Workshop Report: Scoping Certifiable Autonomous Nuclear Robotics

Tech Report: ULCS-20-001

12th of April 2019

Dr Matt Luckcuck,
Department of Computer Science,
University of Liverpool, UK

Contents

1 Introduction

The Scoping Certifiable Autonomous Nuclear Robotics workshop was held on the 12th of April 2019 in Manchester, UK. The aim was to explore the challenges of obtaining permission to use autonomous robotics on the UK nuclear estate. The workshop had attendees from nuclear operators supply chain, the robotics industry, academia; and, crucially, UK's nuclear regulator, the Office for Nuclear Regulation (ONR).

The workshop centred on four concrete case studies concerning nuclear robotics, provided by nuclear operators. One of these case studies was entirely remote operated, while the other three had some semi-autonomous functions. The attendees divided into four discussion groups, each tackling one of the case studies. The discussion groups explored the changes required to design, verification, and regulation approaches with the introduction of (higher-levels of) autonomy and autonomous systems.

The workshop enabled the sharing of concrete case studies of robotic systems being used, or proposed for use, in the nuclear sector. This proved very helpful to the attendees from the nuclear supply chain and academia. The discussion sessions facilitated sharing of the challenges in engineering safe (and increasingly autonomous) robotic systems for the nuclear industry. The variety of different backgrounds among the attendees made these discussions both rich and interesting. A secondary aim of the workshop was to help foster stronger relationships between members of the supply chain, nuclear sites, the ONR, and academia.

The group discussions were intended to provide insights into the verification and certification processes for nuclear robotics, and examine how these might need to change for increasingly autonomous systems. We were particularly interested in barriers to the introduction of autonomy that were common across nuclear operators, verification techniques for autonomous systems, observations about the current use of safety documentation in the nuclear sector, and any areas where collaboration could help reduce barriers or tackle challenges.

Organisation. The workshop was organised by Matt Luckcuck and Michael Fisher, of the Autonomy and Verification Lab at the University of Liverpool[1]. It was funded through the *Robotics and AI for Nuclear (RAIN)* Hub[2] and was strongly supported by the *Office for Nuclear Regulation (ONR)*[3], who contributed 6 participants.

1.1 What is an Autonomous System?

It is important to describe what we mean by an autonomous system in the context of this report. An autonomous system uses artificial intelligence techniques to decide what to do, based on input from its environment, without human intervention. An autonomous system may be *embodied* in a robotic system, so that it can effect its environment.

[1] `https://autonomy-and-verification-uol.github.io`
[2] `https://rainhub.org.uk`
[3] `http://www.onr.org.uk`

It is important to note that autonomy is a spectrum of capabilities and supervision requirements. There are various definitions of levels of autonomy for robotic systems [2], some of which were designed for a particular industrial sector but have been reused in other areas (such as the defintions for on-road cars [5]. These definitions usually describe how a system can range from having no autonomy, to sharing autonomy with a human, up to having full autonomy.

While it makes decisions for itself, an autonomous system can only act within the confines of the actions it has been designed to be able to perform. For example, an autonomous system designed to vacuum floors may decide that once it has finished with this room, it will go and recharge before it finishes the rest of the house. However, it cannot decide to give up cleaning floors and explore the outside world. Similarly, an autonomous system controlling a robot arm designed to safely deconstruct objects might decide which tool to use, but it cannot decide to deconstruct itself.

If we assume that the autonomous system's sensors are providing it with correct information about its environment, the key aspects to look at are the decisions that the autonomous system makes. It can be difficult to ensure that the system's repertoire of actions will always be safe, since actions can have results that the system might not be able to predict. This means that we must ensure the system only chooses each action when (it believes) it is safe to do so. If we imagine a human learning to drive a car, almost every action they can choose *could* be unsafe. The key part of driving is to only choose to perform an action when it is safe to do so.

Autonomy can be implemented using a variety of different artificial intelligence techniques. Some require more up-front work to define the system's available choices, and so are more predictable as a consequence. Others can learn what choices to make from pre-existing data, so can be more flexible and efficient, but also much less predictable. The ability to verify that the system will always choose a safe or secure or ethical action will depend centrally on how the autonomy is implemented. Arguably, for safety-critical systems, the most predictable and analysable techniques should be used for any autonomous components, so that robust verification methods can be used to check the system

2 Case Studies

The workshop centred on four case studies from nuclear operators. Each was presented by a representative from the respective organisation: **Sally Forbes**, from the UK Atomic Energy Authority (UKAEA); **Howard Chapman**, from the National Nuclear Laboratory (NNL); **Andy Melia**, from Sellafield Ltd; and **Andrew Wallwork**, from the Atomic Weapons Establishment (AWE).

2.1 Remote Handling

The UKAEA case study was the MASCOT system, which is a tele-operation system used for remote handling materials in fusion reactors. MASCOT is a pair of robotic arms, where the *slave* arms mirror the actions of the *master* arms as they are manually operated. The master arms receive haptic feedback from the slave end of the system, which provides the operator with some sense of the forces on the grippers.

MASCOT was built to to allow operations inside the JET fusion reactor at the Culham Centre for Fusion Energy (CCFE), UK (while the reactor is off). It enables operators to maintain and clean the inside of the reactor while the residual radiation is still too high for human workers. The UKAEA have adapted various off-the-shelf tools to fit the MASCOT grippers. MASCOT has some semi-autonomous functionality, such as being able to move back into a 'home' position.

2.2 Laser Cutting

Two of the case studies concerned robotic laser cutting systems — systems under development at NNL and Sellafield. The NNL system uses a robot arm with a laser cutting head to perform cutting experiments on radioactive material. The aim of the NNL system is to demonstrate how a robotic system can be more efficient and produce less waste than conventional cutting approaches, while removing human operators from the active material being cut.

The NNL laser cutting system can operate either manually or semi-autonomously. In manual mode, the operator plans out the cutting path; in semi-autonomous mode the cut's start and end points are selected by the operator, but the cutting path is planned by the system. The operator can also check a virtual run of the planned cut before execution. The robot arm can also switch between its laser cutting head and an environment scanner. Hardware interlocks prevent human access to the cutting area either while the laser is active or the robot is moving. To prevent it cutting the containing wall of the room, the laser cutter can only be activated when it is pointing at the cutting area's (reinforced) backplate.

The Sellafield system also has a robot arm with a laser cutting head, but in their system the material to be cut secured to a rotatable table. The Sellafield system is designed to cut waste materials into smaller pieces and transfer those pieces into a waste drum, without direct operator intervention. The waste materials are likely to be significantly contaminated, so implementing a robotic solution helps to reduce human contact with these contaminated materials.

The waste item is scanned before being clamped on to the rotatable table. The scan is used to develop a cutting path plan, which the system uses to control the robot arm and laser cutting head. The cutting plan can be stepped through, or run automatically from start-to-finish. Human access to the cutting room is not

allowed during normal operations. A ‘fuse curtain’ is used to remove power to the system if the laser cuts outside of the safe cutting area.

2.3 Glovebox Operations

The AWE case study uses a robotic arm inside a glovebox[4] (or simply a containment box, since the robot arm removes the need for gloves) to manipulate active materials without human contact. This case study is entirely remote controlled, with an operator using a haptic interface at a workstation physically removed from the glovebox. This removes humans from the potential exposure to various radiological hazards and has the potential to reduce handling mistakes that can occur when using the thick gloves.

Since the robotic arm is only remote controlled, there is no autonomous system currently making decisions. To ensure that humans cannot be physically injured by the robot arm, there will be controls on glovebox access. Further, the remote control system must be designed to prevent the operator breaching the containment of the glovebox, which would be easier to do by accident than with a traditional glovebox because of the power of the robot arm.

[4] In this context, a glovebox is a sealed container with built-in gloves that allow an operator to manipulate objects inside the box without breaking containment.

3 Challenges of Remote-Controlled Robots

The first discussion session focussed on the hazards of the four case studies, described in Sect.2, as they are now. The challenges raised during the discussion were mostly concerned with the introduction and use of remote-controlled robotic systems, because of the low levels of autonomy currently envisaged in the case studies (see Sect.2).

The first concern was the ability of the robotic system to perform its task correctly. Cutting tasks can produce both dust and fumes, which may be chemically or radiologically dangerous, as well as sharp edges, which can be dangerous to humans or the robot itself. Since they are predictable, these sorts of hazards can be designed for.

A less predictable hazard is the robotic system making a mistake: performing a cut incorrectly or mis-handling material. These potential hazards could cause chemical or nuclear emissions, so they need to be mitigated. However, these sorts of hazards (both the predictable and less predictable varieties) are also present with a human performing the task directly. This suggests that the types of hazard mitigation and certification evidence used when there is a human operator provide a useful starting point for a robotic (or even autonomous) system.

There was also concern about the robot itself malfunctioning in some way. For remote-controlled robots, the operator might perform the task incorrectly or give some incorrect instruction. Also, the robot's control system (both hardware and software) could fail. Any of these could cause the robot to damage itself (laser cutting gone wrong, for example) or the physical containment of its environment (a robot arm crashing into a wall, for example). Again, a human performing a task will have similar hazards. An operator might perform a step or the entire task wrong or injure themselves. So looking at how these hazards are mitigated for a human operators would be a good place to start. However, this will not cover all hazards, for example: a human is less likely to be able to damage containment, because the containment was designed with human strength in mind.

Another impact of introducing a robotic system into nuclear sites that was identified, was on the design of the facilities themselves. There was some discussion that introducing a robotic system might expand the system's operating environment to include the working area of the remote-control operator, and the robot's maintenance and storage areas. It could extend further, to how the robot is transported between maintenance, storage, and operating areas. These concerns are, again, present with a human operator performing a task. The hazards involved in its maintenance, storage, and transport of a manual tool would also need to be considered alongside its actual use. However, using a robotic system will require some different considerations because of things like their size and complexity. These considerations may lead to changes in facility design.

Finally, there are several human aspects that will change after the introduction of a robotic system. Each of the case studies takes measures to mitigate the risk of harm to humans by separating the robot from humans, particularly during the robot's operation. Various methods of doing this were suggested, including software systems, hardware interlocks, and workplace protocols. It is likely that several methods will be used in parallel, as defence-in-depth. Even if the system is powered down (or similarly immobilised) for cleaning, maintenance, or upgrades, the staff involved in these tasks may still be exposed to chemical or radiological risks.

A similar concern was that the robot failing (either a mechanical failure of the hardware or a technical failure of the software) brings humans back to the hazardous situation that the robot was introduced to deal with. Fail safe options were discussed, to reduce the risks of recovery or removal of the robot in these situations. Examples include always being able to remove the robot to a safe maintenance areas, and unsafe cuts mechanically triggering the removal of the power supply to the robot.

The approaches to mitigating the identified hazards suggested, and then providing evidence of the efficacy of this mitigation, predominantly focussed on the physical. One of the traditional methods of mitigating a hazard in the nuclear sector has been to contain the hazardous situation in a sealed and constrained environment. But containing a robotic system could have a knock-on effect for the maintenance or upgrade of the system, which is likely to be performed by humans and will likely change over the long-term use of such systems. As previously described, these activities bring humans back into contact with the hazardous situation. Here, the hazards could come from the material the robot is working with or from the robot itself.

The discussions about providing evidence for the safety of the robotic system focussed on demonstrations. Some were physical demonstrations, for example the proposed containment surviving after an impact from the robot; others were simulation demonstrations, of the movement planning for example. While these demonstrations are useful, they are inherently only showing one set of possible behaviours. They show that the containment *may* survive an impact from the robot, and that the movement plans *can* be produced correctly, they are not able to show that the system *will* perform as expected. Simulation-based testing does allow large numbers of repetitions with different parameters, which allows some statistical information to be gathered. Plus, there is evidence to suggest that even low-fidelity simulations can replicate many different physical errors [4]. However, both physical and simulation-based testing are unlikely to be exhaustive.

One gap in the discussion was of the software that is involved in even a remote-controlled robotic system. Hardware interlocks often featured in the discussion of how to secure the robot and its working area. Again, these are useful and should be included; but if the software's safety is ensured as well, it can be part of a defence-in-depth argument for the system's safety. However, the discussions revealed that the default position of most nuclear operators was to not place any safety claims on the software and assume it will fail. Introducing robotic systems, even if they are only remote-controlled, will lead to more (and more sophisticated) software being part of safety-critical systems. Therefore, there is both a necessity and an opportunity for improving the techniques used to specify, design, and verify software used in nuclear systems. This will become even more important with increasing levels of autonomy in these systems, as we discuss in the next section.

4 Challenges for Autonomous Robots

The second discussion session focussed on the challenges of *autonomous* robotic systems. The same case studies were used, but it was imagined that the system was now autonomous – although the level of autonomy was not specified. The discussion was aimed to expose what kinds of autonomy would be used and for what tasks, what the perceived barriers to its introduction are, and the change in hazards and mitigations caused by its use.

The main application area for autonomy in the presented case studies was route/path planning. The two laser cutting case studies (Sect. 2.2) can already plan the cuts they are going to perform. Here, increased autonomy could, for example, allow the system to choose the start and end points of a cut as well as the path. In each of the case studies, the robot arms could be controlled by an autonomous system to plan and execute their movement. While useful, this is clearly not a leap to be made all in one go, nor without careful examination of the particular task(s) that the system is intended to perform.

Examining the task(s) that an autonomous system will perform is key its design being safe, especially in safety-critical situations. A correct and detailed task description is vital to the verification of the system. Tasks that are dangerous to humans or dull and repetitive could be handled by an autonomous system that can take the key decisions needed in the context of that task. For a complex task that contains some dangerous or dull elements, a skilled human operative could cooperate with an autonomous robotic system to make the task easier and safer. It was suggested that starting with decision support systems could be a low risk way of validating a model of the task – since it would not directly perform the actions it is choosing. As the decision support system becomes more competent, it could also improve the operator's trust in autonomy. One obvious barrier is the potential resistance to the introduction of a decision support system, which could lead to incomplete capturing of the task.

If we assume that the task has been correctly described and that the autonomous system implements the task correctly, then a fully autonomous system should reduce errors. There was also discussion that a fully autonomous system could simplify the system's safety case, because of the inherent benefits of it performing a task without deviation or fatigue. But this reliance on the autonomous system increases the hazards if the robot fails (as discussed in Sect. 3). It seems clear that even a fully autonomous system should be designed with fail safes in mind, for example: remote-control as a backup, the ability to remove the robot from its operating environment if it fails, and 'kill switches' to automatically cut the power. This effectively means that the fully autonomous system should keep semi-autonomy (or remote-control) as a backup option. Independent monitoring systems could alert skilled human operatives if any errors are about to occur, remove the need for a skilled human to supervise all the robot's actions.

While the move to greater autonomy brings some new issues, it clearly has the potential to vastly improve efficiency and throughput. Not only can an autonomous robot work *much* faster than a remote-controlled system, but can potentially work 24/7 and even continue working even with limited hardware failures [1].

As mentioned in Sect. 3, the robotic system failing can bring humans back into the hazardous environment (which has possibly been made more dangerous due to the failure). One concern was ensuring the physical reliability of a robotic

system over time, which raised challenges about how the system copes with upgrades and reconfigurations of both hardware and software. This was linked to the system's reliance on its environment not changing, which could cause problems. For example, the MASCOT case study is deployed inside a reactor, which can warp due to the heat and forces of the plasma inside. If the system is not able to adapt then the system's navigation will start producing incorrect plans – plans that, for example, cause the robot to hit a wall. Further, the challenge of system decommissioning or disposal must be considered during the design of the system and analysis of the task it will perform. To avoid this issue increases the hazards for human workers and potentially creates an unnecessary amounts of radioactive 'dead' robots.

Some common barriers to implementing autonomous systems did appear from the discussions, some of which can be resolved with better communication. A large barrier is the mindset of the sector that safety must be hardwired and that no safety claims should (or can) be made about software. Overcoming this challenge requires a difficult culture change. Another difficult challenge is the fear of job losses through automation, which is regularly in the news; the use of autonomous systems should compliment existing skills, allowing human workers to operate more efficiently and safely. There are perceptions that nuclear operators do not want autonomy and that the regulator (the ONR) will not readily give permission for the use of autonomy. This stand-off seems to have stalled development of even semi-autonomous systems in the nuclear sector until very recently. There also seems to be the perception, within the nuclear industry, that all autonomy is unsafe. As mentioned in Sect. 1.1, the amenability of an autonomous system depends heavily on how the autonomy has been implemented. Finally, the potential impact on facility design (Sect. 3) required for safe operation of semi- and fully-autonomous systems might prove difficult and costly for a lot of existing nuclear facilities because of their age.

Assuming that the challenges of using safety-critical software in the nuclear sector have been overcome, we need to ensure that the verification and validation techniques used are as robust as possible. This is because the software is now making decisions that have safety-critical consequences. One idea was to take a mixed-criticality approach; realising that some parts of the system are more safety-critical than others and using the strongest verification and validation techniques there, while still ensuring a baseline of safety and correctness for the less critical components. The intent, here, being to ease the verification and validation workload. Another discussion was about using 'self-certification', where the autonomous system is aware of its own safety and correctness properties and monitors its conformance with them at runtime. This is an ideal technique to be used alongside robust specification and design methods [3].

5 Conclusion

The workshop proved a useful chance to get representatives of nuclear operators, the regulator, and supply chain talking about the introduction of robotics and autonomous systems in to the nuclear sector. Feedback from the attendees indicated a willingness to get to grips with what autonomous robotic systems can do in the nuclear sector, and how they might be built and verified so that they are acceptable to the ONR.

The presentation and discussion of what tasks autonomous robotic systems could be used for, and potential directions for the verification and certification of these systems, seemed to be particularly well received. There was still a mentality of relying on physical mitigations and distrusting any software, which indicates that robust and reliable software design and verification approaches must be used or developed. Hopefully this will build trust while producing reliable software. Thankfully, there was an acceptance in the discussions that this is more of a marathon than a sprint, which allows space for the nuclear industry to incrementally introduce autonomous systems in a safe and controlled way.

The introduction of an autonomous system to the working environment, as with the introduction of any computer-based system, should be handled carefully. It should be a more efficient tool, not a worker replacement. The autonomous system's requirements must be based on an analysis of parts of the task where human workers would most benefit from its introduction and where it will be most effective. The robot must complement the existing skills of the workforce, which requires careful requirements gathering. Further, the autonomous system must be designed so that it fits into the existing workflow, or a new workflow, in a way that is useable for workers. The discussions also identified that there will be new training requirements to ensure that workers are able to effectively use this new tool.

The relevant good practice for the development of autonomous systems is still being developed. To this end, guidelines on how best to design and build autonomous systems so that they are amenable to robust analysis and verification methods will be very useful to developers, as would recommendation for the evidence needed for the regulation of autonomous systems. The guidance for designing, developing, verifying, and using autonomous systems requires a much wider discussion than just one industrial sector. Some of these guidelines and recommendations may be sector-specific, but a large part will be applicable to autonomous systems no matter where they are used. Ethical and social concerns should also be considered to ensure that the use of autonomous systems technologies does not cause (potentially unseen) harm to workers and the public. Involving a wider group to develop and scrutinise such guidelines reduces the risk of gaps or unintentionally bad recommendations.

References

[1] Aitken, J.M., Veres, S.M., Shaukat, A., Gao, Y., Cucco, E., Dennis, L.A., Fisher, M., Kuo, J.A., Robinson, T., Mort, P.E.: Autonomous Nuclear Waste Management. IEEE Intelligent Systems **33**(6), 47–55 (2018). doi:10.1109/MIS.2018.111144814, https://doi.org/10.1109/MIS.2018.111144814

[2] Beer, J.M., Fisk, A.D., Rogers, W.A.: Toward a Framework for Levels of Robot Autonomy in Human-Robot Interaction. Journal of Human-Robot Interaction **3**(2), 74 (jun 2014). doi:10.5898/JHRI.3.2.Beer, http://dl.acm.org/citation.cfm?id=3109833

[3] Fisher, M., Collins, E., Dennis, L., Luckcuck, M., Webster, M., Jump, M., Page, V., Patchett, C., Dinmohammadi, F., Flynn, D., Robu, V., Zhao, X.: Verifiable Self-Certifying Autonomous Systems. In: IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW). pp. 341–348. IEEE (oct 2018). doi:10.1109/ISSREW.2018.00028, https://ieeexplore.ieee.org/document/8539217/

[4] Machin, M., Dufossé, F., Blanquart, J.p., Guiochet, J., Powell, D., Waeselynck, H.: Specifying Safety Monitors for Autonomous Systems Using Model-Checking. In: Bondavalli, A., Di Giandomenico, F. (eds.) Computer Safety, Reliability, and Security, LNCS, vol. 8666, pp. 262–277. Springer International Publishing, Cham (2014). doi:10.1007/978-3-319-10506-2_18, http://link.springer.com/10.1007/978-3-319-10506-2_18

[5] SAE International: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. Tech. rep., Society of Automotive Engineers (2014). doi:10.4271/J3016_201401, https://doi.org/10.4271/J3016_201401