Distinguishing Fact from Fiction in a System of Systems Safety Case

Zoë Stephenson[1], Christian Fairburn[2], George Despotou[1], Tim Kelly[1], Nicola Herbert[2] and Bruce Daughtrey[2]

[1]University of York, UK

[2]BAE Systems, Preston, UK

Abstract Based on our recent experience, 'distinguishing fact from fiction' in relation to System of Systems (SoS) safety has emerged as a pertinent topic in a number of senses. From an analytical perspective, we recognise that it would be a mistake to treat a SoS as 'just another complex system'. The defining properties of a SoS mean that traditional analysis methods may fall short if applied without additional support. On the other hand, we also argue that the structured and comprehensive analysis of a SoS need not be so complex as to be impractical.

We draw on an internal BAE Systems development project, Integrated Aircrew Training (IAT), as an exemplar. IAT interconnects multiple systems and participants – air and ground assets – into a training SoS. As would be expected we have identified a number of sources of complexity in the analysis of this SoS, chiefly the exponential impact of interactions among increasing numbers of system elements on analysis complexity. However, the training domain provides constraints which may be captured as feature models to structure the analysis.

We outline a SoS hazard assessment process and associated safety case approach that are the subject of ongoing research and development and as such, are not yet formally recognised. They acknowledge that the presence of human decision-makers in a SoS means that human factors analysis contributes significantly to SoS safety assessment. We discuss the human element in SoS safety analysis and show how its treatment in the case of IAT has caused us to recognise that augmented-reality training brings with it both novel sources and consequences of human 'error'. In this particular SoS, the 'fact versus fiction' differential also applies to SoS users and the notion of participant 'immersion' is a key area of interest.

1 Introduction

The System of Systems (SoS) term is increasingly used to describe classes of systems, such as the Air Traffic Management (ATM) and Network Centric Warfare (NCW) paradigms, which exhibit a combination of the following characteristics:

- common overall objectives
- multiple elements which are systems in their own right and which demonstrate varying degrees of autonomy
- geographically dispersed elements
- dynamic collaboration of elements to achieve common objectives
- heavy dependency on network communications
- ad-hoc communications networks (Alexander et al. 2004)
- independently developed constituent elements (Despotou et al. 2009)
- emergent behaviours (Maier 1998)
- system complexity.

This definition helps us to differentiate 'fact' from 'fiction' in an analytical sense as it illustrates that the SoS term is about more than just complexity and scale. Consider for example an aircraft. Under the SoS definition presented above, an aircraft would be considered as a complex system but it would be wrong to view an aircraft as a SoS. An aircraft is a complex system whose components perform distinct roles, but these components are not independent systems in their own right in a SoS sense; an aircraft engine provides propulsion and is something of a 'hub' with regard to onboard power generation, however it is co-located with other system elements on the aircraft platform and the way it interacts with these other elements is well defined, predictable and relatively invariable.

Comparing the aircraft complex system with, for example, an ATM or NCW SoS, it is clear that the degree of 'openness' and flexibility with regard to membership and role within the system is quite different. The aircraft engine does not have a 'life of its own' outside the complex system. Aircraft systems are designed to work together in one or a relatively small number of fixed and well defined configurations. The way in which they are integrated takes account of this. Nodes in an NCW network, on the other hand, may or may not always be present and the functions that are performed within an NCW SoS may be allocated in quite different ways across the nodes which are members of the SoS at a particular point in time.

From a safety point of view distinguishing a SoS from complex systems is necessary. Recognition of the fact that a SoS is not just a large complex system and that a complex system such as an aircraft is not a SoS leads us to realise that the safety analyses applied to 'traditional' systems may fall short of exhaustively revealing hazards in a SoS context. Whereas our understanding and the methods used to demonstrate the safety of an aircraft are considered sufficient, it is difficult to identify the hazards in a SoS following the NCW paradigm. The inherent characteristics of SoS affect the ability to understand hazards and apportion safety related requirements (Despotou et al. 2009).

In exploring SoS safety assessment, this paper makes reference to an internal BAE Systems development project, Integrated Aircrew Training (IAT), as an exemplar. IAT is a SoS training system for training and developing combat aircrew. It allows pilots and navigators to be trained by combining Live, Virtual and simulated Constructive (LVC) elements into realistic multiplayer training scenarios

that closely resemble those encountered in operational experience. For the purposes of this paper, 'Live' elements of the IAT SoS are real people operating real equipment (e.g. aircrew flying in fast jet aircraft). 'Virtual' elements of the IAT SoS are real people operating simulated equipment (e.g. aircrew 'flying' a ground based simulator) and 'Constructive' elements of the IAT SoS are simulated people or simulated equipment under computer control (e.g. a computer generated synthetic hostile aircraft or surface to air threat).

IAT provides a common training environment, with the ability to coordinate multiple assets. IAT interconnects multiple systems and participants: aircrew in real aircraft, Virtual participants in ground-based simulators and training devices. Geographically dispersed participants can train side-by-side. IAT can thus be described as a SoS that consists of various interconnected standalone systems. The obvious advantage of adopting such an architecture is that it provides a flexible training system that is able to adapt according to training needs and the required numbers of participants in a training scenario. However, the systems in the IAT SoS are of diverse types and build standards, and will include aircraft, simulators, communications and ground equipment that were not originally designed to be interconnected. Furthermore, the complexity and variability of such a SoS has the potential to make safety analysis problematic.

Having established that SoSs are not the same as complex systems, and that complex systems such as aircraft are not SoSs, we need to establish which existing techniques/approaches can be utilised in support of SoSs, which cannot and what else is required in addition (i.e. distinguishing fact from fiction when moving from complex systems to SoSs). Based on this, the paper will therefore outline a SoS hazard assessment approach, detailing some of the steps we are taking to handle the difficulties of safety analysis for our exemplar, IAT. Despite the fact that a SoS like IAT raises a number of challenges, analysts can still use existing principles to structure the SoS safety lifecycle; however, there can be a certain amount of hazard attributed to the distinct combination of SoS characteristics that will require problem specific approaches (Despotou and Kelly 2010). One of the most challenging parts of this project is the variation in human behaviour; and so we also describe aspects of our approach that are specific to the modelling of how human behaviour contributes to safety. Finally, we conclude with our plans for future development, particularly in the use of automated techniques and in the validation of our proposed method as an adequate and practical response to the 'facts' of SoS analysis.

2 Hazard assessment approach

Rigorous, traceable and comprehensive hazard assessment is required in the production of safety involved systems, regardless of the specific type of system in question. Assessment techniques and processes are well established for this purpose, however knowing what is known about SoSs, it could be a mistake to apply

these existing methods in exactly the same way as they applied in the case of complex systems.

For this paper we treat hazard assessment as a combination of two interrelated concepts: hazard identification, in which the possible hazardous events at the system boundary are discovered, and hazard analysis, in which the likelihood, consequences and severity of the events are determined. The hazard identification process is based on a model of the way in which parts of a system may deviate from their intended behaviour. Examples of such analysis include Hazard and Operability Studies (HAZOP, Kletz 1992), Fault Propagation and Transformation Calculus (Wallace 2005), Function Failure Analysis (SAE 1996) and Failure Modes and Effects Analysis (Villemeur 1992). Some analysis approaches start with possible deviations and determine likely undesired outcomes (so-called inductive approaches) while others start with a particular unwanted event and try to determine possible causes (so-called deductive approaches). The overall goal may be safety analysis, to assess the safety of a proposed system (a design, a model or an actual product) or accident analysis, to determine the likely causes of an incident that has occurred.

The challenge for hazard assessment in a SoS is the inability of current techniques to convincingly account for uncoordinated interactions between the various systems. Such interactions typically involve a coincidence of events that are otherwise not hazardous, and which may (as noted by Alexander 2007) be of rather different types, involving information from multiple domains studied by a diverse range of experts. Without specific techniques to address the complexity, the likelihood is that many hazardous interactions will remain unaddressed.

Approaches that try to address the phenomenon of unknown variation include:

Modelling the variation. A description such as a feature model (Kang et al. 1990) can be created to explicitly detail the configurations that will be encountered. The resulting safety assessment is contingent on every configuration that is subsequently encountered being represented in the feature model, something known as the *oracle hypothesis* (Weiss and Lai 1999). The approach is generally applied in product lines in which the risk of not modelling the variation is greater than the risk of an incorrect model. This use of modelling goes some way towards meeting the suggestion of (Raheja and Moriarty 2006):

> 'System safety needs to pay more attention to hazard analysis on the structure and architecture of the system-of-systems.'

Product-line safety assessment typically generalises from individual safety analyses to produce a configurable analysis result that is then customised to the particular product in question. The process may involve annotation of an existing model, such as fault trees (Dehlinger and Lutz 2005), or it may involve the creation of a new model (Stephenson et al. 2004). In contrast to these relatively simple approaches, (Habli 2009) describes a complex meta-model that relates design variation, context variation, events, consequences and severities. It is important to note that each of these bodies of work assumes the existence of hazard assessment in-

formation relating to particular products that can be used to populate and validate the model.

Focusing the analysis with decision support tools. Tools can cut down a complex search space by highlighting important aspects of the space for further analysis. (Alexander 2007) proposes such a tool for SoS hazard assessment. The tool uses agent-based simulation to suggest likely unwanted events and explain their causes. In cases where the complexity is so high that analysis is impossible without some filtering of the events, such an approach provides useful support; while it does not claim to discover all of the unwanted events, it does provide a tractable starting-point for manual training scenario analysis.

Creating a richer model of causation. (Leveson and Dulac 2005) propose the STAMP accident model and the STPA hazard assessment approach. STAMP is based on systems-theoretic concepts of hierarchical control, internal models of the environment and a classification of control errors. STPA takes that classification as the basis for iterative integrated control system safety assessment. At each design iteration the design is assessed and constraints are derived (equivalent to derived safety requirements) and imposed on further design iterations.

We base our choice of hazard assessment approach on the recognition that the scope of variation in IAT is limited and well-defined due to its reuse of existing training infrastructure and equipment. The key additions with IAT – the augmented reality/synthetic elements – are small modifications to allow for integrated simulation coupled with new technology embodying the majority of the simulation functionality. The deviation from existing end-user functionality is therefore relatively small. In contrast to SoSs that include autonomous equipment, there is significant opportunity for human review in this SoS domain. As additional motivation, we note that the hazard assessment and safety case production for an individual training scenario must be as streamlined as possible.

The approach for hazard assessment and safety case development is outlined in Figure 1. We propose to analyse safety from two perspectives, 'pre-deployment' and 'post-deployment' (discussed in more detail later). We model the limited, well-defined variations between training scenarios as a feature model. The configuration process automatically suggests the corresponding safety case structure, hazard log, derived safety requirements and mitigations, as traceable, validatable links from the configuration specification to explicitly-defined reusable artefacts. At each stage, the automated suggestions are validated using a validation guide that is specific to the type of training scenario being configured. Where mitigating actions alter the configuration, the process is iterated to assess any remaining hazards. To ensure that this process terminates, we impose a constraint that all mitigating actions correspond to a move from one configuration to another that is no less constrained.

To set up the artefacts for this process, we also base our up-front processes on the feature model, as shown in Figure 1. The initial data for each system or agent involved in IAT is derived from a differential analysis, looking in detail at the way

the use of IAT influences a particular task or system. For example, when the IAT SoS is active, cockpit display systems will represent data pertaining to both Live and synthetic agents, also the briefing process may need to include participants who are to 'fly' as Virtuals as well as those who are to fly as Live platforms in the training scenario. A high-level preliminary assessment is performed on the general training scenario scope at a suitable configuration stage. This is chosen so that the groupings correspond to distinct training scenario types, as determined by domain experts. The assessment is based on HAZOP and HAZAN (Hazard Analysis) with customised guidewords for specific flow types. A low-level analysis is performed on exemplars of each training scenario type, to investigate detailed training scenario characteristics. This is generalised into the scope of the preliminary analysis. Where there are mismatches or other issues from generalisation, we perform a specific interaction analysis based on both nominal and abnormal behaviours, and we generalise from all such analyses to inform the overall hazard assessment approach. Finally, where there are specific translations or mappings from one representation or medium to another, we explicitly assess the gap to determine whether it contributes additional deviations that need to be addressed. We recognise throughout this approach that system hazards, SoS hazards, likelihood and severity may all be contingent on the particular configuration options that are selected.

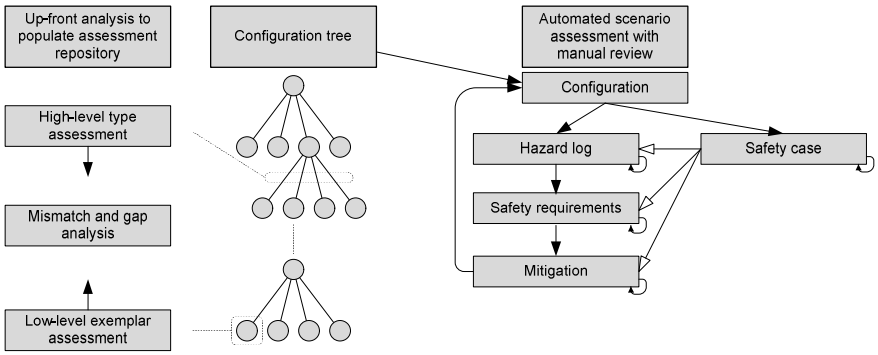

Fig. 1. Configuration-directed hazard assessment and safety-case production

Our initial process definitions have highlighted a number of specific characteristics of the assessment approach, as detailed in the following sections.

2.1 Feature modelling

A typical feature modelling approach represents features – aspects of a product that are interesting to particular stakeholders – and the available combinations of features (Czarnecki and Eisenecker 2000). The original feature modelling ap-

proach (Kang et al. 1990) is based on binary features and logical feature dependencies; we are using an extended feature modelling approach with:

- cardinalities, representing replication of a feature, to handle the variation in the number of participants in a training scenario
- staged configuration, allowing for partial customisation of the feature model
- uncertainty, to represent approximate configurations
- explicit binding processes, to structure the configuration stages and uncertainties into a manageable sequence of decisions.

It is recognised that feature models provide a very useful representation of features and dependencies for a particular product, however care needs to be taken as some types of product feature might be more appropriately represented using a different modelling technique. Typical examples include N×N dependency matrices (Stephenson et al. 2004) or parameter envelopes.

One final concern with the use of the feature model is its central role in deciding what reusable artefacts are included. If the feature model is incorrect, it has consequences for the validity of the hazard assessment, the safety claims and the entire safety case. To offset this concern, we note our assumptions that:

- Many issues with appropriateness of scope and with validity will be uncovered as the safety assessment process explores the configuration space.
- We are able to set up and maintain appropriate feedback paths within the development process to allow any member of the development team to report problems with the feature model.
- We are able to assign appropriate responsibility and authority to control the evolution of the feature model, with associated traceability and change history.
- We are able to set up and maintain similar feedback paths between the customer and the development process to catch any remaining problems that occur.
- We are able to budget for, set up and conduct periodic reviews of the scope and validity of the model.

2.2 Configuration space structure

In our development approach, there is an assumption that the different binding processes are able to communicate the scope of the intended use of the system using a customised feature model. For this to work in practice, it must be able to represent:

- the changes to the configuration over time – as decisions are made, as feasibility studies are conducted, as work is completed to integrate component systems
- the different versions of the product – as new features are developed and enabled for successive releases of the product

- the different customers – each installation will have different equipment available, different subsets of the training needs and so on.

Moreover, these aspects must be managed in the binding processes (the processes that make decisions about the configuration of the model) and the configuration stages as well as in the dependency structure of the feature model. For example, if a later version of the product for a particular customer needs to enable a new type of training scenario, it must be possible to create the appropriate configuration stage within the binding process associated with that deployment; this configuration stage is a rebinding of the configuration that was used for the previous version. There is nothing inherent in this process that ensures that neither of these versions undoes a commitment made in the previous configuration stage, within the development process. Indeed, it may be necessary to unbind internal development commitments in order to produce a product release with the desired features.

Propagating unbinding through previous development stages is especially costly when the feature model and the configuration approach are the basis for the safety assessment and safety case. To limit the effects of these issues, we take the traditional step of aligning feature model structure to business goals (Bayer et al. 1999) – a step well-supported by the use of MODAF (Ministry of Defence Architectural Framework) as it covers both managerial and technical activities –but we also aim to provide specific validation criteria such as measurements of the distances between scenarios with similar objectives, that report on the healthiness of the feature model. We expect that this will be especially valuable when dealing with feature models that carry uncertainty information.

2.3 Pre-deployment hazard assessment

The pre-deployment hazard assessment must take into account the presence of hazards across all of the different agreed possible uses of the deployed system. The assessment is divided into two parts:

- assessment of the non-variable elements of the SoS: the training devices, the network infrastructure, the set of training needs and the process structure
- assessment of variable elements of the SoS: particular objectives for a training scenario, particular training needs, particular equipment used for a training scenario.

We expect the former to be relatively straightforward. For the latter, the goal is to set up appropriate reusable data for the post-deployment assessment process. This includes reusable artefacts such as hazard entries and derived safety requirements, but also specific versions of the assessment process for particular types of training scenario.

2.4 Post-deployment hazard assessment

The suggested hazards from the previous step provide a 'head start' in the hazard assessment of the training scenario and the production of the safety case, augmented by manual review. Particular issues of concern here are:

Incompleteness. There may be hazards in the training scenario that were not associated with the feature model, but which would have been found in a conventional assessment approach. This is a serious concern with any approach that tries to automatically identify hazards in a system. Our aim with the post-deployment assessment is to show that we have supplied appropriate artefacts for reuse and that there are sufficient opportunities for manual review and operational activity to uncover and mitigate additional hazards.

Pessimism. The up-front analysis is likely to be conservative. In certain configurations it will highlight potential hazards that will not be manifest in practice. Similarly, the likelihood of an event may be overestimated or the severity of the consequences may be overstated. We expect to discover these problems during the validation stage and to devise specific guidance on the reduction of pessimism. It is likely that we will introduce mechanisms to measure pessimism in the analysis for particular training scenarios as part of the overall evaluation of safety assessment effectiveness.

System Interactions. The presence of unintended interactions between nominal behaviours complicates the assessment approach. We expect that the constrained scope will lead to a small set of interaction classes to analyse for any given training scenario, generalised from the results of particular assessments. For example, the presence of 'fictional' elements in training scenarios is expected to give rise to a general class of interactions as well as specific issues for particular elements. As part of the analysis of the effectiveness of the safety assessment approach, we intend to provide for feedback and review of interaction analyses.

Multiple Versions. It is desirable to be able to meet a particular training need in a number of different ways, to avoid learning by rote. At the same time, it is important to be able to compare similar training scenarios to ensure that large variations in the safety assessment results are catered for. This naturally leads to the idea of comparing training scenarios for similarity; we are particularly concerned with the possibility that configuration similarity may be quite different from training scenario similarity, and that this may lead to training scenarios that are superficially similar being assessed in similar ways without properly investigating the details of the training scenarios such as the varying demands on human participants.

2.5 Safety case

A number of standards require a system to be accompanied by a safety case, and a SoS should be no different. The safety case communicates an argument, supported by evidence, that a system is acceptably safe in a given operational context. The safety case captures the underlying reasoning and evidence that support claims made about the safety of a system. We believe that the hazard assessment process we have proposed is compatible with the creation of a safety case. We also believe that it will be possible to produce a manageable SoS safety case which takes account of SoS specific issues and which conforms to relevant standards.

Process-based standards such as IEC 61508, DEFSTAN 00-55 and DO-178B list predetermined activities, which when followed by system developers are considered to result in an acceptably safe system. All three standards use a risk-based approach according to which – depending on the consequences of the risk (i.e. catastrophic, major etc.) – the system is assigned a 'safety' level such as Safety Integrity Levels. Safety levels represent quantified risk targets that a system has to meet, based on probability for the occurrence of an accident and/or the severity of that accident. According to the safety level a system is required to achieve, standards specify the means with which the developers will acquire assurance about the system's operation. This also applies to evidence collection, for which there are tables prescribing the testing techniques and methods required (or recommended) for each safety level. By and large, higher assurance levels involve more thorough examination of a system. This can involve additional testing techniques 'cross-checking' the system behaviour, as well as more advanced techniques (such as formal methods) contributing to the overall assurance of claims regarding the system's safe operation.

Use of prescriptive standards provides a specific and easily interpreted way of acquiring assurance about the safe operation of the system. In contrast, goal-based standards such as DEFSTAN 00-56 (Issue 4) require the developer to assure the safety of the delivered system through structured reasoning, with the provision of a safety case. This includes reasoning about using the right (system development) techniques at the right time during system development to support the safety case. This allows potentially greater flexibility, as developers are not instructed to use a specific set of techniques. Although the two categories of standards adopt different philosophies, it is generally recognised that the higher the involved risk is, the more evidence and scrutiny are required. This is a default position in prescriptive standards. In evidence-based standards this principle appears in the form of quantity and quality of evidence required to support a position. DEFSTAN 00-56 (Issue 4) recognises that: 'The quantity and quality of the evidence shall be commensurate with the potential risk posed by the system and the complexity of the system'.

It has thus been identified that there is a requirement for SoSs to be accompanied by a goal-based SoS safety case. One 'fiction' relating to SoS safety cases is that producing and maintaining an SoS safety case for a large, very complex and

changeable SoS will be in all cases an onerous task that borders on the impossible. However, we believe that a maintainable SoS safety case should be achievable through the development of a SoS safety case with a modular design.

The main advantage of adopting a modular approach to the safety case is in its maintainability. Developers can create coherent arguments about an aspect of the system, which can then be integrated into the system safety case. This can be particularly useful for incremental development of the safety case, in parallel with the system. Adopting a modular approach can be useful in the management of change, which is inevitable in development of an incremental safety case. Packaging the safety case in cohesive modules can help in isolating the change and understanding its impact. The effectiveness of a modular safety case (with respect to managing change) depends on the type of change. However, adoption of a modular approach is recommended as it exhibits the following advantages (Despotou and Kelly 2008):

- will contain some of the changes
- does not add any technical overhead when compared to a monolithic safety case
- improves the clarity of the safety case
- allows different stakeholders to isolate the arguments that are most relevant to them.

Furthermore, a modular design will allow isolation of claims about scenarios and the evidence that will need to be produced to support scenario related claims.

In the case of our exemplar, IAT, preliminary safety activities have started to identify the types of evidence that will be considered suitable to support the claims made in the safety case to an acceptable degree of confidence. The identified types of evidence include evidence that are produced by processes common in systems adhering to older standards (such as 00-55) as well as more novel safety analysis methods. Part of the novelty of IAT derives from its SoS characteristics, and therefore novel safety analysis techniques are required.

Another ‘fiction’ relating to SoS safety cases is that the adoption of Def Stan 00-56 Issue 4 will make it difficult to construct a safety case because the onus is on the designers to argue that the right evidence has been produced in support of the safety case. However, the ‘fact’ is that Def Stan 00-56 Issue 4, being a goal-based rather than a prescriptive standard, allows potentially greater flexibility in the types of evidence that can be presented in order to support the safety case and demonstrate that the system of systems is acceptably safe. Therefore the outputs of novel safety analysis techniques can be used as evidence to support the SoS safety case.

All of the safety analysis techniques that have been discussed in this paper will support the SoS safety case by providing evidence that we have identified all IAT hazards, both hazards at traditional system boundaries and from system interactions, and that the streamlined safety analysis process for training scenarios provides a similar hazard assessment capability.

3 Analysis of the human element

The close connections between human 'error', safety and accident prevention are well appreciated in the aviation domain. BAE Systems routinely employs usability assessment, Human Error Prediction (HEP) and Human Reliability Analysis (HRA) techniques in the design and clearance of its complex system aircraft products. Assessment techniques and processes have been established for these purposes, but SoS style products are likely to present new challenges in this area. Endeavouring in a practical way to respect a contextual view on human performance during hazard assessment (see for example Dekker 2002, Hollnagel 1998) brings the need to anticipate and account for circumstances surrounding situations of system use. As we have already described, the defining characteristics of SoSs have the potential to increase complexity in this area but we believe that it is possible to acknowledge and take account of this complexity without making analysis unmanageable. We describe a human factors oriented approach to differential analysis which aims to supplement more traditional human factors analysis methods. This is seen as a key driver for the early stages of SoS hazard identification in a SoS where the human element is considered significant. We briefly show how taking account of 'differences' from the human factors perspective in the case of IAT has led us to recognise that augmented-reality training brings with it both novel sources and consequences of human 'error'.

3.1 Towards human factors methods for SoS hazard identification

Taking into account the significance of human factors to SoS safety, we have identified important roles for human factors analysis in both the IAT SoS hazard identification and hazard analysis processes.

The paper has already introduced the concept of differential analysis and its role in the SoS hazard assessment approach being used for IAT. Considering human users as some of the 'systems' in a SoS architecture, effort has been directed to the development of a human factors oriented approach to differential analysis. This way forward strikes a pragmatic balance with regard to the 'contextual human performance' problem for the purposes of hazard identification, limiting the re-analysis of existing SoS elements but maintaining an appreciation that the interaction of these elements in the novel context of a SoS may lead to hazard.

To be successful, a human factors approach to differential analysis must identify the SoS related 'differences' that may affect humans operating within the SoS but it must also be able to articulate how these differences might influence them. Work is ongoing in this area but in general terms, the 'differences' to which SoS users will be exposed may be conceptualised by taking elements of to-be-performed work and situating these against contributors such as system (software/ hardware/procedural) requirements or other factors of relevance. These 'differ-

ences' are then passed through a model (or models) of human performance to consider influences on human mental and physical behaviour in a traceable, methodical and consistent way (Figure 2). The influences that are identified may then be subjected to deviation analysis by way of HAZOP or similar approaches. Returning to points made earlier in the paper, in the case of IAT it is recognised that we will need to apply this method to take account of work which takes place before, during and also after training in order to take a suitably comprehensive view.

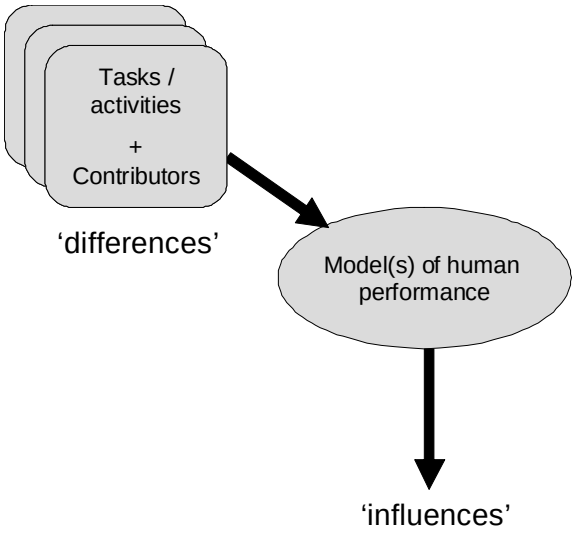

Fig. 2. Human factors oriented differential analysis for hazard identification

3.2 Human factors differential analysis for IAT

The application of a human factors oriented differential analysis to the IAT SoS is helping us to systematically explore the significance of the human element in the processes of configuring the IAT training system for use, using the system within a single training intervention and using the system across multiple training interventions. We are also looking further ahead, where factors such as 'negative training' could conceivably lead to hazard even when the IAT system itself is no longer active. We aim to account for these factors by including explicit modelling of training phases and configuration options, including the 'real mission' phases that occur outside of the training process. The differential approach is in particular helping us to systematically 'unpack' the potentially widespread implications of introducing synthetic elements into a Live training context.

In terms of configuring the IAT system for use, much of the work involved in the authoring of training scenarios will relate to the construction of the augmented reality in which Live and Virtual participants will operate. Where Constructive

elements or synthetic objects are included they must be represented to an appropriate level of fidelity and accuracy. The realism and appropriateness of Constructive entity behaviour must be sufficient both to satisfy training requirements but also to adhere to important real-world rules. Incongruous Constructive entity behaviour may distract other scenario participants away from crucial tasks or perhaps lead them into hazardous situations in more direct ways. Where real-world objects are re-represented in the synthetic elements of a training scenario this must be done reliably. In addition to the reliable composition of training scenarios, scenario authors must also be equipped to understand and verify how a training scenario might evolve or progress over time, taking into account various contingencies. There may be a need to constrain the ways in which a specific training scenario is able to evolve, in effect safeguarding the circumstances in which Live and Virtual participants may be allowed to operate. Taking into account the synthetic 'difference' and its influence across a broad range of human inputs to training system configuration it becomes apparent that quite ordinary or mundane breakdowns in human performance have the potential to translate into significant hazard. For example, slips or mistakes with a keyboard and mouse during training scenario construction might have the power to translate into the incorrect positioning of a real world physical object (high terrain or a tall building) in the synthetic environment. This would influence the behaviour of Constructive entities and Virtual participants at the time of scenario execution. While this type of latent failure may not be a safety critical issue in a simulator, in an augmented reality training system there would be the potential for consequences to propagate across the SoS and for Constructive or Virtual behaviours to also influence the actions of Live participants.

Use of the IAT SoS during a training intervention is ultimately concerned with the participation of human actors in an augmented reality 'gameplay' scenario. There are of course many senses in which human performance should be analysed in this context, but exploring the synthetic 'difference' and its influence on training scenario participants reveals how aircrew behaviour may be affected by a potentially powerful cognitive phenomenon. 'Immersion' is a term commonly used by those working in the domains of gaming and virtual reality. It is colloquially understood as the sense of being 'lost' in the game or virtual environment where players become highly involved in the synthetic elements of their environment and lose awareness of their real world surroundings (Brown and Cairns 2004). In trying to conceptualise this area further, researchers have proposed that immersion is a continuum (Brown and Cairns 2004) and qualitatively distinct notional 'levels' of immersion have been distinguished. In a state of moderate immersion ('Engrossment') a participant is said to become less aware of their surroundings and less self aware. They have started to 'suspend disbelief' in the game. The 'Total' immersive state is said to be achieved if a participant loses awareness of physical reality and in effect becomes 'bought in' to the synthetic elements of their experience, treating them as if they were real. The concept of immersion has close links with other areas of literature including 'Cognitive Absorption' (Agarwal and Karahanna 2000), 'Flow' (Csikszentmihalyi 1990) and of course human attention.

Although it is fair to say that immersion and other related concepts still require some clarification, these notions have attracted the interest of the IAT project and are being borne in mind for safety assessment. The differential analysis approach is helping us to take immersion seriously. Importantly, we take the position that it is not the introduction of augmented reality training that suddenly makes immersion relevant to aircrew safety. After all, training related 'role-play' with aircrew assuming the role of hostile adversaries is commonplace in current training practices and must already invite a certain 'fact versus fiction' partition from a cognitive perspective. We do argue, however, that the introduction of LVC technology has the potential to make immersion a much more significant issue. The participation of ground based Virtual participants purporting to be airborne during a training scenario when they are not and the inclusion of 'fictional' computer generated Constructive entities must serve to increase complexity in this area.

During hazard analysis, much more detailed HRA/HEP work is expected to be required. In the case of IAT we intend to pursue research of the immersion topic in more depth. We aim to understand more about what immersion-related 'error' might look like and the conditions under which immersion should be accounted for in hazard analysis. In turn we aim to appreciate how we might mitigate immersion related routes to hazard and make the IAT SoS tolerant to this phenomenon. We envisage a need to take account of the ways in which Live and Virtual participants may come to hold false hypotheses regarding what is real and what is not during IAT training scenario performance. We need to recognise if or where this could increase risk. Our analyses will also need to appreciate the cognitive work faced by Live participants who must rapidly emerge from augmented reality. We aim to design to support immersive training where this is of benefit, but at the same time we intend to design for graceful emergence from immersion, so that aircrew needing to deal with in-flight emergencies have the attentional capacity to do so without unnecessarily expending resources to determine fact from fiction in a time critical situation.

4 Validation

It should be noted that the process presented here is the subject of ongoing research and as such is not yet formally recognised. We expect its validation to be iterative and content-driven, feeding back information from every use of the process to challenge assumptions and provide further guidance. Our validation of the hazard assessment approach is structured into a number of stages:

- Validation of the underlying models (particularly the feature models) by populating the models with example data. This provides insight into the models for multiple stakeholders, provides a way of generating test-cases that reveal flaws in the models, and ensures that the model definitions match with the specific uses encountered in this domain.

- Validation of the preparatory processes, investigating coverage and estimating effort, particularly where collaboration is needed. The analysis processes are intended to cover humans, technological systems and infrastructure and to set up reusable analysis artefacts, and so the validity of these processes, both in concept and in execution, is especially important.
- Identification of a representative part of the IAT product, covering IAT development, IAT deployment, training scenario development and post-training activity. This ensures that the validation assesses processes that involve a number of different human agents and different levels of 'fiction' – aspects of augmented reality.
- Validation of hazard assessment processes on the identified IAT processes, logging effort data. This covers the complete run of processes from initial scoping through asset generation to streamlined training scenario analysis and feedback.
- Estimation of overall IAT hazard analysis effort, covering up-front processes, per-delivery effort and effort per training scenario.
- Estimation of hazard assessment effectiveness. This will involve feedback from validation stages as well as the post-deployment assessment and safety case construction processes.

It is our intention to document the case for validity of the process: one option is to provide a goal-structured argument to organise the evidence from the validation processes. We are mindful of the need to control complexity, however, and will consider the most manageable alternative. It is recognised that further validation, formalisation and standardisation must take place before this process can be used for certification purposes.

5 Summary and future work

Given that SoSs are not the same as complex systems and that complex systems such as aircraft are not SoSs, this paper sets out a vision for hazard assessment of a complex integrated SoS that involves human elements and a variety of engineered systems. From an analytical perspective, we recognise that it would be 'fiction' to treat a SoS as 'just another complex system'. The defining properties of a SoS mean that traditional analysis methods may fall short if applied without additional support. On the other hand, we also argue that the structured and comprehensive analysis of a SoS need not be so complex as to be impractical. We identify feature modelling as a potential approach to manage the complexity of hazard assessment, and structure the assessment approach so that it may be driven by configurations in the feature model, at least in the abstract. We pay particular attention to the link between the hazard assessment and the safety case, and we investigate some of the human factors issues that arise in the IAT SoS, particularly

the issue of immersion and the 'fact'/'fiction' differential that an augmented reality training SoS brings from a user perspective.

Our future work on this project can be viewed as three complementary strands:

- validation of the approach, particularly its basis in feature modelling and its avoidance of analysis approaches that are targeted only at understanding interactions in systems of systems
- exploration of immersion and its impact on both system design and hazard assessment
- automation of analysis and configuration processes. Candidate processes for automation include the mapping from training scenario setup to features; the mapping of features to hazards, derived requirements and mitigations; the assessment of particular aspects of training scenarios such as network usage; and the assessment of training scenarios using agent-based simulation to discover novel interactions.

References

Agarwal R, Karahanna E (2000) Time flies when you're having fun: cognitive absorption and beliefs about information technology usage. MIS Quarterly 24:665−694

Alexander RD (2007) Using simulation for systems of systems safety analysis. PhD Thesis, University of York

Alexander R, Hall-May M, Kelly T (2004) Characterisation of systems of systems failures. Proceedings of the 22nd International System Safety Conference (ISSC '04)

Bayer J, Flege O, Knauber P et al (1999) PuLSE: a methodology to develop software product lines. Proceedings of the Fifth Symposium on Software Reusability

Brown E, Cairns P (2004) A grounded investigation of game immersion. Proc CHI. ACM Press

Csikszentmihalyi M (1990) Flow: the psychology of optimal experience. Harper and Row, New York

Czarnecki K, Eisenecker U (2000) Generative programming. Addison-Wesley, Reading MA

Dehlinger J, Lutz RR (2005) Software fault tree analysis for product lines. Proceedings Eighth IEEE International Symposium on High Assurance System Engineering.

Dekker S (2002) The field guide to human error investigations. Ashgate, Aldershot, UK

Despotou G, Kelly T (2008) Investigating the use of argument modularity to optimise through-life system safety assurance. Proc 3rd IET Int Conf on System Safety (ICSS). IET

Despotou G, Kelly T (2010) Understanding the safety lifecycle of systems of systems. To appear in: Proc 28th International System Safety Conference (ISSC), Minneapolis

Despotou G, Bennett M, Kelly T (2009) Supporting through life safety assurance of COTS based upgrades. Proc 27th International System Safety Conference (ISSC), System Safety Society

Habli IM (2009) Model-based assurance of safety-critical product lines. PhD Thesis, University of York

Hollnagel E (1998) Cognitive reliability and error analysis method (CREAM). Elsevier, Oxford, UK

Kang KC, Cohen S, Hess J et al (1990) Feature-Oriented Domain Analysis (FODA) feasibility study. Technical report CMU/SEI-90-TR-21

Kletz T (1992) HAZOP and HAZAN: identifying and assessing process industry hazards. Hemisphere Publishing Corporation, Washington

Leveson N, Dulac N (2005) Safety and risk-driven design in complex systems-of-systems. 1st NASA/AIAA Space Exploration Conference

Maier M W (1998) Architecting principles for system of systems. Syst Eng 1:267-284

Raheja D, Moriarty B (2006) New paradigms in system safety. J Syst Saf 42(6)

SAE (1996) ARP-4761 Aerospace recommended practice: guidelines and methods for conducting the safety assessment process on civil airborne systems and equipment, 12th edn. Society of Automotive Engineers

Stephenson Z, de Souza S, McDermid J (2004) Product line analysis and the system safety process. Proceedings of the International System Safety Conference

Villemeur A (1992) Reliability, availability maintainability and safety assessment. John Wiley and Sons, New York

Wallace M (2005) Modular architectural representation and analysis of fault propagation and transformation. Electronic Notes in Theoretical Computer Science 141(3)

Weiss DM, Lai CTR (1999) Software product-line engineering: a family-based software development process. Addison-Wesley, Reading MA