On Specifying for Trustworthiness

Name of Author 1
Institute
City, Country
Email

Name of Author 2
Institute
City, Country
Email

Name of Author 3
Institute
City, Country
Email

Name of Author 4
Institute
City, Country
Email

Name of Author 5
Institute
City, Country
Email

Name of Author 6
Institute
City, Country
Email

Name of Author 7
Institute
City, Country
Email

Name of Author 8
Institute
City, Country
Email

Name of Author 9
Institute
City, Country
Email

Name of Author 10
Institute
City, Country
Email

Name of Author 11
Institute
City, Country
Email

Name of Author 12
Institute
City, Country
Email

ABSTRACT

As autonomous systems are becoming part of our daily lives, specifying for trustworthiness of these systems is crucial. ... In this article, we take a broad view of specification, concentrating on top-level requirements including but not limited to functionality, safety, security and other non-functional properties that contribute to trustworthiness. The main contribution of this article is a set of high-level intellectual challenges related to specifying a trustworthy autonomous system without focussing on how these challenges are actually realized. We also identify their potential uses in a variety of autonomous systems domains. ...

CCS CONCEPTS

• **Computing methodologies** → **Artificial intelligence**; • **Software and its engineering** → **Requirements analysis**; **Software organization and properties**; **Software functional properties**; **Extra-functional properties**.

KEYWORDS

autonomous systems, trust, specification

ACM Reference Format:
Name of Author 1, Name of Author 2, Name of Author 3, Name of Author 4, Name of Author 5, Name of Author 6, Name of Author 7, Name of Author 8, Name of Author 9, Name of Author 10, Name of Author 11, and Name of Author 12. 2021. On Specifying for Trustworthiness. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

2 AUTONOMOUS SYSTEMS DOMAINS AND THEIR UNIQUE CHALLENGES

Each autonomous systems domain brings about unique specification challenges:

2.1 Swarm Robotics

Swarm robotics provides an approach to the coordination of large numbers of robots, which is inspired from the observation of social insects [13]. Robustness, flexibility and scalability are three desirable properties in any swarm robotics system. In [13], a set of criteria is proposed for distinguishing swarm robotic research from multi-robot systems. The individual robots that make up the swarm need to be autonomous. Also, there need to be a large number of robots; a few homogenous groups of robots; relatively incapable or inefficient robots; and robots with local sensing and communication capabilities.

The functionality of a swarm is emergent (e.g. aggregation, coherent ad hoc network, taxis, obstacle avoidance and object encapsulation [15]), and evolves based on the capabilities of the robots and the numbers of robots used. The overall desired behaviours of a swarm are not explicitly coded or engineered in the system, but they are an emergent consequence of the interaction of individual agents with each other and the environment. This evolving functionality across many individual robots poses a challenge in how best to specify a swarm that is trustworthy by design, useful, and acceptable to users?

2.2 Unmanned Aerial Vehicles

A unmanned aerial vehicle (UAV) or drone is a type of aerial vehicle that is capable of autonomous flight without a pilot on board. As UAVs are increasingly being applied in diverse areas of applications [14], such as logistics services, agriculture, emergency response, and security, ensuring their trustworthiness is of utmost importance. Specification of operational environments of UAVs is challenging mainly for two reasons. First, the unclear and uncertain government regulations make it challenging, as these rules may change over time. Second, the complexity and uncertainty of the operational environment of the UAV itself is a challenge. For example, in a parcel delivery service using UAVs in an urban environment, we could mention uncertain flight conditions (e.g. wind gradients), and highly dynamic and uncertain airspace (e.g. other UAVs in operation).

The recent advances in machine learning offer the potential to increase the autonomy of UAVs by allowing them to learn from experience. For example, machine learning can be used to stabilize the flight which can greatly improve performance in gusty urban wind conditions. When one considers the conventional flight controller of a UAV, it can include several measures, such as risetime, overshoot, settling time, and steady-state error. The goal of these specification measures is to ensure that the control system is stable and robust. There is disturbance attenuation against environmental uncertainty; smooth and rapid responses to set-point changes; and steady-state accuracy. In this context, a key challenge is how do we specify a UAV should deal with situations that go beyond the limits of its training?

3 INTELLECTUAL CHALLENGES FOR RESEARCH COMMUNITY

3.1 TAS Functionality

On standards for autonomous systems with evolving functionality

Autonomous systems with *evolving functionality* (i.e. the ability to change in function over time) pose significant challenges to current processes for specifying functionality. Most conventional processes for defining system requirements and characteristics assume that these are fixed and can be defined in a complete and precise manner before the system goes into operation. In this regard, a key limitation is the fact there are no industry standards for specifying evolving functionality, which we discuss using two application areas – swarm robotics and UAVs.

In the field of robotics, several safety standards have been developed by ISO/TC 299 for the non-industrial (service) robotics sector (e.g. ISO 13482 [5], ISO 23482-1/2 [7, 8]), as well as for the industrial robotics sector (e.g. ISO 10218-1/2 [3, 4], ISO/TS 15066 [6]). Different legal and regulatory requirements apply to different robot categories. In service robotics, ISO 13482 covers the hazards presented by the robots and devices for applications in non-industrial environments for providing services. ISO 23482-1/2 standards extends ISO 13482 with guidance and methods that can be used to test personal care robots. On the other hand, in the industrial sector, ISO 10218-1/2 standards provide safety requirements for industrial robots and their integration. Meanwhile, ISO/TS 15066 provides safety requirements for collaborative industrial robot systems and work environment. Although these industry standards focus on ensuring safety of robots at the individual robot level, there are no standards to ensure safety or any other extra-functional property for *swarms*.

Meanwhile, for the airborne systems and in particular for UAVs, several industry standards and regulations have been introduced to ensure their safe operation. DO-178C [12] is the primary standard for commercial avionics software development. It provides software considerations for the production of airborne systems and equipment. On the other hand, DO-254 [11] provides guidance for the development of airborne electronic hardware. ED279 [2] standard provides a framework to support designers when performing a functional hazard assessment process for an unmanned aircraft system. ARP4761 [9] provides guidelines and methods for conducting the safety assessment process on civil airborne systems and equipment. NATO STANAG 4671 [10] is intended to allow military UAVs to operate in other NATO members airspace. As for regulations within the UK, CAP 722 [1] is the primary guidance document for the operation of unmanned aircraft systems. However, none of these standards or regulations provide safety considerations for machine learning components, which is a key limitation.

When one considers the existing industry standards for autonomous systems, they are either implicitly or explicitly based on the V life-cycle model, which moves from requirements through design onto implementation and testing. For systems with the ability to adapt their functionality in response to changes in their environment, or their own internal state, this approach is unlikely to

be suitable. Therefore, there is a need for new standards that better reflect the life-cycle of an autonomous system with evolving functionality.

4 CONCLUSION

ACKNOWLEDGMENTS

This article is a result of the fruitful discussions at the Specifying for Trustworthiness workshop held in conjunction with the Trustworthy Autonomous Systems (TAS) All Hands Meeting. The authors thank all the speakers and fellow participants, and the TAS Hub and EPSRC for their support.

REFERENCES

[1] Civil Aviation Authority. 2020. CAP 722: Unmanned Aircraft System Operations in UK Airspace - Guidance. Online. Retrieved November 1, 2021 from https://publicapps.caa.co.uk/modalapplication.aspx?appid=11&mode=detail&id=415

[2] EUROCAE. 2020. ED-279 - Generic Functional Hazard Assessment (FHA) for UAS/RPAS. Online. Retrieved November 1, 2021 from https://eshop.eurocae.net/eurocae-documents-and-reports/ed-279/

[3] International Organization for Standardization. 2011. ISO 10218-1:2011 Robots and robotic devices — Safety requirements for industrial robots — Part 1: Robots. Online. Retrieved November 1, 2021 from https://www.iso.org/standard/51330.html

[4] International Organization for Standardization. 2011. ISO 10218-2:2011 Robots and robotic devices — Safety requirements for industrial robots — Part 2: Robot systems and integration. Online. Retrieved November 1, 2021 from https://www.iso.org/standard/41571.html

[5] International Organization for Standardization. 2011. ISO 13482:2014 Robots and robotic devices — Safety requirements for personal care robots. Online. Retrieved November 1, 2021 from https://www.iso.org/standard/53820.html

[6] International Organization for Standardization. 2016. ISO/TS 15066:2016 Robots and robotic devices — Collaborative robots. Online. Retrieved November 1, 2021 from https://www.iso.org/standard/62996.html

[7] International Organization for Standardization. 2019. ISO/TR 23482-2:2019 Robotics — Application of ISO 13482 — Part 2: Application guidelines. Online. Retrieved November 1, 2021 from https://www.iso.org/standard/71627.html

[8] International Organization for Standardization. 2020. ISO/TR 23482-1:2020 Robotics — Application of ISO 13482 — Part 1: Safety-related test methods. Online. Retrieved November 1, 2021 from https://www.iso.org/standard/71564.html

[9] SAE International. 1996. Guidelines and methods for conducting the safety assessment process on civil airborne systems and equipment ARP4761. Online. Retrieved November 1, 2021 from https://www.sae.org/standards/content/arp4761/

[10] NATO Standardization Office. 2019. Unmanned Aircraft Systems Airworthiness Requirements (USAR) STANAG 4671. Online. Retrieved November 1, 2021 from https://nso.nato.int/nso/nsdd/main/standards

[11] Inc. RTCA. 2000. DO-254 - Design Assurance Guidance for Airborne Electronic Hardware.

[12] Inc. RTCA. 2011. RTCA/DO-178C Software Considerations in Airborne Systems and Equipment Certification.

[13] Erol Şahin. 2005. Swarm Robotics: From Sources of Inspiration to Domains of Application. In *Swarm Robotics*, Erol Şahin and William M. Spears (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 10–20.

[14] Uchechi Ukaegbu, Lagouge Tartibu, and Modestus Okwu. 2021. Unmanned Aerial Vehicles for the Future: Classification, Challenges, and Opportunities. In *2021 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*. 1–7. https://doi.org/10.1109/icABCD51485.2021.9519367

[15] Alan F. T. Winfield and Julien Nembrini. 2006. Safety in numbers: fault-tolerance in robot swarms. *International Journal on Modelling Identification and Control* 1, 1 (2006), 30–37. https://doi.org/10.1504/IJMIC.2006.008645