KU LEUVEN

Safety in a Changing World

The Case for Human-Centred Dynamic Safety Cases

IEEE Let’s Meet - Bruges

Laure Buysse

Who am I?

- Dual Specialisation in EO and ICT (Engineering Technology)
- TETRA (Technology Transfer) Safety Assurance in Industry 4.0:
  - Cooperative Robot
  - Autonomous Mobile Robot (AMR / AGV)
- FWO PhD Fellowship Strategic Basic Research
- Research:
  - Safety of Autonomous Systems ; Robotics
    - Safety cases ; Hazard analysis techniques
  - Safety of Artificial Intelligence
  - Digital Twins

Autonomous Systems


Hercules Robot
Amazon

Autonomous Taxi

Waymo


Semi-Autonomous Assembly Line
General Motors

A Revolution for Industry

Autonomous systems

- Systems, machines, robots...
- Operating independent based on previous and current inputs

- Boost in Productivity
- Improved Safety
- Less Mundane Task for Humans
- Extend human capabilities
- Cost Efficient
- Multi-Functional

But ...

Industrial robot crushes man to death in South Korean distribution centre

Machine apparently identified man inspecting it as one of the boxes it was stacking

B B C

Tesla Autopilot feature was inv
13 fatal crashes, US regulator say

staff injuries'

30 September 2020

Amazon D
Jeff Bezos' Delivery Dreams

Billions of dollars and a decade later, and Amazon's delivery by drone program still isn't off the ground.

bad delayed
o officials say

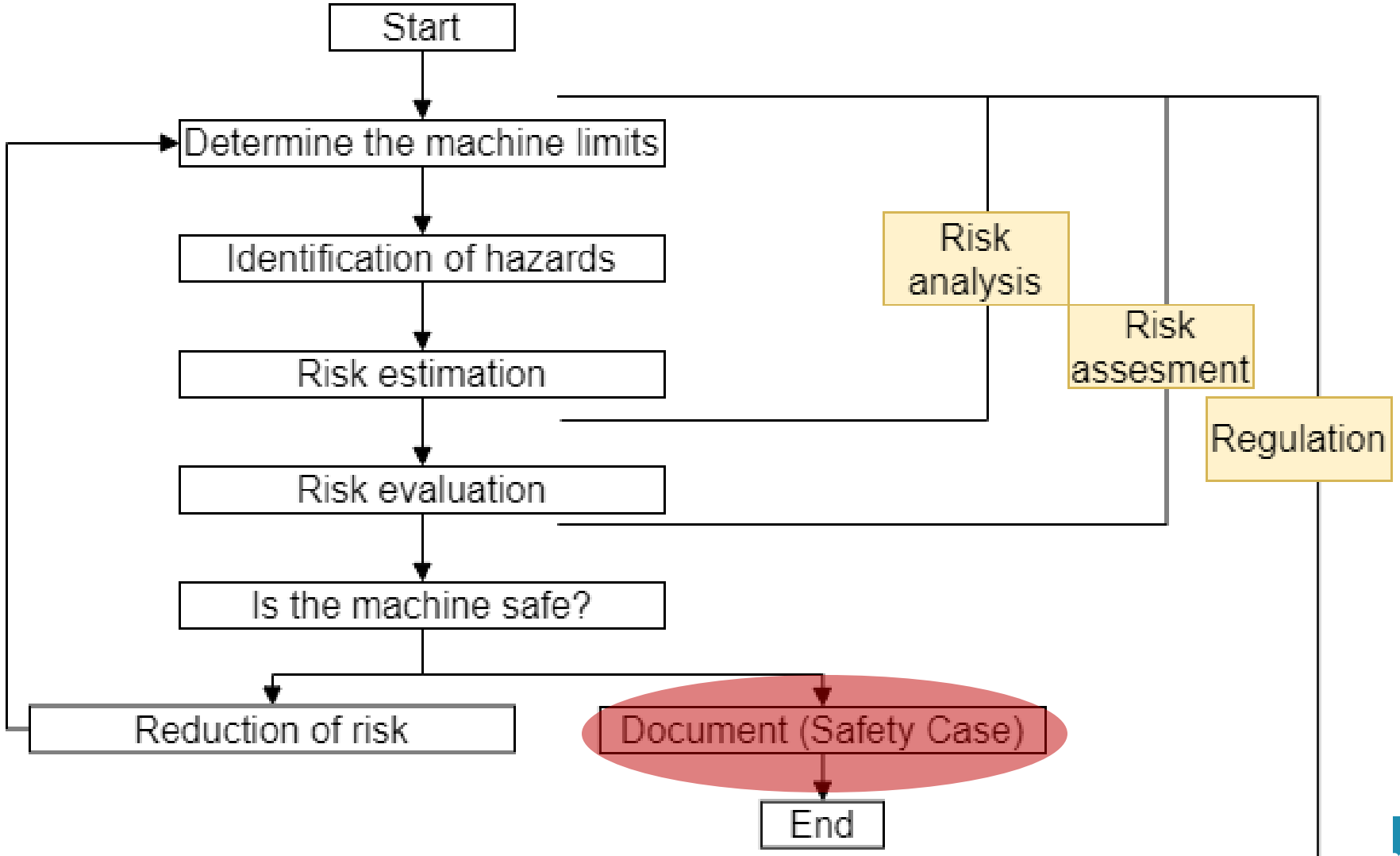
Share Save

e, the robotaxi firm, denies the city's claims its vehicle
ed ambulance which resulted in injured person's
n

Safety (assurance) is still a major roadblock in designing, developing and deploying these systems.

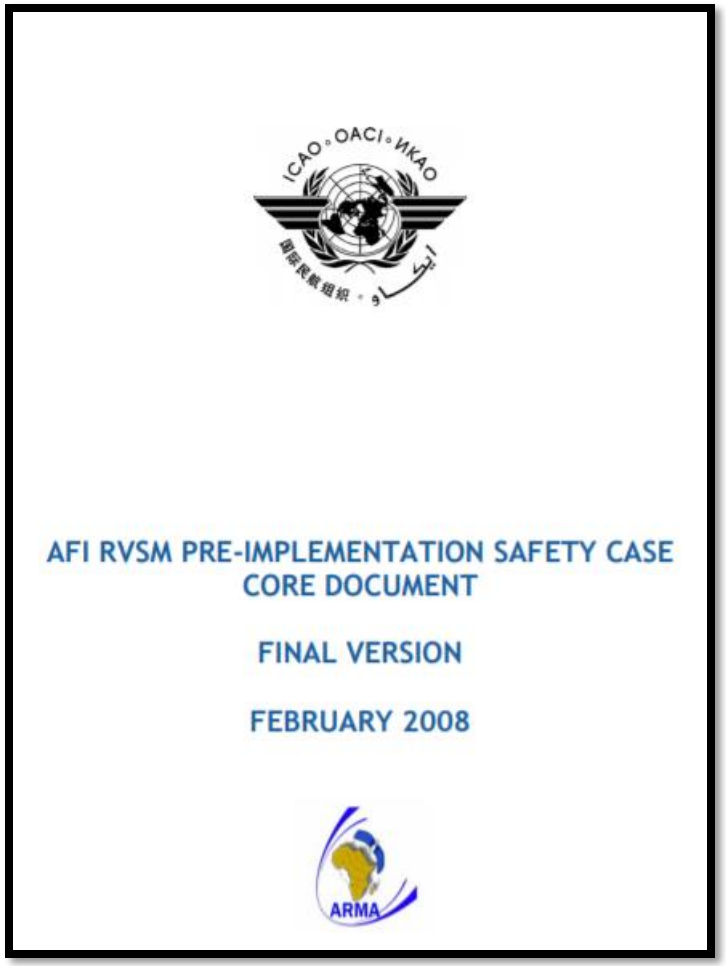
Safety Assurance

How to


Start
Determine the machine limits
Identification of hazards
Risk estimation
Risk evaluation
Is the machine safe?
Reduction of risk
Document (Safety Case)
End
Risk analysis
Risk assesment
Regulation

What is a (static) safety case?

- Contents:
  - Claims about property/system
  - Arguments logically linking evidence and assumptions to the claim(s)
  - Evidence and assumptions supporting the arguments
  - Justification of the choice of top-level claim and the method of reasoning
- Structure not universal (! Sector specific standards !)
- Requires a safety analysis



Safety Cases – They seem perfect ...

- Difference between the *actual*, the *depicted* and the *explained*

The gap can lead to "**a culture of 'paper safety' at the expense of real safety**".*

- Our initial state of belief in safety is based on *predictions* and *assumptions*. We should always be aware of *uncertainty* and the fact that we are designing under *imperfect knowledge*.

C. Haddon-Cave, "*The Nimrod Review: An Independent Review into the Broader Issues surrounding the Loss of the RAF Nimrod MR2 Aircraft XV230 in Afghanistan in 2006*," Report, The Stationery Office, London, UK, Oct. 2009.

Safety Cases – They seem perfect ...

- The content of a safety case contains many different elements, such as
  - System design
  - System configuration
  - Intended environment
  - Identified hazard
  - Risk mitigation principles

All of these can, and often do, **change**, especially when dealing with autonomous systems. Moreover, **modular system** show great variation by default and all **different domains and environments** need to be included.

Dynamic Safety Cases

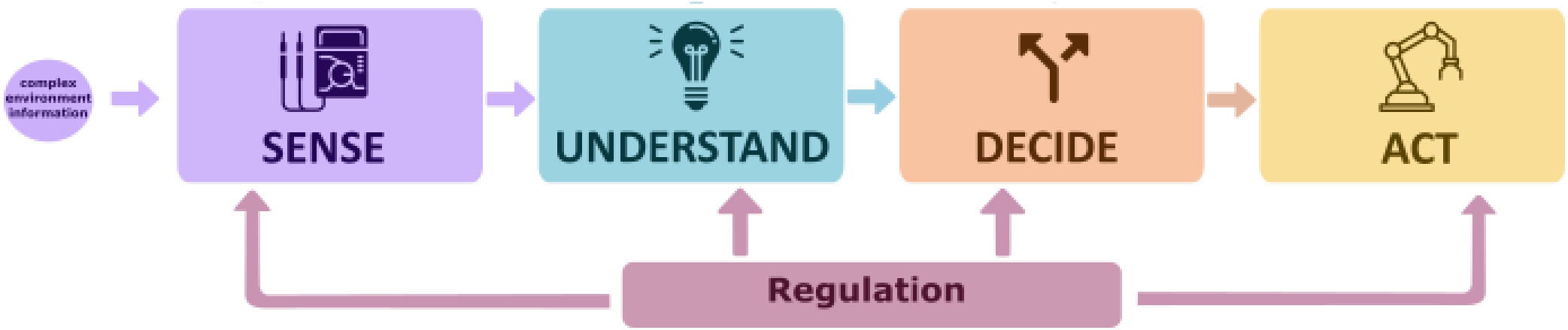
Dynamic Safety Cases

To continuously assess confidence in the validity of the safety case through-life and proactively update the arguments and reasoning of the safety case

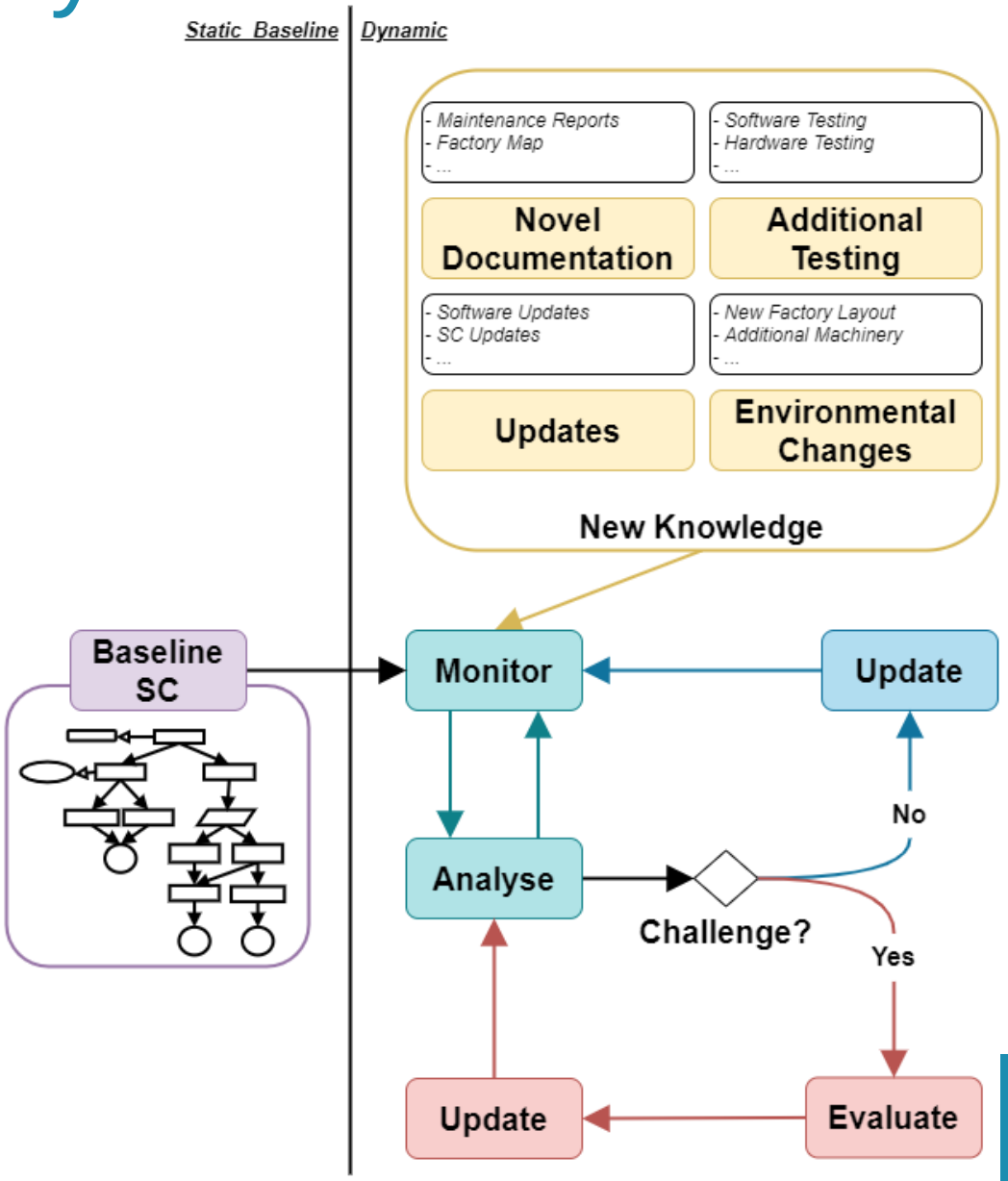
- The fundamental idea behind a DSC is to **reduce the gap between safety at run-time and the safety case as drawn up during design** and development.Especially for autonomous systems, where
  - Exhaustive testing is impossible
  - The design is made under imperfect knowledge
  - the system is prone to emergent behaviour.

- The strength of the extended DSC framework lies in
  - (1) filling in unknowns at run-time wherever possible and
  - (2) keeping track of the assurance deficits or unknowns that cannot be resolved during system design

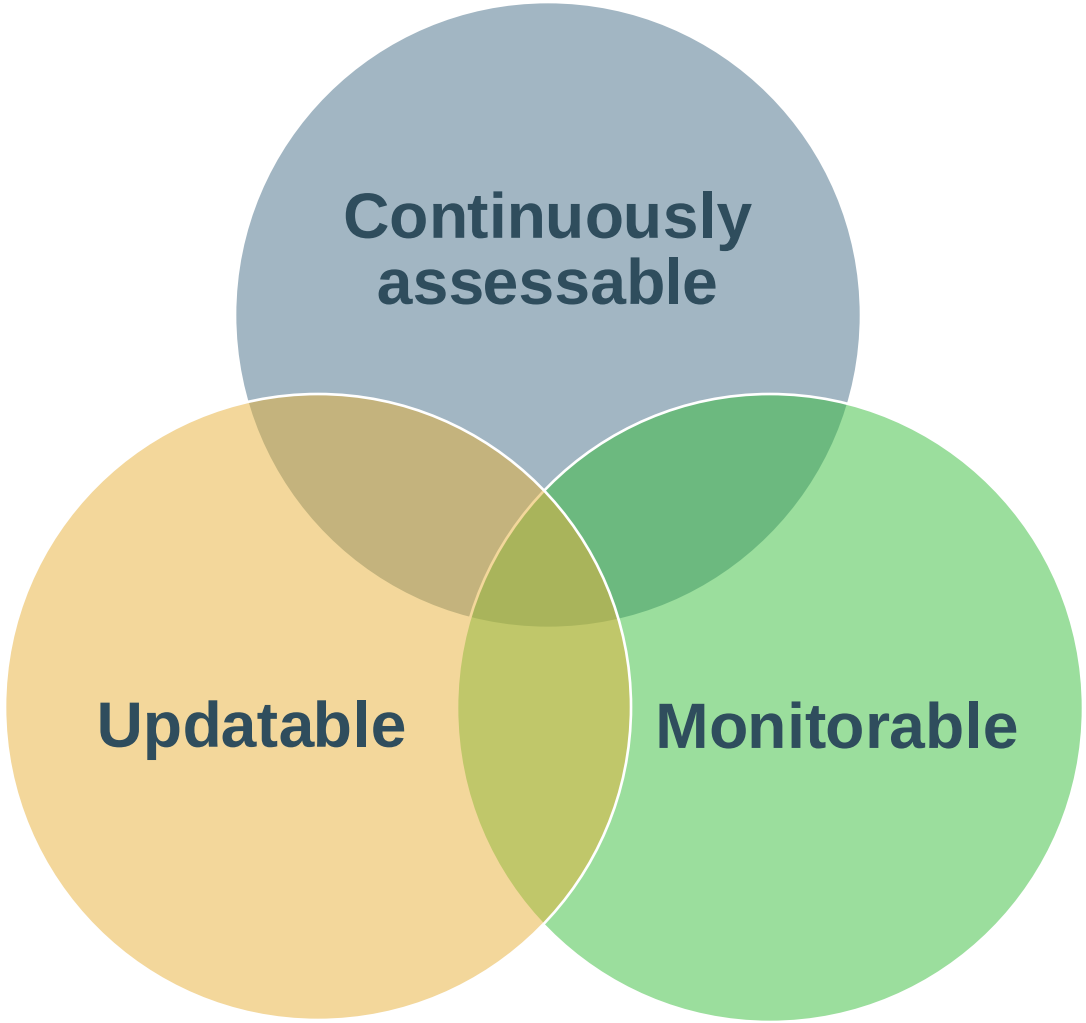
Dynamic Safety Case



Dynamic Safety Cases - Process



Dynamic Safety Cases - Characteristics



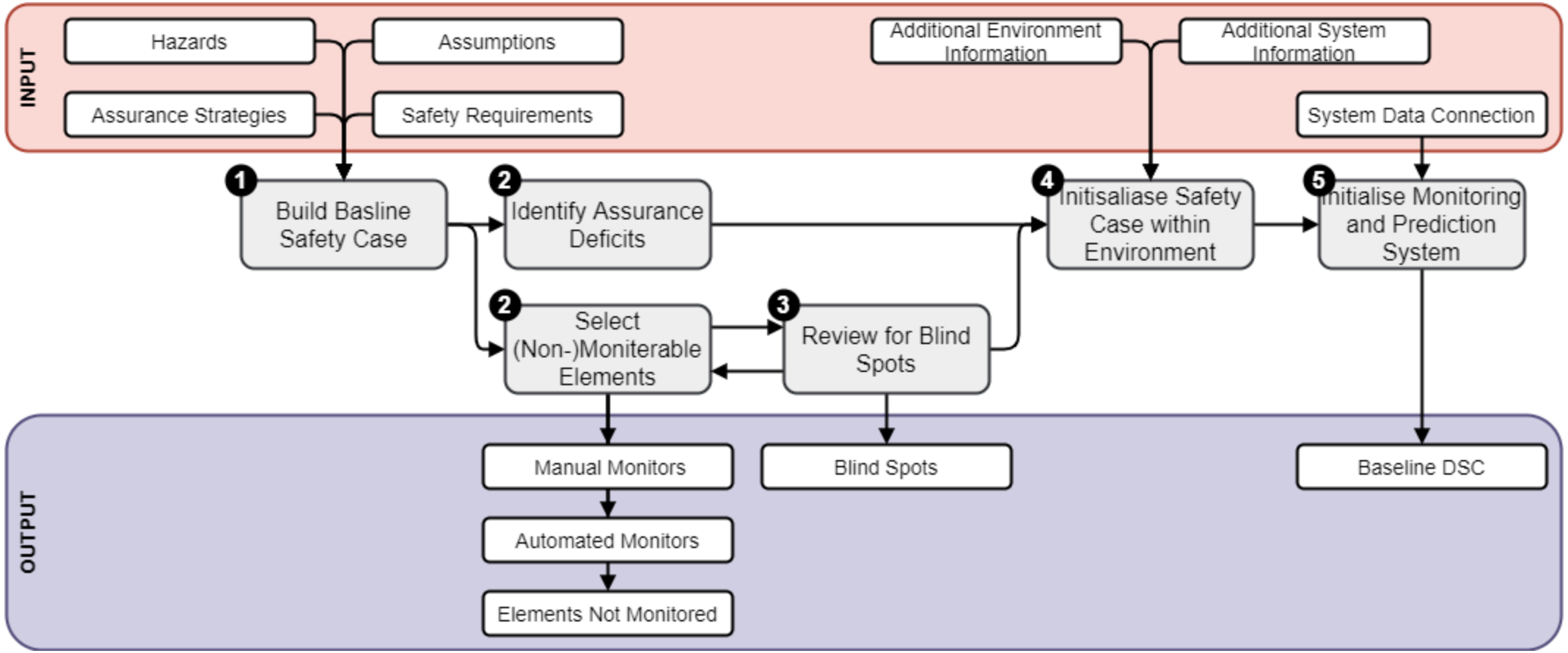
Dynamic Safety Cases – Differentiating Between Systems

Different systems operating in various environments with changing levels of uncertainty and complexity inevitably require a different degree of care at run-time

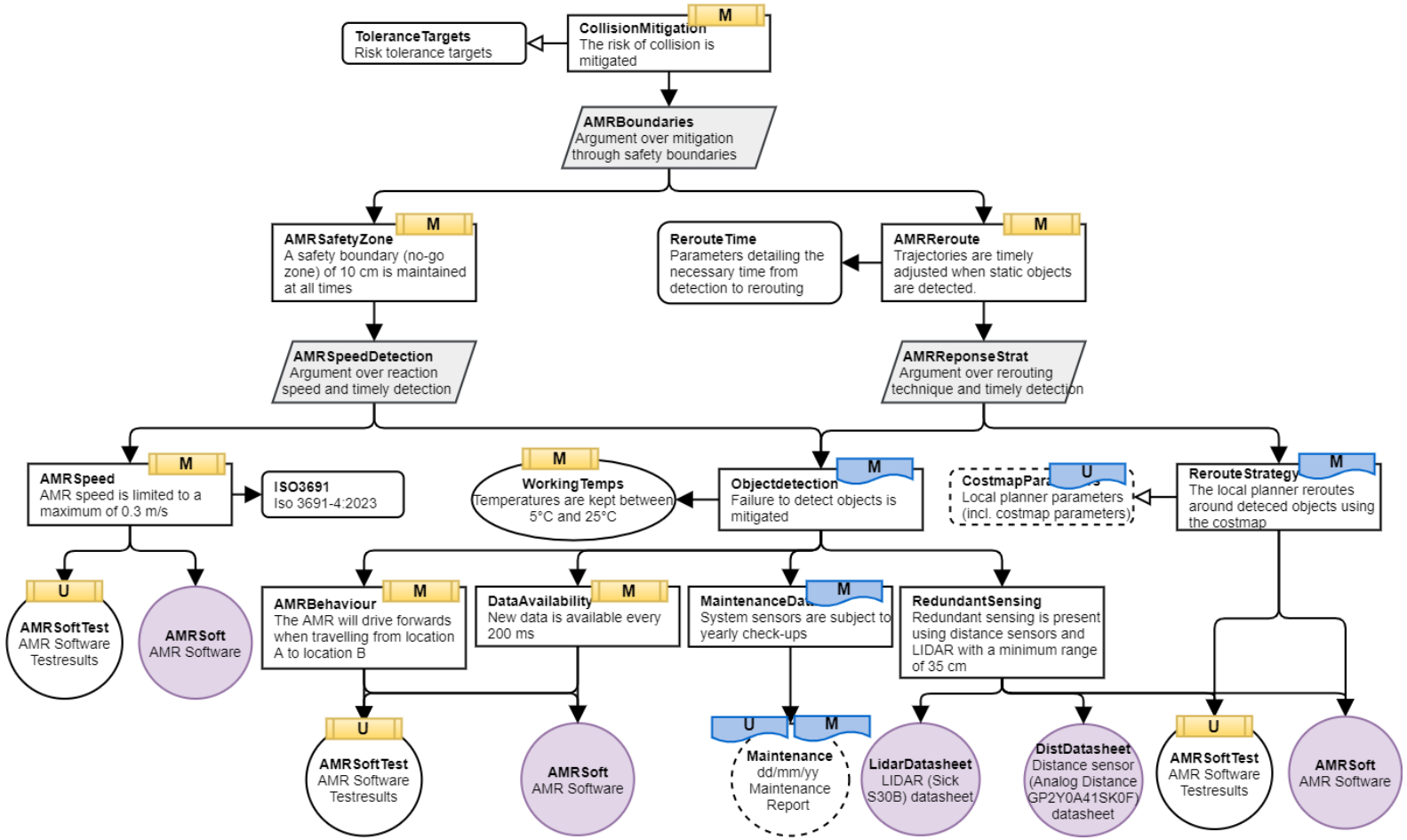
Monitorable			Assessable		
	Setup	Coverage	Representativeness Frequency	Depth	Procedure	Trigger
1	Ad-Hoc	Partial			to case con-ation	Fault detection
2	Systematic	Partial (combining formal and informal)			iited) Guid-	Fault detection ; Change of knowledge
3	Systematic	Sufficient or Full			nsive guid-	Fault detection ; Change in knowledge ; Systematic checks

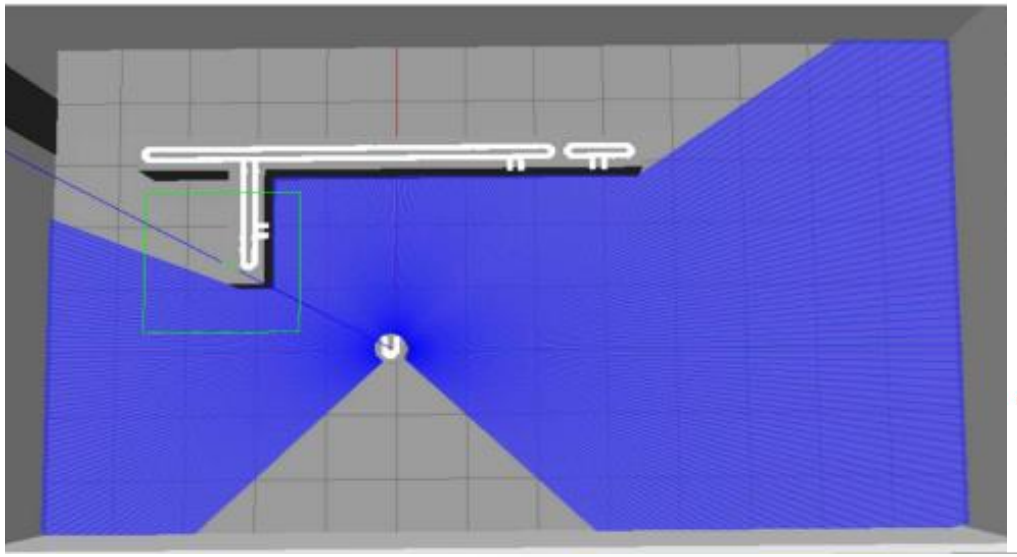
Updateable			
	Content	Procedure	Trigger
1	Single elements	Case to case consideration	Change in system, environment or knowledge
2	Single elements ; Patterns ; Single elements ;	(Limited) Guidance	Change in system, environment or knowledge ; Fault
3	New / Altered argumentation ; Pattern completion ; Single elements	Extensive guidance	Change in system, environment or knowledge; Fault ; Systematic

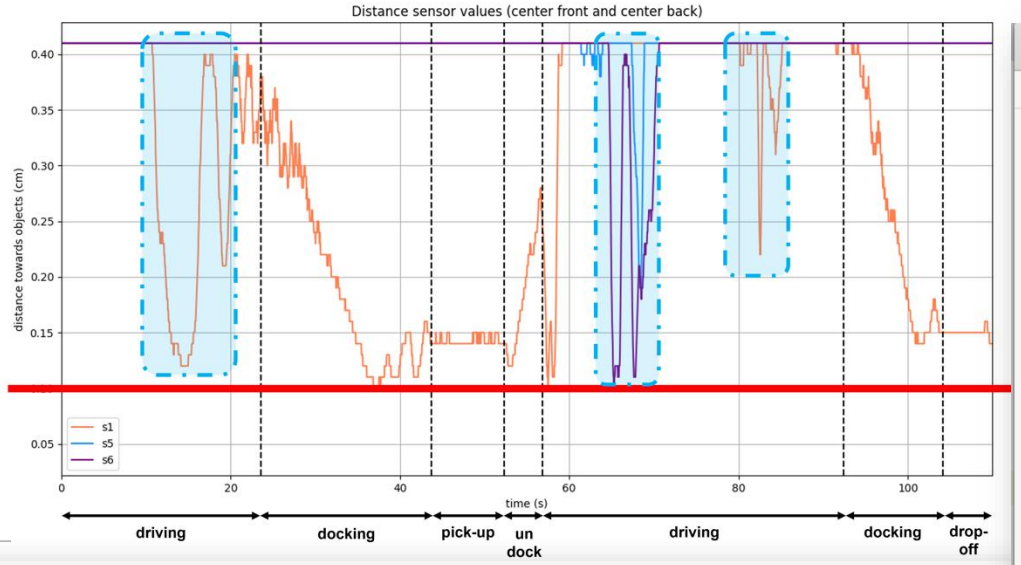
Dynamic Safety Cases – Practical Workflow



Dynamic Safety Cases – Practical Workflow







main.py

File Help

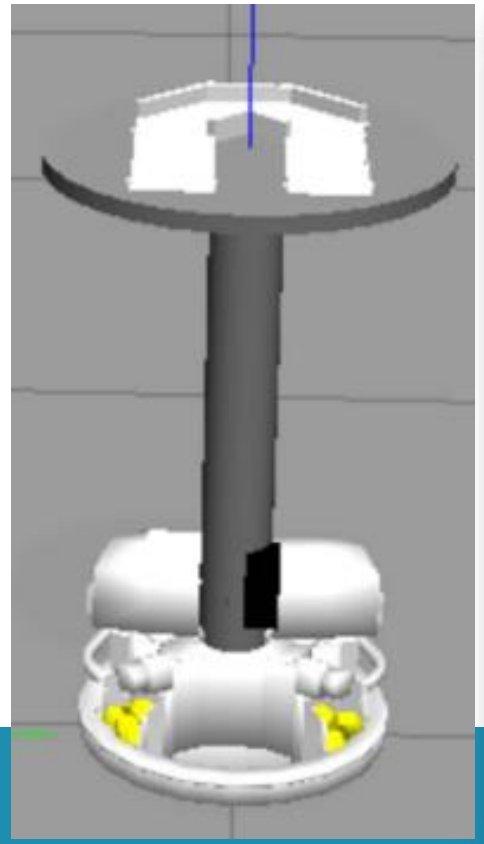
Initiliase | Start | Stop

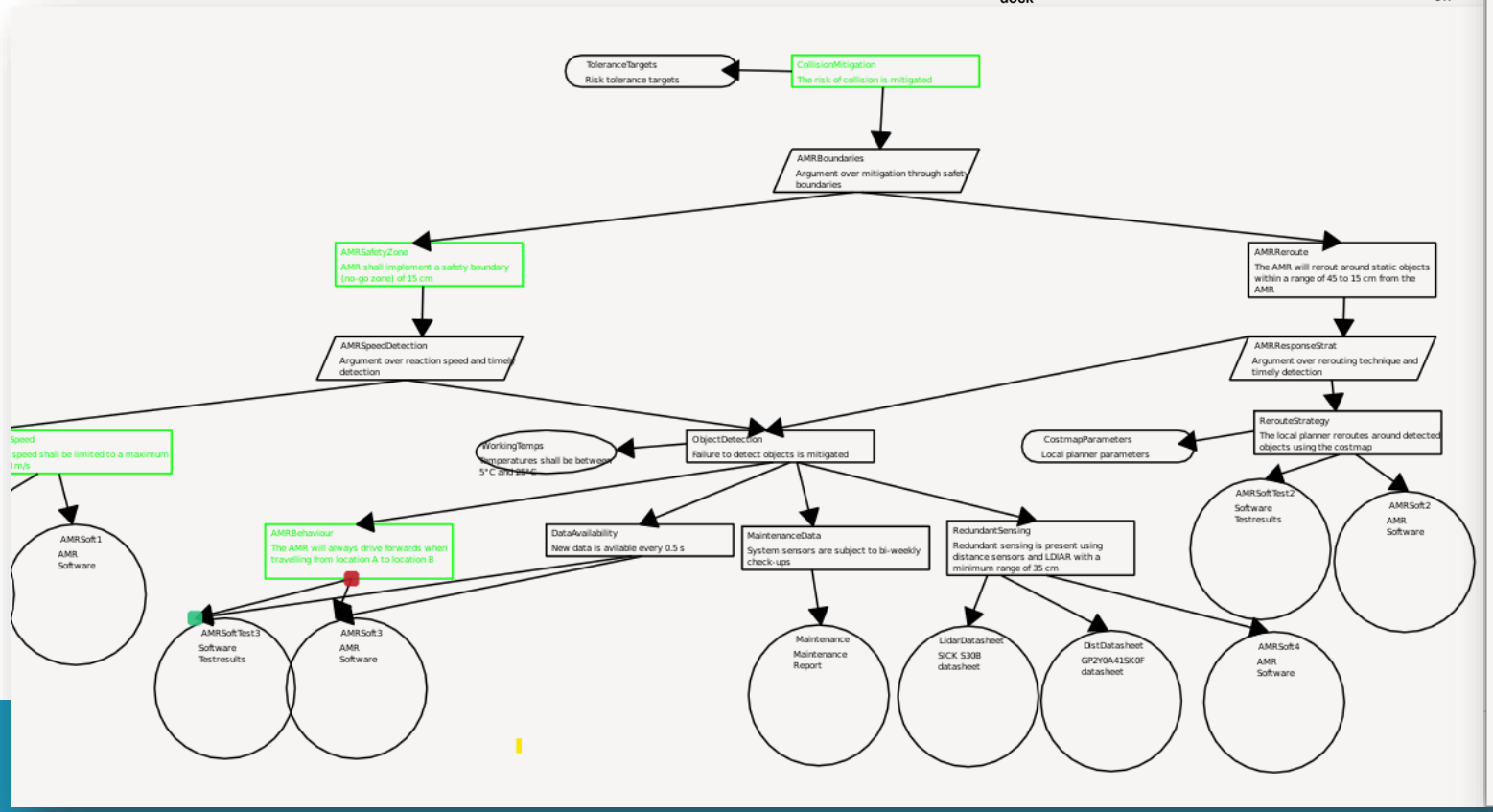
Visualise Safety Case

ID	Logic Formula
CollisionMitigation	{vel > 0.0, bumper: false}
AMRSafetyZone	{vel > 0.0} -> {s1 > 0.15, s2 > 0.15,
AMRSpeed	{vel < 0.3}
AMRBehaviour	{vel_x_lin > -0.01}

Additional Monitors

ID	Logic Formula	Requirement





Conclusion

Conclusion

- Safety first and safety last

- Safety cases as a useful practice to:
  - Reduce risk
  - Amplify confidence
  - Document conformance
  - Bundle all safety related information

- Dynamic safety case as an extension of the safety management system to deal with operational uncertainty inherent to autonomous systems.

Questions?