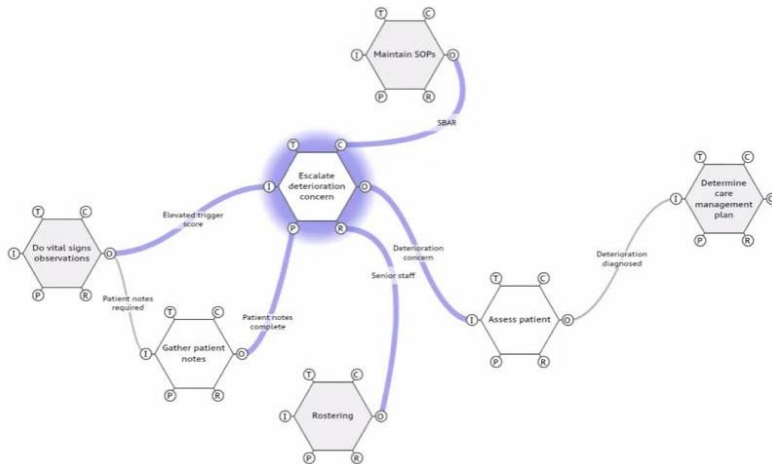


New Safety Analysis Techniques Seminar Report



Richard Trebilcock reports on the SCSC Seminar on New Safety Analysis Techniques which, due to the COVID19 restrictions, was held virtually using Zoom. Richard summarises the four presentations discussing a range of new safety analysis techniques and their application.

This was an afternoon seminar hosted by Mike Parsons with approximately 20 delegates attending from a wide range of industries including Aviation, Defence, Healthcare, Rail and Space. Mike opened the seminar with an overview of the afternoon's proceedings and timetable for the Seminar presentations.

Managing the deteriorating patient – A FRAM analysis

Mark Sujan from Human Factors Everywhere Ltd., discussed the "Management of the Deterioration of Healthcare". In particular, the mis-management of the deterioration of cases, i.e. where patients are admitted into hospital, and their deterioration is not picked up by the medical staff. Retrospectively, it can be seen that there are obvious signals that patient deterioration is taking place but it has gone undetected or not acted upon.

National Early Warning Scores (NEWS) are used to indicate the requirement for early action, standardising the assessment of acute-illness severity in the NHS. These scores, based on an ascending scale in the previous 24 hours, trigger the initiation of a range of clinical responses. However, patients are still dying, and the early warning scores have not delivered as hoped for. Some hospitals are better at managing patient deterioration than others and there is a need to move from failures to interventions. This highlighted the need for a simple support tool, which is where the research started.

The approach was to turn the problem on its head, i.e. rather than look at why some hospitals fail, look at why some hospitals succeed. Also, there was a need to look at Safety II and create an environment where resilient behaviours are created.

Knowing what to look for, in particular, looking at how to improve Resilience Abilities.

The Functional Resonance Analysis Method (FRAM) prompts the analyst to consider a range of factors impacting on an Activity or Function. FRAM is an analysis method, and is not a model. The purpose of FRAM is for an analysis to make sense of an Activity or Function. The analyst builds a framework. Output of the analysis is a list of tables, which help to identify:

- What is the variability?
- What are the reasons for variability?
- Finally, linking variabilities to resilience abilities

Improvement in the pre-alerting for a patient should be achieved through better understanding of what is going on, on an everyday basis, and ensuring resilience is built into the systems and processes.

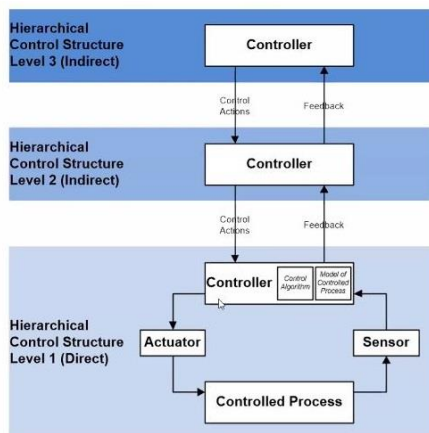
Overview of STAMP & STPA Hazard Analysis

Simon Whiteley from Whiteley Aerospace Safety Engineering & Management Limited, provided an overview of the Systems Theoretic Accident Model Processes (STAMP) & System-Theoretic Process Analysis (STPA) hazard analysis.

Simon started by introducing STAMP, which is an 'Accident Causality Model' based on Systems Thinking, Systems Theory and Control Theory Concepts.

STAMP-based processes comprise four model-based analysis processes:

- Systems Theoretic Early Concept Analysis (STECA);
- Systems Theoretic Process Analysis (STPA);
- Systems Theoretic Process Analysis-Security (STPA-Sec);
- Causal Analysis based on Systems Theory (CAST).



Simon spent time explaining where STAMP-based processes fit in the system lifecycle. He suggested that thinking in terms of control structures helps to organise analyst thinking and understanding. In addition, it provides a natural path to follow and a structure to work to, this is a key reason why STAMP-based assessments are so powerful. Points raised when discussing STAMP:

- Control Loops and control structures, the distance from control to actions.
- It was pointed out that we often have a constrained definition of hazard, which is reflected by our mental model of a process, and subsequently the hazard. We do not always take account of control and process lag times.

STPA is a relatively new hazard analysis technique, which builds on STAMP, a process and accident model using concepts of system and control theory. Moreover, it will give the safety analysts different views on the STPA hazard analysis process. The Ultimate Goal for STPA is to:

- Establish:
 - what the **Control Structure** is intended/required to do
 - how and why it allows, or contributes to, inadequate control, not just failures or “what went **wrong**” but “what went **right**”
- Create and Develop:
 - the changes necessary to **enable** adequate control
 - Safety Requirements / Constraints

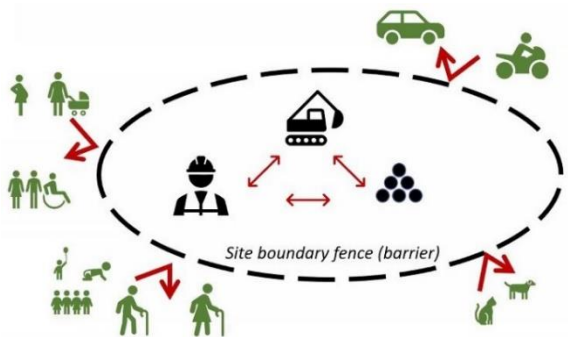
A process overview of STPA is as follows:

- **Foundation Phase:** Considering Losses, Hazard(s) and High Level Safety Constraints (HLSC)
- **Modelling Phase:** to Model Control Structures
- **Analysis Phase:** to identify Unsafe Control Actions, Causal Scenarios & define Safety Constraints
- **Development & Implementation**

When identifying causal scenarios it is necessary to identify the right level of abstraction for the control model. This is an acquired development skill, and for the novice there is the need for a coach.

Environmental Survey Hazard Analysis (ESHA)

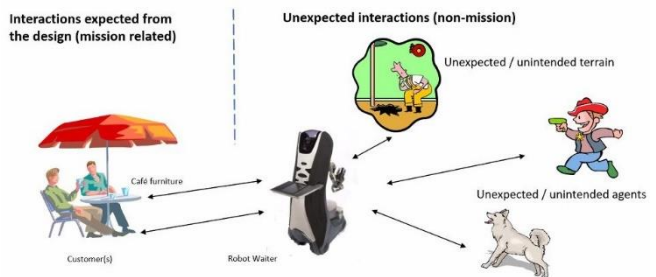
Chris Harper from Bristol Robotics Laboratory, University of the West of England, discussed robotics and how robotic Autonomous Systems (ASs) interact with their environment. He considered the problems presented by bounded and unbounded domains. The bounded domain reduces the number of interactions required by the system. For example, a barrier allows you to cut down the interactions e.g. site boundary fence blocks deflect interactions, reduces the complexity of the problem.



In the field of robotics, with Autonomous Systems moving towards interactions in a much more open domains alongside humans, where we can no longer assume the presence of barriers like steel cages, there are many more features to interact with. Many of the interactions are not eliminated by design. As such, more complex behaviour is required, with more advanced system technology required to achieve intelligent behaviour (i.e. Machine Learning/Artificial Intelligence). As a consequence, it is very much harder to show that a safety analysis is complete.

An Autonomous System (AS) needs to be able to perform Mission and Non-mission tasks/interactions:

- AS must survive in general (without human intervention) as well as perform their mission
- Hence, AS must perform “non-mission” tasks and interactions, as well as mission tasks
- Non-mission tasks are often not covered well by traditional system engineering methods.



In the context of ASs, ESHA Guide Words help designers to classify features of the environment. However, ESHA and Safety Arguments presented problems:

- Conventional systems offer service provision while ASs have to accommodate situation behaviour
- Typical arguments is “acceptable residual risk” which is difficult to define in ASs
- The problem of unbounded domains amplifies existing Safety-Critical Systems methodological problems
 - Expansion of states arising from an increase of AS scope to interactive processes causes ‘combination explosion’
 - Coverage metrics become (even more) questionable
 - What is residual probability of failure if the domain is unbounded / infinite?

There is a need to look for alternative types/styles of argument; a So Far as is Reasonably Practicable (SFAIRP) argument may be more feasible.

There was a requirement for formal classification schemes for environmental features, i.e. an Ontology. These tend to be organised into 2-3 layers:

- **Foundational Ontologies:** ‘Universal’ Abstract Types e.g. continuant, occurrent, physical, abstract
- **Domain Ontologies:** Extended abstract type and higher-level types for domains e.g. system safety, autonomy, environmental domains (e.g. domestic, roadside)
- **Application Ontologies:** Terminology related to a particular application domain e.g. driverless vehicles, healthcare assistive robotics

The Unified Foundational Ontology (UFO) has been selected as the base formalism. The UFO has been modified to support ESHA. While there is still a large programme of work ahead, it was proposed that ESHA may, in the future, be able to support a Safety Case or safety argument.

Dialectic Argumentation

Yvonne Oakshott from Leonardo, presented on behalf of the Assurance Case Working Group (ACWG) and discussed the current interim status of the Dialectic Argument activities. Yvonne covered the following topics:

What is a Dialectic Argument?

Dialectic Argument provides a framework for creating, challenging and questioning assurance cases. A Dialectic Argument is more than just the consideration of counter-evidence. A 'Dialectic Model' is based on a set of moves or responses to each part of the argument.

Who would use a Dialectic Argument?

- Assurance Case Practitioners /Authors
- Reviewers e.g. ISA, Customers, Operators, Regulators
- Various domains e.g. air, sea, land, rail, medical

Why use a Dialectic Argument?

- Assurance arguments, for example Safety Cases:
 - Are by their nature subjective and ultimately require stakeholder acceptance of the subjective position
 - Common Criticisms:
 - May be defeated by facts that destroy the argument
 - Logical fallacies – deceptive appearance of being a good argument
 - Insufficient premise in support of the conclusion
 - Even if the evidence appears to well support the argument, the evidence may not be trustworthy
- Over emphasis on the high level claims can lead to bias and insufficient focus on the wider picture
- Explicit Dialectic argumentation can redress the balance but is little practised

Barriers to use

A number of issues have prevented the adoption of a wider approach to challenging arguments such as Awareness: many practitioners are unaware of dialectic argument and Acceptance: dialectic argument is not routinely used.

The Principles of Dialectic Argument

Challenges can be applied at any level with the assurance case and may defeat parts of the assurance case:

- **Rebuttal:** Evidencing an argument which results in a counter to the conclusions of the assurance case (i.e. refutes, disproves, shows to be false)
- **Undercutting:** The introduction of additional facts (evidence) which challenge the reasoning within the assurance case
- **Undermining:** Defined in some existing work specially raising doubts about the validity of evidence.

Wigmore charts for dialectic arguments are used in law, so there is precedence for their use.

Further developments from SCSC ACWG

Unfortunately, Goal Structuring Notation (GSN) does not provide an accepted way to represent a dialectic discourse. However dialectic arguments can be represented in Structured Assurance Case Meta-model (SACM). The intention is to include dialectic support in version 3 of the GSN community standard. The Objective of the ACWG is to raise the profile of dialectic argumentation.