.J. Chantler, D.A. Robb, S. Ramamoorthy, R. Petrick, S. Vijayakumar, and D. Lane. 2018. The ORCA Hub: Explainable Offshore Robotics through Intelligent Interfaces. In *Explainable Robotic Systems Workshop, HRI'18*.
[6] H.Hastie, F. J. C. Garcia, D. A. Robb, A. Laskov, and P. Patron. 2017. MIRIAM: A Multimodal Chat-based Interface for Autonomous Systems. In *Proc. ICMI'17*. 495.
[7] B. Y. Lim, A. K. Dey, and D. Avrahami. 2009. Why and why not explanations improve the intelligibility of context-aware intelligent systems. In *Proc. CHI'09*. 2119–2129. https://doi.org/10.1145/1518701.1519023
[8] David A. Robb, Francisco J. Chiyah Garcia, Atanas Laskov, Xingkun Liu, Pedro Patron, and Helen Hastie. 2018. Keep Me in the Loop: Increasing Operator Situation Awareness through a Conversational Multimodal Interface. In *Proc. ICMI'18*. https://doi.org/10.1145/3242969.3242974