On Taming Complexity in System Evolution

Dr. Andrew Hussey

Andrew.Hussey@hitachirail.com

Maria Hill

Maria.Hill@hitachirail.com

Abstract

A common feature of Complex Systems is a high degree of interaction between components. This gives rise to features of the system as a whole that are difficult to predict from the properties of the constituent parts. A hallmark of such systems is non-linearity (i.e., causes that have small impact at the component level may produce large system effects). Complex Systems present specific challenges when also Safety-related e.g. causes of hazards often involve multiple failures and there may be reliance on people to keep the system safe (which in itself can introduce new hazards and risks). Such failures may involve the occurrence of events that were not anticipated by the designer of the System, or that were anticipated at a lower rate of occurrence. Algorithms and heuristics for analysis of such Complex Systems exist but the expertise for how and when to apply them is still evolving. We examine an Industrial Case Study involving a Railway Safety-related System, reflecting on which way that System can be considered as "Complex". The Case Study is an update and change of an existing system, which made it difficult to establish the boundary of the system under analysis. We consider techniques applied to analyse that System, such as FFA, FTA and HAZOP as well as measures taken to reduce risk, including conservative bias, layers of protection and reactive fail-safety. We consider the benefits and limitations that were identified and suggest some possible improvements to our approach for the future.

Keywords: Complex Systems, Industrial case study, Railway, Hazard Analysis

1 Introduction

Classical Hazard Analysis is geared towards analysis of greenfields systems but lacks support for analysis of existing "brownfield" systems, especially where the features of interest involve a combination of new added functions, modified function and/or existing functions.

In the classical Hazard Analysis approach, functions of the System under analysis are analysed using a chosen set of guide words. This basic approach is applicable across many techniques, that differ mainly in the identification of the item under analysis (FFA, FMEA, FMECA, HAZOP, etc).

Some of the problems with this approach are the limitation of the scope of the analysis to the model under consideration, limitation of the scope of the hazard identification to the keywords selected and the tendency to focus on the identified "delta" only, in the case of brownfields systems.

This paper considers the suitability of these classical approaches to Complex Systems, especially those where the system is a modification of an existing already operating system (so called "brownfield").

The key contributions and additions to current learning discussed in this paper are as follows:

1. Complexity may arise when an existing system is extended with new function;
2. This complexity is difficult to manage via conventional Hazard Analysis techniques alone, which rely on functional failures;
3. Analysis is necessary based on matching of hazardous properties with the output conditions for system requirements.

2 Acronyms, Abbreviations and Definitions

2.1 Acronyms and Abbreviations

Acronym	Description
FFA	Functional Failure Analysis
FMEA	Failure Modes and Effects Analysis
FMECA	Failure Mode Effects and Criticality Analysis
FTA	Fault Tree Analysis
HAZOP	Hazard and Operability Analysis
ONRSR	Office of the National Rail Safety Regulator
RISSB	Rail Industry Safety and Standards Board
RSNL	Rail Safety National Law
SFAIRP	So Far As Is Reasonably Practicable
SIL	Safety Integrity Level
SRAC	Safety Related Application Condition
STS	Hitachi Rail STS
THR	Tolerable Hazard Rate

2.2 Definitions

Barrier to Escalation

A Mitigation that limits only the consequences of the corresponding Hazard occurring e.g. a protective wall that prevents a chemical escaping from an area of a plant.

Cause

A Safety Risk is considered to be a Cause of a System Hazard if it appears as a node in the corresponding Fault Tree for that System Hazard.

Control

A Mitigation that limits the likelihood of the corresponding Hazard occurring e.g. a device that checks the temperature of a boiler and raises an alarm if it exceeds permitted limits.

Functional Failure Analysis

A Safety analysis conducted by examining Functions of a System via keyword failure prompts (such as "Too much") to determine potential hazardous failure modes of those functions.

Mitigation

Control or Barrier to Escalation.

Requirement

(1) A condition or capability needed by a user to solve a problem or achieve an objective.

(2) A condition or capability that must be met or possessed by a product, service, or product component to satisfy a supplier agreement, standard, specification, or other formally imposed documents.

(3) A documented representation of a condition or capability as in (1) or (2).

Refer to IEEE Standard Glossary of Software Engineering Terminology [IEEE90].

System

The abstraction level related to the Scope of Work under analysis. The System is composed by a set of Subsystems and Interfaces.

System Hazard

A state of a System with the potential for loss of life or injury.

Subsystem

The lower abstraction level, in relation to the Scope of Work under analysis. The Subsystem is a component of the main System.

3 Literature Survey

In his seminal article on the topic of Complex System failure "How Complex Systems Fail", first published in 1998 and still widely referenced today, Dr. Richard I. Cook [Cook98] identifies and discusses 18 core elements of failure in Complex Systems. These focus on the existence of multiple failures giving rise to behaviours that are difficult to anticipate from the behaviour of the parts and the reliance on people to keep a system safe.

The notion that new behaviours can emerge as we combine and extend systems is elaborated further in other more recent work.

According to Scholarpedia [Scholarpedia07] Complexity is "The emergence of traits encompassing the system as whole that can in no way be reduced to the properties of the constituent parts. The intertwining, within the same phenomenon, of large scale regularities and seemingly erratic evolutionary trends. This coexistence of order and disorder raises the issue of predictability of the future evolution of the system at hand on the basis of the record available."

The Waterloo Institute for Complexity [Waterloo19] and Innovation defines Complexity as behaviour arising from the interplay, in densely interconnected systems, between multiplicative causation and positive and negative feedbacks. A signature of such systems is radically disproportional causation (i.e., small causes do not always produce small effects) or what is often called "nonlinearity." Nonlinear systems can undergo sudden flips between stable states or equilibria. A second signature is the "emergence" of structured macroscopic patterns that are the outcome of the independent microscopic interactions of the entities in the system. These macroscopic patterns — be they hurricanes in Earth's atmosphere or boom-bust cycles in global financial markets — often have enormous causal power.

Mathematically, complex adaptive systems are multi-state variable dynamical systems characterized by a moderate degree of structured interactions and interconnections. State variables in these systems are often characterized by heterogeneous parameter sets and updating rules. Spatial and network relationships are often non-uniform and violate mean field theory assumptions.

According to Heslin [Heslin15], the hallmarks of so-called Complex Systems are "a large number of interacting components, emergent properties difficult to anticipate from the knowledge of single components, adaptability to absorb random disruptions, and highly vulnerable to widespread failure under adverse conditions [Dueñas09]."

Similarly, Moses [Moses00] defines Complex Systems as "A system is complex when it is composed of many parts that interconnect in intricate ways." According to this definition, the number and types of parts of a system decide its complexity. The systems become complex when these parts are connected in a nonregular way. The information about the nature of interconnections is usually insufficient and this makes modelling difficult.

Likewise, Rechtin and Maier [Rechtin10] define a Complex System to consist of a set of different elements so connected or related as to perform a unique function not performable by the elements alone. The Complex System acts as "whole is greater than parts," which means a complete knowledge about the parts will not be sufficient to predict the working of a system as a whole.

A common feature of Complex Systems therefore seems to be a high degree of interaction between components so that modelling the System is difficult and causes that have small impact at the component level may produce large system effects). Complex Systems are often Safety-related and causes of failures often involve multiple failures and the reliance on people to keep the system safe. Such failures may involve the occurrence of events that were not anticipated by the designer of the System, or that were anticipated at a lower rate of occurrence.

Fraccascia et al [Fraccascia18] state that a common property of many Complex Systems is resilience, that is, the ability of the system to react to perturbations, internal failures, and environmental events by absorbing the disturbance and/or reorganizing to maintain its functions. Resilience engineering is the ability of a system to sense, recognize, adapt, and absorb variations, changes, disturbances, disruptions, and surprises. The ability to bounce back when hit with unexpected events. The joint ability of a system to resist (prevent and withstand) any possible hazards, absorb the initial damage, and recover to normal operation. However, while Complex Systems in Nature are often resilient (and adaptive), software-based Complex Systems are often not. This is because resilience needs to be built into the software design, and this can only be done if the applicable risks have been identified and mitigated appropriately.

Referring to techniques for modelling and engineering Complex Systems in 2005, Wiles and Watson [Wiles05] argued that the field of Complex Systems currently lacks a library of experience. Shared libraries of algorithms and heuristics exist but without the experience of the Complex Systems modeller in when to apply them.

The famous saying about models is that they are all wrong, but some are useful [Box78], and as such, models play an important role in understanding and taming Complex Systems [Ma'ayan17]. From these models, insightful theoretical rules can be extracted. Nonetheless, the model used as the basis for analysis places inherent limitations on the resolution and applicability of the analysis.

"As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality." – Albert Einstein.

4 Case Study

For reasons of confidentiality, details of the system under analysis cannot be shared. However, from an abstract perspective, the system has the following properties:

- Manages part of the supervision and control functions for locomotives
- Is an extension of an existing system and integrates with that existing system
- Controls must be issued to all connected locomotives in a synchronous manner to avoid potential safety hazards
- The locus of control shifts during the course of a train journey, from one locomotive to another, adding to the possibility for asynchronous behaviours
- The system includes the locomotive-locomotive interface, and new functions added on each locomotive
- Existing isolated locomotive functions can have new consequences when considered in the context of multiple locomotives
- Traditionally fail-safe behaviours, such as Emergency Brake, are not necessarily safe in the context of multiple communicating and synchronising locomotives

Broadly speaking, the system involves two trains A and B that have a wireless connection between them, and that each have their own locomotives, collectively responsible for powering the train. Physically, train A is at the lead of the combined (A+B) train, with train B at the rear. There is a master-slave relationship between the trains but the train B is also able to separate from train A and act independently, this is safe when the conditions for entry to that mode are satisfied, but unsafe otherwise. This leads to an interesting property of the system, which is that there is not a strict fail-safe state for train B when entering or acting in the independent mode, since both failing to stop train B as well as stopping train B

can lead to hazardous consequences (derailment of the train). The fail-safe behaviour of the combined train must be guaranteed as a whole, via a safe stop of the entire train (A + B).

Is the case study System a Complex System? We noted the following properties that are aligned with the definition of Complex System given in Section 3:

1. Difficult to model interactions/interconnections of behaviour, because impact of existing function may not be sufficiently obvious or understood – in this case, to understand the causes of asynchronous controls, it was necessary to understand all the ways in which the existing function could result in unsynchronised stopping of a locomotive, when the train should be operating in synchronous mode.
2. Train drivers are an integral part of the overall function of the System, and perform some of the actions required to assure safe locomotive operation – this is by design.
3. It was not clear from a functional analysis that all causes of risks could be properly identified and an analysis of the system architecture was deemed essential.

5 SFAIRP

In this paper, we reference strongly the principle of SFAIRP as a fundamental cornerstone of our strategy for analysing and mitigating risks arising from complexity.

Under section 46 of the RSNL [RSNL12], Duty Holders are required:

a) to eliminate risks to safety so far as is reasonably practicable; and
b) if it is not reasonably practicable to eliminate risks to Safety, to minimise those risks so far as is reasonably practicable.

The above duties are referred to in this work instruction as the duties to 'ensure Safety SFAIRP'. The persons identified by the RSNL as Duty Holders have a duty of care to ensure Safety SFAIRP.

The concept of SFAIRP is to achieve the best possible Safety outcomes, to the extent that is 'Reasonably Practicable'.

In this context, and under the RSNL (s47), reasonably practicable means that which is, or was at a particular time, reasonably able to be done to ensure safety, taking into account and weighing up all relevant matters including:

a) the likelihood of the hazard or the risk concerned occurring; and
b) the degree of harm that might result from the hazard or the risk; and
c) what the person concerned knows, or ought reasonably to know, about the hazard or risk, and ways of eliminating or minimising the risk; and
d) the availability and suitability of ways to eliminate or minimise the risk; and
e) after assessing the extent of the risk and the available ways of eliminating or minimising the risk, the cost associated with available ways of eliminating or minimising the risk, including whether the cost is grossly disproportionate to the risk.

The ONRSR has published a guideline with respect to understanding the SFAIRP principle [ONRSR16b] – the remainder of this Section summarises key observations from that guideline, which is also closely aligned with the Common Law relating to Duty of Care.

What is reasonably practicable is determined objectively. This means that a Duty Holder must meet the standard of behaviour expected of a reasonable person in the duty holder's position and who is required to comply with the same duty.

There are two elements to what is reasonably practicable. A Duty Holder must first consider what can be done - that is, what is possible in the circumstances for ensuring safety. The Duty Holder must then consider whether it is reasonable, in the circumstances to do all that is possible.

This means that what can be done should be done unless it is reasonable in the circumstances for the Duty Holder to do something less.

Mitigations shall be selected for a particular Safety Risk (i.e. Hazard Cause) based on:

- the objective severity of the risk, in terms of likelihood and consequence;
- the objective reasonably known ways to mitigate that risk;
- the availability/suitability of the objectively known options for reducing risk;

The question of what is reasonably practicable is to be determined objectively, and not by reference to the Duty Holder's capacity to pay or other particular circumstances. A Duty Holder cannot expose people to a lower level of protection simply because it is in a lesser financial position than another Duty Holder.

If a particular Duty Holder cannot afford to implement a reasonably practicable risk control, the Duty Holder should not engage in the activity that gives rise to that hazard or risk.

5.1 Risk Acceptance Strategy

A dual strategy may be used, involving both quantitative and qualitative targets.

Safety risk assessment is provided using the Operator's Risk Assessment matrix/criteria and requires combining the frequency of the occurrence of a hazardous event with the severity of the consequence to establish the level of risk generated by a hazardous event.

Risk associated with hazards is reduced So Far As Is Reasonably Practicable (SFAIRP), taking account of all factors affecting reasonable practicability of mitigations, including level and nature of the risk, technical difficulty and resulting acceptability.

Risks in the first instance may be eliminated where feasible in terms of system scope. Removing a function from consideration means the risk associated with performance of that function is no longer applicable in terms of the System analysis. When elimination is not possible, and where reduction of risk is reasonably practicable, risks may be dealt with through technical mitigations. Where technical mitigation is not possible, non-technical mitigations such as procedures may be introduced provided risk reduction is reasonable. The Operator's acceptability criteria are applied to the residual risk classes.

6 Integrated Customer-oriented Process

In accordance with EN50129, it is the responsibility of the Railway Authority, to outline and document the system (independent of technical realisation), to identify the top-level hazards relevant to the system, to analyse the consequences, to define the risk tolerability criteria, to derive the tolerable hazard rates (i.e. safety targets) for top-level hazards, and to ensure that the resulting risk meets the risk tolerability criteria.

The Railway Authority may be the Operator, or a parent organisation of the Operator. The Railway Authority's Railway Hazard and RAM Risk Log is used to manage top-level railway hazards as well as required procedural/operational mitigations and any other external mitigations that are not part of the System being developed by the Supplier.

The initial safety targets are defined in terms of the following:

1. Safety targets (Tolerable Hazard Rates) for individual top-level railway hazards and risks and/or cumulative risks that have been derived by the Railway Authority from the hazard identification and risk assessment

based on the Railway Authority's criteria for acceptable risks.

2. System-level hazards that have been derived by the Supplier via bottom-up Hazard Analysis activities including Functional Failure Analysis and HAZOP. The functional hazards have been mapped to: (a) corresponding top-level railway hazards to determine whether the System is introducing any new hazards; (b) corresponding System Railway Events to enable assignment of SIL and THR.
3. Safety requirements and corresponding SILs (where applicable) are derived by the Supplier for the System to mitigate the system-level hazards.
4. Safety targets e.g. Safety Integrity Level (SIL) and corresponding Tolerable Hazard Rates (THR) for safety-related subsystems are allocated by conducting relevant analyses to derive subsystem safety requirements as a part of the Safety Apportionment, e.g. Fault Tree Analysis.

The Safety Targets for the System are provided to the Supplier by the Railway Authority. These Safety Targets may be derived by the Railway Authority via a top-down Fault Tree-based analysis, starting from Railway Hazards (these are either already known, or derived as a result of bottom-up feedback from the Supplier Hazard Analysis). For example:

1. "Train collides with another rail vehicle on the mainline"
2. "Train derails on the mainline"

The format of the Fault Trees is as shown in Figure 6-1 for the basic 'AND' and 'OR' gates as well as 'TRANSFER'.

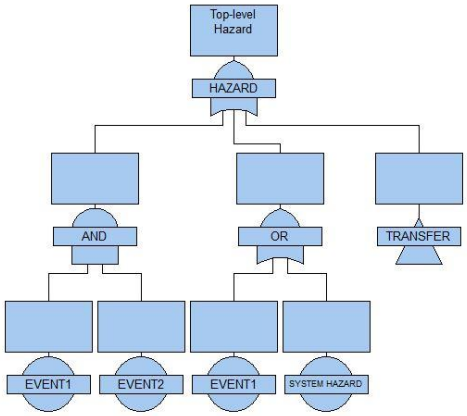

Figure 6-1: Fault Tree Format

As well as these basic symbols, a Fault Tree may combine any of the event and gate symbols given in Tables 6-1 and 6-2.

BASIC EVENT	Basic event for which failure and repair data is available.
CONDITIONAL	Similar to basic event but represents a conditional probability.
UNDEVELOPED	Represents a system event that is yet to be developed.
HOUSE EVENT	Represents definitely operating or definitely not operating (Boolean) events.
TRANSFER	Indicates that part of the fault tree is developed in a different diagram or on another page.

Table 6-1: Primary Event Types

AND	Result event occurs if all input events occur.
OR	Result event occurs if any one of the input events occurs.
COMBINATION	Result event occurs if 'n' of the input events occur.
EXCLUSIVE OR	Result event occurs if one and only one of the input events occurs.
PRIORITY AND	Result event occurs if all input events occur in sequential order from left to right.
INHIBIT	An inhibit gate is essentially an AND gate with an additional conditional event (typically an event external to the configuration represented by the fault tree).
NOT	Result event occurs if the input event does not occur.

Table 6-2: Gate Types

The interaction of the Operator and Supplier whereby the Supplier requires the initial top-down analysis of the System Hazards by the Operator, results in process efficiencies, as from a System Development perspective, the Supplier is the "customer" of the Operator, needing the System Hazards as the input to their System Development process. The resulting process seamlessly integrates the activities of the Supplier and Customer/Operator.

7 System Modelling Techniques

7.1 State transition diagrams with timing parameterisation

State transitions (see Figure 7-1) were used to define the scenarios relating to the operation of trains in synchronous and independent modes. The State Transition diagrams included not only System behaviours, but also external factors, as well as Driver, Operator and Maintainer actions.

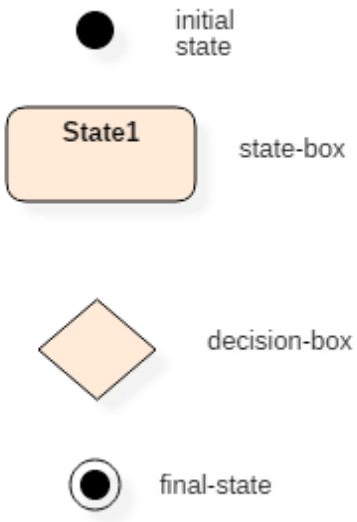

Figure 7-1: State Transition Symbols

An example State Transition diagram follows in Figure 7-2, relating to the case study under consideration.

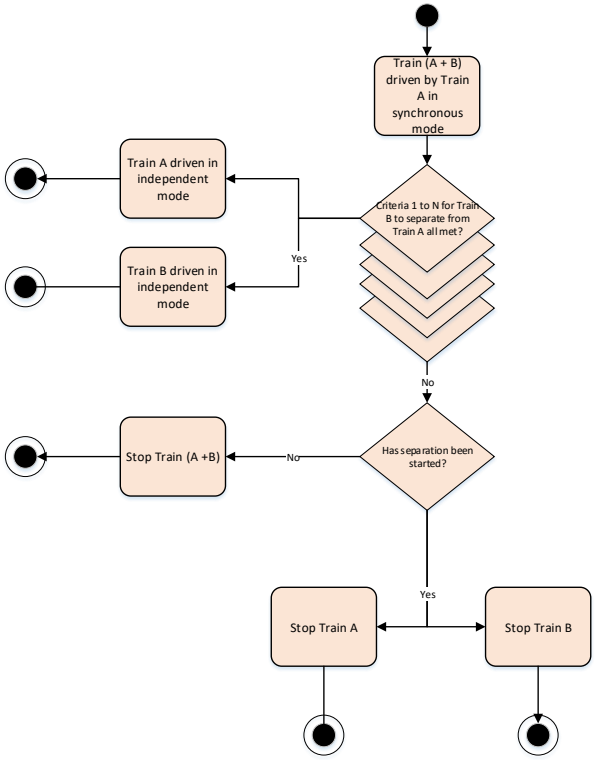

Figure 7-2: Simplified State Transition Diagram for the Case Study System

7.2 Functional requirements

Natural language requirements specification giving atomic predicates that collectively define the behaviour of the System. A weakness with natural language is that logical inconsistencies are relatively easy to introduce into the specification, and inconsistent use of terminology may obscure these logical inconsistencies, or make it difficult to see relationships between collaborating requirements. Some examples of natural language requirements, relating to the case study under consideration, follow:

Requirement Title : No Active Critical Faults

Requirement Text: A Train shall only be allowed to commence a journey when the Train has no active Critical Locomotive faults on start-up.

Requirement Title : Service brake on Train A

Requirement Text: The Train B in a train that is operating in Synchronous mode that detects a Service Brake application on Train A, shall set the Train B throttle to Notch one (1).

8 Analysis Techniques to detect Hazards

8.1 FFA and FTA

An FFA was conducted in a previous phase of the project, based on the over-arching functions of the System. Such functions include:

- driving the integrated train (train A + train B) when in synchronous mode;
- entering independent mode for train B
- driving train A and train B when in independent mode.

Using the outcomes from this FFA as an input, as well as knowledge from Subject Matter Experts, the FTA for the overall railway was generated by the customer, to determine the THR targets for the risks that would be managed by the System.

8.2 HAZOP

A brainstorming meeting was held to identify the preliminary elements and interfaces to be analysed, based on the State Transition Diagrams, preliminary Functional System Requirements and Architecture and knowledge of the current operations and systems.

This was then supplemented and augmented by meetings with the operator, and discussions within the STS Systems Engineering teams responsible for the System Requirements, the STS subsystems teams responsible for the various subsystem and the STS site and lab testing groups.

Based on these inputs the draft Hazard Analysis was created by deriving a list of elements and functions, generating the possible hazards for each element using the generic guidewords from IEC 61882.

Following further iterations of analysis, discussions and reviews, the Hazard Analysis was updated to reflect the finalised understanding of the system as documented in the finalised System Requirements.

The overall objective of the HAZOP was to derive the System and Subsystem requirements necessary for the proper reactive fail-safety of the System.

HAZOP assisted in identifying the hazardous interactions between components of the system, which is a characteristic feature of Complex Systems (refer to Section 3 and references Heslin [Heslin15], Moses [Moses00] and Rechtin and Maier [Rechtin10]). A key objective of such analysis is to identify the disproportional causation which is so characteristic of Complex Systems [Waterloo19].

In addition, a review was made of the existing Hazard Analysis and Safety Requirements to check for applicable functions, causes and mitigations relating to the hazard of asynchronous locomotive behaviour.

Some representative hazards from the analysis follow:

SH-1: Braking on Train B when in synchronous mode	The brake should never be applied by Train B when in synchronous mode (only throttle should be knocked back to notch 1).
SH-2: Brakes not applied on Train B when in independent mode and train A is braking	When operating in independent mode, if the brakes are applied on Train A then the service brake shall also be applied on Train B.

Table 8-1: System Hazard Definition

8.3 Review of existing documents

The existing Hazard Analysis and System Requirements were reviewed:

- Review of the System Hazard & RAM Risk Log to check applicability of existing Hazards, Causes, Controls and SRACs;
- Analysis of the updates to the System Requirements, as well as analysis to check for emergent properties of the System arising from the change in function;
- Analysis of the updates to the interface communications between the different components of the System, relating to the added System function.

9 Mitigating Hazards

Considering that outcomes of system behaviours might not always be fully understood or anticipated, a conservative bias was adopted, assuming the case of a hazard where there was doubt. To avoid the possibility that designed protections might fail, leaving the Train driver exposed to an unsafe behaviour, a layers of protection approach was used, adding mitigations where technically and commercially viable. This is also consistent with the approach advocated for SFAIRP. Further, to avoid possibility of a specific cause being missed, generic mitigations were sought wherever possible. The overall objective of the approach was to achieve

resilience of the system, in the face of perturbations (Fraccascia et al [Fraccascia18]).

Conservative bias:

The simplest solution to a safety risk was employed wherever possible, rather than to search for operationally optimal options. For example, a function for "Safe stop of the train" was devised (stop train A using Full Service Brake, notch 1 applied on train B until stopped then notch 0). This approach is aligned with SFAIRP.

Layers of protection:

LOPA is a risk assessment methodology which uses simplified, conservative rules to define risk as a function of both frequency and potential consequence severity. LOPA is defined as a simplified risk assessment of a one cause - one consequence pair [Willey14]. Conceptually, LOPA is used to understand how a process deviation can lead to a hazardous consequence if not interrupted by the successful operation of a safeguard called an independent protection layer (IPL). An IPL is a safeguard that can prevent a scenario from propagating to a consequence of concern without being adversely affected by either the initiating event or by the action (or inaction) of any other protection layer in the same scenario [Willey14]. In essence, LOPA is concerned with putting in place layers of protection based on Human Factors, Critical alarms, Automated responses, Physical protections and Emergency response. This approach is also similar to the well-known "Swiss Cheese" model propounded by James Reason (refer to Figure 9-1).

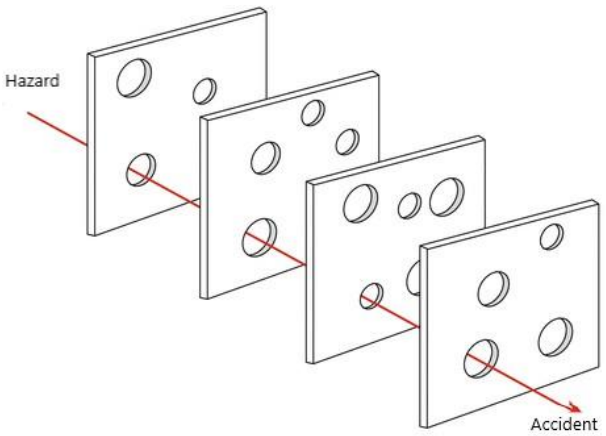

Figure 9-1: Reason's Swiss Cheese

Using an LOPA-oriented approach, the hazards SH-1 and SH-2 shown above in Table 8-1 can be mitigated by a combination of the following measures:

- Checks at start-up and alarms if communications cannot be established between Train A and Train B;
- Alarms to Drivers on Train A and Train B if there is a loss of communications between Train A and Train B;
- In the event of loss of communications, a Safe Stop of both Train A and Train B is initiated;
- Procedures for rail workers when working on track in the vicinity of such merged trains;
- Response procedures in the event of a failure and a stopped train.

As is often the case for functional railway hazards, Physical protections are much less effective than alarms and automated responses, because once a collision between trains has occurred, the consequences can be severe and especially due to the size and weight of freight trains, difficult to reduce.

The use of LOPA has many similarities to SFAIRP, where mitigations are added to address a particular Cause, so long as it is reasonably practicable to do so. The main advantage of applying LOPA was that it stimulated a more thorough search for all practicable mitigations to address system failures.

Reactive fail-safety:

A reactive fail-safety approach was taken to avoid the need to identify all causes of unsafe conditions. This approach was aligned with the Complex System property whereby behaviour is emergent and identification of causes is difficult due to pre-existing function and interactions between humans and the System. To support this reactive fail-safety architecture, the analysis was based on an architectural model of the System.

To avoid relying on identification of every cause, mitigations addressed the unsafe conditions of the System rather than the root causes. For example, "when the Full Service Brake is applied on train A, notch 1 is applied on train B until stopped, then notch 0".

10 Outcomes

10.1 What worked well

FFA and FTA (at the System level) was effective in eliciting an initial set of over-arching hazardous states. Using this initial set of hazards as a guide, the HAZOP could then be conducted in a more directed manner. The initial FFA and FTA guided us towards the risks associated with asynchronous behaviour of train A and train B when operating in Synchronous mode or when failures occurred in Independent mode (e.g. due to a failure to properly transition from Synchronous to Independent mode).

HAZOP was effective in eliciting causes and directing attention towards mitigations that addressed those causes. The HAZOP also gave a second look at the hazards, augmenting the definition of the hazards derived during the FFA. The single cause approach in HAZOP was not a significant problem because when combined with SFAIRP, which advocates a continuous process of mitigation where reasonably practicable, additional mitigations were added even when it could appear that a single item was acceptable according to a risk matrix approach. This approach also resulted in "layers of protection", with less focus on SIL and more focus on overall risk reduction and confidence.

10.2 What worked less well

The analysis is initially constrained by what can be discovered in the preliminary FFA/FTA. This can influence/bias the way in which the HAZOP is then developed. The FFA/FTA guided the analysis towards risks associated with asynchronous behaviours of trains A and B and led us to analyse primarily based on the communications between the onboard systems in train A and train B. This guidance towards a particular model of the System gave structure to the approach but also introduces potential biases. To combat this, we analysed in the HAZOP based on an architecture at the subsystem level, with the hope that this would enable us to discover other risks that could arise. However these emergent behaviours of the overall combined system could be more difficult to recognise when working within the framework of the FFA/FTA. Like most classical Hazard Analysis approaches, HAZOP is also based on use of keywords to guide analysis, and the selection of keywords, and/or their correct application, is critical to deriving a proper outcome from the analysis.

It is also difficult to discover all causes that are already built into the existing system function and operation. This was the major source of complexity that the existing methods did not support or resolve. Using the FFA and HAZOP to guide searches of the existing system function (Requirements and Architecture) also means the searches are potentially biased and could miss an emergent property that was not evident to the FFA/HAZOP analyst. It is not necessarily enough to consider only the updated or changed parts of the System function, but also potentially necessary to consider how existing unchanged function can have a different impact when failures occur in the new modified System.

The reactive fail-safety approach taken enabled us to avoid the need to discover all root causes of unsafe system conditions, but did

not address all availability aspects of the design. The resulting System could be safe, in the sense of reacting appropriately to an unsafe state, but with an unacceptable impact on availability.

11 Future ideas

One possible improvement on the approach taken could be to use Semi-Formal methods – a more formal definition of functional conditions would have enabled us to more easily search the existing system functions, for example, for constraints relating to "all locomotives". A more formal representation of the existing system functions would have been more consistent and therefore less likely that a requirement is overlooked due to a difference in natural language expression.

Expressing SH-1 and SH-2 as Boolean statements, we have:

SH-1: Braking on Train B when in synchronous mode	The brake should never be applied by Train B when in synchronous mode (only throttle should be knocked back). Synchronous → ┐ (SB_{TrainB} ∨ EB_{TrainB})
SH-2: Brakes not applied on Train B when in independent mode and train A is braking	When operating in independent mode, if the brakes are applied on Train A then the service brake shall also be applied on Train B. Independent ∧ (SB_{TrainA} ∨ EB_{TrainA}) → SB_{TrainB}

Table 11-1: Boolean statements

We should see then that an inspection of the System Requirements respects these conditions. The hazard SH-1 may occur as a result of behaviours relating to all locomotives irrespective of mode (or when mode is not specified). Hence any such requirements that can be reduced to the following equation would indicate an unmanaged cause of SH-1:

(trigger conditions) → (SB_{TrainB} ∨ EB_{TrainB})

Similarly for "Independent" mode, SH-2 may occur as a result of behaviours relating to all locomotives irrespective of mode (or when mode is not specified). Any such requirements that do not specify mode may apply when the trains are operating independently. In that case, requirements where the following equation holds would indicate an unmanaged cause of SH-2:

(SB_{TrainA} ∨ EB_{TrainA}) → ┐ SB_{TrainB}

While formalisation of requirements using Boolean algebra is not typically practical, due to limitations in the acceptance and readability of such requirements by the average System developer, use of a more formalised and structured natural language can give similar outcomes. By ensuring consistency of terminology, and mathematical equivalence of natural language statements, the searchability of the System Requirements, for the purpose of brownfields Hazard Analysis, could be much improved.

12 Conclusions

This paper has considered the suitability of classical Hazard Analysis approaches to analysis of a case study Complex System, and in particular where the system is a modification of an existing already operating system (so called "brownfield"). While the application of classical approaches gives benefits, this paper has confirmed some of the problems associated with classical Hazard Analysis approaches, identified in Section 1, such as the limitation of the scope of the analysis to the model under consideration and the risks identified early in the analysis process, limitation of the scope of the hazard identification to the keywords selected and the tendency to focus on the identified "delta" only, as against the original system.

The key contributions and additions to current learning discussed in this paper were:

1. Complexity may arise when an existing system is extended with new functionality;
2. This complexity cannot be managed via conventional Hazard Analysis techniques alone, which rely on functional failures;
3. Analysis is necessary based on matching of hazardous properties with the output conditions for system requirements.

HAZOP assisted in this process of identifying the hazardous interactions between components of the system, which is a characteristic feature of Complex Systems, and corresponding fail-safe reactions.

This paper has also considered the applicability of techniques based on Semi-Formal Methods to resolve some of the limitations of the classical Hazard Analysis approach for brownfields systems, and in particular, to support the activity of matching hazards with system requirements, so as to more effectively elicit causes and corresponding mitigations.

13 References

[Box78] G. E. Box, W. G. Hunter, J. S. Hunter. *Statistics for experimenters: an introduction to design, data analysis, and model building, vol. 1*. New York, NY: Wiley, 1978

[Cook98] R.I. Cook, How Complex Systems Fail (Being a Short Treatise on the Nature of Failure; How Failure is Evaluated; How Failure is Attributed to Proximate Cause; and the Resulting New Understanding of Patient Safety). Chicago, IL: Cognitive Technologies Laboratory, University of Chicago. Copyright 1998, 1999, 2000 by R.I. Cook, MD, for CtL. Revision D (00.04.21), http://web.mit.edu/2.75/resources/rando/How%20Complex%20Systems%20Fail.pdf

[Duenas09] L. Dueñas-Osorio and S. M. Vemuru, Cascading failures in complex infrastructure systems, In *Structural Safety*, Volume 31, Issue 2, March 2009

[Fraccascia18] L. Fraccascia , I. Giannoccaro ,V. Albino, Resilience of Complex Systems: State of the Art and Directions for Future Research, In *Complexity: Vol 2018*, Wiley, 2018

[Heslin15] K. Heslin, Examining and Learning from Complex Systems Failures. https://journal.uptimeinstitute.com/examining-and-learning-from-complex-systems-failures

[Ladyman13] J. Ladyman, J. Lamber, K. Wiesner, What is a Complex System?, In *European Journal for Philosophy of Science*, 2013

[Ma'ayan17] Ma'ayan A. Complex systems biology. Journal of Royal Society Interface 14, 2017

[Moses00] J. Moses, Complexity and flexibility. Quoted in Ideas on Complexity in Systems – Twenty Views, by JM Sussman. February. http://web. mit. edu/esd, 83

[Reason90] J. Reason. "The Contribution of Latent Human Failures to the Breakdown of Complex Systems". Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences. 327 (1241): 475–484, 1990

[Rechtin10] E. Rechtin, & M. W. Maier, The Art of Systems Architecting. CRC Press, UK, ISBN 9781420058529 - CAT# E0440.

[RSNL12] Rail Safety National Law (as applicable to each Australian State)

[ONRSR16b] ONRSR Guideline Meaning of duty to ensure safety so far as is reasonably practicable.17 May 2021

[Scholarpedia07] G. Nicolis and C. Rouvas-Nicolis, Complex systems. Scholarpedia, 2(11):1473, 2007

[Waterloo19] Waterloo Institute for Complexity and Innovation, What are Complex Systems?, https://uwaterloo.ca/complexity-innovation/about/what-are-complex-systems

[Wiles05] J. Wiles and J. Watson, Patterns in Complex Systems Modeling, In *Proceedings of the Sixth International Conference on Intelligent Data Engineering and Automated Learning (IDEAL'05)*, 2005

[Willey14] R. J. Willey, Layer of Protection Analysis, In *Proceedings of the 2014 International Symposium on Safety Science and Technology (ISST2014)*, 2014