

Developing Guidelines for Autonomous Systems in Hazardous Environments

Matt Luckcuck

Department of Computer Science, Maynooth University, Ireland

16th of March 2022



ROBOTICS AND AI IN NUCLEAR

Guidelines for Autonomous Robots

- ▶ “White Paper” – report or guide that ‘provides a basis for further consultation and discussion’

Guidelines for Autonomous Robots

- ▶ “White Paper” – report or guide that ‘provides a basis for further consultation and discussion’
- ▶ *Principles for the Development and Assurance of Autonomous Systems for Safe Use in Hazardous Environments*

Guidelines for Autonomous Robots

- ▶ “White Paper” – report or guide that ‘provides a basis for further consultation and discussion’
- ▶ *Principles for the Development and Assurance of Autonomous Systems for Safe Use in Hazardous Environments*
- ▶ Co-written with help from the Office for Nuclear Regulation (ONR)

Guidelines for Autonomous Robots

- ▶ “White Paper” – report or guide that ‘provides a basis for further consultation and discussion’
- ▶ *Principles for the Development and Assurance of Autonomous Systems for Safe Use in Hazardous Environments*
- ▶ Co-written with help from the Office for Nuclear Regulation (ONR)
- ▶ **Autonomous** system makes decisions requiring intelligence and situational understanding

Guidelines for Autonomous Robots

- ▶ “White Paper” – report or guide that ‘provides a basis for further consultation and discussion’
- ▶ *Principles for the Development and Assurance of Autonomous Systems for Safe Use in Hazardous Environments*
- ▶ Co-written with help from the Office for Nuclear Regulation (ONR)
- ▶ **Autonomous** system makes decisions requiring intelligence and situational understanding
- ▶ Often focussing on autonomous **robots**...

Guidelines for Autonomous Robots

- ▶ “White Paper” – report or guide that ‘provides a basis for further consultation and discussion’
- ▶ *Principles for the Development and Assurance of Autonomous Systems for Safe Use in Hazardous Environments*
- ▶ Co-written with help from the Office for Nuclear Regulation (ONR)
- ▶ **Autonomous** system makes decisions requiring intelligence and situational understanding
- ▶ Often focussing on autonomous **robots**...
 - ▶ Physical

Guidelines for Autonomous Robots

- ▶ “White Paper” – report or guide that ‘provides a basis for further consultation and discussion’
- ▶ *Principles for the Development and Assurance of Autonomous Systems for Safe Use in Hazardous Environments*
- ▶ Co-written with help from the Office for Nuclear Regulation (ONR)
- ▶ **Autonomous** system makes decisions requiring intelligence and situational understanding
- ▶ Often focussing on autonomous **robots**...
 - ▶ Physical
 - ▶ Logical

Nuclear Robotics Certification Workshops

- ▶ Workshops with ONR
- ▶ Open forum for nuclear operators, supply chain, and regulator, plus academia
- ▶ Exploring...
 - ▶ How is the safety of nuclear robots assessed?
 - ▶ What changes with robotics and autonomy?
- ▶ Aiming...
 - ▶ Clarify the questions
 - ▶ Bridge gaps in knowledge

Workshops

First workshop – Sept. 2018

- ▶ Introduce nuclear safety assessment
- ▶ Highlight capabilities and challenges of autonomy
- ▶ Talks from:
 - ▶ Robotics and AI in Nuclear (RAIN) Hub
 - ▶ Assuring Autonomy International Programme (AAIP, University of York)
 - ▶ Office for Nuclear Regulation (ONR)
- ▶ Group discussion
- ▶ Website: tiny.cc/SafetyCaseWorkshop1

Second Workshop – Apr. 2019

- ▶ Scope challenges of autonomous robots in the nuclear industry
- ▶ Four case studies:
 - ▶ UK Atomic Energy Authority
 - ▶ National Nuclear Laboratory
 - ▶ Sellafield
 - ▶ Atomic Weapons Establishment
- ▶ Discussion Sessions
 - ▶ Current Hazards
 - ▶ Future (Autonomy) Hazards
- ▶ Website: tiny.cc/SafetyCaseWorkshop2

White Paper

Scope

- ▶ Good practice for developing autonomous robotic systems amenable to strong V & V
- ▶ Add-on to existing standards and guidance
- ▶ Audience: developers and verifiers of autonomous and robotic systems
- ▶ Authors:
 - ▶ Matt Luckcuck, Louise Dennis, Michael Fisher (RAIN)
 - ▶ Steve Frost, Andy White, Doug Styles (ONR)

**PRINCIPLES FOR
THE DEVELOPMENT
AND ASSURANCE
OF AUTONOMOUS
SYSTEMS FOR SAFE
USE IN HAZARDOUS
ENVIRONMENTS**



Available:

[http://tiny.cc/
AutonomyWhitePaper](http://tiny.cc/AutonomyWhitePaper)

High-Level Recommendations

- 1 System has hardware *and* software

High-Level Recommendations

- 1 System has hardware *and* software
- 2 Hazard assessments to include ethical hazards

High-Level Recommendations

- 1 System has hardware *and* software
- 2 Hazard assessments to include ethical hazards
- 3 Take both a corroborative and a mixed-criticality approach to Verification & Validation

High-Level Recommendations

- 1 System has hardware *and* software
- 2 Hazard assessments to include ethical hazards
- 3 Take both a corroborative and a mixed-criticality approach to Verification & Validation
 - ▶ Mixed-criticality approach

High-Level Recommendations

- 1 System has hardware *and* software
- 2 Hazard assessments to include ethical hazards
- 3 Take both a corroborative and a mixed-criticality approach to Verification & Validation
 - ▶ Mixed-criticality approach
 - ▶ Corroborative approach (*A corroborative approach to verification and validation of human-robot teams* Webster et al. 2019)

High-Level Recommendations

- 1 System has hardware *and* software
- 2 Hazard assessments to include ethical hazards
- 3 Take both a corroborative and a mixed-criticality approach to Verification & Validation
 - ▶ Mixed-criticality approach
 - ▶ Corroborative approach (*A corroborative approach to verification and validation of human-robot teams* Webster et al. 2019)
- 4 Autonomous components should be transparent and verifiable

High-Level Recommendations

- 1 System has hardware *and* software
- 2 Hazard assessments to include ethical hazards
- 3 Take both a corroborative and a mixed-criticality approach to Verification & Validation
 - ▶ Mixed-criticality approach
 - ▶ Corroborative approach (*A corroborative approach to verification and validation of human-robot teams Webster et al. 2019*)
- 4 Autonomous components should be transparent and verifiable
- 5 Tasks and missions should be clearly defined

High-Level Recommendations

- 1 System has hardware *and* software
- 2 Hazard assessments to include ethical hazards
- 3 Take both a corroborative and a mixed-criticality approach to Verification & Validation
 - ▶ Mixed-criticality approach
 - ▶ Corroborative approach (*A corroborative approach to verification and validation of human-robot teams* Webster et al. 2019)
- 4 Autonomous components should be transparent and verifiable
- 5 Tasks and missions should be clearly defined
- 6 Dynamic verification to complement static verification

High-Level Recommendations

- 1 System has hardware *and* software
- 2 Hazard assessments to include ethical hazards
- 3 Take both a corroborative and a mixed-criticality approach to Verification & Validation
 - ▶ Mixed-criticality approach
 - ▶ Corroborative approach (*A corroborative approach to verification and validation of human-robot teams Webster et al. 2019*)
- 4 Autonomous components should be transparent and verifiable
- 5 Tasks and missions should be clearly defined
- 6 Dynamic verification to complement static verification
- 7 Requirements traceability, through development and into deployed system

Three Themes for Autonomous Systems

Three Themes for Autonomous Systems

Assess for Ethics...

- ▶ Consider hazards that aren't just 'safety' or 'security'
- ▶ E.G. machine learning system trained on biased data
- ▶ Avoid these problems being 'baked in' to the system
- ▶ One direction: BS 8611 "Ethical Design and Application of Robots and Robotic Systems"

Three Themes for Autonomous Systems

Assess for Ethics... (as well)

- ▶ Consider hazards that aren't just 'safety' or 'security'
- ▶ E.G. machine learning system trained on biased data
- ▶ Avoid these problems being 'baked in' to the system
- ▶ One direction: BS 8611 "Ethical Design and Application of Robots and Robotic Systems"

Your Sector isn't Special

- ▶ Introducing autonomy bring common challenges
- ▶ Autonomy challenges should be tackled in a sector-agnostic way
- ▶ Autonomous *robotic* systems may be dangerous to physically test in early development
 - ▶ Lean on code analysis and simulation
- ▶ Autonomous system's decisions
 - ▶ Analysable
 - ▶ Cover reactions to unexpected events in environment

Design for Verification

- ▶ Autonomous components (may) make decisions without human approval
 - ▶ Transparent – enable examination
 - ▶ Verifiable – correctness of behaviour
 - ▶ Explainable – understandable by (normal) humans
- ▶ E.G. Formal techniques exist for verifying high-level decisions
- ▶ Ultimately...

Design for Verification

- ▶ Autonomous components (may) make decisions without human approval
 - ▶ Transparent – enable examination
 - ▶ Verifiable – correctness of behaviour
 - ▶ Explainable – understandable by (normal) humans
- ▶ E.G. Formal techniques exist for verifying high-level decisions
- ▶ Ultimately... How will you explain it to the regulator?

Summary

- ▶ Autonomous Systems: New Challenges and New Opportunities

Summary

- ▶ Autonomous Systems: New Challenges and New Opportunities
- ▶ Challenges...
 - ▶ Safety and security
 - ▶ Other ethical issues: privacy, equality, and human autonomy
 - ▶ Many challenges cross-sector

Summary

- ▶ Autonomous Systems: New Challenges and New Opportunities
- ▶ Challenges...
 - ▶ Safety and security
 - ▶ Other ethical issues: privacy, equality, and human autonomy
 - ▶ Many challenges cross-sector
- ▶ Opportunities...
 - ▶ Help with dangerous jobs
 - ▶ Inspect hazardous areas
 - ▶ Enable decisions to be closely analysed

Summary

- ▶ Autonomous Systems: New Challenges and New Opportunities
- ▶ Challenges...
 - ▶ Safety and security
 - ▶ Other ethical issues: privacy, equality, and human autonomy
 - ▶ Many challenges cross-sector
- ▶ Opportunities...
 - ▶ Help with dangerous jobs
 - ▶ Inspect hazardous areas
 - ▶ Enable decisions to be closely analysed
- ▶ Price: consider how the system will be introduced and used, and design for verification
- ▶ Important for regulatory sign-off and worker/public trust

Links

- ▶ Workshop 1: <http://tiny.cc/SafetyCaseWorkshop1>
- ▶ Workshop 2 (including link to report): <http://tiny.cc/SafetyCaseWorkshop2>
- ▶ White Paper: <http://tiny.cc/AutonomyWhitePaper>
- ▶ UoM Policy Blog Post: <http://tiny.cc/UoMBlog>

Thanks