Proceedings of the first ADVISES Young Researchers Workshop

Forskningscenter Risø, Roskilde; Nayebkheil, A.

Publication date:
2005

Document Version
Publisher's PDF, also known as Version of record

RISØ

Risø-R-1516(EN)

Proceedings of the first ADVISES Young Researchers Workshop

Edited by

Hans H.K. Andersen and Asmatullah Nayebkheil

Risø National Laboratory
Roskilde
Denmark
April 2005

Author: Hans H.K. Andersen and Asmatullah Nayebkheil (eds.)
Title: Proceedings of the first ADVISES Young Researchers Workshop
Department: SYS

Risø-R-1516(EN)
April 2005

Abstract (max. 2000 char.): Abstract: The research training network Analysis Design and Validation of Interactive Safety-critical and Error-tolerant Systems (ADVISES) focus on coaching and supervising Post Docs and Ph.D. students. The main objective is to provide a multi-disciplinary research training that can combat the impact of human error during the design, operation and management of safety-critical, interactive systems. Additionally, the exchange of knowledge, practices, tools and experience between adjacent (but still too distinct) disciplines can lead to the efficient integration of complementary research methods. Ultimately, it is hoped that this will contribute to a new and more unified research agenda for the development of safety-critical, interactive systems. One of the training instruments is the Young Researchers Workshop, which is arranged and coordinated by the young researchers themselves. It is an event where the young researchers have opportunity to present their work based on paper submissions. At this first workshop a series of papers, three of these is included in this proceedings. The first paper by Sandra Basnyat and Philippe Palanque, argues that in order to reduce the occurrence of erroneous events in the design of safety-critical interactive systems it is necessary to apply
formal description techniques for task and human error modeling and to extend information usually represented in a standard task model to explicitly express user deviations. The second paper by Bastiaan A. Schupp, emphasize that current approaches to integrating human factors issues in the development of safety critical systems appear is not fully sufficient and argues for creating a safety architecture based on a Safety Modeling Language, which uses barriers to prevent of stop undesired effects. The third paper by Alexandre Alapetite is a position paper on voice recognition in multimodal systems focusing on a case study of anesthesia patient journal. In emergency situations during anesthesia, when doctors and nurses are busy and maybe stressed, the registration process is delayed. This is a problem, because postponing the registration often leads to uncertainty, inaccuracy and other errors. It is argued that multimodal systems in a non-intrusive way can support such activities.

ISSN 0106-2840
ISBN 87-550-3443-8

Contract no.:

Group's own reg. no.:

Sponsorship:

Cover :

Pages: 29
Tables:
References:

Risø National Laboratory
Information Service Department
P.O.Box 49
DK-4000 Roskilde
Denmark
Telephone +45 46774004
bibl@risoe.dk
Fax +45 46774013
www.risoe.dk

Contents

Voice recognition in multimodal systems: the case of anaesthesia patient journal

2005-01-06
Alexandre Alapetite, Ph.D. at Risø National Laboratory;
Systems Analysis Department; Research Programme Safety, Reliability and Human Factors;
DK-4000 Roskilde; Denmark

1 Short summary

During a medical operation with an anaesthesia, the anaesthetic record is important; not only because it is a legal document, but also because it is used during the operation – for example if a new doctor is joining the team – in order to communicate and make available to the anaesthetists what has occurred previously. The fact that the document is an indispensable source of information during the operation is the main reason for maintaining a real-time system: the information entered into the anaesthesia record cannot be just recorded (audio/video) and eventually transcribed.
In operation rooms, registration of patient journal during anaesthesia has been done manually on paper for a long time. Today, some anaesthesia departments have switched to electronic systems. While electronic anaesthesia journal systems are aimed to solve most of the issues encountered with paper-based recording, there is still a room for improvement, especially in emergency situation. In the case of electronic journals, the comments, in particular, are not described as precisely as they could be, due to the use of a keyboard, which is not a convenient input device in such an environment. The aim of this research project is to study how multimodal interfaces, especially with voice interaction, could make recording more accurate, flexible and robust.

2 Rationale

2.1 Context

Voice recognition engines have been significantly improved over the last decade thanks to the introduction of new techniques and increased computer power. When used carefully, as an alternative or a supplement to other more conventional modalities (buttons, touch-screen, keyboard, mouse, etc.), voice input can now be spread in safety-critical environments.

2.2 What is the problem?

In emergency situations during anaesthesia, when doctors and nurses are busy and maybe stressed, the registration process is delayed. This is a problem, because postponing the registration often leads to uncertainty, inaccuracy and other errors.

2.3 Why are we doing this research?

I believe that adding some modalities to existing electronic patient journal systems, like voice, can be beneficial for the quality of recording. Making multimodal interfaces enables practitioners to choose between different ways of registering, depending on the current situation (touch-screen, keyboard, voice, etc.). Indeed, the different modalities do not have the same requirements (using hands, standing close to the machine, noise and light conditions, etc.) and capacities (accuracy, robustness, etc.). Voice input could be very valuable for anaesthesia electronic journal interfaces, as well as for commanding – in some cases – other anaesthesia equipments [**1**]. This could be a good solution for improving recording even in crisis situations, when most of the time the registration is delayed.
Some research has already been done with voice recognition in the medical domain. However, most existing applications are targeted at non-real-time environments where doctors provide dictation, perhaps in an office, where input and subsequent review and correction may be made in batch mode. Consequently, there is little literature about real-time speech input during operations or anaesthesias, when voice recognition is not the primary task.

3 Theoretical background

My approach is based on action research [**2**]. I work in collaboration with a group of users (experts) in Herlev hospital, Copenhagen County in Denmark, and the goal is to improve existing systems, not to validate a theory, or to create a new one. Part of the research will be a loop of prototyping and analysis. Some measurements of the impact of the new system on work procedures and quality of recording will be done and reported.

4 Methodology, main techniques

After having made a very large literature review, I have had several meetings with anaesthesia experts (doctors, nurses, engineers). They agreed on the fact that paper based recording suffers from many problems. On the other side, I have

made interviews in anaesthesia departments where electronic anaesthesia journals are used: even if there is still some improvement needed, most of the paper problems are solved. This seems consistent with an American survey which yielded the following result: “46% of medication errors occur on admission or discharge from a clinical unit/hospital when patient orders are written, and they drop by 90% when they are electronic” [**3**].

I will create some prototypes and test them in an anaesthesia simulation environment, involving doctors and nurses performing simulated anaesthesias. Measurements of the benefits of the prototype compared to existing solutions will be made. Base on those results and feed back from users, prototypes will be updated. This loop will be done several times, if possible.

4.1 Human-computer interface development

An analysis of current practices is needed in order to establish what could be said to the system and how: this is the *phraseology*. Some “Wizard of Oz” experiments are already planed in order to build the grammar, which is the formal basis of what the system can accept, and understand. Free speech, within a limited context, will also be used to allow practitioners to put some more detailed comments in the patient journal.

4.2 Software development

I have spent November 2004 in learning a new technology called XHTML+VoiceXML allowing standard and easy multimodal interaction. There are now various technical possibilities, but I might have to use some redundant voice recognition architecture to face the difficulties to understand speech in a noisy environment, and to be more fault-tolerant.

5 Case study

We are currently working with Herlev University Hospital, Copenhagen County, in Denmark.

5.1 Statistics on paper records

In Herlev, I have made some statistics on already more than 50 anaesthesias, on blank fields that should contain data, and errors about obesity, because those are easy to check: about 200 files have been randomly picked up, and 55 of them could be transcribed into a database by a qualified anaesthesia nurse, while the others where too difficult to be read. Then 27% of the 55 studied journals have provided weight and height of the patient, and among them, 66% have forgotten to check the obesity field even when the body mass index was over 25, which is a commonly accepted threshold for obesity. This first part was finished in July 2004, and more statistical analysis is currently been done on other criteria.

5.2 Mail survey

A survey has been made in Denmark in October 2004. We have sent by paper mail a questionnaire to 47 anaesthesia departments. We got 34 answers showing that 13 of them (⅓) did not use any form of electronic system and 13 used a complete electronic system (⅓). The 19 departments that are running partial or complete electronic systems are using about 12 different systems. On the 29 paper models we received, only 3 of them were identical, while all the others are specific to just one anaesthesia department. This situation appears to have an historical explanation: the different departments have progressively built their own system, with little communication among them.

5.3 Voice recognition and multimodal technologies

No applied research can be made without being in contact with up-to-date technologies. Risø and I have some background in voice interfaces thanks to the European SAFESOUND project on plane cockpits. Moreover, we got in contact with the major Danish voice recognition company early in the project [**4**]. I have analysed some reports and visited some places where using voice input in the medical domain has been successful or has failed [**5**]. In particular, a Philips/Max Manus system has been put into daily use at the radiology department of Vejle Hospital, and tested at the pathology department of Aalborg hospital and the pathology department of Sønderborg hospital. There is also a successful use of voice commands at Hvidovre hospital, with the HERMES technology, but it is in English, instead of Danish, and only for short commands.

6 Discussion

My first subject proposal was targeted at solving some of the issues implied by the use of manual paper recording for patient journals during anaesthesia in Denmark. This was in partnership with Herlev hospital in Copenhagen. It was

aimed to provide a prototype of electronic interface, with at least touch-screen and voice interaction. I have made a proper literature survey, contacted some researchers specialised in anaesthesia environments in various countries, had meetings with medical engineers and doctors. But it was only in September 2004, 7 months after having started some real work within this field, that I discovered that other hospitals in Denmark were already using very efficient electronic systems. Moreover, some statistics that are currently made (not only for my purpose) from paper records, with a high cost of human-hours, could be done more easily and much more efficiently in another hospital (Køge), which is using a full electronic system coupled to a database.

This has changed the focus of my research. As described in the current paper, I will now concentrate on how voice interfaces could improve existing electronic systems (that are only touch-screen based), especially in emergency situations where people usually stop registering. Industrial partners will probably join the project.

7 Collaboration in ADVISES

7.1 Existing collaboration

Interesting collaboration has already been done. In June 2004, I have helped Bastiaan Schupp in designing and building a Web site for improving communication between ADVISES members. In October 2004, I received good mentoring during face-to-face meeting, from Michael Harrison and Peter Wright (University of York, UK), for the preparation of an IT course I was going to give at Roskilde University in Denmark. Then, a close relationship has been established with the IT University of Copenhagen, in the context of preparation for a possible EU-IST project proposal about ambient technologies involving ADVISES nodes (so far, University of York, University of Newcastle upon Tyne, and Risø) and other industrial or university partners.

7.2 Expected collaboration

I would like to find two or three people in ADVISES who are potentially the closest to this project to get relevant advise when needed, and possibly to know where it could be the most relevant to travel to.

References

[**1**] *Are we going to talk with our anaesthesia monitors in the future?* Achim Schmitz. Acta Anaesthesiologica Scandinavica, February 2004, 48(2):255-6. PMID:14995952

[**2**] *Investigation information systems with action research*. Richard Baskerville. Communications of the Association for Information Systems, Volume2, Article 19, October 1999.

[**3**] *Medication reconciliation: a practical tool to reduce the risk of medication errors*. Peter Pronovost. Journal of Critical Care, Volume 18, Issue 4, December 2003, Pages 201-205. doi:10.1016/j.jcrc.2003.10.001

[**4**] Max Manus, http://www.maxmanus.dk

[**5**] Talegenkendelse - muligheder og barrierer for anvendelse til klinisk dokumentation (da-DK). Jens Hvidberg. Master's thesis, Aalborg University, May 2003

- DOI (Digital Object Identifier System): http://www.doi.org
- PMID PubMed (United States National Library of Medicine): http://www.ncbi.nlm.nih.gov/entrez/query.fcgi

A Task Pattern Approach to Incorporate User Deviation in Task Models

Sandra Basnyat & Philippe Palanque

LIIHS - IRIT
118, route de Narbonne
31062, Toulouse, France
basnyat@irit.fr, palanque@irit.fr

8 SUMMARY

In this paper we propose to extend information usually represented in a standard task model to explicitly express user deviations. This is to mitigate the occurrence of erroneous events in the safety-critical interactive systems domain. Due to the complexity of the extended task models, we propose the notion of re-usable task patterns as a means of reducing designers' workload.

9 Introduction

The context of our research is analysis, design and validation of interactive safety-critical and error-tolerant systems. This is performed using formal description techniques for task and human error modelling. Many types of models are used in the design process of these systems, such as the user model, environment model and platform model. We are focusing on the task model because it is known to be fundamental as far as a User Centred Design (UCD) approach is concerned.

Task modelling is usually performed with an error-free perspective. Indeed, it is already complex to deal with standard user behaviour with current notations which are clearly facing difficulties as far as real-life case studies are concerned. It is even more complex to incorporate human error into the task model with current notations lacking dedicated means for taking into account such additional necessary information.

Furthermore, human error plays a major role in the occurrence of accidents in safety-critical systems such as in aviation, railway systems or nuclear power plants [21]. Therefore, we believe, that not only should standard user tasks be considered, but erroneous user behaviour too.

We intend the extended task models to support the design of safer safety-critical systems by, for instance, ensuring that the system can return to a safe state after a problematic event has occurred.

The following section provides a summary of task modelling techniques, human error and previous research addressing the combination of task modelling and human error. We then present where and how our work fills a gap in this current state of the art. Finally we discuss our approach using an illustrative example and conclude with our plans for collaboration.

10 Task modelling

Tasks analysis and modelling is widely accepted as a central element to user centred design approaches [5]. A task model is a representation of user tasks often involving some form of interaction with a system, influenced by its contextual environment. Examples of established task analysis techniques include HTA [1], UAN [6], the GOMS family [8] and CTT [17].

Users perform tasks, which are structured sets of activities [19] in order to achieve higher-level goals. Tasks can be further decomposed resulting in lower level sub goals. This notion of decomposition naturally results in tree-like structures and thus a hierarchical representation of the model. More recently, the ConcurTaskTrees CTT [17] notation has enabled the distinction of abstract, user, interaction and application tasks. It also deals with temporal aspects of interaction when specifying a model. CTT is supported by a tool ConcurTaskTrees Environment (CTTe) for the editing and simulation of task models.

Task modelling notations and tools as yet, do not provide a dedicated means to address the issues of description, representation and analysis of unexpected eventualities. Since the task model influences system design, it is important to understand how to manage and overcome possible erroneous events.

11 CTT

Within this work, we are considering CTT as the most appropriate notation because of its graphical appearance and tool support. CTT is used for building task models of cooperative applications in a hierarchical structure while also denoting temporal operators. The task models can be simulated to study different possible paths of interaction. For further information see [17]. CTT has a number of fallbacks as it does not provide dedicated support for numerous attributes of tasks. For example, context and environmental conditions, artefacts being manipulated, user cognitive workload… Due to this lack of support, it is likely that these attributes affecting the task will not be modelled. Modelling such attributes using CTTe would be extremely cumbersome and make the models almost unreadable.

12 Human ERROR

It has been claimed that up to 80% of all aviation accidents are attributed to human ‘error’ [9, p63]. However, of course this is not guaranteed since the situation depends on the user, the system, the context etc but it gives

an overall idea of the potential impact of not addressing this issue carefully. Interactive systems, particularly those that are safety-critical need to be designed with the eventuality of human error in mind to reduce the likelihood of catastrophes. This means considering erroneous behaviour early in the design process and as well as during the testing phase. Although the term "human error" appears very controversial, theories of human errors such as Rasmussen's [20] SRK, Hollnagel's [17] Phenotypes and Genotypes and Norman's [12] classification of slips can be considered widely acceptable. Using the above mentioned classifications and based fundamentally on the SRK theory, we have produced Human Error Reference Tables (HERTs1). These are used for analysing potential user deviations in task models. They enable the exact identification of precise types of error when analysing human behaviour associated to subtasks of a task model.

13 TASK MODELLING & HUMAN ERRORS

A number of papers deal with similar issues to those we are trying to address. In their paper, Baber and Stanton [2] propose a technique, Task Analysis for Error Identification (TAFEI) which is based on the assumption that interaction is goal-oriented and passes through a series of states. The TAFEI approach consists of three stages. A HTA (although any modelling technique is acceptable), construction of a State Space Diagram (SSD) mapped with the HTA plans and construction of a transition matrix to display state transitions during device use. However, the authors consider as erroneous only the paths in a user's task that are provided by the system but do not support the achievement of the user's goal. We consider human error in a broader sense including those previously described.

Paternò and Santoro [16] suggest that a goal is a desired modification of the state of an application. Their work describes how task models can be used in an inspection based usability evaluation for interactive safety-critical applications. Building upon the HAZOP family of techniques, a set of predefined classes of deviations is identified by guidewords such as "none", "other than"and "ill-timed". This set of guidewords could be considered as minimal.

The Technique for Human Error Assessment (THEA) [18] is aimed at helping designers of interactive systems anticipate human errors resulting in interaction failures. Its foundations lay in human-reliability analysis (HRA) [10] and aims to establish requirements for "error resilient" system design. In their paper, it is noted that errors can be regarded as failures in cognitive processing [18]. The process of analysing a system's vulnerability to human error is performed as follows; 1) pose provided questions, 2) identify possible causal factors of the identified potential problems, 3) identify consequences and their impact on the task, work, user and system. The results are recorded in tables for further analysis. Our work differs in that we intend the resulting task patterns to be applicable to more than one system design, possibly of various domains. This means there will be less repetition of work though an increase in work relating to the identification of patterns, recording of patterns, adaptation and application of patterns.

Our proposal is concrete and grounded on previous work in the field of user error identification and classification.

Also, our proposal to define and exploit task patterns provides a way of coping in an efficient and reliable way with the complexity of task modelling.

We have shown that there is a considerable amount of research in the task modelling domain and in the analysis of human error and user task deviation. However, little research has been dedicated to modelling these two aspects together for error-tolerant systems design. The current state of the art has been summarised in Table 1, which identifies where our work fits in.

14 METHODOLOGY

To date, we have devised a systematic way of taking into account, erroneous user behaviour in task models. This work builds upon previous work in the field of task analysis, task modelling, human error analysis and identification. Our research fits into the bigger picture of task model and system model coherence aiming to support error-tolerant systems design. Thus our work is based on a proposed Modelling Process, see Figure 1.

1 Due to space constraints these tables are not further discussed in the paper.

	DEVIATION/ERROR ANALYSIS TECHNIQUE					
HUMAN ERROR CLASSIFICATIONS	**Name**	THEA [18]	TAFEI [2]	Petri-nets [14]	User deviations in task performance [16]	Error explicit task patterns [15]
Error Classification Name	**Based on**	System modelling	Task modelling	System modelling	Task modelling	Task modelling
	Formal?	Semi-Formal	Formal	Formal	Semi-Formal	Formal
Categories of slips [8]		X	x	x	x	√
Mode error [13]		X	x	x	x	√
GEMs [21]		X	x	x	x	√
Failure modes [21]			x	x	x	√
Common mechanisms [21]		X	x	x	x	√
Phenotypes & Genotypes [7]		X	x	x	x	√
SRK [20]		X	x	x	x	
HAZOP Causes of deviations [11]		√	x	√	√	√
Other		X	√	√	x	x

Table 1: Our contribution with respect to state of the art

In particular, we focus on the two phases highlighted by the dotted line. The task and error analysis phase, and the amalgamation phase.

Due to the space constraints, we only discuss our approach to task modelling, error identification, task patterns and where this fits into the design process.

15 TASK PATTERNS

Task patterns for interactive systems design is a relatively new concept. It aims to solve design problems using existing knowledge of identified patterns and solutions. Task patterns were first introduced by Paternò [Breedvelt et al., 1997] & [Paternò, 1999] as reusable structures for task models. The patterns were described as hierarchical structured task fragments that can be reused to successively build the task model.

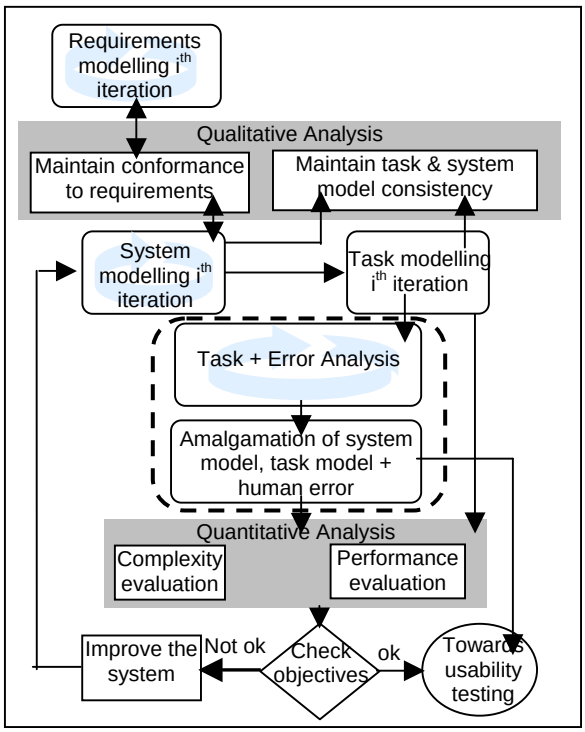

Figure 1: Modified Task & System Model Coherence Diagram

Our approach to task pattern follows the same philosophy but addresses explicitly the error aspect. According to human error theories users' errors are repeated over multiple contexts and multiple systems. For instance post-completion errors occur across a wide variety of systems and are produced by various types of users. Our goal is to provide people responsible for building tasks models with a library of domain independent small task models (that we call task patterns) for them to directly reuse each time they come across an activity available in the library. This process is explained in next section using an illustrative example.

16 ILLUSTRATIVE EXAMPLE

ATMs have often been used to demonstrate task analysis. They are widely used systems demonstrating obvious human-computer interaction. To date we have implemented our approach on an ATM system. It must be noted however, that we have only implemented the skill-based errors of the HERTs and not the rule-based or knowledge-based errors. The ATM system resulted in five task patterns labelled P1, P2, etc. Two of the ATM patterns are shown in Figure 5. After the analysis, we produced an extended task model explicitly expressing possible deviations using the task patterns. It is clear that the model using task patterns to express possible deviations (Figure 6) is far more comprehensible than the model without task patterns incorporating possible deviations (Figure 4).

A simplified process to explicitly express possible user deviations in the task model using task patterns is now described.

Model the task in CTTe, see Figure 2 for a 'standard' task model

Identify and highlight the interactive and cognitive subtasks. We are using the "Insert PIN" subtask in this example.

Depending on the subtasks selected, either:

a. an appropriate task pattern is applied
b. an existing task pattern is adapted and applied or
c. a new task pattern is generated.

For 'c', the HERTs are used to analyse the interactive and cognitive subtasks.

An example of error identification using the HERTs is shown in Figure 2. We can see that the first two analysed errors result in the same outcome, incorrect PIN entry.

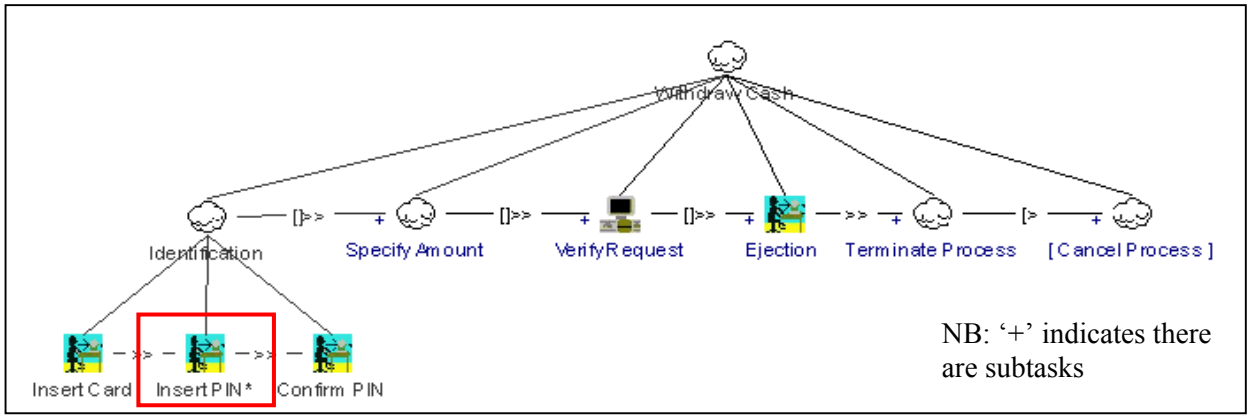

Figure 2: A 'standard' task model for an ATM highlighting 'Insert PIN' subtask for analysis

The subtasks can be remodelled in CTTe to express the errors identified. For example, modelling non-confirmation of the PIN or incorrect PIN entry. Figure 4 illustrates the same ATM task model shown in Figure 2, however this time all possible deviations have been expressed. It is clear that the model is cumbersome and unreadable.

Using the extended task model, we can identify recurring task patterns. The task patterns can be used to remodel the Insert PIN subtask making explicit the possible deviations in a clear manner.

Five task patterns for the Insert PIN subtask were identified:

P1) PIN OK
P2) PIN not OK
P3) Too long entering PIN
P4) Simple Timeout
P5) Timeout on PIN confirmation

See Figure 5 for examples of two task patterns

Figure 6 illustrates the remodelled Insert PIN subtask using the five identified task patterns to clearly express possible deviations.

The remodelled subtasks using task patterns, in this case Insert PIN, can be 'plugged in' to the relevant areas of the original 'standard' task model as illustrated in Figure 6. The result is a legible extended task model which includes both 'standard' and erroneous events using task patterns.

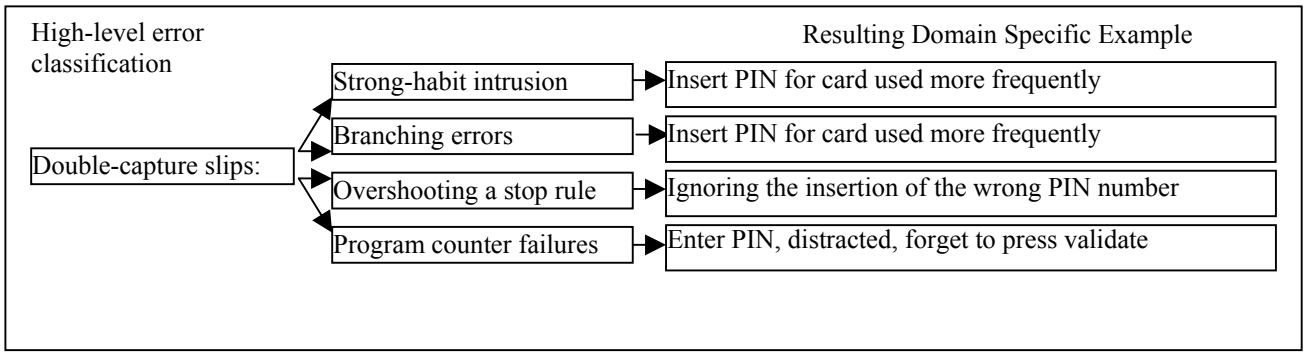

Figure 3: Subset of error analysis for 'Insert PIN' subtask using the HERTs

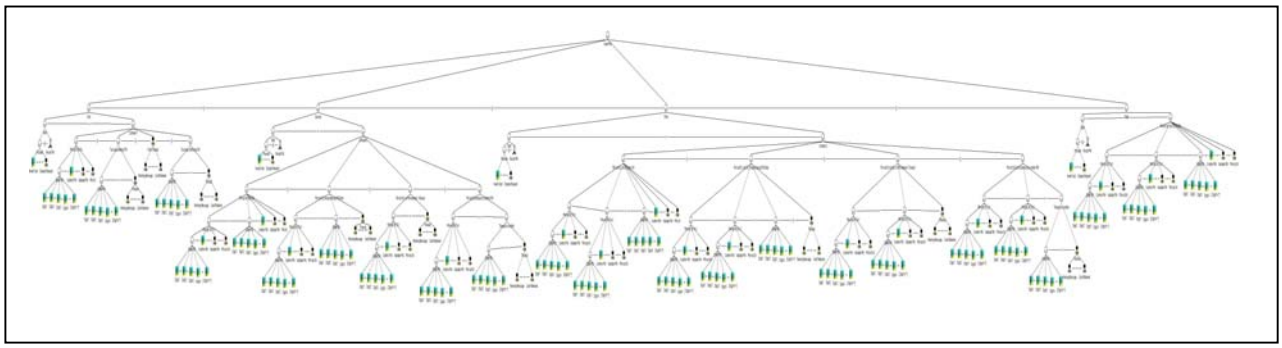

Figure 4: The extended ATM task model explicitly expressing possible deviations

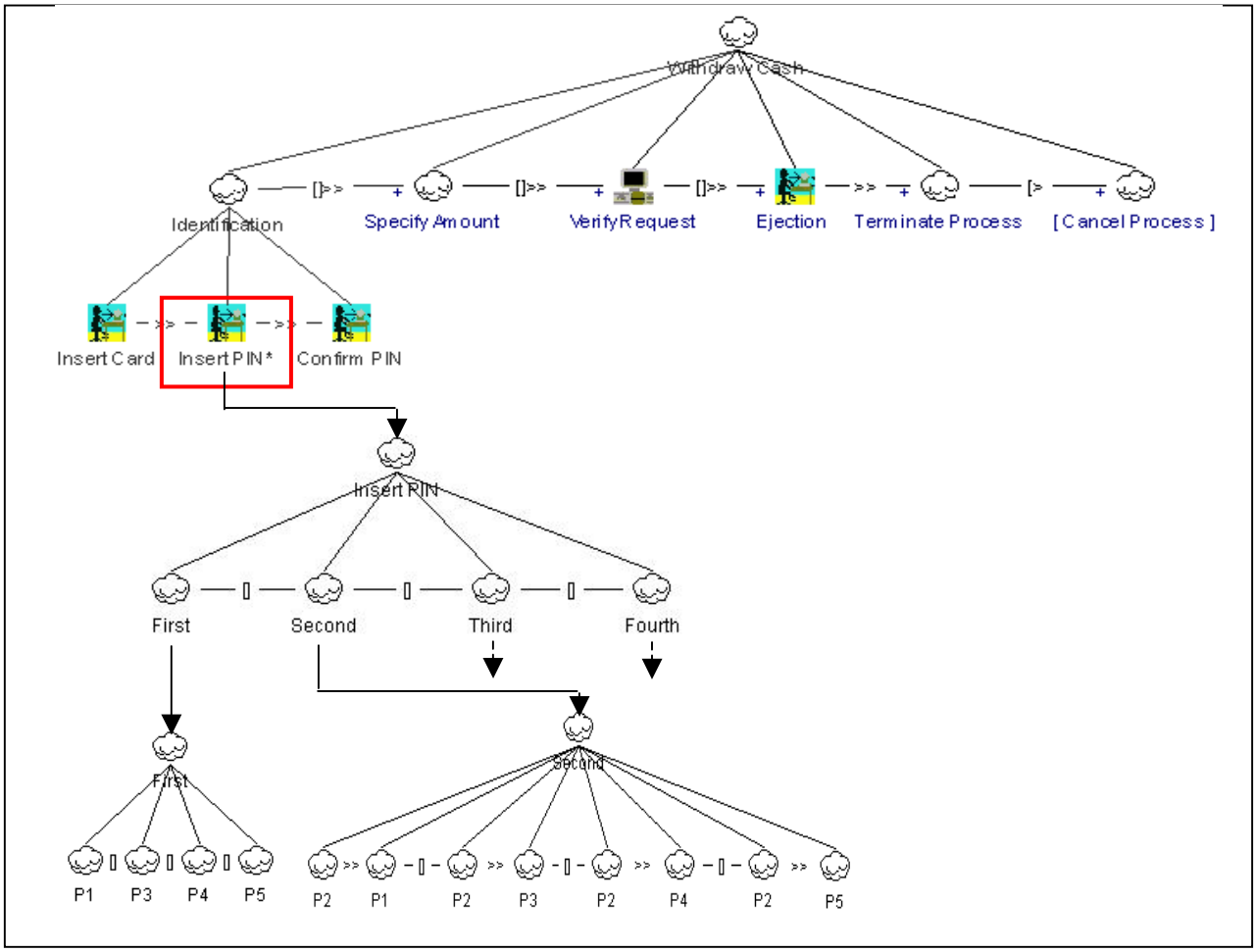

Figure 6: From a standard task model to an extended task model using patterns

17 ONGOING WORK: TOWARDS ERROR-TOLERANT SAFETY-CRITICAL SYSTEMS

We are currently looking into methods for combining human error with task analysis in order to produce re-usable task patterns. The analysis of errors on subtasks has to date been restricted to skill-based errors. We would like to extend this to include knowledge-based and rule-based errors available in the HERTs.

We have identified task patterns for the Insert PIN subtask of an ATM. We plan to try to apply these patterns to a different system also requiring password entry, such as the Windows XP login system to show that the patterns can be applied to various domains. We anticipate that the task patterns may require adaptation. We are aware that the more the patterns are detailed, the less likely they are to be reusable and the more abstract the patterns are, the mode adaptation required.

Therefore, we are also studying methods to support the adaptation process of patterns.

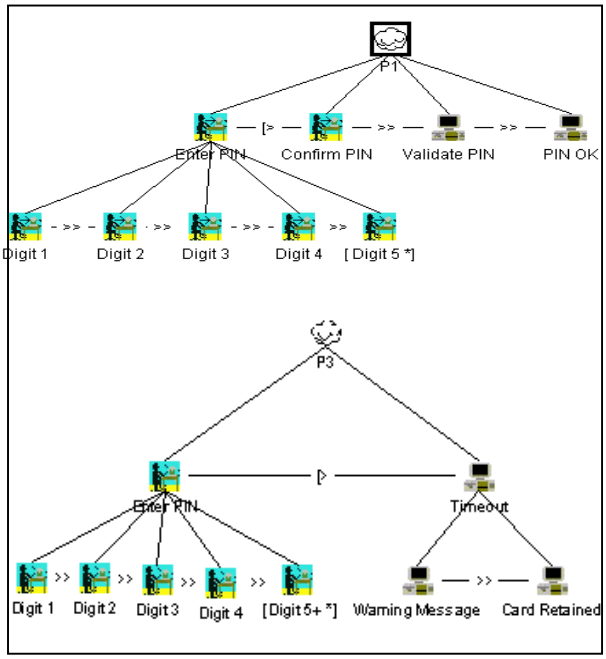

Figure 5: Pattern 1, PIN OK and Pattern 3, Too long entering PIN

18 COLLABORATION IDEAS

With respect to collaboration research with other nodes within the ADVISES network, we anticipate working towards case studies rather than illustrative examples. We are currently working on a case study with the University of Glasgow.

The case study is a report of an accident entitled “Exploding Vessels Under Pressure Accident”. The work will involve the application of several techniques including a fault tree analysis, formal system modelling, safety cases, goal structured notation and System Theory Accident Modelling and Process (STAMP). We are aiming to publish our first joint paper in the Special Issue Of Ergonomics On Command And Control journal.

REFERENCES

1. Annett, J. and Duncan, K. Task Analysis and Training Design, Occupational Psychology. 41, 1967, pp.211-227
2. Baber, C., and Stanton, N. (2004) Task Analysis for Error Identification. In D. Diaper & N. Stanton (Eds.) The Handbook of Task Analysis for Human-Computer Interaction. New Jersey: Lawrence Erlbaum Associates p.367-379
3. Bastide R. & Palanque P. A Visual and Formal Glue between Application and Interaction. International Journal of Visual Language and Computing, Academic Press Vol. 10, No. 5, pp. 481-507. 1999.
4. Breedvelt, I., Paternò, F., Sereriins, C. Reusable Structures in Task Models, Proceedings Design, Specification, Verification of Interactive Systems. Springer Verlag, pp.251-265 (1997).
5. Gullisken, J (1999) User centered design-problems and possibilities: a summary of the 1998 PDC & CSCW workshop**.** ACM SIGCHI, Volume 31, Issue 2 (April 1999) Pages: 25 - 35
6. Hix, D. and Hartson, H. R. Developing User Interfaces. (1993)
7. Hollnagel, E., The Phenotype of Erroneous Actions: Implications for HCI Design. In: Weir, G.R.S. and Alty, J.L., (Eds.), Human-Computer Interaction and Complex Systems, Academic Press. (1991)
8. John B. E. and. Kieras D. E. (1996) The GOMS Family of User Interface Analysis Techniques: Comparison and Contrast. ACM Transactions on Computer-Human Interaction, Vol. 3(4):pages 320--351, December 1996**.**
9. Johnson C.W. (2003) Failure in Safety-Critical Systems: A Handbook of Accident and Incident Reporting. University of Glasgow Press, Glasgow, Scotland, October 2003. P.63.
10. Kirwan, B (1994) A guide to practical human reliability assessment. Taylor and Francis.
11. MOD (1996) HAZOP Studies on Systems Contaning Programmable Electronics. UK Ministry of Defence. Interim Def Stan 00-58, 1996, Issue 1. Available from http://www.dstan.mod.uk/dstan_data/ix-00.htm
12. Norman D.A. The design of everyday things. New York: Currency-Doubleday, 1988.
13. Norman DA (1993). Things that Make us Smart. Addison-Wesley: Reading, MA.
14. Palanque, P & Bastide,R. (1997) Synergistic modelling of tasks, system and users using formal specification techniques. Interacting With Computers, Academic Press, 9, 12, pp. 129-153 1997
15. Palanque, P and Basnyat. S. Task Patterns for Taking into account in an efficient and systematic way both standard and erroneous user behaviours. HESSD 2004. 6th International Working Conference on Human Error, Safety and System Development, 22-27 August 2004, Toulouse, France (within the IFIP World Computing Congress WCC 04).
16. Paternò F. and Santoro C. (2002). Preventing user errors by systematic analysis of deviations from the system task model. International Journal Human-Computer Studies, Elsevier Science, Vol.56, N.2, pp. pp. 225-245, 2002.
17. Paternò, F. Model Based Design and Evaluation of Interactive Applications. Springer Verlag, Berlin (1999)
18. Pocock, S., Fields, B., Harrison, M and Wright, P. (2001) THEA – A Reference Guide. University of York Computer Science Technical Report 336, 2001.

19. Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. & Carey, T. (1994) Human-Computer Interaction. Wokingham, UK: Addison-Wesley.

20. Rasmussen, J. Skills, rules, knowledge: Signals, signs, and symbols and other distinctions in human performance models. IEEE Transactions on Systems, Man, and Cybernetics, 13(3):257-267 (1983)

21. Reason, J. Human Error. Cambridge University Press, 1990.

INTEGRATING HUMAN FACTORS IN THE DESIGN OF SAFETY CRITICAL SYSTEMS

Towards a Safety Architecture

Bastiaan A. Schupp[1]
[1]*University of York, Department of Computer Science, Heslington, York, YO10 5DD, United Kingdom, bastiaan.schupp@cs.york.ac.uk*

Abstract: Human factors contribute to risk in safety critical systems. However, current approaches to integrating human factors issues in the development of safety critical systems appear not fully sufficient. In this paper a we argue for creating a safety architecture. We show how such a architecture can be created using a Safety Modelling Language (SML), which uses barriers to prevent of stop undesired effects. Subsequently we discuss how these barriers can be implemented via various approaches. One of these approaches, the use of activity diagrams is illustrated using a case from the Air Traffic Control domain.

Keywords: Human Factors, Design Methods, Safety, Barriers, Risk, Activity Diagrams, Air Traffic Control

19 Introduction

At the design onset, many safety critical systems are not designed with safety explicitly in mind. Systems are designed, safety is not. Components of systems that are known to be safety critical may be designed up to high standards, formal design methods and verification methods may be used to ensure these high standards. This does not necessarily result in overall system safety or in optimal risk reduction by integration of safety throughout the system. Hence, this design approach is inefficient. It can trigger re-design, it does not allow for risk based evaluation of alternatives and it does not provide for feed back of risk based organisational learning.

In systems that are largely operated by humans these problems are further compounded by the human factor which tends to be harder to manage and understand than the technological. The integration of human factors analysis into systems design is traditionally a difficult problem[1]. Humans are often involved in operating safety systems, can make them fail, and/or are protected by them. More importantly, when humans have to operate a safety system today, its task is most likely complex and hard to automate. Hence the human operator will also face a complex task. This task (and potential failure) will therefore be more difficult to understand for the system designer as well.

A suggested way to improve on some of the design problems of safety critical systems is to create a safety architecture[4]. This is an additional top down perspective on the system that is being designed, which starts with the safety goals, and determines by which functional properties these goals will be reached. In other words, creating the safety architecture allows the design of risk reduction. Such a safety architecture may be particularly useful in further understanding the role of the human in the safety of the system.

One of the obstacles to create a safety architecture is the lack of a suitably expressive language to represent and analyse safety conceptually. Most existing safety methods are for identification or quantification, and do not help in finding solutions. Similar observations are made by Swuste[5] and Harms-Ringdahl[6]. For this reason we are developing an approach that may facilitate the creation of safety architectures. This is based on the previously developed SML[7], which aids designers in conceptually designing risk reduction. It was developed for use in the chemical process domain, but we will show that it can be used in other domains as well, and how to integrate human factors into it.

The format for the remainder of this paper is as follows: Section 2 outlines our approach and formulates our main research objective. Section 3 presents the case study of this paper, which details some of the questions related to that mapping process. It is a simplified case from the air traffic control domain that deals with integrating the human factor in safety systems used in air traffic control. Section 4 presents a brief discussion and concluding remarks.

20 The approach

Central to our approach are barriers and the Safety Modelling Language (SML). Barriers are systems that stop or prevent an unwanted effect; SML uses the Hazard Barrier Target model to develop a safety architecture. For a more detailed discussion of barriers see for instance[8]. As remarked above, this approach constitutes an alternative perspective on the system that is being designed. Systems that are being designed will always have a certain purpose, and can be decomposed in subsystems with sub goals. This is what we can call the production perspective. It defines what the system produces in a hierarchical, top-down manner. For instance it defines that a system is to produce a certain chemical, or to transport people from

A to B. If we decompose this it becomes clear how this production goal will be brought about. Similarly, the safety perspective starts with overall goals, i.e. not to harm the environment, workers, other people which are called primary targets in SML. It also defines the hazards that may cause harm to these targets. As soon as these Hazard Target relations become clear, designers can start thinking of barriers that can prevent the hazards, or defend the targets against their effects. In this manner the designers create the safety architecture in terms of the Hazard Barrier Target Model.
Barriers are often complex socio technical systems, and their function and failure can only be understood when studied in this manner. The socio technical nature of many simple barrier systems such as fire extinguishers or fences may be easily overlooked. A fire extinguisher can only be effectively used if well placed and its user knows how to use it. Similarly, fences have to be in the right place, and people must be instructed not to climb over them. In industry, a fence may be no more then a line on the floor, which sends a message to a person, instead of physically constraining him. Hence a barrier is more then the extinguisher or the fence alone. It also includes a human and organizational component if it is to work. In other barriers humans are more directly involved. Many active barriers, such as the air traffic control system we study in this paper use a detect-decide-deflect sequence. Each of these three activities can be achieved by a human, technology or a combination of the two, facilitated by organisational components.
From the top level, a safety architecture is in our approach decomposed in two manners. Firstly, Barriers are Targets to what we call Functional Hazards. Hence, we can understand barrier failure by looking at these functional hazards, and at why these are caused by the system or its environment. Just as at the primary level, further barriers can be designed which prevent of defend against these functional hazards. We call this principle recursion. Though briefly illustrated in this paper, we will not further explain it here.
The second manner of system decomposition is more important. At the SML level, barriers are considered black boxes. How these function and how these are implemented will not become clear at that level. In this paper we focus at these aspects of a safety architecture. At present we work with the a priori assumption that barriers can be sufficiently understood for design by using three levels of analysis:
The *safety function* level; this concerns the design of role of the barrier in system safety. At this level SML is used to describe the Hazards and Targets and the Barriers that mitigate the risks. It can for instance help to understand how prohibiting smoking may help to reduce fire risks in the greater context of preventing fire in a building. This prohibition constitutes a barrier. As barriers are considered black boxes here, the internal structure and mechanism is ignored at this level.
The *barrier form* level; this concerns how a barrier functions and what its components are, thus what is inside the black box introduced at the safety function level. For example, the components introduced to establish the smoking ban, for instance signs, placement of the signs and enforcement procedures. This level demonstrates which functions the system should provide to allow the barrier to function, also making clear to what extend the barrier is a socio-technical system.
The *embodiment* level; this concerns how the complete system provides the functionality that is required to create the barrier. It sets out how the barrier is implemented conceptually and involves the detailed design of the barrier, and its physical representation and implementation in the safety critical system. For example, the requirement that a non-smoking sign with specified size should be placed at a specified position on every access door, and whom is responsible for enforcing the smoking ban.

This structure thus helps to find out how the barriers should become part of the production system. We are currently working on answering two research questions: How can we properly model barriers to become able to design them abstracted from the system, and how can we use these models to understand the effects on the system as a whole of different barrier configurations at the safety function level? Notice the importance of the human being as part of the barrier here. If the human performs a function in the barrier, the modelling technique must allow modelling that function. This may constrain potential modelling techniques.
To answer the first question a number of modelling approaches are currently being evaluated. These include:

- Structured Analysis and Design Technique (SADT, also called Idef)

- Concurrent Task Trees
- Various UML modelling approaches
- More Formal approaches such as Petri Nets

The use of SADT was discussed in[8]. Concurrent Task Trees seem promising to describe barriers at the form level, especially if humans play a major role in them. However, this approach has not yet been investigated. Formal verification techniques might be particular suited to verify performance and specifications of barriers. In this paper we will further evaluate activity diagrams (part of the UML family) to illustrate and investigate their capacity to model barriers.

21 EN-ROUTE CASE Study2

The above approach is illustrated by means of a case study from the air traffic control domain. Managing air traffic is becoming increasingly difficult due to the increasing traffic volume. At the same time a more flexible traffic control system would allow to save costs, such as lower fuel use. Hence the domain is currently active in researching and developing novel solutions.

Eurocontrol is currently developing a new set of tools that should ease the life of air traffic controllers. These tools will mean an incremental change to current operation, but may also feature in more radical changes in the future, such as the free route airspace concept. The tools under development entail Monitoring Aids (Mona), Safety Nets, Medium Term Conflict Detection (MTCD) and sequencing aids. Mona warns the controller that an aircraft is not at an expected position according to its flight plan, and helps to achieve route changes. Safety Nets are a short term last resort warning system to warn controllers about an imminent (2 minutes) violation of separation (and hence are an important barrier), and sequencing aids facilitate strategic planning of the use of the airspace. These systems are dependent on another novel system, which is called Trajectory Prediction that takes information into account such as the structure of the airspace, the flight plan, current clearances and aircraft performance to predict where the aircraft will be within a certain amount of time.

In this paper the development of Medium Term Conflict Detection (MTCD) is further explored. This is a novel system that warns controllers that two aircraft are on a conflicting course when the conflict may occur in about 20 minutes. Therefore this can most likely be regarded as a barrier.

We start developing this case by creating a safety architecture which allows to study the role of MTCD in overall safety. Subsequently implementation aspects of this architecture are studied using state charts. Some aspects of Air Traffic Control are highly simplified in this study. For instance, the role of Air Traffic Control is assumed only to keep separation distance sufficient, all other tasks are ignored, as is the fact that a strategic and a tactical controller exist. Another simplification is that the safety architecture only supports the problems discussed in this case. In practice it would be far more comprehensive.

21.1 Safety Architecture

Maintaining a safe flight implies maintaining survivable conditions for aircraft passengers. Air Traffic Control (ATC) influences one manner in which passengers will not survive, mid air collisions. Such a collision will have a number of adverse effects on the passengers, for instance extreme forces and rapid loss of cabin pressure. In figure in a SML representation of this top level problem is shown.

2 This case-study is based on various reports as published by Eurocontrol. A key document used here to understand the use of MTCD is[9]. More information used is accessed via http://www.eurocontrol.int/odt/mtcd/, and includes online courses.

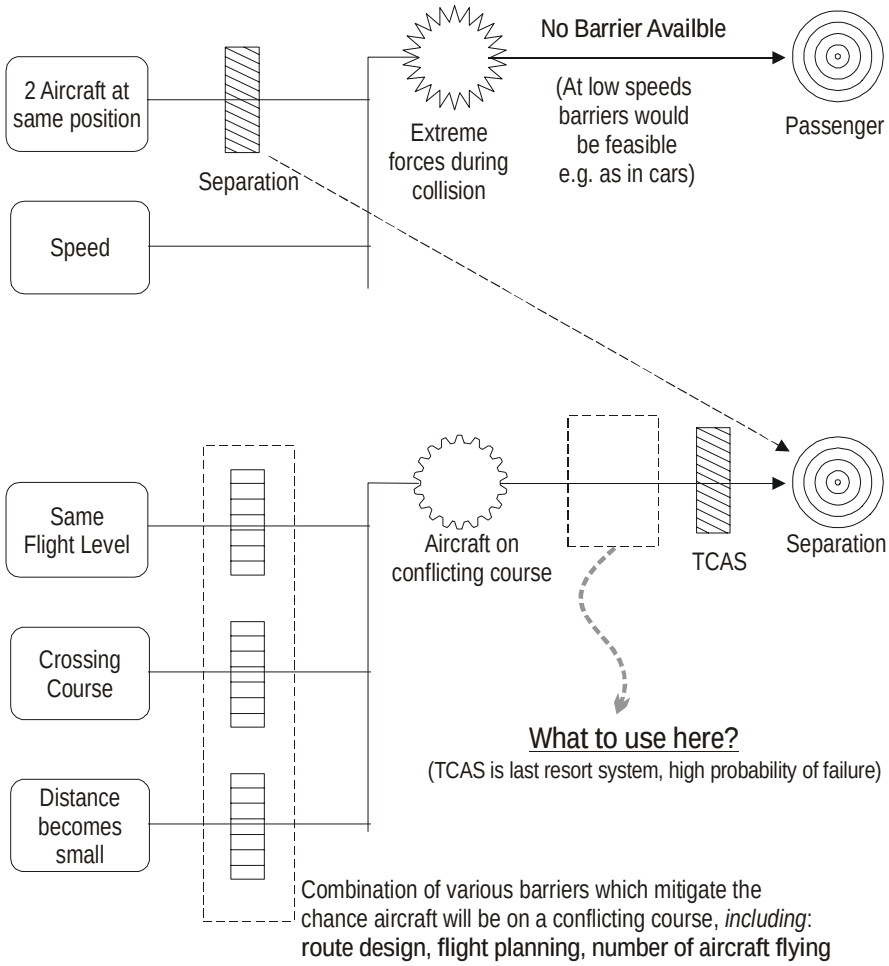

Figure 1. *Part of the safety architecture that defines the separation barrier. A functional hazard to this barrier is aircraft on a conflicting course. A number of mitigative preventive barriers are show, as is TCAS. However, experience shows that this is not enough, and further barriers are required.*

Shown in Figure 1 is the primary hazard to the passengers, extreme forces during a collision. Such a collision occurs when two aircraft are at the same position, and when they are at high speed. Note that during a low speed collision on the ground passengers will not be harmed by extreme forces, though they may be affected by other hazards such as fire, whilst at high speed no barriers are feasible that can mitigate the effect of the forces. In cars barriers such as airbags and seat belts may help to mitigate the effects of an impact, but these are obviously of no benefit in a mid-air collision. Hence, the only available barrier is to maintain sufficient separation.

The major functional hazard to separation distance is aircraft having a conflicting course, as this results in the separation distance being gradually reduced, eventually resulting in a large collision risk. Show in figure 1 is that the chance that this hazard occurs can be mitigated for instance by designing the airspace. However, given current traffic volumes, such conflicts are realistic, and further barriers are required.

One system that can act when a conflict occurs is the Traffic Collision Avoidance System (TCAS) which is implemented in aircraft. It gives a last minute warning and resolution to pilots when a collision is imminent. As this system is considered a last resort system that has a high probability of failure on demand, other systems are needed as well. Safety nets provide a last resort for ATC controllers, but also do not sufficiently reduce the risk.

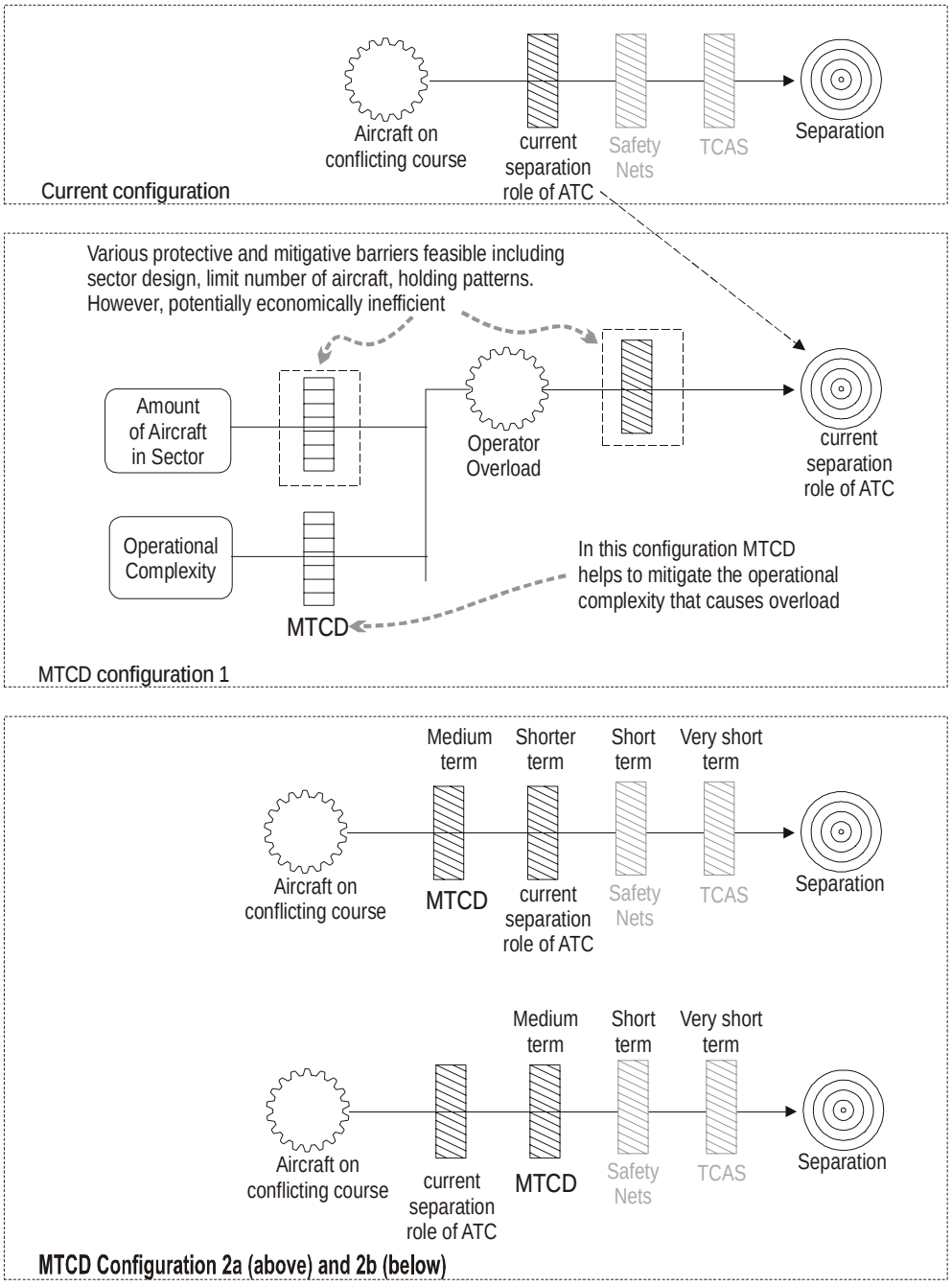

Figure 2. *Two different configurations are shown for protecting the separation barrier against conflicting courses. In the current configuration MTCD is not necessarily used, but can be used to reduce operator overload. In the second configuration MTCD is used directly in resolving conflicts. This may have implications for implementation of MTCD, which is why we refer to the use of MTCD in configuration 2 later in the paper as MTCDR as it in this configuration also must resolve conflicts.*

In figure 2, safety nets and TCAS are greyed, as these are ignored in the remainder of this analysis. Instead we focus at MTCD as a new barrier to complement the existing traditional task of ATC controllers. Today ATC controllers must be vigilant for impending conflicts to maintain separation. Since mid air collisions (or activation of TCAS) occur only infrequently, this current system seems to work fine. However, operator load is increasing which may cause the separation task of ATC to fail more often in the future.

A possible role of MTCD in the safety architecture is to reduce operational complexity, and hence to reduce the number of times that operator overload occurs. It achieves this by helping the operators to search for and identify conflicts, which is a substantial part of their traditional task. This is shown in Figure 2 as MTCD configuration 1. Reduction of overload can also be achieved by reducing the traffic load in a sector, for instance by putting aircraft on hold. Such measures are uneconomical, hence the potential use of MTCD in this way. Notice that calling MTCD a barrier is our wording, not that of Eurocontrol. However, it seems that Eurocontrol strives to use it as a barrier to reduce operator overload, thus as in configuration 1. In a safety study for the future free route airspace concept MTCD is being used to reduce certain risks[10], hence it probably correct to call it a barrier and to give it an explicit role in the safety architecture.
A different MTCD role is to directly help to protect against conflicts. In this manner it will result in a more rigorous protection, as shown in Figure 2, configuration 2a and b. However, if used in this manner MTCD must be implemented differently, as it must do more then just search for and identify conflicts. In this role conflict resolving must also become part of this barrier, as it otherwise would not disappear. This has consequences for the manner how the barrier works, and how it will be implemented. We study these in the next paragraph.

21.2 Implementation of the MTCD configurations

In the previous paragraph we discussed the future need for MTCD to help to maintain sufficient safety. We showed that it can be used in two configurations. One which prevents operator overload, and one that directly defends against conflicts. Here we will study the implementation of these two configurations, to find out how they function, and how these functions become part of the ATC system as a whole. This is achieved at respectively the barrier form level, and the embodiment level, as discussed in paragraph 2. Though we are investigating various modelling techniques, here only the use of activity diagrams is illustrated.
In figure 3, the design of the MTCD system is shown. Three swimlanes are shown, the MTCD system, the ATC operator, and the Pilot. MTCD basically is very simple and has three activities, it receives trajectories from trajectory prediction, it checks whether conflicts exists within these trajectories, and finally reports these on two windows, one for horizontal conflicts, and one for vertical conflicts. In this way it provides information to the operator which eases his work. The operator however has no role in the MTCD barrier.
This becomes different if we where to use MTCD in configuration 2. To work, it must now also involve Resolution (hence the better name would be in this configuration MTCDR). To resolve a conflict the ATC operator and the pilot must have to be involved, because detection alone is not going to stop an accident. Hence, the activities in the ATC operator and Pilot swimlane will be part of the barrier now, stressing its socio technical nature.
In figure 4 the embodiment level is shown. In this paper it is only illustrated how the barrier becomes embodied in the task of the human ATC operators. For this we also greatly simplified the task of these people, just letting them start working, check for and resolve conflicts, and check for the end of their shift. Obviously they have many other tasks, such as handovers from sector to sector, route planning, and so on. These are not displayed here. Also the embodiment of other parts of the MTCD system is not shown, such as the implementation in the computer systems, or in the tasks of the pilots.
Figure 4 makes clear how the task of the controller will change when the different configurations are implemented. In the first configuration part of the search and identification work is now achieved by MTCD, hence the reduced workload. In case of configuration 2, the changes to the tasks of the operator are actually very minor. This is because MTCDR is largely implemented using existing tasks of the ATC operator. Notice that these activity diagrams actually implement two consecutive barriers, MTCDR and manual separation (in different order for respectively configuration 2a and 2b). The ‘new’ tasks that appear in the form level design of MTCDR (Figure 3), can all be implemented by reuse of existing tasks. Of course, there must be a way of achieving this implementation, for instance by training and new procedures. Hence, this training becomes part of the MTCDR barrier as well, but this cannot be shown using activity diagrams.

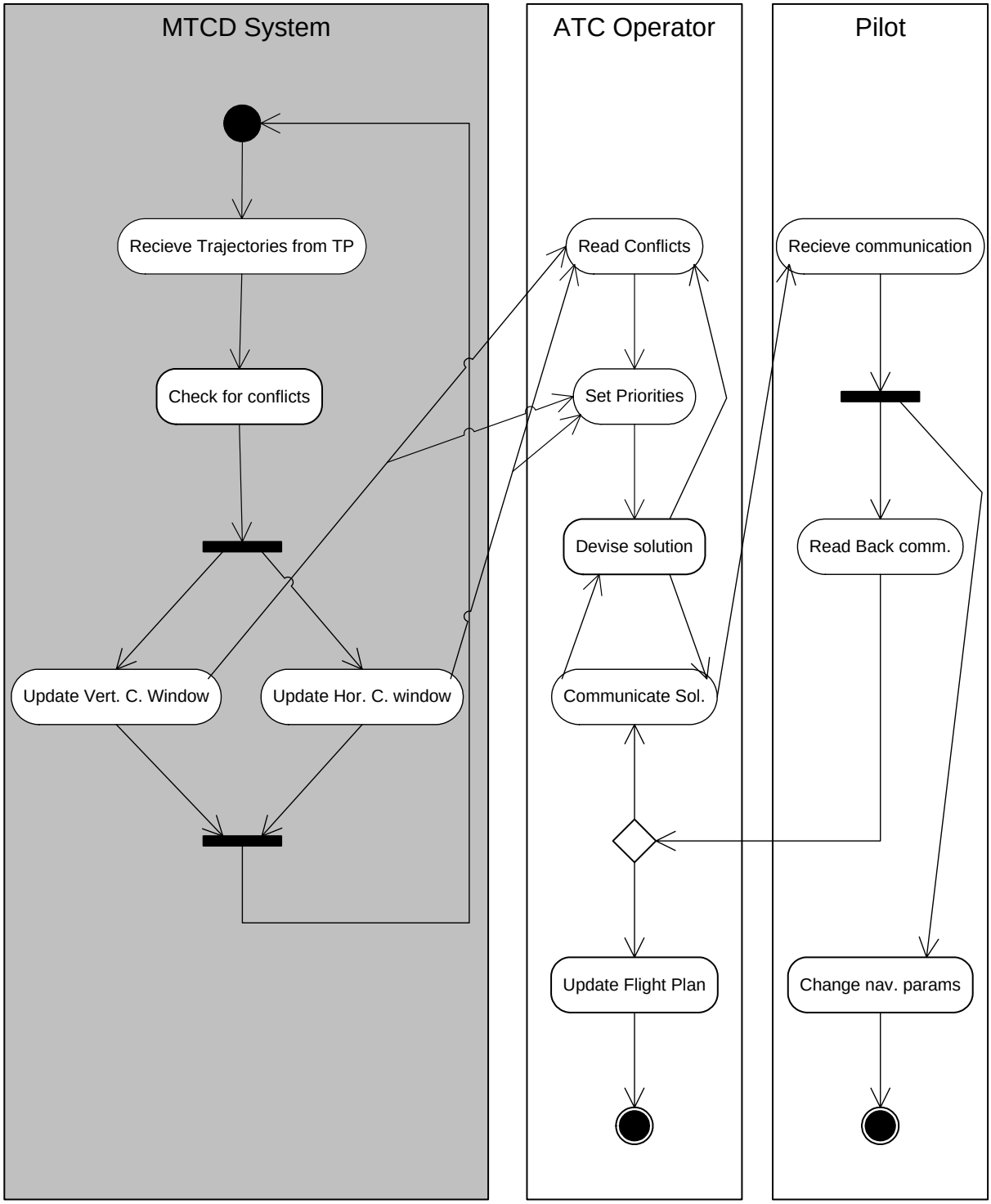

Figure 3. *MTCD(R) activity diagram. The grey swimlane describes what the MTCD barrier is in configuration 1. To use it in configuration 2 (as MTCDR) also the ATC controller and the pilot become part of the barrier, hence the activities in the other two swimlanes must be included as part of the barrier it it is to be used in configuration 2.*

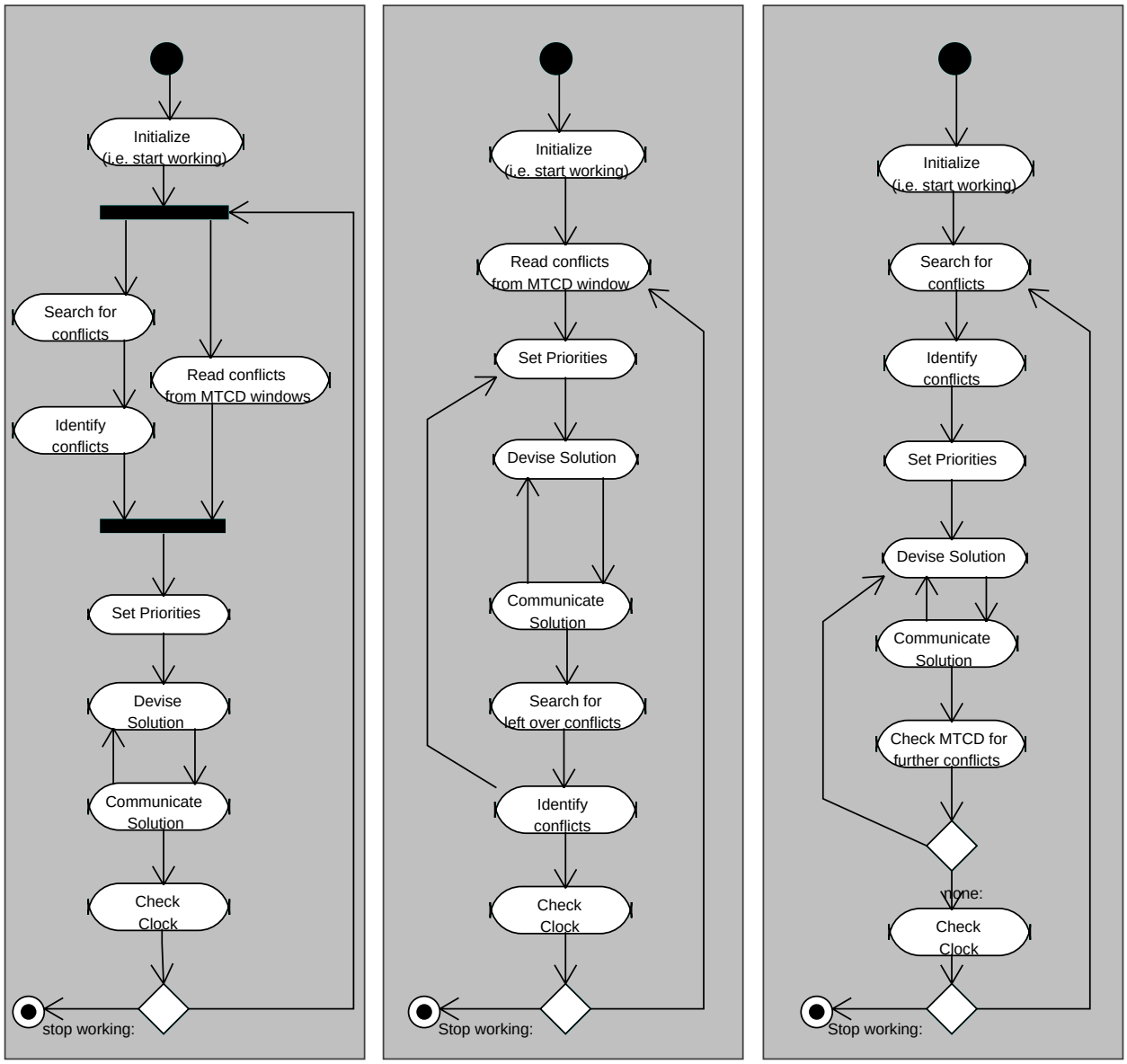

Figure 4. *Here the embodiment of configuration 1, 2a and 2b is shown (from left to right). The figures only show how these barriers are embodied in the work of the air traffic controllers. For this purpose the controllers task is greatly simplified. Here we assume that the controller starts working, resolves conflicts and checks the clock to see whether it is time to go home. Other tasks include planning, optimization of course and handovers to other sectors.*

22 CONCLUDING REMARKS

In this paper we have discussed some aspects of our research into a possible novel approach to design the safety of safety critical systems. Some aspects of a safety architecture have been briefly discussed and illustrated. We described the use of barriers, and how these can be designed at three levels. One level which describes their role in risk reduction, one that describes how they function (form level), and one which describes how they become part of the system (embodiment level). We used SML to define the role in risk reduction, and activity diagrams in the case of form and embodiment level.

Our current research focuses at how to best achieve design at form and embodiment level. Some limitations of the chosen activity diagrams became already apparent, such as that they cannot express everything there is to a barrier (procedures, training), and others have not yet been discussed (system integration, object structure). One of the common problems to modelling techniques (including Idef, other UML and task modelling techniques) is that they seem to have difficulties in coping with concurrency of tasks and activities. In this case the ATC operators would have been busy carrying out many other tasks, and many conflicts would have to be dealt with at the same time. This is difficult to express using the discussed techniques. However, the problem that underlies the creation of MTCD, operator overload, can only be properly understood when the multiplicity of operator tasks can be correctly described. Hence we will divide some attention to research in this area.

REFERENCES

1. Hollnagel, E., *Human Reliability Analysis : Context and Control*. Computers and People Series. 1993, London ; San Diego, CA: Academic Press. xxvi, 326 p.p.
2. Arthur D. Little Inc., American Institute of Chemical Engineers. Center for Waste Reduction Technologies, and American Institute of Chemical Engineers. Center for Chemical Process Safety, *Making Ehs an Integral Part of Process Design*. 2001, New York: CWRT CCPS, American Institute of Chemical Engineers. xvi, 164 p.p.
3. Schupp, B.A., S.M. Lemkowitz, L.H.J. Goossens, A.R. Hale, and H.J. Pasman. *Modeling Safety in a Distributed Technology Management Environment for More Cost-Effective Conceptual Design of Chemical Process Plants*. In Computer-Aided Chemical Engineering; European Symposium on Computer Aided Process Engineering - 12: 2002.ELSEVIER SCIENCE BV: p. 337-42.
4. Nisula, J. *Challenge of Safety Data Analysis Models Wanted*; IFIP WCC 2004 /HESSD: 2004.Kluwer Academic Publishers: p. 224-37.
5. Swuste, P., *Occupational Hazards, Risks and Solutions*, thesis, 1996, Delft University of technology, 217 p.
6. Harms-Ringdahl, L., *Assessing Safety Functions - Results from a Case Study at an Industrial Workplace.* Safety Science, 2003. 41(8): p. 701-20.
7. Schupp, B.A., S.M.L. Lemkowitz, and H.J. Pasman. *Application of the Hazard-Barrier-Target (Hbt) Model for More Effective Design for Safety in a Computer-Based Technology Management Environment*; CCPS ICW: Making Process Safety Pay: the business case: 2001.AIChE/CCPS.
8. Schupp, B.A., S.P. Smith, P. Wright, and L.H.J. Goossens. *Integrating Human Factors in the Design of Safety Critical Systems; a Barrier Based Approach*; Human Error, Safety and Systems Development: 2004.Kluwer Academic Publishers: p. 285-300.
9. Eurocontrol, *Mtcd Shadow Mode Trials at Maastricht Upper Area Control Centre*. 2004, Eurocontrol. p. 51.
10. Eurocontrol, *Safety Assessment of the Free Tour Airspace Concept*. 2001, Eurocontrol. p. 128.

Mission

To promote an innovative and environmentally sustainable technological development within the areas of energy, industrial technology and bioproduction through research, innovation and advisory services.

Vision

Risø's research **shall extend the boundaries** for the understanding of nature's processes and interactions right down to the molecular nanoscale.

The results obtained shall **set new trends** for the development of sustainable technologies within the fields of energy, industrial technology and biotechnology.

The efforts made **shall benefit** Danish society and lead to the development of new multi-billion industries.