The ORCA Hub: Explainable Offshore Robotics through Intelligent Interfaces

Helen Hastie, Katrin Lohan, Mike Chantler, David A. Robb, Ron Petrick, David Lane
Heriot-Watt University
Edinburgh, UK
initial.surname@hw.ac.uk

Subramanian Ramamoorthy, Sethu Vijayakumar
University of Edinburgh
Edinburgh, UK
s.ramamoorthy,sethu.vijayakumar@ed.ac.uk

ABSTRACT

We present the UK Robotics and Artificial Intelligence Hub for Offshore Robotics for Certification of Assets (ORCA Hub), a 3.5 year EPSRC funded, multi-site project with a vision to use teams of robots and autonomous intelligent systems (AIS) on remote energy platforms to enable cheaper, safer and more efficient working practices. The ORCA Hub will research, integrate, validate and deploy remote AIS solutions that can operate with existing and future offshore energy assets and sensors, interacting safely in autonomous or semi-autonomous modes in complex and cluttered environments, co-operating with remote operators. The goal is that through the use of such robotic systems offshore, the need for personnel will decrease. To enable this to happen, the remote operator will need a high level of situation awareness and key to this is the transparency of what the autonomous systems are doing and why. This increased transparency will facilitate a trusting relationship and increase adoption.

KEYWORDS

Autonomous systems, intelligent systems, transparency, trust, situation awareness, explaining robot actions, cognitive load

ACM Reference Format:
Helen Hastie, Katrin Lohan, Mike Chantler, David A. Robb, Ron Petrick, David Lane and Subramanian Ramamoorthy, Sethu Vijayakumar. 2018. The ORCA Hub: Explainable Offshore Robotics through Intelligent Interfaces. In *Proceedings of Explainable Robotic Systems Workshop, ACM Human-Robot Interaction conference.* ACM, New York, NY, USA, 2 pages.

1 INTRODUCTION

The international offshore energy industry currently faces three major challenges of an oil price expected to remain less than $50 a barrel, significant expensive decommissioning commitments of old infrastructure (especially North Sea) and small margins on the traded commodity price per KWh of offshore renewable energy. Further, the workforce is aging as new generations of suitable graduates prefer not to work in hazardous remote environments. Operators, therefore, seek more cost effective, safe methods for inspection, repair and maintenance of their topside and marine infrastructure. Part of this solution is deploying robots and autonomous systems in the air, on the rig and on the water surface and subsea. This will mean fewer staff offshore, reduced costs and increased safety. The long-term industry vision is thus for a completely autonomous offshore energy field, operated, inspected and maintained from the shore.

Explainable Robotic Systems Workshop, ACM Human-Robot Interaction conference, March 2018, Chicago, IL USA
2018.

The hub will investigate key areas in robotic autonomy, mobility, manipulation, sensor processing, autonomous mapping, navigation, multimodal interfaces and human-machine collaboration. It is the latter two that we focus on here. Key to adoption of robotics and AIS is the ability for users to feel confident and trust them enough to deploy them in remote, high stakes, hazardous conditions [6]. Our solution will be to provide effective communication of their world view, which includes explanations of actions, plans and plan failures. Core to our approach is the interpretation and explanation of autonomy models, which include both black-box (e.g. deep neural network) and grey-box models (e.g. Bayesian networks). Such communication will increase transparency, which in turn will aid decision-making and facilitate collaboration including (re)planning and multitasking. Specifically, we will develop interaction techniques to increase transparency in the following ways:

(1) "Why did you do that?": Explain the robot's behaviour models with various scrutability levels, from black-box to rule-based systems [3],
(2) "What are you doing": Explain activity, reporting what the system is actually doing [5],
(3) "What do you see/sense?": Explain the environment and be able to refer to objects in a human-understandable way [9],
(4) "What if you do this instead?": Explain possible outcomes of alternative plans, and collaborative planning [4],
(5) "Do this now": Give instructions/shared autonomy [1, 8].

2 OUR APPROACH

In order to facilitate representation of explanations in context, we will create an adaptive, situated multimodal interface that is in tune with the user's cognitive load. We will adopt a user-centred design approach, capturing from users in-situ operational goals and context, user mental models, terminology, preferred modality (e.g. speech, gesture, augmented reality) and visual conventions. This will inform iterative development of both the user interfaces and associated representations using design probes and both quantitative/qualitative methods.

We propose that for AIS to be truly trusted enough to be integrated into a human-machine team, they need to *interrogatable*, as humans are, providing on-demand explanations of events and behaviours. This could be done through natural language, building on prior work on the MIRIAM System (see Figure 1) [3, 5]; or through visualisation [7] or both. Interruptions may be in the form of calls for help or information.

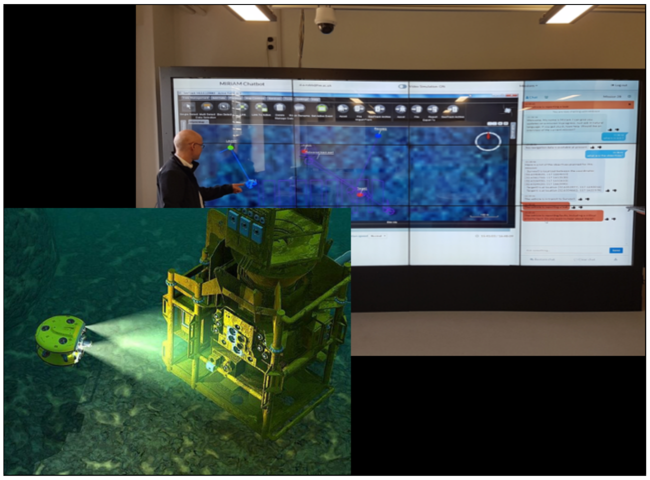

Figure 1: Foreground: an example underwater vehicle, with which the ORCA hub would interface; background: MIRIAM interface [5]

Heavy cognitive load can have negative effects on task completion and decision-making [2]. As the complexity of the offshore scenario increases, we will need a mechanism to manage user cognitive load in order to deliver alerts and explanations with just the right amount of information and in the best modality for that particular user at that particular point in time. Cognitive load refers to the total amount of mental effort being used in the working memory [11] and can vary from person to person. It can be measured subjectively (e.g. through NASA-TLX) or through sensors, such as EEG and pupillary response.

The ORCA interface will provide explanations of the robot perception and robot planning and action including the causal structure of the plan that is being carried out, so that users can interact with the autonomous system in a meaningful way. This approach will be applied to both black-box (e.g. CNN/RNN) and grey box (e.g. Bayesian network) reasoning, and the resulting visual and textual explanations will be presented in a structure and form that exploits hierarchical cognitive chunking and human perceptual (particularly unconscious visual) processing, in order to keep operator cognitive load low.

For black-box models, we will induce causal models directly from robot behaviours and controllers in order to produce queriable causal representations [10]. We will exploit parallel work, which is part of the DARPA XAI COGLE project, to induce programmatic models from observational data of robots and the environment. We will also build tools that directly use the causal planning structures and replanning outputs as the basis for explanation generation, treating high-level planning as an information source for the underlying causal connections between actions, plans, and goals. These representations and mappings will facilitate human understanding of robot perceptions in terms of both an egocentric and birds-eye view of the environment and corresponding reactionary plans.

3 DISCUSSION

Through the ORCA Hub, we propose that a multi-disciplinary approach is needed to increase transparency and trust of complex, multi-system operations and scenarios. This includes Explainable AI, to explain decisions of robots and their environments. We must personalise the explanations for specific users in specific contexts. Understanding the user's mental model of the systems is key to providing just the right information at the right time. If the user does not have a clear mental model then the extra effort of processing and digesting explanations will be worth it for the added benefit. Assessing and evaluating the quality of the explanations will be key. These metrics include intrinsic qualitative and quantitative measures but also extrinsic measures, such as whether the explanations were useful to the user in enabling them to do their job better/faster. Deciding what makes a good explanation will be non-trivial but evaluation techniques can be taken from the field of Natural Language Generation, including metrics such as relevance and Flesch readability score. How certainty of information is portrayed will also be an important question and may depend on the user, with some users preferring to receive all information and hypotheses and some just wanting to hear what the interface can tell them for certain. Balancing information presentation and explanations with the user's actual needs and cognitive load in a dynamic, fast moving environment, will be essential to the successful deployment of robotics and AIS in offshore robotics and elsewhere.

ACKNOWLEDGMENTS

This work was funded by EPSRC (EP/R026173/1, 2017-2021); RAEng/ Leverhulme Trust Senior Research Fellowship Scheme (Hastie /LT-SRF1617/13/37); DARPA COGLE (Ramamoorthy). We acknowledge all the ORCA Hub consortium partners.

REFERENCES

[1] Matthew Crosby, Francesco Rovida, Volker Krueger, and Ronald P. A. Petrick. 2017. Integrating Mission and Task Planning in an Industrial Robotics Framework. In *Proc. of ICAPS'17*. 471–479.

[2] Dennis D Fehrenbacher and Soussan Djamasbi. 2017. Information systems and task demand: An exploratory pupillometry study of computerized decision making. *Decision Support Systems* 97 (2017), 1–11.

[3] Francisco J. Chiyah Garcia, David A. Robb, Xinkun Liu, Atanas Laskov, Pedro Patron, and Helen Hastie. 2018. Explain Yourself: A Natural Language Interface for Scrutable Autonomous Robots. In *Proc. of Explainable Robotic Systems Workshop, HRI'18*.

[4] Christopher Geib, Bart Craenen, and Ronald P. A. Petrick. 2016. Generating Collaborative Behaviour through Plan Recognition and Planning. In *Proc. of the ICAPS 2016 Workshop on Distributed and Multi-Agent Planning (DMAP)*. 98–105.

[5] Helen Hastie, Francisco Javier Chiyah Garcia, David A. Robb, Pedro Patron, and Atanas Laskov. 2017. MIRIAM: A Multimodal Chat-based Interface for Autonomous Systems. In *Proc. of ICMI'17*.

[6] Helen Hastie, Xingkun Liu, and Pedro Patron. 2017. Trust Triggers for Multimodal Command and Control Interfaces. In *Proc. ICMI'17*.

[7] Pierre Le Bras, David A Robb, Thomas S Methven, Stefano Padilla, and Mike J Chantler. 2018. Improving User Confidence in Concept Maps: Exploring Data Driven Explanations. In *Proc. of CHI*. ACM.

[8] Stéphane Mercier and Catherine Tessier. 2008. Some basic concepts for shared autonomy: A first report. *Frontiers in Artificial Intelligence and Applications* 176, 1 (2008), 40–48.

[9] Svetlin Penkov, Alejandro Bordallo, and Subramanian Ramamoorthy. 2017. Physical Symbol Grounding and Instance Learning through Demonstration and Eye Tracking. In *Robotics and Automation, 2017 IEEE International Conference on*.

[10] Svetlin Penkov and Subramanian Ramamoorthy. 2017. Using program induction to interpret transition system dynamics. In *Proceedings of ICML Workshop on Human Interpretability in Machine Learning*.

[11] John Sweller. 1988. Cognitive load during problem solving: Effects on learning. *Cognitive science* 12, 2 (1988), 257–285.