Verification of Autonomous Systems

By Dejanira Araiza-Illan, Michael Fisher, Kevin Leahy, Joanna Isabelle Olszewska, and Signe Redfield

The robotics and autonomous systems communities have seen a significant and rapid increase in both the development of robots and vehicles for commercial use and in using these systems across a wide range of novel applications. As these robots, vehicles, software, and even embedded devices move toward much greater autonomy, techniques for verification, providing much higher confidence than usual, are becoming required. However, the analysis and evaluation processes used for traditional systems must be significantly enhanced to provide increased confidence in this next wave of autonomous systems. The need for well understood and effective verification techniques will become even vital, as we move to commercial applications that rely on complex artificial intelligence technologies, and the utilization of these systems in safety-critical scenarios.

There are a growing number of research developments concerning the verification of complex systems that can all impact upon this problem. These are clearly of relevance for designing, constructing, and deploying autonomous systems, but also have importance to psychology (e.g., social robotics), philosophy (e.g., machine ethics), and law (e.g., certification). Furthermore, constructing autonomous systems without strong behavioral guarantees can lead to serious outcomes and may consequently hold back the widespread adoption of these systems. As the research is currently fragmented and often not well publicized, this technical committee (TC) aims to coalesce this activity, drive the research agenda forward, and instill the necessity for verification firmly within industry, government, standards, and the public.

Digital Object Identifier 10.1109/MRA.2022.3143966
Date of current version: 22 March 2022

Background

Following a number of workshops and collaborations, the Verification of Autonomous Systems Technical Committee (TC-VAS) was established within the IEEE Robotics and Automation Society in 2019. It has now developed into a large community, with a mailing list of over 490 unique individuals (493 unique people and 529 emails, as checked on 23 October 2021), monthly webinars, workshops, and input to standards activities.

What Is "Verification"?

The term *verification* encompasses a range of techniques meant to assess, with varying degrees of strength, whether a particular system matches its required behavior. Essentially, it consists in providing evidence that some system, **S**, matches its requirements, **R**. Since the borders between verification and validation can be blurred, this TC takes an inclusive approach to such classifications. The problem of defining requirements that correctly and completely describe the actual desired behavior is likewise included.

Verification techniques are often categorized into those that have a basis in mathematical logic/proof and those that rely on more empirical approaches. Within each of these classes, there remain many options. For example, within the former (termed *formal verification*), one might employ [8]:

- *Proof*: where a formal proof is carried out to establish that **R**, encoded as a logical formula, follows from the logical description of the behavior of **S**
- *Model checking*: where **R** is exhaustively assessed against a representation of all possible execution paths of **S**
- *Runtime verification*: where **R** is assessed against the system **S** as it is executing [10].

The term *verification* encompasses a range of techniques meant to assess, with varying degrees of strength, whether a particular system matches its required behavior.

All of these options also have probabilistic versions, many of which have become popular in recent years [2], [7]. There is also a range of *empirical* verification approaches, such as:

- *Software testing*: where **R** is checked on a subset of the possible executions of **S**
- *Simulation-based testing*: as above but where a *simulation* of the real environment is used for environmental modeling and interactions [1]

- *In-situ testing*: where R is assessed against the actual working of the system in a real-world environment [3].

Again, there are many variations and options here ranging from the "certain" (in the case of formal proof) to the stochastic. The TC aims to encompass work across all these areas as well as to link and support aspects such as transparency [11] and modularity [6], [9].

What Is "Autonomy"?

Essentially, autonomy is the ability (and often requirement) of a system to make its own decisions and take its own actions. This TC takes an inclusive view on system autonomy, covering full autonomy, where decision making and action is fully within the system's software (and so assessment of why decisions are made becomes crucial); adaptive systems, where decision making and action are driven by (often continuous) interactions with the environment; and automated systems, where decision making and action are prescripted, and so on. Which form of decision making is utilized will also have a strong impact on the effectiveness of any of the verification techniques to which it can be applied.

This TC is concerned with the development of tools and techniques to verify autonomous systems even in such unconstrained and unstructured environments.

Why Is This Important?

As the range of systems that are expected to act on their own expands, the need for verification of these autonomous systems becomes more important. When there is a "human in the loop," i.e., a human providing oversight and control of a system, the key decisions about the system can be delegated to that human, leaving the system analysis to address issues such as reliability. However, once there is a need for the system to make key decisions, much more evidence and confidence in these types of systems will be required. Developing the ability to establish and provide this evidence is then essential not only for engineers but also for all stakeholders such as regulators, the public, and governments, and so for this TC.

If the abilities of systems and the environments in which they are to work remain constrained, then realistic boundaries for system behaviors can be provided. However, once autonomous systems are deployed in hard to predict or unknown environments and we expect them to make key, and sometimes (safety, mission, security) critical decisions, then a much stronger analysis is required. In addition, what requirements might be assessed depend crucially on what is known of the system and its environment. Traditionally, it has been assumed that we can assess (before deployment) all potential issues/concerns and mitigate against these, which might be the case in highly controlled, closed environments. However, with autonomous systems increasingly being used in open, uncontrolled environments and with internal, software behavior able to change in various ways, our ability to predict "everything that might go wrong" is severely limited. Furthermore, stochastic models of complex, unknown environments can never be complete and may have hard to predict errors. Therefore, this TC is concerned with the development of tools and techniques to verify autonomous systems even in such unconstrained and unstructured environments.

TC Organization

Leadership of the TC is provided by five people.

TC Cochairs

- Dejanira Araiza-Illan
- Michael Fisher
- Signe Redfield

TC Junior Cochairs

- Kevin Leahy
- Joanna Isabelle Olszewska

TC Activities

The TC engages in a range of activities:

- monthly hour-long webinars, each comprising of two 20 minute talks, with questions (and answers) and an approximate average number of attendees of 55
- annual meetings to discuss existing activities, collate feedback, and define new activities
- sponsorship of workshops, such as at the International Conference on Intelligent Robots and Systems (IROS) or the International Conference on Robotics and Automation (ICRA)
- input into IEEE standards, such as IEEE P2817 [5] and IEEE P7009 [4].

As an example, the P7009 Standards Working Group, which studies the "Fail-Safe Design of Autonomous Systems," asked this TC for input concerning autonomous systems verification. Thus, a subgroup among TC-VAS was created, met regularly, and subsequently provided the P7009 team with a summary document to be included in their standard.

TC Future Plans

The TC also has plans for a range of future activities:

- a road map highlighting research challenges and developments required over the medium to long term
- a range of educational resources on the "verification of autonomous systems" topic
- a catalog and repository of tools available for some aspect of the verification of autonomous systems
- continued sponsorship of workshops at the Conference on Automation Science and Engineering, IROS, or IRCA, but also a stand-alone event dedicated to the verification of autonomous systems.

Closing Remarks

Individuals can become involved with this TC by visiting our home page at https://www.ieee-ras.org/verification-of-autonomous-systems and selecting the "Join Us" menu option. This will bring up a form that adds new members to our mailing list.

Acknowledgments

Contributions to this article were made by Dejanira Araiza-Illan in her personal capacity. The opinions expressed in this article are the author's own and do not reflect the views of Johnson & Johnson. Distribution Statement A: Approved for public release. Distribution is unlimited. This material is based upon work supported by the United States Air Force under Air Force Contract No. FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Air Force. Signe Redfield is currently a Naval Research Laboratory employee. No laboratory funding or resources were used to produce the result/findings reported in this publication.

References

[1] H. Alghodhaifi and S. Lakshmanan, "Autonomous vehicle evaluation: A comprehensive survey on modeling and simulation approaches," *IEEE Access,* vol. 9, pp. 151,531–151,566, Nov. 2021, doi: 10.1109/ACCESS.2021.3125620.

[2] C. Baier and J.-P. Katoen, *Principles of Model Checking* (Representation and Mind Series). Cambridge, MA, USA: MIT Press, 2008.

[3] A. Bertolino *et al*., "A survey of field-based testing techniques," *ACM Comput. Surv.,* vol. 54, no. 5, pp. 1–39, May 2021, doi: 10.1145/3447240.

[4] M. Farrell et al., "Evolution of the IEEE P7009 standard: Towards fail-safe design of autonomous systems," in *Proc. 2021 32nd Int. Symp. Softw. Rel. Eng.* [Industry Track].

[5] *Guide for Verification of Autonomous Systems,* IEEE P2817. Accessed: Feb. 8, 2022. [Online]. Available: https://www.ieee-ras.org/industry-government/standards/active-projects/verification-of-autonomous-systems

[6] *Standard on Modular Robotics,* ISO 22166. Accessed: Feb. 8, 2022. [Online]. Available: https://committee.iso.org/home/tc299

[7] A. Lavaei, S. Soudjani, A. Abate, and M. Zamani, "Automated verification and synthesis of stochastic hybrid systems: A survey," 2021, arXiv:2101.07491.

[8] M. Luckcuck, M. Farrell, L. A. Dennis, C. Dixon, and M. Fisher, "Formal specification and verification of autonomous robotic systems: A survey," *ACM Comput. Surv.,* vol. 52, no. 5, pp. 1–41, 2020, doi: 10.1145/3342355.

[9] M. Quigley *et al*., "ROS: An open-source robot operating system," in *Proc. ICRA Workshop Open Source Softw*., 2009.

[10] C. Sánchez *et al*., "A survey of challenges for runtime verification from advanced application domains (beyond software)," *Formal Methods Syst. Des.,* vol. 54, no. 3, pp. 279–335, 2019, doi: 10.1007/s10703-019-00337-w.

[11] A. F. T. Winfield *et al*., "IEEE P7001: A proposed standard on transparency," *Frontiers Robot. AI,* vol. 8, p. 665,729, Jul. 2021, doi: 10.3389/frobt.2021.665729.

Dejanira Araiza-Illan, Johnson & Johnson, Singapore, 118222, Singapore. Email: dejanira.araiza.i@gmail.com.

Michael Fisher, University of Manchester, Manchester, M13 9PL, U.K. Email: michael.fisher@manchester.ac.uk.

Kevin Leahy, Massachusetts Institute of Technology Lincoln Laboratory, Lexington, Massachusetts, 02421, USA. Email: kevin.leahy@ll.mit.edu.

Joanna Isabelle Olszewska, University of West of Scotland, Glasgow, G72 0LH, U.K. Email: joanna.olszewska@ieee.org.

Signe Redfield, Naval Research Laboratory, Washington, D.C., 20375, USA. Email: signe.redfield@nrl.navy.mil.

Deadline for RAS Local Chapter Initiative Grants

The RAS Member Activities Board (MAB) awards a limited number of Chapter Initiative Grants to local RAS Chapters for professional development, educational outreach, and other programs. Grant proposals will be reviewed by the MAB at its meeting in late May 2022 and funds up to US$2,000 will be awarded on a competitive basis. The deadline for proposals is 15 April 2022. For submission details, please visit: https://www.ieee-ras.org/chapters/support-for-chapters.

Digital Object Identifier 10.1109/MRA.2022.3143407