Position paper: a vision for the dynamic safety assurance of ML-enabled autonomous driving systems

Requirement Engineering conference

MoDRE 2023

Alvine Boaye Belle[*], Hadi Hemmati[*], Timothy C. Lethbridge[#]

[*]*York University, Toronto, Canada*

[#]*University of Ottawa, Ottawa, Canada*

YORK U

Agenda

- Overview of assurance cases
- High-level overview of the proposed approach
- Part I: providing a higher dynamic safety assurance
- Part II: hazard elicitation and mitigation to increase dynamic safety assurance
- Part III: analysis of the barriers to safety regulation compliance
- Concluding remarks

What is an assurance case?

An assurance case is a document that eases the exchange of information between:

- Various system stakeholders (e.g., suppliers, acquirers)
- And between the operator and regulator, where the knowledge regarding a system's requirements is convincingly conveyed.
- Requirements: safety, security, reliability, etc.

Assurance cases are structured as a hierarchy of claims:

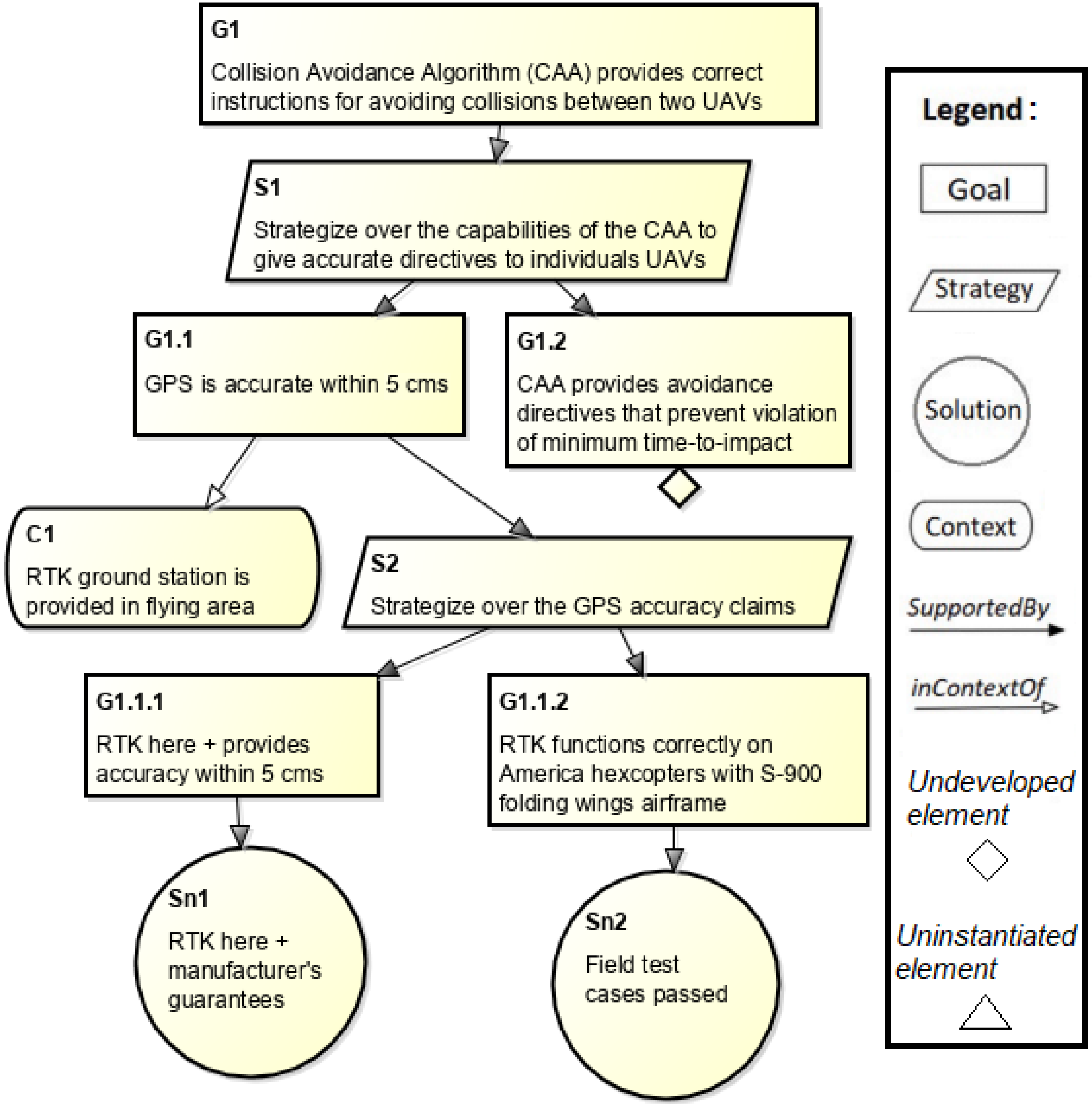
- Lower-level claims draw on concrete evidence, and serve as evidence to justify claims higher in the hierarchy.
- The top claim is a statement such as a system supports non-obvious requirements.

In assurance cases, concrete facts serve as evidence relevant to desirable requirements:

- Algorithms, test results, formal reviews, simulations, resource diagrams and various system artifacts.

How does an assurance case look like?

- An assurance case (e.g., safety case) allows demonstrating that a system will:
  - Satisfy particular requirements (e.g., safety, security)
  - Along with supporting evidence.
- It also allows checking the compliance of systems with standards to support their certification.
- It is represented using various notations such as **GSN** (Goal Structuring Notation)
- It can be assessed using **confidence** and **uncertainty** measures.



G1
Collision Avoidance Algorithm (CAA) provides correct instructions for avoiding collisions between two UAVs
S1
Strategize over the capabilities of the CAA to give accurate directives to individuals UAVs
G1.1
GPS is accurate within 5 cms
G1.2
CAA provides avoidance directives that prevent violation of minimum time-to-impact
C1
RTK ground station is provided in flying area
S2
Strategize over the GPS accuracy claims
G1.1.1
RTK here + provides accuracy within 5 cms
G1.1.2
RTK functions correctly on America hexcopters with S-900 folding wings airframe
Sn1
RTK here + manufacturer's guarantees
Sn2
Field test cases passed
Legend :
Goal
Strategy
Solution
Context
SupportedBy
inContextOf
Undeveloped element
Uninstantiated element

Fig. 1 Partial safety case for UAV Collision Avoidance –adapted from Vierhauser et al. (2019)

Adoption & challenges

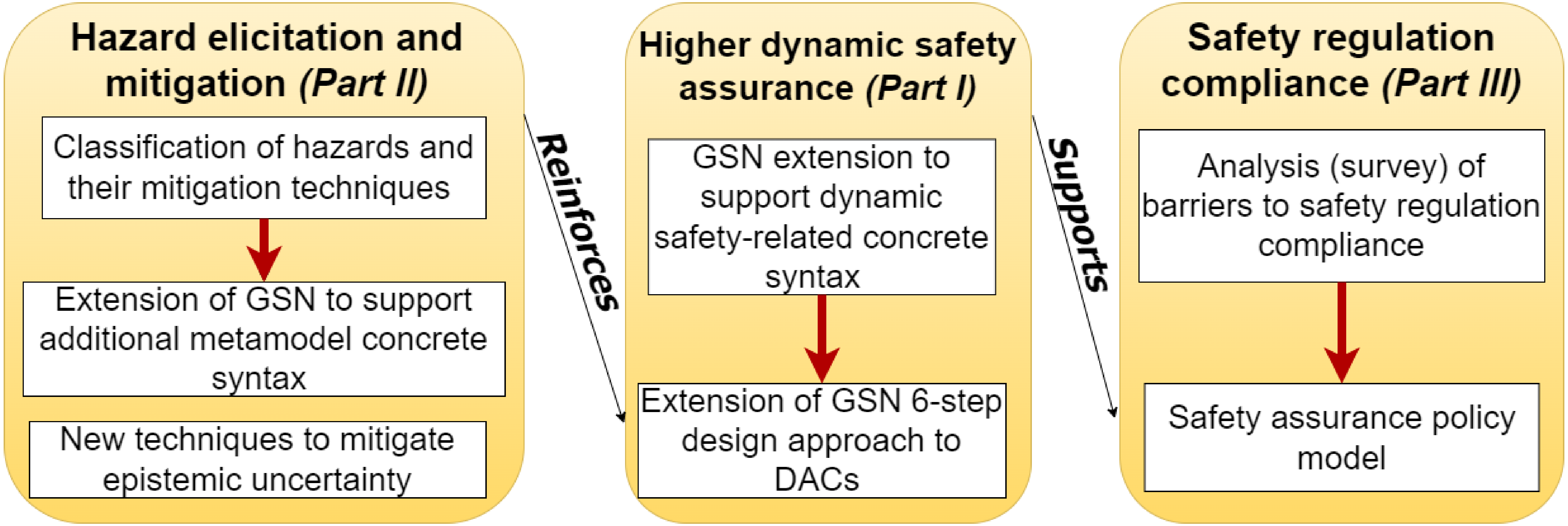
- The popularity and adoption of assurance cases is increasing.
- Assurance cases are mostly used in safety–critical domains to deal with high-risk concerns and demonstrate to stakeholders that safety–critical systems are **safe** according to domain-specific criteria.
- It is usually mandatory that the design authority (manufacturer) develops compelling assurance cases to support that justification and allow regulatory bodies (e.g., NHTSA) to certify such systems.
- The use of assurance cases is also recommended by several international standards such as ISO 26262.
- **But most assurance cases are static i.e., only suitable prior to a system's deployment:**
    - **They may become incorrect, obsolete or even inadequate during the system operation.**

Proposed solution? Assuring that autonomous driving systems (ADSs) are safer throughout their lifecycle

- Focus on safety since it is a life-critical requirement
- Focus on dynamic assurance
- Focus on autonomous driving systems:
    - Their failure could have catastrophic outcomes (e.g., severe injuries, loss of lives).
- Focus on safety standards for the automotive domain.

Source: https://getcruise.com/

High-level overview of the proposed 3-part approach



Hazard elicitation and mitigation *(Part II)*

Classification of hazards and their mitigation techniques

Extension of GSN to support additional metamodel concrete syntax

New techniques to mitigate epistemic uncertainty

Reinforces

Higher dynamic safety assurance *(Part I)*

GSN extension to support dynamic safety-related concrete syntax

Extension of GSN 6-step design approach to DACs

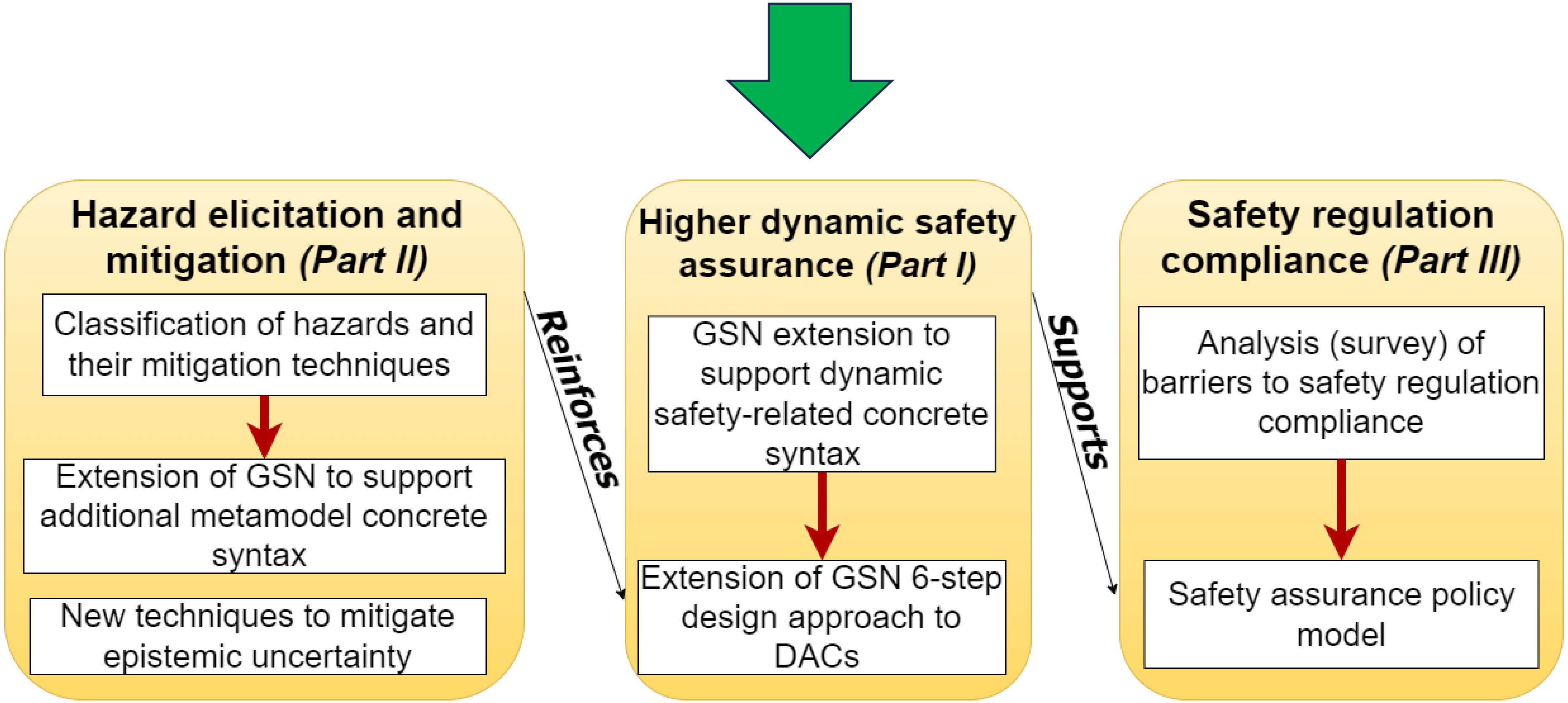
Supports

Safety regulation compliance *(Part III)*

Analysis (survey) of barriers to safety regulation compliance

Safety assurance policy model

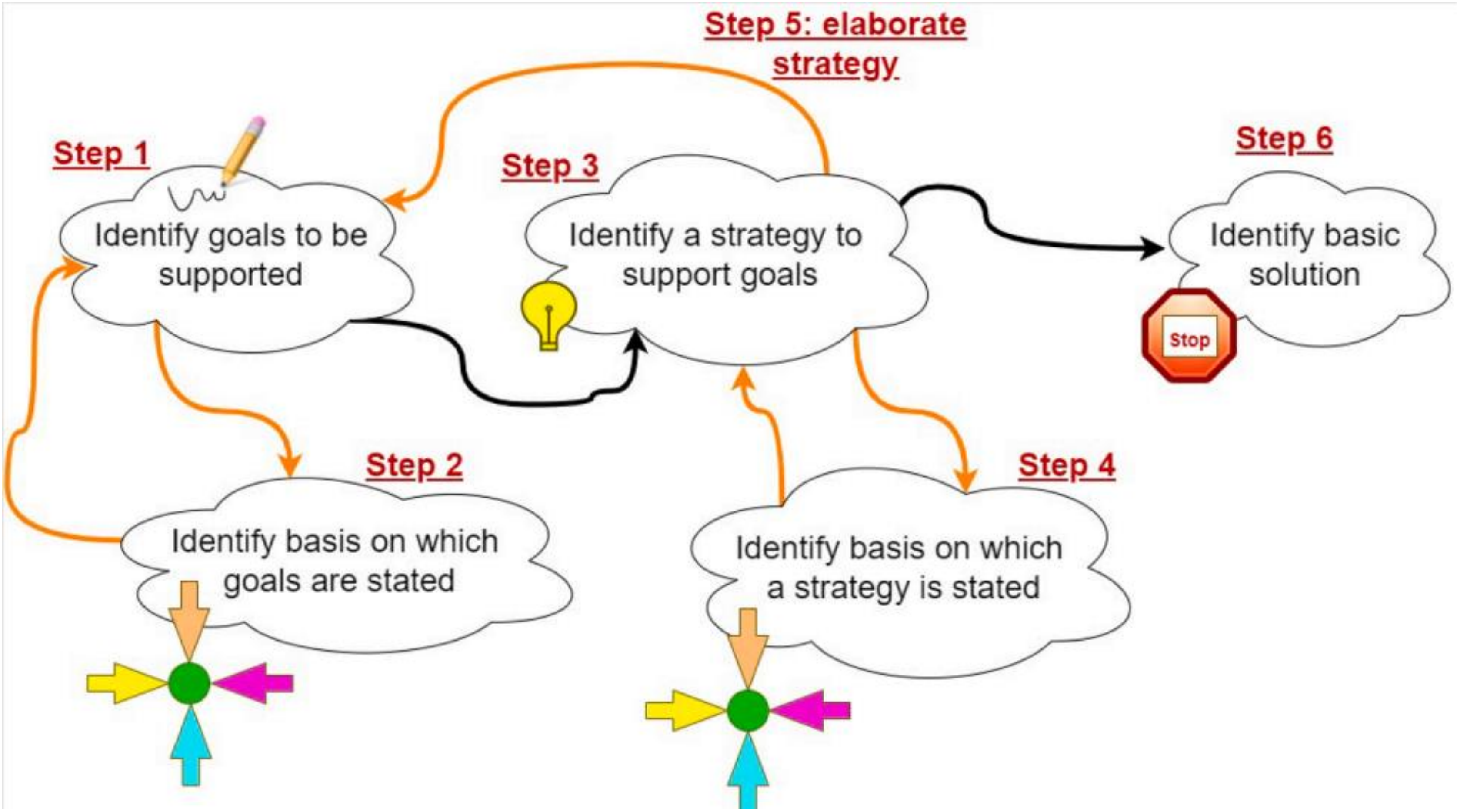
Proposed approach: Part I



Hazard elicitation and mitigation *(Part II)*
Classification of hazards and their mitigation techniques
Extension of GSN to support additional metamodel concrete syntax
New techniques to mitigate epistemic uncertainty
Reinforces
Higher dynamic safety assurance *(Part I)*
GSN extension to support dynamic safety-related concrete syntax
Extension of GSN 6-step design approach to DACs
Supports
Safety regulation compliance *(Part III)*
Analysis (survey) of barriers to safety regulation compliance
Safety assurance policy model

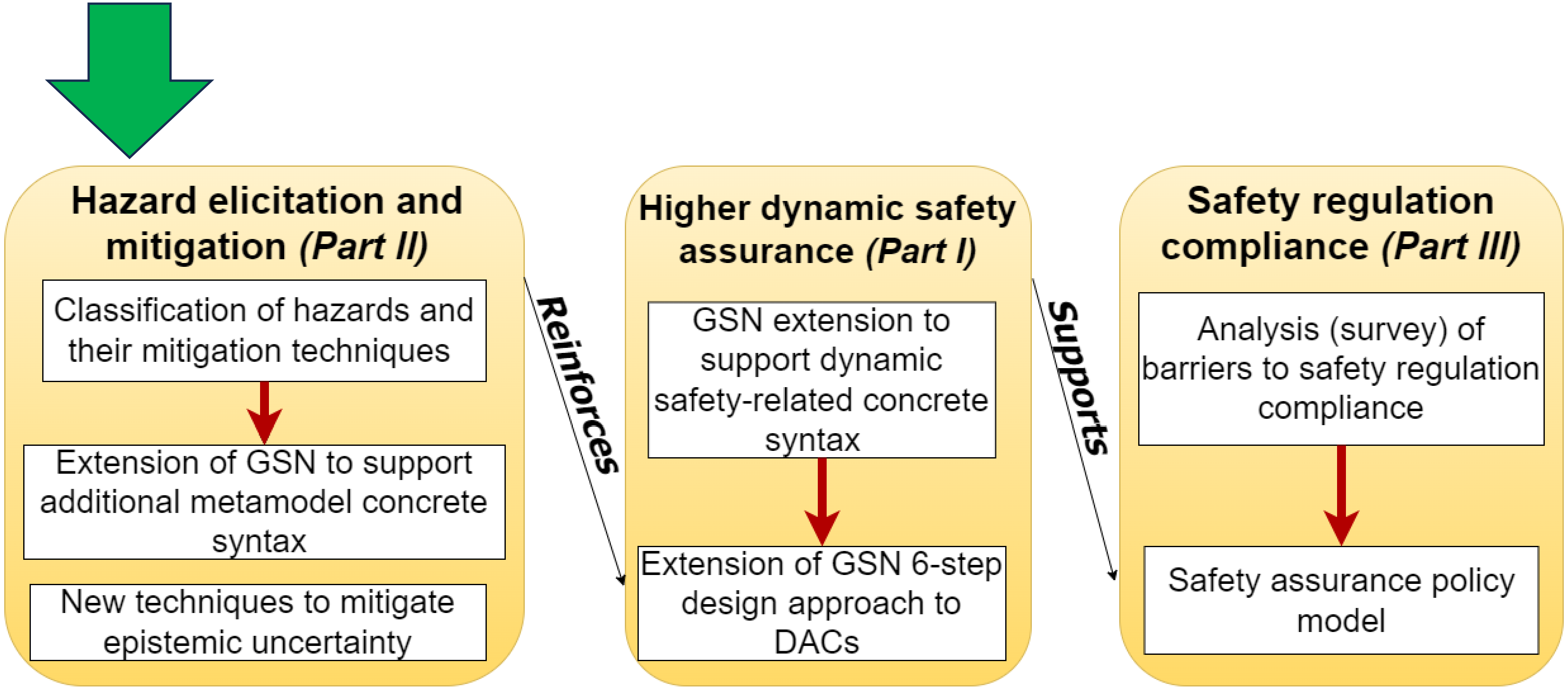
- Investigate if GSN needs to be extended to support the dynamic safety-related concrete syntax.

Proposed approach: Part I (continued)


Step 5: elaborate strategy
Step 1
Identify goals to be supported
Step 3
Identify a strategy to support goals
Step 6
Identify basic solution
Stop
Step 2
Identify basis on which goals are stated
Step 4
Identify basis on which a strategy is stated

- Explore the possibility to extend to Dynamic assurance cases (DACs) the widely used six-step approach that the GSN working group proposed to design static assurance cases.

Proposed approach: Part II



Hazard elicitation and mitigation *(Part II)*

Classification of hazards and their mitigation techniques

Extension of GSN to support additional metamodel concrete syntax

New techniques to mitigate epistemic uncertainty

Higher dynamic safety assurance *(Part I)*

GSN extension to support dynamic safety-related concrete syntax

Extension of GSN 6-step design approach to DACs

Safety regulation compliance *(Part III)*

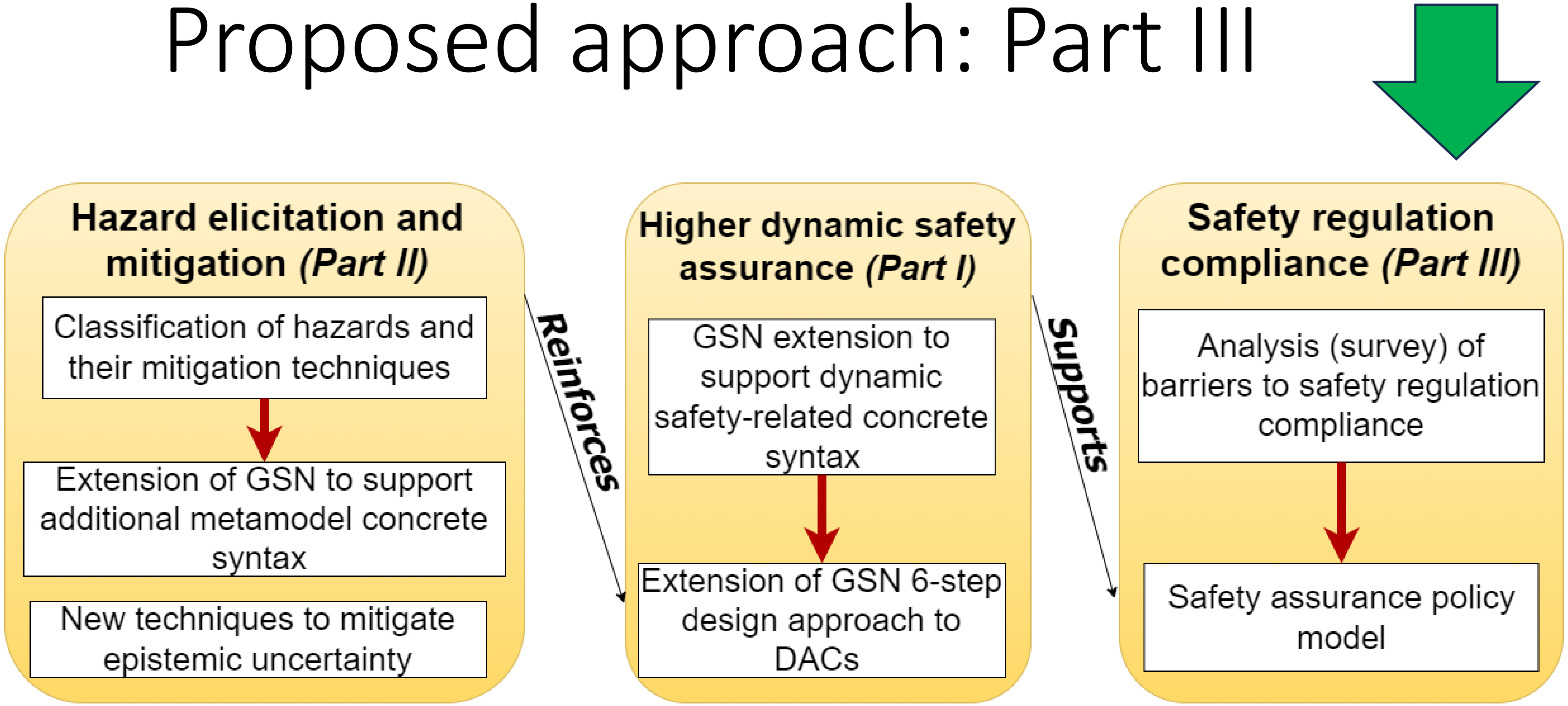
Analysis (survey) of barriers to safety regulation compliance

Safety assurance policy model

Rely on Machine Learning to elicit some of the unforeseen risks (uncertainty) an ADS may face at runtime:

- A DAC may then dynamically update its structure by reasoning away the elicited risks.

Proposed approach: Part III



Hazard elicitation and mitigation *(Part II)*
Classification of hazards and their mitigation techniques
Extension of GSN to support additional metamodel concrete syntax
New techniques to mitigate epistemic uncertainty
Reinforces
Higher dynamic safety assurance *(Part I)*
GSN extension to support dynamic safety-related concrete syntax
Extension of GSN 6-step design approach to DACs
Supports
Safety regulation compliance *(Part III)*
Analysis (survey) of barriers to safety regulation compliance
Safety assurance policy model

- The safety assurance policy model is a model-based representation of assurance policies serving as a basis against which the sufficiency of safety assurance can be established by ADSs manufacturers.
- To better support safety regulation compliance we could rely on such models to make suggestions to improve existing regulations.

Conclusion and future work

- Hazards caused by autonomous vehicles operated by ADSs are sometimes fatal
  - This is likely to lead to corporate failure of manufacturers of these vehicles.
- We therefore propose a novel approach that aims at supporting the dynamic safety assurance of ML-enabled ADSs.
- Our approach has the potential to:
  - Create new knowledge and innovative technology to mitigate edge cases at runtime
  - Support more efficiently the dynamic safety assurance of ADSs
  - **Reduce the mortality rate by yielding safer ADSs ☺.**

Q & A

The proposed approach is still at the proposal phase.

I am therefore **CRAVING** for your suggestions to improve my work.

So, do you have any?

References

1. Mansourov, N., and Campara, D. (2010). System assurance. Elsevier.
2. GSN (Goal Structuring Notation): https://scsc.uk/gsn? [Accessed in September 2023]
3. Vierhauser et al. (2019). Interlocking safety cases for unmanned autonomous systems in shared airspaces. IEEE TSE, 47(5), 899-918.
4. Asaadi, E., Denney, E., Menzies, J., Pai, G. J., & Petroff, D. (2020). Dynamic assurance cases: a pathway to trusted autonomy. Computer, 53(12), 35-46.
5. Ashmore et al. (2021). Assuring the machine learning life cycle. CSUR, 54(5), 1-39.
6. Chechik, M., Salay, R., Viger, T., Kokaly, S., & Rahimi, M. (2019, April). Software assurance in an uncertain world. In FASE (pp. 3-21). Springer.
7. Hawkins et al. Guidance on the assurance of machine learning in autonomous systems (AMLAS). arXiv:2102.01564, 2021.
8. Denney, E., Pai, G., & Habli, I. (2015, May). Dynamic safety cases for through-life safety assurance. In 2015 IEEE/ACM 37th IEEE International Conference on Software Engineering (Vol. 2, pp. 587-590). IEEE.
9. Guerin, J., Delmas, K., & Guiochet, J. (2022, May). Evaluation of runtime monitoring for UAV emergency landing. In 2022 International Conference on Robotics and Automation (ICRA) (pp. 9703-9709). IEEE
10. CRUISE safety report from 2022